

---

## RAPPORT DE STAGE

---

# *Valorisation d'une prospection électrique pour établir une carte de l'épaisseur du sol*

19/09/2013

Salle de conférences de GéoRessources (Université de Lorraine, Faculté des Sciences et  
Technologies à Nancy)



Stage encadré par :

Hocine BOURENNANE

Maud SEGER

Marjorie UBERTOSI

Présenté par : Daril NDJIBA MITOMBO

Année académique 2012 - 2013

## **Remerciements**

Mes remerciements vont, premièrement, à Catherine Hénault, Directrice de l'Unité de Recherche de Science du Sol de l'INRA d'Orléans pour m'avoir accepté au sein de cette unité dont elle a la charge.

Je remercie ensuite Marie-Noël Mistou, responsable du projet CAREX, qui a rendu possible ce stage et qui n'a ménagé aucun effort pour le bon déroulement de ce dernier.

Je remercie également mes encadrants : Hocine Bourennane, Maud Séger et Marjorie Ubertosi pour le soutien, la confiance et l'entière collaboration dont j'ai bénéficié tout au long de ce stage. Merci à Rodolphe Hugard, Technicien de l'UE d'Epoisses, qui a réalisé les prospections électriques et qui a mis à disposition son temps et la documentation en rapport avec ces prospections.

Je remercie particulièrement Denis Baize, Isabelle Cousin pour leurs conseils et le temps qu'ils ont mis à disposition pour répondre à mes questions. Merci à Catherine Pasquier pour l'aide à l'apprentissage du SIG et merci à Alain Couturier et à Birama Sall pour leur soutien en informatique.

Je tiens à remercier toutes les personnes qui ont contribué de près ou de loin à mon intégration et au bon déroulement de mon stage : Sylvaine Parot et Marie Pierre Poussineau pour les aspects administratifs, Joseph Levilain qui a accepté de partager son bureau avec moi, Florian, Caroline, Leslie, Viviane, Maxime, Eva, Nathalie, Eugénie, Estelle...

Enfin, grand merci à l'ensemble du personnel des deux unités INFOSOL et UR SOLS pour l'atmosphère détendue de travail, entrecoupée des pauses cafés, pots de départ, soirée cinéma et restaurant, qui m'ont permis de passer un agréable séjour à Orléans en général et à l'INRA en particulier.

## Sommaire

PRESENTATION DE L'INRA.....	1
<b>1. Introduction générale.....</b>	<b>2</b>
<b>2. Analyse bibliographique .....</b>	<b>3</b>
<b>3. Site et données disponibles.....</b>	<b>5</b>
3.1. Présentation du site de l'Unité Expérimentale d'Epoisses .....	5
3.2. Mesure de la variable auxiliaire : la résistivité électrique .....	6
3.2.1. Résistivité électrique et facteurs d'influence .....	6
3.2.2. Présentation du dispositif ARP.....	6
3.2.3. Prospection électrique réalisée dans le cadre de cette étude.....	7
3.3. Mesure de la variable d'intérêt : Epaisseur de sol selon le plan d'échantillonnage établi à partir des mesures de la résistivité électrique.....	7
<b>4. Approches de spatialisation mises en œuvres.....</b>	<b>8</b>
4.1. Démarche générale .....	8
4.2. Méthodes spatiales et non spatiales mises en œuvre .....	9
4.3. Etape préliminaire : Organisation des données de résistivité sur une grille régulière .....	9
4.4. Estimation de l'épaisseur du sol par l'approche spatiale (Scenarii A et B).....	10
4.4.1. Analyse en composante principale (ACP).....	10
4.4.2. Etude de la structuration spatiale de PCI.....	10
4.4.3. Construction de scenarii basés sur l'approche spatiale.....	11
4.5. Estimation de l'épaisseur du sol par l'approche non spatiale (Scenarii C et D).....	11
4.5.1. Estimation de l'épaisseur du sol à partir de PCI (scénario C).....	11
4.5.2. Estimation de l'épaisseur du sol par choix d'une voie de résistivité (scénario D) .....	11
4.6. Validation croisée .....	11
<b>5. Résultats et discussion.....</b>	<b>14</b>
5.1. Caractérisation pédologique des sols de l'UE d'Epoisses .....	14
5.2. Données de résistivité électrique .....	14
5.3. Estimation de l'épaisseur du sol par l'approche spatiale (Scenarii A et B).....	16
5.3.1. Analyse en composante principale des voies de résistivité .....	16
5.3.2. Estimation de l'épaisseur du sol par krigeage en considérant PCI et S2-PCI comme dérive externe .....	17
5.4. Estimation de l'épaisseur du sol par l'approche non spatiale (Scenarii C et D).....	18
5.4.1. Utilisation de la voie de résistivité la plus corrélée à l'épaisseur du sol.....	18
5.4.2. Utilisation de la composante PCI de la résistivité.....	18
5.5. Cartes de l'épaisseur du sol obtenus par les approches spatiales et non spatiales .....	18
5.6. Validation croisée .....	21
<b>6. Conclusion.....</b>	<b>23</b>
<b>Références bibliographiques.....</b>	<b>24</b>

## **LISTE DES FIGURES**

Figure 1 : Localisation du site d'étude, domaine expérimental de l'INRA d'Epoisses .....	5
Figure 2 : Dispositif ARP (a) lors d'une mesure sur Epoisses et (b) schéma.....	7
Figure 3 : Plan d'échantillonnage pour la mesure de l'épaisseur du sol à partir de la carte de résistivité électrique de la voie 2 de l'ARP .....	8
Figure 4 : Logigramme décrivant les différentes étapes des approches testées.....	13
Figure 5 : Cartes des résistivités mesurées par l'ARP sur les trois voies de mesure : R1 pour les résistivités de la voie 1, R2 pour les résistivités de la voie 2 et R3 pour les résistivités de la voie 3.....	15
Figure 6 : Représentation des variables dans le plan factoriel PC1-PC2.....	17
Figure 7 : Variation des valeurs de PC1 en fonction des valeurs Gaussiennes de l'épaisseur du sol .....	17
Figure 8 : Variogramme expérimental (points verts) et modèles théoriques ajustés (courbe rouge) de PC1..	17
Figure 9 : Cartes d'épaisseur de sol estimée selon les 4 scenarii .....	19
Figure 10 : Comparaison entre les épaisseurs de sol estimées au droit des épaisseurs de sol mesurées pour les 4 scenarii.....	20
Figure 11 : Résultats de la validation croisée : indicateurs statistiques et diagrammes de dispersion entre les épaisseurs estimées et les épaisseurs mesurées .....	21

## **LISTE DES TABLEAUX**

Tableau 1 : Résultats d'analyses pédologiques réalisées à partir des échantillons de sol sur l'UE d'Epoisses.....	14
Tableau 2 : Statistiques descriptives des trois voies de résistivités .....	15
Tableau 3 : Matrice de corrélation sur les données organisées sur la grille 1 m <sup>2</sup> .....	16
Tableau 4 : Corrélation entre la variable d'intérêt (épaisseur du sol) et les variables auxiliaires (résistivité)..	18
Tableau 5 : Comparaison des classes d'épaisseurs prédites versus les classes d'épaisseurs mesurées .....	20

## **PRESENTATION DE L'INRA**

L'Institut National de la Recherche Agronomique « INRA », est un établissement public à caractère scientifique et technologique placé sous la double tutelle des ministères chargés de la Recherche et de l'Agriculture. Créé au lendemain de la seconde guerre mondiale, en 1946, il avait pour mission, à l'origine, de mettre la science et la technologie au service du développement de l'agriculture en améliorant les techniques de production culturale et végétale, afin de répondre aux besoins alimentaires de la population. Aujourd'hui, l'INRA oriente ses travaux sur trois domaines fortement liés : l'alimentation, l'agriculture et l'environnement avec pour but de développer une agriculture compétitive respectueuse de l'environnement, des territoires et des ressources naturelles, et adaptée aux besoins nutritionnels de l'homme ainsi qu'aux nouvelles utilisations des produits agricoles. De ce fait, les sols formant l'interface entre atmosphère, biosphère, lithosphère et hydrosphère, support de toutes récoltes et desquels dépendent la qualité des aliments s'imposent également comme sujet de recherche.

### **Centre de recherche INRA Val de Loire et l'Unité de Recherche de Science du sol**

Le centre INRA Val de Loire est réparti sur trois sites : Orléans, Tours et Bourges, et mène des recherches autour de quatre pôles :

- Dynamique des sols et gestion de l'environnement
- Biologie intégrative des arbres et organismes associés
- Biologie intégrative animale et gestion durable des productions animales
- Santé animal et santé publique

L'Unité de Recherche de Science du Sol (l'UR SOLS) dans laquelle j'ai effectué mon stage est située sur le site d'Orléans (commune d'Ardon). Ses recherches font partie du premier pôle « Dynamique des sols et gestion de l'environnement ».

L'UR SOLS dirigée par Catherine Hénault est une unité propre de recherche du département Environnement et Agronomie de l'INRA. Elle a pour mission principale de développer des travaux de recherches 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, présentés sous la forme d'un projet unique intitulé « Sol'Hyd'Air », ont vocation à contribuer à une meilleure préservation des sols, une meilleure gestion de la ressource en eau et à la protection de la qualité de l'air. L'unité est composée d'une équipe de chercheurs, ingénieurs et techniciens. Les travaux qui y sont réalisés s'appuient sur des activités de laboratoire (simulation de pluie, minéralogie, physique du sol, géochimie...), des activités de terrain (géophysique, cartographie, suivis hydriques...) et des activités d'analyse de données, notamment spatiales et de modélisation du fonctionnement des sols.

## 1. Introduction générale

Le sol est la partie meuble et superficielle de la croûte terrestre. Il résulte de la transformation d'un matériau parental (substratum géologique) et de la dégradation des matières organiques. Le sol n'est donc pas un milieu inerte et stable, mais se forme et se développe sous l'influence conjuguée de plusieurs facteurs environnementaux : nature de la roche mère, relief, climat, activité biologique (Baize et Jabiol, 2011), responsables de leur grande variabilité et leur grande diversité dans l'espace à toutes échelles, de la particule à l'échelle globale. La diversité des sols influe directement sur les propriétés des sols et leur fonctionnement, desquels dépend leur fertilité. Le sol est le support et la source nutritive des cultures. Une bonne connaissance du sol, de ses caractéristiques, de sa variabilité s'avère alors primordial, pour développer une agriculture respectueuse de l'environnement, comme par exemple, dans le cadre de l'agriculture de précision qui a pour objectif d'ajuster au mieux les pratiques agricoles à l'échelle parcellaire (Robert, 1999).

L'INRA réalise des expérimentations agronomiques sur les sites d'unités expérimentales (UE) dont les contextes environnementaux ne sont pas toujours bien connus. Le projet CAREX a donc été mis en place afin de proposer une méthodologie de caractérisation environnementale applicable à toutes les unités expérimentales dans le but de quantifier les stress hydriques et azotés rencontrés par les cultures expérimentées lors des essais. Ce projet s'articule autour de trois grands axes: (1) la caractérisation ou connaissance du sol pour fournir les informations nécessaires à l'établissement de bilans hydriques et azotés, (2) l'adaptation de modèles de suivi des bilans hydriques et azotés avec le double objectif : ajuster les programmes d'intervention culturales et quantifier les stress hydriques et azotés subis par les plantes au cours d'un cycle de végétation et enfin (3) l'instrumentation des sites expérimentaux afin de calibrer et évaluer les modèles à partir de mesure *in situ* sur le sol et la plante et faire des calibrations des modèles si besoin.

Le stage s'inscrit dans le premier axe du projet CAREX et son objectif est **de proposer une approche opérationnelle de cartographie de l'épaisseur du sol, variable nécessaire aux calculs de stocks et donc des bilans hydriques et azotés.**

L'étude a été conduite à partir des mesures de résistivité et de profondeur de sol réalisées sur **le site expérimental de l'UE d'Epoisses, en Bourgogne.**

## 2. Analyse bibliographique

Classiquement, la répartition géographique de la diversité des sols est décrite grâce à l'expertise du pédologue cartographe qui exprime la synthèse de son travail sous forme d'une carte (King, 1984). Le pédologue, s'aidant de l'ensemble des informations disponibles (morphologie du relief, cartes géologiques ou pédologiques préexistantes, photographies aériennes...), décrit le sol à l'aide de sondages à la tarière et d'observation de fosses pédologiques. Il identifie les facteurs de la pédogenèse à l'origine de la différenciation des sols et propose alors un découpage de la couverture pédologique en unités cartographiques homogènes pour des travaux à grande échelle (sur de petites surface), et parfois hétérogènes et complexes (constituées de plusieurs types de sols) pour des travaux à petite échelle (sur de grandes surface) (Jamagne, 1993). Cette méthode implique des coûts importants, liés au nombre d'échantillons jugé nécessaire en fonction de la variabilité même du milieu et la précision recherchée et à la durée nécessaire pour prospecter une couverture pédologique de grande surface. Elle repose principalement sur l'efficacité ou expertise du pédologue cartographe et la difficulté d'expression du résultat de ses travaux à des échelles dites moyennes et petites (King et *al.*, 1994) constituent les principales limites de celle-ci. Dans certaines situations, on ne produit pas une carte pédologique au sens « type de sol », mais plutôt une carte thématique, qui présente uniquement certaines propriétés jugées d'intérêt. Parmi celles-ci, on peut citer : la texture du sol, la profondeur du sol, le degré d'hydromorphie, etc. La profondeur du sol, utilisée pour calculer des stocks d'éléments ou des réserves, constitue une variable d'importance dans de nombreux contextes.

La démarche du pédologue cartographe étant rarement formalisée, elle est donc difficilement reproductible. Dans les années 80-90 se sont alors développés des travaux qui ont eu pour objectif de formaliser cette démarche, en s'appuyant sur des méthodes mathématiques et statistiques (Lagacherie, 1992, par exemple).

Les possibilités d'acquisition de données numériques de façon exhaustive sur une parcelle ou sur l'ensemble d'un site ont permis, grâce à des traitements statistiques et/ou géostatistiques, le développement de nouvelles approches dans le domaine de la modélisation et de la prédiction, dans l'espace, d'une propriété du sol à partir d'un nombre fini d'échantillons (Walter, 1990 ; Bourennane, 1997). Ces nouvelles approches permettent ainsi de traiter des données de nature et de densité différentes. Elles supposent que des liens statistiques existent entre la variable du sol à analyser (variable d'intérêt), et la variable utilisée pour améliorer l'estimation de celle-ci (variable auxiliaire). Parmi les variables auxiliaires utilisées en cartographie pédologique, on peut citer les modèles numériques de terrain (MNT) et les techniques de prospection géophysique (images satellitaires, résistivité électrique). En effet, certains attributs du relief – pente, convexités de la pente – sont liés à l'épaisseur du sol (par exemple Bourennane, 1997; Chartin et *al.*, 2011). La résistivité électrique du sol dépend également de plusieurs propriétés du sol (Friedman, 2005 ; Samouëlian et *al.*, 2005) : on peut citer la teneur en argile, la teneur en eau, la température du sol (par exemple, Campbell et *al.*, 1948), la composition de la solution du sol, la masse volumique (par exemple Besson et *al.*, 2004 ; Séger et *al.*, 2009), la charge en éléments grossiers (par exemple Tetegan et *al.*, 2012). La méthode géophysique de mesure de la résistivité électrique, facile à mettre en place, non destructrice, peut

ainsi permettre de caractériser les hétérogénéités du sol à l'échelle de la parcelle avec une haute résolution de l'ordre du mètre (Bottraud et *al.*, 1984 ; Moeys et *al.*, 2006).

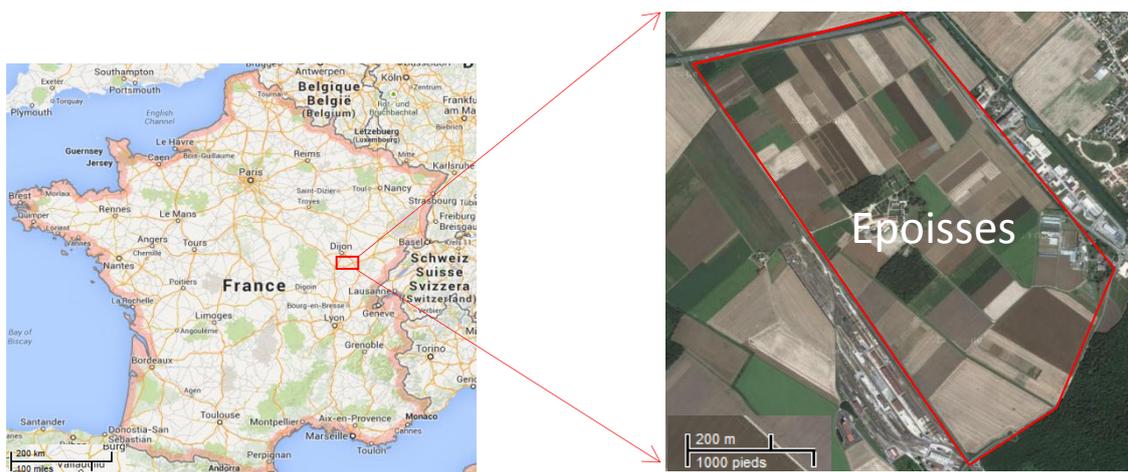
Ainsi, Chéry et *al.*, (1994) ont estimé, par cokrigage, l'épaisseur du sol en utilisant la résistivité électrique, sur une surface de dix hectares. Tetegan et *al.*, (2012) ont utilisé une régression linéaire simple pour estimer la charge en éléments grossiers sur une parcelle de sols caillouteux à partir de données de résistivité électrique. En 2000, Bourennane et *al.*, ont comparé une approche statistique et une approche géostatistique pour la cartographie de l'épaisseur du sol à grande échelle. La profondeur d'apparition du toit calcaire a été également estimée par krigeage en utilisant l'intensité de la pente dérivée d'un MNT et la résistivité électrique comme dérivées externes sur une vingtaine d'hectares (Bourennane et King, 2003). Dans certains cas, la structure spatiale de la variable auxiliaire peut également être analysée et décomposée, et une seule composante peut être utilisée pour estimer la variable d'intérêt. Par exemple, sur une parcelle de 20 hectares, Bourennane et *al.*, (2012) ont amélioré l'estimation de la teneur en eau du sol en utilisant la composante à basse fréquence de la résistivité électrique.

Dans les travaux précités, la variable auxiliaire est toujours utilisée *a posteriori*, et n'est pas employée pour guider l'échantillonnage du pédologue, alors que, concrètement, elle dépend des variables pédologiques. De plus, les travaux présentés ci-dessus testent indépendamment différentes approches géostatistiques. Dans ce cadre, **notre travail a donc consisté à tester et comparer plusieurs approches statistiques et géostatistiques (spatiales et non spatiales) pour estimer, à grande échelle (site de l'UE d'Epoisses), l'épaisseur du sol à partir d'une prospection de résistivité électrique, laquelle a également servi de support au plan d'échantillonnage spatial.** Notre démarche a donc consisté à tester quatre scénarii d'estimation de l'épaisseur du sol. Ces scénarii sont organisés selon deux approches, l'une spatiale (par krigeage avec dérive externe), l'autre non spatiale (par régression linéaire simple). L'épaisseur du sol estimée selon les quatre scénarii a été évaluée par une procédure de validation croisée.

### 3. Site et données disponibles

#### 3.1. Présentation du site de l'Unité Expérimentale d'Epoisses

Le domaine de l'Unité Expérimentale de l'INRA d'Epoisses est situé à 10 km au Sud-Est de Dijon (Côte d'Or), à la bordure occidentale de la vallée de l'Ouche, sur la commune de la Bretenière (Fig. 1). Le domaine compte 120 hectares environ, subdivisés en de nombreuses parcelles expérimentales accueillant différentes variétés d'espèces (maïs, blé, orge, betterave, par exemple). Ces cultures servent de support à des recherches en agroécologie dans des conditions réelles de culture en plein champ.



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

La région se caractérise par un climat de type semi-continental avec une pluviométrie annuelle comprise entre 700 et 800 mm. Sur le plan géologique, la vallée de l'Ouche représente une large plaine alluviale qui s'étale sur une largeur de 3 à 8 km.

La carte pédologique de Dijon, au 1/100000<sup>e</sup> (Chrétien, 1976), indique que la zone d'étude présente deux types de sols : des sols de type rendzines brunifiées et des sols bruns calciques épais 'argilo-limoneux, localement humifères sur cailloutis calcaires des basses terrasses des Tilles et de l'Ouche.

Les rendzines brunifiées, de faible épaisseur, (de 20 à 40 cm) sont constituées d'un horizon reposant directement sur le cailloutis calcaire. Leur teneur en argile est comprise entre 34 et 49 % (41 % en moyenne) avec un taux de matière organique de 5,3 % en moyenne. Leur pH est compris entre 7,0 et 8,1 et la capacité d'échange, entièrement saturée en calcium, oscille entre 25 et 35 mé/100g.

Les sols bruns calciques d'épaisseur variant entre 50 et 100 cm, sont en général argilo-limoneux en surface (37 % d'argile en moyenne) et argileux en profondeur (46 % en moyenne). Leur pH est nettement alcalin (7,8 et 8,1) et leur complexe absorbant toujours saturé.

On note la présence d'une nappe souterraine à 0,5 voire 1,5 m de profondeur en moyenne, qui s'écoule du nord-ouest vers le sud-est avec une faible pente de 0.22 % (1 m/450 m) parallèle à la surface topographique, et un ancien réseau de drainage.

### 3.2. Mesure de la variable auxiliaire : la résistivité électrique

#### 3.2.1. Résistivité électrique et facteurs d'influence

La résistivité électrique  $\rho$  d'un sol est la propriété physique qui détermine sa capacité à s'opposer au passage d'un courant électrique. En termes mathématiques, elle correspond à la résistance ohmique  $R$  (ohm) d'un cylindre de section  $S$  ( $m^2$ ) et de longueur  $L$  (m) unitaire. En pratique, la mesure de la résistivité électrique du sol se fait au moyen de 4 électrodes : un courant électrique artificiel est injecté en surface par deux électrodes conventionnellement appelées A et B, et la différence de potentiel qui en résulte est mesurée par deux autres électrodes conventionnellement appelées M et N. La résistivité électrique mesurée est donnée par l'équation 1 :

$$\rho = R \frac{S}{L} = \frac{\Delta V}{I} * \left\{ \frac{2\pi}{\frac{1}{AM} - \frac{1}{AN} - \frac{1}{BM} + \frac{1}{BN}} \right\} = \frac{\Delta V}{I} K \quad (1)$$

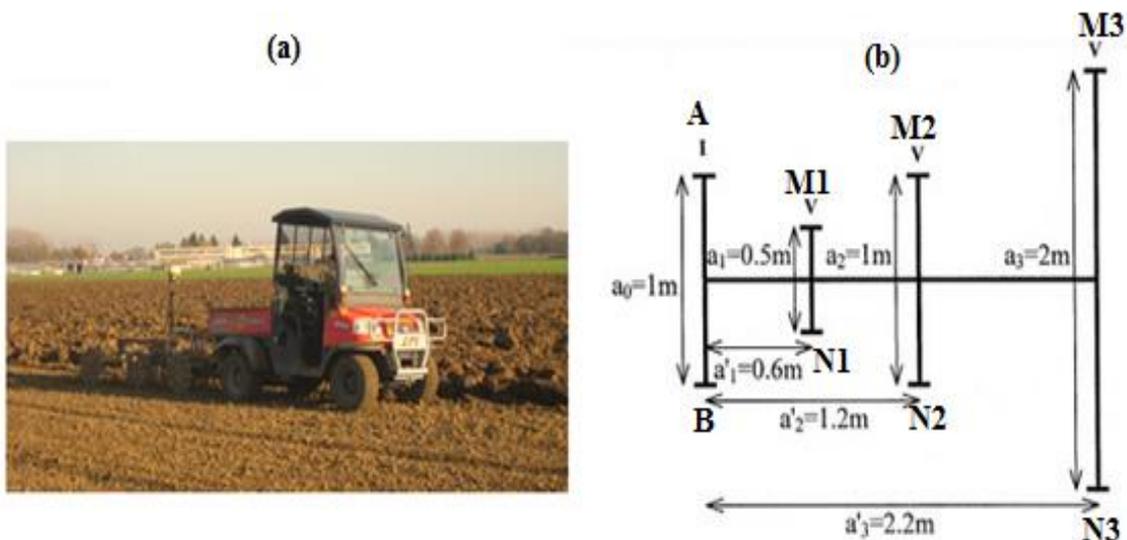
avec  $I$  : intensité du courant injecté (A),  $\Delta V$  : différence de potentiel mesurée entre M et N,  $K$  : facteur géométrique (sans unité) ; AM, AN, BM et BN : distances entre électrodes (m).

La mesure de résistivité électrique est le résultat de l'intégration des hétérogénéités du sol sur l'ensemble du volume prospecté qui dépend de l'écartement entre les électrodes. C'est pourquoi on parle de résistivité apparente. De ce fait, dans le cas d'un sol reposant sur un matériau géologique de nature contrastée, l'épaisseur du profil du sol aura de l'influence sur la mesure de résistivité électrique.

#### 3.2.2. Présentation du dispositif ARP

Le système ARP (Automatic Resistivity Profiler), dit en « Vol de Canard » du fait de la disposition de ses électrodes, est un dispositif multipôle tracté qui a été développé par le centre de recherches géophysique de Garchy (CNRS et Paris VI) (Panissod et *al.*, 1997). Il permet de mesurer la résistivité électrique du sol à l'échelle de la parcelle de manière rapide, car il est tracté par un véhicule. Il est constitué de quatre dipôles, en forme de roues dentées : un dipôle AB permettant l'injection de courant électrique, puis d'une série de trois dipôles récepteurs (M1N1, M2N2 et M3N3) qui mesurent la différence de potentiel électrique. Les écartements entre le dipôle AB et les dipôles récepteurs (M1N1, M2N2 et M3N3) sont respectivement de 0,5 m, 1 m et 2 m. Le dispositif permet donc une mesure de la résistivité apparente du sol sur trois voies (V1, V2 et V3) dont les profondeurs d'investigation sont de l'ordre de la distance d'écartement entre le dipôle d'injection et le dipôle de mesure. Les mesures de résistivité sont intégratrices : la mesure réalisée par la voie 3 intègre les volumes prospectés par la voie 2 et la voie 1. De la même façon, la mesure réalisée par la voie 2 intègre le volume prospecté par la voie 1.

Dans la suite du rapport, les résistivités mesurées par V1, V2 et V3 seront respectivement nommées R1, R2 et R3.



**Figure 2 :** Photographie de l'ARP (a) et (b) schéma du dispositif

Les électrodes sont reliées à un résistivimètre qui permet de gérer l'injection du courant et la mesure de résistivité électrique. Une mesure de résistivité est déclenchée tous les 10 cm ce qui permet un échantillonnage quasi-continu le long du trajet effectué par l'ARP. L'incertitude relative de mesure annoncée par le fabricant étant de  $\pm 2\%$  à  $20 \pm 3^\circ\text{C}$ . Les mesures sont géo-référencées au GPS différentiel (dGPS, Trimble) de précision décimétrique.

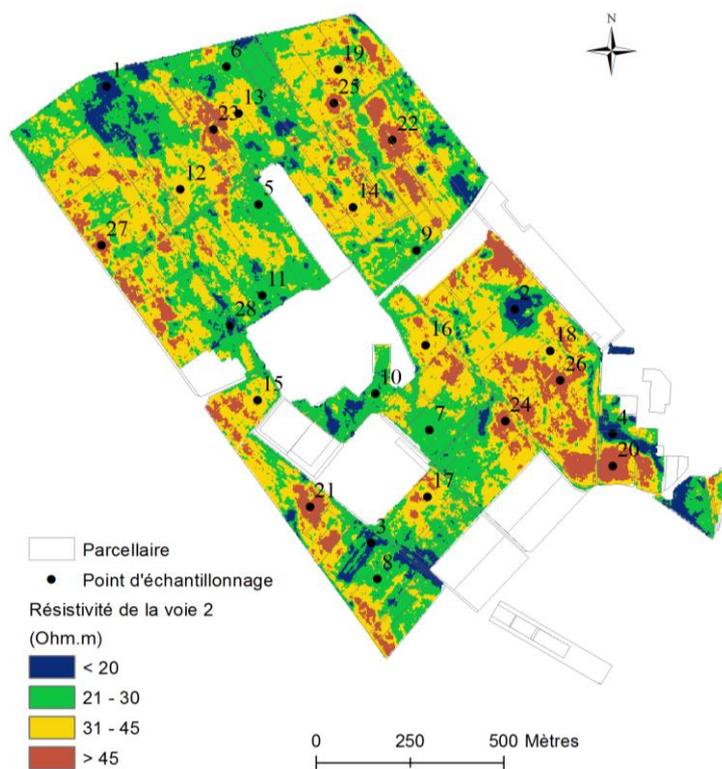
### 3.2.3. Prospection électrique réalisée dans le cadre de cette étude

Les campagnes de prospection électrique ont été réalisées sur le site expérimental en automne 2011 (du 15 au 25 novembre). Cette période se caractérise par de faibles variations de précipitation (2.5 mm), d'évapotranspiration (0.1 - 0.2 mm) et de température, ce qui limite l'impact de ces facteurs sur la mesure de résistivité et nous permet de les considérer comme négligeables. Aucun travail du sol n'a été réalisé pendant cette période de prospection, la structure du sol n'a donc pas évolué.

Les passages d'ARP ont été réalisés tous les 6 m, parfois 12 m. A l'issue de chaque prospection, et avant toute analyse, un traitement est appliqué aux données brutes de résistivité électrique de façon à supprimer les valeurs aberrantes : les valeurs correspondantes aux mesures faites dans les virages, celles égales à 0 ou supérieures à  $10^4$  ohm.m qui correspondent à des problèmes de mesure (mauvais contact entre le sol et l'électrode).

### 3.3. Mesure de la variable d'intérêt : Epaisseur de sol selon le plan d'échantillonnage établi à partir des mesures de la résistivité électrique

L'épaisseur du sol a été mesurée en 28 points de sondage. Le plan d'échantillonnage a été établi sur la base de la carte de résistivité du sol (Fig. 3). La connaissance que nous avons de l'épaisseur du sol sur le site a orienté le choix de la voie 2 pour établir le plan d'échantillonnage. Les 28 points ont été répartis de manière à échantillonner uniformément au sein des quatre classes d'iso-résistivité présenté dans la figure 3.



**Figure 3 :** Plan d'échantillonnage pour la mesure de l'épaisseur du sol à partir de la carte de résistivité électrique de la voie 2 de l'ARP

Chaque sondage a été décrit par un pédologue et la profondeur du sol (limite entre la base du profil de sol et l'apparition du cailloutis calcaire) a été mesurée. Des prélèvements de sol ont été réalisés dans chaque horizon et la composition granulométrique, la capacité d'échange cationique, le taux d'éléments grossiers, et le pH, ont été déterminés au Laboratoire d'Analyses des Sols de l'INRA à Arras.

#### 4. Approches de spatialisation mises en œuvre

##### 4.1. Démarche générale

Les différentes étapes des deux approches testées sont résumées dans le logigramme de la figure 4. A partir de la variable d'intérêt (épaisseur du sol,  $E_p$ ) et de la variable auxiliaire (résistivité électrique,  $R$ ) quatre scénarii (A, B, C et D) ont été mis en œuvre pour aboutir à 4 cartes d'épaisseurs estimées nommées de la même façon.

Ces quatre scénarii s'organisent selon deux approches :

- **Approche spatiale (encadré rouge)** : les cartes des épaisseurs de sol estimées sont réalisées par krigeage à l'aide d'une dérive externe issue des trois voies de la résistivité. Pour utiliser l'information contenue dans les trois voies de la résistivité, on réalise une analyse en composantes principales de ces trois voies. La dérive sera soit la première composante de l'ACP (PC1), soit l'une des composantes de PC1 identifiée grâce à son variogramme.

- **Approche non spatiale (encadré violet)** : les cartes des épaisseurs de sol estimées sont réalisées par des modèles de régression linéaire simple (MRLS) établis entre les épaisseurs de sol mesurées et PC1 d'une part, et les épaisseurs de sol mesurées et la voie de résistivité la mieux corrélée, d'autre part.

#### 4.2. Méthodes spatiales et non spatiales mises en œuvre

Les analyses spatiales par les techniques de géostatistiques procèdent en deux étapes : (1) l'analyse variographique qui est une description de la variabilité moyenne d'un phénomène en fonction de l'échelle spatiale, (2) le krigeage avec dérive externe (KED) qui est une méthode de régression pour estimer les valeurs d'un phénomène, épaisseur de sol, dans notre cas, en tout endroit du site expérimentale de l'UE d'Epoisses. Le KED est utilisé pour l'analyse de deux variables de natures différentes. En effet, il arrive que deux variables mesurées par des procédés différents traduisent le même phénomène et que la variable d'intérêt (épaisseur du sol) soit précise, mais connue en peu de points, alors que la variable auxiliaire (résistivité) est connue à peu près partout sur le site d'étude. La variable auxiliaire sert de guide et dessine les grandes lignes (rôle de dérive) de la variable d'intérêt (Wackernagel, 1995 et Bourennane et *al.*, 1996). On suppose en théorie que la variable d'intérêt  $Z(\mathbf{x})$  peu échantillonnée est en moyenne proportionnelle à la variable auxiliaire  $s(\mathbf{x})$  à une constante  $a_0$  près selon la relation :

$$E [Z(\mathbf{x})] = a_0 + b_1 s(\mathbf{x}) \quad (2)$$

Les analyses non spatiales reposent sur l'établissement de modèle de régression linéaire (MRLS). Le MRLS est une méthode statistique utilisée pour analyser la relation entre une variable à expliquer (ici  $E_p$ , connue en 28 points), et une variable explicative (ici la voie de résistivité). Celui-ci est défini par une équation de la forme :

$$y = \beta_1 + \beta_2 x + \varepsilon \quad (3)$$

avec  $y$  : variable à expliquer,  $x$  : variable explicative,  $\varepsilon$  : résidus du modèle,  $\beta_1$  et  $\beta_2$  : coefficients de régression.

L'ensemble des traitements sont réalisés grâce au logiciel « Isatis » développé et commercialisé par l'entreprise Géovariances.

#### 4.3. Etape préliminaire : Organisation des données de résistivité sur une grille régulière

Nous rappelons que la densité des mesures de résistivité n'est pas homogène sur l'ensemble des parcelles. Les lignes de prospections sont espacées de 6 ou 12 m et le prétraitement des données brutes qui consiste à enlever les valeurs aberrantes (partie 3.2.3) occasionne une perte de l'information. De plus, les scénarii retenus pour l'estimation de l'épaisseur du sol sur la totalité du domaine d'étude nécessitent que la variable auxiliaire soit connue aux nœuds d'une grille régulière couvrant le domaine d'étude. Le jeu de données de résistivité a donc été organisé sur une grille régulière de  $1\text{m}^2$  de résolution, commune, aux enregistrements des trois voies de mesures de résistivité ( $V_1$ ,  $V_2$  et  $V_3$ ). A chaque nœud de la grille on affecte les 3 valeurs de résistivité ( $R_1^*$ ,  $R_2^*$  et  $R_3^*$ ) du voisin le plus proche.

#### 4.4. Estimation de l'épaisseur du sol par l'approche spatiale (Scenarii A et B)

Cette approche se base sur le krigeage avec la première composante principale de l'analyse en composantes principales (ACP) des trois voies de résistivité. Il convient donc, dans un premier temps, de réaliser une ACP et dans un second temps d'analyser la structuration spatiale de la première composante de cette ACP.

##### 4.4.1. Analyse en composante principale (ACP)

En vue de l'analyse en composante principale, les résistivités (R1\*, R2\* et R3\*) organisées sur la grille de 1 m<sup>2</sup> de résolution, ont été transformées dans l'espace Gaussien en de nouvelles variables G\_R1\*, G\_R2\* et G\_R3\* sur lesquelles seront réalisé l'ACP. L'ACP est une méthode de la statistique multivariée, qui consiste à transformer des variables corrélées entre elles en nouvelles variables décorrélées les unes des autres. Les nouvelles variables sont nommées composantes principales (PC). L'ACP étant réalisé sur trois variables de même nature et intrinsèquement corrélées, la première composante principale (PC1) expliquera naturellement la majorité de la variance contenue dans les variables initiales. Aussi, on focalisera dans la suite de l'analyse sur PC1.

##### 4.4.2. Etude de la structuration spatiale de PC1

Cette analyse nécessite de générer un variogramme expérimental qui sera à la suite ajusté par un ou plusieurs modèles théoriques (variogramme théorique). L'ajustement se fait par l'emploi d'un certain nombre de modèles qui sont essentiellement de deux types : des modèles croissants non bornés (modèle linéaire, fonction puissance) et des modèles croissants bornés (modèle sphérique et le modèle exponentiel).

Le variogramme expérimental ( $\gamma^*$ ) est la moyenne des dissemblances. La dissemblance est la moitié du carré de la différence entre deux valeurs  $z_1$  et  $z_2$ , situées en deux points  $x_1$  et  $x_2$ , et pour  $n_h$  paires de points reliés par un vecteur, elle est donnée par l'expression (4).

$$\gamma^*(\mathbf{h}) = \frac{1}{2n_h} \sum_{\alpha=1}^{n_h} [z(\mathbf{x} + \mathbf{h}) - z(\mathbf{x})]^2 \quad (4)$$

Généralement, la dissemblance des valeurs augmente en moyenne en fonction de l'éloignement spatial des points de mesure et atteint fréquemment un palier de variation aux grandes distances. Le comportement aux faibles distances, près de l'origine du variogramme, est une donnée importante, car il indique le degré de continuité de la variable régionalisée, à savoir : différentiable, continue mais non différentiable, ou carrément discontinue. Dans ce dernier cas, on est en présence d'une variable régionalisée donnant lieu à un effet de pépite. Lorsque la dissemblance moyenne des valeurs est constante pour toutes les distances, il y a une absence complète de structuration spatiale des valeurs. Le variogramme permet ainsi de décomposer éventuellement PC1 en plusieurs composantes : S0 (bruit blanc : valeur de la variance pour une distance nulle), S1 (composante des hautes fréquences : structure à courte distance) et S2 (composante des basses fréquences : structure à longue distance).

#### 4.4.3. Construction de scénarii basés sur l'approche spatiale

Deux traitements ont été conduits pour cette analyse spatiale :

- Un KED en utilisant PC1 comme dérive externe (Scénario A).
- Un KED avec S2-PC1 comme dérive externe (Scénario B). On focalisera sur S2 dans ce travail parce

qu'elle représente les grandes tendances contrairement à S1 qui indique les anomalies locales.

A l'issue de ces traitements, les valeurs générées par les analyses spatiales subissent une transformation inverse pour revenir à l'unité de départ (cm).

#### 4.5. Estimation de l'épaisseur du sol par l'approche non spatiale (Scénarii C et D)

L'estimation de l'épaisseur du sol par l'approche non spatiale a été faite en utilisant d'une part PC1, et d'autre part, la voie de résistivité la mieux corrélée aux épaisseurs de sol mesurées.

##### 4.5.1. Estimation de l'épaisseur du sol à partir de PC1 (Scénario C)

Ce scénario consiste à étudier la corrélation entre les valeurs d'épaisseur du sol transformée dans l'espace Gaussien ( $G_{Ep}$ ) et PC1. La transformation dans l'espace Gaussien de l'épaisseur mesurée est due au fait que PC1, a été définie à partir des variables Gaussiennes. Si la corrélation existe, on établit un modèle de régression linéaire simple (MRLS) entre  $G_{Ep}$  et PC1 que l'on applique à la grille de PC1 afin de construire une carte (C). Les valeurs générées par cette analyse subissent également une transformation inverse (partie 4.4.3).

##### 4.5.2. Estimation de l'épaisseur du sol par choix d'une voie de résistivité (Scénario D)

L'estimation débute par la transformation de  $R1^*$ ,  $R2^*$  et  $R3^*$  en valeurs logarithmiques  $\ln(R1^*)$ ,  $\ln(R2^*)$  et  $\ln(R3^*)$ . Cette transformation est importante à faire car souvent elle améliore les corrélations entre  $E_p$  et la résistivité (partie 5.4.1). Ensuite on étudie la corrélation entre l'épaisseur du sol et la résistivité électrique ( $R1^*$ ,  $R2^*$  et  $R3^*$  et leurs logarithmes), en vue du choix de la voie la plus corrélée. La voie la plus corrélée à l'épaisseur du sol servira à établir un modèle de régression linéaire simple (MRLS) entre l'épaisseur du sol et la voie de résistivité choisie. Le modèle de régression sera appliqué à la grille de la voie de résistivité retenue : à partir de ces valeurs, une carte (D) d'épaisseur de sol peut être construite.

#### 4.6. Validation croisée

La validation croisée est l'étape finale commune à l'ensemble des scénarii d'estimation de l'épaisseur du sol. Elle permet de discuter de la qualité des estimations réalisées et de comparer les résultats entre les deux approches. Dans la procédure de validation croisée, une valeur  $z(x_\alpha)$  d'épaisseur mesurée à un point donnée de coordonnées  $(x ; y)$  est retirée du jeu de données et sa valeur  $z^*(x_\alpha)$  et ré-estimée au dit point  $(x ; y)$  à partir des  $(n-1)$  autres valeurs. La procédure est réalisée pour tous les points d'échantillonnage, soit 28 fois dans notre cas.

Pour chaque scénario d'estimation seront définis les indicateurs statistiques suivant :

✓ **l'erreur moyenne (ME)** qui doit être proche de zéro (une moyenne négative (ou positive) représente une sous- (sur-) estimation systématique),

$$ME = \frac{1}{n} \sum_{k=1}^n (\text{Ep. estimée} - \text{Ep. mesurée}) \quad (5)$$

✓ **la racine carrée de l'erreur moyenne quadratique (RMSE)** qui doit avoir la plus petite valeur possible pour une estimation précise et sans biais.

$$RMSE = \sqrt{\frac{1}{n} \sum_{k=1}^n (\text{Ep. estimée} - \text{Ep. mesurée})^2} \quad (6)$$

✓ **l'erreur moyenne absolue (MAE)** indiquant l'ampleur moyenne des erreurs. La MAE est toujours positive, inférieure ou égale à la RMSE.

$$MAE = \frac{1}{n} \sum_{k=1}^n |\text{Ep. estimée} - \text{Ep. mesurée}| \quad (7)$$

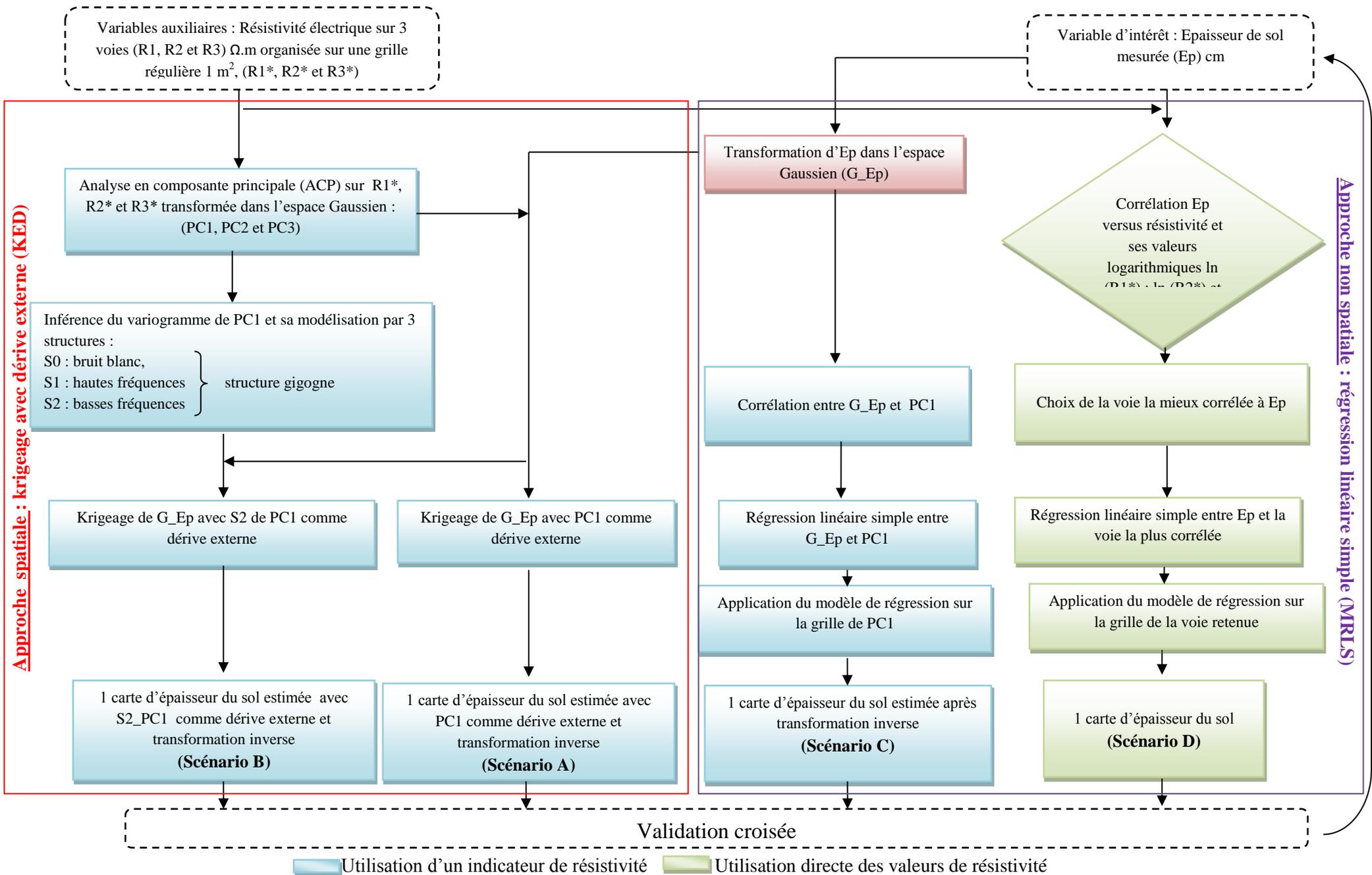


Figure 4 : Logigramme décrivant les différentes étapes des approches testées

## 5. Résultats et discussion

### 5.1. Caractérisation pédologique des sols de l'UE d'Epoisses

La description visuelle à partir des sondages réalisés sur les sols de l'UE d'Epoisses montre qu'ils sont constitués d'un horizon de surface labouré qui repose soit directement sur le cailloutis calcaire pour les sols de faibles épaisseurs, soit, pour les sols plus épais sur un horizon structural. Sur les 28 points d'échantillonnage, réalisés au carottier sur le site d'étude (Fig. 3), l'épaisseur du sol varie entre 23 et 192 cm. Les sols ont été classés en deux groupes sur la base de leur épaisseur (Tableau 1).

✓ d'une part, des sols moins profonds argilo-limoneux superficiels dont l'épaisseur varie de 23 à 50 cm, particulièrement caractérisés par de forte quantité de cailloux de l'ordre de 151 g/kg en moyenne.

✓ d'autre part, des sols profonds argilo-limoneux en surface et argileux en profondeur dont l'épaisseur varie de 51 à 192 cm. Ces sols contiennent de faible quantité de cailloux à l'exception du sol décrit au sondage 23 dont la quantité en cailloux varie de 95,6 à 206 g/kg, et présente une texture argilo-limoneuse sur limons moyens sableux. Il s'agirait d'une observation réalisée dans un paléochénel. Les analyses de sol présentées dans le tableau 1 ne tiennent pas compte de ce sol particulier. On peut constater que les paramètres mesurés varient pour les deux catégories de sol très faiblement, sinon dans les mêmes proportions. Ce qui permet de confirmer que sur les sols de l'UE d'Epoisses, le principal facteur de variabilité des sols est leur épaisseur.

**Tableau 1 :** Résultats d'analyses pédologiques réalisées à partir des échantillons de sol sur l'UE d'Epoisses

Variables	Sols peu profonds (Ep < 50cm)								
	Epaisseur (cm)	CEC cmol/kg	Calcaire totale g/kg	pH	Sables (%)	Limons (%)	Argile (%)	Ca <sup>2+</sup> cmol/kg	Eléments grossiers (%)
Minimum	23	17,4	1,7	5,45	5,5	35,3	34,5	13,5	1
Maximum	50	26,8	119	8,15	18,1	58,7	50,5	57,8	30
Sols profonds (Ep > 50 cm)									
Minimum	51	15,5	1,1	6,53	3,9	29,9	31,9	16,5	1
Maximum	192	26,9	256	8,36	21,8	59,6	54,4	51,4	10

Les rapports de la somme des cations échangeables sur la capacité d'échange cationique (CEC) largement supérieurs à 100 % indiquent que le complexe absorbant est saturé en calcium. Selon le référentiel pédologique (Baize et Girard., 2009) ces sols sont des Calcisols : on parlera de Calcisols leptiques pour les sols peu profonds (appelés rendzines brunifiés dans la carte pédologique) et de Calcisols pachiques pour les sols profonds (appelés sols bruns calciques dans la carte pédologique).

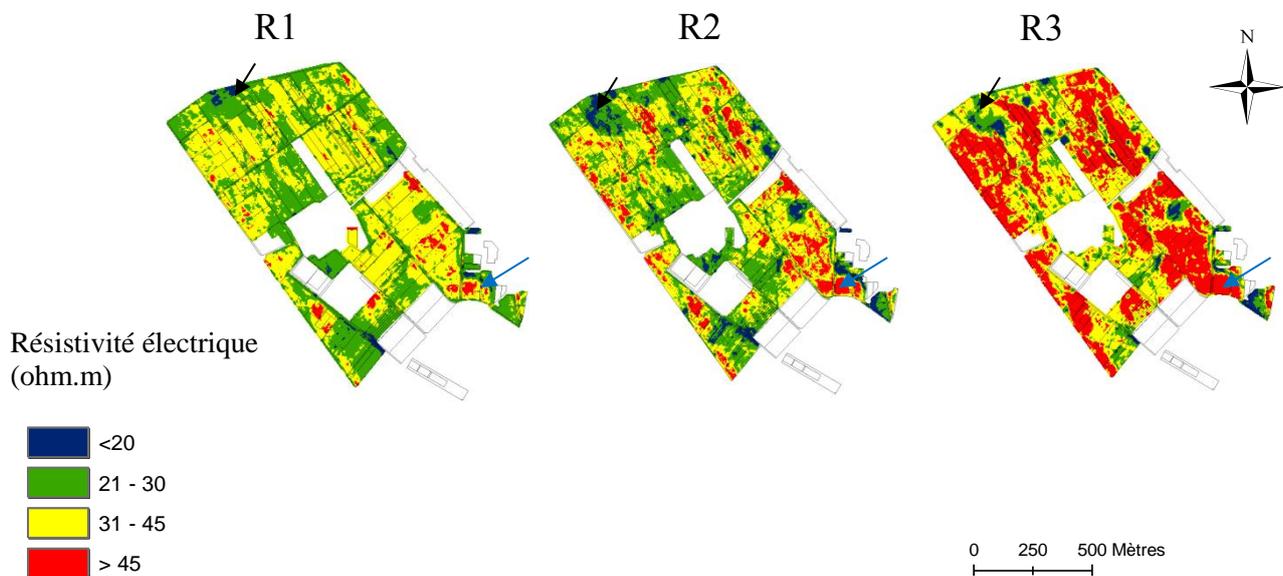
### 5.2. Données de résistivité électrique

L'organisation spatiale des résistivités électriques mesurées (avant organisation sur la grille) par les trois voies de l'ARP (R1, R2 et R3) est présentée dans la figure 5 et les statistiques descriptives sur ces résistivités sont décrites dans la première partie du tableau 2. Les résistivités sont en moyenne plus élevées quand la profondeur d'investigation augmente : les résistivités de la voie 3 (en moyenne égales à 43 ohm.m) sont plus élevées que celles de la voie 2 (en moyenne égales à 33 ohm.m) et que celles de la voie 1 (en

moyenne égales à 32 ohm.m). Cela est due au fait que, plus on prospecte en profondeur et plus le volume prospecté intègre le cailloutis calcaire, plus résistant.

Par ailleurs, on observe une organisation spatiale comparable entre les 3 voies de résistivité. Par exemple, on observe une zone de faible résistivité au nord-ouest du site d'étude (voir flèche noire sur la figure 5) que l'on retrouve sur les trois voies et une zone plus résistante au sud-est (voir flèche bleue). Ceci est aussi dû au caractère « emboîté » des mesures des trois voies.

La carte de résistivité de la voie 1 présente des limites de zone d'isorésistivité qui suivent le contour de certaines parcelles. La voie 1 prospectant à de faibles profondeurs (environ 0 à 50 cm) est impactée par le travail du sol et l'effet des différentes cultures sur le sol. Cet effet parcellaire s'estompe sur la carte de la voie 2 et n'est plus visible sur la carte de la voie 3.



**Figure 5 :** Cartes des résistivités mesurées par l'ARP sur les trois voies de mesure : R1 pour les résistivités de la voie 1, R2 pour les résistivités de la voie 2 et R3 pour les résistivités de la voie 3

Le tableau 2 résume les statistiques descriptives des valeurs de la résistivité des trois voies de l'ARP avant et après organisation sur la grille régulière de 1 m<sup>2</sup> de résolution.

**Tableau 2 :** Statistiques descriptives des trois voies de résistivités

Statistiques descriptives avant organisation sur grille 1 m <sup>2</sup>							
Variables (ohm.m)	Effectif	Minimum	Maximum	Moyenne	g <sub>1</sub>	CV (%)	Ecart-type
R1	1606559	9	625	32	31,5	36	11,58
R2	1606559	2	214	33	1,1	32	10,41
R3	1606559	5	1026	43	14,5	41	17,64
Statistiques descriptives après organisation sur grille 1 m <sup>2</sup>							
R1*	2642360	10	625	30	12,6	28	8,27
R2*	2642360	2	180	30	1,1	33	9,95
R3*	2642360	5	1026	40	5,9	39	15,69

g<sub>1</sub> : coefficient d'asymétrie, CV (%) : coefficient de variation

On remarque que l'organisation de la résistivité sur la grille entraîne une augmentation des effectifs de mesures de résistivité mais ceci n'affecte pas de façon significative les paramètres de position (minimum, maximum et moyenne) et les paramètres de dispersion (g1, CV et écart-type). Que ce soit avant ou après organisation sur la grille, les coefficients d'asymétrie (g1), supérieurs à zéro pour chaque voie, indiquent une dissymétrie vers les plus petites valeurs.

### 5.3. Estimation de l'épaisseur du sol par l'approche spatiale (Scenarii A et B)

#### 5.3.1. Analyse en composante principale des voies de résistivité

Les coefficients de corrélation entre les valeurs de résistivité des trois profondeurs d'investigation sont élevés, supérieures à 0.5 (Tableau 3). Il s'agit là d'un résultat attendu dans le sens où l'on sait que les valeurs de résistivité de la partie inférieure d'un volume, les horizons profonds du sol dans notre étude, intègrent les valeurs enregistrées dans la partie supérieure de ce même volume, les horizons superficiels.

**Tableau 3 :** Matrice de corrélation sur les données organisées sur la grille 1 m<sup>2</sup>

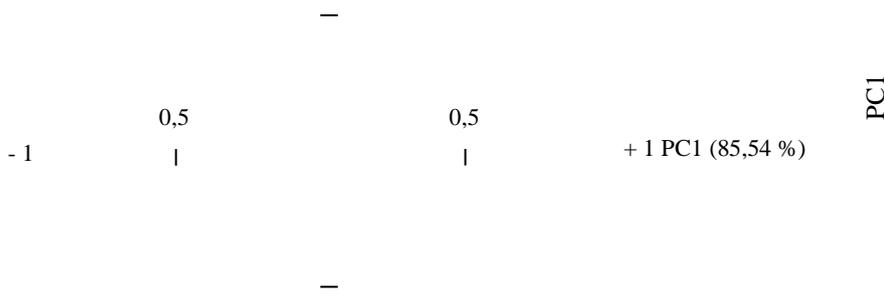
Variables	R1*	R2*	R3*
R1*	1,00		
R2*	0,72	1,00	
R3*	0,58	0,85	1,00

En effet, les coefficients de co-dispersion (résultats non présentés dans ce document) qui décrivent la corrélation entre les valeurs de résistivité mesurées à trois profondeurs différentes en fonction de l'échelle spatiale, sont constants. Ces coefficients indiquent que la corrélation entre les accroissements spatiaux des valeurs de résistivité tend vers la valeur du coefficient de corrélation indiquant ainsi que la variabilité de la résistivité avec la profondeur est indépendante de la distance. Pour toutes ces raisons une analyse en composante principale (ACP) a été conduite sur R1\*, R2\* et R3\* transformées dans l'espace Gaussien (G-R1\*, G-R2\* et G-R3\*) afin de résumer la variance contenue dans ces trois variables corrélées à travers une seule composante principale (PC1) qui explique la part la plus importante de la variance de R1\*, R2\* et R3\*.

Les deux premières composantes issues de l'ACP représentent plus de 97 % de la variance totale des trois variables considérées. Ces dernières sont positivement corrélées à PC1 (Fig. 6). De plus, il existe une forte corrélation négative entre PC1 et l'épaisseur de sol (Fig. 7). Aussi PC1, première composante de l'ACP, peut être utilisée comme variable explicative pour modéliser les variations de l'épaisseur du sol dans notre étude.

PC2 (12,29 %)

$r = -0.81$



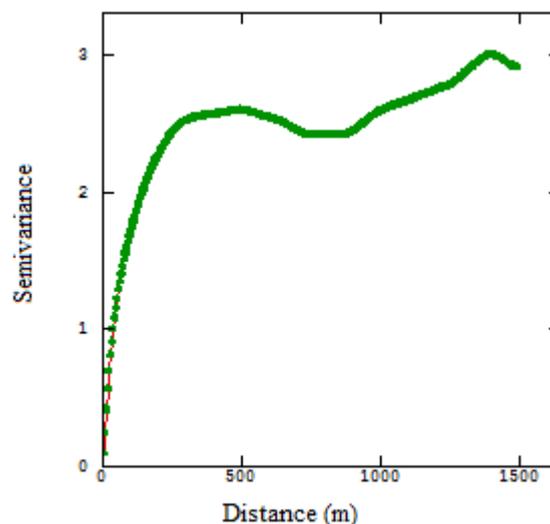
**Figure 6 :** Représentation des variables dans le plan factoriel PC1-PC2

Valeurs Gaussiennes d'épaisseur de sol

**Figure 7 :** Variation des valeurs de PC1 en fonction des valeurs Gaussiennes de l'épaisseur du sol

### 5.3.2. Estimation de l'épaisseur du sol par krigeage en considérant PC1 et S2-PC1 comme dérive externe

L'estimation spatiale de l'épaisseur du sol par krigeage en considérant d'une part PC1 et d'autre part S2 de PC1 (des basses fréquences de PC1) comme dérive externe a nécessité l'estimation et l'ajustement du variogramme expérimental de PC1. Celui-ci a donc été estimé et ajusté (Fig. 8). Le variogramme expérimental est estimé sur une distance de 1,5 km avec un pas de 5 m. L'ajustement du variogramme expérimental a été réalisé à l'aide de trois structures : un effet de pépité (S0) très faible puisque sa valeur est égale à 0,05, et deux modèles exponentiels, de portées respectives 30 et 257 m. Ce variogramme décompose donc PC1 en S0 (effet de pépité : valeur correspondante au bruit blanc), S1 (composante des hautes fréquences : structure à courte distance) et S2 (composante des basses fréquences : structure à longue distance). Ainsi PC1 et S2-PC1 sont respectivement intégrés dans le système d'équations de krigeage afin d'estimer l'épaisseur du sol sur tout le domaine d'étude.



**Figure 8 :** Variogramme expérimental (points verts) et modèles théoriques ajustés (courbe rouge) de PC1

#### 5.4. Estimation de l'épaisseur du sol par l'approche non spatiale (Scenarii C et D)

##### 5.4.1. Utilisation de la voie de résistivité la plus corrélée à l'épaisseur du sol

Afin d'estimer l'épaisseur du sol par régression linéaire simple, nous avons analysé les corrélations entre i) les épaisseurs du sol mesurées à partir des 28 points d'échantillonnage et les résistivités des trois voies de prospection et, ii) les épaisseurs du sol et les valeurs logarithmiques des résistivités. Les coefficients de corrélation indiquent que les relations sont améliorées lorsqu'on considère les valeurs logarithmiques de la résistivité (Tableau 4). Par ailleurs ces coefficients indiquent que la résistivité est inversement proportionnelle à l'épaisseur du sol et que les valeurs logarithmiques de la résistivité de la voie 3 sont la variable explicative la plus pertinente pour modéliser les variations d'épaisseur du sol dans notre cas.

**Tableau 4 :** Corrélation entre la variable d'intérêt (épaisseur du sol) et les variables auxiliaires (résistivité)

Variabes	Coefficient de corrélation
Ep	1
R1*	-0,58
R2*	-0,68
R3*	-0,71
ln (R1*)	-0,60
ln (R2*)	-0,76
ln (R 3*)	-0,84

Ainsi, le modèle indiqué dans l'équation (8) a été établi par régression linéaire simple entre l'épaisseur du sol et les valeurs logarithmiques de la résistivité de la voie 3.

$$Ep = 333.6 - (70.7) X \ln(R3^*) \quad (8)$$

Ce modèle explique 70 % de la variabilité de l'épaisseur de sol ( $R^2_{\text{ajusté}} = 0.70$ , p-value < 0.0001).

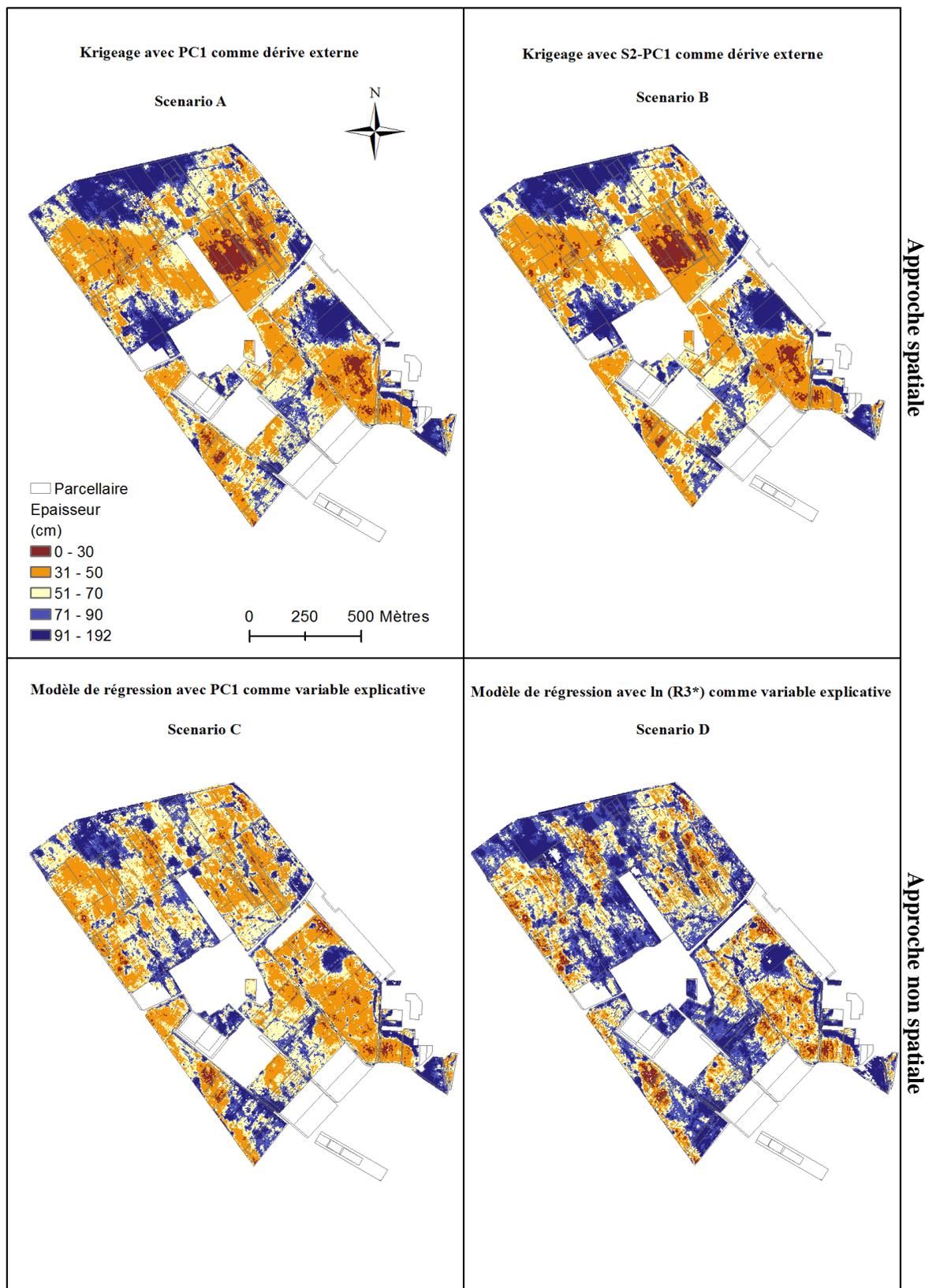
##### 5.4.2. Utilisation de la composante PC1 de la résistivité

L'utilisation de PC1 comme base de l'estimation de l'épaisseur du sol a nécessité l'étude de sa corrélation avec l'épaisseur de sol mesurée. Le modèle proposé dans l'équation (9) a été établi par régression linéaire simple entre les valeurs Gaussiennes de l'épaisseur du sol ( $G_{Ep}$ ) et les valeurs de PC1. Ce modèle explique 65 % ( $R^2_{\text{ajusté}} = 0.65$ , p-value < 0.0001) de variabilité de l'épaisseur du sol.

$$G_{Ep} = 0.23 - (0.24) X PC1 \quad (9)$$

#### 5.5. Cartes de l'épaisseur du sol obtenus par les approches spatiales et non spatiales

La figure 9 regroupe l'ensemble des cartes d'épaisseur estimée par les différentes approches présentées précédemment.



**Figure 9 :** Cartes d'épaisseur de sol estimée selon les 4 scenarii

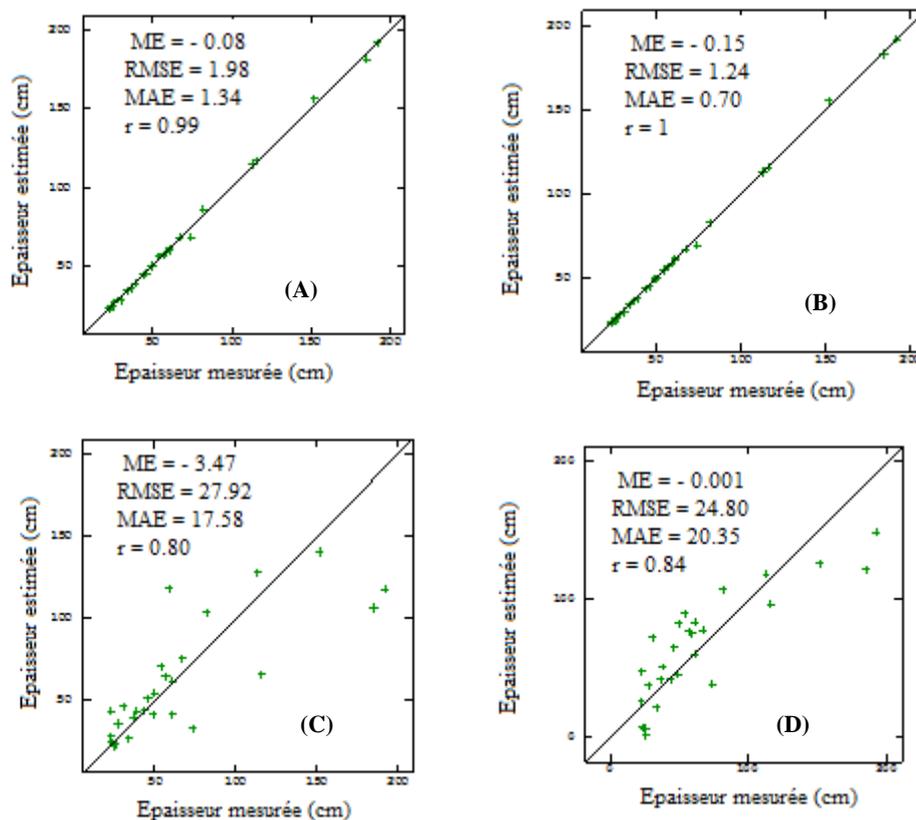
La carte issue du scénario D propose des valeurs d'épaisseur plus élevées que les trois autres cartes, en particulier les cartes issues des scenarii (A) et (B). Ces dernières présentent également un aspect plus lisse, et des zones d'épaisseurs homogènes plus larges. Par exemple, on observe au nord une zone de sols profonds (91-192 cm) et un peu plus bas une zone de sols variant de 0 à 30 cm.

Le tableau 5 indique les proportions des classes d'épaisseur prédites sur chaque carte par les différents scénarii en comparaison à leur équivalent dans le jeu de donnée initial (épaisseurs de sol mesurées). Les méthodes spatiales, en particulier le krigeage avec S2-PC1 comme dérive externe, génèrent des proportions de classes d'épaisseur proches de celles observées dans l'échantillon de la variable d'intérêt (épaisseurs mesurées). En revanche les approches non spatiales semblent surestimer les sols profonds.

**Tableau 5 :** Comparaison des classes d'épaisseurs prédites versus les classes d'épaisseurs mesurées

Classes d'épaisseur (cm)	Méthodes non spatiales		Méthodes spatiales		Epaisseurs mesurées
	MRLS ln (R3*)	MRLS PC1	KED PC1	KED S2_PC1	
0 - 50	20,53	37,57	42,24	43,07	53,57 %
51 - 192	79,47	62,43	57,76	56,93	46,43 %

La figure 10 présente la comparaison entre les épaisseurs estimées par les différents scénarii et les épaisseurs mesurées. Les modèles de régression linéaires simples (MRLS), des scénarii C et D, présentent des MAE (respectivement 17,58 et 20,35 cm) et des RMSE (respectivement 27,92 et 24,80 cm) élevées (Fig 10). Par ailleurs ces modèles ne parviennent pas à reproduire les valeurs extrêmes du jeu de donnée initial (1 à 148 cm pour le scénario D et 23 à 141 cm pour le scénario C).

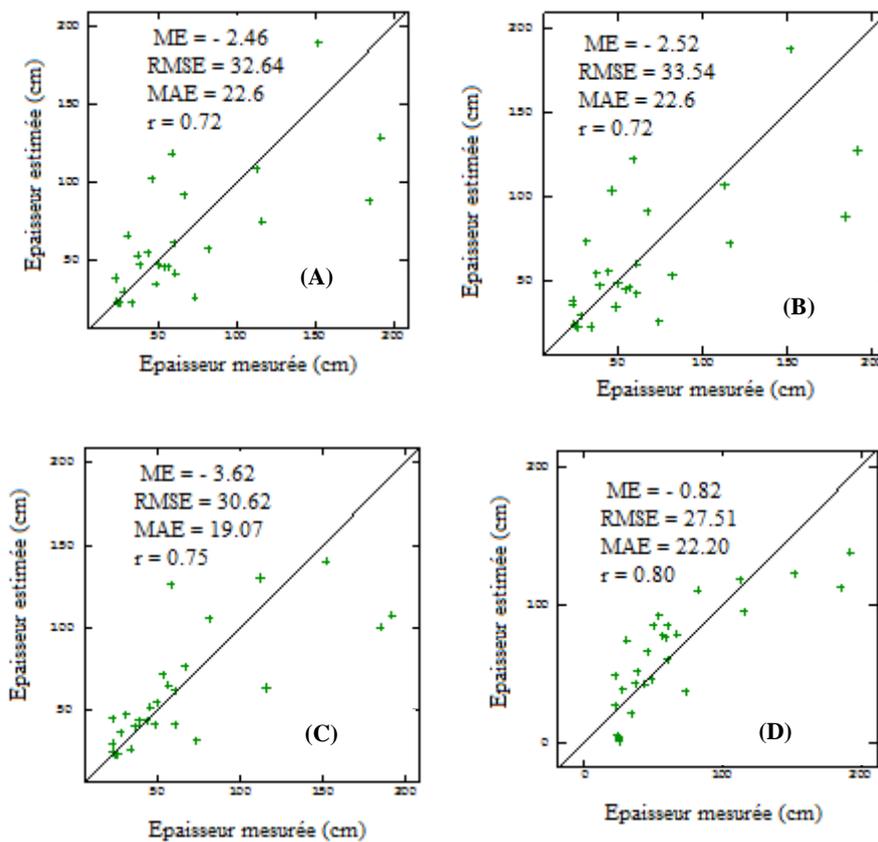


**Figure 10 :** Comparaison entre les épaisseurs de sol estimées au droit des épaisseurs de sol mesurées pour les 4 scénarii

Ceci est particulièrement vrai pour le modèle de régression linéaire où on utilise  $\ln(R3^*)$  comme variable prédictive (scénario D). En outre, ce scénario produit des valeurs d'épaisseur inférieures à celles mesurées sur le terrain. Les modèles à base de RLS nécessitent, lors de la spatialisation, que la variable explicative soit bornée à l'étendue de ses valeurs dans le jeu de données qui a servi à établir le modèle. Au contraire, les modèles spatiaux utilisant le krigeage avec dérive externe présentent des indicateurs statistiques plus faibles et prédisent assez fidèlement les valeurs mesurées en chaque point d'échantillonnage, reproduisant ainsi toute la gamme de variation de l'épaisseur du sol mesurée dans le jeu de donnée initial. Ce résultat est justifié par le fait que le KED étant une méthode d'estimation conditionnelle, l'estimation tient compte des points de mesures initiaux.

### 5.6. Validation croisée

La figure 11 présente les résultats de la validation croisée. Les indicateurs statistiques (ME, RMSE et MAE) sont, dans l'absolu, en faveur de la régression linéaire simple avec  $\ln(R3^*)$  comme variable explicative (scénario D). Cela étant, les valeurs de ces indicateurs sont quasi-équivalentes d'un scénario de cartographie à un autre. La validation croisée confirme aussi que le krigeage avec une dérive externe permet de reproduire toute la gamme de variation de l'épaisseur du sol mesurée dans le jeu de donnée initial. Les épaisseurs estimées par krigeage avec dérive externe varient de 23 à 189 cm contre 0,73 à 149 cm pour les modèles basés sur la régression linéaire simple.



**Figure 11 :** Résultats de la validation croisée : indicateurs statistiques et diagrammes de dispersion entre les épaisseurs estimées et les épaisseurs mesurées

Les résultats de la comparaison des épaisseurs estimées à celles mesurées sur le terrain ainsi que les résultats de la validation croisée (Fig. 10 et Fig. 11) suggèrent de retenir les méthodes spatiales et particulièrement la méthode de krigeage avec S2 de PC1 comme dérive externe pour la spatialisation de l'épaisseur du sol sur l'ensemble du site d'étude. La démarche de validation, dans cette étude, se base sur le même jeu de donnée que celui qui a servi à établir les modèles de prédiction spatiale de l'épaisseur du sol. Pour améliorer la validation il serait opportun de la conduire sur un jeu de données indépendant. Dans notre étude, la taille faible de l'échantillon de la variable d'intérêt ne permet pas la sélection d'un sous ensemble pour réaliser ce type de validation, qualifiée alors de validation externe.

Par ailleurs, la forte corrélation entre l'épaisseur de sol et la résistivité électrique est probablement due en partie au fait que le plan d'échantillonnage a été orienté sur la base des cartes de la résistivité. Il serait donc judicieux d'envisager d'établir un plan d'échantillonnage indépendamment de la résistivité et de reproduire notre démarche, mais cette étude serait alors purement théorique, sans intérêt pratique pour la réalisation d'une carte des sols.

## 6. Conclusion

Deux approches d'estimation de l'épaisseur de sol (variable d'intérêt) via la résistivité électrique (variable auxiliaire) ont été mises en œuvre dans cette étude. Nous rappelons que la prospection géophysique dans le cadre de ce projet intervient à deux niveaux. Dans un premier temps, elle permet d'optimiser le plan d'échantillonnage de la variable d'intérêt et, dans un second temps, elle joue le rôle de variable guide lors de l'interpolation de la variable d'intérêt connue de façon éparse sur le site d'étude.

L'approche non spatiale repose sur un modèle de régression linéaire simple et dans laquelle la voie de résistivité la mieux corrélée à l'épaisseur de sol ou la composante principale (PC1) résumant l'information contenue dans les voies de mesures de résistivité V1, V2 et V3 de l'ARP, sont respectivement considérées comme variables explicatives dans la modélisation et la spatialisation de la variable d'intérêt et l'approche spatiale dans laquelle PC1 et ses basses fréquences (S2-PC1) sont respectivement utilisées comme dérive externe dans le krigeage de la variable d'intérêt.

Les résultats obtenus sur le site expérimental d'Epoisses situé au Sud-Est de Dijon révèlent que les indicateurs statistiques utilisés dans la procédure de validation croisée sont quasi-équivalents. Cela étant, seule l'approche géostatistique permet de reproduire tout le spectre de variation de l'épaisseur du sol mesurée. De plus, par l'approche spatiale, en particulier dans le scénario utilisant S2-PC1 comme dérive externe lors du krigeage, on obtient un rendu cartographique qui en termes d'aspect et de taille des plages cartographiques apparaît plus facilement exploitable dans la perspective d'une utilisation agronomique. La principale limite de l'approche non spatiale développée dans ce travail est que le modèle de régression sous-jacent permettant la cartographie sur tout le domaine d'étude n'est applicable que dans son domaine de définition. Ceci peut rendre l'estimation de la variable d'intérêt impossible sur certaines zones du domaine d'étude. Enfin, l'analyse de la représentativité des classes d'épaisseurs du sol des quatre cartes obtenues montre que les cartes issues des scénarii de l'approche spatiale sont mieux représentatives des classes d'épaisseurs du sol mesurées.

Sur la base de l'ensemble de ces résultats, nous préconisons pour caractériser l'épaisseur du sol sur un domaine de quelques dizaines d'hectares, l'approche spatiale dans laquelle l'information auxiliaire est éventuellement décomposée avant d'être incorporée dans le système de krigeage avec dérive externe sous réserve que l'information auxiliaire soit disponible à tous les nœuds de la grille d'estimation. Cependant, nous recommandons de tester la généralité de l'approche sur une autre Unité Expérimentale et idéalement de valider les résultats actuels à partir d'un jeu de données indépendant.

## **Références bibliographiques**

- BAIZE, D., GIRARD, M.-C. 2009. *Référentiel pédologique 2008*, Editions Quae.
- BAIZE, D. & JABIOL, B. 2011. *Guide pour la description des sols*, Editions Quae.
- BESSON, A., COUSIN, I., SAMOUELIAN, A., BOIZARD, H. & RICHARD, G. 2004. Structural heterogeneity of the soil tilled layer as characterized by 2D electrical resistivity surveying. *Soil and Tillage Research*, 79, 239-249.
- BOTTRAUD, J. C., BORNAND, M. & SERVAT, E. 1984. Mesures de resistivite appliquees a la cartographie en pedologie. *Bulletin de l'Association francaise pour l'etude du sol*, 279-294.
- 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*. Université d'Orléans.
- BOURENNANE, H. & KING, D. 2003. Using multiple external drifts to estimate a soil variable. *Geoderma*, 114, 1-18.
- BOURENNANE, H., KING, D. & COUTURIER, A. 2000. Comparison of kriging with external drift and simple linear regression for predicting soil horizon thickness with different sample densities. *Geoderma*, 97, 255-271.
- BOURENNANE, H., KING, P., CHÉRY, P. & BRUAND, A. 1996. Improving the kriging of a soil variable using slope gradient as external drift. *European Journal of Soil Science*, 473-483.
- BOURENNANE, H., NICOUILLAUD, B., COUTURIER, A., PASQUIER, C., MARY, B. & KING, D. 2012. Geostatistical filtering for improved soil water content estimation from electrical resistivity data. *Geoderma*, 183, 32-40.
- CAMPBLELL, R. B., BOWER, C. A., RICHARDS, L. A. 1948. Change of electrical conductivity with temperature and the relation of osmotic pressure and ion concentration for Soil extracts. *Soil Science Society of America Proceedings*, 13, 66-69.
- CHARTIN, C., BOURENNANE, H., SALVADOR-BLANES, S., HINSCHBERGER, F. & MACAIRE, J.-J. 2011. Classification and mapping of anthropogenic landforms on cultivated hillslopes using DEMs and soil thickness data — Example from the SW Parisian Basin, France. *Geomorphology*, 135, 8-20.
- CHERY, P., DABAS, M., BRUAND, A. & ARROUAYS, D. 1994. Cokriging electrical resistivity for detailed study of soil thickness variability. A case study in Beauce area (France). *Annales Geophysicae*, 12, C462.
- CHRETIEN, J. 1976. Carte pédologique de la France à 1/100.000. Feuille de DIJON. INRA. *Editions, carte et notice*. 218 p.
- FRIEDMAN, S. P. 2005. Soil properties influencing apparent electrical conductivity: a review. *Computers and Electronics in Agriculture*, 46, 45-70.
- JAMAGNE, M., KING, D., DAROUSSIN, J., LE BAS, C., 1993. Evolution et état actuel des programmes européens de connaissance et gestion des sols. *Extrait du bulletin des recherches agronomiques de gembloux*. Vol 28, n° 2-3, 163 pp.
- KING, D. 1984. Analyse de quelques concepts en cartographie des sols basée sur une automatisation des cartes thématiques dérivées. *Agronomie*, 4, 461-472.
- KING, D., DAROUSSIN, J. & JAMAGNE, M. 1994. Proposal for a spatial organization model in soil science (the example of the European Communities Soil Map). *Journal of the American Society for Information Science*, 45, 705-717.
- LAGACHERIE, P. 1992. *Formalisation des lois de distribution des sols pour automatiser la cartographie pédologique à partir d'un secteur pris comme référence*. Cas de la petite région de l'Hérault. Université des Sciences et Techniques du Languedoc.
- MOEYS, J., NICOUILLAUD, B., DORIGNY, A., COQUET, Y. & COUSIN, I. 2006. Cartographie des sols à grande échelle : intégration explicite d'une mesure de résistivité apparente spatialisée à l'expertise pédologique. *Etude et Gestion des Sols*, 13, 269-286.
- 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.
- ROBERT, P. C. (1999). Precision Agriculture: research needs and status in the USA. *Second European Conference on Precision Agriculture, Odense, Denmark*.
- SAMOUELIAN, A., COUSIN, I., TABBAGH, A., BRUAND, A. & RICHARD, G. 2005. Electrical resistivity survey in soil science: a review. *Soil and Tillage Research*, 83, 173-193.

SEGER, M., COUSIN, I., FRISON, A., BOIZARD, H. & RICHARD, G. 2009. Characterisation of the structural heterogeneity of the soil tilled layer by using in situ 2D and 3D electrical resistivity measurements. *Soil and Tillage Research*, 103, 387-398.

TETEGAN, D. M., PASQUIER, C., NICOULLAUD, B., BOURENNANE, H., BESSON, A., DESBOURDES-COUTADEUR, C., BOUTHIER, A., KING, D. & COUSIN, I. 2012. Field-scale estimation of the volume percentage of rock fragments in stony soils by electrical resistivity. *Catena*, 92, 67-74.

WACKERNAGEL, H. 1995. *Multivariate Geostatistics*. Springer, Berlin.

WALTER, 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.

## **Résumé : Valorisation d'une prospection électrique pour établir une carte de l'épaisseur de sol**

Dans le projet cadre du projet CAREX de l'INRA, le travail présenté dans ce document a pour objectif d'étudier et de comparer deux approches (non spatial et spatial) de spatialisation de l'épaisseur du sol (variable d'intérêt), à partir de mesures de la résistivité électrique du sol (variable auxiliaire). Une prospection électrique a été réalisée sur l'ensemble du site expérimental de l'INRA d'Epoisses à partir du dispositif tracté ARP, qui mesure la résistivité pour trois profondeurs d'investigation. Une analyse en composantes principales des trois voies de résistivité montre que le premier facteur (PC1) représente 85 % de la variance. L'approche non spatiale repose sur deux modèles de régression linéaire simple qui utilisent soit la voie 3 de la prospection électrique - voie la mieux corrélée à l'épaisseur de sol – soit PC1, comme variables explicatives dans la modélisation et la spatialisation de la variable d'intérêt. L'approche spatiale utilise soit PC1, soit ses basses fréquences (S2-PC1), comme dérive externe dans le krigeage de la variable d'intérêt. Les résultats obtenus montrent que seule l'approche spatiale (en particulier le krigeage avec S2-PC1 comme dérive externe) permet de reproduire tout le spectre de variation mesurée dans le jeu de donnée initial. Elle fournit également des cartes présentant un aspect plus lisse, plus contrasté avec des zones homogènes plus larges facilitant leur usage.

**Mots clés** : Epaisseur du sol, Résistivité électrique, Cartographie, Régression linéaire, Krigeage avec dérive externe, Analyse en Composantes Principales.

### **Abstract: Soil thickness mapping using electrical resistivity survey**

In the framework of CAREX project, this work aims at studying and comparing two approaches (spatial and non-spatial) to map the soil depth (the target variable) from the soil electrical resistivity (the auxiliary variable). Therefore, an electrical resistivity survey was performed by the ARP mobile device throughout the experimental site of INRA Epoisses. A principal component analysis of the three arrays of electrical resistivity measurements demonstrated that the first principal component (PC1) accounted for 85 % of the total variance. The non-spatial approach was based on two simple linear regression models considering the third electrical resistivity array – that is the best correlated with the soil thickness - or PC1, as independent variables to model and map the target variable. The spatial approach, considers PC1 and its low frequencies (S2-PC1) respectively, as an external drift in kriging of the target variable. The results have shown that the spatial approach (especially kriging with S2-PC1 as an external drift) allowed reproducing the full spectrum of variation measured in the training data. In addition, the spatial approach provided more realistic maps compared to the statistical approach. Indeed, these maps are more contrasted with a smooth aspect and with larger homogeneous areas and would then be helpful for management at the plot scale.

**Keywords**: Soil thickness, Electrical resistivity, Soil mapping, Linear regression, Kriging with external drift, Principal Component Analysis