



HAL
open science

Caractérisation des propriétés du sol par télédétection rapprochée et spatiale en milieux tempérés et arides

Hamouda Aichi

► **To cite this version:**

Hamouda Aichi. Caractérisation des propriétés du sol par télédétection rapprochée et spatiale en milieux tempérés et arides. Sciences du Vivant [q-bio]. Institut National Agronomique de Tunisie, 2009. Français. NNT: . tel-02824212

HAL Id: tel-02824212

<https://hal.inrae.fr/tel-02824212>

Submitted on 6 Jun 2020

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

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

**Institut National Agronomique de
Tunisie**



AGROCAMPUS-OUEST
Sous le seau de l'Université Européenne de
Bretagne



N° ordre : 2009-19
N° Série : D-54

THÈSE en co-tutelle INAT/AGROCAMPUS-OUEST

**Pour obtenir les grades de :
Docteur de l'Institut National Agronomique de Tunisie
Spécialité : Production végétale**

&

**Docteur de l'institut supérieur des sciences agronomiques, agro-alimentaires,
horticoles et du paysage
Spécialité: Science de l'environnement**

Soutenue publiquement à l'INAT par
Hamouda AICHI
le 1^{er} Octobre 2009

**Caractérisation des propriétés du sol par
télédétection rapprochée et spatiale en
milieux tempéré et aride**

Devant le jury composé de

**Mr. Fethi Lebdi
Mr. Mustapha Sanaa
Mr. Christian Walter
Mr. Houcine Khatteli
Mr. Yves Le Bissonais
Mme. Zohra Lili Chabaane
Mr. Youssef Fouad**

**Président
Directeur de thèse
Directeur de thèse
Rapporteur
Rapporteur
Examineur
Examineur**



UMR SAS

« Nous sommes la triste opacité de nos spectres futurs »

Stéphane Mallarmé (poète, 1842-1898)

Abstract

Spectroscopy in the visible-Near infrared (Vis-NIR) in combination with chemometry (As Partial Least Squares Regression (PLSR)) constitutes a promising tool to quantify various soil properties. Four analytical and soil spectral data bases have been constituted. The first corresponds to 22 artificial samples. The second and the third are two regional bases, respectively compounds of 125 samples from the temperate region of Brittany and 89 samples, collected in three distinct sites of the same arid region of the Djerid. The fourth is a compound of 144 samples collected on a given site of the Djerid region (580 ha). With the first base, we have : i) studied qualitatively, the influence of organic carbon and total iron contents on the spectral behaviour, then predicted their contents ii) studied the performance variation of the PLSR models according to the calibration set variation. The Brittany base has been analyzed for the organic carbon (CO), the total nitrogen (N_{tot}), the pH, the Cation Exchange Capacity (CEC), the total iron (F_{tot}) and the five texture fractions (clay (Ar), fine silt (LF), coarse silt (LG), fine sand (SF) and coarse sand (SG)). Then while using the Principal Component Analysis (PCA), as selection procedure of calibration and prediction samples, we calibrated the regional 10 physico-chemical properties models and identified, for each property, which of the spectral ranges : combined Vis-NIR, Vis or NIR, offers the best prediction. The third basis served to study, for the total carbon (C_{tot}) and the electric conductivity (EC), the effect of the calibration mode " specific site " *versus* " regional " on the prediction. The fourth basis, coupled to an ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer) sub-image, served to quantify the C_{tot} and the EC spatially by PLSR-Krigeage. We showed, with the help of the artificial basis, that : i) the CO induces a "decreasing effect" on all the spectra without presenting specific peaks. But F_{tot} have absorptions peaks at 431, 497, 677, 795 and 907 nm, ii) the CO and total iron are well predicted by spectrochemometry iii) the PLSR models integrate the variability of samples and the increase of their strength improves the prediction. In the region of Brittany, predictions are adequate for the CO and the CEC. The NIR offers the best predictions, respectively ($R^2 = 0.831$ and 0.69). The accurate predictions of the CO are assigned to the direct influence of carbons on spectra. Nevertheless the satisfactory prediction of the CEC, the LG and the SG are likely due to their correlation with the CO respectively ($r = 0.94$, -0.65 and 0.55). The adequate CO models at the regional scale, obtained thanks to the PCA selection of samples, open the way to calibrate additional global models. In the Djerid region, C_{tot} " specific site " models are effective for the first two sites and slightly satisfactory for the third R^2 respectively (0.85 , 0.83 and 0.55). However, those of the CE_{1:5}, are slightly satisfactory for the first two sites and effective for the third, R^2 respectively (0.57 , 0.50 and 0.76). For the two properties, the robustness of the " specific site " models is geographically limited. The regional model of

the C_{tot} is sufficiently good ($R^2 = 0.67$). It can be improved if recalibrated on a larger number of samples, collected in further sites. However, the EC model is less efficient ($R^2 = 0.53$). Spatialization of the C_{tot} and EC PLSR models, calibrated on the 144 spectra extracted from the radiometrically corrected sub-image, were much better, after residuals kriging; for the C_{tot} , R^2 respectively (0.78 as opposed to 0.53). For the EC, the improvement of R^2 is least (0.60 against 0.57). But the bias and the RMSE are improved meaningfully, respectively (0.003 against 0.36 dS/ms and 0.55 as opposed to 0.75 dS/ms). Our results offer interesting perspectives for large territories mapping. The spectrochemometry is a promising tool for soil characterization.

Key words : soil characterization, spectrochemometry, visible, near infrared, prediction, spatial quantification, region of Brittany, region of Djerid.

Résumé

La spectrométrie visible proche-infrarouge (Vis-PIR) en combinaison avec la chimiométrie, notamment la régression en moindres carrés partiels (RMCP), constitue un outil prometteur pour quantifier diverses propriétés pédologiques. Quatre bases de données analytiques et pédospectrales ont été constituées. La première correspond à 22 échantillons artificiels. La seconde et la troisième sont deux bases régionales, respectivement formées de 125 échantillons de la région tempérée de Bretagne et de 89 échantillons, collectées dans trois sites distincts de la même région aride du Djérid. La quatrième est formée de 144 échantillons collectés sur un même site de la région du Djérid (580 ha). Avec la première base, nous avons : i) étudié qualitativement, l'influence des teneurs en carbone organique et en fer total sur la réponse spectrale, ii) estimé les teneurs en carbone organique (CO) et en fer total, iii) étudié la variation des performances des modèles RMCP en fonction de la variation du jeu de calibration. La base bretonne a été analysée pour le CO, l'azote total (N_{tot}), le pH, la capacité d'échange cationique (CEC), le fer total et les cinq fractions granulométriques (argile (Ar), limon fin (LF), limon grossier (LG), sable fin (SF) et sable grossier (SG)). Puis en utilisant l'Analyse en Composante Principale (ACP), comme procédure de sélection des échantillons de calibration et de prédiction, nous avons calibré des modèles régionaux des 10 propriétés physico-chimique et identifié, pour chaque propriété, laquelle des fenêtres spectrale : combiné Vis-PIR, Vis et PIR offre la meilleure estimation. La troisième base a servi pour étudier, pour le carbone total et la conductivité électrique, l'effet du mode de calibration « site spécifique » *versus* « régional » sur l'estimation. La quatrième base, couplée à une sub-image ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer) a servi pour quantifier spatialement le C_{tot} et la CE par RMCP-Krigeage. Nous avons montré, à l'aide de la base artificielle, que : i) le CO entraîne un effet « dépresseur » sur tout le spectre sans présenter de pics spécifiques. Le fer total possède des pics d'absorptions à 431, 497, 677, 795 et 907 nm, ii) le CO et le fer total sont bien estimés par spectrochimiométrie iii) les modèles RMCP intègrent la variabilité des échantillons et l'augmentation de leur effectif améliore l'estimation. En région de Bretagne, les prédictions sont adéquates pour le CO et la CEC. Le PIR offre les meilleurs estimations, respectivement ($R^2 = 0.831$ et 0.69). Les bonnes prédictions du CO sont attribuées à l'influence directe des carbones sur les spectres. Par contre les prédictions satisfaisantes de la CEC, du LG et du SG seraient dues à leurs corrélations avec le CO respectivement ($r = 0.94$, -0.65 et 0.55). Les modèles de CO adéquats à une échelle régionale, obtenus grâce à la sélection des échantillons par l'ACP, ouvre la voie pour calibrer des modèles plus globaux. En région du Djérid, les modèles « site spécifique » du C_{tot} sont performants pour les deux premiers sites mais peu satisfaisant pour le troisième : R^2 respectivement (0.85 , 0.83 et

0.55). Alors que ceux de la $CE_{1:5}$, sont peu satisfaisants pour les deux premiers sites et performant pour le troisième, R^2 : respectivement (0.57, 0.50 et 0.76). Pour les deux propriétés, les modèles « site spécifique » sont d'une robustesse géographique limitée. Le modèle régional du C_{tot} est assez satisfaisant ($R^2 = 0.67$). Il peut être amélioré par recalibrage sur un nombre plus important d'échantillons, prélevés sur davantage de sites. Alors que celui de la $CE_{1:5}$ est de performance moyenne ($R^2 = 0.53$). Les spatialisations des modèles RMCP de C_{tot} et de CE calibrés sur les 144 spectres extraient de la sub-image radiométriquement corrigée, ont été davantage meilleures, après krigeage des résidus. Pour le C_{tot} , R^2 est passé de 0.53 à 0.78. Pour la $CE_{1:5}$, l'amélioration de R^2 est moindre (0.60 contre 0.57). Nonobstant le biais et la REQM sont significativement améliorés, respectivement (0.003 contre 0.36 dS/m et 0.55 contre 0.75 dS/m). Nos résultats ouvrent des perspectives intéressantes pour une cartographie sur de larges territoires. La spectrochimimétrie est un bon outil de caractérisation des sols.

Mots clés : caractérisation des sols, spectrochimimétrie, visible, proche infrarouge, prédiction, quantification spatiale, région de Bretagne, région du Djérid.

Dédicaces

Je dédicace ce travail à mon feu papa Salah qui nous a quitté pour un monde meilleur, à ma maman Lilia, mes sœurs : Wafa, Olfa, Chadlia, Malika et Safa, mes frères par alliance : Moez, Hamda, Abdelhak et Tawfik, mes neveux et nièces : Molka, Ines, Khalil, Amen Allah, Loulou, Fares et Hayder pour leur soutien sans faille. Cette thèse est un peu la leur.

Remerciements

Rédiger les remerciements est pour moi un précieux moment pour se remémorer, avec un grand plaisir et une bonne dose de nostalgie, des différents moments, bons et moins bons mais que du bon au final, qui ont marqué ces années de préparation de la thèse. C'est aussi une occasion pour exprimer toute ma gratitude et ma reconnaissance pour toutes les personnes qui ont contribué à l'aboutissement de ce travail.

En premier lieu, j'exprime un vibrant hommage à mon directeur de thèse en France Christian Walter, Professeur à Agrocampus-ouest, pour m'avoir donné ma chance, de m'avoir pris en thèse et de m'avoir soutenu. J'ai profité de son savoir, savoir faire et savoir être. J'ai aussi admiré sa rigueur scientifique, sa grande modestie et son attitude positive. Il a toujours su me rassurer, me reconforter et me donner confiance en moi. Pour cela et pour tout le reste merci et encore bravo.

Je m'incline pour remercier mon co-directeur de thèse en France, Youssef Fouad, je ne peux prétendre exprimer en des mots tout mon estime et ma reconnaissance à sa personne. Je ne peux passer sous silence ses grandes qualités scientifique et humaine. Il a toujours été attentif et sécurisant. C'est en grande partie grâce à lui que ce travail a pu être réalisé.

Je remercie comme il se doit mon directeur de thèse en Tunisie, Mustapha Sanaa, pour son soutien inconditionnel. Il n'a cessé tout le long de ce marathon de me conseiller et de m'orienter. Je suis très sensible à son soutien pour que je boucle ce travail à temps. J'espère honorer sa confiance.

Mes révérences et mes considérations à ma co-directrice de thèse en Tunisie, Zohra Lili Chabaane son grand enthousiasme, sa dynamique et son positivisme me sont très communicatifs. Elle trouve toujours : les bons plans pour que je puisse avancer dans mon travail et les justes mots pour que j'aille de l'avant. Merci infiniment

Je remercie également tous les membres de mon comité de pilotage En France :

Richard Escadafal, Directeur de recherche au Centre d'Etudes Spatiales de Biosphère, auteur d'articles phares que j'ai lu et relu à maintes reprises lors de mon initiation à la discipline. C'était un honneur pour moi de faire sa connaissance et de soumettre mon modeste travail à son critique.

David Causeur, Maître de conférences en Mathématiques appliquées, que je remercie pour son aide précieuse dans les traitements chimiométriques. Je l'ai souvent importuné avec mes visites sans rendez-vous dans son bureau. Mais il m'a toujours réservé un bon accueil pour m'expliquer et discuter de mes résultats.

Hervé Nicolas, professeur en bioclimatologie à Agrocampus-ouest, Dont j'avoue avoir abusé de sa gentillesse. Je le remercie pour sa volonté à toujours se rendre disponible, les nombreux conseils et l'intérêt qu'il a porté à mon travail ont largement contribué au bon déroulement de l'ensemble de ce travail.

Laurence Hubert-Moy, Professeur en géographie à l'université de Rennes 2, que je remercie pour ses conseils directionnels et pour avoir mis à notre disposition le spectroradiomètre du CAREN.

Une mention spéciale à Mme Zahra Thomas, maître de conférences en hydraulique et mécanique des fluides à Agrocampus-Ouest. Pour nos échanges et pour son soutien moral. Son labeur et sa forte personnalité m'ont souvent impressionné. Je lui exprime ma grande estime.

Un clin d'œil à Christophe Cudennec, maître de conférences en hydrologie à Agrocampus-ouest, pour son encouragement.

Je remercie également Didier Michot, maître de conférences en géophysique à Agrocampus-Ouest, pour sa sympathie et l'intérêt qu'il a porté à mon sujet de recherche.

Un tendre remerciement à Pascale Breger, technicienne au laboratoire PSN, je ne peux oublier les bons moments que j'ai passé à sa compagnie à préparer et à analyser les échantillons. Sa grande méticulosité a été indispensable dans la réalisation des mesures spectrales au laboratoire.

Je remercie aussi Mariem ROUGHI technicienne au laboratoire de sciences de l'eau à l'INAT qui a assuré les mesures analytiques des échantillons du sol collectés en Tunisie.

Je remercie également Jean-Claude MARIE, technicien au laboratoire PSN, pour son aide précieuse dans l'aménagement de la chambre noire et pour sa grande disponibilité à chaque fois que je l'ai sollicité.

Je remercie aussi Mme Monique DELABUIS responsable de la bibliothèque de L'UMR SAS pour sa grande disponibilité et son extrême gentillesse.

Le plus fort de mes remerciements est pour Evelyne GRESSIER, secrétaire au laboratoire PSN, pour sa bonne humeur, sa serviabilité, pour les pauses café conviviales, les repas chez elles et les sorties qu'elle nous organisait.

J'adresse mes profonds remerciements à tous mes supérieurs successifs au travail au niveau des CRDA's de Tozeur puis de Siliana : MM Hedi METTICHI, Mabrouk AMRI, Bechir DADI et Mohamed NASRI ainsi qu'à la Direction Générale d'Aménagement et de Conservation des Terres Agricoles : MM Hedi HAMROUNI et Habib FARHAT qui m'ont autorisé à effectuer les séjours en France et qui ont eu un regard attentif et bienveillant sur mon travail.

Une pensée pleine de reconnaissance à mes collègues au travail : Leila LAABIDI, Mohamed Ali TRABELSI, Joneidi HARRATHI, Salah HANI, Lotfi TOUMI, Hedi TOUMI, Moufida ZARAAI, Abdallah ZARAAI, Leila JRIDI, Youssef AZABOU, Aida JRIDI, Chams HARRABI Tahar BOUBAKER, et Safia JRIDI pour leur bonne compagnie ainsi que MM Lahmadi MOUMNI et Samir KAMEL qui ont assuré mon intérim pendant mes séjours à Rennes.

Je remercie également mes amis à Tozeur : Ahlem, Fajra, Sassi et Moujib.

J'adresse mes remerciements les plus sincères à Mr Mohamed Hachicha directeur de recherche à l'INGREF pour son soutien moral et pour son encouragement.

Merci également à Sofian OTHMANI, Mohamed Nour SDADI, Aljia ZARROUG, Zakia BEDOUI, Iftikhar CHETOUI, que j'ai eu le plaisir d'encadrer au cours de leurs stages et projets de fins d'études, leur investissement m'a été précieux dans la collecte des données expérimentales de cette thèse.

Faire une thèse est une cavale en solitaire au milieu de compagnons de fortune ou d'infortune, on ne se rend pas à quel point ils nous sont chers jusqu'au jour où nos chemins se séparent. Cette thèse a été pour moi une occasion pour faire la connaissance de plusieurs personnes en qui j'ai découvert des vrais amis et qui m'ont tous aidé et soutenu : Stéphane FOLLAIN, Cedric LEGOUT, Brice MOURIER, Arnaud LEGOUT, Reza GHAZAVI, Najet MISSI, Ouided ZEKKOUR, Zeineb KASSOUK, Houda BOUDHRAA, Muhammed

GULZAR, Nabihah KHAYATI, Khadija Bou AJILA et Jalel AOUISSI. Je leur souhaite tous bons vents.

En dernier lieu, je tiens à remercier les membres du jury qui malgré leurs responsabilités, me font l'honneur de juger ce travail.

« You shouldn't do science just to improve wealth, do science for the sake of human culture and knowledge. There must be some purpose in life that is higher than just surviving ».

Gerhard Herzberg (1904-1999), Le père de la spectroscopie canadienne, Prix Nobel en 1972.

Sigles et abbréviations

A : Absorbance

ACP : Analyse en composantes principales

AFM: Analyse factorielle multiple

Ar : Argile

assim. : Assimilable.

ASTER : Advanced spaceborne thermic emission and reflection radiometer

B : Bore

b_i : Coefficients du modèle à la longueur d'onde i ,

c : Vitesse de la lumière

Ca : Calcium

CE : Variable Conductivité électrique

CE1:5 : Conductivité électrique mesurée dans un extrait aqueux au 1/5ème

CEC : Capacité d'échange cationique

CI : Carbone inorganique

cmol : Centimole

CO : Carbone organique

Cu : Cuivre

C_{tot} : Carbone total

°C : Degré Celsius

Ø : Diamètre

dS : déciSiemens

E : Champ électrique

E : Energie

E : Est

\mathcal{E}_0 : Erreur résiduelle.

éch. : Échangeable

Fe : Fer

Fetot : Fer total

GPS : Géolocalisation par satellite

h : heure

h : Constante de Plank

ha : hectare

Hz : Hertz

I_0 : Radiation incidente

I_a : Intensité absorbée

I_r : Intensité réfléchie

I_t : Intensité transmise

IR : Infrarouge

ISO : International Standardization organization

j : jour

J : Joule

K : potassium

λ : Longueur d'onde

LF : Limons fins

LG : Limons grossiers

log : logarithme népérien

m : Mètre

M : Champ magnétique

MIR : Moyen infrarouge

Mg : Magnésium

Mn : manganèse

ν : Fréquence

mm : Millimètre

mmol : Millimole

MSC : Multiplicative Signal Correction

N : Nord

NCLM: Nombre de composantes latentes du modèle

NIPALS : NonLinear Iterative Partial Least Squares

NIR : Near infrared

nm : Nanomètre

Ntot : Azote total

O : Ouest

PCA : Principle components analysis

π : π_i

pH : Potentiel d'hydrogène

PLSR : Partial least squares regression

P : Phosphore

PIR : Proche infrarouge

r : Coefficient de corrélation

R : Langage de programmation

R : Réflectance

R² : Coefficient de détermination

REP : Ratio des écarts de prédiction

RCP : Régression en composantes principales

REQM : Racine de l'erreur quadratique moyenne
REQMR : Racine de l'erreur quadratique moyenne relative
RLM : Régression linéaire multiple
RMCP : Régression en moindres carrés partiels
RN : Réseau de neurones
RPD : Ratio of prédiction deviation
RPP : Régression pas à pas
S : Horizon structural
S : Sud
Sg : Horizon structural redoxique
SF : Sables fins
SG : Sables grossiers
SNV : Standard Normal Variate
SRD : Spectroscopie par réflexion diffuse
 σ : Nombre d'onde

θ : Humidité volumique
UTM : Universal Transverse Mercator
UV : Ultra violet
Vis : Visible
Vis-PIR : Visible-Proche infrarouge
vs : *Versus*
W/m²/sr/nm : Watt par mètre carrée par steradian par nanomètre
 X_i : Ensemble de variables explicatives (réflectance à la longueur d'onde i)
Y : Variables à expliquer
 y^* : Valeur estimée
Zn : Zinc

ABSTRACT	3
RESUME	5
SIGLES ET ABBREVIATIONS	13
LISTE DES TABLEAUX	18
LISTE DES FIGURES	19
LISTE DES PHOTOS	22
INTRODUCTION GENERALE	23
CHAPITRE I. ANALYSE BIBLIOGRAPHIQUE	28
« INTRODUCTION A LA SPECTROSCOPIE PAR REFLEXION DIFFUSE (SRD) »	28
1. Principes de la spectroscopie	28
2. Spectroscopie dans le visible et le proche infrarouge	30
2.1. <i>Transitions électroniques (dans le visible et le PIR)</i>	30
2.2. <i>Absorption vibrationnelle (dans le PIR)</i>	31
3. Modes de vibration d'une molécule polyatomique.....	32
3.1. <i>Bandes fondamentales</i>	32
3.2. <i>Bandes harmoniques et bandes de combinaisons</i>	33
4. Réflexion diffuse	35
5. Réflectance spectrale.....	37
6. Acquisition des spectres.....	38
7. Spectroscopie qualitative.....	39
8. Spectroscopie quantitative.....	40
8.1. <i>Régression linéaire multiple</i>	42
8.2. <i>Régression en composantes principales</i>	42
8.3. <i>Régression en moindres carrés partiels</i>	43
8.3.1. <i>Historique</i>	43
8.3.2. <i>Principe</i>	43
8.3.3. <i>Choix du nombre de composantes latentes du modèle RMCP</i>	44
8.3.3.1. <i>Critère d'ajustement ou « fit »</i>	44
8.3.3.2. <i>Critère de « validation croisée »</i>	44
8.3.3.3. <i>Critère de « validation externe »</i>	45
9. Prétraitements des spectres	45
10. Application de la spectroscopie par réflexion diffuse dans l'étude des sols.....	46
11. Facteurs influençant la réponse spectrale des sols.....	46
11.1. <i>Humidité</i>	47
11.2. <i>Rugosité de la surface</i>	48
11.3. <i>Matière organique</i>	48
11.4. <i>Constituants minéraux</i>	49
11.4.1. <i>Minéraux argileux</i>	49
11.4.2. <i>Oxydes et hydroxyde de fer</i>	50
11.4.3. <i>Minéraux carbonatés et sulfatés</i>	51
12. Niveaux de précisions des prédictions spectrométriques des propriétés des sols dans la littérature	51
CHAPITRE II. INFLUENCE DU CARBONE ORGANIQUE ET DU FER TOTAL SUR LES SPECTRES DE REFLECTANCE DU SOL ET ESTIMATION CHIMIOMETRIQUE DE LEURS TENEURS	54
ABSTRACT	54
Résumé	55
1. Introduction	56
2. Matériels et méthodes	58
2.1. <i>Constitution de deux séries d'échantillons artificiels</i>	58
2.2. <i>Acquisition des spectres</i>	62
2.3. <i>Construction du modèle statistique de prédiction</i>	64
3. Résultats	66
3.1. <i>Influence des teneurs en carbone organique et en fer total sur la réponse spectrale</i>	66

3.2. Prédiction du carbone organique et du fer total par régression MCP	70
4. Discussion	73
4.1. Interprétation de la réponse spectrale	73
4.2. Interprétation des performances des modèles calibrés.....	75
5. Conclusion	76
CHAPITRE III. COMPARAISON DES POTENTIELS DU VISIBLE, DU PROCHE INFRAROUGE ET DU COMBINE VISIBLE-PROCHE INFRAROUGE POUR LA PREDICTION SIMULTANEE EN MODE REGIONAL DE PLUSIEURS PROPRIETES DU SOL EN MILIEU TEMPERE (REGION DE BRETAGNE, FRANCE)	77
Abstract.....	77
Résumé	78
1. Introduction	79
2. Matériels et méthodes	81
2.1. Constitution d'une base régionale d'échantillons de sol	81
2.2. Procédure de sélection des échantillons de calibration et de prédiction du modèle	82
2.3. Analyse spectrale des échantillons	83
2.4. Construction du modèle statistique de prédiction	83
2.4.1. Régression en Moindres Carrés Partiels (RMCP)	83
2.4.2. Choix du nombre de composantes latentes des modèles	84
2.4.3. Evaluation des performances des modèles	85
2.4.4. Traitement des données	85
3. Résultats et discussion.....	86
3.1. Statistiques descriptives des propriétés analytiques.....	86
3.2. Corrélations entre les propriétés analytiques.....	87
3.3. Sélection des échantillons de calibration et de validation.....	88
3.4. Performances des modèles.....	91
4. Conclusion	97
CHAPITRE IV. QUALITE D'ESTIMATION DES MODELES SPECTROMETRIQUES « SITE SPECIFIQUE » ET « REGIONAL ». EXEMPLE DE LA TENEUR EN CARBONE TOTAL ET DE LA CONDUCTIVITE ELECTRIQUE DES SOLS, EN MILIEU ARIDE.....	99
Abstract.....	99
Résumé	100
1. Introduction	101
2. Matériels et méthodes	103
2.1. Milieu d'étude.....	103
2.2. Stratégie d'échantillonnage	104
2.3. Caractéristiques des trois sites	108
2.3.1. Site de Chamsa.....	108
2.3.2. Site d'Oudia	109
2.3.3. Site d'Oung.....	109
2.4. Analyses des échantillons.....	109
2.5. Acquisition et traitement des spectres	110
2.6. Calibration des modèles.....	111
2.7. Choix du nombre optimum de composantes latentes des modèles calibrés.....	111
2.8. Evaluation des performances des modèles	112
3. Résultats et discussion	114
3.1. Statistiques descriptives.....	114
3.2. Analyse multivariée des jeux de données.....	116
3.3. Sélection des échantillons de calibration et de validation.....	119
3.4. Performance des modèles « site spécifique »	122
3.5. Performances des modèles établis à l'échelle régionale	128
4. Conclusion	131
CHAPITRE V. QUANTIFICATION SPATIALE DE PROPRIETES DU SOL EN MILIEU ARIDE, PAR COUPLAGE DE DONNEES SPECTROSCOPIQUES VISIBLE-PROCHE INFRAROUGE ACQUISES AU LABORATOIRE ET DE DONNEES D'IMAGE ASTER.....	133
Abstract.....	133

Résumé	134
1. Introduction	135
2. Matériels et méthodes	137
2.1. Site d'étude	139
2.2. Traitement de l'image satellitale ASTER.....	141
2.3. Prélèvement des échantillons sur le terrain.....	142
2.4. Analyses au laboratoire.....	143
2.5. Acquisition des spectres	144
2.6. Correction radiométrique de la sub-image	144
2.7. Extraction des spectres de réflectance au sol à partir de la sub-image corrigée en réflectance relative.....	145
2.8. Spatialisation des données par régression MCP-krigeage.....	145
2.8.1. Première étape : la régression MCP	145
2.8.2. Deuxième étape : le krigeage des résidus	146
2.8.3. Production des cartes thématiques.....	147
3. Résultats et discussion.....	148
3.1. Statistiques descriptives des données analytiques.....	148
3.2. Analyse des spectres des échantillons de sol.....	151
3.3. Spectres de radiances extraits de la sub-image	152
3.4. Spectres extraits de la sub-image, après sa correction radiométrique.....	153
3.5. Performances d'estimation spatiale par régression MCP.....	154
3.6. Analyse variographique des résidus de la prédiction	159
4. Conclusion	170
CONCLUSION GENERALE	172
1. Rappel des objectifs de la thèse.....	172
2. Principaux résultats acquis	174
2.1. Formes des spectres de sol.....	174
2.2. Influence du fer total sur la réponse spectrale.....	174
2.3. Influence du carbone organique sur la réponse spectrale.....	175
2.4. Effets conjugués du fer total et du carbone organique sur la réponse spectrale	175
2.5. Intérêt de l'utilisation des modèles RMCP en modélisation des mesures spectrométriques.....	175
2.6. Intérêt de la procédure d'ACP dans la sélection des échantillons de calibration et de validation des modèles à champ d'application régional.....	176
2.7. Opportunité d'utilisation de la spectrométrie dans la prédiction des différentes propriétés du sol en région tempérée de Bretagne.....	176
2.8. Fenêtre spectrale offrant les meilleures prédictions	178
2.9. Parenté pédogénétique des échantillons	178
2.10. Comparaison entre modèle « site spécifique » et modèle « régional ».....	178
2.11. Prédiction spatiale des propriétés du sol en milieu aride, par couplage des données de spectroscopie visible-proche infrarouge et télédétection satellitale.....	179
3. Perspectives de recherche	180
3.1. Estimation en routine des propriétés du sol.....	180
3.2. Cartographie thématique	183
3.2.1. Co-krigeage	184
3.2.2. Régression-krigeage	185
3.2.3. Classification-krigeage (régression-krigeage par classes radiométriques).....	185
3.3. Surveillance des sols.....	185
BIBLIOGRAPHIE GENERALE	186

Liste des tableaux

Chapitre 1.

Tableau 1. Principaux minéraux argileux avec leurs longueurs d'onde d'absorption correspondantes.....	50
Tableau 2. Résultats de prédiction quantitatives de diverses propriétés du sol utilisant différentes techniques de régression multivariée et la réponse spectrale des échantillons de sols dans le domaines du visible et du proche-infrarouge (D'après une synthèse de Viscarra et al., 2006).	52

Chapitre 2.

Tableau1. Caractéristiques des matériaux ayant servi pour la constitution des deux séries artificielles.	59
Tableau 2. Teneurs en carbone organique et en fer total des deux séries d'échantillons artificielles.	61
Tableau 3. Performance des modèles RMCP pour les deux pédo-variables.	72

Chapitre 3.

Tableau 1. Statistiques sommaires sur les propriétés analytiques des échantillons.....	86
Tableau 2. Matrice des corrélations entre les propriétés analytiques.....	87
Tableau 3. Analyses des échantillons atypiques.....	89
Tableau 4. Statistiques descriptives des échantillons de calibration.	90
Tableau 5. Statistiques descriptives des échantillons de validation.	90
Tableau 6. Performances des modèles aux stades de calibration et de prédiction	92

Chapitre 4.

Tableau 1. Caractéristiques des images satellites.....	105
Tableau 2. Méthodes d'analyse des échantillons, pour les différentes propriétés considérées.	110
Tableau 3. Interprétation des valeurs obtenues pour le ratio REP dans la littérature	112
Tableau 4. Statistiques des variables analytiques à l'échelle de chacun des trois sites.	115
Tableau 5. Statistiques des variables analytiques à l'échelle de la région (Tous sites confondus).	116
Tableau 6. Performances des modèles calibrés en mode « site spécifique », lorsque l'équation de chacun de ces modèles est utilisée pour estimer, les teneurs des deux jeux d'échantillons collectés dans les deux autres sites.	127

Chapitre 5.

Tableau 1. Caractéristiques du capteur ASTER (d'après Yamaguchi et al., 1998).	142
Tableau 2. Statistiques descriptives des deux propriétés analytiques des 144 échantillons de sols collectés.....	148
Tableau 3. Classification des sols selon leurs CE1:5 (d'après : Wang et al., 2008 ; Gherina et al., 2009)	149
Tableau 4. Résultats des performances d'estimation spatiale par régression MCP, appliquée aux 144 spectres en 9 bandes extraits de l'image ASTER après sa correction en réflectance relative.	154

Liste des figures

Chapitre 1.

Figure 1. Représentation schématique des deux champs électrique (E) et magnétique (M) composant une onde électromagnétique (d'après Bonn et Rochon, 1992 dans Girard et Girard, 1999).	28
Figure 2. Diagramme du spectre électromagnétique (d'après Steward, 2006).	30
Figure 3. Absorption de l'énergie par excitation électronique.	31
Figure 4. Bandes d'absorption fondamentales d'une molécule triatomiques non linéaire, exemple de la molécule d'eau (H ₂ O).	33
Figure 5. Exemples de vibrations atomiques : cas d'une chaîne hydrocarbonée (rond noir : atome carbone, rond blanc : atome d'hydrogène) (d'après Bertrand, 2002).	35
Figure 6. Formes de réflexions occasionnées par un faisceau de radiation incidente sur un sol (d'après Viscarra, 2007a).	37
Figure 7. Dispositif expérimental de mesure de la réflectance en conditions contrôlées de laboratoire, d'après (Escadafal et al., 1993 ; Mathieu et al., 1998).	38
Figure 8. Courbes typiques des comportements spectraux de trois matériaux distincts : l'eau, la végétation et le sol (d'après Lillesand et Kiefer, 1987).	40
Figure 9. Influence de la teneur en eau sur la réflectance d'un sol limoneux (d'après Bowers et Hank, 1965).	48
Figure 10. Comparaison des réflectances Vis-PIR de quatre minéraux argileux purs communs dans les sols (d'après Mustard et Sunshine, 1999).	50

Chapitre 2.

Figure 1. Réponses spectrales des deux séries d'échantillons artificiels. (a) série artificielle I (Horizon A+Sg), (b) série artificielle II (Horizon A+S _{« drainé »}). Pour chaque échantillon nous indiquons successivement sa teneur en carbone organique et en fer total.	66
Figure 2. Représentation simultanée des réponses spectrales des deux gammes artificielles. Pour chaque échantillon nous indiquons respectivement, sa teneur en carbone organique et en fer total.	68
Figure 3. Spectres de réflectance des deux séries d'échantillons artificiels après le retrait du continuum. (a) série artificielle I (Horizon A+Sg), (b) série artificielle II (Horizon A+S _{« drainé »}). Chaque échantillon est identifié respectivement, par sa teneur en carbone organique et en fer total.	69
Figure 4. Prédiction du carbone organique par régression MCP pour les deux séries artificielles. (a) et (c) coefficients de la régression MCP en fonction des longueurs d'onde, pour les séries I et II respectivement. (b) et (d) teneurs prédites en fonction des teneurs mesurées par validation croisée, pour les séries I et II respectivement.	70
Figure 5. Prédiction du fer total par régression MCP pour les deux séries artificielles. Série I : (a) et (c) coefficients de la régression MCP en fonction des longueurs d'onde, pour les séries I et II respectivement. (b) et (d) teneurs prédites en fonction des teneurs mesurées par validation croisée, pour les séries I et II respectivement.	71
Figure 6. Prédiction par régression MCP en validation croisée sur l'ensemble des deux séries artificielles : (a) carbone organique et (b) fer total.	73

Chapitre 3.

Figure 1. Localisation des sites de prélèvement des échantillons en différents sites de la région de Bretagne.	82
Figure 2. Sélection des échantillons de calibration et de validation par projection des 125 échantillons de sols dans le premier plan factoriel de l'ACP. (Carré rouge : échantillon	

atypique, carré bleu : échantillon du jeu de calibration, carré vert : échantillon du jeu de validation).....88

Chapitre 4.

Figure 1. Aperçu sur la région d'étude (couverte par une image ASTER) et sur l'emplacement des trois sites prospectés (couverts par trois images QuickBird) : (a) site de Chamsa, (b) site d'Oudia et (c) site d'Oung.....105

Figure 2a. Schéma d'échantillonnage du site de Chamsa (Les épingles représentent les points de collectes des échantillons. Les coordonnées sont en mètre selon la projection UTM).107

Figure 2b. Schéma d'échantillonnage du site d'Oudia (Les épingles représentent les points de collectes des échantillons. Les coordonnées sont en mètre selon la projection UTM).107

Figure 2c. Schéma d'échantillonnage du site d'Oudia (Les épingles représentent les points de collectes des échantillons. Les coordonnées sont en mètre selon la projection UTM).108

Figure 3. Projection des variables pédologiques dans le premier plan factoriel de l'ACP pour les jeux d'échantillons de Chamsa (a), Oudia (b), Oung (c) et pour le jeu d'échantillons régional (d).117

Figure 4. Sélection des échantillons de calibration et de validation par projection des échantillons de sols dans le premier plan factoriel de l'ACP : site de Chamsa (a), site d'Oudia (b), Site d'Oung (c) et base régionale (d). (Carré rouge : échantillon atypique, carré bleu : échantillon du jeu de calibration, carré vert : échantillon du jeu de validation).121

Figure 5. Performances des modèles « site spécifiques » pour la variable Ctot. Valeurs prédites vs mesurées pour les sites de Chamsa (5a), Oudia (5b) et Oung (5c).123

Figure 6. Performances des modèles « site spécifiques » pour la variable CE_{1:5}. Valeurs prédites vs mesurées pour les sites de Chamsa (6a), Oudia (6b) et Oung (6c).126

Figure 7. Performances des modèles régionaux (Valeurs prédites vs mesurées, pour la variable Ctot (7a) et la variable CE_{1:5} (7b)).129

Chapitre 5.

Figure 1. Diagramme de spatialisation des données spectrométriques pour l'obtention de cartes thématiques par modèle de régression MCP-Krigeage.138

Figure 2. Scène de l'image ASTER couvrant la partie septentrionale de la région du Djérid, visualisée en fausses couleurs RVB 321 avec une résolution de 60X60 km. Image prise le 27-10-2006. Le rectangle rouge délimite le site d'étude expérimental, son aperçu est affiché à l'aide de l'outil Google map de Google Earth.140

Figure 3. Schéma d'échantillonnage systématique aux nœuds d'une grille carrée régulière de 200 m de côté. Points échantillonnés, visualisés sur un fond d'image satellite extrait à l'aide de l'outil Google maps de Google Earth.143

Figure 4. Représentation des grandeurs des valeurs mesurées de Ctot (a) et de CE_{1:5} (b) au niveau des points échantillonnés sur le terrain, avec des cercles dont les tailles varient proportionnellement aux valeurs mesurées.150

Figure 5a. Spectres des 144 échantillons collectés sur le terrain, acquis en conditions contrôlées de laboratoire, après séchage et tamisage à 2 mm des échantillons. Des pics d'absorption, liées à la présence de certains composés, sont identifiés à partir de données de la littérature.151

Figure 5b. Spectres de radiance extraits des 144 pixels de la sub-image, correspondants aux nœuds des mailles échantillonnés sur le terrain avant sa correction radiométrique.152

Figure 5c. Spectres extraits de la sub-image, après sa correction radiométrique, aux niveaux des des 144 nœuds de mailles échantillonnés sur le terrain.153

Figure 6a. Valeurs de Ctot (%) estimées par régression MCP, appliquées aux spectres en 9 bandes extraits de la sub-image ASTER après sa correction en réflectance relative, vs valeurs mesurées au laboratoire.155

Figure 6b. Valeurs estimées par régression MCP vs valeurs mesurées au laboratoire dans le cas de la variable $CE_{1:5}$ (dS/m).....	157
Figure 7a. Cartographie du carbone total sur la zone d'étude par application du modèle RMCP, propre à la variable C_{tot} , aux neuf bandes de la sub-image ASTER d'une résolution de 15 m.	158
Figure 7b. Cartographie de la $CE_{1:5}$ sur la zone d'étude par application du modèle RMCP, propre à la variable CE , aux neuf bandes de la sub-image ASTER d'une résolution de 15 m.	159
Figure 8. Variogrammes expérimentaux directionnels (a) et surfacique (b) pour les résidus de la variable C_{tot}	160
Figure 9. Variogrammes expérimentaux directionnels (a) et surfacique (b) pour les résidus de la variable $CE_{1:5}$	161
Figure 10. Krigeage ordinaire des résidus de la régression MCP, obtenue dans le cas de la variable C_{tot}	162
Figure 11. Carte de spatialisation du taux de C_{tot} obtenue par régression MCP-krigeage (issue de la sommation des deux couches rasters : spatialisation de la régression MCP figure (7a) + interpolation des résidus figure (11)).	163
Tableau 5. Résultats des performances d'estimation en appliquant le modèle de la régression MCP puis le modèle de régression MCP- Krigeage.....	164
Figure 12. Valeurs de C_{tot} estimées par RMCP-krigeage au niveau des 144 pixels de la carte générée, correspondants aux points de collecte des échantillons sur le terrain, vs les valeurs mesurées au laboratoire.....	165
Figure 13. Couche d'interpolation par krigeage ordinaire des résidus de la $CE_{1:5}$	166
Figure 14. Carte de spatialisation de la $CE_{1:5}$ obtenue par régression MCP-krigeage.....	167
Figure 15. Valeurs de $CE_{1:5}$ estimées par RMCP-krigeage au niveau des pixels de collecte des échantillons vs valeurs mesurées au laboratoire.	168

Conclusion générale.

Figure 1. Diagramme synoptique pour l'estimation ponctuelle de différentes propriétés du sol à partir de mesures spectrométriques et d'une base de données (pedo-spectrale et analytique) liée à un système d'information géographique. (ACP : Analyse en composantes principales, AFM : Analyse factorielle multiple).....	181
Figure 2. Diverses méthodes d'intégration des mesures spectrométriques dans la perspective d'amélioration de la qualité de la spatialisation des propriétés du sol.	184

Liste des photos

Introduction générale.

Photo 1. Mesures directes, sur le terrain, de la luminance des surfaces du sol (a, b, c, d et e) et de la luminance du panneau de référence (f), en condition d'illumination naturelle.39

Chapitre 1.

Photo 1. Horizons de prélèvement des trois matériaux utilisés dans la constitution des deux séries d'échantillons artificielles. Matériau I : horizon A, Matériau II : horizon Sg et Matériau III : horizon S.59

Photo 2. Matériaux originels après séchage et tamisage à 2 mm placés dans des boîtes de Petri.....60

Photo 3. Séries I et II issues des mélanges à différentes proportions respectivement des matériaux I et II (série I), I et III (série II). Nous indiquons les valeurs min et max ainsi que le pas moyen de variation des teneurs en carbone organique et en fer total.62

Photo 4. Réglage des angles zénithaux des lampes à l'aide d'un pyranomètre (a) et mesure du spectre d'un échantillon en conditions contrôlées de laboratoire.63

Introduction générale

Historiquement les analyses physico-chimiques du sol réalisées au laboratoire constituent la principale source d'informations disponibles pour caractériser le sol, apprécier ses aptitudes, comprendre et modéliser son fonctionnement et le cartographier.

Toutefois, les procédures standards d'analyse au laboratoire sont le plus souvent onéreuses, lentes et exigent un personnel qualifié (Brown *et al.*, 2005). De plus, les analyses classiques génèrent beaucoup de déchets, utilisent des réactifs et des acides qui sont souvent nocifs pour la santé et l'environnement, et rendent les échantillons du sol impropres pour des analyses ultérieures (Malley *et al.*, 1999 ; Chang *et al.*, 2002 ; Dunn *et al.*, 2002).

Les pouvoirs publics incitent de plus en plus à la multiplication des études liées à la gestion des sols (Burrough, 1995). Afin de mieux gérer et préserver la ressource sol, nous connaissons actuellement un fort développement de méthodes d'analyses qui soient à coûts réduits, plus rapides et plus propres (Brown *et al.*, 2005) et qui répondent à des exigences normatives de plus en plus strictes (Thuriès, 2007). Ces méthodes devraient être en mesure de générer en abondance des données fiables, qui sont nécessaires dans la surveillance environnementale, la modélisation du fonctionnement du sol et l'agriculture de précision.

Les développements récents de la spectroscopie appliquée au sol, et notamment de la spectroscopie quantitative qui met en œuvre de nouvelles techniques de chimométrie, confirment que la spectroscopie par réflexion diffuse (SRD) peut être considérée comme un outil fiable, efficient, bon marché et non destructif (Dalal et Henry, 1986 ; Sudduth et Hummel, 1993 ; Ben-Dor et Banin, 1995 ; McCarty *et al.*, 2002 ; Shepherd et Walsh, 2002 ; Dunn *et al.*, 2002 ; Brown *et al.*, 2005). Cette technique permet de caractériser d'une façon simultanée et/ou différée diverses propriétés physiques, chimiques et biologiques du sol (Ben Dor *et al.*, 1997 ; Janik *et al.*, 1998 ; Batten, 1998 ; Dunn *et al.*, 2002 ; McCarty *et al.*, 2002 ; Brown *et al.*, 2005 ; Cécillon *et al.*, 2008). Par ailleurs, la mesure spectroscopique en elle-même est simple, commode et ne produit pas de déchets. Les échantillons restent intacts (Stark, 1986 ; Davies, 1987) et disponibles pour toute utilisation ultérieure. Dès lors la spectroscopie représente une opportunité réelle d'augmenter notre capacité d'observation et d'accès direct aux propriétés du sol (Gomez et Lagacherie, 2007).

Les premiers travaux qui se sont intéressés aux possibilités de l'utilisation de la réflexion diffuse dans l'étude des sols remontent aux années 1960, avec les travaux d'Obukhov et Orlov (1964) en Union Soviétique et Bowers et Hanks (1965) aux Etats-Unis.

Par ailleurs, le nombre des études en la matière a connu une croissance soutenue depuis le lancement du premier satellite commercial d'observation de la terre ERTS-1, lancé le 23 juillet 1972 et rebaptisé ultérieurement LANDSAT-1 (Bonn, 1994). Depuis ce temps, beaucoup de recherches ont porté sur l'utilisation de la spectroscopie par réflexion diffuse Vis-PIR dans les études du sol (Ben-Dor et Banin, 1995 ; Janick et Skjemstad, 1995 ; Janik et *al.*, 1998 ; Ben Dor, 2002 ; Dunn et *al.*, 2002 ; Shepherd et Walsh, 2002 ; Chang et Laird, 2002 ; McCarty et *al.*, 2002 ; Islam et *al.*, 2003 ; Brown et *al.*, 2005 ; Viscarra Rossel et *al.*, 2006).

Un énorme progrès a été réalisé aussi bien en termes de technologies d'acquisition des données que de méthodologies de leurs traitements (Ben Dor, 2002) et nous sommes passés progressivement de la spectroscopie qualitative à la spectroscopie quantitative. Ceci a repoussé des frontières jusque là infranchissables dans l'analyse des propriétés des sols et leurs cartographies thématiques.

Ces recherches ont exploré le potentiel de la technique spectroscopique dans la prédiction de nombreuses propriétés physico-chimiques et biologiques du sol. Sans être exhaustive, nous pouvons citer les propriétés suivantes : le carbone organique (CO), le carbone inorganique (CI), le carbone total (C_{tot}), l'azote total (N_{tot}), le pH, la teneur en eau volumique (θ_m), l'argile (Ar), la capacité d'échange cationique (CEC), le Calcium (Ca) échangeable, le Magnésium (Mg) échangeable, le phosphore (P), le potassium (K), les métaux qui incluent le manganèse (Mn), le zinc (Zn), le fer (Fe), le bore (B), et le cuivre (Cu), *etc.*

Nous pouvons aussi constater que les recherches en spectroscopie quantitative se sont articulées autour des trois axes suivants :

- la prédiction des propriétés à partir de spectres acquis dans des conditions contrôlées de laboratoire,
- la prédiction des propriétés à partir de spectres acquis *in situ* (Houssa et *al.*, 1998 ; Mouazen, et *al.*, 2005 ; Waiser, 2006),
- le couplage de données spectroscopiques de laboratoire ou acquis *in situ* avec les données de spectroradiomètres imageurs et/ou des images de télédétection, pour cartographier à différentes échelles les propriétés du sol (Vaudour et Berthier, 2007 ; Gomez et Lagacherie, 2007). Les images de télédétection sont en fait des ensembles de mesures du rayonnement solaire réfléchi par la surface de la Terre, ordonnées sous une forme matricielle (Deshayes et Maurel, 1990).

Cependant, des auteurs comme Terhoeven-Urselmans *et al.* (2006) et Travis (2006) estiment que l'utilité et les limites de l'utilisation de la spectrométrie Vis-PIR pour la prédiction quantitative des propriétés du sol demeurent insuffisamment étudiées. Beaucoup de questions scientifiques ou de mises en œuvre de la technique font encore l'objet de discussions :

- Comment faut-il choisir les jeux de calibration et de validation ?
- Quel est le domaine spectral qui offre les meilleures prédictions ?
- Comment l'étendue géographique de collecte des échantillons, influence t-elle la qualité des modèles de prédiction ?
- Comment combiner des données spectrométriques acquises au laboratoire et des données des images satellites dans une cartographie thématique ?

L'objectif global de la thèse est d'évaluer l'utilisation d'informations spectrales sur les sols, dans le domaine visible et proche infrarouge (Vis-PIR), pour prédire localement puis dans l'espace des propriétés du sol. Trois objectifs spécifiques ont été retenus :

- (1) analyser la performance des techniques spectrométriques pour caractériser des propriétés analytiques des sols, pour une large gamme de sols incluant des sols limoneux des milieux tempérés (Massif armoricain) et des sols sableux des milieux arides (Djérid tunisien) ;
- (2) analyser la sensibilité des modèles de prédiction au jeu de calibration et à la gamme spectrale considérés ;
- (3) spatialiser des propriétés des sols en couplant des données spectrométriques se référants à des échantillons de sols avec des données de réflectances relatives continues dérivées d'une image satellite.

Nous avons opté pour rédiger ce mémoire sous la forme d'articles scientifiques autonomes, néanmoins présentés selon un ordre logique qui retrace les étapes et l'approche globale selon lesquelles l'ensemble du travail a été conduit. Ce choix de rédaction occasionne des répétitions essentiellement dans la description du matériel et des méthodes, mais offre l'avantage de chapitres compréhensibles indépendamment de l'ensemble. Outre l'introduction et la conclusion générales, cinq chapitres composent ce mémoire :

Chapitre I : *Introduction à l'application de la spectrométrie en science du sol.*

Ce chapitre d'analyse bibliographique a pour objectif d'introduire les bases de la spectroscopie et de la chimiométrie, et de faire l'état de l'art en matière d'utilisation de la spectroscopie pour la caractérisation des sols.

Chapitre II : *Influence du carbone organique et du fer total sur les spectres de réflectance du sol et estimation chimiométrique de leurs teneurs.*

A l'aide de deux jeux d'échantillons artificiels que nous avons préalablement constitués, nous étudions l'influence singulière du carbone organique et du fer total sur les spectres des sols et nous explorons la possibilité de dériver les teneurs de ces constituants à partir des spectres en appliquant des régressions en moindres carrés partiels univariées.

Chapitre III : *Comparaison des potentiels du visible, du proche infrarouge et du combiné visible-proche infrarouge pour la prédiction simultanée en mode régional, (Région tempérée de Bretagne), de plusieurs propriétés du sol en utilisant la procédure couplée ACP/RMCP.*

Dans ce chapitre, nous développons des modèles prédictifs d'extension régionale, qui sont calibrés à partir d'échantillons collectés en divers endroits de la région de Bretagne. La prédiction porte simultanément sur plusieurs propriétés du sol et les modèles sont successivement calibrés dans les domaines spectraux du combiné Vis-PIR, du Vis et du PIR. La sélection des échantillons de calibration et de validation repose sur une procédure exploitant l'analyse en composantes principales. Les résultats de ce chapitre ont fait l'objet d'un poster présenté à l'occasion du colloque international « Pedometrics 2007 » à Tübingen (Allemagne). En outre, ce travail a été soumis à la revue internationale « Biosystems Engineering ».

Chapitre IV : *Qualité d'estimation des modèles spectrométriques « site spécifique » et « régional ». Exemple de la teneur en carbone total et de la conductivité électrique des sols, en milieu aride.*

Dans ce chapitre, pour trois sites de sols nus non mis en culture appartenant à une même région naturelle de la région aride du Djérid (Sud-ouest de la Tunisie), nous étudions la performance des modèles prédictifs en fonction de l'étendue de l'aire géographique de collecte des échantillons de calibration et de validation.

Chapitre V : *Quantification spatiale de propriétés du sol en milieu aride, par couplage de données spectroscopique visible-proche infrarouge acquises au laboratoire et données d'image ASTER.*

La compréhension de la répartition spatiale du carbone total (C_{tot}) et de la conductivité électrique (CE) est un enjeu environnemental majeur en milieu aride. Ce chapitre évalue une procédure géostatistique de régression-krigeage pour cartographier ces deux propriétés.

Chapitre I. Analyse bibliographique

«Introduction à la Spectroscopie par Réflexion Diffuse (SRD)»

Toute l'information que nous possédons sur l'univers qui nous entoure nous vient de la lumière que nous recevons. La spectroscopie est la science de l'étude de l'interaction de la lumière avec les objets (Springsteen, 1998 ; Clark, 1999).

1. Principes de la spectroscopie

La spectroscopie concerne la mesure, l'analyse et l'interprétation des spectres (Homayouni, 2005). Le rayonnement lumineux est composé de deux grandeurs physiques : le champ électrique (E) et le champ magnétique (M). Ces deux grandeurs sont inséparables dans la lumière et interagissent l'une sur l'autre en permanence. En effet, le champ électrique (E) varie en amplitude selon une direction perpendiculaire à la direction de déplacement de la radiation. Le champ magnétique (M) est orienté suivant un angle droit du champ électrique (E) (Figure 1). Les deux champs électrique (E) et magnétique (M) se propagent sans atténuation dans le vide à la vitesse de la lumière 3.10^8 m.s^{-1} . Par contre lorsque le rayonnement électromagnétique traverse un milieu donné l'énergie qu'il transporte et plus ou moins absorbées en fonction de la nature de ce milieu (Girard et Girard, 1999).



Figure 1. Représentation schématique des deux champs électrique (E) et magnétique (M) composant une onde électromagnétique (d'après Bonn et Rochon, 1992 dans Girard et Girard, 1999).

La lumière peut être mesurée en unité de fréquence ou de longueur d'ondes. Selon l'équation (1) :

$$\nu = c/\lambda \quad [1]$$

Où λ est la longueur d'onde (m), ν est la fréquence (Hz) et c est la vitesse de la lumière: $2.99 \cdot 10^8$ m.s⁻¹. Comme c est une constante, la fréquence est donc proportionnelle à l'inverse de la longueur d'onde plus communément appelé nombre d'onde σ . Selon l'équation (2) :

$$\sigma = \lambda^{-1} \quad [2]$$

D'ailleurs, en spectroscopie le rayonnement est souvent caractérisé par le nombre d'onde en cm⁻¹ et non par sa longueur d'onde ou par sa fréquence. La lumière est composée de photons et chaque photon transporte un quantum d'énergie proportionnel à la fréquence (ou longueur d'onde de la lumière) selon l'équation (3) :

$$E = h\nu = hc/\lambda \quad [3]$$

Où E est l'énergie (J) et h est la constante de Plank ($6.626 \cdot 10^{-34}$ J-s).

Cette équation explique l'impact qualitatif de la lumière sur les objets (figure 2) : à de hautes fréquences correspondent des hautes énergies et de ce fait les courtes longueurs d'onde, comme celles du visible, véhiculent plus d'énergie que les grandes longueurs d'onde comme celles du proche infrarouge (PIR) (Umland and Bellama, 1996). L'infrarouge est subdivisé en proche infrarouge (0.7 à 2.5 μ m), moyen infrarouge (2.5 à 25 μ m) et infrarouge lointain (> 25 μ m). Cette classification n'est cependant pas universelle : les bornes varient d'un domaine d'utilisation à l'autre sans que l'on puisse donner raison à qui que ce soit.

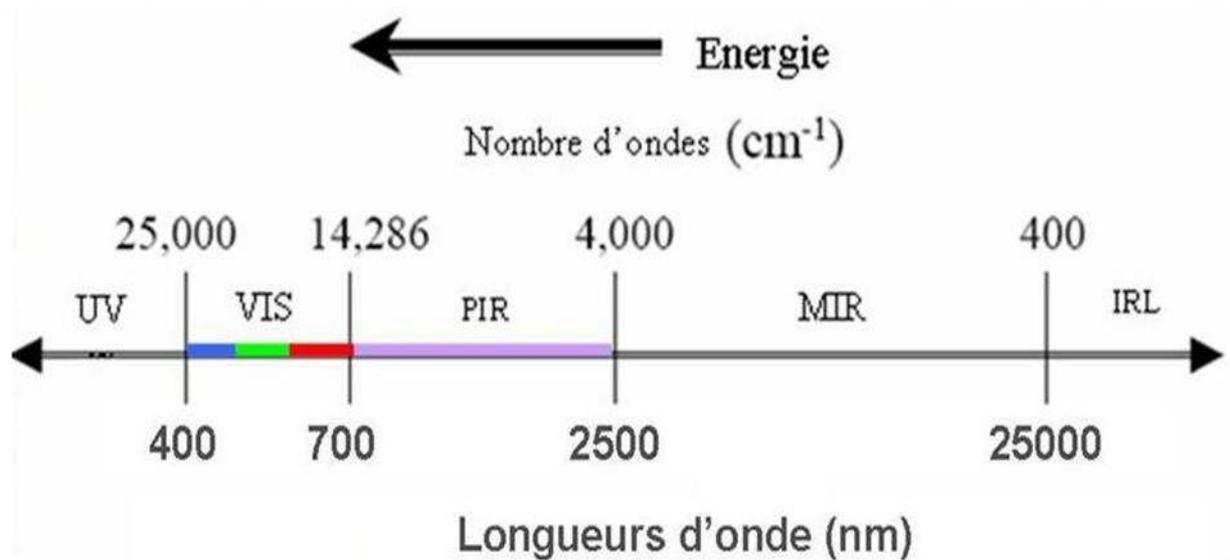


Figure 2. Diagramme du spectre électromagnétique (d'après Steward, 2006).

2. Spectroscopie dans le visible et le proche infrarouge

L'interaction entre le rayonnement électromagnétique, dans les domaines du visible et du proche infrarouge, et la matière fait intervenir deux types de phénomènes physiques qui sont bien expliqués dans le cadre de la mécanique quantique : les transitions électroniques et les absorptions vibrationnelles (Harris et Bertolucci 1978 ; Hollas, 1996).

2.1. Transitions électroniques (dans le visible et le PIR)

Lorsque la matière, notamment le sol, est exposée à des radiations lumineuses dans la région du visible, il peut absorber une quantité spécifique d'énergie lumineuse. Certaines des molécules de ces constituants peuvent subir une excitation électronique, car certains de leurs électrons sont projetés de leur orbitale normale (état fondamental) à une orbitale de niveau supérieur (état excité). L'énergie nécessaire à une excitation électronique doit correspondre à la différence d'énergie entre l'énergie de l'état fondamental et celle de l'état excité (figure 3). Ces transitions nécessitent des énergies importantes et interviennent sous formes de bandes assez larges, dans l'ultraviolet, le visible et une infime partie du proche infrarouge (Escadafal, 1989 ; Visacarra et al., 2007a).

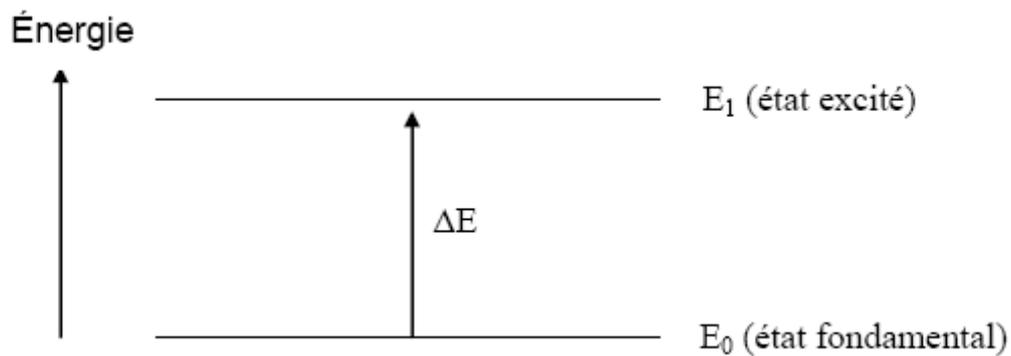


Figure 3. Absorption de l'énergie par excitation électronique.

Dans le cas des échantillons de sol l'absorption électronique est le résultat des transitions des ions métalliques contenus dans les minéraux d'un niveau d'énergie à un autre. Pour l'essentiel, ces phénomènes sont produits par les ions ferreux Fe²⁺ et ferrique Fe³⁺ (Clark, 1999). Ces transferts de charges du fer ferrique débordent des bandes ultraviolet (UV) vers le visible en dessous de 600 nm (Hunt et *al.*, 1971a). Les absorptions électroniques sont souvent la cause de la couleur des minéraux.

2.2. Absorption vibrationnelle (dans le PIR)

Les vibrations des liaisons moléculaires conduisent à des bandes d'absorption dans le domaine de l'infrarouge. Ces bandes d'absorption proviennent de l'interaction entre la composante électrique du rayonnement électromagnétique et les dipôles électriques oscillants des liaisons non symétriques, c'est à dire des liaisons joignant deux atomes distincts (Banwell, 1983). Les liaisons non polaires sont donc « transparentes » dans l'Infrarouge (IR). La mécanique quantique sur la base d'un modèle d'oscillateur anharmonique permet de calculer les niveaux d'énergie vibrationnelle et d'établir les règles de sélection définissant les transitions autorisées entre les différents niveaux de vibration (Bertrand, 2002). Ce modèle est proche de la réalité observée même si le calcul des niveaux d'énergie reste assez complexe et nécessite de faire des approximations. Le lecteur qui souhaite approfondir le sujet peut se reporter à l'article de Lachenal (2006).

3. Modes de vibration d'une molécule polyatomique

3.1. Bandes fondamentales

Les bandes fondamentales correspondent aux transitions entre le niveau vibrationnel de plus basse énergie (état fondamental ou non excité) et le premier état excité. Ces transitions sont observées pour des longueurs d'onde qui sont dans le domaine du moyen infrarouge (MIR).

On dispose pour une molécule à N atomes de $3N$ degrés de liberté puisque chaque atome peut être repéré à l'aide des trois coordonnées cartésiennes (x, y, z), par exemple. La molécule pouvant se déplacer et se déformer dans l'espace, les $3N$ degrés de liberté peuvent se diviser en trois groupes : la translation, la vibration et la rotation. La translation utilise 3 parmi les $3N$ degrés de liberté ainsi que la rotation dans le cas d'une molécule non linéaire. Il en résulte qu'une telle molécule possède $3N-6$ vibrations fondamentales. Par contre, la rotation d'une molécule linéaire ne nécessite que 2 degrés de libertés ce qui dans ce cas conduit à $3N-5$ vibrations fondamentales.

On distingue deux groupes de vibrations, les vibrations de valence (ou d'élongation) et les vibrations de déformation qui peuvent se produire dans le plan ou en dehors du plan (Umland et Bellama, 1996). La figure 4 montre les 3 vibrations fondamentales de la molécule d'eau, à l'état gazeux, ainsi que les nombres d'onde auxquels elles se produisent. Les deux vibrations d'élongation produisent un changement de la longueur de liaison alors que la vibration de déformation modifie l'angle de liaison. Les déformations des molécules non linéaires constituées de trois atomes se font dans le plan. Par contre, pour les molécules d'ordre supérieur ($N > 3$) des déformations hors du plan de la molécule apparaissent. La figure 5 donne un exemple des différents modes de vibrations que l'on peut observer dans le cas d'une molécule constituée de plus de 3 atomes. Le nombre de vibrations fondamentales augmente avec N et on voit apparaître des déformations hors du plan. Dans cette figure 5 sont rappelés les différents modes de vibration observés dans le cas d'une molécule à 4 atomes.

L'intensité des bandes de vibration dépend de la nature et de la polarité de la liaison. Ainsi, une liaison fortement polaire et composée d'atomes différents, absorbe fortement en infrarouge moyen. Par ailleurs, et en règle générale, les liaisons entre atomes légers vibrent à plus haute fréquence que les liaisons entre atomes lourds. Par exemple la liaison C-H vibre à 3000 cm^{-1} alors que la liaison C-O vibre à 1100 cm^{-1} . En outre, les mouvements d'élongation sont plus énergétiques que les mouvements de déformations.

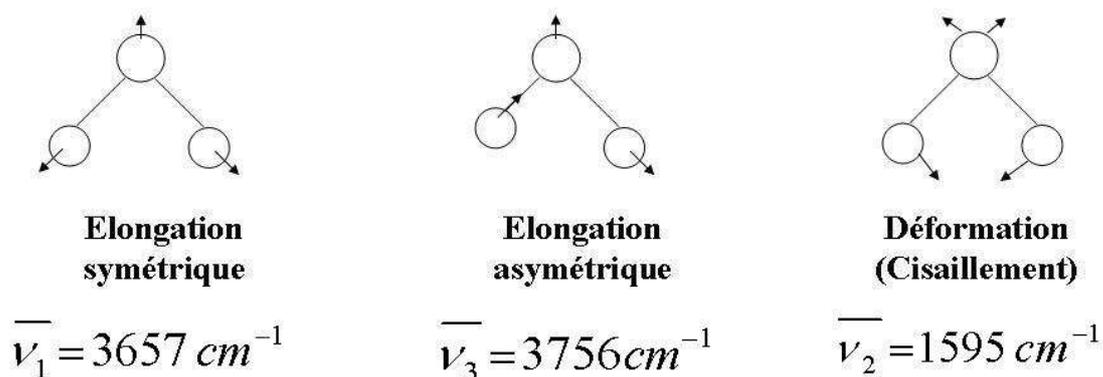


Figure 4. Bandes d'absorption fondamentales d'une molécule triatomiques non linéaire, exemple de la molécule d'eau (H₂O) (Lachenal, 2006).

3.2. Bandes harmoniques et bandes de combinaisons

Ces deux types de bandes apparaissent aussi bien dans le Moyen Infrarouge (MIR) que dans le (PIR). Cependant, ces bandes ont généralement une faible intensité et se trouve masqués, dans le MIR, par les bandes fondamentales. C'est la raison pour laquelle ces bandes sont d'un grand intérêt pour des applications analytiques dans le PIR.

Les absorptions vibrationnelles nécessitent des énergies moins importantes que celles des transitions électroniques et interviennent sous formes de bandes plus étroites au-delà de 2500 nm, mais dont on peut observer des harmoniques à partir de 1300 nm (Escadafal, 1989 ; Clark 1999 ; Visacarra et *al.*, 2007a).

Les bandes harmoniques correspondent à des absorptions vibrationnelles qui se produisent à des fréquences qui sont un multiple entier des fréquences des bandes fondamentales telles que $2\nu_1$, $2\nu_2$, $2\nu_3$ ou $3\nu_1$, $3\nu_2$, $3\nu_3$, *etc.*

Les bandes de combinaisons résultent de l'interaction de deux ou plusieurs modes de vibration pour un même groupe fonctionnel. Dans ces transitions, un photon excite deux ou plusieurs vibrations simultanément. Un photon pour induire ce type de transition, devra avoir une fréquence qui est une combinaison linéaire de deux ou plusieurs fréquences fondamentales et harmoniques, telles que $\nu_1+\nu_2$, $2\nu_1+\nu_2$, $2\nu_1-\nu_2$ ou $\nu_1+\nu_2+2\nu_3$, *etc* (Clark, 1999).

L'intensité des bandes harmoniques dépend de l'anharmonicité et les effets d'anharmonicité d'une liaison sont d'autant plus importants que les déplacements entre les deux atomes sont grands. On observe de grandes amplitudes de la vibration pour des groupements du type H-X où H est un

atome léger tel que l'hydrogène et X un atome plus lourd tel que C, N, O, etc. Dans un groupement contenant un atome d'hydrogène ce dernier va donc avoir un déplacement plus important qu'un atome plus lourd, ce qui est à l'origine d'une plus grande anharmonicité et d'une plus grande intensité de la bande correspondante. Le domaine spectral du PIR contient les premières harmoniques des vibrations des liaisons telles que C-H, O-H et N-H (Weyer et Lo, 2002).

Par ailleurs, l'environnement de la molécule va non seulement influencer la fréquence à laquelle se produit la vibration (observation d'un écart par rapport à la fréquence théorique) mais va aussi avoir un effet sur la largeur de la bande d'absorption. En effet, la largeur d'une raie d'absorption est inversement proportionnelle à la durée de vie de l'état excité. Dans le cas d'un gaz cette durée de vie est grande et les raies sont donc très fines. Par contre, dans les cas des liquides et des solides cette durée de vie est brève en raison des interactions entre molécules ou atomes voisins, ce qui a pour conséquence un élargissement des bandes de vibration. En outre, l'élargissement des bandes vibrationnelles est plus important dans le PIR que dans le MIR. Ainsi, on peut comprendre aisément que dans le cas d'un solide, par exemple, cet élargissement peut conduire à un recouvrement des bandes d'absorption donnant lieu à une bande unique, enveloppe de nombreuses raies. Et dans le cas d'un matériau solide complexe, tel que le sol, qui peut avoir des éléments ayant des fréquences de vibrations voisines, le recouvrement des bandes d'absorptions complique singulièrement l'exploitation qualitative des spectres dans un objectif d'identification.

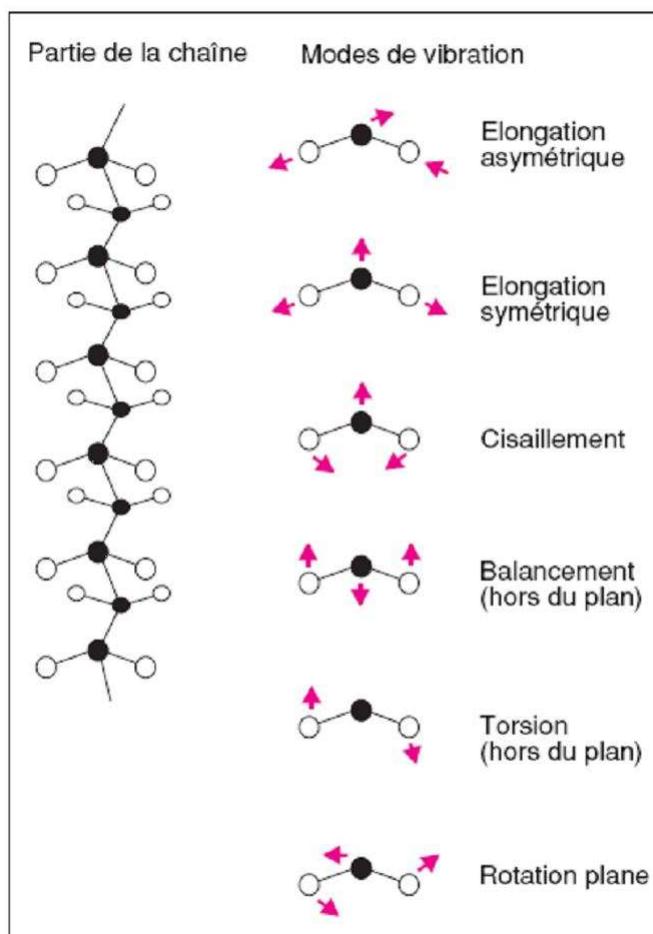


Figure 5. Exemples de vibrations atomiques : cas d'une chaîne hydrocarbonée (rond noir : atome carbone, rond blanc : atome d'hydrogène) (d'après Bertrand, 2002).

4. Réflexion diffuse

Lorsqu'une radiation électromagnétique entre en contact avec de la matière, une première partie de la radiation est absorbée, une seconde est réfléchiée et une troisième est transmise. En vertu du principe de conservation de l'énergie, l'intensité de la radiation incidente (I_0) est telle que :

$$I_0 = I_a + I_r + I_t \quad [4]$$

Où I_a, I_r et I_t sont respectivement les intensités absorbée, réfléchiée et transmise. Ce principe appliqué en spectrométrie permet de calculer la fraction de l'énergie lumineuse absorbée par les molécules à partir des résultats de mesure I_t , ou de I_r (Osborne et *al.*, 1993).

Lorsque la lumière est réfléchiée par un échantillon de matière solide, tel que le sol (Workman et Shenk, 2004), on peut distinguer deux composantes de l'énergie lumineuse réfléchiée (Figure 6). Une composante spéculaire correspondant à la lumière réfléchiée directement en surface de l'échantillon sans pénétrer à l'intérieur de celui-ci. Une composante de réflexion diffuse correspond à la lumière qui pénètre à l'intérieur de l'échantillon, y subit l'absorption moléculaire et aussi le processus de réfraction, diffraction et réflexion multiple entraînant sa dispersion dans toutes les directions et notamment son retour en arrière vers le détecteur de réflexion. Seule la composante diffuse est porteuse d'information sur la composition chimique de l'échantillon.

Le rayonnement qui arrive sur un sol ne rencontre généralement pas une surface plate et lisse, mais une surface rugueuse et irrégulière (Jaquemoud et *al.*, 1993), où la lumière incidente est redirigée selon plusieurs angles après avoir pénétré à travers le matériau (Stewart, 1998). La spectroscopie par réflexion diffuse mesure la lumière réfléchiée, rétro diffusée et dispersée vers tous les angles d'une surface. Lorsque le spectre est mesuré en mode réflectance, les pics d'absorption sont perçus comme des sinuosités au niveau de la courbe de réflectance spectrale. Ainsi, la réflectance diffuse fournit des informations sur la composition moléculaire du matériau objet de l'étude et notamment du sol.

La réflexion spéculaire peut distordre les mesures spectrales portant sur la réflectance diffuse. Les bandes fondamentales les plus intenses sont typiquement beaucoup plus affectées par la réflexion spéculaire que les faibles bandes des harmoniques et des combinaisons qui leur sont liées. C'est pour cette raison que la réflexion spéculaire est négligeable dans le domaine du Vis-PIR, alors qu'elle est importante dans la région du MIR (Ollinger et Griffiths, 1993).

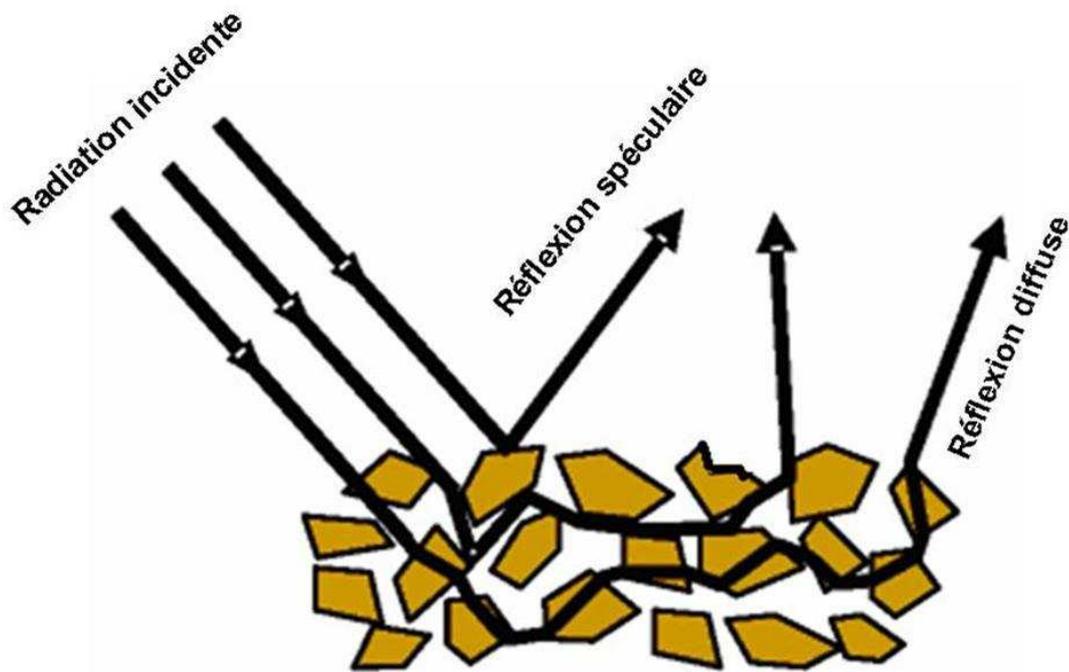


Figure 6. Formes de réflexions occasionnées par un faisceau de radiation incidente sur un sol (d'après Viscarra, 2007a).

5. Réflectance spectrale

La réflectance spectrale exprime le pourcentage de lumière incidente qui est réfléchi, à différentes longueurs d'onde du spectre, par la surface d'un matériau. Elle est définie comme le rapport des mesures de luminances ($W/m^2/sr/nm$) du matériau étudié sur celle d'un panneau blanc de référence (Mutuo et al., 2006). La réflectance est unique pour les objets naturels ou artificiels contenant un seul matériau. La réflectance spectrale est donc un moyen idéal pour l'identification des matériaux, et c'est aussi l'objectif principal de l'application et du développement de la télédétection. Puisque la réflectance de chaque matériau est unique, certains chercheurs préfèrent l'appeler empreinte spectrale ou signature spectrale (Escadafal, 1989). Les signatures spectrales comportent généralement les détails et les informations nécessaires pour qualifier et quantifier certains constituants des matériaux. La signature spectrale peut être considérée comme une fonction continue, de la variable longueur d'onde, qui est mesurable en conditions contrôlées de laboratoire ou directement sur le terrain (Homayouni, 2005 ; Sudduth et Hummel, 1996).

6. Acquisition des spectres

En conditions contrôlées de laboratoire les mesures spectrométriques sont le plus souvent réalisées sur des échantillons de sol remaniés séchés, tamisés à 2 mm puis placés dans des boîtes de Petri. Les mesures de réflectance spectrale sont réalisées à l'aide de spectroradiomètres portables (à usage de laboratoire et de terrain). En fonction des spécifications de l'instrument, ces mesures couvrent généralement le domaine du visible et du proche infrarouge 400-2500 nm ou peuvent être limitées à une fenêtre de ce domaine. Comme schématisé dans la figure 7, L'acquisition des spectres est réalisée, le plus souvent en visée verticale, dans une chambre noire avec des conditions d'illumination standardisées (Deux lampes halogènes, sont fixées de part et d'autre du pistolet du capteur à la même hauteur que celui-ci. Les lampes sont orientées de façon à avoir le maximum d'énergie incidente. Les spectroradiomètres ont des champs de visée réglable. La hauteur à laquelle le capteur devrait être maintenu, est calculée de façon à couvrir toute la surface de l'échantillon à scanner. La réflectance apparente des échantillons est obtenue en mesurant successivement la luminance du panneau blanc de référence, puis celle de l'échantillon. Le panneau blanc de référence est une surface à diffusion lambertienne (ou isotrope) généralement en sulfate de baryum (BaSO_4) dont la réflectance est supérieure à 0.99 dans toute la gamme de longueurs d'onde du capteur.

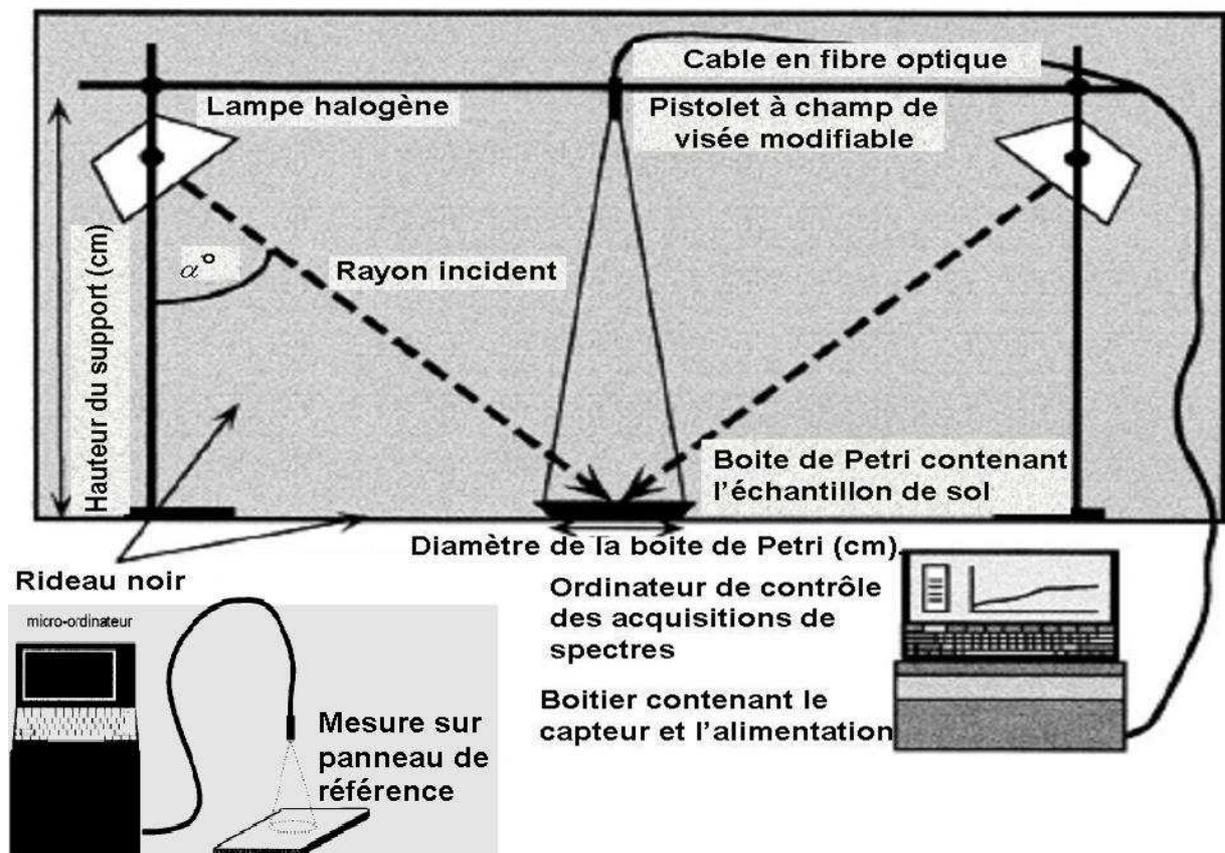


Figure 7. Dispositif expérimental de mesure de la réflectance en conditions contrôlées de laboratoire, d'après (Escadafal et al., 1993 ; Mathieu et al., 1998).

La photo 1 montre des mesures directes, sur le terrain, de la luminance des surfaces du sol (a, b, c, d et e) et de la luminance du panneau de référence (f), en condition d'illumination naturelle. Tout comme au laboratoire, les mesures sur l'échantillon de sol et sur le panneau de référence doivent être faites successivement dans les mêmes conditions de visée verticale, autour du midi solaire (moment de la journée où le soleil est au Zénith). Il est préconisé de faire les mesures par temps clair, où l'éclairement solaire peut être considéré comme stable pendant la durée de la séquence de mesures.

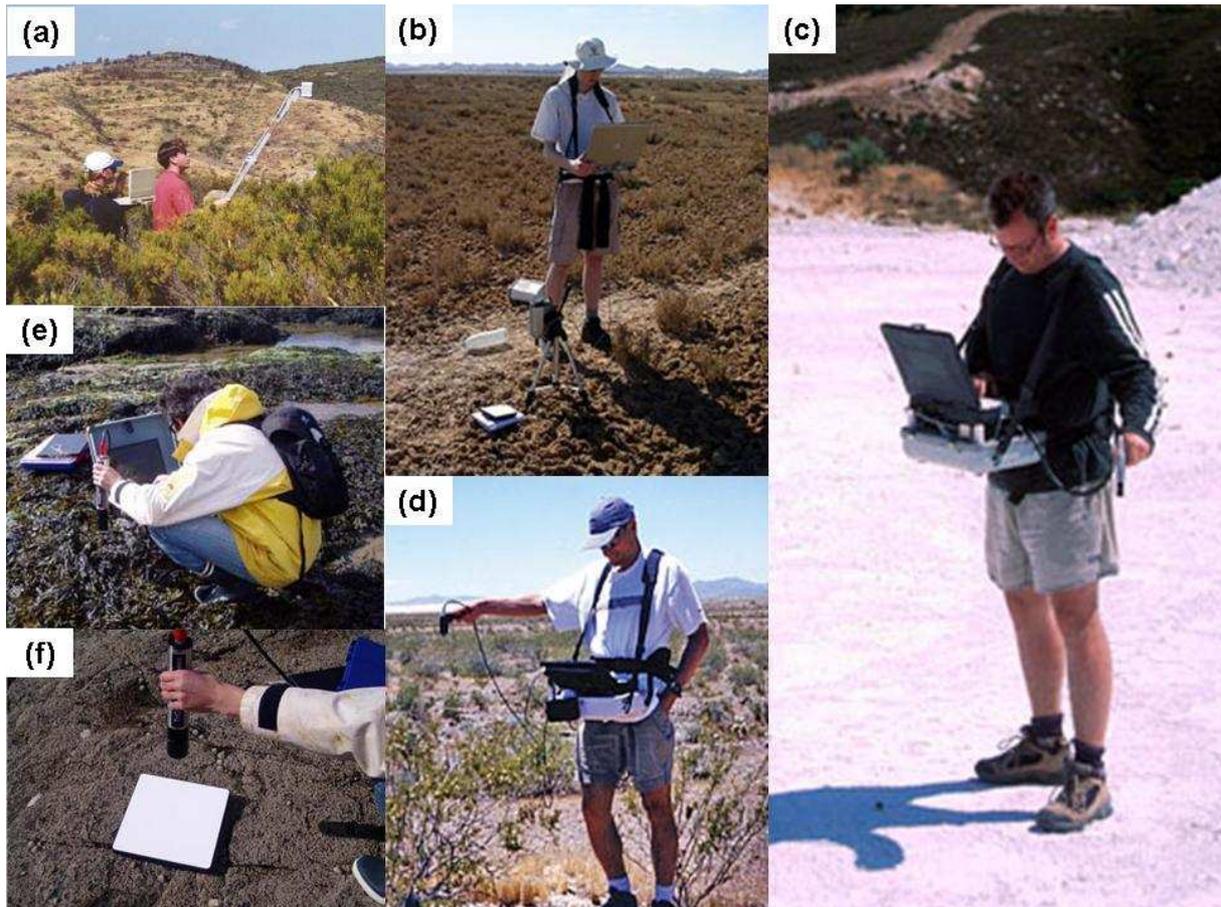


Photo 1. Mesures directes, sur le terrain, de la luminance des surfaces du sol (a, b, c, d et e) et de la luminance du panneau de référence (f), en condition d'illumination naturelle.

7. Spectroscopie qualitative

La mesure et l'analyse des signatures spectrales pour différents types et catégories de matériaux nous permettent de constituer des bibliothèques spectrales. Une bibliothèque spectrale est une collection de réflectance spectrales de matériaux courants. Grâce à des recherches dans le domaine de la spectrométrie et dans le cadre de différents projets,

différentes bibliothèques ont été constituées et sont actuellement accessibles. Certaines d'entre elles, reconnues pour leur qualité, sont publiquement disponibles et sont référencées dans les articles suivants : Clark *et al.*, 1993 ; Grove *et al.*, 1992 ; Elvige, 1990 ; Korb *et al.*, 1996 ; Salisbury *et al.*, 1991 ; Ben-Dor, 2003. Elles comportent généralement une grande variété de matériaux naturels (minéraux et végétaux) ou artificiels (Shippert, 2004). Les spectres fournis par ces bibliothèques peuvent être utilisés pour discerner qualitativement entre les objets étudiés ainsi que dans l'interprétation des informations des images satellites. La figure 8 montre des courbes typiques de réflectances de trois matériaux distincts : l'eau, la végétation et le sol (d'après Lillesand et Kiefer, 1987).

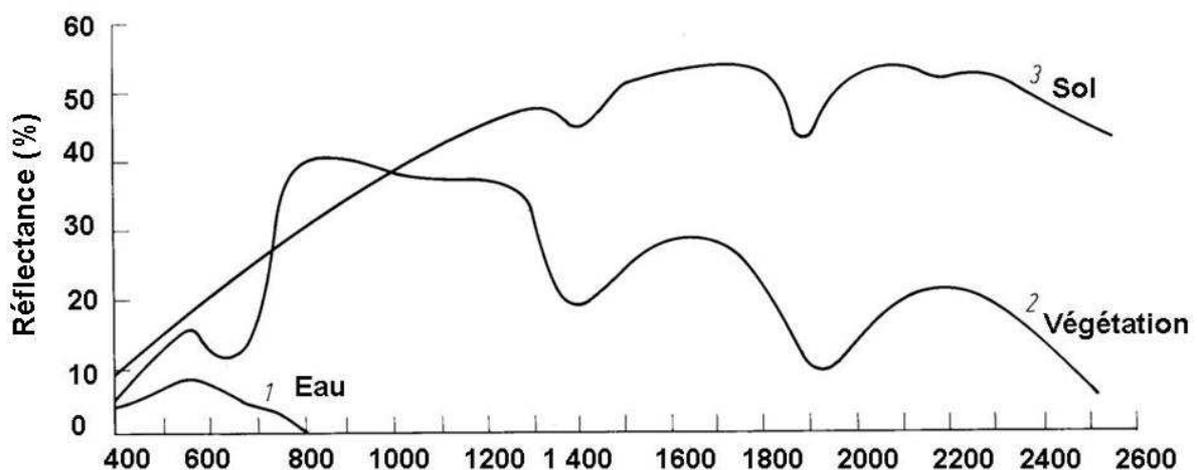


Figure 8. Courbes typiques des comportements spectraux de trois matériaux distincts : l'eau, la végétation et le sol (d'après Lillesand et Kiefer, 1987).

8. Spectroscopie quantitative

Rappelons, tout d'abord, que les bandes d'absorption observées sur des spectres dans le PIR sont dues, essentiellement, aux harmoniques des vibrations d'élongation de l'hydrogène ou une combinaison faisant intervenir les modes de vibration d'élongation et de déformation. Ces bandes sont donc plus larges dans le PIR que dans le MIR et les spectres sont plus complexes. C'est du fait de cette complexité que la spectroscopie PIR a tout de suite tiré profit des techniques de calibration sophistiquées existantes (Chaminade *et al.*, 1998).

L'efficacité de la spectroscopie dépend donc en partie des méthodes statistiques utilisées pour les analyses quantitatives. Le Vis-PIR apparaît a priori comme un mauvais choix pour des analyses quantitatives des composés organo-minéraux, notamment ceux du sol, en raison du grand chevauchement des bandes harmoniques et de combinaisons qui sont

associées aux liaisons C-H, N-H et O-H. Ce recouvrement entraîne une complexité spectrale qui défie la calibration univariée basée sur la sélection d'un seul pic d'absorption.

Cependant, la région du PIR est bien adaptée au développement des applications analytiques. En effet, les harmoniques et les bandes de combinaisons sont moins intenses que les fondamentales correspondantes. Par conséquent, la lumière est moins fortement absorbée dans le PIR que dans le MIR. Elle pénètre de plusieurs millimètres à l'intérieur du produit étudié et il n'est pas alors nécessaire de le diluer pour obtenir des spectres exploitables (Steward, 2006).

En spectroscopie Vis-PIR, assigner des pics d'absorption aux liaisons chimiques ou à certains minéraux peut être fait pour des minéraux purs ou des mélanges purs, mais cela apparaît difficile pour un sol naturel. Cette difficulté résulte du chevauchement des pics et des changements qui peuvent se produire à cause des multiples interactions des liaisons organique et minérale du sol (Travis, 2006). La plupart des bandes d'absorption des composés organiques peuvent être chevauchées par des bandes d'absorptions plus fortes qui sont dues à l'eau et aux composés minéraux, ce qui explique la difficulté de l'interprétation visuelle des spectres des sols acquis dans le domaine du Vis-PIR (Steward, 2006). Et c'est seulement le recours aux techniques chimiométriques qui permet de contourner le problème de la complexité des spectres et de mettre en relation certaines propriétés du sol avec la réponse spectrale.

Selon Geladi (1995), la chimiométrie (*chemometrics*, en anglais) est la « science de l'utilisation des méthodes mathématiques, statistiques et informatiques dans le but d'améliorer l'extraction des informations obtenues à partir des données analytiques ». En spectroscopie, le processus fait référence aux traitements des données spectrales pour prédire les propriétés analytiques (Duckworth, 1998).

Nous pouvons classer les méthodes chimiométriques en fonction de leur nature linéaire ou non linéaire. Les méthodes linéaires font l'hypothèse que l'information utile peut être extraite à partir de combinaisons linéaires des variables prédictives. Les méthodes non linéaires appartiennent au domaine de l'intelligence artificielle et mettent principalement en jeu les réseaux de neurones.

Dans sa forme la plus simple, un modèle linéaire spécifie la relation (linéaire) entre une (ou des) variables à expliquer (réponses) Y et un ensemble de variables explicatives X_i . Le modèle de régression est donné par l'équation (5) :

$$y^* = \sum_i^p b_i X_i + \epsilon_0 \quad [5]$$

Où y^* : la valeur estimée, b_i : le coefficients du modèle à la longueur d'onde i , X_i : la réflectance à la longueur d'onde i et ϵ_0 : erreur résiduelle.

8.1. Régression linéaire multiple

En spectroscopie quantitative appliquée à l'étude des sols, les premiers modèles de prédiction établis sont des régressions linéaires multiples. Ils se sont basés sur une présélection empirique de quelques longueurs d'onde du spectre comme variables explicatives des propriétés analytiques (Al-Abbas *et al.*, 1972 ; Krishnan *et al.*, 1980 ; Dalal et Henry, 1986). La complexité du spectre et les chevauchements des bandes d'absorption rendent ce travail de présélection trop hasardeux. Par ailleurs, le nombre élevé de variables explicatives et la forte corrélation qui existe entre elles rendent cette approche, par régression linéaire multiple, inappropriée dans les études spectroscopiques (Abdi, 2003).

Depuis lors, des approches alternatives comme la Régression sur Composantes Principales, RCP, (Chang *et al.*, 2001 ; Islam *et al.*, 2003), et la Régression en Moindres Carrés Partiels, RMCP, (Janik *et al.*, 1998 ; Reeves *et al.*, 1999 ; Rumpel *et al.*, 2001 ; Dunn *et al.*, 2002 ; Reeves *et al.*, 2002 ; Chang et Laird, 2002 ; Lee *et al.*, 2003 ; Viscarra *et al.*, 2006a), permettant la manipulation de plusieurs centaines de longueurs d'onde, ont été utilisées.

8.2. Régression en composantes principales

La régression en composantes principales (RCP) est une procédure en deux étapes. En premier lieu, elle réduit le nombre de composantes indépendantes nécessaires pour la description des variations le long du même spectre et entre les spectres des différents échantillons. Cette technique permet de réduire les milliers de variables spectrales en quelques composantes principales orthogonales, ce qui supprime la redondance qui caractérise l'information spectrale (Wold *et al.*, 2001 ; Workman et Shenk, 2004). Ensuite, ces composantes principales font l'objet d'une analyse corrélative avec les propriétés analytiques (concentrations mesurées) pour construire les modèles de calibration. La RCP a été l'une des

premières méthodes à être utilisée en conservant toutes les données du spectre dans la prédiction (Chang *et al.*, 2001).

8.3. Régression en moindres carrés partiels

8.3.1. Historique

La régression en moindres carrés partiels (RMCP) est une procédure de régression bilinéaire (Misun, 2002) qui a été développée par Herman Wold dans les années 60, comme technique économétrique (Wold, 1966 ; Desbois, 1999). Elle fut ensuite popularisée grâce à son utilisation en chimie, dans des applications de chromatographie et de spectrographie (Davies, 2001). En effet, en 1966, Herman Wold publie un ouvrage introduisant la régression en moindres carrés partiels (Wold, 1966), alors appelée *NonLinear Iterative Partial Least Squares, NIPALS*. En 1983, Svante Wold et Harold Mertens adaptent NIPALS aux problèmes de régression avec un nombre élevé de variables explicatives (Wold *et al.*, 1983) et appellent PLS cette adaptation de l'algorithme (Vancolen, 2004).

8.3.2. Principe

La RMCP est une technique quantitative de décomposition spectrale étroitement liée à la RCP. Cependant, dans la RMCP, la décomposition est faite d'une manière légèrement différente. Au lieu de décomposer d'abord la matrice spectrale X en un ensemble de vecteurs propres et de scores, et de les régresser contre les Y dans une étape séparée, la RMCP utilise l'information de Y dès le processus de décomposition. Ceci implique que les variables expliquant le mieux les Y seront plus fortement pondérées (Duckworth, 1998). L'idée principale de la RMCP est de donner le plus d'information possible sur Y dans les premières composantes latentes construites. La méthode demande souvent moins de composantes que la RCP (Wold *et al.*, 2001 ; Vancolen, 2004). En outre la RMCP permet de construire des modèles prédictifs lorsque les variables explicatives sont nombreuses et très corrélées (Vancolen, 2004 ; Desbois, 1999), ce qui est le cas en spectroscopie (Sundberg, 1999 ; Viscarra, 2007b). En régression MCP, toutes les variables sont conservées. Les variables importantes dans la prédiction ont des poids élevés et les variables sans importance participent au modèle, mais avec un poids faible (Martens et Næs, 1989 ; Tenenhaus, 1998). En RMCP, le nombre de variables peut être supérieur au nombre d'échantillons du jeu de calibration (Tenenhaus, 1998). Ces qualités expliquent pourquoi, dans ce travail, nous avons fait le choix de mettre en oeuvre cette technique RMCP.

8.3.3. Choix du nombre de composantes latentes du modèle RMCP

Pour un même jeu de données à N échantillons, l'algorithme de la régression MCP permet de construire des modèles de calibration dont le nombre de composantes latentes est compris entre $[1, N-1]$. Il est alors nécessaire d'établir des critères de choix du nombre de composantes latentes du modèle à retenir. Selon le cas l'un de ces trois critères de choix arbitraire peut être utilisé (Vancolen, 2004) :

Critère 1 : Ajustement ou « fit » du modèle sur les données

Critère 2 : Validation interne ou « croisée » sur le même jeu de données

Critère 3 : Validation externe sur un jeu de données test

8.3.3.1. Critère d'ajustement ou « fit »

Pour éviter un surajustement du modèle au jeu de calibration, le modèle retenu ne devra pas avoir un nombre de composantes latentes trop grand, car cela dégradera la prédiction. Donc, il faut un équilibre entre le critère « ajustement » et le critère « prédiction » (interne ou externe). Lors de l'exécution de l'algorithme à chaque nouvel axe (i.e. variable latente) construit, on recalcule la variance des prédictors (X) et des réponses (Y). La stratégie consiste à s'arrêter lorsque le pourcentage de la variance de X est suffisamment grand, pour un gain faible dans le pourcentage de variance de Y . Ce critère est donc basé sur l'ajustement des données et permet de construire un modèle linéaire de dimension n , où n est le nombre de variables latentes retenues pour le modèle.

En termes d'ajustement, il est aussi intéressant de regarder l'évolution des coefficients du modèle de régression linéaire en fonction de la valeur de (n) :

$$\hat{Y}_n = X \hat{\beta}_n \quad [6]$$

Pour une réponse j choisie, l'étendue des valeurs du vecteur $\hat{\beta}_n^j$ reste stable jusqu'à une certaine valeur de n . Au delà de cette valeur n , pour certains prédictors fortement corrélés, les valeurs des coefficients $\hat{\beta}_n^j$ semblent changer de manière drastique. On choisit alors la plus petite valeur n à partir de laquelle ce phénomène apparaît.

8.3.3.2. Critère de « validation croisée »

Après avoir enlevé un individu i (une ligne) aux matrices X et Y (ou 10 % des individus si leur nombre est élevé), on calcule la matrice des coefficients du modèle construit avec les individus restants. Puis, on calcule l'erreur de prédiction faite sur l'individu i . On choisit n (nombre de variables latentes) tel que l'erreur de prédiction soit la plus petite possible.

8.3.3.3. Critère de « validation externe »

Pour utiliser ce critère de choix de n , nous devons posséder, outre le jeu d'échantillons d'apprentissages (X, Y) , un jeu d'échantillons test (X_{test}, Y_{test}) . Différents modèles RMCP basés sur (X, Y) sont calculés pour différentes dimensions. La valeur optimale de n correspond à la plus petite des erreurs de prédictions calculées sur (X_{test}, Y_{test}) .

9. Prétraitements des spectres

En chimiométrie, avant l'ajustement du modèle, nous pouvons avoir recours à une large panoplie de prétraitements mathématiques des données spectrométriques qui s'avèrent parfois, nécessaires. Ces prétraitements ont pour objectif l'amélioration de l'ajustement des modèles de prédiction et d'en assurer la robustesse (Viscarra, 2008a). Le prétraitement des spectres fait donc intervenir divers algorithmes qui permettent de corriger deux problèmes qui peuvent apparaître au niveau du signal acquis. Ces problèmes surviennent soit du fait de la non-linéarité de la réponse spectrale, soit d'un signal bruité à cause d'un rapport signal sur bruit faible (Stevens, 2008). Il n'est évidemment pas exclu que ces deux problèmes puisse apparaître simultanément. La non-linéarité de la réponse spectrale est due aux variations de la diffusion de la lumière et du chemin optique à travers le matériau. Classiquement, pour réduire cet effet de non-linéarité la réflectance R est transformée en absorbance A selon l'équation (7) :

$$A = \log(1/R) \quad [7]$$

Deux algorithmes sont également utilisés, en spectrochimométries par réflectance, à cet effet, l'un connu sous le nom de MSC, pour Multiplicative Signal Correction (Geladi et al., 1985), est utilisé pour corriger les variations de la diffusion de la lumière. L'autre algorithme, utilisé pour éliminer les interférences qui sont dues aux variations de la diffusion de la lumière et du chemin optique, est connu sous le nom de transformation SNV, Standard Normal Variate (Barnes et al., 1989). Dans le cas d'un spectre bruité un filtre de Savitsky-Golay (Savitsky et Golay, 1964) peut être appliqué pour lisser le signal. Des filtres un peu plus sophistiqués basés sur la transformation en ondelettes ont été également utilisés (Viscarra, 2008a).

D'autres prétraitements peuvent être mis en œuvre et nous pouvons citer, en particulier, la dérivée première et seconde du spectre. Ce traitement permet dans certains cas de mieux faire apparaître des changements subtils dans les spectres et à mieux faire ressortir d'éventuels pics d'absorption. En outre, la dérivée 1^{ère} est également utilisée pour corriger le décalage de la ligne de base (ligne qui joint les minima successifs du spectre). En effet, des échantillons d'un même type de sol qui ne se différencient que par la taille de leurs particules auront des spectres

qui seront décalées mais de formes identiques. Ainsi, la dérivée 1^{ère} de ces spectres va donc conduire à une superposition des signaux transformés. Cette transformation va donc permettre de faire ressortir les changements dans les spectres liés à des propriétés du sol autre que la texture.

10. Application de la spectroscopie par réflexion diffuse dans l'étude des sols

Une bonne caractérisation des propriétés optiques des sols est nécessaire pour l'étude des surfaces naturelles ou cultivées (Jaquemoud *et al.*, 1993). Dans un premier temps, les spectres des sols ont été utilisés pour discriminer les sols d'une façon qualitative. Ainsi Obukhov et Orlov (1964) ont défini trois groupes de types de comportement spectraux pour les sols dans la gamme spectrale 400-800 nm. Ils distinguaient :

- les spectres de sol croissant et monotone,
- les spectres à faible fraction réfléchi et à forme concave,
- les spectres à formes sigmoïdale qui traduisent une forte absorption dans le bleu et une forte réflexion dans le vert et le rouge.

De même Condit (1972) a visuellement classifié 160 réponses spectrales d'échantillons de sols en trois groupes de comportements spectraux qui sont similaires à ceux définis par Obukhov et Orlov (1964). Stoner et Baumgardner (1981) ont défini, en travaillant sur un domaine spectral beaucoup plus large 520- 2320 nm, cinq classes de spectres types caractéristiques de la plupart des sols des Etats-Unis.

Par ailleurs, Kasumov *et al.* (1992) signalaient que la réflectance spectrale d'un sol décroît sur tout le domaine du visible, lorsque la teneur en matière organique augmente. Baumgardner *et al.* (1985) puis Irons *et al.* (1989) ont recensé les principaux constituants du sol affectant la réflectance : l'humidité, les minéraux (argiles, oxydes et hydroxydes de fer, carbonates, sulfates, *etc*), la matière organique (humus et substances non humiques), et l'air forment des milieux extrêmement variés, dont les propriétés optiques sont en conséquence presque infinies.

11. Facteurs influençant la réponse spectrale des sols

Les sols sont formés à partir des produits, relativement peu nombreux, de l'altération des roches ; à noter que, plus les sols, sont altérés moins leurs constituants sont nombreux. Les sols sont ainsi constitués de particules de minéraux argileux, de quartz, d'oxydes et hydroxydes de fer, de matières organiques, de carbonates, de feldspaths, *etc*. Ces particules sont assemblées en agrégats de tailles très variables.

Les propriétés spectrales des sols dépendent des propriétés spectrales intrinsèques des matériaux pédologiques qui les constituent (minéraux, matière organique, eau) (Escadafal, 1989, Steward, 2006), de la taille des agrégats (rugosité de sa surface) et de la configuration géométrique de mesure (angle d'incidence et de visée) (Girard et Girard, 1999 ; Workman et Shenk, 2004).

11.1. Humidité

Un sol humide apparaît toujours plus foncé qu'un sol sec (Bedidi *et al.*, 1992 ; Houssa *et al.*, 1996 ; Muller et Décamps, 2000). Ceci s'explique par l'effet de diffusion aux interfaces eau - particules du sol, l'eau agit comme un piège et absorbe le rayonnement incident (Fourty et Baret, 1998). Les harmoniques de l'eau apparaissent à 1200, 1450, 1770 et 1950 nm (Baumgardner *et al.*, 1985). Clark *et al.* (1990) ont montré que les vibrations fondamentales de l'eau se trouvent à 6269, 2738, et 2662 nm avec des harmoniques de l'élongation de la liaison OH au voisinage de 1400 nm et la combinaison de la torsion de la molécule H-O-H avec l'élongation de O-H qui se produit au voisinage de 1900 nm (Clark, 1999).

L'augmentation de la teneur en eau se traduit par une diminution générale de la réflectance à toutes les longueurs d'onde, et l'apparition des bandes d'absorption caractéristiques à 1400 et 1900 nm. Comme l'illustre la figure 9 dans le cas d'un sol limoneux (D'après BOWERS et HANK, 1965).

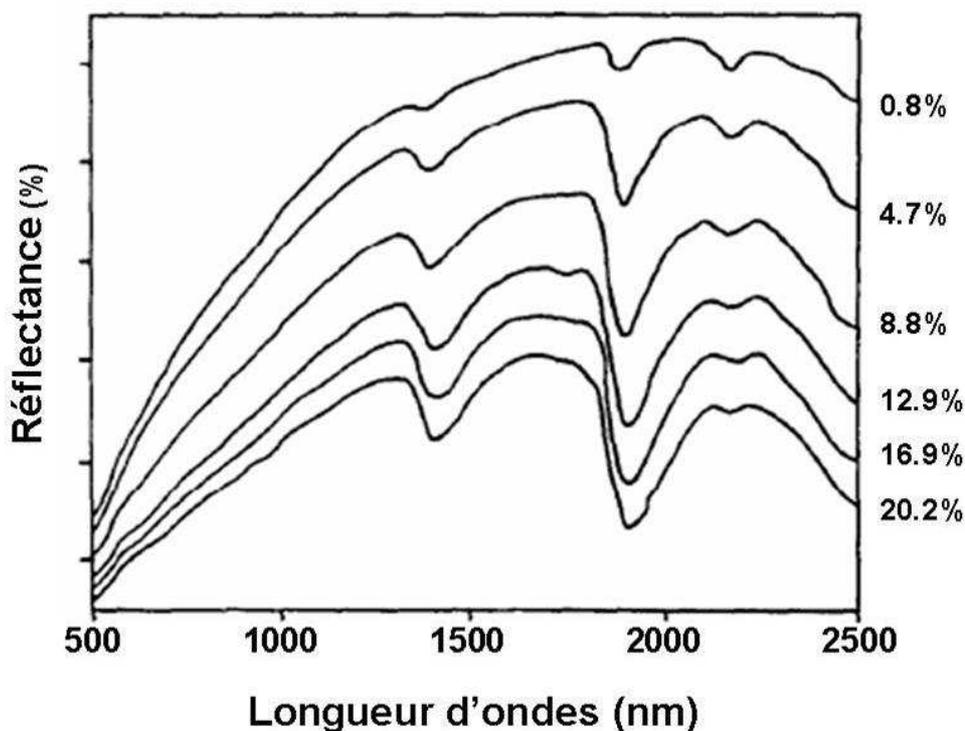


Figure 9. Influence de la teneur en eau sur la réflectance d'un sol limoneux (d'après Bowers et Hank, 1965).

11.2. Rugosité de la surface

Lorsque le *sol* est *rugueux*, la fraction de la radiation diffuse augmente au détriment de la fraction de la radiation réfléchie, ce qui diminue la réflexion (Girard et Girard, 1999).

11.3. Matière organique

La matière organique du sol est une mixture de résidus animal et végétal à divers stade de décomposition, de substances microbiologiquement ou chimiquement néo-synthétisées à partir des résidus et des microorganismes du sol (Shnitzer *et al.*, 1991). Il s'agit donc de mélanges complexes de glucides, peptides, lipides, cires, résines, acides fulviques, acides humiques et humines (Escadafal, 1989). En général, pour tout le spectre du visible et du proche infrarouge la réflectance diminue lorsque les teneurs en matière organique augmente (Stoner et Baumgardner, 1981 ; Galvão *et al.*, 2001 ; Kooistra *et al.*, 2003). Les liaisons chimiques de la matière organique spectralement actives dans le spectre PIR sont les groupements O-H, C-N, N-H, et C=O (Malley *et al.*, 2002). Les composés organiques sont difficiles à caractériser dans la région du Vis-PIR, car cette région du spectre contient seulement des harmoniques des liaisons CH, OH et NH dont les vibrations fondamentales se

situent dans les fréquences inférieures du MIR (Murray *et al.*, 1987, Steward, 2006). Henderson *et al.* (1992) classent les longueurs d'onde importantes pour la prédiction des teneurs en C organique dans trois catégories : *i*) les longueurs d'onde en rapport avec les teneurs en C organique, *ii*) les longueurs d'onde en rapport à la fois avec les teneurs en C organique et avec les teneurs en d'autres constituants du sol, et *iii*) les longueurs d'onde qui sont en rapport avec d'autres propriétés du sol et qui masquent l'effet C organique.

Workman (1998) attribue aux composés organiques des bandes d'absorption dans le PIR. Il s'agit des bandes 1700-1900 et 1100-1300 nm respectivement pour les harmoniques de premier et de second ordres de la liaison C-H, et de la bande 1450-1550 nm pour la première harmonique de la liaison N-H.

Aussi bien Chang *et al.* (2002) que Workman (1998) attribuent une élévation distincte à l'acide humique dans la bande 2300-2350 nm.

Par ailleurs, la matière organique du sol influence les propriétés physiques (structure, porosité), chimiques (agent complexant, phénomène de sorption) et biologiques (source de nutriments pour les plantes et pour les microorganismes du sol). Ainsi, la matière organique influence de façon directe (groupement fonctionnel) et indirecte (tailles et agencements des particules du sol) le spectre de réflectance d'un sol.

11.4. Constituants minéraux

Beaucoup de constituants minéraux influencent le spectre du sol. Il s'agit notamment des minéraux argileux, des oxydes et hydroxydes de fer, des minéraux carbonatés et sulfatés, *etc.*

11.4.1. Minéraux argileux

Chaque minéral argileux possède des longueurs d'onde d'absorption qui lui sont spécifiques (Tableau 1), ce qui lui confère à l'état pur une réponse « empreinte » spectrale singulière (figure 11).

Tableau 1. Principaux minéraux argileux avec leurs longueurs d'onde d'absorption correspondantes.

Minéral	Longueur d'ondes (nm)	Références
Smectite	1400-1900-2200	Hunt et Salisbury, 1970 ; Goetz <i>et al.</i> , 2001
Mica	1400-2200-2600	Hunt et Salisbury, 1970 ; Clark <i>et al.</i> , 1990
Kaolinite	1400-2200	Hunt et Salisbury, 1970

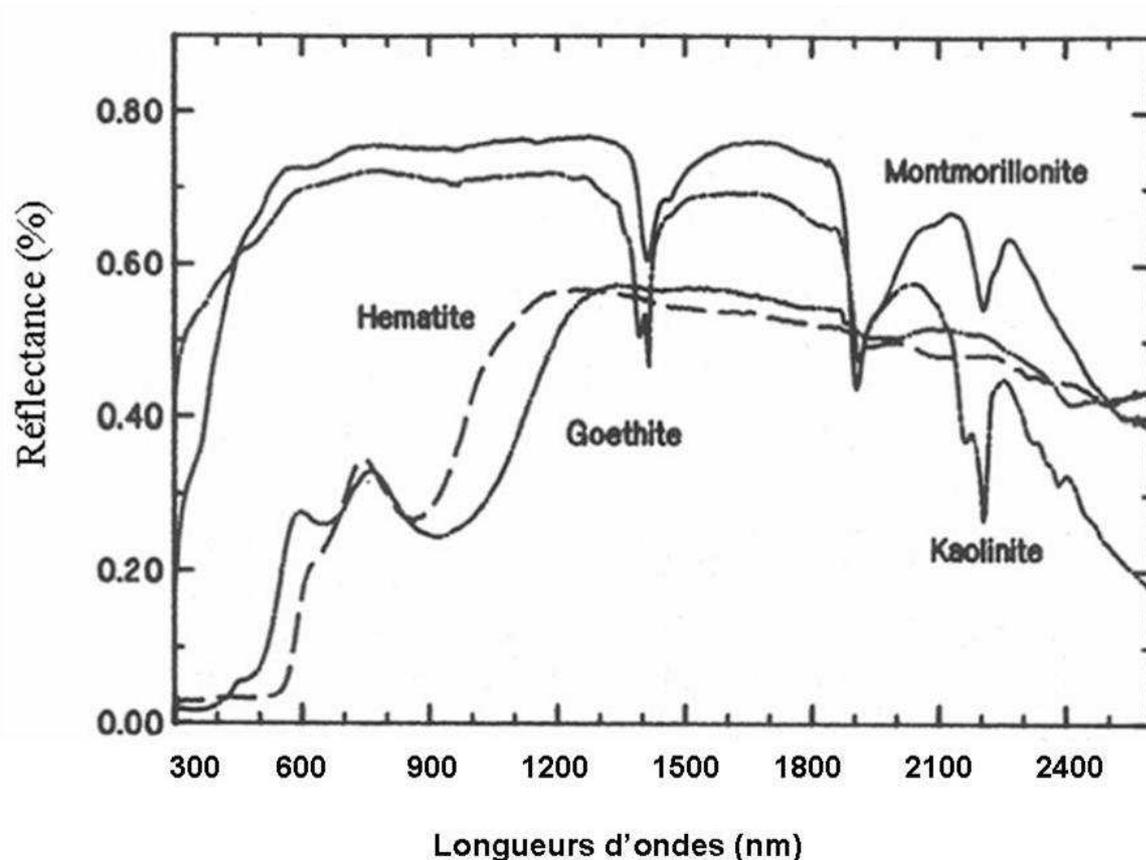


Figure 10. Comparaison des réflectances Vis-PIR de quatre minéraux argileux purs communs dans les sols (d'après Mustard et Sunshine, 1999).

11.4.2. Oxydes et hydroxyde de fer

D'après Baumgardner *et al.* (1985), dans la région du Vis-PIR, les spectres du sol sont le plus influencés par les bandes d'absorption des oxydes du fer situées à 870 et 1000 nm. Travis (2006) affirme que les oxydes de fer présentent des pics d'absorption qui apparaissent au voisinage de 500, 650 nm, et 810 nm. Clark *et al.* (1990), ont montré que des vibrations de combinaison des groupements métal-OH apparaissent entre 2200-2300 nm. De même Gaffey

et al. (1993) ont montré que l'hématite Fe_2O_3 présente trois pics respectivement à 450, 630 et 860 nm. En outre, Scheinost et Scherwtmann (1999) ont montré que la Goethite $\text{FeO}(\text{OH})$ présente trois pics distincts à 493, 660 et 940 nm.

11.4.3. Minéraux carbonatés et sulfatés

Les minéraux les plus considérables sont la calcite (CaCO_3), la dolomite $\text{Ca.Mg}(\text{CO}_3)_2$ et le gypse ($\text{CaSO}_4, 2\text{H}_2\text{O}$). Les bandes d'absorption observées pour la calcite sont dues à l'ion planaire CO_3^{2-} . Les bandes d'absorption les plus fortes se produisent dans la région du MIR à 14700, 11400, 9407 et 7067 nm, mais il y a plusieurs harmoniques et combinaisons de bandes qui se situent dans le PIR (Clark *et al.*, 1990). On distingue deux bandes relativement intenses à 2500 - 2550 nm et 2300-2350 nm, et des bandes plus faibles situées à 2120 -2160 nm, 1970-2000 nm et 1850-1870 nm (Clark *et al.*, 1990). Le gypse, qui est très sensible à l'eau, présente trois pics d'absorption à 1450, 1950 et 2350 nm (Girard et Girard, 1999).

12. Niveaux de précisions des prédictions spectrométriques des propriétés des sols dans la littérature

Bien que les minéraux (argiles, oxydes de fer), la matière organique et l'eau constituent les principaux facteurs qui influencent la réponse spectrale « chromophores » (Ben dor *et al.*, 2002 ; Malley *et al.*, 2004), beaucoup de propriétés du sol comme la teneur en carbone organique (CO) (Ben-Dor et Banin, 1990a), la capacité d'échange cationique (CEC) (Kariuki *et al.*, 2003), la texture (Sørensen et Dalsgaard, 2005), les métaux lourds (Kemper et Sommer, 2002), les sels (Farifteh *et al.*, 2007), les nutriments biodisponibles : Na (sodium), P (phosphore), K (potassium), S (soufre), Mg (magnesium), Fe (fer), Mn (manganèse), *etc* (Islam *et al.*, 2003) ont été prédites à partir des mesures spectrométriques.

Le tableau 2, extrait d'une synthèse bibliographique compilée par Viscarra *et al.* (2006), montre que la prédiction en utilisant l'outil spectroscopique a été testée dans le cas de diverses propriétés du sol. Ces études ont utilisé diverses fenêtres spectrales, diverses méthodes statistiques de calibration, diverses tailles de jeux d'échantillons de calibration et de validation. Certaines de ces études ont procédé à une validation externe des modèles calibrés à l'aide de jeux d'échantillons indépendants des jeux de calibration, alors que d'autres ont appliqué seulement une validation croisée. Toutefois elles montrent toutes que la spectrométrie est un outil prometteur pour l'étude des sols.

Tableau 2. Résultats de prédiction quantitatives de diverses propriétés du sol utilisant différentes techniques de régression multivariée et la réponse spectrale des échantillons de sols dans le domaines du visible et du proche-infrarouge (D'après une synthèse de Viscarra et al., 2006).

Propriété du sol	Gamme spectrale (nm)	Méthodes statistiques	N _{cal} /N _{val}	REQM	R ²	Auteurs
Acide (ech) ; cmol/kg	400-2498	RCP (11)	30/119	24.4	0.65	Chang et al. (2001)
Biomasse (N) ; mg/kg	1100-2300	RMCP (8)	180	-	0.71	Reeves et Mc Carty (2001)
Biomasse (N) ; g/kg	1000-2498	RMCP (6)	120/59	-	0.79	Reeves et al. (1999)
Biomasse ; mg/kg	400-2498	RCP (9)	30/119	389.7	0.60	Chang et al. (2001)
C (inorg.) ; g/kg	1100-2498	RMCP (19)	177/60	-	0.87	Mc Carty et al. (2002)
C (total) ; g/kg	1100-2498	RMCP (16)	76/32	0.15	0.96	Chang et Laird (2002)
C (total) ; g/kg	1100-2498	RMCP (7)	120/59	-	0.96	Reeves et al. (1999)
C (total) ; g/kg	400-2498	RMCP (7)	30/119	0.79	0.87	Chang et al. (2001)
Ratio C/N	400-2498	RMCP (6)	76/32	-	-	Chang et Laird (2002)
CEC ; cmol(+)/kg	700-2500	RCP	121/40	-	0.67	Islam et al. (2003)
CEC ; cmol(+)/kg	400-2498	RCP (8)	30/119	38.2	0.81	Chang et al. (2001)
Ca ; mmol/kg	700-2500	RCP	121/40	-	0.72	Islam et al. (2003)
Ca (ech) ; cmol(+)/kg	400-2498	RCP (12)	30/119	40	0.75	Chang et al. (2001)
Carbonate (%)	1000-2500	RLM	22/47	-	0.69	Ben Dor et Banin (1995)
Cu (mg/kg) ;	400-2500	RMCP	310	-	0.82	Cozzolino et Moron (2003)
CE (µS/cm) ;	400-2400	RPP	15/10	-	0.65	Shibusawa et al. (2001)
Fe (libre) ; (%)	700-2500	RCP	121/40	-	0.49	Islam et al. (2003)
Fe (mg/kg) ;	400-2500	RMCP	311	-	0.90	Cozzolino et Moron (2003)
K(assim.) ; (mg/kg)	400-1100	RN	41	-	0.80	Daniel et al. (2003)
K(éch.) ; (mg/kg)	400-2498	RCP (13)	30/119	61.4	0.64	Chang et al. (2001)
Besoin en chaulage (t/ha)	700-2500	RMCP	188	-	0.73	Janik et al. (1998)
Mg (éch.) ; mg/kg	400-2498	RCP (9)	30/119	12.8	0.68	Chang et al. (2001)
Mg ; mmol(+)/kg	700-2500	RCP	121/40	-	0.59	Islam et al. (2003)
Mg ; g/kg	400-2500	RMCP	315	-	0.90	Cozzolino et Moron (2003)
N(NO ₃) ; mg/100g	400-2400	RPP	15/10	-	0.54	Shibusawa et al. (2001)
N (total) ; %	1100-2500	RLM	72/48	-	0.92	Dalal et Henry (1986)
N (total) ; mg/kg	1100-2498	RMCP	120/59	-	0.95	Reeves et al. (1999)

N (total) ; g/kg	400-2498	RMCP	76/32	0.04	0.86	Chang et Laird (2002)
Na (éch.) ; cmol/kg	400-2498	RCP	30/119	1.3	0.09	Chang et al. (2001)
Carbone (org.) ; %	1100-2500	RLM	72/48	-	0.93	Dalal et Henry (1986)
Carbone (org.) ; %	700-2500	RCP	121/40	-	0.68	Dalal et Henry (1986)
Carbone (org.) ; g/kg	1100-2498	RMCP	177/60	-	0.82	McCarty et al. (2002)
Matière organique ; %	400-1100	RN	41	-	0.86	Daniel et al. (2003)
P (assim.) ; mg/kg	400-1100	RN	41	-	0.81	Daniel et al. (2003)
pH	1100-2498	RMCP (11)	120/59	-	0.73	Reeves et al. (1999)
pHeau	700-2500	RCP	121/40	-	0.70	Islam et al. (2003)
pHeau	400-2498	RCP (13)	30/119	0.55	0.57	Chang et al. (2001)
Argile ; %	700-2500	RCP	121/40	-	0.75	Islam et al. (2003)
Argile ; %	400-2498	RCP (12)	30/119	4.06	0.67	Chang et al. (2001)
Sable ; %	400-2498	RCP (8)	30/119	9.51	0.84	Chang et al. (2001)
Limon ; %	400-2500	RMCP	317	-	0.80	Islam et al. (2003)
Θ ; %	1100-2500	RLM	72/48	-	0.97	Dalal et Henry (1986)
Θ ; g/g	400-2498	RCP (8)	30/119	0.005	0.84	Chang et al. (2001)

Nota. RCP : Régression en composantes principales, RLM : Régression linéaire multiple, RN : Réseau de neurones, RPP : Régression pas à pas. RMCP : Régression en moindres carrés partiels, éch. : Échangeable, assim. : Assimilable. Θ : humidité.

Chapitre II. Influence du carbone organique et du fer total sur les spectres de réflectance du sol et estimation chimiométrique de leurs teneurs

H. Aïchi ^{1,2*}, Y. Fouad¹, M. Sanaa², Z. Lili Chabaane², S. Follain³, P. Breger¹, D. Causeur ⁴, C. Walter ¹

¹UMR SAS Agrocampus Rennes-INRA, 65 rue de St Brieuc, CS 84215, 35042 Rennes, France

²Institut National Agronomique de Tunisie, Cité Mahrajène 1002 Tunis, Tunisie

³UMR LISAH SupAgro-INRA-IRD, 2 place Viala Bat.24 34060 Montpellier Cedex 1

⁴Agrocampus Rennes-Laboratoire de Mathématiques Appliquées, 65 rue de St Brieuc, CS 84215, 35042 Rennes, France

Abstract

Soil spectral response is simultaneously influenced by a large number of factors including : wetness, surface roughness, various mineral constituents, organic matter, *etc.* Using “artificial” soil samples, we have : i) studied qualitatively, the influence of organic carbon and total iron on the spectral response, ii) predicted organic carbon and total iron contents, by fitting Partial Least Squares (PLS) regression models to reflectance spectra, iii) and tested the sensitivity of PLS models to calibration sets.

Two sets of 11 samples were constituted, by mixing in different proportions an organic carbon rich material respectively with a total iron rich material and a total iron poor material. For these two sets, contents in organic carbon and in total iron vary meaningfully, whereas the other parameters influencing the spectral response are similar. Spectra have been acquired in controlled conditions between 400 and 950 nm

We showed that : i) The depressing effect of Organic carbon has repercussions on the entire spectral range, without showing any specific absorption peak. By contrast, total iron does have specific absorption peaks. The continuum removal procedure made it possible to highlight five of the total iron peaks centred at 431, 497, 677, 795 and 907 nm. ii) Both organic carbon and total iron are accurately predicted by the PLS Regression models, with coefficients of determination that are superior to 0,9 and root mean square errors that are close to zero, iii) the PLS Regression models seem to integrate the variability of samples. Moreover, the increase of the size of calibration samples contributes to the accuracy of prediction.

Key words : spectroscopy, PLS regression, organic carbon, total iron, artificial soil samples.

Résumé

La réponse spectrale du sol est influencée simultanément par plusieurs facteurs tels que l'humidité, la rugosité, les constituants minéraux, la matière organique, *etc.*

À l'aide d'échantillons « artificiels », nous avons : i) étudié qualitativement, l'influence des teneurs en carbone organique et en fer total sur la réponse spectrale, ii) estimé successivement les teneurs en carbone organique et en fer total, par ajustement de modèles de Régressions en Moindres Carrés Partiels (RMCP), iii) et étudié la variation des performances des modèles RMCP en fonction de la variation du jeu de calibration. Les échantillons artificiels sont composés de deux séries de 11 échantillons chacune qui ont été obtenues en mélangeant en différentes proportions un matériau riche en carbone organique, avec respectivement un matériau riche en fer et un matériau pauvre en fer. Pour ces deux séries, les teneurs en carbone organique et en fer total varient significativement alors que les autres paramètres qui influencent la réponse spectrale demeurent similaires. Les spectres ont été acquis dans des conditions contrôlées dans la bande spectrale 400-950 nm. Nous avons montré que : i) le carbone organique entraîne un effet « dépresseur » sur tout le spectre de réflectances apparentes, sans présenter de pics spécifiques. La normalisation des spectres par retrait du continuum montre des pics d'absorptions spécifiques au fer total, centrés sur les longueurs d'onde suivantes : 431, 497, 677, 795 et 907 nm, ii) aussi bien le carbone organique que le fer total sont bien prédits par des modèles RMCP, avec des coefficients de détermination qui sont supérieurs à 0,9 et des racines de l'erreur quadratiques moyennes qui sont proches de zéro, iii) les modèles RMCP semblent intégrer la variabilité des échantillons et l'augmentation de l'effectif des échantillons de calibration améliore la qualité de prédiction.

Mots clés : spectrométrie, Régression MCP, carbone organique, fer total, échantillons de sols artificiels.

1. Introduction

La couleur est l'un des traits pédologiques les plus étudiés. Elle représente une information directement accessible sur le terrain (Krishnan et *al.*, 1981 ; Ben Dor et *al.*, 1997 ; Post et *al.*, 2000). Par ailleurs, elle est retenue comme critère de base à différents niveaux de la classification des sols (Al Abbas et *al.*, 1972 ; Stoner et Baumgardner, 1981, Henderson et *al.*, 1989 ; Schulze et *al.*, 1993). La variation spatio-temporelle de la couleur du sol est fort révélatrice de l'occurrence des processus pédogénétiques tels que la décarbonatation, le lessivage, l'acidification, l'érosion, l'hydromorphie, la salinisation (Bedidi et *al.*, 1992 ; Mathieu et *al.*, 1998 ; Girard et Girard, 1999 ; Blavet et *al.*, 2000 ; Galvão et *al.*, 2001).

Les pédologues ont cherché de longue date à mieux approcher la couleur et ce pour s'affranchir de l'appréciation subjective de l'examineur. L'utilisation de la charte Munsell est une forme de codification de la couleur qui est utilisée à cette fin. Désormais la couleur peut être mesurée à l'aide des spectroradiomètres (Escadafal et *al.*, 1993). L'apparition de nouvelles générations de spectroradiomètres hyperspectral portable, avec une résolution radiométrique très fine, ouvre de nouvelles applications à la spectrométrie dans la caractérisation des couvertures pédologiques (Ben Dor et Banin, 1995 ; Girard et Girard 1999 ; Weidung et *al.*, 2002 ; Batchily et *al.*, 2003, Brown et *al.*, 2006), le suivi de leurs évolutions et l'estimation rapide, non destructive et simultanée d'un nombre important des propriétés pédologiques (Dalal et Henry 1986, Ben Dor et *al.*, 1997 ; Mathieu et *al.*, 1998 ; Weidung et *al.*, 2002 ; Mc Bratney et *al.*, 2002 ; Batchily et *al.*, 2003, Viscarra et *al.*, 2006).

Les facteurs qui influencent la couleur du sol et par conséquent la réponse spectrale, en tant que mesure de cette dernière, sont nombreux. Il s'agit notamment de :

- L'humidité : Pour des faibles teneurs en eau, plus le sol est humide, plus la réflectance diminue (Bedidi et *al.*, 1992 ; Muller et Décamps 2000). L'eau agit d'une manière généralisée sur tout le spectre. Un sol mouillé apparaît toujours plus foncé qu'un sol sec, ceci s'explique par l'effet de diffusion aux interfaces eau - particules du sol, l'eau agit comme une piège pour le rayonnement incident (Fourty et Baret, 1998). Toutefois, à partir d'un seuil critique de teneur en eau, nous notons une inversion dans la relation humidité-réflectance ; la réflectance augmente quand la teneur en eau augmente (Weidung et *al.*, 2002). L'eau présente des bandes d'absorption spécifiques situées autour de 950, 1150, 1450, 1950, 2350 nm. Les pics d'absorption à 950 et 1150 nm sont assez faibles alors que ceux à 1450 et 1950 nm sont toujours visibles (Weidung et *al.*, 2002).

- La rugosité : La réflectance augmente quand la taille des particules de la matrice du sol diminue ; le rayonnement réfléchi augmente d'une façon exponentielle avec la diminution de la taille des particules jusqu'à la taille de 400 μm (Dalal et Henry, 1986). Lorsque les particules sont de grande taille la part de la lumière rétrodiffusée augmente au détriment de la lumière réfléchie ce qui diminue la réflectance (Girard et Girard, 1999). Certains auteurs mentionnaient des variations de réflectance selon des teneurs plus ou moins grandes en argile ou en sable (Stoner et Baumgardner, 1981 ; Henderson et *al.*, 1992 ; Houssa et *al.*, 1996), mais elles sont alors dépendantes de la teneur en eau de ces éléments, puisqu'ils présentent des propriétés de rétention qui sont différentes (Girard et Girard 1999).

- Les constituants minéraux du sol : La couleur du sol est fortement liée à la couleur de la matrice minérale (Henderson et *al.*, 1989 ; Bedidi et *al.*, 1992 ; Leone et Sommer, 2000 ; Ben-Dor, 2002 ; Batchily et *al.*, 2003). Les propriétés spectrales des minéraux sont complexes et résultent de deux processus : les transitions électroniques et les vibrations ioniques. Les transitions électroniques sont liées aux ions métalliques et sont dues au transfert des électrons d'un niveau d'énergie à un autre. Pour l'essentiel, ces phénomènes sont produits par les ions ferreux Fe^{2+} ou ferrique Fe^{3+} . Les oxy-hydroxides de fer présentent des absorptions très caractéristiques dans le domaine Visible-Proche Infrarouge (Vis-PIR). Les processus vibratoires correspondent à la vibration de l'anion entier à l'intérieur du réseau cristallin (OH^- , CO_3^- , SO_4^-). Ils se traduisent surtout par des bandes spectrales d'absorption dans le proche infrarouge (PIR) et l'infrarouge thermique (Hunt et Salisbury, 1970 ; Clark, 1999 ; Ben-Dor et Banin, 1995). Bien que peu abondants comparés aux autres constituants du sol, les oxy-hydroxides de fer, sont responsables des grandes ondulations des spectres des sols (Henderson et *al.*, 1992) et des couleurs qui en découlent (Escadafal et *al.*, 1993). Les oxydes de fer sont des agents pigmentants surtout pour les sols ayant des teneurs faibles en matière organique (Torrent et *al.*, 1983).

- La matière organique : Dans le spectre du Vis-PIR, la matière organique présente un effet dépresseur sur la réflectance (Dalal et Henry, 1986). La matière organique intervient non seulement par sa teneur mais aussi par son stade de décomposition. Plus le sol est riche en humus mature et stable, plus sa couleur est foncée (Ben Dor et *al.*, 1997). La réponse spectrale de la matière organique ainsi que ces pics d'absorption évoluent au cours du temps en fonction de la disparition de certains composés chimiques et l'apparition des produits de la décomposition. Selon Krishnan et *al.* (1980), la matière organique présente un pic d'absorption à 800 nm, or d'après les travaux de Ben Dor et *al.* (1997), ce pic n'est observé que lorsque la matière organique est mature. D'après les travaux de Henderson et *al.* (1992),

ce sont les acides humiques qui sont le plus responsables de l'effet dépressur, les acides fulviques semblent avoir un effet moins important.

Dans notre étude, nous nous intéressons particulièrement à deux paramètres d'intérêt agronomique et environnemental, la teneur en carbone organique et la teneur en fer. Ces deux constituants, variables dans l'espace et dans le temps (Torrent et *al.*, 1983 ; Mathieu et *al.*, 1993 ; Leone PA, et Sommer S, 2000 ; Batchily et *al.*, 2003), sont particulièrement pertinents dans l'évaluation de la qualité des sols (Krishnan et *al.*, 1981, Escadafal et *al.*, 1993, Ben Dor et *al.*, 1997, Fourty et Baret 1998). En effet, l'état de dégradation des sols est corrélé négativement à la teneur en carbone organique du sol, de même que les sols lessivés sont appauvris en fer (Baize et Girard, 1995).

Les objectifs de notre travail consistaient à :

- (i) - étudier qualitativement l'influence des teneurs en carbone organique et en fer total sur la réponse spectrale.
- (ii) - prédire successivement les teneurs en carbone organique et en fer total, par ajustement de modèles de régressions en moindres carrées partielles aux spectres de réflectance apparente.
- (iii) - étudier comment varie la performance des modèles RMCP en fonction de la variation du jeu de calibration.

2. Matériels et méthodes

2.1. Constitution de deux séries d'échantillons artificiels

Comme il est difficile de disposer d'une collection de sols naturels où les variables d'intérêts, à savoir la teneur en carbone organique et la teneur en fer total, varient significativement alors que les autres facteurs qui influencent la réponse spectrale (texture, humidité, constituants minéralogiques, *etc.*) demeurent similaires, nous avons procédé à la constitution de deux séries d'échantillons « artificiels ».

Pour cela, trois matériaux de même texture, mais différents par leurs teneurs en matière organique et en fer total (tableau 1) ont été prélevés sur un même versant développé sur granite de la région de Montours en Bretagne (France). Le sol en place est un Néoluvisol formé sur un placage limoneux d'origine éolienne d'une épaisseur de l'ordre du mètre (Follain, 2005).

Tableau1. Caractéristiques des matériaux ayant servi pour la constitution des deux séries artificielles.

Matériau	Origine	Couleur à l'état		Teneurs (%)		Texture
		Sec	Humide	Carbone ¹	Fer total ²	
1	Horizon A	10YR 6/3	10 YR 3/3	2.28	2.03	Limoneuse
2	Horizon Sg	10YR 7/2	10 YR 6/3	0.10	1.89	Limoneuse
3	Horizon S	10YR 7/6	10 YR 6/6	0.11	2.63	Limoneuse

¹par combustion selon la norme ISO 10694 ²par dosage HF selon la norme ISO 11885

Le premier matériau a été prélevé dans un horizon riche en matière organique. Il s'agit d'un horizon superficiel organo-minéral labouré LA (15 premiers centimètres). La couleur à l'état sec de cet horizon, selon le code Munsell, est 10YR 3/3. Le second matériau a été prélevé dans un horizon appauvri en fer. Il s'agit d'un horizon structural déferrifié rédoxique (Sg), situé à une profondeur entre 40 et 60 cm, engorgé temporairement d'eau. La couleur du second matériau, à l'état sec, est gris clair (10YR 7/2). Le troisième matériau a été prélevé, dans la même fosse que le premier matériau dans un horizon structural bien drainé (S), donc relativement plus riche en fer que le second matériau, à une profondeur entre 40 et 60 cm. La couleur du troisième matériau, à l'état sec, est ocre (10YR 7/6). La photo 1 montre les horizons de prélèvement de ces trois matériaux, décrits de façon détaillée par Follain (2005).

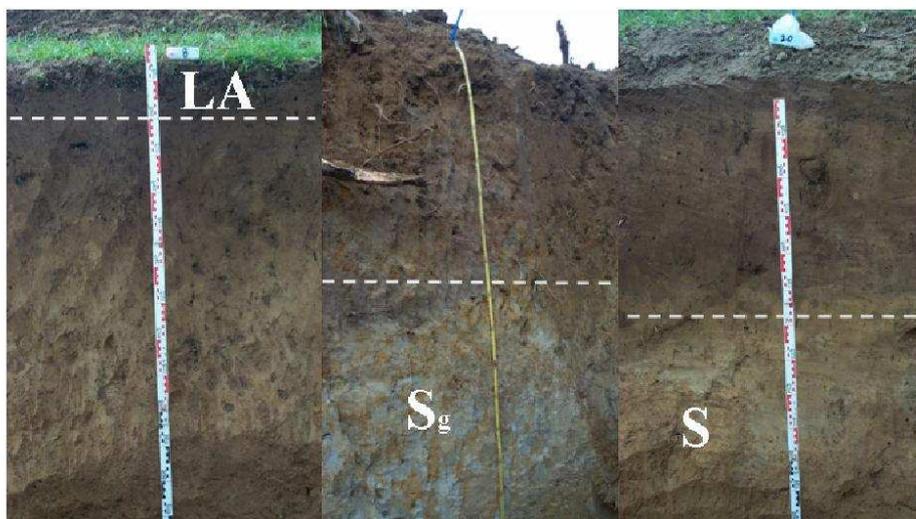


Photo 1. Horizons de prélèvement des trois matériaux utilisés dans la constitution des deux séries d'échantillons artificielles. Matériau I : horizon A, Matériau II : horizon Sg et Matériau III : horizon S.

Les trois matériaux ont été séchés et tamisés à 2 mm. La photo 2 montre ces trois matériaux placés dans des boîtes de Petri après séchage et tamisage.



Photo 2. Matériaux originels après séchage et tamisage à 2 mm placés dans des boites de Petri.

Ensuite, nous nous sommes fixés, les teneurs en carbone organique théoriques suivantes : 0.1 ; 0.2 ; 0.4 ; 0.6 ; 0.8 ; 1.0 ; 1.2 ; 1.4 ; 1.6 ; 1.8 et 2.0 %, pour les deux séries artificielles. Pour cela, sur la base de 100 g d'échantillon de sol artificiel, nous avons calculé les masses de prise de terre, qu'il faut à chaque fois mélanger : des matériaux **I** et **II** pour constituer la première série d'échantillons. Nous avons procédé de la même façon pour le mélange des matériaux **I** et **III**, pour constituer la deuxième série d'échantillons artificiels.

Une fois les mélanges réalisés nous avons mesuré les teneurs exactes en carbone organique et en fer total : chaque échantillon a fait l'objet d'une analyse pour sa teneur en carbone organique par combustion dans un analyseur CHN (Flash EA 1112) selon la norme ISO 10694 (AFNOR, 1996), et pour sa teneur en fer total par dosage à l'acide fluorhydrique HF selon la norme ISO 11885 (AFNOR, 1996). Trois répétitions de mesure sont effectuées sur chaque échantillon.

De la sorte, nous avons obtenu nos deux séries d'échantillons artificiels composées chacune de 11 échantillons (Photo 3). Le tableau 2 relate les teneurs en carbone organique et en fer total mesurées, de ces deux séries d'échantillons artificiels.

Tableau 2. Teneurs en carbone organique et en fer total des deux séries d'échantillons artificielles.

Numéro D'échantillon	Série I (mélange à différentes proportions des horizons A et Sg)		Série II (mélange à différentes proportions des horizons A et S)	
	Carbone organique ¹ %	Fer total ² %	Carbone organique ¹ %	Fer total ² %
1	0,14	1,89	0,13	2,62
2	0,26	1,91	0,24	2,56
3	0,46	1,92	0,48	2,50
4	0,66	1,93	0,66	2,44
5	0,93	1,94	0,91	2,41
6	1,13	1,95	1,11	2,35
7	1,26	1,97	1,30	2,29
8	1,60	1,99	1,51	2,23
9	1,81	2,00	1,84	2,17
10	1,95	2,01	1,94	2,11
11	2,19	2,02	2,16	2,05
Etendue de la série	2,06	0,13	2,03	0,57
Pas moyen entre deux échantillons	0,204	0,013	0,202	0,056

¹par combustion selon la norme ISO 10694

²par dosage HF selon la norme ISO 11885

Comparées à des échantillons de sol naturels, ces deux séries d'échantillons artificiels présentent les avantages suivants :

- les deux séries couvrent des gammes assez larges de teneurs en carbone organique et de teneur en fer total avec des pas moyens de variation qui sont faibles. Il nous aurait fallu prospecter des champs spatiaux très importants pour collecter des échantillons de sol qui embrassent cette variabilité de teneur en carbone et en fer.
- l'ensemble des échantillons appartenant aux deux séries ont des textures et des matrices minérales très proches. Ils ont aussi une matière organique de même nature avec la même dynamique d'évolution récente. Pour des échantillons de sols naturels, il est difficile de pouvoir disposer d'échantillons qui diffèrent par leurs taux de carbone, mais avec un type de matière organique analogue.
- Les échantillons des séries I et II ont la même origine de matériau. Par contre, ils diffèrent significativement par leurs teneurs en fer total dans les faibles teneurs en carbone organique.

2.2. Acquisition des spectres

Les échantillons sont séchés, broyés à l'aide d'un broyeur planétaire puis tamisés à une fraction < 400 µm (Dalal et Henry, 1986). Au moyen de ce tamisage fin, nous escomptons réduire l'influence de la rugosité de l'état de surface sur la réponse spectrale. Les échantillons ont été placés dans des boîtes de Pétri (He et *al.*, 2005) de 6.4 cm de diamètre intérieur, préalablement peintes en noir pour éviter toute source d'énergie autre que celle en provenance de l'échantillon de sol lui-même. Ensuite, ils sont passés à l'étuve, pendant 24 h, à une température de 45°C pour homogénéiser leur t aux d'humidité (Demattê, 2002).

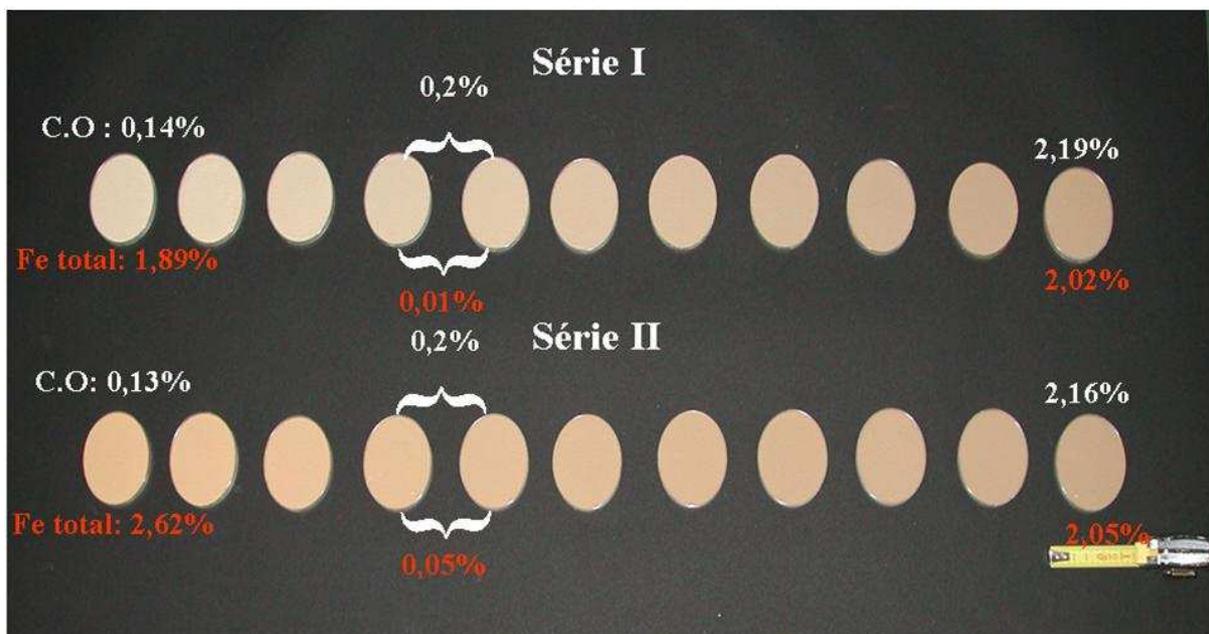


Photo 3. Séries I et II issues des mélanges à différentes proportions respectivement des matériaux I et II (série I), I et III (série II). Nous indiquons les valeurs min et max ainsi que le pas moyen de variation des teneurs en carbone organique et en fer total.

Les mesures de réflectance spectrale sont réalisées à l'aide d'un spectroradiomètre portable RAMSES-ARC (Tri OS Optical Sensors). Bien que les mesures soient acquises dans la bande spectrale 309-1130 nm, les spécifications du constructeur indiquent que les mesures sont effectives dans la bande spectrale 320-952 nm. L'échantillonnage spectral est de 3.3 nm et la résolution spectrale est de 0.3 nm. Comme la série des UV ne présente aucun intérêt pour notre étude, nous n'avons conservé que la bande spectrale 400-950 nm.

L'acquisition des spectres a été réalisée, en visée verticale, dans une chambre noire avec des conditions d'illumination standardisées (photo 4). Deux lampes halogènes, d'une

puissance de 300 W chacune, ont été fixées de part et d'autre du capteur et à la même hauteur que celui-ci. A l'aide d'un pyranomètre, positionné à la place de l'échantillon, nous avons orienté les lampes de façon à avoir le maximum d'énergie incidente.

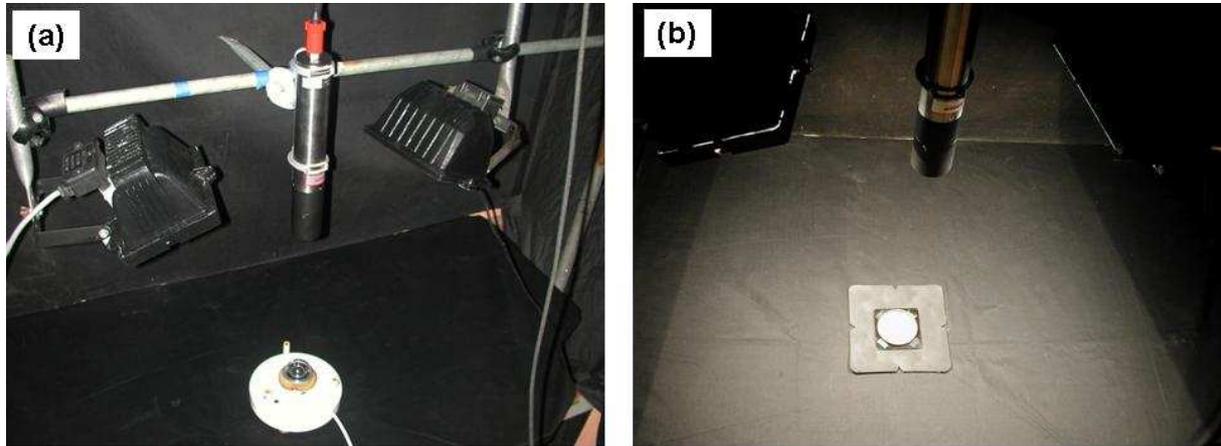


Photo 4. Réglage des angles zénithaux des lampes à l'aide d'un pyranomètre (a) et mesure du spectre d'un échantillon en conditions contrôlées de laboratoire.

Le spectroradiomètre a un champ de visée de 7° . La hauteur à laquelle le capteur devrait être maintenu pour couvrir toute la surface de l'échantillon est de 52.3 cm. Cependant, nous avons fixé l'objectif du capteur à une hauteur de 45.8 cm, ce qui présente un retrait de 4 mm de la circonférence de la boîte, pour éviter l'effet de bordure (ombre portée par le contour de la boîte qui parasite le signal). Toutefois, nous avons maintenu une bonne surface d'intégration du signal par le capteur : $\pi \cdot (2.8)^2$ soit 24.61 cm^2 . A titre indicatif, cette surface représente plus que 12 fois la surface considérée par Kooistra et *al.* (2003). En effet, il est préconisé de considérer une surface d'intégration du signal qui soit importante pour que les données spectrométriques acquises ne soient pas biaisées.

La réflectance apparente des échantillons est obtenue en mesurant successivement la luminance de la cible de référence, puis celle de l'échantillon. La cible de référence est un spectralon® ($23 \times 23 \text{ cm}^2$) dont la réflectance est supérieure à 0.99 dans toute la gamme comprise entre 400 et 1000 nm. Pour chaque échantillon, nous avons acquis 4 spectres en tournant à chaque fois l'échantillon d'un angle de 90° . C'est la moyenne des 4 mesures qui a été considérée par la suite. De la sorte, nous nous sommes affranchis de l'effet directionnel dû à l'hétérogénéité de la surface (Karnieli et *al.*, 1999) et nous avons compensé les éventuels défauts de symétrie du dispositif d'illumination.

Pour étudier qualitativement l'influence du fer sur la réponse spectrale, nous avons eu recours à la technique de la normalisation des spectres de réflectance apparente par retrait du continuum (*continuum remove*, en anglais) (Bullard et White, 2002). Il s'agit d'une technique communément utilisée dans les applications géologiques (Clark et Swayze, 1995). Cette technique permet en repérant où se situent les pics d'absorption, d'identifier la présence de certains minéraux ayant des pics d'absorption bien définis. D'après les travaux de Mougnot (1990), cités par Girard et Girard (1999), les minéraux riches en fer comme l'hématite, la goethite, la montmorillonite et la jarosite, présentent des pics d'absorptions provoqués par les processus de transition électronique dans les bandes 430, 650, 850 et 930 nm. Le continuum est défini comme un signal de fond non lié à des propriétés spécifiques d'absorption (Clark et Roush, 1984 ; Kruse et *al.*, 1988 ; Green et Craig, 1985). Il correspond à une enveloppe qui est ajustée à tout le spectre en reliant les maxima locaux par des segments de droites. Le continuum est retiré en divisant le spectre original par cette enveloppe (Lagacherie et *al.*, 2008).

2.3. Construction du modèle statistique de prédiction

En spectrométrie, le nombre de variables explicatives qui correspond au nombre de longueurs d'onde est très important d'autant plus que ces variables sont très corrélées (Desbois, 1999). Il n'est pas donc opportun d'ajuster à ces données des modèles de régression multiple (Abdi H, 2003). Pour contourner ce problème et pouvoir développer des modèles de régressions linéaires multiples, les pédométriciens ont toujours été obligés d'avoir recours à des techniques empiriques de manipulation du spectre : transformation de la réflectance en absorbance, dérivation première et seconde du signal, moyenne mobile et élimination de bandes spectrales (Krishnan et *al.*, 1981 ; Dalal et Henry 1986 ; Demetriades-Shah et *al.*, 1990 ; Ingleby et Crowe, 2000) ou à des procédures algorithmiques : régression pas à pas (Weidung et *al.*, 2002) pour sélectionner quelques longueurs d'onde, généralement 2 à 5 parmi des centaines.

Il existe des techniques statistiques plus appropriées pour ajuster un modèle à des données spectrométriques. Il s'agit notamment de la régression en moindres carrés partiels RMCP. Cette méthode statistique permet de construire des modèles prédictifs lorsque les variables explicatives sont nombreuses et très corrélées (Sundberg, 1999).

La RMCP présente l'avantage de retenir toutes les longueurs d'onde comme variables explicatives (Tenenhaus, 1998), nul besoin de sélections préliminaires et donc d'écartier de prime abord des longueurs d'onde qui pourraient être prédictives. La régression en Moindres

Carrés Partiels (RMCP) a été développée à l'origine par Herman Wold dans les années 60, comme technique économétrique (Desbois, 1999), puis fut popularisée auprès des chimistes dans des applications de chromatographie et de spectrographie (Davies, 2001).

En régression MCP, toutes les variables sont conservées. Les variables importantes dans la prédiction ont des poids élevés et les variables de moindre importance participent au modèle, mais avec un poids faible (Tenenhaus, 1998). De même, le nombre de variables peut être supérieur au nombre d'échantillons du jeu de calibration. La régression MCP univariée relie une variable à expliquer Y à un bloc de variables explicatives X (Tenenhaus, 1998) et ce, en passant par des variables latentes qui sont des combinaisons linéaires des variables explicatives initiales (Desbois, 1999). Les variables latentes expliquent le plus possible la covariance entre chacune des variables explicatives X_i et Y (Sundberg, 1999). Les variables latentes sont présentées comme un compromis entre la structure interne des variables explicatives et leur lien avec la variable à expliquer (Wold et *al.*, 2001).

Dans notre cas, les X_i sont les 170 valeurs de réflectance à différentes longueurs d'onde, la variable à expliquer est tour à tour la teneur en carbone organique et la teneur en fer des échantillons. Le nombre d'échantillons de calibration du modèle est de 11.

Le choix du nombre de composantes latentes du modèle (NCLM) a été effectué en appliquant une validation croisée (Tenenhaus, 1998). Le modèle qui présente la racine de l'erreur quadratique moyenne (REQM) la plus faible est retenue. En outre, la performance du modèle est évaluée en observant le coefficient de détermination R^2 .

Les données ont été traitées sous R. R est un langage de programmation et un environnement pour les calculs statistiques et les représentations graphiques (Ihaka et Gentleman, 1996). Un projet R est une implémentation libre du langage S (Brostaux, 2002) utilisable à travers le logiciel de statistiques S-Plus® 6.1 (commercialisé par la société Insightful, [http : //www.splus.com/](http://www.splus.com/)).

3. Résultats

3.1. Influence des teneurs en carbone organique et en fer total sur la réponse spectrale

La figure 1 montre les spectres moyens obtenus pour les deux séries artificielles. D'une façon générale, les courbes de réflectance sont continues et croissantes sur tout le spectre du visible et du proche infrarouge. Les séries artificielles I et II présentent deux allures générales distinctes (figures 1a et 1b), mais au sein d'une même série, nous avons la même forme de spectre.

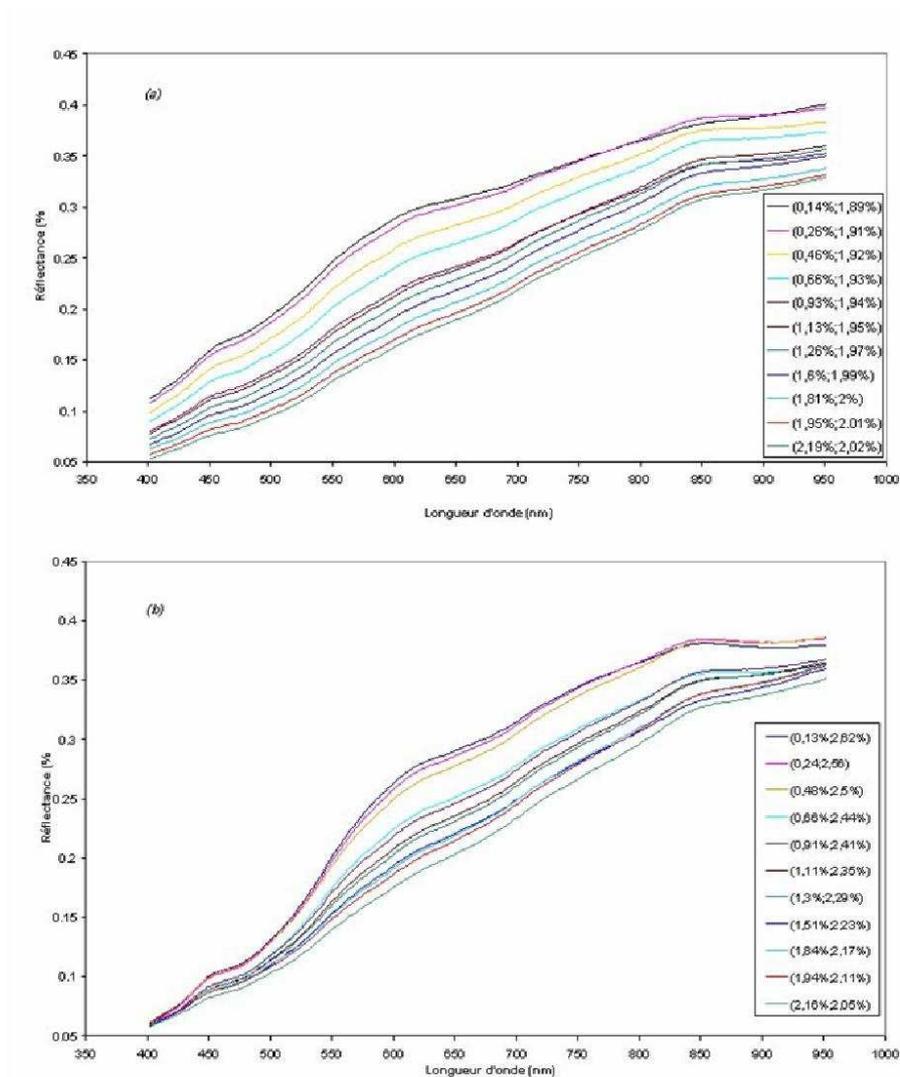


Figure 1. Réponses spectrales des deux séries d'échantillons artificiels. (a) série artificielle I (Horizon A+S_g), (b) série artificielle II (Horizon A+S_{« drainé »}). Pour chaque échantillon nous indiquons successivement sa teneur en carbone organique et en fer total.

Pour la série artificielle I (figure 1a), sur toute la bande spectrale 400-680 nm, l'ordre des courbes de réflectance est lié à la teneur en carbone organique des échantillons par une relation inverse : lorsque la teneur en carbone augmente, la réflectance diminue. La courbe est translatée vers le bas, mais sa forme reste quasi inchangée.

Pour la série artificielle II (figure 1b), plus riche en fer pour les faibles teneurs en carbone organique, les courbes de réflectance se présentent en fuseau plus compact. Elles sont plus serrées sur tout le spectre avec un regroupement des courbes dans la bande 400-440 nm. Ce regroupement s'observe aussi vers la bande 920-950 nm, mais de façon moins prononcée. En comparant les réponses spectrales de la série I à ceux de la série II (figure 1), nous constatons que les réponses de la série II présentent une pente accentuée dans la bande 480-600 nm ce qui correspond à une forte absorption de la lumière dans les bandes vert et bleu. Par contre, sur le reste du spectre, nous n'observons pas de différence notable entre les pentes des courbes des deux séries.

Pour la série II, nous remarquons que nous atteignons un palier de réflectance au voisinage de 850 nm pour les teneurs en carbone organique inférieure à 0,48 %, alors que pour les teneurs qui sont supérieures, les courbes continuent de croître.

La figure 2 représente simultanément les réponses spectrales des deux séries artificielles I et II. Elle nous permet de constater que les courbes de réflectance de ces deux séries s'enchevêtrent et il est difficile de trouver des fenêtres spectrales où l'ordre des courbes de réflectance est corrélé avec la teneur en carbone organique. Nous constatons aussi que les courbes de réflectance de la série II sont enveloppées par les courbes de réflectance de la série I.

La figure 2 nous permet en outre de comparer entre les échantillons des deux séries artificielles. A titre d'exemple, en comparant les courbes des deux échantillons appartenant respectivement à la série I et à la série II, ayant les couples de valeurs de carbone organique et de fer total (0,66 % ; 1,93 %) dont le spectre est en rouge gras et (0,13 % ; 2,62 %) dont le spectre est en bleu gras, nous constatons que dans la bande 400-560 nm, la réponse spectrale de celui qui est plus riche en fer total se situe en dessous : c'est « l'effet fer » qui prime sur « l'effet carbone organique », car nous remarquons que le deuxième échantillon est significativement plus riche en fer total une fois comparé au premier échantillon. Par contre, au delà de 560 nm, l'effet carbone organique masque l'effet fer.

De même lorsque nous comparons deux échantillons ayant la même teneur en carbone organique et qui diffèrent par leurs teneurs en fer total soient le même échantillon de la série I qui a le couple de valeurs (0,66 %, 1,93 %) et l'échantillon de la série II qui a le couple de valeurs (0,66 %, 2,62 %) dont le spectre est en vert gras, nous constatons que la courbe de celui qui est plus riche en fer total (appartenant à la série II) se situe en dessous de celui qui est moins riche. Ceci reste vrai pour tout le spectre. L'écart entre les deux courbes de réflectance est surtout important dans la bande 400-480 nm, c'est là où le fer absorbe le plus la lumière. De 600 à 850 nm, les courbes sont quasi-parallèles.

En dehors des bandes où le fer présente une absorption importante, la relation entre l'effet carbone et l'effet fer est de type « additif ». En effet, le fer entraîne une absorption de la lumière qui vient s'ajouter à l'absorption induite par le carbone organique.

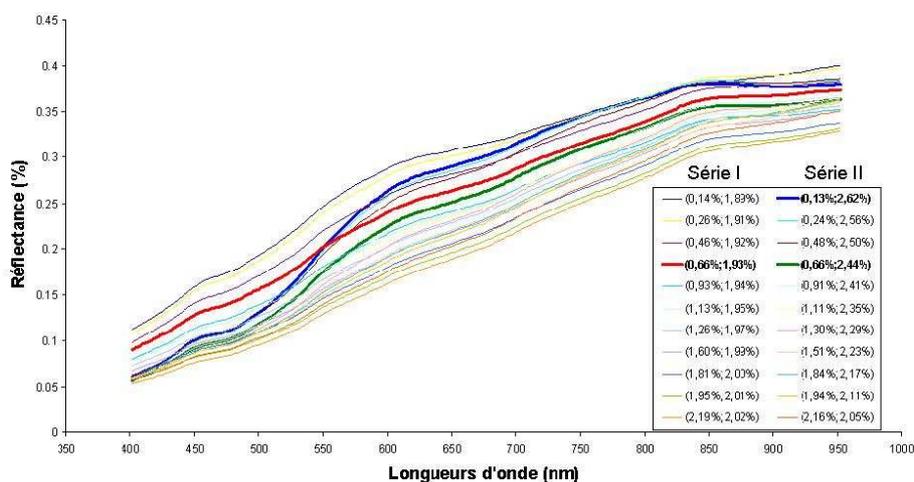


Figure 2. Représentation simultanée des réponses spectrales des deux gammes artificielles. Pour chaque échantillon nous indiquons respectivement, sa teneur en carbone organique et en fer total.

Les figures 3a et 3b montrent les spectres de réflectance des deux séries après le retrait du continuum. Nous remarquons que les courbes des deux séries présentent la même allure générale. Nous remarquons aussi que pour les deux séries, nous avons les mêmes pics d'absorption. Ces pics sont respectivement centrés sur les longueurs d'onde 431, 497, 677, 795 et 907 nm. Les pics d'absorption les plus importants sont situés dans le visible à 431 et à 497 nm. Au niveau de ces deux pics, il est aisé de constater que les spectres de la série I sont bien ordonnés en fonction de leurs teneurs en fer total. Plus l'échantillon est riche en fer total plus son pic d'absorption est important. Ce constat demeure vrai pour les courbes de la série

Il excepté pour les deux échantillons les moins riches en fer total et les plus riches en carbone organique où cet ordre n'est pas respecté.

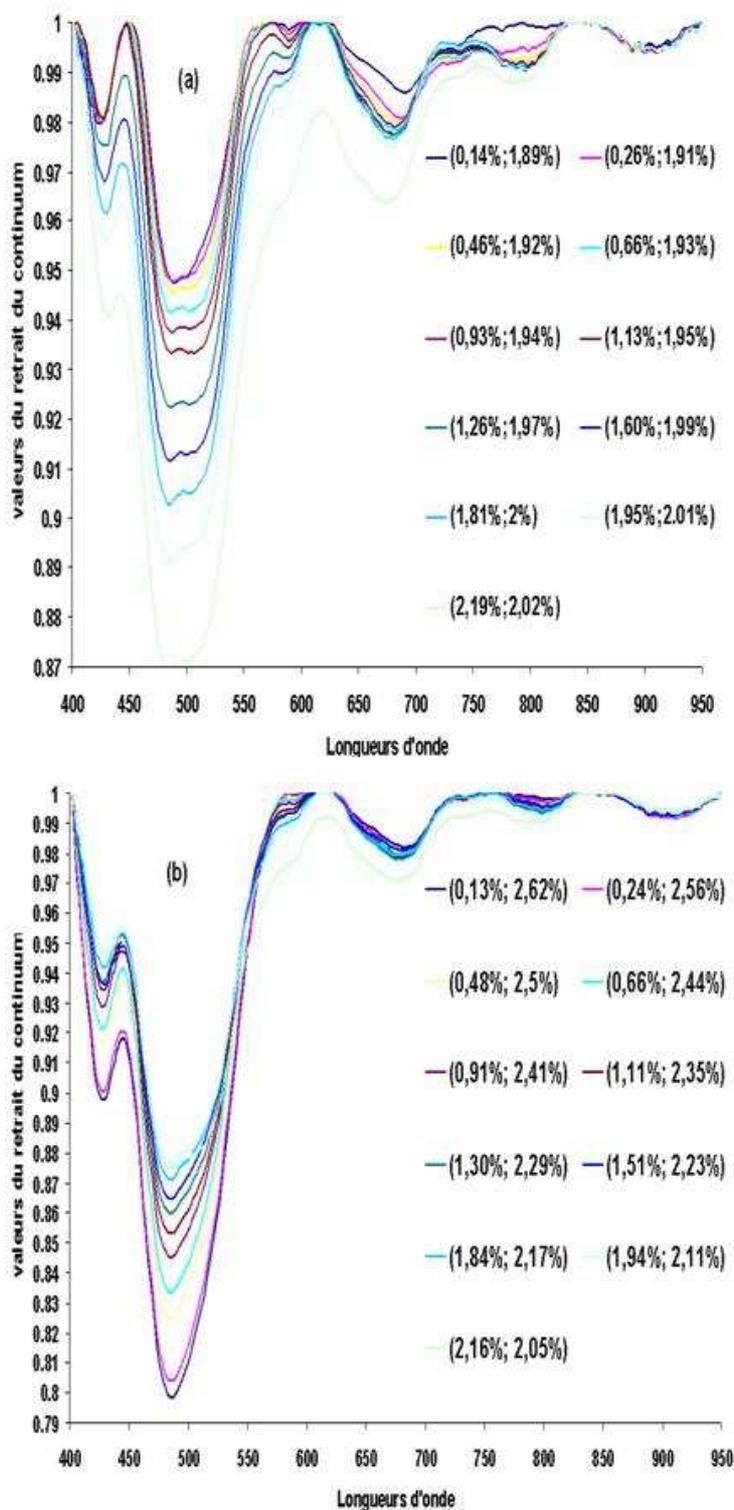


Figure 3. Spectres de réflectance des deux séries d'échantillons artificiels après le retrait du continuum. (a) série artificielle I (Horizon A+Sg), (b) série artificielle II (Horizon A+S« drainé »).

Chaque échantillon est identifié respectivement, par sa teneur en carbone organique et en fer total.

3.2. Prédiction du carbone organique et du fer total par régression MCP

Les figures 4 et 5 représentent respectivement, les résultats de la prédiction du carbone organique et du fer total par régression MCP pour les deux séries « artificielle ». Les figures 4a et 4c représentent les coefficients de la régression MCP, en fonction des longueurs d'onde, des modèles prédictifs des teneurs en carbone organique calibrés pour chacune des deux séries d'échantillons. Les figures 5a et 5c représentent les coefficients de la régression MCP, en fonction des longueurs d'onde, des modèles prédictifs des teneurs en fer total. Nous constatons que la régression MCP affecte un coefficient « poids » important aux longueurs d'onde les plus déterminantes dans la prédiction de la variable analytique considérée. Par contre, si une longueur d'onde est considérée sans importance particulière dans la prédiction, le modèle lui attribue un coefficient qui est faible en valeur absolue.

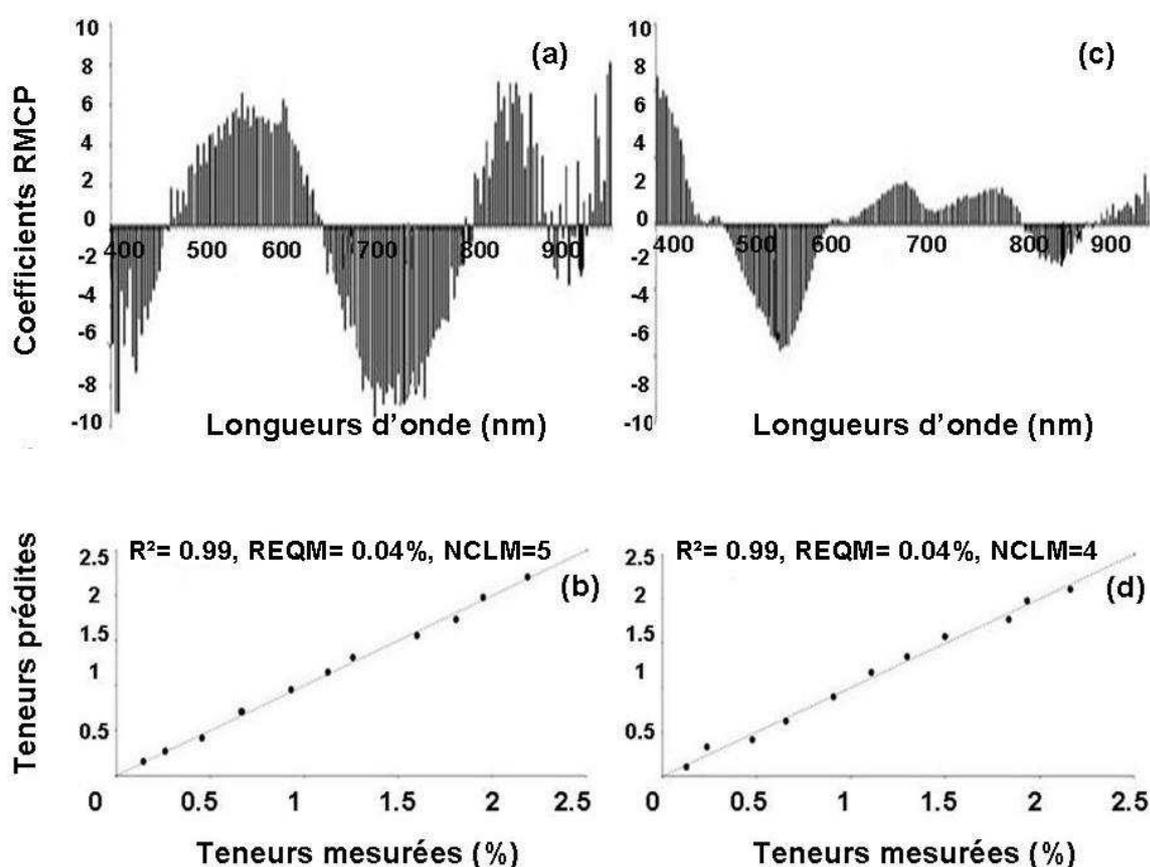


Figure 4. Prédiction du carbone organique par régression MCP pour les deux séries artificielles. (a) et (c) coefficients de la régression MCP en fonction des longueurs d'onde, pour

les séries I et II respectivement. (b) et (d) teneurs prédites en fonction des teneurs mesurées par validation croisée, pour les séries I et II respectivement.

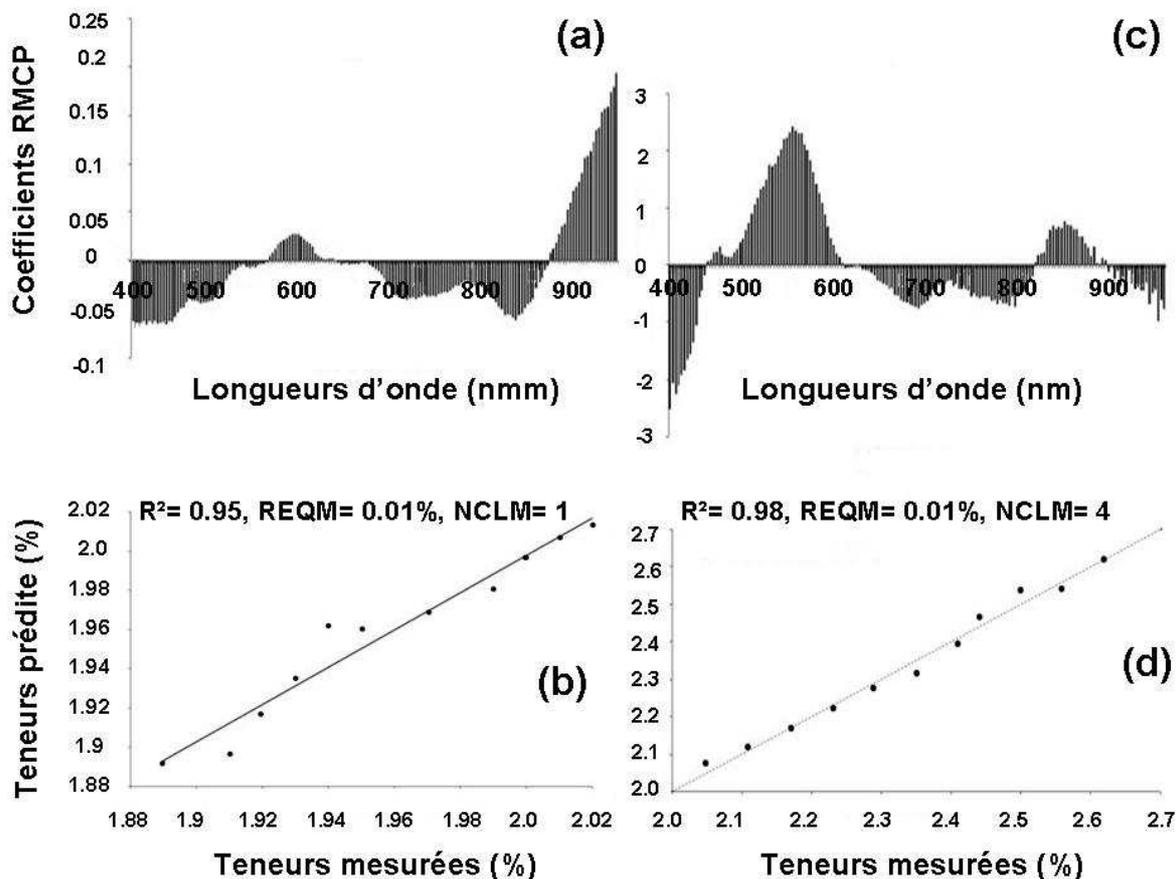


Figure 5. Prédiction du fer total par régression MCP pour les deux séries artificielles. Série I : (a) et (c) coefficients de la régression MCP en fonction des longueurs d'onde, pour les séries I et II respectivement. (b) et (d) teneurs prédites en fonction des teneurs mesurées par validation croisée, pour les séries I et II respectivement.

En examinant les figures 4a, 4c, 5a et 5c, nous constatons qu'aussi bien pour la série I que pour la série II et aussi bien pour la variable carbone organique que pour la variable fer total, deux longueurs d'onde voisines présentent des coefficients de régression de même ordre de grandeur. La variation de la valeur absolue des coefficients se fait d'une façon progressive et l'inversion dans la relation de corrélation (corrélation positive ou négative) se fait par bande spectrale.

En comparant les coefficients de la régression MCP des deux séries d'échantillons artificiels qui représentent deux jeux de calibrations indépendants, nous remarquons dans le cas de la prédiction des teneurs en carbone organique (figures 4a et 4c) et dans le cas de la prédiction des teneurs en fer total (figures 5a et 5c), que les coefficients de régressions sont totalement différents. Nous constatons aussi que les longueurs d'onde les plus discriminantes (ayant les coefficients les plus élevés en valeur absolue) varient d'un jeu de calibration à un autre et d'une pédo-variable à expliquer à une autre.

Les figures 4b et 4d représentent les teneurs en carbone organique mesurées par combustion en fonction des teneurs prédites par les modèles RMCP, calibrés pour chacune des deux séries. Alors que les figures 5b et 5d représentent les teneurs en fer total mesurées par la méthode HF en fonction des teneurs prédites par les modèles RMCP, calibrés pour chacune des deux séries. Ces figures montrent que la calibration de modèle de type RMCP permet d'avoir des fortes corrélations entre les teneurs mesurées analytiquement et les teneurs prédites. Ceci reste vrai pour les deux séries d'échantillons et pour les deux pédo-variables considérées.

Le tableau 3 récapitule les coefficients de détermination et les REQM qui ont été obtenus.

Tableau 3. Performance des modèles RMCP pour les deux pédo-variables.

	Pédo-variables					
	Carbone organique			Fer total		
	R ²	REQM	NCLM	R ²	REQM	NCLM
Série I	0,99	0,04 %	5	0,95	0,01 %	3
Série II	0,99	0,04 %	4	0,98	0,02 %	4

Nota. R²: Coefficient de détermination, REQM: Racine de l'erreur quadratique moyenne, NCLM: Nombre de composantes latentes du modèle.

Nous notons que non seulement les différents modèles RMCP calibrés présentent de bonnes qualités de prédiction avec des coefficients de détermination supérieurs à 0.9 et des racines de l'erreur quadratiques moyennes très faibles, mais également que leurs performances sont stables, car en changeant d'un jeu de calibration à un autre et d'une pédo-variable à une autre les modèles calibrés sont de qualité similaire.

La figure 6 représente la qualité de prédiction des modèles globaux du carbone organique et du fer total lorsque les modèles sont calibrés sur l'ensemble des 22 échantillons. Nous constatons que la qualité de la prédiction est légèrement supérieure à celle obtenue précédemment avec les jeux d'échantillons séparés.

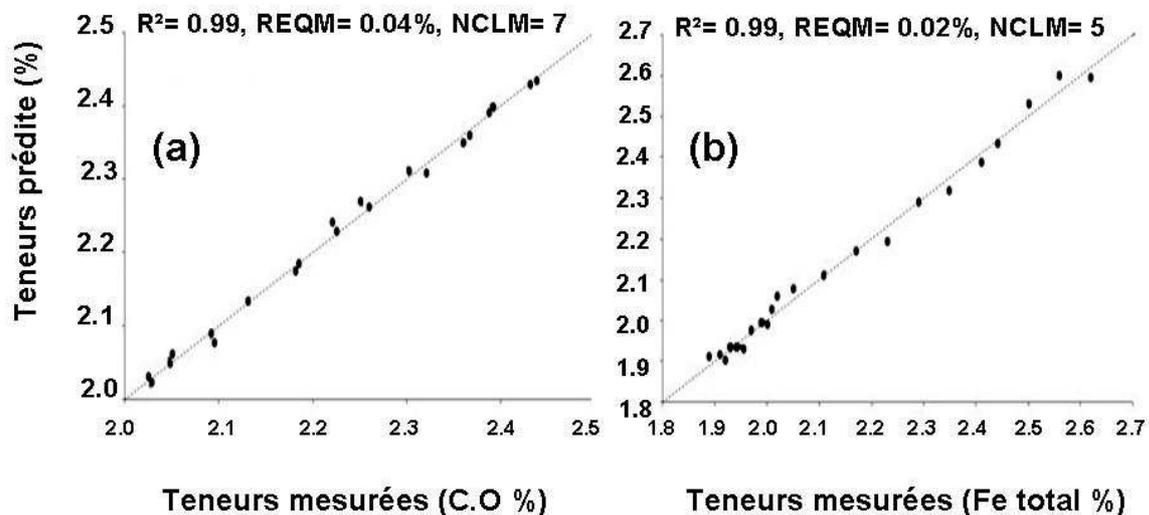


Figure 6. Prédiction par régression MCP en validation croisée sur l'ensemble des deux séries artificielles : (a) carbone organique et (b) fer total.

4. Discussion

4.1. Interprétation de la réponse spectrale

Les courbes de réflectance des sols sont continues et croissantes sur tout le spectre de 400-950 nm ; la réponse spectrale change graduellement et une forte corrélation existe entre les bandes voisines. Cette redondance spectrale a été déjà notée par Weidung et *al.* (2002) qui ont mis en relation la réponse spectrale avec le taux d'humidité du sol. En règle générale, les sols riches en carbone organique ont une courbe concave, dans la bande spectrale 400-950 nm, alors que les sols les moins riches ont une courbe plutôt convexe. Ces observations des formes classiques des spectres ont été relevées par des auteurs tels qu'Escadafal et *al.* (1993), Weidung et *al.* (2002) et Demattê et *al.* (2002).

Le fer présente des pics d'absorption pouvant être décelées par la simple visualisation des spectres de réflectance apparente. Il s'agit notamment de la bande 400-440 nm et de la bande 930-950 nm. La normalisation des spectres par retrait du continuum a permis

d'identifier les pics d'absorptions de fer d'une façon beaucoup plus précise. Ces pics sont centrés sur les longueurs d'onde suivantes : 431, 497, 677, 795 et 907 nm. Les pics d'absorption les plus importants sont situés dans le visible à 431 et à 497 nm. Pour la longueur d'onde 497 nm, la relation linéaire inverse entre la teneur en fer total et la réflectance normalisée est évidente.

En dehors de ces deux pics, le fer entraîne un effet « dépresseur » qui est généralisé sur tout le spectre. En effet, pour deux échantillons ayant la même teneur en carbone organique, celui qui est plus riche en fer total réfléchit moins la lumière, ceci restant vrai sur tout le spectre. Les courbes de réflectance des deux échantillons présentent la même forme. Elles sont quasi-parallèles sur toute la gamme spectrale considérée, excepté au niveau des bandes spectrales où le fer présente des pics d'absorption. C'est le cas des bandes spectrales 400-460 et 920-950 nm, où l'écart entre les deux courbes augmente significativement. Cela rejoint les travaux de Rezend et *al.* (1980) qui avaient observé une absorption du fer dans la bande spectrale 440 à 460 nm sur des courbes de réflectance d'oxisols. De même Girard et Girard (1999) mentionnent que le fer présente une absorption entre 860 et 910 nm. Par comparaison aux échantillons de la série artificielle I, les échantillons de la série artificielle II, plus riches en fer total dans les faibles teneurs en carbone organique, présentent une pente accentuée dans la bande 480-600 nm.

La même observation a été faite par Bédidi et *al.* (1992) dans la bande 400-700 nm pour des sols latéritiques riches en oxy-hydroxydes de fer. Par contre dans leurs travaux Hunt (1971) et Bullard J et White K (2002) attribuent au fer deux pics d'absorption situés à 550 nm et à 850 nm. De même d'après les travaux de Bonn et Escadafal (1996) et ceux de Madeira et *al.* (1997). Les sols ou les roches riches en oxyde de fer ferrique (Fe^{3+}) présentent une absorption proche de 550 nm pour l'hématite (rouge) et proche de 480 nm pour la goethite. Chen et *al.* (2000) qui ont cartographié la teneur en carbone organique de la surface d'une parcelle de 115 ha en Georgie (EUA), préconisent de soustraire l'effet fer des spectres avant de modéliser la variation des teneurs en carbone organique.

Dans notre domaine spectral d'investigation de 400 à 950 nm, nous n'avons pas pu déceler des pics d'absorption spécifiquement attribuables au carbone organique. Le carbone organique entraîne un effet « dépresseur » sur tout le spectre, sans présenter de bande spécifique comme le fer. Certains auteurs qui avaient comme objectif la caractérisation de la matière organique (Ben Dor et *al.*, 1997 ; Fourty et Baret, 1998 ; Dematté et *al.*, 2004) ont exploré un domaine spectral plus large que le nôtre (400-2600 nm) et ont noté des pics d'absorption liés aux liaisons carbone au voisinage de 760, 900, 2080, 2350 et 2500 nm.

Selon Coûteaux et *al.* (2003), ces pics sont spécifiques de certains produits de la décomposition de la matière organique (cellulose, lignine, amidon, cire, acide humique, *etc.*). Par ailleurs, la nature de la matière organique et le stade de sa décomposition sont déterminants dans l'observation de ces pics. En effet, Krishnan et *al.* (1980) ont constaté que le carbone organique présente un pic à 800 nm. Or, d'après les travaux de Ben Dor et *al.* (1997), ce pic n'est observable que lorsque la matière organique est humifiée.

Pour des teneurs élevés en carbone organique, l'effet du carbone organique masque ou distord l'effet du fer. Ce qui expliquerait le fait qu'après retrait du continuum, les spectres des deux échantillons les plus riches en carbone organique et les moins riches en fer total de la deuxième série ne sont pas ordonnés en fonction de leur teneur en fer total comme c'est le cas pour le reste des échantillons de la même série. En effet, Baumgardner et *al.* (1970) qui ont étudié des sols de l'Indiana (EUA), soulignaient que pour des taux élevés de matière organique qui dépassent les 2 %, l'effet carbone peut masquer l'effet des autres constituants minéraux sur les spectres des sols. De même, He et *al.* (2005), estiment que la matière organique a une composition chimique complexe et que sa forte propriété d'absorption peut masquer les propriétés d'absorption des autres constituants minéraux du sol.

4.2. Interprétation des performances des modèles calibrés

Soulignons que nous sommes persuadés que les hautes performances des modèles RMCP que nous avons calibrés, seraient en partie attribuables à l'homogénéité des échantillons de calibration. En effet, ces performances sont nettement supérieures à ce que l'on observe sur des échantillons de sols naturels qui se différencient entre eux par de nombreux paramètres. En effet, les prédictions des constituants du sol par spectrométrie peuvent échouer si les échantillons sont collectés à partir d'une large aire géographique ou à partir de différents contextes morphopédologiques (Henderson et *al.*, 1992 ; Schulze et *al.*, 1993).

Henderson, et *al.* (1992) ainsi que Chen et *al.* (2000) expliquent cet échec par la différence entre la parenté des matériaux. Morra et *al.* (1991) signalent qu'en spectrométrie VIS-PIR, il est nécessaire d'utiliser une population d'échantillons qui soit bien définie et bien homogène. Ces mêmes auteurs indiquent que le degré d'homogénéité des échantillons de calibration est une limite inhérente à l'utilisation de l'outil spectrométrique et ce quel que soit le constituant que l'on souhaite prédire.

Dans notre cas, en considérant au départ chacune des séries séparément, nous analysons une population homogène où la réponse spectrale est influencée par les deux variables d'intérêts. Par la suite, même dans le cas du modèle global, qui est calibré sur l'ensemble des échantillons des deux séries, la population reste quasi-homogène en raison d'une texture et d'une matrice minérale très proches. L'amélioration des performances par rapport aux deux séries considérées séparément, est dans ce cas liée à l'augmentation de la taille de la population. Toutefois, il est intéressant de constater que les modèles type RMCP intègrent la variabilité des réponses spectrales en tant que données d'entrée du modèle et c'est ce qui en fait un outil très intéressant pour étudier des populations très hétérogènes. En effet, Ludovic et *al.* (1995), soulignaient qu'en modélisation « le modèle doit suivre les données et non l'inverse ».

5. Conclusion

La spectroscopie est un outil fort intéressant pour l'étude des sols. Elle peut être mise à profit pour caractériser qualitativement des couvertures pédologiques, par simple examen de la forme des spectres (richesses relatives en fer total, en carbone organique, *etc*). De même des mesures spectroscopiques diachroniques, effectuées sur une même couverture, permettraient d'apprécier la tendance évolutive de la qualité des sols sans avoir recours à des mesures analytiques de laboratoire, sauf en certaines périodes pour avoir des mesures de référence.

L'ajustement de modèles de régressions en moindres carrés partiels (RMCP) aux données spectrométriques permet de prédire les teneurs en certains constituants du sol. Cette technique est peu coûteuse, rapide et fiable.

Les modèles type RMCP intègrent implicitement l'hétérogénéité des échantillons de calibration. De même la qualité de prédiction s'améliore lorsque l'effectif des échantillons de calibration augmente.

Chapitre III. Comparaison des potentiels du visible, du proche infrarouge et du combiné visible-proche infrarouge pour la prédiction simultanée en mode régional de plusieurs propriétés du sol en milieu tempéré (Région de Bretagne, France)

H. Aïchi^{1,2*}, Y. Fouad¹, C. Walter¹, Z. Lili Chabaane², M. Sanaa², R.A. Viscarra Rossel³, R. Roy⁴ et P. Launeau⁴,

¹Agrocampus Rennes-UMR SAS, 65 rue de St Brieuc, CS 84215, 35042 Rennes, France

²Institut National Agronomique de Tunisie, Tunisie

³Australian Centre for Precision Agriculture, Faculty of Agriculture, Food & Natural Resources, McMillan Building A05, University of Sydney, NSW 2006

⁴UMR-CNRS 6112 Laboratoire Planétologie et Géodynamique, Faculté des Sciences et des Techniques, Université de Nantes, France

Abstract

Spectroscopy is promising in simultaneously quantifying several pedological properties. Our objective is triple : i) Calibrating regional predictive models for 10 physico-chemical properties, ii) using the Principal Components Analysis (PCA), applied to the 10 properties to select samples of calibration and of prediction, iii) Comparing, for every property, between models fitted to the spectra of combined Vis-NIR, Vis and NIR. 125 samples, collected from the superficial horizon in the region of Brittany (France), have been analyzed for : the organic carbon (CO (%)), the total nitrogen (Ntot (%)), the pH1:5(H₂O), the cationic exchange capacity (CEC (cmol/kg), the total iron (Fetot (%)) and the five texture fractions (clay (Ar (%)), fine silt (LF (%)), coarse silt (LG (%)), fine sand (SF (%)) and coarse sand (SG (%)). Through the use of the ParLeS software (3.1 version), spectroscopic and analytic data have been related by partial least square regressions. In the prediction stage, models of Ntot, pH, Fetot, LF and SF, are unsatisfactory for the three spectra range. For the CO and the CEC, predictions are adequate for the three spectral domains, although the NIR offers the best predictions, respectively (Coefficient of determination $R^2 = 0.831$, Ratio of prediction deviation RPD = 2.42) and ($R^2 = 0.69$, RPD = 1.81). For the LG and the SG, for the three spectral domains, the RPD are between 1.8 and 1.4. These models can be used to appraise these two fractions. Finally for the Ar, the REP of the three spectral domain models, are between 0.99 and 1.2. These models can be used to differentiate between the rich and poor clay samples. The correct predictions of the CO can be assigned to the direct influence of carbon bonds on spectra. However, the satisfactory predictions of the CEC, the LG and the SG are likely due to their correlation with the CO ($r = 0.94$, -0.65 and 0.55). The adequate models of CO at a regional scale, obtained by using PCA as samples selection procedure, open the way to calibrate universal models.

Key words : spectroscopy, visible, near-infrared, PLS, PCA, organic carbon.

Résumé

La spectrométrie est prometteuse pour quantifier simultanément plusieurs propriétés pédologiques. Notre objectif est triple : i) Calibrer des modèles régionaux prédictifs de 10 propriétés physico-chimique, ii) utiliser l'Analyse en Composante Principale (ACP), appliquée sur les 10 propriétés comme procédure de sélection des échantillons de calibration et de prédiction, iii) Comparer, pour chaque propriété, entre les modèles ajustés aux spectres du : combiné Vis-PIR, Vis et PIR. 125 échantillons, collectés de l'horizon superficiel en région de Bretagne (France), ont été analysés pour : le carbone organique (CO (%)), l'azote total (Ntot (%)), le pH_{1.5} (H₂O), la capacité d'échange cationique (CEC (cmol/kg)), le fer total (Fetot (%)) et les cinq fractions granulométriques (argile (Ar (%)), limon fin (LF (%)), limon grossier (LG (%)), sable fin (SF (%)) et sable grossier (SG (%))). A l'aide du logiciel ParLeS (version 3.1), les données spectrométriques et analytiques ont été liées par des régressions en moindres carrés partiels. Au stade prédiction, les modèles de Ntot, pH, Fetot, LF et SF, tous spectres confondus, sont insatisfaisants. Les prédictions sont adéquates pour le CO et la CEC dans les trois domaines spectraux, quoique le PIR offre les meilleurs prédictions, respectivement (coefficient de détermination $R^2 = 0.831$, ratio des écarts de prédiction REP = 2.42) et ($R^2 = 0.69$, REP = 1.81). Pour les trois domaines spectraux les REP du LG et du SG sont comprises entre 1.8 et 1.4. Ces modèles peuvent être utilisés pour estimer ces deux fractions. Enfin pour l'Ar les REP des modèles des trois domaines spectraux, sont comprises entre 0.99 et 1.2. Ces modèles peuvent être utilisés pour différencier entre les échantillons riches et pauvres en argiles. Les bonnes prédictions du CO peuvent être attribuées à l'influence directe des liaisons carbonées sur les spectres. Par contre les prédictions satisfaisantes de la CEC, du LG et du SG seraient dues à leurs corrélations avec le CO respectivement ($r = 0.94$, -0.65 et 0.55). Les modèles de CO à échelle régionale adéquats, obtenus grâce à la sélection des échantillons par l'ACP, ouvre la voie pour calibrer des modèles plus globaux voir universel.

Mots clés : spectrométrie, visible, proche infrarouge, RMCP, ACP, carbone organique.

1. Introduction

La spectrométrie est une technique quantitative prometteuse pour l'évaluation simultanée de plusieurs propriétés physico-chimique et biologique du sol (Ehsani et *al.*, 1999 ; Chang et *al.*, 2001, Coûteaux, 2003, Brown et *al.*, 2005, Terhoeven-urselmans et *al.*, 2006). Cette discipline est issue de la science de télédétection (Escadafal et *al.*, 1993, Ben-Dor, 2002), la possibilité de la réaliser directement grâce aux spectroradiomètres portables a permis son utilisation dans la quantification des différentes propriétés des sols (Ben Dor et *al.*, 1997 ; Mathieu et *al.*, 1998 ; Blavet et *al.*, 2000 ; Weidung et *al.*, 2002 ; Batchily et *al.*, 2003 ; Mouazen et *al.*, 2007).

La spectroscopie est basée sur les propriétés d'absorption de la lumière par les constituants du sol. Dans le visible et le proche infrarouge, ces absorptions correspondent à des harmoniques et à des combinaisons associées à des bandes fondamentales qui se trouvent dans le moyen infrarouge (Chang et *al.*, 2001 ; Brown et *al.*, 2005 ; Cozzolino et Moron, 2006 ; Brunet et *al.*, 2007). Les spectres des sols sont aussi influencés par sa structure physique : la taille, la forme des particules et leur arrangement spatial ainsi que les pores interstitiels conditionnent le cheminement de la lumière à travers le sol et par conséquent les parts d'énergies absorbée et réfléchi (Bower et Hanks, 1965 ; Dalal et Henry, 1986). Une récapitulation des facteurs qui influencent la réflectance du sol est donnée par Ben-Dor et *al.* (1999).

Plusieurs auteurs s'accordent sur le fait que les mesures spectroradiométriques sont faciles à mettre en œuvre, rapides, répétitives, non destructives et peu onéreuses (Dalal et Henry 1986 ; Ben Dor et Banin, 1995 ; Morra et *al.*, 1991 ; Kooistra et *al.*, 1993 ; Chang et *al.*, 2001 ; Islam et *al.*, 2003 ; Mouazen et *al.*, 2007 ; He et *al.*, 2007). Par ailleurs, les échantillons ne nécessitent pas une préparation particulière avant les mesures des spectres de réflectances (Wetterlind et *al.*, 2008). Les relations entre la réflectance spectrale et les propriétés du sol telles que les fractions granulométriques, l'humidité, les oxydes de fer, les teneurs en carbonates, les teneurs en matières organiques ont été établis à l'aide de modèles statistiques dans plusieurs études (Krishnan et *al.*, 1980; Kosmas et *al.*, 1984; Ben-Dor et Banin, 1995a, b ; Jarmer et Schütt, 1998 ; Jarmer, 2005 ; Farifteh et *al.*, 2008). A ces titres la pédo-spectrométrie peut être utilisée comme une méthode alternative ou auxiliaire aux méthodes d'analyses physico-chimiques voir biologiques conventionnelles du laboratoire (Couillard et *al.*, 1997). Toutefois, des auteurs comme Terhoeven-Urselmans et *al.* (2006) estiment que l'utilité et les limites d'utilisation de la spectrométrie Vis-PIR dans la prédiction quantitative des propriétés du sol restent insuffisamment explorées.

En effet, les aspects relatifs à : i) la détermination de la portée spatiale (champ géographique) d'application des modèles calibrés, ii) la procédure de sélection des échantillons de calibration et de prédiction, iii) le choix de la fenêtre spectrale à utiliser dans la prédiction n'ont pas été largement débattus dans la littérature.

Primo, le plus souvent les modèles calibrés sont à portée géographique limitée. Ils ne peuvent pas être utilisés pour des prédictions en dehors du contexte pédologique de leurs calibrations (Brown et al., 2006). Islam et al. (2003) affirment que la calibration des modèles prédictifs sur un nombre limité d'échantillons ou sur des échantillons assez homogènes peut limiter le champ d'application du modèle à son contexte pédologique de calibration. Par ailleurs, Sudduth et Hummel (1996) affirment que la qualité de prédiction des modèles spectrométrique se détériore lorsque l'étendue géographique de l'origine des échantillons de calibration s'élargit. De surcroît, Sorensen et Dalsgaard (2005) mentionnent que l'utilisation des modèles calibrés pour la prédiction des teneurs d'échantillons ayant des caractéristiques organo-minérales différentes des échantillons de calibration peut se solder par un échec et ce malgré que la teneur en le constituant à prédire soit dans la gamme des teneurs du jeu de calibration

Secondo, la procédure standard de sélection des échantillons consiste à subdiviser les échantillons, dont on peut disposer, en une série de calibration et une série de prédiction en utilisant un générateur de nombres randomisés (Brown et al., 2005) qui fournit un mécanisme de sélection aléatoire. Par ailleurs, il est plus ou moins convenu de départager les échantillons dans les proportions : deux tiers voir trois quarts de la population des échantillons pour la calibration et un tiers voir un quart de la population des échantillons pour la prédiction. Ces fractions ne paraissent pas être basées sur des fondements empiriques (Mc Bratney et al., 2003). La sélection des échantillons par randomisation ignore l'impact des aléas du choix des échantillons sur les performances de prédiction. Aucune garantie n'est offerte quant à la similarité entre les échantillons de calibration, sur lesquelles le modèle a été construit, et les échantillons de prédiction. Par conséquent, les performances du modèle seront aussi aléatoires. Ainsi une bonne ou une mauvaise qualité de prédiction peut être un artéfact de la procédure de sélection des échantillons beaucoup plus qu'une représentation réelle de la précision et de la fiabilité du modèle. Des auteurs comme Shenk et Westerhaus (1991) et Dardenne et al. (2000) signalent qu'il est nécessaire que les échantillons de prédiction soient similaires à ceux de la calibration. En outre, Ludwig et al. (2002) qui ont utilisé la spectrométrie PIR pour la prédiction, entre autre, des teneurs en carbone organique et en azote total, se sont rendus compte que le choix qu'ils ont fait de sélectionner les échantillons de calibration par simple randomisation n'était pas le meilleur car aussi bien le jeu de

calibration que le jeu de prédiction présentait des échantillons atypiques qui dégradent la qualité de prédiction.

Tertio, dans la littérature nous constatons que pour prédire les diverses propriétés physico-chimique du sol des auteurs comme : Ben-Dor et Banin (1994), Demattê et *al.* (2004), Brunet et *al.* (2007) utilisaient le combiné Vis-PIR. Alors que des auteurs comme : Dalal et Henry (1986), Ben Dor et Banin (1990), Viscara et McBratney (1998), Ehsani et *al.* (1999), Hummel et *al.* (2001), Sorensen et Dalsgaard (2005), He et *al.* (2007) utilisaient seulement le PIR. A ce titre il est opportun de mener pour un même jeu d'échantillon et pour les mêmes propriétés pédologiques une étude comparative des performances des domaines spectraux : Vis, PIR et combiné Vis-PIR.

Dans cette étude nous nous sommes mis devant la difficulté de constituer une base régionale d'échantillon de sol avec une large variabilité des propriétés analytiques ce qui permet de couvrir la diversité des sols rencontrés dans la région et ce pour que les modèles calibrés peuvent servir pour des prédictions à une échelle régionale.

Dés lors l'objectif de ce travail est triple. Nous nous proposons de : i) Calibrer des modèles régionaux de prédiction de 10 propriétés chimiques et physiques du sol, ii) utiliser la procédure d'Analyse en Composante Principale (ACP), appliquée sur les 10 variables analytiques, comme une procédure raisonnée de sélection des échantillons de calibration et de prédiction. iii) Comparer, pour chaque pédo-variable, entre les performances des modèles de prédiction, ajustés aux données spectrométrique acquises successivement dans les spectres du : combiné Vis-PIR, Vis et PIR.

2. Matériels et méthodes

2.1. Constitution d'une base régionale d'échantillons de sol

Pour couvrir la variabilité des sols de la région de Bretagne (France), un total de 125 échantillons, a été collectés de l'horizon superficiel dans des sites ayant des contextes pédologiques très différents et spatialement distants. La figure 1 montre les sites de prélèvement de ces échantillons.

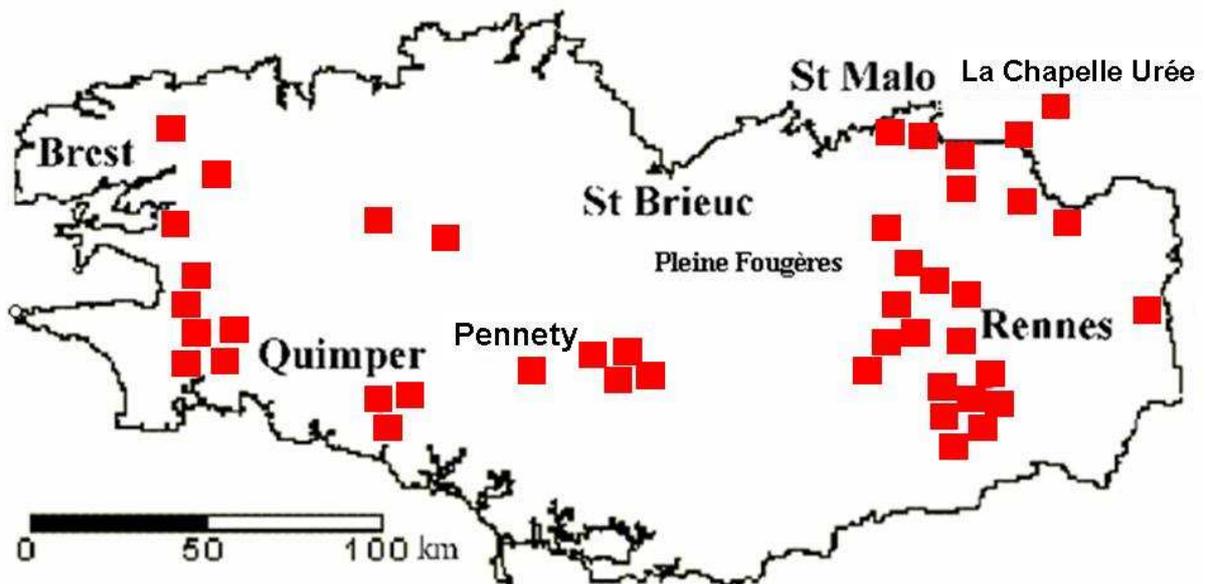


Figure 1. Localisation des sites de prélèvement des échantillons en différents sites de la région de Bretagne.

Les échantillons ont été séchés et tamisés à 2 mm puis analysés au laboratoire selon des normes analytiques standard. Les analyses ont intéressé les 10 paramètres analytiques suivants : la teneur en carbone organique (CO (%)), la teneur en azote total (Ntot (%)), le pH_{1.5} (H₂O), la capacité d'échange cationique (CEC (cmol/kg)), la teneur en fer total (Fetot (%)) ainsi que les cinq fractions granulométriques (argile (Ar (%)), limon fin (LF (%)), limon grossier (LG (%)), sable fin (SF (%)) et sable grossier (SG (%)).

2.2. Procédure de sélection des échantillons de calibration et de prédiction du modèle

Pour sélectionner les échantillons de calibration et de prédiction, nous nous sommes inspirés de la procédure utilisée par Islam et *al.* (2003). Toutefois au lieu d'appliquer une Analyse en Composante Principale (ACP) sur les spectres de réflectance des échantillons, comme ces auteurs l'ont fait. Nous avons appliqué une analyse en composante principale sur les 10 variables analytiques, pour les quelles les échantillons ont été caractérisés. Les échantillons ont été projetés dans le premier plan factoriel de l'ACP. De la sorte, la similarité qui existe entre les échantillons a été explorée. Nous avons pu d'une part identifier les échantillons atypiques et d'autre part partager le reste des échantillons en un jeu de calibration et un jeu de prédiction.

2.3. Analyse spectrale des échantillons

Les données spectrométriques ont été acquises à l'aide d'un spectroradiomètre de type GER 3700 dans la gamme 300-2500 nm (Mélédère et *al.*, 2003). Comme la série des Ultra-Violet UV 300-400 nm ne présente aucun intérêt pour notre étude, elle a été tronquée des spectres. Seul le domaine spectral du Vis-PIR 400-2500 nm a été retenu soit 576 canaux. L'acquisition des spectres a été réalisée, en visée verticale, dans une chambre noire avec des conditions d'illumination standardisées. Deux lampes halogènes, d'une puissance de 300 W chacune, ont été fixées de part et d'autre du capteur et à la même hauteur que celui-ci. A l'aide d'un pyranomètre, positionné à la place de l'échantillon, nous avons orienté les lampes de façon à avoir le maximum d'énergie incidente.

Les échantillons ont été séchés et tamisés à 2 mm puis placés dans des boîtes de Pétri de 6,4 cm de diamètre intérieur, préalablement peintes en noir pour éviter toute source d'énergie autre que celle en provenance de l'échantillon du sol lui-même. Ensuite, ces échantillons ont été mis à l'étuve, pendant 24 h, à une température de 45 °C pour homogénéiser leur taux d'humidité (Demattê, 2002).

La réflectance apparente des échantillons est obtenue en mesurant successivement la luminance de la cible de référence puis celle de l'échantillon. La cible de référence est un spectralon® (23x23 cm²). Pour chaque échantillon, nous avons relevé 40 spectres, 10 spectres par position, en tournant après chaque 10 mesures l'échantillon d'un angle de 90°. C'est le spectre moyen des 40 mesures qui a été considérée dans les traitements. De la sorte, nous nous sommes affranchis de l'effet directionnel dû à l'hétérogénéité de la surface (Karnieli et *al.*, 1999) et nous avons compensé les éventuels défauts de symétrie du dispositif d'illumination.

2.4. Construction du modèle statistique de prédiction

2.4.1. Régression en Moindres Carrés Partiels (RMCP)

En spectrométrie, le nombre de variables explicatives qui correspond au nombre de longueurs d'onde est très important d'autant plus que ces variables sont très corrélées (Desbois, 1999). Il n'est pas donc opportun d'ajuster à ces données des modèles de régressions multiples (Abdi, 2003). Pour contourner ce problème et pouvoir développer des modèles de régressions linéaires multiples, les pédométriciens ont toujours été obligés d'avoir recours à des techniques empiriques de manipulation du spectre : transformation de la

réflectance en absorbance, dérivation première et seconde du signal, moyennement et élimination de bandes spectrales, (Krishnan et *al.*, 1980 ; Dalal et Henry 1986 ; Ingleby et Crowe, 2000) ou à des procédures algorithmiques : régression pas à pas (Weidung et *al.*, 2002) pour sélectionner quelques longueurs d'onde, généralement, 2 à 5 parmi des centaines.

D'ores et déjà, il existe des techniques statistiques plus appropriés pour ajuster un modèle à des données spectrométriques, Il s'agit notamment de la régression en moindres carrés partiels RMCP. La régression MCP présente deux avantages majeurs :

- i) toutes les variables sont conservées, les variables importantes dans la prédiction ont des poids élevés et les variables sans importance participent au modèle, mais avec un poids faible (Tenenhaus, 1998).
- ii) le nombre de variables peut être supérieur au nombre d'échantillons du jeu de calibration (Davies, 2001).

La régression MCP univariée relie une variable à expliquer Y à un bloc de variables explicatives X (Tenenhaus, 1998) et ce en passant par des composantes latentes. Les composantes latentes sont des combinaisons linéaires des longueurs d'onde initiales (Desbois, 1999). Elles sont complètement décorrelés et expliquent le plus possible la covariance entre chacune des variables explicatives X_i et Y (Sundberg, 1999). A cet effet elles sont présentées comme un compromis entre la structure interne des variables explicatives et leur lien avec la variable à expliquer (Druilhet et Mom, 2006).

Pour chaque pédo-variable trois modèles de Régression en Moindres Carrés Partiels univarié (RMCP1) ont été calibrés et comparés et ce en utilisant successivement : le spectre du combiné Vis-PIR (400-2500 nm), uniquement le spectre du Vis (400-700 nm) et uniquement le spectre du PIR (701-2500 nm).

2.4.2. Choix du nombre de composantes latentes des modèles

Le choix du nombre de composantes latentes a été à chaque fois effectué en appliquant une validation croisée sur le jeu de calibration (Tenenhaus, 1998). Le modèle qui présente la racine de l'erreur quadratique moyenne (REQM) la plus faible, en validation croisée, est retenu et son équation est utilisée au stade prédiction.

2.4.3. Evaluation des performances des modèles

En s'inspirant des travaux de Gauch et *al.* (2003), Islam et *al.* (2003), Brown et *al.* (2005), Viscarra et *al.* (2006b), Brunet et *al.* (2007), l'adéquation de chaque modèle calibré a été évaluée en calculant les paramètres statistiques suivants :

- le coefficient de détermination (R^2)
- le biais (erreur moyenne)
- la racine de l'erreur quadratique moyenne (REQM)
- le ratio des écarts de prédiction (REP). (Ratio of Prediction Deviation (RDP) en anglais)

Ces paramètres sont définis comme suit :

$$R^2 = \left[\frac{\sum_{i=1}^n (y_i - \bar{Y})(y_i^* - \bar{Y}^*)}{\sqrt{\sum_{i=1}^n (y_i - \bar{Y})^2 \sum_{i=1}^n (y_i^* - \bar{Y}^*)^2}} \right]^2 \quad [1]$$

$$Biais = \frac{1}{n} \sum_{i=1}^n (y_i^* - y_i) \quad [2]$$

$$REQM = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i^* - y_i)^2} \quad [3]$$

$$REP = \sigma y_i^* / \sigma \mathcal{E}_i \quad [4]$$

y_i : valeur mesurée, y_i^* : valeur estimée, \bar{Y} : moyenne des valeurs mesurées, \bar{Y}^* : moyennes des valeurs prédites, σy_i^* : écart-type des valeurs prédites, $\sigma \mathcal{E}_i$: écart-type des erreurs de prédiction.

2.4.4. Traitement des données

Les données spectrométriques et analytiques ont été mises en forme et importées, sous forme de fichier texte (format txt), dans le logiciel ParLeS (version 3.1). ParLes est un logiciel libre d'analyses chimiométriques, dédié aux traitements des données spectrométrique. Une description détaillée de ce logiciel et de la panoplie des pré-traitements et traitements qu'il offre est reportée par Viscarra, R (2008a).

Avant de les mettre en relation avec les données des propriétés analytiques, les spectres acquis en mode réflectance (R) ont été transformés en absorbance A selon l'équation 5 :

$$A = \log(1/R) \quad [5]$$

3. Résultats et discussion

3.1. Statistiques descriptives des propriétés analytiques

Les statistiques descriptives des propriétés analytiques des 125 échantillons collectés sont succinctement présentées dans le tableau 1.

Tableau 1. Statistiques sommaires sur les propriétés analytiques des échantillons

Propriétés	Effectif	Minimum	Maximum	Moyenne	Ecart-type
CO (%)	125	0.82	8.90	1.92	1.19
Ntot (‰)	123	0.09	6.07	1.36	0.98
pH _{1.5} (H ₂ O)	125	4.46	8.03	6.26	0.75
CEC (cmol/Kg)	125	6.03	26.10	8.84	3.14
Fetot (%)	124	0.27	0.78	0.38	0.10
Ar (%)	125	11.20	32.80	16.22	3.45
LF (%)	125	11.80	41.10	23.06	4.56
LG (%)	125	9.30	50.90	33.87	11.75
SF (%)	125	4.30	20.00	11.91	2.62
SG (%)	125	3.40	57.40	14.92	12.25

Encore faut-il signaler que deux échantillons n'ont pas été analysés pour leurs teneurs en azote total et un échantillon n'a pas été analysé pour sa teneur en fer total.

D'après le tableau 1, nous constatons que la variabilité au sein des 125 échantillons est très importante et ce pour l'ensemble des propriétés analytiques. A titre indicatif, pour le carbone organique les échantillons couvre une large gamme de teneur qui varie entre 0.821 et 8.901 %. Ainsi la valeur maximale est plus que dix fois plus importante que la valeur minimale. La moyenne des teneurs en carbone est de 1.922 % avec un écart type de 1.192 %. Ces valeurs sont compatibles avec les teneurs en carbone organiques que nous pouvons rencontrer dans la région (Lemercier et *al.*, 2006).

3.2. Corrélations entre les propriétés analytiques

Le tableau 2 résume les corrélations, issus de l'ACP, qui existent entre les 10 propriétés analytiques.

Tableau 2. Matrice des corrélations entre les propriétés analytiques

	CO	Ntot	pH _{1.5}	CEC	Fetot	Ar	LF	LG	SF	SG
	(%)	(‰)	(H ₂ O)	(cmol/kg)	(%)	(%)	(%)	(%)	(%)	(%)
CO (%)	1.00									
Ntot (‰)	0.57	1.00								
pH _{1.5} (H ₂ O)	-0.44	-0.09	1.00							
CEC (cmol/kg)	0.94	0.55	-0.41	1.00						
Fetot (%)	0.47	0.12	-0.42	0.62	1.00					
Ar (%)	0.30	0.38	-0.13	0.39	0.50	1.00				
LF (%)	-0.10	0.14	0.07	-0.05	0.33	0.46	1.00			
LG (%)	-0.65	-0.42	0.51	-0.67	-0.55	-0.37	0.14	1.00		
SF (%)	0.14	0.26	0.01	0.11	-0.12	-0.28	-0.38	-0.21	1.00	
SG (%)	0.55	0.19	-0.48	0.52	0.29	-0.04	-0.55	-0.86	0.21	1.00

Dans notre jeu d'échantillon régional, les corrélations les plus remarquables sont les suivantes :

- La teneur en carbone organique est fortement corrélée, positivement, à la CEC ($r = 0.94$). De même une corrélation négative existe entre la teneur en carbone organique et la fraction de limon grossier ($r = -0.65$).
- L'azote total est moyennement corrélé à la teneur en carbone organique ($r = 0.57$). L'azote est aussi moyennement corrélé à la CEC ($r = 0.55$). Le carbone et l'azote constituent les deux éléments majeurs de la matière organique du sol et les teneurs en carbone organique et en azote organique sont le plus souvent hautement corrélés. Toutefois dans le cas d'espèce nous avons mesuré l'azote total (azote organique + azote minérale) et non l'azote organique. À l'échelle de la région l'azote total est peu corrélé au carbone organique à cause des épandages réguliers d'azote minérale pour les besoins des cultures intensives.
- Le fer total est le plus corrélé à la CEC ($r = 0.62$).
- Les fractions : argile, limon fin et sable fin ne montrent aucune inter-corrélation qui soit remarquable. Dans la littérature la fraction argile serait beaucoup plus corrélée à la teneur en carbone organique et à la CEC
- Le limon grossier est négativement corrélé à la CEC ($r = -0.67$).
- La fraction de sable grossier est fortement corrélée, négativement, à la fraction de limon grossier ($r = -0.86$).

3.3. Sélection des échantillons de calibration et de validation

La figure 2 représente les résultats de la projection des 125 échantillons dans le premier plan factoriel de l'ACP.

Facteur II : (21.66%)

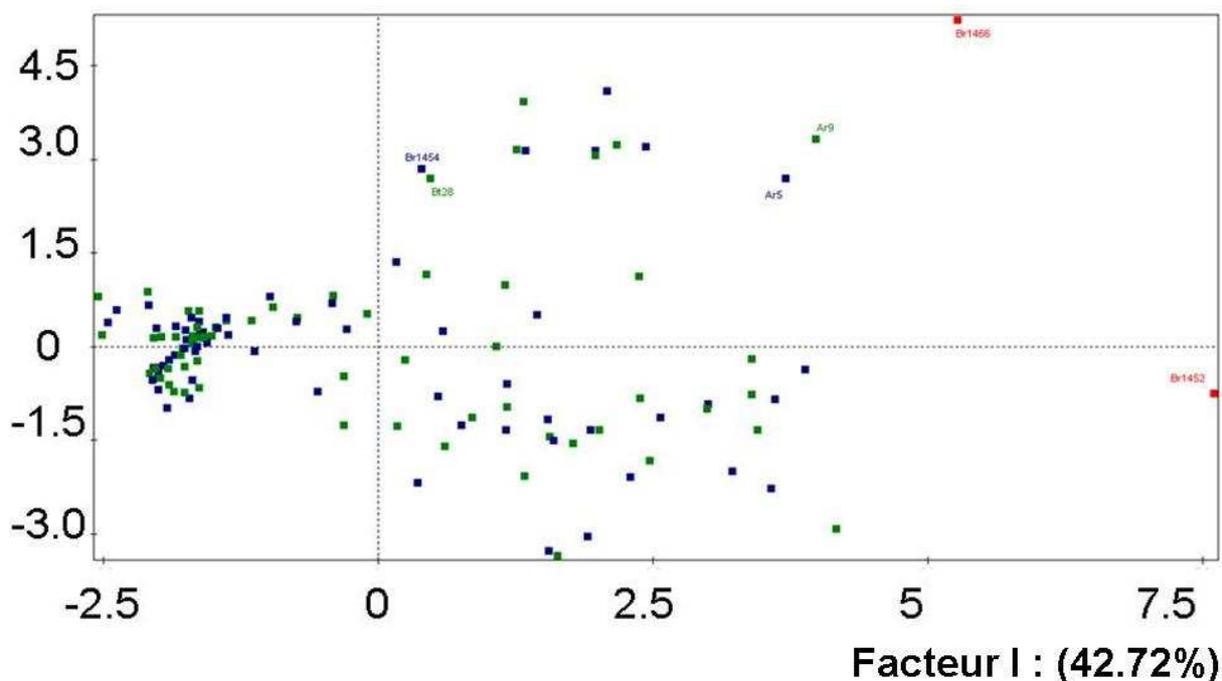


Figure 2. Sélection des échantillons de calibration et de validation par projection des 125 échantillons de sols dans le premier plan factoriel de l'ACP. (Carré rouge : échantillon atypique, carré bleu : échantillon du jeu de calibration, carré vert : échantillon du jeu de validation).

Nous pouvons noter que le premier plan factoriel explique 64.38 % de la variabilité analytique des échantillons. Rappelons que les axes factoriels sont des combinaisons linéaires des variables initiales (les propriétés analytiques). Le premier axe factoriel explique 42.72 % alors que le deuxième axe explique 21.66 % de cette variabilité. D'après cette figure, deux échantillons semblent être atypiques. Il s'agit de l'échantillon Br1452 et de l'échantillon Br1466. Ces deux derniers sont très loin du centre du plan factoriel et très loin du reste des échantillons. A cet effet ils ont été soustraits de la base des échantillons.

Le tableau 3 reprend les analyses de ces deux échantillons en juxtaposition avec la moyenne des 125 échantillons. Il est évident que l'échantillon Br1452 se différencie par ces propriétés chimiques car il présente des valeurs de carbone organique, d'azote total et de CEC qui sont très en dessus de la moyenne des 125 échantillons.

De même pour l'échantillon Br1466 qui se différencie par ces propriétés physique car il a des valeurs très en dessus de la moyenne pour la fraction argileuse et très en dessous de la moyenne pour les fractions de limon grossier et de sable fin.

Tableau 3. Analyses des échantillons atypiques

Propriétés	Moyenne	Br1452	Br1466
CO (%)	1.92	8.90	3.08
Ntot (‰)	1.36	6.07	3.74
pH _{1:5} (H ₂ O)	6.26	5.38	4.89
CEC (cmol/Kg)	8.84	26.1	16.25
Fetot(%)	0.38	-	0.7
Ar (%)	16.22	18.3	32.8
LF (%)	23.06	17.7	29.9
LG (%)	33.87	22.2	13.4
SF (%)	11.91	14.7	4.3
SG (%)	14.92	27.1	19.6

Le reste des échantillons a été regroupé en couples de points en fonction de leur proximité dans le premier plan factoriel et ce en examinant visuellement les distances euclidiennes qui les séparent, puis répartis l'un dans le jeu de calibration et l'autre dans le jeu de prédiction. C'est le cas par exemple, comme montré dans la figure 2 des couples des échantillons (Ar5 et Ar9) et (Br1454 et Bt28). Après avoir retiré les deux échantillons atypiques, nous nous sommes trouvés avec 62 échantillons pour le jeu de calibration et 61 échantillons pour le jeu de prédiction. Les deux jeux sont analytiquement similaires et ils couvrent les mêmes gammes de teneurs pour l'ensemble des propriétés analytiques. En effet, d'après les travaux de Tillman et *al.* (2001) cités par Terhoeven-Urselmans et *al.* (2006) la qualité de prédiction est médiocre lorsque les prédictions sont faites sur des échantillons qui sont en dehors de la gamme des échantillons de calibration. De même la précision de prédiction se détériore avec l'élargissement de la gamme des échantillons de calibration.

Les tableaux 4 et 5 montrent qu'à l'aide de l'ACP nous avons pu partager l'ensemble des échantillons en un jeu de calibration (formé de 62 échantillons) et un jeu de prédiction (formé de 61 échantillons) qui sont significativement similaires pour la totalité des propriétés analytiques. Comme l'atteste les valeurs des minimums et des maximums, les deux jeux de données couvrent, pour la totalité des paramètres, quasiment les mêmes gammes de valeurs.

Les deux jeux présentent, pour les 10 pédo-variables, presque les mêmes paramètres de position à savoir les moyennes et les mêmes paramètres de dispersion à savoir l'écart-type. Par conséquent leurs coefficients de variations sont de mêmes ordres de grandeurs.

Ces coefficients sont d'ailleurs quasi identiques pour les variables : carbone organique, Azote total, CEC, Fer total, limons fins, limons grossiers et sable grossiers. Islam et *al.* (2003) qui ont utilisé la même procédure de sélection des échantillons à savoir L'ACP mais qui l'ont appliqué sur les données des spectres au lieu de l'appliquer sur les données des attributs de leurs échantillons de sols se sont trouvés avec un jeu de calibration et un jeu de prédiction qui sont dans la même gamme pour certaines variables et dans des gammes différentes pour certaines autres. D'où l'intérêt d'appliquer l'ACP sur les variables analytiques au lieu des spectres dans la procédure de la sélection des échantillons de calibration et de prédiction.

Tableau 4. Statistiques descriptives des échantillons de calibration.

Libellé de la variable	Effectif	Minimum	Maximum	Moyenne	Ecart-type	Coefficient de variation
CO (%)	62	0.87	4.25	1.81	1.00	1.81
Ntot (‰)	62	0.10	3.09	1.30	0.86	1.50
pH _{1.5} (H ₂ O)	62	4.56	7.45	6.27	0.71	8.72
CEC (cmol/Kg)	62	6.03	15.79	8.61	2.65	3.24
Fetot (%)	62	0.28	0.78	0.38	0.10	3.82
Ar (%)	62	11.2	28.9	16.20	3.51	4.60
LF (%)	62	11.8	39.7	23.20	4.54	5.10
LG (%)	62	9.3	50.9	33.59	11.33	2.96
SF (%)	62	5.9	17.7	11.96	2.35	5.07
SG (%)	62	4.3	56.7	15.02	11.75	1.27

Tableau 5. Statistiques descriptives des échantillons de validation.

Libellé de la variable	Effectif	Minimum	Maximum	Moyenne	Ecart-type	Coefficient de variation
CO (%)	61	0.82	4.72	1.89	1.02	1.85
Ntot (‰)	59	0.09	3.43	1.31	0.86	1.52
pH _{1.5} (H ₂ O)	61	4.46	8.03	6.28	0.77	8.05
CEC (cmol/Kg)	61	6.09	16.55	8.66	2.68	3.22
Fetot (%)	61	0.27	0.69	0.37	0.09	3.99
Ar (%)	61	11.2	26.7	15.93	2.68	5.92
LF (%)	61	12.4	41.1	22.89	4.51	5.07
LG (%)	61	10.7	50.2	34.69	11.93	2.90
SF (%)	61	6.5	20	11.92	2.72	4.37
SG (%)	61	3.4	57.4	14.54	12.82	1.13

Le recours à L'ACP a rendu possible :

- l'étude des corrélations inter variables analytiques.
- la visualisation de la ressemblance et de la dissemblance des échantillons en fonction de leurs propriétés analytiques et par conséquent l'identification et par la suite la soustraction des échantillons atypiques de la base des échantillons
- le partage raisonné des échantillons, en un jeu de calibration et un jeu de prédiction. A ces titres l'ACP peut être utilisée comme une procédure standard de sélection des échantillons.

En effet, Morra et *al.* (1991) affirment que l'une des limites inhérentes à l'utilisation des prédictions spectrométriques de n'importe quel constituant du sol, est la nécessité de calibrer le modèle sur un jeu d'échantillons analytiquement bien défini ou d'utiliser une population d'échantillons qui soit homogène.

3.4. Performances des modèles

Le tableau 6 récapitule les performances des modèles aux stades de calibration (validation croisée) et de prédiction. Notons que le Ratio des écarts de prédictions REP est défini comme le rapport entre : L'écart type des valeurs analytiques mesurées au laboratoire pour le jeu de données (selon le cas : jeu de calibration ou jeu de prédiction) et l'écart type des valeurs des erreurs résiduelles du même jeu d'échantillons (calibration ou prédiction) (Starr et *al.*, 1981)

Tableau 6. Performances des modèles aux stades de calibration et de prédiction (manque les unités)

		Calibration					Prédiction				
		NCL	R ²	REQM	Biais	REP	R ²	REQM	Biais	REP	
CO (%)	Vis-PIR	4	0.664	0.58	-0.003	1.74	0.806	0.454	-0.053	2.27	
	Vis	8	0.723	0.53	-0.016	1.90	0.739	0.53	-0.029	1.93	
	PIR	4	0.664	0.58	-0.011	1.74	0.831	0.427	-0.037	2.42	
Ntot (‰)	Vis-PIR	8	0.493	0.637	0.022	1.37	0.278	0.812	0.083	1.09	
	Vis	8	0.644	0.523	0.000	1.67	0.175	0.958	0.127	0.93	
	PIR	4	0.376	0.693	0.010	1.26	0.177	0.824	0.049	1.08	
pH _{1.5} (H ₂ O)	Vis-PIR	6	0.573	0.481	-0.002	1.51	0.319	0.669	0.019	1.18	
	Vis	8	0.516	0.53	-0.012	1.36	0.269	0.721	0.034	1.09	
	PIR	17	0.179	0.828	0.07	0.88	0.145	1.075	-0.274	0.73	
CEC (cmol/Kg)	Vis-PIR	3	0.608	1.672	-0.000	1.60	0.674	1.54	-0.118	1.76	
	Vis	2	0.627	1.623	-0.004	1.65	0.671	1.54	-0.06	1.76	
	PIR	3	0.618	1.643	-0.007	1.63	0.691	1.50	-0.119	1.81	
Fetot(%)	Vis-PIR	1	0.411	0.078	0.000	1.30	0.314	0.079	0.004	1.22	
	Vis	4	0.444	0.076	-0.001	1.34	0.302	8.69	-8.28	0.31	
	PIR	6	0.432	0.077	-0.008	1.31	0.303	0.083	-0.009	1.15	
Arg(%)	Vis-PIR	11	0.424	2.8	-0.15	1.26	0.535	2.73	0.37	0.99	
	Vis	8	0.136	3.6	-0.098	0.98	0.576	2.25	0.93	1.20	
	PIR	9	0.179	3.59	-0.03	0.99	0.246	3.00	-0.06	0.90	
LF(%)	Vis-PIR	3	0.162	4.22	0.01	1.08	0.058	4.69	0.37	0.97	
	Vis	8	0.454	3.42	-0.13	1.34	0.071	5.03	0.69	0.90	
	PIR	7	0.291	4.11	-0.04	1.11	0.13	4.87	0.197	0.93	
LG(%)	Vis-PIR	15	0.721	6.174	-0.012	1.85	0.691	7.23	-0.778	1.66	
	Vis	4	0.476	8.25	0.02	1.39	0.616	7.47	-1.102	1.61	
	PIR	12	0.730	6.041	0.070	1.89	0.623	8.39	-1.34	1.43	
SF(%)	Vis-PIR	4	0.291	2.032	0.007	1.17	0.152	2.61	-0.36	1.05	
	Vis	4	0.082	2.31	0.010	1.03	0.118	2.57	0.148	1.07	
	PIR	6	0.191	2.25	0.030	1.06	0.076	26.26	-23.02	0.46	
SG(%)	Vis-PIR	3	0.528	8.08	0.05	1.47	0.574	8.39	0.145	1.54	
	Vis	4	0.496	8.36	0.08	1.42	0.516	8.94	0.65	1.45	
	PIR	3	0.590	7.53	0.07	1.57	0.596	8.15	0.214	1.59	

Viscarra et *al.* (2006b) considèrent que lorsque la valeur de ce ratio est <1.0 le modèle est de qualité médiocre. Il n'est pas recommandé de l'utiliser pour des prédictions ultérieures. Alors qu'une valeur de ce ratio comprise entre 1.0 et 1.4 indique que le modèle est à moitié recommandable. Il peut être utilisé pour différencier grossièrement entre les échantillons ayant des teneurs faibles et les échantillons ayant des teneurs élevées. De même Sayes et *al.* (2005) estiment qu'une valeur inférieure à 1.5 indique une inadéquation du modèle et par conséquent ce dernier est peu recommandé pour être utilisé dans des prédictions ultérieures. Selon Viscarra et *al.* (2006b) lorsque ce ratio est compris entre 1.4 et 1.8, le modèle peut être utilisé pour faire des estimations et établir des corrélations. Alors que des valeurs de ce ratio comprises entre 1.8 et 2.0 indiquent que le modèle est performant. Par ailleurs, Sinnaeve et *al.* (2001) estiment que lorsque ce ratio est supérieur à 2, le modèle offre une bonne qualité de prédiction.

Nous constatons que les performances des modèles de prédiction, varient d'une variable analytique à une autre. Il est vraisemblable que les modèles spectrométriques régionaux que nous avons calibré pour les variables : teneurs en azote total, pH, teneurs en fer total, teneurs en limons fins et teneurs en sables fins, tous spectres confondus, sont de qualité médiocres.

Ceci étant, pour la variable Azote total He et *al.* (2007) ont obtenu un coefficient de corrélation égale à 0.92 entre les teneurs mesurés au laboratoire et les teneurs prédites par spectrométrie dans le PIR. Toutefois les 125 échantillons qu'ils ont utilisé pour la calibration et la prédiction ont été prélevés sur une même parcelle à sol limoneux ayant une superficie de 160 m² ce qui laisse entrevoir la très forte homogénéité des échantillons.

Dalal et Henry (1986) ont conclu qu'il n'était pas possible de prédire correctement les teneurs en carbone organique et en azote total lorsque la gamme de couleur des sols considérés devient large. Signalons aussi que Morra et *al.* (1991) qui ont estimé en utilisant le PIR les teneurs en carbone total et en azote total dans les fractions limoneuses et argileuses du sol, affirment que les équations développées en considérant seulement la fraction limoneuse sont plus adéquates pour la détermination des concentrations d'azote total.

Ehsani et *al.* (1999) ont étudié l'effet singulier de l'azote minéral sur la réflectance dans le PIR. Pour cela ils ont enrichi deux parcelles la première à texture limoneuse et la seconde à texture argileuse avec dix doses croissantes d'azote minérale selon les bandes de labour puis ont collecté les échantillons de sol et calibré des modèles de prédiction, ont conclu que les modèles prédictifs de l'azote minérale sont robustes lorsque le jeu de calibration et le jeu de

prédiction sont collectés dans la même parcelle et qu'ils échouent complètement lorsque le jeu de calibration et le jeu de prédiction sont collectés sur deux parcelles différentes. Ces auteurs préconisent la calibration de modèle de type site spécifique.

Du moment que les ions nitrate sont polaires, ils présentent des vibrations dans le domaine du PIR (Vogt et Finlayson-Pitts, 1994) ce qui permettrait de prédire les teneurs en cet élément par spectrométrie mais malheureusement, beaucoup d'autres sels tel que les carbonates ont aussi des harmoniques d'absorption dans la même gamme spectrale. Ainsi les vibrations de la liaison de l'azote peuvent être occultées par l'absorption due aux vibrations des liaisons carbonées (Ehsani, 1999). Martin et *al.* (2002) reportaient que la détermination de l'azote total pourrait être indirecte, due à sa forte corrélation au carbone total, plutôt que directe due à l'absorption spécifique de l'azote. Alors que Brunet et *al.* (2007) affirment dans leurs résultats que la détermination de l'azote par spectrométrie peut être directe ou indirecte en fonction de la texture des échantillons de calibration. Dans notre cas l'échec, de prédiction de l'azote total, peut être expliqué par la forte hétérogénéité des échantillons.

Dans leur travail Islam et *al.* (2003) ont relativement bien prédit les valeurs du $\text{pH}_{1.5}(\text{H}_2\text{O})$ ($R^2 = 0.63$ dans le Vis et $R^2 = 0.70$ dans le Vis-PIR) et ce bien que leurs échantillons couvraient une gamme très large de pH [3.7-9.7]. Toutefois ces auteurs n'ont apporté aucune explication à ces bons résultats. De même Chang et *al.* (2001) ont obtenu des résultats de prédiction des titres de $\text{pH}_{1.5}(\text{H}_2\text{O})$ et de $\text{pH}_{1.2}(\text{CaCl}_2)$ qui étaient moyennes ($R^2 = 0.55$ et 0.56 pour la première et la deuxième méthode de mesure). Cependant là aussi aucune interprétation n'est donnée aux résultats qu'ils ont obtenus. Dans notre cas nous estimons que le pH en tant que mesure de l'activité des ions hydrogène dans la solution du sol n'influence pas directement le spectre du sol. Par ailleurs, la large gamme des pH des sols de la région qui couvre les classes des sols : acide, peu acide, neutre et basique explique notre échec dans la prédiction.

Seuls les formes des oxydes de fer sont spectralement actives à cause des transitions électroniques des cations de fer ferreux (Fe^{2+}) et ferrique (Fe^{3+}) d'un niveau d'énergie à un autre (Rezend et *al.*, 1980 ; Bédidi et *al.*, 1992, Bullard et *al.*, 2002 ; Ben-Dor, 2002) or dans notre cas nous avons mis en relation les spectres des sols avec les mesures du fer total, ce qui explique l'échec de prédiction.

Pour la variable carbone organique, au stade prédiction, le modèle calibré sur le PIR ($R^2 = 0.831$, REQM = 0.42 %, Biais = -0.03 % et REP = 2.42) est plus performant que le modèle calibré sur le combiné Vis-PIR ($R^2 = 0.806$, REQM = 0.45 %, Biais = -0.05 et REP =

2.27). Par ailleurs, le modèle calibré sur le Vis ($R^2 = 0.739$, REQM = 0.53 %, Biais = -0.02, REP = 1.93) est aussi qualifié de performant. Les bonnes prédictions obtenues pour le carbone organique s'expliquent par les vibrations des liaisons carbones dont les fondamentales se situent dans le MIR et dont les harmoniques de premier et de second ordre (plus faible absorption de lumière) se situent respectivement dans le PIR et le Vis (Brown et al., 2005 ; Viscarra, 2006a). Ceci explique plausiblement pourquoi le modèle du PIR est plus performant que les modèles calibrés respectivement sur le Vis-PIR et sur le Vis. Néanmoins des auteurs comme Krishnan et al. (1980) et Cozzolino et Moron (2006) ont trouvé que le Vis est plus opportun que le PIR pour la prédiction des teneurs en carbone organique.

Pour la variable CEC, au stade prédiction, les modèles calibrés sur les trois domaines spectraux sont qualifiés de performants puisqu'ils présentent des valeurs de REP qui sont de 1.76, 1.76 et 1.81 respectivement dans le combiné Vis-PIR, le Vis et le PIR. Comme pour le carbone, parmi les trois domaines d'investigation, le domaine du PIR semble être le plus prédictif. Les bonnes prédictions obtenues pour la CEC ont été escomptées vu la forte corrélation qui existait entre la CEC et la teneur en carbone organique. En effet, l'utilisation de la spectroscopie pour la prédiction des attributs inorganiques pourrait paraître inappropriée. Nonobstant Sorensen et Dalsgaard (2005) ont montré que des bonnes prédictions peuvent être obtenues pour les attributs inorganiques s'ils sont corrélés à la composante organique du sol.

Pour les variables limons grossiers et sables grossiers les modèles calibrés sur les trois domaines spectraux présentent au stade de prédiction des valeurs de REP comprises entre 1.8 et 1.4. En conséquence, ces modèles peuvent être utilisés pour faire des estimations des fractions de limons grossiers et de sables grossiers contenues dans le sol.

Pour la variable teneur en argiles, les modèles calibrés présentent au stade prédiction, des valeurs de REP de 0.99, 1.2 et 0.9 respectivement pour le combiné Vis-PIR, le Vis et le PIR. Ces modèles peuvent être utilisés pour discriminer grossièrement entre les échantillons de sols de la région, riches et pauvres en fraction argileuse. Lagacherie et al. (2008) ont mieux estimé les teneurs en argiles en pré-traitant les spectres par la technique du retrait du continuum (*continuum remove, en anglais*) et en appliquant une régression linéaire simple sur une seule longueur d'onde liée à l'absorption des minéraux argileux présents dans leur région d'étude à savoir : l'illite, la Kaolinite et la Montmorillonite. Dans leurs cas la relation est directe elle concerne l'influence des vibrations des liaisons des groupements hydroxyles de ces minéraux phyllosilicatés sur la réponse spectrale (Ben-Dor, 2002) et non l'influence de la fraction argileuse sur la réponse spectrale. De même Levin et al. (2007) ont calculé un indice

des particules fines, basé sur la sélection de 3 longueurs d'onde liées aux propriétés d'absorption des minéraux argileux et ont obtenu des bons résultats de prédiction. Toutefois, leurs échantillons de calibration sont issus d'un seul type de sol : sol sableux dunaire. Viscarra et Mc Bratney (1998) qui ont bien prédit les teneurs en argile du sol se sont basés lors de la calibration de leur modèle sur un jeu d'échantillons artificiels qu'ils ont constitué en mélangeant à différentes proportions du sable inerte avec le matériau d'un sol argileux. Ces derniers ont toutefois bien précisé que leurs échantillons ne présentaient pas une large différence dans leurs matrices minérales. Dès lors, une fois rattaché à son contexte de calibration le champ d'application de leur modèle dans la prédiction des teneurs en argile est très restreint. Nous sommes persuadé que nos résultats concernant la prédiction des teneurs en argile qui sont à moitié encourageant peuvent être améliorés. Pour cela nous devons commencer par caractériser les minéraux argileux qui existent à l'échelle de la région, pour mettre, par la suite, en relation la réflectance en ces longueurs d'onde avec les teneurs en fraction argileuse du sol à l'aide de régression simple voir multiple.

Dans cette étude, les bonnes prédictions obtenues dans le cas du carbone organique peuvent être attribuées à l'influence directe des liaisons carbonées sur les spectres des sols. Par contre, l'une des explications qui peut être donnée à la bonne prédiction de la CEC serait sa forte corrélation avec le carbone ($r = 0.94$). En effet, d'après Clark et *al.* (1987), la spectrométrie permettrait de déterminer la concentration en certains cations lorsqu'ils sont associés à des molécules organiques ou à des molécules inorganiques hydratées.

Nous avons constaté aussi que seuls les fractions granulométriques, qui étaient corrélés à la variable carbone à savoir le limon grossier et le sable grossier ont été bien prédites. Par ailleurs, ces deux fractions granulométriques sont fortement corrélées entre elles ($r = 0.86$). Ces deux fractions seraient indirectement liées à la réponse spectrale des sols de la région par le biais de la matière organique qui serait concentrée dans ces deux fractions. Ben-Dor et Banin (1995) soutiennent l'hypothèse que la spectrométrie peut être utilisée pour prédire des attributs du sol qui n'influencent pas directement le comportement spectrale et ce *via* leurs inter-corrélation avec des propriétés du sol qui sont spectralement actives.

Pour le reste des fractions granulométriques notamment les limons fins et les sables fins, la qualité de prédiction était mauvaise. Girard et Girard (1999) estiment que dans le domaine des longueurs d'onde du Vis, il ne peut pas y avoir d'interaction des fractions granulométriques du sol avec le rayonnement, les tailles des argiles étant trop grandes mais celle-ci interviennent dans le PIR. Les limons ou les sables ont des tailles trop grandes pour intervenir sur sa réflectance dans le Vis ou le PIR. Girard et Girard (1999) présumant que si

certain auteurs mentionnaient des variations de réflectance selon les teneurs plus ou moins grandes en argiles ou sables, elles seraient alors dépendantes de la teneur en eau de ces éléments. Par ailleurs, Bowers et Hanks (1965) ont démontré que les particules du sol n'influencent la réponse spectrale du sol que lorsqu'ils présentent une taille inférieure à 400 μm . De même Al-Abbas et *al.* (1972) qui ont mis en relation les teneurs en matière organique et en argiles de 197 échantillons de sol, prélevés à la surface d'une parcelle de 25 ha, avec leurs luminances multi-spectrales affirment ne peuvent pas être certains si la relation entre la réflectance relative et la teneur en argiles qu'ils ont pu établir est directe ou qu'il s'agit d'une relation indirecte qui résulte dans leur cas de la corrélation entre la teneur en argiles et la teneur en matière organique.

4. Conclusion

La spectrométrie est désormais communément reconnu comme un outil de quantification rapide, non destructif et simultané des propriétés spectralement actives et des propriétés qui leurs sont corrélées. Les courbes de réflectances spectrales des sols sont croissantes et monotones dans le visible alors qu'elles sont ascendantes dans certaines fenêtres du PIR et descendantes sur certaines autres. A ce titre le domaine du PIR est plus indiqué pour la prédiction des propriétés du sol nous l'avons au moins vérifié pour le carbone organique et pour la CEC. La spectrométrie peut être utilisée pour des analyses sur des grandes séries d'échantillons prélevés sur des sites différents, sous réserve que les échantillons de calibration et de prédiction soient similaires. L'utilisation de la procédure d'ACP dans la sélection des échantillons est très recommandée à cet effet.

Vu les enjeux agronomique et environnemental que représente le carbone organique dans le sol et son influence directe sur les propriétés physique, chimique et biologique du sol, nous constatons que les données sur les stocks de carbone dans le sol et sur leurs évolutions sont de plus en plus demandés par non seulement les pédologues mais aussi par des scientifiques d'autres disciplines. Les bonnes prédictions que nous avons obtenues pour cette variable et pour les variables qui lui sont corrélés à l'échelle de notre région constituent un vrai challenge qui ouvre la voie devant le développement de modèles spectrométriques beaucoup plus globaux voir universels.

Les mesures analytiques de laboratoires demeurent des mesures de références. Toutefois, la vraie appréciation de la qualité de prédiction obtenue par spectrométrie, doit être laissée au jugement du praticien et de l'application qu'il va en faire avec les valeurs prédites qu'il aura à sa disposition.

Par ailleurs, nous estimons que la spectrométrie peut être très utile dans des applications comme :

- l'agriculture de précision pour l'optimisation de l'épandage spatiale des fertilisants,
- la surveillance de la qualité des sols dans le temps, une chronoséquence de mesures spectrométriques permettrait d'apprécier la tendance évolutive de la qualité des sols sans avoir recours à des analyses de laboratoire sauf en certaines périodes pour avoir des mesures de référence ;
- la cartographie à différentes échelles de la variabilité des propriétés du sol par intégration des outils de spectrométries et des outils de géostatistiques.

Chapitre IV. Qualité d'estimation des modèles spectrométriques « site spécifique » et « régional ». Exemple de la teneur en carbone total et de la conductivité électrique des sols, en milieu aride.

H. Aïchi^{1,2*}, Y. Fouad², M. Sanaa¹, Z. Lili Chabaane¹, H. Nicolas², C. Walter².

¹*Institut National Agronomique de Tunis, Avenue Charles Nicole Cité Mahrajène 1002 Tunis, Tunisie.*

²*Agrocampus Ouest - UMR 1069 SAS, 65 rue de St Briec, CS 84215, 35042 Rennes, France.*

Abstract

Visible-Near infrared spectroscopy in association with chemometric approaches is nowadays recognized as a promising tool for predicting various soil properties. We coupled spectroscopy to PLS algorithm to predict two interesting properties in arid zone : total carbon (C_{tot}) and electrical conductivity (CE). In three distinct sites, covered by three QuickBird satellite images (Chamsa : 2800 ha, Oudia : 2800 ha and Oung : 2548 ha), belonging to a given arid region (Djerid : 5600 km², south western Tunisia), we collected respectively 30, 29 and 30 samples according to a parsimonious sampling strategy, integrating radiometric image data. We compared those two soil properties the prediction performance of two calibration procedures : “ site specific ” *versus* “ regional ”. For C_{tot} , specific site models were effective for the first two sites and showed poor performance for the third. Prediction performances of these models were respectively (R^2 (coefficient of determination) = 0.85, Bias (mean error) = -0.01%, RMSE (Root Mean Squared Error) = 0.14% and RPD (Ratio of Prediction Deviation) = 2.53 ; R^2 = 0.83, Bias = -0.01%, RMSE = 0.64% and RPD = 1.64 ; R^2 = 0.55, Bias = 0.04%, RMSE = 0.43% and RPD = 0.76). For EC, site specific models showed poor performance for the first two sites and appeared effective for the third. Prediction performances of these models were respectively (R^2 = 0.57, Bias = 0.12 dS/m, RMSE = 0.54 dS/m and RPD = 0.32 ; R^2 = 0.51, Bias = 0.37 dS/m, RMSE = 2.5 dS/m and RPD = 1.42 ; R^2 = 0.76, Bias = 9.77 dS/m, RMSE = 12.44 dS/m and RPD = 2.01). For both variables, prediction performances declined noticeably when a site specific model was used to predict samples collected on the two other sites. This fact testifies a limited geographical robustness of these specific site models. The regional model of C_{tot} was acceptable (R^2 = 0.67, Bias = 0.18 dS/m, RMSE = 0.93 dS/m and RPD = 1.72). Its quality can be improved by recalibrating the model on a larger sample set collected from other sites. The regional model of the EC showed medium performance (R^2 = 0.53, Bias = 4.59 dS/m, RMSE = 10.75 dS/m and RPD = 1.44).

Key words : spectroscopy, chemometry, arid zone, site specific model, regional model, total carbon, electrical conductivity, QuickBird.

Résumé

La spectrométrie visible, proche-infrarouge en association avec les approches chimiométriques sont aujourd'hui reconnues comme prometteurs pour estimer diverses propriétés physico-chimiques du sol. Nous avons, couplé la spectrométrie à la régression en moindres carrés partiels (RMCP) pour l'estimation de 2 propriétés d'intérêt en milieu aride : le carbone total (C_{tot}) et la conductivité électrique (CE). Dans trois sites distincts, couverts par trois images satellites QuickBird (Chamsa : 2800 ha, Oudia : 2800 ha et Oung : 2548 ha), d'une même région aride (Djérid : 5600 km², Sud-Ouest tunisien), nous avons respectivement collecté 30, 29 et 30 échantillons selon une stratégie d'échantillonnage parcimonieuse, intégrant les données radiométriques des images. Nous avons étudié pour ces 2 propriétés, l'effet du mode de calibration « site spécifique » *versus* « régional » sur les performances des modèles d'estimation. Pour le C_{tot} , le modèle « site spécifique » est performant pour les deux premiers sites et peu satisfaisant pour le troisième. Les performances d'estimation de ces modèles sont respectivement (R^2 (coefficient de détermination) = 0.85, Biais (erreur moyenne) = -0.01 %, REQM (Racine de l'erreur quadratique moyenne) = 0.14% et REP (Ratio des Ecart de Prédiction) = 2.53 ; R^2 = 0.83, Biais = -0.01%, REQM = 0.64 % et REP = 1.64 ; R^2 = 0.55, Biais = 0.04 %, REQM = 0.43 % et REP = 0.76). Pour la $CE_{1:5}$, le modèle « site spécifique » est peu satisfaisant pour les deux premiers sites et performant pour le troisième. Les performances d'estimation de ces modèles sont respectivement (R^2 = 0.57, Biais = 0.12 dS/m, REQM = 0.54 dS/m et REP = 0.32 ; R^2 = 0.50, Biais = 0.37 dS/m, REQM = 2.5 dS/m et REP = 1.42 ; R^2 = 0.76, Biais = 9.77 dS/m, REQM = 12.44 dS/m et REP = 2.01). Dans le cas des deux propriétés, nous avons constaté une nette détérioration des performances d'estimation lorsque le modèle spécifique à un site est utilisé pour estimer les teneurs des échantillons prélevés sur les deux autres sites. Ce fait témoigne d'une robustesse géographique limitée des modèles « site spécifique ». Le modèle régional du C_{tot} est assez satisfaisant (R^2 = 0.67, Biais = 0.18 %, REQM = 0.93 % et REP = 1.72). Sa qualité peut être améliorée en recalibrant le modèle sur un nombre plus important d'échantillons prélevés, sur davantage de sites. Le modèle régional pour la $CE_{1:5}$ est de performance moyenne (R^2 = 0.53, Biais = 4.59 dS/m, REQM = 10.75 dS/m et REP = 1.44).

Mots clés : spectroscopie, chimiométrie, zone aride, modèle site specific, modèle régional, carbone total, conductivité électrique, QuickBird.

1. Introduction

Il y a un intérêt manifeste et croissant pour l'utilisation de la Spectroscopie par Réflexion Diffuse dans le Visible-Proche Infrarouge (SRD Vis-PIR) en tant qu'outil rapide, non polluant (car non consommateur de réactifs) et peu onéreux pour la caractérisation des propriétés du sol (Ben Dor et Banin, 1995 ; Stenberg et *al.*, 1995 ; Janick et *al.*, 1998 ; Reeves et *al.*, 1999, 2000, 2001, 2002 ; Confalonieri et *al.*, 2001 ; Reeves et Mc Carty 2001 ; Dunn et *al.*, 2002 ; Mc Carty et *al.*, 2002 ; Shepherd et Walsh, 2002 ; Stenberg et *al.*, 2002 ; Islam et *al.*, 2003, 2004 ; Brown et *al.*, 2006 ; Genot et *al.*, 2007). Pour certaines propriétés du sol, la fiabilité d'estimation reportée dans de nombreuses études, suggère que cette technique pourrait remplacer dans les prochaines années, pour des applications particulières, les méthodes analytiques standards de laboratoire (Mc Carty et *al.*, 2002 ; Reeves et *al.*, 2002).

L'utilisation de la SRD Vis-PIR pour caractériser les sols (Stoner et Baumgardner, 1981 ; Baumgardner et *al.*, 1985), les minéraux du sol (Hunt, 1977 ; Clark et Roush, 1984 ; Clark, 1999) ainsi que la matière organique du sol (Baumgardner et *al.*, 1985 ; Henderson et *al.*, 1992) remonte aux travaux pionniers de Obukhov et Orlov (1964) et ceux de Bowers et Hanks (1965). Toutefois, la première génération de ces travaux avait pour objectif d'étudier les propriétés spectrales des sols pour interpréter l'information des images satellites (Baumgardner et *al.*, 1985 ; Escadafal, 1989 ; Ben-Dor, 2002). En effet, jusqu'à peu, l'analyse spectrale des sols était seulement mise à profit dans la description qualitative et la classification des sols, par le biais de la couleur (Bauer et *al.*, 1980 ; Stoner et Baumgardner, 1981 ; Blavet et *al.*, 2003).

Les progrès technologiques dans la fabrication des spectroradiomètres portables ont rendu cet outil financièrement plus accessible. Par la même, le développement de logiciels commerciaux de chimiométrie a offert de nouveaux outils statistiques pour la quantification des propriétés du sol à partir des mesures spectrales (Friedman, 2001 ; Wold et *al.*, 2001 ; Brown et *al.*, 2006).

L'estimation fiable des teneurs en carbone total (C_{tot}) et de la conductivité électrique ($CE_{1:5}$) du sol par la SRD Vis-PIR est d'un intérêt majeur, particulièrement en milieu aride. En effet, plusieurs études ont souligné que la spectrométrie est un bon outil d'estimation de ces deux propriétés indicatrices de la qualité du sol (Islam et *al.*, 2004 ; Farifteh et *al.*, 2008), particulièrement en milieu aride.

Cependant les aspects méthodologiques relatifs, entre autre à l'origine et à l'homogénéité du jeu des échantillons n'ont pas été exhaustivement étudiés (Brunet et *al.*, 2007). En effet, dans la plupart des études jusque là entreprises, dans des contextes pédologiques différents, les échantillons de calibration et de validation ont été le plus souvent sélectionnés sans considération particulière quant à l'échelle géographique de leur collecte, qu'elle soit en mode « site spécifique » ou « régional » (Brown et *al.*, 2006).

Néanmoins, Sudduth et Hummel (1993) ont remarqué dans leurs investigations que la qualité d'estimation se détériore lorsque l'origine des échantillons s'étend géographiquement, d'où l'existence d'un effet de la parenté des échantillons et de l'étendue géographique sur la qualité du modèle d'estimation. Nous pouvons constater que beaucoup d'études ont été réalisées sur des jeux de données hétérogènes, qui incluent une large gamme de types de sols (Ben Dor et Banin, 1995 ; Chang et *al.*, 2001 ; Shepherd et Walsh, 2002). Plusieurs auteurs considèrent que la crédibilité de l'estimation par la SRD Vis-PIR est contestable, lorsque nous étudions un jeu d'échantillons hétérogène (Sudduth et Hummel, 1996 ; Brown et *al.*, 2005). Des auteurs comme Genot et *al.* (2007) estiment que la calibration doit être conduite sur un jeu d'échantillons représentatif de la population concernée par les futures estimations. Ceci implique préalablement une évaluation de la diversité de la population pour les propriétés objets des estimations.

Dès lors, davantage de travaux concernant l'effet de l'hétérogénéité des jeux des échantillons sur la fiabilité de l'estimation sont nécessaires pour appliquer plus « globalement » la technique de la SRD Vis-PIR à la caractérisation des sols. En effet, dans la pratique, les modèles calibrés ont une utilité limitée à moins qu'ils puissent être appliqués à de nouveaux échantillons, en particulier à des échantillons relevant d'une aire géographique restreinte et qui présentent des similarités avec les échantillons du jeu de calibration (Brown et *al.*, 2006).

En spectrométrie, les recherches actuelles s'intéressent au développement de modèles d'estimation des propriétés du sol qui soient applicables à l'échelle d'un bassin versant (Berzaghi et *al.*, 2000), voire d'une région entière (Brown et *al.*, 2006). L'objectif ultime d'aboutir à des modèles d'estimation des propriétés du sol, qui seraient globaux et par conséquent d'une utilité universelle, a également été suggérée par de nombreux chercheurs (Barthes et *al.*, 2006 ; Viscarra, 2008).

La calibration est une procédure statistique qui relie l'information spectrale aux propriétés des échantillons mesurées par les méthodes analytiques standards. Ces équations de calibration sont ensuite utilisées pour estimer ces propriétés pour d'autres échantillons.

L'estimation des propriétés du sol par la SRD Vis-PIR requiert le développement d'un modèle de calibration (Genot et *al.*, 2007). Celui-ci peut être élaboré en mode site spécifique ou régional.

Dans le premier cas (mode site spécifique), une équation d'estimation est établie et utilisée pour chaque site de collecte des échantillons, alors que dans le second cas (mode régional) une seule et unique équation d'estimation est construite sur un jeu d'échantillon collecté dans différents sites appartenant à une même région.

Lors du développement d'un modèle d'estimation, il est nécessaire de calibrer le modèle en utilisant un jeu d'échantillon dit jeu de calibration et de le valider sur un deuxième jeu, dit de validation, composé d'échantillons qui n'ont pas servi à sa calibration. C'est alors qu'il est possible de déterminer comment le modèle serait applicable en dehors du jeu de données de calibration, ce qui est l'objectif ultime de l'exploration de modèles spectraux d'estimation.

L'objectif de ce travail est d'étudier l'effet du mode de calibration (site spécifique vs régional) sur la qualité d'estimation de deux propriétés du sol. L'étude a été réalisée en milieu aride à l'aide de trois jeux d'échantillons collectés dans trois sites à sols nus, ayant des faciès pédo-paysagiques distincts, mais appartenant à une même région naturelle à savoir la région aride du Djérid (Sud-Ouest tunisien). Les propriétés pédologiques étudiées ont été le carbone total (C_{tot}) et la conductivité électrique ($CE_{1:5}$). Nous avons, successivement produit pour ces deux propriétés trois modèles « site spécifique ». Leurs performances ont été évaluées *intrasite* (sur des échantillons prélevés sur le même site) et *intersite* (sur les échantillons prélevés sur chacun des deux sites restants) et comparées à la qualité d'estimation d'un modèle régional (calibré sur les échantillons prélevés simultanément sur les trois sites).

2. Matériels et méthodes

2.1. Milieu d'étude

La figure 1 donne un aperçu sur la région (image ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer)) et sur la répartition spatiale des trois sites locaux prospectés (images QuickBird).

La région du Djérid, d'une superficie d'environ 5 600 km², se situe entre 7,5° et 8,6° de longitude Est et 33,7° et 34,5° de latitude Nord (Figure 1). Elle occupe un milieu présaharien,

de transition entre deux domaines sensiblement différents, à savoir le domaine atlasique au nord et le domaine de la plateforme saharienne au sud (Coque, 1962 ; Rabiaa, 1984). Ces deux domaines présentent de forts contrastes d'ordres structural, topographique, climatique, hydrique et végétal (Mamou, 1976 ; Bryant, 1999).

La région fait partie de l'étage bioclimatique aride inférieur (Kamel, 2007). Cette aridité est confirmée par une évapotranspiration supérieure à 1500 mm/an (Mamou, 1989) avec une différence de température de 20°C entre le mois le plus chaud et le mois le plus froid. Les pluies sont rares mais réparties entre les saisons d'automne, d'hiver et du printemps. La pluviométrie ne dépasse guère 90 mm/an. Il s'agit d'un climat continental, mettant en évidence l'influence du Sahara et du Grand Erg oriental (Rabiaa, 1984). Par ailleurs, les vents sont très fréquents dans la région, avec en moyenne 120 j/an de vents forts. Le vent dominant est de direction Nord-Est. Il souffle surtout en automne et au printemps (Kamel, 2007). Les types de sols les plus fréquents au niveau de la région sont les Gypsosols, Lithosols, Regosols, Salisols, Fluvisols, et les sols hydromorphes.

2.2. Stratégie d'échantillonnage

Face à l'étendue de la zone étudiée d'une part, et à la modestie des moyens de projection du terrain d'autre part, nous nous sommes servis d'une image satellite de type ASTER (Figure 1) (60x60 km²) qui couvre, la quasi totalité de la région et de trois images de type QuickBird qui couvrent trois secteurs locaux distincts et distants, situés au sein de cette même région. Ces images ont été mises à profit pour visualiser et explorer la variabilité du terrain et pour élaborer, en conséquence, un schéma d'échantillonnage parcimonieux. En effet, nous nous sommes basés sur les résultats d'une classification non supervisée des 3 images, pour établir un schéma d'échantillonnage sur le terrain qui soit fiable, moins intensif et plus facile à mettre en œuvre qu'un schéma d'échantillonnage systématique régulier conventionnel.

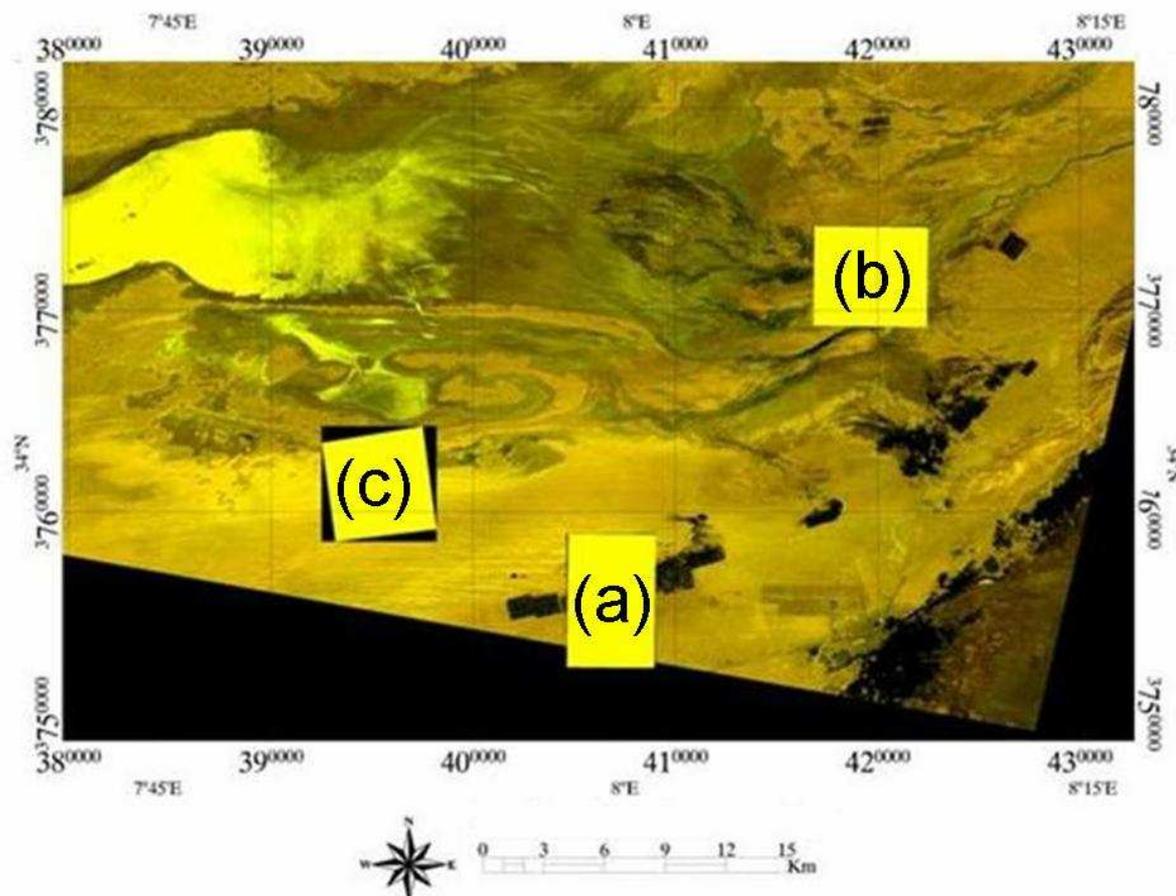


Figure 1. Aperçu sur la région d'étude (couverte par une image ASTER) et sur l'emplacement des trois sites prospectés (couverts par trois images QuickBird) : (a) site de Chamsa, (b) site d'Oudia et (c) site d'Oung.

Le tableau 1 donne les caractéristiques de ces images qui sont déjà corrigées radiométriquement et calées géométriquement par le fournisseur des images.

Tableau 1. Caractéristiques des images satellites.

Scène couverte	Capteur	Date d'acquisition	Coin haut gauche	Coin bas droit	Superficie (ha)
Région du Djérid	ASTER	27/10/2006	(7°42'E ; 34°31'N)	(8°13'E ; 33°52'N)	360.000
Site de Chamsa	Quick Bird	29/05/2005	(8°09'E ; 34°10' N)	(8°15'E ; 34°05'N)	2800
Site d'Oudia	Quick Bird	18/02/2004	(7°96'E ; 33°96'N)	(8°01'E ; 33°90'N)	2880
Site d'Oung	Quick Bird	18/02/2004	(7°83'E ; 34°00'N)	(7°89'E ; 33°96'N)	2548

Les coordonnées planimétriques de ces images ont été converties en référence à la projection Transverse Universelle de Mercator (*Universal transverse Mercator (UTM), en anglais*). X (Abscisse, direction Est), Y (ordonnée, direction Nord). Les paramètres de projection sont les suivants : (Projection : UTM, Fuseau : 32 N, Ellipsoïde : Clarke 1880).

Pour tester la fiabilité du géoréférencement des images, nous avons relevé sur chacune d'elles les coordonnées de quelques points de contrôle sur le terrain (croisements de routes et de pistes). Nous nous sommes, assurés que les coordonnées relevées sur le terrain par un récepteur de Géolocalisation Par Sattelite (GPS) Trimble Geo XM[®] et les coordonnées des images étaient identiques, à la précision du positionnement GPS près (5 mètres).

Le schéma d'échantillonnage a été établi pour chacun des trois sites couverts par les trois images Quickbird sous le logiciel de traitement d'image satellitale ENVI 4.3[®] ENvironment for Visualizing Image (ITT, 2007) qui est dédié à la visualisation et aux traitements des images satellites. Ce schéma se décline comme suit : i) classification non supervisée de l'image produisant des classes radiométriques distinctes, ii) habillage de l'image classée avec une grille géographique UTM produisant une spatio-carte au 1 : 20 000, iii) à la lumière des classes radiométriques produites, supposées correspondre à des états de surfaces différents, nous avons tracé des transects passant par les différentes classes radiométriques et qui couvrent uniformément la totalité de la surface de chaque site et iv) sur ces transects, nous avons fixé des points d'échantillonnages, une trentaine pour chaque site, qui sont régulièrement espacés. Pour se repérer sur le terrain, nous nous sommes déplacés à l'aide des spatio-cartes. Les points préfixés d'échantillonnage ont été repérés à l'aide d'un récepteur GPS Trimble Geo XM[®] d'une précision planimétrique de 5 m. Au niveau de chaque point préfixé, nous avons collecté un échantillon composite formé de huit sous-échantillons prélevés à la surface du sol (10 premiers centimètres du sol) dans huit quartiers différents dans un rayon de 8 m.

Les figures 2a, 2b et 2c montrent respectivement les plans d'échantillonnage des sites de Chamsa, Oudia et Oung. Les punaises montrent les points de collecte des échantillons.

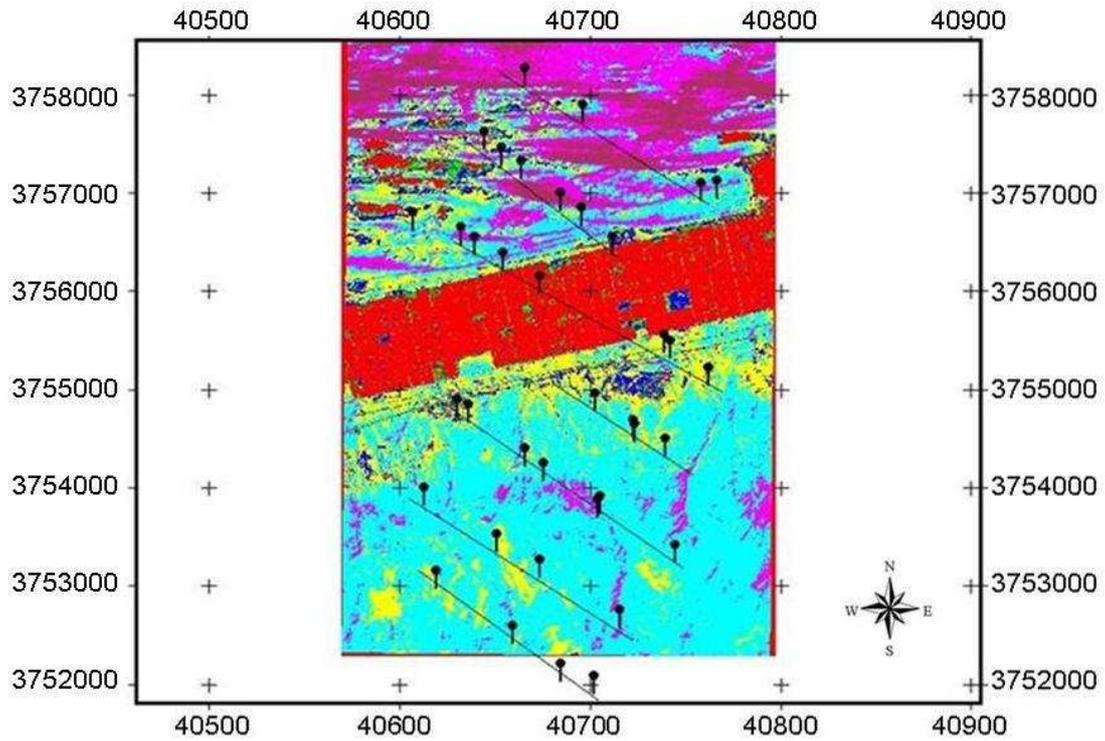


Figure 2a. Schéma d'échantillonnage du site de Chamsa (Les punaises représentent les points de collectes des échantillons. Les coordonnées sont en mètre selon la projection UTM).

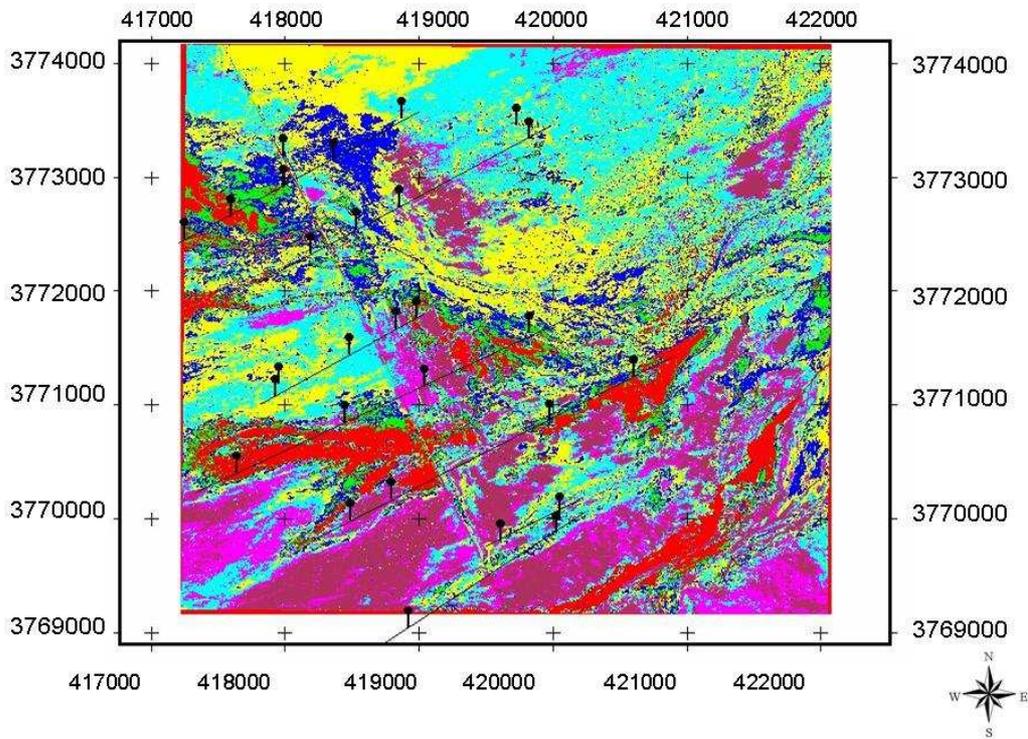


Figure 2b. Schéma d'échantillonnage du site d'Oudia (Les punaises représentent les points de collectes des échantillons. Les coordonnées sont en mètre selon la projection UTM).

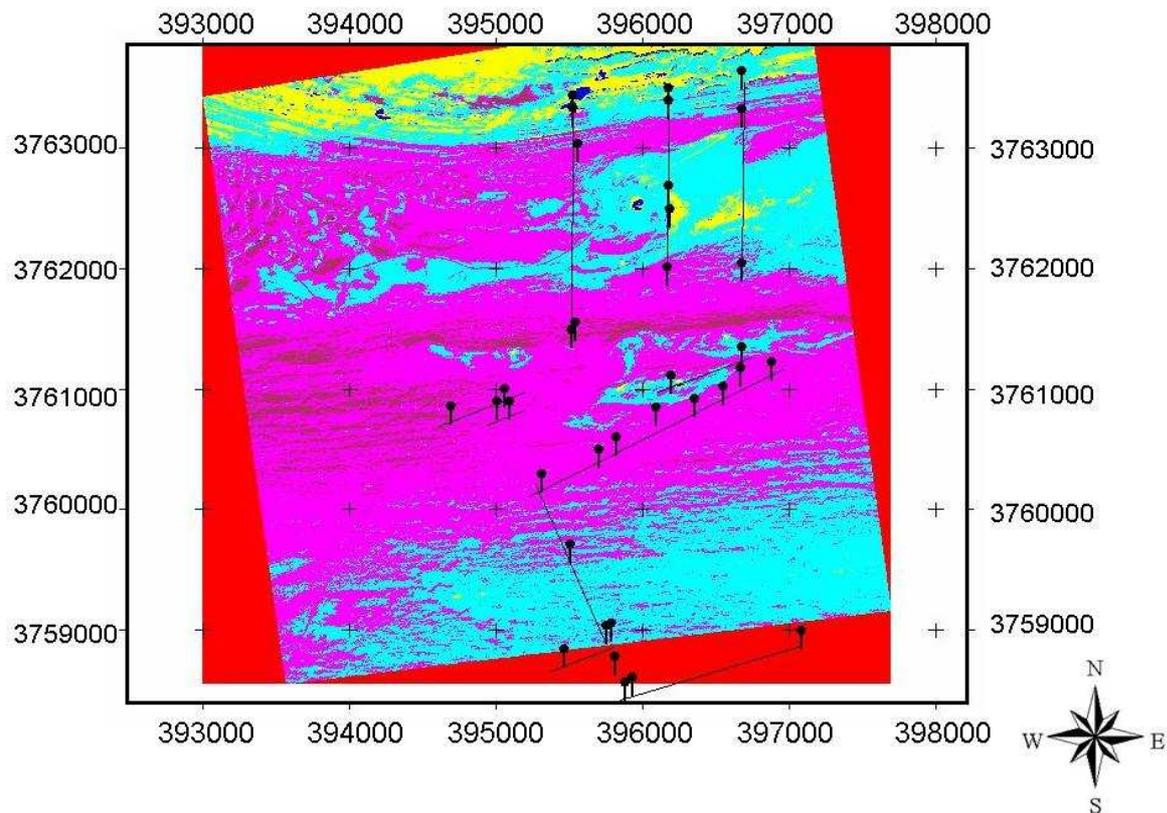


Figure 2c. Schéma d'échantillonnage du site d'Oudia (Les punaises représentent les points de collectes des échantillons. Les coordonnées sont en mètre selon la projection UTM).

La campagne de collecte des échantillons sur le terrain s'est déroulée au cours du mois de novembre 2007 en période sèche. Au terme de cette campagne, nous avons pu constituer une base pédo-spectrale régionale composée de 89 échantillons collectés à partir des 3 sites locaux, à raison de 30 échantillons pour le premier site, dit site de Chamsa, 29 pour le second, dit site d'oudia et 30 pour le troisième, dit site d'Oung.

2.3. Caractéristiques des trois sites

2.3.1. Site de Chamsa

Le site prospecté de Chamsa est couvert par une image Quick Bird d'une superficie de 2800 ha. Il est situé au sud du chott El Gharsa (Figure 1, rectangle (a)). Ce site se trouve dans la partie amont du bassin versant du chott, sur un versant de direction Sud/Nord. Les sols de ce secteur sont des sols minéraux bruts d'apport éolien évoluant sur croûte gypseuse (Melis, 2003). Par endroit, les matériaux sableux sont déblayés et la croûte gypso-calcaire affleure en

surface en Reg (état de surface pierreux, suite à l'arrachage des particules fines, sous l'effet du vent et des événements pluvieux torrentiels).

2.3.2. Site d'Oudia

Le site d'Oudia est également couvert par une image Quick Bird d'une superficie de 2800 ha. Il est situé au nord-est (Figure 1, rectangle (b)) du chott El Gharsa. Ce site est connu comme un couloir de vent de direction Est-Ouest (Melis, 2003). D'un point de vue pédopaysagique, le secteur d'Oudia peut être assimilé à un bas de versant en rupture de pente limité par les berges d'un oued : le niveau le plus bas est occupé par le lit majeur de cet oued saisonnier (Oued « El Melah » qui veut dire salée en arabe), qui descend en pente douce du Djebel Metlaoui (situé beaucoup plus au nord). Il s'agit aussi d'une zone de déflation, où les pertes en matières solides sont importantes par arrachage éolien, mais où les apports d'alluvions sont aussi importants (Melis, 2003). Les matériaux originels de ce site sont des produits allochtones, issus d'un déplacement et d'une répartition en nappe d'épandage. Il est aussi opportun de signaler que ce secteur constitue une zone d'épandage non contrôlé des eaux de rejet de phosphate de l'unité de raffinement de Metlaoui, riches en particules fines de la fraction granulométrique des argiles. Au niveau de ce secteur, les sols sont des sols minéraux bruts d'apport éolien, appelé localement Erg, et des sols minéraux bruts d'apports alluvial et fluvial.

2.3.3. Site d'Oung

Ce dernier est couvert par une image Quick Bird d'une superficie de 2548 ha. Il se situe à l'ouest de la région, sur la rive sud du Chott El Gharsa (Figure 1, rectangle (c)). Les rives du Chott sont délimitées par des croûtes gypseuses fossilisées sous forme de terrasses couvertes par des nebkas plus ou moins éparses. Ces mêmes croûtes sont par endroit couvertes par des dunes de sables de plusieurs mètres de hauteur, dénotant de la prépondérance de l'érosion éolienne sur l'érosion hydrique (Melis, 2003). Au niveau de ce site, nous distinguons essentiellement des sols minéraux bruts d'apport éolien, des sols calcimagnésiques à croûte gypseuse affleurante par endroit, ainsi que des sols peu évolués pré-chotteux et chotteux, de texture fine et à hydro-halomorphie prononcée.

2.4. Analyses des échantillons

Les échantillons collectés ont été séchés à l'air libre et tamisés à 2 mm avant d'être analysés au laboratoire. Le tableau 2 consigne les méthodes d'analyses qui ont été utilisées.

Tableau 2. Méthodes d'analyse des échantillons, pour les différentes propriétés considérées.

propriétés	Méthodes	Norme AFNOR
C_{tot}	Par combustion	ISO 10694
N_{tot}	Par combustion	ISO 13878
$pH_{eau\ 1:5}$	pH-mètre de laboratoire	ISO 10390
$CE_{1:5}$	Conductivimètre de laboratoire	
RS	Pesage du résidu sec	
Sables Grossiers ($2000 < \varnothing < 1000\ \mu m$)	Tamissage	
Sables Moyens ($1000 < \varnothing < 200\ \mu m$)		
Sables fins ($200 < \varnothing < 0.063\ \mu m$)		
Éléments Fins ($\varnothing < 0.063\ \mu m$)		

Ces analyses ont porté sur : le carbone total (C_{tot}) et l'azote total (N_{tot}) mesurés par combustion à l'aide d'un analyseur élémentaire, respectivement selon les normes ISO 10694 et ISO 13878 ; le $pH_{eau\ 1:5}$ mesuré selon la norme ISO 10390 ; la Conductivité Electrique ($CE_{1:5}$) mesurée à l'aide d'un conductivimètre de laboratoire et le Résidu Sec (RS). Nous avons aussi calculé les fractions granulométriques suivantes : Sables Grossiers (SG) : $2000 < \varnothing < 1000\ \mu m$, Sables Moyens (SM) $1000 < \varnothing < 200\ \mu m$, Sables Fins (SF) $200 < \varnothing < 63\ \mu m$ et Éléments Fins (EF) $\varnothing < 63\ \mu m$.

2.5. Acquisition et traitement des spectres

Les spectres ont été acquis à l'aide d'un spectroradiomètre de la société ASD (Analytical Spectral Devices, Boulder Co), modèle « Field spec pro FR ». L'instrument a la capacité de détecter la lumière dans la gamme spectrale 350-2500 nm avec une résolution constante de 1 nm sur tout le domaine spectral. Comme le bruit dans le domaine 350-400 nm est important (Wetterlind et *al.*, 2008), la fenêtre de l'UV a été tronquée du spectre, d'autant plus que nous ne nous intéressons pas à ce domaine. Par ailleurs, pour des raisons de parcimonie des données et pour homogénéiser notre base pédo-spectrale globale, les spectres ont été ré-échantillonnés au même pas d'échantillonnage spectral du spectroradiomètre GER 3700. En effet, les échantillons de la région de Bretagne (cf, chapitre 3) ont été acquis avec un spectroradiomètre de type GER 3700.

Pour chaque échantillon, les acquisitions des spectres ont été réalisées dans 4 positions à raison de 10 acquisitions par position et en faisant à chaque fois tourner l'échantillon d'un angle de 90° , soit un total de 40 acquisitions par échantillon. Par la suite, ces 40 spectres ont été moyennés pour obtenir un spectre moyen (Seiler et *al.*, 2007).

2.6. Calibration des modèles

Dans une première phase, nous avons calibré des modèles en mode site spécifique. Pour cela, nous avons considéré chaque jeu d'échantillons issus d'un même site à part. Dans un second temps, nous avons calibré des modèles en mode régional. Pour cela, nous avons considéré l'ensemble des trois jeux d'échantillons originaires des trois sites locaux, qui forment ensemble la base de données régionale. Une procédure d'Analyse en Composantes Principales (ACP), appliquée sur toutes les variables analytiques disponibles caractérisant les échantillons de sols, a été à chaque fois utilisée pour :

- i) explorer la variabilité statistique des échantillons de chaque site pris séparément ainsi que pour le jeu de données résultant du regroupement de l'ensemble des échantillons des 3 sites. Ce dernier jeu d'échantillons est dit régional,
- ii) détecter les échantillons atypiques (Shepherd et Walsh, 2002 ; Islam et *al.*, 2003 ; Cozzolino et Moron, 2006),
- iii) sélectionner et affecter les échantillons dans le jeu de calibration et le jeu de validation,

En effet, les échantillons appartenant à chaque population ont été projetés dans le premier plan factoriel de l'ACP. De la sorte, la similarité qui existe entre les échantillons a été explorée. Nous avons pu d'une part identifier les échantillons atypiques qui ont été écartés, de prime à bord, du jeu de données et d'autre part regrouper le reste des échantillons en couples de points en fonction de leur proximité dans le premier plan factoriel, par simple examen visuel des distances euclidiennes qui les séparent, puis affecter aléatoirement l'un dans le jeu de calibration et l'autre dans le jeu de validation.

Une Régression en Moindres Carrés Partiels (RMCP) avec validation croisée a été appliquée à l'aide du logiciel Splus 6.1[®], pour relier les données analytiques de référence et les données spectrométriques acquises dans le spectre du Vis-PIR.

2.7. Choix du nombre optimum de composantes latentes des modèles calibrés

Le choix du nombre optimum de composantes latentes des modèles (NCLM) RMCP a été effectué en appliquant une validation croisée (Tenenhaus, 1998 ; Vancolen, 2004 ; Cozzolino et Moron, 2006). Le modèle, qui présentait la racine de l'erreur quadratique moyenne (REQM) en validation croisée la plus faible, est retenu. En outre, afin d'éviter un sur-ajustement du modèle qui peut conduire à une chute des performances en validation (Vancolen, 2004), le

NCLM a été toujours arbitrairement maintenu inférieur à la moitié du nombre d'échantillons du jeu de calibration (He et *al.*, 2007).

2.8. Evaluation des performances des modèles

Les performances des modèles calibrés en mode site spécifique ont été évaluées en deux étapes. Dans la première étape, le modèle est appliqué pour estimer les teneurs analytiques du jeu d'échantillons de validation collecté sur le même site d'apprentissage du modèle : jeu de données indépendant sélectionné par la procédure de l'ACP, dont les échantillons n'ont pas été utilisés pour la calibration du modèle. Dans la seconde étape, le modèle est appliqué pour estimer les teneurs analytiques des échantillons des deux autres sites locaux. Pour cela, nous avons tour à tour utilisé les équations de calibrations construites à l'échelle de chaque site pour estimer les teneurs analytiques des deux jeux d'échantillons prélevés sur les deux autres sites.

L'évaluation des performances des modèles s'est basée sur les critères statistiques suivants : le coefficient de détermination R^2 , le biais, la Racine de l'Erreur Quadratique Moyenne (REQM), ainsi que le Ratio des écarts d'estimation REP, défini comme le rapport entre l'écart type des valeurs prédites par le modèle et l'écart type des erreurs résiduelles d'estimation (Starr et *al.*, 1981 ; Islam et *al.*, 2004).

Le paramètre REP donne une information quantitative sur la qualité du modèle et des classes ont été établies permettant de connaître les valeurs pour lesquelles le modèle calibré peut être raisonnablement, ou non, utilisé pour des prédictions ultérieures. Cette classification résulte de l'expertise acquise au cours des diverses études pedospectrométriques antérieures. Le tableau 3 récapitule les interprétations associées à ces classes.

Tableau 3. Interprétation des valeurs obtenues pour le ratio REP dans la littérature

Valeur du ratio	Qualité du modèle calibré	Auteurs
REP < 1	Le modèle est de qualité médiocre	(Viscarra, 2006)
1 < REP < 1.5	Le modèle est peu recommandable pour des estimations ultérieures	(Sayes et <i>al.</i> , 2005 ; Viscarra, 2005)
1.5 < REP < 1.8	Le modèle peut être utilisé pour des estimations quantitatives	(Viscarra, 2006)
1.8 < REP < 2	Le modèle est performant	(Viscarra, 2006)
REP > 2	Le modèle est excellent	(Sinnaeve et <i>al.</i> , 2001 ; Brown et <i>al.</i> , 2005 ; Viscarra, 2006)

Il ressort de ce tableau qu'à partir d'un REP > 1.5, le modèle calibré peut être utilisé pour des estimations quantitatives.

Nous rappelons, ci-après, les équations de calcul des différents paramètres intervenant dans les critères statistiques que nous avons utilisé

$$R^2 = \left[\frac{\sum_{i=1}^n (y_i - \bar{Y})(y_i^* - \bar{Y}^*)}{\sqrt{\sum_{i=1}^n (y_i - \bar{Y})^2 \sum_{i=1}^n (y_i^* - \bar{Y}^*)^2}} \right]^2 \quad [1]$$

$$Biais = \frac{1}{n} \sum_{i=1}^n (y_i^* - y_i) \quad [2]$$

$$REQM = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i^* - y_i)^2} \quad [3]$$

$$REP = \sigma y_i^* / \sigma \mathcal{E}_i^* \quad (\text{Starr et al., 1981}). \quad [4]$$

y_i : valeur mesurée, y_i^* : valeur estimée, \bar{Y} : moyenne des valeurs mesurées, \bar{Y}^* : moyennes des valeurs prédites, σy_i^* : écart-type des valeurs prédites, $\sigma \mathcal{E}_i^*$: écart-type des erreurs de prédiction.

3. Résultats et discussion

3.1. Statistiques descriptives

Le tableau 4 ci-dessous, consigne les statistiques concernant les variables analytiques à l'échelle de chaque site spécifique, alors que le tableau 5 montre les mêmes statistiques lorsque nous considérons la base régionale des échantillons, tous sites confondus.

En examinant les tableaux 4 et 5, nous notons les faits suivants :

- les teneurs en carbone total les plus élevées sont au niveau du site d'Oudia. En effet, la moyenne des teneurs en carbone total est de 3.3 % pour ce site, alors qu'elle est respectivement de 0.9 et 1.0 % pour les sites de Chamsa et d'Oung. Les teneurs en azote sont quasi-nulles pour les trois sites. Ceci s'explique par le fait que les aires de collecte des échantillons sont des sols nus non mis en culture,
- les trois sites présentent des sols alcalins à pH basique, nettement supérieur à 7. D'après la distribution des fractions granulométriques, les trois sites sont formés par des sols à textures très grossières.
- d'après les valeurs de la $CE_{1:5}$ et du RS, qui sont deux mesures qui renseignent sur la même propriété du sol à savoir la salinité, le site le plus salé est le site d'Oung. En effet, il s'agit d'un site pré-chotteux situé en bordure du chott. Par ailleurs, les trois sites couvrent des gammes de teneur différentes. En effet, le site de Chamsa contient des échantillons dont les $CE_{1:5}$ sont comprises entre 0.80 et 7.10 dS/m alors que les $CE_{1:5}$ des échantillons collectés à Oudia sont comprises entre 0.88 et 53.3 dS/m et ceux collectés à Oung sont comprises entre 0.81 et 74.1 dS/m. Les sels sont en fait, redistribués par les vents à partir des bassins évaporites que sont les Chotts. Plus le site est proche du chott plus sa salinité est forte (Bryant, 1999). D'après Bryant (1999), les sels de ce chott sont formés pour l'essentiel par le gypse ($CaSO_4, 2H_2O$), l'halite (NaCl) et la carnallite ($KMgCl_3, 6H_2O$).
- aussi bien pour la variable carbone total que pour la variable $CE_{1:5}$, les teneurs mesurées s'étalent sur des larges gammes qui couvrent les variabilités rencontrées dans la région. Les moyennes régionales sont sensiblement différentes des moyennes relatives à chacun des trois sites locaux.

Tableau 4. Statistiques des variables analytiques à l'échelle de chacun des trois sites.

Localité de Chamsa (effectif = 30)				
Libellé de la variable*	Moyenne	Ecart-type	Minimum	Maximum
C _{tot} (%)	0.91	0.57	0.20	3.15
N _{tot} (‰)	0.00	0.01	0.00	0.06
pH _{eau 1:5}	7.70	0.17	7.38	8.09
CE _{1:5} (dS/m)	2.29	1.13	0.80	7.10
RS (g/l)	1.64	1.46	0.06	7.53
SG (%)	41.82	11.57	10.30	65.40
SM (%)	25.15	7.80	13.50	42.40
SF (%)	27.56	10.04	15.10	57.00
EF (%)	5.80	4.41	0.20	19.60
Localité d'Oudia (effectif = 29)				
Libellé de la variable	Moyenne	Ecart-type	Minimum	Maximum
C _{tot} (%)	3.26	1.51	0.87	5.34
N _{tot} (‰)	0.09	0.16	0.00	0.94
pH _{eau 1:5}	7.75	0.17	7.27	8.21
CE _{1:5} (dS/m)	8.06	11.43	0.88	53.30
RS (g/l)	9.48	14.97	0.71	75.74
SG (%)	13.96	14.05	0.90	50.20
SM (%)	37.83	16.93	5.90	71.80
SF (%)	42.93	17.09	17.00	85.00
EF (%)	5.26	7.22	0.20	32.60
Localité d'Oung (effectif = 30)				
Libellé de la variable	Moyenne	Ecart-type	Minimum	Maximum
C _{tot} (%)	1.01	0.57	0.35	2.43
N _{tot} (‰)	0.00	0.01	0.00	0.04
pH _{eau 1:5}	7.64	0.19	7.21	7.99
CE _{1:5} (dS/m)	14.79	18.32	0.81	74.10
RS (g/l)	11.01	13.77	0.29	60.67
SG (%)	29.93	18.25	4.00	65.30
SM (%)	31.41	12.03	13.70	60.10
SF (%)	34.71	13.67	16.70	69.60
EF (%)	3.95	4.40	0.60	20.30

* C_{tot} (%) : pourcentage de carbone total, N_{tot} (‰) : pourcentage d'azote total, pH_{eau 1:5} : potentiel d'hydrogène mesuré dans un extrait au 1/5, CE_{1:5} (dS/m) : Conductivité électrique mesurée dans un extrait au 1/5, RS : Résidu sec, SG (%) : fraction de sables grossiers, SM (%) : fraction de sables moyens, SF (%) : fraction de sables fins, EF (%) : fraction d'éléments fins.

Tableau 5. Statistiques des variables analytiques à l'échelle de la région (Tous sites confondus).

Région (effectif = 89)				
Libellé de la variable	Moyenne	Ecart-type	Minimum	Maximum
C _{tot} (%)	1.72	1.47	0.20	5.34
N _{tot} (%)	0.03	0.10	0.00	0.94
pH _{eau 1:5}	7.69	0.18	7.21	8.21
CE _{1:5} (dS/m)	8.38	13.50	0.80	74.10
RS (g/l)	7.35	12.43	0.06	75.74
SG (%)	28.73	18.74	0.90	65.40
SM (%)	31.39	13.76	5.90	71.80
SF (%)	34.97	15.21	15.10	85.00
EF (%)	5.00	5.54	0.20	32.60

* C_{tot} (%) : pourcentage de carbone total, N_{tot} (%) : pourcentage d'azote total, pH_{eau 1:5} : potentiel d'hydrogène mesuré dans un extrait au 1/5, CE_{1:5} (dS/m) : Conductivité électrique mesurée dans un extrait au 1/5, RS : Résidu sec, SG (%) : fraction de sables grossiers, SM (%) : fraction de sables moyens, SF (%) : fraction de sables fins, EF (%) : fraction d'éléments fins.

3.2. Analyse multivariée des jeux de données

La figure 3 montre la projection des variables pédologiques dans le premier plan factoriel de l'ACP pour les quatre jeux d'échantillons.

Dans le cas du jeu d'échantillon propre au site de Chamsa (Figure 3a), nous constatons que le premier plan factoriel de l'ACP explique 45 % de la variabilité qui existe au sein des échantillons. En fait, le premier axe factoriel explique 28.9 % de cette variabilité alors que le second en explique 22.63 %. Les variables initiales contribuent dans la formation de ce premier axe dans l'ordre d'importance suivant : SF, SG, RS, CE_{1:5}, pH, EF, C_{tot}, N_{tot} et SM. Toutefois les variables de texture SF et SG sont les plus déterminantes. Ces mêmes variables initiales contribuent dans la formation du second axe dans l'ordre d'importance suivant : CE_{1:5}, RS, SM, N_{tot}, C_{tot}, EF, pH, SF et SG. Par contre, pour ce deuxième axe ce sont les variables chimiques CE_{1:5}, RS, N_{tot} et C_{tot} qui sont les plus déterminantes dans sa formation. Par ailleurs, nous constatons l'existence d'une forte corrélation entre les variables CE_{1:5} et RS ce qui est en accord avec le fait que ces deux variables sont directement liées à la même propriété pédologique, la salinité. Une forte corrélation existe aussi entre le C_{tot} et les EF, ce qui signifie que pour ce site l'essentiel du carbone total (organique et minéral) existe dans la fraction < 63 µm.

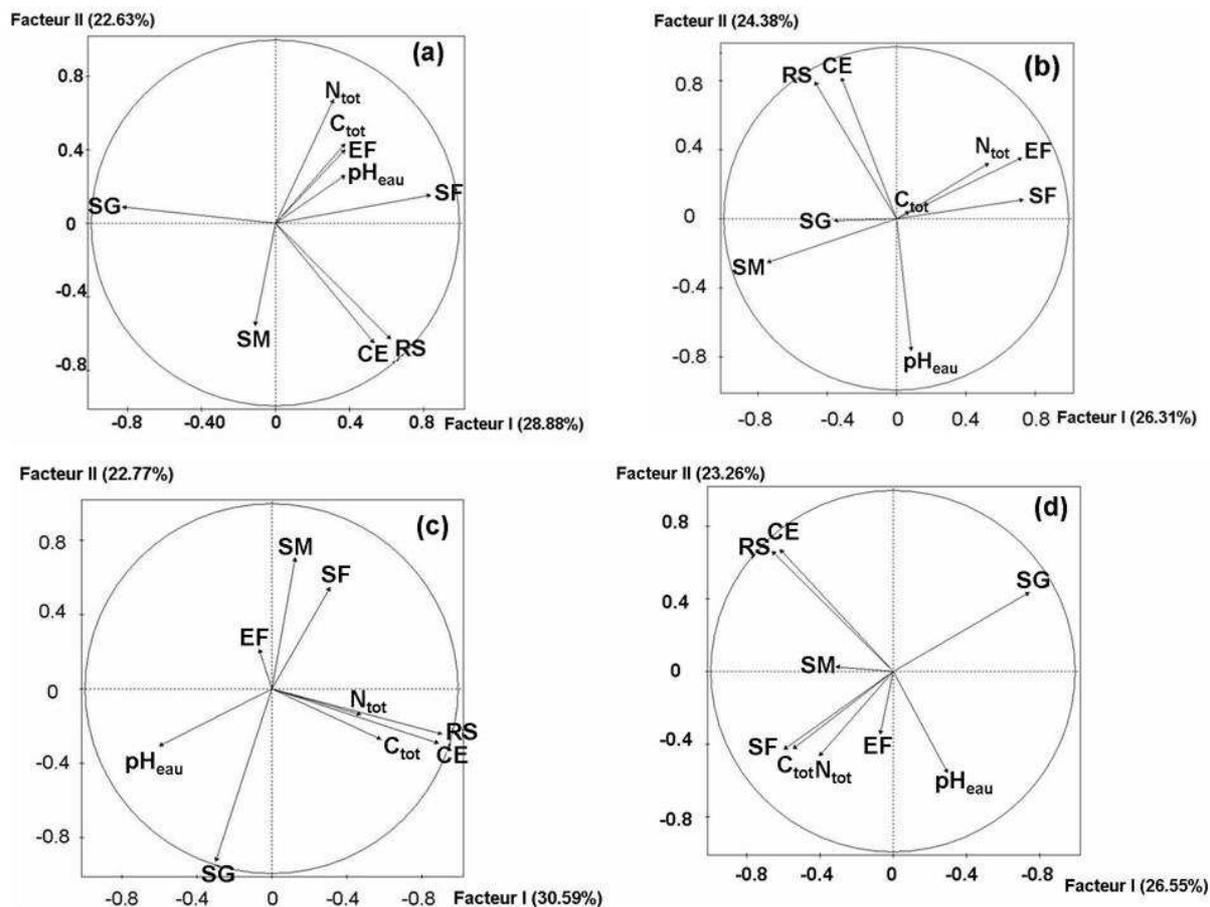


Figure 3. Projection des variables pédologiques dans le premier plan factoriel de l'ACP pour les jeux d'échantillons de Chamsa (a), Oudia (b), Oung (c) et pour le jeu d'échantillons régional (d).

Dans le cas du jeu d'échantillon propre au site d'Oudia (Figure 3b), nous constatons que le premier plan factoriel de l'ACP explique 50.7 % de la variabilité qui existe au sein des échantillons. En effet, le premier axe factoriel explique 26.3 % de cette variabilité alors que le second explique 24.4 %. Les variables initiales contribuent dans la formation de ce premier axe dans l'ordre d'importance suivant : SM, SF, EF, N_{tot} , RS, SG, $CE_{1:5}$, pH et C_{tot} . Néanmoins, les variables les plus déterminantes sont les SM, les SF et les EF. L'importance de ces variables dans la formation du second axe est dans l'ordre suivant : pH, $CE_{1:5}$, RS, EF, SM, N_{tot} , SF, C_{tot} et SG. Par ailleurs, nous constatons l'existence d'une forte corrélation positive entre les variables $CE_{1:5}$ et RS d'une part et entre les variables C_{tot} , N_{tot} et EF d'autre part.

Dans le cas du jeu d'échantillons propre au site d'Oung (Figure 3c), nous constatons que le premier plan factoriel de l'ACP explique 53.4 % de la variabilité qui existe au sein des échantillons. En effet, le premier axe factoriel explique 30.6 % de la variabilité alors que le

second axe en explique 22.8 %. Les variables initiales contribuent dans la formation de ce premier axe dans l'ordre d'importance suivant : RS, CE_{1:5}, pH, C_{tot}, N_{tot}, SF, SG, SM et EF. Avec une prépondérance pour les cinq premières variables. Quant au second axe ces variables initiales interviennent dans l'ordre suivant : SG, SM, SF, pH, CE_{1:5}, C_{tot}, RS, EF, et N_{tot}. Par ailleurs, dans le cas de ce site, nous remarquons l'existence d'une forte corrélation positive entre les variables initiales N_{tot}, RS, CE_{1:5} et C_{tot}. Ces corrélations inter variables s'expliquent par le fait que l'essentiel des sels est intimement lié aux carbonates, d'autant plus que le site est colonisé par une végétation naturelle hydro-halophyte (salicornes).

Dans le cas du jeu de données régionales (Figure 3d) (constitué par le regroupement des trois jeux propres à chaque site), le premier plan factoriel explique 49.8 % de la variabilité qui existe au sein des échantillons de la région. Le premier axe factoriel explique 26.55 % de cette variabilité alors que le second axe explique 22.6 %. Les variables initiales entre dans la composition de ce premier axe factoriel dans l'ordre d'importance suivant : RS, CE_{1:5}, SF, C_{tot}, SG, N_{tot}, pH, SM et EF. Par contre, l'importance de ces variables initiales dans la formation du second axe est dans l'ordre suivant : CE_{1:5}, RS, SG, pH, N_{tot}, C_{tot}, SF, SM et EF. Par ailleurs, à l'échelle régionale, nous constatons l'existence d'une forte corrélation entre les variables RS et CE_{1:5} d'une part et entre le C_{tot} et la fraction SF d'autre part.

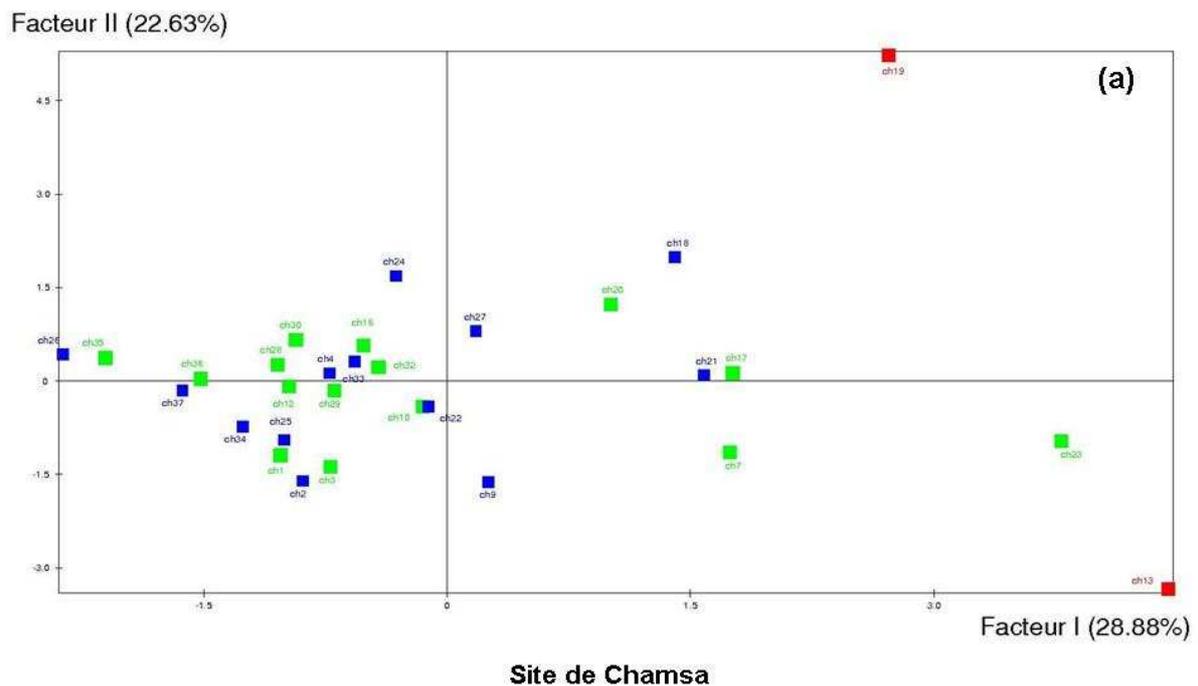
Ainsi, d'une façon générale, nous constatons que les corrélations entre les propriétés pédologiques varient d'un site à l'autre ce qui montre la pédo-diversité des trois sites échantillonnés et ce bien qu'ils appartiennent à la même région naturelle. Toutefois, les variables C_{tot} et N_{tot} ainsi que les variables RS et CE_{1:5} sont toujours fortement corrélées et ce quel que soit le site. Ces fortes corrélations sont vérifiées aussi lorsque l'ACP est appliquée à une échelle régionale. En outre, nous pouvons constater que le pH et la CE_{1:5} sont corrélés négativement sur les trois sites prospectés. Ceci s'explique par l'effet tampon que joue l'ion bicarbonaté (CO₃⁼) (omniprésence du calcaire dans toute la région) dans la variation du pH. En effet, la CE_{1:5} peut augmenter mais le pH demeure stable.

D'une façon globale en observant l'ACP propre à chaque site, et en particulier le premier axe de ces ACP, nous pouvons nous rendre compte des propriétés qui conditionnent le plus l'évolution pédogénétique de chaque site. Ainsi pour le site de Chamsa, considérée comme zone d'arrachage par l'effet éolien, ce sont les variables de texture SG et SF qui départagent les échantillons collectés. Alors que pour le site d'Oudia, exposé au vent et considéré comme zone d'accumulation, ce sont également les fractions granulométriques qui vont différencier les échantillons quoique dans ce cas ce sont surtout les EF, les SF et les SM

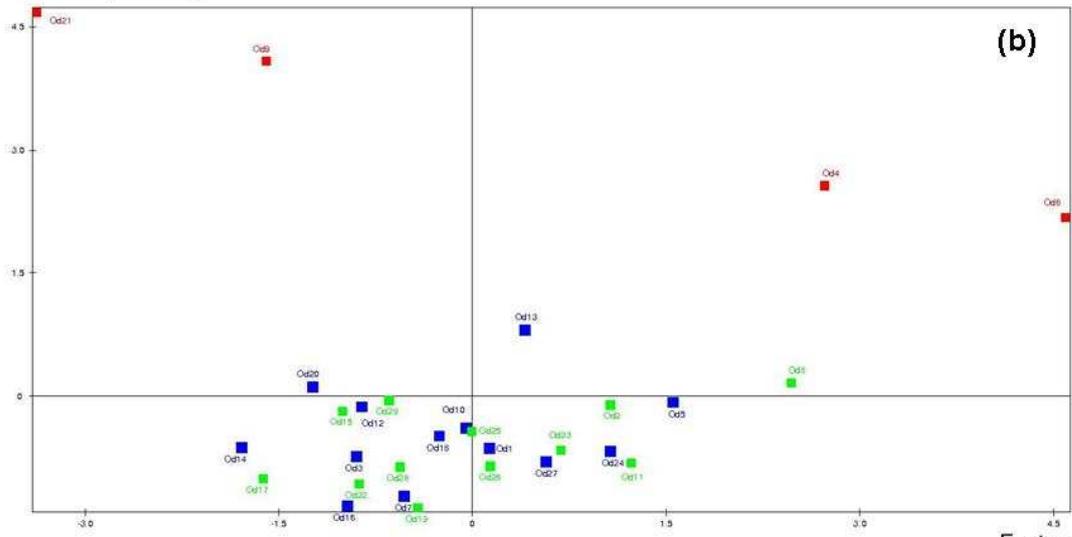
qui interviennent. Quant au site pré-chauffé d'Oung : ce sont la conductivité électrique et le pH qui discriminent le plus entre les échantillons.

3.3. Sélection des échantillons de calibration et de validation

La figure 4 identifie les jeux de calibration et de validation sélectionnés par la procédure de l'ACP réalisée sur les échantillons des 3 sites pris séparément (Figures 4a, 4b et 4c) et sur la totalité des échantillons (Figure 4d, base régionale). Dans chaque cas de figure, les échantillons ont été scindés en un jeu de calibration et un jeu de validation après leurs projections dans le premier plan factoriel de l'ACP. Les échantillons atypiques qui se trouvent loin du centre du plan factoriel, et à l'écart du peloton des échantillons, ont été à chaque fois éliminés des bases des échantillons. De plus, nous avons repéré sur ces figures les échantillons atypiques par des Carrés rouges, les échantillons affectés dans le jeu de calibration par des carrés bleus et les échantillons du jeu de validation par des carrés verts.



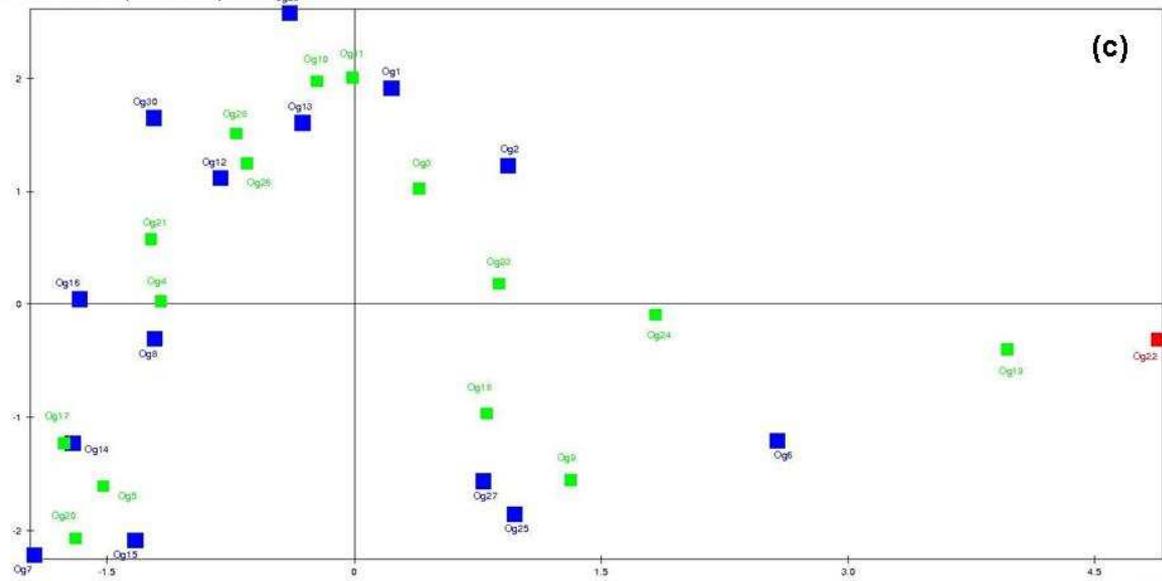
Facteur II (24.38%)



Facteur I (26.31%)

Site d'Oudia

Facteur II (22.77%)



Facteur I (30.59%)

Site d'Oung

distinguent en conséquence, par une texture plus fine. En effet, le site d'oudia est situé dans le couloir du vent dominant de direction Nord-Est (Kamel, 2007) et il est considéré comme une zone d'accumulation (dépôts des particules fines). Alors que le site d'oung est un site préchotteux qui reçoit les dépôts alluvionnaires. Les échantillons de ces deux sites sont départagés selon l'axe des ordonnées de l'ACP, eu égard aux valeurs de leurs CE. Les échantillons d'Oung qui présentaient la gamme de CE la plus large présente la dispersion la plus importante selon ce second axe de l'ACP.

Il découle de ces résultats que la similarité entre les échantillons issus d'un même site est en général supérieure à celle entre les échantillons issus de deux sites différents, et ce bien qu'ils appartiennent à la même région naturelle.

Ces observations confirment aussi que le vent et la présence du chott sont deux facteurs qui influencent directement la nature de la couverture pédologique rencontrée dans la région (Bryant, 1999 ; Melis, 2003). Par ailleurs, un échantillon considéré comme atypique à l'échelle d'un site donné, peut devenir normal et modélisable lorsqu'on passe à une échelle régionale. C'est le cas par exemple des échantillons ch13 et ch19 collectés sur le site de Chamsa (Figures 4a et 4d). Ceci indique qu'un échantillon collecté sur un site quelconque peut présenter des propriétés pédologiques qui sont différentes du reste des échantillons collectés sur le même site, mais s'apparente bien à une population d'échantillons collectées à une échelle spatiale plus large. Par contre d'autres échantillons demeurent atypiques même à l'échelle régionale. C'est le cas des échantillons Od9 et Od21 collecté sur le site d'Oudia et Og22 collecté sur le site d'Oung (Figure 4b, 4c et 4d).

3.4. Performance des modèles « site spécifique »

La figure 5, montre les valeurs prédites vs mesurées pour la variable C_{tot} pour les sites de Chamsa (5a), Oudia (5b) et Oung (5c)

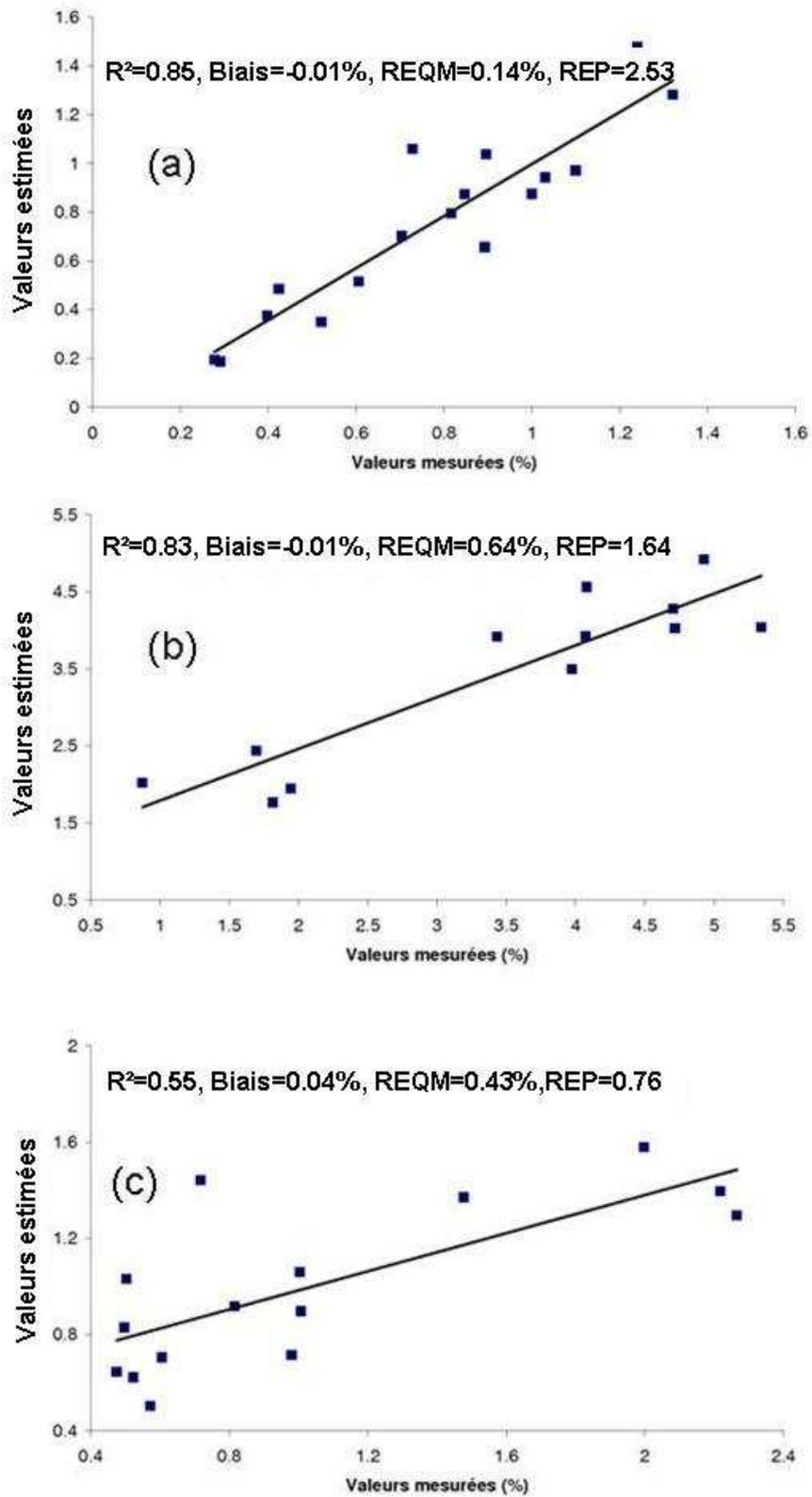


Figure 5. Performances des modèles « site spécifiques » pour la variable C_{tot} . Valeurs prédites vs mesurées pour les sites de Chamsa (5a), Oudia (5b) et Oung (5c).

Pour la variable C_{tot} , nous avons obtenu des bonnes estimations pour le site de Chamsa (Figure 5a) et le site d'Oudia (Figure 5b) et une estimation assez satisfaisante dans le cas du site d'Oung (Figure 5c). En effet, pour le site de Chamsa, les critères d'évaluation de la performance du modèle présente les valeurs suivantes : $R^2 = 0.85$, Biais = -0.01 %, REQM = 0.14 % et REP = 2.53. De même, pour le site d'Oudia les critères d'évaluation de la performance du modèle présente les valeurs suivantes : $R^2 = 0.83$, Biais = -0.01 %, REQM = 0.64 % et REP = 1.64.

Les performances obtenues pour les deux premiers sites (Chamsa et Oudia) suggèrent qu'il est possible d'utiliser les équations de ces modèles pour l'estimation des teneurs en C_{tot} pour d'autres échantillons collectés sur ces mêmes sites.

Par contre, le modèle associé au site d'Oung est peu satisfaisant et son REP < 1, (voir Tableau 3), suggère qu'il ne peut être recommandé pour des estimations quantitatives : $R^2 = 0.55$, Biais = 0.04 %, REQM = 0.43 % et REP = 0.76. Toutefois, Il peut être utilisé pour des appréciations grossières des teneurs en C_{tot} et pour discerner entre les échantillons, collectés sur ce même site, qui pourraient être riches en C_{tot} et ceux qui le seraient moins.

Les moins bonnes performances du modèle du site d'Oung par rapport aux deux autres, s'expliquent par la forte salinité de ce site (Oung) pré-chotteux. En effet, il est démontré que les sels influencent d'une façon directe la réponse spectrale des sols (Mougenot, 1993 ; Escadafal, 1989 ; Ben Dor, 2002). En forte concentration, les sels peuvent perturber sinon masquer l'effet des autres paramètres qui influencent la réponse spectrale, en particulier le C_{tot} .

La figure 6, montre les valeurs prédites vs mesurées pour la variable $CE_{1:5}$. Dans le cas de la variable $CE_{1:5}$, nous avons obtenu une qualité d'estimation peu satisfaisante pour les sites de Chamsa (Figure 6a) et d'Oudia (Figure 6b) et une bonne estimation pour le site d'Oung (Figure 6c). En effet, pour le site de Chamsa, le modèle calibré présente les performances suivantes : $R^2 = 0.57$, Biais = 0.12 dS/m, REQM = 0.54 dS/m et REP = 0.32. De même pour le site d'Oudia, les critères d'évaluation de la performance du modèle présentent les valeurs suivantes : $R^2 = 0.51$, Biais = 0.37 dS/m, REQM = 2.5 dS/m et REP = 1.42. Par contre pour le site d'Oung, les critères d'évaluation de la performance du modèle présentent les valeurs suivantes : $R^2 = 0.76$, Biais = 9.77 dS/m, REQM = 12.4 dS/m et REP = 2.01. Pour ce dernier site, bien que les valeurs du biais et de la REQM soient relativement élevées, ces valeurs demeurent acceptables. En effet, rappelons que les échantillons collectés sur ce site présentent une conductivité électrique moyenne de 14.79 dS/m et un écart type de 18.32 dS/m (Tableau 4).

Ainsi en se basant sur les valeurs du R^2 et du REP, Ce modèle peut être utilisé pour des estimations futures des conductivités électriques d'échantillons collectés sur ce même site. Par ailleurs, nous estimons que la faible relation entre la $CE_{1:5}$ et la réflectance, pour les sites de Chamsa et d'Oudia, comparés au site d'Oung, est probablement attribuable aux faibles concentrations en sels des deux premiers sites comparés au troisième. Les sels influencent le plus les spectres des sols lorsque leurs concentrations sont élevées. En effet, dans une expérience réalisée en conditions de laboratoire sur une série d'échantillons de sols sableux mélangés à différentes proportions pondérales d'halite (NaCl), Mougnot (1993) a montré que l'effet de ce sel ne se fait sentir sur le comportement spectral du sol qu'à partir d'un taux pondéral de 10 à 15 %.

Le tableau 6 résume les performances des modèles qui ont été calibrés en mode « site spécifique », lorsque l'équation de chacun de ces modèles est utilisée pour estimer, les teneurs des deux jeux d'échantillons collectés dans les deux autres sites.

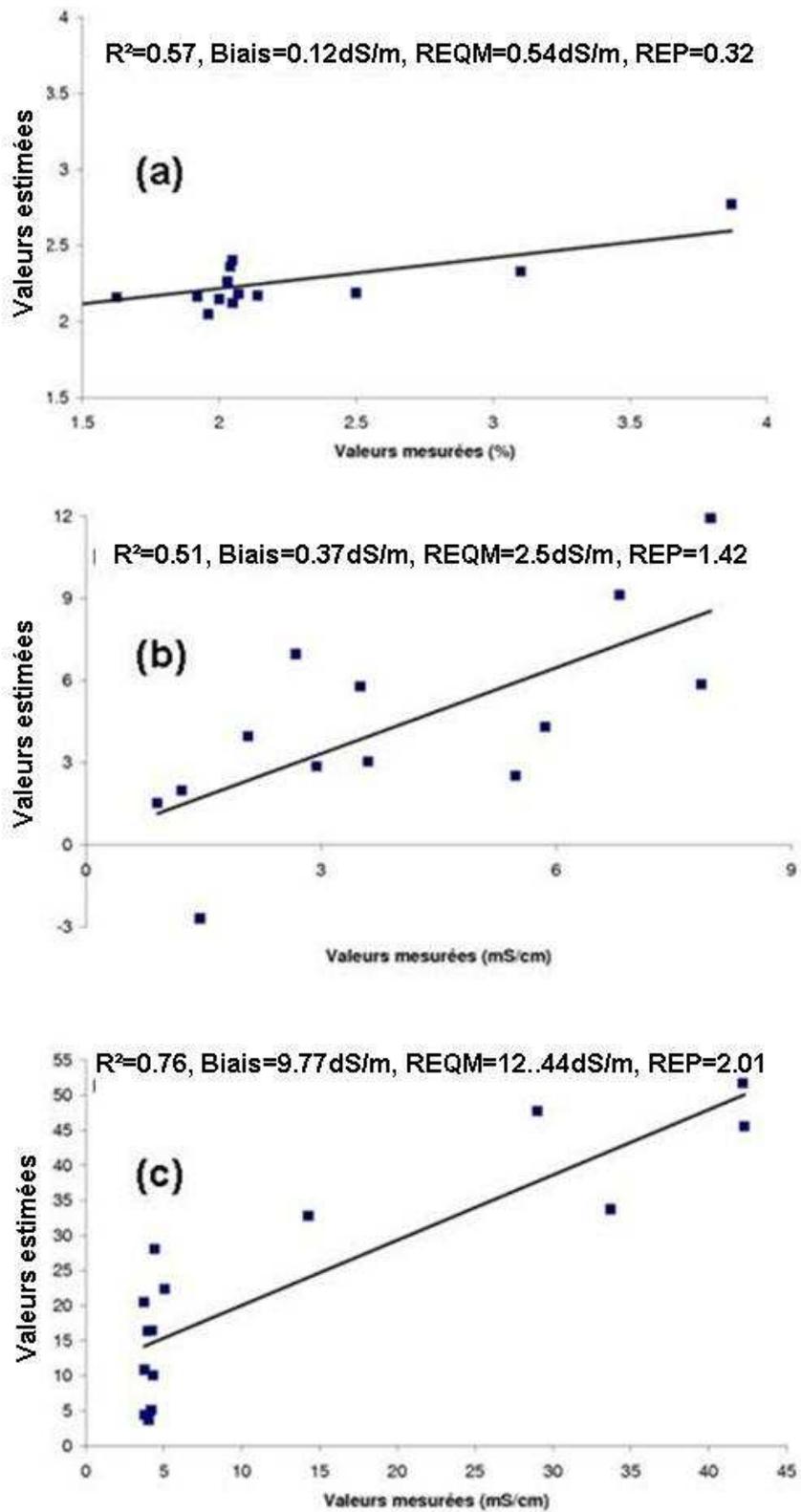


Figure 6. Performances des modèles « site spécifiques » pour la variable $CE_{1.5}$. Valeurs prédites vs mesurées pour les sites de Chamsa (6a), Oudia (6b) et Oung (6c).

Tableau 6. Performances des modèles calibrés en mode « site spécifique », lorsque l'équation de chacun de ces modèles est utilisée pour estimer, les teneurs des deux jeux d'échantillons collectés dans les deux autres sites.

	Variable C_{tot}					
	Ch/Od	Ch/Og	Od/Ch	Od/Og	Og/Ch	Og/Od
R ²	0.51	0.20	0.02	0.22	0.07	0.50
Biais (%)	-3.74	-0.38	-0.01	-1.41	0.92	-1.56
REQM (%)	4.66	2.01	7.7	0.3	0.21	0.42
REP	0.55	0.91	0.37	0.65	0.80	1.09
	Variable CE					
	Ch/Od	Ch/Og	Od/Ch	Od/Og	Og/Ch	Og/Od
R ² (%)	0.04	0.00	0.03	0.00	0.01	0.00
Biais (dS/m)	-2.21	-12.51	-3.05	-2.96	47.29	36.91
REQM (dS/m)	1.21	4.06	2.26	3.81	11.7	8.05
REP	0.07	0.06	0.97	0.4	0.02	0.5

Ch : Chamsa, Od : Oudia et Og : Oung. Site(i)/site(j) (par exemple Ch/Od) ; Equation du modèle de Chamsa utilisée pour estimer les teneurs des échantillons d'Oudia.

En examinant l'ensemble des critères d'évaluation (R², Biais, REQM et REP), nous constatons qu'aussi bien dans le cas de la variable C_{tot} que dans le cas de la variable CE_{1:5}, les modèles « site spécifique » ne sont pas performants lorsqu'ils sont utilisés pour l'estimation des teneurs d'échantillons collectés en dehors du site auquel appartenaient les échantillons de calibration.

Rappelons que pour ces deux variables les sites couvraient des gammes de teneurs qui sont significativement différentes. En outre, dans le cas de la variable CE_{1:5}, les modèles spécifiques aux sites de Chamsa et d'Oudia, calibrés en mode site spécifique étaient déjà peu performants pour estimer le jeu de validation issu du même site, et donc à *fortiori* peu performants en estimation des teneurs des échantillons prélevés sur les deux autres sites.

Le modèle spécifique au site d'Oung était performant en validation en mode site spécifique. Par contre, bien que son étendue de concentration de CE_{1:5} couvre celles des deux autres sites, il n'en demeure pas moins que la différence des propriétés pédologiques entre ces sites conduit à un échec d'estimation. En outre, d'après les travaux de Tillman et *al.* (2001), cités par Terhoeven-Urselmans et *al.* (2006), la qualité d'estimation est médiocre lorsque les estimations sont faites sur des échantillons qui sont en dehors de la gamme des échantillons de calibration.

Ainsi, nous montrons que les modèles « site spécifique » sont applicables aux sites dont sont issus les échantillons de calibration et échouent dans l'estimation des teneurs des échantillons qui sont prélevés dans un autre site ou qui sont en dehors de l'étendue de la gamme des teneurs du jeu de calibration ou qui présentent des propriétés pédologiques qui soient différents.

3.5. Performances des modèles établis à l'échelle régionale

La figure 7, montre les valeurs prédites vs mesurées dans le cas des modèles régionaux de C_{tot} (7a) et de $CE_{1:5}$ (7b).

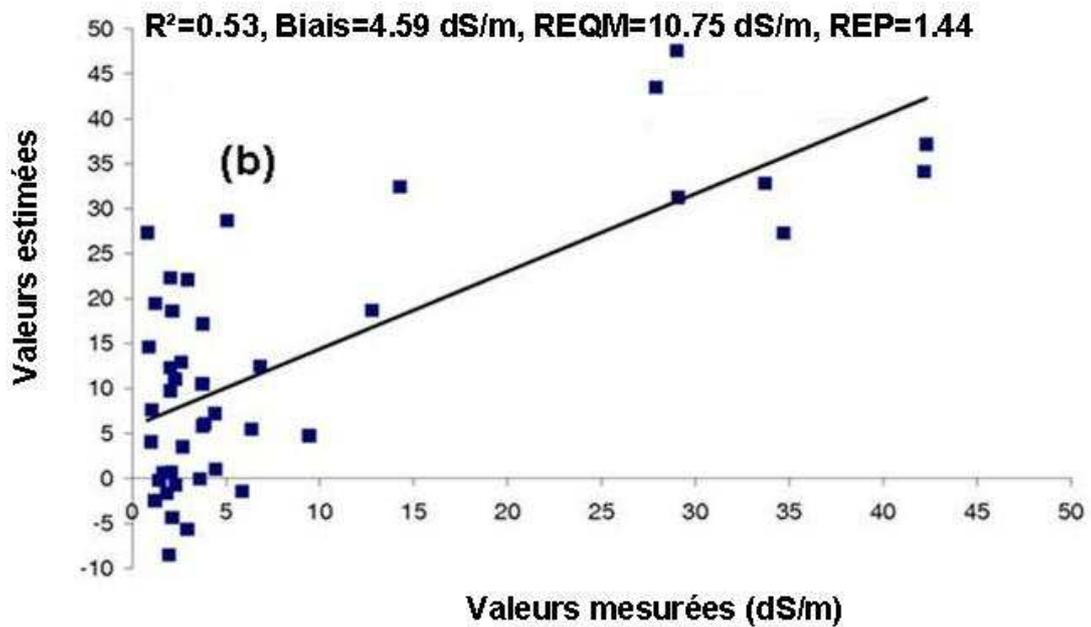
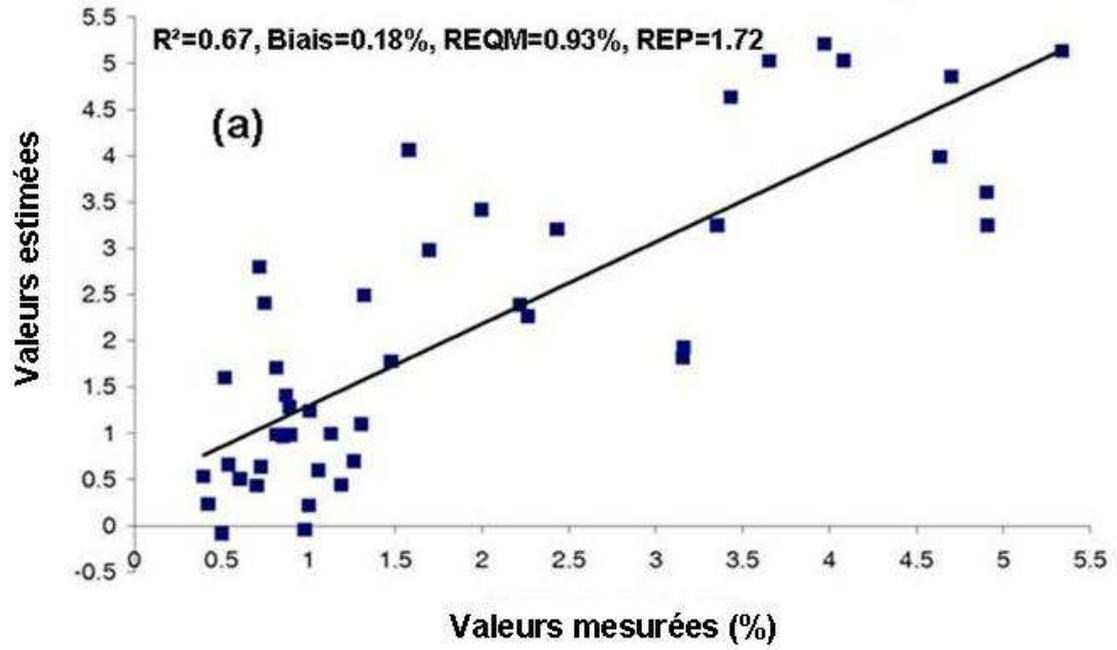


Figure 7. Performances des modèles régionaux (Valeurs prédites vs mesurées, pour la variable C_{tot} (7a) et la variable $CE_{1:5}$ (7b)).

Ces modèles présentent respectivement, pour la variable C_{tot} (Figure (7a)) et la variable $CE_{1:5}$ (Figure (7b)), les performances d'estimation suivantes : ($R^2 = 0.67$, Biais = 0.18 dS/m, REQM = 0.93 dS/m et REP = 1.72) et ($R^2 = 0.53$, Biais = 4.59 dS/m, REQM = 10.75 dS/m et REP = 1.44).

Pour la variable C_{tot} , le modèle régional est moins performant que les modèles sites spécifiques de Chamsa et d'Oudia (Figures (5a) et (5b)), mais il est plus performant que le modèle site spécifique d'Oung (Figure (5c)). Par contre pour la variable $CE_{1:5}$, le modèle régional est plus performant que les modèles sites spécifiques de Chamsa et d'Oudia (Figures (6a) et (6b)), mais il est moins performant que le modèle site spécifique d'Oung (Figure 6c).

Par ailleurs, ces deux modèles régionaux sont incontestablement plus performants que les trois modèles « site spécifique » lorsque chacun de ces derniers est utilisé pour estimer les teneurs en C_{tot} et de $CE_{1:5}$ des échantillons collectés dans les deux autres sites.

Il apparaît pour ces deux variables que les modèles calibrés en mode régional sont capables d'intégrer la variabilité intrinsèque des échantillons de sols, prélevés dans l'ensemble des trois sites. Par conséquent lorsque l'échantillon objet de l'estimation n'appartient pas au site spécifique pour lequel le modèle d'estimation a été calibré, il est plus prudent d'utiliser un modèle d'estimation qui soit à portée régionale, que d'utiliser un modèle « site spécifique » calibré pour un site différent du site de collecte de l'échantillon objet de l'estimation.

Dans le cas de la variable C_{tot} , les résultats d'estimation obtenus avec le modèle régional sont prometteurs. Nous sommes confiants que la qualité d'estimation peut être améliorée davantage en incorporant dans le jeu de calibration un nombre plus important d'échantillons de sols collectés dans ces trois sites ainsi que dans d'autres sites spécifiques appartenant à la même région géographique. Tout comme Cozzolino et Moron (2006) et Genot et al. (2007), nous sommes convaincus qu'en spectrométrie les modèles qui peuvent être calibrés présentent un besoin régulier en recalibration et en contrôle de la qualité d'estimation.

Dans le cas de la variable $CE_{1:5}$, le modèle régional calibré peut être utilisé pour discerner grossièrement entre les échantillons riches en sels et les échantillons qui le sont moins. Notre échec dans la calibration d'un modèle régional performant, pourrait être attribuable aux faibles et moyennes concentrations en sels des échantillons issus des sites de

Chamsa et d'Oudia. Or, comme il est signalé plus haut, à faible concentration, les sels n'influencent pas significativement le comportement spectral des sols et ainsi les concentrations en sels ne peuvent pas être dérivées de la réponse spectrale.

Par ailleurs, d'après Rao et *al.* (1991), les sols salés ont le plus souvent des réflectances élevées dans le Vis et le PIR. Toutefois, lorsque les sols sont sableux, il est difficile de discriminer les sols salés car tout comme les sels, les cristaux de quartz ont aussi des réflectances relativement élevées dans le Vis et le PIR (Verma et *al.*, 1994). Ainsi, dans un contexte de milieu aride saharien, il est difficile de calibrer un modèle d'estimation de la conductivité électrique à étendue régionale. Toutefois pour les sites préchotieux à forte salinité, il est possible de calibrer un modèle spécifique du site, qui soit performant.

4. Conclusion

La procédure d'ACP a permis de visualiser et de mettre en évidence l'existence d'un effet « site ». Chaque site a ses spécificités pédologiques.

En milieu aride, l'étude confirme l'intérêt de la SRD Vis-PIR pour une détermination fiable du C_{tot} , particulièrement lorsque le modèle est calibré en mode site spécifique pour des sites peu affecté par la salinité (trait pédologique qui caractérise la plupart des aridosols).

L'échec d'estimation, en dehors des sites d'apprentissage des modèles, témoigne d'une robustesse géographiquement limitée des modèles calibrés en mode site spécifique.

Lorsque nous disposons d'un modèle site spécifique performant, il vaut mieux utiliser ce modèle, autrement il faut utiliser un modèle régional. En effet, ce dernier même s'il n'est pas aussi précis qu'un modèle site spécifique, il permet d'offrir des estimations satisfaisantes du fait qu'il intègre une variabilité d'échantillons plus large. Ainsi pour une analyse en grande série des teneurs en C_{tot} , il est plus opportun d'utiliser les modèles calibrés en mode site spécifique, alors que pour d'autres applications comme la cartographie à échelle régionale, il est plus opportun d'utiliser le modèle calibré en mode régional.

Pour la variable $CE_{1:5}$, nos résultats confirment l'intérêt de la SRD Vis-PIR dans l'estimation de cette variable, lorsque les échantillons sont riches en sels notamment dans le cas du site pré-chotieux. Par contre pour les sites où la salinité du sol est moins prononcée, nous estimons qu'un travail préalable doit être fait pour identifier la nature des sels présents dans le sol. Ces éléments chimiques influencent d'une manière directe le comportement

pedospectrale. Il est alors tout à fait concevable que pour un type de sol donné, une même concentration en sels solubles peut conduire à des spectres de sols différents, en fonction notamment de la nature de ces sels et de leurs concentrations relatives. Il est alors dans ce cas plus pertinent que la modélisation spectrométrique porte sur la concentration en ces éléments simples plutôt que sur la $CE_{1:5}$.

Chapitre V. Quantification spatiale de propriétés du sol en milieu aride, par couplage de données spectroscopiques visible-proche infrarouge acquises au laboratoire et de données d'image ASTER

« Qui voit de haut, voit bien ; qui voit de loin voit juste »

Victor Hugo, L'art d'être grand père, l'exilé satisfait, 1877

Abstract

The spatial quantification of soil properties is required to manage and monitor soil resources. Our aim is to set up a soil properties mapping approach, by means of Partial Least Squares Regression-kriging (PLSR-kriging), while combining data contained in the 9 Visible-Near infrared (Vis-NIR) spectral bands of an ASTER sub-image and a collection of 144 soil samples spectra acquired in laboratory in the 400-2500 nm range. This approach is tested in the Djerid arid region (Southwest of Tunisia) to map total carbon (totC) and the electric conductivity ($EC_{1:5}$) over 580 ha of bare soils. On these samples, collected in nodes of a square grid of 200 m of side, dried and sifted to 2 mm, we measured the totC and the $EC_{1:5}$, then acquired spectra. For each property, we calibrated a PLSR model, based on the 144 spectra extracted from the sub-image, previously corrected radiometrically with the empirical line method. This model, when applied to the 9 bands of the sub-image, generates a first prediction map of the property. Residues have been calculated at sampled grid nodes. The residue kriging produces a raster layer, which, when added to the first map, offers a final prediction map. Residual kriging has improved PLSR prediction accuracy. Indeed, for the totC we obtain respectively ($R^2 = 0.78$ and $RMSE = 0.16\%$) against to ($R^2 = 0.53$ and $RMSE = 0.52\%$). For the EC, although R^2 improvement is slight 0.60 as opposed to 0.57, the remainder performance criteria have significantly improved : bias (0.003 as opposed to 0.36dS/m) and RMSE (0.55 as opposed to 0.75 dS/m). Furthermore the spatial distribution of these quantifications has a physical significance. From our results emerge interesting perspectives for soil properties mapping on large territories.

Key words : spatial quantification, PLSR, ASTER, total carbon, electric conductivity.

Résumé

La quantification spatiale des propriétés pédologiques est nécessaire pour gérer et surveiller la ressource sol. Notre objectif est de développer une approche de cartographie des propriétés du sol par Régression en Moindres Carrées Partiels-krigeage (RMCP-krigeage), en combinant des informations issues des 9 bandes visible proche infrarouge (Vis-PIR) d'une sub-image ASTER et d'une collection de 144 spectres d'échantillons de sol acquis au laboratoire dans le Vis-PIR (400-2500 nm). Cette approche est testée, dans la région aride du Djérid (Sud-Ouest tunisien) pour cartographier, le carbone total (C_{tot}) et la conductivité électrique ($CE_{1.5}$) sur 580 ha de sols nus. Sur ces échantillons, collectés aux nœuds d'une grille carrée de 200 m de côté, séchés et tamisés à 2mm, nous avons mesuré le C_{tot} et la $CE_{1.5}$, puis acquis les spectres. Pour chaque propriété, nous avons calibré un modèle de RMCP, basé sur les 144 spectres extraits de la sub-image, radiométriquement corrigé par la méthode de la ligne empirique. Ce modèle, lorsqu'il est appliqué aux 9 bandes de la sub-image, génère une première carte d'estimation de la propriété. Les résidus ont été calculés aux niveaux des nœuds des mailles échantillonnées. Le krigeage des résidus produit une couche raster dont la sommation avec la première carte offre une carte finale d'estimation. Le krigeage des résidus a amélioré la qualité d'estimation par régression MCP. En effet, pour le C_{tot} nous obtenons respectivement ($R^2 = 0.78$ et $REQM = 0.16$ %) contre ($R^2 = 0.53$ et $REQM = 0.52$ %). Pour la CE, bien que l'amélioration de R^2 soit moindre (0.60 contre 0.57), les autres critères de performance sont significativement améliorés, notamment le biais (0.003 contre 0.36 dS/m) et la REQM (0.55 contre 0.75 dS/m). Par ailleurs, la distribution spatiale de ces quantifications possède bien un sens physique. Nos résultats font émerger des perspectives intéressantes pour cartographier les propriétés du sol sur de larges territoires

Mots clés : quantification spatiale, RMCP, ASTER, carbone total, conductivité électrique.

1. Introduction

La cartographie des propriétés physico-chimique du sol est un véritable outil d'information et de communication (Carré, 2002). La compréhension de la répartition spatiale des propriétés du sol est un enjeu majeur, en lien notamment avec la surveillance et l'aménagement des sols. La spectrométrie offre une méthode alternative pour l'estimation d'une large gamme de propriétés du sol (Lagacherie et al 2008). Pour nombre de ces propriétés, il est démontré que les données spectroradiométriques acquises au laboratoire permettent de bonnes estimations (Viscarra et al., 2006).

Bien que celles-ci soient ponctuelles, elles permettent de générer une masse importante de données d'entrée pour les modèles d'interpolation (Bilgili et al., 2007). Toutefois leur spatialisation par simple recours aux techniques géostatistiques, telle que réalisée par exemple par Bilgili et al. (2007), peut conduire à des estimations peu précises (Bonn, 1998). En général, lorsque l'interpolation se base sur uniquement des données ponctuelles, les cartes générées peuvent être très lissées quand la résolution d'échantillonnage est insuffisante pour bien appréhender les variations spatiales à courte distance (Walter, 2002 ; Douaoui et al., 2006 ; Minasny et Mc Bratney, 2007).

D'autres travaux se sont intéressés à l'application de la spectroscopie spatiale pour estimer des propriétés de sols, notamment les travaux de Baptista et Netto (2001) qui ont prédit spatialement les manifestations d'érosion, Chabrillat et al. (2002) les minéraux argileux, Ben-Dor et al. (2002, 2004) la matière organique et la salinité, Gomez et Lagacherie (2007) le carbone organique, Crouvi et al. (2006) les dépôts alluviaux en milieu aride.

Toutefois, tous ces derniers travaux de spatialisation ont eu pour support des images hyperspectrales. Comparés aux capteurs multispectraux, les capteurs hyperspectraux se distinguent par un nombre plus important de bandes spectrales, ce qui permet d'acquérir un spectre plus complet et laisse entrevoir une estimation plus précise des propriétés du sol (Stevens et al., 2006). Néanmoins, en raison de coûts élevés des campagnes de mesures avec les capteurs hyperspectraux, il est difficile de concevoir que ces études se multiplient en sciences du sol dans le futur proche.

La spatialisation est le processus d'estimation des valeurs d'une propriété donnée à un endroit non enquêté, par la prise en compte d'observations voisines ou par modélisation des processus spatiaux (Walter, 2002). Lorsqu'elle est appliquée à l'ensemble d'un site d'étude, elle est aussi connue sous le nom d'interpolation spatiale ou de cartographie (Hengl et al.,

2004). La cartographie des propriétés du sol, sur une large superficie avec une précision et un coût abordable, nécessite de développer et de promouvoir des méthodes spécifiques qui utilisent les informations disponibles et qui minimisent l'effectif des échantillons à un nombre qui soit raisonnable (Lagacherie et Voltz, 2000). Partant du fait que les propriétés spectrales de surface sont liées à des propriétés de sols, le couplage des données de télédétection rapprochée (Spectrométrie Visible et Proche Infrarouge (Vis-PIR)) et de données de télédétection spatiale continues dans l'espace (image satellite), offre une opportunité d'accroître notre capacité d'acquisition et d'extraction des propriétés du sol pour les quantifier d'une façon spatiale (Gomez et Lagacherie, 2007).

En effet, les images sont des ensembles de mesures du rayonnement solaire réfléchi par la surface de la Terre, ordonnées sous une forme matricielle, et couvrant un domaine spectral s'étendant du visible au moyen infrarouge (Escadafal, 1993). La combinaison de la spectroscopie avec la technologie d'acquisition de l'information spatiale sur de grands secteurs, peut se voir comme une imagerie spectroscopique (Homayouni, 2005) qui peut être mise à profit dans la quantification des attributs du sol.

Des techniques hybrides ont été développées pour combiner deux approches conceptuelles de modélisation et de spatialisation des propriétés du sol différentes : (a) l'interpolation se basant seulement sur les données ponctuelles de la variable d'intérêt au niveau des points d'observation ; (b) l'interpolation basée sur la régression de la variable d'intérêt sur des variables auxiliaires à information exhaustive. Ces techniques peuvent offrir une meilleure estimation que les techniques reposant sur une seule approche (Knotters et *al.*, 1995 ; Bishop et Mc Bratney, 2001 ; Bourennane et King, 2003 ; Hengl et *al.*, 2004 ; Lloyd 2005 ; Yemefack et *al.*, 2005). L'une de ces techniques hybrides d'interpolation est la régression-krigeage qui a été initiée par Odeh et *al.* (1994,1995). Elle utilise en premier lieu une régression de la variable d'intérêt sur les informations auxiliaires qu'elle combine avec un krigeage ordinaire des résidus (Hengl et *al.*, 2004). L'avantage de la régression-krigeage est la possibilité d'étendre la méthode à une large gamme de techniques de régression et de permettre une interprétation séparée des deux composantes interpolées (Hengl et *al.*, 2007). Ceci rend la méthode facile à utiliser et très attractive (Carré et Girard, 2002 ; Finke et *al.*, 2004 ; Baxter et Oliver, 2005 ; Herbest et *al.*, 2006 ; Simbahan et *al.*, 2006 ; Douaoui et *al.*, 2006 ; Minasny et Mc Bratney, 2007). Dans plusieurs cas, la régression combinée au krigeage, s'est montrée supérieure à la simple utilisation des techniques géostatistiques, car engendrant des résultats plus précis (Hengl et *al.*, 2007). La régression-krigeage a été

extensivement utilisée, car il s'agit d'une méthode non seulement simple sur le plan conceptuel, mais aussi robuste (Minasny et Mc Bratney, 2007).

Ce travail s'inscrit dans le cadre général de développement de méthodes d'estimation spatiale des propriétés du sol.

Son objectif est de développer une approche de cartographie des propriétés du sol par régression-krigeage en combinant des informations issues d'une image satellite multi-bandes et une collection de spectres d'échantillons de sol acquis en conditions contrôlés de laboratoire. Cette approche est testée à titre d'exemple dans le cas de deux variables déterminantes dans l'évolution pédogénétique des sols en milieu aride à savoir le taux de carbone total et la conductivité électrique du sol. Le secteur d'étude est d'une superficie de 580 ha, comprenant des sols nus, situés dans la région naturelle aride du Djérid (Sud-Ouest tunisien).

2. Matériels et méthodes

La méthode de spatialisation développée et utilisée dans ce travail est fondée sur une régression en Moindres Carrés Partiels-Krigeage (régression MCP-Krigeage). La démarche est synthétisée dans la figure 1, qui résume les différentes étapes de cette approche.

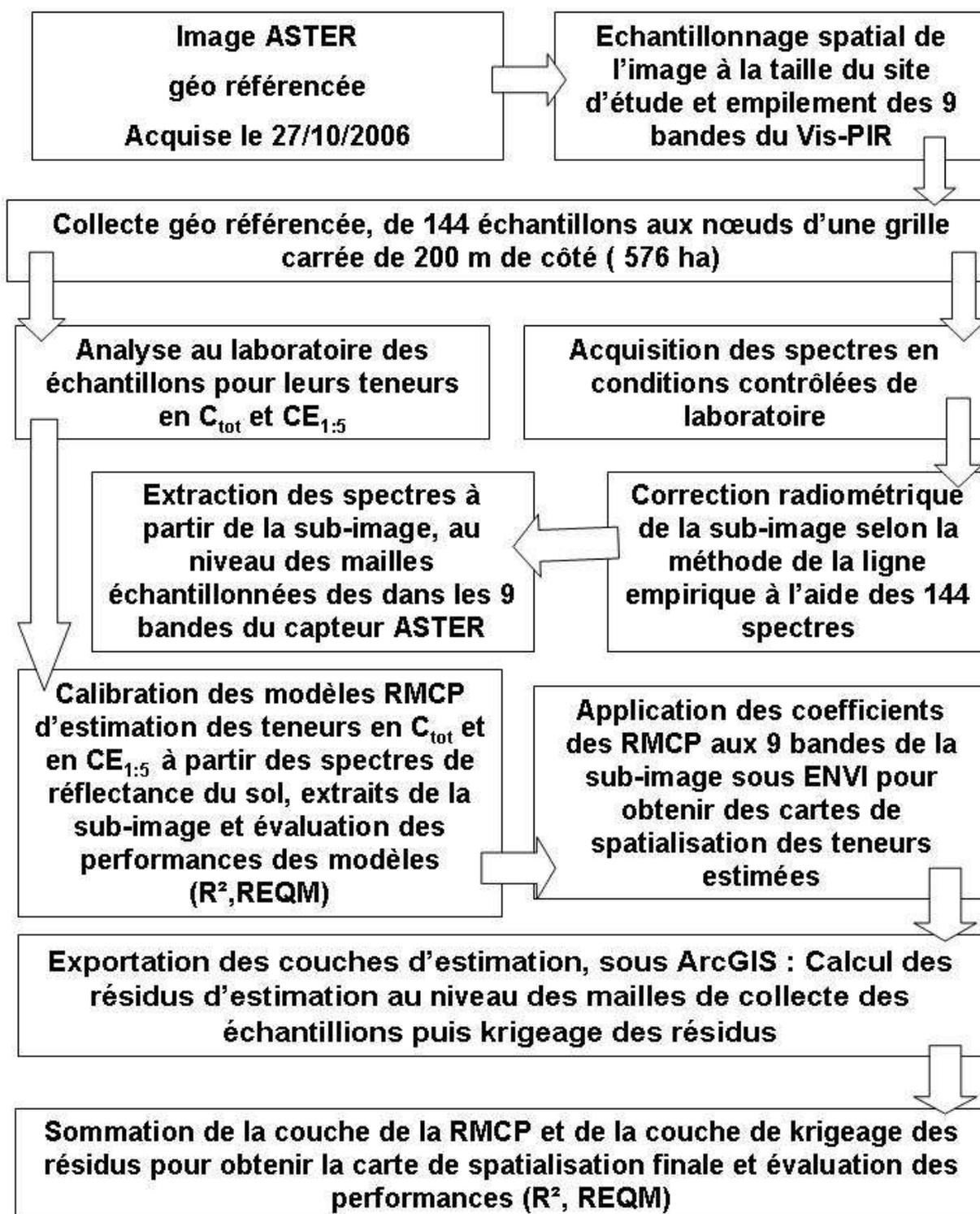


Figure 1. Diagramme de spatialisation des données spectrométriques pour l'obtention de cartes thématiques par modèle de régression MCP-Krigeage.

2.1. Site d'étude

La figure 2 montre la scène de l'image qui a été utilisée dans ce travail. Il s'agit d'une image acquise avec le capteur ASTER le 27 octobre 2006. Elle est visualisée en fausses couleurs RVB 321. La scène couvre la partie septentrionale de la région naturelle du Djérid. Les coordonnées respectivement de son coin haut gauche et de son coin bas droit sont : (34° 31'45.67" N, 7° 42'27.30" E) et (33° 52'48.75" N, 8° 13'6.62" E). Sur le plan structural, la région constitue une zone de transition entre le système atlasique au nord et le début du système saharien ou la plate-forme saharienne au sud (Atlas, 1996). Nous distinguons trois grandes unités topographiques et géomorphologiques :

- les deux chaînes montagneuses de direction Ouest-Est (occupant le nord et le nord est de la scène (figure 2)),
- les piémonts où s'étendent des glacis fossilisés par des croûtes et des encroûtements gypseux,
- les dépressions fermées ou chotts : chott El Gharsa au nord (occupant ici le centre de la scène) et chott Djérid au sud (non couvert par la scène). Leurs bassins versants englobent toute la région et imposent un régime d'écoulement endoréique (Bryant, 1999). La dépression de chott El Gharsa est d'une superficie de 620 km² (Barbieri et *al.*, 2006). Elle s'étend de l'est vers l'ouest et elle est le siège de dépôts de sable d'apport éolien et d'évaporites qui précipitent sur place (Swezey, 1996 ; 2003).

Le climat est aride. Il est caractérisé par une température moyenne annuelle de 22°C, alors que les précipitations moyennes annuelles ne dépassent pas 100 mm. En outre, les vents sont très fréquents dans la région. En effet, nous enregistrons en moyenne 120 j/an de vents forts. Le vent dominant est de direction nord-est, il souffle surtout en automne et au printemps (Atlas, 1996). L'aridité, la faiblesse du couvert végétal, l'action du vent, l'influence des dépressions fermées (halomorphie et redistribution du gypse) ainsi que la proximité de la nappe d'eau de la surface du sol, sont autant de facteurs qui conditionnent l'organisation de la couverture pédologique dans la région, formée pour l'essentiel de sols calcimagnésiques à texture sableuse (El Fkih, 1965).

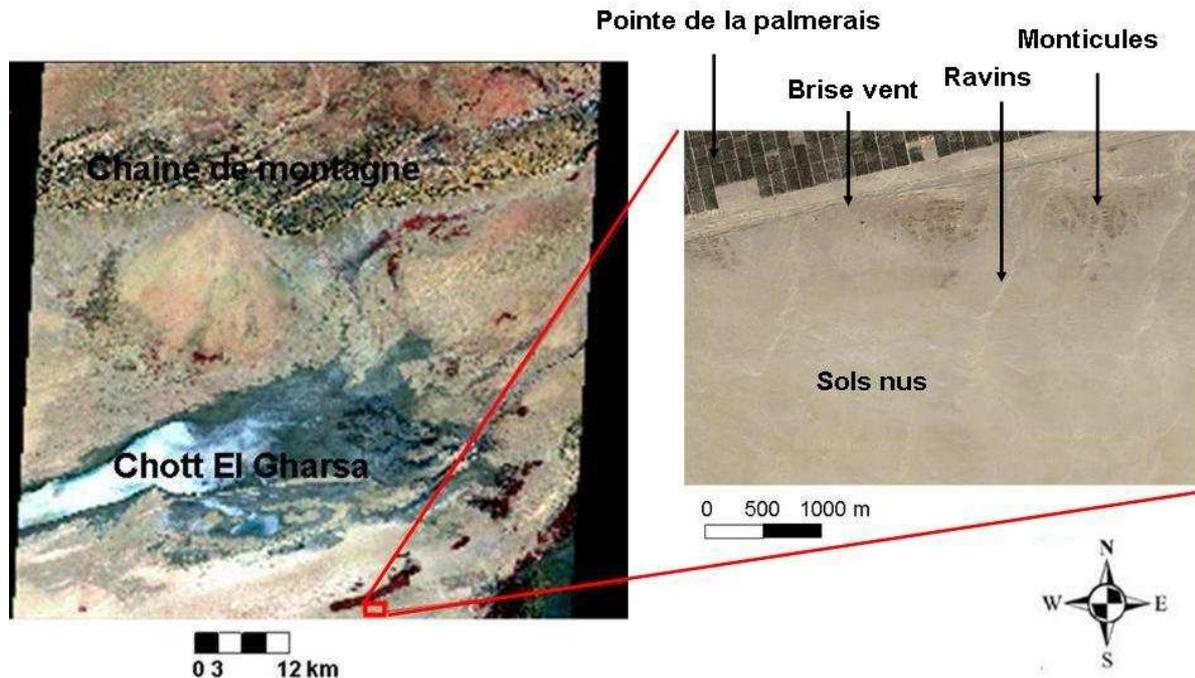


Figure 2. Scène de l'image ASTER couvrant la partie septentrionale de la région du Djérid, visualisée en fausses couleurs RVB 321 avec une résolution de 60X60 km. Image prise le 27-10-2006. Le rectangle rouge délimite le site d'étude expérimental, son aperçu est affiché à l'aide de l'outil Google map de Google Earth.

L'étude a été réalisée dans un site particulier de la région à savoir le site de Chamsa. La sub-image (Portion de l'image initiale couvrant la région d'intérêt pour l'étude), rééchantillonnée spatialement à partir de l'image initiale, couvre 580 ha. Ce site fait partie du bassin versant du chott El Gharsa et se situe au sud de ce chott sur un versant en pente générale douce (< 1 %) mais régulière, de direction Sud/Nord. Il est composé d'un haut de versant, d'un mi versant et d'un bas de versant.

La partie aval de ce site, le haut de versant (partie sud de la sub-image) correspond à un plateau. La couverture pédologique de cette partie est un Lithosol à matériel parental graveleux et sableux reposant sur un encroûtement gypseux Y_m (10 YR 8/2). Son état de surface correspond à un Reg à graviers siliceux arrondis et à cailloux gypso-calcaires.

Le terrain de la partie intermédiaire à mi-versant est vallonné. Il est occupé par des formations de sable siliceux jaune rougeâtre (7,5 YR 7/6) à structure particulière sous forme de Nebkhas : il s'agit de petits amoncellements en buttes buissonnantes de hauteur décimétrique de sables et pseudo sables gypseux d'apport éolien autour des touffes de végétation naturelle (Job, 1992). Les sols sont des sols peu évolués xériques, pour lesquels le

facteur limitant est la sécheresse. Ces sols peu profonds reposent sur une croûte gypso-calcaire. Le terrain de cette partie est raviné, mais le réseau hydrographique est peu dense dénotant la prépondérance de l'érosion éolienne sur l'érosion hydrique. Au niveau des talwegs sans végétation affleure la croûte gypso-calcaire en Lithosols regosoliques (état de surface pierreux dénudé de végétation (Démangeot et Bernus, 2001 ; Baize et Girard, 2009), légèrement recouverte par un film éolien.

Au niveau de la partie nord du site (bas du versant) se dresse une chaîne de monticules à dômes formés de grès tabulaires désagrégés à sables siliceux et carbonatés, reposant sur un horizon d'argile rouge (5YR 6/4). Le sol est constitué de détritiques de grès mélangé à l'argile rouge. Il s'agit de sols peu évolués colluvio-alluvionnaires à caractère salin (carbonaté). Dans la partie Nord-Ouest du site, ces structures monticulaires sont limitées par la pointe de la palmeraie (*Phoenix dactylophera*) dite Ibn Chabbat qui apparaît sur l'image en rouge (en fausse couleur).

2.2. Traitement de l'image satellitale ASTER

Nous disposons d'une image satellite géo-référencée acquise avec le capteur ASTER. Il s'agit d'un système à radiomètre imageur multibandes installé sur la plate-forme spatiale Terra en 1999 (Ramsey et al., 1999). Le capteur ASTER comporte 14 bandes ayant une étendue spectrale qui va de 0,52 μm à 11,65 μm (Luleva, 2007). Il est muni de trois radiomètres séparés (Abrams, 1997), couvrant respectivement les domaines spectraux du visible, du proche infrarouge, du moyen infrarouge et du thermique (tableau 1). Les images ASTER se caractérisent par une haute résolution spectrale dans le domaine du proche infrarouge (6 bandes) et par une haute résolution spatiale dans le visible (15 m). Le capteur ASTER est caractérisé par une répétitivité de 16 jours et fournit une scène de l'ordre de 60 x 60 km². Il parcourt la même orbite que le capteur Landsat, mais 7h 30 minutes plus tard (Chikhaoui et Bonn, 2004).

Tableau 1. Caractéristiques du capteur ASTER (d'après Yamaguchi et *al.*, 1998).

Caractéristiques	Visible et Proche infrarouge	Proche infrarouge	Thermique
Bandes (nm)	B1 : 520-600 (Nadir)	B4 : 1600-1700	B10 : 8120-8470
	B2 : 630-690 (Nadir)	B5 : 2140-2180	B11 : 8470-8820
	B3 : 760-860 (Nadir)	B6 : 2180-2220	B12 : 8920-9270
	B3B : 760-860	B7 : 2230-2280	B13 : 1025-1095
	(± incliné de 24°)	B8 : 2290-2360	B14 : 1095-1165
		B9 : 2360-2430	
Résolution spatiale en (m)	15	30	90

L'image ASTER utilisée a été commandée auprès du Centre de données EROS (Earth Observing System Data Gateway). L'image a été corrigée par le Centre EROS des effets atmosphériques et des déformations géométriques liées au capteur (El Quaedy et *al.*, 2006). Elle est déjà géoréférencée selon la projection UTM (Universal Transverse Mercator), Zone 32N. L'image a été acquise le 27 octobre 2006. Son identifiant est : Granule_Id = "ASTL1B 0610271018220611290709". Elle couvre la partie septentrionale de la région du Djérid. Les traitements d'image ont été réalisés à l'aide du logiciel ENVI[®]4.3. Les cinq bandes du thermique, sans intérêt pour l'étude, ont été de prime abord supprimées. Ensuite, l'image a été re-échantillonnée spatialement à la taille de notre site d'étude produisant une sub-image qui a été utilisée dans les traitements ultérieurs. Pour être superposables aux trois bandes du visible, la taille des pixels des 6 bandes du proche infrarouge a été re-échantillonnée de 30 à 15 m en utilisant la méthode du plus proche voisin.

2.3. Prélèvement des échantillons sur le terrain

La campagne de collecte des échantillons sur le terrain s'est déroulée au cours de la première quinzaine du mois de novembre 2007 en période sèche. Nous avons effectué un échantillonnage systématique aux nœuds d'une grille régulière, ayant une maille de 200 x 200 m², soit un échantillon tous les 4 ha.

Nous nous sommes munis d'un récepteur de Géolocalisation Par Satellite (GPS) Trimble Geo XM[®] d'une précision planimétrique de 5 m pour nous rendre aux points préfixés d'échantillonnage. Un échantillon composite formé de 8 sous-échantillons de sol a été à chaque fois prélevé (Jarmer et *al.*, 2008) à la surface du sol (10 premiers centimètres du sol) au niveau de placettes d'un rayon de 8 m. Nous avons collecté 144 échantillons, couvrant ainsi une superficie de 576 ha. La figure 3 montre le schéma d'échantillonnage, sur un fond

d'image satellite extraite à l'aide de l'outil Google map de Google Earth, qui a été respecté sur le terrain.

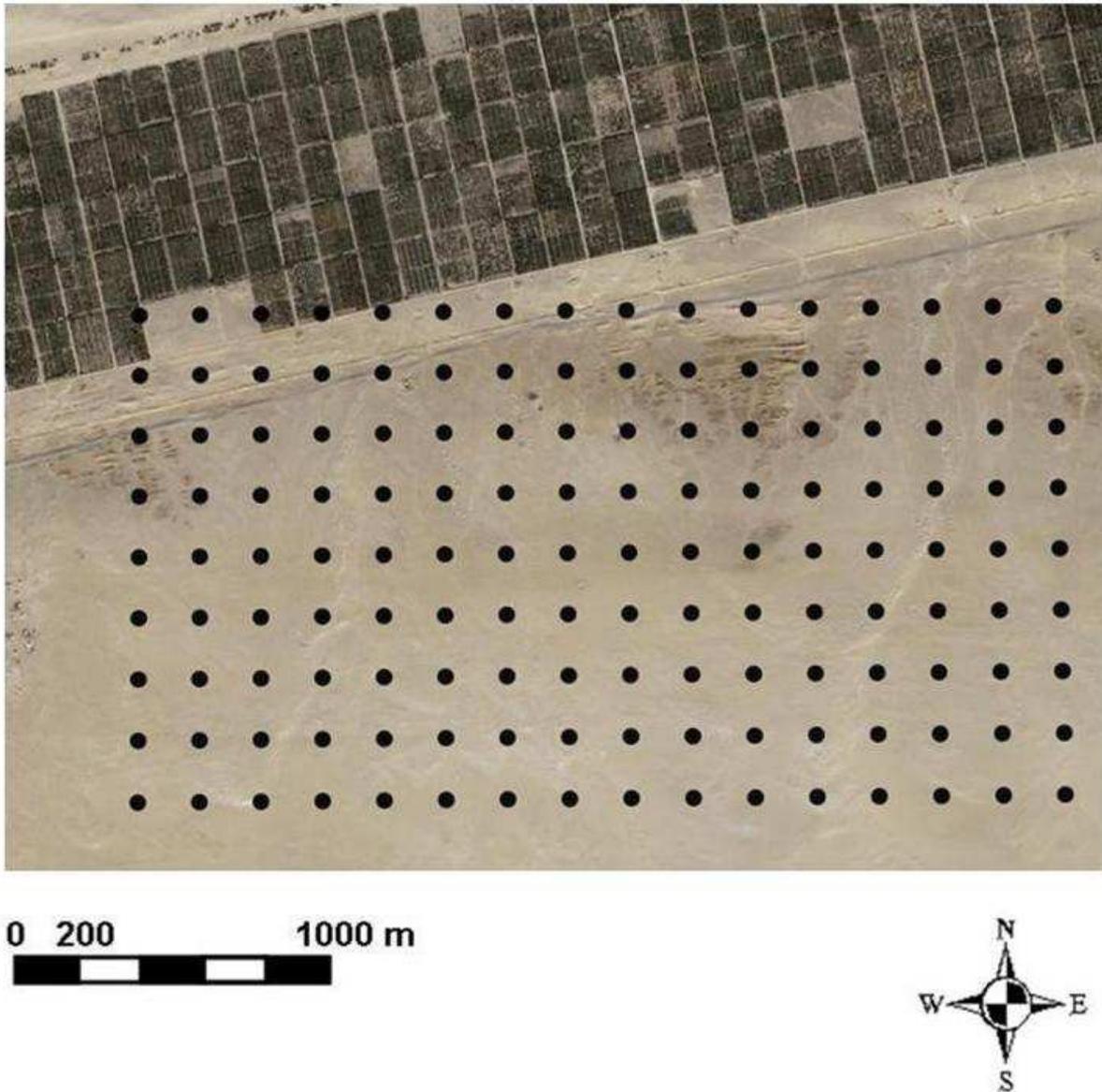


Figure 3. Schéma d'échantillonnage systématique aux nœuds d'une grille carrée régulière de 200 m de côté. Points échantillonnés, visualisés sur un fond d'image satellite extraite à l'aide de l'outil Google maps de Google Earth.

2.4. Analyses au laboratoire

Les échantillons ont été emmenés au laboratoire, séchés à l'air libre, tamisés à 2 mm et broyés, pour être analysés pour leurs teneurs en C_{tot} par combustion à l'aide d'un analyseur élémentaire HCN (Norme ISO 10694) et pour leurs conductivités électriques ($CE_{1:5}$) mesurées dans un extrait aqueux dilué au 1/5^{ème} (USSL, 1954).

2.5. Acquisition des spectres

L'acquisition des spectres des 144 échantillons collectés s'est déroulée dans les conditions contrôlées de laboratoire. Les mesures ont été faites avec un spectroradiomètre portable ASD FieldSpec Pro Fr. Ce spectroradiomètre est doté de trois détecteurs qui opèrent dans le visible et le proche infrarouge soit entre 350 nm et 2500 nm. Ce dernier permet un balayage spectral continu avec un pas d'échantillonnage de 10 nm, entre les longueurs d'onde de 350 et 1000 nm, et de 30 nm, entre les longueurs d'onde de 1000 et 2500 nm. Le système rééchantillonne les spectres à un pas de 1 nm, ce qui permet d'obtenir un spectre avec 2151 bandes (El Quaedy et *al.*, 2006). La fenêtre de l'UV, sans intérêt pour l'étude, a été tronquée du spectre.

2.6. Correction radiométrique de la sub-image

La correction radiométrique de la sub-image a été réalisée avec le logiciel de traitement d'images ENVI®4.3 (ITT, 2007). Les spectres de réflectance des 144 échantillons géoréférencés ont été utilisés pour corriger radiométriquement la sub-image ASTER. Pour cela, nous avons utilisé la méthode de la ligne empirique (Chabrillat et *al.*, 2002 ; Vaudour et Berthier, 2007). Cette méthode consiste à établir une régression linéaire multiple reliant les 144 spectres mesurés au laboratoire et les 144 valeurs des comptes numériques (luminances) des pixels de la sub-image ASTER qui leurs sont appariées. L'équation suivante montre comment le gain (pente de la droite) et le décalage (en anglais « offset » : ordonné à l'origine (Bannari et *al.*, 2007)) de la ligne empirique sont calculés pour une bande donnée :

Réflectance (spectre au labo) = gain * luminance (données de l'image ASTER) + offset (Di et Rundquist, 1994). [1]

La correction de la sub-image par la méthode de la ligne empirique permet en outre de s'affranchir de l'effet des perturbations atmosphériques (Xu et Hang, 2008) et de l'effet du bruit du capteur embarqué. Par ailleurs, l'effet atmosphérique, s'il n'est pas soustrait, peut diminuer ou augmenter la réflectance des objets au sol ce qui introduit un biais dans la quantification des propriétés du sol (Ben Dor et Levin, 2000).

2.7. Extraction des spectres de réflectance au sol à partir de la sub-image corrigée en réflectance relative

Une fois que la sub-image ait été radiométriquement corrigée selon la méthode de la ligne empirique, nous avons procédé, sous le logiciel ENVI®4.3, à l'extraction des spectres simulés dans les neuf bandes du capteur ASTER et ce au niveau des pixels de la sub-image qui correspondaient aux placettes géoréférencées, lieux de prélèvement des échantillons sur le terrain. En effet, chaque pixel fournit le spectre complet d'une surface élémentaire dans les différentes bandes du capteur (Stevens et al., 2006). Nous avons extrait de la sub-image multi-spectrale 144 spectres correspondants à la réflectance relative au sol en 9 bandes, des 144 placettes échantillonnées sur le terrain.

2.8. Spatialisation des données par régression MCP-krigeage

2.8.1. Première étape : la régression MCP

Les 144 spectres extraits de la sub-image ASTER radiométriquement corrigée en réflectance relative ont été mis en relation respectivement avec les teneurs mesurées de C_{tot} et de $CE_{1:5}$ en utilisant le modèle statistique de Régression en Moindres Carrés Partiel MCP (*Partial Least Squares Regression, PLSR*, en anglais). En effet, d'après Hill et Shùtt (2000), pour utiliser au mieux les images satellites en cartographie, il est indispensable de disposer de modèles qui convertissent l'information spectrale contenue dans les pixels de l'image en une information thématique. Les coefficients du modèle propre à chaque propriété pédologique ont été appliqués sous ENVI®4.3 aux 9 bandes du Vis-PIR de la sub-image (Figure 1).

Pour chacune de ces deux propriétés, le résultat de la spatialisation est une couche raster de distribution des teneurs de la propriété objet de l'étude. Ces deux couches d'information ont été exportées sous le logiciel de système d'information géographique ArcGis®9.2 (ESRI, 2006). Les couches visualisées sur une échelle de compte numérique de 0 à 255 ont été multipliées par les termes $y_{max}/255$ respectivement de C_{tot} et de CE , spatialement estimées, pour avoir une lecture directe des valeurs des teneurs estimées sous le logiciel ArcGIS®9.2.

2.8.2. Deuxième étape : le krigeage des résidus

L'analyse de la variabilité spatiale de l'erreur a été entreprise à l'aide du logiciel ArcGIS® 9.2 et plus particulièrement du module *geostatistical analyst*. Ce dernier permet d'explorer la variabilité des données, d'étudier la variographie et de spatialiser par krigeage.

Pour chacune des deux variables, à l'aide de l'outil de statistiques zonales du module *spatial analyst* d'ArcGIS® 9.2, nous avons extrait les valeurs estimées du C_{tot} et de $CE_{1:5}$ au niveau des 144 pixels échantillonnés. Par la suite, afin de déterminer les résidus, nous avons calculé la différence entre la valeur de la variable mesurée par analyse au laboratoire et la valeur de la variable estimée par régression MCP. La formule appliquée pour calculer les résidus est la suivante :

$$\xi_i = y_i - y_i^* \quad [2]$$

où ξ_i est le résidu d'estimation pour le $i^{ème}$ échantillon, y_i est la valeur mesurée pour le $i^{ème}$ échantillon et y_i^* est la valeur estimée par le modèle pour le $i^{ème}$ échantillon.

Une première évaluation des performances d'estimation a été faite en calculant pour chaque variable le coefficient de détermination R^2 , le biais et la racine de l'erreur quadratique moyenne REQM entre les valeurs mesurées au laboratoire et les valeurs estimées par régression MCP. Par ailleurs, nous avons aussi calculé la racine de l'erreur quadratique moyenne relative REQMR (Selige, 2006, Farifteh et al., 2007) définie comme le rapport entre la racine de l'erreur quadratique moyenne et la moyenne de la série des valeurs de la propriété d'intérêt mesurée au laboratoire, ainsi que le ratio entre l'étendue et la REQM (Wetterlind et al., 2008). Ces critères sont définis selon les équations 3 à 7 :

$$R^2 = \left[\frac{\sum_{i=1}^n (y_i - \bar{Y})(y_i^* - \bar{Y}^*)}{\sqrt{\sum_{i=1}^n (y_i - \bar{Y})^2 \sum_{i=1}^n (y_i^* - \bar{Y}^*)^2}} \right]^2 \quad [3]$$

$$Biais = \frac{1}{n} \sum_{i=1}^n (y_i^* - y_i) \quad [4]$$

$$REQM = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i^* - y_i)^2} \quad [5]$$

$$REQMR = 100. \frac{REQM}{\bar{Y}} \quad [6]$$

$$Ratio = Etendue / REQM \quad [7]$$

Où : y_i est la valeur mesurée, y_i^* est la valeur estimée, \bar{Y} est la moyenne des valeurs mesurées et \bar{Y}^* est la moyenne des valeurs prédites.

L'interpolation par krigeage ordinaire des résidus s'est faite après le calcul des paramètres des variogrammes expérimentaux. Les couches de krigeage ont été rasterisées à la taille d'extension de la sub-image avec une résolution pixellaire de 15 m pour être à la même taille que les couches de spatialisation des équations des régressions MCP propres à chaque variable (Figure 1).

2.8.3. Production des cartes thématiques

Pour chacune des deux variables, les deux couches rasters, la première issue de la spatialisation de l'équation de la régression MCP et la seconde issue de l'interpolation par krigeage des résidus ont été additionnées pour produire la carte d'estimation de la propriété d'intérêt par régression MCP-krigeage (Knotters et *al.*, 1995 ; Bishop et Mc Bratney, 2001 ; Carré, 2002 ; Douaoui, 2005).

Afin d'apprécier les niveaux de précision des cartes produites, les valeurs d'estimation contenues aux niveaux des pixels des points d'échantillonnage ont été de nouveau extraites et les erreurs résiduelles entre les valeurs mesurées au laboratoire et les valeurs estimées par régression MCP-krigeage ont été calculées. Une deuxième évaluation des performances d'estimation a été faite en calculant pour chaque variable les critères statistiques cités plus haut.

3. Résultats et discussion

3.1. Statistiques descriptives des données analytiques

Le tableau 2 présente les statistiques descriptives des deux propriétés analytiques d'intérêt, en l'occurrence le taux de carbone total (C_{tot}) et la conductivité électrique, après leur analyse au laboratoire.

Tableau 2. Statistiques descriptives des deux propriétés analytiques des 144 échantillons de sols collectés.

	Min	Max	Etendue	Moyenne	Ecart-type
C_{tot} (%)	0.12	2.38	2.26	0.70	0.34
$CE_{1:5}$ (dS/m)	0.07	4.93	4.86	0.90	0.86

Le carbone total (C_{tot}) varie dans une étendue comprise entre 0.12 et 2.38 %, avec une moyenne de 0.7 % et un écart-type de 0.34 %. Ce carbone est pour l'essentiel présent dans le sol sous forme minérale. En effet, dans ce site la végétation, formée d'herbacées à faible masse végétale, est très éparse. Elle couvre moins de 1 % de la surface du sol. Par ailleurs, l'analyse des teneurs en carbone minéral effectuée pour 31 échantillons collectés dans le même site de Chamsa (Cf. chapitre 4), mais sur une superficie plus large de 2800 ha contre 580 ha pour ces 144 échantillons, a montré que le carbone minéral représente en moyenne 96.5 % du carbone total. Ces valeurs élevées pour un horizon de surface, s'expliquent par la nature carbonatée du matériau originel sous-jacent, désagrégé par l'effet conjugué du vent (abrasion) et de l'amplitude thermique entre le jour et la nuit, à l'origine d'une thermoclastie (Coudé-Gaussen, 1990). Une néoformation de carbonates en surface en raison de l'aridité du milieu peut également être mise en avant (Drake, 1997 ; Bryant, 1999).

La mesure de la CE dans un extrait aqueux de rapport sol/eau égal à 1/5, permet de classer les sols selon leur degré de salinité (tableau 3).

Tableau 3. Classification des sols selon leurs CE_{1:5} (d'après : Wang et al., 2008 ; Gherina et al., 2009)

CE _{1:5} (ds/m)	
0-1	Sol non salé
1-2	Sol peu salé
2-4	Sol modérément salé
4-8	Sol salé
8-16	Sol fortement salé
>16	Sol très fortement salé

Dans notre cas, la valeur minimale mesurée est de 0.07 dS/m alors que la valeur maximale est de 4.93 dS/m. Nous balayons donc les classes des sols non salés jusqu'à celle des sols salés. Par ailleurs, la moyenne est de 0.9 dS/m et l'écart type est de 0.86 dS/m. La valeur élevée de la moyenne, dans le cas de sols non mis en irrigation, témoigne d'une salinité primaire élevée due à la présence de sels : Na⁺, K⁺, Mg²⁺, Ca²⁺, Cl⁻, CO₃⁻ et SO₄⁼ (Bryant, 1999). Cette salinité d'origine géologique et éolienne peut s'accroître si le sol est irrigué d'autant plus que les ressources en eau qui peuvent être mobilisées pour l'irrigation sont également chargées en sels (Mamou, 1976 ; Kamel, 2007).

La figure 4 indique en cercles pleins les endroits qui ont été échantillonnés sur le terrain, selon une grille régulière de 200 x 200 m². D'après Hengl et al. (2007) et Minasny et Mc Bratney, (2007) la technique de la régression-krigeage est étroitement tributaire de la qualité des données d'entrée. Si les données ont été échantillonnées en utilisant un schéma d'échantillonnage biaisé ou non représentatif, les estimations peuvent être de mauvaise qualité. Nous avons présenté ces cercles à des tailles qui varient en fonction des grandeurs des valeurs mesurées. Les figures 4a et 4b montrent successivement comment les valeurs mesurées de C_{tot} et de CE_{1:5} varient d'un point d'échantillonnage à un autre. Pour ces deux propriétés, nous remarquons que les valeurs mesurées présentent un gradient en se déplaçant du sud vers le nord, c'est-à-dire en se déplaçant vers le chott (Figure 2). D'après cette figure, nous pouvons noter l'existence d'une corrélation spatiale entre les deux variables C_{tot} et CE_{1:5}. En effet, la concentration en ions calcium Ca²⁺ et bicarbonate CO₃⁼ contenus dans les calcites CaCO₃ du sol, a un effet direct sur la CE du sol.

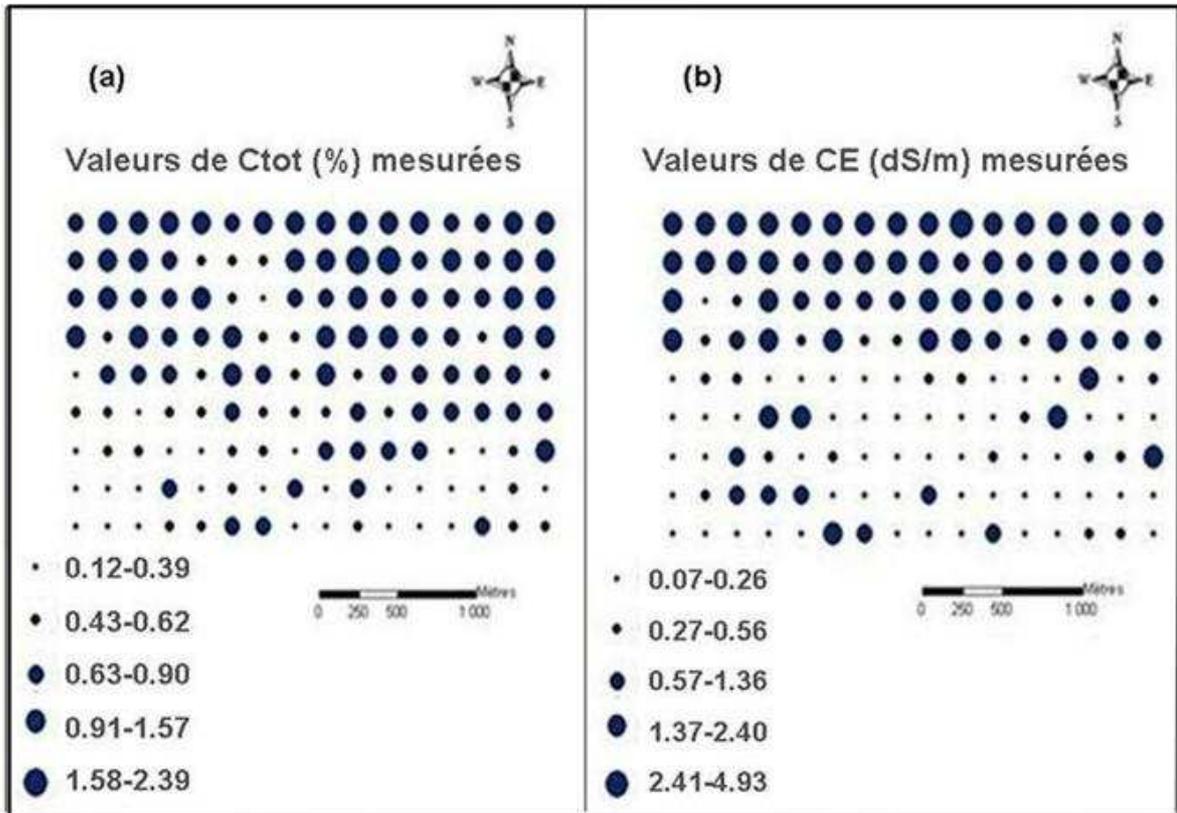


Figure 4. Représentation des grandeurs des valeurs mesurées de C_{tot} (a) et de $CE_{1:5}$ (b) au niveau des points échantillonnés sur le terrain, avec des cercles dont les tailles varient proportionnellement aux valeurs mesurées.

3.2. Analyse des spectres des échantillons de sol

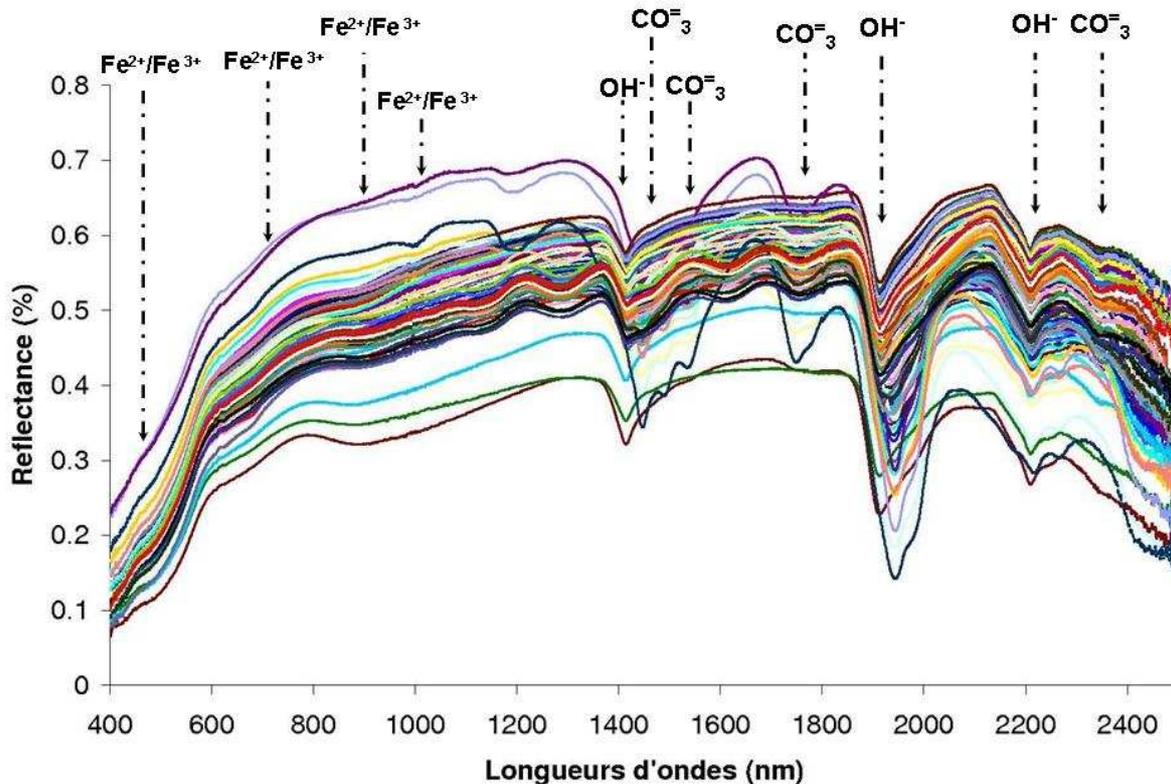


Figure 5a. Spectres des 144 échantillons collectés sur le terrain, acquis en conditions contrôlées de laboratoire, après séchage et tamisage à 2 mm des échantillons. Des pics d'absorption, liées à la présence de certains composés, sont identifiés à partir de données de la littérature.

La figure 5a montre la réponse spectrale des échantillons du sol dans la gamme 400-2500 nm avec une résolution spectrale de 1 nm. Ces spectres sont continus et croissants sur tout le domaine du visible (400-700 nm). Les changements de pente des spectres du sol, qui se situent aux voisinages de 440, 550, 630 et 860 nm, correspondent aux bandes d'absorption du fer (Richter et *al.*, 2009). Dans le proche infrarouge, le spectre présente différentes évolutions passant par une croissance de pente faible, suivi d'un palier puis d'une décroissance. Par ailleurs, nous notons la présence de bandes d'absorption, plus ou moins larges en fonction des échantillons, aux voisinages de 1400 et 1900 nm, qui sont causés par le groupement OH de l'eau libre ou intégrée dans les réseaux cristallins des sels hygroscopiques (Mougenot, 1993, Dematté et *al.*, 2004). Une autre bande d'absorption au voisinage de 2200 nm est reconnue comme liée à l'absorption de l'ion hydroxyle contenu dans les minéraux argileux

particulièrement la kaolinite (Demattê et *al.*, 2004, Lagacherie et *al.*, 2008). Enfin, une série de bandes d'absorptions situées entre 1600 et 2500 nm sont attribuables à l'ion bicarbonate CO_3^- (Gaffey, 1986). Les valeurs relativement élevées de la réflectance témoignent de la nature quartzique des sables (Girard et Girard, 1999 ; Barnes, 2003) qui forment l'essentiel de la couverture pédologique dans le site d'étude.

3.3. Spectres de luminances extraits de la sub-image

La figure 5b montre les spectres de radiances extraits de la sub-image ASTER avant sa correction radiométrique. Avant cette correction radiométrique, l'information de chaque pixel des neuf bandes empilées de la sub-image est exprimée en compte numérique sur une échelle de 0 à 255.

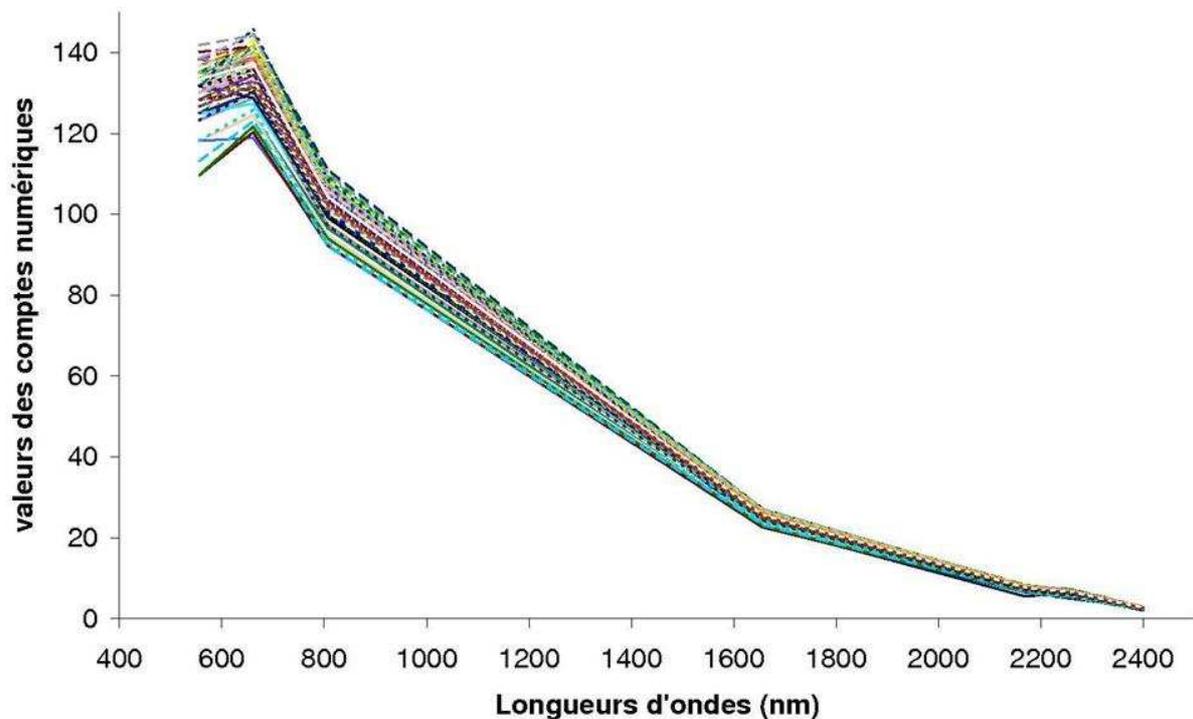


Figure 5b. Spectres de luminance extraits des 144 pixels de la sub-image, correspondant aux nœuds des mailles échantillonnés sur le terrain avant sa correction radiométrique.

3.4. Spectres extraits de la sub-image, après sa correction radiométrique

La figure 5c montre les spectres extraits de la sub-image après sa correction radiométrique. L'information est convertie en réflectance relative sur une échelle continue de 0 à 1.

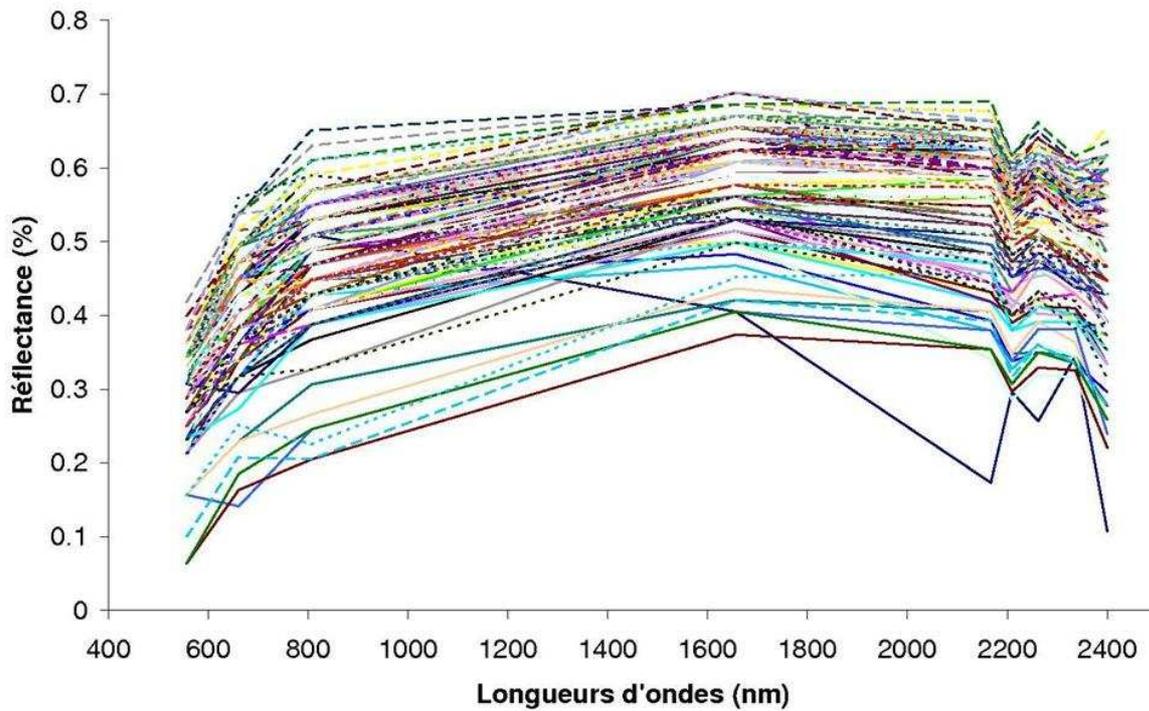


Figure 5c. Spectres extraits de la sub-image, après sa correction radiométrique, aux niveaux des des 144 nœuds de mailles échantillonnés sur le terrain.

Ces spectres, qui constituent une information dégradée limitée à neuf bandes, présentent une allure générale très similaire à celle des spectres acquis au laboratoire.

Les spectres extraits de la sub-image, après sa correction en réflectance relative, ont été utilisés dans l'établissement des équations de régressions MCP successivement pour la variable C_{tot} et pour la variable $CE_{1.5}$. La correction de l'image par la méthode de la ligne empirique en utilisant des spectres d'échantillons de sols permet de disposer pour tous les pixels de l'image d'une réflectance relative qui présente l'avantage de ne plus contenir qu'une signature caractéristique du matériau objet de l'étude (Méléder et *al.*, 2003), à savoir ici le sol.

3.5. Performances d'estimation spatiale par régression MCP

Le tableau 4 récapitule les résultats des performances d'estimation spatiale par régression MCP

Tableau 4. Résultats des performances d'estimation spatiale par régression MCP, appliquée aux 144 spectres en 9 bandes extraits de l'image ASTER après sa correction en réflectance relative.

		R ²	Biais	REQM	REQMR	Etendu/REQM
Régression PLS	C _{tot} (%)	0.53	0.27 %	0.52%	74.77 %	4.29
	CE (dS/m)	0.57	0.36 dS/m	0.75 dS/m	82.98 %	6.44

Dans le cas de la variable C_{tot}, l'équation de la régression MCP obtenue est :

$$C_{tot}^* = -1.24 R(b1) + 0.28 R(b2) - 0.59 R(b3) - 1.29 R(b4) + 0.34 R(b5) - 0.57 R(b6) - 0.22 R(b7) - 0.47 R(b8) - 0.06 R(b9) + 2.49. [8] ; R^2 = 0.53 \text{ et } REQM = 0.52 \text{ \%}.$$

Nous constatons que les 9 bandes spectrales contribuent dans l'estimation des teneurs en C_{tot} dans l'ordre d'importance décroissant suivant : b4 (1600-1700 nm), b1 (520-690 nm), b3 (760-860 nm), b6 (2180-2220 nm), b8 (2290-2360 nm), b5 (2140-2180 nm), b2 (630-690 nm), b7 (2230-2280 nm) et b9 (2360-2430 nm).

Dans ses travaux sur des minéraux carbonatés (calcite, dolomite et aragonite), Gaffey, (1986) affirme que ces minéraux présentent une série de bandes d'absorption dans la région 1600-2550 nm, dues à la vibration de l'ion carbonate (CO₃⁼). De même, lorsque ces minéraux carbonatés contiennent des inclusions de fer dans leurs compositions chimiques, ils présentent en sus des bandes d'absorption dans le visible. Ceci concorde avec nos résultats puisque les bandes les plus prépondérantes dans l'estimation sont : la bande 4 qui couvre la fenêtre spectrale 1600-1700 nm, suivie des bandes 1 et 3 qui couvrent respectivement les fenêtres spectrales 520-690 nm et 760-860 nm. Ces deux fenêtres sont identifiées comme des bandes d'absorption, par transfert d'énergie, pour le fer (Hunt et al., 1971 ; Morris et al., 1985 ; Richter et al., 2009). En effet, en milieu aride, le fer qui se libère des minéraux primaires s'oxyde et s'associe à la matrice minérale le plus souvent sous forme d'un revêtement pigmentant qui cimente les particules du sol. Ce phénomène pédogénétique est connu sous le nom de « rubéfaction » (Ben Dor et al., 2006, Richter et al., 2009). Rappelons que, pour notre site d'étude, les trois parties, amont, intermédiaire et avale, étaient respectivement de couleur : 10YR 8/2 (jaunâtre), 7.5YR (jaune rougeâtre) et 5YR 6/4 (rougeâtre) ce qui reflète

l'omniprésence des hydroxydes et oxydes de fer (goethite : $\text{FeO}(\text{OH})$ (jaune) et hématite Fe_2O_3 (rouge)).

La figure 6a montre, dans le cas de la variable C_{tot} , les valeurs estimées vs mesurées, lorsque nous estimons les valeurs de carbone total en appliquant l'équation du modèle de régression MCP développé pour cette variable.

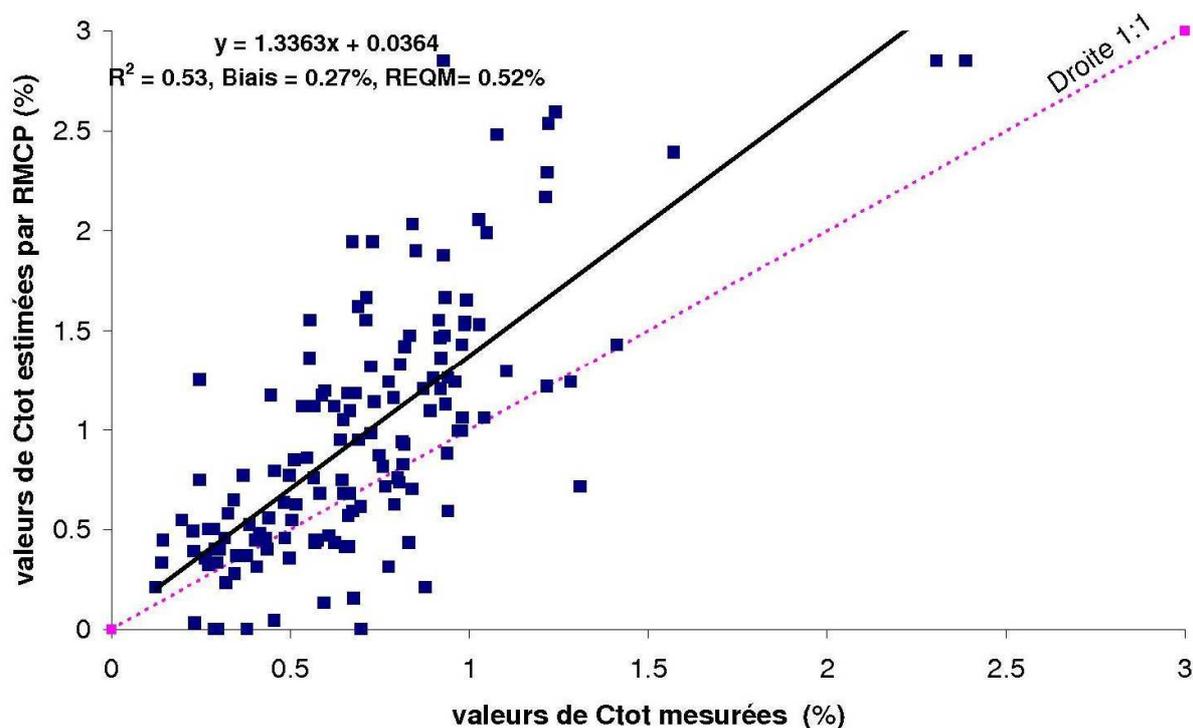


Figure 6a. Valeurs de C_{tot} (%) estimées par régression MCP, appliquées aux spectres en 9 bandes extraits de la sub-image ASTER après sa correction en réflectance relative, vs valeurs mesurées au laboratoire.

En principe, le biais doit être proche de zéro pour une estimation non biaisée. De même pour une estimation précise, la REQM doit être faible (Bourennane, 2000). Dans notre cas, nous constatons que ce modèle présente un coefficient de détermination $R^2 = 0.53$ et une racine de l'erreur quadratique moyenne REQM de 0.52 % et un biais positif de 0.27 %. Nous constatons que la droite ajustée au nuage des points se situe au dessus de la première bissectrice ce qui signifie que le modèle surestime les valeurs mesurées. Cette surestimation pourrait s'expliquer par la présence de constituants du sol autres que les carbonates dont les vibrations occasionnent une augmentation de l'énergie absorbée. En effet, Hunt et Salisbury (1976) ont démontré que les sols carbonatés présentaient plusieurs bandes d'absorption à

2250, 2350, 2380 et 2465 nm dues à la vibration interne de l'ion bicarbonaté CO_3^- , alors que les sols riches en gypse présentaient une large bande d'absorption due aux molécules d'eau dans la région 1330-1650 nm. Dans le cas d'espèce, les sols sont calcimagnésiques avec une présence à la fois des ions cristallins CO_3^- et SO_4^- , ce qui expliquerait la surestimation du modèle.

Dans le cas de la variable $\text{CE}_{1.5}$, l'équation de la régression obtenue est :

$$\text{CE}^* = 0.14 \text{ R}(b1) + 1.48 \text{ R}(b2) - 1.08 \text{ R}(b3) - 3.37 \text{ R}(b4) + 1.67 \text{ R}(b5) - 3.4 \text{ R}(b6) - 9.25 \text{ R}(b7) - 0.63 \text{ R}(b8) + 0.59 \text{ R}(b9) + 4.55. [9] ; R^2 = 0.57 \text{ et } \text{REQM} = 0.56 \text{ dS/m.}$$

Nous constatons que les 9 bandes spectrales contribuent dans l'estimation de la CE dans l'ordre d'importance décroissant suivant : b7 (2230-2280 nm), b6 (2180-2220 nm), b4 (1600-1700 nm), b5 (2140-2180 nm), b2 (630-690 nm), b3 (760-860 nm), b8 (2290-2360 nm), b9 (2360-2430 nm) et b1 (520-690 nm). Les sols salés ont souvent une forte concentration en Na^+ , Mg^{2+} , Ca^{2+} , Cl^- et / ou SO_4^- en solution du sol (Howari et *al.*, 2002). Dans notre région d'étude, Bryant (1999) a démontré l'existence de l'ensemble de ces ions en solution. La précipitation de ces ions entraîne la formation de divers sels qui influencent la réponse spectrale. Il s'agit notamment : de l'halite (NaCl), du gypse ($\text{CaSO}_4 \cdot 2\text{H}_2\text{O}$), du sulfate de sodium (Na_2SO_4) et des sels carbonatés tels que le bicarbonate de sodium (NaHCO_3), la calcite CaCO_3 et la dolomite ($\text{CaMg}(\text{CO}_3)_2$). Pour les sels peu solubles susceptibles de se former, l'ordre de précipitation en fonction de leurs produits de solubilités, est le suivant : la calcite (CaCO_3), la magnésite (MgCO_3) puis le gypse (CaSO_4) (Braudeau et Hachicha, 1998). Par ailleurs, Howari et *al.* (2002) ont démontré que les croûtes de sels de l'halite, du gypse, du bicarbonate de sodium et de sulfate de sodium sont spectralement actives entre 1000 et 2336 nm. De même, Weng et *al.* (2008) ont montré que des sels comme NaCl et MgCl_2 influencent les spectres des sols dans les bandes spectrales 1931-2123 nm et 2153-2253 nm. Comme signalé précédemment, les minéraux carbonatés comme la calcite et la dolomite sont actives dans la bande spectrale 1600-2550 nm (Gaffey, 1986). Ceci concorde avec nos résultats puisque les bandes spectrales les plus déterminantes dans l'estimation de la CE sont respectivement les bandes 7, 6 et 4 qui intéressent la fenêtre spectrale entre 1600 et 2280 nm.

La figure 6b montre, dans le cas de la variable $\text{CE}_{1.5}$, les valeurs estimées vs mesurées, lorsque nous estimons les valeurs de $\text{CE}_{1.5}$ en appliquant l'équation du modèle de régression MCP développé pour cette variable. Ce modèle présente un coefficient de détermination $R^2 =$

0.57 et une racine de l'erreur quadratique moyenne REQM de 0.75 dS/m, le biais étant de 0.36 dS/m.

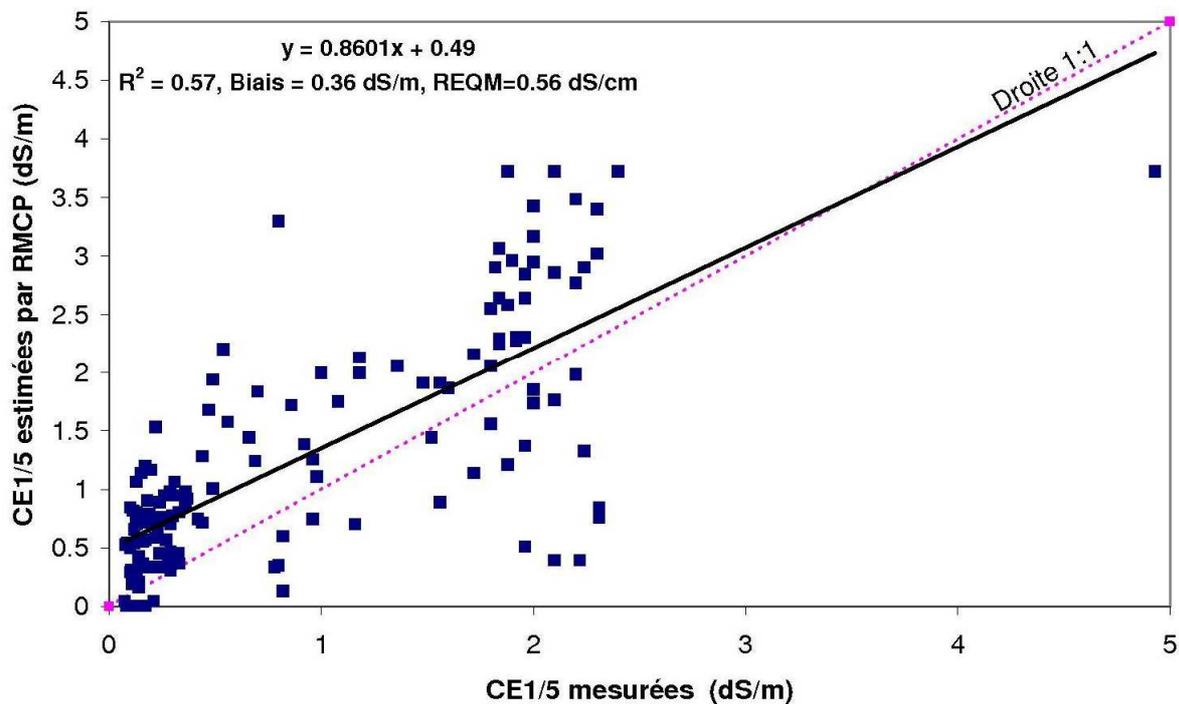


Figure 6b. Valeurs estimées par régression MCP vs valeurs mesurées au laboratoire dans le cas de la variable CE_{1:5} (dS/m).

Nous notons que la qualité d'estimation est moyenne. Ceci peut s'expliquer par le fait suivant : dans le sol, les sels peuvent se trouver sous forme dissoute dans la solution du sol ou sous forme cristallisée lorsque la solution est saturée (Servant, 1975). Or, la CE_{1:5} est une mesure de la concentration des divers sels dissous. Dans le cas de notre zone d'étude, Bryant (1999) a identifié les sels solubles suivants : Na⁺, K⁺, Mg²⁺, Ca²⁺, Cl⁻ et SO₄⁼ en solution du sol. Les sels calciques : carbonatés et sulfatés sont des sels faibles (peu solubles). Ils influencent la réponse spectrale mais ils se dissolvent dans la solution du sol d'une façon partielle. Ils ne sont pas totalement pris en considération dans la mesure de la conductivité électrique dans un extrait au 1/5^{ème}, d'où une surestimation potentielle du modèle. Par ailleurs, Barnes *et al.* (2003) affirment que la croûte blanche qui se forme en surface des sols affectés par les sels modifie leur comportement spectral. La réflectance devient élevée dans le spectre du visible et du proche infrarouge (Barnes *et al.*, 2003). Cependant, les sols riches en sables quartziques peuvent aussi avoir une réflectance élevée dans le visible et le proche infrarouge. Ce fait rend difficile, en milieu aride à sols sableux comme dans notre cas d'étude, la différenciation des sols salés, des sols quartziques non salés.

La figure 7a montre la cartographie des teneurs en C_{tot} obtenue en appliquant l'équation du modèle de la régression MCP, propre à la variable C_{tot} , aux neuf bandes de la sub-image ASTER.

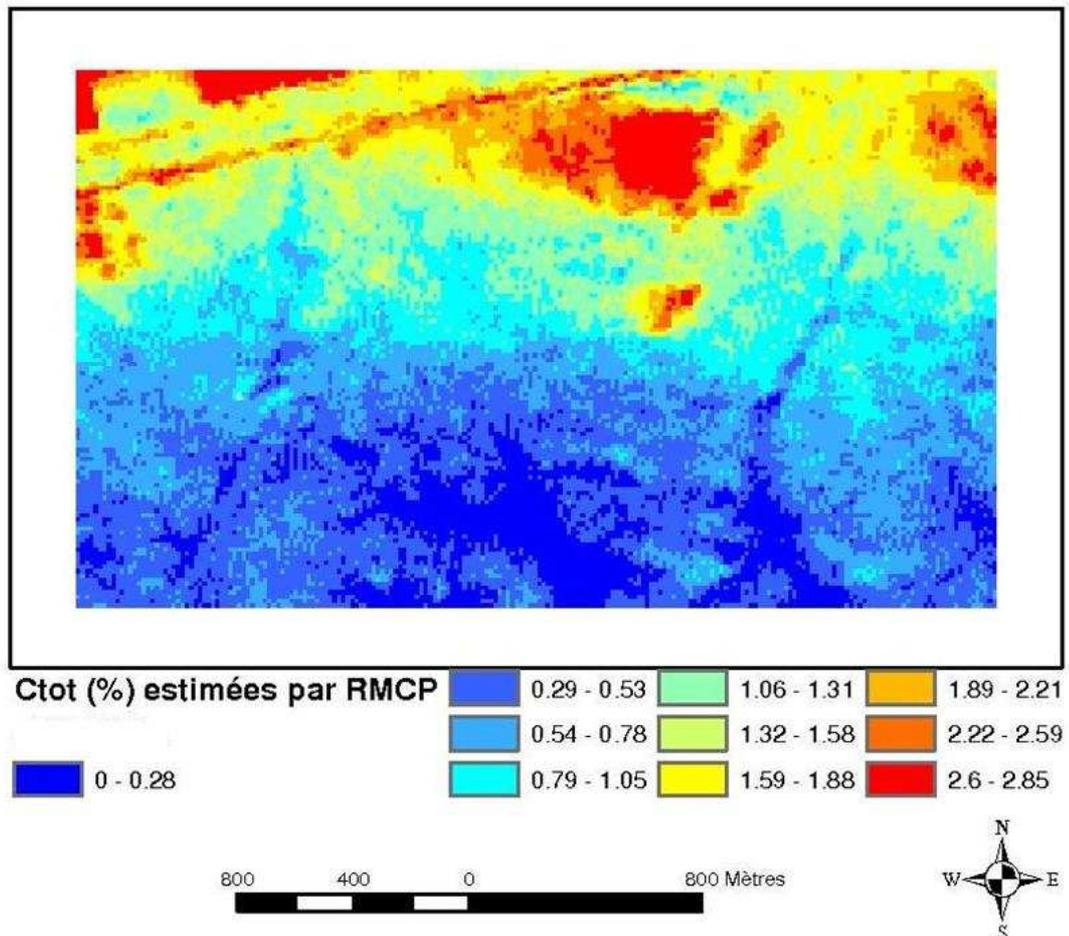


Figure 7a. Cartographie du carbone total sur la zone d'étude par application du modèle RMCP, propre à la variable C_{tot} , aux neuf bandes de la sub-image ASTER d'une résolution de 15 m.

Nous constatons que la spatialisation du modèle spectrométrique de régression MCP propre à la variable C_{tot} génère des classes de teneurs estimées qui sont fidèles aux structures physiques observées sur le terrain. En effet, dans la partie méridionale, les teneurs sont faibles alors que dans la partie septentrionale où la croûte calcimagnésique est affleurante, les teneurs sont fortes. La plage à forte teneur en C_{tot} qui se dégageait de cette spatialisation a une signification physique puisqu'elle circonscrit d'une part les monticules et d'autre part la piste qui est ouverte dans un litholosol gypso-calcaire et délimité par une palissade formée de palmes sèches (riche en carbone organique) utilisé comme un brise-vent pour protéger la palmeraie contre l'ensablement.

La figure 7b montre la cartographie des valeurs de la CE en appliquant les coefficients du modèle de la RMCP, propre à la variable $CE_{1:5}$, aux neuf bandes de la sub-image ASTER.

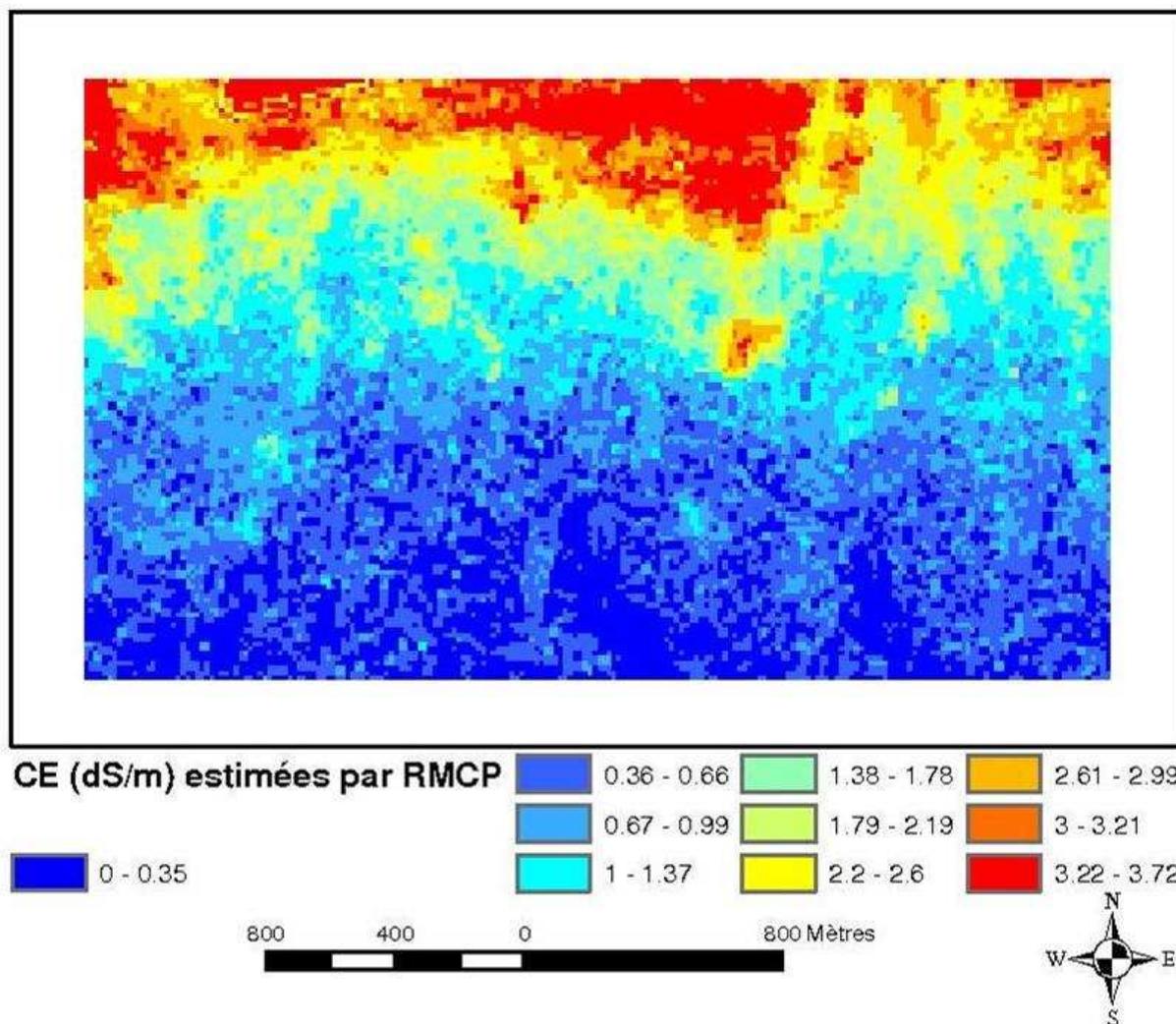


Figure 7b. Cartographie de la $CE_{1:5}$ sur la zone d'étude par application du modèle RMCP, propre à la variable CE, aux neuf bandes de la sub-image ASTER d'une résolution de 15 m.

Nous constatons la présence d'un gradient sud nord. La CE augmente lorsque nous nous déplaçons du sud vers le nord, c'est-à-dire dans le sens du Chott.

3.6. Analyse variographique des résidus de la prédiction

La figure 8a montre le variogramme expérimental moyen pour les résidus de la variable C_{tot} . Ce variogramme a été calculé avec un pas moyen équivalent au pas d'échantillonnage sur le terrain, soit 200 m, jusqu'à une distance de 2000 m. Le meilleur ajustement est un modèle anisotrope sphérique avec un effet de pépite de $0.038 (\%)^2$ et un palier de $0.208 (\%)^2$. La portée majeure de ce modèle est de 1985 m selon l'orientation Est-Ouest alors que la portée mineure

est de 1324 m selon l'orientation Nord-Sud. Ceci indique que la variation spatiale est plus importante dans la direction Nord-Sud. Le faible effet de pépite qui représente 18 % de la variance totale exprime une variabilité locale faible.

La figure 8b montre le variogramme surfacique du site d'étude pour les résidus du C_{tot} . Le variogramme est anisotrope, la semi variance $\delta(h)$ est plus importante dans la direction Sud/Nord ce qui indique que la dissimilarité spatiale est la plus importante dans cette direction

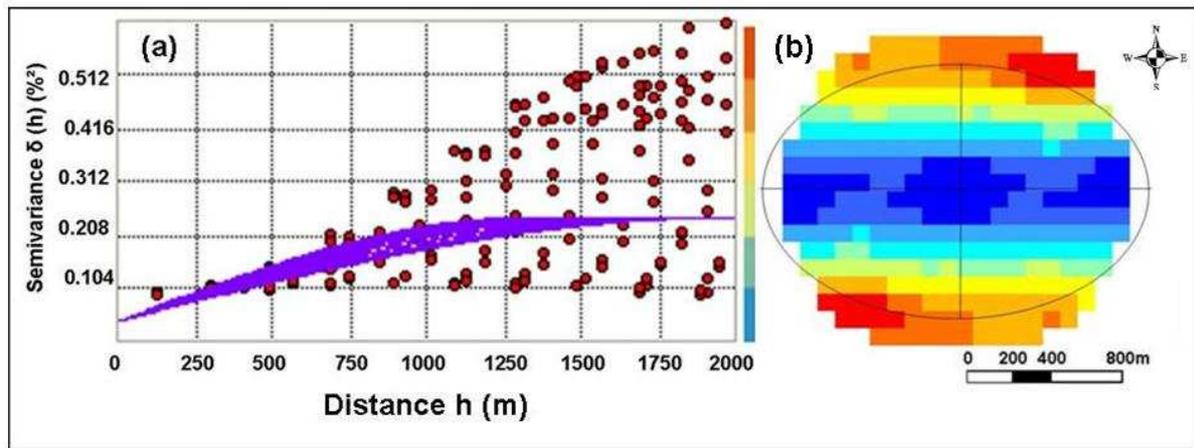


Figure 8. Variogrammes expérimentaux directionnels (a) et surfacique (b) pour les résidus de la variable C_{tot} .

La figure 9 montre le variogramme expérimental pour les résidus de la variable CE. Ce variogramme a été calculé avec un pas moyen d'échantillonnage de 200 m jusqu'à 1800 m. Le meilleur ajustement est un modèle anisotrope sphérique avec un effet de pépite de $0.31 (\%)^2$ et un palier de $0.45 (\%)^2$. La portée majeure de ce modèle est de 1772 m selon la direction Est/Ouest alors que la portée mineure est de 1190 m selon la direction Nord/Sud. Ceci indique que la variation spatiale est plus importante dans la direction Nord/Sud. La valeur élevée de l'effet de pépite qui représente près de 69 % de la variance totale, exprime une variabilité locale importante (Bourennane, 1997 ; Douaoui, 2006).

La figure 9b montre le variogramme surfacique des erreurs d'estimation de la $CE_{1.5}$. La semi-variance est relativement élevée et présente une anisotropie selon un gradient de variation maximale Nord/Sud.

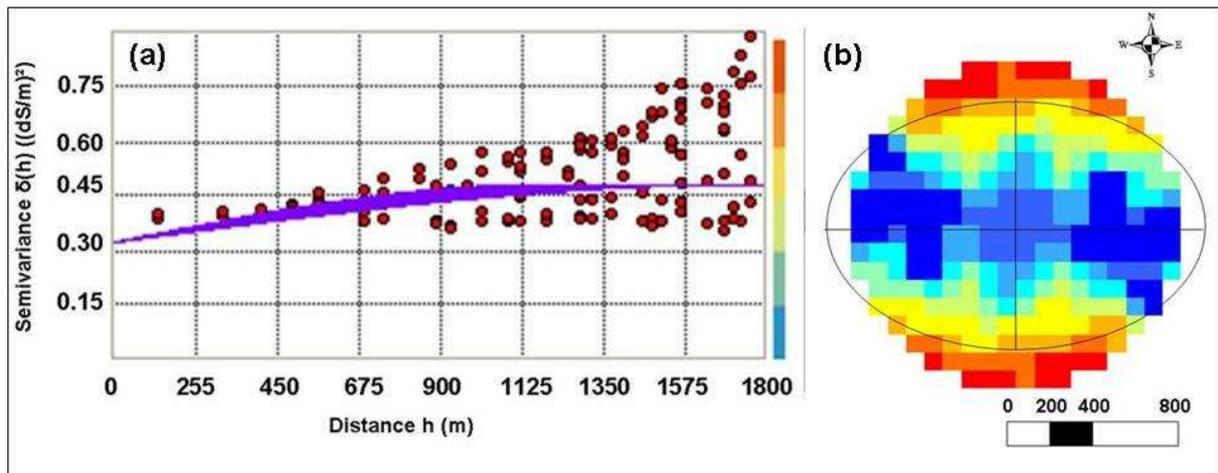


Figure 9. Variogrammes expérimentaux directionnels (a) et surfacique (b) pour les résidus de la variable $CE_{1.5}$.

La figure 10 présente la couche d'interpolation par krigeage des résidus de la régression MCP, dans le cas de la variable C_{tot} : les valeurs des erreurs présentent un gradient général de direction Sud/Nord. Rappelons que les valeurs mesurées aux points d'échantillonnage sont élevées au nord et faibles au sud. Ceci est cohérent avec ce que nous avons observé sur la figure 6a. En effet, le modèle surestime les fortes teneurs en C_{tot} avec des écarts à la 1^{ère} bissectrice relativement importants. Par ailleurs, ces fortes erreurs observées dans la partie nord pourraient s'expliquer par la présence du gypse et de l'argile qui forment à côté des carbonates les structures monticulaires gréseuses et qui influencent directement les propriétés spectrales de la surface du sol.

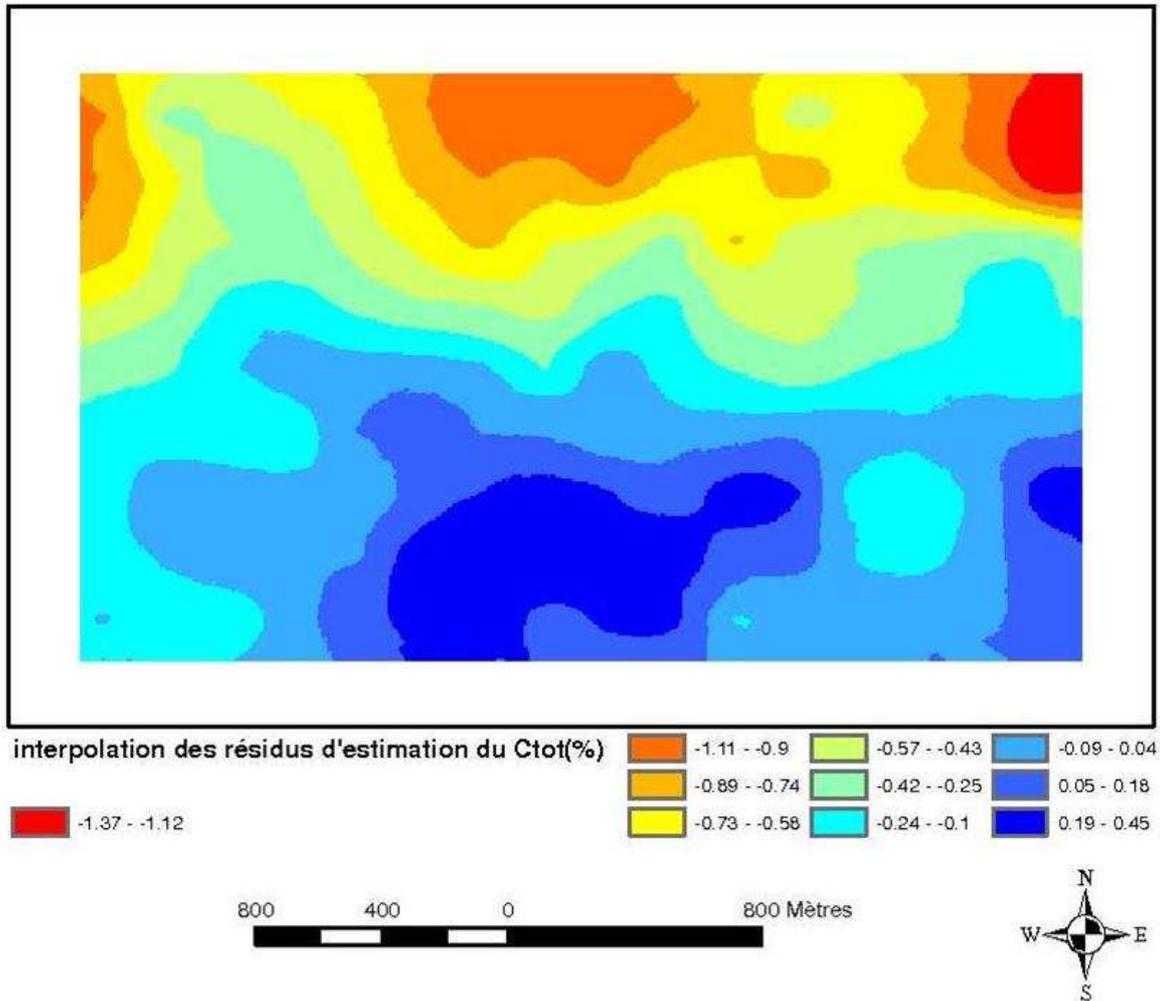


Figure 10. Krigage ordinaire des résidus de la régression MCP, obtenue dans le cas de la variable C_{tot} .

La figure 11 montre la carte de spatialisation du taux de C_{tot} , obtenue par régression MCP-krigeage par sommation des deux couches rasters (spatialisation de la régression MCP (figure 7a) + interpolation des résidus (figure 11)).

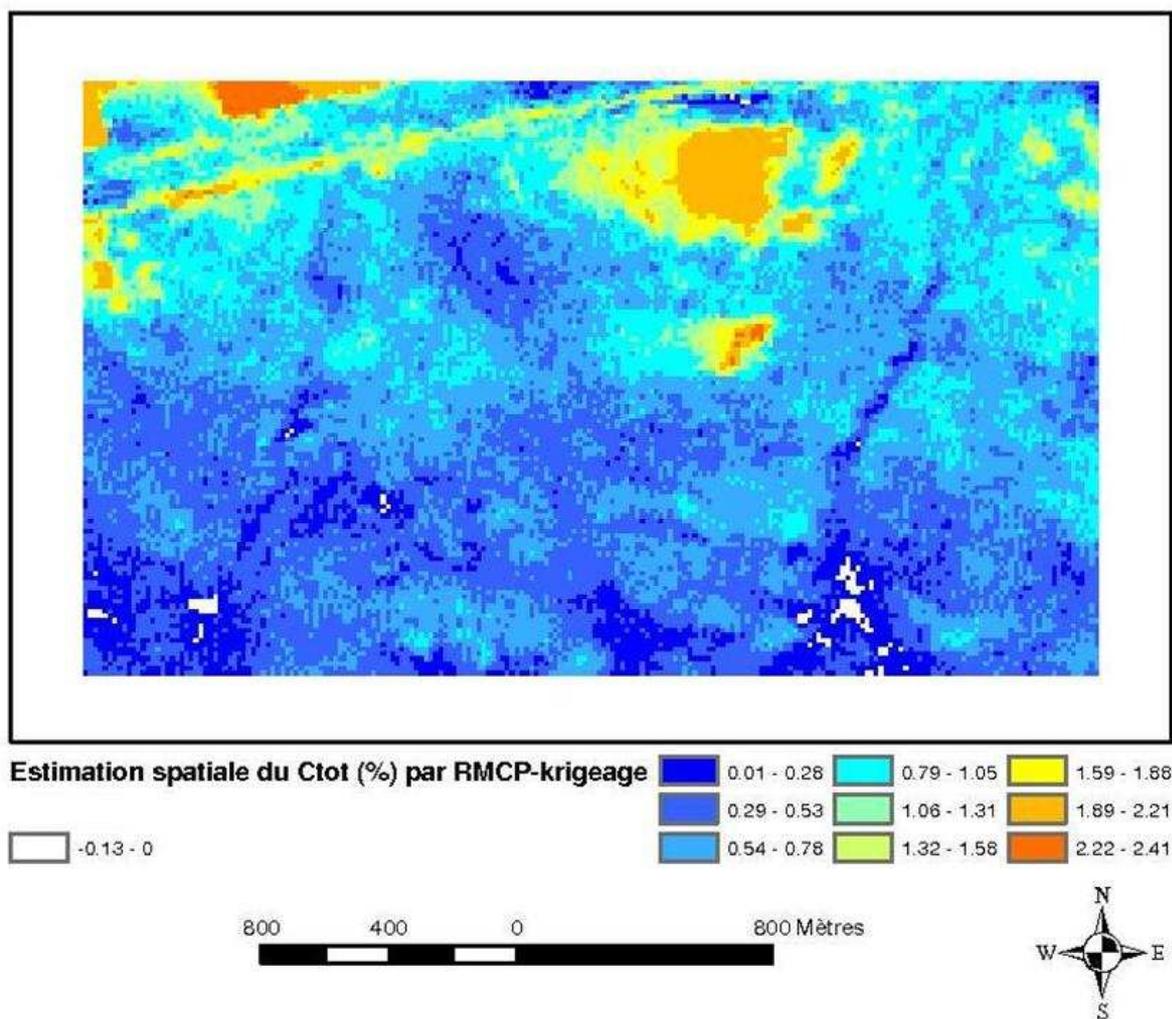


Figure 11. Carte de spatialisation du taux de C_{tot} obtenue par régression MCP-krigeage (issue de la sommation des deux couches rasters : spatialisation de la régression MCP figure (7a) + interpolation des résidus figure (11)).

D'une façon qualitative, nous remarquons que la carte obtenue présente des plages spatiales continues, mais contrastées, dont la forme concorde avec celles des structures physiques qui façonnent le milieu (monticules, réseau physiographique, plages de sables, piste et brise-vent, limites de la palmeraie). Les taux de C_{tot} les plus élevés se concentrent au niveau du monticule. La nature calcimagnésique de son matériau originel explique ces fortes valeurs. Les valeurs pixellaires les plus faibles (qui sont ici négatives) se situent au niveau des ravins qui constituent des sols déblayés et décarbonatés, par l'effet mécanique et chimique des eaux de ruissellement, légèrement couvert par un film éolien. Dans le cas d'une régression suivie d'un krigeage, un problème commun est que cette méthode peut produire des valeurs en dehors de la gamme physique (i.e valeurs négatives) et ces aires peuvent être masquées ou corrigées (Goovaerts, 1997 ; Hengl et *al.*, 2004). Nous avons délibérément gardé ces valeurs négatives car la classe de ces valeurs est spatialement entièrement ceinturée par la classe des plus

faibles valeurs positives estimées et par conséquent elle peut être facilement confondue avec cette dernière. Par ailleurs, cette classe des valeurs estimées les plus faibles représente physiquement le réseau physiographique.

La figure 12 montre les valeurs de C_{tot} estimées par RMCP-Krigeage au niveau des pixels de collecte des échantillons vs les valeurs mesurées au laboratoire. Cette figure permet d'apprécier la qualité d'estimation de la carte d'estimation des taux de C_{tot} par la méthode de la régression MCP-krigeage. Nous notons que l'application du krigeage des résidus comme traitement postérieur à l'application de la régression MCP aux bandes de la sub-image satellitaire a significativement amélioré la qualité de l'estimation. En effet, nous obtenons un coefficient de détermination $R^2 = 0.78$, un biais de 0.003 % et une REQM = 0.16 %. Par ailleurs, le tableau 5 rappelle les résultats des performances d'estimation par régression MCP-Krigeage (Tableau 4) et expose les résultats des performances d'estimation par régression MCP-Krigeage dans le cas des deux variables cartographiées. Dans le cas de la variable C_{tot} le ratio étendue/REQM est passé de 4.29 à 13.77.

Tableau 5. Résultats des performances d'estimation en appliquant le modèle de la régression MCP puis le modèle de régression MCP- Krigeage

		R ²	Biais	REQM	REQMR	Etendu/REQM
Régression MCP	C_{tot} (%)	0.53	0.27 %	0.52 %	74.77 %	4.29
	CE (dS/m)	0.57	0.36 dS/m	0.75 dS/m	82.98 %	6.44
Régression MCP-krigeage	C_{tot} (%)	0.78	0.003 %	0.16 %	23.32 %	13.77
	CE (dS/m)	0.60	-0.02 dS/m	0.55 dS/m	61.17 %	8.74

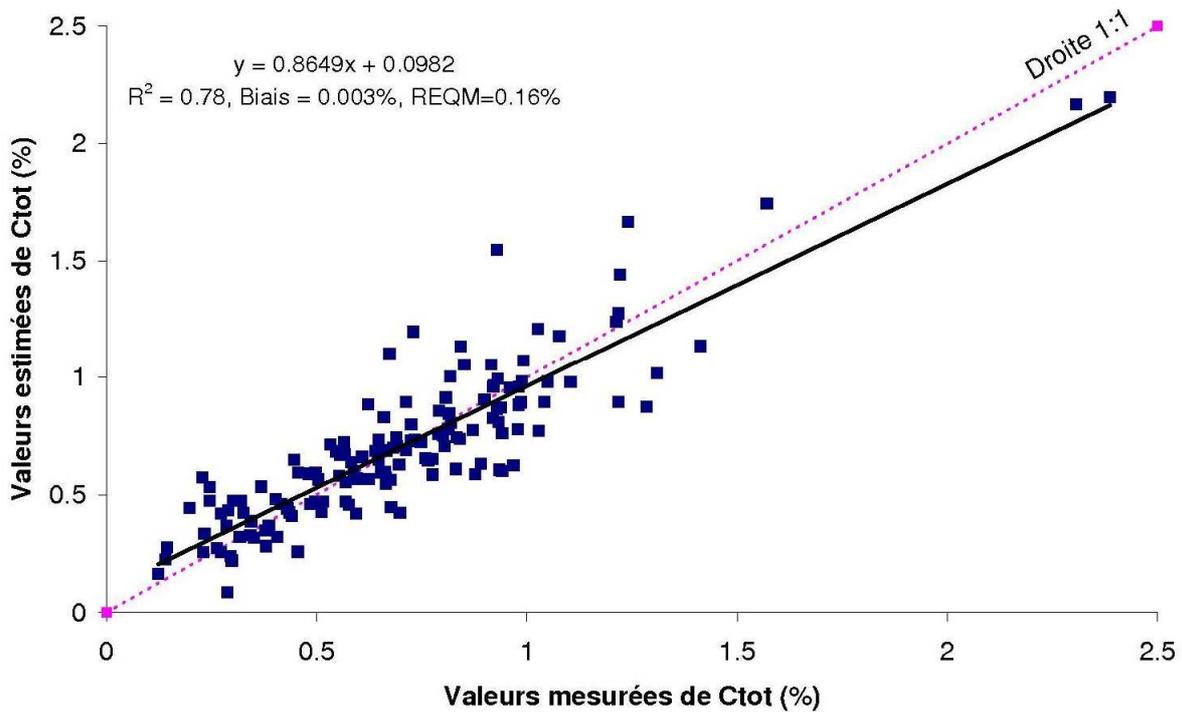


Figure 12. Valeurs de C_{tot} estimées par RMCP-krigeage au niveau des 144 pixels de la carte générée, correspondants aux points de collecte des échantillons sur le terrain, vs les valeurs mesurées au laboratoire.

La figure 13 présente la couche d'interpolation par krigeage des résidus de la $CE_{1:5}$. Nous notons que, comme pour le C_{tot} , l'erreur résiduelle en valeur absolue présente également un gradient général Sud/Nord qui est associé à des structures locales dont l'extension spatiale est conditionnée par la portée du variogramme. Toutefois, la distribution spatiale de l'erreur pour la CE est caractérisée par des transitions graduelles. Ceci pourrait s'expliquer par une micro-régionalisation qui par ailleurs, n'a pas été décelée dans le variogramme.

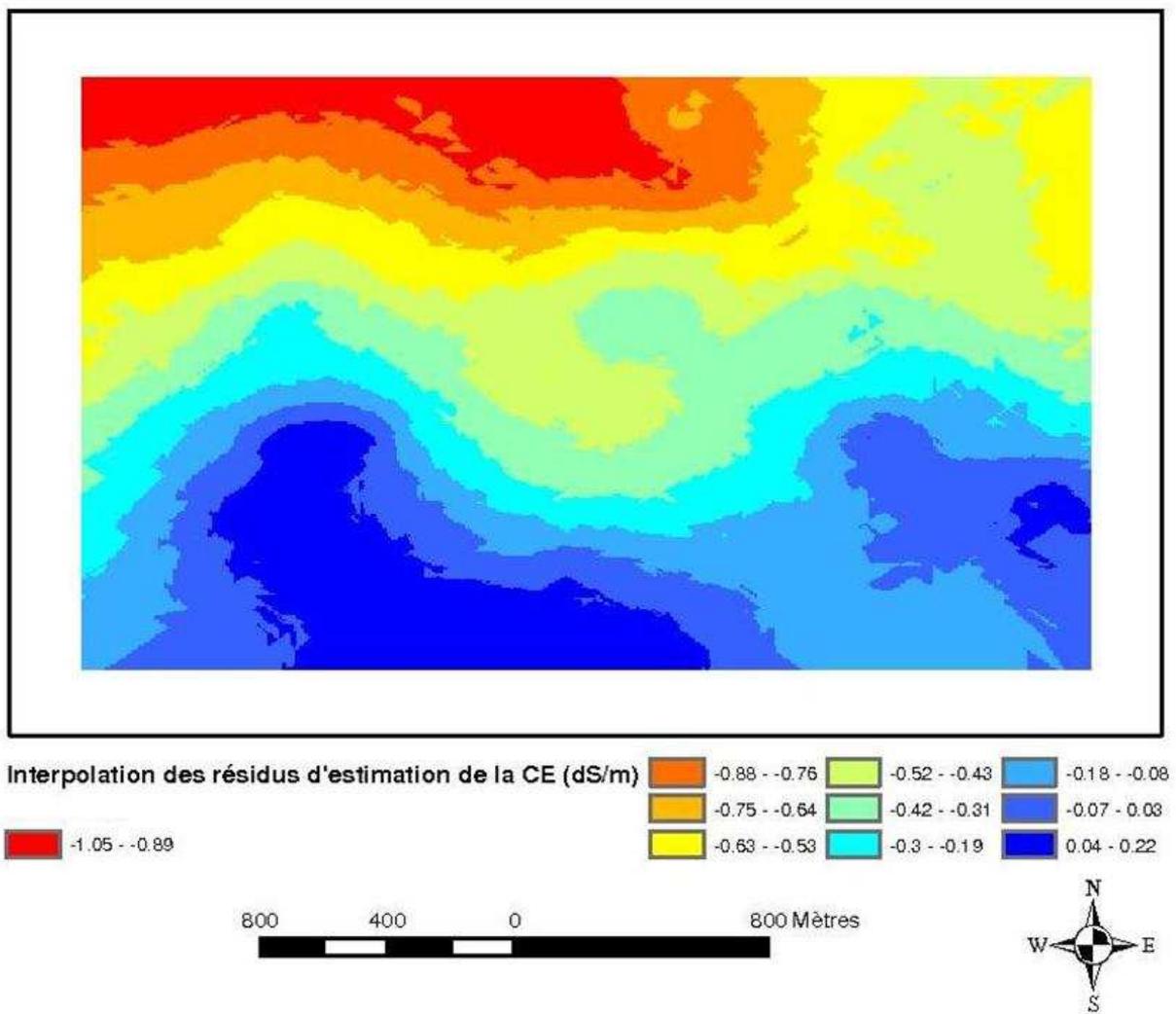


Figure 13. Couche d'interpolation par krigeage ordinaire des résidus de la $CE_{1.5}$.

La figure 14 montre la carte de spatialisation de la $CE_{1.5}$ obtenue par régression MCP-krigeage. Qualitativement, nous remarquons que tout comme pour la carte de spatialisation du C_{tot} , les plages spatiales sont continues. Par contre, les limites de l'espace occupé par les monticules sont moins nettes. Ceci indique que ces derniers ne se distinguent pas par une classe de CE particulière et qu'ils se confondent pour cette propriété pédologique avec d'autres structures du paysage. Nous notons l'existence d'un gradient de direction Sud/Nord. Toutefois, les pixels sont beaucoup plus bariolés, ce qui dénote d'une distribution diffuse de la propriété pédologique en question et de sa variabilité à des échelles locales. En effet, selon Mc Bratney et *al.* (2003) l'un des facteurs limitants pour la précision d'une carte numérique serait l'existence de variations locales difficiles à appréhender même en augmentant la résolution de l'échantillonnage.

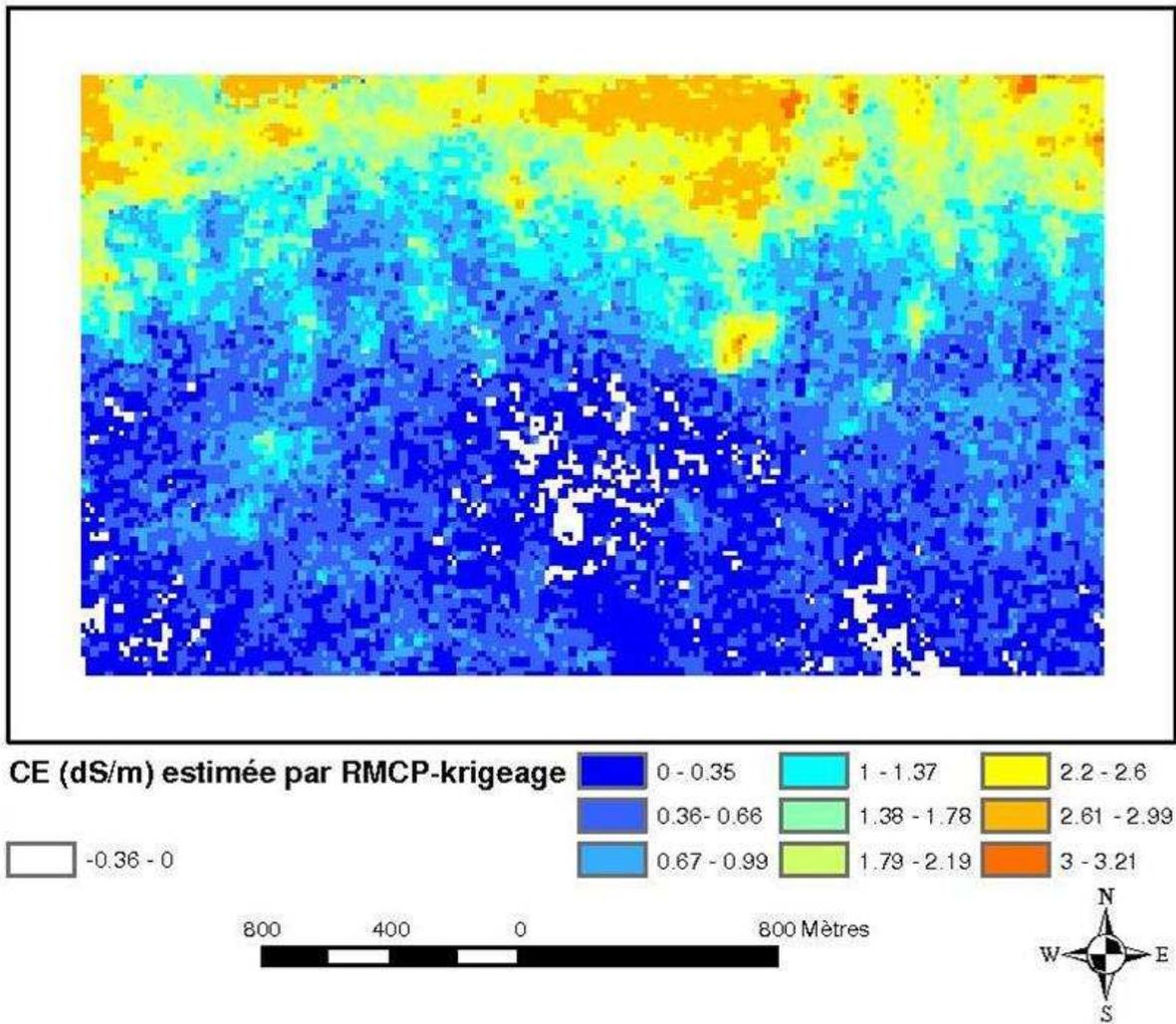


Figure 14. Carte de spatialisation de la $CE_{1.5}$ obtenue par régression MCP-krigeage.

La figure 15 montre les valeurs de $CE_{1.5}$ estimées par RMCP-krigeage au niveau des pixels de collecte des échantillons vs les valeurs mesurées au laboratoire.

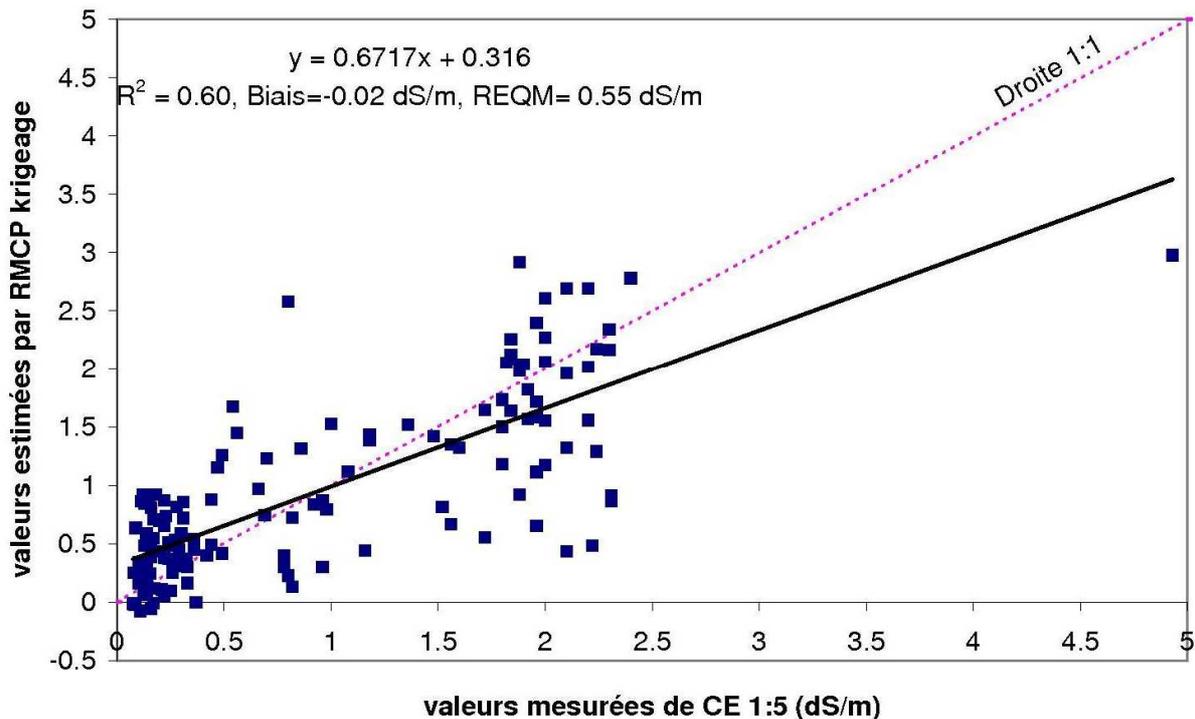


Figure 15. Valeurs de $CE_{1:5}$ estimées par RMCP-krigeage au niveau des pixels de collecte des échantillons vs valeurs mesurées au laboratoire.

La comparaison des résultats obtenus par RMCP-krigeage avec ceux de la simple application de la régression MCP montre que la qualité de l'estimation est du même ordre ($R^2 = 0.57$ contre $R^2 = 0.60$). Toutefois, tous les autres critères statistiques des performances du modèle sont significativement améliorés. En effet, comme nous pouvons noter dans le tableau 4, le biais est passé de 0,36 dS/m à 0,003 dS/m, la REQM est passé de 0.75 à 0.55 dS/m, le REQMR est passé de 83.0 % à 61.2 % et le ratio étendue/REQM est passé de 6.44 à 8.74 (Cf. tableau 5). La REQM et la REQMR qui sont des indicateurs appropriés de l'amplitude de l'erreur d'estimation (Jarmer et al., 2008) sont faibles, ce qui indique que la carte produite est utile, d'autant plus que la distribution spatiale de ces quantifications possède bien un sens physique.

Par ailleurs, s'il est incontestable que la qualité de l'estimation de la CE s'est relativement améliorée, elle reste cependant inférieure à ce que nous avons obtenu pour le C_{tot} . Cette moins bonne qualité d'estimation pour la CE pourrait être attribuable à la nature intrinsèque de cette propriété pédologique. En effet, la CE est une propriété à distribution spatiale diffuse qui demeure variable dans le temps (Conel et al., 1988 ; Mougenot, 1993), a *fortiori* dans le cas des sols nus, sans manifestations spécifiques de surface affectés par les sels, où la CE à la surface du sol est très conditionnée par la pluie (dilution/concentration) et les

vents (redistribution des sels à partir des chotts voisins et assèchement du sol) (Job, 1992 ; Mougenot, 1993). Le fait que la conductivité électrique présente une distribution diffuse laisse supposer l'existence d'une micro-régionalisation non détectée, dans notre cas, à cause du pas d'échantillonnage (Witterlind et al., 2008). L'application d'un échantillonnage avec des pas différents pour l'amont, la transition et l'aval, aurait donné des résultats différents (Bourennane, 1997). La technique de régression MCP-krigeage pourrait produire des cartes plus précises si la densité d'échantillonnage avait été plus grande. En effet, la précision d'estimation est très sensible à la densité des points d'observation (Lagacherie et Voltz, 2000). Douaoui (2005) qui a cartographié la salinité à une échelle semi-régionale, pour des sols argileux, à partir de mesures ponctuelles avec un pas d'échantillonnage de 250 m a conclu que l'effet de pépite important du variogramme expérimental montrait l'existence d'une variabilité locale importante qui ne peut être cernée que par un pas d'échantillonnage beaucoup plus serré ou le cas échéant par l'intégration de données de télédétection.

Farifteh et al. (2007) en travaillant sur des images hyperspectrales (Hymap) couvrant des zones moins importantes que la nôtre (35 ha), autour d'un lac en Australie, ont obtenu de bonnes estimations de la CE (R^2 entre 0.8 et 0.96) avec une régression MCP. Ces bonnes performances sont d'une part liées à une meilleure résolution spatiale (3.5 m) et spectrale (128 bandes) et d'autre part à la forte densité de l'échantillonnage sur les différentes zones (entre 1060 et 3069 mesures par zone d'une superficie de 9 ha). Toutefois, le coût d'une telle cartographie thématique est très élevé. Notre approche s'avère mieux adaptée pour une cartographie sur des étendues plus vastes, surtout si nous souhaitons nous placer dans une perspective de suivi et de surveillance des propriétés des sols. En effet, nous avons montré que dans le cas du C_{tot} , nous obtenons des résultats très satisfaisants avec une densité d'échantillonnage beaucoup moins contraignante (144 échantillons pour 580 ha). De plus, si cette même approche aboutit pour la CE à des valeurs moins précises, il n'en demeure pas moins que la carte obtenue et les structures spatiales qu'elle montre, sont cohérentes avec la géomorphologie du milieu physique. En effet, *de facto*, la partie haute située en amont est moins salée que la partie basse située en aval.

4. Conclusion

A la lumière de nos résultats, nous avons pu montrer l'intérêt de l'utilisation du capteur ASTER pour la spatialisation des propriétés du sol en milieu aride. La méthode que nous avons mise en œuvre pour décrire quantitativement la variabilité des sols est peu contraignante et de loin moins coûteuse qu'une mission organisée avec un spectroradiomètre imageur qui ne permet qu'une cartographie à grande échelle. Cette méthode combine la spatialisation d'une régression MCP établie à partir des spectres extraits de l'image corrigée en réflectance relative à l'aide de spectres acquis au laboratoire avec une interpolation des erreurs résiduelles d'estimation calculées au niveau des points d'observation. A l'aide de cette méthode, nous avons pu générer des cartes thématiques de quantification présentant les performances d'estimation suivante $R^2 = 0.78$ et 0.60 , $REQM = 0.16 \%$ et 0.55 dS/m respectivement pour le C_{tot} et la $CE_{1:5}$. Par ailleurs, cette approche ne nécessite pas des mesures spectrométriques de terrain et encore moins de faire coïncider ces mesures au sol avec les dates de passage du capteur comme c'est le plus souvent le cas avec les spectroradiomètres imageurs.

Nos résultats font émerger des perspectives intéressantes pour la cartographie de propriétés du sol sur de larges territoires. Néanmoins, ces résultats s'inscrivent dans un contexte spécifique :

- un contexte pédologique bien précis, des surfaces de sol nues et sèches d'un milieu aride saharien ;
- deux propriétés retenues en raison de leur intérêt dans la zone d'étude (C_{tot} , $CE_{1:5}$).

Notre méthode doit donc être testée dans le même contexte climatopédologique pour d'autres variables d'intérêt comme les fractions granulométriques, la teneur en fer, teneur en gypse, *etc.* Il faudrait également considérer d'autres contextes pédoclimatiques notamment en milieu tempéré, avec d'autres variables qui seraient pertinentes pour ces milieux, par exemple les teneurs en carbone organique, argile, humidité du sol, *etc.* De plus, l'application de cette approche en milieu tempéré, où les sols sont le plus souvent couverts par une couverture végétale, permettrait de vérifier si nous pouvons nous affranchir de l'effet du couvert végétal sur la réponse spectrale pour pouvoir accéder directement aux propriétés du sol et les quantifier. Par ailleurs, notre méthode qui a été appliquée pour une image multi-spectrale peut *a fortiori* s'appliquer à des images hyperspectrales qui offrent un nombre plus important de bandes caractérisant mieux la signature spectrale des surfaces physiques étudiées.

Nos résultats sont encourageants et ouvrent des perspectives très intéressantes pour l'utilisation de la télédétection dans la quantification spatio-temporelle des propriétés des sols, notamment dans les régions non couvertes par des études cartographiques et/ou d'accès difficile.

La contrainte majeure à une plus large utilisation de la Régression MCP-krigeage est le fait que le praticien doit effectuer les diverses étapes de son traitement à l'aide de plusieurs logiciels (Hengl et *al.*, 2007) notamment de traitement d'image de télédétection, de spectrométrie, de statistiques et de SIG. L'offre d'une interface commune entre ces logiciels pourrait contribuer à faciliter l'usage de cette technique.

Conclusion générale

« *Une conclusion, c'est quand vous en avez assez de penser* »

Herbert Albert Fisher (historien, 1865-1940)

1. Rappel des objectifs de la thèse

La spectroscopie par réflexion diffuse dans le visible et le proche infrarouge est une technique qui est de plus en plus utilisée en sciences du sol. Elle est décrite comme une technique d'analyse quantitative alternative aux analyses physico-chimiques conventionnelles du laboratoire. Cette technique qui repose sur le principe physique d'interaction de la lumière avec la matière est surtout attractive pour sa rapidité et son faible coût. Les recherches actuelles en pédo-spectroscopie intéressent deux domaines d'intérêts qui sont à la fois indépendants et complémentaires. Il s'agit en l'occurrence de la prédiction ponctuelle des propriétés des échantillons de sol par opposition à la prédiction spatiale. Ces deux types de prédictions peuvent intéresser d'une façon simultanée ou différée diverses propriétés du sol et ce à partir d'une même mesure spectrale.

La baisse constante du coût de l'instrument de mesure (spectroradiomètre) associée au développement et à la promotion de nouveaux logiciels de chimiométrie laisse suggérer que la spectroscopie puisse dans les années à venir, pour certaines propriétés du sol, supplanter les techniques conventionnelles d'analyse du sol et devenir des méthodes de références qui répondent mieux aux exigences normatives (reproductibilité des résultats).

Toutefois au jour d'aujourd'hui, bien que les avantages qu'offre la spectroscopie soient incontestables, les pédométriciens n'ont pas statué sur plusieurs questions primordiales liées notamment aux protocoles d'acquisition des données et de leurs traitements. Des recherches supplémentaires sont nécessaires avant de pouvoir standardiser les étapes de préparation, d'acquisition des spectres et de traitement chimiométriques des données. Sur un autre plan, d'autres approches de spatialisation des propriétés du sol qui incorporent les données spectroscopiques, faciles à mettre en œuvre et peu onéreuses doivent être développées et promues.

L'objectif global de cette thèse était la prédiction à différentes échelles et pour différentes fenêtres spectrales des propriétés physico-chimiques du sol par télédétection rapprochée (spectroscopie) et spatiale en milieux tempéré et aride. Pour cela nous avons utilisé l'information spectrale dans les domaines visible et proche infrarouge (Vis-PIR) obtenue

sur des échantillons de sols, préalablement sélectionnés et caractérisés, pour développer des modèles spectrochimométriques de prédiction de diverses propriétés du sol dans deux contextes pédologiques différents. Puis nous avons développé une méthode de spatialisation des propriétés du sol par couplage des prédictions spectrométriques avec les données continues d'une image satellite acquise avec un capteur appartenant à la nouvelle génération des capteurs multi-spectraux à savoir le capteur ASTER.

Pour y parvenir, nous avons conduit le travail comme suit :

Dans une première étape nous avons commencé par constituer deux séries d'échantillons artificiels pour étudier l'influence singulière de deux paramètres qui influencent la réponse spectrale « chromophores » et nous avons testé la robustesse du modèle chimométrique que nous envisageons d'utiliser ultérieurement dans les prédictions à savoir la régression en moindres carrés partiels RMCP.

Dans une seconde étape, dans un contexte de milieu tempéré (la région de Bretagne, en France), nous avons calibré des modèles à étendue régionale de prédiction de diverses propriétés du sol, dans diverses fenêtres spectrales Vis, PIR et Vis-PIR et discuté de la performance de prédiction en fonction des propriétés du sol et du domaine spectral. Par ailleurs, pour sélectionner les échantillons de calibration et de validation nous avons utilisé la procédure d'Analyse en Composantes Principales (ACP).

Dans une troisième étape, dans un contexte de milieu aride (la région du Djérid, en Tunisie), nous avons calibré pour deux propriétés pertinentes pour la surveillance environnementale de ce type de milieu à savoir le carbone total (dont la fraction minérale est prépondérante) et la conductivité électrique, des modèles « site spécifique » puis des modèles à étendue régionale et nous avons étudié la variation des performances des modèles prédictifs en fonction de l'étendue de l'aire géographique de collecte des échantillons de calibration et de validation.

Dans une dernière étape, et toujours dans ce contexte de milieu aride (la région du Djérid, en Tunisie), nous avons utilisé une approche géostatistique hybride à savoir la régression en moindres carrés partiels-krigeage pour mettre en œuvre une méthode qui permette de coupler les données spectrométriques de laboratoire et les données continues, dérivées des 9 bandes du visible et du proche infrarouge d'une sub-image satellite ASTER pour spatialiser les teneurs en carbone total et la conductivité électrique.

2. Principaux résultats acquis

A l'aide des deux séries d'échantillons artificiels que nous avons constitué nous avons pu étudier les aspects suivants :

- les formes des spectres des sols ;
- l'Influence qualitative du carbone organique sur la réponse spectrale ;
- l'Influence qualitative du fer total sur la réponse spectrale ;
- l'effet conjugué du fer total et du carbone organique sur la réponse spectrale ;
- l'intérêt de l'utilisation de la RMCP en modélisation des mesures spectrométriques.

2.1. Formes des spectres des sols

Les courbes de réflectance des sols sont continues et croissantes sur tout le spectre du visible et du proche infrarouge. La réponse spectrale change graduellement d'une longueur d'onde à une autre. Une forte corrélation existe entre les bandes voisines. Cette redondance spectrale a déjà été notée par Weidung et al. (2002) qui ont mis en relation la réponse spectrale avec le taux d'humidité du sol. En règle générale, les sols riches en carbone organique ont une courbe concave, dans le domaine du visible et du proche infrarouge, alors que les sols les moins riches ont des courbes plutôt convexes. Ces observations des formes classiques des spectres ont été relevées par des auteurs tels qu'Escadafal et al. (1993), Weidung et al. (2002) et Demattê et al. (2002).

2.2. Influence du fer total sur la réponse spectrale

Le fer présente des pics d'absorption pouvant être décelés par la simple visualisation des spectres de réflectance apparente. Il s'agit notamment de la bande 400-440 nm et de la bande 930-950 nm. La normalisation des spectres par retrait du continuum a permis d'identifier les pics d'absorptions de fer d'une façon beaucoup plus précise. Ces pics sont centrés sur les longueurs d'onde suivantes : 431, 497, 677, 795 et 907 nm. Les pics d'absorption les plus importants sont situés dans le visible à 431 et à 497 nm. Pour la longueur d'onde 497 nm, la relation linéaire inverse entre la teneur en fer et la réflectance normalisée est évidente. En dehors de ces deux pics, le fer entraîne un effet « dépresseur » qui est généralisé sur tout le spectre. Cela rejoint les résultats d'investigation de Rezend et al. (1980) qui avaient observé une absorption du fer dans la bande spectrale 440 à 460 nm sur des courbes de réflectance d'Oxisols. De même Girard et Girard (1999) mentionnent que le fer présente une absorption entre 860 et 910 nm.

2.3. Influence du carbone organique sur la réponse spectrale

Dans notre cas, sur le domaine spectral d'investigation 400-950 nm, le carbone organique a entraîné un effet « dépresseur » sur tout le spectre, sans présenter de bande spécifique comme le fer. Certains auteurs qui avaient comme objectif la caractérisation de la matière organique (Ben Dor *et al.*, 1997 ; Fourty et Baret, 1998 ; Demattê *et al.*, 2004) ont exploré un domaine spectral plus large que le nôtre (400-2600 nm) et ont noté des pics d'absorptions liés aux liaisons carbone au voisinage de 760, 900, 2080, 2350 et 2500 nm. Selon Coûteaux *et al.* (2003), ces pics sont spécifiques de certains produits de la décomposition de la matière organique (cellulose, lignine, amidon, cire, acide humique, *etc.*). Par ailleurs, la nature de la matière organique et le stade de sa décomposition sont déterminants dans l'observation de ces pics. En effet, Krishnan *et al.* (1980) ont constaté que le carbone organique présente un pic à 800 nm. Or, d'après les travaux de Ben Dor *et al.* (1997), ce pic n'est observable que lorsque la matière organique est humifiée.

2.4. Effets conjugués du fer total et du carbone organique sur la réponse spectrale

Nous avons constaté que pour des teneurs élevées en carbone organique, l'effet du carbone masque ou distord l'effet du fer. Ceci expliquerait le fait qu'après retrait du continuum, les spectres des deux échantillons les plus riches en carbone organique et les moins riches en fer de la deuxième série ne sont pas ordonnés en fonction de leur teneur en fer comme c'est le cas pour le reste des échantillons de la même série. En effet, Baumgardner *et al.* (1970) qui ont étudié des sols de l'Indiana (EUA), soulignaient que pour des taux élevés de matière organique qui dépassent les 2 %, l'effet carbone peut masquer l'effet des autres constituants minéraux sur les spectres des sols. De même, He *et al.* (2005) estiment que la forte propriété d'absorption de la matière organique peut masquer les propriétés d'absorption des autres constituants minéraux du sol. Par contre, pour des valeurs faibles de teneurs en CO, nous avons constaté que la relation entre l'effet carbone et l'effet fer est de type « additif ». En effet, le fer entraîne une absorption de la lumière qui vient s'ajouter à l'absorption induite par le carbone organique.

2.5. Intérêt de l'utilisation des modèles RMCP en modélisation des mesures spectrométriques

Les modèles RMCP sont à la base, prédictifs et non explicatifs : nous avons pu montrer que ce type de modèle statistique intègre la variabilité des réponses pédo-spectrométriques en tant que données d'entrée du modèle, ce qui en fait un outil très intéressant pour étudier

des populations très hétérogènes. En effet, Ludovic et *al.* (1995) soulignaient qu'en modélisation « le modèle doit suivre les données et non l'inverse ».

La régression MCP affecte un coefficient « poids » important aux longueurs d'onde les plus déterminantes dans la prédiction de la variable analytique considérée. Par contre, si une longueur d'onde est considérée sans importance particulière dans la prédiction, le modèle lui attribue un coefficient qui est faible en valeur absolue. Nul besoin de présélectionner les longueurs d'onde les plus prédictives, comme c'est le cas pour les régressions multiples.

2.6. Intérêt de la procédure d'ACP dans la sélection des échantillons de calibration et de validation des modèles à champ d'application régional

La procédure d'analyse en composantes principales que nous avons adoptée a rendu possible :

- l'étude des corrélations inter variables analytiques ;
- la visualisation de la ressemblance et de la dissemblance des échantillons en fonction de leurs propriétés analytiques et par conséquent l'identification et par la suite la soustraction des échantillons atypiques de la base des échantillons ;
- la répartition des échantillons en un jeu de calibration et un jeu de validation qui sont analytiquement similaires. Ils couvrent les mêmes gammes de teneur pour l'ensemble des propriétés analytiques. Ainsi, nous pouvons brasser la pédo-diversité de toute la région.

A ce titre l'ACP peut être utilisée comme une procédure standard de sélection des échantillons pour construire le modèle. En effet, Morra et *al.* (1991) affirment que l'une des limitations inhérentes à l'utilisation des prédictions spectrométriques de n'importe quel constituant du sol, est la nécessité de calibrer le modèle sur un jeu d'échantillons analytiquement bien défini ou d'utiliser une population d'échantillons qui soit homogène. De même d'après les travaux de Tillman et *al.* (2001), cités par Terhoeven-Urselmans et *al.* (2006), la qualité de prédiction est médiocre lorsque les prédictions sont faites sur des échantillons qui sont en dehors de la gamme des échantillons de calibration. De même, la précision de prédiction se détériore avec l'élargissement du champ spatial d'origine des échantillons de calibration.

2.7. Opportunité d'utilisation de la spectrométrie dans la prédiction des différentes propriétés du sol en région tempérée de Bretagne

Les performances des modèles de prédiction, varient d'une variable analytique à une autre. Nous constatons que les modèles spectrométriques régionaux que nous avons calibrés

pour les variables : teneur en azote total, pH, teneur en fer total, teneurs en limons fins et sables fins, tous spectres confondus, sont de qualité médiocres.

Dans cette étude, les bonnes prédictions obtenues dans le cas du carbone organique peuvent être attribuées à l'influence directe des liaisons carbone sur les spectres des sols. Par contre, l'une des explications qui peut être donnée à la bonne prédiction de la CEC serait sa forte corrélation avec le carbone ($r = 0.94$). En effet, d'après Clark et *al.* (1987), la spectroscopie permettrait de déterminer la concentration en certains cations lorsqu'ils sont associés à des molécules organiques ou à des molécules inorganiques hydratées. Nous avons constaté aussi que, seules les fractions granulométriques, qui étaient positivement corrélées à la variable carbone à savoir les variables limons grossiers et sables grossiers ont été bien prédites. Par ailleurs, ces deux fractions granulométriques sont fortement corrélées entre elles ($r = 0.86$). Ces deux fractions seraient indirectement liées à la réponse spectrale des sols de la région par le biais de la matière organique qui serait concentrée dans ces deux fractions. Ben-Dor et Banin (1995) soutiennent l'hypothèse que la spectroscopie peut être utilisée pour prédire des attributs du sol qui n'influencent pas directement son comportement spectral et ce *via* leurs inter-corrélation avec des propriétés du sol qui sont spectralement actives.

Pour le reste des fractions granulométriques notamment les limons fins et les sables fins, la qualité de prédiction était mauvaise. Girard et Girard (1999) estiment que dans le domaine des longueurs d'onde du visible, il ne peut pas y avoir d'interaction des fractions granulométriques du sol avec le rayonnement. Lorsqu'ils sont comparés aux longueurs d'onde, les limons ou les sables ont des tailles trop grandes pour intervenir sur sa réflectance dans le visible ou le PIR. Par contre les tailles des argiles sont trop grandes dans le visible mais elles interviennent dans le PIR. Girard et Girard (1999) présumant que si certains auteurs mentionnent des variations de réflectance selon les teneurs plus ou moins grandes en argiles ou sables, elles seraient alors dépendantes de la teneur en eau de ces éléments. Par ailleurs, Bowers et Hanks (1965) ont démontré que les particules du sol n'influençaient la réponse spectrale du sol que lorsqu'elles présentaient une taille inférieure à 400 μm . De même, Al-Abbas et *al.* (1972) qui ont mis en relation les teneurs en matière organique et en argile de 197 échantillons de sol, prélevés à la surface d'une même parcelle d'une superficie de 25 ha, avec leurs luminances multi-spectrales, affirment ne peut pas être certains que la relation entre la réflectance relative et la teneur en argiles qu'ils ont pu établir était directe ou s'il s'agissait d'une relation indirecte qui résulte dans leur cas de la corrélation entre la teneur en argiles et la teneur en matière organique.

2.8. Fenêtre spectrale offrant les meilleures prédictions

Pour les variables que nous avons pu prédire correctement comme le carbone organique et la capacité d'échange cationique, le domaine du PIR est le plus indiqué pour la prédiction. En effet, les spectres des sols sont généralement monotones et croissants sur tout le domaine du visible. Par contre dans le domaine du proche infrarouge, ils présentent des tronçons de courbes successives, plus ou moins larges, qui sont croissants ou décroissants. Dans le PIR, les spectres sont qualifiés d'ondulés. Ceci s'explique par le fait que le PIR est le domaine des absorptions vibrationnelles de l'énergie lumineuse. Beaucoup d'harmoniques des liaisons des groupements fonctionnels et notamment les liaisons carbone se trouvent dans ce domaine (Bertrand, 2002 ; Lachenal, 2006).

2.9. Parenté pédogénétique des échantillons

En milieu aride, nous constatons que les corrélations entre les propriétés pédologiques varient d'un site spécifique à un autre ce qui montre la pédo-diversité des trois sites échantillonnés qui présentaient des micro-contextes pédologiques spécifiques et ce, bien qu'ils appartiennent à la même région naturelle.

Nous avons montré que la similarité des propriétés analytiques entre les échantillons issus d'un même site est en général supérieure à celle entre les échantillons issus de deux sites différents. La parenté pédogénétique est plus forte pour des échantillons collectés sur un site où agissent, les mêmes facteurs de pédogenèse.

Nous avons aussi montré, en confondant les trois jeux de données à sites spécifiques dans une même base de données à portée régionale, qu'un échantillon considéré comme atypique à l'échelle d'un site donné, peut devenir « normal » et modélisable lorsqu'on passe à une échelle régionale. Ceci indique qu'un échantillon collecté sur un site quelconque peut présenter des propriétés pédologiques qui sont différentes du reste des échantillons collectés sur le même site, mais s'apparente bien à une population d'échantillons collectée à une échelle spatiale beaucoup plus large.

2.10. Comparaison entre modèle « site spécifique » et modèle « régional »

Les modèles calibrés en mode site spécifique notamment pour le C_{tot} , sont plus performants que les modèles calibrés en mode régional. Toutefois par essence, ces modèles ont un champ spatial d'application qui est limité. Les modèles à site spécifique peuvent

trouver des applications dans l'analyse en grandes séries d'échantillons collectés sur le même site, ainsi que pour des analyses diachroniques. Ils sont pour ces raisons forts utiles pour les systèmes de surveillance dans le temps de la qualité des sols et pour l'agriculture de précision.

Les modèles RMCP régionaux que nous avons pu calibrer, aussi bien pour la variable C_{tot} que pour la variable CE, semblent être capables d'intégrer la variabilité intrinsèque des échantillons de sols. Toutefois, ils sont moins performants que les modèles à sites spécifiques. La qualité de ces modèles, notamment pour le C_{tot} peut être davantage améliorée. Pour cela nous sommes persuadés qu'il faut alimenter la base de données avec un nombre plus important d'échantillons de sols pour étoffer la palette des divers micro-contextes pédoclimatiques rencontrés dans la région et recalibrer à chaque fois le modèle jusqu'à tendre vers la qualité des modèles à site spécifique. Tout comme Cozzolino et Moron (2006) et Genot et al. (2007), nous pensons qu'en spectrométrie les modèles qui peuvent être calibrés présentent un besoin continu en recalibration et en contrôle de la qualité de prédiction.

Aussi bien dans le cas de la variable C_{tot} que de la variable CE, lorsque l'échantillon objet de la prédiction n'appartient pas au site spécifique pour lequel le modèle de prédiction a été calibré, il est plus prudent d'utiliser un modèle de prédiction qui soit à portée régionale que d'utiliser un modèle à site spécifique calibré pour un site différent du site de collecte de l'échantillon objet de la prédiction.

Dans un contexte de milieu aride saharien, il est difficile de calibrer un modèle de prédiction de la conductivité électrique à étendu régional : même en faible quantité les sels sont sous formes de cristaux fins dissimulés dans la matrice du sol et remaniés par les vents avec les sables quartziques eux même très réfléchissants. Par contre, les sites préchoteux à forte salinité, où les états de surface observés sont les croûtes salines ou encroûtements, des efflorescences salines, des dalles de sels et des sels en aiguilles, il est possible de calibrer un modèle à site spécifique qui soit performant.

2.11. Prédiction spatiale des propriétés du sol en milieu aride, par couplage des données de spectroscopie visible-proche infrarouge et télédétection satellitale.

La méthode de spatialisation que nous avons développée est fondée sur une régression en moindres carrés partiels-krigeage (MCP-Krigeage). Nous avons pu montrer l'intérêt de l'utilisation des données du capteur ASTER pour la spatialisation des propriétés du

sol en milieu aride. La méthode que nous avons mis en œuvre pour décrire quantitativement la variabilité des sols est peu contraignante et de loin moins coûteuse qu'une mission organisée avec un spectroradiomètre imageur d'autant plus qu'elle permet une cartographie à haute résolution spatiale sur une grande surface d'étude.

Cette méthode combine la spatialisation d'une régression MCP établie à partir des spectres extraits de l'image corrigée en réflectance relative à l'aide de spectres acquis au laboratoire avec une interpolation des erreurs résiduelles d'estimation calculées au niveau des points d'observation. Elle ne nécessite pas des mesures spectrométriques de terrain et encore moins de faire coïncider ces mesures au sol avec les dates de passage du capteur comme c'est le plus souvent le cas avec les spectroradiomètres imageurs.

Nos résultats montrent que le couplage des traitements des données spectrométriques par régression MCP aux outils de géostatistiques peut être pertinent pour spatialiser des propriétés du sol. Dans notre cas la méthode a été testée, à titre d'exemple, pour deux propriétés pédologiques qui conditionnent l'évolution des sols en milieu aride à savoir le C_{tot} et la CE. Le krigeage des résidus a amélioré la qualité d'estimation par régression MCP. En effet, pour le C_{tot} nous obtenons respectivement ($R^2 = 0.78$ et $REQM = 0.16\%$) contre ($R^2 = 0.53$ et $REQM = 0.52\%$). Pour la CE, bien que l'amélioration de R^2 soit moindre (0.60 contre 0.57), les autres critères de performance sont significativement améliorés : le biais (0.003 contre 0.36 dS/m) et la REQM (0.55 contre 0.75 dS/m). Par ailleurs, la distribution spatiale de ces quantifications possède bien un sens physique. Des résultats que nous avons obtenus émergent des perspectives intéressantes pour la cartographie des propriétés du sol à petite échelle.

3. Perspectives de recherche

La spectroscopie est un outil intéressant pour l'étude des sols. Elle peut être mise à profit pour caractériser qualitativement et quantitativement les couvertures pédologiques. Cet outil peut être mis à profit dans plusieurs champs d'applications.

3.1. Estimation en routine des propriétés du sol

Ceci est concevable grâce aux avantages qu'elle présente en termes de facilité de préparation des échantillons pour l'analyse, la reproductibilité des résultats, le faible coût des mesures, la rapidité de l'obtention des estimations. Les modèles régionaux de carbone organique et de la CEC calibrés en milieu tempéré et les modèles « site spécifique »

performants pour la variable C_{tot} et pour la variable CE, en milieu aride, peuvent être utilisés pour des estimations en grande série pour des échantillons appartenant aux mêmes espaces géographiques que les échantillons de calibration de ces modèles.

En perspective, la calibration des modèles peut se faire selon plusieurs options en fonction de la nature des données disponibles. Nous pouvons concevoir la constitution d'une base de données (pédospectrale et analytique) liée à un système d'information géographique. Cette base va stocker toutes les données qui peuvent être disponibles sur les échantillons collectés : coordonnées géographiques, types de sols, analyses physiques, chimiques et biologiques et analyses spectrales. Elle sera élargie en continu, au fur et à mesure de la multiplication des campagnes de collecte des échantillons. Elle sera aussi mise à jour à chaque fois qu'il y aura de nouvelles données disponibles sur les échantillons qui sont déjà stockés. Le schéma 1 récapitule les diverses méthodes possibles de sélection des échantillons.

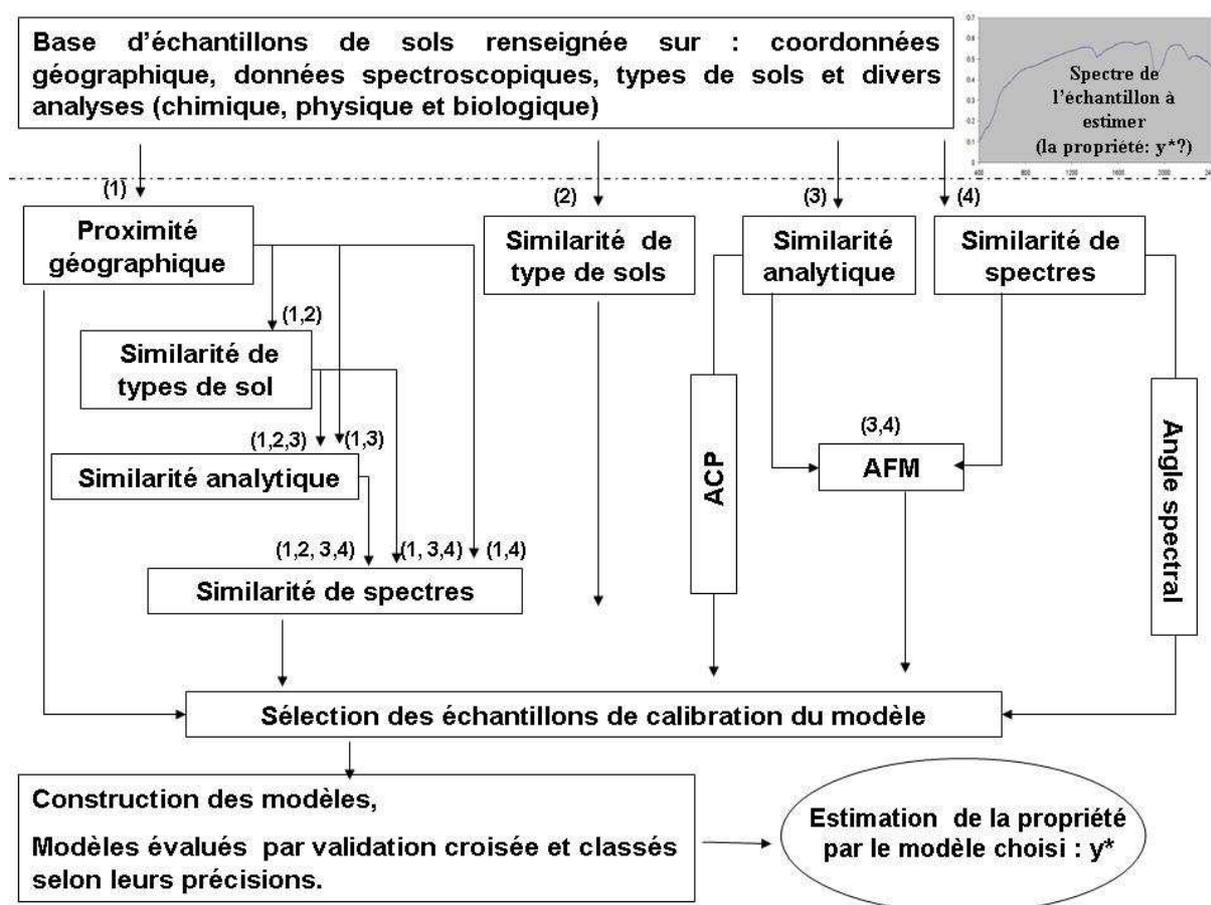


Figure 1. Diagramme synoptique pour l'estimation ponctuelle de différentes propriétés du sol à partir de mesures spectrométriques et d'une base de données (pedo-spectrale et analytique) liée à un système d'information géographique. (ACP : Analyse en composantes principales, AFM : Analyse factorielle multiple).

La sélection des échantillons de calibration du modèle d'estimation d'une propriété « P » pour un échantillon « e » peut se faire selon l'une des méthodes suivantes :

- *la proximité géographique* : dans notre base de données nous allons sélectionner pour calibrer le modèle, les échantillons qui sont spatialement proches du point de collecte de l'échantillon e . Selon le cas, différentes distances seuils peuvent être fixées pour décider quant aux échantillons à intégrer dans le jeu de calibration du modèle ;

- *le type de sol* : la base peut être interrogée, pour identifier les échantillons ayant une similarité de type de sol avec l'échantillon objet de l'estimation. Le modèle de calibration va être construit uniquement sur ces échantillons.

- *la similarité analytique* : dans le cas où nous disposons de données sur certaines propriétés analytiques concernant l'échantillon e objet de l'estimation pour sa propriété P , nous pouvons sélectionner dans notre base de données les échantillons ayant une similarité analytique avec l'échantillon objet de l'estimation. Pour cela, il faut projeter l'échantillon objet de l'estimation en tant qu'individu inactif dans le premier plan factoriel d'une Analyse en Composantes Principales (ACP) appliquée sur les variables analytiques de la base. Cet échantillon ne va pas participer à la formation des axes de l'ACP, mais va être uniquement projeté sur ces axes afin de repérer les échantillons qui lui sont proches dans ce plan. Ces derniers vont permettre de calibrer le modèle.

- *la similarité spectrale* : nous pouvons sélectionner dans la base de données, les échantillons ayant une similarité spectrale avec l'échantillon objet de l'estimation. Ceci est possible, entre autres, en calculant un angle spectral, tel que défini par Kruse et *al.* (1993), un angle seuil peut être défini et utilisé comme test d'acceptation ou de rejet des échantillons.

Il faut aussi signaler qu'en fonction de la taille de la base de données et de la nature des données qui y sont stockées, l'ensemble de ces différentes méthodes ou certaines d'entre elles peuvent être appliquées en cascade. Ainsi, si la base est large nous pouvons commencer par présélectionner les échantillons qui sont spatialement proches de notre échantillon objet de l'estimation. Puis, au sein de cette sous-base sélectionner les échantillons ayant une similarité de type de sol avec l'échantillon objet de l'estimation. Ensuite, étudier la similarité analytique qui existe entre cette série d'échantillons et l'échantillon objet de l'estimation afin de sélectionner les échantillons qui lui sont analytiquement similaires. Une dernière sélection va être faite en considérant la similarité spectrale. Au final, les échantillons

de la base qui auront répondu successivement à ces différentes exigences vont être utilisés pour calibrer le modèle.

Nous pouvons imaginer aussi, faire abstraction de la proximité géographique et de la similarité des types de sols. En effet, ces données peuvent faire défaut ou l'effectif des échantillons contenus dans la base de données qui répondraient à ces critères de choix n'est pas important. Nous pouvons alors combiner la sélection en se basant sur la similarité analytique et spectrale à l'aide d'une Analyse Factorielle Multiple (AFM). Ce genre d'analyse permet de projeter les échantillons de la base selon leur proximité analytique et spectrale d'une façon simultanée. A chaque modèle de prédiction est associée une performance de prédiction, déterminée en validation croisée (R^2 , REQM, REP). Les modèles calibrés selon ces différentes méthodes vont être classés selon leurs performances de prédiction en validation croisée. Le modèle qui est le plus performant va être utilisé pour l'estimation.

3.2. Cartographie thématique

La méthode que nous avons mise en œuvre pour spatialiser les propriétés du sol en se basant sur les données radiométriques continues d'une sub-image multispectrale ASTER et des spectres d'un jeu d'échantillons de sols mesurés en conditions contrôlées de laboratoire est facile à mettre en œuvre. Elle permet une cartographie à haute résolution spatiale et sur de large territoire. Cette méthode fait émerger des perspectives intéressantes pour la cartographie des propriétés du sol sur de larges territoires. Toutefois, elle a été testée dans un contexte pédologique bien précis : milieu aride, sol nu assez homogène. Elle doit donc être expérimentée dans le cas d'une couverture pédologique beaucoup plus variable, pour d'autres propriétés pédologiques, ainsi que dans le cas d'autres milieux notamment en milieu tempéré.

La spatialisation des propriétés du sol par les techniques géostatistiques telles que le krigeage ordinaire, se base sur un nombre limité de mesures directes. Ceci offre l'avantage de la précision des mesures au niveau des points observés. Toutefois, l'inconvénient majeur de cette méthode de spatialisation est la faible densité d'échantillonnage. En perspective d'une amélioration de la cartographie, les mesures spectrométriques peuvent être intégrées dans les démarches de spatialisation. La figure 2 récapitule les diverses méthodes envisageables.

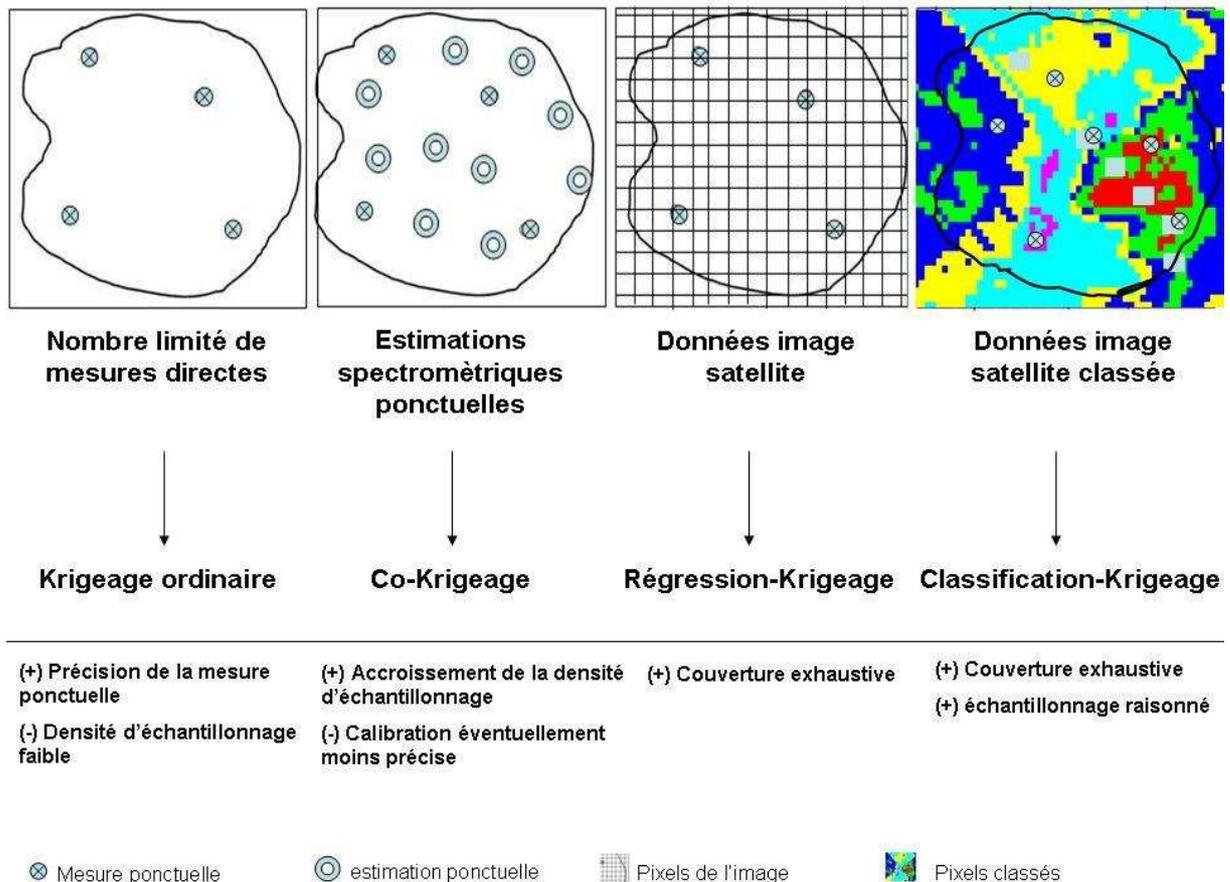


Figure 2. Diverses méthodes d'intégration des mesures spectrométriques dans la perspective d'amélioration de la qualité de la spatialisation des propriétés du sol.

3.2.1. Co-krigeage

Les estimations spectrométriques sont dans ce cas considérées comme des variables auxiliaires. Pour cela, nous pouvons utiliser les points pour lesquels la propriété a été mesurée pour calibrer le modèle spectrométrique. Ce dernier va servir pour estimer la propriété en différents points supplémentaires de l'espace à cartographier. Ces estimations vont permettre d'augmenter la densité d'échantillonnage et elles vont être utilisées comme variable auxiliaire pour une cartographie par co-krigeage. L'inconvénient de cette méthode serait le fait que la précision de la carte produite est tributaire de la performance du modèle spectrométrique.

3.2.2 Régression-krigeage

Comme nous l'avons testé (Cf. chapitre 5), il est possible de coupler les données spectrométrique de laboratoire, ou encore les mesures spectrométriques directes de terrain, avec les données spectrométriques spatialement continues extraites d'une image satellite.

3.2.3. Classification-krigeage (régression-krigeage par classes radiométriques)

La classification radiométrique de l'image permet en premier lieu de raisonner l'échantillonnage sur le terrain. Un modèle spectrométrique peut être calibré pour chaque classe radiométrique. Pour cela il faut considérer les échantillons collectés dans l'espace couvert par chacune de ces classes à part. Des régressions-krigeage par classes radiométriques vont être appliquées. Cette même classification peut être affinée en utilisant les données issues des modèles numériques de terrain.

3.3. Surveillance des sols

La spectroscopie présente un vrai potentiel pour être utilisée dans la surveillance diachronique des propriétés du sol. De simples mesures spectrométriques in situ acquises au niveau des mêmes points d'échantillonnage à différentes dates, peuvent permettre d'estimer la valeur d'une même propriété au cours du temps. Ceci permet d'étudier les variations temporelles et de dresser des bilans. Cette surveillance peut aussi être spatiale : une chrono-séquence d'acquisition d'images satellites et de mesures spectrométriques de laboratoire, sur des échantillons de sols collectés aux mêmes endroits à des dates différentes, permettent de générer des cartes de spatialisation des propriétés du sol à différentes dates.

Bibliographie générale

Abdi H 2003 *Partial Least Squares (PLS) Regression*. In M. Lewis-Beck, A. Bryman, T. Futing (Eds) : Encyclopedia for Research methods for the Social Sciences. Thousand Oaks (CA) : Sage.

Abrams M 1997 The Advanced Spaceborne Thermal Emission and Reflection Radiometer ASTER : data products for the high spatial resolution imager on NASA's Terra platform. *Int. J. Remote Sensing* 2, 847-859.

AFNOR 1996 *Qualité des sols, recueil de normes françaises* 534 p.

Al Abbas AH, Swain PH and Baumgardner MF 1972 Relating organic matter and clay content to the multispectral radiance of soils. *Soil Sci.* 114, 477-485.

Atlas du gouvernorat de Tozeur 1996. Direction générale de l'aménagement de territoires, Ministère de l'environnement et d'aménagement de territoire, Tunisie 57p.

Baize D et Girard MC (coord) 1995 *Référentiel pédologique*. INRA Editions 332 p.

Baize D et Girard MC (coord) 2009 *Référentiel pédologique 2008*. Editions Quae, Versailles 405 p.

Bannari A, El-Harti A, Haboudane D, Bachaoui E, El-Ghmari A 2007 Intégration des variables spectrales et géomorphométriques dans un SIG pour la cartographie des aires exposées à l'érosion. *Revue Télédétection* 7 (1-2-3-4), 327-342.

Banwell CN 1983 *Fundamentals of Molecular Spectroscopy*. McGraw-Hill, New-York 338 p.

Baptista GM and Netto MJS 2001 RCGb index : A tool for mapping the degree of weathering in the tropical soils in Brazil. In OG Green (Ed.), *Proceeding of the Tenth JPL Airborne Earth Science Workshop* : JPL Publication.

Barbieri R, Stivaletta N, Marinangeli L and Ori GG 2006 Microbial signatures in sabkha evaporite deposits of Chott el Gharsa (Tunisia) and their astrobiological implications. *Planetary and Space Science* 54, 726-736

Barnes EM, Sudduth KA, Hummel JW, Lesch SM, Corwin DL, Yang C, Daughtry CST and Bausch WC 2003 Remote and ground-based sensor techniques to map soil properties. *Photogrammetric Engineering and Remote Sensing* 69 (6), 619-630.

Barnes RJ, Dhanoa MS and Lister SJ 1983 correction to description of standard normal variate (SNV) and De-Trend (DT) Transformations in practical spectroscopy with applications in food and beverage analysis. *Journal of Near Infrared Spectroscopy* 1, 185-186.

Barrett LR 2002 Spectrophotometric color measurement in situ in well drained sandy soils. *Geoderma* 108, 49-77.

Barthès BG, Brunet D, Ferrer H, Chotte JL and Feller C 2006 Determination of total carbon and nitrogen content in a range of tropical soils using near infrared spectroscopy: influence of replication and sample grinding and drying. *J. Near Infrared Spectrosc.* 14, 341-348.

Bauer ME, Vanderbilt VC, Robinson BF, Daughtry CST 1980 Spectral properties of agricultural crops and soils measured from space, aerial, field and laboratory sensors, Proc. 14th ISPRS Congress, Hamburg (FRG) 23(B-7), 56-73.

Baumgardner MF, Kristof SJ, Johannsen CJ and Zachary AL 1970 The effects of organic matter on multispectral properties of soils. *Proc. Indiana Acad. Sci.* 79, 413-422.

Baumgardner MF, Silva LF and Stoner ER 1985 Reflectance properties of soils. *Advances in Agronomy* 38, 1-45.

Batchily AK, Post DF, Breckenfeld DJ 2003 Spectral reflectance and soil morphology characteristics of Santa Rita experimental range soils. *USDA Forest Service Proceedings RMRS-P* 30, 175-182.

Batten GD 1998 Plant analysis using near infrared reflectance spectroscopy : the potential and the limitations. *Aust. J. Exp. Agric.* 38, 697-706.

Baxter SJ and Oliver MA 2005 The spatial prediction of soil mineral N and potentially available N using elevation. *Geoderma* 128, 325-339.

Bedidi A, Cervelle B, Madeira J and Pouget M 1992 Moisture effects on visible spectral characteristics of lateritic soils. *Soil Science Am. J.* 153(2), 2129-2141.

Ben-Dor E 2002 Quantitative remote sensing of soil properties. *Advances in Agronomy* 75.

Ben-Dor E and Banin A 1990 Near-infrared reflectance analysis of carbonate concentration in soils, *Applied Spectroscopy* 44(6), 1064-1069.

Ben-Dor E and Banin A 1994 Visible and near-infrared (0.4-1.1 μ m) analysis of arid and semiarid soils, *Remote Sensing of Environment* 48 (3), 261-274.

Ben-Dor E and Banin A 1995a Near-infrared analysis (NIRA) as a method to simultaneously evaluate spectral featureless constituents in soils. *Soil Sci.* 159(4), 259-270.

Ben-Dor E and Banin A 1995b Near-infrared analysis as a rapid method to simultaneously evaluate several soil properties. *Soil Sci. Soc. Am. J.* 59(2), 364-372.

Ben-Dor E and Irons JR and Epema GF 1999 Soil reflectance. In Rencz AN (ed.) *Remote Sensing for the Earth Sciences. Manual of Remote Sensing. Vol. 3. 3rd Ed.* John Wiley & Sons, New York, Chichester, Weinheim, Brisbane, Singapore pp. 111-189.

Ben-Dor E, Inbar Y and Chen Y 1997 The reflectance spectra of organic matter in the visible near infra red and short wave infrared region (400-2500 nm) during a controlled decomposition process. *Remote sensing environment* 61, 1-15.

Ben-Dor E and Levin N 2000 Determination of surface reflectance from raw hyperspectral data without simultaneous ground data measurements : A case study of the GER 63-channel sensor data acquired over Naan, Israel. *International Journal of Remote Sensing* 21(10), 2053-2074.

Ben-Dor E, Levin TN, Singer A, Karnieli A, Braund O and Kidron GJ 2006 Quantitative mapping of the soil rubification process on sand dunes using an airborne hyperspectral sensor. *Geoderma* 131, 1-21.

Ben Hassine, Jelassi K, Harrathi J, Aïchi H et Marzouki M 2002 Suivi de l'évolution de la salinité des sols et de la nappe dans le périmètre irrigué de Chamsa (Tozeur). Ministère de l'agriculture de l'environnement et des ressources hydrauliques, Direction générale de l'aménagement et de la conservation des terres agricoles. Etude spéciale ES 322. 27p.

Bertrand D 2002 La spectroscopie proche infrarouge et ses applications dans les industries de l'alimentation animale. INRA Prod. Anim. 15, 209-219.

Berzaghi P, Shenk JS and Westerhaus MO 2000 Local prediction with near infrared multi-product databases. J.Near Infrared Spectrosc. 8, 1-9.

Bilgili AV, Van Es HM, Akbas F, Akis R, Hively WD 2007 Integration of Visible-Near Infrared Reflectance Spectroscopy and (Co) Kriging for Estimations of Soil Variables. Pedometrics 2007. Biannual Conference of Commission 1.5 Pedometrics, Division 1 of the International Union of Soil Sciences (IUSS). August 27 to 30, 2007. p27.

Bishop T and McBratney A 2001 A comparison of prediction methods for the creation of field-extent soil property maps. Geoderma 103 (1-2),149-160.

Blavet D, Mathe E and Leprun JC 2000 Relations between soil colour and waterlogging Duration in a representative hillside of the West African granito-gneissic bedrock. Catena 39, 187-210.

Bonn F 1994 Télédétection de l'environnement dans l'espace francophone. Presse de l'université du Québec. La télédétection appliquée aux sols. In Bonn F, précis de télédétection, PUQ/ AUPELF, Québec 2, 92-136.

Bonn F et Escadafal R 1996 La télédétection appliquée aux sols. In : précis de télédétection vol 2: Application thématiques universités francophone-UREF. Presse de l'université du Québec / AUPELF, pp 91-136.

Bourennane H 1997 Etude des lois de distribution spatiale des sols de Petite Bauge. Application à la cartographie d'un horizon par couplage de méthodes morphométriques et géostatistiques. Thèse université Orléans. 299p.

Bourennane H and King D 2003 Using multiple external drifts to estimate a soil variable. Geoderma 114, 1-18.

Bourennane H, King D and 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.

Bowers SA and Hanks RJ 1965 Reflection of radiant energy from soils. *Soil Sci.* 100: 130-138.

Braudeau E et Hachicha M 1998 Guide de suivi de la salinité des eaux et du sol. *Sols de Tunisie, Bulletin de la Direction des Sols n°18, Tunis*, 240 p.

Brostaux Y 2002 Introduction à l'environnement de programmation statistique R. <http://cran.r-project.org/>.

Brown DJ, Brickelmeyer RS, and Miller PR 2005 Validation requirements for diffuse reflectance soil characterization models with a case study of VNIR soil C predictions in Montana. *Geoderma* 129: 251-267.

Brown DJ, Shepherd KD, Walsh MG, Mays MD and Reinsch TG 2006 Global soil characterization with VNIR diffuse reflectance spectroscopy. *Geoderma* 132, 273-290.

Brunet D, Barthès BG, Chotte JL and Feller C 2007 Determination of carbon and nitrogen contents in Alfisols, Oxisols and Ultisols from Africa and Brazil using NIRS analysis : Effect of samples grinding and set heterogeneity. *Geoderma* 139, 106-117.

Bryant RG 1999 Application of AVHRR to monitoring a climatically sensitive playa. Case study : Chott El Djerid, southern Tunisia. *Earth surface processes and landforms.* 24, 283-302. John Wiley & Sons, Ltd.

Bullard JE and White K 2002 Quantifying iron oxide coatings on dune sands using spectrometric measurements: An example from the Simpson-Strzelecki Desert, Australia. *Journal of Geophysical Research* 107(B6), 2125-2138.

Burrough PA 1995 Spatial data quality and error analysis issues : GIS Functions and environmental modelling. In *GIS and Environmental modelling : progress and research issues*, MF Goodchild et al (Eds). *Gis world Inc, Fort Collins*.

Cécillon L, Cassagne N, Czarnes S, Gros R, Vennetier M, Brun JJ 2008 Variable selection in near infrared spectra for the biological characterization of soil and earthworm casts. *Soil biology and biochemistry* 40, 1975-1979.

Chaminade P, Baillet A and Ferrier D 1998 Data treatment in near infrared spectroscopy. *Analusius Magazine* 26 (4), 33-38.

Chang CW and Laird DA 2002 Near-infrared réflectance spectroscopic analysis of soil C and N. *Journal of Soil Science* 167(2), 110-116.

Chang CW, Laird DA, Mausbach MJ and Hurburgh CR 2001 Near-infrared réflectance spectroscopy – principal components regression analyses of soil properties. *Soil Sci. Soc. Am. J.* 65, 480-490.

Carré 2002 Cartogénèse des sols et changements d'échelle, application dans la région de La Rochelle sur une base de données pédologiques de plusieurs milliers d'observations. Thèse de l'IINAP-G 367p.

Carré F and Girard MC 2002 Quantitative mapping of soil types based on regression kriging of taxonomic distances with landform and land cover attributes. *Geoderma* 110, 241-263.

Chabrilat S, Goetz A, Krosley L and Olsen H 2002 Use of hyperspectral images in the identification and mapping of expansive clay soils and the role of spatial resolution, *Remote sensing of Environment* 82, 431-445.

Chen F, Kissel DE, West LT and Wayne A 2000 Field-scale mapping of surface soil organic carbon using remotely sensed imagery. *Soil Sci. Soc. Am. J.* 64, 746-753.

Chikhaoui M, Bonn F et Merzouk A 2004 Cartographie de la fraction argileuse du sol dans le Rif marocain à l'aide du capteur ASTER et de l'analyse géostatistique. *Revue internationale de Géomatique.* 14 (3-4), 359-377.

Clark DH, Mayland HF and Lamb RC 1987 Mineral analysis of forage with near infrared reflectance spectroscopy. *Agronomy Journal* 79, 485-490.

Clark RN 1999 Spectroscopy of rocks and minerals, and principles of spectroscopy. In : Rencz N (Ed.), *remote sensing for the earth sciences: Manual of remote sensing*. 3^{ed} Jhon Wiley and Sons, New York, p. 3-52.

Clark RN and Roush TL 1984 Reflectance spectroscopy : Quantitative analysis techniques for remote sensing applications : *Journal of Geophysical Research* 89 (7), 6329-6340.

Clark RN and Swayze GA 1995 Mapping minerals, amorphous materials, environmental materials, vegetation, water, ice, and snow, and other materials: the USGS Tricoder Algorithm: in Summaries of the fifth Annual JPL Airborne Earth Science Workshop, JPL Publication 95 (1), 39-40.

Clark RN, King TVV, Klejwa M and Swaze GA 1990 High spectral resolution reflectance spectroscopy of minerals. *Journal of Geophysical Research*. 95: 12653-12680.

Condit HR 1972 Application of characteristic vector analysis to the spectral energy distribution of daylight and the spectral reflectance of American soils, *Appl. Opt.* 11, 74-86.

Conel JE, Vane G, Green RO, Alley RE, Carere V, Gabell A, Bruegge CJ 1988 Airborne visible / infrared imaging spectrometer (aviris) : inflight radiometric calibration and the determination of surface reflectance. In : *Proceedings of the 4th Int. Coll. on Spectral Signatures of Objects in Remote Sensing*. 18-22 January, Aussois, France, 293-297.

Confalonieri M, Fornasier F, Ursino A, Boccardi F, Pintus B and Odoardi M 2001 The potential of near infrared reflectance spectroscopy as a tool for the chemical characterization of agricultural soils. *J. Near Infrared Spectrosc.* 9, 123-131.

Coque R 1962 *La Tunisie présaharienne, etude géomorphologique*. Edition Masson Paris 448

Coudé-Gaussen G 1990 Les régions sources de poussières au Sahara. *Sécheresse* 1(2), 134-141.

Couillard A, Turgeon AJ, Westerhaus MO and Shenk JS 1997 Determination of soil separates with near infrared reflectance spectroscopy. *J. Near Infrared Spectrosc.* 4, 201-212.

Coûteaux MM, Berg B and Rovira P 2003 Near infrared reflectance spectroscopy for determination of organic matter fractions including microbial biomass in coniferous forest soils. *Soil Biology and Biochemistry* 35, 1587-1600.

Cozzolino D and Moron A 2006 Potential of near infrared reflectance spectroscopy and chemometrics to predict soil organic carbon fractions. *Soil and tillage research* 85, 78-85.

Crouvi O, Ben-Dor E, Beyth M, Avigad D, Amit R 2006 Quantitative mapping of arid alluvial fan surfaces using field spectrometer and hyperspectral remote sensing *Remote Sensing of Environment* 104,103-117.

Dalal RC and Henry RJ 1986 Simultaneous determination of moisture, organic carbon, and total nitrogen by near infrared reflectance spectrophotometry. *Soil Sci. Soc. Am. J.* 50, 120-123.

Dardenne P, Sinnaeve G and Baeten V 2000 Multivariate calibration and chemometrics for near infrared spectroscopy : which method ?. *Journal of Near Infrared spectroscopy* 8 (4), 229-237.

Davies AMC 2001 Uncertainty testing in PLS regression. *Spectroscopy Europe* 13(2), 16-19.

Davies AMC, Dennis C, Grant A, Hall MN and Robertson A 1987 Screening of tomato puree for excessive mould content by near infrared spectroscopy : a preliminary evaluation. *J. Sci. Food Agric.* 39, 349-355.

Demangeot J and Bernus E 2001 *Les milieux désertiques*. Armand Colin 295 p.

Demattê JAM 2002 Characterization and discrimination of soils by their reflected electromagnetic energy. *Pesq. Agropec .bras. Brasilia* 37(10), 1445-1458.

Demattê JAM, Campos RC, Alves MC, Fiorio PR, Nanni MR 2004 Visible-NIR reflectance: a new approach on soil evaluation. *Geoderma* 121, 95-112.

Demetriades-Shah TH, Steven MD, and Clark JA 1990 High resolution derivative spectra in remote sensing. *Rem.Sens.Env.* 33, 55-64.

Desbois D 1999 Introduction à la régression sur variables latentes : la procédure PLS de SAS. *Le cahier des techniques de l'INRA* 41 : 81p.

Deshayes Michel et Maurel P 1990 L'image spectrale et son contenu. *Option mediterraneenes, série A n°4. La télédétection en a griculture.* p11-20.

Di L and Rundquist DC 1994 A one step algorithm for correction and calibration of AVHRR level 1b data. *Photogrammetric Engineering and remote sensing* 60 (2), 165-171.

Douaoui A 2005 Variabilité Spatiale de la salinité et sa relation avec certaines caractéristiques des sols de la plaine du Bas-Chélif. Apport de la géostatistique et de la télédétection. Thèse Doct. d'état, INA Alger. 233p.

Douaoui A, Nicolas H, Walter C 2006 Detecting salinity hazards within a semiarid context by means of combining soil and remote-sensing data. *Geoderma* 134, 217-230.

Drake NA 1997 Recent Aeolian origin of surficial gypsum crusts in southern Tunisia: geomorphological, archeological and remote sensing evidence. *Earth Surface Processes and Landforms*, 22, pp. 641-656.

Druilhet P and Mom A 2006 PLS regression: a signal to noise ratio approach", *Journal of Multivariate Analysis* 97, 1313-1329.

Duckworth JH 1998 Spectroscopic Quantitative Analysis. In : Workman, J. and Springsteen, A. (Eds.), *Applied Spectroscopy – A Compact Reference for Practitioners*, Academic Press, Chestnut Hill, MA, pp. 93-107.

Dunn BW, Beecher HG, Batten GD, Ciavarella S 2002 The potential of near infrared reflectance spectroscopy for soil analysis – a case study from the Riverine Plain of south-eastern Australian. *Australian Journal of Experimental Agriculture*, 42: 607-614.

El Fekih M 1965 Etude pédologique des oasis continentale du Djerid (Sedada, Degache, Tozeur). *Direction des Sols. publ. N°ES 289*. 138 p. Tunis.

El Qayedy J, Taj-Eddine K, Bonn F, Chikhaoui M et Witam O 2006 Caractérisation lithologique du haut Atlas Marocain à l'aide des données d'ASTER et de mesures spectrales de terrain. *Télédétection. Contemporary publishing international*. 6 (2), 153-175. Éditions Scientifiques GB

Elvidge CD 1990 Visible and near infrared reflectance characteristics of dry plant materials. *Int J. Remote Sens.* 11, 1775-1795.

Ehsani MR, Upadhayaya SK, Slaughter D, Shafii S and Pelletier M 1999 A NIR Technique for Rapid Determination of Soil Mineral Nitrogen. *Precision Agriculture* 1, 217-234 Kluwer Academic Publishers. Manufactured in The Netherlands.

Escadafal R 1989 Caractérisation de la surface des sols arides par observations de terrain et par télédétection. Thèse de Doctorat, université Pierre et Marie curie, Etudes et thèses, ORSTOM, Paris, 317p.

Escadafal R, Gouinaud C, Mathieu R, Pouget M 1993 Le spectroradiomètre de terrain : un outil de la télédétection et de la pédologie. *Cah. Orstom sér. pédol.* 28(1), 15-29.

ESRI 2006 ArcGis 9.2 tutorials.

Farifteh J, Farshad A, George RJ 2006 Assessing salt-affected soils using remote sensing, solute modelling, and geophysics *Geoderma* 130, 191-206.

Farifteh J, Van der Meer F, Atzberger C, Carranza EJM 2007 Quantitative analysis of salt-affected soil reflectance spectra : A comparison of two adaptive methods (PLSR and ANN) *Remote Sensing of Environment* 110, 59-78.

Follain S 2005 Effet du réseau bocager sur l'organisation des sols. Redistribution des sols et stockage en carbone organique. Thèse de l'Université ENSA de Rennes. 241p.

Fourty T et Baret F 1998 Restitution du contenu en matière organique et en eau d'un sol à partir de ses caractéristiques spectrales : Développement de modèles empiriques et semi-analytiques. Projet PAAGE. INRA, Avignon. 45 p.

Finke PA, Brus DJ, Bierkens MFP, Hoogland T, Knotters M and De Vries F 2004 Mapping groundwater dynamics using multiple sources of exhaustive high resolution data. *Geoderma* 123 (1-2), 23-39.

Friedman JH 2001 Greedy function approximation : a gradient boosting machine. *Ann. Stat.* 29 (5), 1189-1232.

Fujisida H, Sakuma F, Ono A and Kudho M 1998 Design and pre-flight performance of ASTER instrument protoflight model. *IEEE Transactions on Geoscience and Remote Sensing* 36, 1152-1160.

Gaffey SJ 1986 Spectral reflectance of carbonate minerals in the visible and near infrared (0.35–2.55 microns): Calcite, aragonite, and dolomite. *Am. Mineral.* 71, 151-162.

Gaffey SJ, McFadden LA, Nash D and Pieters CM 1993 Ultraviolet, visible, and near-infrared reflectance spectroscopy : Laboratory spectra of geologic materials. pp 43-73. In CM Pieters and PAJ Englert (ed.) *Remote geochemical analysis elemental and mineralogical composition*. University of Cambridge Press, Cambridge.

Galvão LS, Pizarro MA and Epiphanyo JCN 2001 Variations in reflectance of tropical soils: spectral chemical composition relationships from AVIRIS data. *Remote sensing environment* 75, 245-255.

Gauch HG, Hwang JTG and Fick GW 2003 Model evaluation by comparison of model-based predictions and measured values. *Agron. J.* 95 (6), 1442-1446.

Geladi P 1995 An overview of multivariate spectral data analysis. In : Davies AMC and Williams P (eds). *Near Infrared spectroscopy : The future waves*. NIR publications, Chichester, UK.

Genot V, Gilles C, Dardenne P, Bah B, Legrain X et Bock L 2007 Etude de la performance de la spectroscopie proche infrarouge pour la mesure de caractéristiques du sol nécessaires au conseil de fumure. Actes des 9^{ème} journées nationales de l'étude des sols. Angers 3-5 Avril 2007 J.P. Rossignol (ed) Angers.

Gherina SA, Douaoui A, Hartani T, Vincent B 2009 Application de la télédétection dans la connaissance de la salinité des sols de la plaine du Bas-Chéliff. *Economie d'eau en systèmes irrigués au Maghreb*. Actes du quatrième atelier régional du projet Sirma. Cirad, Montpellier, France.

Girard MC et Girard CM 1999 *Traitement de donnée de télédétection*. DUNOD, Paris, 529 p.

Goetz AF, Chabrilat S and Lu Z 2001 Field reflectance spectroscopy for detection of swelling clays at construction sites. *Field Anal. Chem. Technol.* 5,143-155.

Gomez C et Lagacherie P 2007 Spatialisation de propriétés du sol par Spectroscopie Visible et Proche Infrarouge. Résumé de Communication orale. 6^{ème} rencontres HélioSPIR. Colloque

« La spectroscopie proche infrarouge en Sciences du Sol » 25 et 26 Octobre 2007, Grenoble (France).

Goovaerts P 1997 Geostatistics for Natural Resources Evaluation. Oxford Univ. Press, New York.

Green AA and Craig MD 1985 Analysis of aircraft spectrometer data with logarithmic residuals : in proceedings, AIS workshop, 8-10 April, 1985, JPL Publication., 85-41, Jet propulsion laboratory, Pasadena, California. p111-119.

Grove CI, Hook SJ and Paylor ED 1992 Laboratory reflectance spectra for 160 minerals 0.4-2.5 micrometers : JPL Publication 92-2, Jet Propulsion Laboratory, Pasadena, CA.

Harris D and Bertolucci M 1978 Symmetry and Spectroscopy, Dover, New York,

He Y, Huang M, Garcia A, Hernandez A and Song H 2007 Prediction of soil macronutrients content using near-infrared spectroscopy. Computers and Electronics in Agriculture 58, 144-153.

He Y, Song H, Pereira AG and Gomez AH 2005 Measurement and analysis of soil nitrogen and organic matter content using near-infrared spectroscopy techniques. Journal of Zhejiang University Sci. 6B(11), 1081-1086.

Henderson TL, Baumgardner MF, Franzmeier DP, Scott DE and Coster DC 1992 High dimensional reflectance analysis of soil organic matter. Soil Sci. Soc. Am. J. 56, 865-872.

Henderson TL, Szilagyi A, Baumgardner MF, Chen CT and Landgrebe DA 1989 Spectral band selection for classification of soil organic matter content. Soil Science Society of America Journal 53, 1778-1784.

Hengl T, Heuvelink GBM, Stein A 2004 A generic framework for spatial prediction of soil variables based on regression-kriging. Geoderma. 120, 75-93.

Hengl T, Heuvelink GBM, Rossiter DG 2007 About regression-kriging : From equations to case studies. Comp. Geosci. 33,1301-1315.

Herbst M, Diekkrüger B and Vereecken H 2006 Geostatistical co-regionalization of soil hydraulic properties in a micro-scale catchment using terrain attributes. *Geoderma* 132, 206-221.

Hill J and Schütt B 2000 Mapping complex patterns of erosion and stability in dry Mediterranean ecosystems. *Remote Sensing of Environment* 74, 557-569.

Hollas JM 1996 *Modern Spectroscopy*, John Wiley & Sons, Chichester, 3^e éd.

Homayouni S 2005 *Caractérisation des scènes urbaines par analyse des images hyperspectrales*. Thèse de l'Ecole Nationale Supérieure des Télécommunications de Paris. 174p.

Houssa R, Pion JC and Yéssou H 1996 Effects of granulometric and mineralogical composition on spectral reflectance of soils in a Sahelian area. *ISPRS Journal of photogrammetry and remote sensing* 51, 284-298.

Howari FM, Goodell PC and Miyamoto S 2002 Spectral Properties of Salt Crusts Formed on Saline Soils. *J. Environ. Qual.* 31, 1453-1461.

Hunt GR and Salisbury JW 1970 Visible and near-infrared spectra of minerals and rocks : I. Silicate minerals. *Modern Geol.* 1, 283-300.

Hunt GR and Salisbury JW 1976 Visible and near infrared spectra of minerals and rocks XII: Metamorphic rocks. *Modern Geology* 5, 219-228.

Hunt GR 1977 Spectral signatures of particulate minerals, in the visible and near-infrared, *Geophysics* 42 (3), 501-513.

Hunt GR, Salisbury JW and Lenhoff CJ 1971 Visible and near-infrared spectra of minerals and rocks: III. Oxides and hydroxides. *Modern Geol.* 2, 195–205.

Hummel JW, Sudduth KA, Hollinger SE 2001 Soil moisture and organic matter prediction of surface and subsurface soils using an NIR soil sensor. *Computers and Electronics in Agriculture.* 32, 149-165.

Ihaka R and Gentleman R 1996 R: A Language for Data Analysis and Graphics. *Journal of Computational and Graphical Statistics.* 5(3), 299-314.

Ingleby HR, Crowe TG 2000 Reflectance models for predicting organic carbon in Saskatchewan soils. *Canadian Agricultural engineering*. 42(2), 57-63.

Irons JR, Weismiller RA and Petersen GW 1989 Soil reflectance. In Asrsr G (Ed.) *Theory and applications of optical remote sensing*. John Wiley, New York. pp 66-106.

Islam K, Singh B, McBratney AB 2003 Simultaneous estimation of several soil properties by ultra-violet, visible and near infrared reflectance spectroscopy. *Australian Journal of Soil Research*. 41, 1101-1114.

Islam K, Singh B, Schwenke G and McBratney AB 2004 Evaluation of Vertosol soil fertility using ultra-violet, visible and near-infrared reflectance spectroscopy ; 3rd Australian New Zealand Soils Conference, 5-9 December 2004, University of Sydney, Australia. Published on CDRom. Website www.regional.org.au/au/asssi/

ITT 2007 ENVI 4.3 Software by Visual Information Solutions. [http : //www.itvis.com](http://www.itvis.com) (Consulté le 20 Mars 2007).

Janik LJ and Skjemstad JO 1995 Characterization and analysis of soils using midinfrared partial least-squares. II. Correlations with some laboratory data. *Aust. J. Soil. Res.* 33, 637-650.

Janik LJ, Merry RH and Skjemstad JO 1998 Can mid infrared diffuse reflectance analysis replace soil extractions ? *Aust. J. Exp. Agric.* 38, 681-696.

Jarmer T 2005 The use of reflectance spectroscopy and satellite data for detection of pedo-chemical properties in semi-arid and arid regions of Israel (in German). Ph.D Dissertation, Trier University. 182 pp.

Jarmer T and Schütt B 1998 Analysis of iron contents in carbonate bedrock by spectroradiometric detection based on experimentally designed substrates. In Schaeppmann M, Schläpfer D and Itten K (eds.) 1st EARSeL Workshop on Imaging Spectroscopy. Remote Sensing Laboratories, University of Zurich, Switzerland, 6-8 October 1998. EARSeL, Paris. pp. 375-382.

Jarmer T, Vohland M, Lilienthal H and Schnug E 2008 Estimation of some chemical properties of an agricultural soil by spectroradiometric measurements. *Pedosphere*. 18(2), 163-170.

Jaquemoud S, Baret F and Hanocq JF 1993 Modélisation de la réflectance spectrale et directionnelle des sols : application au concept de droite des sols. Cah. Orstom, sér. Pédol. 28(1), 31-43.

Job JO 1992 Les sols salés de l'oasis d'El Guettar (Sud tunisien) Thèse de doctorat 161p, Université Montpellier II- Sciences et Techniques du Languedoc.

Kariuki PC, Van Der Meer FD and Verhoef PNW 2003 Cation exchange capacity (CEC) determination from spectroscopy. International journal of remote sensing 24(1), 161-167.

Knotters M, Brus D and Voshaar J 1995 A comparison of kriging, co-kriging and kriging combined with regression for spatial interpolation of horizon depth with sensed observations. Geoderma 67 (3-4), 227-246.

Kamel S 2007 Caractérisation hydrodynamique des aquifères du Djérid (Sud-Ouest tunisien). Thèse de doctorat en géologie. Université de Tunis El Manar. 230 p.

Karnieli A, Kidron GJ, Glaesser C and Ben Dor E 1999 Spectral characteristics of cyanobacteria soil crust in semiarid environments. Remote Sensing Environment 69, 67-75.

Kasumov OK, Nabieva TA, Tereshenkov OM 1992 Study of the spectral characteristics of soils of the Azerbaijan SSR, Sov. J. Remote Sens. 9(4), 608-617.

Kemper T and Sommer S 2002 Estimate of heavy metal contamination in soils after a mining accident using reflectance spectroscopy. Environmental Science and technology 36(12), 2742-2747.

Krishnan P, Alexander JD, Butler BJ, Hummel JW 1980 Reflectance technique for predicting soil organic matter. Soil Science Society. Am. J. 44, 1282-1285.

Kruse FA, Calvin WM, and Seznec O 1988 Automated extraction of absorption features from Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) and Geophysical Environmental Research Imaging Spectrometer (GERIS) data : In Proceedings of the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) performance evaluation workshop, JPL Publication, 88-38, p. 62-75.

Kruse FA, Lefkoff AB and Dietz JB 1993 Expert System-Based Mineral Mapping in northern Death Valley, California/Nevada using the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS): Remote Sensing of Environment, Special issue on AVIRIS, May-June 1993, 44, 309-336.

Kooistra L, Wanders J, Epema GF, Leuven RSEW, Wehrens R and Buydens LMC 2003 The potential of field spectroscopy for the assessment of sediment properties in river floodplains. *Analytica Chimica Acta* 484, 189-200.

Korb AR, Dybwad P, Wadsworth W and Salisbury JW 1996 Portable FTIR spectroradiomètre for field measurements of radiance and emissivity, *Applied Optics* 35, 1679-1692.

Kosmas CS, Curi N, Bryant RB and Franzmeier DP 1984 Characterisation of iron oxide minerals by second-derivative visible spectroscopy. *Soil Sci. Soc. Am. J.* 48, 401-405.

Lachenal G 2006 Introduction à la spectroscopie infrarouge. In: *La spectroscopie infrarouge et ses applications analytiques*. Bertrand D et Dufour E coordonnateur. Tec & Doc (Eds). France pp 35-75.

Lagacherie P and Voltz M 2000 Predicting soil properties over a region using sample information from a mapped reference area and digital elevation data : a conditional probability approach. *Geoderma* 97, 187-208.

Lagacherie P, Baret F, Feret JB, Netto JM and Robbez-Masson JM 2008 Estimation of soil clay and calcium carbonate using laboratory, field and airborne hyperspectral measurements. *Remote Sensing of Environment* 112, 825-835.

Lee SW, Sanchez JF, Mylavarapu RS and Choe JS 2003 Estimating chemical properties of Florida soils using spectral reflectance. *Trans. ASAE* 46, 1443-1453.

Lloyd CD 2005 Assessing the effect of integrating elevation data into the estimation of monthly precipitation in Great Britain. *Journal of Hydrology* 308(1-4), 128-150.

Leiffield J 2006 Application of diffuse reflectance FT-IR spectroscopy and partial least squares regression to predict NMR properties of soil organic matter. *British Society of Soil Science, European Journal of Soil Science* 57, 846–857.

Lemercier B, Walter C, Schwartz C, Saby N, Arrouays D and Follain S 2006 Suivi des teneurs en carbone organique et en phosphore extractible dans les sols de trois régions françaises. *Etude et Gestion des Sols* 13(3), 165-179.

Leone PA and Sommer S 2000 Multivariate analysis of laboratory spectra for the assesement of soil development and soil degradation in the southern Apennines (Italy). *Remote sensing of environnement* 72, 346-359.

Levin N, Kidron GJ and Ben-Dor E 2007 Surface properties of stabilizing coastal dunes : combining spectral and field analyses. *Sedimentology* 54, 771-788.

Lillesand TM and Kiefer RW 1987 Earth ressource satellites. Chapter 9 in *Remote sensing and image interpretation*. Second edition. Jhon Wiley, New York. 612p

Ludwig B, Khanna P, Bauhus J, Hopmans P 2002 Nearinfrared spectroscopy of forest soils to determine chemical and biological properties related to soil sustainability. *Forest Ecol. Manage.* 171, 121-132.

Luleva MI 2007 Identification of soil property variation using spectral and statistical analysis on field and ASTER data. A case study of Tunisia. Thesis 111p. International Institute for Geo-Information Science and Earth Observation, Enschede, The Netherlands.

Ludovic L, Morineau A et Piron M 1995 *Statistique exploratoire multidimensionnelle*. Ed Dunod, Paris, 437p.

Madeira NJS, Bédidi A, Cervelle B, Pouget M and Flay N, 1997 Visible Spectrometric Indices of Hematite (Hm) and Goethite (Gt) Content in Lateritic Soils : the application of a Thematic Mapper (TM) image for soil-mapping in Brasília, Brazil", *Int. J. Remote Sensing* 18, 2835-2852.

Mamou A 1976 *Contribution à l'étude hydrogéologique de la presqu'île de Kébili*. Thèse de troisième cycle. Université Pierre et Marie curie. Paris

Mamou A 1989 *Caractéristiques, évaluation et gestion des ressources en eau du sud tunisien*. Thèse doctorat Es sciences université de Paris sud. France. 426 p.

McCarty GW, Reeves VB, Follett RF and Kimble JM 2002 Mid-infrared and Near-infrared diffuse reflectance spectroscopy for soil carbon measurement. *Soil Sci. Soc. Am. J.* 66, 640-646.

Melis MT 2003 Campagna di misure radiometriche e osservazioni per l'interpretazione di immagini da satellite nella regione di Tozeur. Università degli Studi di Cagliari. 111p.

Mougenot B 1993 Effet des sels sur la réflectance et télédétection des sols sales. *Cah. Orstom sér. Pédol.* 28(1), 45-54.

Malley DF, Eilers RG, Rempel E, Van Wallegghem M and Martin PD 1999 Rapid analysis of nutrient and salt concentrations in hog manure by near infrared spectroscopy. P 633-641. In manure management 22-24 June 1999. Saskatoon, SK, Canada.

Malley DF, Martin PD and Ben-Dor E 2004 Application in analysis of soils. *Agron. Monogr.* 44, 729-785.

Mamou A 1976 Contribution à l'étude hydrogéologique de la presqu'île de Kébili. Thèse de troisième cycle. Université Pierre et Marie curie. Paris

Mamou A 1989 Caractéristiques, évaluation et gestion des ressources en eau du sud tunisien. Thèse doctorat Es sciences université de Paris sud. France. 426 p.

Martens H and Næs T 1989 *Multivariate Calibration*. John Wiley & Sons, Chichester, UK.

Martin PD, Malley DF, Manning G and Fuller L 2002 Determination of soil organic carbon and nitrogen at the field level using near-infrared spectroscopy. *Can. J. Soil Sci.* 82, 413-422.

Mathieu R, Pouget M, Cervelle B and Escadafal R 1998 Relationships between satellite-Based Radiometric Indices Simulated Using Laboratory Reflectance Data and Typic Soil Color of an Arid Environment. *Remote sensing of environment* 66, 17-28.

Mc Bratney AB, Mendonça Santos ML and Minasny B 2003 On digital soil mapping. *Geoderma* 117, 3-52

Mc Bratney AB, Minasny B, Cattle SR and Vervoort RW 2002 From pedotransfer functions to soil inference systems. *Geoderma* 109, 41-73.

Mc Carty GW, Reeves VB, Follett RF and Kimble JM 2002 Mid-infrared and near infrared diffuse reflectance spectroscopy for soil carbon measurement. *Soil Sci. Soc. Am. J.* 66, 640-646.

Méléder V, Launeau P, Barillé L and Rincé Y 2003 Cartographie des peuplements du microphytobenthos par télédétection spatiale visible-infrarouge dans un écosystème conchylicole. *Comptes Rendus Biologies* 326, 377-389.

Melis MT 2003 Campagna di misure radiometriche e osservazioni per l'interpretazione di immagini da satellite nella regione di Tozeur. Università degli Studi di Cagliari. 111p.

Misun K 2002 Quantification of soil organic carbon using Mid and Near-drift spectroscopy. Master of science thesis, 65p. Texas A and M University.

Morra MJ, Hall MH and Freeborn LL 1991 Carbon and Nitrogen Analysis of Soil Fractions Using Near-Infrared Reflectance Spectroscopy *Soil Sci. Soc. Am. J.* 55, 288-291.

Morris RV, Lauer HV, Lawson CA, Gibson EK Jr, Nace GA and Stewart C 1985 Spectral and other physicochemical properties of submicron powders of Hematite (α -Fe₂O₃), Maghematite (γ -Fe₂O₃), Magnetite (Fe₃O₄), Goethite (α -FeOOH), and Lepidocrocite (γ -FeOOH). *J. Geophys. Res.* 90, 3126-3144.

Mouazen AM, De Baerdemaeker J and Ramon H 2005 Towards development of on-line soil moisture content sensor using a fibre-type NIR spectrophotometer. *Soil & Tillage Research*, 80(1-2), 171-183.

Mouazen AM, Maleki MR, De Baerdemaeker J and Ramon H 2007 On-line measurement of some selected soil properties using a Vis-NIR sensor *Soil & Tillage Research* 93, 13-27.

Mougenot B 1993 Effet des sels sur la réflectance et télédétection des sols sales. *Cah. Orstom sér. Pédol.* 28 (1), 45-54.

Muller E and Décamps H 2000 Modeling soil moisture reflectance. *Remote sensing of environment* 76, 173 -180.

Murray I and Williams PC 1987 Chemical principles of near-infrared technology. p 17 in *Near-Infrared Technology in the Agricultural and Food Industries*. PC Williams and K Norris, ed. Am. Assoc. Cereal Chem., Inc., St. Paul, MN.

Mustard JF and JM Sunshine 1999 Spectral analysis for earth science : Investigations using remote sensing data. p. 251-306. In A.N. Rencz (ed.) Remote sensing for the Earth sciences : Manual of remote sensing. John Wiley & Sons, New York.

Mutuo PK, Shepherd KD, Albrecht A, Cadisch G 2006 Prediction of carbon mineralization rates from different soil physical fractions using diffuse reflectance spectroscopy. Soil biology and biochemistry 38, 1658-1664.

Obukhov AI and Orlov DS 1964 Spectral reflectivity of the major soil groups and possibility of using diffuse reflection in soil investigations. Sov. Soil Sci. 2, 174-184.

Odeh I, McBratney A and Chittleborough D 1994 Spatial prediction of soil properties from landform attributes derived from a digital elevation model. Geoderma 63 (3-4), 197-214.

Odeh I, McBratney A and Chittleborough D 1995 Further results on prediction of soil properties from terrain attributes : heterotopic cokriging and regression-kriging. Geoderma 67 (3-4), 215-226.

Olinger JM and Griffiths PR 1993 Effects of sample dilution and particle-size morphology on diffuse reflectance spectra of carbohydrate systems in the NIR and MIR. Applied Spectroscopy 47(6), 687-694.

Osborne BG, Fearn T and Hindle PH 1993 In : Practical NIR spectroscopy with applications in food and beverage analysis 2nd Ed. Longman Scientific and Technical.

Post DF, Fimbres A, Mathias AD, Sano EE, Accioly L, Batchily AK and Ferreira LG 2000 Predicting soil albedo from soil color and spectral reflectance data. Soil Science Society of America Journal. 64, 1027-1034.

Rabiaa MC et Zargouni F 1990 Cartographie du couvert sédimentaire dans le chott Jerid et ses environs (Sud tunisien) par télédétection Éd. AUPELF-UREF. John Libney Eurotext. Paris p 253-263.

Ramsey MS, Stefanov WL and Christtensen PR 1999 Monitoring world-wide land cover changes with ASTER : Preliminary results from the phoenix azlter site. In Proceeding of international Conference on Applied Geologic Remote Sensing, p 237-244, Vancouver BC.

Rao BR, Dwivedi RS, Venkataratnam L, Ravishankar T, Thammappa SS, Bhargawa GP, Singh AN 1991 Mapping the magnitude of sodicity in part of the Indo-Gangetic plains of Uttar Pradesh, Northern India using Landsat-TM data. *Int. J. Remote Sensing* 12 (3), 419-425.

Reeves JB, Mc Carty GW and Meisinger JJ 1999 Near infrared spectroscopic for the analysis of agricultural soils. *J Near infrared spectroscopy* 7(3), 179-193.

Richter N, Jarmer T, Chabrilat S, Oyonarte C, Hostert P and Kaufmann H 2009 Free Iron Oxide Determination in Mediterranean Soils using Diffuse Reflectance Spectroscopy. *Soil Sci. Soc. Am. J.* 73, 72-81.

Reeves JB, McCarty GW and Meisinger JJ 2000 Near infrared spectroscopy for the determination of biological activity in agricultural soils. *J Near infrared spectrosc.* 8 (3), 161-170.

Reeves III JB, McCarty GW and Reeves VB 2001 Mid-infrared diffuse reflectance spectroscopy for the quantitative analysis of agricultural soil. *J. Agric. Food Chem.* 49, 766-772.

Reeves III JB, McCarty GW and Mimmo T 2002 The potential of diffuse reflectance spectroscopy for the determination of carbon inventories in soils. *Environmental Pollution* 116, 277-284.

Rezend SB 1980 Geomorphology, mineralogy and genesis of four soils on genesis in southeastern Brazil. PhD Thesis. Purdue Univ. Microfilms. Ann Arbor, Mich.

Rumpel C, Janik LJ, Skjemstad JO and Kogel-Knabner I 2001 Quantification of carbon derived from lignite in soils using mid-infrared spectroscopy and partial least squares. *Organic Geochemistry* 32, 831-839.

Sanchez-Maranon M, Delgado G, Delgado R, Perez MM and Melgosa M 1997 Spectrometric and visual color measurements of disturbed and undisturbed soil samples. *Soil science* 160 (4), 291-303.

Savitzky A and Golay M 1964 Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry* 36, 1627-1639.

Sayes W, Mouazen AM and Ramon H 2005 Potential for on-site and on-line analysis of hog manure using visual and near-infrared reflectance spectroscopy. *Biosyst. Eng.* 91(4), 393-402.

Scheffer F and Schachtschabel P 2002 Textbook of Soil Science (in German). Spektrum Akademischer Verlag, Heidelberg. 593pp.

Scheinost AC and Schwertmann U 1999 Color identification of iron oxides and hydroxysulfates: Use and limitations. *Soil Sci. Soc. Am. J* 63, 1463-1471.

Schnitzer MHR, Schulten P Schuppli and Angers DA 1991 Extraction of organic matter from soils with water at high pressures and temperatures. *Soil. Sci. Am. J* 55, 102-108.

Schulze DG, Nagel JL, Van Scoyoc GE, Henderson TL, Baumgaedner MF and Scott DE 1993 Significance of organic matter in determining soil colors., In J.M. Bigham and E.J.Ciolkosz (ed.) soil color. SSSA spec. publ. 31, SSSA, Madison, WI, p. 71-90.

Selige T, Böhner J and Schmidhalter U 2006 High resolution topsoil mapping using hyperspectral image and field data in multivariate regression modeling procedures. *Geoderma* 136 (1-2), 235-244.

Seiler B, Kneubühler M, Wolfgramm B and Itten KI 2007 Quantitative Assessment of Soil Parameters in Western Tajikistan using a Soil Spectral Library Approach. In 10th Intl. Symposium on Physical Measurements and Spectral Signatures in Remote Sensing (eds Schaepman ME, Liang S, Groot NE and Kneubühler M), Intl. Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 36 (7/C50), 451-455.

Servant J 1975 Contribution à l'étude pédologique des terrains halomorphes. L'exemple des sols salés du Sud et du Sud-Ouest de la France. Thèse Doctorat. Es Sciences Naturelles, ENSA Montpellier, 194 p.

Shepherd KD and Walsh MG 2002 Development of reflectance spectral libraries for characterization of soil properties. *Soil Sci. Soc. Am. J.* 66, 988-998.

Shenk JS and Westerhaus MO 1991 Population Definition, Sample Selection, and Calibration Procedures for Near Infrared Reflectance Spectroscopy Crop Science Society of America Sci 31, 469-474.

Simbahan GC, Dobermann A, Goovaerts P, Ping J, Haddix ML 2006 Fine-resolution mapping of soil organic carbon based on multivariate secondary data. *Geoderma* 132, 471-489.

Shippert P 2004 Why use hyperspectral imagery. *Photogramm. Eng. Remote Sens.* 70, 377-380.

Sinnaeve G, Herman JL, Baeten V, Dardenne P and Frankinet M 2001 Performances of an on board diode array NIR instrument for the analysis of fresh grass. *Journée thématique AFMEX, Appareils embarqués de mesure de la biomasse, Rennes, France.*

Sorensen LK and Dalsgaard S 2005 Determination of Clay and Other Soil Properties by Near Infrared Spectroscopy *Soil Sci. Soc. Am. J.* 69, 159-167.

Springsteen A 1998 Reflectance Spectroscopy : An Overview of Classifications and Techniques. In : Workman, J. and Springsteen, A. (Eds.), *Applied Spectroscopy – A Compact Reference for Practitioners*, Academic Press, Chestnut Hill, MA, pp. 194-223.

Sudduth KA and Hummel JW 1993a Soil organic matter, CEC and moisture sensing with a portable NIR spectrophotometer. *Trans. ASAE* 36, 1571-1582.

Sudduth KA and Hummel JW 1993b Portable, near-infrared spectrophotometer for rapid soil analysis. *Transactions of the ASAE* 36 (1), 185-193.

Sudduth KA and Hummel JW 1996 Geographic operating range evaluation of a NIR soil sensor. *Trans. ASAE* 39 (5), 1599-1604.

Sundberg R 1999 Multivariate calibration, direct and indirect regression methodology. *Board of the Scandinavian Journal of Statistics* 26, 162-207.

Stark E and Luchter K 1986 Near-infrared analysis (NIRA) a technology for quantitative and qualitative analysis. *Applied Spectroscopy Reviews* 22 (4), 335-399.

Starr C, Morgan AG and Smith DB 1981 An evaluation of near-infrared reflectance analysis in some plant breeding programmes. *Journal of Agricultural Science* 97, 107-18.

Stenberg B, Jonsson A and Bôrjesson T 2002 Near infrared technology for soil analysis with implications for precision agriculture. P 279-284. In A.M.C Dvies and R.K. Cho (ed.)

Proceedings of the 10th International Conference on near infrared spectroscopy. NIR Publication, Chichester.

Stevens A, Wesemael B, Vandenschrick G, Touré S and Tychon B 2006 Detection of Carbon Stock Change in Agricultural Soils Using Spectroscopic Techniques. *Soil Sci. Soc. Am. J.* 70, 844-850.

Steward GC 2006 Diffuse reflectance spectroscopy for the characterization of calcareous glacial till soils. Montana State University. Thesis of Master of Land Resources and Environmental Sciences. 145 p.

Stewart JM 1998 Color and Solar Transmittance in Applied Spectroscopy : A Compact Reference for Practitioners. Workman, J. and Springsteen, A. (ed.), Academic Press, Massachusetts, pp. 344-382.

Swezey CS, 1996 Structural controls on Quaternary depocentres within the Chotts Trough region of Southern Tunisia. *J. African Earth Sci.* 22, 335-347.

Swezey CS 2003 The role of climate in the creation and destruction of continental stratigraphic records : an example from the northern margin of the Sahara Desert. In : Swezey CS, Climate Controls on Stratigraphy, SEPM Special Publication. 77, 207-225.

Stoner ER and Baumgardner MF 1981 Characteristic variations in reflectance of surface soils. *Soil Sci. Soc. Am. J.* 45, 1161-1165.

Stenberg B, Jonsson A and Bôrjesson T 2002 Near infrared technology for soil analysis with implications for precision agriculture. p 279-284. In AMC Davies and RK Cho (ed.) Proceedings of the 10th International Conference on Near infrared spectroscopy. NIR Publication, Chichester.

Tenenhaus M 1998 La régression PLS théorie et pratique. Edition technip, Paris, 251 pp.

Terhoeven-Urselmans T, Michel K, Helfrich M, Flessa H and Ludwig B 2006 Near-infrared spectroscopy can predict the composition of organic matter in soil and litter *J. Plant Nutr. Soil Sci.* 169, 168-174.

Thuriès L 2007 Prédiction des transformations d'apports organiques agro-industriels aux sols. Résumé de Communication orale. 6ème rencontres HélioSPIR. Colloque « La spectroscopie proche infrarouge en Sciences du Sol » 25 et 26 Octobre 2007, Grenoble (France).

Travis HW 2006 In situ characterization of soil properties using Visible Near infrared Diffuse Reflectance Spectroscopy. Texas A & M University. Thesis 75p.

Travis HW, Morgan CLS, Brown DJ and Hallmark CT 2007 In Situ Characterization of Soil Clay Content with Visible Near-Infrared Diffuse Reflectance Spectroscopy. Soil Sci. Soc. Am. J. 71, 389-396

Torrent J, Schewertmann U, Fechter H and Alferez F 1983 Quantitative relationships between soil color and hematite content. Soil science 136, 354-358.

Udelhoven T, Emmerling C and Jarmer T 2003 Quantitative analysis of soil chemical properties with diffuse reflectance spectrometry and partial least-square regression : A feasibility study. Plant and Soil 251, 319-329.

Umland JB and Bellama JM 1996 General Chemistry, 2nd edition. Brooks Cole Publishing Company, Pacific Grove, CA. pp 224-234.

USSL 1954 Diagnosis and improvement of saline and alkali soils. United States Salinity Laboratory Staff US Department of Agriculture, Handbook 60, U.S. Print. Office, Washington DC.

Vancolen S 2004 La régression PLS Diplôme Postgrade en Statistique. Groupe de Statistique Université de Neuchâtel Suisse, 33p.

Vaudour E et Berthier L 2007 Prédiction spatiale des teneurs en carbone organique des sols par couplage de télédétection et de spectroscopie visible-proche infrarouge 6èmes rencontres HélioSPIR Colloque « La spectroscopie proche infrarouge en Sciences du Sol » 25 et 26 Octobre 2007, Grenoble_France. p9.

Verma, KS, Saxena RK, Barthwal AK and Deshmukh SN 1994 Remote sensing technique for mapping salt affected soils. International Journal of remote sensing 15(9), 1901-1914.

Viscarra RA and McBratney AB 1998 Laboratory evaluation of a proximal sensing technique for simultaneous measurement of soil clay and water content *Geoderma* 85, 19-39

Viscarra RA, Walvoort DJJ, McBratney AB, Janik LJ and Skjemstad JO 2006a. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma* 131(1-2), 59-75.

Viscarra RA, McGlynn RN and McBratney AB 2006b Determining the composition of mineral-organic mixes using UV-vis-NIR diffuse reflectance spectroscopy. *Geoderma* 137, 70-82.

Viscarra RA 2007a An introduction to diffuse reflectance spectroscopy. SINO-GERMAN Summer School, Yichang, China. [http // : WWW.tinyurl.com](http://WWW.tinyurl.com).

Viscarra RA 2007b An introduction to chemometrics. SINO-GERMAN Summer School, Yichang, China. [http // : WWW.tinyurl.com](http://WWW.tinyurl.com).

Viscarra RA 2008a ParLeS : Software for chemometric analysis of spectroscopic data. *Chemometrics and Intelligent Laboratory Systems* 90, 72–83.

Viscarra RA 2008b The spectroscopy group and the development of a global spectral library. *IUSS Bulletin* 133, 30-32.

Vogt R and Finlayson-Pitts J 1994 A diffuse reflectance infrared Fourier transform spectroscopic_DRIFTS. Study of the surface reaction of NaCl with gaseous NO₂ and HNO₃. *The Journal of Physical Chemistry* 98 (14), 3747-3755.

Walter C 2002 Analyse spatiale des sols en vue de leur gestion précise et de leur surveillance. *Dimplôme d'habilitattion à diriger des recherches*. ENSA Rennes. 96p.

Wang W, Kobayashi T, Yasutake D, Kitano M, Cho H, Araki T and Yoshikoshi H 2008 Experiments on the control of salinity and sodicity in surface-irrigated fields in the upper Yellow river valley, objectives and methodology. *J. Fac. Agr. Kyushu Univ.* 53 (1), 251-256.

Weidung L, Baret F, Xingfa G, Qingxi T, Lanfen Z and Bing Z 2002 Relating soil surface moisture to reflectance. *Remote sensing of environment* 81, 238-246.

Weng Y, Gong P and Zhu Z 2008 Reflectance spectroscopy for the assessment of soil salt content in soils of the Yellow River Delta of China. *International Journal of Remote Sensing* 29 (19), 5511-5531.

Wetterlind J, Stenberg B, Söderström M 2008 The use of near infrared (NIR) spectroscopy to improve soil mapping at the farm scale. *Precision Agric.* 9, 57-69. Springer Science+Business Media, LLC 2008.

Weyer LG and Lo SC 2002 Spectra structure correlations in the Near-infrared. In : *Handbook of Vibrational Spectroscopy*, John Wiley and Sons Ltd., pp. 1817-1837.

Wold H 1966 Estimation of principal component and related models by iterative least squares, *Multivariate Analysis*, ed. P.R. Krishnaiah, Newyork : Academic Press, 391-420.

Wold S, Martens H and Wold H 1983 The multivariate calibration problem in chemistry solved by the PLS method, in *Proc. Conf. Matrix Pencils*, Ruhe A and Kågström B (Eds), March 1982, *Lecture Notes in Mathematics*, Springer Verlag, Heidelberg. p. 286-293.

Wold S, Sjöström M and Eriksson L 2001 PLS-regression : a basic tool of chemometrics. *Chemometrics Intelligence Laboratory Systems* 58, 109-130.

Workman J 1998 Ultraviolet, Visible and Near-Infrared Spectroscopy. In : Workman, J. and Springsteen, A. (Eds.), *Applied Spectroscopy – A Compact Reference for Practitioners*, Academic Press, Chestnut Hill, MA, p. 29-47.

Workman J and Shenk J 2004 Understanding and using the near-infrared spectrum as an analytical method. p. 3-10. In. Roberts C.A., Workman J and. Reeves III JB (ed.) *Near-infrared spectroscopy in agriculture*. ASA, CSSA, and SSSA, Madison, WI.

Xu JF and Huang JF 2008 Empirical Line Method Using Spectrally Stable Targets to Calibrate IKONOS Imagery. *Soil Science Society of China Pedosphere* 18(1), 124-130.

Yamaguchi Y, Kahle AB, Tsu H, Kawakami T and Pniel M 1998 Overview of Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER). *IEEE Transactions on Geoscience and Remote Sensing* 36, 1062-1071.

Yemefack M, Rossiter DG, Njomgang R 2005 Multi-scale characterization of soil variability within an agricultural landscape mosaic system in southern Cameroon. *Geoderma* 125 (1-2), 117-143.