HAL open science

Caractérisation et quantification des minéraux argileux dans les sols expansifs par spectroscopie infrarouge aux échelles du laboratoire et du terrain.

Camille Truche

▶ To cite this version:

Camille Truche. Caractérisation et quantification des minéraux argileux dans les sols expansifs par spectroscopie infrarouge aux échelles du laboratoire et du terrain.. Planète et Univers [physics]. Université Paul Sabatier - Toulouse III, 2010. Français. NNT : . tel-00594021

HAL Id: tel-00594021

https://theses.hal.science/tel-00594021

Submitted on 18 May 2011

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.

Université de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :
Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)

Discipline ou spécialité :

Minéralogie - spectroscopie infrarouge

Présentée et soutenue par :
Camille Truche

le : lundi 13 décembre 2010

Titre :

Caractérisation et quantification des minéraux argileux dans les sols expansifs par spectroscopie infrarouge aux échelles du laboratoire et du terrain

---

JURY

Didier Beziat	Professeur UPS,Toulouse	Président
Jocelyne Brendlé	Professeur ISM, Mulhouse	Rapporteur
Patrick Launeau	Professeur, Université de Nantes	Rapporteur
Jean Vaunat	Professeur ETSCCP, Barcelone	Rapporteur
Frédéric Villieras	Directeur de recherche CNRS, Nancy	Examinateur

---

Ecole doctorale :
Sciences de l'Univers, de l'Environnement et de l'Espace (SDU2E)

Unité de recherche :
Laboratoire des Mécanismes et Transferts en Géologie

Directeur(s) de Thèse :

François MARTIN, José DARROZES et Anne BOURGUIGNON

Rapporteurs :

J. Brendlé, P. Launeau, J. Vaunat

AVANT-PROPOS

Ce mémoire ne serait pas ce qu'il est sans la collaboration, l'aide et le soutien de nombreuses personnes. Il m'a permis de vivre une expérience très enrichissante au niveau professionnel et humain.

Je remercie en premier lieu, José Darrozes (LMTG, Toulouse), Anne Bourguignon (BRGM, Orléans) et George Delpont (BRGM, Toulouse) pour m'avoir proposé ce sujet de recherche ainsi que François Martin et Philippe Roubichou. Ils ont participé à l'encadrement de ce travail de thèse. Je remercie également la région Midi-Pyrénées et le BRGM pour avoir subventionné mon projet de recherche pendant ces trois années.

Je remercie tous particulièrement Madame Jocelyne Brendlé (ISM Mulhouse), Monsieur Patrick Launeau (LPG, Nantes) et Monsieur Jean Vaunat (ETJCCP, Barcelone) qui ont accepté de juger ce travail et d'en être rapporteurs. Je tiens également à remercier Monsieur Béziat (LMTG, Toulouse), Monsieur Villieras (LEM, Nancy) et Monsieur Grauby (CINaM, Marseille) d'avoir accepter de participer au jury de cette thèse.

J'exprime ici toute ma reconnaissance envers Sabine Petit (Hydr'Asa, Poitiers) pour l'intérêt qu'elle a manifesté pour ce travail. Ces conseils avisés et ses connaissances m'ont permis d'avancer dans les moments les plus difficiles.

Je remercie tous ceux sans qui cette thèse ne serait pas ce qu'elle est aussi bien par les discussions que j'ai eu la chance d'avoir avec eux, leurs suggestions et contributions. Je pense ici en particulier Michel Thibaut, Jannick Ingrin, Philippe de Parceval, Pierre Micoud (LMTG, Toulouse) et Isabelle Bouroullec (BRGM, Toulouse).

Mes remerciements vont à l'ensemble des thésards avec qui j'ai partagé des moments inoubliables que au laboratoire ou ailleurs. Irina, Sophie, Guilhem, Jérémy (qui n'est pas resté longtemps, mais qui a vécu le pire !) merci pour la bonne humeur dans le bureau (le meilleur du LMTG) et pour avoir partagé mes joies et mes peines tout au long de ces trois années. Merci, entre autre, pour les grands moments musicaux, les « gossip », l'apprentissage du russe et les grandes (et longues) discussions !! Merci aussi à tous ceux qui ne sont pas dans le bureau (quoi que…) : Elena, Teresa, Laurent, Ana, Cyrielle, Alisson, Sylvaine, Quentin, Jérémy, Tristan, FIX, Emilie, Julien, Thomas, Aymen et tout les autres… Qu'aurais-je fait sans vous !! Et puis courage : vous allez la finir aussi votre thèse (si ce n'est pas encore fait) !!

Je remercie mes parents, Mano et Michel, et mes frères et sœurs, Laurent, Violaine, Irène, Vincent, Perrine, pour leur irremplaçable et inconditionnel soutien. Ils ont été présents pour écarter les doutes et partager les joies.

Enfin, milles merci à ceux qui, malheureusement, ont un ami qui a fait une thèse sur les « argiles gonflantes » : Sylvie, Mathieu, Emma, Popo, Pauline, Clem… Merci à Nico pour prendre soin de moi.

RESUME ETENDU

Ce travail tend à l'amélioration de la détection et de la quantification des minéraux argileux, plus particulièrement des minéraux argileux « gonflants », dans la partie superficielle des sols. L'étude est menée à l'échelle locale et à l'échelle globale pour accompagner la démarche des régions pour la prévention du risque de retrait-gonflement des argiles à l'origine de dommages aux constructions.

La première partie présente les principes généraux et la base de données utilisés dans la thèse. Une approche utilisant la spectrométrie infrarouge à différentes échelles (laboratoire et terrain) et dans différents domaines de longueur d'ondes (proche et moyen infrarouge) permet d'analyser le comportement des pics d'absorption caractéristiques des minéraux argileux dans un mélange polyphasé de type sol. Pour cela, la construction d'une bibliothèque de pôles purs est fondamentale. Elle a permis la construction de mélanges synthétiques pouvant ensuite être comparés aux sols naturels.

Dans une deuxième partie, ces données ont permis de tester les capacités et les limites des méthodes classiquement utilisées pour l'analyse des paramètres physiques (pH, CEC, …) et minéralogiques (teneur en argile ou en sable, …) qui caractérisent les sols. Les méthodes statistiques telles que l'analyse en composantes principales (ACP) et la régression par moindres carrés partiels (PLSR) ont montré qu'elles permettent la construction de modèles de bonne qualité. En revanche, leur robustesse est largement dépendante de la qualité et de la légitimité de la collection de données utilisée pour construire le modèle vis-à-vis des échantillons analysés. Ces méthodes statistiques ont été confrontées à une méthode s'intéressant plus particulièrement à la forme et à la position des pics d'absorption caractéristiques des minéraux argileux. Cette dernière a mis en évidence la quantification de la kaolinite à l'échelle du laboratoire, la perte d'information importante lors du passage à l'échelle du terrain et la difficulté à différencier l'information spectrale provenant des illites et des smectites.

Les limites de ces deux premières approches nous ont conduits à nous intéresser plus finement au signal infrarouge par une approche mathématique basée sur la corrélation croisée en ondelettes. Cette méthode utilise la transformée en ondelettes continue des spectres infrarouges et la corrélation croisée en ondelettes. La transformée en ondelettes permet d'extraire les informations portées par le spectre infrarouge présents à différentes résolutions (fréquences). Les caractéristiques invisibles par les autres méthodes des spectres sont ainsi mises en évidence. La corrélation croisée permet de quantifier l'écart entre deux signaux, l'un connu et l'autre à caractériser. Cette méthode a été testée sur des signaux synthétiques composés de gaussiennes puis sur des spectres infrarouges de mélanges de minéraux argileux. Les résultats obtenus montrent que la quantification des principaux minéraux argileux et plus particulièrement de la smectite est possible.

Ainsi, cette méthode basée sur l'analyse en ondelettes des spectres infrarouges est la plus prometteuse pour la quantification fine de la teneur en smectite des sols, à l'échelle locale comme à l'échelle globale, dans les études de prévention des risques de retrait et de gonflement des sols.

RESUME COURT

Ce travail tend à l'amélioration de la détection et de la quantification des minéraux argileux, et plus particulièrement des minéraux argileux « gonflants », dans la partie superficielle des sols. Cette étude est menée à l'échelle locale et globale pour accompagner la démarche des régions pour la prévention du risque de retrait-gonflement des argiles. La première partie présente les principes généraux de la spectrométrie infrarouge et la base de données utilisés dans la thèse. Nous confrontons ensuite deux approches classiquement utilisées dans les sciences du sol : Les méthodes statistiques (PLSR et ACP) et l'étude des figures d'absorption caractéristiques des minéraux argileux. Elles montrent principalement la quantification de la kaolinite dans un mélange multiphasé et dans un sol. Les limites de ces méthodes pour la quantification de la smectite amènent au développement d'une nouvelle méthode basée sur la transformée en ondelettes continue et la corrélation croisée en ondelette. Cette méthode apparaît être la plus prometteuse pour la quantification fine de la teneur en smectite des sols, à l'échelle locale comme à l'échelle globale, dans les études de prévention des risques de retrait et de gonflement des sols.

ABSTRACT

The purpose of this study is the detection and quantification of clay minerals, especially swelling clay minerals, in surface soils part. This study is carried out at local and global scale, in order to support swelling soil hazards approach of the engineering offices. The first part presents general principles of infrared spectroscopy and the database used in the thesis. Then, we compare two frequently used methods in soil sciences: statistic methods (PLSR and PCA) and absorption figures analysis featuring clay minerals. They mainly provide kaolinite quantification in multiphase mixtures and in soil. The limits of smectite quantification of both methods lead to a new method development based on continuous wavelet transform and cross correlation. This method is the most efficient and promising one for weak quantification of swelling minerals, at local scale as well as at global scale, in studies prevention of swelling soil hazard.

TABLE DES MATIERES

I. CHAPITRE 1 : INTRODUCTION ET CONTEXTE D'ETUDE 19

I.1. LA PREVENTION DU RISQUE DE RETRAIT GONFLEMENT EN FRANCE : GENERALITE 19

I.2. PROBLEMATIQUE 23

I.3. CONTEXTE GEOLOGIQUE ET GEOMORPHOLOGIQUE 26

I.3.1 L'HISTOIRE DU BASSIN CENTRAL AQUITAIN ET DE SES FORMATIONS MOLASSIQUES 28

I.3.1.1 L'extension Mésozoïque 28

I.3.1.2 Tectonique Campano-Oligocène 30

I.3.1.3 Rifting Oligo-Miocène 31

I.3.2 LES FORMATIONS QUATERNAIRES 33

I.4. DEVELOPPEMENT METHODOLOGIQUE 37

II. CHAPITRE 2. LA SPECTROSCOPIE INFRAROUGE : CARACTERISATION ET BASE DE DONNEES 41

II.1. THEORIE DU RAYONNEMENT ELECTROMAGNETIQUE 41

II.1.1 ORIGINES DES BANDES D'ABSORPTION 42

II.1.2 INTERPRETATION DES SPECTRES 44

II.1.2.1 Dans le moyen infrarouge 44

II.1.2.2 Dans le proche infrarouge 45

II.1.3 LES PERTURBATIONS ENGENDREES PAR L'ATMOSPHERE SUR LE RAYONNEMENT ELECTROMAGNETIQUE 46

II.2. INSTRUMENTS DE MESURE UTILISES EN SPECTROSCOPIE 47

II.2.1 LA SPECTROSCOPIE INFRAROUGE A TRANSFORMEE DE FOURIER 47

II.2.1.1 Principe 47

II.2.1.2 Dans le domaine de l'infrarouge moyen (4000 - 400 cm^{-1}) : la transmission/ absorption. 48

II.2.1.3 Dans le domaine du proche infrarouge (9000 – 4000 cm^{-1}) : la réflexion diffuse 48

II.2.2 SPECTROMÈTRE ASD FIELD SPEC FR® (ROY, 2007) 49

II.2.3 CAPTEUR MULTISPECTRAL ASTER 50

II.3. MOYENS DE VALIDATION DES DONNEES ACQUISES 52

II.3.1 LA DIFFRACTION DES RAYONS X 52

II.3.2 ANALYSES GEOTECHNIQUES 54

II.3.3 CHIMIE DES ELEMENTS MAJEURS 55

II.4. BASE DE DONNEES 57

II.4.1 LES MINERAUX ARGILEUX 57
II.4.1.1 Structure des minéraux argileux 57
II.4.1.2 Classification des minéraux argileux les plus fréquents dans les sols de l'étude 58
II.4.2 LES COLLECTIONS 60
II.4.2.1 Les minéraux argileux de référence 60
II.4.2.2 Les mélanges synthétiques 66
II.4.2.3 Les sols 71

III. CHAPITRE 3 : TRAITEMENT DES DONNEES 83

III.1. APPROCHES STATISTIQUES 83
III.1.1 L'ANALYSE EN COMPOSANTES PRINCIPALES 84
III.1.1.1 Démarche générale. 84
III.1.1.2 Résultats 87
III.1.1.3 Discussion, conclusion 93
III.1.2 LA REGRESSION PAR MOINDRES CARRES PARTIELS (PLSR : PARTIAL LEAST SQUARES REGRESSION) 93
III.1.2.1 Démarche générale 94
III.1.2.2 Démarche sur les échantillons de l'étude 97
III.1.2.3 Résultats 98
III.1.2.4 Discussion, conclusion 104
III.2. APPROCHE BASEE SUR LA FORME DES ABSORPTIONS CARACTERISTIQUES DES MINERAUX ARGILEUX 105
III.2.1 PUBLICATION 1: CLAY MINERALOGY OF MIDI-PYRÉNÉES AREA (FRANCE) BY INFRARED SPECTROSCOPY: DISAPPOINTED RESULTS (SOUMIS À CLAY MINERALS) 105
III.2.1.1 Résumé en Français 105
III.2.1.2 Abstract 107
III.2.2 INTRODUCTION 107
III.2.3 GEOLOGICAL SETTING 109
III.2.4 MATERIALS AND METHODS 111
III.2.4.1 Sampling 111
III.2.4.2 XRD measurements 112
III.2.4.3 Estimation of the soil mineralogy 112
III.2.4.4 Infrared laboratory and field measurements 113
III.2.4.5 Spectral data analysis 115
III.2.5 RESULTS 116

III.2.5.1 Reference and soil sample composition 116
III.2.5.2 Synthetic mixtures in the MIR range with FTIR spectrometer 117
III.2.5.3 Synthetic mixtures in the NIR range with FTIR spectrometer 119
III.2.5.4 Synthetic mixtures with ASD spectrometer 119
III.2.5.5 Soils under laboratory conditions with ASD spectrometer 120
III.2.5.6 Soils analyses under natural conditions with ASD spectrometer 121
III.2.6 DISCUSSION AND CONCLUSIONS 122
III.2.7 ACKNOWLEDGEMENTS 123
III.2.8 REFERENCES 124
III.3. CONCLUSION 127

IV. CONTRIBUTION DE L'ANALYSE EN ONDELETTES COUPLEE A LA CORRELATION CROISEE A LA SPECTROMETRIE INFRAROUGE 133

IV.1. INTRODUCTION ET ETAT DE L'ART 133
IV.2. PRINCIPE GENERAL 135
IV.2.1 L'ANALYSE MULTI-RESOLUTION 135
IV.2.1.1 La transformée de Fourier 135
IV.2.1.2 La transformée de Gabor 135
IV.2.1.3 La transformée en ondelettes 136
IV.2.2 LA TRANSFORMEE EN ONDELETTES CONTINUES 137
IV.2.3 LES ONDELETTES 138
IV.2.4 CHOIX DE L'ONDELETTE 141
IV.3. CORRELATION 142
IV.3.1 CORRELATION CROISEE 143
IV.3.2 CONE D'INFLUENCE ET INTERVALLE DE CONFIANCE 144
IV.4. PROBLEMATIQUE 144
IV.5. ANALYSE EN ONDELETTES ET CORRELATION CROISEE DES SIGNAUX INFRAROUGES 145
IV.5.1 METHODOLOGIE 145
IV.5.1.1 Approche générale 145
IV.5.1.2 Seuil 146
IV.5.1.3 Normalisation 146
IV.5.2 ANALYSE DES GAUSSIENNES 147
IV.5.2.1 Influence de l'amplitude des gaussiennes 147

IV.5.2.2 Influence de la variance des gaussiennes 149
IV.5.2.3 Influence de la position des gaussiennes 151
IV.5.3 CARACTERISATION DE FIGURES DE MELANGES DE GAUSSIENNES 153
IV.5.3.1 Présence de la gaussienne inconnue dans la bibliothèque de référence 153
IV.5.3.2 Absence de la gaussienne « inconnue » dans la bibliothèque de référence 155
IV.5.4 QUANTIFICATION DE MELANGES A DEUX MINERAUX ARGILEUX 156
IV.5.4.1 Matériel 157
IV.5.4.2 Méthode utilisée 159
IV.5.4.3 Résultats 160
IV.5.4.4 Application de la méthode aux mélanges à deux minéraux : discussion et conclusion 166
IV.5.5 ANALYSE EN ONDELETTES ET CORRELATION CROISEE : DISCUSSION ET CONCLUSION 167
IV.5.5.1 Méthodes de normalisation 167
IV.5.5.2 Application à la caractérisation des argiles « gonflantes » 167

V. CONCLUSIONS GENERALES ET PERSPECTIVES 175

VI. BIBLIOGRAPHIE 181

VII. ANNEXE 1 : ANALYSES PAR DIFFRACTION DES RAYONS X DES MINERAUX ARGILEUX DE REFERENCE 191

VIII. ANNEXE 2 : DONNEES DE DRX ACQUISE SUR LES SOLS DE LA REGION MIDI-PYRENEES SUIVANT LES UNITES MORPHO PEDOLOGIQUES PRELEVEES. 201

VIII.1. LES SOLS DE LA MOLASSE 201
VIII.2. LES SOLS DES PLAINES ER DES TERRASSES QUATERNAIRES 206
VIII.3. LES SOLS DES PLATEAUX MARNO-CALCAIRE 210

IX. ANNEXE 3 : MODELISATION PLSR 213

X. ANNEXE 4 : PROGRAMME 217

Chapitre 1

Introduction et contexte d’étude

I. CHAPITRE 1 : INTRODUCTION ET CONTEXTE D'ETUDE

I.1. LA PREVENTION DU RISQUE DE RETRAIT GONFLEMENT EN FRANCE : GENERALITE

En France métropolitaine, la vulnérabilité des constructions individuelles sur certains sols argileux a été mise en évidence à l'occasion de la sécheresse exceptionnelle de l'été 1976 et a pris une réelle ampleur lors des périodes sèches des années 1989-1991 et 1996-1997, puis dernièrement au cours de l'été 2003 (figure I-1). C'est ce que l'on appelle le phénomène de retrait-gonflement des argiles. Le risque associé est devenu en France, dans le cadre du régime des catastrophes naturelles, la deuxième cause d'indemnisation via les assurances, par les montants engagés.

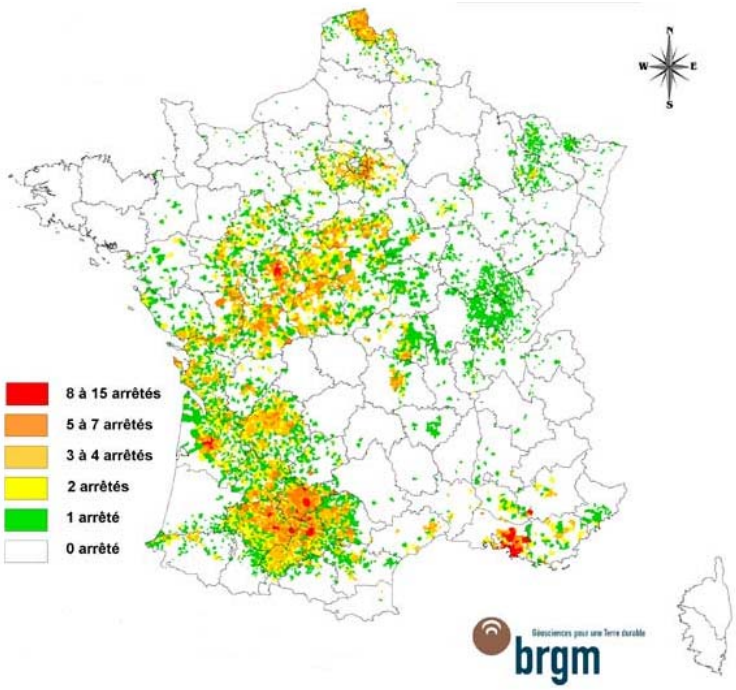


Figure I-1 : Répartition des mouvements de terrain différentiels consécutifs à la sécheresse et à la réhydratation des sols d'après le nombre d'arrêtés de reconnaissance de l'état de catastrophe naturelle pris par commune à la date du 15 aout 2006 (source : MEDD (www.prim.net - extraction du 22/08/2006)).

La nécessité du développement de nouveaux outils pour la prévention de ce risque est liée à la limite des outils actuels vis-à-vis du caractère diffus de ce risque. En effet, c'est sur les habitations individuelles que les effets du retrait et gonflement de certaines formations géologiques argileuses affleurantes sont les plus évidents. Ils se traduisent par des tassements différentiels du sol et par la perte de contact entre les fondations et le terrain d'assise ce qui provoquent des désordres au niveau des constructions tels que, la fissuration des structures enterrées ou aériennes, la distorsion des ouvertures, le décollement des éléments composites ou l'étirement des canalisations (figure II-2). Dans la région Midi-Pyrénées, l'importante croissance démographique de la région observée depuis le milieu des années 90, s'accompagne d'une

expansion géographique des villes importantes et de l'apparition d'une urbanisation diffuse sur l'ensemble du territoire (figure I-3) et plus particulièrement le long des axes Toulouse- Cahors et Toulouse- Albi.

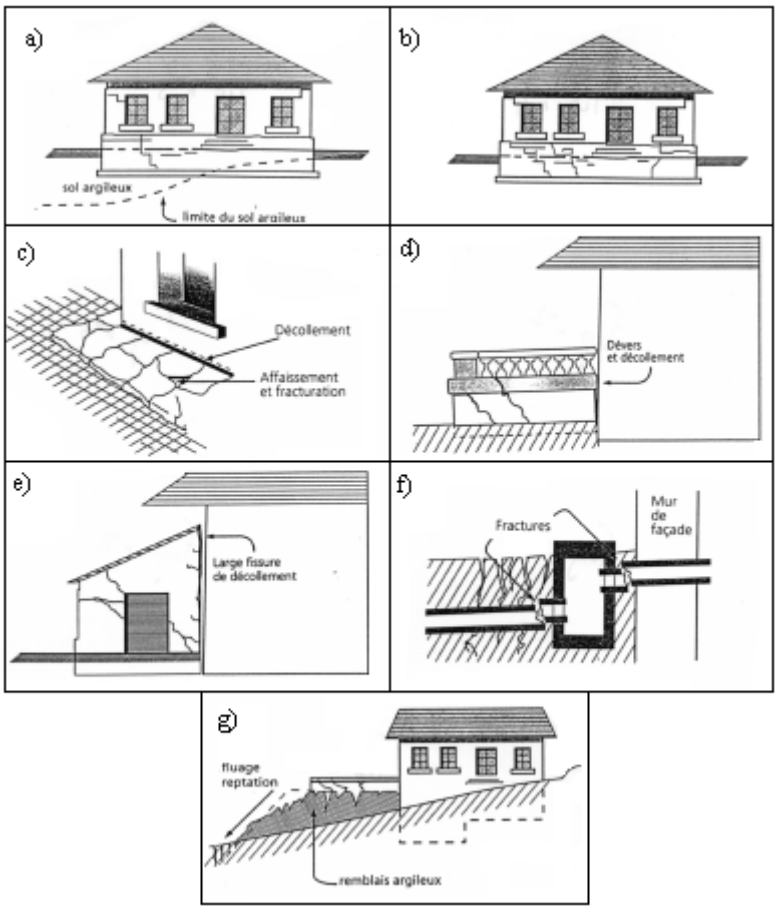


Figure I-2 : Exemple de dommages occasionnés par les phénomènes de retrait et de gonflement des sols. a) désordres partiels dus à la variation d'épaisseur du sol; b) désordres dus à l'ensemble du soubassement et de l'ossature ; c), d), e) et f) désordres affectant un dallage extérieur, une terrasse, un appentis, une conduite enterrée ; g) aggravation par la sécheresse de désordres affectant un remblai argileux (phénomène de solifluxion). (Source : GUIDE DE PREVENTION « sécheresse et construction », Ministère de l'Environnement, Direction de la Prévention des Pollutions et des Risques, Délégation aux Risques majeurs.)

Ce type d'urbanisation est d'autant plus sensible aux phénomènes de retrait-gonflement qu'il est souvent localisé sur les coteaux molassiques qui sont généralement des sols potentiellement « à risque » (zone d'aléa moyen, figure I-4). D'autre part les études préalables à ce type d'habitats sont limitées par l'application de l'article L. 231-2 du Code de la construction et de l'habitation (CCH), qui n'impose pas au constructeur de procéder systématiquement à des études de sols préalables à la signature des contrats de construction.

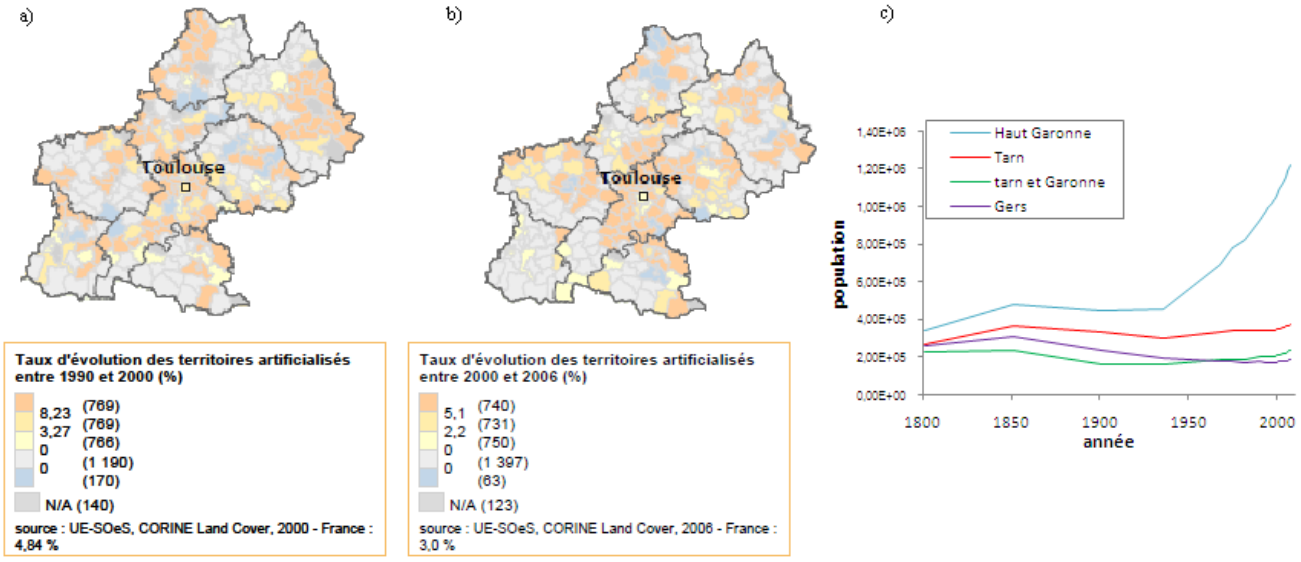


Figure I-3 : Evolution des territoires artificialisés en Midi-Pyrénées pour les périodes de a) 1990-2000 et b) 2000-2006 (source : Union européenne – SOeS, Corine Land Cover, 2006) et c) estimation de la population au 1 er janvier pour les principaux départements constituant le bassin centrale aquitain entre 1801 et 2008 (source : INSEE et IAURIF).

Bien que ce ne soit pas le seul facteur expliquant l'apparition de ce phénomène, il existe une corrélation étroite entre la répartition des sinistres et la nature géologique des formations superficielles qui a amené le BRGM, dés le milieu des années 1990 (Chassagneux D. 1995) , à élaborer des développements méthodologiques en vue de cartographier l'aléa retrait-gonflement des argiles, d'abord à l'échelle communal, puis à l'échelle départementale laquelle s'est révélée plus adaptée pour établir de manière relativement rapide et homogène, les documents de base nécessaires à l'élaboration d'une véritable politique de prévention de ce risque naturel.

La cartographie actuelle des risques de retrait-gonflement des sols vise à délimiter les zones pouvant être exposées à ce risque naturel, puis à édicter des mesures de protection vis-à-vis de ce risque. Elle se caractérise par deux étapes majeures :

- La première constitue la phase technique permettant d'estimer la probabilité de survenance du phénomène donné en un endroit donné, probabilité appelée aléa ;
- La deuxième, qui constitue la phase réglementaire qui se pose plus en termes d'aménagement, détermine quelles sont les meilleures réponses à apporter aux problèmes mis en évidence par la première phase. Ces cartes doivent être précises. Elles constituent des documents d'aide à la décision et servent de base à l'élaboration de documents à valeur juridique tels que les Plan de Prévention des Risques institués par la loi Barnier du 2 février 1995, et qui ont pour vocation de réduire les

dommages lors de catastrophes naturelles en maîtrisant l'aménagement du territoire dans les zones exposées à un risque (Summers, Lewis et al. 2009).

La phase technique de réalisation de la carte d'aléa se base principalement sur les cartes géologiques et sur les observations de terrain des géologues. La démarche consiste tout d'abord à établir une cartographie départementale des formations argileuses et marneuses affleurantes et sub-affleurantes, à partir de la synthèse des cartes géologiques du département au 1/50 000. Les formations ainsi identifiées, font ensuite l'objet d'une hiérarchisation quant à leur **s**usceptibilité vis-à-vis du phénomène de retrait-gonflement. Cette caractérisation est établie sur la base de trois critères principaux: la caractérisation lithologique de la formation, la composition minéralogique de sa phase argileuse, et son comportement géotechnique, ce qui conduit à l'établissement d'une **carte départementale de susceptibilité au retrait-gonflement**. La carte d'aléa (figure I-4) est alors établie à partir de cette carte synthétique des formations argileuses et marneuses, en tenant compte de la susceptibilité des formations identifiées ainsi que de la probabilité d'occurrence du phénomène en relation avec la localisation des sinistres recensés sur les différentes communes. Ainsi, l'élaboration de la carte de susceptibilité reposent pour une grande part sur l'étude de documents existants, qu'il s'agisse de travaux réalisées par le BRGM ou des rapports d'expertise disponibles dans les archives des bureaux d'études, de la banque de données des sous-sols ou des organismes d'assurances. L'autre part des données consiste en des **expertises géotechniques ponctuelles** permettant d'actualiser l'information géologique acquise sur le terrain. Durant cette phase de travail, des échantillons sont prélevés sur les différentes formations argileuses. Les analyses géotechniques et diffractométriques de ces échantillons apportent un complément d'information indispensable aux données déjà récoltées, en particulier pour les formations argileuses pour lesquelles n'existe aucune information.

L'inconvénient majeur de la méthode mise en œuvre pour définir les zones sujettes aux effets de retrait-gonflement, réside dans le fait que les cartes géologiques sont rarement optimales pour caractériser et délimiter de manière précise les contours des affleurements de matériaux argileux. En outre, les formations superficielles n'y font pas souvent l'objet d'une description détaillée. Il en résulte une certaine imprécision qui apparaît clairement dans le cas des molasses, dont la lithologie séquentielle implique de nombreuses variations latérales de faciès, que les cartes géologiques au 1 / 50 000 ne peuvent pas traduire dans le détail.

De plus, les campagnes de terrain visant à compléter l'information des cartes géologiques sont souvent confrontées à des problèmes logistiques et techniques. Elles fournissent des échantillons sur lesquels sont réalisées des analyses géotechniques et de diffraction des rayons X (DRX) qui complètent l'analyse minéralogique ainsi que la détermination du potentiel de gonflement des sols. **Ces prélèvements sont ponctuels**. Ils constituent des points de contrôle qui créent une information de terrain hétérogène et discontinue. Enfin, la précision de la carte d'aléa qui, partant d'une information valide à l'échelle 1/ 50 000, aboutit à un document valide à une échelle entre 1/50000 et 1/100 000. Cependant l'utilisation future de la carte sera l'élaboration de cartes de risques par découpage communal, soit à une échelle bien supérieure de

celle de validité des données. Aussi, **ces cartes sont peu adéquates pour la prévention du risque de retrait et de gonflement des sols.**

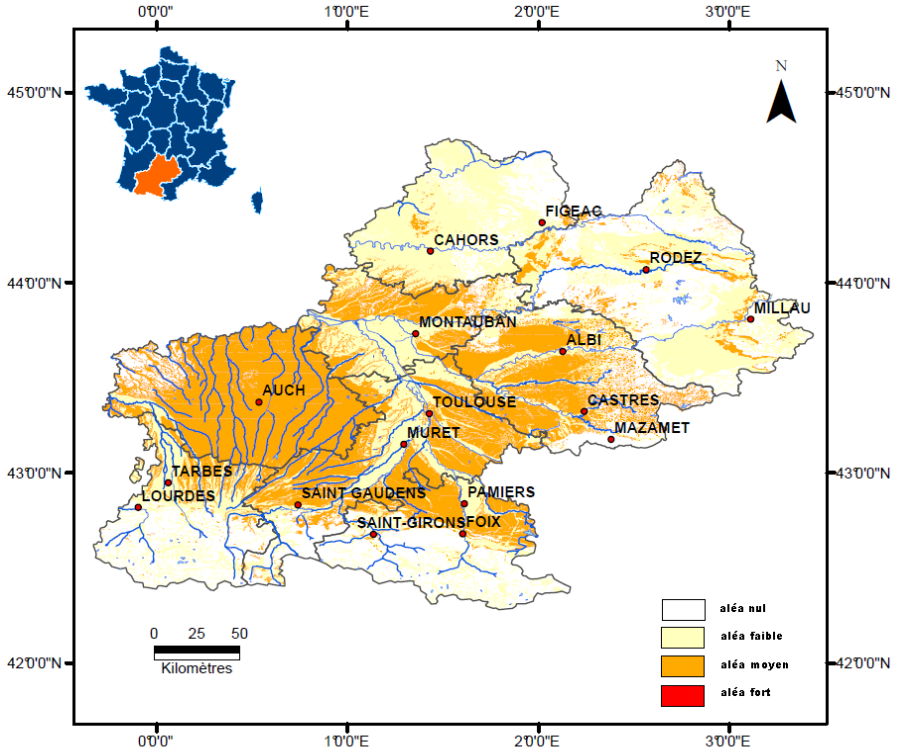


Figure I-4 : Carte de l'aléa retrait-gonflement des argiles pour la région Midi-Pyrénées. (Source : BRGM, (Norie A. 2001; Delpont G. 2002; Delpont G. 2002; Delpont G. 2005; Delpont G. 2006; Bouroullec I. 2009))

I.2. PROBLEMATIQUE

De nombreux facteurs de prédisposition tels que la géomorphologie, le contexte hydrogéologique, la végétation, les actions anthropiques ou les défauts de construction permettent de caractériser la susceptibilité du milieu vis-à-vis de l'aléa retrait-gonflement des sols, sa répartition spatiale et donc l'occurrence de ce phénomène. Cependant, comme le montre le descriptif de la procédure d'établissement des cartes de susceptibilité au retrait-gonflement présenté précédemment, ces différents facteurs ne sont généralement pris en compte que ponctuellement. En effet, leur influence à l'échelle départementale est souvent difficile à mettre en évidence ou bien jugée non discriminante. Ainsi**, la nature des formations argileuses du département ou de la région reste donc le critère le plus utilisé pour la cartographie du risque de retrait et de gonflement des argiles.**

D'autre part, dans le cadre du projet ARGIC (Analyse du Retrait-Gonflement et de ses Incidences sur les Constructions : ANR-05-PRGCU-005 ; (Vincent M. 2009)), une étude comparative a permis d'évaluer la pertinence des différentes techniques classiquement utilisées pour la caractérisation de la susceptibilité d'un

sol au retrait-gonflement . Il a été défini qu'en matière de prévention du phénomène de retrait-gonflement des argiles, les études géotechniques préalables doivent s'intéresser à :

- reconnaître et identifier les sols présents au droit de la construction,
- caractériser leur répartition spatiale et repérer d'éventuelles hétérogénéités verticales et horizontales,
- identifier la présence éventuelle d'argiles gonflantes et en déterminer le potentiel de retrait-gonflement en cas de variation des conditions hydriques,
- repérer l'existence d'éventuels facteurs d'aggravation susceptibles de modifier localement la teneur en eau dans le sol sur la parcelle.

Cette reconnaissance géotechnique passe par la réalisation d'essais géotechniques en laboratoire et éventuellement, par une caractérisation semi-quantitative de la composition minéralogique de la phase argileuse qui peut être faite par diffraction des rayons X. Cependant, la prise en compte de cette dernière est la plus onéreuse et la plus difficile à mettre en œuvre en termes de spatialisation du fait de son caractère discontinu. La production de carte proposant une cartographie continue de la variation de la minéralogie de la phase argileuse permettrait d'accroître la qualité et la quantité d'information nécessaire aux décideurs pour délimiter les zones potentiellement risquées.

Ainsi, en parallèle à cette étude, **ce travail de thèse a pour objectif le développement de nouvelles méthodes permettant de répondre aux préoccupations présentées dans le projet ARGIC en terme de reconnaissance et d'identification des argiles gonflantes des sols et de la caractérisation de leur répartition spatiale.** Elles doivent faciliter à la fois une analyse rapide et économique des sols mais également permettre un maillage de cartographie des sols plus approprié afin d'associer les grandes échelles, utilisées ponctuellement, et les études plus globales et qualitatives telles que les cartes d'aléa.

Pour cela, de nombreuses études ont montré que la **spectrométrie infrarouge** pouvait être reliée à certaines propriétés et caractéristiques des sols. En outre, les résolutions spectrales relativement hautes des différents appareils de mesures ainsi que les connaissances apportées sur l'assignation des bandes spectrales, par les minéralogistes et les spécialistes dans ce domaine, ont largement contribué à la généralisation de l'utilisation de ces méthodes en sciences du sol. Cependant, la nature hétérogène des sols crée de nombreuses difficultés dans l'obtention des propriétés à partir des spectres de réflectance et son utilisation reste encore difficile à appliquer pour des bureaux d'étude spécialisés dans l'analyse des sols. Elle n'est pas, à l'heure actuelle, intégrée dans les procédures de caractérisation des argiles contrairement à l'analyse par diffraction des rayons X qui est pourtant une méthode plus lourde à mettre en œuvre en terme de temps d'analyse et de coût et dont l'intérêt est plus qualitatif que quantitatif. Nous avons donc cherché **les possibilités d'intégration de ces méthodes spectrométriques à des procédures routinières de bureau d'étude.**

Ainsi, dans ce manuscrit, nous présentons **dans quelles mesures la spectrométrie infrarouge peut constituer une méthode opérationnelle de localisation et de quantification des zones riches en argiles dites « gonflantes »** tout en apportant une information homogène et continue aux différentes échelles intéressant les décideurs. Cela revient à répondre à deux questions principales:

- comment rendre la spectrométrie infrarouge applicable par des bureaux d'études pour caractériser et quantifier les minéraux argileux des sols?

- comment faire le lien entre les différentes échelles utilisables en spectrométrie infrarouge?

Plusieurs objectifs intermédiaires ont été définis :

- créer des bibliothèques spectrales contenant des spectres de pôles purs, de mélanges expérimentaux et de sols enregistrés en laboratoire. Elles contiennent le matériel de base pour étudier, à partir de leur réponse spectrale, les figures d'absorption et les variations intéressantes afin de différentier et quantifier les principales argiles.

- développer une méthode permettant de relier l'information spectrale à la teneur en différents minéraux argileux (principalement de la famille des illites, des smectites, et des kaolinites).

- vérifier que les relations établies puissent être utilisées à partir de données acquises sur le terrain.

- vérifier que la méthode développée est intéressante pour les trois échelles d'analyse choisie (laboratoire, terrain, image).

- évaluer la capacité des images multispectrales ASTER pour détecter des variations de la composition de la phase argileuse des sols en associant les résultats obtenus sur le terrain à l'analyse d'image. Cette étape n'a pas été traitée dans la thèse.

Le bassin central aquitain, situé dans la région Midi-Pyrénées, est choisi comme zone d'étude. De part sa situation géologique, il fait partie en France des zones les plus fortement impactées par les phénomènes de retrait-gonflement des sols (figure I-1 et I-5). D'autre part, la région Midi-Pyrénées a fait l'objet en octobre 2007 et pour une durée de deux ans, d'un projet multithématique autour des problématiques liées à la présence d'argiles « gonflantes ». Ce projet est conduit en collaboration entre plusieurs partenaires publics et privés tels que LMTG (Laboratoire des Mécanismes de Transfert en Géologie), BRGM, 2GH Entreprise, SMABTP-SOCABAT, IUT Génie Civil, LRPC. Il s'inscrit dans la continuité de projets de recherche antérieurs sur la thématique du retrait-gonflement des sols.

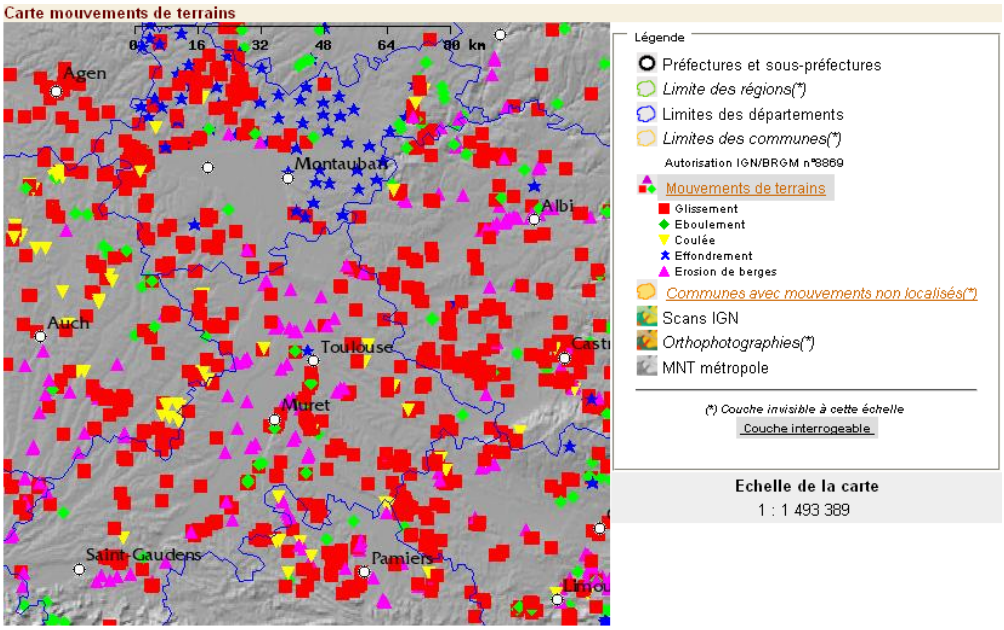


Figure I-5 : Recensement des mouvements de terrains observés dans le bassin central aquitain (extrait: base BDMVT gérée par le BRGM).

I.3. CONTEXTE GEOLOGIQUE ET GEOMORPHOLOGIQUE

Cette étude est focalisée sur la partie centrale du bassin aquitain (figure I-6) constituée majoritairement par les départements de la Haute Garonne, du Gers, du Tarn et du Tarn-et-Garonne. Ces départements sont en effet les plus touchés par les phénomènes de retrait-gonflement des argiles par le fait de la nature des sols, de la densité de population qui se traduit par une importante urbanisation diffuse. Or, l'habitat isolé est généralement soumis à une réglementation moins sévère quant à la prévention de ce risque. Le Bassin Aquitain est un bassin flexural d'avant-pays, qui se développe au front du chevauchement frontal Nord Pyrénéen (FNP, figure I-10a.). Il est très déformé dans sa partie sud, que l'on appelle avant-pays plissé, correspondant à la zone sous pyrénéenne. Il est limité au sud par le chevauchement frontal Nord Pyrénéen à vergence nord. Ce bassin dans sa partie plissée, est constitué par des sédiments détritiques avec des conglomérats syntéctoniques d'âge Eocène. Ils sont surmontés en discordance par des molasses post-tectoniques.

Le paragraphe suivant contient une synthèse sur le bassin d'Aquitaine, sur les formations molassiques et les sols de la région. Elle met en évidence une échelle inadaptée des documents existants pour la prévention du risque de retrait-gonflement des argiles au vu de la répartition discontinue et souvent de faible étendue des formations présentant un risque.

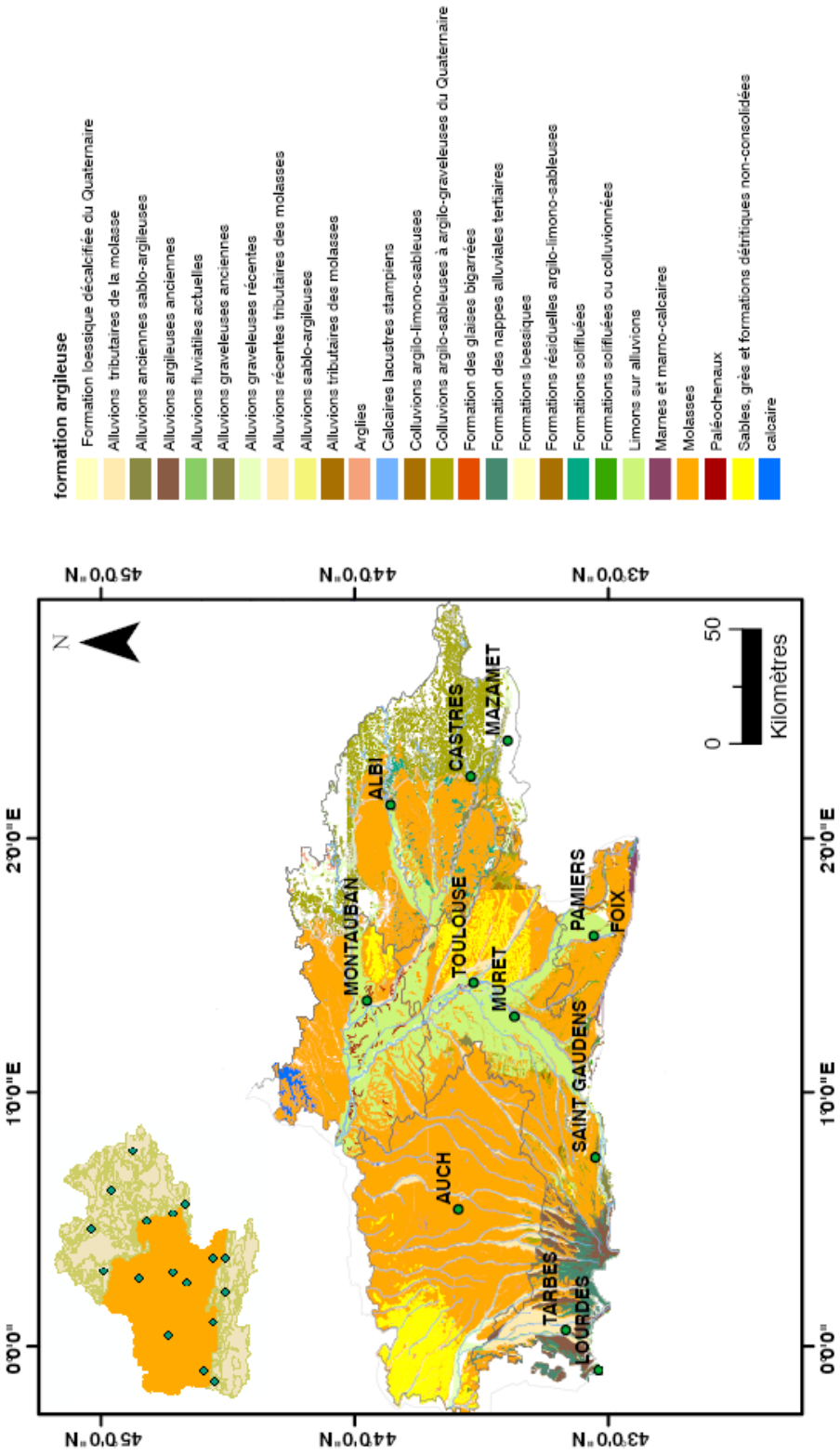


Figure I-6 : Formations superficielles argileuses du bassin central Aquitain (Source : BRGM). Coordonnées géographique du système de référence WGS 84.

I.3.1 L'histoire du Bassin central aquitain et de ses formations molassiques

Le Bassin aquitain occupe le quart sud-ouest de la France. Il est délimité au nord par le massif armoricain (Vendée), à l'est par le massif central, et au sud par la chaîne des Pyrénées.

I.3.1.1 L'extension Mésozoïque

L'histoire du bassin aquitain commence dès le Permien avec **une période de distension fin-hercynienne** qui traduit le démantèlement de la Pangée. Les accidents tardi-hercyniens qui limitent des bassins subsidents, sont réutilisés en failles normales. Les événements extensifs mésozoïques conduisent progressivement, à l'Ouest, à l'amincissement de la zone nord-pyrénéenne au Trias et, à l'Est, au rifting de l'océan Liguro-Piémontais au Lias. Cet océan se refermera suite aux déformations alpines.

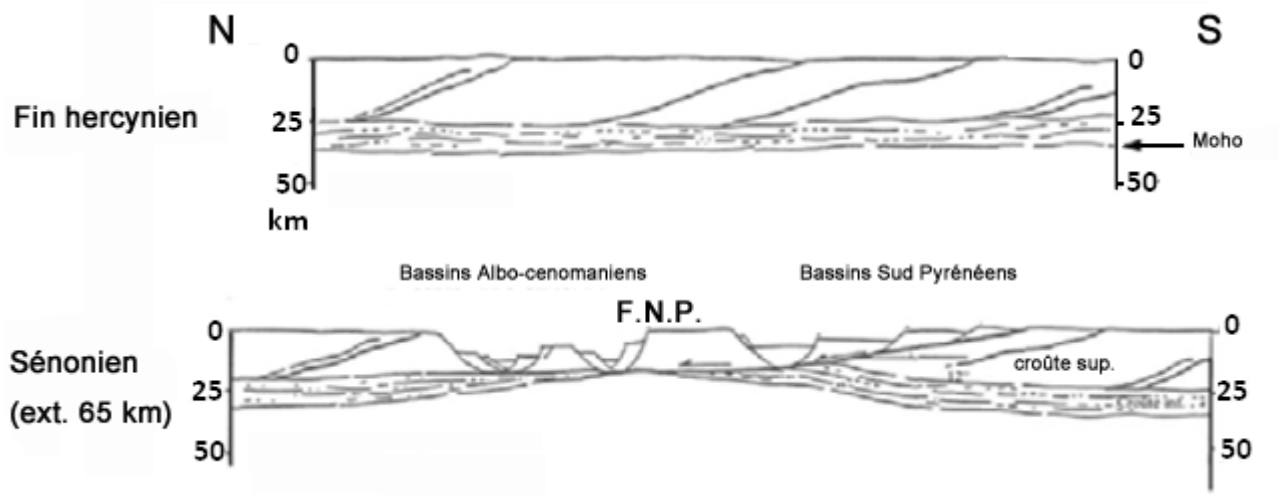


Figure I-7: Inversion des chevauchements hercyniens en failles normales lors de l'extension crétacée (inférieur et moyen), engendrant un amincissement crustal (65 km). Cette phase est relayée par l'extension albo-cénomanienne avec l'ouverture de bassin en « pull apart » liée à la migration de la plaque ibérique le long de la Faille Nord Pyrénéenne (FNP) qui joue donc en décrochement sénestre (Specht 1989)

L'apparition du basin aquitain se localise à l'emplacement actuel des Pyrénées (figure I-7), après le comblement des fossés permiens. Il est limité au Sud, par une série de grandes failles d'origine hercynienne parallèles à la future chaîne pyrénéenne et peut être divisé en deux parties par une ligne allant d'Arcachon à Toulouse (flexure celtaquitaine : (Dercourt 1997)). Au sud de cette ligne existe un bassin subsident à série complète reposant en discordance sur le Paléozoïque. Celui-ci est caractérisé au Trias supérieur par des séries évaporitiques à l'origine, plus tard, de la tectonique salifère. Au nord de cette ligne, une plate-forme peu subsidente se forme ; elle alterne, à partir de la fin du Trias, sédimentation en eaux profondes et périodes d'érosion.

Au Jurassique inférieur, une transgression forte fait apparaître une série évaporitique épaisse (800m) sur l'ensemble du bassin subsident et une plate forme jusqu'au massif central (Souque 2002). Jusqu'à la fin du Jurassique, c'est un bassin, en eaux généralement peu profondes, dont l'ensemble est émergé exception faite de quelques petits bassins.

Progressivement, la paléogéographie est transformée par la progression vers le nord de l'ouverture atlantique et par l'ouverture du Golfe de Gascogne qui lui est associé. En effet, au crétacé inférieur (Aptien), la séparation entre l'Ibérie et l'Europe (Ziegler 1988) précède l'ouverture océanique du Golfe de Gascogne, ouverture qui conduit à un mouvement relatif senestre (figure I-8) entre ces deux plaques le long d'une faille transformante reconnue comme étant la faille Nord Pyrénéenne (Le Pichon 1970; Choukroune 1978; Daignières 1981; Daignières 1982). L'absence d'ophiolites dans les Pyrénées suggère que l'extension, estimée à 100 km (Boillot 1984), n'y a pas atteint le stade de l'océanisation. Suit l'ouverture de l'Atlantique nord à l'Aptien-Albien (110 Ma).

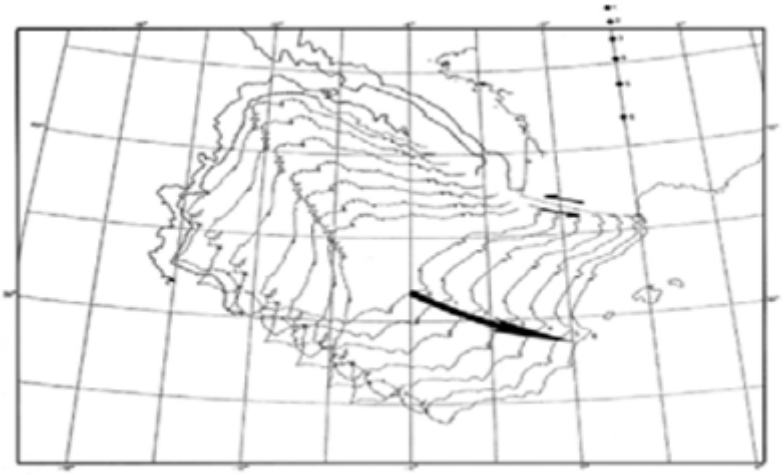
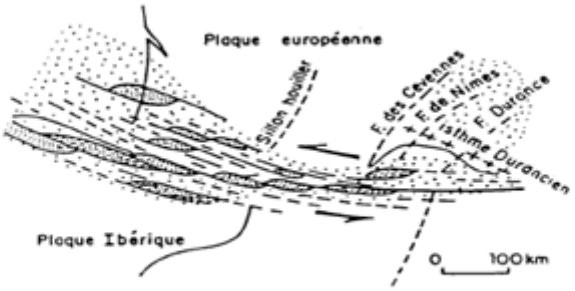


Figure I-8: Modèle d'ouverture du Golfe de Gascogne (Choukroune 1973). Les losanges montrent la migration du pôle de rotation vers le sud lors de l'ouverture du golfe de Gascogne. La faille Nord-Pyrénéenne joue alors en décrochement sénestre.

A l'Albien, l'apparition de fortes épaisseurs de turbidites dans le Bassin aquitain traduit une forte érosion, donc une reprise des reliefs et une première crise compressive dans les Pyrénées.

Par ailleurs, de l'Albien moyen au Cénomanien, dans le contexte de décrochements sénestres E-W limités au Sud par la Faille Nord-Pyrénéenne (FNP), on constate l'ouverture de bassins en « pull-appart » qui se mettent en place dans les flyschs (figure I-9). Cet épisode est associé à un métamorphisme HT restreint aux abords immédiats de la FNP (Choukroune 1978; Debroas 1987).



En pointillé lâche: domaine marin; pointillé: principaux bassins de sédimentation;
figuré serré: individualisation des sillons de flysch albiens. La forme losangique des bassins a été schématisée.

Figure I-9: Reconstruction du domaine pyrénéen albien (Choukroune 1978)

Durant l'extension, le mouvement relatif de la plaque ibérique par rapport à la plaque européenne est matérialisé par une rotation antihoraire dont le centre, initialement au Nord de la France, migrerait progressivement vers le Sud au cours du temps (figure I-8) (Choukroune 1973). Cette migration du pôle tient compte des observations dans le Golfe de Gascogne et permet le passage progressif d'un contexte extensif à un contexte compressif dans la zone pyrénéenne.

I.3.1.2 Tectonique Campano-Oligocène

Le passage de la phase extensive à la phase compressive débute dans les Pyrénées orientales au Sénonien (80-85 Ma) et correspond au blocage du système transformant senestre nord-pyrénéen dû à la collision des plaques ibérique et européenne.

A la fin du Crétacé, la collision est limitée à la partie orientale de la chaîne, la collision progresse ensuite vers l'ouest, atteint les Pyrénées centrales à l'Yprésien, et les Pyrénées occidentales avant le Lutétien (Choukroune 1973). La compression est tout d'abord localisée au niveau de la FNP et gagne des zones de plus en plus externes durant le Tertiaire, la plaque européenne passant sous la plaque ibérique du Paléocène à l'Eocène (Sibuet 1971; Boillot 1977).

Lors de la croissance de la chaîne, la plaque continentale se flexure en réponse à la surcharge des reliefs des zones internes et s'accompagne de l'apparition du sillon Nord-Pyrénéen au début du tertiaire (Danien) à l'ouest, avec développement progressif vers l'est. Ce sillon orogénique à sédimentation très épaisse, marine de type flysch, est connu sur toute la bordure Nord-Pyrénéenne et constituera la future zone sous–Pyrénéenne.

A la fin de l'Eocène, l'ensemble du domaine pyrénéen est soumis à un raccourcissement nord-sud. A l'Ilerdien, après une transgression marine majeure, le sillon Nord-Pyrénéen commence à se combler localement à l'est, avec passage d'une sédimentation de plate-forme carbonatée à une sédimentation fluviatile et lacustre connue sous le nom des « molasses de Carcassonne ». Le milieu marin est ensuite repoussé de plus

en plus à l'ouest. En réponse à l'orogènie pyrénéen, un bassin flexural se développe au nord des Pyrénées. La subsidence s'étale de la fin du Crétacé à l'Eocène terminal dans le bassin flexural aquitain.

I.3.1.3 Rifting Oligo-Miocène

Après la structuration des unités « alpines », donc post Eocène supérieur, débute l'ouverture du bassin provençal par rotation du bloc corso-sarde dans un contexte de bassin arrière-arc. La subduction en cause est l'actuelle subduction du sud-est de l'arc appenino-calabrais. L'extention entre le bloc corso-sarde et la plaque européenne débute à l'Oligo-Aquitanien pour atteindre la séparation crustale au Burdigalien (Souque 2002).

Dans l'avant pays nord-est pyrénéen, les structures extensives tertiaires se surimposent aux structures de chevauchement-plissement. Les accidents pyrénéens sont réactivés en extension. Les fossés oligo-miocènes sont limités à l'Ouest par des failles majeures à pendage vers l'est. Ces failles se localisent à l'aplomb de rampes de la nappe des Corbières et deviennent listriques en profondeur pour se raccorder aux surfaces de décollement éocènes (Gorini 1991).

La partie centrale du bassin aquitain qui nous intéresse plus particulièrement, comprend une grande partie des affleurements tertiaires et s'inscrit à l'intérieur de l'auréole de terrains secondaires. On peut y distinguer une lisière externe de pays molassiques à dominante calcaire, et une partie interne, vaste plaine découpée en terrasse aux pieds des coteaux par les grandes vallées de la Garonne et de ses affluents. Les formations molassiques ou « molasses » qui affleurent sur l'ensemble du bassin (figure I-10), sont le résultat final de l'évolution par diagenèse des sédiments issus du démantèlement des Pyrénées, déposés entre le Lutétien moyen et le Tortonien (50 et 7 Ma), dans le bassin subsident qui s'étend au-delà du piedmont français des Pyrénées.

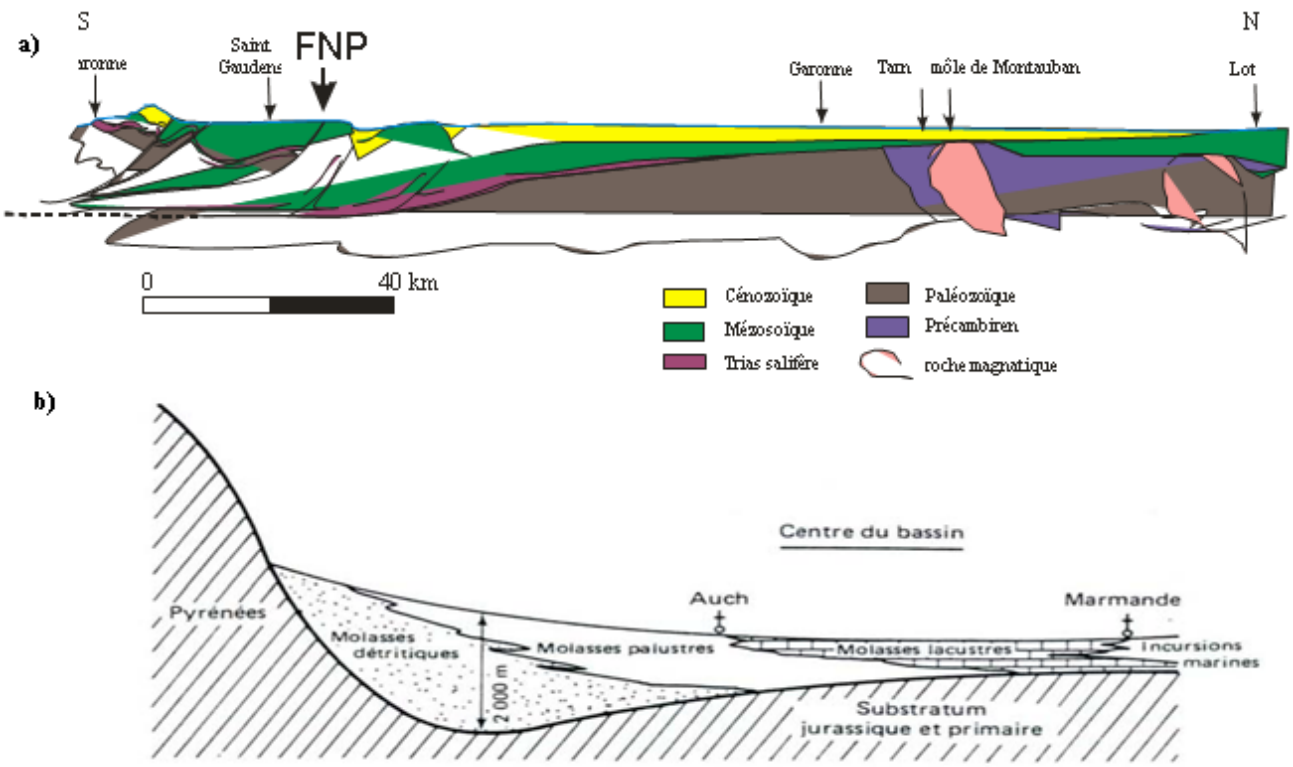


Figure I-10: Schémas synthétiques de la structure centre-orientale du Bassin aquitain : a) coupe structurale du Bassin aquitain entre Saint-Gaudens et la rivière du Lot prés de Montauban modifiée d'après(Winnock 1974) ; b) Variation horizontale d'ensemble des faciès molassiques (source : Laboratoire Régional des Ponts et Chaussés). Les coupes sont positionnées sur la figure 15.

Ces molasses sont souvent représentées dans un diagramme triangulaire (figure I-11) dont les trois pôles correspondent au quartz, aux carbonates et aux argiles. A l'intérieur de ce diagramme apparaissent des plages de faciès ayant des proportions déterminées des ces trois éléments.

Les molasses détritiques sont caractérisées par la présence de niveaux durs de grés et de poudingues. Le diagramme triangulaire fait bien ressortir pour ce type de molasse la répartition entre les faciès détritiques et palustres (pôle argile).

Les molasses palustres sont constituées des formations Oligocènes et Miocènes et représentent le groupe qui occupe la plus grande surface dans le Bassin aquitain. Globalement, le diagramme triangulaire montre une réduction importante des faciès détritiques et l'apparition du faciès calcaire et un fort enrichissement en argile. Ces molasses regroupent un ensemble de roches sédimentaires continentales dans lesquelles les proportions des éléments grossiers détritiques, des carbonates et des argiles sont extrêmement variables.

Les molasses lacustres sont caractérisées par le développement de niveaux calcaires assez continus que l'érosion met en relief lorsque les calcaires sont isolés dans un ensemble marneux.

Figure I-11: Représentation triangulaire schématique des faciès des molasses et des trois grands types de molasses (Enjalabert 1960).

I.3.2 Les formations quaternaires

Surmontant les molasses, les formations quaternaires sont constituées des alluvions de la Garonne et de ses affluents. La faible résistance des molasses à l'érosion a permis le creusement d'importantes vallées, dont l'évolution s'est produite surtout par migrations latérales vers l'Est pour les cours d'eau issus du versant Nord des Pyrénées et vers le Nord pour ceux issus de l'Ouest du Massif Central. De nombreux auteurs (Hubschman 1975 a; Hubschman 1975 b; Icole 1980; Icole 1982 a; Icole 1982 b) ont décrit l'étagement de ces différents dépôts fluviatils et ont distingué cinq terrasses dans le paysage. La chronologie relative donnée ici est dérivée d'observation de terrain, de datation paléontologique (Capdeville 1997; Chalard 1997 a; Chalard 1997 b) et anthropologique (figure I-12) et quelques rares datations au C14 pour l'Holocène (Hubschman 1975 b).

échelle des temps géologiques	Glaciations	terrasse	Altitudes moyennes (par rapport à l'étiage et au sud de Toulouse)	Archéologie
Holocène	Post glaciaire	T1	~0 (Lit actuel)	-
Pléistocène sup.	Würm	T2	10-15 m	Paléolithique supérieur et moyen
Pléistocène moyen	Riss-Mindel ?	T3	22-30 m/ 55-80 m	Paléolithique moyen / Paléolithique ancien
Pléistocène inf. terminal	Günz ?	T4	90-110 m	Paléolithique ancien ?
Pléistocène moyen et basal	Donau ?	T5	120-150 m	?

Figure I-12: Eléments de datation des terrasses de la Garonne obtenus à partir de l'activité humaine Paléolithique (modifié d'après Chalard et al., 1997b)

Lorsque la série des formations fluviatiles est complète, on constate les successions des alluvions suivantes (figure I-13) :

- **Hauts niveaux** (T5), Ils se rattacheraient aux phases glaciaires du Donau; les lambeaux culminants (il s'agit des dépôts *Ft* de la carte géologique de la région) qui dominent la Garonne ainsi que les autres terrasses et sont situés à 150 m au dessus du lit actuel. Les principaux points d'affleurement sont localisés à Puyjaudran et dans la forêt de Bouconne.

- **Hautes Terrasses** (T4), Elles seraient contemporaines de la fin de la glaciation du Günz et se présentent en plusieurs paliers d'altitude relative. En moyenne elles se situent à une altitude de 90 m par rapport au niveau d'étiage du fleuve ; elles affleurent sur le plateau de Rieumes et dans une partie de la forêt de Bouconne

- **Moyennes Terrasses** (T3), Elles sont datées de la fin du Mindel et du Riss et possèdent une remarquable unité géomorphologique, clairement visible au sein du système des terrasses. C'est sur ce type de dépôt alluvial que se situe la plupart des industries acheuléennes (Paléolithique inférieur) du bassin de la Garonne.

- **Basses Terrasses** (T2), toujours en plusieurs paliers polygéniques qui sont datés du Würmien, elles sont également bien marquées dans le paysage et surplombent la vallée d'une vingtaine de mètres ;

- **Basses Plaines** (T1), s'abaissant d'une quinzaine de mètres jusqu'au niveau de l'étiage actuel. Elles sont post-Würmiennes et proviendraient de la débâcle glaciaire « Wurmienne » (Younger Dryas,

environs 10000 ans). Ces alluvions récentes sont des matériaux détritiques en provenance des massifs pyrénéens. C'est un mélange très grossier, hétérogène, de sables, graviers, galets et blocs provenant des massifs rocheux anciens (Massif Central et Pyrénées). Elles peuvent être suivies de manière quasi continue sur l'ensemble de la région.

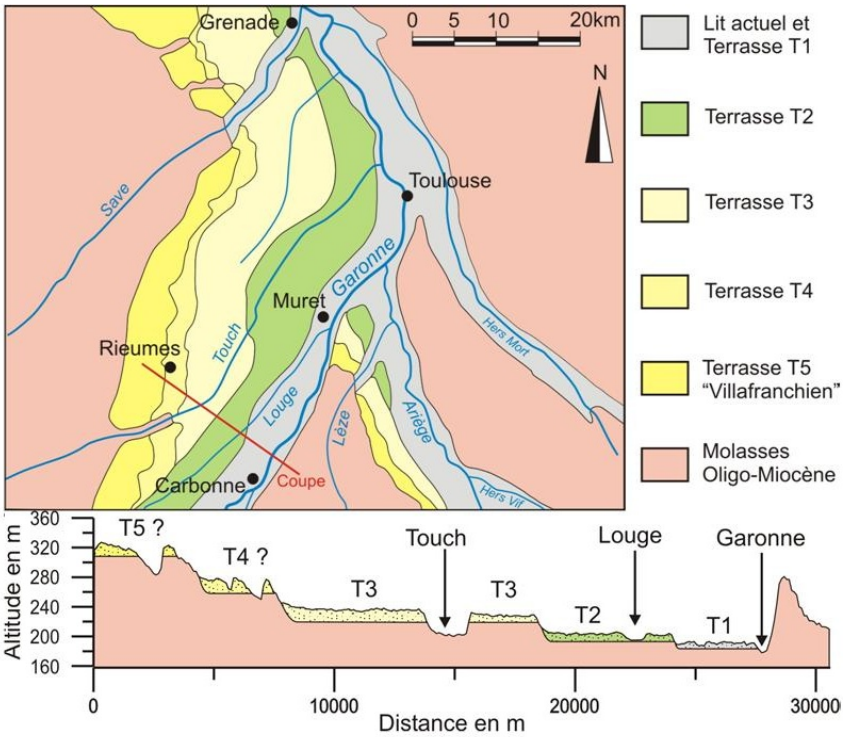


Figure I-13: Carte géologique et coupe des terrasses (entre Carbonne et Rieumes) (Maire E. (soumis)).

Les sols observés dans notre zone d'étude sont fortement hétérogènes est sont généralement décrits en grands ensembles morpho-pédologiques illustrés par la figure I-14 réalisée d'après les données de la chambre d'agriculture Midi-Pyrénées. On distingue rapidement :

- **Les basses plaines** qui sont en général constituées de sols de compositions hétérogènes à dominante de sols limoneux et à composante caillouteuses d'autant plus marquée que l'on se rapproche du piémont pyrénéen ; cela n'exclut pas la présence de bancs plus argileux notamment dans les anciens méandres.

- **Les terrasses alluviales** qui forment de vastes surfaces planes constituées par des sols localement appelés Boulbène. Ce sol se caractérise par un horizon de surface limoneux, battant, à très faible stabilité structurale, recouvrant des horizons d'accumulation d'argiles. Plusieurs types de boulbènes sont distingués, selon la profondeur des couches argileuses et selon la profondeur du cailloutis.

- **Les vallées secondaires** qui entaillent les terrains molassiques du tertiaire (Gascogne et Lauraguais) sont formées par des sols argileux plus au moins calcaires avec peu ou pas de cailloutis.

- **Les coteaux molassiques tertiaires** caractérisés majoritairement par le « terrefort » qui est un sol argileux, calcaire ou calcique, dont la composition varie suivant le relief et la lithologie et par des molasses sableuses à argileuses.

La spécificité de ces grands ensembles, dont la lithologie implique de nombreuses variations de faciès, souligne la problématique de notre étude. Ainsi, des variations d'ensemble des molasses peuvent être saisies à l'échelle régionale. Cependant, les formations molassiques et les formations superficielles montrent une forte hétérogénéité à grande échelle et apparaissent comme un matériau difficile à appréhender dans le détail. Les investigations ponctuelles (sondage, prélèvement, essais géotechniques…) réduisent les possibilités d'erreur mais ne peuvent pas les supprimer. De plus, si l'on réduit la maille, les mesures deviennent rapidement onéreuses. Ainsi, la localisation des zones sujettes au retrait-gonflement des argiles se heurte à des problèmes causés par les variations rapides de faciès. Pour être complète, la reconnaissance doit se faire à deux niveaux d'étude : un niveau ponctuel (continuité des faciès sous fondation…) et un niveau général (plan de prévention sur une commune…).

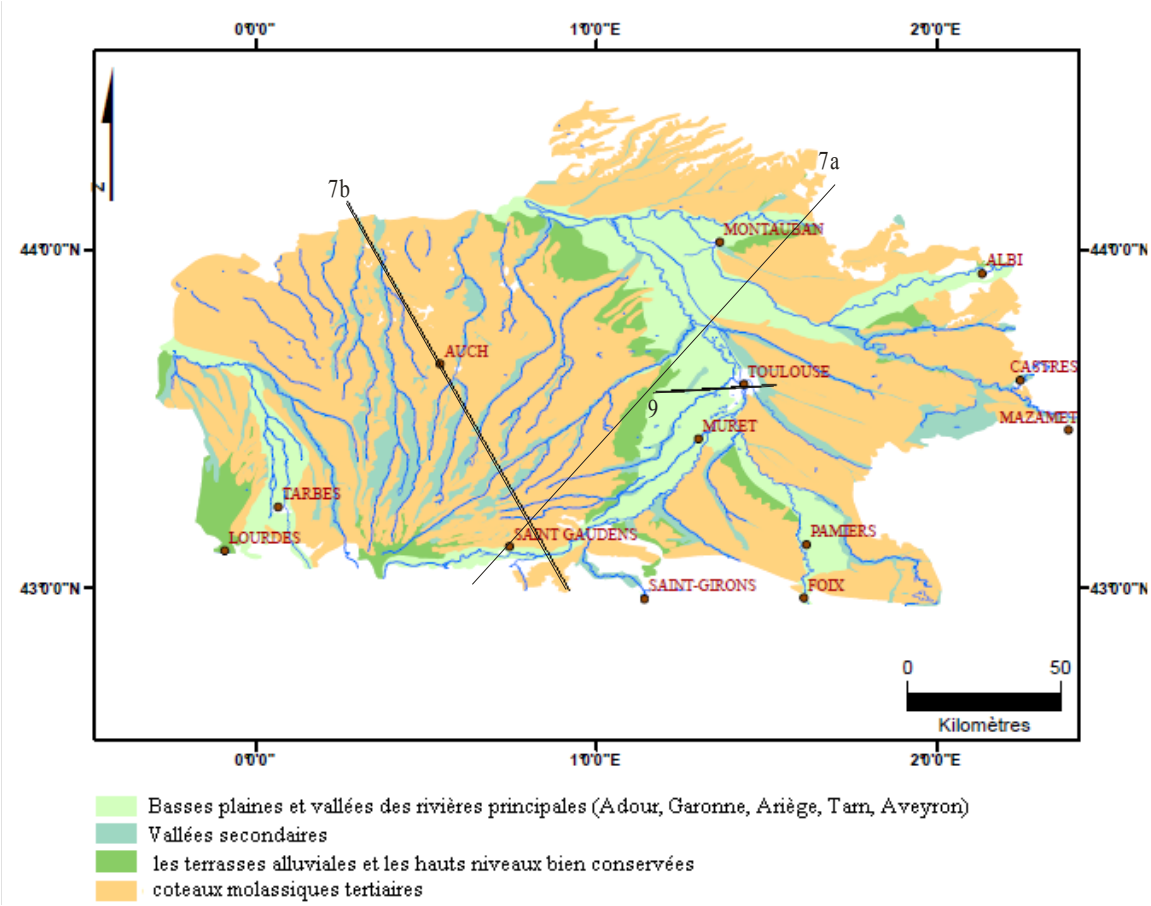


Figure I-14 : Grand ensemble morpho-pédologiques du Bassin central aquitain (source : Chambre Régionale d'Agriculture de Midi-Pyrénées). Système de coordonnées utilisé : coordonnées géographique du système de référence WGS 84.

I.4. DEVELOPPEMENT METHODOLOGIQUE

Ce manuscrit comporte quatre principaux chapitres. Il débute par la présentation du contexte d'étude et de la méthodologie suivie (chapitre 1). **Le chapitre 2** apporte des éléments de compréhension de spectroscopie infrarouge et fait le point sur les différentes méthodes et échelles d'analyses utilisées. Il présente également la base de données réalisée pour cette étude. Elle est constituée d'une collection de données spectrales de minéraux de référence et de mélanges synthétiques et d'une collection de sols prélevés dans la région Midi-Pyrénées. Ces collections ont été enregistrées par différents spectromètres aux différentes échelles de notre étude. La composition des échantillons qui constituent ces collections ont été validés par des analyses par diffraction des rayons X et des analyses géochimiques des éléments majeurs.

Le **chapitre 3** constitue une étude comparative des possibilités d'estimation de la composition minérale des sols à partir de deux méthodes classiquement utilisées pour l'analyse des propriétés des sols à partir de données spectrométriques. Dans un premier temps, nous présentons les capacités et les limites des méthodes statistiques telles que l'Analyse en Composantes Principales (ACP) et la Régression par Moindres Carrées Partiels (en anglais, PLSR). Ces méthodes ont été appliquées sur les spectres enregistrés dans le proche et le moyen infrarouge des échantillons des différentes collections. Dans un deuxième temps, nous détaillons le potentiel de la méthode du « continuum removal » et de l'analyse de la surface des figures d'absorption caractéristiques des phyllosilicates dans le cadre d'une utilisation des résultats par les bureaux d'étude.

L'analyse quantitative est fondamentale pour la localisation des zones riches en smectites. Le **chapitre 4** propose une utilisation originale de l'analyse en ondelettes couplée à la corrélation croisée appliquée aux données de spectrométrie infrarouge. Il montre que l'on est capable de réaliser une validation quantitative de la composition argileuse des sols indépendante des méthodes de validation utilisées dans le précédent chapitre (DRX, géochimie).

Le **chapitre 5** est une conclusion du travail réalisé dans la thèse et propose des perspectives pour l'utilisation de la méthode basée sur l'analyse en ondelettes ainsi que pour l'utilisation des principaux résultats de cette étude pour l'analyse des données multispectrales.

CHAPITRE 2.

La spectrométrie infrarouge :

Caractérisation et base de données

II. CHAPITRE 2. LA SPECTROSCOPIE INFRAROUGE : CARACTERISATION ET BASE DE DONNEES

Le principe de base de la spectrométrie repose sur la mesure de l'interaction entre un rayonnement électromagnétique et la matière à différentes fréquences. Elle permet d'obtenir des informations cristallochimiques sur la matière à partir de son interaction avec le rayonnement incident. Selon la fréquence de ce rayonnement (ultraviolet, visible, infrarouge…), l'interaction matière-rayonnement concerne divers types de niveaux d'énergie de la matière. Dans cette étude nous nous intéressons plus spécifiquement à la gamme de longueurs d'ondes de l'infrarouge. Un bref rappel des notions générales est fait, avant de présenter les approches utilisées pour l'enregistrement des spectres infrarouges aux différentes échelles d'analyse. Sont ensuite présentées les applications de ces techniques à la caractérisation des argiles, ainsi que les différents échantillons que nous avons étudiés en laboratoire, sur le terrain et par imagerie multispectrale. Les matériels utilisés dans cette thèse et les limites associées à chaque échelle d'analyse sont décrits. La diffraction des rayons X ainsi que des données de chimie ont constitué des données de référence et sont également présentées dans ce chapitre.

II.1. THEORIE DU RAYONNEMENT ELECTROMAGNETIQUE

Le principe de la spectrométrie infrarouge a été souvent détaillé (King P.L. 2004; Roy 2007). Le rayonnement infrarouge (IR) est une radiation de nature électromagnétique, correspondant à la partie du spectre comprise entre 12 800 cm^{-1} et 10 cm^{-1} (0.78 à 1000 µm). La figure II-1 présente la partie du spectre électromagnétique correspondant à l'infrarouge. La fenêtre spectrale de l'IR, se décompose habituellement en 3 parties : le proche, le moyen et le lointain IR. Nous nous intéresserons plus particulièrement à la région du proche (PIR) et du moyen infrarouge (MIR), qui s'étendent respectivement de 12 800 à 4000 cm^{-1} (0.78 à 2.5 µm) et de 4000 à 400 cm^{-1} (2.5µm à 25 µm). Dans ce manuscrit, nous utiliserons indifféremment la **périodicité spatiale** (longueur d'onde λ) ou la **périodicité temporelle** (fréquence ν) pour caractériser les ondes.

Les différents instruments généralement utilisés en spectrométrie de laboratoire couvrent les domaines du PIR (0.78 à 2.5 µm) et du MIR (2.5µm à 25 µm), tandis que la spectrométrie spatiale (spectrométrie de terrain et télédétection) couvre le visible et le proche infrarouge (VNIR = 25 000 à 7000 cm^{-1} soit 0.4 à 1.4 µm), l'infrarouge à courte longueur d'onde (SWIR= 7000 à 4000 cm-1 soit 1.4 à 2.5 µm) et pour certains capteurs, l'infrarouge thermique (TIR = 1000 à 700 cm^{-1} soit 10 à 14 µm). Les subdivisions différentes du domaine infrarouge suivant le champ de recherche sont principalement liées aux caractéristiques des spectromètres utilisés. D'autre part l'unité spectrale diffère suivant ces champs de recherche. Généralement, la minéralogie utilise les nombres d'onde (cm^{-1}) alors que la télédétection utilise les longueurs d'ondes (µm ou nm).

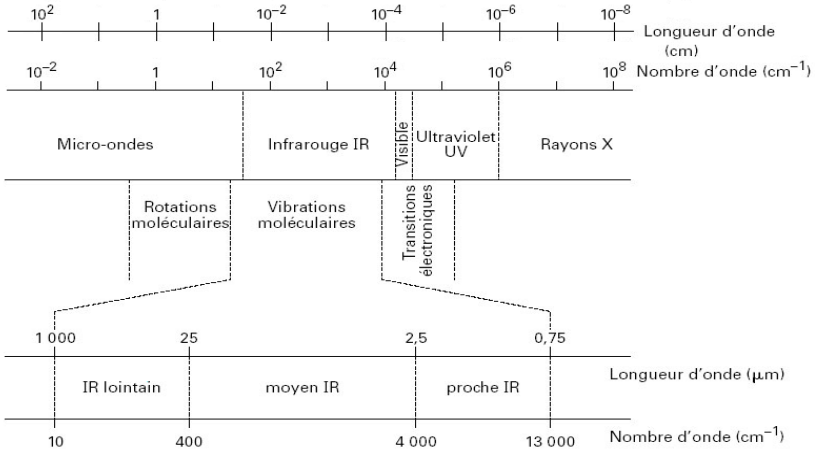


Figure II-1: Les grands domaines spectraux du rayonnement électromagnétique avec un focus sur le domaine de l'infrarouge. Les équivalences entre les échelles des longueurs d'onde et des nombres d'onde sont présentées.

L'examen de cette vaste gamme de longueurs d'ondes permet l'accès à une multitude de propriétés physiques et chimiques des matériaux observés. En effet, l'intensité du rayonnement incident étant connue, la mesure de l'intensité du rayonnement transmis, diffusé ou réfléchi, chargé d'information sur la matière exposée au rayonnement est possible. Une application de la spectrométrie est l'analyse de la cristallochimie des matériaux. Elle passe par la compréhension des bandes d'absorptions attribuables aux différentes propriétés physiques et chimiques de la matière. Elles peuvent donc être utilisées pour aider à l'identification des différents minéraux qui constituent ces matériaux.

II.1.1 Origines des bandes d'absorption

Le principe de base de la spectrométrie repose sur la mesure de l'interaction entre un rayonnement électromagnétique et la matière à différentes fréquences. L'objet de ce paragraphe est d'expliciter cette relation entre la matière et le rayonnement, qui véhicule une énergie liée à sa fréquence, de manière à permettre une meilleure compréhension des traitements et des analyses réalisés dans cette étude. Une description complète des processus synthétisés dans cette partie et appliqués à des domaines similaires à ceux de notre étude (géochimie et télédétection) existe dans l'ouvrage de King (2004) ainsi que dans l'article de (Hunt 1977).

Lorsqu'une molécule est irradiée par un rayonnement IR, elle peut absorber partiellement et sélectivement ce rayonnement et ainsi passer d'un état d'énergie E_1 vers l'état d'énergie supérieur E_2. Dans le cas de la spectroscopie infrarouge, le rayonnement incident entraine une modification de l'énergie vibrationnelle et rotationnelle de la molécule. Comme les liaisons des molécules ne sont pas parfaitement élastiques, les liaisons chimiques qui la composent doivent être considérées comme des oscillateurs anharmoniques (figure II-2). En effet, l'excitation des harmoniques supérieures demande,

progressivement, de moins en moins d'énergie supplémentaire et conduit à la dissociation de la molécule. On observe donc un rapprochement des niveaux vibrationnels quand l'énergie augmente. Ils ne sont donc plus équidistants comme dans le cas de l'oscillateur harmonique et la première harmonique a une fréquence qui est légèrement inférieure à deux fois la fondamentale.

Ainsi, lorsqu'une liaison oscille à la même fréquence que la composante électrique de l'onde IR, celle-ci transmet son énergie à la liaison. Plus précisément, un rayonnement de fréquence donnée étant absorbé il conduit à l'état excité, caractérisé par une amplitude de vibration plus élevée. De ce fait, l'énergie du rayonnement IR se retrouve diminuée après l'interaction, ce qui conduit à l'apparition d'une bande d'absorption à cette fréquence. **L'énergie absorbée (ν_o) est donc caractéristique de chacune des liaisons chimiques du minéral analysé.** Notons que s'il n'y a pas de dipôle permanent, c'est-à-dire si les charges électriques sont parfaitement symétriques, il n'y aura pas de couplage possible avec l'onde électromagnétique. Il n'y aura donc aucune absorption d'énergie (c'est le cas par exemple pour la molécule de KBr, généralement utilisée pour diluer les échantillons dans les méthodes de transmission). Les molécules non polaires sont « transparentes » dans l'IR, on parle de transition inactive. Une transition active possède donc nécessairement un dipôle dont le module fluctue avec la distance interatomique.

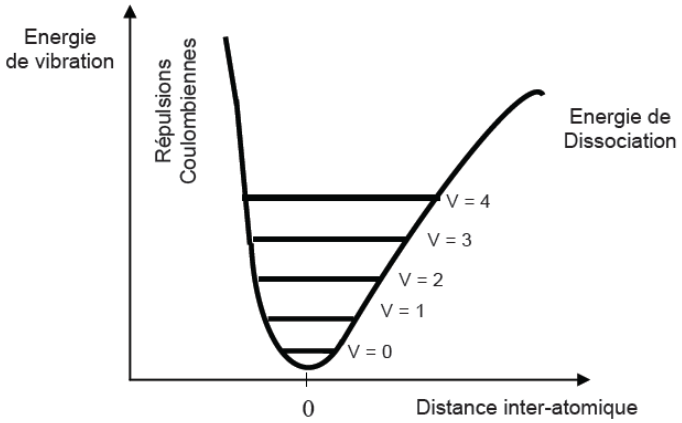


Figure II-2: Courbe de potentiel anharmonique d'une molécule diatomique: évolution de l'énergie potentielle de vibration *v* en fonction de la distance inter-atomique.

Le spectre IR est ainsi constitué de nombreuses bandes d'absorption. Il représente pour chaque longueur d'onde λ, le rapport des intensités transmises avec ($I_{éch}$) et sans (I_0) échantillon. On peut ainsi calculer la transmittance *T* :

$$\mathbf{T(\lambda) = I_{éch} / I_0} \qquad \textbf{équation 1}$$

Ainsi que l'absorbance *A* :

$$\mathbf{A(\lambda) = \log 1/ T(\lambda)} \qquad \textbf{équation 2}$$

Par conséquent, **tout minéral possède une signature qui lui est propre**. C'est pourquoi cette technique est très souvent utilisée pour caractériser des échantillons formés de plusieurs composants, tels que des roches ou des sols. L'étude des spectres permet de tirer deux types d'informations :

Des informations qualitatives : l'identification d'un composé est recherchée à partir de sa signature spectrale (celle-ci dépendant des niveaux d'énergie « sondés » par le rayonnement), par attribution des absorptions aux différentes liaisons chimiques en présence.

Des informations quantitatives : l'intensité de la bande d'absorption rend compte de la concentration du groupe chimique caractéristique de cette absorption. Dans des conditions idéales, c'est-à-dire dans le cas où un seul composé est responsable de l'absorption, en utilisant une gamme de longueurs d'onde appropriée et en maîtrisant le chemin optique, elle peut être modélisée par **la loi de Beer-Lambert** (éq. 3) qui fournit une relation entre l'absorbance A (λ) d'un échantillon et sa concentration molaire C :

$$\mathbf{A(\lambda) = \log 1/ T(\lambda) = \varepsilon(\lambda) * C * d} \qquad \textbf{équation 3}$$

Avec $\boldsymbol{\varepsilon(\lambda)}$: coefficient d'extinction molaire, fonction de la longueur d'onde λ

d : longueur du chemin optique

Cependant, cette relation se complique dès que plusieurs composés sont présents dans un mélange car les absorbances des différents composés se combinent alors. En première approximation, l'absorbance mesurée est considérée comme étant la somme des absorbances dues à chaque composé. Cette hypothèse suppose que chaque absorption du spectre de mélange n'est due qu'à un seul composé du mélange. De plus, des phénomènes physiques tels que la diffusion dans le milieu influencent également les absorptions. Ainsi, le modèle de Beer-Lambert traduit bien le lien qui existe entre la grandeur mesurée (le spectre) et la grandeur cherchée (la ou les concentrations des composés du mélange), mais le caractère approximatif de ce modèle appliqué à la spectrométrie, rend la démarche analytique complexe, car elle nécessite la connaissance des spectres d'interaction.

II.1.2 Interprétation des spectres

II.1.2.1 Dans le moyen infrarouge

Les spectres IR contiennent des informations pertinentes sur les caractéristiques physiques et chimiques des échantillons analysés. Ils sont composés de bandes d'absorption qui vont être utilisées pour déterminer la structure de la molécule étudiée ou du mélange à identifier. Les bandes d'absorption dans le MIR (800 – 4000 cm^{-1}) résultent des modes de vibration, ceux-ci, pouvant être attribués à des groupements fonctionnels fondamentaux. L'interprétation des spectres est empirique et consiste le plus souvent, en la comparaison des résultats obtenus avec ceux de la bibliographie. Elle est basée principalement sur l'analyse des spectres IR de molécules isolées. Une étude complète des

spectres d'absorbance des phyllosilicates a été réalisée par (Farmer 1974) et de nombreux auteurs ont largement par la suite, enrichie cette bibliographie notamment pour l'étude des phyllosilicates (Petit 1995; Janik 1998; Joussein 2001; Madéjova 2001; Bishop, Madejova et al. 2002; Madéjova 2003; Pironon 2003).

La position d'une bande dépend à la fois de la nature du vibrateur (et donc de la liaison) mais aussi de son environnement. Les liaisons entre atomes légers vibrent à plus haute fréquence que les liaisons entre les atomes plus lourds. Lorsque la masse diminue, la fréquence ν aussi et μ augmente, selon la loi de Hooke :

$$\mathbf{N = 1/\ 2\pi \sqrt{(k/\mu)}} \qquad \textbf{équation 4}$$

Avec : μ = $(m_1.m_2) / (m_1 + m_2)$

k : constante de la force de liaison

m_1 et m_2 : masse des atomes liés

L'intensité d'une bande dépend de la concentration, mais aussi de la nature et de la polarité de la liaison. De même, de nombreux facteurs externes ou internes à la molécule ont une influence sur la fréquence de vibration : taille et électronégativité des atomes environnants, liaison hydrogène, etc….

II.1.2.2 Dans le proche infrarouge

On nomme harmoniques les fréquences $n\nu_0$ multiples de la fréquence fondamentale ν_0 d'un oscillateur. Nous avons vu précédemment que l'anharmonicité des vibrations modifie l'expression de la fréquence d'une transition. Elle a aussi pour conséquence de provoquer l'apparition de **bandes dites harmoniques** correspondant à des transitions, de fréquences 2ν, 3ν,… d'intensité faible par rapport à la bande de fréquence ν.

Une autre conséquence de cette anharmonicité est le couplage de deux vibrations de fréquence ν_1 et ν_2 en une vibration de fréquence somme $\nu_1 + \nu_2$ ou différence $/\nu_1 - \nu_2/$. On parle alors de **combinaisons** respectivement additive ou soustractive des bandes de fréquence ν_1 ou ν_2.

Le spectre d'absorption enregistré dans ce domaine est donc complexe, les bandes d'absorption de toutes les liaisons chimiques se chevauchant. L'interprétation du spectre est de fait difficile. Néanmoins, elle offre une grande potentialité pour les problèmes de spatialisation de l'information spectrale et ses applications se multiplient, notamment en science du sol du fait de l'utilisation de spectromètre travaillant dans ce domaine de longueurs d'ondes (Ben-Dor and Banin 1995; Goetz 2001; Viscarra Rossel, Cattle et al. 2009).

II.1.3 Les perturbations engendrées par l'atmosphère sur le rayonnement électromagnétique

Le rayonnement électromagnétique peut être transmis par différentes sources : des sources artificielles (quartz halogène par exemple) ou le soleil dans les domaines du visible, du proche et du moyen infrarouge et le sol dans les domaines du thermique et des micro-ondes. En laboratoire, les appareils utilisés fonctionnent avec des sources artificielles. Cela permet notamment de s'affranchir des problèmes engendrés par la traversée de l'atmosphère par le rayonnement électromagnétique. Dans le cas de la spectrométrie de terrain et de la télédétection, le soleil est utilisé comme source du rayonnement électromagnétique. Lors de son passage dans l'atmosphère, le rayonnement solaire, émis ou rétrodiffusé par les objets observés au sol, subit plusieurs processus qui ont pour conséquence d'appauvrir le rayonnement solaire par les différentes couches atmosphériques (figure II-3). Ils sont liés aux réactions qui se produisent entre des parties spécifiques du spectre électromagnétique solaire et les différents gaz atmosphériques et leurs concentrations aux différentes élévations. Ainsi, on distingue trois processus majeurs suivant la longueur d'onde du rayonnement et les caractéristiques physiques des gaz et des particules atmosphériques. La **diffusion ou dispersion** et l'**absorption** du rayonnement par des composantes gazeuses et des particules solides et liquides, ainsi que la **réflexion** par les nuages (particules solide et liquides) et les aérosols. Des fenêtres atmosphériques apparaissent ainsi dans les zones ne présentant pas ou peu d'absorption atmosphérique. Elles définissent ainsi les domaines de longueurs d'ondes dans lesquels les mesures sont possibles.

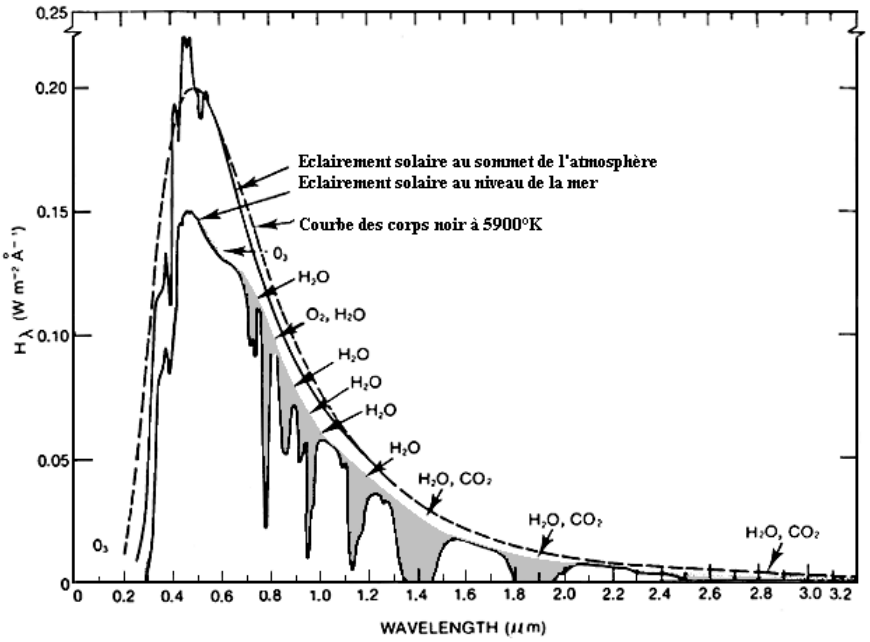


Figure II-3: Courbes d'éclairement solaire modifié par les absorptions atmosphériques dans le visible et le proche infrarouge. Les zones grisées correspondent aux domaines de longueurs d'ondes où se produit une absorption du rayonnement par l'atmosphère et où aucune information spectrale n'est disponible lors de mesures in situ ou par télédétection.

II.2. INSTRUMENTS DE MESURE UTILISES EN SPECTROSCOPIE

Dans ce paragraphe, sont présentés les instruments de mesures qui nous ont permis d'enregistrer les spectres infrarouges des échantillons prélevés et synthétisés dans un environnement contrôlé en laboratoire ou directement sur le terrain. Nous présentons également les données du capteur multispectral ASTER.

II.2.1 La spectroscopie infrarouge à transformée de Fourier

II.2.1.1 Principe

Pour notre étude, nous avons utilisé un spectromètre à transformée de Fourier Perkin- Elmer Nicolet 5700, dont le schéma général de principe est décrit en figure II-4 (LMTG, UMR5563 CNRS).

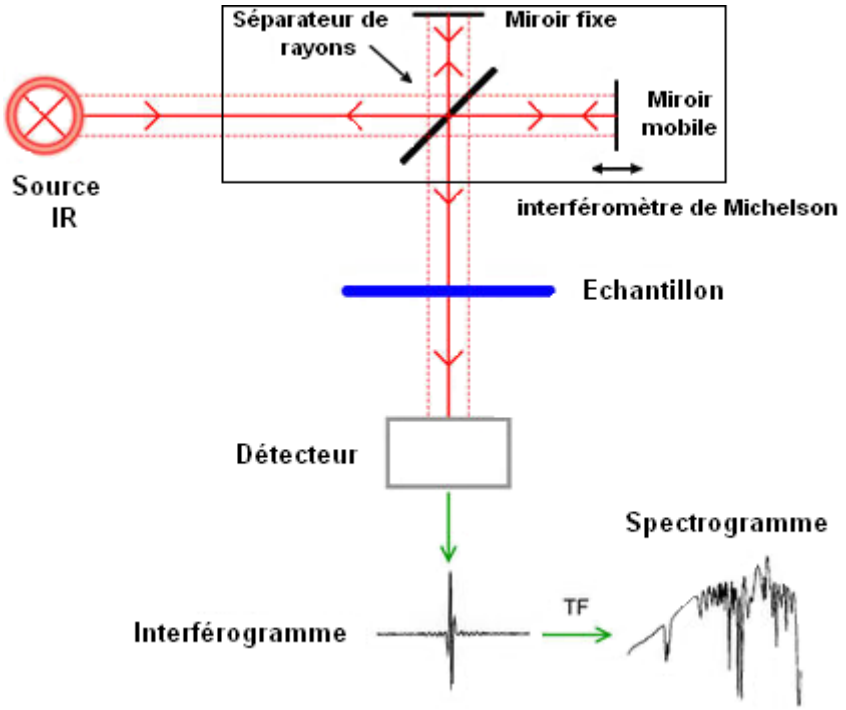


Figure II-4: Schéma de principe de l'analyse par spectroscopie d'absorption infrarouge

Dans ce type d'appareillage, le faisceau infrarouge provenant de la source est dirigé vers l'interféromètre de Michelson qui va moduler chaque longueur d'onde du faisceau à une fréquence différente. Dans l'interféromètre le faisceau arrive sur la séparatrice. La radiation de la source est alors divisée en deux faisceaux par un séparateur de faisceaux (Beamsplitter en quartz). Un des faisceaux parcourt un chemin optique fixe, l'autre un chemin optique de longueur variable à cause d'un miroir mobile, avant d'être recombinés, de traverser l'échantillon et de frapper le détecteur. Quand la différence de chemin optique entre les faisceaux correspond à un multiple entier de la longueur

d'ondes d'une bande, on obtient une interférence constructive. Une interférence négative est obtenue lorsque la différence correspond à un multiple entier impair du quart de la longueur d'onde. L'ensemble de ces interférences positives et négatives transmises à travers un signal complexe produit un interférogramme à partir duquel le spectre est calculé par transformée de Fourier. L'avantage de ce dispositif est d'avoir accès instantanément à la densité optique pour toute une gamme de longueurs d'ondes (Avantage de Fellgett). D'autre part ce type d'appareil permet d'avoir une résolution spectrale et une précision spectrale élevées, de travailler avec des signaux faibles et d'obtenir le spectre rapidement et avec un bon rapport signal sur bruit.

En regard des différentes réponses de la matière à l'excitation par un rayonnement IR, nous avons utilisé dans cette étude deux modes de spectrométrie IR.

II.2.1.2 Dans le domaine de l'infrarouge moyen (4000 - 400 cm^{-1}) : la transmission/ absorption.

Nous avons utilisé la spectrométrie en mode transmission (TR) où le faisceau IR émis traverse l'échantillon. La différence entre l'intensité d'émission et de réception permet de calculer l'absorbance. Cette méthode nécessite d'utiliser un support transparent en infrarouge tel que le bromure de potassium (KBr). La méthode consiste à broyer quelques milligrammes (0.5 à 1 mg) de l'échantillon en présence de poudre de KBr sec, dans un mortier en agate. Le mélange est ensuite comprimé dans une presse hydraulique, puis mis à l'étuve durant 24h afin de supprimer l'humidité résiduelle.

La radiation infrarouge est produite par chauffage d'une source qui est souvent un filament de Nernst (alliage d'oxyde de Zr, Th et Ce). La séparatrice de l'interféromètre est en KBr.

II.2.1.3 Dans le domaine du proche infrarouge (9000 – 4000 cm^{-1}) : la réflexion diffuse

Nous avons utilisé la spectroscopie en mode réflexion diffuse qui est très sensible aux vibrations d'élongation ou de déformation des groupements atomiques des phases superficielles. Dans ce mode, le faisceau interagit plusieurs fois avec l'échantillon au travers de phénomènes combinés de diffusion, réflexion ou d'absorption sur des particules orientées de façon aléatoire (figure II-5). Cela permet d'approcher le mode d'échantillonnage utilisé sur le terrain. L'accessoire qui permet ces mesures est constitué d'un jeu de miroirs qui dirigent et récupèrent le rayonnement sous une géométrie biconique. Cet instrument nous permet d'analyser directement les échantillons en poudre, sans prétraitement préalable. L'échantillon est placé dans une coupelle, la poudre est compactée et la surface est égalisée à l'aide d'une spatule. De cette manière, une légère rugosité aléatoire est donnée à la surface. Pour normaliser ces spectres, un miroir en or dont la surface est lisse et présente un comportement

spéculaire, est utilisé comme référence. Cette différence de comportement se traduit sur le spectre par des faibles valeurs de réflectance en sortie.

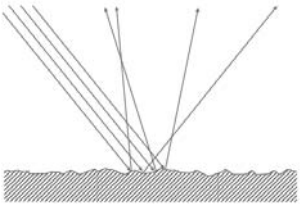


Figure II-5: Interaction faisceau matière: réflexion diffuse.

Dans ce cas nous avons utilisé un faisceau de lumière blanche émis par un filament en tungstène qui permet de travailler dans la gamme de longueurs d'ondes du proche infrarouge (9000-4000 cm^{-1}). La séparatrice est en CaF_2. Les analyses réalisées dans cet intervalle sont comparables aux domaines couverts par les instruments utilisés sur le terrain et en télédétection.

II.2.2 Spectromètre ASD Field Spec FR® (Roy, 2007)

Pour le spectromètre de terrain, le spectromètre ASD FieldSpec FR® est choisi (figure II-6). Il constitue un lien parfait entre le spectromètre de laboratoire utilisé en mode réflectance et celui de la plupart des satellites « optiques ». Il permet de faire des enregistrements suivant deux configurations de mesures différentes. D'une part, **en laboratoire** les échantillons sont analysés à l'aide d'une sonde qui contient à la fois la source d'éclairement artificielle (lampe halogène) et la fibre optique. Cette sonde est ensuite plaquée à l'échantillon pour analyser sa composition. D'autre part, il utilise le soleil comme source d'éclairement sous sa **configuration de terrain**. Une fibre optique permet de récupérer le signal réfléchi par la surface. Les mesures se font suivant une visée verticale et des conditions d'éclairement idéales (mi-journée, ciel clair).

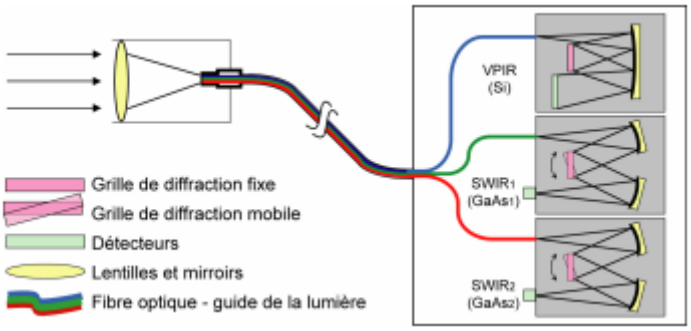


Figure II-6 Chemin optique et caractéristiques des détecteurs du spectromètre ASD FieldSpec FR® (MacArthur 2006).

Sous ces deux configurations, le rayonnement électromagnétique est récupéré et dirigé vers trois détecteurs (figure II-6). Le premier est en silice et couvre la gamme de longueurs d'ondes comprise

entre 0.35 et 1 µm (domaine du visible et du proche IR en imagerie). Une grille de diffraction fixe disperse la lumière qui est ensuite enregistrée par le détecteur en fonction des différentes longueurs d'ondes. L'intervalle compris entre 1.00 et 2.50 µm (domaine du moyen infrarouge en imagerie) est observé par deux détecteurs InGaAs. La lumière est alors dispersée à l'aide d'une grille mobile. Ainsi, les deux détecteurs récupèrent le rayonnement pour des longueurs d'ondes différentes.

L'avantage du laboratoire est que l'épaisseur de l'atmosphère est négligeable tandis qu'avec le soleil, celle-ci empêche de travailler dans les gammes de longueurs d'ondes situées autours de 1400 et 1900 nm où la vapeur d'eau fait écran (Cf. II.1.3). Ainsi, l'utilité première du spectromètre ASD FieldSpec FR® est l'analyse in-situ ; Elle permet à la fois d'étudier de manière quasi continue l'évolution de la composition du sol sur une parcelle et de créer des points de calibration lors de l'utilisation de données de capteurs hyperspectraux ou multispectraux. Cependant, son utilisation en condition de laboratoire ajoute une étape fondamentale pour l'étude du saut d'échelle entre l'analyse en laboratoire (spectromètre IRTF) et l'analyse in-situ. Elle permet de ne tenir compte que du changement de résolution spatiale et spectrale en limitant les effets liés à l'état de surface de l'échantillon (rugosité, pierrosité, humidité,…) et en supprimant les effets liés aux absorptions atmosphériques (figure II-3) présents lors de l'analyse in-situ.

II.2.3 Capteur Multispectral ASTER

Le capteur multispectral ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer) est un des cinq instruments embarqués sur la plateforme TERRA- EOS (Earth Observing System) lancée en 1999. Il est né d'une coopération entre la NASA, le ministère Japonais de l'Economie, du Commerce et de l'Industrie (METI) et l' « Earth Remote Sensing Data Analysis Center » (ARSDAC). Le système imageur permet d'obtenir des données de forte résolution spatiale dans le domaine du visible, du proche infrarouge, du moyen infrarouge et de l'infrarouge thermique sur une surface d'environ 60 km par 60 km. Le capteur ASTER est lui-même assez complexe puisqu'il est constitué de trois instruments indépendants (Abrams 1995) :

- Un instrument enregistre la réflectance dans les 3 bandes spectrales du Visible et Proche Infrarouge (VNIR).
- Un deuxième instrument enregistre la réflectance dans les 6 bandes spectrales de l'Infrarouge Moyen (SWIR).
- Un troisième instrument enregistre la réflectance dans les 5 bandes spectrales du Thermique Infrarouge (TIR).

Le tableau II-1 détaille les 15 bandes spectrales correspondant à chaque instrument.

D'autre part, l'instrument réalisant les mesures dans le VNIR est constitué de deux télescopes et systèmes d'acquisition. Le premier télescope vise au nadir et capte les données selon les trois bandes

spectrales du VNIR (bandes 1, 2 et 3N). Le deuxième télescope vise vers l'arrière et capte les données selon une bande spectrale (3B). Ce système permet de produire des Modèles Numériques de Terrain (MNT) de haute résolution spatiale (15 à 30m). Il existe depuis octobre 2009, une base de données de MNT ASTER (GDEM, disponible sur http://www.ersdac.or.jp/GDEM) qui ont une résolution spatiale de 30 m.

date de lancement	18 décembre 1999 à Vandenberg, Californie, USA		
passage à l'équateur	10 : 30 AM (du nord vers le sud)		
orbite	705 km d'altitude, héliosynchrone		
période de revisite	16 jours		
résolution	15 à 90 m		
instruments	**VNIR**	**SWIR**	**TIR**
bandes	1 - 3	4 - 9	10 - 14
résolution spatiale	15 m	30 m	90 m
Fauchée	60 km	60 km	60 km
option	images stéréoscopiques, modèles de terrain		

Tableau 1: Conditions d'acquisition des données ASTER et leurs caractéristiques spectrales (modifié d'après l'ASTER USER's guide).

La localisation des bandes spectrales utilisées par ASTER est représentée sur la courbe de transmission atmosphérique (figure II-7). Grâce à sa fine résolution spatiale et l'importante couverture du spectre électromagnétique, le satellite ASTER présente un grand potentiel pour la cartographie géologique de part sa qualité spectrale, sa résolution spatiale mais aussi son coût qui présente le meilleur rapport qualité/ prix du marché.

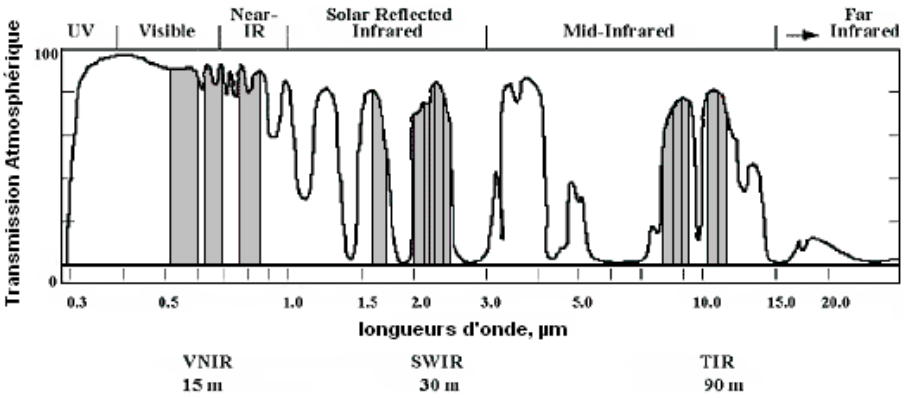


Figure II-7: Localisation des bandes spectrales ASTER sur le spectre de transmission atmosphérique.

Deux niveaux de données sont disponibles pour les images ASTER. Les données du niveau L1A contiennent les valeurs brutes du signal transmis par le satellite en valeurs numériques. Les coefficients de conversion de la radiance (intensité d'un rayonnement) au niveau du capteur et les corrections géométriques sont fournis dans les métadonnées, mais aucune transformation n'est appliquée. Ce niveau de données est particulièrement intéressant pour la création des Modèles

Numériques de Terrain. Dans le cas des données au niveau 1B, les valeurs numériques du niveau 1A sont converties en valeur de radiance et des transformations géométriques ont été appliquées pour enregistrer l'image dans le système de coordonnées géographiques. Du fait que ces données aient été ré-échantillonnées, elles sont moins appropriées pour la création du Modèle Numérique de Terrain. En revanche, le niveau 1B peut être utilisé pour l'analyse spectrale sans prétraitement.

II.3. MOYENS DE VALIDATION DES DONNEES ACQUISES

II.3.1 La diffraction des rayons X

Ce paragraphe donne les grands principes de la méthode de diffraction des rayons X. Une présentation détaillée est disponible dans la thèse de (Hubert 9 déc. 2008).

La diffraction des rayons X est une méthode adaptée à l'identification de la composition minéralogique des assemblages argileux. Elle consiste à appliquer un rayonnement d'une longueur d'onde du domaine des rayons X (entre 0.1 et 10 nm) sur un échantillon orienté ou non. L'analyse des phases cristallines est rendue possible grâce aux périodicités de l'arrangement atomique des cristaux. Ces périodicités sont dues à un empilement de plans identiques dans un cristal et sont décrites par des longueurs correspondant aux distances entre les plans d'empilement. Cette distance entre les plans réticulaires est nommée distance réticulaire (figure II-8). Ainsi, lorsque le rayonnement pénètre le cristal, une partie de l'énergie est absorbée entrainant l'excitation des atomes et l'émission de radiations par les plans atomiques. Le faisceau diffracté mesuré résulte de l'ensemble des interactions constructives des rayons émis par les atomes d'un même plan. Selon la loi de Bragg :

$$2d \sin\theta = n\lambda \qquad \textbf{équation 5}$$

n nombre entier correspondant à l'ordre de diffraction

λ longueur d'onde du rayonnement

d distance réticulaire

θ angle de diffraction

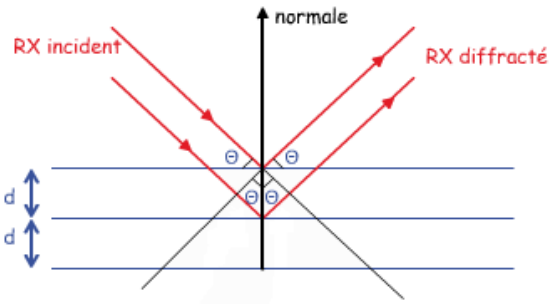


Figure II-8: Schéma de diffraction des rayons X par une famille de plan réticulaire. d est la distance réticulaire, θ est l'angle de Bragg.

Après diffraction on obtient un diffractogramme du faisceau diffracté en fonction de l'angle de diffraction 2θ ou de la distance réticulaire d (en Å). L'identification des phases présentes dans un échantillon se fait par comparaison de son diffractogramme (positions et intensités des raies diffractées) avec des diffractogrammes de phases connues.

Deux méthodes d'investigation par diffraction des rayons X sont classiquement appliquées par les laboratoires d'analyse. Dans un premier temps une analyse sur poudre totale fournit les éléments d'une semi-quantification de tous les minéraux présents et notamment la fraction totale regroupant minéraux phyllosilicates. D'autre part, l'analyse de la fraction argileuse (inférieure à 2 µm) extraite du matériau est préparée sous forme de lames orientées. Elle fournit une bonne estimation des divers composants argileux grâce à l'application de trois diagnostiques : « normal », avec l'éthylène glycol et enfin après chauffage (500°C).

Les estimations de la concentration des composants minéraux peuvent être obtenues en utilisant la méthode du « Peak-height Ratio » et donnent uniquement des indications sur les contenus. Dans cette méthode, la distance verticale entre l'apex de chaque pic intéressant et la ligne de base est mesurée et multipliée par le facteur correctif correspondant. Les résultats de chaque multiplication sont ensuite totalisés pour les différents minéraux identifiés, puis leurs pourcentages respectifs dans le matériau sont établis suivant une simple règle de trois. En général, l'intensité est mesurée sur le pic de référence (réflexion 001).

Dans notre étude la première méthode, sur poudre totale, est utilisée pour estimer la proportion du contenu en minéraux phyllosilicatés à partir de la mesure de l'intensité de la réflexion commune (020) vers 4,15Å. Cette intensité est multipliée par un facteur 20 (Marcoen 2001). La plupart des minéraux non argileux, qui sont généralement présents en petites quantités en association avec les minéraux argileux, sont enlevés lors de prétraitements chimiques (décarbonatation, suppression de la matière organique à H_2O_2). Les minéraux argileux sont ensuite étudiés à partir des résultats de l'analyse sur lame orientée en multipliant l'intensité mesurée par un facteur correctif tenant compte de la forme du pic et donc de l'état de cristallinité du minéral ainsi que de paramètres liés à la technique (Thorez 1975; Thorez 1976.) :

Minéral	réflexion utilisée	facteur multiplicatif correctif (J.Thorez 1975, 1976)
illite	10 Å (EG)	1.0
chlorite	14 Å (EG)	0.34
vermiculite	[14 Å (EG) - 14 Å (Ch)]	0.34
smectite	[10 Å (Ch) - 10 Å (EG)]	1.0
kaolinite	7 Å (EG)	0.7

Tableau 2: Caractérisation des principaux minéraux argileux et facteurs correctifs associés.

En outre, une identification et une quantification univoque des phases dans un mélange de plusieurs composés, sont soumises à plusieurs sources d'erreurs qui sont relatives à l'instrumentation utilisée, aux méthodes de préparation et de prétraitements, ainsi qu'aux caractéristiques du minéral (Brindley 1980). Les facteurs inhérents à ce dernier sont difficilement contrôlables. Ils sont liés d'une part à la variabilité chimique et structurale intrinsèque du minéral argileux et d'autre part à l'orientation préférentielle des particules argileuses qui est fonction de la taille, de la densité et de la morphologie des particules et enfin à la similarité de structure entre certains minéraux argileux qui se traduit par des réflexions très proches qui peuvent se superposer.

Les analyses par DRX ont été, soit fournies par le BRGM (données traitées dans le cadre de la réalisation des cartes d'aléa retrait-gonflement des départements de la région Midi-Pyrénées), soit réalisées au LMTG. Elles portent sur des minéraux de référence et sur des sols de la région Midi-Pyrénées. Dans le premiers cas, la diffraction des rayons X permet une vérification de la mono-minéralité de ces minéraux dits « purs ». Dans le deuxième cas, la diffraction des rayons X sert de méthode de référence. Elle est utilisée comme moyen de détermination et de semi-quantification des phases minérales argileuses contenues dans les sols. Dans le but d'améliorer cette « quantification » des minéraux argileux, les analyses par DRX réalisées au LMTG sont associées aux données de chimie des éléments majeurs (Cf II.3.3).

II.3.2 Analyses géotechniques

Ce type d'analyse décrit le comportement géotechnique du matériau vis-à-vis du retrait-gonflement. Ils produisent une estimation de la sévérité des déformations différentielles verticales potentielles sur le site.

Essai	Paramètres déterminés	référence
Limites d'Atterberg	Plasticité, Consistance	ASTM D 4318
Limite de liquidité	Indice de plasticité	ASTM D 4318
Limite de plasticité	Indice de liquidité	ASTM D 4318
Limite de retrait	Coefficient de retrait	ASTM D 427 XP P94-060.1.
Fraction argileuse	pourcentage de passant au tamis de 2 µm	ASTM D 422
Capacité d'échanges cationiques	CEC (meq / 100g)	Chapman (1965)
Gonflement libre	Potentiel de gonflement	Holtz et Gibbs (1956)
Gonflement à volume constant	Indice de gonflement	Lambe (1960)
Gonflement sous charge constante	Indice de gonflement	ASTM D 4546
Coefficient d'extensibilité linéaire (Cole)	Déformation linéaire d'un sol naturel séché d'une succion initiale de 33kPa à l'état sec	Brasher et al. (1966)
Essai au bleu de méthylène	Capacité d'absorption du sol	AFNOR NF P 18-592

Tableau 3: Principaux tests géotechniques pour l'identification des sols gonflants en laboratoire (Nowamooz 19 déc. 2007).

La palette d'essais utilisée dans la bibliographie varie en fonction de l'objectif et de la zone d'étude. En effet, l'homogénéisation des méthodes entre les différents pays n'est pas toujours faite. Les plus fréquemment utilisées sont résumées dans le tableau II-3.

Du fait du prix des analyses géotechniques et du temps nécessaire pour les réaliser, un minimum d'échantillonnage est fait pour chaque projet de construction. Typiquement, un échantillon est prélevé entre 1 et 3m de profondeur (profondeur de fondation) pour chaque parcelle de maison individuelle. De plus, les essais les plus fréquemment utilisés par les bureaux d'étude sont finalement : l'indice de plasticité (I_p), la valeur de bleu (VBS) ainsi que la granulométrie.

L'indice de plasticité est calculé à partir des limites d'Atterberg qui mettent en évidence l'influence de la teneur en eau sur la consistance du matériau fin. Cet indice correspond à la différence entre la limite de liquidité et la limite de plasticité du matériau. Il représente donc l'étendue du domaine plastique et donne une indication sur l'aptitude du matériau argileux à acquérir de l'eau. On considère généralement (Chassagneux D. 1995) que la susceptibilité d'une argile vis-à-vis du retrait-gonflement varie en fonction de l'indice de plasticité comme indiqué dans le tableau II-4.

L'essai au bleu de méthylène permet d'évaluer la surface spécifique d'échange d'un matériau argileux, ce qui constitue un bon indicateur de sa susceptibilité au phénomène de retrait-gonflement. Il consiste à mesurer la capacité d'adsorption en bleu de méthylène, c'est-à-dire la quantité de ce colorant nécessaire pour recouvrir d'une couche mono-élémentaire les surfaces externes et internes de toutes les particules argileuses présentes dans 100 g de sol. On appelle cette quantité, la valeur de bleu, notée VB et exprimée en gramme de bleu pour 100 g de matériau.

indice de plasticité	valeur au bleu	susceptibilité	Note
$I_p < 12$	< 2.5	faible	1
$12 \leq I_p < 25$	2.5 à 6	moyenne	2
$25 \leq I_p < 40$	6 à 8	forte	3
$I_p \geq 40$	> 8	très forte	4

Tableau 4: Barème d'évaluation de la susceptibilité du retrait-gonflement en fonction de l'indice de plasticité et de la valeur de bleu.

Enfin, la granulométrie correspond à la répartition pondérale en argile (< 2 µm), limon (2 à 20 µm puis 20 à 50 µm) et sable (50 à 200 µm puis 0.2 à 2 mn) de l'échantillon. Elle permet de déterminer la texture et de préciser le comportement des sols.

II.3.3 Chimie des éléments majeurs

L'objectif est d'une part de vérifier la « pureté » des minéraux et d'autre par d'avoir une estimation de la composition des sols par une autre méthode que la diffraction des rayons X. Pour cela, la chimie des éléments majeurs des minéraux argileux de référence et de certains sols a été faite par ICP-AES au Service d'Analyse des Roches et Minéraux au Centre de Recherche Pétrographique et

Géologique de Nancy. L'équipement comprend quatre parties : le dispositif de nébulisation, la torche à plasma, le dispositif dispersif et un détecteur. Pendant l'ionisation, chaque élément va émettre un spectre lumineux caractéristique dû aux mouvements intercouches de ses électrons. Ces spectres vont être triés par longueurs d'ondes par l'intermédiaire d'un réseau de diffractions (système dispersif) puis les rayons seront ensuite transformés en intensité par l'intermédiaire d'un phototube (détecteur). L'intensité mesurée est proportionnelle à la concentration de l'élément considéré, la conversion s'effectuant grâce au passage de solutions étalon.

Associées à une identification de la composition minéralogique faite par diffraction des rayons X et aux données de l'analyse chimique des minéraux, les données de l'analyse chimique des sols permettent une quantification précise comparativement à l'estimation faite à partir des données de diffraction des rayons X. l'approche consiste à considérer le bilan de masse entre les compositions chimiques mesurées des échantillons étudiés (minéraux de référence, sols) et les compositions chimiques classiquement considérées des minéraux monophasiques (tableau II-5), en principe, de même nature que ceux identifiés par la diffraction des rayons X dans les échantillons analysés. La principale limite à cette approche est liée à la composition chimique de certains minéraux tels que les smectites, qui est très variable. En effet, dans notre cas, une composition moyenne fixe pour les éléments structuraux est utilisée (Tableau II-5). D'autre part, la quantification est faite en prenant pour hypothèses que le potassium (K) renseigne uniquement sur la proportion d'illite et que les minéraux interstratifiés, difficile à prendre en compte dans ces analyses, sont supposés négligeables. Enfin, ces analyses sont d'autant plus précises que la phase est dominante dans les mélanges.

	SiO_2	Al_2O_3	Fe_2O_3	FeO	MnO	MgO	CaO	Na_2O	K_2O	TiO_2	P_2O_5	CO_2	H_2O
Quartz	100	—	—	—	—	—	—	—	—	—	—	—	—
Kaolinite	46,55	39,5	—	—	—	—	—	—	—	—	—	—	13,96
Illite	54,01	17,02	—	1,85	—	3,11	—	—	7,26	—	—	—	12,03
Chlorite	30,28	17,13	—	15,09	—	25,39	—	—	—	—	—	—	12,11
Vermiculite	11,92	43,48	—	12,82	—	9,4	—	—	4,5	—	—	—	17,87
Smectite	43,77	18,57	—	—	—	—	1,02	1,13	—	—	—	—	36,09
Plagioclases	66,57	18,83	—	—	—	—	10,36	5,72	—	—	—	—	—
Calcite	—	—	—	—	—	—	56,03	—	—	—	—	44	—
Dolomite	—	—	—	—	—	21,86	30,41	—	—	—	—	47,7	—
Goethite	—	—	89,86	—	—	—	—	—	—	—	—	—	10,14
Hématite	—	—	100	—	—	—	—	—	—	—	—	—	—
Magnétite	—	—	68,97	—	—	—	—	—	—	—	—	—	—
Rutile	—	—	—	—	—	—	—	—	—	100	—	—	—
Apatite	—	—	—	—	—	—	55,07	—	—	—	41,8	—	0,59

Tableau 5: Données de l'analyse chimique de minéraux argileux utilisés pour la quantification minéralogique.

II.4. BASE DE DONNEES

Une présentation simple des différentes familles d'argiles est faite dans la première partie de ce paragraphe afin de se familiariser avec ce type de matériel. Les références utilisées comme support sont essentiellement (Brindley 1980; Moore 1989; Bouchet 2000). La présentation des différentes collections utilisées dans la thèse est ensuite faite.

II.4.1 Les minéraux argileux

Un résumé de la structure générale des minéraux argileux et des caractéristiques majeures des principales familles de phyllosilicates est présenté de manière à donner le contexte nécessaire pour la compréhension des figures spectrales associées à ces minéraux.

II.4.1.1 Structure des minéraux argileux

Les minéraux argileux appartiennent à la famille des phyllosilicates. Leur structure est un empilement de couches tétraédriques et octaédriques qui forme les feuillets. Les tétraèdres sont liés les uns aux autres par des atomes d'oxygène. Ils forment ainsi un réseau planaire d'hexagones d'atomes d'oxygène. Les octaèdres sont liés par un côté. L'agencement des couches permet de distinguer trois principaux types de minéraux :

- Type 1 :1 ou T –O (1 couche octaédrique et 1 couche tétraédrique) équidistance voisine de 7 Å
- Type 2 :1 ou T -O- T (1 couche octaédrique entre 2 couches tétraédriques) équidistance variable entre 9 et 15 Å, suivant le remplissage ou non de l'espace interfoliaire ;
- Type 2 :1 :1 ou T –O-T –O est similaire au type précédent mais l'espace interfoliaire est occupé par une couche d'octaèdres. L'équidistance caractéristique est alors de 14 Å environ.

Chimiquement, les tétraèdres sont constitués de quatre oxygènes (trois planaires et un apical) encadrant généralement un atome de Si^{4+} ou Al^{3+}, et plus occasionnellement Fe^{3+} et Cr^{3+}. Idéalement, une demi-maille élémentaire est définie par trois octaèdres. Quoi qu'il en soit, l'électroneutralité de la maille est réalisée avec 6 charges positives obtenues par le remplissage cationique de deux ou trois cavités octaédriques. On peut ainsi distinguer deux grands groupes parmi les familles de minéraux argileux : les minéraux dioctaédriques (deux cavités sur trois sont remplies par des cations trivalents) et les minéraux trioctaédriques (les trois cavités sont remplies par des cations divalents). Les différentes substitutions possibles dans le tétraèdre ou dans l'octaèdre créent un déficit de charge cationique dans la structure. Ce déficit, s'il existe, est compensé au plus prés par la charge du remplissage interfoliaire, soit par des cations (Ca^{2+}, Mg^{2+}, Na^{+}, K^{+}, NH_4^{+}), soit par une couche octaédrique (cas des chlorites), qui agit également sur la distance entre les feuillets (figure II-9).

II.4.1.2 Classification des minéraux argileux les plus fréquents dans les sols de l'étude

L'illite $(K, H_2O)_2\ Si_8\ (Al, Fe, Mg)_{4.6}\ O_{20}\ (OH)_4$

Les illites appartiennent à la famille des phyllosilicates 2 :1. Le feuillet élémentaire est composé d'une couche octaédrique d'aluminium comprise entre deux couches tétraédriques de silicium. Dans les couches tétraédriques de silicium un ion Si^{4+} sur quatre est remplacé par un ion Al^{3+}. Le déficit de charge qui en résulte est compensé dans l'espace interfoliaire par des ions potassium K^+ qui assurent des liaisons fortes entre les feuillets.

Les smectites $(OH)_4\ Si_8\ (Al_{10/3}, Mg_{2/3})\ O_{20}, nH_2O$

Les smectites, dont la montmorillonite et la bentonite, appartiennent à la famille des phyllosilicates 2 :1 gonflant. Le feuillet élémentaire est caractérisé par l'empilement de 2 couches tétraédriques séparées par une couche octaédrique. Les atomes de silicium constituent majoritairement les tétraèdres mais des cations (Al, Fe) peuvent se substituer au Si dans ces tétraèdres. Les couches octaédriques sont quant à elles composées d'Al, Mg, Fe^{2+}, Fe^{3+},… (Leger 18 septembre 1997). Les smectites ont, de par leur structure, la nécessité d'équilibrer leur charge dans l'espace interfoliaire par l'insertion de cations dans cet espace. Ces cations, peu liés à la structure, sont susceptibles d'être échangés par d'autres cations. C'est ce que l'on appelle la capacité d'échange cationique (CEC). Cette structure permet aussi l'insertion, entre les feuillets de molécules telles que l'eau, le glycérol, ou d'autres molécules organiques. Cette insertion se traduit par un gonflement de la structure. Ces variations sont dépendantes de la nature de la molécule, des conditions de température, de pression et de la nature du cation compensateur.

Les vermiculites $(Mg,Ca)_{0,7}(Mg,Fe,Al)_6(Al,Si)_8O_{22}(OH)_4.8H_2O$

Les vermiculites appartiennent à la famille des phyllosilicates 2 :1. Elles constituent une famille proche de celle des smectites, mais elles sont caractérisées par un déficit de charge plus important. Elles sont majoritairement trioctaédriques. Le déficit de charge est essentiellement dû aux substitutions tétraédriques et la compensation est assurée dans l'espace interfoliaire par des cations (Mg^{2+} principalement) et des couches d'eau.

Les chlorites $(OH)_4\ (Si\ Al)_8(Mg, Fe)_6O_{20}$

Les chlorites appartiennent à la famille des phyllosilicates 2 :1 :1. Il est nécessaire de citer les chlorites du fait que leurs propriétés se rapprochent de celles des argiles et que ce minéral peut être présent en quantité non négligeable dans les sols les moins évolués de la région. Leur structure est cependant plus complexe, et se compose d'un feuillet 2 :1 entre lequel s'insère un feuillet de brucite.

La kaolinite (Si_4O_{10}) Al_4 $(OH)_8$

Les kaolinites appartiennent à la famille des phyllosilicates 1 :1. Le feuillet élémentaire est composé d'une couche tétraédrique de silicium et d'une couche octaédrique d'aluminium. D'un feuillet de kaolinite à l'autre des ions hydroxyles OH^- des couches octaédriques sont en vis-à-vis d'un plan contenant des ions d'oxygène O^{2-} de la couche tétraédrique ; dans ce cas, les liaisons interfoliaires résultent de l'effet composé de liaisons hydrogène et de forces de Van der Waals (figure II-9).

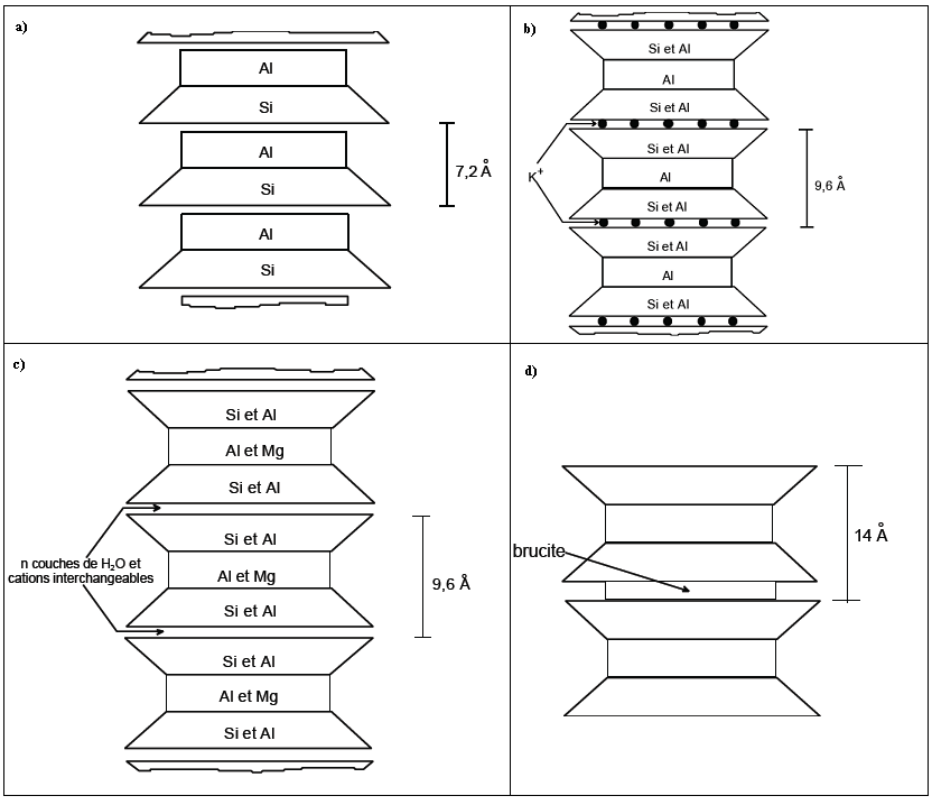


Figure II-9: Schéma de la particule de a) kaolinite, b) illite, c) smectite et d) chlorite

Les caractéristiques de ces phyllosilicates sont résumées dans le tableau II-6 suivant :

nom	type	diamètre d'une particule (µm)	surface spécifique en m2 / g	C.E.C en meq/ 100g
Kaolinite	1:1	0.1 - 4	10 - 30	3 - 15
Illite	2:1	0.1 - 1	100 - 175	25 - 40
Smectite (montmorillonite)	2:1	0.1	700 - 840	80 - 100
Vermiculite	2 :1	0.1	760	100 - 150
Chlorite	2:1:1	0.1	20	5 - 15

Tableau 6: Caractéristiques des différentes familles d'argile.

Les minéraux interstratifiés

Ils résultent de l'empilement aléatoire ou ordonné suivant l'axe c, de plusieurs espèces de minéraux argileux (deux en général). Les combinaisons possibles entre les différents minéraux sont extrêmement nombreuses, mais seulement quelques unes représentent la majorité des interstratifiés naturels. Ces phyllosilicates sont difficiles à mettre en évidence par la spectrométrie infrarouge. Ils ne sont pas pris en compte dans notre étude.

II.4.2 Les collections

Des mélanges de minéraux argileux et des sols, ainsi que les assemblages minéralogiques correspondants, ont été caractérisés en laboratoire et sur le terrain. Cette partie présente les données de caractérisation associées et les principales figures utilisées pour identifier les phyllosilicates dans les études de laboratoire et de terrain. Les propriétés spectrales sont associées aux données de chimie et de diffraction des rayons X. Elles sont nécessaires pour vérifier la composition chimique des minéraux de référence et la structure cristalline des minéraux. Ce diagnostic est fondamental pour l'analyse des figures spectrales caractéristiques des minéraux argileux dans les mélanges réalisés à partir de ces minéraux de référence et les sols (chapitre 3, 4 et 5).

II.4.2.1 Les minéraux argileux de référence

Une étape de sélection des minéraux argileux précède leur reconnaissance spectrale. Plusieurs minéraux représentant les trois principales familles d'argiles (illite, smectite et kaolinite) rencontrées dans les sols de la région et provenant de différentes sources, ont été collectés, analysés par diffraction des rayons X et par analyses chimiques, et enfin sélectionnés suivant qu'ils étaient constitués d'une phase minérale unique ou largement majoritaire. Les minéraux sélectionnés sont :

- une kaolinite naturelle commercialisée par la société « AGS minéraux » (Clérac, France), une kaolinite naturelle Ska commercialisée par la Société des Kaolins d'Armor et une kaolinite purifiée SKb commercialisée par la Société des Kaolins de Beauvoir.

- deux illites naturelles l'une provenant du Puy en Velay en France et l'autre commercialisée par la société Argiletz. Une troisième illite sans référence a été testée. Elle est nommée illite USA.

- trois smectites provenant respectivement du Wyoming (USA), de la société CECA (Honfleur, France) et de la société ABM (France).

La composition chimique de ces minéraux et les résultats de leurs analyses par diffraction des rayons X sont résumés dans les tableaux II-7et II-8.

	Kaolinite SKb	Kaolinite Ska	Kaolinite Clérac	Smectite ABM	Montmorillonite Wyoming	Bentonite CeCa	Illite LRPC	illite Velay	illite USA
SiO_2	46,85	71,28	44,65	48,34	58,80	52,59	47,98	49,46	53,51
Al_2O_3	35,84	19,76	36,77	15,38	18,07	15,14	13,02	19,18	11,97
Fe_2O_3	1,22	0,68	0,73	3,72	3,41	5,92	5,16	6,38	2,35
MnO	0,03	0,01	0,00	0,03	0,01	0,13	0,07	0,041	0,009
MgO	0,21	0,21	0,07	5,23	2,24	4,03	5,17	3,52	4,66
CaO	0,06	< L.D.	0,12	2,35	1,10	1,49	7,50	1,14	3,14
Na_2O	0,04	0,06	0,05	0,08	2,07	2,95	0,16	0,28	0,08
K_2O	2,08	1,70	0,30	1,20	0,43	0,50	3,26	7,79	6,14
TiO_2	0,05	0,18	1,27	0,31	0,15	0,67	0,69	0,71	0,99
P_2O_5	0,28	0,03	< L.D.	0,07	0,05	0,16	0,15	0,44	0
LOI=H_2O	12,29	6,64	14,51	22,22	13,72	17,25	16,97	11,23	12,3
Total	98,96	100,55	98,47	98,92	100,05	100,84	100,12	100,16	101,16

Tableau 7: Chimie des éléments majeurs des minéraux argileux de référence (analyses réalisées par le Service d'Analyse des Roches et des Minéraux, CRPG, Nancy). ***(L.D. = limite de détection)***

nom	phase principale	autres minéraux
Kaolinite SKb	85% kaolinite	5% quartz _ 10% illite
Kaolinite Ska	50% kaolinite	45% quartz _ 5% illite
Kaolinite Clérac	99% kaolinite	traces de quartz et d'illite
Smectite ABM	92% smectite	5%illite _ 2%quartz, pyrite, microcline
Montmorillonite Wyoming	99%smectite	traces de quartz, goethite, plagioclase
Bentonite CeCa	95% Interstratifiés illite/smectite	2%quartz _ traces de pyrite, zéolite, plagioclase

Tableau 8: Semi-quantification sur diffractogramme de RX (Source : BRGM Orléans).

Cette base de données de minéraux de référence est complétée par d'autres phyllosilicates généralement minoritaires dans les sols (chlorite, vermiculite, micas) et par des minéraux non argileux tels que de la calcite, du quartz, des feldspaths (plagioclase, albite) et des oxydes de fer (hématite, goethite). Ces minéraux servent pour les méthodes d'analyses qui s'intéressent au spectre dans son ensemble, telles que les méthodes statistiques (chapitre 3) ou la décomposition en ondelettes (chapitre 4). Pour ces minéraux, un diagnostic par DRX a été réalisé.

Pour la réalisation des mélanges synthétiques les minéraux suivant ont été utilisés. Ils sont choisis du fait de leur caractère quasi « pur » (Cf tableau II-8):

- la kaolinite de Clérac (Annexe 1 : diffractogrammes).

- la smectite de la société ABM (les Argiles du Bassin Méditerranéen) et la montmorillonite provenant du Wyoming (Annexe 1: Diffractogrammes).

- l'illite LRPC et l'illite du Puy en Velay (Annexe 1 : Diffractogrammes).

Les autres minéraux argileux ont servi de manière complémentaire pour étudier l'influence de l'utilisation de minéraux complexes (interstratifiés ou phase pluri-minérales) sur les relations fondamentales établies sur la base des minéraux précédemment cités.

Spectrométrie infrarouge

Comme nous l'avons montré dans les paragraphes précédents, la spectroscopie infrarouge est extrêmement sensible aux variations de la structure cristallographique et de la composition chimique des minéraux. Ainsi les compositions

chimiques variables des minéraux argileux présentés dans le tableau II-7 vont avoir un impact important sur la forme générale des spectres. Cette information disponible dans les spectres de minéraux est résumée dans un diagramme de « signatures spectrales ». Ce diagramme fourni une référence pour l'interprétation des figures spectrales de mélanges simples (mélange de deux ou trois minéraux argileux) et de données plus complexes (sols).

Le moyen IR est constitué des bandes de vibrations fondamentales (cf. II.1.2.1) et une importante documentation facilite la lecture des spectres soit de manière générale (Farmer 1974), soit par des études ponctuelles (Petit 1995; Janik 1998; Petit 1998; Madéjova 2001; Bishop, Madejova et al. 2002; Madejova 2002; Madéjova 2003; Viscarra Rossel 2006; Janik mars-avril 2007). La figure II-10 illustre les caractéristiques des différentes familles d'argile et les différents minéraux au sein d'une même famille à partir de leurs spectres infrarouge.

Dans le domaine d'élongation des liaisons hydroxyles (υ_{OH}) entre 3800 et 3600 cm^{-1}, la variabilité au sein d'une même famille peut venir de la présence d'une autre phase (présence de kaolinite dans le cas de l'illite LRPC et de l'illite USA qui se traduit par l'apparition d'un pic à 3697 cm^{-1}), du degré de cristallinité (défaut de cristallinité de la bentonite CECA qui se traduit par des bandes d'absorptions mal définies (pic à 1037 cm^{-1})), ou plus généralement, de la chimie de la couche octaédrique et de la couche tétraédrique (présence de trois ou quatre pics d'absorption dans la zone d'élongation des OH pour les kaolinites (3620, 3652, 3669 et 3697 cm^{-1}) et position du pic d'absorption (proche de 3620 cm^{-1}) des groupements OH pour les smectites, par exemple).

Le domaine de déformation et d'élongation des liaisons Si-O et de déformation des liaisons hydroxyles (δ_{OH}) entre 1300 et 400 cm^{-1} est également sensible à ces paramètres. Ainsi la présence de quartz dans les kaolinites Ska et SKb apparaissent notamment sous la forme d'un doublet proche de 780 cm^{-1}. Les différences dans l'arrangement des feuillets se traduisent par des changements de forme et de position des pics d'absorption. Enfin, différentes substitutions dans les sites octaédriques peuvent expliquer la présence de pics supplémentaires entre les minéraux d'une même famille.

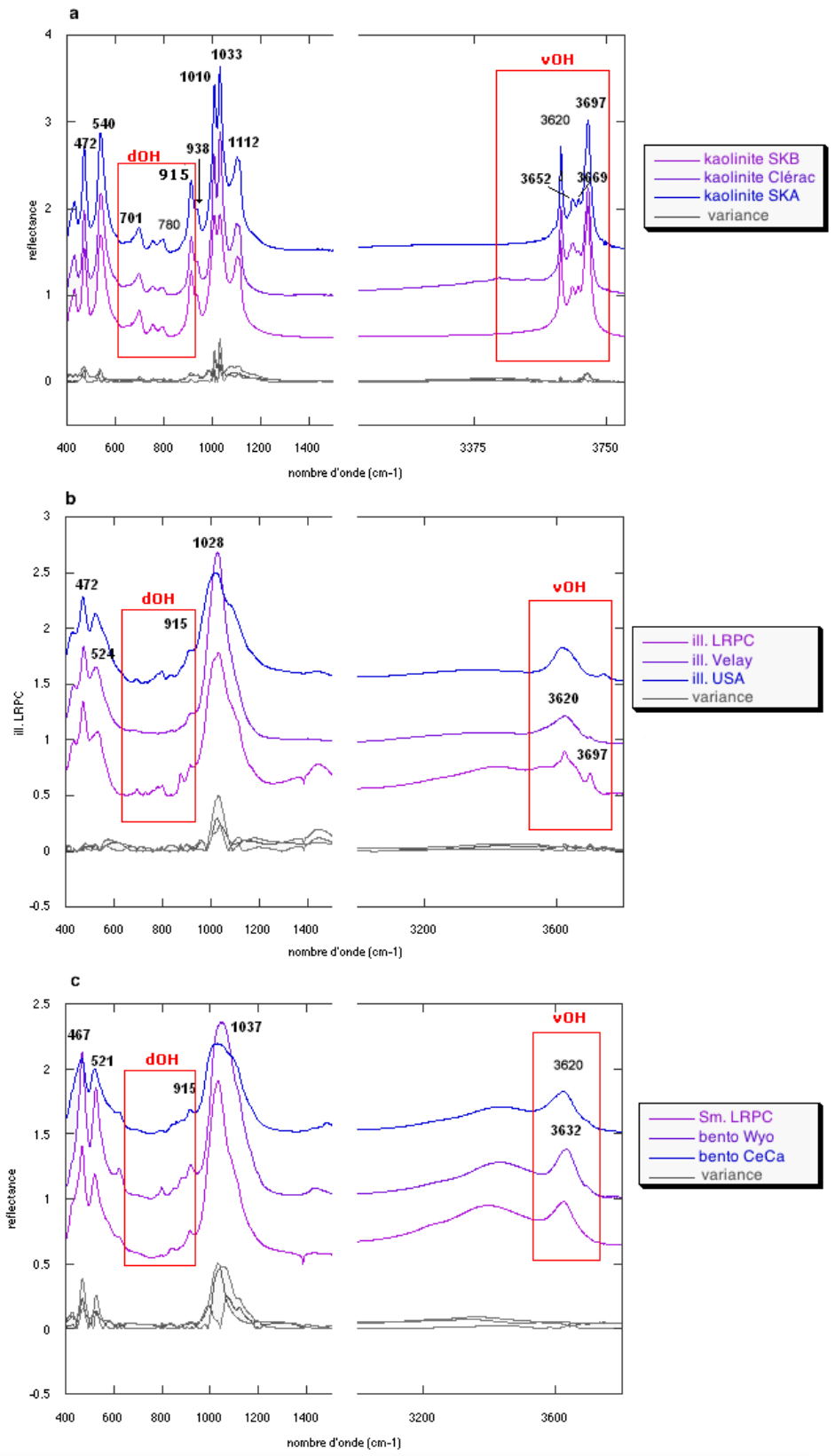
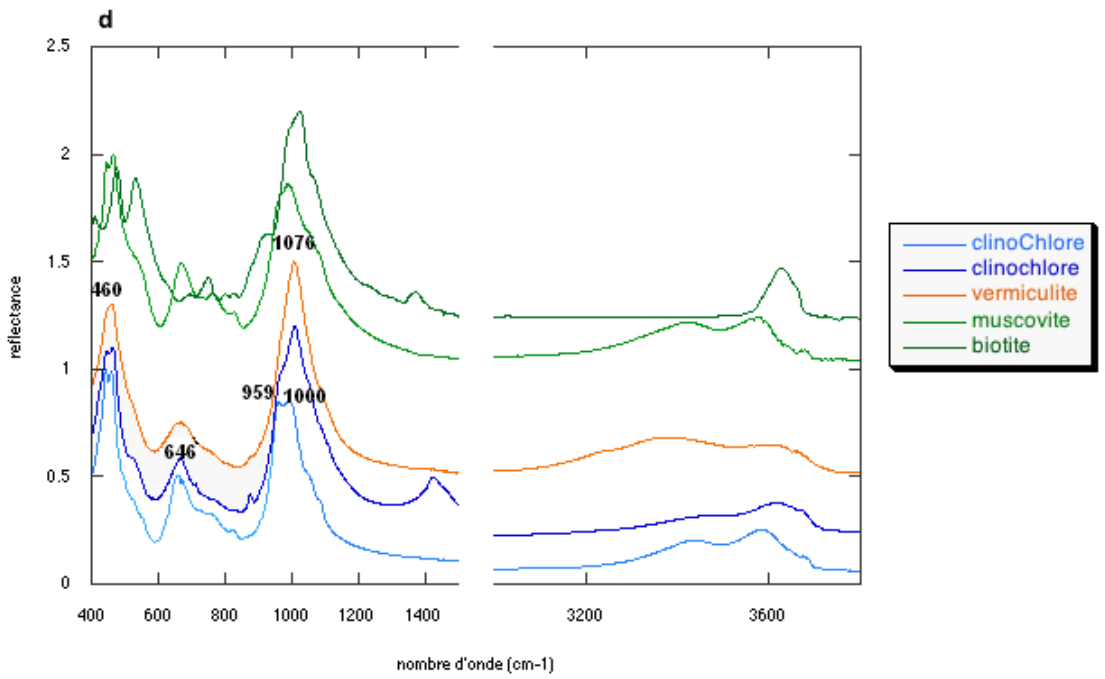


Figure II-10: Réflectance des minéraux argileux de référence dans le moyen infrarouge (400-4000 cm^{-1}). Les spectres sont regroupés par familles sauf pour la vermiculite, les micas et les clinochlorites (d). La variance dans les familles de kaolinite (a), illite (b) et smectite (c), est représentée par l'écart type entre chaque spectre pour chaque nombre d'onde (spectres grisés).



Suite de la figure II-10

Le proche IR représente le domaine qui nous intéresse plus particulièrement du fait de la possibilité d'analyse in-situ et de calibration des images satellites. Elle inclut des spectres de minéraux enregistrés en laboratoire à l'aide du spectromètre FTIR et ceux enregistrés avec le spectromètre ASD FieldSpec FR® en condition de laboratoire (Cf. II.2.2 et II.4.2.3). La librairie créée dans le PIR et rassemblant les spectres enregistrés en réflexion diffuse avec le spectromètre IRTF est la plus fournie du fait de la qualité des mesures en termes de résolution et de temps d'analyse. Cependant, les spectres enregistrés dans le proche infrarouge contiennent moins de pics que les spectres correspondants, enregistrés dans le moyen infrarouge. Cela est dû à la nature des deux régions spectrales où le proche infrarouge consiste en plusieurs superpositions de pics de combinaison et d'harmonique des pics présent dans le MIR.

Les liens établis avec les analyses réalisées avec le spectromètre de terrain ASD FieldSpec FR® ont pour but d'ouvrir notre travail sur l'analyse de terrain. La figure II-11 fait un parallèle entre les spectres des minéraux de référence enregistrés avec ces deux spectromètres.

La bande d'absorption proche de 1400 nm (noté 1a et 1b, figure II-11) est lié à la première harmonique de la déformation des liaisons O-H de l'eau ($2\nu_{H2O}$) et de l'élongation des liaisons métal-O-H ($2\nu_{OH}$). La position exacte de la bande dépend donc de la nature du cation qui est lié à la liaison hydroxyle. De plus, la présence d'eau aura tendance à élargir ce pic, diminuant par la même occasion les possibilités de détermination des autres liaisons responsables de la bande d'absorption.

Le pic proche de 1900 nm (noté 2a et 2b, figure II-11) est le résultat de la combinaison de la vibration d'élongation et de déformation de la molécule d'eau ($(\upsilon + \delta)_{H2O}$). Ce pic est bien marqué

dans le spectre des smectites, plus faible dans le spectre de l'illite et quasiment absent dans le spectre de la kaolinite.

Le pic d'absorption proche de 2200 nm (noté 3a et 3b, figure II-11) est le résultat de la combinaison de la vibration de déformation des liaisons métal-OH et la vibration d'élongation de la liaison OH ($(\upsilon + \delta)_{OH}$). les trois petits pics visibles entre 2500 et 2300 nm (noté 4a et 4b, figure II-11) sur le spectre de la kaolinite sont le résultat de la combinaison de l'élongation des groupements OH avec des vibrations de déformation de la structure (Petit 1999).

Globalement, la forme et la taille de ces figures d'absorption caractéristique des minéraux argileux est similaire sur les spectres enregistrés avec les deux spectromètres. Cependant, certains pics d'absorption sont moins détaillés sur les spectres enregistrés avec le spectromètre de terrain. C'est le cas par exemple du doublet apparaissant à environ 1400 nm sur les spectres de kaolinite. Les pics d'absorption de ce doublet correspondent aux harmoniques des vibrations d'élongation des liaisons Al_2OH. Ils sont bien différentiés avec le spectromètre IRTF (figure II-11b) alors que le pic à 1393 nm n'apparait que sous la forme d'un épaulement avec le spectromètre de terrain (figure II-11a). La même constatation est faite sur le doublet apparaissant à 2200 nm.

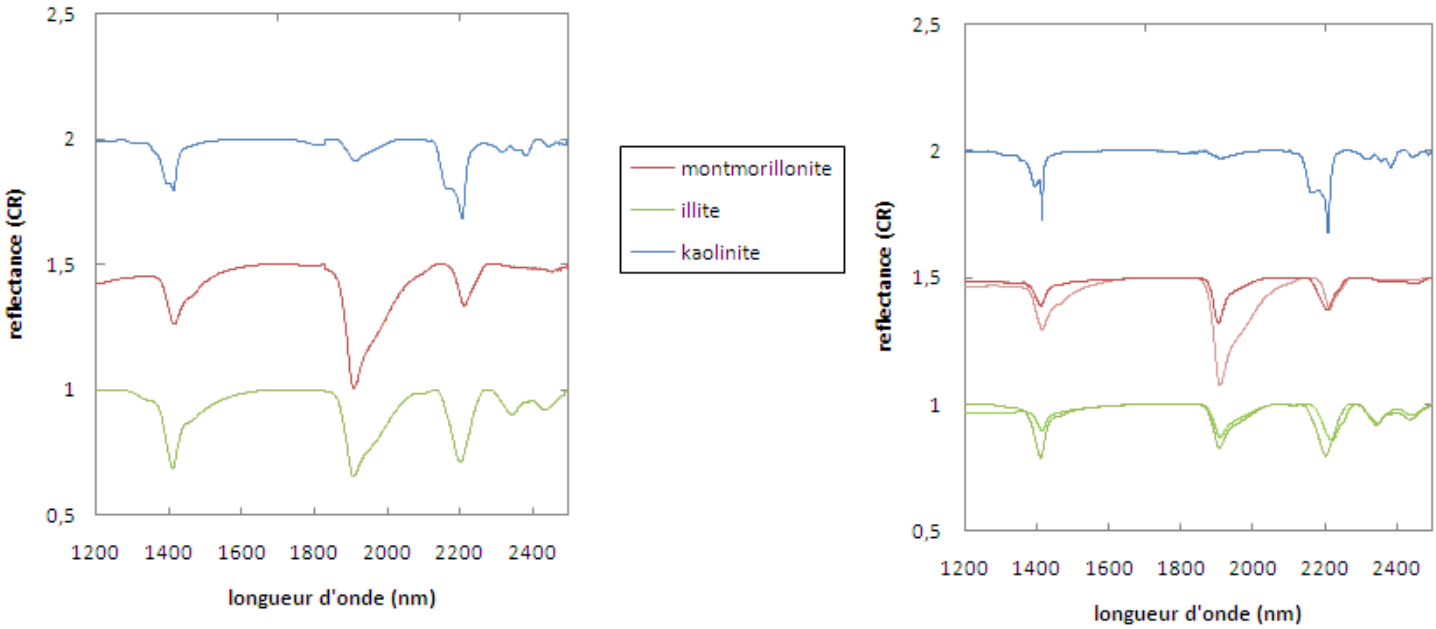


Figure II-11: Spectres de réflectance des différentes familles de phyllosilicates après retrait du continuum, enregistrés dans le visible et le PIR avec le spectromètre de terrain ASD (a) et dans le PIR avec le spectromètre IRTF en réflexion diffuse (b).

Parmi les minéraux non argileux, les minéraux tels que le quartz ou les feldspaths, qui apparaissaient suivant plusieurs pics dans le MIR (entre 2000 et 1800 cm^{-1}), sont quasi transparents dans le NIR (ils ne présentent pas de figures d'absorptions marquées). En revanche les oxydes de fer et particulièrement la goethite FeO(OH) et l'hématite (Fe_2O_3) présentent une composition minéralogique

caractérisée par la présence des ions ferriques (fer trivalent Fe^{3+}) qui sont responsables d'une absorption dans le visible. Enfin, la calcite présente un pic d'absorption majeur dû à la vibration de la molécule CO_3, à environ 2320 nm.

II.4.2.2 Les mélanges synthétiques

Le but est d'avoir des mélanges suffisamment simples pour acquérir une connaissance des spectres d'interaction liés à la superposition des spectres d'absorption des différentes espèces, sachant que l'additivité des absorbances est un cas idéal. Dans cette étude, les mélanges sont une base théorique intéressante pour calibrer les modèles lors des analyses statistiques, établir des relations fondamentales basées sur les caractéristiques des figures d'absorption propres à ces minéraux et définir les capacités de l'analyse en ondelette.

Ainsi, des mélanges sont constitués à partir de deux puis trois minéraux correspondant aux trois familles de minéraux argileux les plus fréquentes dans les sols (kaolinite, illite et smectite). Les minéraux sont réduits en poudre à l'aide d'un broyeur à bille en agate. Les poudres fines ainsi formées sont mélangées dans les proportions voulues (figure II-12) et re-broyées dans un mortier en agate de manière à favoriser le mélange et l'homogénéité des échantillons. En effet, la taille des particules influence la quantité d'absorption de l'échantillon et l'intensité des pics d'absorption de son spectre. La différence de taille de grain entre les différents constituants du mélange complique alors l'analyse des spectres de mélange dans lesquels l'échantillon le plus grossier devient prédominant (Crowley 1988).

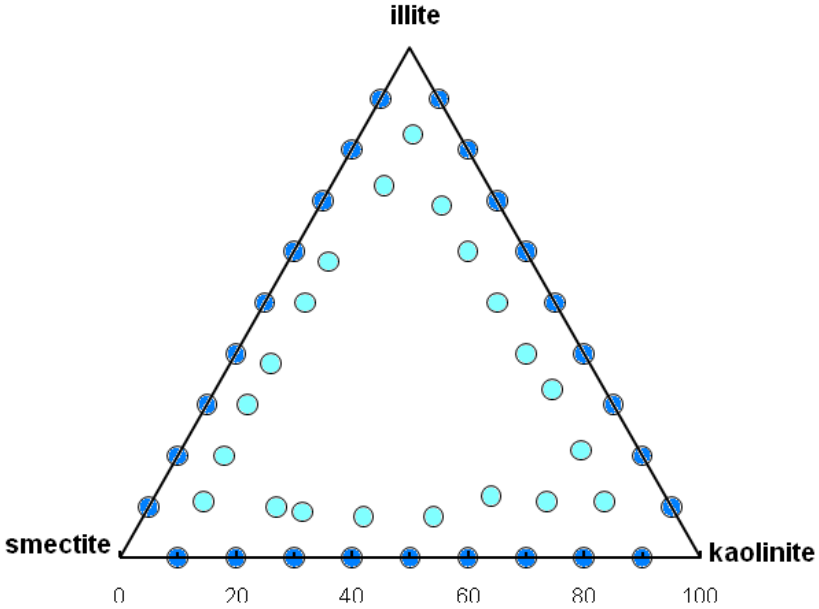


Figure II-12: Répartition des mélanges synthétiques dans le diagramme ternaire illite, kaolinite, smectite

Nous avons présenté la variabilité des spectres en fonction de la composition chimique de quelques minéraux argileux. Dans le cas des mélanges, le spectre résultant est une combinaison non linéaire des spectres des différents éléments du mélange.

Dans le moyen IR, la figure II-13 montre la variabilité spectrale de mélanges comportant de l'illite (du Puy en Velay et du LRPC), de la kaolinite (de Clérac) et la smectite (de la société ABM) en proportions variables. Les spectres correspondant montrent des propriétés intermédiaires aux pôles purs, ces propriétés variant non-linéairement en fonction de la quantité relative de chacun des pôles.

Pour les mélanges contenant à la fois de l'illite et de la smectite, les propriétés spectrales sont assez similaires. L'apparition d'un pic à 3696 cm^{-1} et de petits pics entre 600 et 900 cm^{-1} est principalement liée au faible pourcentage de kaolinite présent dans la smectite LRPC (Ivelay/Slrpc_10/90 ; figure II-13a).

En revanche, pour les mélanges contenant de la kaolinite, les spectres sont globalement dominés par les caractéristiques de ce minéral. La forme du spectre et l'intensité des bandes d'absorption sur l'ensemble du domaine infrarouge, augmentent fortement avec la teneur en kaolinite. D'après Joussein et al. (2001) cette augmentation de l'intensité des pics d'absorption suggère que les coefficients d'absorption des pics de la kaolinite sont plus importants que ceux de la smectite ou de l'illite.

Lorsque la teneur en smectite et en illite est largement majoritaire (> 80%) dans le mélange avec la kaolinite, on observe également que l'intensité du pic à 3620 cm^{-1} est plus importante que celle du pic à 3696 cm^{-1}. D'autre part, dans la région entre 1200 et 400 cm^{-1}, l'augmentation de la kaolinite se traduit par la séparation du pic d'absorption de la liaison Si-O en plusieurs pics et par l'apparition d'un épaulement à 938 cm^{-1} lié à la liaison AlAlOH (Madejova 2002). Dans le mélange smectite/kaolinite, on peut également noter la migration du pic d'absorption de la liaison Al-O-Si de 541cm^{-1} à 520 cm^{-1} lors de l'augmentation de la teneur en kaolinite.

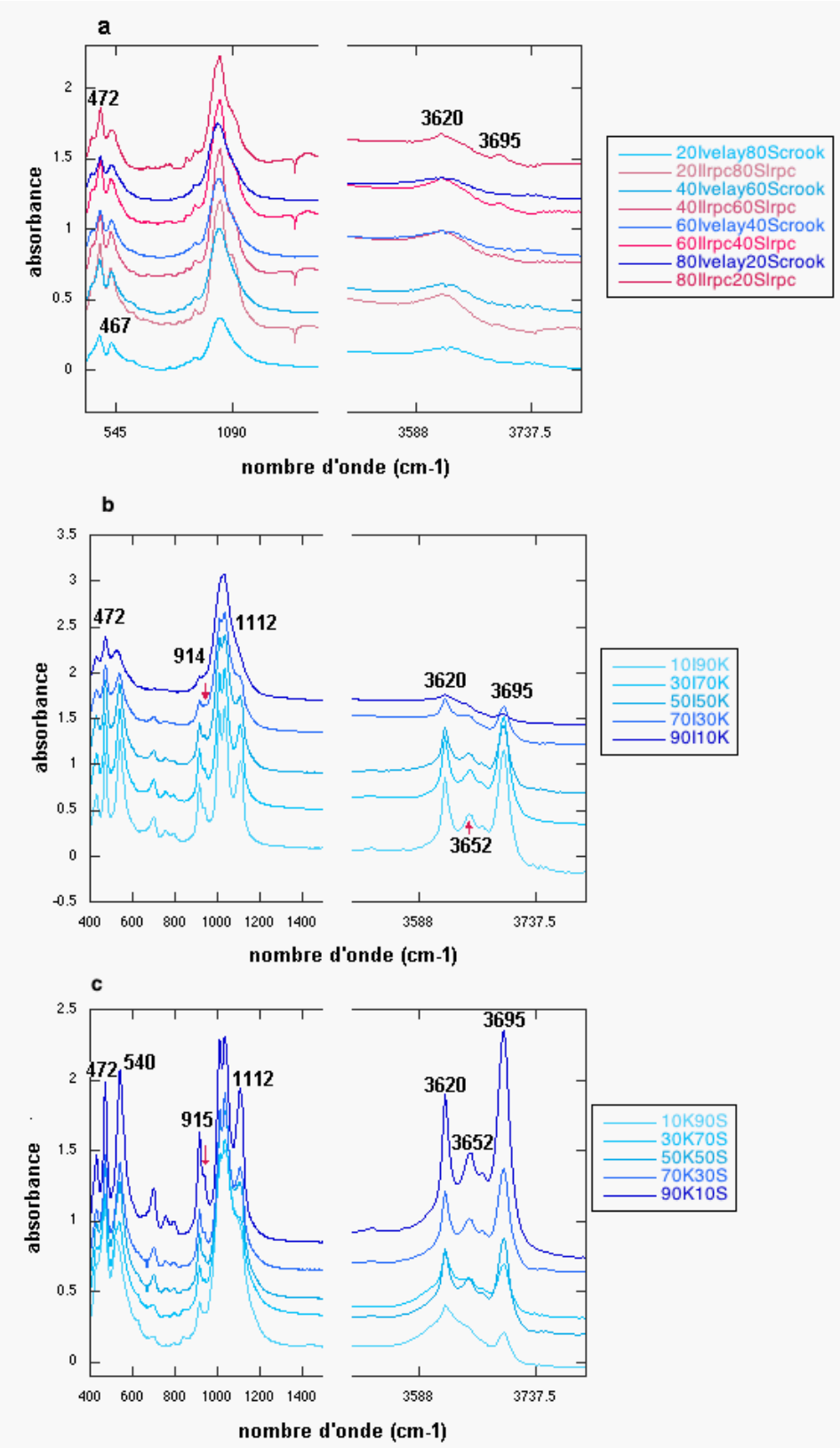


Figure II-13: Exemple de spectres de réflectance de mélanges mécaniques pondéraux d'illite / smectite (a), illite / kaolinite (b), smectite / kaolinite (c) enregistrés dans la gamme du MIR avec le spectromètre IRTF en transmission. L'ensemble des séries allant de 10% à 90% avec un mélange tout les 10%. Les mélanges sont affichés par pas de 20% pour plus de clarté.

Dans le proche IR, l'analyse des mélanges se fait sur une plus grande quantité d'échantillon que dans le MIR (quelques grammes contre quelques milligrammes), améliorant ainsi la représentativité de ce qui est mesuré. Les spectres enregistrés en laboratoire avec le spectromètre ASD FieldSpec FR® et avec le spectromètre IRTF ont des valeurs de réflectance comparables du fait de la similitude de préparation des échantillons et de l'utilisation d'une source artificielle pour le spectromètre de terrain. Cela permet, dans un premier temps, de limiter l'impact de l'atmosphère sur le spectre infrarouge. Ainsi la principale différence entre les deux types de mesure vient de la différence de résolution spatiale et spectrale. En effet, dans les deux cas, le mélange analysé est préparé dans les mêmes conditions, avec un temps identique de séchage en étuve (24h à 90°C).

La figure II-14 présente les spectres de mélanges mécaniques d'illite, de smectite et de kaolinite dans des proportions variables, enregistrés avec le spectromètre de terrain (colonne de gauche) et avec le spectromètre IRTF en mode réflexion diffuse (colonne de droite). Elle montre l'évolution des caractéristiques des figures d'absorption propres à chacun de ces minéraux telles que présentées précédemment.

L'augmentation de la teneur en kaolinite se traduit principalement par l'apparition d'un doublet à 2200 nm et à 1400 nm (noté 3 et 1 sur la figure II-14), ainsi que par l'apparition de petits pics entre 2300 et 2400 nm (zoom figure II-14 c, d, e et f). L'augmentation de la teneur en smectite se traduit principalement par l'augmentation de la profondeur du pic à 1900 nm (noté 2 sur la figure II-14).

L'illite en revanche est beaucoup plus difficile à mettre en évidence dans les mélanges car les figures d'absorption de ce minéral ne sont pas dominantes par rapport à celles de la smectite ou de la kaolinite. Cependant, un décalage de la bande d'absorption de la combinaison de la vibration de déformation des liaisons métal-O-H et la vibration d'élongation de la liaison O-H, est observé de 2213 nm vers 2208 nm lorsque la teneur en illite augmente (voir zooms sur les figures II-14 a et b).

Ces différentes constatations se retrouvent dans des mélanges contenant les trois minéraux (illite, smectite et kaolinite). Notamment, le pic à 1900 nm évolue avec la teneur en smectite et un doublet à 2200 nm et 1400 nm apparaît lors de l'augmentation de la teneur en kaolinite (figures II-14 g et h).

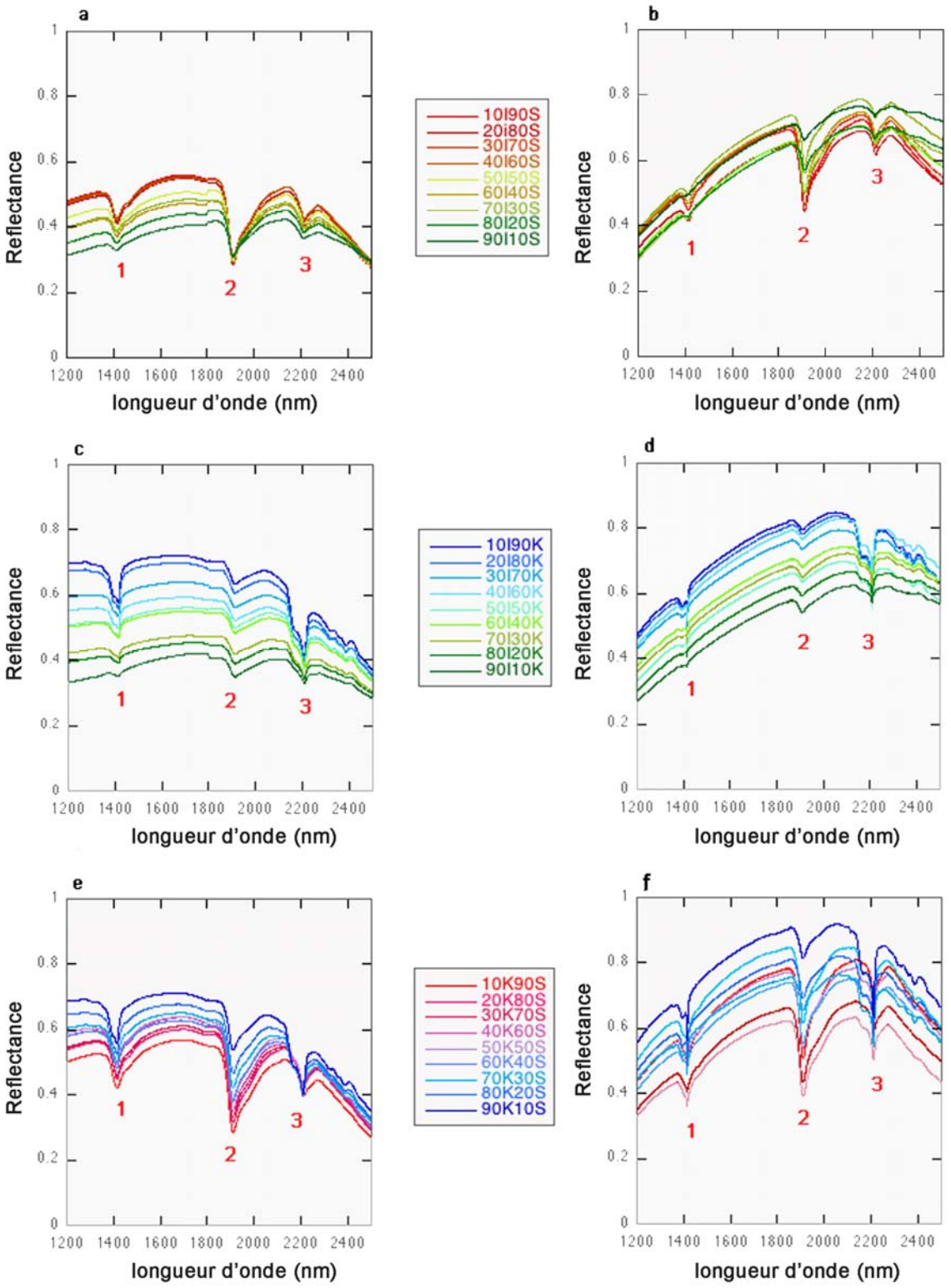
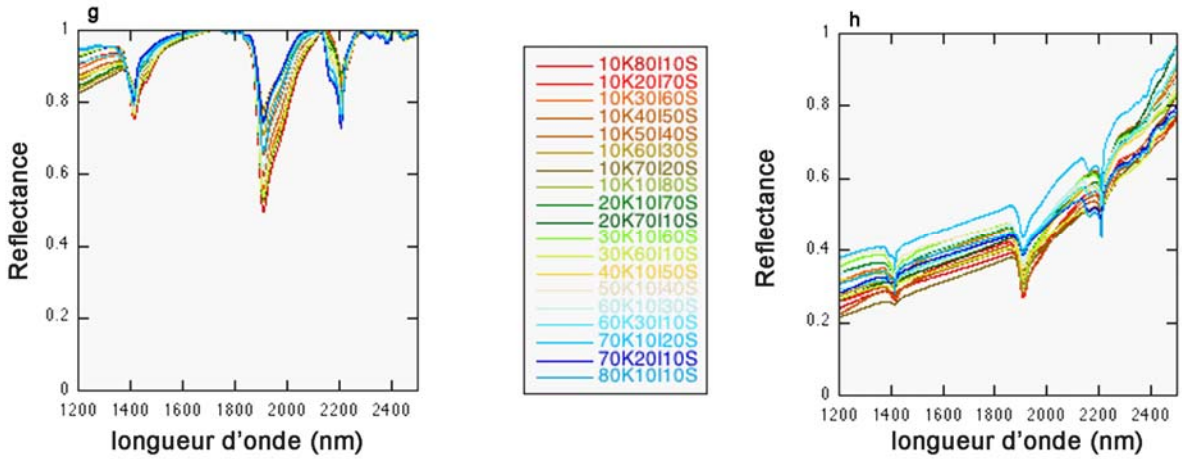


Figure II-14: Spectres de réflectance des mélanges synthétiques (en %) d'illite (I) et/ou smectite (S) et/ou de kaolinite (K) enregistrés dans le visible et le PIR avec le spectromètre de terrain ASD (colonne de gauche) et dans le PIR avec le spectromètre IRTF en réflexion diffuse (colonne de droite). Les chiffres (en rouges) soulignent l'apparition d'un doublet.



Suite de figure II-14

II.4.2.3 Les sols

La collection de sols est constituée de sols de surface (0 à 10 cm de profondeur) provenant de différentes terres cultivées de la région Midi-Pyrénées et prélevés en juin 2007 et 2008. La collection inclut des sols issus de différents reliefs (coteaux molassiques, terrasses et plaines alluviales et plateaux marno-calcaires principalement) et de différents matériaux parents (figure II-15). Ce sont ces paramètres qui influencent fortement la répartition spatiale des sols. Ainsi, bien qu'ils n'aient pas été formellement classés, ils sont rattachés à des unités pédologiques (définies par les pédologues par la chambre d'agriculture de Midi-Pyrénées) relativement larges tels que les basses plaines, les coteaux molassiques, les plateaux marno-calcaires ou les hautes terrasses qui tiennent compte de ces paramètres. Le quartz (présent sous forme sableuse) est généralement majoritaire dans les sols de la région et provient de l'érosion du Massif central et des Pyrénées.

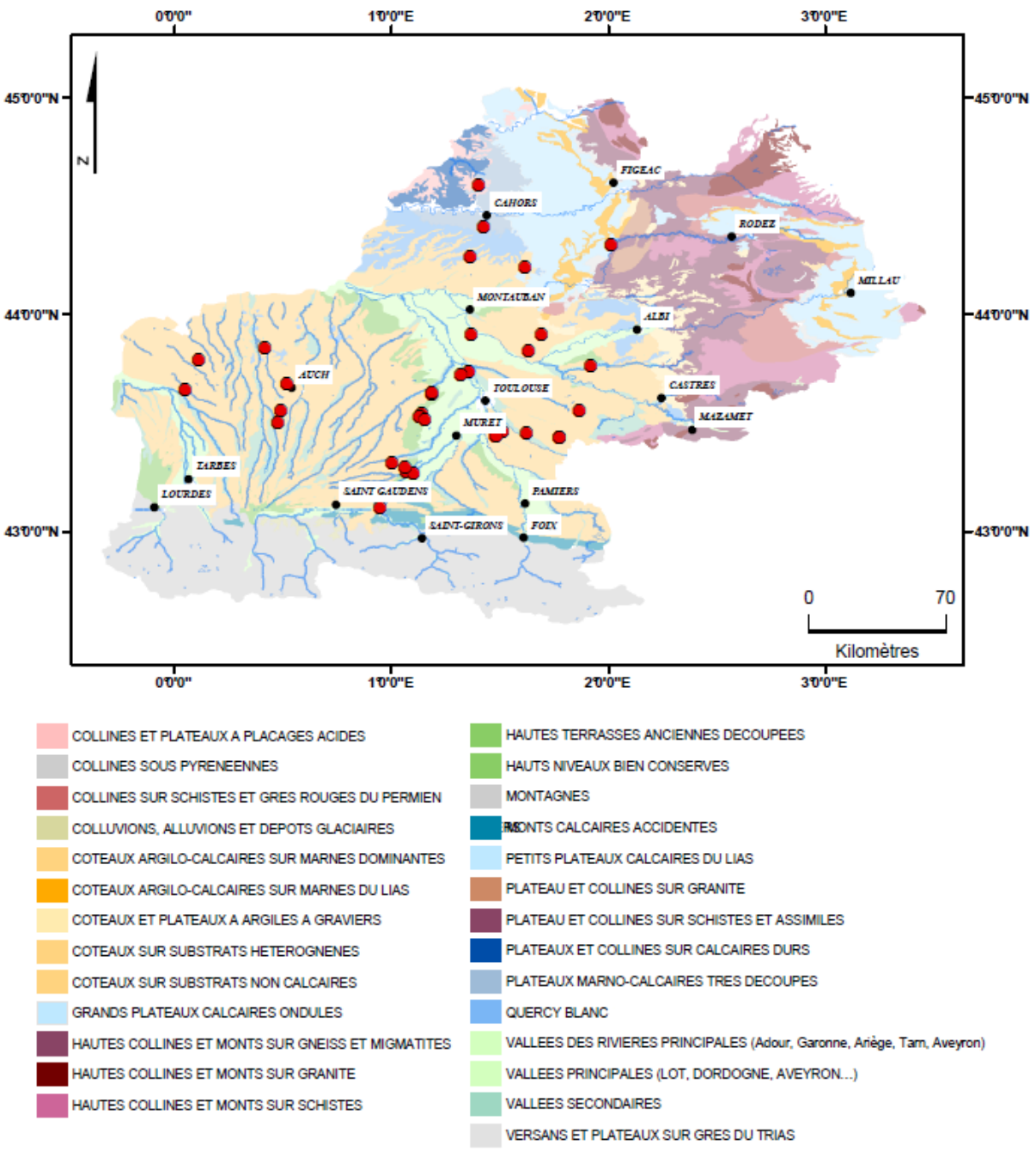


Figure II-15: Localisation des points d'échantillonnage (en rouge) réalisés dans la région Midi-Pyrénées en juin 2007 et 2008 vis-à-vis des données de pédologie disponible (source : Chambre Régionale d'Agriculture de Midi-Pyrénées)

Sur le terrain, les mesures spectrales ont été faites avec le spectroradiomètre portable ASD FieldSpec® (Cf. II.2.2). Dans un premier temps, la mesure d'un panneau de référence est acquise. Il s'agit d'un Spectralon® qui est un réflecteur parfait puisqu'il ne présente pas d'absorption dans la gamme de longueurs d'ondes considérées et réfléchit le rayonnement dans toutes les directions. Dans

un deuxième temps, le spectre d'échantillon est acquis en luminance puis exportés automatiquement en réflectance par le logiciel intégré à cet instrument.

Lors de la campagne de juin 2007, réalisée avant la thèse, les sols ont été prélevés après localisation du site d'échantillonnage au moyen d'un GPS d'une précision de 5 m. les mesures ont toutes été faites le même jour (conditions météo et d'éclairage optimales). Le spectromètre était posé sur un trépied à quelques centimètres au-dessus de l'échantillon. Ce positionnement a permis d'observer, en visée verticale, une surface circulaire d'un rayon d'environ 1 à 2 cm. Dans ce type de mesures, les échantillons sont broyés. Les bords des porte-échantillons sont couverts par de la feutrine noire afin d'éviter qu'ils jouent le rôle de réflecteurs très lumineux (figure II_15). Cette technique limite les effets liés à la rugosité, à l'humidité et à la pierrosité mais ne peut pas être considérée comme une méthode d'analyse in-situ au sens strict.

Lors de la deuxième campagne de terrain, réalisée en juin 2008 durant la thèse, les mesures ont été réalisées in-situ (figure II-15). Les points de mesures ont été choisis dans des terres cultivées caractérisées par des périodes de sol nu (cultures saisonnières) et des tailles suffisamment importantes pour être incluses dans un pixel de l'image ASTER. Ainsi les mesures spectrales obtenues permettent également une utilisation comme point de calibration pour la classification des pixels des images correspondant à la zone étudiée. Ils ont été localisés au moyen d'un GPS d'une précision de 5 m. Les conditions climatiques et l'inclinaison du soleil n'étant pas favorable au moment de certaines mesures, certains spectres in-situ ont été retirés. Ainsi certains sols ont été enregistrés uniquement en laboratoire.

Figure II-16: Différence de mode opératoire pour la mesure de la réflectance des sols lors des campagnes de 2007 et 2008. A gauche, une sonde auto éclairé permet de travailler de tout temps. A droite, la mesure est réalisée en lumière naturelle (soleil).

En laboratoire, les sols sont broyés dans un broyeur à bille, et passés dans un tamis à 200 µm afin d'en retirer les particules les plus grossières (fragments de roches, de litières ou de racines). Ils sont ensuite placés en étuve à 90° pendant 24 h avant les mesures spectrales (figure II-17). Le spectre

de sol résulte de superpositions d'absorptions de composés organiques et inorganiques. Les sols sont caractérisés par une série de propriétés telles que la teneur en eau, la teneur en matière organique (M.O.), la taille des particules et la nature minéralogique, qui sont susceptibles d'influencer leur réponse spectrale. L'ensemble de ces propriétés a été revu par de nombreuses publications de synthèse (Baumgardner 1986; Ben-Dor 2004).

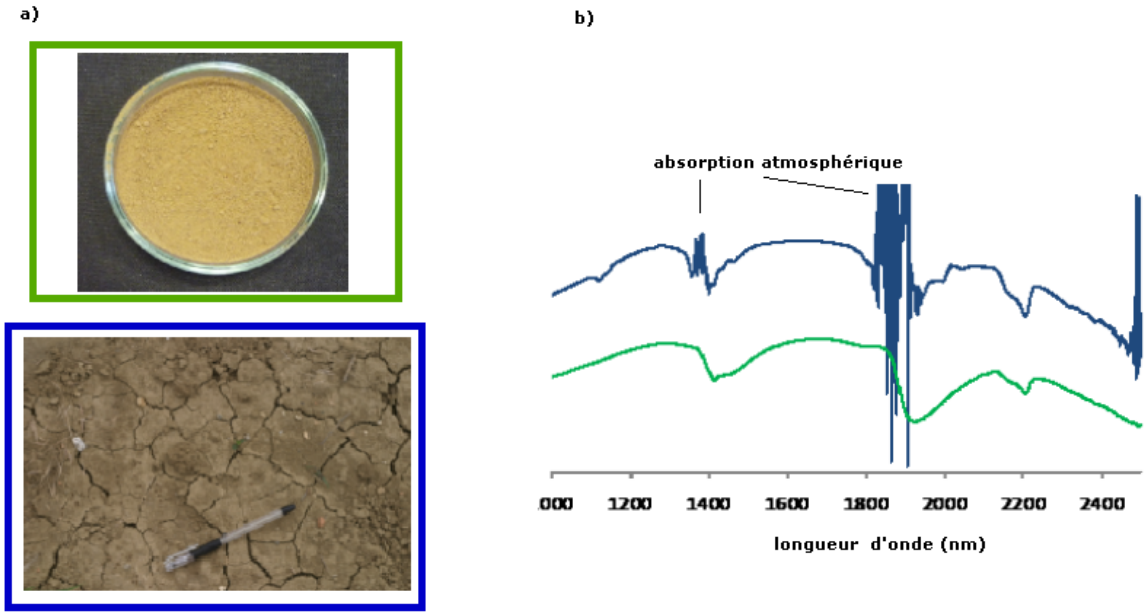


Figure II-17: a) Echantillon de sol broyé et séché comparé à la surface type de l'affleurement d'où il a été prélevé. B) Spectres ASD correspondant aux surfaces examinées en a).

De manière générale, les sols sont des mélanges plus complexes que les mélanges synthétiques parfaitement calibrés, présentés précédemment. Ils sont constitués d'un assemblage de particules de nature, de tailles de formes et d'arrangements très variés, alors que les spectres de minéraux et de mélanges de minéraux argileux ont été déterminés sur des poudres calibrées.

La figure II-18 présente les spectres de sols enregistrés en conditions de laboratoire suivant ces ensembles. Chacun de ces spectres montrent l'influence des minéraux argileux (pic d'absorption à 1400, 1900 et 2200 nm). Cependant, de même que dans les diagrammes ternaires (figure II-19), il n'apparaît pas de similitude particulière à l'intérieur d'une même classe. Cela s'explique par le fait qu'aucune cohérence particulière n'est trouvée entre les ensembles géomorphologiques et la composition argileuse des sols (figure II-16, tableau II-9).

Une cartographie beaucoup plus fine est ainsi nécessaire pour suivre la variabilité des sols, en terme de minéralogie des argiles, et le rôle important de la spectrométrie pour la précision de ces grands ensembles apparaît alors clairement.

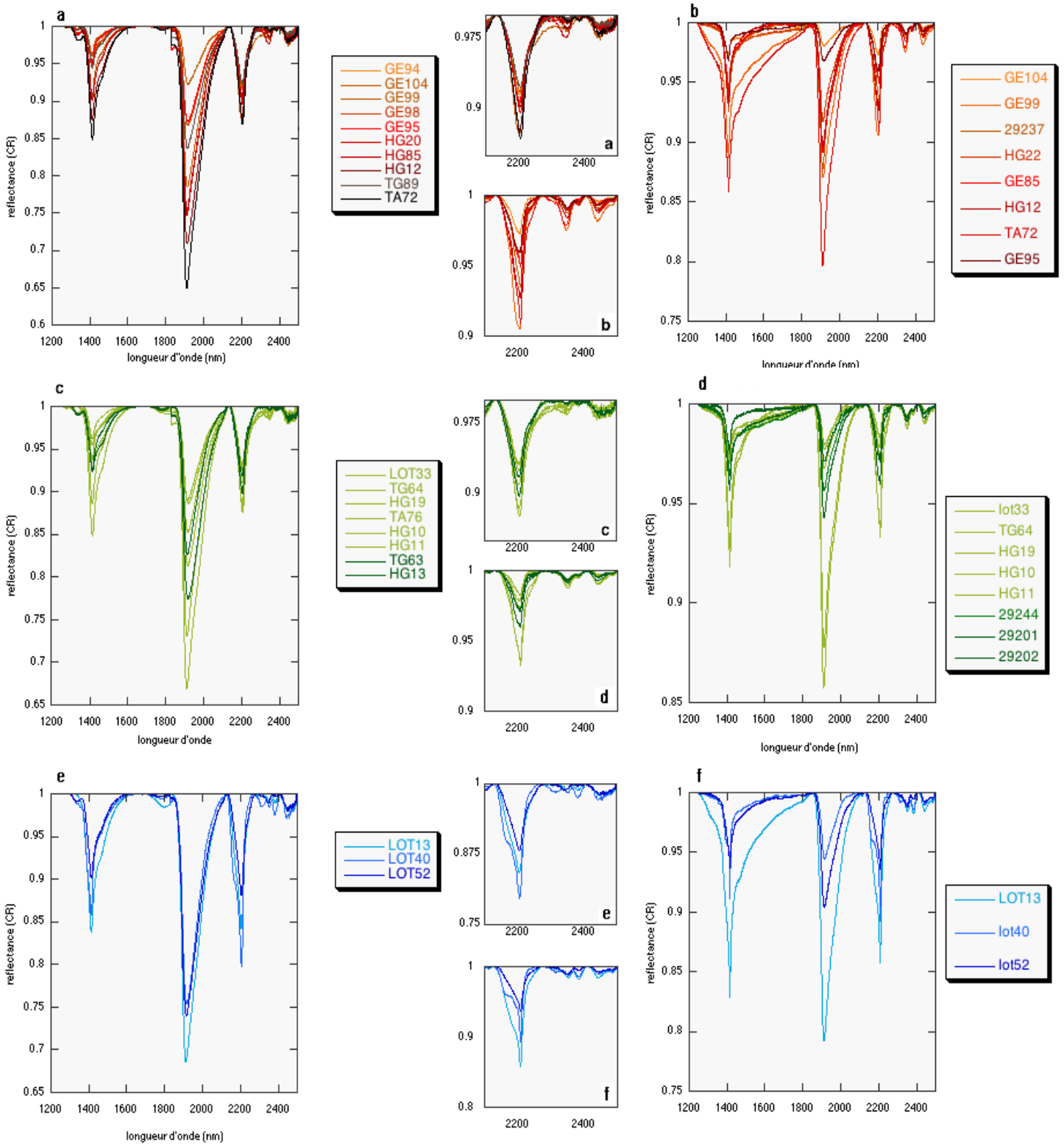


Figure II-18: Spectres de réflectance des sols après retrait du continuum, enregistrés dans le visible et le PIR avec le spectromètre de terrain ASD (a, c, e) et dans le PIR avec le spectromètre IRTF en réflexion diffuse (b, d, f). Les sols sont regroupés suivant qu'ils proviennent de la molasse (a, b), de plaines et terrasses quaternaires (c, d) ou de plateaux marno-calcaires (e, f).

Les analyses chimiques réalisées sur ces sols permettent d'étudier leur répartition suivant leur appartenance à différentes unités géomorphologiques en fonction des éléments chimiques majeurs les plus présents dans les différents sols (figure II-19). Dans les diagrammes en SiO_2, les sols sont largement « tirés » vers ce pôle. Ce glissement vers le pôle SiO_2 témoigne de l'abondance de la fraction détritique quartzeuse dans les sols. La teneur en calcium est relativement variable suivant les différents ensembles. Les sols des plateaux et des avant-causses en sont riches ainsi que certains sols de molasses. Cependant, la distinction des matériaux par les teneurs en éléments majeurs reste difficile à appréhender. Aucun assemblage typique ne permet de distinguer les différents ensembles géomorphologiques.

La chimie des éléments argileux montrent que les illites et les smectites choisies ont des compositions relativement variables. Elles ne se distinguent vraiment que par leur teneur en potassium (diagramme en bas à droite). Les autres éléments chimiques ont des teneurs similaires dans les illites et dans les smectites. Les kaolinites sont bien regroupées autour du pôle alumineux à part la kaolinite Ska qui est largement influencée par le pôle SiO_2.

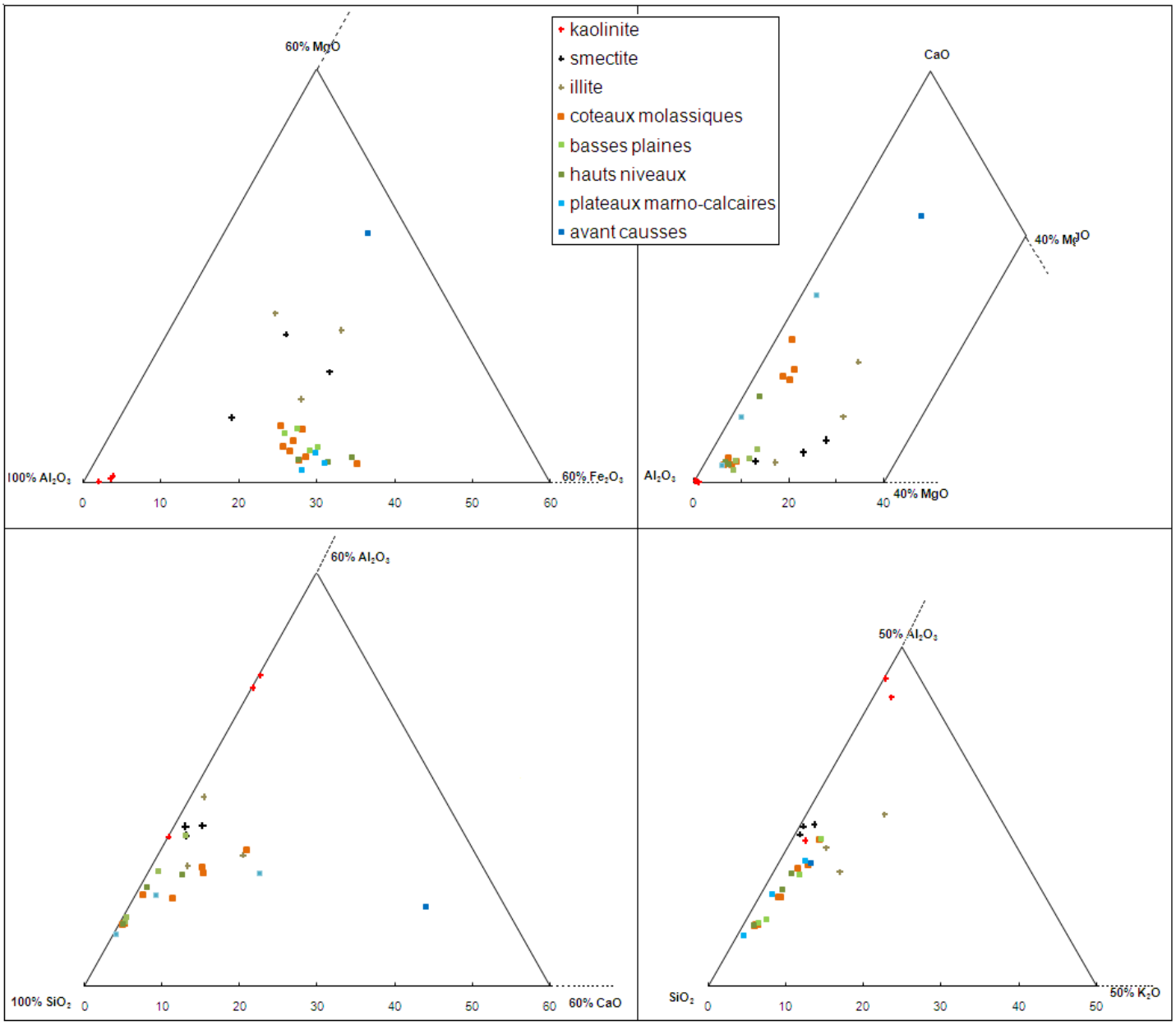


Figure II-19: Diagrammes ternaires des compositions chimiques de sols prélevés dans la région Midi-Pyrénées et de kaolinites, d'illites et de smectites. Les sols sont groupés suivant leur appartenance à différentes domaines géomorphologiques. Les teneurs sont exprimées en %.

Une quantification pour les 21 sols prélevés dans le cadre de cette étude, a été faite à partir des analyses chimiques associées aux données de diffraction des rayons X (Cf. paragraphe analyse chimique des éléments majeurs). Les résultats sont présentés dans le tableau II-9. Les données pour chaque sol sont détaillées dans l'annexe 2. Nous disposons de données de sols complémentaires fournis par le BRGM dans le cadre de la réalisation des cartes d'aléa retrait-gonflement des argiles des départements de la région Midi-Pyrénées. De manière générale, la teneur en quartz est majoritaire dans la plupart des sols et varie entre 20 et 70 %. Parmi les minéraux argileux considérés, l'illite est la plus représentée et varie entre 10 et 30%. La teneur en kaolinite est généralement inférieure à 15% sauf

dans les sols provenant du département du lot. Les teneurs calculées pour la smectite dans les sols sont généralement faibles voir nulles.

	C org	*M.O.*	quartz	illite	kaolinite	smectite/ vermiculite	calcite/dolomite	goethite/ oxyde de fer	rutile	feldspath
29237	*0,62*	*1,07*	27	38	9	13	9	4	0,51	0
46008	*0,26*	*0,45*	69	11	13	4	0	3	0,56	0
46002	*0,67*	*1,15*	21	26	8	22	19	4	0,59	0
HG14	*1,13*	*1,94*	32	40	9	8	0	3	0,63	8
HG19	*1,02*	*1,75*	20	39	13	14	1	6	0,74	9
GE94	*0,74*	*1,27*	64	19,5	2	8	0	3	0,98	3
TA76	*1,23*	*2,12*	53	30	5	2	0	4	0,63	6
TA72	*0,69*	*1,19*	47	29	12	7	1	3	0,63	1
TG89	*0,71*	*1,22*	60	24	9	0	0	2,5	0,86	4
TG63	*1,06*	*1,82*	41	27	12	8	0	4	0,79	7
LOT10	*1,51*	*2,60*	11	21	5	6	50	2	0,35	4
LOT40	*1,06*	*1,82*	51	16	25	0	4	4	0,57	0
LOT52	*1,37*	*2,36*	24	28	10	12	19	5	0,58	2
GE98	*0,65*	*1,12*	66	17	6	5	1	4	0,82	0
TG64	*0,94*	*1,62*	68	25	4	0	0,4	1,6	0,6	1
GE99	*1,36*	*2,34*	8	34	2	16	10	6	0,69	24
HG10	*1,23*	*2,12*	55	21	3	7	0	3	0,8	11
LOT13	*0,8*	*1,38*	37	26	17	8	6	6	0,8	0
LOT33	*0,55*	*0,95*	30	30	20	0	8	4	0,65	8
HG22	*0,62*	*1,07*	40	30	10	5	6	3	0,64	6
HG13	*0,89*	*1,53*	58	16	5	8	0	4	0,94	9

Tableau 9: Résultats de la quantification minéralogique par l'association des données de la DRX et des données de l'analyse chimique pour les sols échantillonnées dans la région. Les données de carbone organique (C_{org}) et le pourcentage de matière organique (M.O.) associé sont également présentés.

CHAPITRE 3

Traitement des données

III. CHAPITRE 3 : TRAITEMENT DES DONNEES

Dans ce paragraphe, sont présentés les principaux outils classiquement utilisés pour l'examen des mesures spectrales effectuées sur les échantillons et les surfaces. Ils sont de deux types suivant que l'on se base sur l'étude de quelques bandes d'absorptions bien individualisées ou sur l'analyse de la forme globale des spectres.

Dans le premier cas, une normalisation des spectres de réflectance par le retrait du continuum (Clark 1984) permet d'individualiser les figures d'absorption particulières se produisant à des longueurs d'onde spécifiques et d'utiliser des caractéristiques géométriques (profondeur, symétrie, aire) de ces pics pour obtenir de l'information sur le contenu des échantillons (Kariuki 2004).

Dans le second cas, des approches statistiques telles que l'analyse en composantes principales et la régression par moindres carrés partiels sont couramment utilisées. Elles ont montré que l'analyse quantitative de certaines propriétés des sols telles que la teneur en CaCO3 et en argile était possible (Viscarra Rossel 2006). Parallèlement, en minéralogie, des études ont montré qu'il était possible de caractériser les différents minéraux argileux susceptibles de composer la fraction argileuse des sols à partir de leur spectre enregistré dans le moyen infrarouge (Petit 1995; Petit 1998; Madéjova 2001; Madéjova 2003) et dans le proche infrarouge (Post and Noble 1993; Petit 1999).

Nous avons donc cherché à voir dans quelles mesures ces méthodes sont capables de distinguer les différents minéraux de la phase argileuse. Dans ce paragraphe, les méthodes et les résultats obtenus pour l'étude de l'argile des sols sont présentés. Par l'application de ces méthodes, les possibilités et les limites pour la caractérisation des minéraux argileux présents dans les sols sont mises en évidence. Particulièrement l'analyse de l'aire des figures d'absorption après normalisation des spectres par la méthode du « continuum removal » apportent les résultats les plus intéressants et les plus directement exploitables dans notre objectif de méthode opérationnelle. Cette dernière partie a fait l'objet d'un article dans « Clay Minerals », revue de rang A.

III.1. APPROCHES STATISTIQUES

La quantification des composants d'un mélange est une procédure difficile à mettre en œuvre que ce soit par des méthodes classiques telles que la Diffraction des Rayons X (Thorez 1975; Bish 1993; Bouchet 2000; Taubaso 2004) ou par des méthodes utilisant la spectrométrie (Matteson 1993; Bertaux 1998; De La Fuente 2000; Udelhoven 2003; Yitagesu, van der Meer et al. 2009).

L'objectif de notre approche est de caractériser les échantillons par spectrométrie infrarouge pour prévoir leur appartenance à un groupe connu, préalablement référencé et inclus dans une bibliothèque de minéraux et de mélanges de référence. Ces groupes de références sont obtenus par mesure spectrale de nombreux échantillons semblables et forment des nuages de points dans l'espace

multidimensionnel. Les spectres sont ainsi analysés suivant des méthodes statistiques et plus particulièrement suivant l'analyse qualitative de type Analyse en Composantes Principales et l'analyse quantitative de type régression PLS (Partiel Least squares).

III.1.1 L'analyse en composantes principales

III.1.1.1 Démarche générale.

L'Analyse en Composante Principale (ACP) est une méthode statistique qui vise à extraire l'information utile présente dans un jeu de données important. Elle a notamment été développée et formalisée dans les années 30 par Harold Hotelling (Hotelling 1936). Elle est particulièrement adaptée à l'étude exploratoire des données spectrales fortement redondantes. En effet, elle permet de représenter la structure des données de l'analyse, de vérifier l'existence de similarités aux niveaux des longueurs d'onde caractéristiques des minéraux de référence, des mélanges et des sols et enfin, de détecter les individus pouvant introduire une erreur dans l'analyse du fait de leur comportement extrême.

Le principe de l'ACP consiste à rechercher les directions dans l'espace des *m* variables qui représentent le mieux la dispersion des *n* individus (équation 1). Les directions de plus grandes dispersions sont des directions intéressantes et la variabilité associée à ces directions correspond à de l'information. Pour cela une matrice de données prédictives comprenant *n* lignes représentant les individus et *m* colonnes représentant les variables (dans notre cas, n étant le nombre d'échantillon et m l'absorption en chaque longueur d'onde) est remplacée par une nouvelle matrice, dérivé de X, comprenant le même nombre de lignes que X mais un nombre de colonnes bien inférieur. Ainsi, elle remplace les données spectrales d'origine (les variables prédictives), fortement redondantes dans le cas de données de spectrométrie, par des variables synthétiques, les composantes principales. Ces composantes principales sont des combinaisons linéaires des variables d'origine. Elles correspondent aux directions principales d'allongement du nuage des facteurs. En effet, dans l'espace original, les individus peuvent être représentés par un nuage de points dans un espace à m dimensions. L'ACP consiste en une rotation effectuée de manière à placer les nouveaux axes dans la direction de plus grande dispersion du nuage de points (figure III-1). Le deuxième axe est orthogonal au premier et tient compte de la plus grande dispersion résiduelle, et ainsi de suite. Les composantes principales sont donc ordonnées par ordre décroissant de variance expliquée. Elles contiennent la quasi-totalité de l'information, et ont l'avantage d'être non corrélées, ou orthogonales entre elles. De cette manière, on évite d'avoir la même « information » dans plusieurs composantes principales. Les coordonnées des points représentatifs des spectres dans ce nouveau repère sont les coordonnées factorielles.

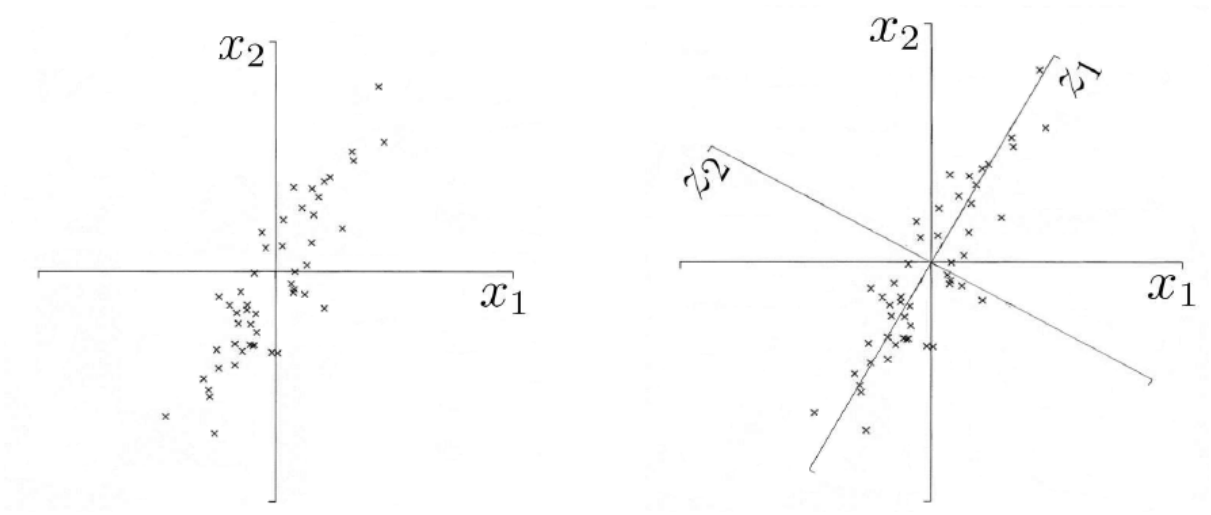


Figure III-1: Changement de repère orthogonal et définition des nouvelles composantes principales (z_1 et z_2)

Plusieurs études en science du sol utilisent cette méthode. Elle apporte une vision globale et plus compacte de la structure de données multivariables et d'identifier s'il existe des données aberrantes. Viscarra Russel (2006) utilise les ACP pour comparer des spectres de mélanges « théoriques » avec des spectres de sols « naturels ». Il trouve ainsi des similarités mais montre que les ACP mettent principalement en évidence des différences implicites entre les mélanges et les sols. En 2009, il étudie les différences entre des spectres enregistrés sur le terrain et des spectres enregistrés en laboratoire pour caractériser l'influence des conditions de mesures et de calibration des échantillons en laboratoire. Ainsi, il montre que les différences les plus remarquables peuvent provenir de la différence de taille des particules, de l'homogénéité des mélanges par rapport à l'hétérogénéité des sols, d'erreurs dans la procédure expérimentale, d'un bruit d'analyse marqué, etc.

En spectrométrie, une difficulté vient de la dimension élevée des spectres infrarouges et une étape de réduction de la taille des données est souvent nécessaire avant la classification ou l'analyse de ces données. L'ACP permet cette réduction en diminuant le nombre de valeurs propres utilisées pour construire les vecteurs propres et les composantes principales (Linker 2008). Dans ce cadre là, seul les valeurs propres de forte entropie représentant le mieux la base de données infrarouge, sont conservées. Ainsi, la majorité des études qui utilisent cette technique en sciences du sol, l'utilise en prétraitement de méthodes plus complexes tels que la régression par moindres carrées partiels (PLSR) (Viscarra Rossel, Cattle et al. 2009; Savvides 2010).

Une formulation algébrique de L'ACP est la suivante :

Les données sont structurées dans une matrice X à n lignes et m colonnes :

$$X_{(n,m)} = \begin{bmatrix} X_{1,1} & \cdots & X_{1,m} \\ \vdots & \ddots & \vdots \\ X_{1,n} & \cdots & X_{n,m} \end{bmatrix}$$ **Equation 1**

Ainsi, chaque individu *i* est représenté par le vecteur $u_i = [x_{i,1}, \ldots. X_{i,m}]$ et chaque variable ***X j*** est représentée par le vecteur $v_j = [x_{1,j}, \ldots. x_{n,j}]$. On peut calculer pour chaque variable sa moyenne X_j et son écart type σ (X_j). Le vecteur **g** des moyennes des variables constitue le centre de gravité du nuage de points et sera l'origine du nouveau repère.

Les moyennes des variables servent ensuite à calculer les covariances et les corrélations entre les différents individus. La covariance entre les variables X_i et X_j est donnée par :

$$cov\left(X_i, X_j\right) = \frac{1}{n} \times \sum_{m=1}^{n} [(X_{i,m} - \bar{X}_i) \times (X_{j,m} - \bar{X}_j)]$$ **Equation 2**

La matrice de covariance Q de dimension m * m est constituée des covariances entre les spectres i et j :

$$Q = cov(X_i, X_j)$$ **Equation 3**

La variance et les écarts types de chaque spectre permettent ensuite de calculer la corrélation entre deux spectres :

$$Cor\left(X_i, X_j\right) = \frac{\frac{1}{n} \times \sum_{m=1}^{n} (X_{i,m} - \bar{X}_i)}{\sigma_i} \times \frac{(X_{j,m} - X_j)}{\sigma_j}$$ **Equation 4**

La matrice de corrélation S entre les spectres i et j est alors :

$$S = cor(X_i, X_j)$$ **Equation 5**

Le polynôme caractéristique de la matrice de covariance Q permet alors de calculer les valeurs propres :

$$\det(Q \ \lambda * Id) = 0$$ **Equation 6**

On obtient alors les valeurs propres associées aux vecteurs propres :

$$Q \text{ v } V_i = \lambda * V_i$$ **Equation 7**

Et les composantes principales associées :

$$Y_i^k = \sum_{j=1}^{M} \left| V_k^j * \left(V_k^j - \overline{X_i}\right) \right|$$ **Equation 8**

III.1.1.2 Résultats

Les analyses en composantes principales présentées ci-dessous ont été réalisées avec le logiciel Unscrambler® de chez CAMO.

Analyse des spectres enregistrés dans le MIR et le PIR

En préalable à l'Analyse en Composantes Principales, la dérivée des spectres par filtre Savitsky-Golay (Savitzky and Golay 1964) est calculée afin d'améliorer la localisation des bandes spectrales en faisant ressortir les maxima et minima et les largeurs de bandes tout en diminuant le bruit.

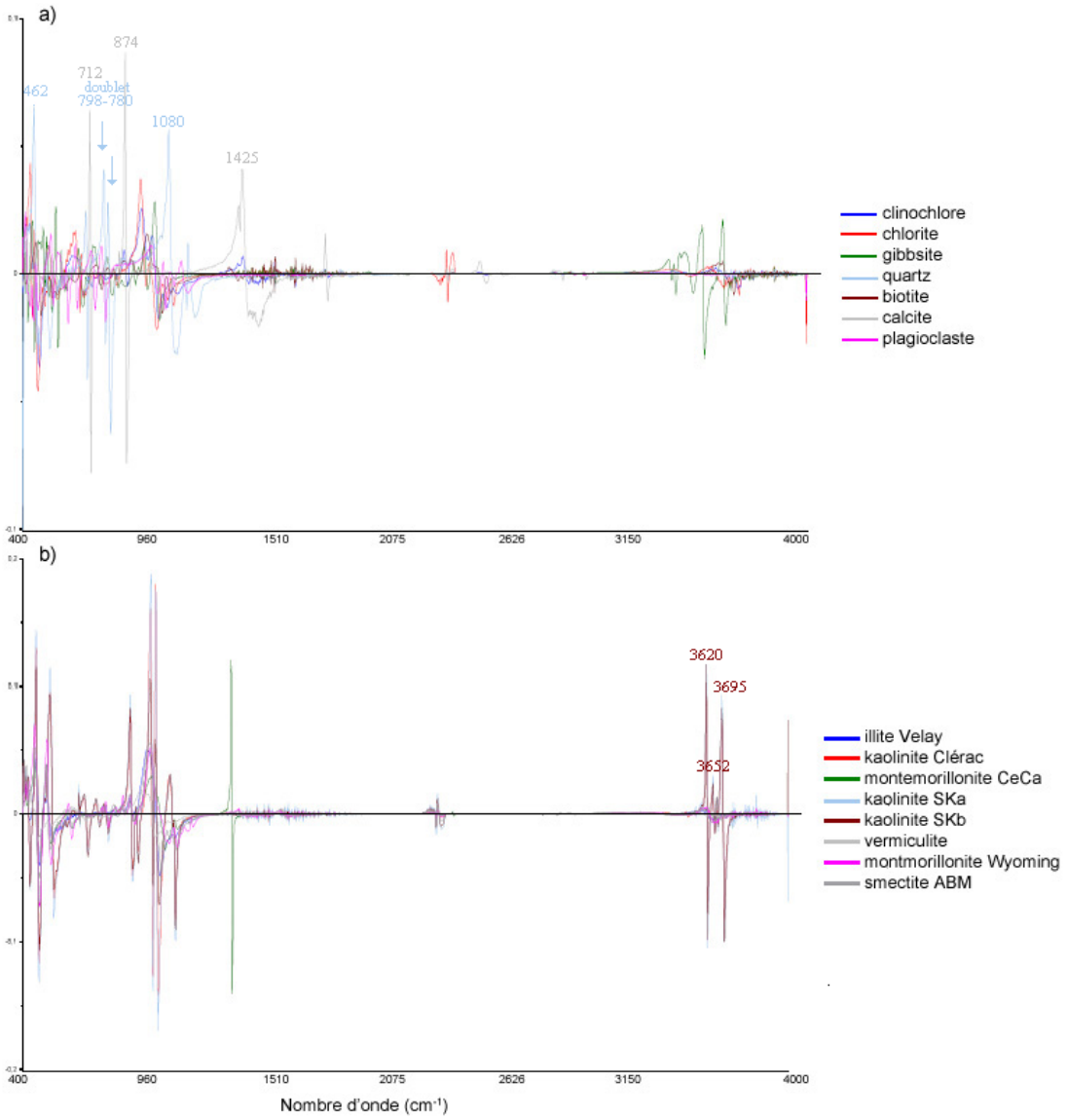


Figure III-2: Spectres de dérivée des minéraux non argileux (a) et argileux (b) de référence enregistrés dans le MIR et utilisés dans l'ACP présentée ci-dessous. On peut noter que les minéraux non argileux (a) présentent des figures d'absorption bien différentiées notamment le quartz et la calcite (spectre bleu et gris). En revanche les minéraux argileux (b) présentent quasiment les mêmes pics caractéristiques ce qui les rendent difficile à différencier.

Dans le moyen infrarouge (MIR), la figure III-2a souligne que les spectres de dérivées des minéraux de référence non argileux ont des pics d'absorption relativement bien distincts notamment dans le domaine des liaisons Si (entre 400 et 1500 cm^{-1}). Plus particulièrement, le quartz qui est le composant majoritaire des sols de la région, et la calcite qui peut être en proportion non négligeable dans les sols, présentent des bandes d'absorption bien individualisées (voir annotation figure III-2). En revanche, les spectres de dérivée des minéraux de référence argileux (figure III-2b) sont similaires et le domaine des liaisons OH (entre 3500 et 4000 cm^{-1}) est le plus intéressant pour les caractériser. Particulièrement, la kaolinite est caractérisée par trois principales figures d'absorption à 3620,3652 et 3695 cm^{-1} qui la distinguent de l'illite et de la smectite.

Les pôles purs ont une influence beaucoup plus importante dans l'analyse en composantes principales que les mélanges et les sols. Leur variance résiduelle est en revanche généralement faible (ils sont donc bien expliqués par le modèle). Ce sont donc eux qui guident, à un dégré plus important que les autres, le modèle. Afin de limiter ce biais, les minéraux n'étant pas parfaitement purs tels que la kaolinite Ska ou la montmorillonite CeCa (Cf. chapitre 2) et ayant une influence significative sur le modèle ou une variance résiduelle (variance non expliquée) importante, sont considérés comme étant des valeurs aberrantes (outliers) et sont retirés de l'analyse. Il en va de même pour les mélanges.

Malgré ce biais, le plan factoriel, composé des deux premières composantes principales, reflète relativement bien la structure de notre base de données (figure III-3). Deux minéraux peuvent donner un sens à l'axe PC1. Le quartz dans la partie gauche et la kaolinite, dans une moindre mesure, dans la partie droite. L'axe PC1 renseigne principalement la description des échantillons. Ainsi, ces deux minéraux sont représentatifs de la structure des données dans le plan factoriel. La smectite et l'illite ne sont pas discriminés mais ils se positionnent le long de l'axe PC2. Ils apportent un deuxième niveau (moindre que le premier) d'explication de la répartition des échantillons dans le plan factoriel. Les mélanges synthétiques se répartissent entre les différents pôles purs. Les sols sont excentrés et regroupés sur le cadran supérieur gauche du diagramme. Ils sont associés aux pôles purs non argileux et notamment au quartz qui a, de la même manière que la kaolinite, une forte influence sur le modèle. Cette disposition traduit la prédominance de la « fraction non argileuse » des sols.

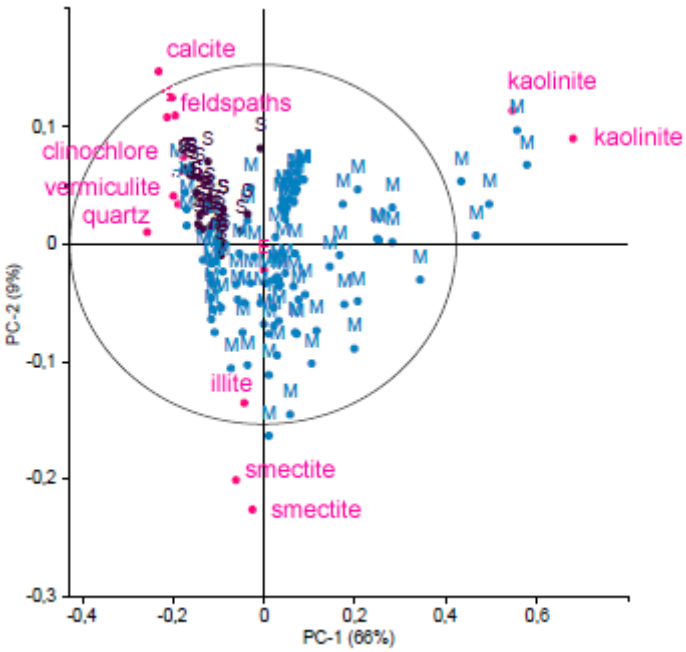


Figure III-3: Plan factoriel de l'ACP réalisée sur les spectres enregistrés dans le MIR avec un spectromètre IRTF (en mode transmission), des minéraux (rose), des mélanges expérimentaux de kaolinite, illite, smectite et quartz (bleu) et des sols (violet). L'ellipse correspond à l'ellipse de confiance de Hotelling à 95%.

Dans le PIR, les spectres de dérivée des minéraux argileux (figure III-4) mettent en évidence le doublet caractéristique de la kaolinite aux alentours de 2200 nm (résultat de la déformation d'une liaison métal-OH associée à l'élongation de la liaison O-H) et de 1400 nm (résultat de la déformation de la liaison Al-OH associée à l'élongation de la liaison O-H). Ils sont soulignés par les flèches rouges sur la figure III-4.

Les spectres de dérivée des minéraux de la famille de smectite sont caractérisés par un pic d'absorption intense autour de 1900 nm caractéristique de ces minéraux riches en eau (courbe bleue, rouge et grise, figure III-4). De manière générale, les spectres de dérivée des illites sont moins marqués.

Les minéraux tels que le quartz et les feldspaths ne possèdent pas de figures d'absorptions majeures dans ce domaine de longueur d'onde. Les oxydes montrent des figures d'absorption localisées dans le domaine du visible (entre 450 et 900 nm) qui n'est pas échantillonné par le spectromètre IRTF en réflexion diffuse. Seul le spectre de dérivée de la calcite met en évidence les vibrations de combinaisons et les harmoniques des vibrations fondamentales de la molécule CO3 aux environ de 1850 et 2160 nm et entre 2300 et 2500 nm.

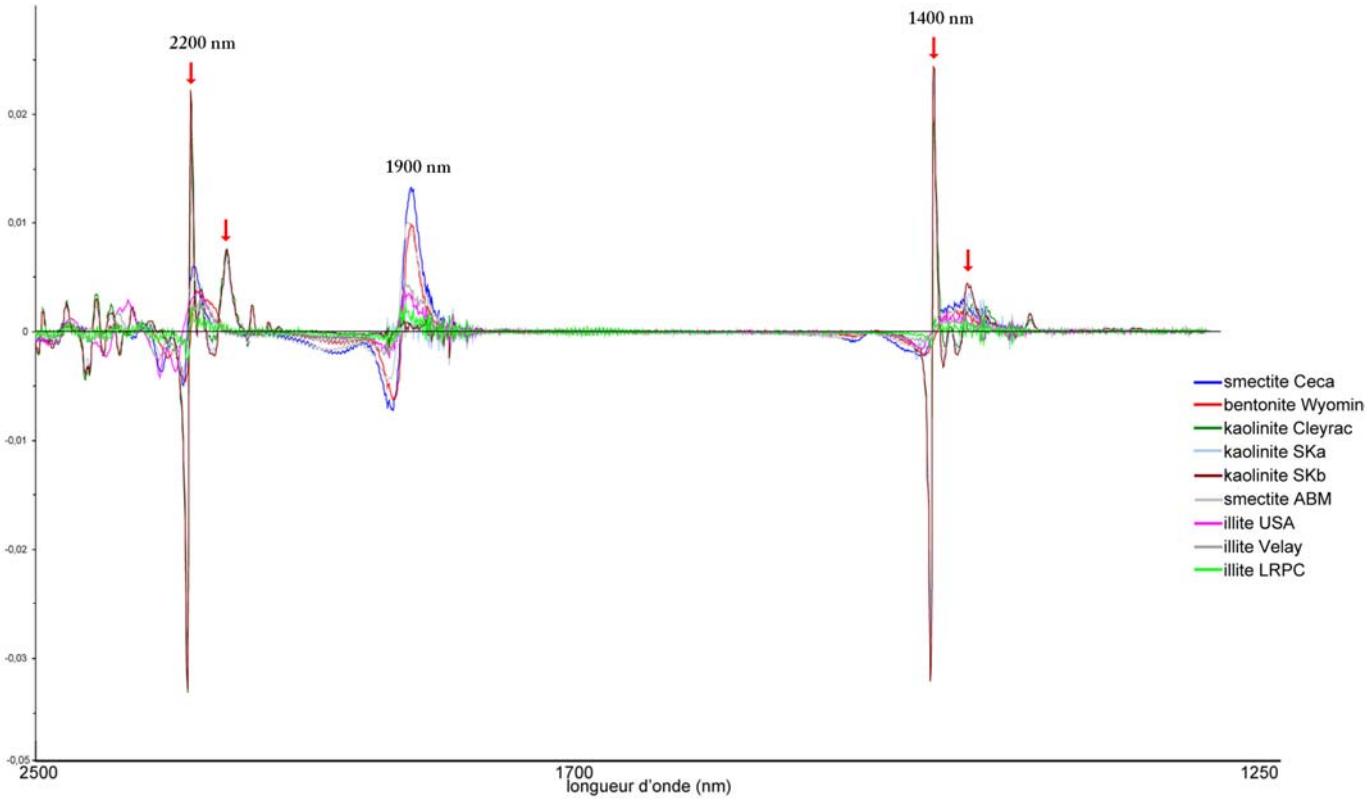


Figure III-4: Spectres de dérivées des minéraux de références argileux, enregistrés dans le PIR avec le spectromètre IRTF en mode réflexion diffuse.

De la même manière que pour l'ACP réalisée sur les données enregistrées dans le moyen infrarouge, les valeurs aberrantes (individu ayant une forte influence sur le modèle et/ou variance résiduelle importante) sont enlevées. Une exception est faite pour les minéraux de référence dont la forte influence est justifiée par le caractère extrême de leur spectre dû à leur mono-minéralité.

Le plan factoriel porté par les deux premières composantes principales donne une vision biaisée de la structure de notre base de données (figure III-5). Les mélanges se répartissent entre les trois pôles argileux (illite, smectite, kaolinite). L'axe PC1 porte l'information permettant de caractériser la présence de kaolinite, tandis que les illites et les smectites se répartissent le long de l'axe PC2. Le quartz est un facteur moins bien représenté par l'axe PC1 que dans l'ACP précédente (figure III-3). Cela s'explique par l'absence de pic d'absorption caractéristique de ce minéral dans le proche infrarouge (PIR). Cependant, le positionnement des échantillons de sols est comparable à celui observé dans le moyen infrarouge (MIR) et peut également être expliqué par le caractère sableux (forte teneur en quartz des sols) dominant des échantillons.

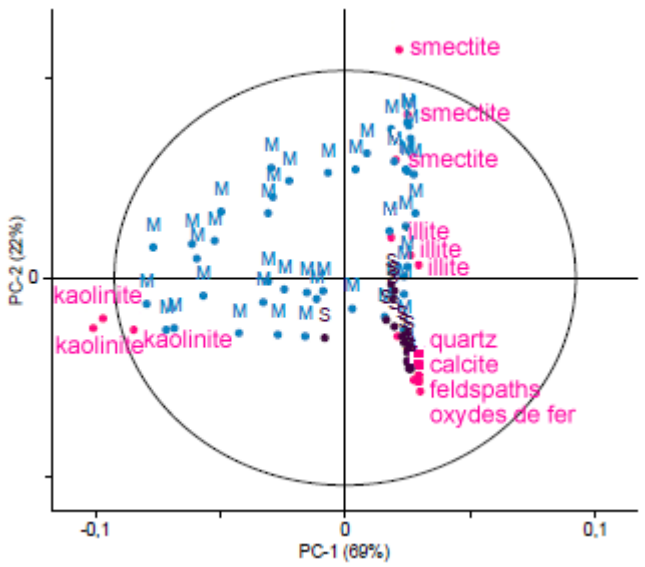


Figure III-5: Plan factoriel de l'ACP réalisée sur les spectres enregistrés dans le PIR avec un spectromètre IRTF en mode réflexion diffuse sur les minéraux (rose), les mélanges expérimentaux de kaolinite, illite, smectite (bleu) et les sols (violet).

Comparaison des spectres enregistrés en laboratoire avec les spectres de terrain

L'ACP est également un bon moyen de vérifier :

- d'une part, la compatibilité des données provenant du spectromètre IRTF en mode réflexion diffuse et celle provenant du spectromètre de terrain
- d'autre de part, la compatibilité des données enregistrées dans les conditions de laboratoire et de terrain par le spectromètre de terrain.

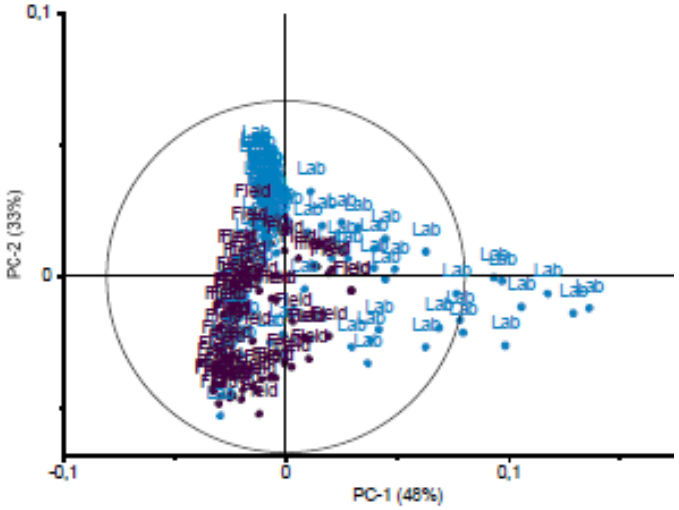


Figure III-6: Plan factoriel de l'ACP réalisée sur les spectres proche infrarouge enregistrés en laboratoire avec le spectromètre IRTF en mode réflexion diffuse (bleu) avec le spectromètre de terrain (violet).

Dans le premier cas, l'ACP est réalisée sur les dérivées premières des spectres de mélanges et de sols, enregistrés dans le PIR, à l'aide du spectromètre IRTF et du spectromètre de terrain en condition de laboratoire (lumière artificielle, impact limité de l'absorption atmosphérique). La figure III-6 montre que les deux jeux de données se répartissent de manière similaire dans le plan factoriel des deux premières composantes. La dispersion plus importante des échantillons enregistrés avec le spectromètre IRTF (échantillons notés « Lab ») peut s'expliquer par la plus grande précision et intensité des bandes d'absorption sur les spectres enregistrés avec ce spectromètre. Cependant, malgré les différences de résolution spatiale et spectrale des deux spectromètres, les deux jeux de données peuvent être associés.

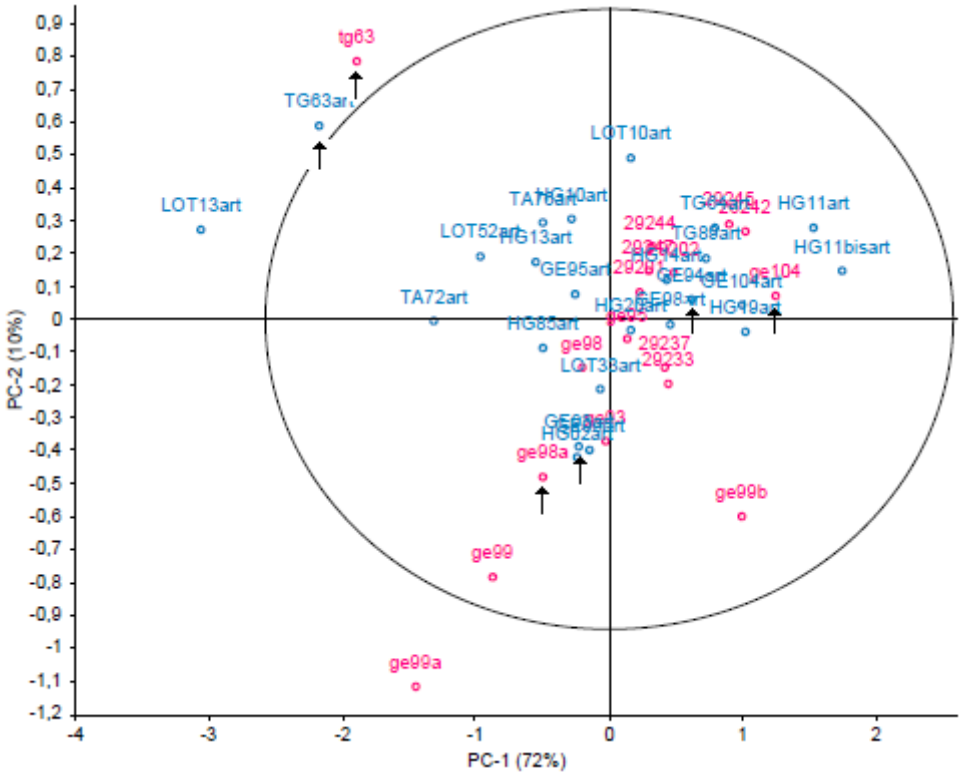


Figure III-7: Plan factoriel de l'ACP réalisée sur les spectres proche infrarouge enregistrés avec le spectromètre ASD en condition de terrain (rose) et en condition de laboratoire (bleu). Les longueurs d'onde correspondant aux différents domaines d'absorption atmosphérique ont été préalablement enlevées.

Dans le second cas, l'Analyse en Composantes Principales est réalisée sur les spectres de dérivées des sols, enregistrés dans le visible et le PIR, soit en condition de laboratoire soit sur le terrain, avec le spectromètre ASD FieldSpec FR®. La variance résiduelle des échantillons enregistrés en condition de terrain est plus importante que celle trouvée pour les échantillons enregistrés en condition de laboratoire. Cela s'explique par (i) la perte d'information liée aux absorptions atmosphériques lors de mesure in-situ (Cf. II.1.3), (ii) du changement de l'état de surface de l'échantillon (humidité, rugosité,…) du fait du changement d'échelle et de préparation de l'échantillon. En effet, en laboratoire, quelques grammes d'échantillon préalablement séchés et broyés sont utilisés. Cependant, bien qu'il y ait un décalage entre les données enregistrées en condition de laboratoire et en condition de terrain, la distribution des deux jeux de données est

similaire dans le plan factoriel des deux premières composantes (figure III-7). D'après Viscarra Rossel (Viscarra Rossel, Cattle et al. 2009) la différence observée peut être attribuée notamment à des différences d'albédo, de teneur en eau dans les sols et à la taille des particules. Il est intéressant de voir que les spectres d'un même échantillon de sols enregistré suivant les deux méthodes, se placent de la même manière dans le plan factoriel. C'est le cas pour les échantillons TG63, GE98 et GE104 (repérés par des flèches noires sur la figure III-7). Malheureusement le nombre de sols communs aux deux échelles d'analyse est limité. Plusieurs échantillons ont été retirés de l'analyse ACP (variabilité résiduelle trop importante) et plusieurs sols ont été mesurés uniquement en laboratoire du fait des mauvaises conditions météorologiques lors des campagnes de mesures sur le terrain. Ainsi, la représentativité de ces trois sols n'est pas vérifiée.

III.1.1.3 Discussion, conclusion

L'analyse en composantes principales permet de capturer, dans les premières composantes, l'information des bandes spectrales qui expriment la variation la plus importante. Cette information est indépendante de celle portée par les bandes d'absorption (composition minéralogique de l'échantillon). La première composante explique donc la variation des propriétés qui influencent le plus le modèle tout en restant trop approximative pour pouvoir mettre en évidence l'évolution d'une propriété bien précise telle que la teneur en smectite. Ainsi, si les résultats montrent des variations significatives dans les minéraux et les mélanges, ils ne renseignent pas sur la composante argileuse des sols.

Cette approche renseigne donc sur la structure générale de notre jeu de données et permet de faire des hypothèses sur le sens des relations entre les différentes données, tels que le caractère sableux de nos sols, ou la prédominance d'un des minéraux de référence (illite, smectite, kaolinite) dans les mélanges argileux. Cependant, elle ne permet pas d'en appréhender la complexité et son intérêt pour l'analyse de la minéralogie de la composante argileuse des sols est limité. En ce sens, l'ACP peut être considérée comme une méthode descriptive et non explicative.

III.1.2 La régression par moindres carrés partiels (PLSR : Partial Least Squares Régression)

La régression par moindres carrés partiels (PLSR) est une technique de modélisation statistique quantitative détaillée notamment par Wold et al. (Wold, Sjöström et al. 2001). De même que l'ACP, elle permet d'analyser des données fortement corrélées, mais utilise pour cela une base constituée d'un ensemble de données auxquelles correspondent des valeurs de la grandeur à prédire, mesurées par des méthodes de référence (notion d'apprentissage). Elle fait partie des méthodes de régression factorielle utilisées de manière générale pour décrire la variabilité des variables observées, au moyen de variables latentes. Elle constitue la méthode statistique la plus couramment utilisée dans les analyses spectrométriques et a été appliquée par de nombreux auteurs pour l'identification et la quantification de certaines propriétés des sols (Udelhoven 2003; Viscarra Rossel 2008).

III.1.2.1 Démarche générale

La démarche générale du développement d'une analyse utilisant la PLSR sur des données spectrométriques consiste à effectuer un étalonnage (ou échantillonnage étalon) qui permet de mettre en relation une valeur y de référence avec des mesures spectrométriques x. Le modèle statistique associé à la régression linéaire est donné par :

$$\mathbf{y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_m x_{im} + \varepsilon_i} \qquad \text{(équation 9)}$$

A	y_i	Valeur observée d'indice i,
	$x_{i1}, x_{i2}, \ldots, x_{im}$	Variables explicatives
	$\beta_0, \beta_1, \ldots, \beta_m$	Coefficients de régression
	ε_i	Terme d'erreur

La première étape de l'analyse consiste à choisir un ensemble d'échantillons représentatifs, ou étalons, de ceux qui seront analysés par la suite. La composition de ces échantillons est soit connue au moyen d'une analyse par une méthode de référence (DRX, analyses chimiques), soit expérimentale dans le cas des mélanges. Ainsi, nous disposons au départ, d'une matrice d'étalonnage **X** formée de *n* lignes et de *m* colonnes avec *n* le nombre de spectres formant la collection de référence et *m* le nombre de points de mesures (longueurs d'onde). Elle regroupe ainsi, les variables explicatives (valeur d'absorption à chaque longueur d'onde dans notre cas) et le vecteur **y** formé par les *n* valeurs de référence (teneur en smectite de chaque échantillon par exemple). Cette collection de références doit être suffisamment importante pour être représentative de la variabilité du paramètre que l'on cherche à modéliser. A partir de la collection d'étalonnage représentée par la matrice **X** et le vecteur **y**, nous cherchons ensuite à estimer les paramètres du modèle défini par l'équation (9).

Cependant, la structure des données spectrométriques est en général particulière. En effet, le nombre d'observations *n* est souvent largement inférieur au nombre de variables *m* et les collections de données spectrométriques sont fortement redondantes. La même information dont la variabilité est liée au même phénomène physique, apparaît dans plusieurs régions du spectre (par exemple, plusieurs régions spectrales peuvent souvent être associées à un même groupement chimique). On parle alors de **données colinéaires**. Dans ces conditions, l'utilisation de la PLSR est intéressante car elle permet tout comme l'ACP, de remplacer la matrice de données prédictives **X** par une nouvelle matrice **T** dérivée de **X**, comprenant le même nombre de ligne que **X**, mais un nombre de colonnes très inférieur à *m*. **T** est donc une nouvelle matrice dont les colonnes forment des **variables latentes (artificielles)**, obtenues par combinaison linéaire des variables d'origine, sur laquelle la régression est appliquée. Contrairement à l'ACP, le calcul des composantes de T se fait dans le cas de la PLSR, en tenant compte de la variable y. Ainsi, les variables latentes (artificielles) sont obtenues par un processus itératif qui maximise la valeur absolue de la covariance entre X et les concentrations y.

A partir de là, il est possible d'établir un modèle prédictif dont on peut faire varier la complexité en changeant le nombre de termes apparaissant dans l'équation de régression. Cependant, il est souvent très difficile de déterminer le meilleur compromis entre la complexité et la fiabilité des modèles. Le nombre **k** de composantes de la régression est un indicateur de cette complexité. Sur les données ayant servi à l'étalonnage, les erreurs résiduelles diminuent toujours en fonction de cette complexité. Ainsi, si la valeur de k est trop faible, on risque d'omettre des phénomènes d'interactions ou des dépendances de spectres et d'obtenir **un modèle avec un faible pouvoir explicatif**. En revanche, l'augmentation du nombre de composantes peut amener à inclure dans le modèle des variables latentes faiblement significatives. Cela conduit à un sur apprentissage du modèle (over fit, en anglais) qui ne décrit alors plus un rapport fondamental (la teneur en smectite *versus* les données spectrales) mais par exemple, le bruit contenu dans les données spectrales. Le risque est d'avoir un **bon modèle (très forte corrélation avec les données d'entrée) avec une capacité prédictive faible**.

Le nombre *k* de variables latentes (nombre de facteurs) définit donc la dimension du sous espace dans lequel opère la régression. La validation des modèles prédictifs en facilite la détermination. Elle consiste à appliquer les modèles sur la collection de validations (données indépendantes de l'étalonnage), pour des valeurs croissantes de *k*. On peut également utiliser la méthode de la validation croisée si l'on ne dispose pas d'un jeu de données suffisant pour créer deux ensembles indépendants. Cette dernière consiste à recommencer plusieurs fois l'étape de validation sur un petit nombre, chaque fois différent, de données issues de la collection de la calibration. On peut alors comparer l'évolution de l'erreur résiduelle liée à l'étalonnage avec celle obtenue dans les essais de validation, en fonction du nombre de facteurs utilisés (*k*). Cette erreur est désignée par le terme anglais Root Mean Square Error (**RMSE**, racine carrée de l'erreur quadratique moyenne). Il est généralement considéré qu'à la valeur minimale de cette erreur correspond la dimension « idéale » du modèle. Après que l'on a jugé que le modèle de régression était satisfaisant, on peut l'appliquer pour doser les échantillons de sols dont la valeur de référence (teneur en smectite) est inconnue.

Une synthèse non exhaustive des articles récents utilisant l'analyse PLSR dans leurs études est faite dans le tableau III-1. Elle met en évidence que la principale difficulté lors de l'utilisation de la PLSR vient de l'échantillonnage de données. En effet, les données doivent être suffisamment nombreuses et indépendantes pour permettre une validation réellement représentative et ainsi autorise l'utilisation des modèles pour prédire des données « inconnues ». La plupart des modèles sont établies sur des sols ou des échantillons propres à un certain environnement et certains auteurs (Shepherd and Walsh 2002; Brown 2006) montrent l'importance de l'utilisation d'une bibliothèque spectrale globale. Ainsi dans beaucoup d'études, seuls les résultats de la calibration et/ou de la validation croisée sont reportés (Janik 1998; Reeves, McCarty et al. 2001; Reeves, McCarty et al. 2002; Udelhoven 2003) et seulement peu d'études vont jusqu'à utiliser les modèles pour la prédiction sur un ensemble de données inconnues. Dans le cas d'étude géotechnique ce problème de l'échantillonnage est fondamental et compromet l'apport de telles méthodes.

article	objectif de l'étude	échantillons analysés	domaine IR	prétraitement	Analyses statistiques
Nanni and Demattê, 2006	estimation des propriétés des sols tropicaux brésiliens	368 sols tropicaux brésiliens	PIR	sélection de bandes spectrales caractéristiques d'après observations et bibliographie	régression linéaire multiple (SAS système): établissement de corrélation entre données spectrales et caractéristiques des sols
Viscarra Rossel and al., 2006	modélisation des mélanges minéraux / matières organiques	mélanges illite/kaolinite/smectite/goethite/50-50mix acide humique et fulvique/quartz	UV-vis-PIR	compression des données et données aberrante: **ACP;** Algorithme de débruitage utilisant les **ondelettes; dérivée première.**	modèle: **PLSR 1 / p**rédiction: **bootstrap aggregation-PLSR/** comparaison mélange/sol: **ACP**
Viscarra Rossel and al., 2009	estimation de la couleur et de la composition des sols; prédiction de teneur en argile des sols	10 profils de sol dérivant de matériaux parents différents: 249 échantillons; libraire de sols d'origine variable: 1361 sols	Vis-PIR	comparaison des spectres par **ACP et T-test.** utilisation en prétraitement d'un **filtre Stavisky-Golay avec les dérivées 1ères**	Retrait du Continuum et localisation de figures d'absorption caractéristiques / **PLSR avec Bootstrap aggregation** pour prédiction du contenu en argile.
Brown and al., 2006	évaluer la précision des modèles empiriques de caractérisation des sols à partir d'un set de données indépendante de sols	sols de différents pays, climats, enregistrés avec différents appareils	Vis-PIR		boosted regression trees versus PLSR
Savvides et al., 2010	prédiction de la CEC. Mise en évidence de la dépendance à l'échelle	échantillonnage emboité (hiérarchisé): au total: 288 pts	vis-NIR	réduction du nombre de variable: **ACP**	an**alyse de la covariance emboitée (hiérarchisée)**: corrélation DRS / CEC à différentes échelles / **PLSR**: prédiction de la CEC
McBratney et al. 2006	intégration de données issues des statistiques multivariables dans des fonctions de pedotransferts	116 ech. Provenant d'un champ de 16 ha.	MIR		PLSR; Bootstrap method+ bootstrat aggregating
Shepherd and Walsh, 2002	développement de librairie pour l'évaluation des sols. test: changement de taille du set de calibration; détection d'outliers lors de la prédiction de nouveaux échantillons	1000 sols de surface d'Afrique de l'est et du sud	Vis-PIR	**modèle linéaire graphique**: hypothèse d'indépendance des propriétés des sols.	**MARS:** calibration des propriétés des sols sur la réflectance; arbre de classification
Reeves et al.,2001	détermination de la composition des sols provenant de différentes profondeur, localisation, terre cultivé et taux de fertilisation	180 sols	MIR		**statistiques descriptives, PLSR**
Haaland et Melgaard, 2002	présentation de famille de calibration venant augmenter les capacités des analyses PLS classiques				
Odlare et al., 2005	variation spatiale de la teneur en argile, du pH et du carbone total dans les sols	99 ech. Prélevés sur une parcelle de 200*160 m	NIR	PCA	**Géostatistiques**
Janik et al. 2007	détermination du contenu en eau des sols	96 ech. de sols provenant de 36 sites localisés en Australie	MIR		PLSR; cross validation
Sørensen et Dalsgaard, 2005	prédiction de la teneur en argile, et en carbone total	plus de 800 ech provenant de tout le Danemark	NIR	dérivée première	PLSR; cross validation
Goetz et al. 2001	modélisation de la teneur en smectite et du potentiel de gonflement des sols	178 sols du Colorado	NIR	dérivée seconde	PLSR

Tableau 10 : Synthèse bibliographique non exhaustive des principaux articles récents utilisant la PLSR ou des méthodes statistiques similaires dans le domaine des sciences du sol.

III.1.2.2 Démarche sur les échantillons de l'étude

Les régressions par moindres carrés partiels de cette étude ont été réalisées avec le **logiciel Unscrambler® de chez CAMO**. Trois collections de données ont été constituées pour la calibration des modèles. La première collection est composée des minéraux de référence argileux (« pôles purs ») et non argileux et des mélanges synthétiques réalisés à partir de ces minéraux argileux de référence (kaolinite, illite, smectite). La deuxième est constituée des minéraux, des mélanges et des sols pour lesquels les teneurs en kaolinite, illite et smectite sont connues par les résultats de l'analyse par DRX couplée aux analyses chimiques. Enfin, la troisième collection se compose des mélanges et des sols (les différentes collections sont détaillées dans le chapitre 1). La première collection permet d'établir des modèles sur une base de données calibrées en laboratoire. La deuxième collection a pour but de tester si les modèles restent opérationnels lors de l'ajout d'échantillons plus complexes (les sols). Enfin la dernière collection permet de vérifier si une amélioration est possible en ne gardant que des données spectralement proches, celles des mélanges représentant grossièrement la fraction argileuse des sols. Il est ainsi possible d'étudier séparément l'influence des minéraux de référence et des sols dans la construction des modèles, afin de comprendre l'influence de la disparité des données. Les collections sont analysées dans le MIR puis dans le PIR (incluant les spectres provenant du spectromètre IRTF en réflexion diffuse et ceux du spectromètre de terrain) afin d'analyser l'influence de la gamme spectrale et des conditions d'enregistrement des spectres (résolution spectrale et spatiale, état initial de l'échantillon lors de la mesure…) sur la qualité des modèles.

Les modèles sont calibrés sur la **dérivée première des spectres** enregistrés dans le domaine du MIR et du NIR. La dérivée des spectres permet de faire ressortir les figures d'absorption les plus caractéristiques en mesurant le taux local de modification de l'absorbance en fonction de la longueur d'onde (Ben-Dor and Banin 1995; Chang, Laird et al. 2001; Sorensen 2005). Les données sont ensuite centrées, réduites puis transformées au cours de la calibration PLSR en un petit nombre de variables latentes combinées avec les spectres et les valeurs de kaolinite ou illite ou smectite dans les modèles de calibration. L'algorithme PLS2, qui constitue une variante à l'algorithme utilisé précédemment, est également testé. Il paraît plus intuitif pour l'analyse de mélange car les trois minéraux argileux sont utilisés simultanément comme variable Y et non à tour de rôle comme dans le premier modèle. Cependant il semble que l'algorithme initial surpasse cet algorithme en termes d'interprétation et de prédictibilité du modèle. Du fait de la taille relativement réduite des collections de calibration (entre 60 et 117 individus suivant la gamme spectrale d'analyse), **la méthode de la validation croisée a été choisie**. Elle permet de déterminer le nombre de facteurs (composantes) retenus pour les modèles de calibration. 10 facteurs (composantes) ont été testés. **On sélectionne ensuite le nombre de facteurs pour lequel la RMSE est minimale**. Si tous les individus de la validation croisée sont correctement identifiés et placés dans un groupe unique par types de produit, le modèle est alors prêt à être utilisé pour la détermination de l'identité d'échantillon inconnu.

Enfin, il est à noter que, de la même manière que pour l'ACP, des valeurs aberrantes (outliers) mises en évidence lors de la calibration par leur important degré d'influence et/ ou leur importante variable résiduelle, sont retirées du modèle.

III.1.2.3 Résultats

Les valeurs mesurées dans le MIR et dans le PIR avec le spectromètre IRTF et le spectromètre de terrain sont comparées par une régression aux valeurs prédites. Les résultats de la calibration et de la validation des modèles établis pour la quantification de la teneur en kaolinite, en illite et en smectite sont présentés en détail dans l'annexe 3 et sont résumés dans le tableau III-2.

Afin de démontrer de quelle manière les interprétations tirées du tableau III-2 sont faites, l'exemple de la kaolinite est choisi. Le nombre de facteurs sélectionnés pour l'établissement des modèles établies pour la prédiction de la teneur en kaolinite est dépendant de la RMSE obtenue. Le nombre de facteurs pour lequel la RMSE est minimale est choisi. Suivant la collection de données et la gamme de longueurs d'ondes étudiée, ce nombre varie de 1 à 7 facteurs (composantes). La précision de la prédiction (définie par la RMSE) est en moyenne proche de 7% pour des teneurs en kaolinite allant de 10 à 90% dans les mélanges et de 0 à 25% pour les sols. Les écarts entre la RMSE de calibration et de validation sont faibles (RMSE de validation moyenne de 8%). Les coefficients des régressions linéaires établies entre les valeurs mesurées et prédites sont supérieurs à 0.79 dans le cas de la calibration et de la validation (valeurs entre parenthèses) ce qui est comparable aux valeurs trouvées dans la littérature.

L'utilisation de la collection de données n'incluant pas de minéraux de référence (uniquement des mélanges et des sols) améliore la précision des prédictions (RMSE de 5.6% contre 7.2 % pour la collection de données composée de minéraux (pôles purs), de mélanges et de sols) sur l'ensemble des gammes de longueur d'onde (figure III-8). Cela s'explique par l'importance des mélanges dans la collection de calibration, par leur nombre et par le fait qu'ils portent la variation de concentration des différentes espèces étudiés. Les minéraux de référence (pôles purs) ont un signal spectral fort mais la variabilité de leur concentration se limite à présent (100%) ou absent (0%). La présence de sols dans la collection de calibration n'influence que très peu la modélisation. Cela semble être dû soit à la teneur en minéraux argileux qui est trop faible pour être significative vis-à-vis de la concentration dans les mélanges, soit parce que les sols ne complexifient pas la collection de calibration au regard de la quantification de la kaolinite.

minéral	set de données	domaine infrarouge et spectromètre utilisé	nb d'individu utilisé	facteur utilisé pour la PLSR	RMSE	R^2
kaolinite	minéraux, mélanges, sols	MIR (IRTF en transmission)	121	2	7,79 *(8,21)*	0,88 *(0,86)*
		PIR (IRTF en réflexion diffuse)	92	2	6,7 *(7,39)*	0,94 *(0,93)*
		vis-PIR (ASD)	75	3	7,02 (8,83)	0,92 (0,88)
	mélanges, sols	MIR (IRTF en transmission)	104	4	5,47 *(6,8)*	0,93 *(0,89)*
		PIR (IRTF en réflexion diffuse)	72	1	5,83 *(6,05)*	0,95 *(0,95)*
		vis-PIR (ASD)	74	7	5,41 *(9,01)*	0,96 *(0,89)*
	minéraux, mélanges	MIR (IRTF en transmission)	101	2	7,32 *(7,83)*	0,91 *(0,89)*
		PIR (IRTF en réflexion diffuse)	63	1**(3)**	7,81 *(8,42)* **(4.4)**	0,94 *(0,94)* **(0,91)**
		vis-PIR (ASD)	72	4	10,39 *(14.14)*	0,87 *(0,79)*
smectite	minéraux, mélanges, sols	MIR (IRTF en transmission)	124	4	14,19 *(17,15)*	0,76 *(0,66)*
		PIR (IRTF en réflexion diffuse)	90	6	7,72 *(10,92)*	0,84 *(0,76)*
		vis-PIR (ASD)	87	1	15,55 *(16,53)*	0,696 *(0,67)*
	mélanges, sols	MIR (IRTF en transmission)	107	7	10,74 *(13,96)*	0,84 *(0,72)*
		PIR (IRTF en réflexion diffuse)	72	8	4,68 *(8,43)*	0,96 *(0,89)*
		vis-PIR (ASD)	74	5	5,51 *(8,29)*	0,96 *(0,90)*
	minéraux, mélanges	MIR (IRTF en transmission)	103	4	14,1 *(17,7)*	0,77 *(0,65)*
		PIR (IRTF en réflexion diffuse)	63	4 **(2)**	11,68 *(14,73)* **(3,8)**	0,85 *(0,77)* **(0.96)**
		vis-PIR (ASD)	72	8	4,14 *(10)*	0.98 *(0.73)*
illite	minéraux, mélanges, sols	MIR (IRTF en transmission)	116	3	14,09 *(15,72)*	0,7 *(0,63)*
		PIR (IRTF en réflexion diffuse)	92	10	6,95 *(12,84)*	0,93 *(0,76)*
		vis-PIR (ASD)	84	10	5,857 *(18,57)*	0,96 *(0,57)*
	mélanges, sols	MIR (IRTF en transmission)	102	9	8,58 *(11,66)*	0.89 *(0.79)*
		PIR (IRTF en réflexion diffuse)	72	6	7,77 *(11,29)*	0,88 *(0,75)*
		vis-PIR (ASD)	74	8	6,21 *(12,45)*	0,94 *(0,77)*
	minéraux, mélanges	MIR (IRTF en transmission)	95	4	12,31 *(14,79)*	0,80 *(0,72)*
		PIR (IRTF en réflexion diffuse)	63	7 **(2)**	13,1 *(19,29)* **(3,9)**	0,84 *(0,68)* **(0.99)**
		vis-PIR (ASD)	72	10	5,89 *(25,54)*	0.97 *(0.37)*

Tableau 11: Statistiques de la calibration et de la validation croisée (entre parenthèse) pour la prédiction de la teneur en kaolinite, smectite et illite utilisée dans les données de calibration. Les valeurs indiquées en rouge sont issues d'un travail de Viscarra Rossel, les modèles étant établis sur une collection de données incluant des minéraux argileux et non argileux typiques des sols de la région étudiée, des mélanges de minéraux argileux et des mélanges plus complexes incluant en plus des composés organiques (Acides fulviques et humiques) et de la goethite et du quartz (Viscarra Rossel 2006).

Les valeurs indiquées en rouge dans le tableau III-2 sont tirées d'une étude réalisée par Viscarra Rossel en 2006 dans laquelle il établi des modèles de prédiction de la teneur en illite, smectite et kaolinite à partir d'une collection de calibration similaire (minéraux, mélange et sols) à ceux utilisés dans cette études. La meilleure précision (RMSE) obtenue sur les modèles peut s'expliquer par la complexité des mélanges utilisés ainsi que par la taille de la collection de calibration et enfin par la méthode de validation utilisée (collection de données indépendantes dans leur étude, validation croisée dans notre étude). En effet, les mélanges étudiés dans notre étude représentent la fraction argileuse des sols alors que des mélanges comprenant des teneurs en kaolinite, illite, smectite, goethite, quartz et matière organique variables sont utilisés dans leur étude. Ainsi, dans notre cas, toute les variations présentent dans les sols ne sont pas forcement présentent dans les exemples ayant servis à l'apprentissage du modèle ou bien la variation d'une concentration d'un composé secondaire est en fait constant sur tout les individus de la base d'étalonnage. Par exemple, la calcite est présente uniquement dans les minéraux de référence et la variabilité de sa concentration se limite donc à présente (100%) ou absente (0%). Cela démontre bien l'influence de la collection de calibration sur la précision des modèles et par là même la limitation que cela représente pour l'utilisation de ces modèles en routine sur des échantillons différents de ceux des données d'entrée du modèle.

Comparaison des résultats de la cross-validation.

Afin de comparer les capacités prédictives de chaque collection de données, enregistré à différentes résolutions spatiales et spectrales dans différentes gammes de longueurs d'ondes, pour les différents minéraux argileux (kaolinite, illite et smectite), les valeurs de RMSE et de R^2 sont représentées dans les graphiques des figure III-8 (pour la kaolinite) III-9 (pour la smectite) et III-10 (pour l'illite). Les modèles établis pour la prédiction de la teneur en illite et en smectite s'ajustent bien aux données (Annexe 3). Ils présentent une RMSE de calibration inférieur à 10 % pour l'ensemble des collections de calibration utilisées mais l'écart entre la calibration et la validation croisée est beaucoup plus important que dans le cas de la kaolinite et un nombre généralement plus important de facteurs (composantes) est utilisé pour la modélisation (Tableau III-2). D'autre part des différences importantes sont obtenues suivant les différentes longueurs d'ondes et les différentes collections utilisées dans la modélisation (figure III-9 et III-10), ce qui traduire un manque de stabilité du modèle au regard des résultats très proches obtenus dans le cas de la kaolinite. Globalement, le PIR permet de prédire la teneur en illite et en smectite avec la meilleure précision (RMSE).

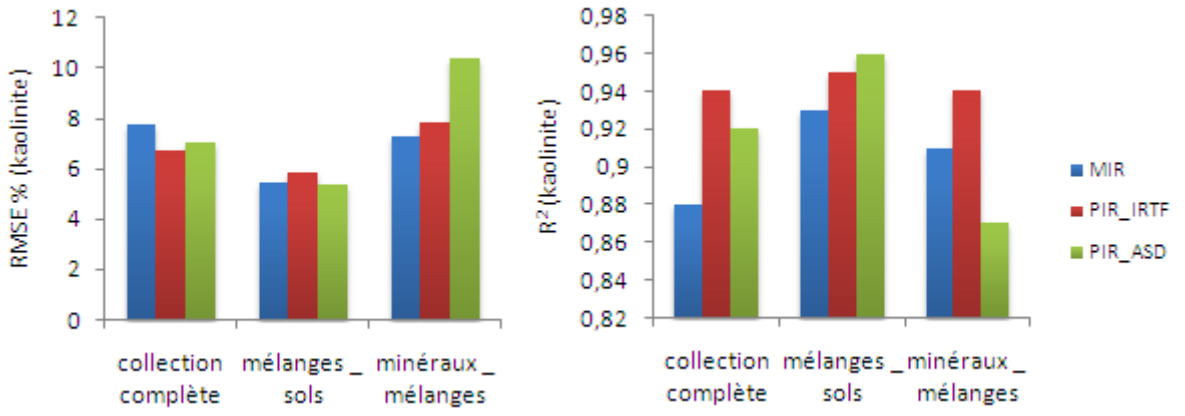


Figure III-8: Comparaison entre les prédictions de la teneur en kaolinite obtenues pour les différentes collections dans le domaine spectral MIR et dans le NIR, d'après la valeur de RMSE (a) et la valeur de R^2

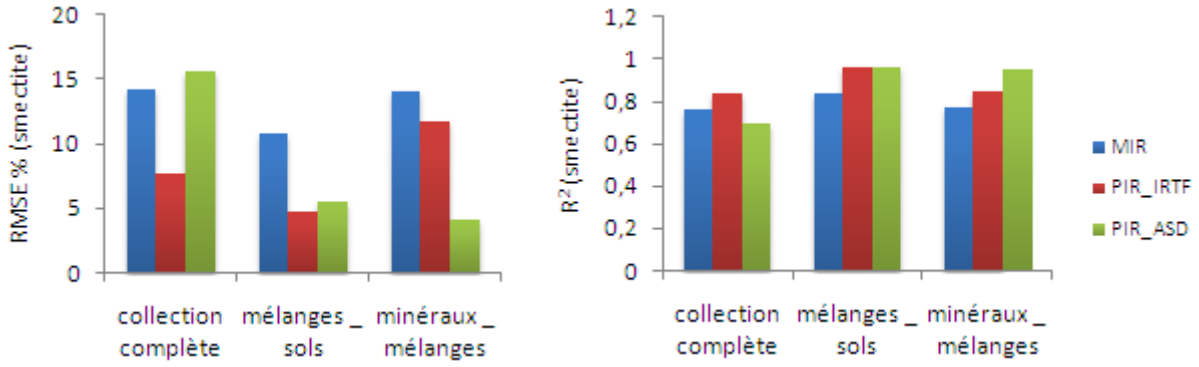


Figure III-9: Comparaison entre les prédictions de la teneur en smectite obtenues pour les différentes collections dans le domaine spectral MIR et dans le NIR, d'après la valeur de RMSE (a) et la valeur de R^2

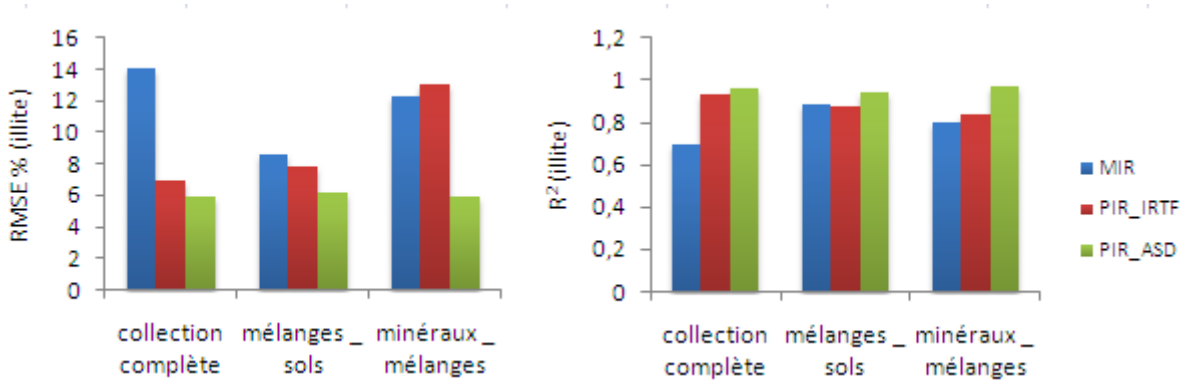


Figure III-10: Comparaison entre les prédictions de la teneur en smectite obtenues pour les différentes collections dans le domaine spectral MIR et dans le NIR, d'après la valeur de RMSE (a) et la valeur de R^2

Les figures III-11 et III-12 montrent que la prédiction de la teneur en kaolinite d'échantillon de sols réalisée à partir de ces modèles, donne des résultats proches des valeurs mesurées malgré les faibles teneurs en kaolinite des sols de la région Midi-Pyrénées (moins de 30%). La précision des prédictions est sensiblement la

même pour les différents sets de données. Dans le MIR les coefficients de régression sont de 0.63 pour la collection complète, 0,56 pour la collection associant les mélanges synthétiques et les sols et 0.64 pour la collection associant les minéraux argileux et non argileux de référence et les mélanges synthétiques de minéraux argileux et/ou de quartz. Dans le PIR les coefficients sont en moyenne de 0.58. L'erreur moyenne de prédiction est proche de 10%.

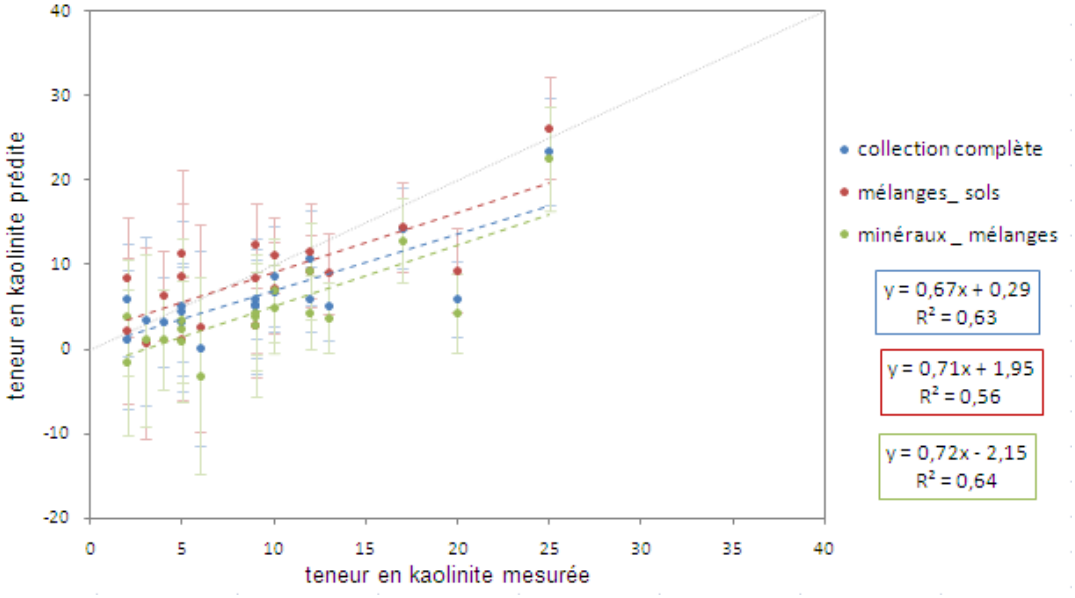


Figure III-11: Prédiction de la teneur en kaolinite des sols réalisée à partir des modèles établis sur les différentes collections de données enregistrées dans le MIR.

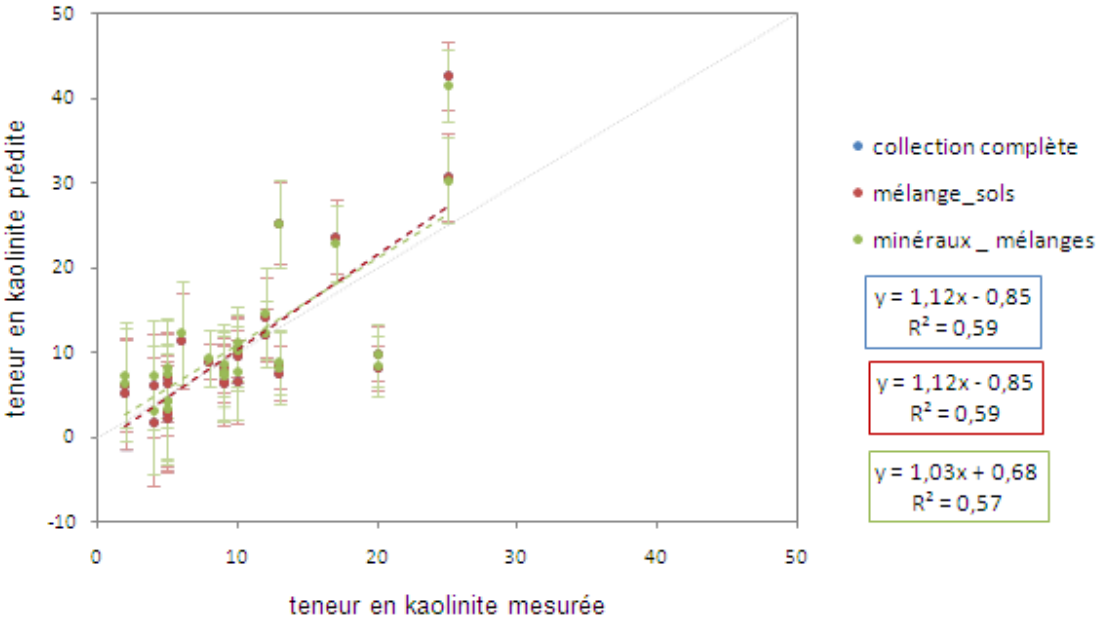


Figure III-12: Prédiction de la teneur en kaolinite des sols réalisée à partir des modèles établis sur les différentes collections enregistrées dans le PIR avec le spectromètre IRTF en réflexion diffuse. La régression réalisée pour la collection complète et pour la collection associant les mélanges et les sols se superposent.

En revanche les figures III-13 et III-14 montrent que les modèles établis pour la prédiction de l'illite et de la smectite ne sont pas robustes. En effet, lors du passage à la phase de prédiction des teneurs en illite et en smectite dans les sols de la région. Des taux de réussite assez faibles sont observés comme le montre les faibles coefficients de corrélation qui existent entre les teneurs prédites et les teneurs. L'utilisation des différents domaines spectraux (PIR, MIR) n'apporte aucune amélioration quant à la prédiction. Si l'on confronte nos résultats à ceux de Viscarra Rossel et al. (2008) on peut constater que ses modèles sont beaucoup plus robustes que les nôtres et permettent effectivement la prédiction des teneurs en kaolinite, illite et smectite dans les mélanges synthétiques réalisés dans son étude, avec une erreur comprise entre 10 et 20%. Cependant il est important de noter qu'aucune prédiction de kaolinite, illite et smectite n'a été testée sur des sols « naturels » comme c'est le cas dans notre étude. Il semble donc que ce changement de complexité entre les mélanges synthétiques et mélanges naturels (sols) engendre une forte augmentation des nombres de paramètres à prendre en compte dans les modèles de prédiction.

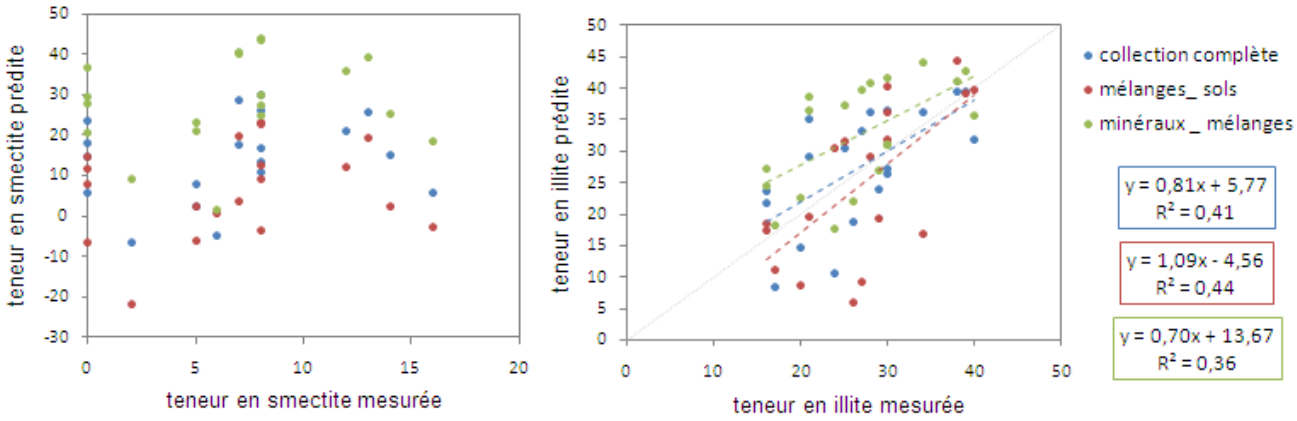


Figure III-13: Prédiction de la teneur en smectite (graphique de gauche) et en illite (graphique de droite) des sols réalisée à partir des modèles établis sur les différentes collections de données enregistrées dans le MIR.

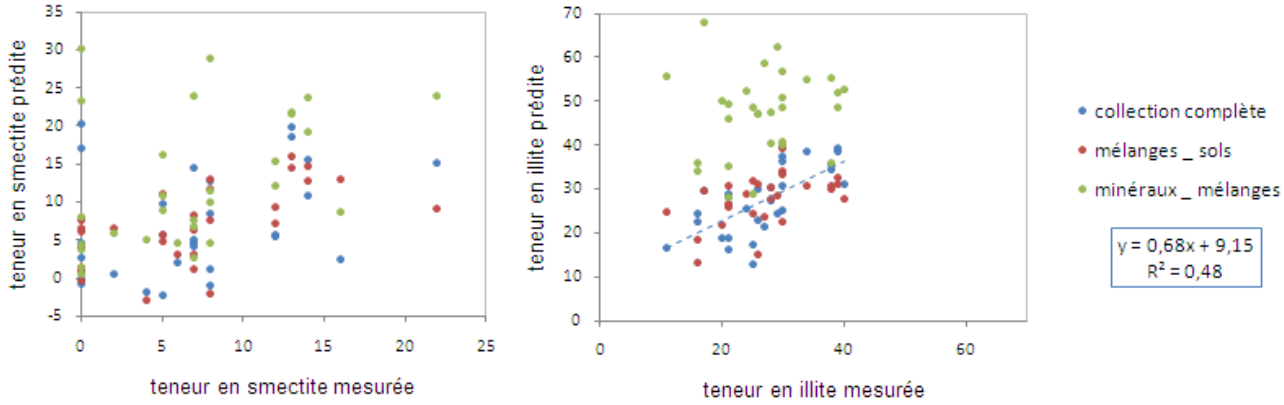


Figure III-14: Prédiction de la teneur en smectite (graphique de gauche) et en illite (graphique de droite) des sols réalisée à partir des modèles établis à partir de la collection complète de données (bleu), la collection regroupant les données de mélanges et de sols (en rouge) et la collection associant les minéraux de référence et les mélanges (en vert). Les données ont été enregistrées dans le PIR avec le spectromètre IRTF en réflexion diffuse.

III.1.2.4 Discussion, conclusion

L'application de la méthode de régression par moindres carrés partiels (PLSR) montre de fortes potentialités mais pose aussi certains problèmes tels que l'inter-corrélation des variables. Cette situation aboutit à un modèle très incertain. En effet, les variables sont par nature très corrélées car nous mesurons un spectre qui possède une certaine continuité : le signal mesuré à une longueur d'onde donnée est très dépendant de ceux mesurés aux longueurs d'ondes voisines. De plus, un trop grand nombre de variables explicatives peut s'avérer néfaste: il augmente la probabilité de capter des perturbations et entraine une complexification du modèle, ce qui induit une augmentation du nombre de variables latentes. L'information non utile contenue dans les spectres rend la construction des variables latentes moins pertinentes. Elle ajoute de l'erreur en prenant en compte des zones spectrales qui ne sont pas reliées à la concentration des minéraux considérés.

Ainsi la construction de la collection de calibration est fondamentale. Dans notre cas, l'apprentissage des modèles est essentiellement basé sur les mélanges synthétiques. Ce sont eux qui portent l'essentiel de l'information suivant une large gamme de concentration. Cependant, si la kaolinite présente des bandes d'absorption spectrales caractéristiques qui permettent de la différencier dans un mélange de minéraux argileux, la smectite et l'illite ont des spectres de réflectance beaucoup plus proches dans le MIR et de le PIR (Cf. chapitre 1). Il est alors difficile de différencier ces deux minéraux argileux car la quantité d'information identique dans un spectre d'illite et de smectite est largement supérieure à la quantité d'information spécifique aux spectres de l'illite et de la smectite. La PLSR n'est donc pas un outil adapté pour faire ressortir ces différences qui se retrouvent noyées dans l'incertitude de la méthode notamment dans le cadre de l'étude des sols qui ont des teneurs cumulées en illite et en smectite généralement inférieures à 40%.

D'autre part les modélisations et les prédictions de la teneur en kaolinite des sols sont de qualité équivalente quelque soit la gamme de longueurs d'ondes étudiées. Cela indique que la collection de calibration est correcte pour les deux gammes de longueurs d'ondes.

Bien que les résultats soient encourageants pour la quantification de la kaolinite, l'utilisation en routine de la régression par moindres carrés partiels n'est pas simple car elle nécessite d'avoir de fort échantillonnages et une bonne connaissance du formalisme des modèles pour pouvoir interpréter les résultats obtenus. Ainsi, son utilisation par les bureaux d'études est difficilement envisageable. En effet, l'établissement de la collection de calibration entraine plusieurs difficultés. Elle doit inclure un nombre important de mélanges argileux ou de sols pour lesquels des analyses chimiques ou de DRX sont nécessaires pour connaître la composition minéralogique de la fraction argileuse. Ces méthodes n'étant pas quantitatives au sens strict, une erreur importante est à prendre en compte dans le modèle. D'autre part la calibration requière un jeu de données plus important et plus diversifié afin de mieux cibler les paramètres susceptibles d'interférer sur les modèles.

De plus le processus de sélection des variables explicatives en fonction de leurs niveaux de signification est empirique et peut conduire des erreurs importantes dans l'estimation des coefficients de régression. Les variables estimées de cette façon peuvent n'avoir aucune valeur prédictive et peuvent induire des erreurs de prévisions importantes.

III.2. APPROCHE BASEE SUR LA FORME DES ABSORPTIONS CARACTERISTIQUES DES MINERAUX ARGILEUX

Cette partie constitue la suite des travaux de Kariuki (Kariuki 2004) et de Lagacherie (Lagacherie, Baret et al. 2008) . Elle se base également sur plusieurs études en minéralogie dans le domaine du MIR(Joussein 2001; Madejova 2002).

Comme nous l'avons présenté dans le chapitre 1 paragraphe I.2., l'absorption de la lumière est reliée aux propriétés des milieux dans lesquels elle passe par la loi de Beer-Lambert, selon laquelle la concentration et les coefficients d'absorption d'un matériel déterminent le niveau d'absorption. Elle permet d'obtenir une relation linéaire entre la concentration de l'élément absorbant et sa contribution au log (1/R) à la longueur d'onde absorbée. Les paramètres de la figure d'absorption (position, aire, profondeur, largeur et assymetrie) sont contrôlés par la structure cristalline particulaire.

III.2.1 Publication 1: Clay mineralogy of midi-Pyrénées area (France) by infrared spectroscopy: disappointed results (soumis à Clay Minerals)

TRUCHE CAMILLE[1]*, DARROZES JOSE[1], PETIT SABINE[2], BOURGUIGNON ANNE[3], ROUBICHOU PHILIPPE[3], MARTIN FRANÇOIS[1]

[1] LMTG - UMR CNRS / IRD / UPS, Toulouse, 31400, France
[2]UMR6532 CNRS, HydrASA, Université de Poitiers, 86022, France
[3]BRGM, Orleans, 42000, France

III.2.1.1 Résumé en Français

Introduction, cadre géologique, matériels et méthodes

La région Midi-Pyrénées est caractérisée par une forte sensibilité au risque de retrait et de gonflement des sols du fait à la fois, de la nature argileuse de ces formations superficielles mais également de l'augmentation de l'urbanisation diffuse durant ces cinquante dernières années (sécheresse de 1989, 1992 puis 2003). Ce type d'urbanisation se traduit par une augmentation du nombre d'habitations potentiellement concernées par ce risque, et par une augmentation de la demande de prévention. Dans cette étude nous proposons l'utilisation de la spectrométrie infrarouge pour répondre aux besoins des bureaux d'études en terme de méthodologie de reconnaissance et de quantification des minéraux argileux des sols. L'étude se base sur des minéraux naturels de référence, des mélanges synthétiques réalisés à partir de ces minéraux et sur des

sols échantillonnés dans la région. Les mesures spectrométriques ont été réalisées dans la gamme de longueurs d'ondes du MIR et du PIR avec différents spectromètres. Pour chaque échantillon nous avons mesuré :

- Dans le MIR, l'aire des pics correspondant aux vibrations fondamentales des liaisons OH des minéraux argileux.
- Dans le PIR, après retrait du continuum, l'aire du pic à 4520 cm^{-1} (2200 nm). Ce pic est le résultat de la combinaison de l'élongation de la liaison OH et de la déformation de la liaison Al-OH. Elle est présente dans le spectre de tous les minéraux et elle n'est pas influencée par les vibrations de la molécule d'eau, ce qui assure sa présence lors des mesures in-situ.

(Les échantillons et la méthode sont décrits avec plus de précision dans le chapitre 1).

Résultats

En laboratoire, quelque soit la gamme infrarouge étudiée, il est possible de mettre en évidence le caractère dominant de la kaolinite dans les mélanges et les sols. Cela se traduit par une relation exponentielle entre l'aire des pics considérés et la teneur en kaolinite des mélanges et des sols (figure 5, 6, 7 et 8). Aucune relation permettant de distinguer l'illite de la smectite n'est trouvée.

Les mêmes résultats ont été obtenus avec le spectromètre IRTF en réflexion diffuse et le spectromètre de terrain en condition de laboratoire (utilisation d'une source de lumière artificielle). En revanche l'utilisation in-situ du spectromètre de terrain ne permet pas d'établir des corrélations entre l'aire du pic d'absorption à 4520 cm^{-1} (2200 nm) et la teneur en kaolinite (figure 9).

Discussion et conclusion

Cette étude permet tout d'abord de souligner la limite à l'utilisation du pic d'absorption à 4520 cm^{-1} (2200 nm) pour l'analyse de la teneur en argile des sols (Lagacherie, Baret et al. 2008) puisque tous les minéraux argileux ne l'influencent par de la même manière.

D'autre part cette étude montre que de la même manière que pour les études statistiques, le choix d'un échantillonnage représentatif est très important. Nous montrons ainsi qu'il est possible d'obtenir une relation régionale pour l'analyse de la teneur en kaolinite dans les sols. En effet, les minéraux argileux des sols de la région ont généralement la même histoire géologique.

Nous recommandons aux bureaux d'études de faire les mesures en laboratoire sur des échantillons calibrés. En effet, les analyses peuvent ainsi se faire par tout temps et la précision sur les résultats est beaucoup plus avantageuse en terme d'analyse de la fraction argileuse. En effet, l'analyse in-situ des sols se fait sur des échantillons non calibrés et les paramètres extérieur à la composition minéralogique de l'échantillon (humidité, rugosité…) gènent l'analyse.

Enfin, l'apport d'une telle étude pour les bureaux d'études porte sur l'amélioration de la cartographie des zones riches en kaolinite et donc en illite plus smectite par rapport à la carte des formations argileuses

actuellement utilisée pour la réalisation des cartes d'aléa. La spectrométrie apporte ainsi une information spatialisée complémentaire aux analyses ponctuelles actuellement utilisées (analyses géotechniques et DRX).

III.2.1.2 Abstract

Study of clay minerals has been required to cope with the damages caused by expansive soil hazard on infrastructures, especially in the Midi-Pyrenees area (Southern France) which is subject to an increasing urbanization on this type of soil. As part of a regional project concerning the search of technical solution for the improvement of high-risk areas, the feasibility of middle (MIR) and near infrared (NIR) spectroscopy for laboratory, and field characterization of clay mineralogy is evaluated. An operational method based on spectral feature measurement is preferred to sophisticated multivariable statistical techniques to produce homogeneous and continuous mineralogical information. From experimental reference situations towards in situ observations, an evaluation of the method capabilities to determine the relative abundances of the clay minerals from NIR and MIR soil spectra is made. Therefore, some correlations are established comparing the spectra recorded in these two ranges of analysis, on the basis of the information on structural OH groups in clay minerals. In laboratory, the mineralogical quantification of synthetic mineral mixtures is studied in the NIR and the MIR. We show that a primary analysis of absorption features of NIR and MIR spectra allows a rough quantification of kaolinite content. The combination of laboratory and field measurements is considered. The applicability of the IR spectroscopy for the spatialization of the results will then be discussed.

III.2.2 Introduction

The important demographic pressure and the rising urbanization increase the vulnerability of territories. The economic development results in a growing occupation of risk areas leading to a potential danger for people and buildings during natural disasters due to swelling soil, or landslides. More particularly, Midi-Pyrenees in Southern France is faced with a growing diffuse urbanization linked to the development of aeronautic and spatial activities and to the attraction that represents rural landscape. This type of urbanization associated to the specific clay composition of soil, and repeated summer droughts is a challenge for the prevention of damage from swelling soils. At regional scale, the mapping of high-risk areas is based on maps of clay rich formation constructed using existing geological and soil maps from scaled varying from 1/50000 to 1/200000. They are specified by punctual X-ray diffraction and geotechnical analyses as plasticity index or blue methylene values VBS. These regional maps do not satisfy the needs in terms of prevention of local risks. At local scale, very punctual geotechnical studies are led by engineering consulting firms.

To cost-effectively improve the present prevention possibilities, laboratory and field infrared spectroscopy has been proposed as a rapid and inexpensive method of measuring and mapping clay content. Since many years, spectroscopic techniques (visible (VIS), near (NIR) and mid infrared (MIR) spectroscopy) have been considered as possible alternatives to enhance conventional laboratory methods of soil analysis

(Janik et al., 1998; Goetz et al., 2001). This process is non-destructive and therefore allows the preservation of the basic integrity of the soil system. Moreover, these spectroscopic techniques are fast and inexpensive and do not require long time pre-processing of samples or use of chemical extractants. Intense fundamental molecular vibrations bands which appear in the MIR range (4000 to 400 cm^{-1}) allow the identification of clay minerals. The determination of the specific mineral chemistry is consequently possible (Petit et al., 1995; Madejová and Komadel, 2001; Madéjová, 2003; Wilson, 1994). In NIR domain (12500 to 4000 cm^{-1}), overtones and combinations of the fundamental vibrations due to the stretching and bending of OH groups dominate. Since two decades, NIR spectroscopy has being developed for soil property analysis like clay and organic matter content, organic carbon, pH and cation exchange capacity (Ben-Dor et al., 1997; Viscarra Rossel and McBratney, 1998; Odlare et al., 2005; Sorensen and Dalsgaard, 2005; Brown et al., 2006; McBratney et al., 2006; Brown, 2007; Waiser et al., 2007). Thus, NIR spectroscopy appears as a major tool in soil sciences for precision agriculture (Shepherd and Walsh, 2002; Odlare et al., 2005; Waiser et al. 2007) and for prediction of swelling soil (Janik et al., 1998; Kariuki et al., 2003; Kariuki et al., 2004; Brown et al., 2006). In this study, two spectrometers have been tested under different conditions in the MIR and NIR ranges to choose the best compromise between method performance, analysis rapidity and professional needs.

The infrared data processing by a simple and reproducible method is privileged. Multivariable statistical methods allow quantitative spectral analysis of soil using near (NIR) and middle (MIR) infrared. Among these, partial least-square regression (PLSR) is the most usually considered. Many studies show that predictions of the amount of clay minerals are possible with this technique (Goetz et al., 2001; Viscarra Rossel et al., 2006). However, the creation of a robust empirical model and the use of this method as a routine are widely dependent on the quality and the size of the calibration set. Thus, it requires a large soil sampling with X-ray diffraction and geotechnical measurements. The need of this *a priori* knowledge limits the feature of the approach within the framework of an operational use. Concurrently with the statistical approach, continuum removal (CR) technique and measurement of specific absorption band surface (Kariuki et al., 2003; Kariuki et al., 2004) is also used for estimating clay content. The relation between the integrated intensity of absorption features in the characteristic wavelength and the concentration of the chemical group responsible for the absorption is used.

For preventing swelling soil damage, estimations of clay minerals relative abundances at large scales allow to identify the high-risk areas to focus geotechnical analyses. Some studies have already showed that it is possible to have information about the clay content of soil (Ben-Dor and Banin 1995; Odlare 2005; Lagacherie, Baret et al. 2008). In this paper, the combination of laboratory and field measurements is considered for providing a characterization and a quantification of the main clay minerals. Indeed, the swelling behavior of soil sample is greatly influenced by the amount and type of clay minerals in soil. The applicability of the IR spectroscopy for the spatialization of the results will then be discussed.

III.2.3 Geological setting

Our study area is located in the South-East of the Aquitaine basin in Southern France. The Aquitaine basin is composed of different geomorphological units related to geological characteristics. In the purpose of the present study, three main types have been distinguished; (i) a limestone plateau, (ii) a molassic area and (iii) a detrital quaternary sedimentary zone. Our study area is mainly located in the molassic area which covers 25000 km^2. “Mollasse” formation results from the diagenetic evolution of sediments issued from the dismantling of the Massif Central and of the Pyrénées during and after Alpine orogenesis. The materials were deposited in the subsidence basin following a succession of coarse and clastic rhythmic deposits. This is repeated many times during the Tertiary. The ages of basin deposits are ranged from the Eocene to the Miocene. This area is composed by almost all of the outcropping Tertiary. In the North area, there are some molasses with a dominance of limestone, but the most important part of this area is a vast tertiary fluvial floodplain cut in terraces by the quaternary valleys of the Garonne River and its tributaries. The soil map of Midi-Pyrenees area together clay characterization are given in figure 1.

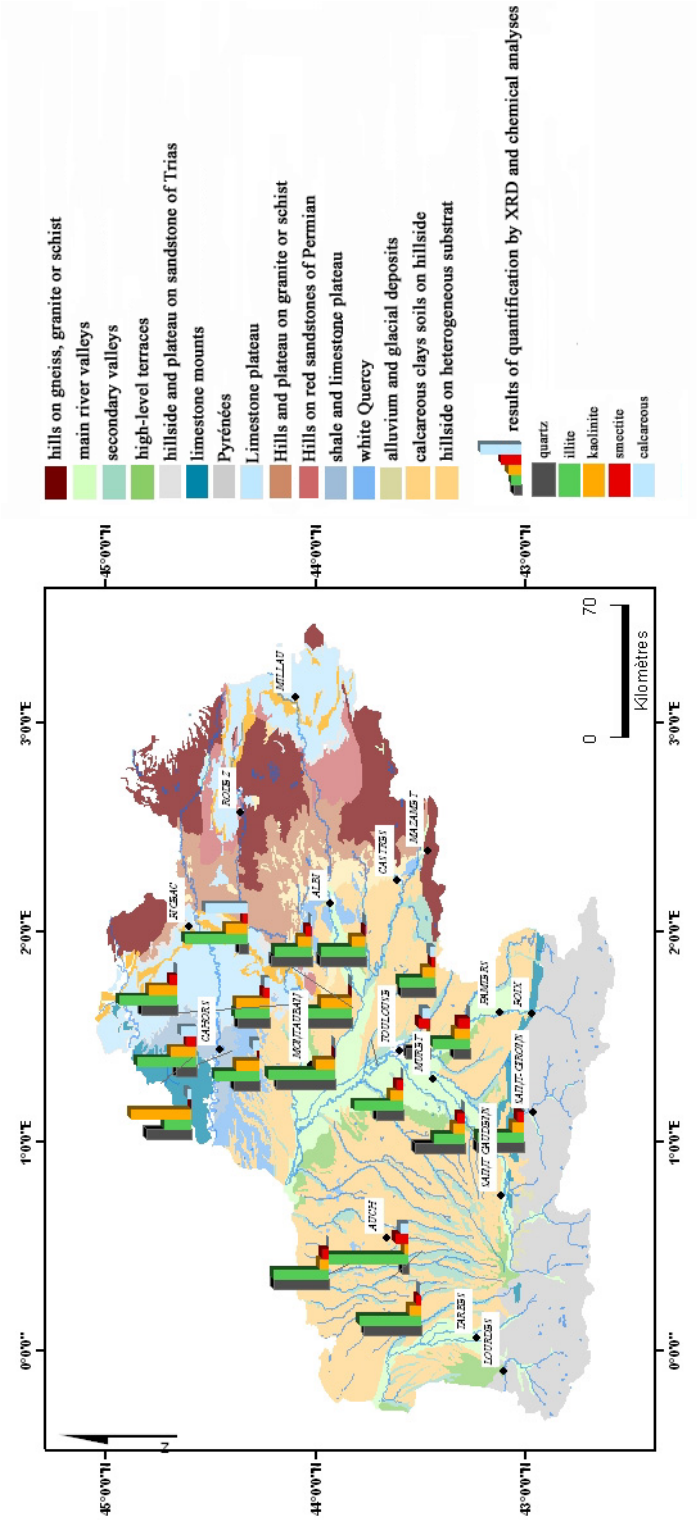


Figure III-15: Soil map of Midi-Pyrénées (southern France) provided by the "Chambre Regional d'agriculture de Midi-Pyrenees». The north of the region is mainly represented by limestone plateau. The center is mainly composed by oligo-miocene molasses and alluvion quaternary deposits. The result of chemical and XRD analysis of soil samples are figurated as bar diagram.

III.2.4 Materials and methods

III.2.4.1 Sampling

As the most common soil clay minerals, kaolinite, illite and smectite, were used in this study to prepare double and triple mechanical mixtures. The standard clay minerals used were the illite from "le Puy en Velay" (France), a smectite from ABM society (France), and a natural kaolinite from "AGS minéraux" (Clérac-France). Their mineralogical composition is given in table 1. Mechanical mixtures were prepared by mixing these standard minerals in various molar ratios to get homogeneous mixtures.

Standard minerals	Main phase	Secondary phases
illite « le Puy en Velay »	illite	
kaolinite (Clérac, France)	kaolinite	trace of quartz and illite
smectite(ABM)	smectite	≈ 2% illite, ≈ 2% quartz, trace of pyrite and microcline

Table 12 : Mineralogical of the standards used from chemical and X-ray diffraction data.

Particle size distributions of standard clay minerals were measured by laser granulometry using a VASCO-2 Laser grain size analyzer from Cordouan Technologies, with a range of particle size from 2 nm to 6 µm. Powders were dispersed in distilled water and the particles were disaggregated by ultrasonication. Ten repeated measurements were undertaken for each sample. The median diameter of the particles (d_{50}) was around 450 nm for illite Velay and kaolinite Clérac and around 475 nm for smectite ABM. These equivalent particle sizes allow minimizing any spectral predominance of one component over another due to particle size difference.

Twenty-one surface soil samples coming from sites across central basin of Midi-Pyrénées (figure 1) are also used. Samples were collected in June 2007 and 2008 during periods of drought with low soil moisture contents. Each site was precisely located using a GPS device. Soil data include soil from molasses, terraces and alluvial plain, and limestone plateau.

III.2.4.2 XRD measurements

XRD powder patterns of bulk samples were measured with a CPS120 curved detectors diffractometer equipped with a Co source and a curved detector to attain the mineralogical composition of the soils. XRD measurements (natural, after solvation with ethylene glycol, and after heating to 490°C) of oriented deposits of the clay fraction (less than 2 µm) were obtained with a diffractometer (goniometer) theta/2theta INEL G3000 (source Cu) equipped with a SiLi diode. The clay fraction was extracted by settling from suspension.

III.2.4.3 Estimation of the soil mineralogy

Respective amounts of clay minerals were obtained by crossing X-ray diffraction and chemical analysis (major-elements) data (figure 1, table 2). The observed soil samples do not have a wide range of variability in their mineralogy. Quartz is present in most of bulk soil samples and its content varies between 10 and 70 %. Among the considered clay minerals, illite is the most abundant and its content varies between 10 and 40 %. Kaolinite content is generally lower than 15 % and smectite is present in few soil samples with content generally low. The values of illite, kaolinite and smectite in table 2 are the content of these minerals in bulk soil. However, assuming that these minerals constitute 100% of the soil clay content makes it possible to recalculate the percentage of each of these clay minerals in the clay fraction (less than 2µm) generally considered as being the clay fraction of soil. This weak variability of the clay fraction in soil makes the differentiation difficult with a lake in extreme representative type of soil.

soil name	quartz	illite	kaolinite	smectite	calcite/dolomite	goethite/ oxyde de fe	feldspath	total	argile
29237	27	38	9	13	9	4	0	*100*	60
46008	69	11	13	4	0	3	0	*100*	28
46002	21	26	8	22	19	4	0	*100*	56
HG14	32	40	9	8	0	3	8	*100*	57
HG19	20	39	13	14	1	6	9	*102*	66
GE94	65	19	2	8	0	3	3	*100*	29
TA76	53	30	5	2	0	4	6	*100*	37
TA72	47	29	12	7	1	3	1	*100*	48
TG89	61	24	9	0	0	2	4	*100*	33
TG63	42	27	12	8	0	4	7	*100*	47
LOT10	12	21	5	6	50	2	4	*100*	32
LOT40	51	16	25	0	4	4	0	*100*	41
LOT52	24	28	10	12	19	5	2	*100*	50
GE98	66	17	6	5	1	4	0	*99*	28
TG64	68	25	4	0	0,4	1,6	1	*100*	29
GE99	8	34	2	16	10	6	24	*100*	52
HG10	55	21	3	7	0	3	11	*100*	31
LOT13	37	26	17	8	6	6	0	*100*	51
LOT33	30	30	20	0	8	4	8	*100*	50
HG22	40	30	10	5	6	3	6	*100*	45
HG13	58	16	5	8	0	4	9	*100*	29

Tableau 13: Characterization of soil samples from chemical and X-ray diffraction analysis.

III.2.4.4 Infrared laboratory and field measurements

A Nicolet 5700 FTIR spectrometer, having a spectral resolution of 4 cm^{-1} in the MIR and the NIR domains, was used. In the MIR range (4000 to 400 cm^{-1}) this spectrometer operates in transmission mode. It is equipped with a KBr beam splitter and a DTGS detector. Samples were ground in order to obtain a fine powder with a constant roughness, then diluted in a KBr matrix with a 0.5 mg sample 80 mg of KBr ratio, and pressed during 10 min at 10 tons to obtain pellets of 0.5 cm diameter. In the NIR range (9000 to 4000 cm^{-1}), the spectrometer was used on diffuse reflection mode (DRIFT) with a Smart Diffuse reflectance accessory. The spectrometer was equipped with a CaF_2 beam splitter and an InGaAs detector. Spectra are measured relative to a slick gold surface in a H_2O and CO_2 purged environment without any dilution in the KBr. The NIR and MIR spectra of the three standard clay minerals are given in figure 2.

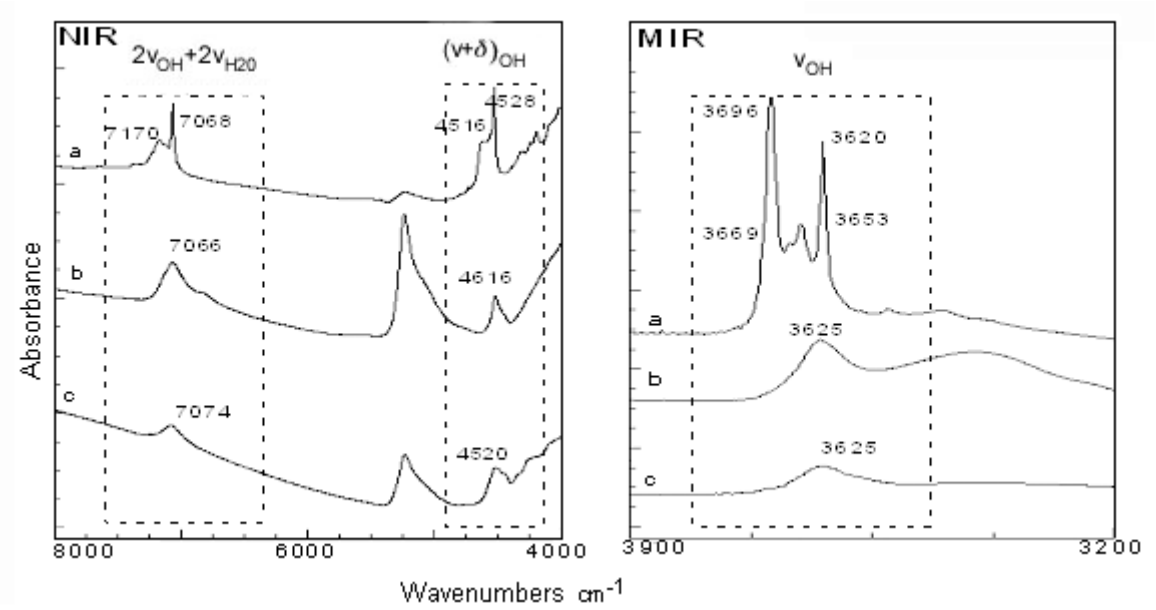


Figure III-16: NIR and MIR spectra of (a) kaolinite from Clérac (France), (b) smectite from ABM (France) (c) illite from The Puy en Velay (France), used as standards of minerals for synthetic mixtures.

An ASD pro FR Portable Spectro-radiometer having a spectral resolution of 10 nm and a 2 nm sampling -2500 interval in the 350nm (28600 to 4000 cm^{-1}) domain was also used. A white Labsphere's Spectralon® panel provided the absolute reflectance factor both for field and laboratory measurements. In laboratory, it uses a lamp as a light source which allows the angle of incidence to be accurately adjusted while removing the atmospheric disturbances. Powder samples were analysed without any dilution in the KBr matrix after grinding and drying to minimize the effect of variations in soil moisture and roughness on reflectance. For field measurements, reflectance spectra were recorded from sun light with an optical fiber. The ASD spectrometer operates under natural conditions in the NIR range with the same resolution as in laboratory. A comparison of spectra recorded in the 4000-8000 cm^{-1} range with the Nicolet 5700 and ASD spectrometers are given in figure 3. To minimize external factors, such as atmospheric effects and solar illumination, the average of 10 spectra is recorded at each measurement. Many spectra (around 35%) obtain under bad conditions of analysis (wind perturbation, light conditions …) were disregarded due to their poor quality.

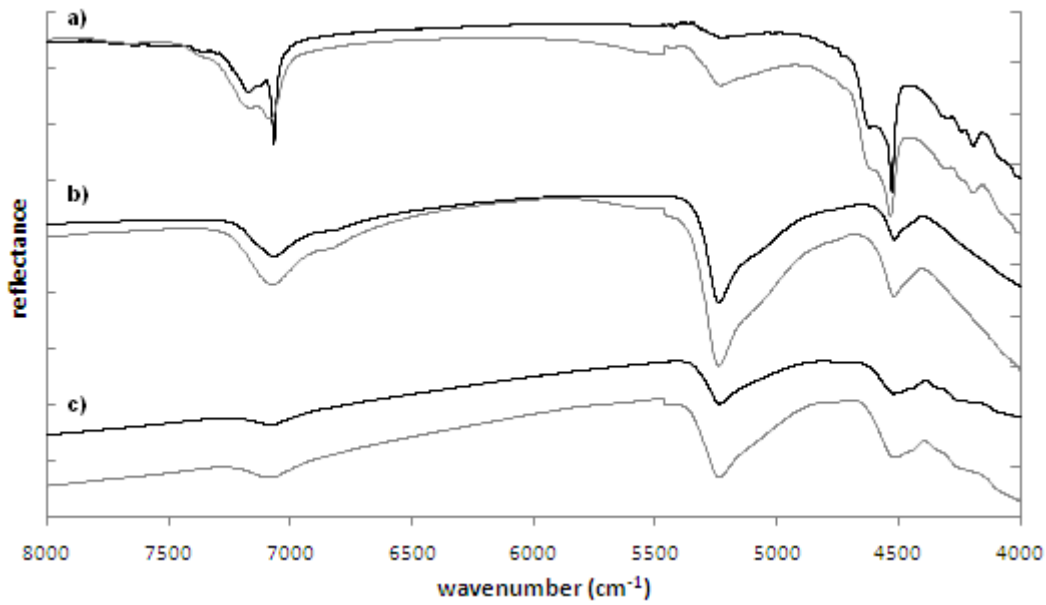


Figure III-17: NIR spectra of a) kaolinite from Clérac (France), b) smectite from ABM (France) and c) illite from le Puy en Velay (France), recorded with FTIR spectrometer on diffuse reflection mode (black) and with ASD spectrometer (grey) under laboratory condition.

III.2.4.5 Spectral data analysis

Among the different types of phyllosilicates vibrations, the structural OH group vibrations are the most relevant for this study because they can be observed in both the NIR and MIR ranges. In the MIR, the OH stretching vibration bands (νOH) are located in the 3700 to 3500 cm^{-1} range. In the NIR, the combination of OH absorption band [($\nu + \delta$) OH] is at around 4520 cm^{-1} in aluminous phyllosilicates. OH combination bands are effectively observed for kaolinite, illite and smectite used as standard (figure 2). This band appears in a "window" in the atmospheric spectrum. The integrated intensity of the νOH vibration bands, noted AOH in the MIR and of the A (ν+δOH in the NIR reflect the amount of phyllosilicates and can be used for their quantification. AOH and A(ν+δOH were measured in arbitrary unit (a.u.), by integrating the depth of the band after baseline removal for the spectra recorded with the Nicolet 5700 spectrometer and after continuum removal for the spectra recorded with the ASD spectrometer. The continuum is a convex hull fit over the top of a spectrum using straight-line segments that connect local spectra maxima. This procedure is illustrated in figure 4.

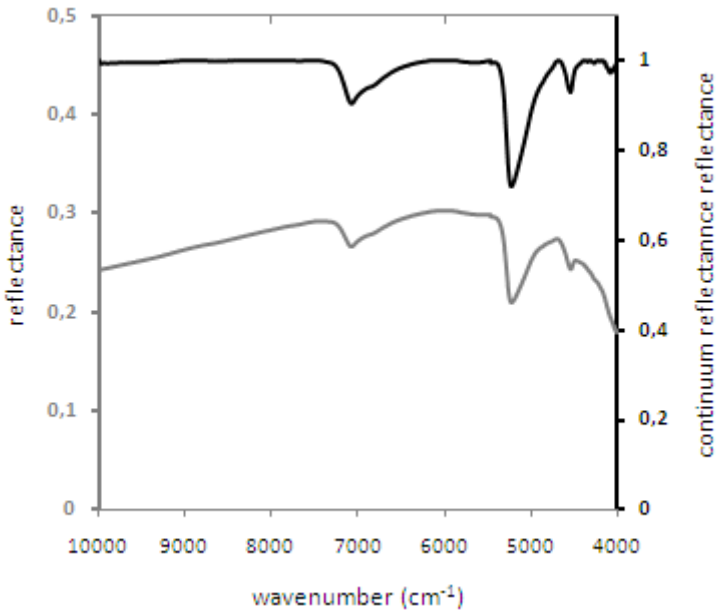


Figure III-18: continuum removal procedure. a) The standard illite spectrum (grey). b) the continuum removed spectrum (black).

III.2.5 Results

III.2.5.1 Reference and soil sample composition

The results of the quantification from major elements analysis of clay mineral, associated with the mineral identification by XRD analyses are presented in table III-3. These results are used to obtain the reference mineral purity or to identify other mineral phases.

	Main phase	secondary phases
Illite USA	95% illite	5% quartz
Illite Velay	100% illite	
Kaolinite Clérac	99% kaolinite	Trace of quartz and illite
Smectite ABM	95% smectite	2% illite, 2% quartz, trace of pyrite and microcline
Montmorillonite Wyoming	99%smectite	Traces of quartz, goethite and plagioclase
Bentonite CeCa	95% Interstratified illite/smectite	2% quartz , trace of pyrite and plagioclase
"illite-kaolinite"	85% illite	15% kaolinite, trace of interstratified illite/smectite and chlorite

Tableau 14: Results for quantification of the different phases present in the reference mineral from chemical and X-ray analysis. Note the significant portion of kaolinite (15%) of the last reference sample.

Results of XRD associated with chemical quantitative analysis for soil samples are summarized in table III-4. The observed soil samples do not have a wide range of variability in their mineralogy. Quartz is present in most of bulk soil samples and its content varies between 20 and 70 %. Among the considered clay minerals, illite is the most abundant and its content varies between 10 and 30 %. Kaolinite content is generally lower than 15 % and smectite is present in few soil samples with content generally low. The values of illite, kaolinite and smectite in table III-4 are the content of these minerals in bulk soil. However, assuming that these minerals constitute 100% of the soil clay content makes it possible to recalculate the percentage of each of these clay minerals in the fine earth fraction (less than 2µm) generally considered as being the clay fraction of soil. This weak variability of the clay fraction in soil makes the differentiation difficult with a lake in extreme representative type of soil.

	C org	*M.O.*	quartz	illite	kaolinite	smectite/ vermiculite	calcite/ dolomite	goethite/ oxyde de fer	rutile	feldspath
29237	*0,62*	*1,07*	27	38	9	13	9	4	0,51	0
46008	*0,26*	*0,45*	69	11	13	4	0	3	0,56	0
46002	*0,67*	*1,15*	21	26	8	22	19	4	0,59	0
HG14	*1,13*	*1,94*	32	40	9	8	0	3	0,63	8
HG19	*1,02*	*1,75*	20	39	13	14	1	6	0,74	9
GE94	*0,74*	*1,27*	64	19,5	2	8	0	3	0,98	3
TA76	*1,23*	*2,12*	53	30	5	2	0	4	0,63	6
TA72	*0,69*	*1,19*	47	29	12	7	1	3	0,63	1
TG89	*0,71*	*1,22*	60	24	9	0	0	2,5	0,86	4
TG63	*1,06*	*1,82*	41	27	12	8	0	4	0,79	7
LOT10	*1,51*	*2,6*	11	21	5	6	50	2	0,35	4
LOT40	*1,06*	*1,82*	51	16	25	0	4	4	0,57	0
LOT52	*1,37*	*2,36*	24	28	10	12	19	5	0,58	2
GE98	*0,65*	*1,12*	66	17	6	5	1	4	0,82	0
TG64	*0,94*	*1,62*	68	25	4	0	0,4	1,6	0,6	1
GE99	*1,36*	*2,34*	8	34	2	16	10	6	0,69	24
HG10	*1,23*	*2,12*	55	21	3	7	0	3	0,8	11
LOT13	*0,8*	*1,38*	37	26	17	8	6	6	0,8	0
LOT33	*0,55*	*0,95*	30	30	20	0	8	4	0,65	8
HG22	*0,62*	*1,07*	40	30	10	5	6	3	0,64	6
HG13	*0,89*	*1,53*	58	16	5	8	0	4	0,94	9

Tableau 15: Results for quantification of the main phases present in soil samples from chemical and X-ray analysis.

III.2.5.2 Synthetic mixtures in the MIR range with FTIR spectrometer

From literature, the νOH absorption coefficient of kaolinite is about five times higher than that of smectite (Joussein et al., 2001; Madejova et al., 2002). In the MIR domain, this observation implies that the MIR spectroscopy is very sensitive for the detection of kaolinite. This is in agreement with the data obtained

for smectite/ kaolinite mixtures (figure III-18). It is also in accordance with the data obtained for illite/ kaolinite mixtures and for tertiary mixtures. For each mixture, the data set has a similar exponential behavior with some very good correlation coefficients (R^2 > 0.86), introduced on figure III-18. In spite of the measurement uncertainty related to this approach, which consists to consider OH groups as a whole without taking in to account the different vibrations band included in the spectrum area, it is possible to use these relations for having accurate kaolinite quantification in our mixtures. The dispersion of measures associated with the increase of the kaolinite content in mixtures can be probably explained by particle size of kaolinite and inhomogeneous dispersion of sample during preparation of KBr pellets, which are at the origin of scattering effect of the infrared beam.

Illite and smectite, which are both 2:1 phyllosilicates, have very similar chemical structures. Their absorption coefficients are very close and they cannot be directly distinguished through using surface of the *ν*OH absorption band (AOH).

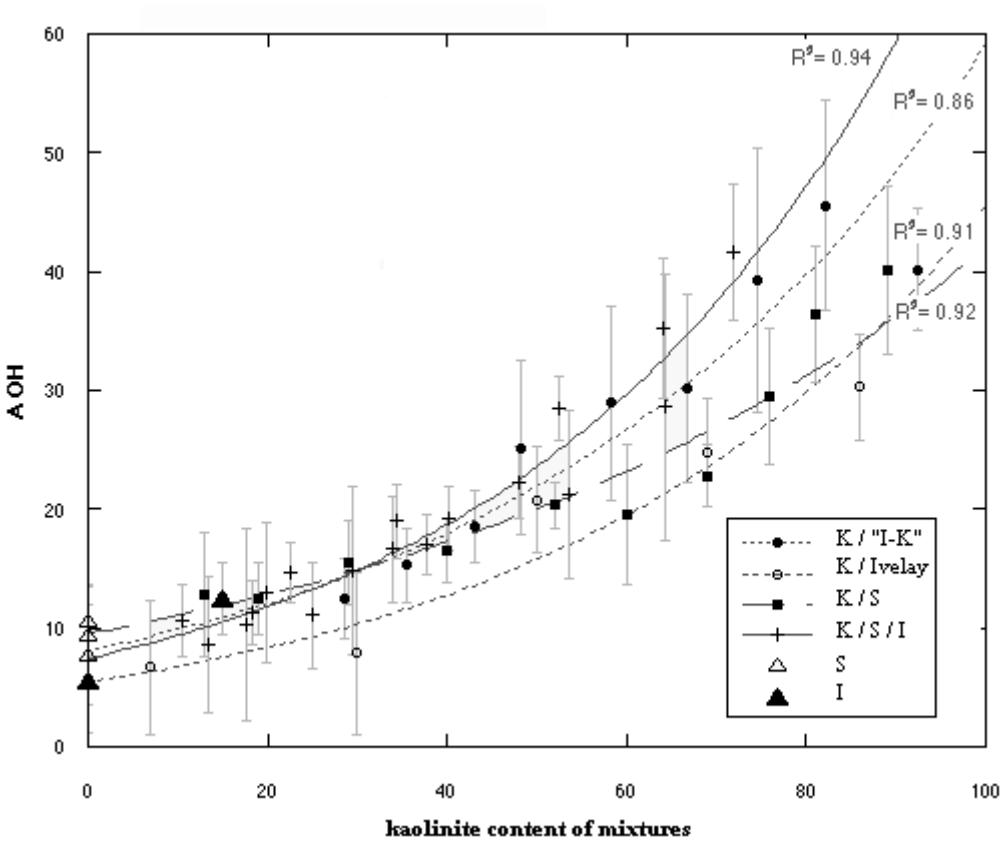


Figure III-19: Average surface of the νOH bands (a.u.) estimated from mixture spectra recorded in the MIR domain (transmission mode with FTIR spectrometer) versus kaolinite content (%) of these mixtures.K,I and S indicate kaolinite, illite and smectite. "I-K" indicates the illite which contain 15% of kaolinite. Error bars (expressed in %) are used on graph to indicate the surface variability. Its corresponds to the standart error from 10 spectra recorded in different dates. Best fits correspond in exponential relation. Coefficients of determination (R^2) were range from 0.86 to 0.94.

III.2.5.3 Synthetic mixtures in the NIR range with FTIR spectrometer

Figure III-19 illustrates the dependence of the kaolinite content of double and triple amount mixture systems versus the total absorption surface in the 4700 to 4400 cm^{-1} (2130 to 2270 nm) range. It shows that quantification of kaolinite in a mixture is also possible using the combination bands of OH in the NIR domain contrary to quantification of illite and smectite. The presented relations are more sensitive to the nature of minerals mixed with the kaolinite than in the MIR domain. Nevertheless, an estimation of the kaolinite content is possible and the quantification and value error is lower than 5%.

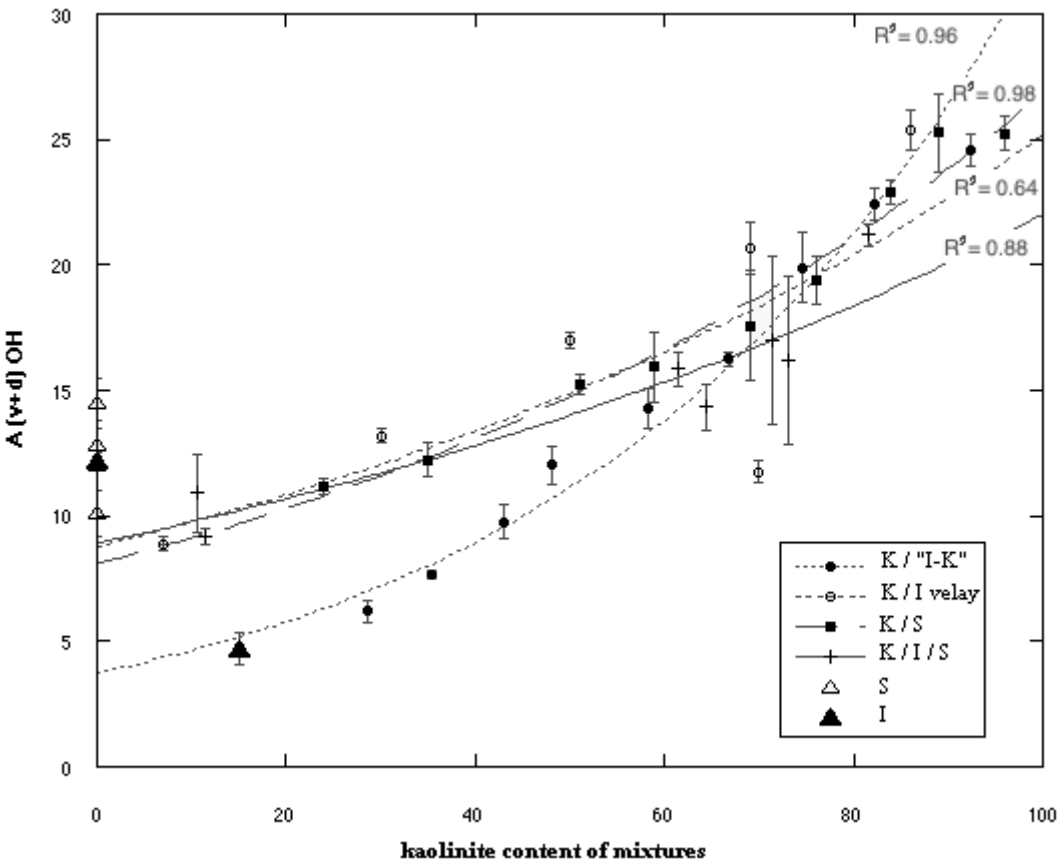


Figure III-20: Average surface of the OH bands (a.u.) estimated from various mixture spectra in the NIR (diffuse reflectance mode with FTIR spectrometer) versus kaolinite content (%) of these mixtures. K,I and S indicate kaolinite, illite and smectite. "I-K" indicates the illite which contain 15% of kaolinite. Error bars are used on graph to indicate the surface variability. Its corresponds to the standart error from 10 spectra recorded in different dates. Best fits correspond in exponential relation. Coefficients of determination (R^2) were range from 0.64 to 0.98.

III.2.5.4 Synthetic mixtures with ASD spectrometer

In order to prepare the *in situ* analyses, the spectra of double and triple amount mixture systems are recorded with the field spectrometer used under laboratory condition. The spatial and spectral resolutions decrease in comparison with those of spectra recorded with the FTIR spectrometer, but the moisture and the

roughness of the sample is the same. Although the OH bands are broader, the data obtained for synthetic mixtures can be described by exponential relations as illustrated in figure III-20. The influence of the kaolinite content in clay mixtures is also obvious at this scale. In figure III-20, kaolinite / illite velay mixtures were not able to be measured because of the lack of illite Velay (several grams are required for measurement). Error bars are not indicating because all the spectra used for this analysis were recorded the same day. The data presented in figure III-20 are calculated on the average spectrum from 10 spectra to take into account the variability lie with the spectrometer.

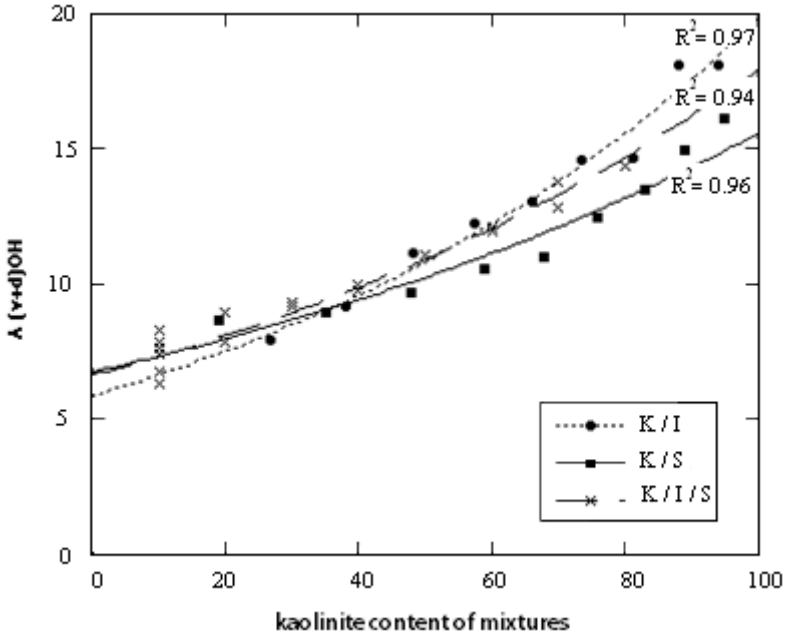


Figure III-21: Surface of the combination band (a.u.) estimated from mixture spectra in the NIR (ASD field spectrometer) versus the kaolinite content (%) of these mixtures. K,I and S indicate kaolinite, illite and smectite. Each data is measured on mean spectra from 10 spectra recorded in the same date.The variability recorded for 10 spectra is weak due to the stability of the measurement in laboratory.

III.2.5.5 Soils under laboratory conditions with ASD spectrometer

In the NIR domain, the surface of the OH band around 4520 cm^{-1} depends on the kaolinite content in the synthetic mixtures which are very close to 100% clay (Figure III-20). Compared to the laboratory experiments conducted, the range of clay content for the field plots was significantly less and the kaolinite content in soils is range between 0 and 25% (table 2). In order to highlight the clay fraction, the influence of the kaolinite content on the OH combination band is first tested from the spectra of fine-earth fraction (less than 2 µm) of soil. For this analysis, the FTIR spectrometer in DRIFT mode is appropriate due to the low amount of soil collected after settling. The influence of kaolinite on the OH combination band calculated from the bulk soil spectra was highlighted and compared with the influence of clay content (without distinction between kaolinite, illite and smectite) on the OH combination band calculated from the bulk soil spectra (Lagacherie et al., 2008).

Figure III-21a shows the results for the detection of kaolinite tested on fine-earth fraction of soil and directly on the bulk soil. In the fine-earth fraction (less than 2 µm), the variation of the OH band surface was explained by the content of kaolinite (solid line; $R^2 = 0.82$). In the bulk soil, due to the complexity of the absorption band and the weak content in kaolinite in soils of Midi-Pyrénées, the kaolinite content does not influence the surface of the OH bands (dotted line; $R^2 = 0.47$). However, this non-attendance of kaolinite content influence on the OH band surface measured on bulk soil spectra can be explained by the sampling process. The soil samples are composed of less than twenty percents of kaolinite, so it is not possible to verify the relation establish on fine earth fraction soil spectra. As a comparison, Figure III-21b shows the results for the detection of bulk soil clay content. In spite of the weak clay content of soil samples (< 40%), our results agree with those of Lagacherie et al., 2008 ($R^2 = 0.66$).

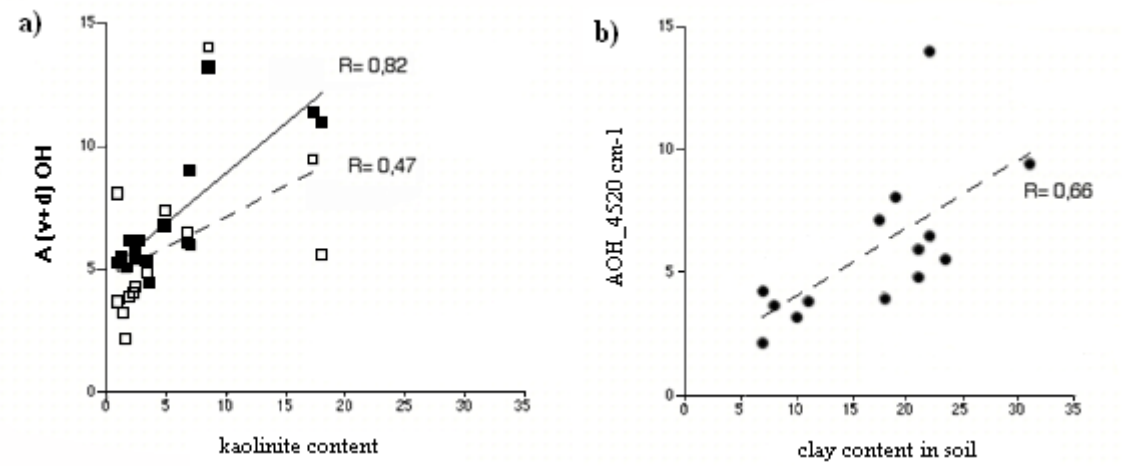


Figure III-22: Comparison of the surface of the combination band around 4520 cm^{-1} versus a) kaolinite content in fine-earth fractions (black square, $R^2 = 0.82$) and bulk soil (white square, $R^2 = 0.47$) and, b) the global clay content (without disctinction between illite, smectite and kaolinite) in soil ($R^2 = 0.66$). The kaolinite content is calculed combining XRD analyses of oriented samples and chemical analysis(Table 2). The clay content in soil is provided by XRD analysis of powder. The results for kaolinite quantification are little significant due to the sampling poverty and are given for information.

III.2.5.6 Soils analyses under natural conditions with ASD spectrometer

The bulk soil samples were analyzed under both laboratory and natural conditions with the ASD field spectrometer. These two sets of data were compared for control the influence of scale measurement. However, some spectra have been recorded only under laboratory condition. On field, measurements, which are sensitive to wind perturbation, clouds and solar illumination, were suppress due to the poor quality of their spectra. Thus, there is generally, more “under laboratory condition” point than “under field condition” point. In figure III-22, the surface of the OH band is related to the total clay content in soil. The results obtained directly in the field, were the same as those obtained under laboratory conditions. Thus, surface band

measurements were robust in relation to change of scale but the lower value of the coefficient of correlation (0.69) reflects the limits of this method for the analysis on bulk soil.

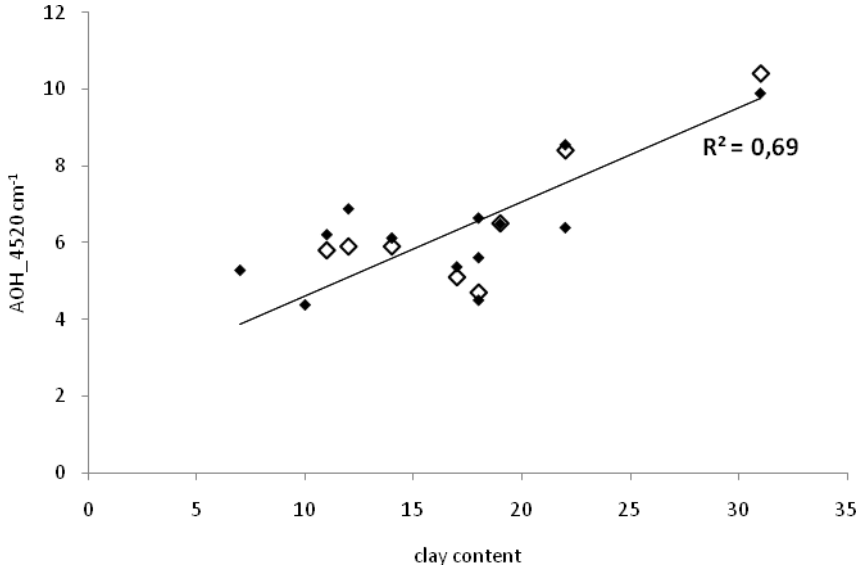


Figure III-23: Surface of the combination band (AU) around 4520 cm-1,measured on soil spectra recorded with ASD field spectrometer under laboratory condition (black diamond) and under field condition (diamond), versus clay content (%). Black line is the linear regression for in-situ measurement. It is characterize by the coefficient of linear correlation R^2.

III.2.6 Discussion and conclusions

In previous studies, the surface of the [(υ+δ) OH] absorption band (integrated intensity) around 4520 cm^{-1} was used to obtain the clay content in soil (Lagacherie et al., 2008). It works well because this band is present in every aluminous phyllosilicates (e.g. Madejová, 2003). However, this approach is correct provided that all the phyllosilicates contained in clay influence in the same way this absorption band. In the present study, the surface of the absorption band in the MIR and the NIR domain increases exponentially with the content of kaolinite in a mixture of clay minerals. Indeed, we show that the exponential relation established in the MIR domain for kaolinite _ smectite mixtures (e.g. Joussein et al., 2001) is validated for all studied mixtures and is transposable in the NIR domain. This is due to the υOH absorption coefficient of kaolinite which is higher than that of smectite and illite. This makes possible a quantitative estimation of kaolinite content in synthetic mixtures with FTIR spectrometer and ASD field spectrometer.

The method is not adapted for discriminating between illite and smectite. It can be explained by the complexity of the OH band in the MIR domain and the OH combination band in the NIR domain. This band is, in fact, composed of several overlapping bands corresponding to several absorption band from clay minerals and from other minerals which may have bands positioned at the same wavenumber (like muscovite, palygorskite, alunite, gypsum).

Some studies are dedicated to the relation between the swelling behavior of soil and near infrared spectra. This approach is similar to ours. However, the swelling soil behavior is a result of geotechnical behavior and not of any particular mineralogical composition. On the other hand, IR spectra result in an accurate mineralogical composition. Thus, it seems more relevant to correlate data from IR spectra with mineralogical composition than swelling behavior.

The advantage of the spectral measurements acquired in laboratory is that the chemical composition of minerals and the modal composition of the mixture are perfectly controlled. On field, the ideal conditions are not combined. The characteristics of the observation, the number of minerals, their chemical composition and the interactions which result from it can vary from a spectrum to the other one. For example, an increase of the moisture and of the organic matter causes a general decrease in reflectance values in the NIR spectrum (e.g. Baumgardner et al., 1986; Demattê et al., 2004; Kariuki et al., 2004). Because of the erratic behavior of soils, it is difficult to extract information on the particle or aggregate size from the recorded spectra. Although the use of a continuum limits the impact of these parameters insulating absorption bands, their influence on the NIR spectrum impacts the surface of the OH absorption bands. This lowers the accuracy of the method used in the present study for in-situ analyses. This is the reason why kaolinite content has a stronger impact on the surface of the OH bands in laboratory than with on field measurements.

As a conclusion, the use of IR spectroscopy is tested at different scales and under different measurement conditions. This method yields fast and inexpensive informations about the clay composition of the fine-earth soil fraction. This method also allows the quantification of the kaolinite, so showing the non-linear contribution of the various constituents of a mixture in absorption bands. It provides spatial information about the soil clay composition by a method based on the Beer Lambert law. At regional scale, NIR field spectroscopy provides localised check points for calibration of multispectral or hyper-spectral data. Meanwhile, in order to extend this study, the use of wavelet analysis is tested with the aim to discriminate between the content of illite and smectite in soils (e.g. Ge et al., 2007).

III.2.7 Acknowledgements

This research has been conducted within the multi -thematic project about the development of infrared spectroscopy method for swelling soils the characterization, funded by the Midi-Pyrénées region, in southern France. It is a joint project between the Midi-Pyrénées region, the “Bureau de Recherches Geologiques et Minières” (BRGM; National office for geologic and Minning) and the “centre national de la recherche scientifique” (CNRS). Some Funding come from a European project (SUDOE) about DO SMS.

The authors would like also to thank Michel Thibaut from LMTG, Toulouse and Claude Fontaine from HydrASA, Poitiers for a part of the XRD analyses. They thank Jean-Claude Soula for its constructive advices.

III.2.8 References

Baumgardner, M. F., Silva,L.F., Biehl, L.L., Stoner, E.R., 1986. Reflectance Properties of Soils. Advances in Agronomy, Academic Press. 38, 1- 44.

Ben-Dor, E., Inbar Y., and Chen, Y., 1997. the reflectance spectra of organic matter in the visible near-infrard and short wave infrared spectra (400-2500 nm) during a controlled decomposition process. Remote Sensing of Environment. 61, 1-15.

Ben-Dor, E., Kindel, B., Goetz, A.F.H., 2004. Quality assessment of several methods to recover surface reflectance using synthetic imaging spectroscopy data. Remote Sensing of Environment. 90, 389 - 404.

Bishop, J.L., Pieters, C.M., Edwards, J.O., 1994. Infrared spectroscopic analyses on the nature of water in montmorillonite. Clays Clay Miner. 42 (6), 702- 716.

Brown, D.J., 2007. Using a global VNIR soil- spectral library for local soil characterization and landscape modeling in a 2nd-order Uganda watershed. Geoderma 140, 444- 453.

Brown, D.J., Sherpherd, K.D., Walsh, M.G., Dewayne Mays, M., Reinsch, T.G., 2006. Global soil characterization with VNIR diffuse reflectance spectroscopy. Geoderma. 132, 273- 290.

Clark, R.N., Roush, T.L., 1984. Reflectance Spectroscopy: Quantitative Analysis Techniques for Remote Sensing Applications. J. Geophys. Res. 89 (NB7), 6329- 6340.

Clark, R.N., King, T.V.V., Klejwa, M., Swayze, G.A., Vergo, N., 1990. High spectral resolution reflectance spectroscopy of minerals. J. Geophys. Res., -Solid Earth Planets 95 (B8), 12653-12680.

Demattê, J.A.M., Campos, R.C., Alves,M.C., Fiorio, P.R., Nanni, M.R., 2004."Visible-NIR reflectance: a new approach on soil evaluation. Geoderma 121 (1-2), 95-112.

Ge, T., Morgan, C.L.S., Thomasson, J.A., Waiser, T., 2007. A new perspective to near-infrared spectroscopy: a wavelet Approach. ASABE 50 (1), 303-311.

Goetz, A.F.H., Chabrillat, S., Lu, Z.,2001. Field reflectance spectrometry for detection of swelling clays at construction sites. Field Analytical Chemestry and Technology **5** (3), 143 - 155.

Janik, L.J., Skjemstad, J.O., Merry, R.H.,1998. Can mid infrared diffuse reflectance analysis replace soil extractions? Aust. J. of Exp. Agric. 38 (7), 681-696.

Joussein, E., Petit, S., Decarreau, A., 2001. A new method to determine the ratio of clay minerals in mixtures by IR spectroscopy. C.R. Acad. Paris / Earth and Planetary Sciences **332**, 83 - 89.

Kariuki, P.C., Van der Meer, F., Siderius, W., 2003. Classification of soils based on engineering indices and spectral data. International Journal of Remote Sensing 24, 2567-2574.

Kariuki, P.C., Woldai, T., Van der Meer, F., 2004. The role of remote sensing in mapping swelling soils. Asian Journal of Geoinformatics 5 (1), 1- 11.

Lagacherie, P., Baret, F, Feret, J.B., Netto, J.M., Robbez-Masson, J.M., 2008. Estimation of soil clay and calcium carbonate using laboratory, field and airborne hyperspectral measurements. Remote Sensing of Environment 112 (3) 825- 835.

Madejová, J., 2003. FTIR techniques in clay mineral studies. Vibrational Spectroscopy 31, 1 - 10.

Madejová, J., Keckés, J., Palkova, H., Komadel, P., 2002. Identification of components in smectite / kaolinite mixtures. Clay Minerals 37, 377- 388.

Madejová, J., Komadel, P., 2001. Baseline studies of the clay minerals society source clays: infrared methods. Clays Clay Miner. 49 (5), 410 - 432.

Marcoen, J.M., Tessier, D., Thorez, J., Monjoie, A., Schroeder, Ch., 2001. manuel relatif aux matières naturelles pour barrières argileuses ouvragées pour C.E.T et réhabilitaion de dépôtoirs en Région Wallonne. Dir. Gén. Ress. Natur. Environ., Jambes.

McBratney, A.B., Minasny B., Viscarra Rossel, R.A., 2006. Spectral soil analysis and inference systems: A powerful combination for solving the soil data crisis. Geoderma 136, 272- 278.

Moore, D.M., Reynolds, R.C., 1997. X-Ray Diffraction and the identification and analysis of Clay Minerals. Oxford University Press, 378.

Odlare, M., Svensson, K., Pell, M., 2005. Near infrared reflectance spectroscopy for assessment of spatial soil variation in an agricultural field. Geoderma 126, 193- 202.

Petit, S., Robert, J-L., Decrreau, A., Besson, G., Grauby, O. Martin, F., 1995. Apport des méthodes spectroscopiques à la caractérisation des phyllosilicates 2 : 1. [contribution of spectroscopic methods to 2 : 1 clay characterization]." Bull. Centres Rech. Explor.-Prod. elf Aquitaine 19 (1), 119- 147.

Shepherd, K.D., Walsh, M.G., 2002. Development of Reflectance Spectral Libraries for Characterization of Soil Properties. Soil Sci Soc Am J 66 (3), 988- 998.

Sorensen, L.K., Dalsgaard, S., 2005. Determination of clay and other soil properties by near infrared spectroscopy. Soil Sci. Soc. Am. 69, 159- 167.

Viscarra Rossel, R.A., mcBratney, A.B., 1998. Laboratory evaluation of a proximal sensing technique for simuyltaneous measurement of soil clay and water content. Geoderma 85**,** 19- 39.

Viscarra Rossel, R.A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J., Skjemstad, J.O., 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. Geoderma 131**,** 59- 75.

Waiser, T.H., Morgan, C.L.S., Brown, D.J., Hallmark, C.T., 2007. In situ characterization of soil clay content with near infrared diffuse reflectance spectroscopy. Soil Sci. Soc. Am. 71 (2), 389- 396.

III.3. CONCLUSION

Dans ce chapitre, nous démontrons principalement les limites de la quantification des principaux minéraux argileux (illites, smectite et kaolinite) sur la base de données issue de la spectrométrie infrarouge, dans le cas d'une étude globale des spectres (méthodes statistiques) et de l'étude des pics caractéristiques des minéraux argileux.

L'analyse spectrométrique des mélanges synthétiques et des sols a montré que l'identification des minéraux argileux est une tâche compliquée du fait de la ressemblance des spectres représentatifs des minéraux des principales familles d'argiles. En effet les spectres présentent les mêmes bandes d'absorption qui se différencient généralement par de subtiles modifications dans leur forme générale.

L'étude des différents mélanges synthétiques dans la gamme du proche infrarouge a démontrée que la teneur en kaolinite est quantifiable que ce soit par des méthodes statistiques telles que la régression par moindres carrés partiels (PLSR) ou par l'analyse de la surface de la bande d'absorption à 4500 cm^{-1} (2200 nm) produite par la vibration de la molécule Al-OH.

En revanche, l'illite et la smectite qui sont identifiables dans les spectres PIR par une absorption à 4500 cm^{-1} (2200 nm), ne sont pas dissociables l'une de l'autre. Lorsqu'elles sont majoritaires dans le mélange, elles sont dissimulées par l'absorption plus intense de la kaolinite si celle-ci est présente dans le mélange ou dans le sol. Ainsi, l'absorption produite par la vibration de la molécule Al-OH n'est pas un critère suffisamment discriminant lors de l'analyse de la surface des bandes d'absorption liées aux phyllosilicates.

L'étude statistique des différents mélanges synthétiques dans la gamme du moyen infrarouge (MIR) permet de construire des modèles plus robustes que dans le proche infrarouge (PIR) mais amène à des conclusions similaires. De la même manière, l'analyse de la surface des bandes d'absorption liées aux vibrations des groupements OH dans le MIR, permet de quantifier la teneur en kaolinite des mélanges synthétiques mais ne permet pas la discrimination de l'illite et de la smectite. Cette conclusion est en accord avec plusieurs travaux de minéralogie réalisés dans cette gamme de longueurs d'ondes.

Dans tous les cas, l'étude des sols in situ est compliquée par le fait que l'aire des absorptions s'atténue fortement en raison de la présence importante de quartz dans les sols qui masque l'expression des autres phases minérales, contrairement à l'analyse en laboratoire qui porte principalement sur la fraction argileuse inférieur à 2 µm.

Finalement, l'application de la spectrométrie infrarouge au travers des méthodes statistiques et de l'analyse de la surface des bandes d'absorption caractéristiques des minéraux argileux par les bureaux d'études n'est pas envisageable. En effet, le PLSR donne des résultats intéressants mais la création de modèles robustes et l'utilisation en routine de ce type de méthode est largement dépendante de la qualité et de la taille de la bibliothèque de calibration. Elle nécessite un important échantillonnage de sols adapté à la zone d'étude pour lesquels des analyses minéralogiques par diffraction des rayons X et/ ou par analyses

chimiques sont nécessaires. Cette connaissance *a priori* limite la fonctionnalité de l'approche dans le cadre d'une utilisation opérationnelle. L'analyse de la surface des bandes d'absorption est quant à elle trop imprécise pour l'identification et la quantification des minéraux argileux. Elle peut donc uniquement être utilisée en appui aux méthodes classiques (analyses géotechniques principalement) dans un but de spatialisation des informations ponctuelles délivrées par ce type de méthode.

En conclusion, la spectrométrie infrarouge a le potentiel de satisfaire l'augmentation de la demande d'information sur la composition minéralogique des sols. Cependant les limitations des méthodes analytiques actuelles appellent à l'évaluation d'autres approches.

A partir de la spectroscopie de laboratoire et de terrain, les méthodes de décomposition des spectres et de modélisation des figures d'absorption des différents composants ont récemment émergé et semblent prometteuses pour la quantification. Une des approches qui semblent particulièrement adaptée est l'analyse d'ondelette

.

CHAPITRE 4

Contribution de l'analyse en ondelettes couplée à la corrélation croisée à la spectrométrie proche infrarouge.

IV. CONTRIBUTION DE L'ANALYSE EN ONDELETTES COUPLEE A LA CORRELATION CROISEE A LA SPECTROMETRIE INFRAROUGE

La spectroscopie infrarouge donne un signal composé de multiples pics qui reflètent la composition minéralogique d'un sol. Les limites des précédentes méthodes appellent à l'utilisation et à l'évaluation de nouvelles méthodes analytiques. L'une des possibilités est d'analyser ce signal en termes de fréquence. En effet, à partir de données de terrain et de laboratoire, les méthodes de décomposition des spectres en fréquences et de modélisation des figures d'absorption de ces spectres ont récemment montré des résultats prometteurs pour la quantification des constituants des sols. Parmi ces méthodes, l'analyse en ondelettes apparait être particulièrement bien adaptée aux données de la spectroscopie. Ainsi, le développement d'une nouvelle approche utilisant les ondelettes est proposé pour l'identification et la quantification des différents constituants des mélanges de minéraux et des sols.

IV.1. INTRODUCTION ET ETAT DE L'ART

Certains modèles tels que le Modèle Gaussien Modifié (MGM) (Sunshine 1990) repose sur le fait que les variations du spectre dans le visible et le proche infrarouge sont composées de bandes d'absorptions assimilables à des gaussiennes (Clark 1984). Ce modèle fonctionne en superposant un ensemble de gaussiennes au continuum choisi. Le nombre de gaussiennes utilisées pour une déconvolution est conditionné par le nombre de bandes d'absorption présentes dans le spectre. Ainsi, la complexité minéralogique augmente le nombre de gaussiennes. Il existe deux approches concernant l'utilisation du MGM. La première consiste à initialiser le centre d'une gaussienne en fonction de la position du maximum de l'absorption pour une composition chimique choisie *a priori* de façon arbitraire. La seconde est développée par (Clenet 2009). Il considère que la position initiale de la gaussienne peut ne pas traduire la composition chimique réelle du minéral, elle doit être dans une certaine mesure, « libre » de se déplacer pour modéliser le spectre de la façon la plus réaliste possible autour d'une position de référence. Au sein d'une même absorption, la profondeur des différentes gaussiennes sera plus intimement liée aux proportions relatives de chacun des minéraux. Il existe donc une multitude de combinaisons possibles. La principale contrainte que cette méthode rencontre est la dépendance de tous les paramètres qui permettent de définir les bandes d'absorptions. En effet, l'utilisateur les rentre en guidant l'information et peut ainsi biaiser le résultat final. D'autre part cette approche nécessite une importe connaissance *a priori* sur la minéralogie des échantillons analysés.

Pour ma part, j'ai choisi l'approche fréquentielle qui semble nécessiter moins d'hypothèses quant à la composition de l'échantillon, et pour laquelle de récentes études en spectroscopie de laboratoire ont montré qu'elle offrait de nombreux avantages par rapport aux précédentes approches spectrales (chapitre 3). Un aperçu non exhaustif de ces études est ainsi présenté dans ce paragraphe.

De nombreuses études utilisent la transformée en ondelettes dans une étape préliminaire aux méthodes de calibration multivariables telles que la PLSR (Jouan-Rimbaud 1997; Ge 2007; Bilgili 2008). Elle permet de

mettre en avant l'information utile contenue dans les données spectrales, mais également de filtrer le bruit (généralement hautes fréquences) et les informations non utiles contenues dans les spectres (plus basses fréquences). Ainsi, en choisissant les coefficients en ondelettes appropriés, un modèle spectral peut être établi par régression sur ces coefficients vis-à-vis de la concentration chimique des éléments. Blackburn et al. (2007) montrent ainsi que les résultats obtenus en utilisant l'analyse en ondelettes pour présélectionner les paramètres à modéliser, sont comparables aux résultats de la PLRS lorsque les données de calibration et de prédictions sont similaires en concentration, en revanche l'analyse en ondelettes surpasse la PLSR lorsque les données sont très différentes ((Blackburn 2007). Ces méthodes apportent donc une amélioration certaine pour les méthodes de régression, mais ne supprime pas pour autant les limites propres aux méthodes statistiques et l'importante quantité de pré requis nécessaires pour un fonctionnement et une utilisation correcte des résultats de ces méthodes.

La robustesse des méthodes utilisant l'analyse en ondelettes est également démontrée pour la suppression de la partie du signal spectroscopique liée à la ligne de base (Mittermayr 2001), au bruit ou encore à la réflectance spéculaire pour permettre l'estimation de la concentration des composants grâce à la préservation de figures d'absorption fines autrement perturbées (Cai 2001). Ainsi, la capacité de l'analyse en ondelettes pour supprimer le bruit haute fréquence et son indifférence aux variations de la ligne de bases des données spectrales est souvent exploitée dans l'analyse quantitative de données de télédétection pour l'extraction de caractéristiques spectrales pertinentes. Cela permet de travailler directement sur les bandes d'absorption ainsi mises en évidence, mais également de développer des méthodes de classification permettant de regrouper les pixels dont les spectres présentent les mêmes bandes d'absorption (Zhang 2005; Gendrin 2006).

Quelques études utilisent l'analyse en ondelette dans le but de quantifier certaines propriétés des sols à partir de données de spectroscopie infrarouge (Ge 2006; Jahn 2006; Blackburn 2007; Ge 2007). Ainsi Jahn (2006) définit de fortes corrélations entre le volume des bandes d'absorption calculé sur les spectres de sols déconvolués par l'analyse en ondelette, et les concentrations en nitrate des sols. Cependant, leurs résultats sont limités dans le cas de sols calcaires, car le pic d'absorption caractéristique des carbonates, proche de 1450 cm^{-1}, tend à éclipser la figure d'absorption des nitrates porche de 1370 cm^{-1}. Ainsi une calibration de leur méthode en fonction du type de sols est nécessaire pour préciser la quantification des teneurs en nitrate.

Enfin, certaines études utilisent l'analyse en ondelettes dans le cadre d'un suivi temporel de certaines zones à partir de données de télédétection (Galforf 2008)

Ces études montrent ainsi la capacité des ondelettes à grouper les spectres de même compositions minéralogiques, à détecter des matériaux présents en faible concentration et à séparer les minéraux présentant des bandes d'absorption différentes dans une même gamme de longueurs d'ondes, sans nécessité de connaissance *a priori*.

IV.2. PRINCIPE GENERAL

IV.2.1 L'analyse multi-résolution

IV.2.1.1 La transformée de Fourier

En 1807 Fourier proposa le principe de la transformée du même nom, qui repose sur le fait que toute fonction périodique peut être représentée comme la somme de sinusoïdes ou d'harmoniques dont on fait varier d'une part les amplitudes en les multipliant par des coefficients, et d'autre part les phases en les décalant de manière à ce qu'elles s'additionnent ou se soustraient. L'analyse de Fourier nous enseigne ainsi qu'un signal quelconque peut s'écrire comme la somme de sinusoïdes de fréquences et d'amplitudes variables. Un signal est entièrement caractérisé pour l'ensemble des amplitudes des sinusoïdes, qui forme ce que l'on appelle sa transformée de Fourier.

La transformée de Fourier est porteuse de toutes les informations disponibles sur le signal analysé. Ainsi, si elle n'a que peu de valeurs faibles pour des valeurs élevées de fréquences, alors elle varie lentement. Inversement, si elle prend des valeurs importantes pour les hautes fréquences, le signal contient une quantité non négligeable de hautes fréquences, et donc varie rapidement, au moins dans certaines zones. C'est précisément là que nous touchons du doigt l'une des limites importantes de l'analyse de Fourier usuelle. La transformée de Fourier est incapable de localiser les portions du signal dans lesquelles les variations sont rapides, ni celles où elles sont lentes.

Ainsi, dans une transformée de Fourier, les principaux problèmes théoriques sont les suivants :

- L'absence totale de la localisation temporelle engendre une difficulté à donner une interprétation physiquement acceptable aux signaux ou phénomènes étudiés.
- La non causalité nécessite la connaissance de tout le signal pour pouvoir en calculer sa transformée et conduit à l'impossibilité d'un traitement en temps réel.

Le signal est donc étudié soit en fonction du temps, soit en fonction des fréquences qu'il contient, sans possibilité de conjuguer les deux analyses. En effet, un support éphémère en temps et un support éphémère en fréquence ne peuvent coexister : c'est le principe d'incertitude Heisenberg-Gabor.

IV.2.1.2 La transformée de Gabor

La transformée de Gabor, ou transformée de Fourier à « fenêtre glissante » ou fenêtrée, présentée par Gabor (1946), permet de localiser l'analyse autour d'un temps *t* défini à l'aide d'une fenêtre symétrique, et de décomposer un signal en une combinaison linéaire temps-fréquence.

Gabor suggérait de rendre locale l'analyse de Fourier en s'aidant de fenêtres. Une fenêtre est une fonction régulière, lentement variable, et bien localisée (elle est nulle en dehors d'une certaine zone). En multipliant la fonction étudiée par une fenêtre, on obtient une version « locale », dont on peut déterminer le

contenu fréquentiel par une analyse de Fourier classique. On renouvelle alors l'opération en déplaçant la fenêtre d'analyse (figure IV-1, a). L'ensemble de ces transformées de Fourier ainsi localisées, forme la transformée de Gabor du signal et fournit donc une analyse fréquentielle locale. Elle est définie par les équations 1 (pour l'analyse) et 2 (pour la synthèse) suivantes:

$$SFT(v,b) = \int_{-\infty}^{+\infty} f(t).g_{v,b}(t)dt = \int_{-\infty}^{+\infty} f(t).g(t-b).e^{-t2\pi vt}\,dt$$ Equation 1

$$f(t) = \iint_{R^2} SFT(v,b).G_{v,b}(t).dvdb \text{ avec } g(t) = e^{-\pi.t^2}$$ Equation 2

SFT se traduit par « Short Fourier Transform ». Elle représente la transformée de Gabor, $g_{v,b}(t)$, et joue le rôle de la réponse impulsionnelle d'un filtre sélectif à la fréquence v. La SFT est interprétée comme une batterie de filtres de même dimension, décalés en fréquence.

Cependant, dans ce type d'analyse, la résolution temporelle est fixée par le choix de la taille de la « fenêtre » qui peut prendre différentes formes (rectangulaires, de Hamming, de Hanning, de Blackman, de Barlette, etc). Elle ne permet donc pas, dans le même temps, une analyse précise en temps et en fréquences (figure IV-1a).

Afin de remédier à cette incertitude, une solution possible pour faire l'analyse temps/fréquences fût apportée par la transformation en ondelettes sur des signaux non stationnaires pour lesquels la transformée de Fourier s'avère peu efficace.

IV.2.1.3 La transformée en ondelettes

Les caractéristiques multi-résolutions de la transformée en ondelettes permettent de conserver un nombre d'oscillations constant dans une enveloppe (fenêtre) que l'on peut contracter et dilater à volonté (figure IV-1b). Elle ne nécessite donc pas de connaissances *a priori* du signal. Les ondelettes que nous utilisons par la suite correspondent à l'analyse temps/échelles. Dans l'approche temps/échelles, la largeur spatiale des fonctions *g* est inversement proportionnelle à la fréquence. **Ainsi, ces fonctions ont une taille de support proportionnelle à la résolution spatiale liée à leur fréquence, et s'adaptent parfaitement au signal analysé.**

La transformée en ondelettes agit donc comme un « *zoom mathématique* » sur le signal à analyser puisqu'en faisant varier l'échelle d'analyse du signal on peut extraire les détails présents à différentes résolutions. De cette façon, elle permet d'étudier l'information contenue dans un signal en mettant en évidence des caractéristiques invisibles au premier abord.

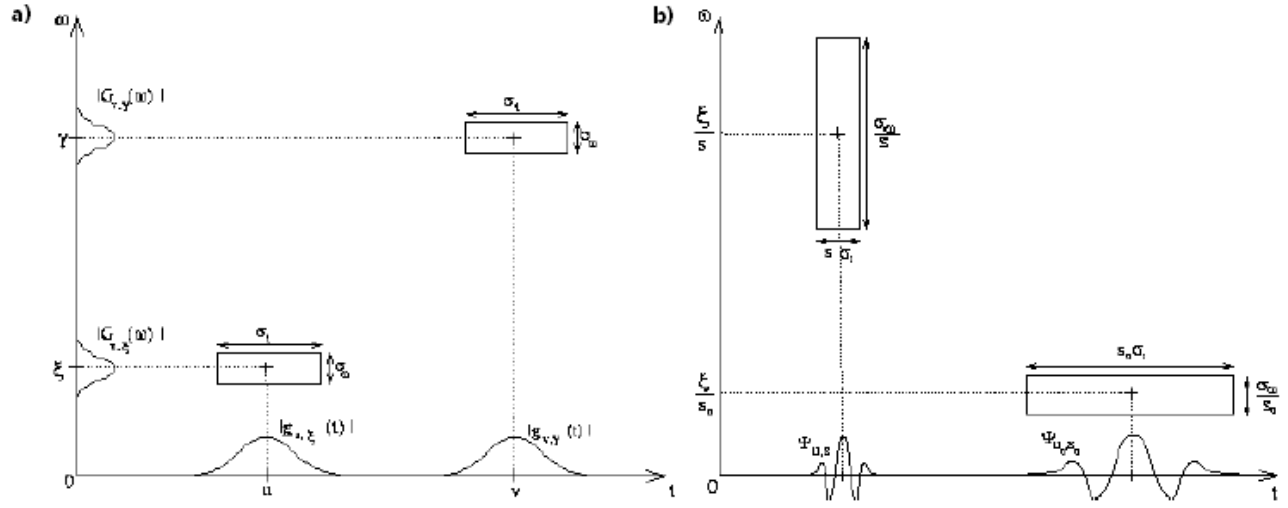


Figure IV-1: Différence entre a) les transformée de Gabor et b) les transformées en ondelette (Diou 2000). Dans les basses fréquences, la transformée en ondelettes est précise en fréquence, mais moins sur le temps. La tendance est inversée en haute fréquence. Cela traduit l'adaptation des ondelettes au signal étudié contrairement à la transformée de Fourier à fenêtre qui conserve une résolution d'analyse fixe quelle que soit la fréquence observée.

IV.2.2 La transformée en ondelettes continues

Depuis les travaux de Grossman (Grossman 1984), la transformation en ondelettes est très rapidement apparue comme un outil performant pour résoudre des problèmes relevant de différents domaines d'application (Mayer,1990), comme en sismique (Ricker 1953), en tectonique (Ouillon 1995; Ouillon 1996; Darrozes 1998), en hydrologie (Labat 2002), en climatologie (Torrence and Compo 1998; Torrence 1999), en hydraulique (Ge,2010), en statistiques (Bigot 2009), en géophysique (Gholami 2009) et en analyse d'images de fabriques minérales (Gaillot 1999).

Il s'agit d'un groupe de fonctions complexes ayant des propriétés mathématiques variables qui permettent de décomposer les données suivant les différentes fréquences qui les caractérisent (Mallat 2001). Elles offrent la possibilité d'une représentation des signaux faisant apparaître simultanément des **informations temporelles** (localisation dans le temps, durée) **et fréquentielles.** De ce fait, elles facilitent l'identification des caractéristiques physiques de la source du signal.

Il existe deux types de transformée en ondelettes :

- La transformée en ondelettes discontinue qui est une représentation compacte des données, et qui est particulièrement intéressante pour la réduction du bruit et la compression de données.

- La transformée en ondelettes continue qui est intéressante pour l'extraction et la caractérisation temps/ fréquence de figures particulières dans un signal.

La transformée en ondelettes discrète n'a pas été choisie car elle ne donne pas des coefficients à tous les temps pour toutes les échelles. Ainsi, si l'on veut estimer la régularité en un temps to, il se peut que l'on ne

dispose pas de coefficients à la date choisie pour toutes les échelles. On doit donc utiliser pour certaines échelles des coefficients correspondants à des temps voisins. Ceci pose le problème de non invariance par translation temporelle de la DWT : une faible translation du signal peut induire de fortes variations des coefficients en ondelettes. Or, nous analysons dans cette étude des pics d'absorption très proches dans le temps (illite smectite résultent de bandes superposées).

Ainsi, dans cette étude nous utilisons des méthodes d'analyse basées sur le formalisme mathématiques de la transformation en ondelettes continue. En effet, dans un premier temps le but de notre étude est la caractérisation parfaite de notre signal spectral et, dans un second temps, la quantification du contenu spectral. La technique de la transformée en ondelettes continues permet de mettre en évidence les caractéristiques d'un signal au cours du temps à une ou plusieurs échelles données. Cet outil est adapté à la description des phénomènes transitoires, intermittents et multi-échelles.

IV.2.3 Les ondelettes

La transformée en ondelettes remplace les sinusoïdes de la transformée de Fourrier par une famille de translations et dilatations d'une même fonction « mère » : l'ondelette. Cette fonction « mère », notée ψ, est régulière et bien localisée en temps. Elle satisfait la condition d'admissibilité de moyenne nulle (équation 3), centrée au voisinage de zéro :

$$\int_{-\infty}^{+\infty} \Psi(t)dt = 0 \qquad \text{Equation 3}$$

L'équation 3 montre que la fonction « mère » possède un ou plusieurs moments égaux à zéro. Cette « oscillation » est à l'origine du terme d'ondelette. A partir de cette fonction « mère », on construit par dilatation a, où a représente le facteur d'échelle, et translation b (figure IV-2), les fonctions de base, notés $\psi_{a,b}(t)$ (équation 4).

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{a}} \times \Psi\left(\frac{t-b}{a}\right) \quad avec\ a \in R^{+}, b \in R \qquad \text{Equation 4}$$

Ainsi, les paramètres de translation et de dilatation constituent les deux arguments de la transformée en ondelettes.

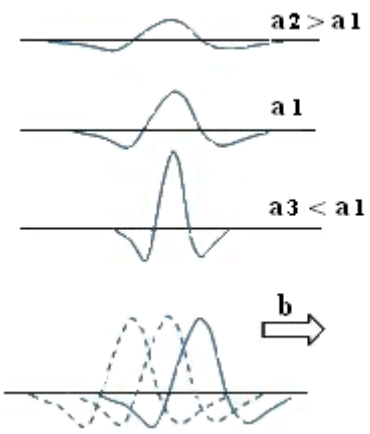


Figure IV-2: Variable d'une fonction ondelette 1D. *a*, aussi appelé facteur d'échelle, est le paramètre de dilatation lié à la fréquence tandis que *b* est le facteur de translation.

Si le centre de fréquence de la fonction mère, ψ, est υ_0, alors le centre de la fonction dilatée est en υ_0 / s (s étant l'échelle). La transformée en ondelettes a une résolution temps / fréquence qui dépend de l'échelle, *s*, de l'analyse. Par conséquent, la résolution fréquentielle absolue est d'autant plus mauvaise que l'échelle associée à une fréquence plus élevée (figure IV-3).

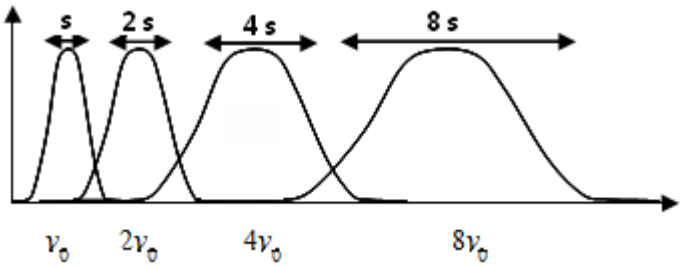


Figure IV-3: filtres associés à la transformée continue en ondelettes. La détérioration de la résolution apparait lorsque l'on va vers les basses fréquences, avec υ_0 la fréquence fondamentale de l'atome et s l'échelle.

La transformée en ondelettes $C_f(a,b)$ (équations 5 et 6) d'une fonction *f* d'énergie finie, permet donc de caractériser les périodicités des fluctuations en fonction du temps. Elle représente l'ajustement à chaque pas de temps et pour chaque périodicité entre la fonction d'ondelette et la série temporelle. Ainsi plus l'ajustement est précis, plus la transformée en ondelettes est forte.

On utilise **le spectre de puissance d'ondelettes $C_f(a,b)^2$** pour quantifier l'importance de la variabilité de la série temporelle expliquée par la transformée d'ondelettes (équation 5). Celui-ci est calculé directement à partir de la transformée en ondelettes en fonction du temps et des périodicités.

$$C_f(a,b) = \int_{-\infty}^{+\infty} f(t).\Psi^*_{a,b}(t)dt \qquad \text{Equation 5}$$

Cette transformée peut être représentée comme une convolution :

$$C_f(a,b) = \int_{-\infty}^{+\infty} f(t) \cdot \frac{1}{\sqrt{a}} \cdot \Psi^*\left(\frac{t-b}{a}\right) dt = f(b) \times \frac{1}{\sqrt{a}} \Psi^*\left(\frac{-b}{a}\right)$$

Equation 6

En décomposant une série temporelle dans l'espace temps-fréquence, on est donc en mesure de déterminer à la fois les modes dominants de variabilité et les tendances de variation de ces modes dans le temps. Cela se traduit par des valeurs de coefficient d'ondelettes fortes pour les échelles considérées alors qu'un signal constant engendrera à l'inverse un coefficient nul (figure IV-4). L'ondelette, $\psi(\upsilon)$, dans l'espace de Fourier, ayant une fréquence centrale υ_0 est une bande équivalente [υ_0-B/2,υ_0+B/2], on peut dire que celle-ci possède les caractéristiques d'un filtre passe bande (au sens large). Cependant peu d'études ont permis d'obtenir des résultats quantitatifs. Les premières études furent celles de Gendrin (Gendrin 2006) qui permet, à parti de données hyperspectrales provenant de Mars, d'estimer les teneurs en calcium sur les pyroxènes et celles de Jahn (Jahn 2006) qui montre la caractérisation des nitrates présents dans des sols.

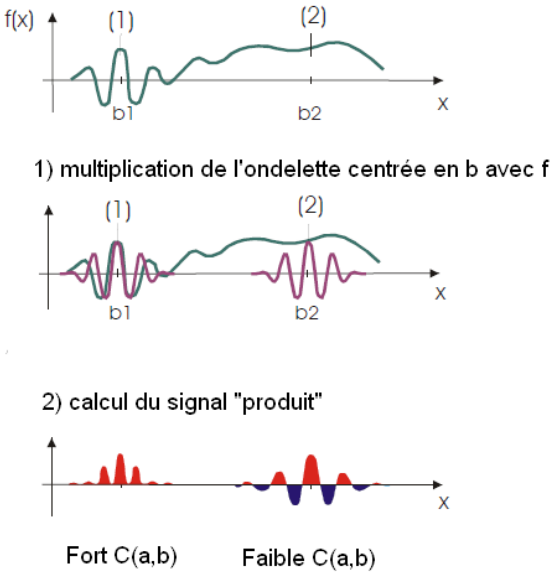


Figure IV-4: Comportement du coefficient d'ondelettes. Le coefficient d'ondelette C_f (a, b) d'un signal f(x) dépend de la forme de celui-ci au voisinage de la position b. 1) Quand f(x) est irrégulière autour d'une position (b= b1) et que ces variations sont de « fréquence » comparable à celle de l'ondelette, le produit signal-ondelette, donc C_f (a, b_1), est généralement beaucoup plus grand. 2) Quand au contraire, f(x) est à peu près constant (2) autour d'une position (b= b2), le produit de f(x) par ψ(x) est faible et se traduit par un coefficient C_f (a, b_2) très petit (Meyer 1987). Les coefficients permettent donc de localiser les zones porteuses d'informations.

L'intensité de ces coefficients est directement dépendante de l'échelle d'analyse. Plus l'échelle sera grande plus la réponse sera grande. Ceci pose un problème si l'on désire faire une analyse multi-échelle c'est-à-dire, comparé les échelles entre elles afin de ne conserver que les coefficients les plus significatifs. Très rapidement Torrence et Compo (1998) se sont rendus compte de ce problème et ont proposé une méthode de

normalisation (équation 7) qui est actuellement la règle dans l'analyse multidimensionnelle. Ainsi, chaque échelle peut être comparée aux autres.

$$\sum_{k=0}^{N-1} \left| \hat{\Psi}(s\omega_k) \right|^2 = N \qquad \text{Équation 7}$$

Avec N le nombre de points du spectre, k l'index des fréquences qui varie de 0 à N-1 et s le facteur d'échelle.

Cette normalisation (équation 7), consiste, pour chaque échelle, à calculer la fonction ondelette de telle manière à avoir une énergie maximale unité. Pour cela, elle est pondérée par les amplitudes des coefficients de Fourier. Cela assure que la transformée en ondelette à chaque échelle est directement comparable aux autres.

IV.2.4 Choix de l'ondelette

Les fonctions « mères », ψ, se divisent en deux grandes catégories :

- la première a simplement une représentation dans l'espace réel. Ces ondelettes sont généralement utilisées pour la détection des transitions brutales d'un signal (détection de signaux transitoires, analyse fractale, etc).
- La seconde a une représentation complexe. On parle aussi d'ondelettes analytiques. elles permettent de séparer l'amplitude et la phase du signal ce qui permet d'observer l'évolution temporelle des composantes fréquentielles. C'est cette seconde famille qui sera utilisée par la suite afin de définir l'évolution, non pas en temps mais en nombre d'ondes, des composantes fréquentielles.

Plus spécifiquement, l'ondelette analytique dite de Morlet est choisie. Cette fonction possède une décroissance exponentielle en temps et en distance qui lui assure une localisation optimale en temps et en fréquence. C'est une ondelette régulière, symétrique et complexe (l'information de phase est conservée).

L'ondelette de Morlet (figure IV-5) est définie par la fonction « mère » suivante (équation 8) :

$$\Psi_0(t) = \frac{1}{2}\, \pi^{-\frac{1}{4}} \exp^{i\,\omega_0 t} \exp^{-\frac{t^2}{2}} \qquad \text{Équation 8}$$

Cette ondelette n'est pas parfaitement admissible. En effet, elle ne vérifie pas complètement le critère d'admissibilité car sa transformée de Fourier ne s'annule pas à l'origine. La solution proposée par Morlet est de choisir t_o et f_o tels que $5 < w_0 t_0 < 6$, avec $w_0 = 2\pi f_0$, de manière à imposer une valeur nulle à la fréquence zéro.

De plus, l'ondelette de MORLET permet d'avoir accès à la partie imaginaire qui donne accès au calcul du déphasage. **Ce déphasage quantifie le décalage entre le pic de l'ondelette et celui du signal inconnu.** Lorsque ce décalage est trop grand il revient à 0 (cycle complet) et il ne sera donc connu qu'à 2kπ prés. Nous verrons plus loin les problèmes posés par cette incertitude. De plus, elle correspond à une fonction oscillante dont l'enveloppe est de forme gaussienne. Elle est donc adaptée à l'étude des variations des spectres

infrarouge composés de bandes d'absorption assimilables à des gaussiennes (Clark 1984). Enfin, l'ondulation sinusoïdale permet de supprimer les petits pics parasites.

Finalement, l'aspect statistique via un formalisme simple proposé par Torrence et Compo (1998) permet de calculer l'intervalle de confiance à 95% (IC_{95}) sans passer par des simulations type Monte-Carlo pour quantifier cet intervalle (voir paragraphe 1.3.2.).

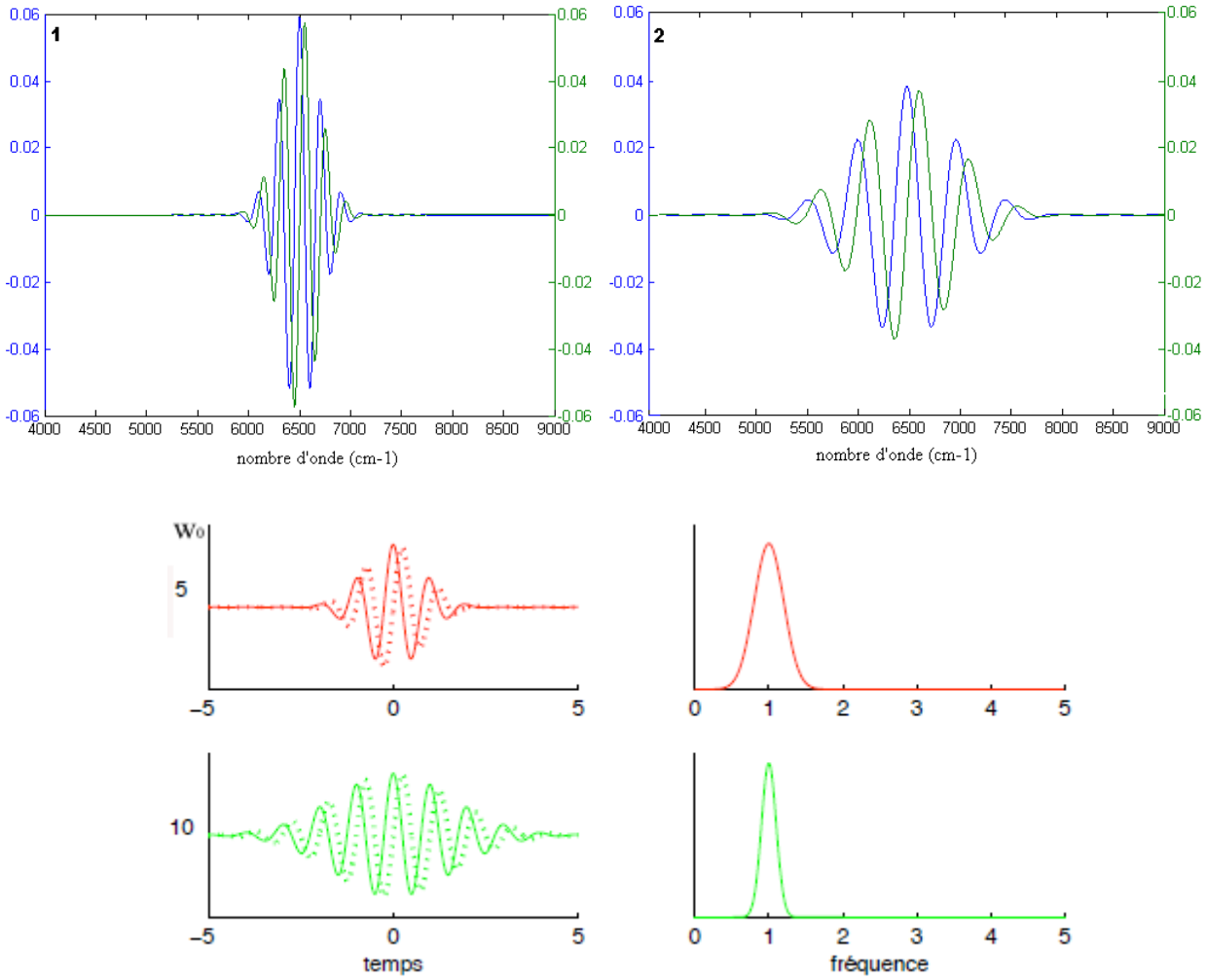


Figure IV-5: Ondelette de Morlet. La figure du haut montre l'effet de la dilatation sur la partie réelle (en bleu) et la partie imaginaire (en vert) de l'ondelette de Morlet avec $a_1 < a_2$. La figure du bas montre l'influence de la vitesse angulaire fondamentale ω0 = 5 (en rouge), et 10 (en vert) avec les réponses fréquentielles correspondantes (modifié d'après (Cnockaert Aug 2007)).

IV.3. CORRELATION

La corrélation choisie dans ce travail est **la corrélation croisée à partir d'analyse en ondelettes**. Cette corrélation croisée est utilisée pour examiner les relations qui existent dans l'espace temps fréquence entre deux séries temporelles. Elles correspondent à une représentation « espace-échelle » des interactions entre deux signaux et peut s'écrire sous la forme :

$$W2C_{xy} = CWT_{fx} . CWT_{fy}^{*} \qquad \text{Equation 9}$$

Où,

$$W2C_{xy} = \frac{1}{D}\int_{-D/2}^{+D/2} CWT_{fx}(a,b)\,CWT_{fy}(b+\tau,a)\,db \qquad \text{Équation 10}$$

Avec $W2C_{xy}$ les coefficients de la corrélation croisée, CWT_{fx} et CWT_{fy} les coefficients d'ondelettes des deux signaux à corréler f(x) et f(y) et τ le décalage et D la longueur totale du signal.

IV.3.1 Corrélation croisée

La méthode est basée sur l'interaction de deux signaux, l'un connu et l'autre à caractériser. Le terme « croisée » indique qu'un signal de référence est utilisé lors de la caractérisation. Si le signal de référence est une réplique du signal à caractériser (inconnu), on parle alors **d'auto-corrélation**. La corrélation croisée (équation 9) est ici construite à partir de deux transformées en ondelettes continues des deux signaux f(x) et f(y). Elle met en évidence leur puissance commune et leur phase relative.

Ainsi, Pour deux séries temporelles $f_x(t)$ et $f_y(t)$ données, il est possible de construire le spectre de corrélation croisée (équation 9) à partir de leurs transformées en ondelettes respectives CWT_{fx} et CWT_{fy} (CWT_{fy}^* étant le conjugué complexe de CWT_{fy}). Le spectre de corrélation croisée (équation 10) est un complexe dont les maxima locaux, exprimés sous forme de puissance, fournissent les échelles caractéristiques auxquelles des événements cohérents ont lieu. L'avantage de cette méthode est qu'elle compare les signaux à toutes les échelles et non pas simplement sur l'ensemble du spectre comme les analyses classiques. L'interprétation des cartes de résultats, aussi appelé **scalogramme** (l'axe des x correspondant aux nombres d'ondes et l'axe des y aux échelles) des coefficients est relativement aisée puisqu'elle utilise les conventions classiques :

-1 pour des signaux anti-corrélés

+1 pour les signaux corrélés

0 pour les signaux non corrélés

La principale difficulté qui découle de la corrélation croisée est rattachée à l'amplitude des signaux considérés. En effet, la corrélation croisée consiste à faire le produit scalaire entre deux signaux. Si l'amplitude d'un des signaux est plus importante que celle du deuxième signal, alors ce signal à forte amplitude va augmenter les coefficients résultants. Ce phénomène de forte variation d'amplitude est une caractéristique des spectres infrarouges car il met en résonnance les différentes liaisons OH au sein des minéraux. Selon le type de liaison, l'excitation sera plus ou moins énergétique et cela se matérialisera par des amplitudes différentes sur les spectres infrarouges (Clenet 2009). Ainsi, plus le signal à une intensité forte, plus le coefficient de la corrélation croisée sera élevé.

Dans le cas de la spectroscopie infrarouge la localisation des bandes d'absorption ne varie pas dans le temps, et la corrélation croisée est réduite au produit scalaire du spectre « inconnu » à caractériser avec les spectres de référence.

Linker (2004) utilise la technique de corrélation croisée pour comparer un spectre « inconnu » à des spectres contenus dans une bibliothèque de référence. Cependant, il ne passe pas préalablement par une décomposition en ondelette et montre ainsi que les résultats obtenus sont relativement pauvre par rapport aux résultats des approches statistiques (PLSR, Principal Component Regression). Il montre également que cela est d'autant plus vrai que les bandes d'absorptions étudiées présentent un fort taux de recouvrement (Linker 2004).

IV.3.2 Cône d'influence et intervalle de confiance

La transformée en ondelettes présente des effets de bord proportionnels à l'échelle de l'ondelette d'analyse. Ces effets sont dus au fait qu'elle n'est pas parfaitement localisée en nombre d'ondes (axe des x). C'est pourquoi il s'avère pratique de calculer le cône d'influence (COI) où ces effets de bords sont négligeables ou inexistants (Torrence and Compo 1998). Pour déterminer ce cône, on localise, pour chaque échelle, la zone où l'on a une chute de spectre de puissance d'un facteur e^{-2} qui correspond à une discontinuité situé en limite du signal. Les coefficients situé en dehors du COI ne sont pas pris en compte dans les statistiques de nos analyses.

La distribution théorique du spectre de puissance d'une corrélation croisée (équation 11) de deux signaux ayant les spectres de « fond » suivant P^x_k et P^y_k (avec k, l'index des fréquences de Fourier) est définie par Torrence et Compo (1998).

$$D\left(\frac{|W_n^X(s)W_n^Y(s)|}{\sigma_X\sigma_Y} < p\right) = \frac{Z_\nu(p)}{\nu}\sqrt{P_k^X P_k^Y} \qquad \text{Equation 11}$$

Où $Z_\nu(P)$ représente l'intervalle de confiance (IC_p) associé à une probabilité *p*.

Tous les coefficients de la corrélation croisée situés en dehors de cette ligne des *p%* de signification, n'ont pas de valeur et sont référencé comme « Not a Number » (NaN). De cette manière ils ne sont pas pris en compte dans l'étape de quantification.

IV.4. PROBLEMATIQUE

Les méthodes classiques de traitement des données de spectrométrie infrarouge présentées dans le chapitre précédent donnent des résultats intéressants pour l'étude de certaines propriétés des sols (teneur en kaolinite, teneur en argile, etc.) mais ne sont pas suffisantes pour différencier et pour quantifier des principaux minéraux argileux et plus spécifiquement l'illite et la smectite.

De plus, l'ensemble de ces méthodes utilisent une connaissance *a priori* soit de la composition des échantillons (DRX), soit de la structure chimique des minéraux (connaissance des liaisons chimiques de chaque minéraux pour connaître le nombre de vibrations responsables d'un même pic). Il est donc nécessaire

de développer des techniques plus complexes pour accéder à ces informations tout en se dégageant des contraintes imposées par les connaissances *a priori*.

Notre objectif est d'utiliser l'analyse en ondelettes de façon automatique sur des jeux de données dont nous ignorons les caractéristiques spectrales. Les figures d'absorption des principaux minéraux argileux (illite, smectite et kaolinite) se superposent et compliquent les relations entre la réflectance spectrale des mélanges ou des sols et la concentration en minéraux argileux. L'analyse en ondelettes a la capacité de capturer plus d'informations contenues dans les spectres que les approches précédemment développées, et ouvre la perspective de développer une méthode robuste pour la détermination de la teneur en smectite des sols.

Les bandes d'absorption portées par les spectres enregistrés dans le proche (PIR) et le moyen (MIR) infrarouge peuvent être considérées comme des perturbations locales d'amplitude finie rapidement atténuée et n'ayant donc que des effets de courtes longueurs d'ondes. L'intérêt de l'utilisation des ondelettes vient alors de leurs propriétés d'analyse locale et multi-échelles des signaux qui permettent l'identification et l'extraction de structures ondulatoires dominantes d'échelles différentes.

IV.5. ANALYSE EN ONDELETTES ET CORRELATION CROISEE DES SIGNAUX INFRAROUGES

IV.5.1 Méthodologie

IV.5.1.1 Approche générale

Le signal spectral est particulièrement complexe à traiter puisqu'il est affecté par de nombreux paramètres liés à la fois aux composés analysés et aux conditions d'observation. Cependant, les variations des spectres proche-infrarouges sont composées de bandes d'absorption assimilables à des gaussiennes (Clark 1984). Ainsi, l'utilisation dans un premier temps de gaussiennes permet d'étudier le comportement de l'analyse en ondelettes couplée à la corrélation croisée sur ce type de signal. La méthode développée est donc présentée et illustrée sur des exemples synthétiques avant d'être testée sur des exemples naturels. Les exemples synthétiques sont composés de gaussiennes d'écart-types, de positions et d'amplitudes variables.

Tout d'abord, les signaux (gaussiennes ou spectres infrarouges) de référence et à caractériser (appelé par la suite signal « inconnu ») sont décomposés par une transformée en ondelettes continue.

La fonction de corrélation croisée en ondelettes entre les spectres des signaux mesure ensuite la similitude (ou dissemblance) entre les signaux à chaque échelle, déterminant ainsi la nature et la composition du signal inconnu. **On obtient ainsi pour chaque échelle *a*, le déphasage moyen τ_m, pour lequel le coefficient de corrélation croisée en ondelettes C(a, τ_m) correspond à la moyenne des coefficients significatifs**. Si les composants des signaux sont ressemblants, alors le déphasage doit être faible (et même nul dans le cas de l'auto-corrélation), tandis que le coefficient de corrélation sera lui maximal, $C_{max}(a, \tau)$ et devra être proche de la valeur 1. Par la suite, ces deux paramètres (C(a, τ_m) *et* τ_m) sont représentés dans les mêmes graphiques.

La phase montre le décalage entre deux signaux. Cette phase correspond au logarithme de la partie imaginaire de l'ondelette (ou phase = atan2 (partie Imaginaire, partie Réelle)). Ainsi lorsqu'un pic est décalé, plus « étalé », ou plus dissymétrique, un décalage de phase apparaît (phase non nulle). On peut ainsi visualiser les échelles pour lesquelles nous avons une bonne ressemblance, caractérisée simultanément par un fort coefficient de corrélation moyen (voire maximal) et un faible déphasage (voire nul).

Le programme développé dans cette thèse est présenté dans l'annexe 4.

IV.5.1.2 Seuil

Les coefficients significatifs permettant de calculer le coefficient de corrélation croisée moyen C(a, τ_m), sont sélectionnés au moyen d'une valeur seuil arbitraire, mais généralement supérieur à celui fixé par IC_{95}, permettant de localiser le calcul aux zones les plus similaires. En effet, la ligne des 95% de signification, définie par Torrence et Compo (1998) classiquement utilisée dans ce type d'analyse, correspond à une mesure statistique pas suffisamment sélective pour la comparaison de signaux fortement redondants tels que les spectres infrarouges de minéraux argileux. Cependant, comme nous le verrons plus loin, il ne s'agit pas du point clef pour la quantification des argiles.

IV.5.1.3 Normalisation

Les coefficients de la corrélation croisée sélectionnés sont ensuite normalisés pour corriger les effets d'amplitude qui apparaissent lorsque les signaux corrélés des amplitudes différentes. La corrélation croisée étant le produit scalaire des deux signaux, elle est d'autant plus importante si l'un des signaux à une amplitude importante. Différentes normalisations ont été testées.

Par le spectre de référence

Par cette normalisation, les coefficients sont généralement faibles car seul le spectre de référence de plus forte amplitude donnera une réponse maximale. De plus, même s'il existe un spectre de référence identique au signal « inconnu », s'il ne possède pas la plus forte amplitude il sera « vu » avec de faibles coefficients. Cette normalisation ne permettant aucune discrimination a été abandonnée.

Coef. Normalisé = coef. Analysé / coef. Maximal obtenu pour l'auto-corrélation de la référence choisie

Par le spectre « inconnu »

Pour cette normalisation, nous considérons qu'il ne peut pas y avoir de meilleure corrélation que celle du spectre inconnu par lui-même (auto-corrélation). Ainsi les coefficients obtenus lors de la corrélation croisée sont normalisé par rapport aux coefficients obtenus lors de l'auto-corrélation :

Coef. Normalisé = coef. Analysé / coef auto-corrélation du spectre inconnu

Nous pouvons avoir ici des coefficients supérieurs à 1 si l'un des spectres de référence possède une amplitude bien supérieure au spectre à analyser. De ce fait, son intérêt est limité puisqu'il est difficile de dire si nous avons trouvé le spectre optimal.

Par le rapport des coefficients maximaux

Pour cette normalisation, nous utilisons le rapport entre les coefficients maximum d'auto-corrélation du spectre inconnu d'une part et du spectre de référence considéré d'autre part. Ce rapport est différent pour chaque corrélation croisée puisque le spectre de référence change. Dans ce cas, le coefficient maximum sera égal à 1 et correspondra au spectre homologue de celui analysé.

Coef normalisé = coef analysé / (coef max auto-corrélation de la réf. / coef. max. spectre inconnu)

L'avantage de cette normalisation est qu'elle ne dépend pas de la présence du spectre inconnu dans la bibliothèque de référence.

IV.5.2 Analyse des gaussiennes

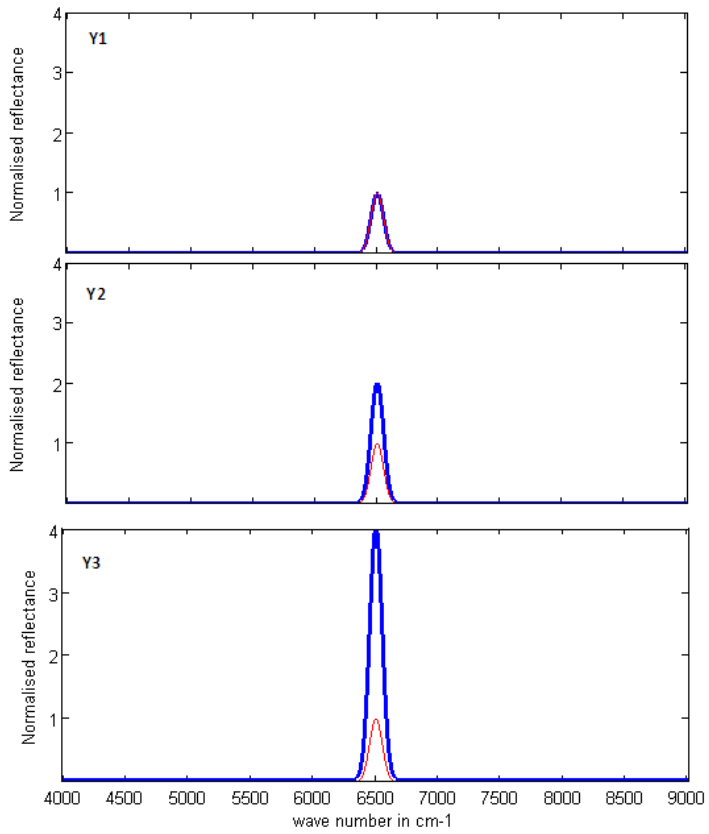
IV.5.2.1 Influence de l'amplitude des gaussiennes

Pour tester l'influence de l'amplitude sur la corrélation croisée, trois gaussiennes centrées sur un même nombre d'ondes (6500 cm^{-1}) et de même écart type (50 cm^{-1}) sont utilisées. Seule leur amplitude varie d'un facteur allant de 1 à 4 (figure IV-6). Les méthodes de normalisation des coefficients de corrélation croisée par le spectre « inconnu » et par le rapport entre les coefficients maximaux du spectre « inconnu » et du spectre de référence considéré sont utilisées. Les tableaux 6 a et b montrent les résultats de ces deux normalisations. La normalisation par l'auto-corrélation du spectre inconnu montre que c'est systématiquement la gaussienne Y3 (de plus forte amplitude) qui est majoritaire. Cela démontre clairement que cette normalisation n'est pas adaptée aux fortes variations d'intensité des spectres d'absorption IR. En revanche, la méthode du rapport des maxima montre dans tout les cas un maximum centré sur l'auto-corrélation. Le coefficient diminue ensuite au fur et à mesure que l'on s'éloigne de l'amplitude optimale.

Les tableaux 6a et 6b donnent également les valeurs angulaires du déphasage entre les signaux corrélés. Du fait que tous les spectres soient centrés sur le même nombre d'ondes (6500 cm^{-1}), nous observons que le déphasage reste faible même lorsque les rapports d'amplitude augmentent. La valeur du déphasage est égal à 0 à chaque fois qu'il y a auto-corrélation, et ce quelque soit la méthode de normalisation utilisée.

Dans ces tableaux, la matrice de corrélation se lit de la manière suivante:

Pour la ligne n°1 nous comparons le spectre Y1 avec les *n* colonnes de la ligne 1 c'est-à-dire Y1 versus Y1 puis Y1 versus Y2 etc.



a)

Normalisation par le spectre inconnu			
coef. (norm) > seuil	Y1	Y2	Y3
Y1	0.89283	1.2302	1.8288
Y2	NaN	0.89283	1.2302
Y3	NaN	NaN	0.89283
phase	Y1	Y2	Y3
Y1	0	-1.5467e-010	-3.407e-010
Y2	NaN	0	1.1675e-010
Y3	NaN	NaN	0

b)

normalisation par le rapport entre le spectre inconnu et le spectre de référence			
coef.(norm) > seuil	Y1	Y2	Y3
Y1	0.25113	0.12556	0.062782
Y2	0.12556	0.25113	0.12556
Y3	0.062782	0.12556	0.25113
phase	Y1	Y2	Y3
Y1	0	-3.059e-010	-1.6145e-010
Y2	3.059e-010	0	1.4445e-010
Y3	1.6145e-010	-1.4445e-010	0

Figure IV-6: Effet de l'amplitude et des méthodes de normalisation de la corrélation croisée. Les graphiques présentent les différents signaux impliqués. Les tableaux fournissent la matrice de corrélation des différents signaux et la matrice des valeurs angulaires du déphasage entre les signaux corrélés pour la normalisation par le spectre inconnu (a) et pour la normalisation par le rapport des coefficients d'auto-corrélation des spectres considérés (b).

Un deuxième exemple de l'influence de l'amplitude des gaussiennes sur les valeurs des coefficients de la corrélation croisée est présenté dans les figure IV-7 a et b.

Dans ce cas deux séries de trois gaussiennes sont utilisés (figure IV-7a). Les gaussiennes de la première série sont centré sur 6500 cm^{-1} et ont un écart type de 50 cm^{-1}. Les gaussiennes de la deuxième série sont centrées sur 5500 cm^{-1} et on un écart type de 100 cm^{-1}. L'évolution de l'amplitude est linéaire dans le premier cas (l'amplitude est doublée entre chaque gaussiennes) et non linéaire dans le second cas (Y5 est le double de Y4 et Y6 vaut 5 fois Y4).

La figure IV-7b montre que le coefficient moyen est maximal dans le cas de l'auto-corrélation, puis diminue suivant une exponentielle décroissante avec l'augmentation de l'intensité. Dans le cas de la première série (courbe bleu), qui présente une évolution linéaire de l'amplitude des gaussiennes, trois possibilités de rapport d'amplitude existent, 1 (auto-corrélation), 2 (Y1 versus Y2 et Y2 versus Y3) et 4 (Y1 versus Y3). Dans le cas de la deuxième série (courbe verte), les rapports d'amplitudes possibles sont : 1 (auto-corrélation), 2 (Y4 versus Y5), 5 (Y4 versus Y6) et 2.5 (Y5 versus Y6). Le décalage entre les résultats obtenus pour les

deux séries montre que la position et/ou la variance influence la valeur du coefficient moyen mais ne modifie pas la forme de la loi de régression.

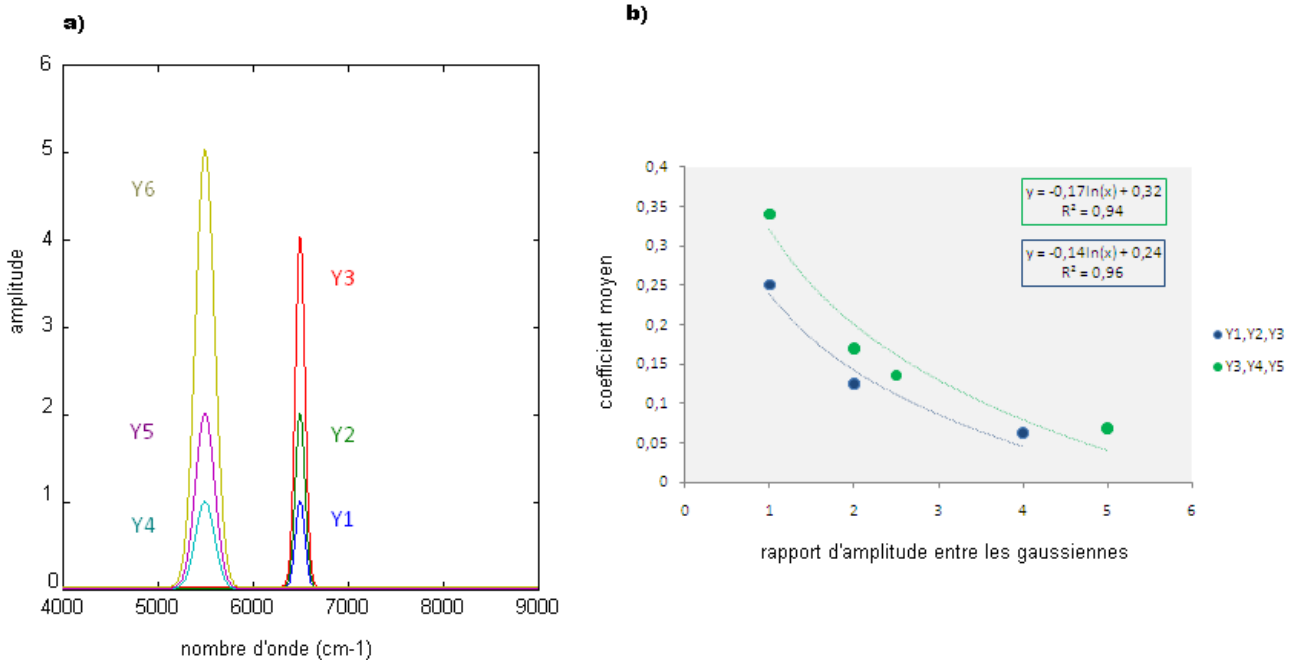


Figure IV-7: a) représentation des différents signaux impliqués dans la corrélation croisée, b) évolution du coefficient moyen en fonction de l'augmentation ou de la diminution de l'amplitude entre la gaussienne inconnue et la gaussienne de référence.

L'utilisation de la normalisation par le rapport des maxima permet ainsi de corriger le biais que nous avions avec la normalisation par le spectre inconnu lorsque l'intensité du pic était trop grande par rapport au spectre inconnu.

IV.5.2.2 Influence de la variance des gaussiennes

Dans ce paragraphe, l'effet de la variation de l'écart type des gaussiennes corrélées sur les coefficients de corrélation est analysé. Une série de gaussiennes centrées sur 6500 cm^{-1} et d'écart type croissant (de 25 cm^{-1} à 550 cm^{-1}) est utilisée. Les gaussiennes étant toute centrées sur le même nombre d'ondes, le déphasage entre les gaussiennes varie très peu et n'est pas analysé dans cet exemple.

La figure IV-8 présente une sélection de quatre gaussiennes d'écart type variable et le scalogramme de leur transformée en ondelettes. Elle montre que plus les gaussiennes ont un écart type important, plus leur signal est localisé dans les basses fréquences (haut du scalogramme).

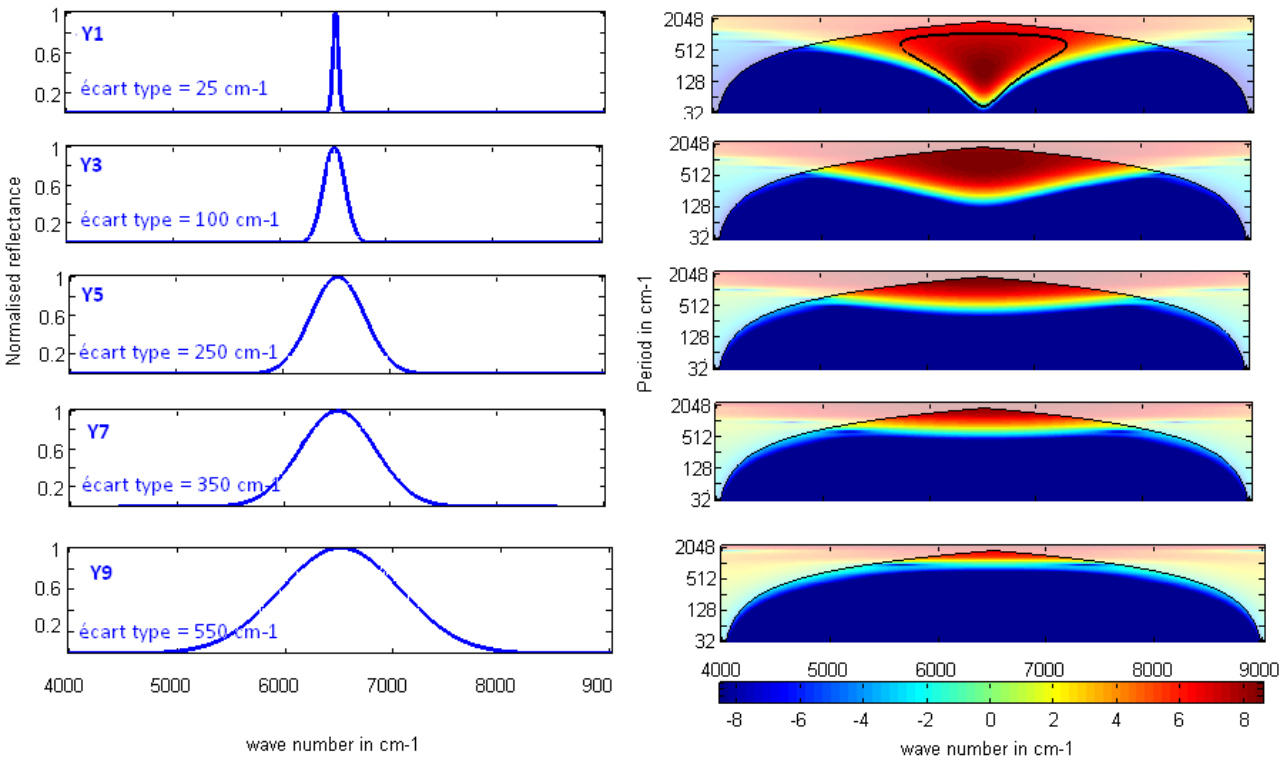


Figure IV-8: illustration de l'effet de la variabilité des gaussiennes sur le scalogramme associé. L'augmentation de la variance se traduit par un décalage du signal vers les basses fréquences (haut du scalogramme). Le trait fin aux extrémités des scalogrammes marquent la limite où les effets de bord deviennent importants (COI définit dans le paragraphe 1.3.2). Le trait épais souligne les zones statistiquement significatives ($IC_{95\%}$). Il n'est présent que dans le cas de la gaussienne Y1. Dans les autres cas, les coefficients ne dépassent pas la barre imposée par $IC_{95\%}$.

La figure IV-9 montre l'influence de l'écart type ainsi que la position du signal. Dans un premier temps, le coefficient moyen augmente lorsque l'écart type augmente, et ce jusqu' à une valeur critique d'environs 150 cm^{-1}. Au-delà de cette valeur elle diminue. Ce comportement s'explique par le fait que lorsque l'écart type des gaussiennes augmente, elles sont constituées de plus basses fréquences. Par conséquent, le nombre de coefficients pris en compte pour le calcul de la valeur moyenne diminue tout comme leur intensité.

L'expérience a été mené sur des gaussiennes localisé à 6500 cm^{-1} et 5500 cm^{-1} pour des seuils différents (0.6, 0.8 et 0.9). Les courbes sont quasi parallèles, excepté pour le spectre centré sur 5000 cm^{-1}. Dans ce cas, plus l'écart type des gaussiennes augmente, plus l'on s'éloigne des autres courbes, car les effets de bord commencent à influencer le calcul des coefficients même si ces effets sont encore inférieurs à la limite imposée par le (COI). Dans tout les cas, le coefficient maximal est constant quelque soit l'écart type ou le seuil utilisé.

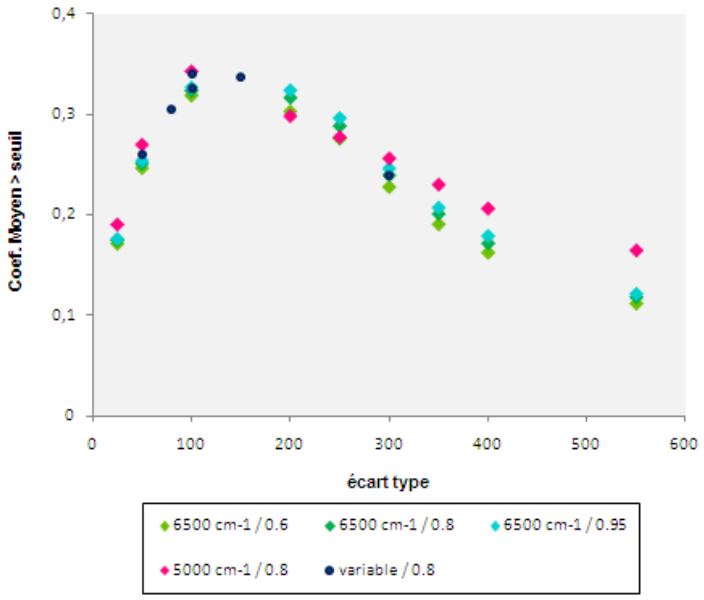


Figure IV-9:Evolution des coefficients en fonction de l'écart type pour différentes séries de gaussiennes. Les séries sont centrées sur différents nombres d'ondes (6500 cm^{-1}, 5500 cm^{-1}ou variable) et pour différent seuil.

Il sera donc impossible d'étudier des signaux ayant un écart type (variance) trop important, à moins de recourir à la méthode du « zero padding » qui consiste à ajouter des zéros aux extrémités du signal afin d'avoir un signal de $^{2n+1}$ et ainsi d'atténuer les effets de bord en augmentant la résolution de l'analyse.

IV.5.2.3 Influence de la position des gaussiennes

L'exemple précédent met en évidence l'influence de la localisation de la gaussienne (figure IV-9). Ainsi, dans ce paragraphe, plusieurs séries de gaussiennes sont utilisées pour analyser le comportement des coefficients de corrélation et du déphasage lié au décalage des gaussiennes en fonction de la distance séparant la gaussienne « inconnue » et la gaussienne de référence. Quatre séries sont considérées (figure IV-10 a). Les séries 1 et 2 sont constituées de gaussiennes d'écart type de 100 cm^{-1}, réparties entre 6500 et 7000 cm^{-1}. La série 1 contient 10 gaussiennes décalées d'une distance allant entre 25 cm^{-1} entre les premières gaussiennes à 100 cm^{-1} entre les dernières gaussiennes de la série. La série 2 contient 10 gaussiennes décalées d'une distance de 50 cm^{-1}. Les séries 3 et 4 sont constituées de 6 gaussiennes respectivement centrées sur 6300, 6400, 6500, 6600, 6700 et 6800 cm^{-1}. Dans la séries 3 les gaussiennes ont un écart type de 100 cm^{-1}. Dans la séries 4 les gaussiennes ont un écart type de 25 cm^{-1}.

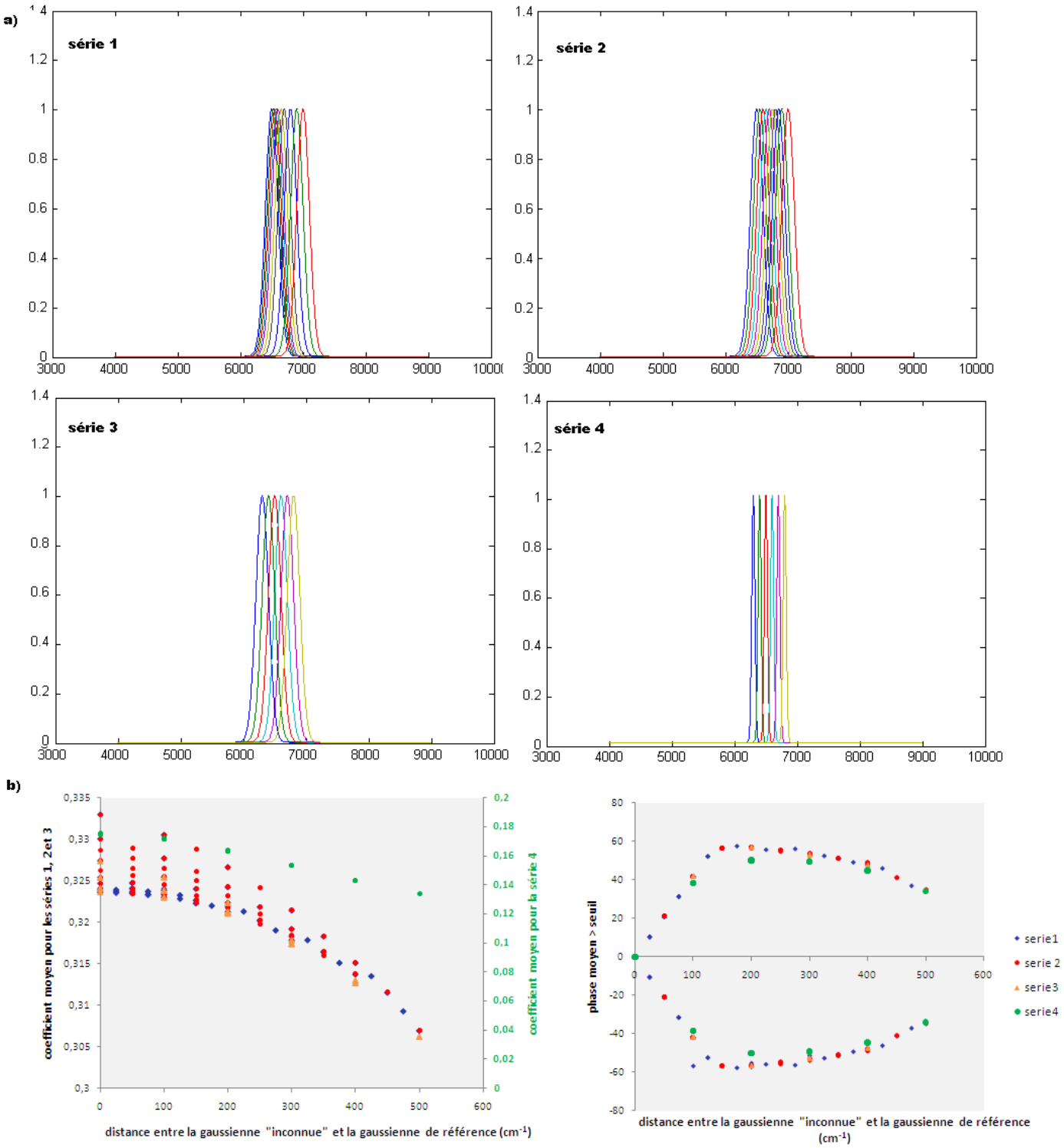


Figure IV-10: a) présentation des différents signaux impliqués dans la corrélation croisée. Entre chaque série, la localisation, le pas de décalage et l'écart type des gaussiennes changent. b) : Effet de la distance entre la gaussienne « inconnu » et la gaussienne de référence sur l'évolution du coefficient de corrélation croisée et du déphasage.

La figure IV-10b montre que quelque soit l'écart type et la position des gaussiennes, ainsi que le pas de décalage, l'évolution de la phase moyenne suit toujours le même schéma. Elle vaut zéro dans le cas de l'auto-corrélation puis sa valeur absolue augmente progressivement lorsque les gaussiennes « s'éloignent ». **Le signe positif ou négatif décrit simplement le sens de déplacement de la gaussienne inconnue par rapport à la gaussienne de référence**. Cette augmentation est détectée jusqu'à ce que la distance entre la gaussienne inconnue et la gaussienne de référence atteigne 150 cm^{-1}. Au delà de cette distance, la phase enregistrée n'a plus de signification.

L'évolution du coefficient est principalement dépendante de l'écart type des gaussiennes considérées. Une évolution exponentielle est enregistrée sauf pour la série 4 ou la relation est linéaire. Ainsi, plus la distance entre la gaussienne inconnue et la gaussienne de référence augmente plus la valeur du coefficient moyen diminue. L'interaction entre les deux séries sera alors plus faible et le déphasage en valeur absolue augmentera jusqu'à la distance critique de 150 cm^{-1}. Ce comportement s'explique par la diminution de la zone de chevauchement entre les deux gaussiennes (plus les gaussiennes sont rapprochés plus la zone de recouvrement est importante).

Contrairement au coefficient, l'évolution de la phase (entre 0 et 150 cm^{-1}) est linéaire. Ainsi, la valeur de la phase est un indice de la distance du décalage. Cette caractéristique permet non seulement de travailler sur la dissymétrie de pic, mais également sur la reconnaissance de la gaussienne de référence la plus proche, dans le cas ou la gaussienne inconnue n'est pas dans bibliothèque de référence.

Dans tous les cas, le couple coefficient et phase associée facilite l'identification de la gaussienne inconnue (coefficient maximal et phase égale 0).

IV.5.3 Caractérisation de figures de mélanges de gaussiennes

IV.5.3.1 Présence de la gaussienne inconnue dans la bibliothèque de référence

L'objectif est de travailler sur la quantification des teneurs en minéraux d'un mélange par le biais de la transformée en ondelettes couplée à la méthode de corrélation croisée. Ce paragraphe est donc un préambule à l'analyse de spectres infrarouge de mélanges de minéraux argileux. Dans cet exemple, des spectres de mélanges de gaussiennes sont réalisés à partir de deux gaussiennes « pures » Y1 et Y2 (figure IV-11a). Y1 est centrée sur 6500 cm^{-1}, son écart type est de 200 cm^{-1}. Y2 est centrée sur 7000 cm^{-1}, son écart type est de 100 cm^{-1}. Quatre mélanges sont ensuite réalisés à partir de ces gaussiennes sur la base de l'hypothèse des mélanges linéaires. Ainsi, nous avons :

- Y3 = 80*Y1 + 20*Y2
- Y4 = 60*Y1 + 40*Y2
- Y5 = 40*Y1+ 60*Y2
- Y6 = 80*Y1 + 20*Y2

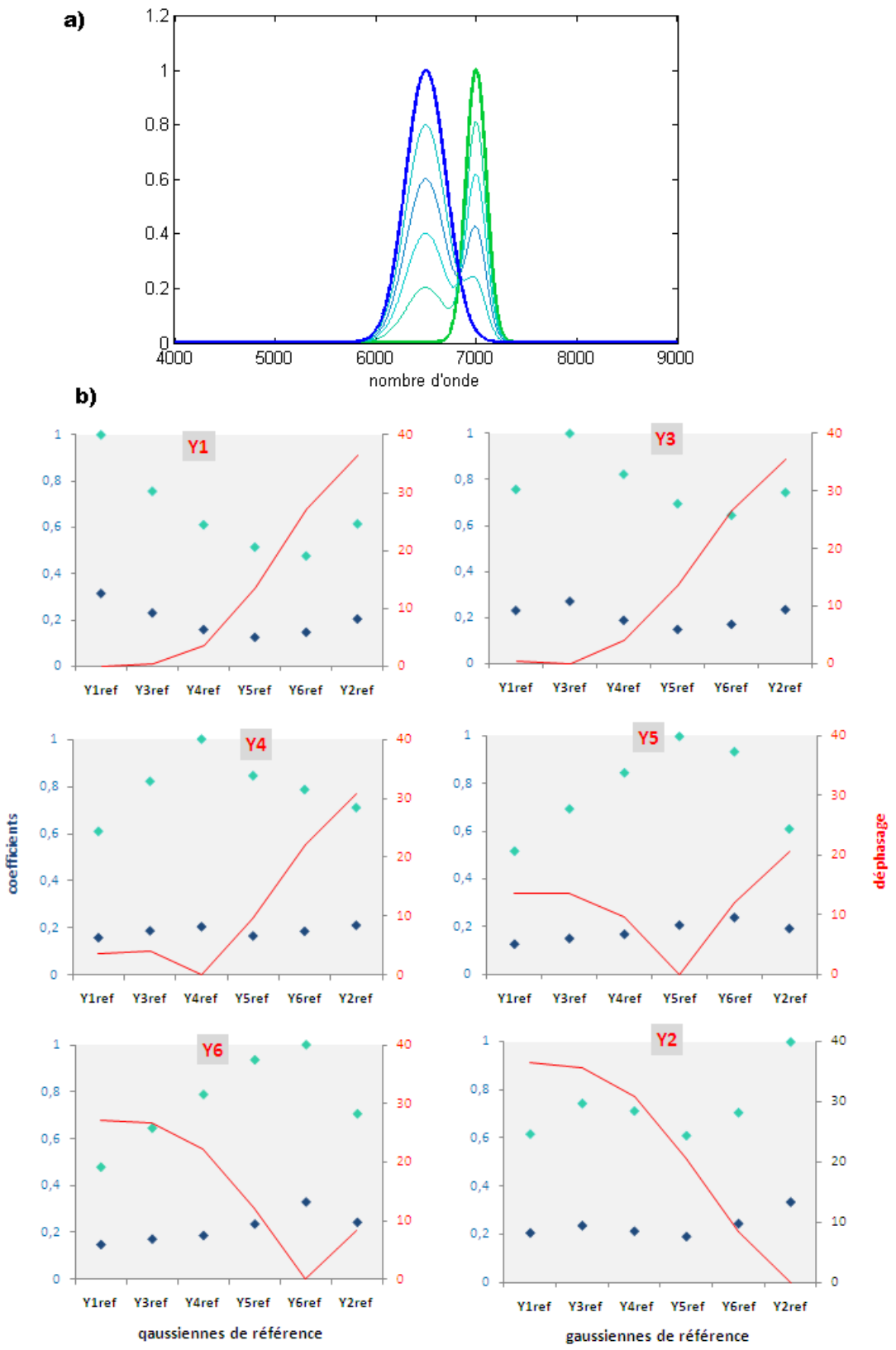


Figure IV-11: a)Présentation des différents signaux impliqués dans la corrélation croisée. Les Gaussiennes Y1 (en bleu) et Y2 (en vert) représentent « deux pôles purs ». b) évolution de la valeur de la phase (rouge), du coefficient maximal (bleu clair) et du coefficient moyen (bleu foncé) lors de la corrélation de Y1 et Y2 pour les pôles purs et Y3, Y4, Y5, Y6 (notée en rouge) avec chaque gaussiennes de la bibliothèque de référence.

La figure IV-11b montre que le coefficient de corrélation maximal (bleu clair) et moyen (bleu foncé) évoluent de manière comparable. Leur valeur est maximale dans le cas de l'auto-corrélation puis il diminue lorsque la différence entre le spectre « inconnu » et le spectre de référence augmente. Ainsi, dans le diagramme Y4 (figure IV-11b) qui présente la corrélation entre la gaussienne Y4 et l'ensemble des gaussiennes de référence (figure IV-11a), la valeur du coefficient est maximale pour l'auto corrélation de Y4. Il est plus faible pour la corrélation entre Y4 et Y3 d'un coté et Y4 et Y5 de l'autre et il est minimal dans le cas de la corrélation entre Y4 et Y1 d'une part et Y4 et Y2 d'autre part. Ainsi, la composition de Y4 est bien intermédiaire entre celle de Y3 et de Y5. La même analyse peut être faite pour chaque diagramme en se basant sur l'évolution de la phase.

Il est donc possible d'identifier la gaussienne « inconnue » si elle est dans la bibliothèque de référence. La quantification ainsi établie est une quantification « absolue » donnée par l'auto-corrélation.

L'évolution du coefficient montre bien le passage du pôle Y1 majoritaire au pôle Y2 majoritaire mais cette évolution n'est pas linéaire par rapport à la quantité d'Y1 ou d'Y2 dans le mélange. Ceci est d'autant plus vrai que la différence des pôles purs est marquée en termes d'écart type, de position et d'amplitude. Dans la figure IV-11b, le pôle Y1 (centré sur 6500 cm^{-1} et d'écart type supérieur à celui de Y2) influence plus le mélange que le pôle Y2.

En revanche, la phase montre un comportement moins influencé par cette différence et indique plus clairement le mélange le plus proche du spectre « inconnu ». Lorsque le spectre inconnu est présent dans la bibliothèque de référence le déphasage est un marqueur plus robuste et plus simple à interpréter que les coefficients qui sont très sensibles aux effets de bords et à la forme du spectre de mélange. Il est donc important d'utiliser une ondelette à partie réelle et complexe comme l'ondelette de Morlet pour pouvoir identifier et quantifier le spectre « inconnu ».

IV.5.3.2 Absence de la gaussienne « inconnue » dans la bibliothèque de référence

Les exemples précédents ont montré qu'il est possible d'identifier la gaussienne inconnue si elle est dans la bibliothèque de référence et de réaliser une quantification absolue.

De la même manière, il est possible d'identifier la gaussienne la plus proche de la gaussienne inconnu quand elle n'est pas dans la bibliothèque mais la quantification sera alors « relative ». Le terme de quantification relative indiquant que l'on ne peut donner la composition absolue du mélange mais la fourchette des spectres de référence dans lequel le spectre « inconnu » est inclus.

Ainsi, le tableau IV-1 présente les valeurs des coefficients moyens et du déphasage obtenu pour la corrélation entre des mélanges de gaussiennes « inconnu » de composition variable (colonne) et des mélanges de gaussiennes de références (ligne). Les gaussiennes Y1, Y2, Y3, Y4, Y5 et Y6 sont les même que dans l'exemple précédent (figure IV-11b).

a)

coefficient	Y1 ref.	Y3 ref.	Y4 ref.	Y5 ref.	Y6 ref.	Y2 ref.
Y1	**0,3167**	0,2316	0,1591	0,1253	0,1473	0,2054
90Y1*10Y2	0,2734	0,2591	0,1778	0,1396	0,1626	0,2254
Y3	0,2316	**0,2718**	0,1889	0,1488	0,1712	0,2357
70Y1*30Y2	0,1922	0,2277	0,1984	0,1561	0,1771	0,2414
Y4	0,1591	0,1889	0,2062	0,1664	0,1866	**0,2125**
50Y1*50Y2	0,1356	0,1614	0,1791	0,1814	0,2041	0,1917
Y5	0,1253	0,1488	0,1664	0,2045	**0,2361**	0,1902
30Y1*70Y2	0,1299	0,1527	0,1700	0,2151	0,2864	0,2093
Y6	0,1473	0,1712	0,1866	0,2361	**0,3318**	0,2438
10Y1*90Y2	0,1734	0,2000	0,2151	0,2294	0,2903	0,2855
Y2	0,2054	0,2357	0,2125	0,1902	0,2438	**0,3329**

b)

déphasage	Y1 ref.	Y3 ref.	Y4 ref.	Y5 ref.	Y6 ref.	Y2 ref.
Y1	**0**	0,3709	3,6595	13,5919	27,2800	36,5347
90Y1*10Y2	0,1780	0,4970	4,2013	13,9675	27,2833	36,2851
Y3	0,3709	**0**	4,0430	13,6069	26,7252	35,5783
70Y1*30Y2	1,2309	1,0875	2,5852	12,3246	25,1385	33,8808
Y4	3,6595	4,0430	**0**	9,6635	22,1908	30,8345
50Y1*50Y2	7,7218	7,9414	4,1356	5,5088	17,7990	26,3706
Y5	13,5919	13,6069	9,6635	**0**	12,0319	20,5613
30Y1*70Y2	20,8221	20,5366	16,0188	6,4936	5,8748	13,5398
Y6	27,2800	26,7252	22,1908	12,0319	**0**	8,4141
10Y1*90Y2	32,5194	31,7219	27,0680	16,7643	4,6898	3,6088
Y2	36,5347	35,5783	30,8345	20,5613	8,4141	**0**

Tableau 16 : valeur moyenne des coefficients et des phases enregistrée lors de la corrélation entre des spectres inconnus absents de la bibliothèque de référence.

Le tableau IV-1 montre ainsi que dans le cas de spectres « inconnu » non inclus dans la bibliothèque de référence (ligne surligné en rouge) la valeur minimale de la phase et maximale du coefficient indiquent bien le mélange de référence qui a la composition la plus similaire au spectre « inconnu ». Ainsi le spectre composé de 50 % de Y1 et 50% de Y2 est compris entre Y4 eY5 soit entre 60 % de Y1 et 40% de Y2 et 40 % de Y1 et 60% de Y2.

IV.5.4 Quantification de mélanges à deux minéraux argileux

Ce paragraphe décrit différents cas concrets de mélanges réalisés à partir de minéraux argileux (illite, smectite et kaolinite). Ces mélanges « synthétiques » construits en laboratoire seront ainsi analysés par l'analyse en ondelettes couplée à la méthode de corrélation croisée.

IV.5.4.1 Matériel

Les spectres de mélanges à deux minéraux argileux (illite Velay/smectite LRPC, illite Velay/kaolinite Clérac, smectite LRPC/kaolinite Clérac) analysés dans le chapitre 3 par les méthodes statistiques (PLSR et ACP) et par l'étude de l'aire du pic caractéristique des minéraux argileux proche de 4500 cm^{-1}, sont utilisés dans cette application. L'ensemble de ces spectres de mélanges ont été enregistrés dans la gamme du proche IR (4000 à 9000 cm^{-1}) à l'aide du spectromètre Infrarouge à Transformée de Fourier (IRTF) en mode réflexion diffuse. Cette échelle correspond en effet à une étape intermédiaire entre le laboratoire et le terrain, tant en terme de résolution spatiale et spectrale qu'en terme de procédure d'analyse des échantillons (cf. chapitre 2).

L'analyse porte plus particulièrement sur la gamme de longueurs d'ondes comprise entre 4000 et 5000 cm^{-1} et sur le pic proche de 4500 cm^{-1}. En effet, ce pic d'absorption est commun aux spectres des trois minéraux argileux considérés (illite, smectite et kaolinite) et il est présent dans les spectres enregistrés in-situ en lumière naturelle. Enfin, il correspond à un canal du satellite multispectral ASTER, ce qui permet par la suite d'envisager une nouvelle d'analyse des sols.

La figure suivante présente les spectres des différents mélanges utilisés sur l'ensemble de la gamme du proche infrarouge du spectromètre IRTF (4000 à 9000cm^{-1}). Les zones encadrées correspondent à la gamme de longueurs d'ondes plus particulièrement analysées dans cette étude.

Ainsi, la figure IV-12a correspond aux spectres qui composent la bibliothèque de spectres de référence nommée « mélanges illite/smectite ». Les spectre de mélanges d'illite et de smectite sont en proportion allant de 10 à 90% par pas de 10%, et sont corrigés par la méthode du continuum removal. Le spectre vert épais correspond au mélange constitué de 90% d'illite et 10% de smectite ; le spectre rouge épais correspond au mélange constitué de 10% d'illite et 90% de smectite.

La figure IV-12b correspond aux spectres qui composent la bibliothèque nommée « mélanges illite/kaolinite ». Les spectres de mélanges d'illite et de kaolinite sont en proportion allant de 10 à 90% par pas de 10%, et sont corrigés par la méthode du continuum removal. Le spectre vert épais correspond au mélange constitué de 90% d'illite et 10% de kaolinite ; le spectre bleu épais correspond au mélange constitué de 10% d'illite et 90% de kaolinite

La figure IV-12c correspond aux spectres qui composent la bibliothèque nommée « mélanges kaolinite/smectite ». Les spectres de mélanges de kaolinite et de smectite sont en proportion allant de 10 à 90% par pas de 10%, et sont corrigés par la méthode du continuum removal. Le spectre bleu épais correspond au mélange constitué de 90% de kaolinite et 10% de smectite ; le spectre rouge épais correspond au mélange constitué de 10% de kaolinite et 90% de smectite.

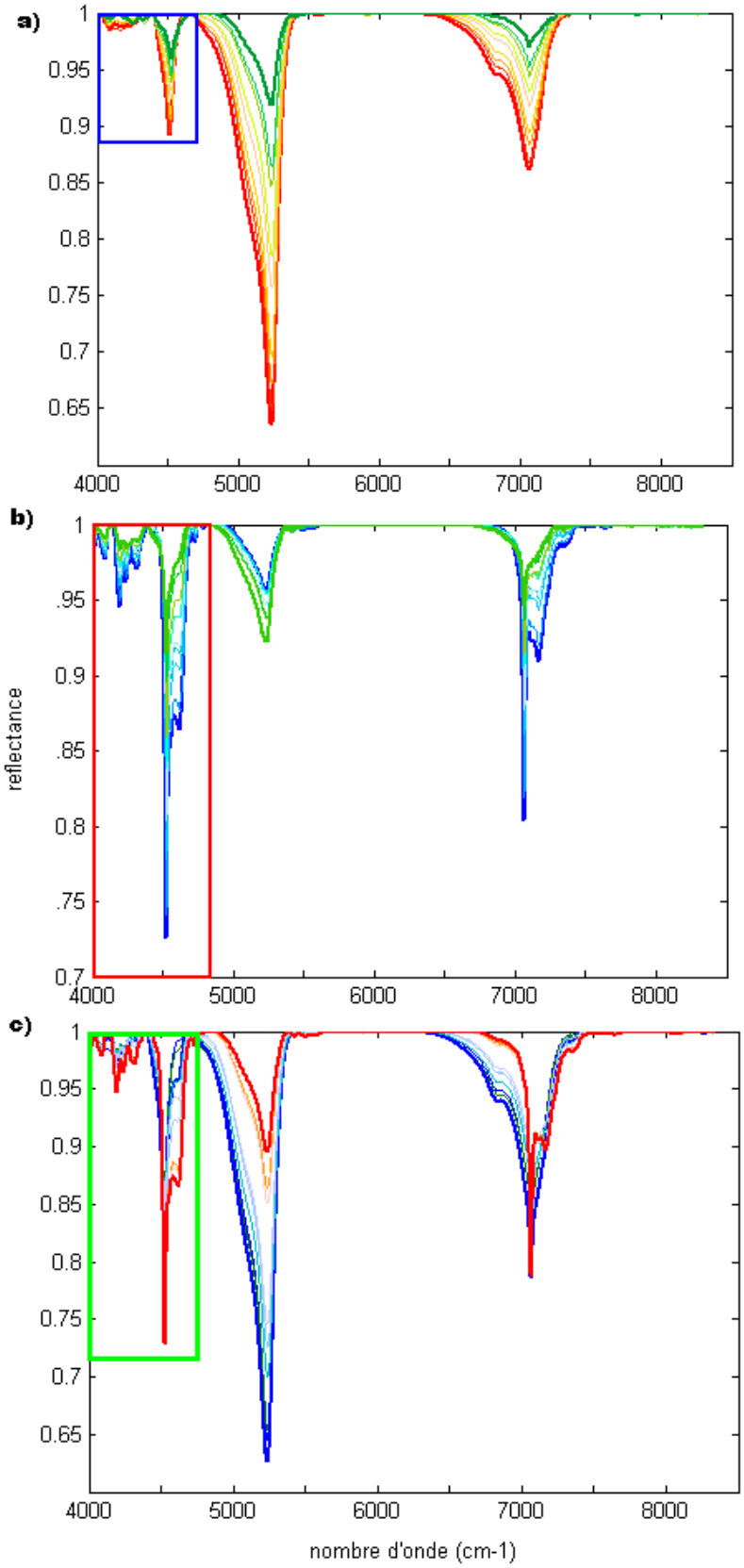


Figure IV-12: a) spectres de la bibliothèque de spectres de référence nommée « mélanges illite (vert)/smectite (rouge) ». b) spectres de la bibliothèque nommée « mélanges illite/kaolinite (bleu) ». c) spectres de la bibliothèque nommée « mélanges kaolinite/smectite ».

IV.5.4.2 Méthode utilisée

Afin de montrer que la quantification de la teneur en un minéral argileux dans un mélange est possible par la méthode d'analyse en ondelettes couplée à la corrélation croisée, nous comparons la teneur en smectite, illite ou kaolinite (suivant les mélanges étudiés) à la valeur de déphasage, pour chaque série de corrélation. Le terme « série de corrélation » correspond à la corrélation d'un des spectres de la bibliothèque de référence (qui joue alors le rôle de spectre "inconnu") avec l'ensemble des spectres de la bibliothèque. Nous nous plaçons alors dans le cas d'une bibliothèque contenant le spectre « inconnu ».

La figure IV-13 présente un exemple de séries de corrélation entre en spectre "inconnu" en rouge (10I90S) et les spectres de référence constitués de mélange d'illite et de smectite en proportion allant de 10% à90% par pas de 10%. La première étape présente le cas de l'auto-corrélation, puis progressivement un décalage apparaît entre le "spectre inconnu" (en rouge) et le spectre de référence avec lequel il est comparé (en bleu). Le déphasage moyen mesuré entre le coefficient moyen de deux spectres considérés, est une mesure globale de ce décalage.

Cette analyse a été faite pour chacun des mélanges à deux minéraux présentés précédemment.

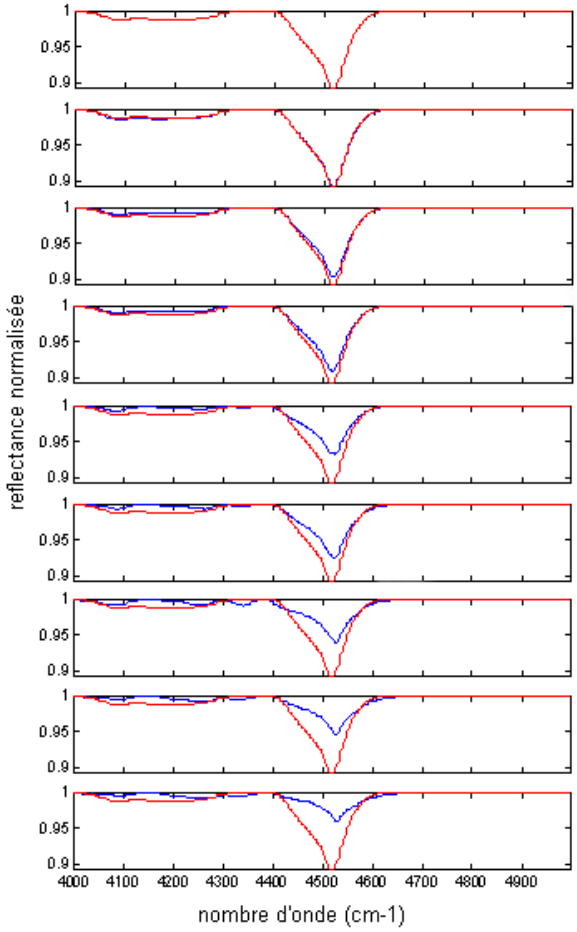


Figure IV-13: Exemple d'une série de corrélations réalisée entre un spectre « inconnu » (en rouge) et les spectres de la bibliothèque de référence (en bleu). Pour chaque corrélation croisée, la valeur de la phase moyenne est enregistrée. Elle mesure le décalage entre le spectre « inconnu » et un spectre de référence.

IV.5.4.3 Résultats

Mélanges illite/ smectite

La figure IV-14 présente les mélanges d'illite et smectite dans la gamme de nombre d'onde allant de 4000 à 4800 cm^{-1} avant et après retrait du continuum. Ils constituent les spectres de référence de la bibliothèque « mélange illite/smectite ».

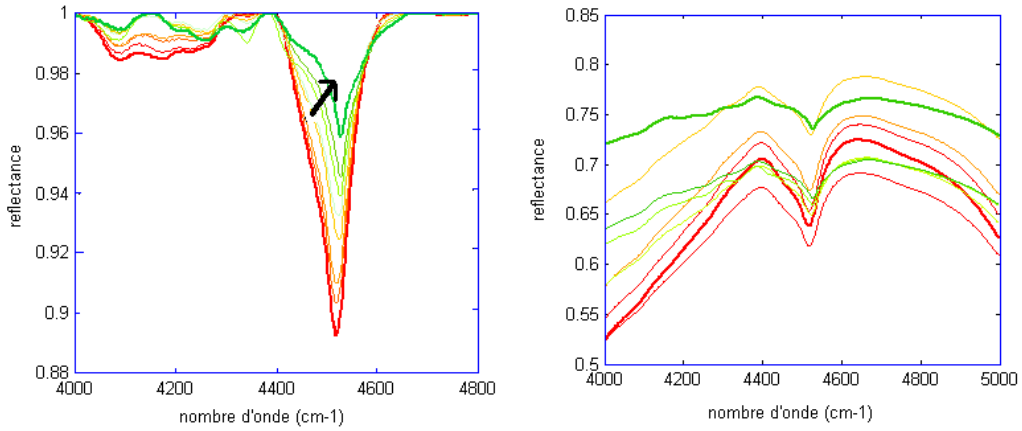


Figure IV-14: spectres d'absorption illite/smectite pour l'intervalle 4000-4800 cm^{-1} corrigés du continuum removal (CR) (a) et non traités (b). La flèche noire indique le sens du déphasage.

La bibliothèque contient neuf spectres donc neuf séries de corrélation sont réalisées. Ainsi, chaque courbe de la figure IV-15 correspond à l'évolution de la valeur du déphasage entre un des spectres de mélanges de la bibliothèque (précisé en légende) et l'ensemble des spectres de cette bibliothèque « mélange illite/smectite ». L'évolution du déphasage est donnée en fonction de la teneur en illite du spectre de référence correspondant.

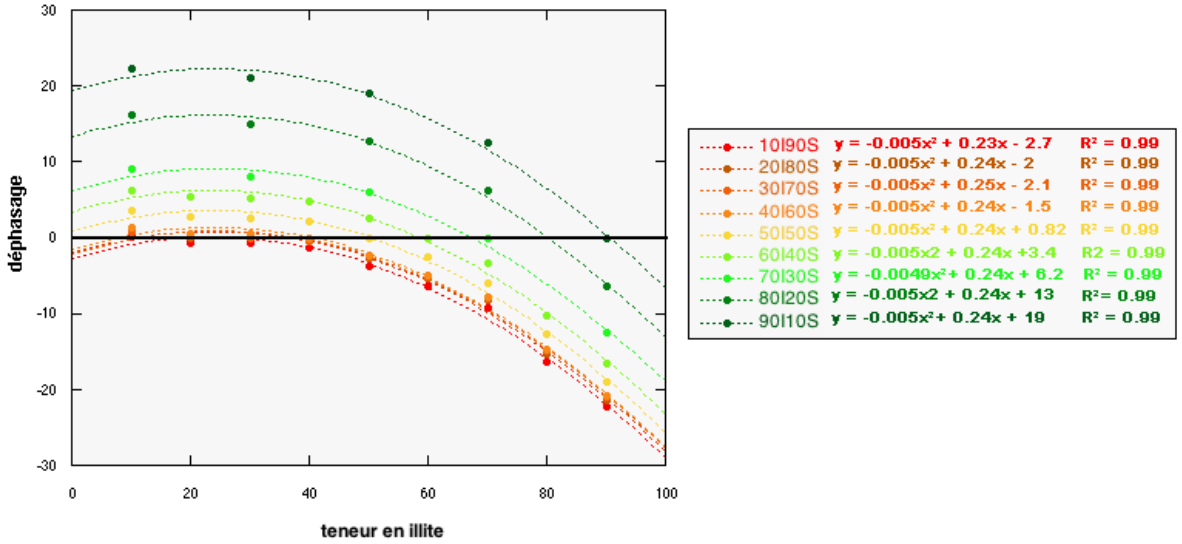


Figure IV-15: Résultats obtenus par corrélation croisée sur les spectres corrigés du CR. Les coefficients de corrélation non linéaire et les équations non linéaires caractéristiques sont donnés en légende. les paramètres a et b des équations des courbes de tendance sont constant alors que c témoigne du changement de composition.

Dans le cas **de l'auto-corrélation,** c'est-à-dire lors de la corrélation croisée du spectre « inconnu » avec lui-même (contenu dans la bibliothèque de référence), la valeur du déphasage est de 0. On peut donc déterminer le pourcentage du mélange de manière précise sur la ligne de déphasage 0° (ligne noire épaisse, figure IV-15).

Pour **l'ensemble des séries**, le déphasage en valeur absolue entre le spectre « inconnu » et le spectre de référence varie peu, lorsque la teneur illite est majoritaire dans le mélange de référence. En revanche, il évolue rapidement lorsque la teneur en illite diminue dans les spectres de référence. Cette évolution non linéaire est liée à la forme non symétrique du pic de l'illite par rapport à celui de la smectite et au fait que la position du pic d'absorption évolue dans le même sens que la dissymétrie du pic lorsque la teneur en smectite augmente dans le mélange (figure IV-14).

Les valeurs négatives du déphasage moyen s'expliquent par le sens dans lequel se décale le pic d'absorption lorsque la teneur en illite augmente (flèche noire, figure IV-14). Lorsque le décalage se fait vers la droite (comme c'est le cas ici), le déphasage est négatif.

D'autre part, l'évolution non linéaire explique le décalage entre chaque série de corrélation. En effet, nous observons bien un décalage de plus en plus important entre deux séries lorsque la teneur en illite dans le spectre « inconnu » augmente (évolution suivant une verticale). Cela montre l'aspect symétrique de la matrice de corrélation avec un R^2 égale à 1 (auto-corrélation) dans la diagonale, ce qui est classique. Cela se matérialise également par une matrice de déphasage symétrique mais dont la diagonale donne un déphasage nul.

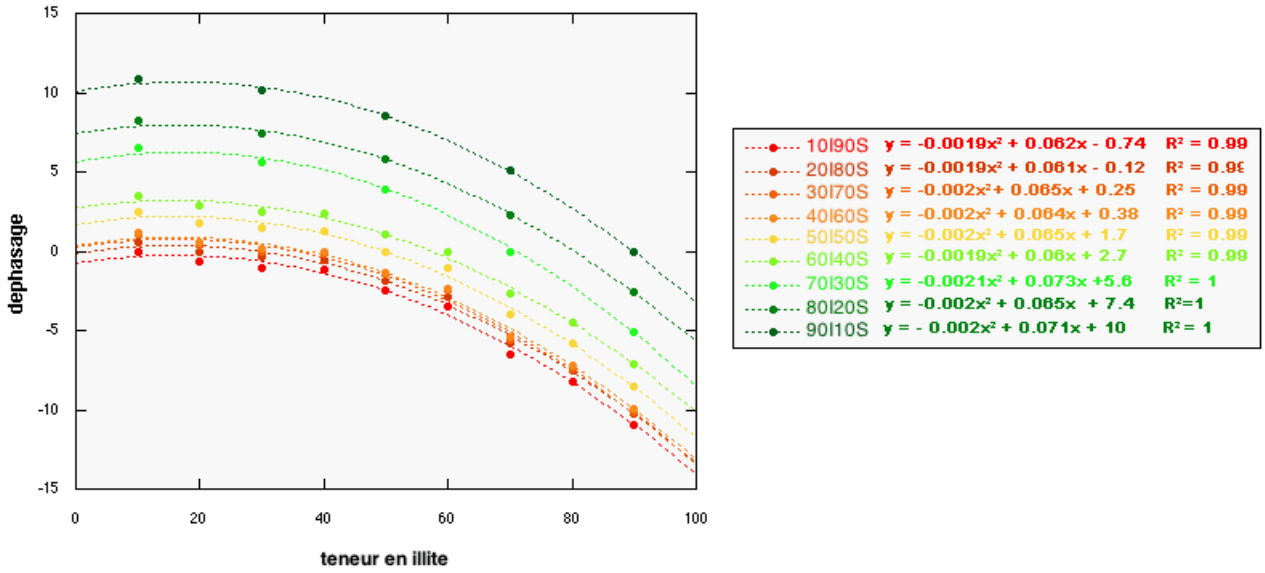


Figure IV-16: Résultats obtenus par corrélation croisée sur les spectres non traités. Les coefficients de corrélation non linéaire et les équations non linéaires caractéristiques. Notez la grande constance des paramètres a et b alors que c témoigne du changement de composition.

Le fait de travailler sur des spectres non corrigé du CR ne change pas ce résultat (figure IV-16). En effet, le même type de relations non linéaire est trouvé pour les différentes séries. Cependant une augmentation plus importante du déphasage est observée dans le cas des spectres corrigés par le CR (figure IV-15) pour une même analyse croisée. Cela tend à prouver que la méthode du CR est un peu trop « grossière » et induit un décalage plus important du pic ce qui constitue un artéfact de la méthode.

Cette similitude entre les résultats s'explique par le fait que la méthode d'analyse des ondelettes couplée à la corrélation croisée est insensible à la ligne de base. En effet, la ligne de base des spectres infrarouges constitue une très basse fréquence dans le signal et est automatiquement supprimé par le COI (cône d'influence de l'ondelette). Cela constitue un avantage majeur par rapport aux autres méthodes. Par la suite nous travaillerons donc uniquement avec les spectres bruts.

Enfin, lorsque la teneur en illite est faible (<20%, figure IV-15 et IV-16) la courbe est quasiment horizontale, ce qui rend la quantification délicate car un bruit de mesure pourrait affecter la pente locale. Cependant, la méthode des ondelettes peut minimiser cet effet en prenant une fréquence minimale d'analyse supérieure aux hautes fréquences du bruit expérimental.

Mélanges illite/kaolinite

La figure IV-17 présente les mélanges d'illite et de kaolinite dans la gamme de nombre d'onde allant de 4000 à 4800 cm^{-1} avant et après retrait du continuum.

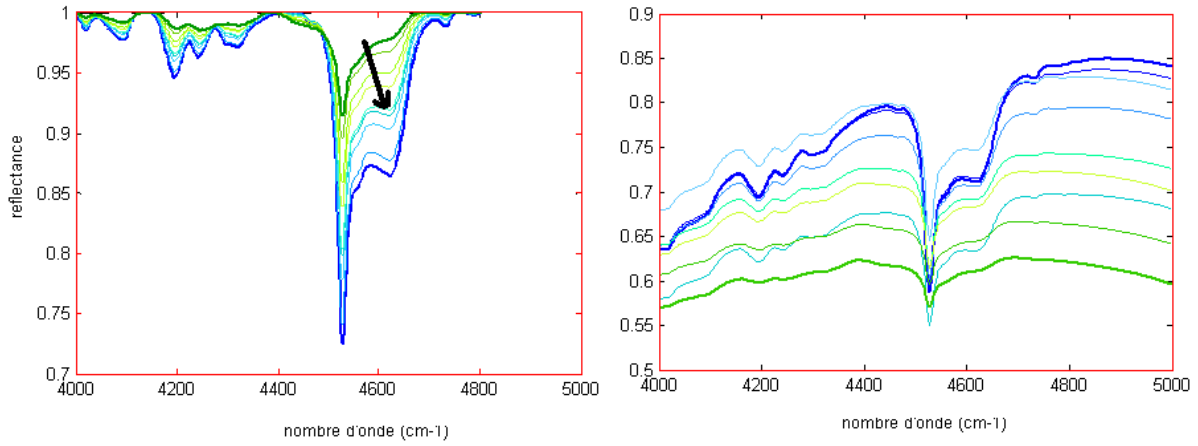


Figure IV-17: Zooms sur le spectre d'absorption illite/kaolinite pour l'intervalle 4000-4800 cm^{-1} présenté en figure de la figure IV-12b. a) spectres corrigés du continuum removal (CR; b)) spectres non traités. La flèche noire indique le sens du déphasage.

De la même manière que pour la relation illite/smectite, la figure IV-17 montre que les mélanges d'illite et de kaolinite présentent une évolution non linéaire, croissant avec l'augmentation en illite du mélange (ou la diminution en kaolinite). Lorsque la teneur en illite est inférieure à 50% dans le mélange, la valeur du déphasage varie peu. En revanche, lorsque la kaolinite devient majoritaire, l'augmentation du déphasage est rapide (figure IV-18a).

Cette évolution non linéaire explique le décalage entre chaque série de corrélation (pour une teneur en illite donnée, le long d'une verticale). En effet, nous observons bien un décalage de plus en plus important entre deux séries lorsque la teneur en illite augmente dans le spectre « inconnu ».

L'évolution de la phase des valeurs négatives vers les valeurs positives (contrairement à la figure précédente) s'explique par le sens dans lequel évolue le pic d'absorption lors de l'augmentation de la teneur en illite dans le mélange (flèche noire, figure IV-17). En effet, dans le cas des mélanges illite/smectite, lorsque la teneur en illite augmente, le décalage se produit vers la droite (figure IV-14) alors que dans le cas des mélanges illite/kaolinite, il se produit vers la gauche (figure IV-17).

Lorsque la teneur en illite dans les spectres de référence diminue, les courbes sont plus horizontales et la quantification devient plus délicate. Le CR semble plus efficace que le simple filtrage de basse fréquence des ondelettes. Cependant, les résultats sont très bons dans les deux cas (figure IV-18b).

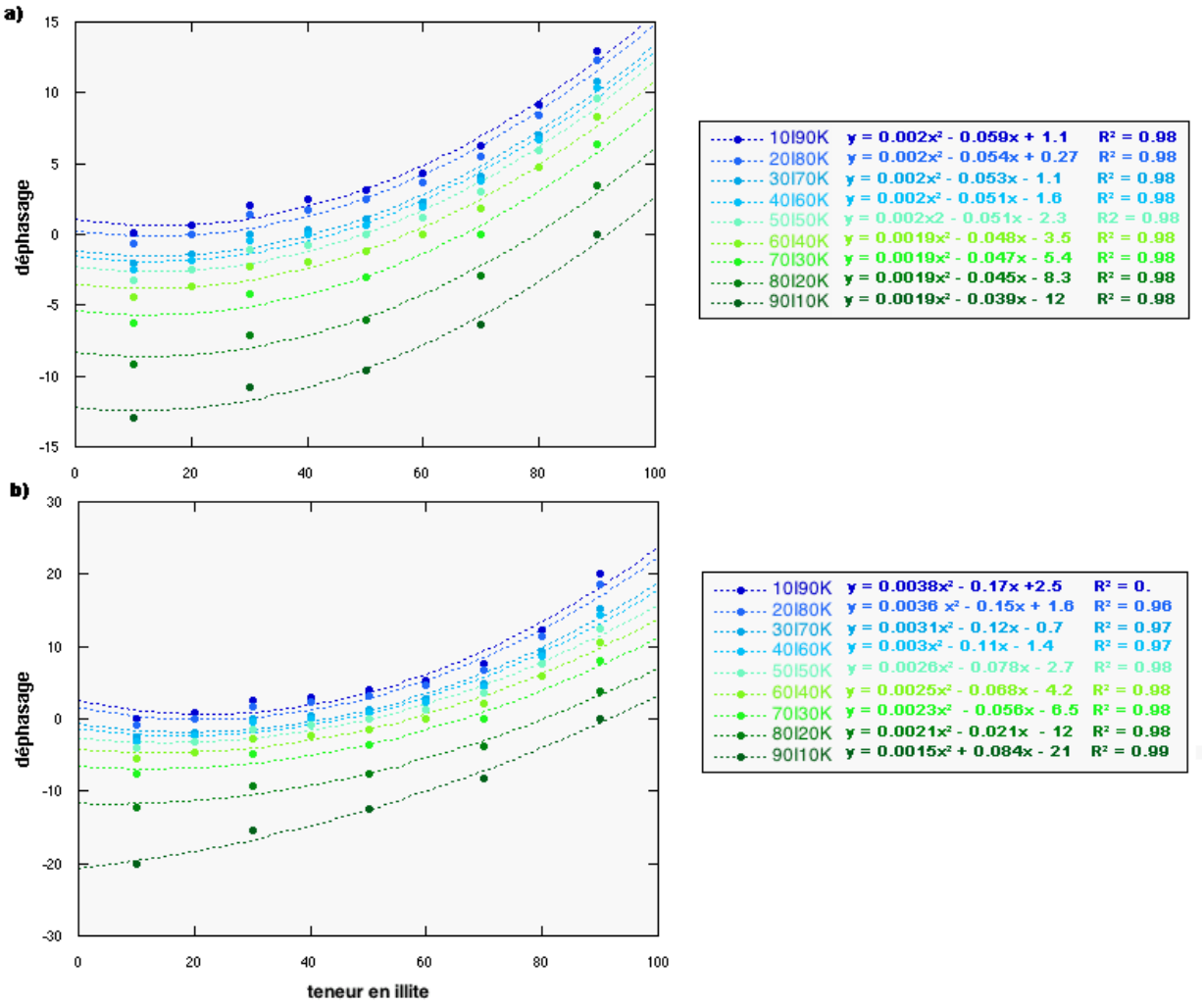


Figure IV-18: Résultats obtenus par corrélation croisée sur les spectres corrigés du CR (a), et sur les spectres non traités (b). Les coefficients de corrélation non linéaire et les équations non linéaires caractéristiques. Les paramètres a et b sont constants alors que c témoigne du changement de composition. Chaque couleur (du bleu au rouge) correspond à une série de corrélation réalisée entre un des spectres de la bibliothèque de référence et l'ensemble des spectres de référence. L'évolution du déphasage est la même pour chaque série.

Mélanges kaolinite/smectite

La figure IV-19 présente les mélanges de kaolinite et de smectite dans la gamme de nombre d'onde allant de 4000 à 4800 cm^{-1} avant et après retrait du continuum.

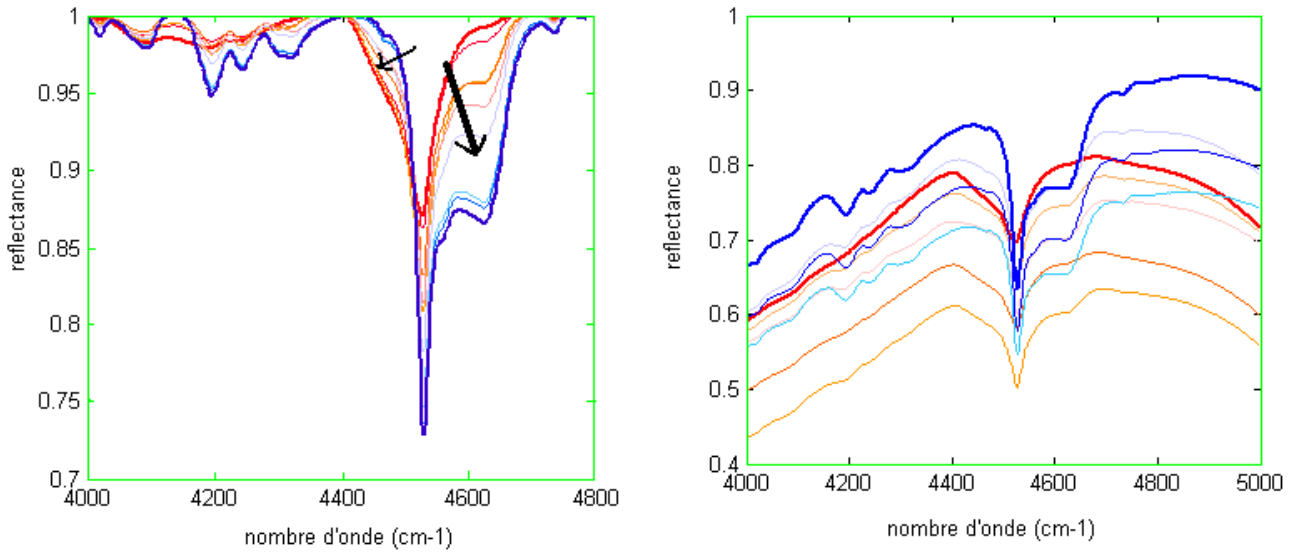


Figure IV-19: Zooms sur le spectre d'absorption illite/smectite pour l'intervalle 4000-4800 cm^{-1} présenté en figure de la figure IV-12c. a) spectres corrigés du continuum removal (CR); b) spectres non traités. La flèche noire indique le sens du déphasage.

Contrairement aux relations misent en évidence précédemment, la figure IV-20 montre que dans le cas des mélanges de kaolinite et de smectite, l'évolution de la valeur du déphasage entre le spectre « inconnu » et les spectres de références se fait de manière linéaire. Cette évolution linéaire du déphasage s'explique par la forme symétrique du pic d'absorption de la smectite vis-à-vis du pic d'absorption de la kaolinite (figure IV-19). En effet, le déphasage mesuré correspond à une valeur moyenne calculée sur l'ensemble des coefficients significatifs. La figure IV-19 montre ainsi que lorsque la teneur en smectite augmente dans le mélange, le pic d'absorption à 4500 cm^{-1} se décale vers la gauche. A l'inverse, lorsque la teneur en kaolinite augmente dans le mélange, le pic se décale vers la droite. Ainsi, l'évolution moyenne est bien toujours vers la droite (dans le sens négatif) mais est « ralentie » par la composante (prise en compte dans le calcul de la phase moyenne) amenée par la symétrie du pic d'absorption de la smectite (décalage inverse au décalage principal).

La figure IV-20b montre que la méthode d'analyse en ondelettes couplée à la corrélation croisée des spectres non traités présente des variations assez proches de celles obtenues pour les données spectrales corrigées du CR (figure IV-20a). Si nous observons les pentes de ces spectres, elles sont situées entre celles des mélanges illite/kaolinite et celles des mélanges smectite/illite. Les coefficients de corrélation sont excellents et appuie l'utilisation des spectres non traités, évitant ainsi une étape de correction des spectres qui peut s'avérer parfois assez délicate.

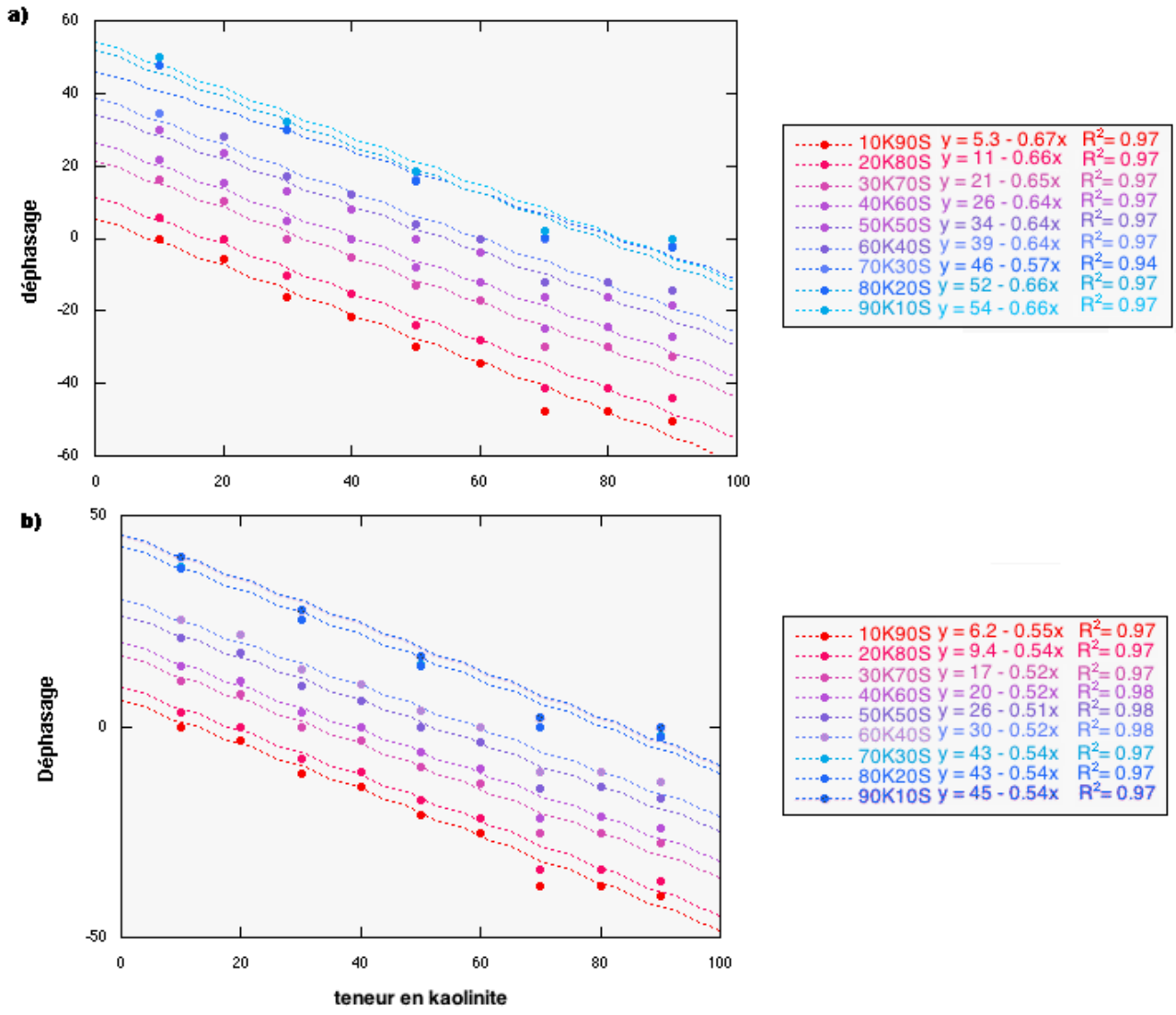


Figure IV-20: Résultats obtenus par corrélation croisée sur les spectres corrigés du CR (a), et sur les spectres non traités (b). Les coefficients de corrélation non linéaire et les équations non linéaires caractéristiques. Notez la grande constance des paramètres a et b alors que c témoigne du changement de composition. Chaque couleur (du bleu au rouge) correspond à une série de corrélation réalisée entre un des spectres de la bibliothèque de référence et l'ensemble des spectres de référence. L'évolution du déphasage est la même pour chaque série.

Le décalage entre les différentes droites (séries) est bien constant sauf pour les trois dernières séries (« 70K30S », « 80K20S », « 90K10S » en rouge). Dans ce cas, une erreur expérimentale lors de la réalisation du mélange de 70% de kaolinite avec 30% de smectite, ou lors de l'enregistrement du spectre peut expliquer le comportement de ce mélange.

Finalement, ce cas est idéal pour la quantification de la smectite.

IV.5.4.4 Application de la méthode aux mélanges à deux minéraux : discussion et conclusion

La méthode de la corrélation croisée entre la transformée en ondelettes d'un spectre « inconnu » et les transformées en ondelettes des spectres de référence a été appliquée à l'analyse de mélange binaire d'illite et/ou de smectite et/ou de kaolinite. Cette méthode donne accès à deux types de données. Elle permet le calcul du **coefficient moyen** et du **déphasage moyen** à partir des coefficients significatifs de la corrélation (compris entre 0,8 et 1).

Dans le cadre de l'analyse du pic d'absorption proche de 4500 cm^{-1} des spectres des mélanges de minéraux argileux, la valeur du déphasage a été choisie. En effet, lors de la comparaison de gaussiennes, ce paramètre s'est relevé plus sensible à de faibles différences entre les signaux comparés que le coefficient moyen. Il est donc plus adapté à l'étude de signaux fortement corrélés tels que les spectres proches infrarouges de l'illite et de la smectite.

Les résultats montrent que, en plus de la quantification de la kaolinite, la différentiation et la quantification de l'illite et de la smectite dans un mélange est possible. Cette quantification s'appuie sur une hypothèse préalable sur la composition du mélange « inconnu ». Chaque série de corrélation présentée dans cette analyse est faite entre un spectre « inconnu » et une bibliothèque contenant des spectres de mélanges de même composition que le spectre « inconnu ». Néanmoins, lorsque aucune connaissance préalable ne permet de conditionner le choix de la bibliothèque de référence, le spectre « inconnu » doit être comparé à une bibliothèque de référence contenant l'ensemble des spectres de mélanges des séries « illite/smectite », « kaolinite/smectite » et « illite/kaolinite ».

L'utilisation de cette hypothèse sur la composition du mélange « inconnu » a permis de mettre en évidence l'influence de la forme du pic sur la valeur du déphasage. En effet, les relations établies pour les séries de corrélation des mélanges illite/smectite et illite/kaolinite sont de types polynomiaux, alors que celles établies pour les mélanges kaolinite/smectite sont linéaires. Dans les deux premiers cas, le déphasage est unidirectionnel (figures IV-14 et IV-17) et la forme polynomiale s'explique par un changement plus lent de la forme du pic d'absorption lorsque l'illite devient majoritaire dans le mélange. Ainsi, dans ces deux cas, la quantification de l'illite (et donc de la kaolinite et de la smectite suivant le type de mélange) est plus délicate lorsqu'elle est majoritaire dans le mélange. Dans le cas des mélanges kaolinite/smectite, le déphasage est bidirectionnel. L'évolution du pic dans la direction de déphasage principale est limitée par une deuxième direction de déphasage qui apparaît lorsque la teneur en smectite augmente dans le mélange (figure IV-19). Cette influence de la forme du pic sur les relations mises en évidence pour la quantification de la teneur en illite, smectite et kaolinite dans les mélanges est très importante. En effet, la forme des pics d'absorption est dépendante de la composition chimique des minéraux responsables de ces pics. Ainsi, suivant les minéraux utilisés pour la réalisation des mélanges de référence, l'évolution du déphasage ne se fera probablement pas suivant la même relation.

IV.5.5 Analyse en ondelettes et corrélation croisée : discussion et conclusion

IV.5.5.1 Méthodes de normalisation

Pour quantifier les différentes argiles présentes dans un mélange par l'analyse en ondelettes couplée à la méthode de corrélation croisée, il est nécessaire de s'affranchir du problème des amplitudes caractéristique de ce type de méthode. Pour cela nous avons étudié différentes normalisation basée sur les spectres de références, sur le spectre inconnu ou sur une approche combinée.

Dans le cas d'une normalisation par les spectres de référence, chaque signal est normalisé par auto-corrélation du spectre de référence ce qui donne des coefficients égaux à 1 pour un spectre parfait correspondant à la référence de plus grande amplitude. Ces coefficients servent à la normalisation. En effet, ils font ressortir les signaux de la librairie qui ont des fortes amplitudes. C'est pourquoi cette approche a été abandonnée.

Dans le cas d'une normalisation par le spectre inconnu, tous les spectres sont normalisés par les coefficients d'auto-corrélation du spectre inconnu. Dans ce cas les signaux de forte amplitude ont des coefficients supérieurs à 1, ce qui rend difficile l'interprétation de la carte des coefficients de la corrélation.

Enfin, dans le cas d'une normalisation par une approche combinée, le signal est normalisé par sa fonction d'auto-corrélation puis divisé par le rapport des amplitudes maximales entre le spectre de référence et le spectre inconnu. Cette dernière normalisation diminue l'effet de l'amplitude sans toutefois l'annuler complètement (cf. paragraphe IV-5.2.2).

IV.5.5.2 Application à la caractérisation des argiles « gonflantes »

Principaux résultats

Lorsque le spectre inconnu existe dans la bibliothèque de référence, sa détection est absolue et les cartes de coefficients de la corrélation peuvent être utilisées de manière efficace. Les coefficients égaux à 1 montrent une correspondance entre le signal « inconnu » et l'un des éléments de la bibliothèque de référence, ce qui donne accès à la quantification absolue du spectre « inconnu ».

Lorsque le spectre inconnu n'a pas d'équivalent dans la bibliothèque de référence, une quantification relative où le déphasage le plus faible indique le ou les mélanges de composition la ou les plus proches de celle du spectre « inconnu » est possible.

La quantification de l'illite, de la smectite et de la kaolinite dans un mélange est alors dépendante de la qualité de la bibliothèque de référence. Notamment, des améliorations doivent intervenir au niveau de la détermination et de l'optimisation du choix des pôles de mélange, avec par exemple l'introduction de la notion de mélanges itératifs. En effet, l'analyse en ondelettes suivant l'approche développée dans ce chapitre s'avère très efficace, mais nécessite des bibliothèques spectrales suffisamment représentatives de différentes compositions chimiques de chaque minéral argileux. Lorsqu'il s'agit de mélanges intimes, il est nécessaire

d'identifier un à un les différents mélanges, et il faut alors les introduire comme référence dans la bibliothèque spectrale.

Nous avons également montré dans ce chapitre que l'analyse par corrélation croisée en ondelettes est peu influencée par le continuum (CR). Cette étape de l'analyse des données spectrales peut donc être évitée sans changement sensible de l'identification et de la quantification des minéraux argileux**.** Cela est particulièrement important dans le cas de l'analyse in-situ. En effet, le continuum, s'appuie sur des points particuliers du spectre. Cela peut être problématique lorsque les bandes d'absorption rapprochées tendent à écraser les maxima qui les séparent. Ainsi, appliquer un CR altère la mesure des bandes d'absorption particulièrement lorsque la pente locale de la portion du spectre est forte.

Enfin, la méthode développée permet de comparer entre eux des spectres tout en évitant d'introduire une connaissance *a priori* sur les critères spectraux qui les définissent. Contrairement aux méthodes statistiques et à l'étude des pics d'absorption caractéristiques, elle ne nécessite pas une étude préalable de la composition par diffraction des rayons X ou par analyses chimiques des éléments majeurs. Dans sa mise en application, elle permet de traiter rapidement un grand nombre de données tout en fournissant des résultats cohérents pour fournir des documents d'aide à la décision pour les acteurs de la prévention des risques naturels.

Perspectives

Un premier résultat pour la quantification des teneurs en illite, smectite et kaolinite dans des mélanges contenant ces trois minéraux, montre que, contrairement à l'ajout d'illite ou de smectite, **l'ajout de kaolinite dans un mélange n'en contenant pas initialement, entraine une modification importante des relations établies précédemment**. Ainsi lorsque l'on compare l'évolution du déphasage entre un spectre de mélange illite/smectite avec les spectres de la bibliothèque de « mélange illite/smectite », et l'évolution du déphasage entre un spectre de mélange illite/ smectite plus 10% de kaolinite par rapport à la même bibliothèque, on constate que les relations établies suivent une même loi polynomiale, alors que le décalage vertical imposé par l'ajout de kaolinite (axe secondaire en rouge) ne permet plus une quantification de la teneur en illite ou en smectite (figure IV-21a). En revanche, Les figure IV-21b et c montrent que l'ajout de 10% de smectite à des mélanges illite/kaolinite et de 10% d'illite à des mélanges smectite/kaolinite ne modifie que très peu l'évolution des déphasages mis en évidence lors de la corrélation croisée avec une bibliothèque de mélanges « illite/kaolinite » ou une bibliothèque de mélanges « kaolinite/smectite ». L'influence de la kaolinite sur les valeurs de déphasage s'explique par la forme prédominante des pics d'absorption des spectres de ce minéral vis-à-vis de la forme des pics d'absorption des spectres contenant de l'illite ou de la smectite (cf. chapitre 2).

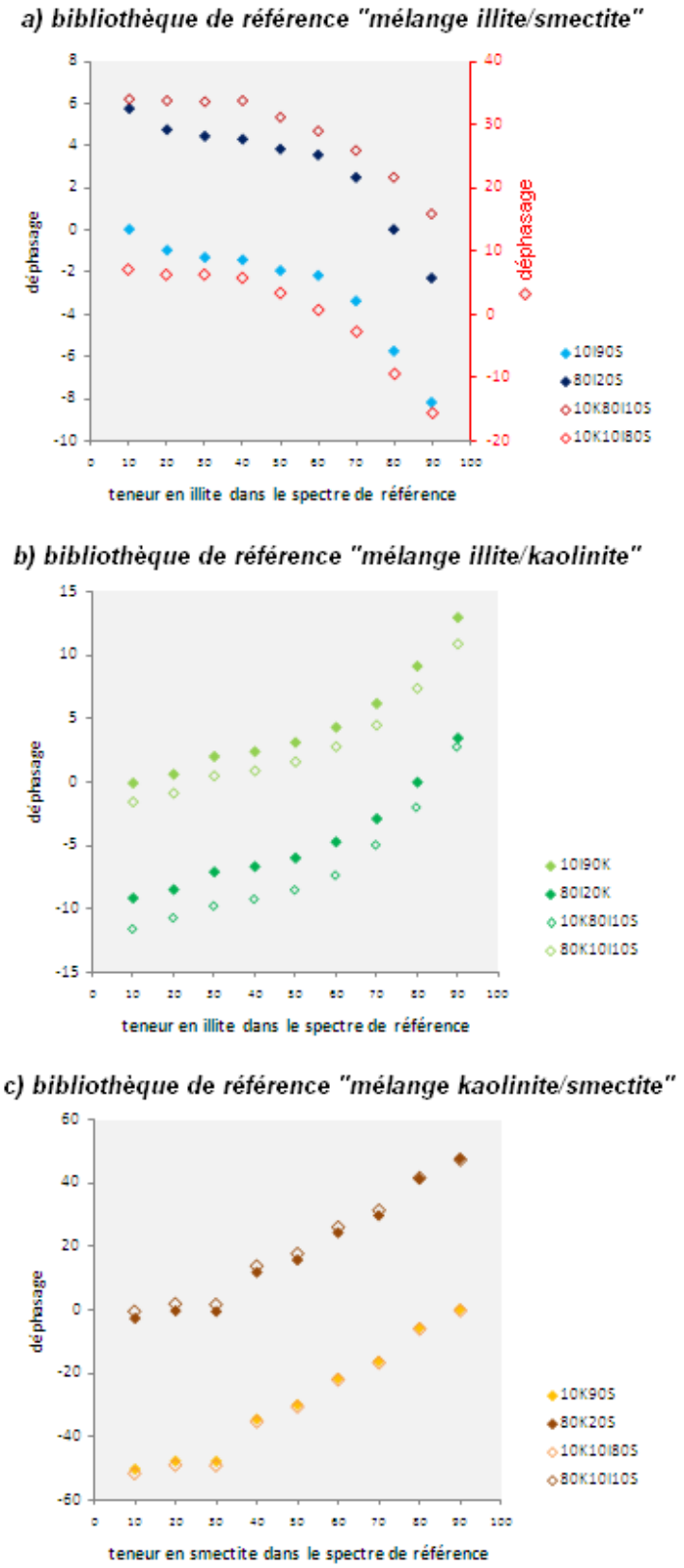


Figure IV-21: Evolution du déphasage lors de corrélations croisées entre des spectres « inconnu » contenant de l'illite, de la smectite et de la kaolinite et a) la bibliothèque de référence « mélange illite/smectite » (figure IV-14), b) la bibliothèque de référence « mélange illite/kaolinite (figure IV-17) et c) la bibliothèque de référence « mélange kaolinite/smectite » (figure IV-19). Les légendes indiquent les spectres « inconnu » considérés ainsi que leur composition.

Ces résultats préliminaires montrent que la quantification des teneurs en illite, smectite et kaolinite dans des mélanges synthétiques réalisés à partir de ces trois minéraux par la méthode de corrélation croisée pourra être étudiée de la même manière que pour les mélanges à deux minéraux. Cependant l'influence particulière de la kaolinite, déjà mise en évidence dans le chapitre 3 par l'étude de l'aire des pics d'absorption, devra être étudiée avec attention.

D'autre part, par le biais de l'étude des mélanges à deux minéraux, nous avons montré que la forme du pic d'absorption étudié, dépendant par ailleurs de la structure chimique des minéraux responsables de ce pic, joue un rôle fondamental dans l'établissement des relations basées sur les valeurs du déphasage. Cela implique que lors de la quantification des minéraux des mélanges étudiés, plusieurs spectres correspondant aux différentes compositions chimiques prises par un minéral soient inclus dans la bibliothèque de référence, de manière à prendre en compte la meilleure combinaison possible pour le mélange. Cela sera également fondamental pour l'étude des mélanges à trois minéraux et des sols.

Finalement, l'influence d'autres paramètres non étudiés dans ce chapitre devra être étudiée. C'est le cas notamment des pics d'absorption proches du pic considéré, tels que le pic d'absorption lié à la présence de calcite lors de l'analyse du pic proche de 4500 cm^{-1}. Ces pics peuvent en effet avoir une influence sur le pic d'absorption propre aux minéraux argileux.

De même les paramètres extérieurs tels que l'humidité, la rugosité ou l'éclairement doivent être pris en compte. Cependant, ces paramètres sont souvent basse fréquence puisqu'ils tendent à diminuer la réflectance général du spectre. Ainsi de la même manière que pour le CR, ils peuvent facilement être exclus de l'analyse.

Enfin, les spectres infrarouges provenant d'autres capteurs (spectromètre de terrain, capteur hyperspectraux, multispectraux) pourront être analysés par cette méthode afin d'étudier l'influence des changements de gamme de longueurs d'ondes ainsi que de la résolution spatiale et spectrale sur les résultats obtenus en laboratoire.

Du point de vue de la méthode, une seule fonction ondelette a été testée pour la déconvolution des spectres. Même si son choix est parfaitement justifié, la quantification pourrait peut être, être améliorée en utilisant plusieurs fonctions « mères ». En effet, plusieurs études montrent que suivant la famille d'ondelettes choisie, des variations importantes sont observées (Blackburn 2007). Cependant, le temps de calcul serait fortement alourdi par ce type d'approche et ne permettrait pas de répondre à notre objectif de fournir un outil facilement utilisable pour quantifier les teneurs en argile des sols expansifs. Il est cependant clair qu'en recherche fondamentale cette voie doit être explorée.

Conclusion

Le travail présenté dans ce chapitre montre que l'analyse en ondelettes tient ses promesses pour la quantification de la teneur en smectite, illite et kaolinite, mais des améliorations et des ajouts doivent encore être apportés au programme pour affiner la quantification. Ainsi, malgré ces défauts de jeunesse, cette analyse a démontré tout son potentiel. Elle présente aussi l'avantage d'être transposable très facilement vers d'autres problématiques comme par exemple l'évolution d'un contaminant dans les sols. Cependant, le cas des argiles est sûrement le plus redoutable en raison de la forte similitude des réponses spectrales des minéraux argileux. Il est donc évident qu'en améliorant cette approche sur les argiles, cela rendra d'autant plus simple la transposition et l'utilisation de la méthode dans d'autres domaines environnementaux.

CONCLUSIONS GENERALES ET PERSPECTIVES

V. CONCLUSIONS GENERALES ET PERSPECTIVES

Ce travail a pour but l'amélioration de la détection et de la quantification des minéraux argileux, et plus particulièrement des minéraux argileux « gonflants », dans la partie superficielle des sols à partir de mesures de spectrométrie infrarouge. L'étude a pour objectif d'accompagner la démarche des régions pour la prévention du risque de retrait-gonflement des argiles de l'échelle locale et à l'échelle globale. Les méthodes testées et développées dans ce manuscrit visent ainsi à trouver **dans quelle mesure la spectrométrie infrarouge peut constituer une méthode opérationnelle d'indentification, de localisation et de quantification des zones riches en argiles « gonflantes »** tout en apportant une information homogène et continue aux différentes échelles intéressant les décideurs.

Au cours de ce travail, nous avons pu tester les capacités et les limites de méthodes classiquement utilisées en science du sol. Nous avons ainsi démontré que les méthodes statistiques telle que l'Analyse en Composantes Principales (ACP), et les méthodes de régression par moindres carrés partiels (PLSR) permettent d'établir des modèles de prédiction des teneurs en kaolinite, illite et smectite. Cependant, les modèles sont fortement dépendants des données de calibration et présentent donc une corrélation quasi-parfaite avec celles-ci. L'utilisation des modèles est donc largement limitée par la dépendance des modèles à un certain nombre de paramètres (qualité et quantité des données, qualité de l'échantillonnage, présence d'information redondante dans les données, etc.) qui conditionnent la qualité de l'information servant de base à la construction des modèles. Ainsi, **nous avons démontré que les méthodes statistiques ne peuvent être utilisées dans une application opérationnelle.**

Nous avons ensuite travaillé sur une approche basée sur la forme et l'aire des pics d'absorption des spectres de minéraux argileux et de sols. Les relations établies dans plusieurs études antérieures pour la quantification de la kaolinite à partir de données spectrales acquises dans le moyen infrarouge ont été validées en laboratoire pour plusieurs types de mélanges contenant de la kaolinite et dans plusieurs domaines dans le proche et le moyen infrarouge.

Le domaine, plus spécifique, du proche infrarouge (4000 à 10000 cm^{-1} ou 1000 à 2500 nm), qui constitue un continuum entre le laboratoire, le terrain et le satellite, a été plus spécifiquement étudié. En effet, à l'heure actuelle, la majorité des spectromètres de terrain utilisés pour l'analyse in-situ des sols, enregistrent la réflectance dans cette gamme de longueurs d'onde. Ce domaine est également utilisé dans les applications de télédétection qui pourront être développées suite à ce travail pour des images provenant du satellite ASTER. Il est l'un des meilleurs candidats pour une application opérationnelle.

L'étude des absorptions des spectres de minéraux argileux et de sols et plus particulièrement le pic d'absorption proche de 4500 cm^{-1} (2200nm), a montré au travers de diverses études sa potentialité en science du sol avec la quantification des teneurs en argiles (Lagacherie et al. 2008) et en minéralogie avec une quantification suivant une loi exponentielle du rapport kaolinite/smectite (Petit et al. 2003). Notre étude a confirmé l'existence de cette loi exponentielle avec le rapport kaolinite/smectite mais aussi kaolinite/illite.

Cette constatation montre que la détermination par une relation linéaire des argiles dans un sol ne fonctionne que pour de faibles quantités de kaolinite. Dans le cas contraire, la présence de kaolinite dans les sols engendre une augmentation exponentielle de l'aire de ces pics qui tend à masquer l'influence des autres argiles. Cela remet en cause les relations linéaires classiquement utilisées. De plus, cette méthodologie, en plus de mettre en avant les difficultés liées à la perte d'information et de qualité du signal lors du passage du laboratoire au terrain, ne s'est pas montrée satisfaisante pour la différentiation et la quantification de l'illite et de la smectite dans les mélanges « synthétiques » et à fortiori dans les sols.

Finalement, Aucune de ces deux premières méthodes (statistiques et aire des pics) ne permet d'automatiser et de simplifier en temps et en coût d'analyses la quantification de la teneur en minéraux gonflants des sols. La spectrométrie apporte donc une information spatialisée complémentaire aux analyses ponctuelles actuellement utilisées (analyses géotechniques et DRX) mais ne permet en aucun cas de les remplacer. Nous avons ainsi développé une méthode originale travaillant directement sur la comparaison de signaux infrarouges dans leur ensemble. L'analyse en ondelettes par sa capacité d'analyse multi-échelle, est l'outil mathématique qui s'est avéré être le plus adapté aux traitements des spectres infrarouges.

Par le biais des transformées en ondelettes analytiques continues des spectres infrarouges couplées aux méthodes de corrélation croisée, l'identification des différents constituants argileux présents dans les spectres infrarouges a été réalisée en s'appuyant sur les coefficients d'ondelettes et sur le déphasage.

Nous avons montré qu'il était possible d'établir des corrélations entre une bibliothèque de référence et un spectre « inconnu » principalement sur la base du déphasage entre les bandes d'absorption propres à chaque minéral et ce, sans connaissance *a priori* sur les données spectrales. De plus, l'étape délicate de la correction des spectres de leur continuum a été supprimée. En effet, les résultats obtenus étaient comparables, voire plus précis, lorsque l'on évitait l'étape du « continuum removal ». Les résultats obtenus montrent que cette méthode permet la quantification « absolue » des proportions d'illite, de smectite et de kaolinite dans un mélange lorsque la bibliothèque de référence contient le spectre inconnu. Dans le cas contraire, elle permet d'obtenir une quantification « relative » en déterminant une fourchette de composition possible pour le spectre « inconnu » à partir des spectres de références présentant le déphasage le plus faible possible. On perçoit alors que la qualité de la quantification est directement dépendante de la qualité des bibliothèques de référence.

Ainsi, l'outil développé dans ce travail constitue une base solide pour systématiser la caractérisation et la quantification des argiles dans les sols. Il doit permettre à terme de diminuer, voire de remplacer, les analyses telles que les analyses géotechniques ou les analyses par diffraction des rayons X. Elles sont actuellement la base des études menées par les bureaux d'études dans le cadre de la prévention du risque de retrait-gonflement des argiles.

Finalement, **les fondements d'une nouvelle méthode d'analyse des données de spectrométrie infrarouge ont été apportés par cette thèse** et de nombreuses possibilités de développement se dégagent de ce travail.

Dans le cadre de la quantification des minéraux argileux « gonflants » présent dans les sols, plusieurs paramètres doivent être mieux pris en considération. C'est le cas par exemple de la composition chimique des minéraux de référence. Nous avons utilisé dans cette étude une seule illite, une seule smectite et une seule kaolinite. Or, ces minéraux peuvent intégrer dans leur structure différents éléments. Par exemple, il existe différentes smectites en fonction des substitutions Al-Fe telles que la beidellite (Al) ou la nontronite (Fe).

Les changements de résolution liés aux changements d'échelle doivent également être étudiés. Ces changements ont été abordés dans ce travail vis-à-vis des mesures en laboratoire et *in situ*. En revanche, l'échelle régionale, importante pour la réalisation des cartes de prévention des risques, n'a pu être abordée. Si le passage du laboratoire au terrain s'est avéré assez simple du fait que la résolution spectrale et spatiale change peu, il n'en est pas de même pour le passage à l'échelle régionale qui engendre une perte d'homogénéité entre le pixel et les autres échelles en termes de rugosité de surface, d'humidité, de composition minérale et de présence de végétation. Ces paramètres modifient fortement la réponse spectrale. De plus, la résolution spectrale joue aussi un rôle important. Si dans un premier temps des signatures spectrales proches du laboratoire peuvent être conservées en utilisant l'imagerie hyperspectrale, le passage à l'opérationnel imposera l'utilisation de capteurs multispectraux par définition plus dépouillés (résolution spectrale plus faible) mais bien meilleur marché.

Un autre développement non négligeable est la confrontation des analyses *in situ* et de la cartographie régionale par satellites aux données de mouvement de sols afin de quantifier l'influence de la teneur en argiles « gonflantes » sur les phénomènes de retrait-gonflement des sols. Les développements actuels des cartes de déplacements des sols (projets européens TERRAFIRMA, DO SMS) montrent que l'on est capable de faire de la surveillance sur l'aléa mouvement de terrain. Le libre accès aux données RADAR dans le cadre du programme GMES permettra de généraliser ce type de données.

Un autre aspect qui peut être regardé est le seuil modal minimal d'argiles « gonflantes » nécessaire pour que l'aléa soit réel et surtout voir quels sont les autres paramètres qui contrôlent les phénomènes de retrait gonflement comme par exemple la succion des racines des végétations arbustives, la déstructuration du sol par les sécheresses extrêmes, etc.

Finalement, cette méthode pourrait appuyer dans le futur des études telles que le suivi environnemental d'exploitations minières ou l'impact de l'agriculture intensive sur les sols, et de manière plus générale, sur le suivi de polluant dans les sols. En effet, le programme développé dans cette thèse est évolutif, et applicable à de nombreux autres domaines. De plus, du point de vue méthodologique, il sera plus simple à mettre en place pour l'analyse des divers polluants dont les spectres sont plus spécifiques que ceux des minéraux argileux

.

BIBLIOGRAPHIE

VI. BIBLIOGRAPHIE

Abrams, M., Hook, S., (1995). "Simulated ASTER data for geologic studies." Transactions on Geoscience and Remote Sensing **33**: 692-699.

Baumgardner, M. F., Silva, L.F., Biehl, L. L., Stoner, E.R. (1986). Reflectance Properties of Soils. Advances in Agronomy, Academic Press. **Volume 38:** 1-44.

Ben-Dor, E. and A. Banin (1995). "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., Kindel, B., Goetz, A.F.H. (2004). "Quality assessment of several methods to recover surface reflectance using synthetic imaging spectroscopy data." Remote Sensing of Environment **90**: 389 - 404.

Bertaux, J., Fröhlich, F., Ildefonse, P. (1998). "Multicomponent analysis of FTIR spectra : quantification of amorphous and crystallized mineral phases in synthetic and natural sediments." JOURNAL OF SEDIMENTARY RESEARCH **68**(3): 440-447.

Bigot, J., Van Bellegem (2009). "Log-density decomvolution by wavelet Thresholding." Scandinavian Journal of Statistics **36**(4): 749-763.

Bilgili, A. V., Hively, W.D, van Es, H., Reeves, J., Gaston, L. (2008). Wavelet analysis of soil reflectance for the characteirzation of soil properties. Proceedings of the Southern Regional Cooperative Soil Survey Conferences, Gainesville, Florida.

Bish, D. L., Post, J. E. (1993). "Quantitative mineralogical analysis using Rietveld full-pattern fitting method." The American mineralogist **78**(9-10): 932-940.

Bishop, J., J. Madejova, et al. (2002). "The influence of structural Fe, Al and Mg on the infrared OH bands in spectra of dioctahedral smectites." Clay Minerals **37**(4): 607-616.

Blackburn, G. A. (2007). "Wavelet decomposition of hyperspectral data: a novel approach to quantifying pigment concentrations in vegetation." Int. J. Remote Sens. **28**(12): 2831-2855.

Boillot, G., Capdevila, R., (1977). "The Pyrenees: Subduction and collision ? ." Earth Planet. Sci. Lett. **35**: 151-160.

Boillot, T. (1984). Les marges continentales actuelles et fossiles autour de la France. Paris, ed. Masson.

Bouchet, A., Meunier, A., Sardini, P. (2000). Minéraux argileux: structure cristalline, identification par diffraction des rayons X. Pau.

Bouroullec I., D. G., Plat E., Nowicki L., Ghyselinck M. (2009). Cartographie de l'aléa retrait-gonflement des sols argileux dans le département de l'Aveyron. BRGM/RP57409-FR**:** 124p., 47 ill., 5 ann., 3cartes h.-t.

Brindley, C. W., Brown, C., (1980). Crystal Structures of Clay Minerals and their X-ray Identification. London.

Brown, D. J., Sherpherd, K.D., Walsh, M.G., Dewayne Mays, M., Reinsch, T.G. (2006). "Global soil characterization with VNIR diffuse reflectance spectroscopy." Geoderma **132**: 273 - 290.

Cai, T., Zhang, D., Ben-Amotz, D. (2001). "Enhanced chemical classification of Raman Images Using Wavelet Transformation." Applied spectroscopy **55**: 1124-1130.

Capdeville, J.-P., Chalard, P. Jarry, M., Millet, D., O'YL, W. (1997). "« Le gisement d'En Jacca -La Sauvegarde à Colomiers (Haute-Garonne) : nouvelles données. »." Paléo **9**: 69-99.

Chalard, P., Jarry, M., Capdeville, J.P., Colonge, D., Lacombe, S.,Servelle, C. (1997 a). « Lotissement du Château à Fonsorbes (Haute-garonne). DFS d'évaluation archéologique ». S. M.-P. AFAN-GSO, UMR 5608, Toulouse., SRA Midi-Pyrénées, UMR 5608, Toulouse.**:** 88.

Chalard, P., Jarry, M., Capdeville, J.P., O'YL, W., Servelle, C. (1997 b). « En Jacca – La Sauvegarde. Colomiers (Haute-Garonne). » DFS de sauvetage urgent, AFAN-GSO, , AFAN-GSO, SRA Midi-Pyrénées, UMR 5608, Toulouse. **:** 88.

Chang, C.-W., D. A. Laird, et al. (2001). "Near-Infrared Reflectance Spectroscopy-Principal Components Regression Analyses of Soil Properties." Soil Sci. Soc. Am. J. **65**(2): 480-490.

Chassagneux D., S. L., Mouroux P., avec la coll. de Ducreux G.H. (1995). "cartographie de l'aléa retrait gonflement des sols (sècheresse/ pluie) dans la région de Manosque (Alpes de Haute-Provence). Echelle communale et départementale. Approche méthodologique." Rapport BRGM 38695: 90.

Choukroune, P. (1973). "La brèche de Lherz dite "d'explosion liée à la mise en place des lherzolites" est une brèche sédimentaire d'âge cénozoïque." C.R. Acad. Sci Paris **277**: 2621-2624.

Choukroune, P., Mattauer, M., (1978). "Tectonique des plaques et Pyrénées: sur le fonctionnement de la faille transformante nord-pyrénéenne; comparaison avec des modèles actuels." Bull. Soc. géol. Fr. **7**(20): 698-700.

Clark, R. N., and T. L. Roush (1984). "Reflectance Spectroscopy: Quantitative Analysis Techniques for Remote Sensing Applications." Journal of Geophysical Research **89**(B7): 6329-6340.

Clenet, H. (2009). Télédétection hyperspectrale:minéralogie et pétrologie, Application au volcan Syrtis Major (MARS) et à l'ophiolite d'Oman. Dynamique terrestre et planétaire. Toulouse, Université de Toulouse. **Thèse:** 367.

Cnockaert, L., Schoentgen, J., Ozsancak, Auzou, P., Grenez, F. (Aug 2007). Effect of Intensive Voice Therapy on Vocal Tremor for Parkinson Speakers. Proceedings of the 8th Annual Conference of the International Speech Communication Association, Anvers, Belgique, Interspeech 2007.

Crowley, J. K., Vergo, N. (1988). "Near-infrared reflectance spectra of mixtures of kaolin-group minerals: use in clay mineral studies." Clays and Clay Minerals **36**(4): 310-316.

Daignières, M., Gallart, J., Banda, E. (1981). "Lateral variation of the crust in the North-Pyrenean Zone." Annales Géophysicae **37**: 435-456.

Daignières, M., Gallart, J., Banda, E., Him, A., (1982). "Implications of the seismic structure for the orogenic evolution of the Pyrenees range." Earth Planet. Sci. Lett. **57**: 88-100.

Darrozes, J., Gaillot, P., Courjault-Radé, P. (1998). "2D propagation of a sequence of aftershocks combining anisotropic wavelet transform en GIS." Phys Chem Earth **3**: 303-308.

De La Fuente, S., Cuadros, J., Linares, J. (2000). "Quantification of mixed-layer illite-smectite in glass matrices by Fourier-Transform infrared spectroscopy". Clays and Clay Minerals **48**(2): 299_ 303.

Debroas, E.-J. (1987). "Modèle de bassin triangulaire à l'intersection de décrochements divergents pour le fossé albo-cénomanien de la Ballongue (zone nord-pyrénéenne, France)." Bull. Soc. géol. Fr. **8**(5): 887-898.

Delpont G., B. A., Ghyselinck M., (2006). Cartographie de l'aléa retrait-gonflement des argiles dans le département du Lot. BRGM/RP-55122-FR**:** 107 p., 22 ill., 5 ann., 3 cartes h.-t.

Delpont G., R. C., Vincent M., (2002). Cartographie de l'aléa retrait-gonflement des argiles dans le département de la Haute Garonne. BRGM/RP-51894-FR**:** 84p., 13 tab., 11 fig., 3 ann., 3 cartes h-t.

Delpont G., R. C., Vincent M., (2002). Cartographie de l'aléa retrait-gonflement des argiles dans le département de Tarn et Garonne. BRGM/RP-51893-FR**:** 84p., 13 tab., 14 fig.

Delpont G., S. E., Ghyselinck M., (2005). Cartographie de l'aléa retrait-gonflement des argiles dans le département du Tarn. BRGM/RP-RP-53531-FR**:** 113 p., 22 ill., 6 ann., 3 carte h.-t.

Dercourt, J. (1997). Géologie et géodynamique de la France Outre-mer et européenne. Paris.

Diou, C. (2000). Contribution à l'intégration sur silicium de la transformée en ondelettes, Application au traitement d'images. Electronique, Optronique et Systèmes. Montpellier Montpellier II _Sciences et Techniques du languedoc. **Doctorat:** 178.

Enjalabert, H. (1960). Les pays aquitain. Le modelé et les sols.

Farmer, V. C. (1974). The Infrared Spectra of Minerals. London, Monograph 4. Mineralogical Society.

Gaillot, P., Darrozes, J., Bouchez, J-L. (1999). "wavelet transform: a future of rock fabric analysis?" Journal of structural geology **21**: 1615-1621.

Galforf, G. L., Mustard, J.F., Melillo, J., Gendrin, A., Cerri, C.C., Cerri, C.E.P. (2008). "Wavelet analysis of MODIS time series to detect expansion and intensification of row-crop agriculture in Brazil." Remote Sensing of Environment **112**: 576-587.

Ge, Y., Morgan, C.L.S, Thomasson, J.A., Waiser, T. (2007). "A new perspective to near-infrared reflectance spectroscopy: a Wavelet Approach." American Society of Agricultural and Biological Engineers **50**(1): 303_ 311.

Ge, Y., Thomasson, J.A. (2006). "Wavelet incorporated spectral analysis for soil property determination." American Society of Agricultural and Biological Engineers **49**(4): 1193-1201.

Gendrin, A., Langevin, Y., Bibring, J.-P., Forni, O. (2006). "A new method to invastigate hyperspectral image cubes: An application of the wavelet transform." Journal of geophysical research **111**.

Gholami, A., Siahkoohi, H.R., (2009). "A two-step wavelet based regularization for linear inversion of geophysical data." Geophysical Prospecting **57**(5): 847-862.

Goetz, A. F. H., Chabrillat, S., Lu, Z. (2001). "Field reflectance spectrometry for detection of swelling clays at construction sites." Field Analytical Chemestry and Technology **5**(3): 143 - 155.

Gorini, C., Vialard, P., Déramond, J. (1991). "Modèle d'inversion négative: la tectonique extensive post-nappe du fossé de Narbonne-Sigean (Corbières, Sud de la France)." C.R. Acad. Sci. **312**: 1013-1019.

Grossman, A., Morlet, J. (1984). "Decomposition of Hardy functions into square integrable wavelets of constant shape." SIAM Journal of Mathematical Analysis **15**: 723-726.

Hotelling, H. (1936). "RELATIONS BETWEEN TWO SETS OF VARIATES." Biometrika **28**(3-4): 321-377.

Hubert, F. (9 déc. 2008). Modélisation des diffractogrammes de minéraux argileux en assemblages complexes dans deux sols de climat tempere. Implications mineralogique et prédologiques. Terre solide et enveloppe superficielle. Poitiers, Faculté des Sciences Fondamentales et Appliquées. Université de Poitiers**:** 223 p.

Hubschman, J. (1975 a). ""Les terrasses récentes de la Garonne et leur évolution. In: Le piemont quaternaire de la Garonne (sud-est du bassin Aquitain, piémont central nordpyrénéen."" Bulletin de l'Association Française pour l'Etude du Quaternaire **44-45**: 37-147.

hubschman, J. (1975 b). ""L'évolution des nappes alluviales antérissiennes de la Garonne dans l'avant-pays molassique. In: le piémont quaternaire de la Garonne (sud-est du bassin Aquitain, piémont central nord-pyrénéen."" Bulletin de l'Association Française pour l'Etude du Quaternaire **44-45**: 148-160.

Hunt, G. R. (1977). "Spectral signatures of particulate minerals in the visible and near infrared." Geophysics **42**(3): 501 -513.

Icole, M. (1980). "" L'altération au sein des alluvions quaternaires du piémont occidental nord-pyrénéen: éléments de corrélation." in: CHALINE (J.) dir.- Problèmes de stratigraphie quaternaire en France et dans les pays limitrophes. Actes de la Tabble Ronde de Dijon." Bulletin de l'Association Française pour l'Etude du Quaternaire **1**: 312-313.

Icole, M., Rieucau, L. (1982 a). ""les actions fluviatiles dans le bassin de la Garonne et ses affluents." In: Préhistoire de Midi-Pyrénées. Extraits de La préhistoire Française." Ed. Du CNRS: 5-8.

Icole, M., Rieucau, L. (1982 b). ""Les limons quaternaires et les dépôts de pente dans le bassin de la Garonne et ses affluents." In: Préhistoire de Midi-Pyrénées. Extraits de la Préhistoire Française." Ed. Du CNRS: 5-8.

Jahn, B. R., Linker, R., Upadhyaya,S.K., Shaviv, A., Slaughter, D.C., Shmulevich,I. (2006). "Mid-Infrared Spectroscopic Determination of Soil Nitrate Content." Biosystems Engineering **94**(4): 505-515.

Janik, L. J., Merry, R.H., Forrester, S.T., Lanyon, D.M., Rawson, A. (mars-avril 2007). "Rapid prediction of soil water retention using mid infrared spectroscopy." Soil Science Society of America **71**(2): 507 - 514.

Janik, L. J., Skjemstad,J. O., Merry,R. H. (1998). "Can mid infrared diffuse reflectance analysis replace soil extractions?" Australian Journal of Experimental Agriculture **38**(7): 681-696.

Jouan-Rimbaud, D., Walczak, B., Poppi, R. J., de Noord, O. E., Massart, D. L. (1997). "Application of Wavelet Transform To Extract the Relevant Component from Spectral Data for Multivariate Calibration." Analytical Chemistry **69**(21): 4317-4323.

Joussein, E., Petit, S., Decarreau A. (2001). " A new method to determine the ratio of clay minerals in mixtures by IR spectroscopy." C.R. Acad. Paris, Sciences de la Terre et des Planètes / Earth and Planetary Sciences **332**: 83 - 89.

Kariuki, P. C. (2004). Spectroscopy and swelling soil. Department of Applied Earth Sciences. Delft, University of Technology**:** 160 p.

King P.L., R. M. S., Swayze G.A. (EDITORS) (2004). Infrared Spectroscopy in Geochemistry, Exploration Geochemistry, and Remote Sensing., Mineralogical Association of Canada.

Labat, D., Ababou, R., Mangin, A. (2002). "Multiresolution cross-analysis of rainfall rates and karstic spring runoffs." Comptes rendus Geosciences **334**(8): 551-556.

Lagacherie, P., F. Baret, et al. (2008). "Estimation of soil clay and calcium carbonate using laboratory, field and airborne hyperspectral measurements." Remote Sensing of Environment **112**(3): 825-835.

Le Pichon, X., Sibuet, J.-C., Bonnin, J. (1970). "La faille nord-pyrénéenne, faille transformante liée à l'ouverture du Golfe de Gascogne." C.R. Acad. Sci Paris **271**: 1941-1944.

Leger, C.-L. (18 septembre 1997). Etude expérimentale des solutions solides octaédriques ternaires dans les smectites. Poitiers**:** 156.

Linker, R. (2008). Computer and computing technologies in agriculture. Soil classification via Mid-Infrared spectroscopy. S. Boston.

Linker, R., Kenny A., Shaviv, A., Singher, L., Shmulevich, I. (2004). "Fourirer transform Infrared-attenuated total reflection nitrate determination of soil pastes using principal cimponent regression, partial least squares, and cross-correlation." Applied spectroscopy **58**(5): 516-520.

MacArthur, A., MacLellan, C., Malthus, T.J. (2006). What does a spectroradiometer see? Remote Sensing and Photogrammetry Society, Cambridge, UK.

Madéjova, J. (2003). "FTIR techniques in clay mineral studies." Vibrational Spectroscopy **31**: 1 - 10.

Madejova, J., Keckés,J. , Palkova, H., Komadel, P. (2002). "Identification of components in smectite / kaolinite mixtures." Clay Minerals **37**: 377 _ 388.

Madéjova, J., Komadel, P. (2001). "Baseline studies of the clay minerals society source clays: infrared methods." Clays and Clay Minerals **49**(5): 410 - 432.

Maire E., D. J., Courjault-Radé P., Christophoul F., Guerrero N. ((submitted)). "« Assessment of the holocen fluvial erosion in the Garonne river basin using mathematical morphology on a digital elevation model"." Geodyn. acta.

Mallat, S. (2001). Une exploration des signaux en ondelettes, Ecole Polytechnique.

Marcoen, J. M., Tessier, D., Thorez, J., Monjoie, A. Schroeder,Ch. (2001). "manuel relatif aux matières naturelles pour barrières argileuses ouvragées pour C.E.T et réhabilitaion de dépôtoirs en Région Wallonne." Dir. Gén. Ress. Natur. Environ., Jambes.

Matteson, A., Herron, M.M. (1993). Quantitative mineral analysis by Fourier transform infrared spectroscopy. SCA.

Meyer, Y., Jaffard, S., Rioul, O. (1987). "L'analyse en Ondelette." Pour la science (French edition of Scientific American) **11**: 28-37.

Mittermayr, C. R., Tan, H.W., Brown S.D (2001). "Robust calibration with respect to background variation." Applied spectroscopy **55**(7): 827-833.

Moore, D. M., Reynolds, R.C. (1989). "X-Ray Diffraction and the identification and analysis of Clay Minerals." Oxford University Press: 378.

Norie A., C. J. p., Vincent M. (2001). Cartographie de l'aléa retrait-gonflement des argiles dans le département du Gers. BRGM/RP-51251-FR.

Nowamooz, H. (19 déc. 2007). Retrait/Gonflement des sols argileux compactes et naturels. Génie Civil - Hydrosystèmes - Géotechnique, Institut National Polytechnique de Lorraine**:** 223.

Ouillon, G., Castaing C.,Sornette, D. (1996). "Hierachical geometry of faulting." Journal of Geophysical Research **101**(B3): 5477-5487.

Ouillon, G., Sornette, D., Castaing C., (1995). "Organisation of joints and faults from 1 cm to 100 km scales revealed by optimized anisotropic wavelet coefficient method and multifractal analysis." Nonlinear Processes in Geophysics **2**: 158-177.

Petit, S., Madejova, J., Decareau, A., Martin, F. (1999). "Characterization of octahedral substitutions in kaolinites using near infrared spectroscopy." Clays and Clay Minerals **47**(1): 103_108.

Petit, S., Righi, D., Madejova, J., Decareau, A., (1998). "Layer charge estimation of smectites using infrared spectroscopy." Clay Minerals **33**: 579 - 591.

Petit, S., Robert, J-L., Decrreau, A., Besson, G., Grauby, O. and Martin, F. (1995). "apport des méthodes spectroscopiques à la caractérisation des phyllosilicates 2 : 1. [*contribution of spectroscopic methods to 2 : 1 clay characterization*]." Bull. Centres Rech. Explor.-Prod. elf Aquitaine **19**(1): 119- 147.

Pironon, J., Pelletier, M., De Donato, P., Mosser-Ruck, R. (2003). "Characterization of smectite and illite by FTIR spectroscopy of interlayer NH_4^+ cations." Clay Minerals **38**: 201-211.

Post, J. L. and P. N. Noble (1993). "The near-infrared combination band frequencies of dioctahedral smectites, micas, and illites." Clays and Clay Minerals **41**(6): 639-644.

Reeves, J., G. McCarty, et al. (2002). "The potential of diffuse reflectance spectroscopy for the determination of carbon inventories in soils." Environmental Pollution **116**(Supplement 1): S277-S284.

Reeves, J. B., G. W. McCarty, et al. (2001). "Mid-infrared Diffuse Reflectance Spectroscopy for the Quantitative Analysis of Agricultural Soils." Journal of Agricultural and Food Chemistry **49**(2): 766-772.

Ricker, N. (1953). "The form and laws of propagation of seismic wavelets." Geophysics **18**: 10-40.

Roy, R. (2007). Influence des altérations de surface naturelle sur la reconnaissance des roches par télédétection V-IR: application à la cartographie de l'ophiolite d'Oman et au programme d'étude des nouvelles AOC des Muscadet de la région nataise. sciences de la terre, télédétection. Nantes, Université de Nantes _ faculté des sciences et des techniques**:** 480.

Savitzky, A. and M. J. E. Golay (1964). "Smoothing and Differentiation of Data by Simplified Least Squares Procedures." Analytical Chemistry **36**(8): 1627-1639.

Savvides, A., Corstanje, R., Baxter, S.J., Rawlins, B.G., Lark, R.M., (2010). "The relationship between diffuse spectral reflectance of the soil and its cation exchange capacity is scale-dependent " Geoderma **154**: 353-358.

Shepherd, K. D. and M. G. Walsh (2002). "Development of Reflectance Spectral Libraries for Characterization of Soil Properties." Soil Sci. Soc. Am. J. **66**(3): 988-998.

Sibuet, J.-C., Le Pichon, X. (1971). Structure gravimétrique du golfe de Gascogne et le fossé marginal nord-espagnol, in Histoire Structurale du golfe de Gascogne, "Publications de l'institut Français du Pétrole", Collection Colloques et Séminaires.

Sorensen, L. K., Dalsgaard, S. (2005). "Determination of clay ond other soil properties by near infrared spectroscopy." Soil Science Society of America **69**: 159- 167.

Souque, C. (2002). Magnétisme structural dans les chaines de chevauchement-plissement: développements analytiques et exemples d'utilisation dans les corbières, Université de Cergy-Pontoise**:** 350.

Specht, M. (1989). Tectonique de chevauchement le long du profil ECORS-Pyrénées: Un modèle d'évolution de prisme d'accrétion continental. geologie. Brest, Université de Brest**:** 353.

Summers, D., M. Lewis, et al. (2009). "Visible near-infrared reflectance spectroscopy as a predictive indicator of soil properties." Ecological Indicators **In Press, Corrected Proof**.

Sunshine, J. M., Pieters, C.M., Pratt, S.F. (1990). "Deconvolution of mineral absorption bands: an improved approach." Journal of geophysical research **95**(B5): 6955-6966.

Taubaso, C., Dos Santos Afonso, M., Torres Sánchez, R. M. (2004). "Modelling soil surface charge density using mineral composition." Geoderma **121**(1-2): 123-133.

Thorez, J. (1975). Phyllosilicates and clay minerals: a laboratory handbook for their x-ray diffraction analysis. Dison, Lelotte.

Thorez, J. (1976.). Practical identification of clay minerals. A handbook for teachers and students in

clay mineralogy. . Belgique., Dison.

Torrence, C. and G. P. Compo (1998). "A Practical Guide to Wavelet Analysis." Bulletin of the American Meteorological Society **79**(1): 61-78.

Torrence, C., Webster, P.J. (1999). "Interdecadal changes in the ENSO-Monsoon system." Journal of Climate **12**: 2679-2690.

Udelhoven, T., Emmerling, C.,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**(2): 219-329.

Vincent M., C. R., Fleureau J.-M., Cu Y.J., Jacquard C., Kazmierczak J.-B., Masrouri F., Tessier D., Alimi-Ichola I., Magnan J.-P., Blanchard M., Fabre D., Pantet A., Audiguier M., Plat E., Souli H., Taibi S., Tang A.-M., Morlock C., Maison T., Mrad M., Bréda N., Djeran-Maigre I., Duc M., Soubeyroux J.-M., Denis A., Proust D., Geremew Z., le Roy S., Dumont M., Hemmati S., Nowamooz H., Coquet Y., Pothier C., Makki L., Chrétien M., Fontaine C. (2009). Rapport de synthèse final du projet ARGIC (Analyse du retrait - gonflement et de ses Incidences sur les Constructions). Projet ANR-05-PRGCU-005. l. B. Rapport BRGM/RP-57011-FR en partenariat entre le Centre de Géosciences, le LMSSMat, le CERMES, Fondasol, l'INERIS, le LAEGO, l'INRA, le LGCIE, le LCPC, Météo-France, le GHYMAC et l'université de Poitiers**:** 92 p., 29 ill., 6 tabl., 39 ann.

Viscarra Rossel, R. A. (2008). "ParLes: Software for chemometric analysis of spectroscopic data." Chemometrics and intelligent laboratory systems **90**: 72-83.

Viscarra Rossel, R. A., S. R. Cattle, et al. (2009). "In situ measurements of soil colour, mineral composition and clay content by vis-NIR spectroscopy." Geoderma **150**(3-4): 253-266.

Viscarra rossel, R. A., McGlynn, R.N., McBratney, A.b. (2006). "determining the composition of mineral-organic mixes using UV-vis-NIR diffuse reflectance spectroscopy." Geoderma **137**: 70-82.

Viscarra Rossel, R. A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J., Skjemstad, J.O. (2006). "Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties." Geoderma **131**: 59 - 75.

Winnock, E. (1974). "Le bassin d'Aquitaine", In Géologie de la France. Paris.

Wold, S., M. Sjöström, et al. (2001). "PLS-regression: a basic tool of chemometrics." Chemometrics and intelligent laboratory systems **58**(2): 109-130.

Yitagesu, F. A., F. van der Meer, et al. (2009). "Quantifying engineering parameters of expansive soils from their reflectance spectra." Engineering Geology **105**(3-4): 151-160.

Zhang, X., Younan, N.H., O'Hara, C.G. (2005). "Wavelet Domain Statistical Hyperspectral Soil Texture Classification." IEEE Transactions on geoscience and remote sensing **43**(3): 615-618.

Ziegler, P. A. (1988). evolution of the Artic-North-Atlantic and the Western Tethys, Association of Petroleum Geologists.

ANNEXES

VII. ANNEXE 1 : ANALYSES PAR DIFFRACTION DES RAYONS X DES MINERAUX ARGILEUX DE REFERENCE

Kaolinite Cleyrac

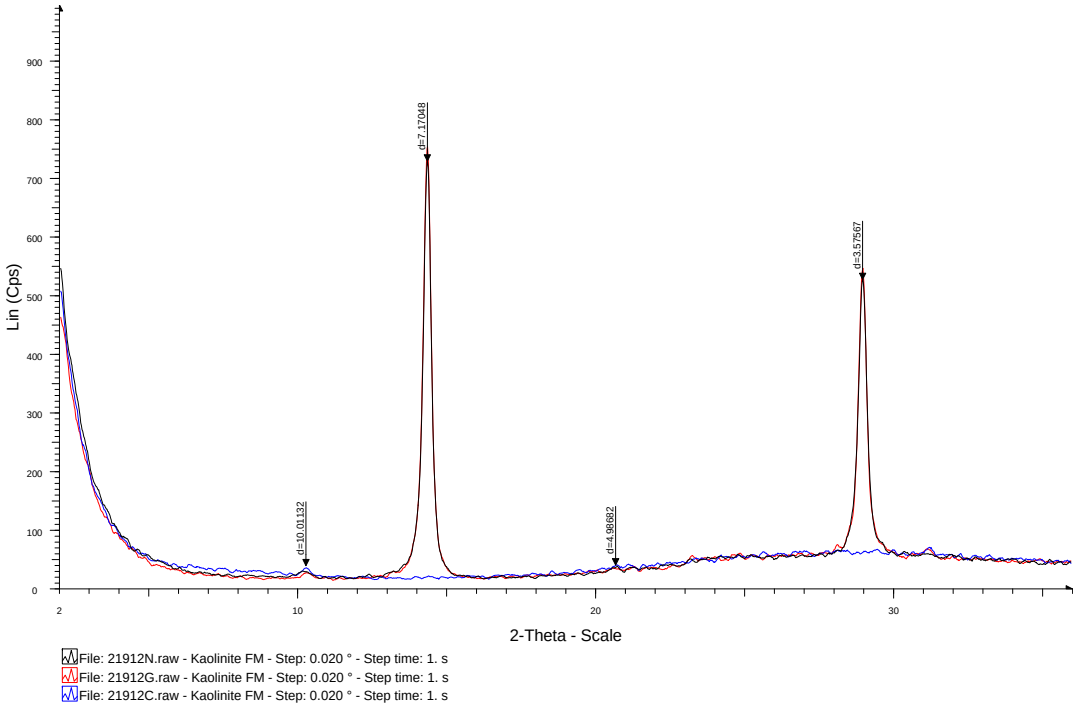
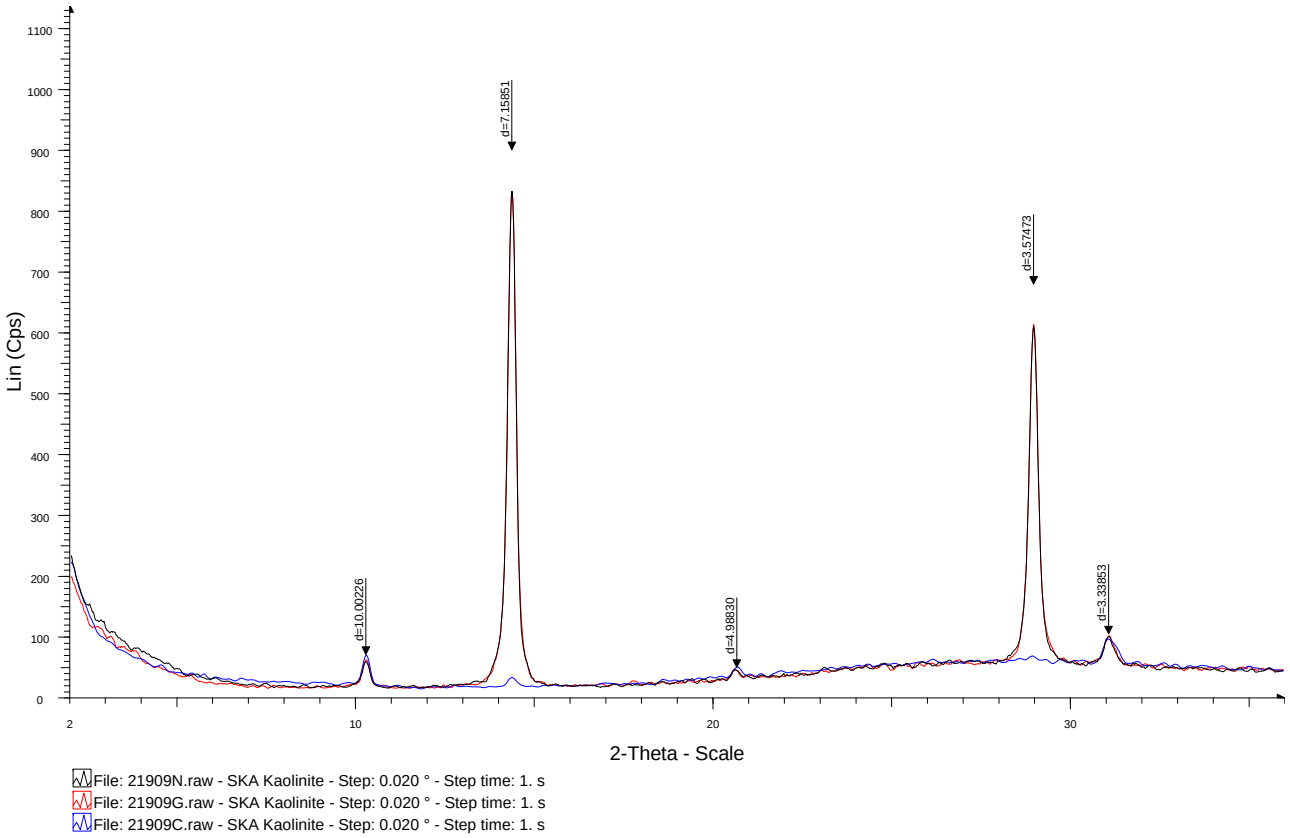


Figure VII-1 : Diffractogramme Normal, glycolé et chauffé de la kaolinite de Cleyrac.

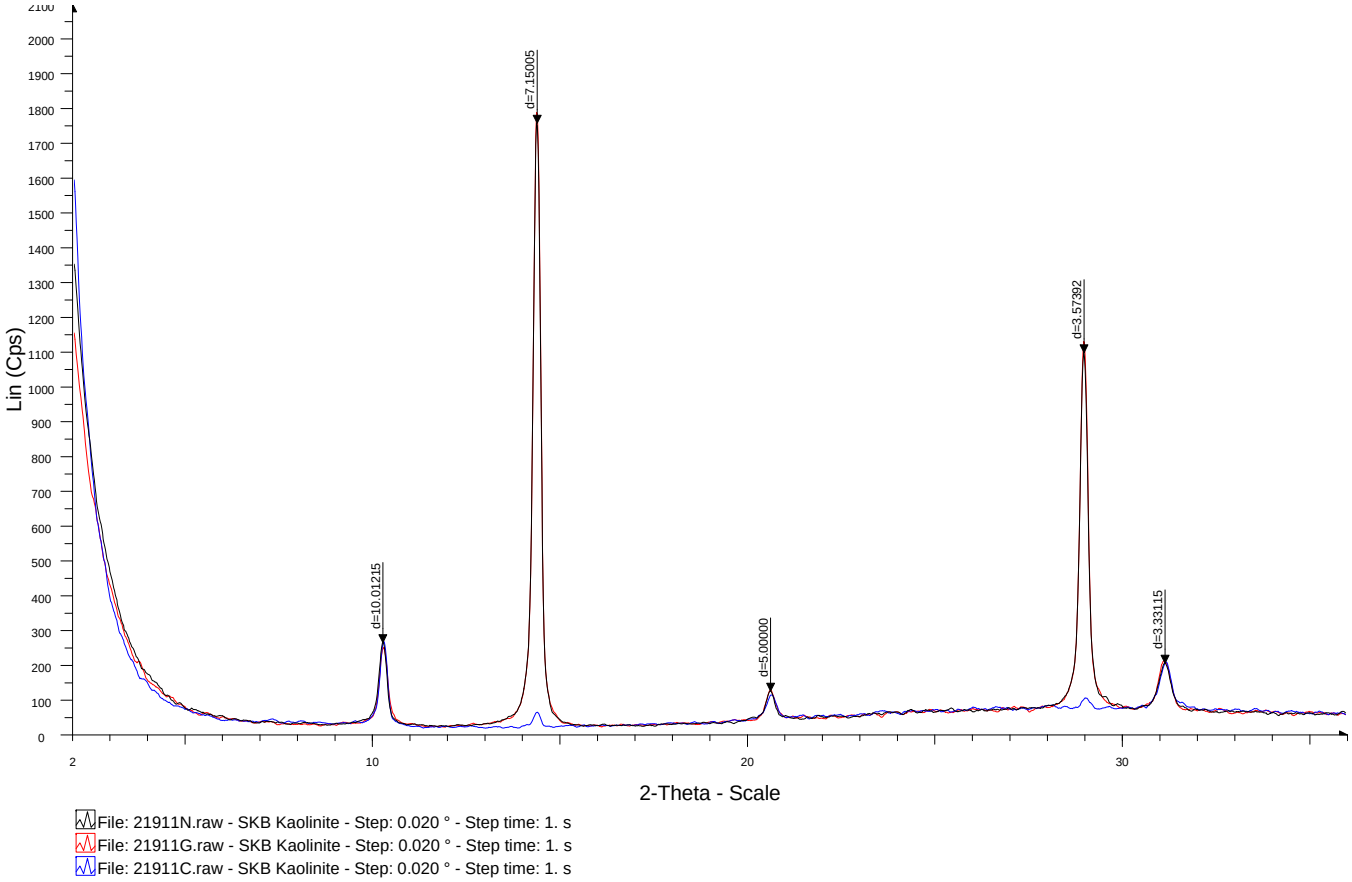
- quartz en traces (~ 1%)
- gibbsite probable en traces
- fraction phylliteuse très abondante et représentée par (sur base 100) :
    - illite et/ou micas : ~ 1%
    - kaolinite : ~ 99%

Kaolinite Ska



- quartz présent (~ 45%)
- fraction phylliteuse présente et représentée par (sur base 100) :
    - illite et/ou micas : ~ 3%
    - kaolinite : ~ 97%

Kaolinite Skb



- quartz en traces (~ 5%)
- fraction phylliteuse très abondante et représentée par (sur base 100) :
    - illite et/ou micas : ~ 11%
    - kaolinite : ~ 89%

Smectite ABM

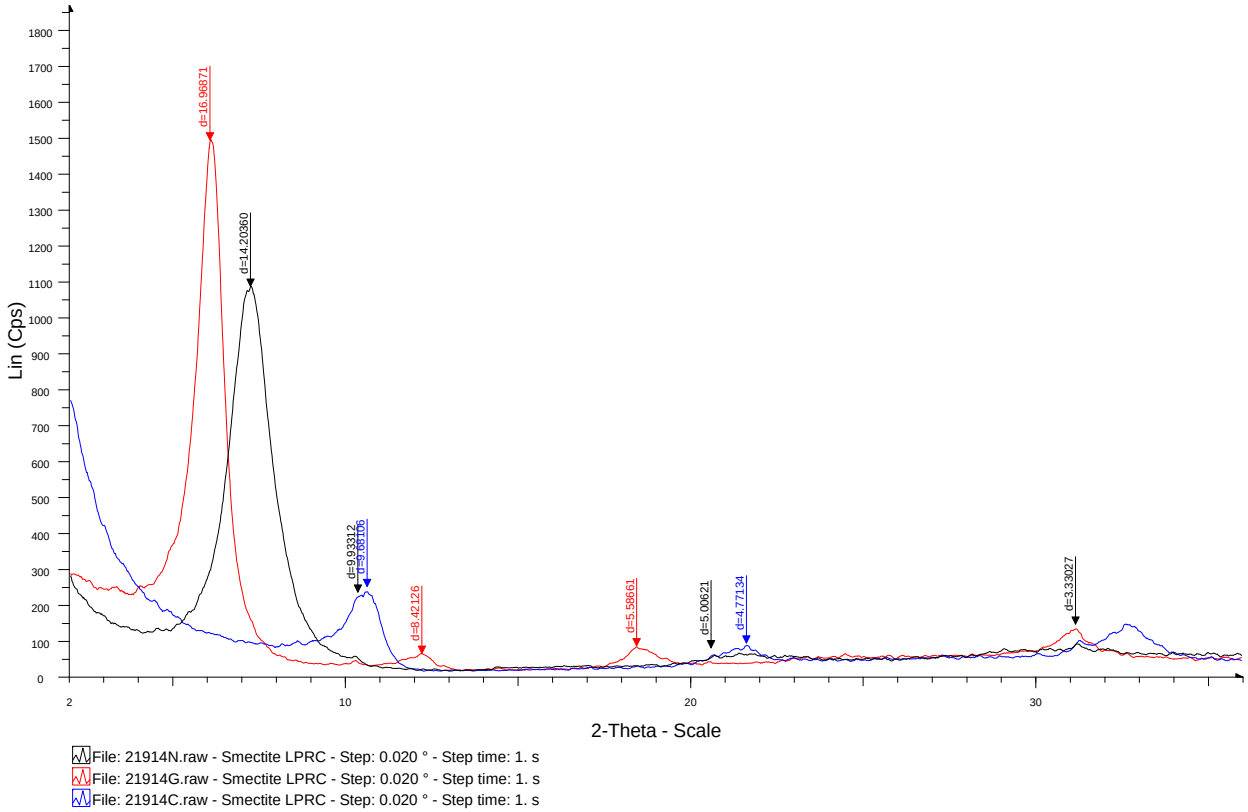


Figure VII-2 : Diffractogramme Normal, glycolé et chauffé de la smectite ABM

- quartz en traces (~ 2%)
- pyrite en traces
- microcline probable en traces
- fraction phylliteuse très abondante et représentée par (sur base 100) :
    - illite et/ou micas : < 1%
    - smectite : ~ 100%

30K8

A.B.M

ARGILES DU BASSIN MEDITERRANEEN

MONTMORILLONITES - SMECTITES - SEPIOLITES - KAOLINITES - ILLITES

FICHE TECHNIQUE ARGILE VERTE MONTMORILLONITE SARDE NATURELLEMENT PURE

Lot n° 20080301 P
Nom INCI : MONTMORILLONITE
N° CAS 1318-93-0
EINECS 215-288-5
Formule : $(Na, Ca)O_{0,33} (Al, Mg)_2 Si_4 O_{10} (OH)_2, nH_2O$

ELEMENTS MAJEURS (Moyenne)	Pourcentages
SIO2	58 < x < 62
AL2O3	15 < x < 18
Fe2O3	3 < x < 4
MgO	5 < x < 6
K2O	0,5 < x < 1
Na2O	0,1 < x < 0,3
TIO2	0,1 < x < 0,3
CaO	3 < x < 4
MnO	Traces
P2O5	Traces

Analyse Bactériologique moyenne	Quantité par g
Mésophiles viables	< 5 000
Pathogènes	absence
Levures moisissures :	< 80

METAUX LOURDS	ppm
ARSENIC	< 1
PLOMB	< 15
SELENIUM	< 1
MERCURE	< 1

Nature de l'argile : Silicate naturel, complexe d'aluminium, fer et magnésium .
Constituant majeur : Smectite (Montmorillonite), à plus de 97 %.
Aspect (séchée) : Poudre vert clair
Capacité d'échange cationique : Supérieur à 120 Meq par 100 grs
Pouvoir Adsorbant (Sulfate de strychnine) : Supérieur à 423 mg/g
Grît (Autre que Montmorillonite) : inférieur à 1,90 %
Perte au Feu (1000°C) : 10.60 %

Présentation :

Cette argile se présente sous forme de poudre : < 77 µm

Il est à noter que s'agissant d'une argile, donc d'un minéral naturel, l'analyse chimique varie d'un point à un autre de la carrière sans que le minéral lui même change en quoi que ce solt. Une argile est avant tout caractérisée par son spectre de diffraction et sa capacité d'échange cationique. Ces valeurs sont donc indicatives et peuvent varier à plus ou moins 10%.

Sarl ARGILES DU BASSIN MEDITERRANEEN au Capital de 7 622 €
Z.A. du Bas Jasson - 985, chemin du Puits de la Commune, F-83250 LA LONDE
Tél.: 04 94 66 62 94 - Fax : 04 94 66 67 94 - www.argiles-abm.com - E-mail : argiles@wanadoo.fr
Siret 412 481 970 000 17 - TVA Intracommunautaire : FR 44412481970 - APE 523E

Smectite Wyoming

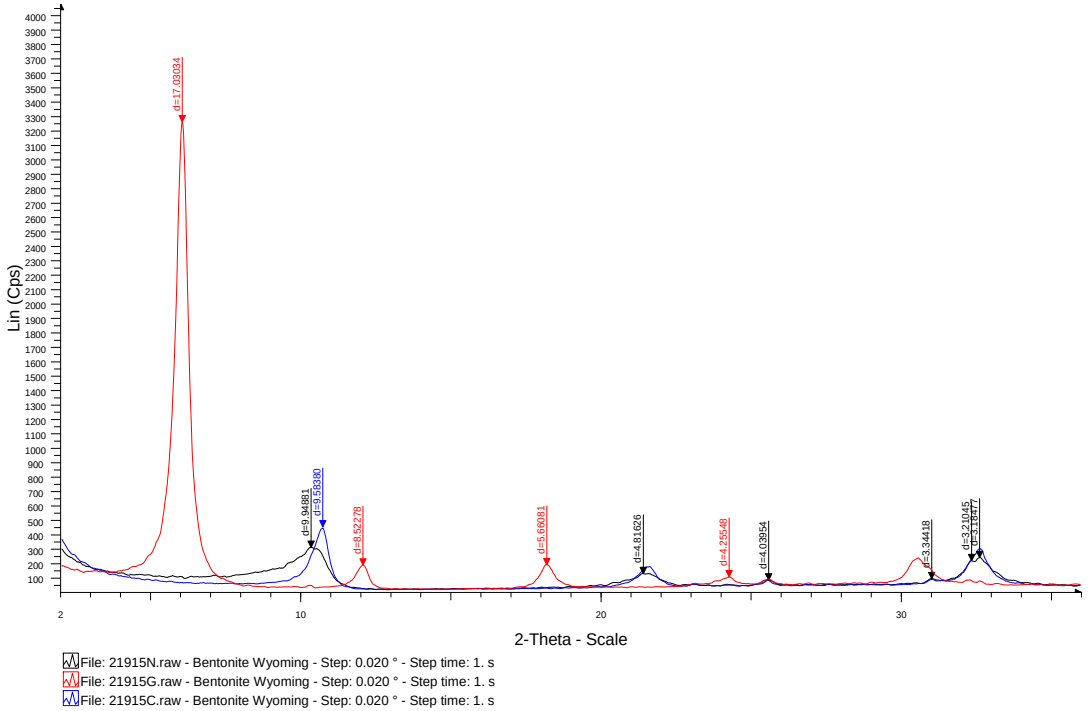
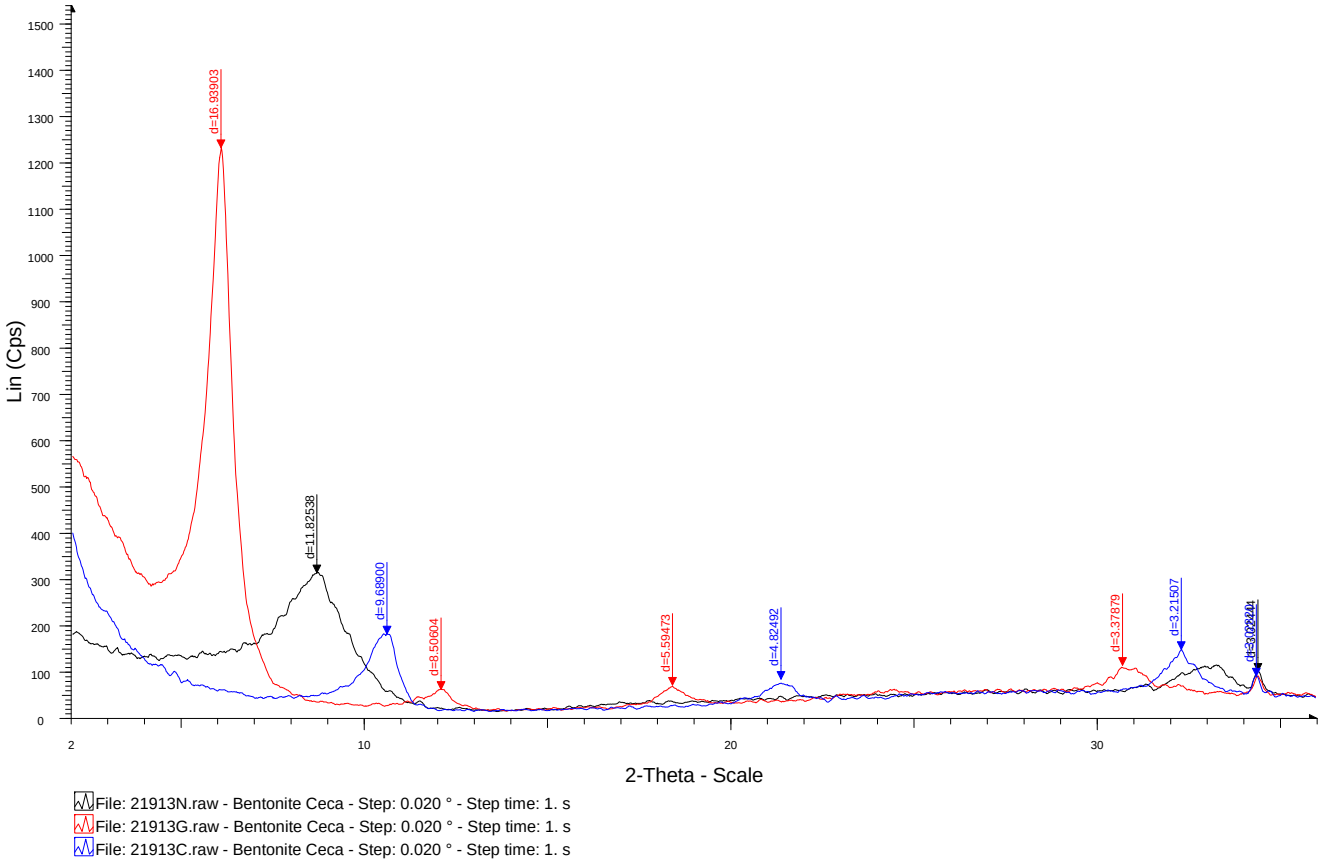


Figure3 : Diffractogramme Normal, glycolé et chauffé de la smectite Wyoming.

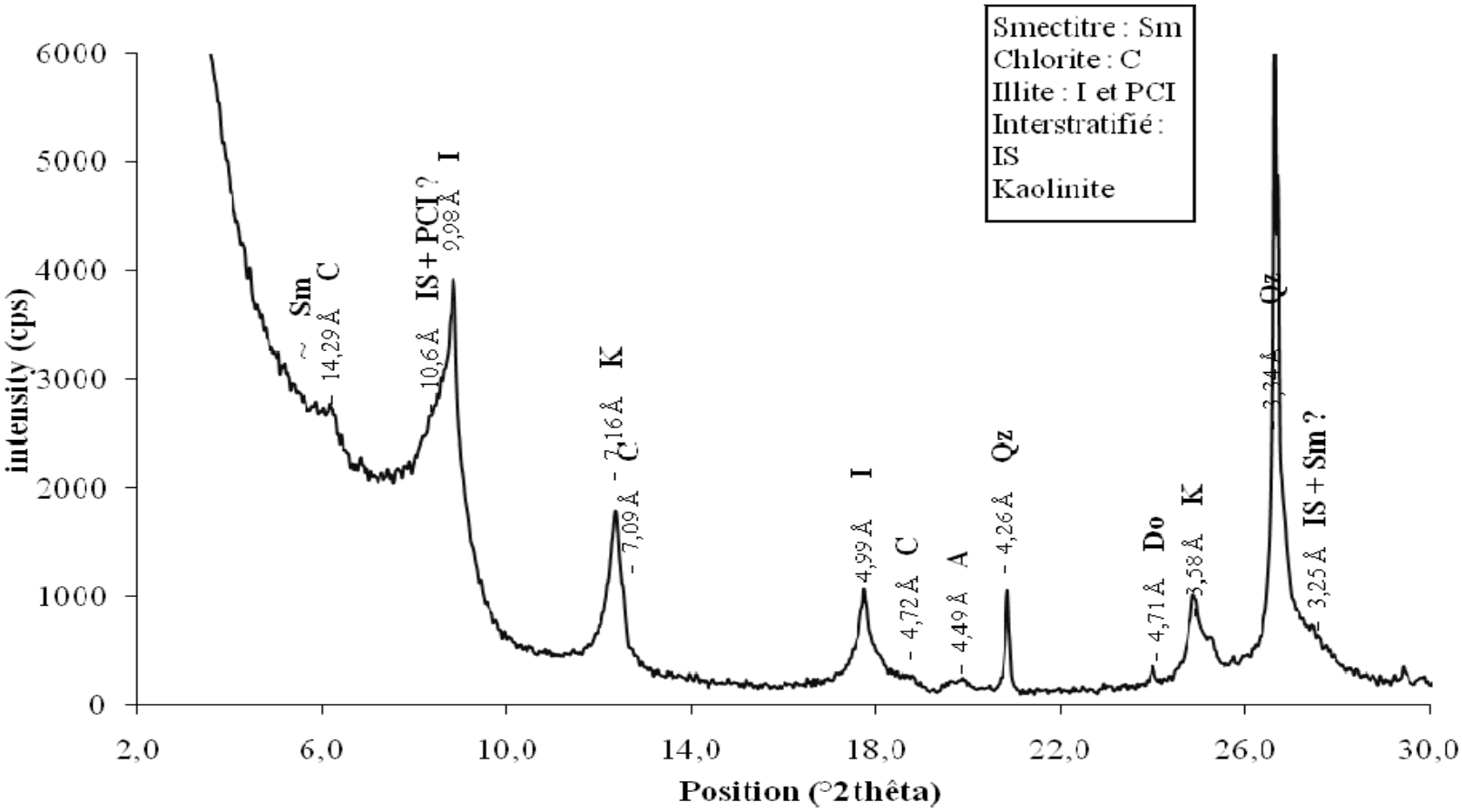
- quartz en traces (~ 2%)
- microcline en traces
- plagioclase (albite sur le diagramme) en traces
- cristobalite probable en traces
- goethite probable en traces
- fraction phylliteuse très abondante et représentée par (sur base 100) :
    - illite et/ou micas : < 1%
    - smectite : ~ 100%

Smectite CECA

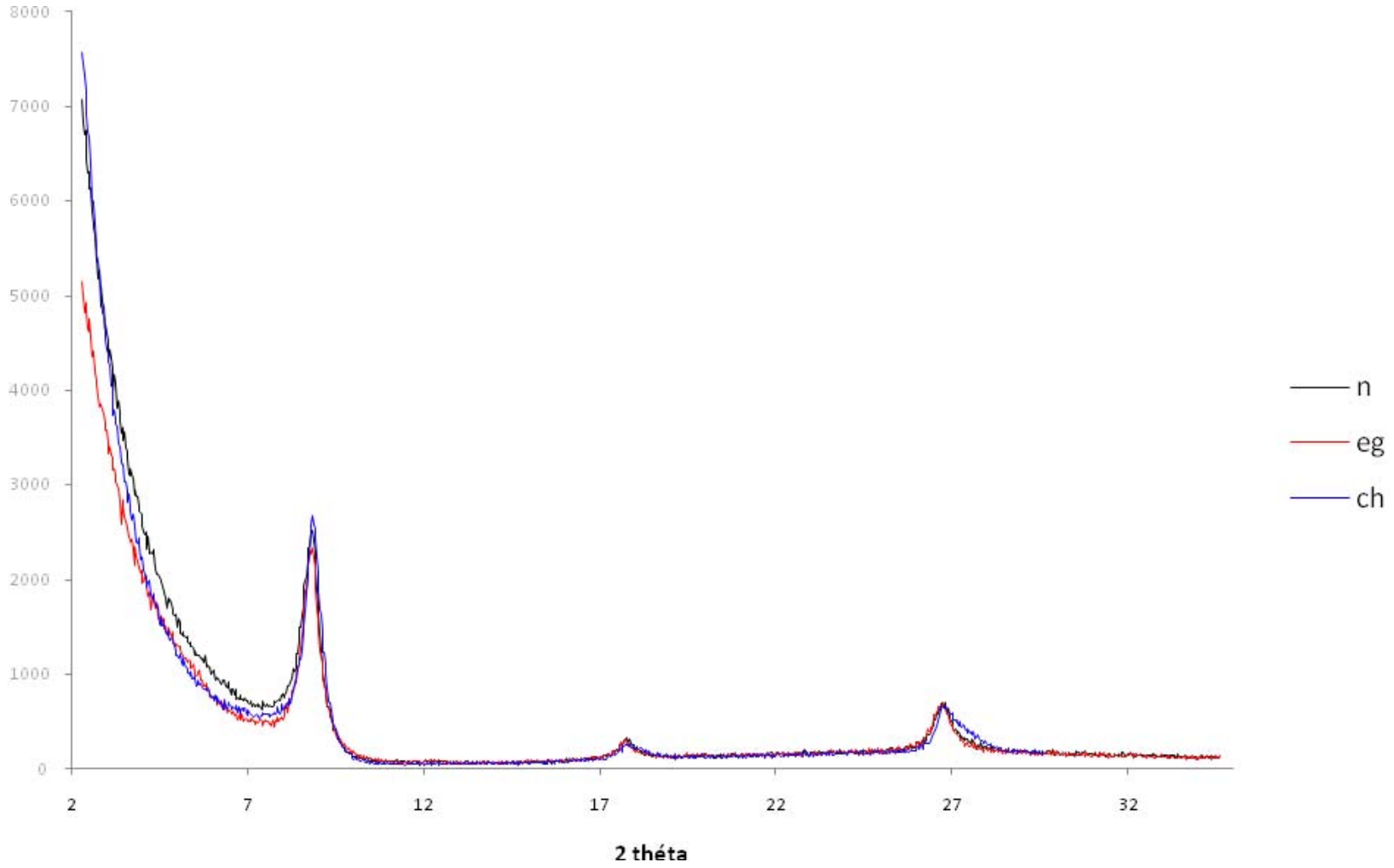


- quartz en traces (~ 2%)
- plagioclase (anorthite sur le diagramme) en traces
- zéolite (clinoptilolite sur le diagramme) en traces à faible
- pyrite en traces
- fraction phylliteuse abondante et représentée par (sur base 100) :
    - interstratifié smectite/illite ~ 100%

illite LRPC


Smectitre : Sm
Chlorite : C
Illite : I et PCI
Interstratifié :
IS
Kaolinite
intensity (cps)
Position (°2thêta)
6000
5000
4000
3000
2000
1000
0
2,0
6,0
10,0
14,0
18,0
22,0
26,0
30,0
~ Sm
14,29 Å C
10,6 Å IS + PCI ?
9,98 Å I
7,16 Å K
7,09 Å C
4,99 Å I
4,72 Å C
4,49 Å A
4,26 Å Qz
4,71 Å Do
3,58 Å K
3,34 Å Qz
3,25 Å IS + Sm ?

illite du Puy en Velay



VIII. ANNEXE 2 : DONNEES DE DRX ACQUISE SUR LES SOLS DE LA REGION MIDI-PYRENEES SUIVANT LES UNITES MORPHO PEDOLOGIQUES PRELEVEES.

VIII.1. LES SOLS DE LA MOLASSE

Gers 94

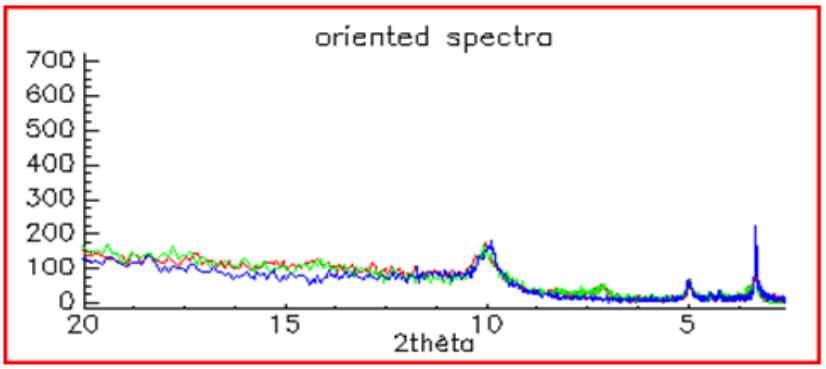


Figure VIII-1 : diffractogramme sur lame orientée (normal : rouge ; éthylène glycol : vert ; chauffé : bleu)

Semi-quantification sur DRX :
- illite : 80%
- kaolinite : 10%
- interstratifiés : 10%

Gers 95

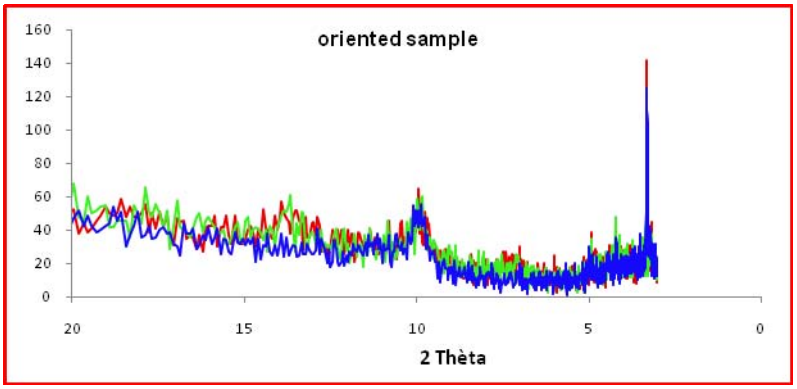


Figure VIII-2 : diffractogramme sur lame orientée (normal : rouge ; éthylène glycol : vert ; chauffé : bleu)

Semi-quantification sur DRX :
- illite : 55%
- kaolinite : 10%
- vermiculite: 35%

Gers 104

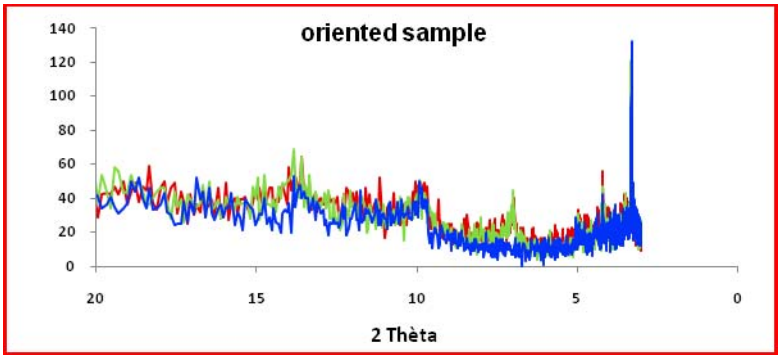


Figure VIII-3 : diffractogramme sur lame orientée (normal : rouge ; éthylène glycol : vert ; chauffé : bleu)

Semi-quantification sur DRX :
- illite : 40%
- kaolinite : 20%
- chlorite : 40%

Gers 99

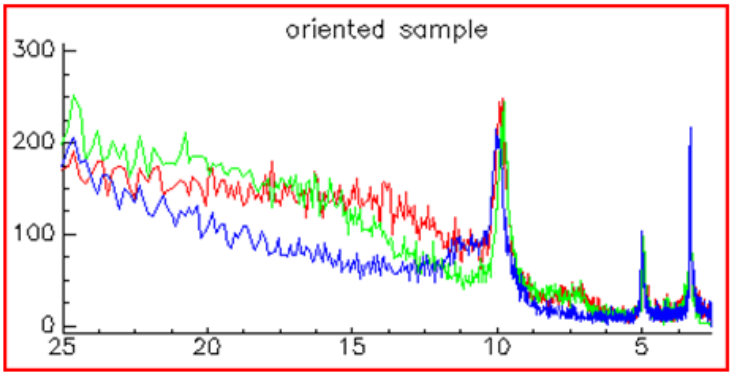


Figure VIII-4 : diffractogramme sur lame orientée (normal : rouge ; éthylène glycol : vert ; chauffé : bleu)

Semi-quantification sur DRX :
- illite : 80%
- kaolinite : 10%
- interstratifiés (C/M) : 10%

Gers 98

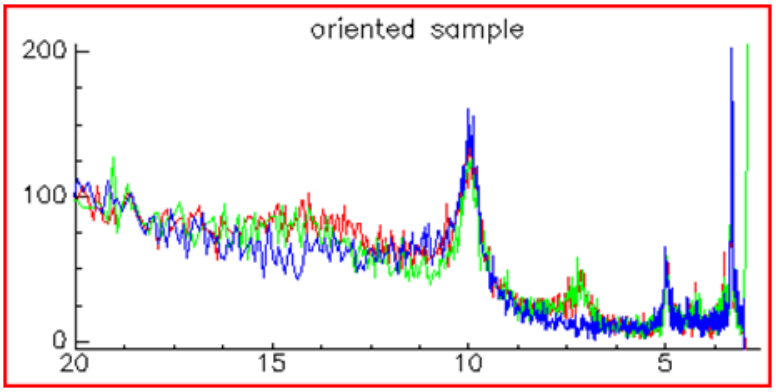


Figure VIII-5 : diffractogramme sur lame orientée (normal : rouge ; éthylène glycol : vert ; chauffé : bleu)

Semi-quantification sur DRX :
- illite : 75%
- kaolinite : 15%
- interstratifiés : 10%

Haute Garonne 12

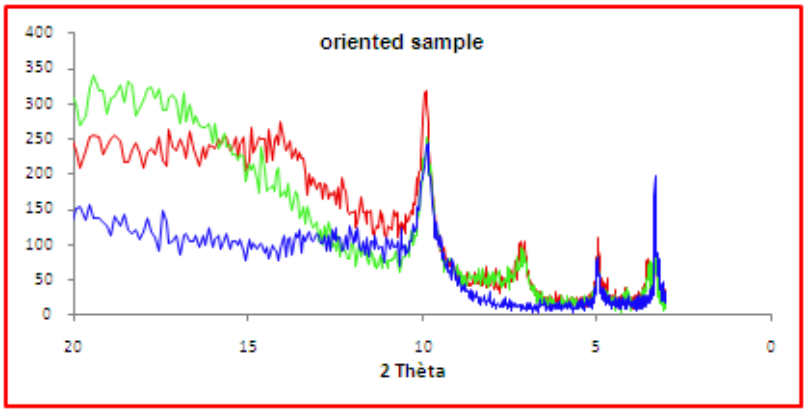


Figure VIII-6 : diffractogramme sur lame orientée (normal : rouge ; éthylène glycol : vert ; chauffé : bleu)

Semi-quantification sur DRX :
- illite : 65%
- kaolinite : 15%
- smectite : 10%

Haute Garonne 14

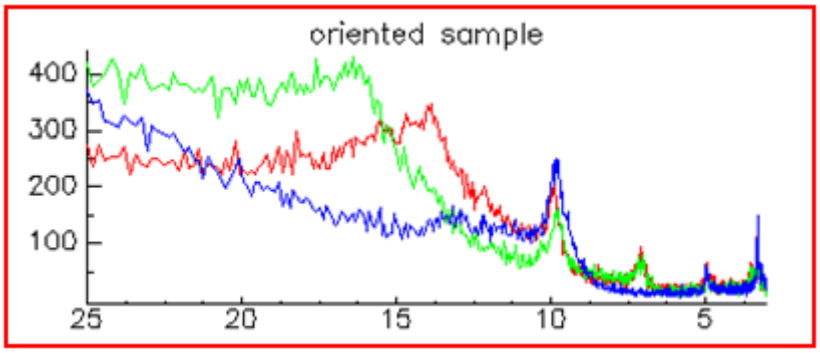


Figure VIII-7 : diffractogramme sur lame orientée (normal : rouge ; éthylène glycol : vert ; chauffé : bleu)

Semi-quantification sur DRX :

- illite : 60%
- smectite : 25%
- kaolinite : 15%
- interstratifiés : traces

Haute Garonne 22

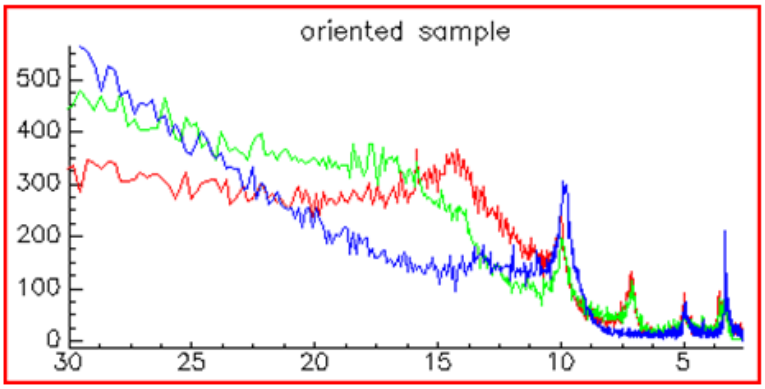


Figure VIII-8 : diffractogramme sur lame orientée (normal : rouge ; éthylène glycol : vert ; chauffé : bleu)

Semi-quantification sur DRX :

- illite : 45%
- smectite : 25%
- kaolinite : 15%
- interstratifiés (C/M) : 15%

Tarn et Garonne 85

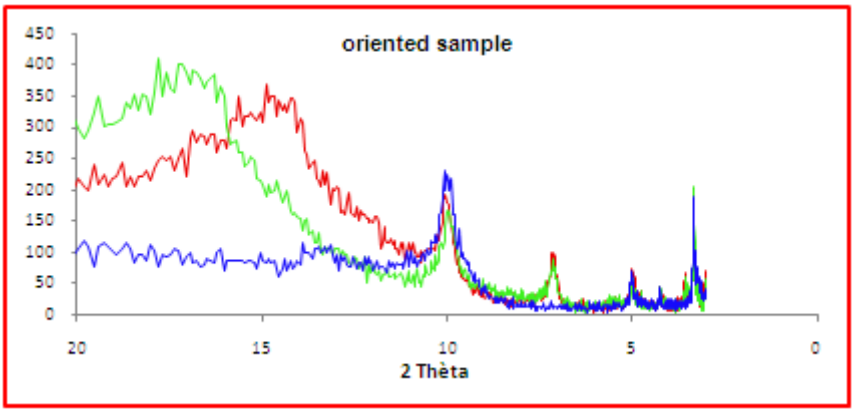


Figure VIII-9 : diffractogramme sur lame orientée (normal : rouge ; éthylène glycol : vert ; chauffé : bleu)

Semi-quantification sur DRX :
- illite : 60%
- smectite : 20%
- kaolinite : 20%

Tarn et Garonne 89

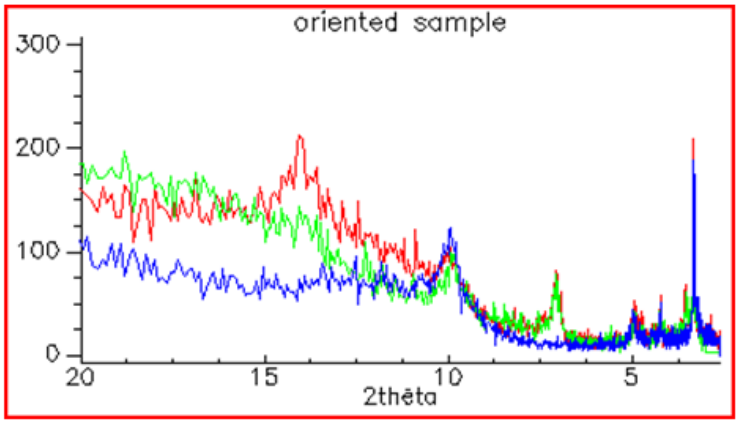


Figure VIII-10 : diffractogramme sur lame orientée (normal : rouge ; éthylène glycol : vert ; chauffé : bleu)

Semi-quantification sur DRX :
- illite : 65%
- kaolinite : 25%
- interstratifiés (C/V): 15%

Tarn 72

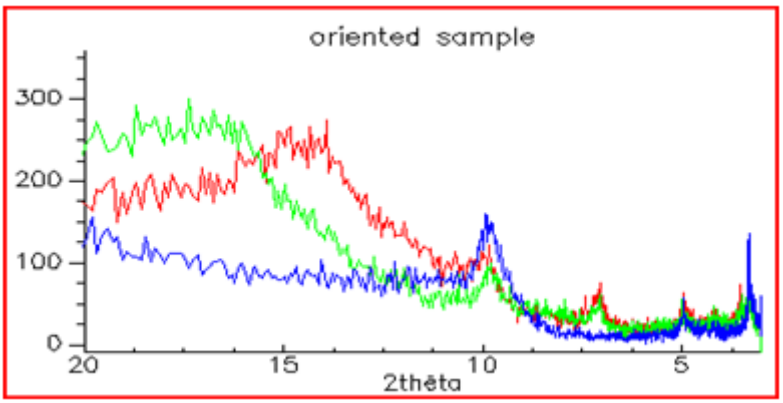


Figure VIII-11 : diffractogramme sur lame orientée (normal : rouge ; éthylène glycol : vert ; chauffé : bleu)

Semi-quantification sur DRX :
- illite : 40%
- smectite 40 %
- kaolinite : 20%

VIII.2. LES SOLS DES PLAINES ER DES TERRASSES QUATERNAIRES

Haute Garonne 10

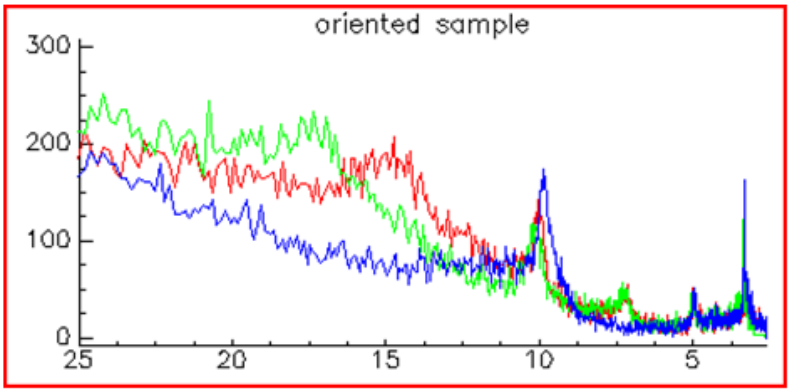


Figure VIII-12 : diffractogramme sur lame orientée (normal : rouge ; éthylène glycol : vert ; chauffé : bleu)

Semi-quantification sur DRX :
- illite : 55%
- smectite : 30 %
- kaolinite : 15%
- interstratifiés : traces

Haute Garonne 11

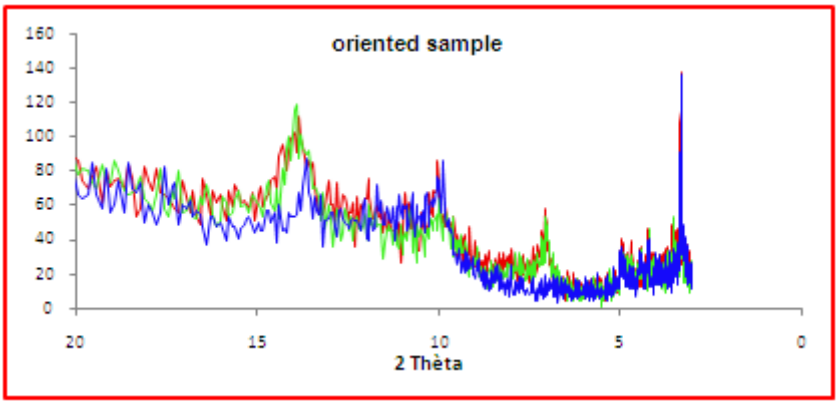


Figure VIII-13 : diffractogramme sur lame orientée (normal : rouge ; éthylène glycol : vert ; chauffé : bleu)

Semi-quantification sur DRX :
- illite : 30%
- kaolinite : 15%
- interstratifiés (V/C) : 65%

Haute Garonne 13

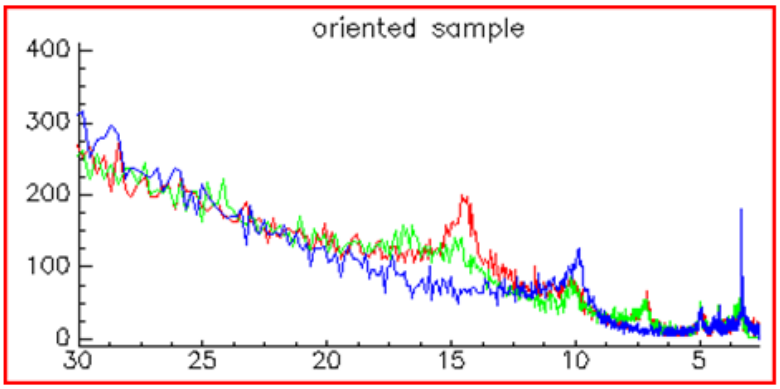


Figure VIII-14 : diffractogramme sur lame orientée (normal : rouge ; éthylène glycol : vert ; chauffé : bleu)

Semi-quantification sur DRX :
- illite : 35%
- smectite : 10 %
- kaolinite : 15%
- vermiculite
- : 40%

Haute Garonne 19

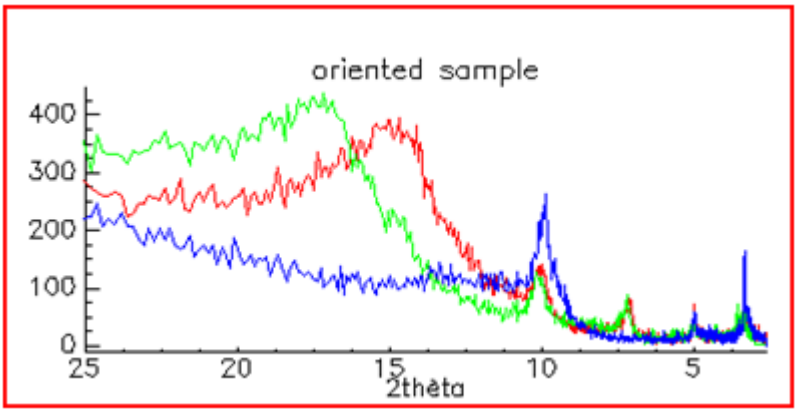


Figure VIII-15 : diffractogramme sur lame orientée

(normal : rouge ; éthylène glycol : vert ; chauffé : bleu)

Semi-quantification sur DRX :
- illite : 40%
- smectite : 40%
- kaolinite : 20%
- interstratifiés : traces

Tarn 76

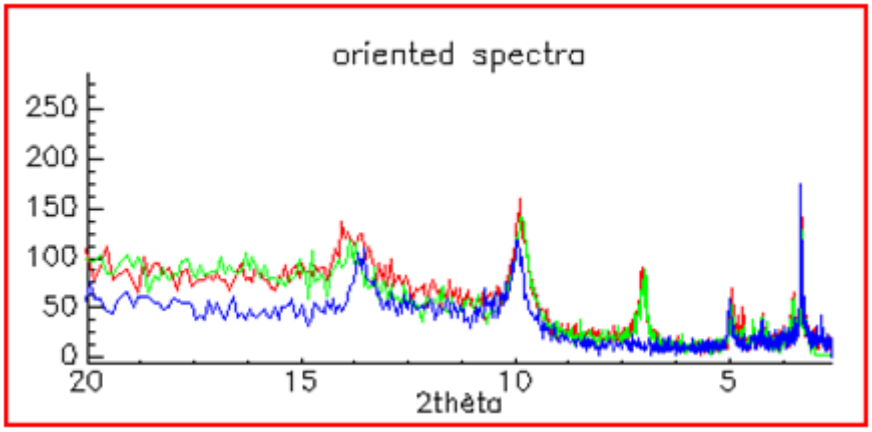


Figure VIII-16 : diffractogramme sur lame orientée

(normal : rouge ; éthylène glycol : vert ; chauffé : bleu)

Semi-quantification sur DRX :
- illite : 50%
- kaolinite : 20%
- Chlorite : 30%

Tarn et Garonne 63

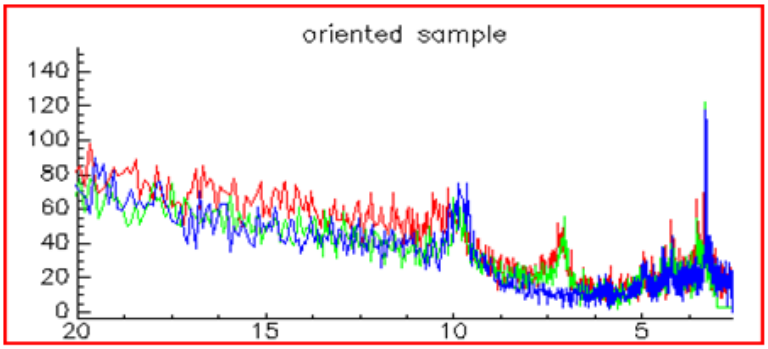


Figure VIII-17 : diffractogramme sur lame orientée (normal : rouge ; éthylène glycol : vert ; chauffé : bleu)

Semi-quantification sur DRX :
- illite : 60%
- kaolinite : 40%
- interstratifiés : traces

Tarn et Garonne 64

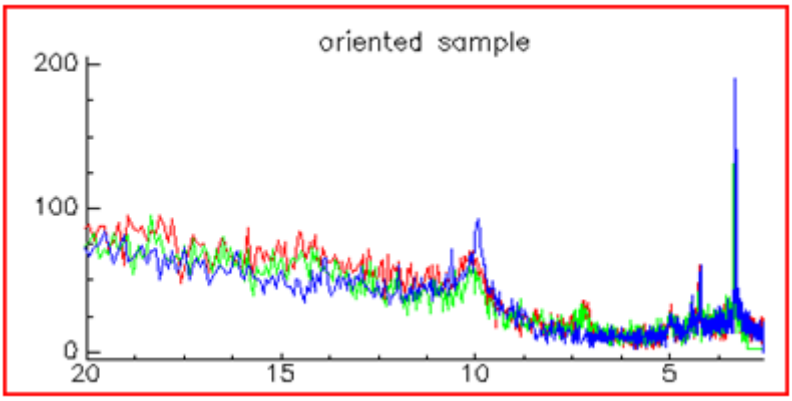


Figure VIII-18 : diffractogramme sur lame orientée (normal : rouge ; éthylène glycol : vert ; chauffé : bleu)

Semi-quantification sur DRX :
- illite : 75%
- kaolinite : 25%

Lot 33

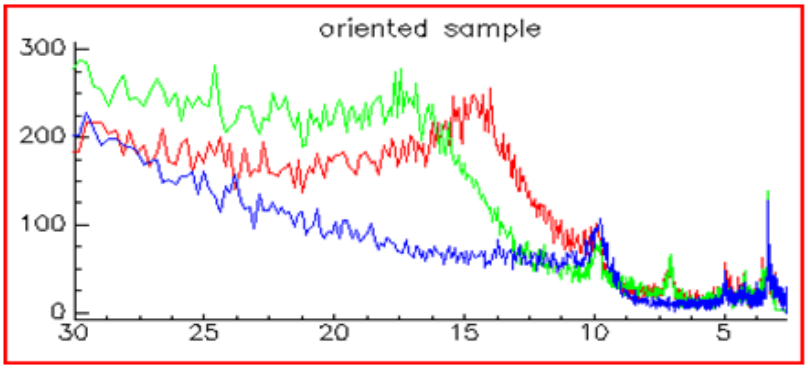


Figure VIII-19 : diffractogramme sur lame orientée (normal : rouge ; éthylène glycol : vert ; chauffé : bleu)

Semi-quantification sur DRX :
- illite : 50%
- smectite : 20%
- kaolinite : 30%
- interstratifiés : traces

VIII.3. LES SOLS DES PLATEAUX MARNO-CALCAIRE

Lot 13

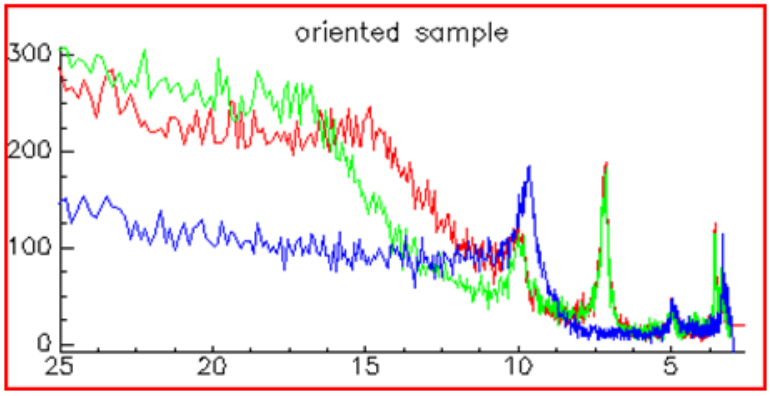


Figure VIII-20 : diffractogramme sur lame orientée (normal : rouge ; éthylène glycol : vert ; chauffé : bleu)

Semi-quantification sur DRX :
- illite : 40%
- kaolinite : 40%
- smectite : 20%

Lot 40

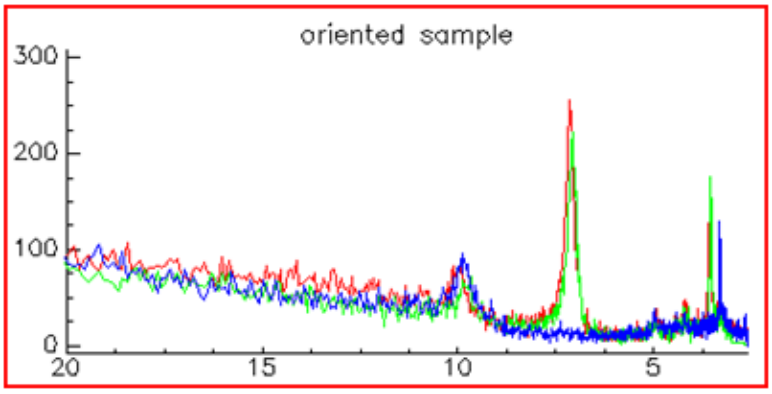


Figure VIII-21 : diffractogramme sur lame orientée

(normal : rouge ; éthylène glycol : vert ; chauffé : bleu)

Semi-quantification sur DRX :
- illite : 30%
- kaolinite : 70%

LOT 52

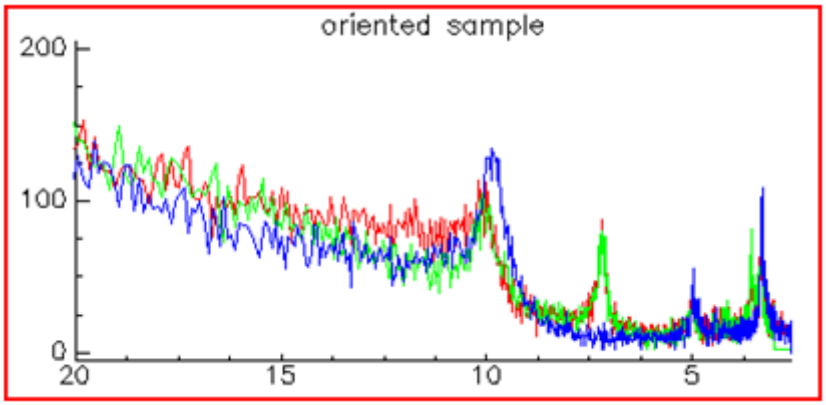
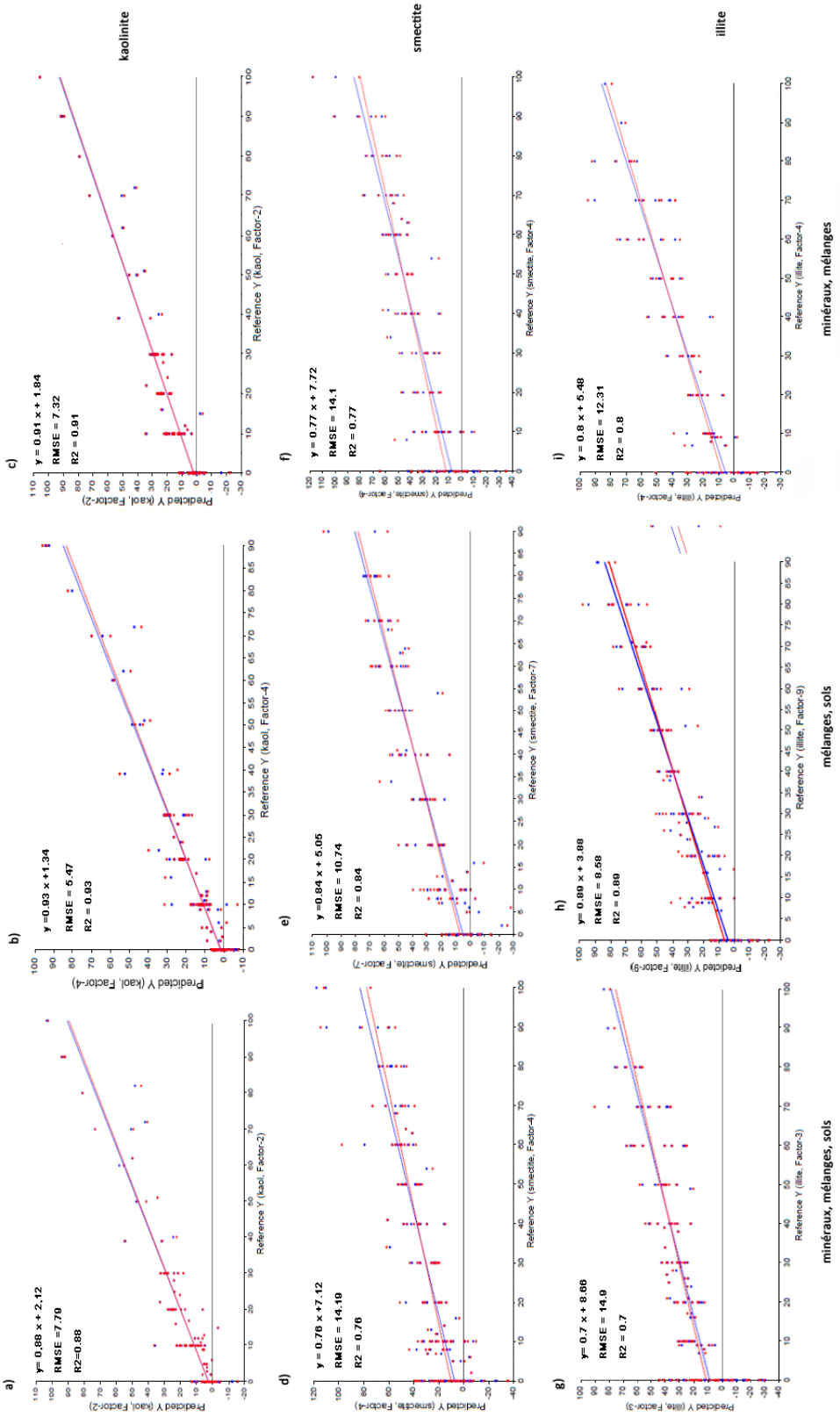


Figure VIII-22 : diffractogramme sur lame orientée

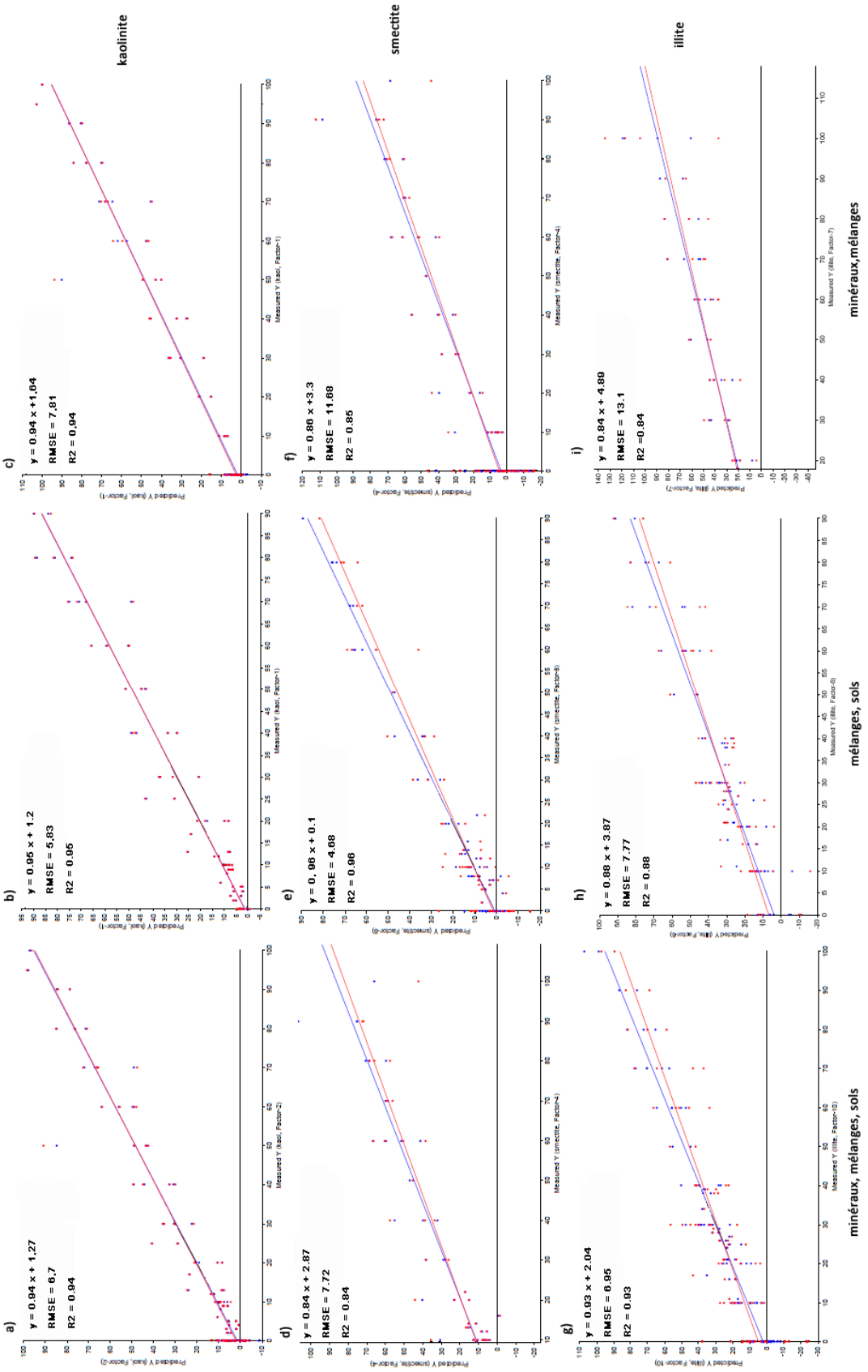
(normal : rouge ; éthylène glycol : vert ; chauffé : bleu)

Semi-quantification sur DRX :
- illite : 70%
- kaolinite : 30%

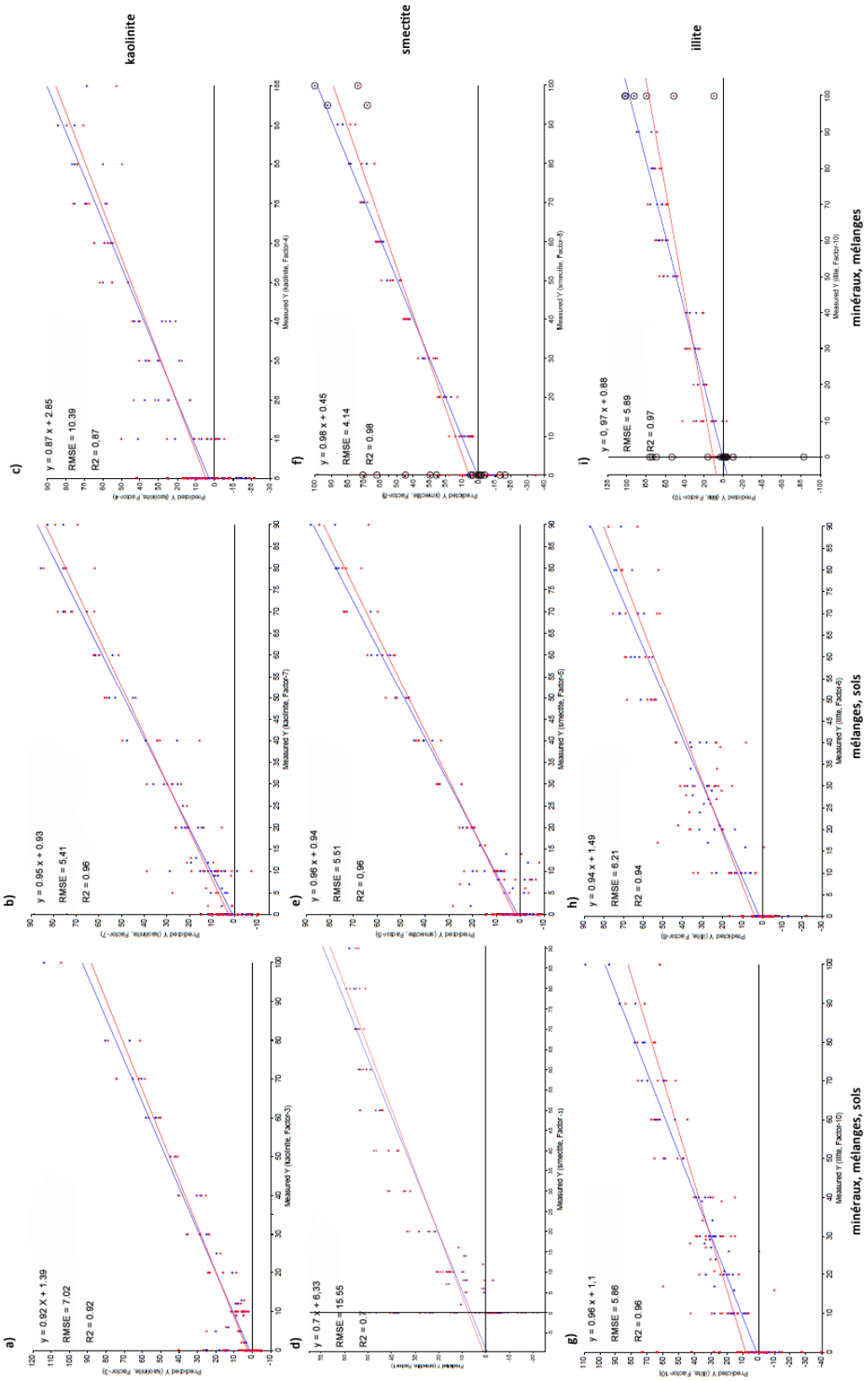
IX. ANNEXE 3 : MODELISATION PLSR



Modèles établies sur les dérivées de spectres MIR pour la prédiction de la kaolinite (a, b, c), la smectite (d, e, f) et l'illite (g, h, i) soit sur l'ensemble des données de calibration (a, d, g) soit sur uniquement sur les sols et les mélanges (b, e, f) ou les minéraux de référence et les mélanges (c, f, i)



Modèles établies sur les dérivées de spectres DRIFT pour la prédiction de la kaolinite (a, b, c), la smectite (d, e, f) et l'illite (g, h, i) soit sur l'ensemble des données de calibration (a, d, g) soit sur uniquement sur les sols et les mélanges (b, e, f) ou les minéraux de référence et les mélanges (c, f, i).



Modèles établies sur les dérivées de spectres ASD pour la prédiction de la kaolinite (a, b, c), la smectite (d, e, f) et l'illite (g, h, i) soit sur l'ensemble des données de calibration (a, d, g) soit sur uniquement sur les sols et les mélanges (b, e, f) ou les minéraux de référence et les mélanges (c, f, i). Les individus entourés correspondent aux minéraux.

X. ANNEXE 4 : PROGRAMME

```
%% do cross-correlation an cohenrence between
% two FTIR spectrums as a function of
% both spectral wavelenght in nm or microm or wave number cm-1
% *** Requires Matlab wavelet toolbox ********************
% d  spectral lenght
% y1,y2     Datasets to be compared
% nscales   No. of scales to use - not more than length(d/2)...
% Outputs;
% W3c     : Wavelet Cross Corelation Coherence; a matrix of values between
% W2C     : Wavelet Cross Corelation ; a matrix of values
%           -1          anticorrelation,
%            zero       no correlation, and
%            1          positive correlation
% derived from reference paper;
%*****************************************************************
% Cooper, G.R.J., and Cowan, D.R., 2008.
% Comparing Time Series using Wavelet Based Semblance Analysis
% Computers & Geosciences v.34(2) p.95-102.
%*****************************************************************
% Grinsted, Moore, Lapland, Jevrejeva, 2004
% Application of Cross Wavelet transform and wavelet Coherence to
% geophysical Time series.
% Nonlinear Processes in Geophysics, 11, 561-566
%*****************************************************************
% Torence and Compo, 1998
% A practical guide for wavelet analysis
% Bull. Amer. Meteorol. Soc., 79, 61-78
%
% Camille Truche and Darrozes José 2009
% ******************************************************

%% Initiation of global variables
clear
warning off all
global fig1, global fig2, global fig3, global dS; global aWxy;
global accWxy; global accWxy3;
global wave1; global wave2; global wave3; global wave;
global coi1, global coi2; global coi3; global coi;
%*****************************************************************
global sig95; global nbx; global sigma; global sig952;
global sigma2, global sigma1; global sigma3; global period, global scale,
global period1, global scale1,  global sig951;
global period2, global scale2,  global sig952;
global period3, global scale3,  global sig953;
global signif; global signif1; global signif2; global signif3;
%****************wavelet transform results
global d1, global d2; global d3; % d1 reference pectrum , d2 unknow spectrum
global name; global moycoefcohetot; global coico; global tt;
global meandx, global meandx2; global moycoeftot; global coixy;
global WaveletCrossCor; global WaveletCrossCor3;
global Rsq; global Rsq3; global nscale;global GRAPHICINT;
global Dj; global SEUIL; global s0; global J1; global scale;
global fourier_factor; global k0; global var2; global power;
global C; global NORMALISATION;global visualisation;
global daughterREAL; global daughterIMA; global AR1COEF;
global meancoefscale; global maxcoefscale; global ETC;
global meancoefscaleX; global maxcoefscaleX; global ECX;global tot;
global coef_max; global SEUILMAX;
```

```
GRAPHICINT=1;   % No intermediary plot result
%% Load of libary and unknow data
%global WaveletCrossCor2;
[filename, pathname] = uigetfile('*.txt', 'Pick a text file with yours data');
visualisation=0  %% wavelet form in the real space are not displayed
%answer=input('Would you work with the entire spectrun (0) or only with the peak
à 2200nm (1 default)? -> ');
answer=1;
if (answer==0)
        Xref=strcat(pathname,'ref_x.txt')
        Yref=strcat(pathname,'ref_value.txt')
        Yname=strcat(pathname,'ref_nametot.txt')
else
        Xref=strcat(pathname,'ref_x_clay_2200.txt')
        %Yref=strcat(pathname,'ref_value_jose2.txt')
        %Yname=strcat(pathname,'ref_name_jose2.txt')

        Yref=strcat(pathname,'ref_value_soil_2200.txt')
        Yname=strcat(pathname,'ref_name_soil_2200.txt')
end
%% Variable initialisation
Yuser=strcat(pathname,filename)
filenameX='ref_x.txt';
Xuser=strcat(pathname,filenameX)
Xreference=load(char(Xref));
Yreference=load(char(Yref));
name=importdata(char(Yname),' ');
%Yreference=1-Yreference;
y2=load(Yuser);
%y2=1-y2*(1/max(y2));

position = findstr('ref_value_G1Bbis.txt', Yuser);
if position >0
    x2=load(Xuser);
else
    x2=Xreference;
end
[nby,nbx]=size(Yreference);
meancoefscaleX=zeros(nbx,1);
maxcoefscaleX=zeros(nbx,1);
SEUIL=0.8                     % threshold value of cross correlation

dx=(max(Xreference)-min(Xreference))/nbx;  % step
k0=6; %optimal value define by Torrence et Compo, 1998
Dj=1/24;                                % octave per scale
s0=16*dx;                                % starting scale
J1=fix((log(nbx*dx/s0)/log(2))/Dj)-fix(s0)+2      % number of scale
scale = s0*2.^((0:J1)*Dj);
[a nscales]=size(scale);
fourier_factor = (4*pi)/(k0 + sqrt(2 + k0^2))
%%% initial nscales= round((log2(nbx*.17)/(Dj))+dx/2);

offsetnois=0; out=zeros(nbx);
nby=24; % number of pole
WaveletCrossCor=zeros(nscales,nbx,nby+2);
WaveletCrossCor3=zeros(nscales,nbx,nby+2);
Rsq=zeros(nscales,nbx,nby+2);Rsq3=zeros(nscales,nbx,nby+2);
Rsq3=zeros(nscales,nbx,nby+2);Rsq3=zeros(nscales,nbx,nby+2);
moycoeftot=zeros(nby,1); moycoefcohetot=zeros(nby,1);

test=1;
```

```
y=zeros(nby,nbx);
y(1,:)=(Yreference(1,:)); y(2,:)=Yreference(2,:); y(3,:)=Yreference(3,:);
y(4,:)=Yreference(4,:); y(5,:)=Yreference(5,:);y(6,:)=Yreference(6,:);
y(7,:)=(Yreference(7,:));y(8,:)=(Yreference(8,:));y(9,:)=(Yreference(9,:));
y(10,:)=(Yreference(10,:));y(11,:)=(Yreference(11,:));y(12,:)=(Yreference(12,:));
y(13,:)=(Yreference(13,:));y(14,:)=(Yreference(14,:));y(15,:)=(Yreference(15,:));
y(16,:)=(Yreference(16,:));y(17,:)=(Yreference(17,:));y(18,:)=(Yreference(18,:));
y(19,:)=(Yreference(19,:));y(20,:)=(Yreference(20,:));y(21,:)=(Yreference(21,:));
y(22,:)=(Yreference(22,:));y(23,:)=(Yreference(23,:));y(24,:)=(Yreference(24,:));
%%%%%%%%%y2=y2*400; %%%%%%%%%%%%%%%%%%%%%%


y3=flipud(y);  % changing libraray order to disminush the spectrum combination
effect
d2=[x2',y2'];
[d2,dx2]=formatts(d2);   % verification of the format
meandx2=mean(dx2); meandx=(min(dx2)-0.002);
visualisation=0; % for morlet wavelet representation put 1
%calculation of the wavelet transform of unknown spectrum
cxwt(d2,meandx);
wave2=wave; coi2=coi; signif2=signif; scale2=scale; period2=period;
sig952=sig95;


%% wavelet calculation for each pole correlated to unknown spectrum
for indice=1:nby
%% plot initial spectrums
    indice
   %name(indice)
    d1=[Xreference',y(indice,:)'];
    [d1,dx]=formatts(d1);   % verification of the format
    d3=[Xreference',y3(indice,:)'];
    [d3,dx]=formatts(d3);
    tlim=[min(d1(1,1),d2(1,1)) max(d1(end,1),d2(end,1))];
    AR1COEF=[ar1nv(d1(:,2)+whitenoise(nbx,0.3,0.3)) ar1nv(d2(:,2)+
whitenoise(nbx,1.3,1.2))];
    if any(isnan(AR1COEF))
        error('Automatic AR1 estimation failed. Specify them manually (use the
arcov or arburg estimators).');
    end
    plotcolpos=1;
    if GRAPHICINT==0 fig1=figure(); end
    fig2=figure();
    figure(fig2);

    subplot(3,1,3)
    miniX=min(min(d1(:,1),d2(:,1))); %if (miniX > min(d2(:,1)))
miniX=min(d2(:,1)); end
    maxiX=max(max(d1(:,1),d2(:,1))); %if (maxiX < max(d2(:,1)))
maxiX=max(d2(:,1)); end
    miniY=min(min(d1(:,2),d2(:,2))); %if (miniY > min(d2(:,2)))
miniY=min(d2(:,2)); end
    maxiY=max(max(d1(:,2),d2(:,2))); %if (maxiY < max(d2(:,2)))
maxiY=max(d2(:,2)); end

    plot(d1(:,1), d1(:,2),'b', d2(:,1),d2(:,2),'r');
    axis([miniX  maxiX  miniY  maxiY+0.0000005])
    ylabel('Normalised reflectance');
    xlabel('wave number in cm-1');
    set(findobj(gca,'Type','line','Color',[0 0 1]),...
    'Color','blue',...
```

```
    'LineWidth',2)
    title(['PS of library ',   name{indice},  ' in blue and Unknow ',filename, '
in red']);

%% continuous wavelet transform calculation
visualisation=0;
cxwt(d1,meandx);
wave1=wave; scale1=scale; scale2=scale; scale3=scale;
period1=period; period2=period; period3=period; coi1=coi; signif1=signif;
sig951=sig95;
%signal 1 from the library plot phase, magnitude power spectrum
%*************************************************************************
plotcxwt(tlim, plotcolpos, d1(:,1),meandx, sigma, wave,period,signif,coi,sig951);

%*************************************************************************
title(['PS of   ' name{indice} ' ']); %Power Spectrum
set(gca,'xlim',tlim);
%*************************************************************************
%signal 2 continous wavelet transform calculation


plotcolpos=2;
        if GRAPHICINT==0 figure(fig1); end
%*************************************************************************
plotcxwt(tlim, plotcolpos, d2(:,1),meandx, sigma2, wave2, period2, signif2,
coi2,sig952);

%*************************************************************************
%signal 1 plot phase, magnitude power spectrum of unknown signal


cxwt(d3,meandx);
wave3=wave; coi3=coi; signif3=signif;
sig953=sig95;


title(['PS of  ' filename ' ']);
set(gca,'xlim',tlim);
fig1=figure();
figure(fig1);
 %*************************************************************************
%% Wavelet Cross Correlation transform (W2C)
% The XWT finds regions in time frequency space where
% the time series show high common power.
subplot(2,1,1);
w2c(indice)
title(['W2C transform: ' name{indice} '-' filename ' coefmoy= '
num2str(moycoeftot(indice)) ' coef max= ' num2str(coef_max) ' '] )
if (indice==1)
    total2=tot;
else
    total2=[total2 ; tot];
end


if indice==nby
    for(i=1:nby) WaveletCrossCor3(:,:,i)=WaveletCrossCor(:,:,abs(nby+1-i)); end
end

%% Wavelet coherence (WTC)
% The WTC finds regions in time frequency space where the two
```

```
% time series co-vary (but does not necessarily have high power).


subplot(2,1,2);
wcoherence(indice)

title(['WTC: ' name{indice} ' - '  filename ' coefmoy =
',num2str(moycoefcohetot(indice)) ' coef max= ' num2str(coef_max) ' ' ] );

figure(fig1);
end
fig2b= figure();
[yhisto xhisto]=hist(total2',15);
    miniY=min(min(hist(yhisto))); %if (miniY > min(d2(:,2))) miniY=min(d2(:,2));
end
    maxiY=max(max(hist(yhisto))); %if (maxiY < max(d2(:,2))) maxiY=max(d2(:,2));
end
    plot(xhisto,sum(yhisto'))
    axis([ 0  1+0.0000005 miniY  maxiY])
    ylabel('count');
    xlabel('correlation coef.')
fig3=figure();
figure(fig3);
  subplot(3,1,1)
    miniX=min(d1(:,1)); %if (miniX > min(d2(:,1))) miniX=min(d2(:,1)); end
    maxiX=max(d1(:,1)); %if (maxiX < max(d2(:,1))) maxiX=max(d2(:,1)); end
    miniY=min(min(meancoefscaleX)); %if (miniY > min(d2(:,2)))
miniY=min(d2(:,2)); end
    maxiY=max(max(meancoefscaleX)); %if (maxiY < max(d2(:,2)))
maxiY=max(d2(:,2)); end

    plot(d1(:,1), (meancoefscaleX/maxiY)');
    axis([miniX  maxiX  0 1+0.0000005])
    ylabel('Mean coef by Wave number');
    xlabel('wave number in cm-1')
    title(['statistical analysis Mean', ' in red']);


  subplot(3,1,2)
    miniY=min(min(ECX)); %if (miniY > min(d2(:,2))) miniY=min(d2(:,2)); end
    maxiY=max(max(ECX)); %if (maxiY < max(d2(:,2))) maxiY=max(d2(:,2)); end

    plot(d1(:,1), (ECX/maxiY)');
    axis([miniX  maxiX  miniY  1+0.0000005])
    ylabel('Standart deviation of coef by Wave number');
    xlabel('wave number in cm-1')

fig4=figure();

figure(fig4);
  subplot(3,1,1)
    miniX=min(d1(:,1)); %if (miniX > min(d2(:,1))) miniX=min(d2(:,1)); end
    maxiX=max(d1(:,1)); %if (maxiX < max(d2(:,1))) maxiX=max(d2(:,1)); end
    miniY=min(min(meancoefscaleX)); %if (miniY > min(d2(:,2)))
miniY=min(d2(:,2)); end
    maxiY=max(max(meancoefscaleX)); %if (maxiY < max(d2(:,2)))
maxiY=max(d2(:,2)); end
    Indic=find(isnan(meancoefscaleX));
    meanscaleX(Indic)=0.0;
    plot(d1(:,1),(meancoefscaleX/maxiY)');
    axis([miniX  maxiX  0  1+0.0000005])
```

```
    ylabel('Mean coef by Wave number');
    xlabel('wave number in cm-1')
    title(['statistical analysis Mean', ' in red']);


  subplot(3,1,2)
    miniX=min(min(ECX)); %if (miniY > min(d2(:,2))) miniY=min(d2(:,2)); end
    maxiX=max(max(ECX)); %if (maxiY < max(d2(:,2))) maxiY=max(d2(:,2)); end
    plot(d1(:,1), (ECX/maxiX)');
    axis([ miniX  maxiX 0  1+0.0000005 ])
    ylabel('Standart deviation of coef by Wave number');
    xlabel('wave number in cm-1')

  subplot(3,1,3)
    miniY=min(min(maxcoefscaleX)); %if (miniY > min(d2(:,2))) miniY=min(d2(:,2));
end
    maxiY=max(max(maxcoefscaleX)); %if (maxiY < max(d2(:,2))) maxiY=max(d2(:,2));
end
    miniX=min(period); %if (miniX > min(d2(:,1))) miniX=min(d2(:,1)); end
    maxiX=max(period);
    plot(period, (maxcoefscaleX/maxiY));
    axis([miniX  maxiX  0  1+0.0000005])
    ylabel('Mean coef by scale');
    xlabel('scale in cm-1')

%% calculation of families pourcentage

pourcent=zeros(1,nby);
pourcent2=zeros(1,nby);
pourcent3=zeros(1,nby);
compteur=0;
compteur2=0;
compteur3=0;
WaveletCrossCor(:,:,nby+1)=0.00;
WaveletCrossCor3(:,:,nby+1)=0.00;
Rsq(:,:,nby+1)=-0.001;
coef1= mean(moycoeftot);
[meantheta,R,sigma,confangle,kappa]=anglemean(accWxy);
coef2=mean(moycoefcohetot);
[meantheta,R,sigma,confangle,kappa]=anglemean(aWxy);

s0=2;
a= (max(period)-min(period))/max(log2(coixy));
b=min(period);
aa=((max(period)-min(period))/nscales);
bb=min(period);
cpt=0;


for i=1:nscales
       for j=1:nbx
           for k=1:nby

            if (aa*i+bb)<(a*log2(coixy(1,j))+b)

               if (WaveletCrossCor(i,j,k)) > (WaveletCrossCor(i,j,nby+1))
                   %higest coef selected
                   WaveletCrossCor(i,j,nby+1)=WaveletCrossCor(i,j,k);
                   WaveletCrossCor(i,j,nby+2)=k; % type of mineral

                end
```

```
                end
                 if (WaveletCrossCor3(i,j,k)) > (WaveletCrossCor3(i,j,nby+1))

                     %higest coef selected
                     WaveletCrossCor3(i,j,nby+1)=WaveletCrossCor3(i,j,k);

                     WaveletCrossCor3(i,j,nby+2)=k; % type of mineral
                     compteur3=compteur3+1;

                      end
                 end
                 if (Rsq(i,j,k)) > (Rsq(i,j,nby+1))
                    %idem for coherence
                     Rsq(i,j,nby+1)=Rsq(i,j,k);
                     Rsq(i,j,nby+2)=k;

                 end
                  if (Rsq3(i,j,k)) > (Rsq3(i,j,nby+1))
                    %idem for coherence
                     Rsq3(i,j,nby+1)=Rsq3(i,j,k);
                     Rsq3(i,j,nby+2)=k;

                 end

             end

        end



Yticks = 2.^(fix(log2(min(period))):fix(log2(max(period))));

correction=(length(y2)-sum(y2))./(length(y2)-sum(y'))



%% graphical representation of mineral phases

%histogram of mineralogical class for cross correlation and coherence
figure(3000);clf;
subplot(2,1,1);
histo=zeros(nscales,nbx);
histo1=WaveletCrossCor(:,:,nby+2);
histo=WaveletCrossCor(:,:,nby+1);
HI1=zeros(1,nbx*nscales);
HI=zeros(1,nbx*nscales);
HI1=histo1(1,:);
HI=histo(1,:);
for(i=2:nscales), HI1=[HI1 histo1(i,:)]; HI =[HI histo(i,:)]; end
COEF=find(HI>SEUIL);
n=hist(HI1(COEF),[1:nby]);
n=n.*correction;
n=n/sum(n)*100.0;
%n1=n.*correction./sum(n.*correction)*100;

histo3=zeros(nscales,nbx);
histo31=WaveletCrossCor3(:,:,nby+2);
histo3=WaveletCrossCor3(:,:,nby+1);
HI31=zeros(1,nbx*nscales);
HI3=zeros(1,nbx*nscales);
HI31=histo31(1,:);
```

```
HI3=histo3(1,:);
for(i=2:nscales), HI31=[HI31 histo31(i,:)]; HI3 =[HI3 histo3(i,:)]; end
COEF3=find(HI3>(SEUIL));
subplot(2,1,1);
n3=hist(HI31(COEF3),[1:nby]);
n3=n3.*correction;
n3=n3/sum(n3)*100.0; % sense fedcba
n3=fliplr(n3);% sense abcdef
n
n3
nmoy=(n+n3)/2.0
bar([n; (n+n3)/2.0; n3]',1.5);

text(1,n(1),'A');text(2,n(2),'B');
text(3,n(3),'C');text(4, n(4),'D');
text(5, n(5),'E');text(6, n(6),'F');
text(7, n(7),'E');text(8, n(8),'F');
text(9, n(9),'E');text(10, n(10),'E');
text(11, n(11),'E');text(12, n(12),'E');
text(13, n(13),'E');text(14, n(14),'E');
text(15, n(15),'E');
text(16, n(16),'E');text(17, n(17),'E');
text(18, n(18),'E');text(19, n(19),'E');
text(20, n(20),'E');
text(21, n(21),'E');text(22, n(22),'E');
text(23, n(23),'E');text(24, n(24),'E');

%*********************************************************************
return;
figure(3001);clf;
subplot(2,1,1);
histoC=zeros(nscales,nbx);
histo1C=Rsq(:,:,nby+2);
histoC=Rsq(:,:,nby+1);
HI1C=zeros(1,nbx*nscales);
HIC=zeros(1,nbx*nscales);
HI1C=histo1(1,:);
HIC=histo(1,:);
for(i=2:nscales), HI1C=[HI1C histo1(i,:)]; HIC =[HIC histo(i,:)]; end
COEFC=find(HIC>SEUIL);
nC=hist(HI1C(COEFC),[1:nby]);
nC=nC.*correction;
nC=nC/sum(nC)*100.0;
n1C=nC.*correction./sum(n.*correction)*100;
histo3C=zeros(nscales,nbx);
histo31C=Rsq3(:,:,nby+2);
histo3C=Rsq3(:,:,nby+1);
HI31C=zeros(1,nbx*nscales);
HI3C=zeros(1,nbx*nscales);
HI31C=histo1C(1,:);
HI3C=histoC(1,:);
for(i=2:nscales), HI31C=[HI31C histo3C(i,:)]; HIC =[HIC histoC(i,:)]; end
COEF3C=find(HI3C>SEUIL);
n3C=hist(HI31C(COEF3C),[1:nby]);

n3C=n3C.*correction;
n3C=n3C/sum(n3C)*100.0;
bar([nC; n3C; (nC+n3C)/2.0]');
text(1,nC(1),'A');text(2,nC(2),'B');
text(3,nC(3),'S');text(4, nC(4),'D');
text(5, nC(5),'E');text(6, nC(6),'F');
```

```
text(7, nC(7),'G');text(8, nC(8),'H');
text(9, nC(9),'I');text(10, nC(10),'j');
text(11, nC(11),'K');text(12, nC(12),'L');
text(13, nC(13),'M');text(14, nC(14),'N');
text(15, nC(15),'O');
text(16, nC(16),'P');text(17, nC(17),'Q');
text(18, nC(18),'R');text(19, nC(19),'S');
text(20, nC(20),'O');
text(21, nC(21),'P');text(22, nC(22),'Q');
text(23, nC(23),'R');text(24, nC(24),'S');
pause()

bar(1:nby,pourcent,'r');
text(1,n(1),'A');text(2,n(2),'B');
text(3,n(3),'C');text(4, n(4),'D');
text(5, n(5),'E');text(6, n(6),'F');
text(7, n(7),'G');text(8, n(8),'H');
text(9, n(9),'I');text(10, n(10),'J');
text(11, n(11),'K');text(12, n(12),'L');
text(13, n(13),'M');text(14, n(14),'N');
text(15, n(15),'O');
text(16, n(16),'P');text(17, n(17),'Q');
text(18, n(18),'R');text(19, n(19),'S');
text(20, n(20),'O');
text(21, n(21),'P');text(22, n(22),'Q');
text(23, n(23),'R');text(24, n(24),'S');

subplot(2,1,2);
bar(1:nby,pourcent2,'b');
text(1,pourcent(1),'A');text(2,pourcent(2),'B');
text(3,pourcent(3),'C');text(4,pourcent(4),'D');
text(5,pourcent(5),'E');text(6,pourcent(6),'F');
text(7,pourcent(7),'G');text(8,pourcent(8),'H');
text(9,pourcent(9),'I');text(10,pourcent(10),'J');
text(11,pourcent(11),'K');text(12,pourcent(12),'L');
text(13,pourcent(13),'M');text(14,pourcent(14),'N');
text(15,pourcent(15),'O');text(16,pourcent(16),'P');
text(17,pourcent(17),'Q');text(18,pourcent(18),'R');
text(19,pourcent(19),'S');
text(20,pourcent(20),'O');text(21,pourcent(21),'P');
text(22,pourcent(22),'Q');text(23,pourcent(23),'R');
text(24,pourcent(24),'S');
figure(3001);
subplot(2,1,1);
% the better coef of the n reference spectrum and for each pixel is selected
H=imagesc(Xreference',log2(period),WaveletCrossCor(:,:,nby+1)); axis xy; axis
tight; title('Global W2C'); ylabel('Period cm-1');
hold on

 HCB=colorbar;%#ok
        set(gca,'YLim',log2([min(period),max(period)]), ...
            'YDir','normal', 'layer','top', ...
            'YTick',log2(Yticks(:)), ...
            'YTickLabel',num2str(Yticks'), ...
            'layer','top')
        ylabel('Period in cm-1')
hold off


% coefficient correlation map for cross correlation and coherence
subplot(2,1,2);
```

```
H=imagesc(Xreference',log2(period),Rsq(:,:,nby+1));axis xy; axis tight;
title('Global Coherence'); ylabel('Period cm-1');
hold on

 HCB=colorbar;%#ok
        set(gca,'YLim',log2([min(period),max(period)]), ...
            'YDir','normal', 'layer','top', ...
            'YTick',log2(Yticks(:)), ...
            'YTickLabel',num2str(Yticks'), ...
            'layer','top')
        ylabel('Period in cm-1')
  hold off
  figure(3003);
subplot(2,1,1);

% mineralogical repartition map for cross correlation and coherence
H=imagesc(Xreference',log2(period),WaveletCrossCor(:,:,nby+2));
axis xy; axis tight; title('W2C repartition map of clay minerals');
ylabel('Period cm-1');
hold on

 HCB=colorbar;%#ok
        set(gca,'YLim',log2([min(period),max(period)]), ...
            'YDir','normal', 'layer','top', ...
            'YTick',log2(Yticks(:)), ...
            'YTickLabel',num2str(Yticks'), ...
            'layer','top')
        ylabel('Period in cm-1')
hold off

subplot(2,1,2);
H=imagesc(Xreference',log2(period),Rsq(:,:,nby+2));
axis xy; axis tight; title('Coherence repartition map of clay minerals');
ylabel('Period cm-1');
hold on

 HCB=colorbar;%#ok
        set(gca,'YLim',log2([min(period),max(period)]), ...
            'YDir','normal', 'layer','top', ...
            'YTick',log2(Yticks(:)), ...
            'YTickLabel',num2str(Yticks'), ...
            'layer','top')
        ylabel('Period in cm-1')
  hold off
```

AUTEUR : Camille TRUCHE

TITRE : Caractérisation et quantification des minéraux argileux dans les sols expansifs par spectroscopie infrarouge aux échelles du laboratoire et du terrain.

DIRECTEURS DE THESE :

François MARTIN, Professeur à l'Université Paul Sabatier, Toulouse III

José DARROZES, Maître de conférences à l'Université Paul Sabatier, Toulouse III

Anne BOURGUIGNON, Ingénieur BRGM, Orléans

LIEU ET DATE DE SOUTENANCE : Université Paul Sabatier, LMTG, le 13 décembre 2010

RESUME :

Ce travail tend à l'amélioration de la détection et de la quantification des minéraux argileux, et plus particulièrement des minéraux argileux « gonflants », dans la partie superficielle des sols. Cette étude est menée à l'échelle locale et à l'échelle globale pour accompagner la démarche des régions pour la prévention du risque de retrait-gonflement des argiles. La première partie présente les principes généraux de la spectrométrie infrarouge et la base de données utilisés dans la thèse. Nous confrontons ensuite deux approches classiquement utilisées dans les sciences du sol : Les méthodes statistiques (PLSR et ACP) et l'étude des figures d'absorption caractéristiques des minéraux argileux. Elles montrent principalement la quantification de la kaolinite dans un mélange multiphasé et dans un sol. Les limites de ces méthodes pour la quantification de la smectite amènent au développement d'une nouvelle méthode basée sur la transformée en ondelettes continue et la corrélation croisée en ondelette. Cette méthode apparaît être la plus prometteuse pour la quantification fine de la teneur en smectite des sols, à l'échelle locale comme à l'échelle globale, dans les études de prévention des risques de retrait et de gonflement des sols.

MOTS CLES : Minéraux argileux, spectroscopie proche et moyen infrarouge, PLSR, ACP, transformée en ondelettes continue, corrélation croisée

---

ABSTRACT:

The purpose of this study is the detection and quantification of clay minerals, especially swelling clay minerals, in surface soils part. This study is carried out at local and global scale, in order to support swelling soil hazards approach of the engineering offices. The first part presents general principles of infrared spectroscopy and the database used in the thesis. Then, we compare two frequently used methods in soil sciences: statistic methods (PLSR and PCA) and absorption figures analysis featuring clay minerals. They mainly provide kaolinite quantification in multiphase mixtures and in soil. The limits of smectite quantification of both methods lead to a new method development based on continuous wavelet transform and cross correlation. This method is the most efficient and promising one for weak quantification of swelling minerals, at local scale as well as at global scale, in studies prevention of swelling soil hazard.

KEY WORDS: Clay minerals, near and middle infrared spectroscopy, PLSR, PCA, continuous wavelet transform, cross correlation

---

DISCIPLINE ADMINISTRATIVE: Sciences de la Terre

INTITULE ET ADRESSE DU LABORATOIRE:

Laboratoire des Mécanismes et Transferts en Géologie (LMTG, nouvellement GET), UMR 5563

Université de Toulouse, CNRS, IRD, OMP, 14 Avenue Edouard Belin, F-31400 Toulouse