



**HAL**  
open science

**Obtenir une représentation en continu de la lithologie et de la minéralogie. Exemples d'application du traitement statistique de données de diagraphie aux structures sédimentaires en régime de convergence de plaques (Leg ODP 134, 156 et 160)**

Alain Rabaute

► **To cite this version:**

Alain Rabaute. Obtenir une représentation en continu de la lithologie et de la minéralogie. Exemples d'application du traitement statistique de données de diagraphie aux structures sédimentaires en régime de convergence de plaques (Leg ODP 134, 156 et 160). Géologie appliquée. Université Montpellier II - Sciences et Techniques du Languedoc, 1998. Français. NNT : . tel-00425334

**HAL Id: tel-00425334**

**<https://theses.hal.science/tel-00425334>**

Submitted on 20 Oct 2009

**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.

ACADEMIE DE MONTPELLIER

UNIVERSITE MONTPELLIER II  
— SCIENCES ET TECHNIQUES DU LANGUEDOC —  
I.S.T.E.E.M.

**T H E S E**

présentée à l'Université de Montpellier II Sciences et Techniques du Languedoc pour obtenir  
le diplôme de DOCTORAT

SPECIALITE : *Physique et Chimie de la Terre*  
Formation Doctorale : *Structure et Evolution de la Lithosphère*  
Ecole Doctorale : *Géosciences*

**Obtenir une représentation en continu de la  
lithologie et de la minéralogie.**

**Exemples d'application du traitement statistique de données de  
diagraphie aux structures sédimentaires en régime de convergence de  
plaques (Leg ODP 134, 156 et 160)**

par

**Alain RABAUTE**

Soutenue le 26 Juin 1998 devant la commission d'examen composée de :

Michel SEURET	Université Montpellier II	Professeur	Sédimentologie	Président
Bruce NELSON	University of Washington	Professeur	Géochimie	Rapporteur
Henri BOUGAULT	IFREMER	Directeur de Recherche	Géochimie	Rapporteur
Fouad BADRAN	C.N.A.M.	Maître de Conférence	Statistiques	Rapporteur
Louis BRIQUEU	Université Montpellier II	Directeur de Recherche	Géochimie	Directeur de thèse
Jean-Pierre DELHOMME	Schlumberger R.P.C.	Expert Technique	Géophysique	Examineur
Philippe PEZARD	Université Aix-Marseille III	Chargé de Recherche	Géophysique	Examineur

# Table des matières

<b>Introduction</b>	<b>ix</b>
<b>I La diagraphie</b>	<b>1</b>
<b>1 La diagraphie nucléaire</b>	<b>3</b>
1.1 Introduction . . . . .	3
1.2 Désintégration radioactive et statistiques . . . . .	3
1.3 Détection de la radioactivité gamma naturelle . . . . .	4
1.3.1 Introduction . . . . .	4
1.3.2 Principe . . . . .	7
1.3.3 Mesure de la radioactivité naturelle totale . . . . .	10
1.3.4 Mesure des concentrations en U, K and Th . . . . .	10
Article 1 : “Analysis of natural gamma-ray spectra obtained from sediment cores with the shipboard scintillation detector of the Ocean Drilling Program : example from Leg 156” (1997) Peter Blum, Alain Rabaute, Pierre Gaudon, James F. Allan, <i>Proceedings of The Ocean Drilling Program, Scientific Results</i> , Vol. 156, 183–195 . . . . .	12
1.4 Interaction entre rayons gamma et formation . . . . .	26
1.4.1 Principe . . . . .	26
1.4.2 Mesure de la densité totale ou masse volumique . . . . .	29
1.4.3 Mesure du facteur d’absorption photoélectrique . . . . .	30
1.5 Interactions entre neutrons et formation . . . . .	30
1.5.1 Principe . . . . .	30
1.5.2 La spectrométrie de rayonnement gamma . . . . .	33
1.5.3 Mesure de l’indice d’hydrogène apparent ou porosité neutron . . . . .	42
<b>2 La diagraphie électrique</b>	<b>45</b>
2.1 Principes . . . . .	45
2.1.1 Résistance et résistivité . . . . .	45
2.1.2 Le poids de l’empirisme . . . . .	46
2.2 La diagraphie de résistivité : les dispositifs focalisés . . . . .	46
2.2.1 Le Laterolog <sup>TM</sup> . . . . .	47
2.2.2 La focalisation sphérique . . . . .	47
2.3 Les sondes à induction . . . . .	48
2.3.1 La diagraphie de conductivité . . . . .	48
2.3.2 Induction ou électrode ? . . . . .	51
2.3.3 La diagraphie de susceptibilité magnétique . . . . .	51

2.4	Imagerie de la paroi du trou . . . . .	53
2.4.1	Principe . . . . .	53
2.4.2	Traitement des données et production de l'image . . . . .	54
<b>3</b>	<b>La diagraphie magnétique</b> . . . . .	<b>57</b>
3.1	La résonance magnétique nucléaire . . . . .	57
3.1.1	Principes . . . . .	57
3.1.2	Paramètres influençant la mesure . . . . .	59
3.1.3	Une mesure de la porosité indépendante de la lithologie . . . . .	62
3.2	La mesure de l'induction magnétique totale . . . . .	62
3.2.1	Principe . . . . .	63
3.2.2	Application . . . . .	63
<b>II</b>	<b>Le traitement statistique des données de diagraphies</b> . . . . .	<b>65</b>
<b>4</b>	<b>Qualité des données et préparation à l'analyse</b> . . . . .	<b>67</b>
4.1	Introduction . . . . .	67
4.2	La population et l'échantillon . . . . .	67
4.3	La qualité des données . . . . .	68
4.3.1	La mesure de diagraphie . . . . .	68
4.3.2	La précision . . . . .	68
4.3.3	La reproductibilité . . . . .	69
4.4	La représentation des données . . . . .	70
4.4.1	Introduction . . . . .	70
4.4.2	La détermination des vecteurs propres . . . . .	71
4.4.3	L'Analyse en Composantes Principales . . . . .	73
4.4.4	L'analyse factorielle . . . . .	74
<b>5</b>	<b>La classification de données multivariées</b> . . . . .	<b>75</b>
5.1	Introduction . . . . .	75
5.2	Le diagramme X-Y . . . . .	75
5.3	Les méthodes de classification . . . . .	77
5.3.1	La classification hiérarchique . . . . .	77
5.3.2	La classification itérative non-hiérarchique ou classification en nuées dynamiques . . . . .	80
5.3.3	Introduction des probabilités d'appartenance dans le classement en nuées dynamiques . . . . .	81
<b>6</b>	<b>Donner un sens aux classes individualisées</b> . . . . .	<b>85</b>
6.1	Introduction . . . . .	85
6.2	Les informations directes . . . . .	85
6.2.1	Les valeurs et vecteurs propres de l'ACP . . . . .	85
6.2.2	La composition des barycentres . . . . .	86
6.3	Les modèles déterministes . . . . .	87
6.3.1	La méthode des moindres carrés (" <i>least-squares</i> ") . . . . .	89
6.3.2	L'évaluation quantitative de l'erreur . . . . .	90
6.3.3	Modèles non linéaires : l'algorithme de l'inversion totale . . . . .	91
6.4	Les modèles statistiques : les réseaux de neurones . . . . .	92

6.4.1	Introduction . . . . .	92
6.4.2	Présentation . . . . .	93
6.4.3	Le réseau à rétropropagation du gradient ou Perceptron Multi-Couches . . . . .	96
6.4.4	Les réseaux à compétition (récurents) . . . . .	101
6.5	Remarque générale sur les réseaux de neurones . . . . .	107
<b>III Retrouver l'information géologique à partir des diagraphies</b>		<b>109</b>
<b>7</b>	<b>Reconstruire un log minéralogique</b>	<b>111</b>
7.1	Puits 833B, arc volcanique des Vanuatu, Leg ODP 134 . . . . .	113
	Article 2 : "Characterisation of oceanic sediments and basalts from Hole 833B, New Hebrides Island arc (Vanuatu) from downhole measurements" (1997) Alain Rabaute, Louis Briquieu, Peter K. Harvey, and Henri Mercadier, <i>Scientific Drilling</i> , Vol. 6, Number 2, 91-101. . . . .	113
7.2	Sites 947 et 948, prisme d'accrétion de la Barbade, Leg ODP 156 . . . . .	125
7.2.1	Inversion minéralogique de données géochimiques obtenues par analyse en fluorescence X sur des échantillons carottés . . . . .	125
	Article 3 : "Combination of cluster analysis and total inversion of geochemical data as an advantageous way of inferring mineralogy : Example from the North-Barbados accretionary prism, Leg ODP 156" Alain Rabaute, Louis Briquieu, and Mohamed Ramadan, soumis à <i>Geophysical Research Letters</i> . . . . .	125
7.2.2	Utilisation d'un Perceptron Multi-Couches pour reconstruire une minéralogie en continu à partir de mesures LWD : exemple du puits 948A . . . . .	137
	Article 4 : "Inferring mineralogy with neural network classifier : Example using Logging-While-Drilling data from Hole 948A, ODP Leg 156" Alain Rabaute, en préparation . . . . .	137
7.3	Le puits 967E, Mont sous-marin Erathosthenes, Leg ODP 160 . . . . .	153
7.3.1	Introduction . . . . .	153
7.3.2	Géologie du site . . . . .	153
7.3.3	Histoire tectonique . . . . .	155
7.3.4	Interprétation des diagraphie géochimiques et géophysiques, Site 967, puits 967E . . . . .	157
7.4	Le forage BALAZUC-1, Programme <i>Géologie Profonde de la France</i> . . . . .	159
7.4.1	Introduction et cadre géologique . . . . .	159
7.4.2	L'inversion minérale . . . . .	161
7.4.3	Résultats . . . . .	163
<b>8</b>	<b>Retrouver les paramètres pétrophysiques des formations traversées</b>	<b>169</b>
8.1	Dynamique du prisme d'accrétion nord-Barbade, Leg ODP 156 . . . . .	171
	Article 5 : "Circulation et surpression de l'eau interstitielle dans le prisme d'accrétion nord-Barbade : résultats du Leg ODP 156", Pierre Labaume, Pierre Henry, <b>Alain Rabaute</b> , et l'équipe scientifique du Leg ODP 156, <i>C. R. Acad. Sci. Paris</i> , Tome <b>320</b> , Série IIa, 1995, pp. 977-984. . . . .	171
	Article 6 : " <i>In Situ</i> determination of permeability from downhole measurements analysis", André Revil, Delphine Fitzenz, Alain Rabaute, Mohamed Ramadan, Louis Briquieu, soumis à <i>Water Resources Research</i> . . . . .	181

8.2	Structure du réseau poreux dans les formations carbonatées . . . . .	207
	Article 7 : “Pore space components and their partitioning in carbonate rocks” Alain Rabaute, T.S. Ramakrishnan, Stefan Luthi, Marina Polyakov, en attente de la levée de la confidentialité sur les données. Sera soumis à un journal du type <i>Journal of Sedimentary Research</i> . . . . .	207
	<b>Synthèse et perspectives</b>	<b>225</b>
	<b>Annexe A : Outils de diagraphie et paramètres mesurés</b>	<b>229</b>
	<b>Annexe B : Glossaire de quelques termes particuliers utilisés en diagraphie</b>	<b>232</b>
	<b>Annexe C : Tableaux récapitulatifs (éléments majeurs et traces)</b>	<b>234</b>
	Data Report : “Trace and major elements concentrations from ODP Leg 156 deep sea cores : a data report” Alain Rabaute, Peter Blum, Pierre Gaudon, and James F. Allan, <i>Non publié</i> . . . . .	234
	<b>Annexe D : Caractéristiques diagraphiques des minéraux</b>	<b>250</b>
	<b>Annexe E : Présentations et posters</b>	<b>252</b>
	<b>Bibliographie</b>	<b>260</b>

# Table des figures

1.1	Exemple de spectre gamma naturel. . . . .	6
1.2	L'outil NGT pour la mesure de la radioactivité gamma naturelle. . . . .	8
1.3	Profondeur d'investigation latérale calculée pour l'outil NGT . . . . .	9
1.4	Les trois types d'interactions rayonnements gamma–formation en fonction de l'énergie incidente. . . . .	26
1.5	Schéma de l'outil HLDT servant à mesurer la densité et le facteur photoélectrique de la formation. . . . .	28
1.6	Evolution du rapport $Z/A$ pour les éléments chimiques communément rencontrés dans les formations géologiques. . . . .	29
1.7	Domaines d'énergie des neutrons avec les principales interactions associées . . . .	31
1.8	Schéma de la chaîne d'outils composant le <i>Geochemical Logging Tool</i> ou GLT. . .	34
1.9	Exemple de spectres gamma de capture mesurés par le GST, et de spectres standards, isolés pour chaque élément, et mesurés en laboratoire. . . . .	36
1.10	Détail du cycle de mesure du GST en mode capture-tau. . . . .	37
1.11	Schéma de l'outil CNT utilisé pour mesurer l'indice d'hydrogène apparent ou porosité neutron. . . . .	43
2.1	Les principes de focalisation des lignes de courant dans un puits de forage. . . . .	48
2.2	Schéma de l'outil Dual Laterolog ou DLL. . . . .	49
2.3	Principe des sondes à induction. . . . .	50
2.4	Principes de mesure utilisés dans la sonde dipôle–dipôle servant à la mesure de la susceptibilité magnétique de la formation. . . . .	52
2.5	Principes et schéma des outils électriques d'imagerie de micro-résistivité. . . . .	53
3.1	Etapes d'une mesure par résonance magnétique nucléaire. . . . .	58
3.2	Principes de la résonance magnétique nucléaire. . . . .	60
3.3	Traitement du signal de précession pour obtenir la courbe de distribution $T_2$ . . .	61
4.1	Cycle d'acquisition des données en diagraphie. . . . .	69
4.2	Illustration du biais de mesure en diagraphie à partir de la mesure de la micro-résistivité par l'outil FMI. . . . .	70
4.3	Reproductibilité de la mesure de diagraphie . . . . .	71
5.1	Exemple de diagramme x-y permettant un premier partitionnement des données. . . .	76
5.2	Diagramme permettant de séparer les différents types de minéraux argileux en fonction de leur teneur en Al et Th. . . . .	78
5.3	Exemple de dendrogramme obtenu par analyse hiérarchique par groupe. . . . .	79

5.4	Méthode utilisée pour le choix du nombre optimum de barycentres de départ pour le partitionnement d'un jeu de données. . . . .	81
5.5	Exemple de classification en nuées dynamiques utilisant la logique floue. . . . .	83
6.1	Informations immédiate donnée par l'Analyse en Composantes Principales. . . . .	86
6.2	Première interprétation de l'information directe donnée par la classification en nuées dynamiques. . . . .	88
6.3	Exemples de fonctions d'activation utilisées dans les réseaux de neurones. . . . .	94
6.4	Schéma général d'un réseau neuronal de type Perceptron Multi-Couches. . . . .	97
6.5	La descente de gradient sur la surface d'erreur dans un réseau de neurones à rétropropagation du gradient. . . . .	100
6.6	Le problème du surapprentissage d'un réseau de neurones. . . . .	101
6.7	L'apprentissage par compétition entre neurones dans un réseau de neurones interconnecté. . . . .	103
6.8	Exemples de topologies que peut prendre la carte auto-organisatrice de Kohonen. . . . .	104
6.9	Trois formes de la fonction caractérisant le pas d'apprentissage dans le cas de la carte auto-organisatrice. . . . .	105
6.10	Structuration de la carte auto-organisatrice au cours de la phase d'apprentissage. . . . .	106
7.1	Localisation des 4 sites forés sur le mont sous-marin Eratosthènes durant le Leg ODP 160. . . . .	154
7.2	Démonstration de l'effet perturbateur du tubage sur la mesure du fer par le GLT. . . . .	157
7.3	Les évaporites messiniennes vues par les diagraphies. . . . .	158
7.4	Données de diagraphie significatives dans le puits 967E, Leg ODP 160. . . . .	160
7.5	Coupe géologique réalisée lors du forage BA-1. . . . .	161
7.6	Résultats de l'inversion minérale sur le forage BA-1. . . . .	166
7.7	Utilisation des propriétés pétrophysiques pour interpréter les niveaux où l'inversion minérale à échoué. . . . .	167
7.8	Validation des résultats de l'inversion minérale. . . . .	168



# Liste des tableaux

1.1	Abondances et activité gamma relative des éléments radioactifs naturels . . . . .	5
1.2	Abondances en Th et U de quelques minéraux lourds. . . . .	7
1.3	Section efficace de capture des neutrons thermiques, et efficacité normalisée en fonction de l'abondance moyenne des principaux éléments rencontrés dans les formations géologiques. . . . .	33
1.4	Rapports entre les contributions (yields) de différents éléments mesurés par le GST, et utilisés comme indicateurs pour différentes propriétés de la formation. . .	39
1.5	Répétabilité de l'outil GST dans des conditions de mesure normales, donnée pour les divers éléments mesurés. . . . .	40
6.1	Résultats d'une Analyse en Composantes Principales. . . . .	86
6.2	Exemple d'information directe donnée par la classification itérative non-hiérarchique par groupes. . . . .	87
7.1	Composition des barycentres après une analyse en nuées dynamiques sur les données du forage BA-1. . . . .	162
7.2	Paragenèses de départ affectées aux groupes individualisés par l'analyse en nuées dynamiques pour le forage BA-1. . . . .	163
7.3	Compositions des minéraux utilisés pour l'inversion minérale sur le forage BA-1. .	163
A.1	Outils et principaux paramètres mesurés par diagraphie nucléaire. . . . .	230
A.2	Outils et principaux paramètres mesurés par diagraphie électrique. . . . .	230
A.3	Outils et principaux paramètres mesurés par diagraphie magnétique. . . . .	230
A.4	Résolution verticale et pas d'échantillonnage de quelques outils de diagraphie. . .	231
D.1	Caractéristiques diagraphiques des principaux minéraux des formations sédimentaires. . . . .	251



# Introduction

Que de temps a passé depuis l'idée (géniale, peut-on le dire!) de Conrad Schlumberger en 1912, d'utiliser les propriétés électriques intrinsèques des roches pour identifier la lithologie du sous-sol! En fait, c'était il y a 70 ans qu'aidé de son frère Marcel, il mène son idée jusqu'au bout : le 5 Septembre 1927, le premier "*carottage électrique*" est effectué à Pechelbronn, en Alsace. Il s'agissait de mesurer en continu la résistivité électrique des parois du puits de forage. Cet enregistrement apporte immédiatement des informations inédites, en permettant notamment une différenciation qualitative des différentes unités lithologiques, la connaissance des fluides présents dans la formation, et donc une évaluation de sa porosité. Ces mesures rendent possible les corrélations puits à puits, donc le suivi d'un horizon lithologique entre les sondages, et une image en trois dimensions de la géologie du sous-sol. Malgré des progrès constants et souvent spectaculaires, les techniques de diagraphies ne se sont imposées au monde pétrolier que vers les années 50, à une époque où la précision des mesures réalisées dans les forages, les économies de temps et d'argent qu'elles permettent de réaliser par rapport à un carottage en continu, finissent par convaincre les pétroliers.

Le monde académique a, quant à lui, beaucoup plus longtemps douté (et doute encore) de la capacité de diagraphies à apporter une information géologique complémentaire à l'observation des carottes et même originale. Cela provient sans doute de plusieurs facteurs inhérents à la technique de mesure. Les méthodes de mesure par diagraphie reposent en effet sur des principes physiques plus ou moins complexes, faisant intervenir des domaines aussi variés que la physique nucléaire, électrique, le magnétisme, ou l'acoustique. De plus, la diagraphie ne donne que rarement un résultat direct (d'où l'engouement pour l'imagerie de micro-résistivité des parois du puits par l'outil FMS à travers la communauté scientifique. Ces mesures sont directement interprétables quelques minutes seulement après leur acquisition), et les données sont soumises à des traitements préalables, transformations, convolutions, calibrations, dans lesquels l'empirisme a souvent un poids non négligeable sur la donnée finale. Et même alors, la quantité de données, le nombre impressionnant de variables mesurées qu'il est difficile d'appréhender simultanément, la présentation rébarbative sous forme de serpents ("*wiggly lines*" disent les anglo-saxons moqueurs) où seul l'œil averti sait séparer les variations réelles du bruit de la mesure, sont autant de barrières qui ont contribué à affubler bien injustement la diagraphie d'une définition tenant en deux mots : "*boîte noire*". La première partie de ce manuscrit traite donc assez simplement des principes physiques servant de fondations à la mesure de diagraphie, dont il est à mon avis indispensable d'avoir connaissance pour apprécier la justesse et la précision de la mesure, et surtout pour savoir l'utiliser en relation avec d'autres mesures effectuées de façon plus conventionnelle, sur échantillons par exemple. De plus, détailler les principes de base permet de s'apercevoir du manque de robustesse de certains traitements effectués (par exemple, le modèle de fermeture permettant d'obtenir les pourcentages d'oxydes à partir des mesures spectrales du GST), ou de l'archaïsme de certaines méthodes utilisées pour la calibration. Il serait sans doute possible de les améliorer en utilisant des méthodes moins rigides, et donnant une idée plus précise de l'erreur

sur la donnée obtenue.

Cette étude tente par des exemples de montrer que cette technique largement utilisée maintenant dans le monde pétrolier, bien que complexe à mettre en oeuvre (j'ai pu m'en rendre compte lors de ma participation aux deux Leg ODP 156 et 160), peut donner de très bons résultats lorsqu'elle est appliquée à des objectifs scientifiques : on peut citer le programme *Géologie Profonde de la France*, qui a réuni autour d'une problématique scientifique commune le Bureau de Recherches Géologiques et Minières, le ministère de la Recherche, les Universités et le Centre National de la Recherche Scientifique, avec la collaboration de nombreuses sociétés comme Schlumberger ou la CGG. Cet ouvrage comporte une étude de l'utilité des données de diagraphies obtenues dans le forage Balazuc-1 au cours de ce programme pour retrouver une information lithologique et minéralogique. Aux Etats-Unis, le programme de forage *Cajon Pass*, en Californie, ou bien encore le fameux *KTB, Continental Drilling Program* en Allemagne, et puis bien sûr le méga-programme international *Ocean Drilling Program*, ont introduit la diagraphie dans leurs objectifs. ODP a récemment (Leg 156, 1994) utilisé la technique de pointe de Logging-While-Drilling (diagraphie en cours de forage) avec succès dans les environnements tectoniquement actifs (prismes d'accrétion), et effectue même maintenant des campagnes axées sur la diagraphie (Leg 171A).

Dans la seconde partie de ce travail, je détaille différentes méthodes de statistique multivariée et d'inversion minérale pouvant être utilisées pour extraire l'information géologique et minéralogique de la mesure par diagraphie. Comprendre le modèle que l'on utilise est au moins aussi important que l'acquisition de la mesure. A cette occasion, je tiens à rendre un hommage tout particulier au Professeur Peter Harvey et son équipe du Borehole Research Group de l'University of Leicester. Sa maîtrise des outils statistiques ainsi que sa connaissance approfondie des méthodes d'inversion minérale, et la disponibilité avec laquelle il m'a appris à les utiliser, ont été pour moi une aide formidable. Cette compréhension des méthodes d'analyse de données multivariées devient cruciale lorsque l'on utilise des techniques avancées comme les réseaux de neurones, où la phase d'entraînement conditionne de façon importante le résultat que l'on obtient. Il est aussi extrêmement important de quantifier et de *rendre compte* des erreurs associées aux mesures de diagraphies, ainsi que les erreurs associées aux solutions obtenues avec ces méthodes d'analyse. Souvent en effet, des modèles approximatifs ont été développés par l'industrie pétrolière, pour déterminer la porosité, la saturation, le volume d'argiles. Par exemple, on fait souvent rimer volume d'argiles avec un rayonnement gamma naturel important, ce qui est bien sûr réducteur, la concentration en aluminium par exemple étant un bien meilleur indicateur de l'argilosité d'une formation. Une mesure de la porosité utilisée extensivement dans les diagraphies scientifiques est la porosité neutron, mesurant principalement les atomes d'hydrogène car cet élément a le meilleur rapport abondance/section de capture des neutrons thermiques. On associe classiquement l'hydrogène aux fluides présents dans la formation, mais qu'en est-il de la contribution au signal du fluide adsorbé à la surface des grains, ou des  $\text{OH}^-$  du réseau cristallin des minéraux argileux ? quelle est l'erreur associée à la mesure dans ce cas ? Et peut-on voir dans cet enregistrement la contribution d'autres éléments moins abondants, mais dont la section de capture est plusieurs ordres de grandeur supérieure à celle de l'hydrogène (Gd, Ba, Cl) ? Retrouver une minéralogie en continu à partir des diagraphies de la minéralogie peut aider à répondre à ces questions. La troisième partie, présentée pour une grande part sous forme d'une compilation d'articles, donne quelques résultats qui tentent de rendre compte de l'aide précieuse que les diagraphies peuvent apporter à la connaissance à différentes échelles, en continu et *in situ* des propriétés des formations géologiques.

Dans les pages suivantes, le symbole <sup>™</sup> signale une Marque Déposée de la société Schlumberger.

Première partie

La diagraphie



# Chapitre 1

## La diagraphie nucléaire

### 1.1 Introduction

Dans la nature, les phénomènes radioactifs se classent en 3 grandes classes : les radiations  $\alpha$ ,  $\beta$  et  $\gamma$ . Celles-ci sont provoquées par la modification de l'état d'un atome ou de sa désintégration. Les radiations  $\alpha$  et  $\beta$  consistent en particules chargées, les premières étant identiques au noyau d'hélium (2 protons et 2 neutrons), les secondes étant des électrons rapides, positifs (positrons) ou négatifs. De ce fait, elles vont surtout interagir dans la formation avec les électrons des atomes qui la composent. Ces particules  $\alpha$  et  $\beta$  vont rapidement perdre de l'énergie au profit des électrons qu'elles rencontrent, et n'auront donc qu'une incidence spatiale limitée. Les radiations  $\alpha$  et  $\beta$  ne sont pas utilisées en diagraphie pour cette raison.

Par contre, les radiations  $\gamma$  sont des paquets d'énergie électromagnétique (photons) très énergétique. Ils correspondent, dans un noyau atomique, à la transition d'un état quantique à un autre, qui lui est inférieur. Ce changement d'état se manifeste sous la forme d'un photon dont l'énergie est égale à la différence entre les énergies des deux états quantiques. Les énergies des rayons gamma sont de l'ordre du  $keV$  ( $10^3$  eV) ou du  $MeV$  ( $10^6$  eV). Contrairement aux radiations  $\alpha$  et  $\beta$ , les rayons gamma ont un rayon d'action beaucoup plus étendu. Ils sont à l'origine de trois types d'outils de diagraphie nucléaire :

- ceux qui mesurent la radioactivité gamma naturelle
- ceux producteurs de rayons gamma, à l'aide notamment de sources radioactives ( $^{137}\text{Cs}$  ou  $^{60}\text{Co}$ ), et qui utilisent les interactions de ces rayonnements avec les atomes de la formation
- ceux qui bombardent la formation de neutrons rapides produits par une source à neutrons, et utilisent les interactions de ces neutrons avec la formation.

Les principes de ces trois types d'outil ainsi que la physique qui s'y rapporte sont expliqués en détail dans les pages suivantes.

Auparavant, il est nécessaire de rappeler certaines propriétés importantes des radiations nucléaires sur lesquelles est basé le principe même de la mesure de ces radiations.

### 1.2 Désintégration radioactive et statistiques

La désintégration radioactive d'un noyau atomique est un processus qui dépend seulement du **temps**, et, ce qui est important, qui est complètement **indépendant de facteurs extérieurs**, tels que par exemple la désintégration d'autres noyaux à proximité, ou bien la concentration en éléments radioactifs. Si l'on considère une population  $N_0$  d'atomes radioactifs identiques, le

nombre  $dN$  se désintégrant dans un intervalle de temps  $dT$  sera

$$dN = -\lambda \cdot N \cdot dt$$

où  $\lambda$  est la *constante de désintégration* propre à l'atome considéré, soit sa probabilité de se désintégrer par unité de temps. On en déduit la loi générale de la désintégration radioactive

$$N_t = N_0 e^{-\lambda t}$$

où  $N_t$  est le nombre d'atomes restant à un instant  $t$ ,  $N_0$  le nombre initial d'atomes. On peut définir la *demi-vie*, qui est le temps nécessaire à la décroissance radioactive de la moitié d'une population  $N_0$  présente à un instant donné, comme étant

$$t_{1/2} = \frac{0,693}{\lambda}$$

La désintégration radioactive peut être décrite par la loi de distribution binomiale. Soit un évènement discret qui a une probabilité  $P$  de se produire lors d'une observation. La loi de distribution binomiale décrit la probabilité  $P_x$  que cet évènement se produise  $x$  fois lors de  $z$  observations. Dans notre cas,  $z$  est le nombre de particules se désintégrant, qui est très grand comparé à la probabilité  $P$  d'observer la désintégration d'un isotope dans une unité de temps. Cette condition permet la simplification de la loi binomiale qui devient

$$P_x = \bar{\mu}^x \frac{e^{-\bar{\mu}}}{x!} \quad (1.1)$$

$P_x$  donne alors la probabilité de voir  $x$  désintégrations dans un temps donné alors que l'on attend un nombre moyen de  $\bar{\mu}$  désintégrations. L'équation 1.1 est une *distribution de Poisson*. Elle s'applique très bien pour le comptage de phénomènes nucléaires tel que la désintégration radioactive, où un grand nombre d'atomes participent à un grand nombre d'observations, bien que le nombre de cas positifs (désintégrations) soit faible. Lorsque le taux de comptage est élevé, la loi de distribution devient normale.

La déviation standard des phénomènes obéissant à la loi de Poisson est directement proportionnelle à la moyenne

$$\sigma = \sqrt{\bar{\mu}}$$

On voit bien que le moyen de réduire l'erreur statistique de comptage sera d'augmenter le nombre de phénomènes observés, soit par l'utilisation de compteurs plus efficaces, soit en augmentant le temps de comptage, c'est à dire en diminuant la vitesse de l'outil.

## 1.3 Détection de la radioactivité gamma naturelle

### 1.3.1 Introduction

Dans la nature, il existe seulement 3 éléments possédant des isotopes qui contribuent à la radiation gamma naturelle, car leur demi-vie est comparable à l'âge de la planète :  $^{40}\text{K}$ , dont la demi-vie est de  $1,3 \cdot 10^9$  an,  $^{238}\text{U}$ , avec  $4,4 \cdot 10^9$  an, et  $^{232}\text{Th}$ , avec  $1,4 \cdot 10^9$  an. Le fait que ces éléments soient présents dans les formations géologiques, et souvent en relation directe avec la minéralogie ou la lithologie (Deer et al., 1978; Grim, 1968; Pettijohn, 1975), a suscité très tôt un grand intérêt pour la diagraphie de radioactivité naturelle de type gamma. D'abord utilisée comme indicateur de la richesse en argile d'une formation, son champ d'application s'est élargi avec l'introduction de la spectrométrie qui permet de donner une estimation des teneurs en U, Th



	K	Th	U
Abondances relatives dans la croûte terrestre <sup>+</sup>	2.59%	≈12 ppm	≈3 ppm
Radiation gamma par unité de poids et unité de temps*	1	1300	3600

<sup>+</sup>Serra (1984), Serra et al. (1980)

\*Adams and Weaver (1958)

TAB. 1.1 – Abondances et activité gamma relative des éléments radioactifs naturels

et K. Dès lors, elle devient une aide pour l'interprétation géochimique, la détection de circulations de fluides, ou la reconnaissance du minéral argileux. Néanmoins, même si la radioactivité gamma totale mesurée dans une formation (SGR ; voir Tableau A.1) sert toujours comme indicateur de l'argilosité d'une formation, Herron and Herron (1996) ont montré que Al était un bien meilleur indicateur de cette propriété.

On connaît bien maintenant les énergies caractéristiques des pics d'émission gamma de ces isotopes. Néanmoins, les rayons gamma provenant de la formation devront traverser celle-ci, de même que la boue de forage, avant d'atteindre le détecteur, et perdront de l'énergie du fait de l'*effet Compton* (collisions des photons haute-énergie avec les électrons présents dans l'espace traversé). Ainsi, les énergies d'émission uniques des différents isotopes radioactifs naturels vont être dégradées (bien que l'aspect gaussien du comptage reste préservé), et l'on va observer un *spectre* d'émission plus ou moins complexe selon la concentration des éléments radioactifs présents. Un tel spectre est présenté Figure 1.1. Le <sup>40</sup>K émet un seul pic d'énergie à 1460 keV. Par contre, uranium et thorium sont chacun à l'origine d'une chaîne de désintégration faisant intervenir de nombreux isotopes d'autres éléments, ayant une demi-vie très courte (de la milli-seconde à la semaine), et menant à un isotope stable du plomb (respectivement <sup>206</sup>Pb et <sup>208</sup>Pb). La Figure 1.1 montre les trois principaux pics d'émission contribuant à la série de <sup>238</sup>U, principalement <sup>214</sup>Bi (à 610, 1120 et 1764 keV), et ceux contribuant à la chaîne du <sup>232</sup>Th, <sup>208</sup>Tl et <sup>228</sup>Ac (respectivement à 584 et 2615 keV, et à 912 et 966 keV).

Quantitativement, le **potassium** est de loin l'élément le plus abondant parmi les trois éléments radioactifs naturels (sa concentration est beaucoup plus élevée (1000 à 10000 fois plus) dans la croûte terrestre que l'uranium et le thorium (Tableau 1.1) ; on mesure son abondance en *pourcent* alors que uranium et thorium sont mesurés en *ppm*), mais sa contribution à la radioactivité totale par rapport à son poids va être faible (Tableau 1.1). En fait, la contribution des trois éléments est à peu près du même ordre de grandeur. Le potassium est présent dans les grès par les feldspaths potassiques, et dans les évaporites (sylvite entre autres). Il apparaît aussi comme ion de substitution dans l'espace inter-feuillets du réseau cristallin d'une classe de minéraux argileux très commune, à structure proche de celle des micas, à laquelle appartiennent l'illite, la glauconite et la pyrophyllite. Le thorium et l'uranium seront plus rares. Des différences majeures de chimisme et de réactivité permettent de mettre en relation leur présence avec des lithologies ou des minéraux particuliers.

Le **thorium** est présent à la valence 4+ dans la nature, et a un degré d'oxydation très stable, et de fait, ses composés sont tous insolubles. Il sera très peu mobile, facilement hydrolysé, et aura tendance à se concentrer dans les minéraux et roches résiduels lors d'un lessivage. Les sources de thorium sont surtout les roches magmatiques plutoniques (où le rapport Th/U est relativement constant à 3,5–4) dans des minéraux lourds comme la monazite (la variété *chéralite* en contient jusqu'à 33%), l'allanite, le zircon, le sphène, l'épidote ou la thorite (voir Tableau 1.2). Dans les sédiments, le thorium sera d'origine détritique et présent dans les bauxites, les bentonites

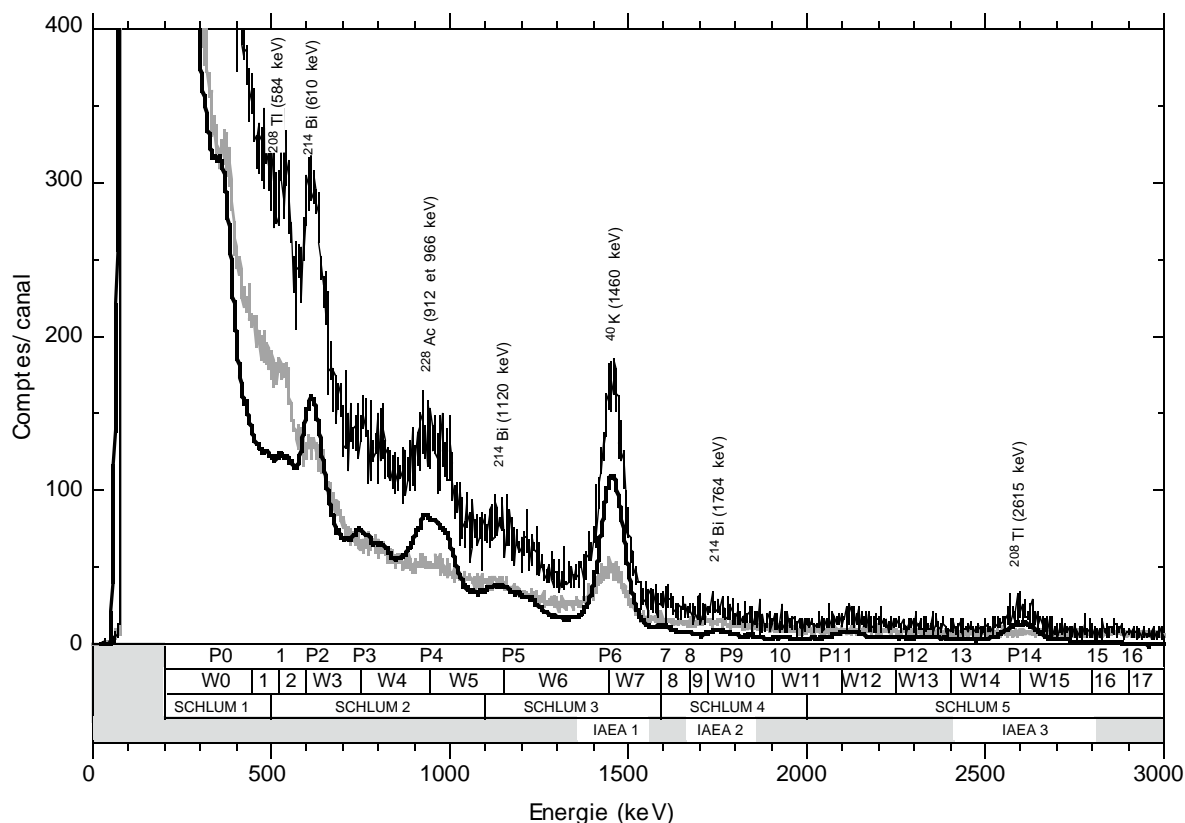


FIG. 1.1 – Exemple de spectre gamma naturel (adapté de Blum et al. (1997)). Sont reportés les isotopes majeurs contribuant au rayonnement gamma et leur énergie. P1-16 sont les pics identifiés dans l'étude Blum et al. (1997), W1-17 sont les fenêtres de comptage indentifiées par calcul du minimum relatif entre ces pics pris deux à deux. SCHLUM 1-5 sont les 5 fenêtres classiquement utilisées par Schlumberger pour l'analyse spectrale du rayonnement gamma naturel, tandis que IAEA 1-3 sont celles définies et utilisées par l'International Atomic Energy Agency. Le spectre original non corrigé acquis pendant 4 heures dans 2048 canaux est identifié par la courbe noire ; le bruit de fond acquis dans les mêmes conditions est la courbe grisée ; la courbe en gras représente le spectre mesuré corrigé du bruit de fond, et ensuite lissé par fenêtre glissante (30 canaux).

(roche composée à 90% d'argile (smectite) formée par l'altération de formations cinéritiques), les formations volcano-sédimentaires. Les argiles, dont presque tous les types en contiennent dans des proportions variables mais significatives (quelques dizaines de ppm ; voir Hassan et al. (1976)), piègent le thorium sous forme de cation de substitution dans les espaces inter-feuillets. Il sera plus abondant dans les argiles d'origine continentale (kaolinite) que dans les argiles d'origine marine (glauconite).

L'**uranium** sera facilement oxydé en ion uranyle  $UO_2^{2+}$ , cet ion étant très soluble, l'uranium sera facilement mobilisable. L'ion uranyle se complexera dans des conditions réductrices avec la matière organique dans les schistes argileux ("shales") ou bien précipitera sous certaines conditions de pH sous forme d'uraninite  $UO_2$ . Il pourra se trouver en concentrations non négligeables dans les minéraux lourds de type zircon ou monazite, dont le réseau cristallin est assez solide pour supporter les effets de sa désintégration (voir Tableau 1.2), et sera un constituant majeur du xénotime, un phosphate d'yttrium, qui peut en contenir jusqu'à 4%. Dans les roches sédimentaires, l'uranium sera souvent associé à la présence de matière organique ou de minéraux

Nom	Type	Th (ppm)	U (ppm)
Thorite	ThSiO <sub>4</sub>	25–63	
Monazite	variété chérolite (Ce,La,Th)PO <sub>4</sub>	2.5x10 <sup>4</sup> –25x10 <sup>4</sup>	500–3000
Allanite	aluminosilicate de (Ca,Ce,Th,Fe,Mg)	500–5000	30–700
Zircon	ZrSiO <sub>4</sub>	100–2500	300–3000
Sphène	CaTiSiO <sub>4</sub>	100–600	100–700
Epidote	aluminosilicate de (Ca,Fe,Al)	50–500	20–50

TAB. 1.2 – Abondances en Th et U de quelques minéraux lourds (d’après Serra et al. (1980)).

lourds détritiques (“shales”, volcanosédimentaire). Il sera présent dans une moindre mesure dans les minéraux argileux, les phosphates, les dolomies.

### 1.3.2 Principe

**Le détecteur** La Figure 1.2 montre un schéma de l’outil NGT (voir Tableau A.1) conçu et utilisé par sch (1984). Il comprend un détecteur à scintillation dont le principe de fonctionnement est aussi schématisé Figure 1.2. Le détecteur est classiquement un cristal d’iodure de sodium dopé au thallium, NaI(Tl). Cette combinaison possède de bonnes propriétés d’absorption des rayons gamma, ainsi qu’une fréquence de scintillation suffisamment rapide (0,23  $\mu$ s) pour autoriser la spectroscopie à des taux de comptage élevés. Tout rayon gamma provenant de la formation entrant en contact avec les atomes du cristal provoque une avalanche d’électrons secondaires qui sont finalement capturés par les impuretés du réseau, en l’occurrence le thallium. La capture est la cause d’une émission de photons. Chaque émission est amplifiée par un tube photomultiplicateur et transformée en une impulsion électrique. On mesure alors le nombre de coups par unité de temps. L’avantage d’utiliser ce cristal est de pouvoir faire de la spectroscopie gamma (voir Figure 1.1), car la valeur de l’impulsion électrique est directement proportionnelle à l’énergie du rayon gamma incident. Ceci permettra dans le cas de la diagraphie par rayonnement gamma induit (voir Section 1.5.2) de retrouver la source du rayon gamma induit, et donc sa nature.

**Calibration** La radioactivité gamma est mesurée en **API** ou **gAPI**, unité développée par l’American Petroleum Institute à partir de standards dans le but d’éliminer l’influence de la taille et de la sensibilité du détecteur à scintillation lors de la mesure. Une formation radioactive artificielle a été construite à Houston, Texas. Cette formation est composée de deux couches faiblement radioactives A entourant une troisième couche très radioactive B (dont la composition est approximativement 4% K, 24 ppm Th, and 12 ppm U). Une unité API correspond au 1/200<sup>e</sup> de la différence de radioactivité entre les couches A et B, la couche B étant calibrée pour donner 200 API. Sur le terrain, l’outil est étalonné par rapport à une source contenant de la monazite au moyen de laquelle on contrôle les taux de comptage dans chaque fenêtre.

**Corrections** Le signal que reçoit le NGT est affecté par le diamètre du trou (voir Section 1.4.2) et la profondeur d’invasion. Le type de boue de forage utilisée joue un rôle majeur : sa densité peut induire une diffusion par effet Compton des rayons gamma (boue à barytine), alors que les boues à KCl utilisées pour stabiliser les murs du puits perturbent le signal du potassium. Leurs effets ont été modélisés par Wahl (1983) et étudiés par Ellis (1982). Ellis (1982) propose une méthode pour corriger le signal en calculant ces effets pour chaque fenêtre de mesure et en les

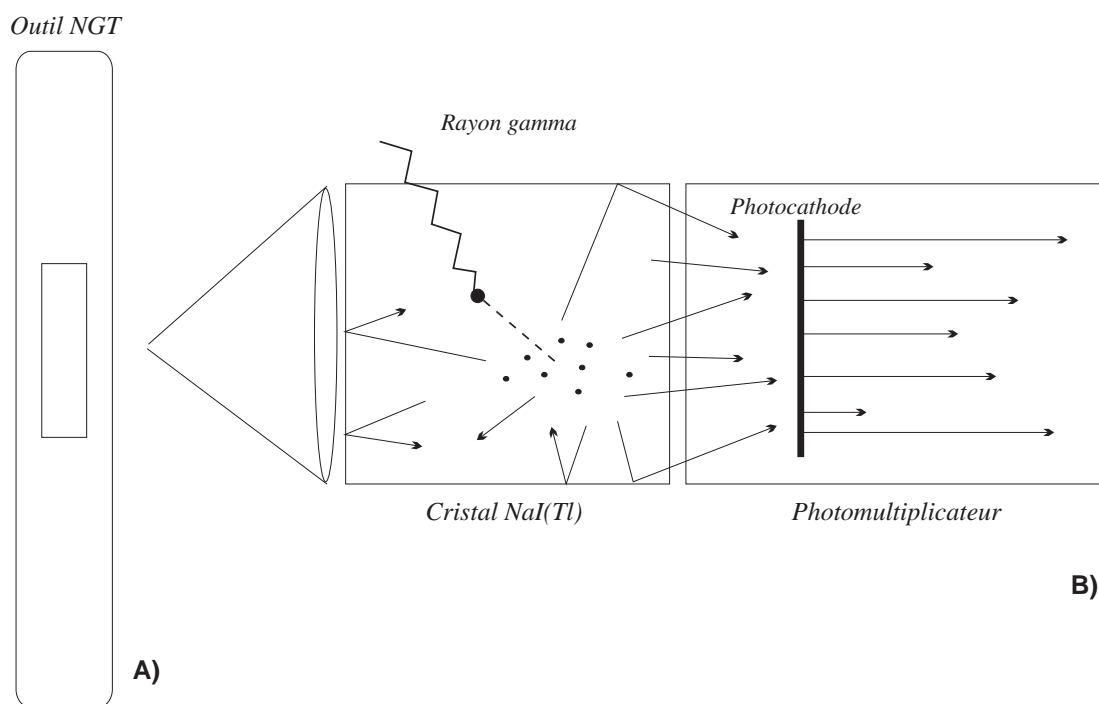


FIG. 1.2 – A. L’outil NGT pour la mesure de la radioactivité gamma naturelle. B. Détail du détecteur.

supposant constants sur tout l’intervalle mesuré. On a aussi vu précédemment (Section 1.2) que l’émission de rayons gamma est aléatoire, aussi choisit-on un temps de mesure suffisant pour que la variation statistique ne dépasse pas 2,5%. Une autre source de perturbation est l’apparition de pics d’interférence proches du pic principal d’un des trois éléments radioactifs naturels, d’où l’importance de la définition des limites des fenêtres de mesure pour minimiser les erreurs sur les abondances finales (voir Section 1.3.4 ; voir Blum et al. (1997) en fin de section 1.3.4 pour une discussion sur ce problème).

**Rayon d’investigation et résolution verticale** Desbrandes (1968), et plus récemment Wahl (1983) ont cherché à évaluer ces deux paramètres importants de la sonde gamma. En particulier, Wahl (1983) a utilisé la méthode de simulation numérique Monte Carlo pour déterminer jusqu’à quelle profondeur effective l’outil NGT “voit” dans la formation. Il considère d’abord la réponse pour un seul élément responsable du rayonnement gamma naturel, avant d’étudier la réponse d’une formation complexe incorporant U, Th et K. La figure 1.3 montre le signal intégré et normalisé  $J(r)$  produit par une formation de radioactivité connue, en fonction de la profondeur d’investigation. Pour le potassium, on remarque que 90% du signal idéal (pic discret à 1,46 MeV ; courbe pointillée) provient d’un cercle de 15 cm de rayon. Dans le cas où l’on considère tout le signal après effet Compton (courbe en gras), la profondeur d’investigation augmente de quelques centimètres. La courbe intermédiaire représente le signal provenant seulement de la fenêtre de mesure W3 et correspondant uniquement à l’émission du  $^{40}\text{K}$ . La profondeur d’investigation augmente lorsque l’on considère les basses énergies du spectre d’émission car les rayons provoquant ces impulsions ont subi plusieurs collisions, et donc ont plus de chance de venir de plus loin dans

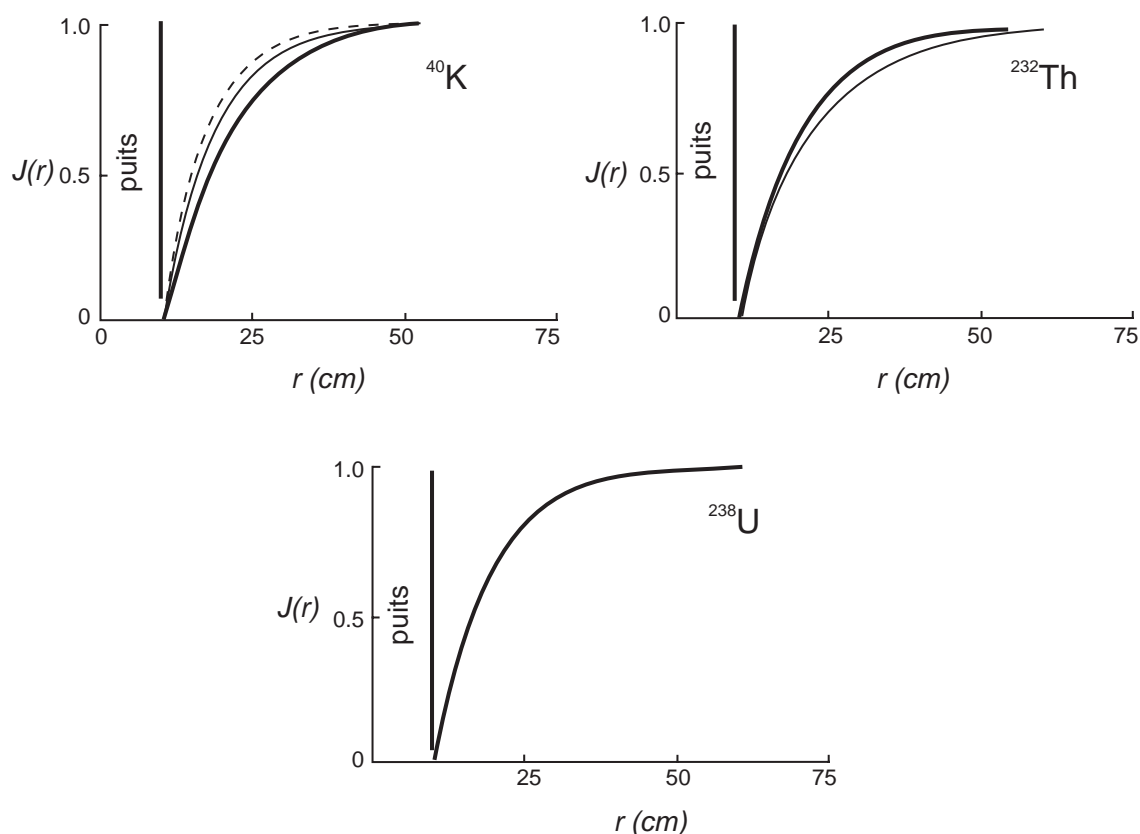


FIG. 1.3 – Modélisation de la réponse de la sonde gamma en fonction de la distance latérale, utilisant la méthode Monte Carlo (d'après Wahl (1983)).  $J(r)$  représente la contribution cumulée de la formation en fonction de la profondeur latérale. Courbe pointillée : réponse calculée pour un signal discret à 1,46 MeV. Courbe en gras : réponse calculée avec effet Compton en considérant tout le spectre d'énergie. Courbe continue : même chose mais en limitant le signal à la seule fenêtre de mesure correspondant au pic majeur pour l'élément considéré (W3, 1100–1600 keV pour K, et W5, 2000–3000 keV pour Th)

la formation. Figure 1.3 montre le même graphique pour U et Th. Pour le thorium, la profondeur d'investigation est plus grande si l'on considère seulement le signal provenant de la fenêtre W5 contenant le pic d'émission majeur. Pour cet élément, la plupart du signal provient de la zone basse énergie du spectre, produite surtout par les collisions entre le signal haute énergie de la fenêtre W5 et la formation, d'où un effet inverse à celui du potassium. Pour une formation contenant les trois isotopes, la profondeur d'investigation dépendra aussi, bien que plus faiblement, de la densité de la formation, ainsi que du diamètre du puits. En général, on considère qu'elle est comprise entre 5 et 25 cm.

La *résolution verticale* peut être définie comme la réponse d'un détecteur en fonction de la profondeur dans le cas où la radioactivité est confinée dans un banc infiniment mince (Wahl, 1983). Elle va être fonction de l'énergie du rayonnement, le diamètre du puits, la densité de la formation, et va être influencée par la boue de forage. Généralement, la résolution verticale sera voisine de 50 cm.

L'effet Compton est la principale source de perturbation des rayons gamma. Cependant, dans le cas où la formation contient des éléments à fort numéro atomique ( $>20$ ), l'effet photoélectrique (voir Section 1.4.1) peut devenir non-négligeable. Wahl (1983) démontre que l'effet photoélec-

trique affecte surtout la partie basse énergie du spectre gamma et que les variations du signal gamma total dues à la lithologie sont de l'ordre de 1% dans les formations ayant un index photoélectrique  $P_e$  (voir Section 1.4.3) inférieur ou égal à 5.

### 1.3.3 Mesure de la radioactivité naturelle totale

Le Tableau A.1 récapitule les différents paramètres mesurés par l'outil NGT détaillés ci-dessous :

- SGR (*Spectral Gamma Ray*) : le rayonnement gamma total, corrigé du bruit de fond et des interférences citées plus haut. Il est exprimé en unités API.
- CGR (*Computed Gamma Ray*) : c'est le rayonnement gamma total SGR sans la contribution de l'uranium. Il est exprimé en unités API.

SGR est utilisé principalement pour le calage en profondeur des profils diagraphiques entre eux ou avec les carottes, ou bien pour de la corrélation puits à puits. On profite de l'information spectrale pour calculer le signal "hors-uranium" CGR en ne considérant que les contributions de K et Th. Ce paramètre donne une information plus exacte sur le volume d'argile présent dans les formations car la présence d'uranium n'est pas forcément corrélée avec l'argilosité (voir Section 1.3.1).

### 1.3.4 Mesure des concentrations en U, K and Th

Le principe de mesure est expliqué en détail par Serra et al. (1980). L'amplitude du voltage de sortie du photomultiplicateur couplé au cristal de NaI du détecteur du NGT est directement proportionnelle à l'énergie du rayon gamma incident (Section 1.3.2), de même que la quantité de rayons gamma émise est proportionnelle aux concentrations d'éléments radioactifs qui leur ont donné naissance (Section 1.2). Les impulsions sont classées selon leur énergie par un analyseur multi-canaux dans un histogramme de fréquence comprenant 5 classes W1 à W5 (Figure 1.1) couvrant tout le spectre gamma de 0,18 à 3,0 MeV. La procédure servant à calculer les abondances relatives en U, Th et K à partir de la mesure des 5 fenêtres, est détaillée dans Ellis (1987). Elle passe par la construction d'une matrice  $\bar{A}$  reliant les taux de comptage de chaque fenêtre avec les concentrations en U, Th et K, à partir de formations artificielles calibrées, soit

$$W_i = A_i Th + B_i U + C_i K + r_i$$

où  $Th$ ,  $U$  et  $K$  sont les concentrations élémentaires dans la formation,  $r_i$  est un facteur correspondant aux erreurs statistiques, et  $W_i$  est le taux de comptage de la fenêtre  $i$ . Les coefficients d'étalonnage  $A_i$ ,  $B_i$  et  $C_i$  pour chaque fenêtre  $i$  sont obtenus dans le puits d'étalonnage décrit plus haut (voir Section 1.3.2) et forme la matrice  $\bar{A}$ . Sous forme matricielle, cela donne

$$\begin{bmatrix} W_1 \\ W_2 \\ W_3 \\ W_4 \\ W_5 \end{bmatrix} = \bar{A} \times \begin{bmatrix} Th \\ U \\ K \end{bmatrix}$$

Les résidus  $r_i$  pour chaque fenêtre seront les taux de comptage mesurés comparés à une mesure non-bruitée.

$$\sum_{i=1}^5 r_i^2 = \sum_{i=1}^5 \frac{1}{W_i} (W_i - A_i Th - B_i U - C_i K)^2 = r^2$$

Les moindres carrés faisant intervenir une pondération en  $1/W_i$  sont classiquement utilisés pour minimiser la somme du carré des résidus  $r^2$  (Bevington, 1969) : La solution est de la forme (Ellis, 1987)

$$\begin{bmatrix} \text{Th} \\ \text{U} \\ \text{K} \end{bmatrix} = \bar{m} \times \begin{bmatrix} W_1 \\ W_2 \\ W_3 \\ W_4 \\ W_5 \end{bmatrix}$$

où les éléments de  $\bar{m}$  sont une combinaison des éléments de  $\bar{A}$ . La matrice des poids  $1/W_i$  peut représenter le taux de comptage moyen attendu dans une formation type, ou bien peut être calculée pour chaque pas de mesure sur la base des taux de comptage mesurés (l'erreur sur  $W_i$  est en  $\sqrt{W_i}$ , et la variance est en  $W_i$ ).

A l'occasion de la campagne océanographique 156 d'Ocean Drilling Program, Blum et al. (1997) ont acquis 79 spectres du rayonnement gamma naturel complets (0,1 à 3000 keV) à partir de sections carottées à travers la zone de décollement du prisme d'accrétion Nord-Barbade. Parallèlement, les concentrations en U, Th et K ont été mesurées par fluorescence X et ICP-MS sur des échantillons représentatifs prélevés sur les mêmes sections carottées, en vue de calibration. L'étude qui suit, publiée dans le volume *Scientific Results* 156 d'ODP, a été menée dans le but de discuter la meilleure façon d'extraire l'information spectrale contenue dans les spectres de rayonnement gamma naturel.

### Article 1 :

Shipley, T.H., Ogawa, Y., Blum, P., and Bahr, J.M. (eds), 1997  
*Proceedings of the Ocean Drilling Program, Scientific Results*, Vol. 156, pp 183–195

Analysis of natural gamma-ray spectra obtained from sediment cores with  
the shipboard scintillation detector of the Ocean Drilling Program :  
example from Leg 156

Peter Blum, **Alain Rabaute**, Pierre Gaudon, and James F. Allan





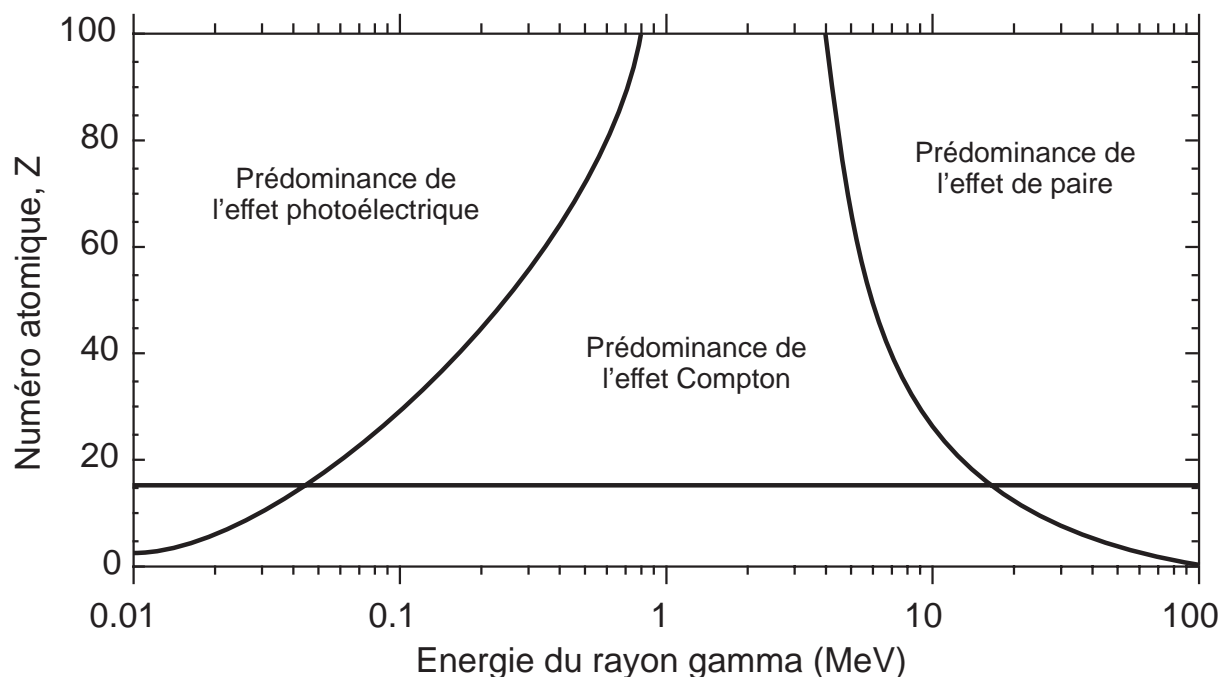


FIG. 1.4 – Domaines de prédominance des trois types d'interactions entre les rayons gamma et la formation en fonction de l'énergie du rayonnement incident et le numéro atomique de l'élément réagissant avec le rayonnement gamma. La ligne horizontale représente un numéro atomique de 16, qui est la valeur moyenne relative des éléments le plus souvent rencontrés dans les formations géologiques.

## 1.4 Interaction entre rayons gamma et formation

### 1.4.1 Principe

L'outil HLDT (voir Tableau A.1) de Schlumberger utilise les interactions entre les rayons gamma et les atomes de la formation pour mesurer la *densité électronique*  $\rho_e$  de la formation, à partir de laquelle on pourra ensuite calculer la *masse volumique*  $\rho_b$ , désignée RHOB, et le *facteur* ou *index d'absorption photoélectrique*  $P_e$ , désigné PEF en interprétation. Pour comprendre le principe de la mesure, il est nécessaire de revenir sur les principales interactions entre un rayonnement gamma et les atomes de la formation.

La Figure 1.4 récapitule les trois principales interactions entre un rayonnement gamma et la matière : l'*effet photoélectrique*, l'*effet Compton*, et l'*effet de paire*. Les énergies du rayonnement gamma émis par les sources radioactives en diagraphie étant toujours inférieures à 1 MeV, on peut éliminer immédiatement la troisième interaction. Les deux facteurs qui vont décider entre effet Compton et effet photoélectrique sont l'énergie du rayon gamma incident et le numéro atomique moyen de la formation considérée. Les deux interactions font intervenir les électrons des atomes de la formation.

**L'effet Compton** Ce sera l'effet dominant pour des énergies gamma supérieures à 100 keV. Le rayon gamma heurte un électron  $e^-$ , lui transférant une partie de son énergie initiale  $E^0$ , et ricoche d'un angle  $\theta$  avec une nouvelle énergie  $E'$ . Dans une formation composée d'atomes de poids atomique  $A$  et de numéro atomique  $Z$ , on définit un coefficient d'absorption linéaire, qui

sera dépendant de la *section efficace macroscopique de capture Compton*  $\Sigma_{Co}$ , définie comme :

$$\Sigma_{Co} = \sigma_{Co} \frac{N_{Av}}{A} \rho_b Z \quad (1.2)$$

où  $\sigma_{Co}$  est la section efficace de capture Compton pour l'élément de poids atomique  $A$  et de numéro atomique  $Z$  (Ellis, 1987).  $\sigma$  est un paramètre très important dans toute les interactions radiation/matière. Il représente la surface apparente (efficace) de l'atome vis-à-vis du rayonnement. L'unité utilisée pour décrire la section de capture est le *barn*. Il s'exprime bien sûr en unité de surface classique et 1 barn =  $10^{-24}$  cm<sup>2</sup>. L'atténuation d'un rayon gamma due à l'effet Compton sera une fonction de la masse volumique  $\rho_b$  et du rapport  $Z/A$  (Equation 1.2). Ce rapport est considéré comme constant pour les éléments rencontrés dans les formations géologiques et vaut approximativement  $\frac{1}{2}$ . La formule générale d'atténuation des radiations gamma est donnée par

$$\Phi = \Phi_i e^{-\rho_b \frac{Z}{A} N_{Av} \sigma_{Co} h} = \Phi_i e^{-\mu \rho_b h} \quad (1.3)$$

$\Phi_i$  est le flux de radiation gamma entrant, et  $h$  est l'épaisseur de matériau provoquant l'atténuation. Pour l'outil HLDT par exemple,  $h$  sera la distance source/détecteur.  $\mu$  est le *coefficient d'absorption massique* exprimé en cm<sup>2</sup>g<sup>-1</sup> qui est égal à

$$\mu = \frac{Z}{A} N_{Av} \sigma$$

et est dépendant de l'énergie du rayonnement sortant de la source. La relation 1.3 sera importante quand on voudra calculer la masse volumique de la formation RHOB (voir Section 1.4.2).

**L'effet photoélectrique** Lorsque l'énergie du rayon gamma tombe en dessous de 100 keV, par suite des multiples interactions Compton, l'effet dominant va être l'effet photoélectrique (voir Figure 1.4). Dans ce cas, le rayon gamma va percuter un atome de la formation et lui transférer *totalemment* son énergie. Si l'énergie du rayon gamma est suffisante, un électron est arraché à son orbite et interagit à son tour avec les atomes de la formation. L'électron manquant dans le réseau électronique de l'atome perturbé va être remplacé par un électron d'une couche électronique, cette transition s'accompagnant de l'émission d'un rayon X basse énergie (moins de 100 keV). De même que pour l'effet Compton, l'effet photoélectrique est caractérisé par une *section efficace de capture photoélectrique* notée  $\sigma_{Pe}$ . Contrairement à  $\sigma_{Co}$ ,  $\sigma_{Pe}$  varie beaucoup avec l'énergie, et est aussi très dépendant du numéro atomique  $Z$  du matériau absorbant selon la relation

$$\sigma_{Pe} = 12.1 \frac{Z^{4.6}}{E^{3.15}}$$

La relation générale d'atténuation 1.3 peut aussi être exprimée pour  $\sigma_{Pe}$  :

$$\Phi = \Phi_i e^{-\rho_b \frac{Z}{A} N_{Av} \sigma_{Pe} h} \quad (1.4)$$

**Le détecteur** L'outil HLDT est schématisé Figure 1.5. La source radioactive naturelle <sup>137</sup>Cs envoie un flux constant de rayonnement gamma de 662 keV dans la formation. Le matériau absorbant protège les détecteurs des radiations directes venues de la source. Deux détecteurs à scintillation classiques (voir Section 1.3.2) sont placés à des distances différentes de la source. Le premier, (détecteur "F" pour *Far*) placé loin de la source, est principalement influencé par la réponse de la formation. Il donne une mesure non corrigée de la densité de la formation grâce à deux fenêtres de mesure (LS) situées dans la zone du spectre résultant de l'effet Compton, ainsi

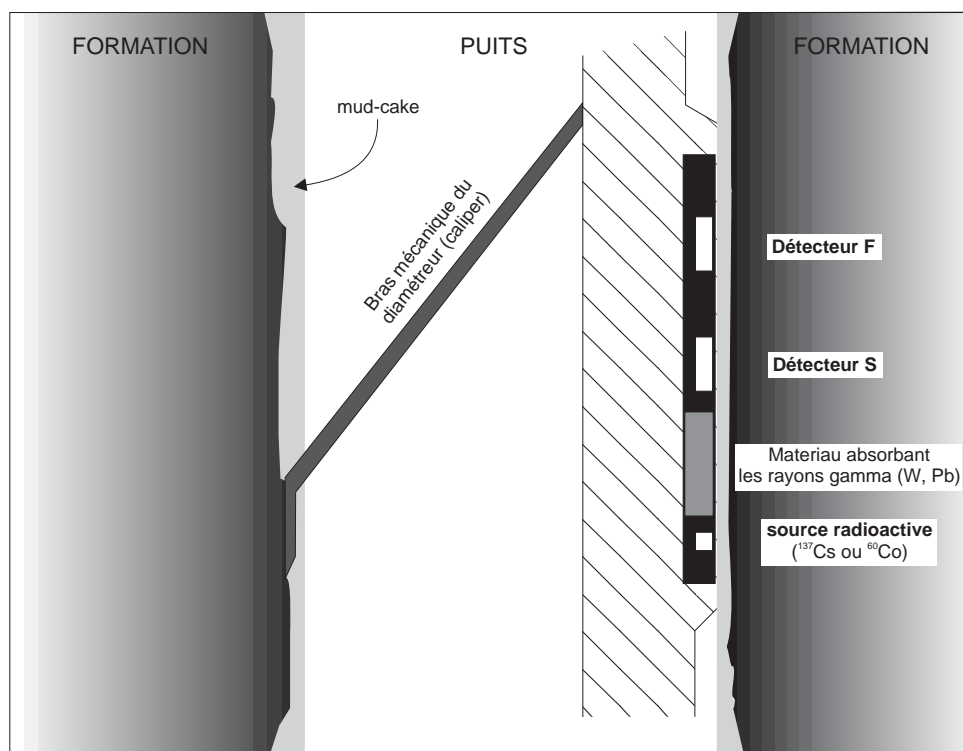


FIG. 1.5 – Principe de l'outil HLDT utilisant les interactions entre rayons gamma émis et atomes de la formation pour mesurer densité et facteur photoélectrique (d'après Ellis et al. (1983)). 'S' est pour "Short-spacing" (détecteur près de la source), 'F' est pour "Far" (détecteur loin de la source). Le "mud-cake" est une mince pellicule se formant sur les parois du puits due au filtrat laissé par la boue de forage dans les formations poreuses des parois du puits; généralement le mud-cake a une perméabilité très faible. A noter que le mud-cake est le plus souvent absent dans les puits Ocean Drilling Program dans lesquels on utilise une boue à base d'eau de mer.

qu'une mesure du facteur d'absorption photoélectrique par une troisième fenêtre appelée LITH et située dans la partie basse énergie du spectre. Il est équipé d'une fenêtre transparente au beryllium pour laisser passer en quantité les gamma basse énergie (résultant des multiples interactions du rayonnement source avec les électrons de la formation). La profondeur d'investigation du second détecteur, placé plus près de la source (détecteur "S" pour *short-spacing*), est faible et sa mesure principalement influencée par la présence du mud-cake (voir Figure 1.5) ou d'une zone envahie par le fluide de forage. Les détecteurs fonctionnent en mode proportionnel, ce qui permet de limiter la détection à des fenêtres d'énergie bien précises pour chaque détecteur.

**Corrections** La différence  $\Delta\rho$  des mesures données par les deux détecteurs est appelée la *correction de densité*, et est utilisée pour corriger la valeur obtenue par le détecteur F. La correction  $\Delta\rho$  est proportionnelle à l'épaisseur du mud-cake et au contraste de densité entre celui-ci et la formation.  $\Delta\rho$  ne donne donc pas simplement une mesure de l'épaisseur du mud-cake, et la pertinence de la correction appliquée devra être vérifiée par le biais d'autres paramètres (porosité neutron, résistivité à faible profondeur d'investigation). La mesure dans la fenêtre LITH permet de corriger des effets photoélectriques résiduels la mesure Compton (Ellis et al., 1983).

Le diamètre du puits est aussi un paramètre fondamental de la mesure. L'outil HLDT est pourvu d'un bras extensible qui plaque le bloc source-détecteur contre la paroi du puits (Fi-

gure 1.5). Dans le cas où le diamètre du puits dépasse 8 pouces (taille standard utilisée lors de la calibration de l'outil), et tant que le bloc de mesure reste en contact avec la formation, Ellis et al. (1983) démontrent que la correction de densité à appliquer à la mesure reste inférieure à  $0,02 \text{ g/cm}^3$ . Dans le cas où le contact est perdu (diamètre trop important), la mesure devient très vite inutilisable.

**Calibration** Plusieurs outils HLDT ont été calibrés à Paris et Houston grâce à des formations de densité connue. Ces formations sont soit naturelles et homogènes (calcaire, dolomie, dolérite), soit artificielles utilisant diverses combinaisons de silice pure, d'aluminium, de magnésium, et de carbure de silicium. La précision obtenue est de  $0,01 \text{ g/cm}^3$ . On procède de même pour la calibration de la mesure du facteur d'absorption photoélectrique. Le  $P_e$  des formations standard est calculé d'après le rapport  $Z/A$  de leurs constituants. La partie basse-énergie du spectre mesuré est le rapport entre la fenêtre LITH et les fenêtres LS. Bertozzi et al. (1981) montre que ce rapport est une fonction simple de  $P_e$  et d'autres paramètres constants ou indépendants de l'outil, déterminés durant la calibration.

### 1.4.2 Mesure de la densité totale ou masse volumique

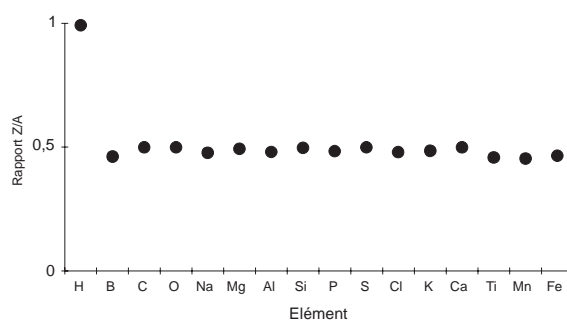


FIG. 1.6 – Evolution du rapport  $Z/A$  pour les éléments chimiques communément rencontrés dans les formations géologiques.

On a vu (Section 1.4.1) que la masse volumique de la formation, appelée abusivement densité, est étroitement liée à la densité électronique et au rapport  $Z/A$ . Ce rapport vaut  $1/2$  pour la plupart des éléments constituant des roches, sauf pour l'hydrogène, pour lequel il vaut 1 (Figure 1.6). La relation d'atténuation (Equation 1.3) peut donc être utilisée pour calculer la masse volumique  $\rho_b$  en assimilant celle-ci à la densité électronique. On obtient une relation linéaire de la forme

$$\rho_e = A - B \ln N$$

où  $N$  correspond aux coups reçus par le détecteur F. Pour tenir compte de l'effet perturbateur en présence d'eau, on a calibré la mesure pour des roches sédimentaires (calcaire, dolomie) saturées en eau. Dans ce cas, la *densité apparente mesurée* notée  $\rho_{log}$ , est directement assimilable à la masse volumique ou densité totale de la formation  $\rho_b$  suivant la relation

$$\rho_{log} = 1,07\rho_e - 0,188 = \rho_b$$

Par contre, certains minéraux (oxydes de fer, évaporites) peuvent perturber la mesure s'ils sont présents dans la formation en quantité suffisante.

### 1.4.3 Mesure du facteur d'absorption photoélectrique

Le facteur ou index d'absorption photoélectrique  $P_e$  donné par l'outil est défini empiriquement par

$$P_e = \left[ \frac{Z}{10} \right]^{3.6} \quad (1.5)$$

Il est fortement sensible à la présence d'élément à fort numéro atomique dans la formation. Il est classiquement utilisé pour tracer la présence de fer ou de baryum (quoique le baryum provienne le plus souvent de la boue de forage et soit donc corrigé ; voir Section 1.4.1).

Si l'index photoélectrique est un paramètre linéaire lorsque l'on considère des fractions *élémentaires*, il ne l'est plus lorsque l'on considère des *combinaisons d'éléments*. Il est nécessaire de définir un nouveau paramètre  $U$ , exprimé en section de capture/cm<sup>3</sup> (barnes/cm<sup>3</sup>), qui indique l'absorption photoélectrique d'un volume de matériau composé de combinaisons d'éléments. On le définit comme

$$U = P_e \rho_e = \sum_i V_i i P_{e,i} \rho_{e,i}$$

$$U = \sum_i V_i U_i$$

où  $\rho_{e,i}$  est la densité électronique du composé  $i$ ,  $P_{e,i}$  son index photoélectrique, et  $V_i$  sa fraction volumique (Ellis, 1987). Dans la plupart des cas, une bonne approximation de  $P_e \rho_e$  est  $P_e \rho_b$ .

## 1.5 Interactions entre neutrons et formation

### 1.5.1 Principe

De même que pour les rayons gamma, les interactions entre des neutrons artificiellement émis et la matière sont utilisées pour déterminer certains paramètres physiques et chimiques importants de la formation. Ces réactions dépendent toutes de l'énergie des neutrons considérés. La Figure 1.7 indique les différents intervalles d'énergie et les réactions leur correspondant. On distingue

- les réactions dites modératrices, c'est à dire celles dont l'énergie du neutron incident est réduite après la collision, et qui comprennent les réactions élastiques et inélastiques,
- les réactions d'absorption, où le neutron est absorbé par le noyau de l'élément atteint, donnant lieu soit à une absorption simple soit à une activation, avec émission d'un rayonnement gamma dit respectivement instantané ou retardé.

Le principe physique de ces multiples réactions est détaillé dans les sections qui suivent.

Il est utile ici de définir quelques paramètres qui vont intervenir dans les interactions entre les neutrons et la formation. La *section macroscopique de capture des neutrons thermiques* qui est le produit de la section microscopique de capture  $\sigma_i$  avec le nombre d'atomes par cm<sup>3</sup>, soit

$$\Sigma_i = N \sigma_i = \frac{N_{Av} \rho_b}{A} \sigma_i$$

où  $N_{Av}$  est le nombre d'Avogadro,  $\rho_b$  est la densité totale,  $A$  est le poids atomique.  $\Sigma_i$  est exprimé en cm<sup>-1</sup>, mais une unité spéciale a été définie, l'unité de capture  $cu$ , qui vaut 1000 fois  $\Sigma_i$  (Ellis, 1987).

On définit aussi (Ellis, 1987)

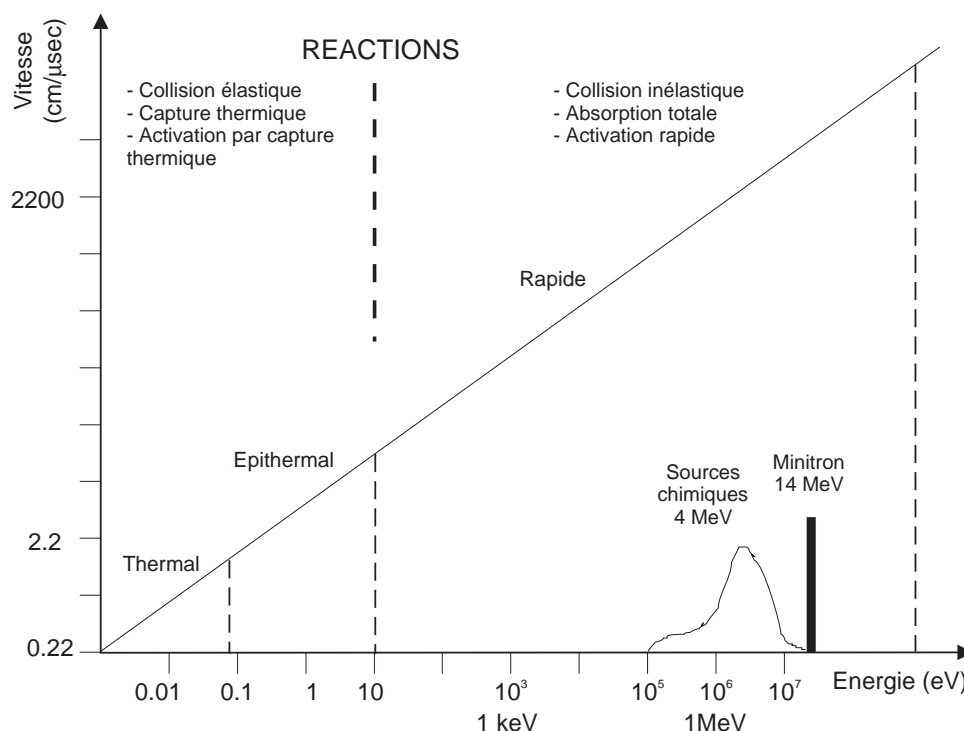


FIG. 1.7 – Les différents domaines d'énergie des neutrons utilisés en diagraphie, avec les réactions qui leur sont associées (d'après Ellis (1988))

- la *longueur de ralentissement*  $L_s$ , qui traduit la distance à parcourir par le neutron pour passer au stade thermique et liée en théorie à la section macroscopique de capture des neutrons thermiques de la formation  $\Sigma_{form}$ , mais qui sera aussi fonction de la section macroscopique de capture du puits de forage  $\Sigma_{BH}$  dans le cas où l'outil n'est pas en contact avec la formation (le bras-ressort (voir Figure 1.8) permet de plaquer l'outil contre la formation).
- la *longueur de diffusion thermique*  $L_d$ , qui représente la distance qu'un neutron doit parcourir depuis le moment où il passe à l'état thermique jusqu'à sa capture finale.  $L_d$  est lié entre autres à  $\Sigma_{form}$ .
- la *longueur de migration*  $L_m$ , qui s'exprime en  $L_m^2 = L_s^2 + L_d^2$ , et qui est donc la combinaison du chemin parcouru durant la phase de ralentissement et de la distance jusqu'à la capture du neutron. Ce paramètre permet de prédire la réponse de l'outil de mesure de la porosité neutron, discutée en Section 1.5.3.

### 1.5.1.1 Réactions modératrices

**La collision inélastique** Cette réaction implique des neutrons de très haute énergie (dizaine de MeV), produits par minitron (voir Figure 1.7). Seule une partie de l'énergie du neutron incident sera absorbée par le noyau atteint, provoquant l'excitation de ce dernier, puis son retour instantané à l'état stable avec l'émission d'un rayonnement gamma caractéristique de l'élément rencontré. La collision inélastique est la seule capable de produire un rayonnement gamma à partir du *carbone* et de l'*oxygène*. Le rapport entre ces deux éléments est en théorie très différent dans la formation et dans les fluides et est utilisé par les pétroliers pour tracer la présence

d'hydrocarbures dans la formation (Ellis, 1987; Ellis et al., 1987; Westaway et al., 1980) et pour évaluer la saturation en eau. Si, Ca, Fe et S produisent aussi un rayonnement gamma sous l'effet d'une collision inélastique, mais dans une moindre mesure, et leur abondance est plutôt mesurée grâce à un autre mécanisme, la *capture thermique* (voir plus loin).

**La collision élastique** Contrairement à la précédente, cette réaction implique des neutrons d'énergie plus faible (plusieurs MeV), généralement produit par une source chimique. Le système neutron-élément voit son énergie cinétique rester égale au cours de la collision, le neutron perd seulement une partie de son énergie, et passe progressivement au stade épithermique, puis thermique (voir Figure 1.7). On montre (Ellis, 1986; Ellis et al., 1987) que l'efficacité d'un atome à ralentir le neutron est d'autant plus grande que sa masse se rapproche de celle du neutron. Pour cette raison, l'hydrogène est l'élément qui a une section efficace de capture neutronique—donc une efficacité à ralentir les neutrons—maximale (100%—le neutron peut perdre la totalité de son énergie lors de la collision avec l'atome d'hydrogène). Suivent d'autres éléments légers comme le carbone (15,8%), l'oxygène (12%), et plus loin le silicium (7%), le chlore ou le sodium. Ce type de réaction est utilisé pour évaluer la porosité directement relative à la présence d'atomes d'hydrogène. C'est la *porosité neutron* NPHI (voir Section 1.5.3). L'outil détecte les neutrons dans les domaines épithermal et thermal résultant de leurs multiples interactions avec les atomes de la formation.

### 1.5.1.2 Les réactions d'absorption

**avec des neutrons rapides** Le noyau atteint absorbe le neutron, devenant ainsi un isotope instable et excité. C'est l'*activation*, qui sera dite *rapide* car produit d'un neutron de haute énergie. Le noyau repasse à un état non excité par émission d'un rayonnement gamma dit *instantané*. Pour revenir à un état stable non excité, le noyau doit encore passer par plusieurs étapes caractérisées par l'émission d'un autre rayon gamma, dit *retardé*. Un exemple est l'excitation de l'atome de  $^{28}\text{Si}$  (Ellis et al., 1987), qui en absorbant un neutron rapide, va devenir un atome de  $^{28}\text{Al}$ , lequel excité et instable, va repasser sous la forme  $^{28}\text{Si}$  en plusieurs étapes. L'activation de Si par des neutrons rapides (d'énergie supérieure à 4,5 MeV) va créer une interférence dans la mesure de l'aluminium. Cette interférence sera évitée par l'utilisation d'une source délivrant des neutrons de plus faible énergie (voir Section 1.5.2.2).

**avec des neutrons ayant atteint le stade thermique : la capture thermique** Lorsque le neutron atteint le stade thermique, il diffuse dans la formation et peut être capturé par les atomes ayant une section efficace de capture des neutrons thermiques importante. Le Tableau 1.3 donne la section efficace de capture des principaux éléments rencontrés dans les formations géologique (Ellis, 1987; Hertzog et al., 1989). Une exception est le Gd, terre rare intermédiaire, qui bien que très peu abondant relativement aux autres éléments cités, contribue au spectre mesuré du fait de sa très grande section efficace de capture des neutrons thermiques. On mesure aussi le taux d'absorption des neutrons thermiques, qui est le reflet de l'abondance en Cl, élément le plus couramment rencontrés (présent dans les eaux de formation salines) et ayant une forte section de capture des neutrons thermiques ( $\sigma = 33,2\text{ barns}$ ). Un taux d'absorption rapide indique la présence d'eau salée, alors que faible, il indique une eau douce ou la présence d'hydrocarbures (Ellis et al., 1987; Ellis, 1987).

**avec des neutrons ayant atteint le stade thermique : l'activation neutronique** Certains éléments présents dans la formation, et indispensables pour donner un bonne évaluation

Elément	Section efficace de capture des neutrons thermiques (barns)	Efficacité par rapport à l'abondance moyenne	Principales raies gamma pour les éléments activés par les neutrons thermiques (keV)
Si	0,16	0,23	
Ca	0,43	0,21	3084, 4071
Fe	2,55	0,23	
S	0,52	0,03	
Ti	6,10	0,025	320, 929
K	2,10	0,054	
Al	0,23	0,026	1779
Gd	49000	0,16	
Na	0,53	0,020	1369, 2754
Mg	0,06	0,020	844, 1014
H	0,33		
Cl	33,20		

TAB. 1.3 – Section efficace de capture des neutrons thermiques, et efficacité normalisée en fonction de l'abondance moyenne pour chacun des principaux éléments rencontrés dans les formations géologiques. L'efficacité normalisée est le produit d'une abondance moyenne de l'élément par sa section de capture neutronique, divisé par le poids atomique de l'élément (d'après Ellis (1987) et Hertzog et al. (1989)).

de la lithologie, ne produisent pas assez de rayons gamma instantanés de capture pour pouvoir être détectés. C'est le cas de Al, qui est crucial pour la détermination des argiles (Herron and Herron, 1996). Par contre, lorsque le noyau stable de  $^{27}\text{Al}$  absorbe un neutron thermique, il se transforme en  $^{28}\text{Al}$  instable. De la même façon que pour l'activation rapide (voir plus haut),  $^{28}\text{Al}$  se désintègre en  $^{28}\text{Si}$ , en émettant un rayonnement gamma *retardé*, qui peut être détecté et permettra de mesurer l'abondance en Al.

## 1.5.2 La spectrométrie de rayonnement gamma

### 1.5.2.1 Introduction

Un schéma de la chaîne d'outils *Geochemical Logging Tool* ou GLT est donné Figure 1.8. Elle combine un NGT classique (voir Section 1.3.2), servant à mesurer l'abondance en K, U et Th par analyse du spectre gamma naturel (voir Section 1.3.4), un *Activation Aluminum Clay Tool* ou AACT qui utilise les neutrons produits par la source radioactive  $^{252}\text{Cf}$ , d'un *Compensated Neutron Tool* ou CNT pour mesurer l'abondance en Al, et un *Gamma-ray Spectrometry Tool* ou GST, qui mesure l'abondance en Si, Ca, Fe, Ti, S et Gd. On utilise le spectre gamma naturel mesuré par le NGT dans la mesure de Al (voir section suivante).

### 1.5.2.2 La mesure de l'abondance en Al par activation neutronique

Les neutrons rapides provenant de la source au  $^{252}\text{Cf}$  du CNT sont ralentis par collision élastique dans la formation, arrivent au stade thermique, et sont absorbés par le seul isotope naturel de l'aluminium,  $^{27}\text{Al}$ . Cette réaction provoque l'*activation* de l'aluminium en  $^{28}\text{Al}$ , qui se désintègre d'abord en  $^{28}\text{Si}$  avec une demi-vie de 2,3 minutes et émission d'une particule  $\beta$ , puis le silicium revient à un état stable par émission d'un rayonnement gamma à 1779 keV.



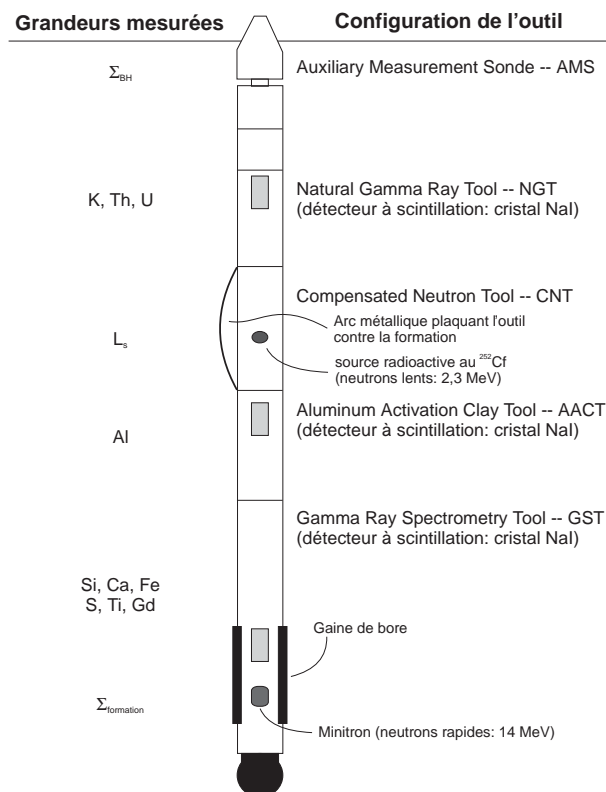


FIG. 1.8 – Schéma de la chaîne d'outils composant le *Geochemical Logging Tool* ou GLT, montrant les différents outils combinés ainsi que les différentes grandeurs mesurées (adapté de Hertzog et al. (1989)). Le CNT servant de source à neutrons pour l'AACT est exactement le même que celui utilisé pour mesurer la porosité neutron (voir Section 1.5.3).

La source  $^{252}\text{Cf}$  est utilisée ici préférentiellement à la source Am-Be, classiquement utilisée en diagraphie, car elle permet d'atténuer la création de rayons gamma résultant de l'absorption des neutrons rapides par les atomes de silicium présents dans la formation (voir Section 1.5.1.2). L'énergie moyenne des neutrons produit par le  $^{252}\text{Cf}$  est 2,3 MeV (à comparer avec les 4,5 MeV pour l'Am-Be). Ainsi, seuls 12% des neutrons vont avoir une énergie suffisante pour déclencher l'absorption rapide (Scott and Smith, 1973). Le degré d'interférence provenant de cette réaction est aussi dépendant de la concentration en silicium de la formation. La contribution du Si au spectre mesuré par le GST (voir Section 1.5.2.3) sert à corriger la contribution du  $^{28}\text{Si}$  dans la mesure de Al. Une autre interférence concerne le manganèse dont l'activation produit deux raies caractéristiques à 1811 et 847 keV, dont le rapport des amplitudes est connu et constant. La capacité de résolution du cristal NaI(Tl) étant trop faible pour séparer la raie à 1779 keV de Al et celle à 1811 de Mn, on utilise la contribution à 847 keV pour évaluer l'erreur sur la mesure de Al à 1779 keV. Pour cela, Schlumberger a modifié quelque peu la méthode d'analyse spectrale en ajoutant 3 fenêtres de mesure, l'une positionnée sur le pic à 847 keV (W6), donnant la contribution de Mn, et les deux autres de part et d'autre de façon à évaluer le bruit de fond.

**Le détecteur** L'outil AACT est un NGT classique (voir Section 1.3.2) composé d'un cristal NaI(Tl) couplé à un photomultiplicateur, et d'un collecteur multi-canaux. Le détecteur enregistre le rayonnement gamma naturel issu de la formation, plus le rayonnement gamma induit

à 1779 keV par l'activation de l'aluminium. On remarque que le pic tombe dans la fenêtre W4 (SCHLUM 4 sur la Figure 1.1). Lorsque les possibles interférences avec Si et Mn ont été corrigées, la contribution de Al est mesurée en soustrayant le spectre naturel mesuré par le NGT de celui mesuré par le AACT (Scott and Smith, 1973; Hertzog et al., 1989).

**Corrections environnementales** Les conditions de forage ainsi que la nature de la formation influencent le taux de comptage de l'outil AACT. L'absorption des neutrons dans le puits affecte le flux total de neutrons dans la formation (Scott and Smith, 1973; Hertzog et al., 1989). Un autre facteur perturbateur est la présence dans la formation d'éléments absorbeurs de neutrons thermiques, comme le bore par exemple, modifiant la section macroscopique de capture neutronique de la formation  $\Sigma_{form}$ . Le diamètre du puits devra aussi être pris en compte, puisque plus il augmente, plus le volume de formation effectif concerné autour de l'outil AACT est faible. Hertzog et al. (1989) ont réalisé une simulation Monte Carlo pour évaluer l'influence de paramètres tels que  $\Sigma_{BH}$ ,  $\Sigma_{form}$ , le diamètre du puits et la longueur de ralentissement  $L_s$  sur le taux de comptage pour l'aluminium. En faisant varier ces paramètres dans un intervalle de valeurs auxquelles on peut s'attendre dans les conditions de mesure *in situ*, Hertzog et al. (1989) montrent que le taux de comptage, à concentration en Al constante, peut varier d'un facteur 3. Un algorithme de correction a été développé qui utilise la mesure de  $L_s$  extrait du rapport entre les comptages des deux détecteurs N et F de mesure des neutrons épithermiques (voir Figure 1.11) de l'outil CNT, ainsi que  $\Sigma_{form}$  mesuré par le GST et  $\Sigma_{BH}$  qui est estimé à partir de la salinité de la boue de forage mesuré par l'*Auxiliary Measurement Sonde* de la chaîne GLT (voir Figure 1.8).

**Répétabilité de la mesure et calibration de l'outil** Hertzog et al. (1989) montrent que la répétabilité déduite de multiples passages de l'outil dans un puits est en accord avec l'erreur calculée à partir des erreurs statistiques sur des taux de comptage et provenant des outils, après propagation à travers les algorithmes de correction des abondances et des effets environnementaux. Pour l'aluminium, on obtient une répétabilité de 0,1 à 0,8% (erreur à un sigma sur la valeur mesurée), dans un intervalle de concentration de 0 à 15%. La variation dépend presque exclusivement de la statistique des taux de comptage, qui dépend beaucoup des conditions de mesure.

La calibration de l'outil est faite régulièrement en laboratoire, en utilisant une formation artificielle de concentration en Al connue, pour un flux de neutrons correspondant à une source précise de  $^{252}\text{Cf}$ , en utilisant un outil de référence. On active la formation pendant une période de temps suffisamment longue pour qu'elle arrive à saturation. Le comptage est effectué pendant un intervalle connu suffisant pour arriver à une bonne statistique. Le taux de comptage mesuré peut alors être corrigé par rapport à un maximum théorique atteint par l'outil de référence. La précision obtenue pour l'AACT en laboratoire est d'environ 3,5%. Dans les conditions de mesure *in situ* (en tenant compte des corrections environnementales et des statistique des taux de comptage), une précision de 12% sur la mesure peut être raisonnablement attendue (0,1 à 1,8% pour une concentration en Al dans la formation variant de entre 0 et 15%) (Schweitzer et al., 1987, 1988).

### 1.5.2.3 La mesure de l'abondance en Si, Ca, Fe, S, Ti, Gd

Cette mesure est effectuée par l'outil GST, dont la Figure 1.8 donne un schéma simplifié. Il est composé d'un minitron et d'un détecteur classique à scintillation (cristal NaI(Tl) couplé à un photomultiplicateur + analyseur multicanaux). Le minitron est un appareil électromécanique opérant à haute tension qui, par le biais d'un bombardement d'une cible de tritium par des atomes

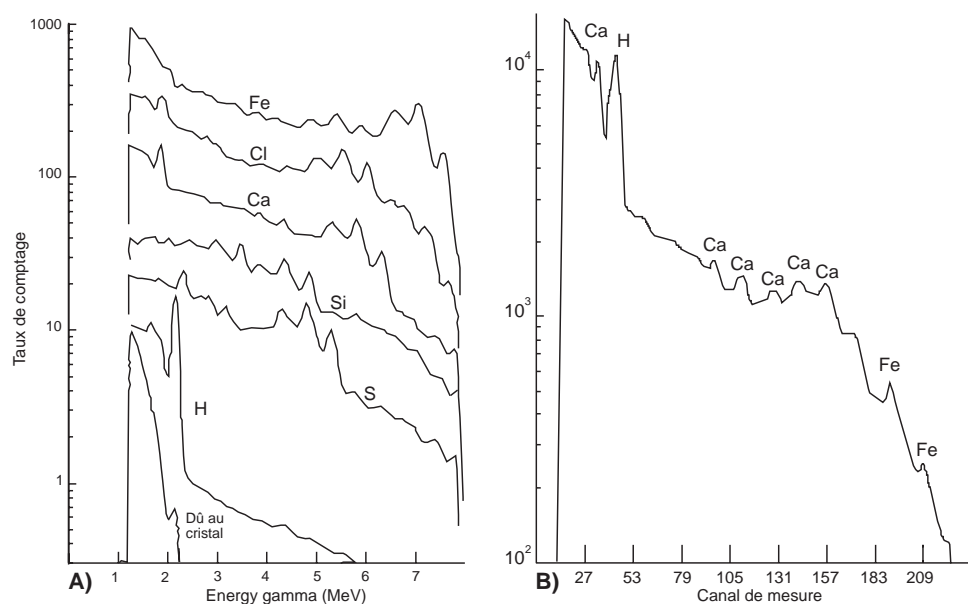


FIG. 1.9 – **A.** Spectres standards isolés pour chaque élément. Ces spectres sont obtenus en laboratoire. La contribution due au cristal est non négligeable et provient des réactions d'activation causées directement par les neutrons de haute énergie dans le cristal du détecteur. **B.** Exemple de spectre gamma de capture accumulé par le GST. Le spectre est mesuré sur 256 canaux, allant de 1 à 8 MeV. De nombreux effets perturbateurs (vibrations et interactions dans le cristal, effet Compton, faible résolution en énergie du détecteur NaI(Tl)) dégrade ce spectre et explique la largeur des pics (d'après (Grau et al., 1980)).

de deutérium, produit des neutrons à très haute énergie (14 MeV). Ce dispositif, contrairement à une source radioactive, permet un contrôle précis du flux de neutrons.

Le GST identifie les rayons gamma résultant des interactions entre les neutrons et les noyaux des atomes de la formation. L'outil peut être mis en œuvre suivant trois modes distincts : le mode *inélastique*, le mode *capture-tau* ou capture thermique, et le mode *activation* (Ellis, 1987). Dans les trois modes, les neutrons sont émis cycliquement dans la formation sous forme pulsée. La mesure se fait pendant le laps de temps entre les impulsions.

Le mode inélastique utilise les réactions de collision inélastique (voir Section 1.5.1.1). Il est le seul apte à produire un rayonnement gamma à partir du carbone et de l'oxygène. Le spectre obtenu est comparé avec des spectres étalons établis en laboratoire pour chaque élément capable de donner un rayonnement gamma instantané détectable (C et O principalement, mais aussi Ca, Fe, Si et S). L'acquisition du spectre intervient immédiatement après l'impulsion de neutrons, sans attendre que les neutrons aient perdu trop d'énergie, de façon que la majorité du rayonnement gamma mesuré provienne des réactions inélastiques. L'intérêt de ce type de mesure provient du contraste important existant entre le rapport C/O de l'eau et celui du pétrole (Westaway et al., 1980; Ellis, 1987). Les pétroliers utilisent beaucoup le diagramme C/O en fonction de la porosité dans les formations relativement propres (exemptes de recristallisation ou d'argiles) pour déterminer la saturation en hydrocarbures (Westaway et al., 1980). Le mode activation sert à mesurer l'oxygène et fait aussi intervenir des neutrons de très haute énergie. L'oxygène est activé en <sup>16</sup>N, qui se désintègre en oxygène en émettant un rayonnement gamma d'énergie 6,13 MeV (Ellis, 1987). Étant donné que l'oxygène nécessite des neutrons de 11 MeV pour être activé, la profondeur d'investigation de ce type de mesure sera faible, et cet "oxygen log" sera surtout utilisée comme détecteur de la limite eau-hydrocarbure dans le puits de forage. Le programme

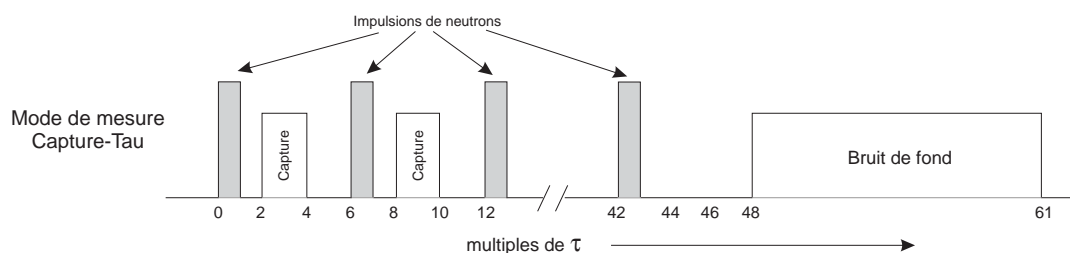


FIG. 1.10 – Détail du cycle de mesure du GST en mode capture-tau. Ce mode utilise  $\tau$  comme unité de temps pour définir l'intervalle entre impulsion et mesure. Cet intervalle change à chaque cycle à cause des variations de  $\tau$ . Un cycle complet équivaut à  $62 * \tau \mu\text{sec}$  (d'après Westaway et al. (1980)).

de forage Ocean Drilling Program n'utilise pas ces deux modes d'opération dans ses puits, aussi pour plus d'information, on peut de référer à Westaway et al. (1980) ou à Hertzog and Plasek (1979).

Le mode *capture-tau* ou capture thermique est beaucoup plus intéressant du point de vue géologique (voir Figure 1.10). Un spectre gamma de capture est accumulé dans l'intervalle entre deux impulsions de neutrons, mais cette fois en prenant soin de laisser un moment entre la fin de l'impulsion et l'acquisition proprement dite. Ce délai permet aux neutrons de perdre leur énergie, d'arriver au stade thermique, et de pouvoir être capturés par les noyaux des atomes de la formation. L'intervalle de temps entre deux impulsions de neutrons est constamment modifié au cours de la diagraphie selon le paramètre  $\tau$ , qui décrit la décroissance de population des neutrons thermiques dans le temps (Westaway et al., 1980). Ce déclin est lié au pouvoir absorbant de la formation et s'exprime (Ellis, 1987) par

$$\tau = \frac{1}{v\Sigma_{form}}$$

soit pour des neutrons thermiques ( $v = 0,22 \text{ cm}/\mu\text{sec}$ )

$$\tau = \frac{4550}{\Sigma_{form}}$$

$\tau$  est mesuré par le GST pendant l'acquisition du spectre gamma de capture. Les intervalles de temps entre impulsion et acquisition sont des multiples de  $\tau$ . De cette façon, on minimise le signal provenant du puits de forage (surtout dû aux réactions inélastiques), ce qui permet d'obtenir un bon rapport signal-bruit (Westaway et al., 1980; Hertzog and Plasek, 1979). H, Cl et Fe sont trois éléments à forte section de capture des neutrons. Les deux premiers sont présents surtout dans le fluide de forage, tandis que Fe provient du corps de l'outil. Pour minimiser leurs effets sur le signal, une gaine de bore (voir Figure 1.8) a été rajoutée autour du GST. Le bore a aussi une forte section de capture des neutrons et réduit le nombre des réactions de capture induites par Cl ou Fe, augmentant la contribution de la formation dans le spectre mesuré.

**L'analyse spectrale** Figure 1.9.B montre un spectre de capture typique accumulé par le GST. Celui-ci a été mesuré dans une formation carbonatée saturée en eau douce (Grau et al., 1980). Les pics correspondant aux différents éléments contribuant au spectre sont visibles sur le bruit de fond. La largeur des pics est due aux interférences qui perturbent l'enregistrement du rayonnement gamma (vibrations et interactions dans le cristal, effet Compton, faible résolution en

énergie du détecteur). La mesure de la radiocativité gamma naturelle, où la proportionnalité entre désintégration et comptage est vérifiée (voir Section 1.2), est simple dans son principe. De même, le rayonnement gamma retardé, servant à la détermination de l'abondance en Al est le produit d'une source  $^{252}\text{Cf}$  d'intensité connue, et l'on pourra calculer les paramètres nécessaires pour transformer les intensités détectées en concentrations. Par contre, on ne peut pas directement calibrer le rayonnement gamma de capture dans le cas où un minitron est utilisé en raison de l'instabilité du flux de neutrons.

Durant la mesure, les spectres sont accumulés dans le GST et à intervalles réguliers, un ajustement par la méthode des moindres carrés pondérés entre le spectre mesuré et une bibliothèque de spectres standards (Grau and Schweitzer, 1987) est effectué. Le principe en est simple mais ingénieux : on suppose d'abord que le spectre mesuré est une combinaison linéaire de tous les spectres correspondant à chaque élément. Cela donne

$$p = s^{\text{Ca}} x_{\text{Ca}} + s^{\text{Si}} x_{\text{Si}} + \dots$$

où  $p$  est un vecteur colonne représentant le nombre de canaux,  $s^{el}$  sont les vecteurs normalisés représentant les standards pour chaque élément, et  $x_{el}$  sont les inconnus (les "yields", qui peuvent être décrits comme la proportion du spectre total attribuée à chaque élément). Représenté sous forme matricielle, on a

$$\begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_{200} \end{bmatrix} = \begin{bmatrix} s_1^{\text{Ca}} & s_1^{\text{Si}} & \dots & s_1^{\text{S}} \\ s_2^{\text{Ca}} & s_2^{\text{Si}} & \dots & s_2^{\text{S}} \\ \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots \\ s_{200}^{\text{Ca}} & s_{200}^{\text{Si}} & \dots & s_{200}^{\text{S}} \end{bmatrix} \begin{bmatrix} x_{\text{Ca}} \\ x_{\text{Si}} \\ \vdots \\ x_{\text{S}} \end{bmatrix}$$

soit

$$p = S \cdot x \quad (1.6)$$

Comme il n'existe aucune solution directe de ce système, le principe consiste à introduire un vecteur d'erreur  $\epsilon$ . Il faut alors trouver une solution  $x$  pour que la somme pondérée du carré des erreurs, qui peut être écrit sous la forme  $\epsilon^T W \epsilon$ , soit minimale. Cela revient à minimiser  $(p - S \cdot x)^T W (p - S \cdot x)$  dont la solution est

$$x = (S^T W S)^{-1} S^T W \cdot p$$

$x$  est la matrice  $E$  des estimateurs, composée d'autant de lignes que d'éléments inclus dans l'estimation, chaque ligne contenant 200 termes représentant la contribution d'un élément à un canal de mesure. La solution déterminée est ensuite utilisée dans l'équation 1.6 pour recalculer un spectre estimé  $\hat{p}$ . Une description plus précise de ce calcul matriciel, avec une discussion sur les effets perturbateurs dus au cristal ou à l'électronique, peut être trouvée dans Hertzog (1978) et Grau et al. (1980). Grau et al. (1984) donne la technique utilisée pour séparer la contribution de la formation de celle du puits de forage dans les spectres mesurés.

La base de données de standards, dont un exemple est donné par la Figure 1.9.A, est obtenue en laboratoire pour chaque élément excité par le GST (Grau et al., 1980). Les rapports entre certains "yields" (voir Tableau 1.4) peuvent déjà être utilisés pour décrire qualitativement des propriétés macroscopiques de la formation (Westaway et al., 1980; Grau et al., 1980) avant que les "yields" ne soient utilisés pour calculer les concentrations..

Yield Ratio	Interaction	Nom	Label
C/O	Inélastique	Carbon-Oxygen ratio	COR
Cl/H	Capture	Salinity-indicator ratio	SIR
H/(Si + Ca)	Capture	Porosity-indicator ratio	PIR
Fe/(Si + Ca)	Capture	Iron-indicator ratio	IIR
Si/(Si + Ca)	Capture et inélastique	Lithology-indicator ratio	LIR

TAB. 1.4 – Rapports entre les contributions (yields) de différents éléments mesurés par le GST, et utilisés comme indicateurs pour différentes propriétés de la formation.

**Des yields aux pourcentages d'oxides** Calculer des concentrations relatives est aisé à partir des yields en divisant chaque yield  $Y_i$  par le facteur de sensibilité spectrale lui correspondant  $S_i$  (Grau et al., 1980; Hertzog et al., 1989). Ces facteurs sont des constantes propres aux outils déterminées en laboratoire. Ils rendent compte des différences entre les éléments concernant par exemple la section efficace de capture des neutrons thermiques, l'efficacité de transport du rayonnement gamma ou sa détection par l'outil. Ces concentrations relatives sont liées aux concentrations absolues  $W_i$  par un facteur de normalisation appelé  $F$ , tel que

$$W_i = F \cdot \frac{Y_i}{S_i} \quad (1.7)$$

$F$  étant une fonction complexe de pratiquement tous les paramètres physiques de l'environnement de mesure (formation et puits de forage), il est impossible de la calculer directement. Pour tourner la difficulté, un *modèle de fermeture* a été développé (Grau and Herron, 1987; Grau and Schweitzer, 1989; Grau et al., 1989; Hertzog et al., 1989; Schweitzer et al., 1988).  $F$  est calculé de façon que la somme de tous les oxides soit égale à l'unité. Le problème dans cette approche, évoqué par Hertzog et al. (1989), est que l'on ne prend pas en compte les éléments C, O, Na et Mg, qui peuvent contribuer significativement à la fraction totale mesurée, et qui ne sont pas mesurés directement par l'outil GST. Pour pallier à l'absence de la mesure de C et O, Hertzog et al. (1989) proposent que chacun des éléments mesurables soit associé à un oxide ou carbonate et un seul. De plus, Schweitzer et al. (1988) montrent qu'une façon d'améliorer encore le résultat est de ne considérer les éléments présent dans la matrice et pas ceux du fluide de forage ou contenu dans les pores. Ceci permet de ne pas avoir à se préoccuper de H et Cl, présent en quantité dans le puits de forage. Ce modèle compense automatiquement les variations de volume du puits en augmentant ou en diminuant  $F$ . On écrit l'équation de fermeture à résoudre à chaque niveau de mesure comme

$$F \left[ \sum_i X_i \frac{Y_i}{S_i} \right] + X_K W_K + X_{Al} W_{Al} = 1.0 \quad (1.8)$$

où  $X_i$  est le rapport entre le pourcentage de l'oxyde considéré et celui de l'élément  $i$ , et  $W_K$  et  $W_{Al}$  respectivement la fraction d'Al mesurée par activation et la fraction de K déterminée par analyse du spectre de radioactivité naturelle. Les  $X_i$  sont aisés à déterminer. Pour Si par exemple

$$X_i = \frac{W_{SiO_2}}{W_{Si}} = 2,139$$

Une fois  $F$  déterminé à partir du modèle de fermeture, l'équation 1.7 est utilisée pour calculer les fractions absolues pour chaque élément.

Elément (poids%)	Incertitude sur la concentration = $1\sigma$ (poids%)	Intervalle des concentrations (poids%)
Si	6,4 à 2,3	25 à 47
Ca	4,2 à 1,6	0 à 40
Fe	0,3 à 0,8	1 à 5
Ti	0,1 à 0,3	0,1 à 0,8
Gd	0,4 à 1 ppm	0 à 10 ppm
S	environ 1,3	

TAB. 1.5 – Répétabilité de l'outil GST, dans des conditions de mesure normales, donnée pour les différents éléments mesurés, sous forme d'une déviation standard autour de la moyenne calculée à partir des résultats de plusieurs passages de l'outil et de la mesure de l'incertitude théorique entachant une mesure par capture neutronique (d'après Hertzog et al. (1989), Grau et al. (1990) et van den Oord (1991)).

**Validité du modèle de fermeture** Pour vérifier la cohérence du modèle avec la minéralogie, Schweitzer et al. (1988) ont calculé les  $F \left[ \sum_i X_i \frac{Y_i}{S_i} \right]$  pour quelques minéraux communs (pyrite, minéraux argileux, quartz, calcite, ...) rencontrés dans les formations géologiques. Ils montrent que la relation de fermeture est satisfaite avec une erreur de  $\pm 5\%$  pour la plupart des minéraux. Par contre, le modèle ne donne pas de bons résultats en présence de pyrite, d'ankérite ( $\text{CaFe}[\text{CO}_3]_2$ ), de dolomite (Schweitzer et al., 1988; Hertzog et al., 1989). L'analyse du facteur photoélectrique PEF et la comparaison des contributions de Fe et S permet de différencier la pyrite de l'anhydrite. L'ankérite et la dolomite contiennent des concentrations en Mg non négligeables, élément qui n'est pas pris en compte par le modèle, ce qui entraîne une erreur pouvant atteindre 46% (Hertzog et al., 1989). Une manière de corriger pour la présence de Mg est de comparer la mesure du PEF donnée par l'outil HLDT avec une estimation du PEF calculée d'après les concentrations élémentaires données par le modèle de fermeture (Schweitzer et al., 1988; Hertzog et al., 1989). Si la fraction de l'élément  $i$  est  $Wt_i$ , on écrit d'après l'équation 1.5

$$P_e = \sum_i Wt_i \left( \frac{Z_i}{10} \right)^{3.6} \quad (1.9)$$

Le modèle de fermeture ne considérant que les éléments présents dans la matrice, il faut au préalable corriger la mesure du PEF de l'effet des fluides. Dans notre cas, le fluide est l'eau, et on peut écrire (Hertzog et al., 1989) :

$$P_{e,\log} = W_{\text{fluide}} P_{e,\text{fluide}} + (1 + W_{\text{fluide}}) P_{e,\text{matrice}} \quad (1.10)$$

où  $W_{\text{fluide}} = \phi/\rho_b$  et  $P_{e,\text{fluide}}$  est compris entre 0,36 (eau douce) et 0,80 (eau salée à 120 ‰) (Ellis, 1987). Quant à  $P_{e,\text{matrice}}$ , on le génère à partir des fractions données par le GST

$$P_{e,\text{matrice}}^{\text{estim}} = \sum_i P_{e,i} W_i + P_{e,K} W_K + P_{e,Al} W_{Al} + P_{e,O} W_O \quad (1.11)$$

On suppose que  $W_O$  est constant et égal à 0,52. Si la formation ne contient pas ou peu de Mg, on peut écrire  $P_{e,\text{matrice}}^{\log} = P_{e,\text{matrice}}^{\text{estim}}$ . Si Mg est présent dans la formation en quantité suffisante (plusieurs pourcents), la relation de fermeture va surestimer le facteur de normalisation  $F$ , et par voie de conséquence les fractions issues de la mesure du GST ainsi que  $P_{e,\text{matrice}}^{\text{estim}}$ . Il faut donc calculer un  $F'$  et introduire un terme tenant compte de Mg dans l'équation 1.8 :

$$F' \left[ \sum_i X_i \frac{Y_i}{S_i} \right] + X_{Mg} W_{Mg} + X_K W_K + X_{Al} W_{Al} = 1.0 \quad (1.12)$$

De l'égalité des  $P_e$ , on écrit

$$F' \left[ \sum_i P_{e,i} \frac{Y_i}{S_i} \right] + P_{e,Mg} W_{Mg} + P_{e,K} W_K + P_{e,Al} W_{Al} + P_{e,O} W_O = P_{e,matrice}^{log} \quad (1.13)$$

$W_{Mg}$  peut alors être obtenu (Hertzog et al., 1989) :

$$W_{Mg} = \left[ \frac{X_i}{P_{e,matrice}^{estim} - P_{e,MgCO_3}} \right] (P_{e,matrice}^{estim} - P_{e,matrice}^{log}) \quad (1.14)$$

Les concentrations corrigées sont obtenues en multipliant les concentrations initiales par  $F'/F$ .

On peut utiliser le même système pour corriger de la présence de Na, mais il sera impossible de distinguer entre Na et Mg. Dans le cas où Na est présent, mais non déterminé, Hertzog et al. (1989) estiment cette approximation comme induisant peu d'erreur sur le calcul des fractions pondérales des autres éléments par le modèle de fermeture étant donné que  $X_{Na}$  est faible (si l'on choisit  $Na_2O$  comme oxyde).

**Précision et répétabilité de la mesure** Les incertitudes sur la mesure du GST ont été établies par Grau et al. (1990). Ces auteurs ont d'abord montré que les erreurs sur les concentrations calculées ne dépendaient que faiblement des concentrations, mais plutôt d'autres paramètres tels que statistique de comptage, nature du cristal, ou sensibilité de l'outil. Schweitzer et al. (1982) ont montré que le mode de mesure capture-tau était assez peu sensible aux variations moyennes ( $\pm 50\%$ ) du diamètre de puits ou à une invasion équivalente de la formation par le fluide de forage. La profondeur d'investigation du GST est estimée à 30 cm. Pour évaluer la répétabilité de la mesure, ils ont appliqué un bruit statistique artificiel à des spectres simulés et ont calculé la déviation standard des concentrations obtenues. Ils ont simultanément calculé ces incertitudes à partir de relations théoriques relevant de la physique de la mesure par capture neutronique. Hertzog et al. (1989) ont enfin comparé les concentrations mesurées par plusieurs passages successifs de l'outil. Les trois méthodes d'évaluation se sont accordées à moins de 2% près et donnent les erreurs sur les concentrations présentées Tableau 1.5.

On évalue la précision de la mesure du GST en comparant les mesures effectuées sur échantillons prélevés sur carottes aux mesures de diagraphies. Cette méthode ne donne toutefois qu'une vérification qualitative car plusieurs incertitudes interviennent, parmi elles la difficulté de corréler en profondeur une mesure de diagraphie et un échantillon carotté, ou le fait que la mesure sur carotte peut ne pas être représentative, du fait d'hétérogénéités dans la formation, du volume infiniment plus grand mesuré par l'outil (la résolution verticale du GST est limitée à 1 mètre du fait d'un lissage des données par fenêtre glissante englobant 7 niveaux de mesure—1 mètre—effectué lors du traitement des données brutes pour améliorer le rapport signal-bruit des données brutes).

Grau et al. (1990) démontrent que dans des conditions de mesure normales, les erreurs absolues sur les concentrations sont de 2 à 3 % pour Si et Ca, 1 % pour S, et de l'ordre de quelques dixièmes de pourcent pour Fe et Ti. Pour le Gd, l'erreur est environ de 0,2 ppm. Grau et al. (1990) montrent que des conditions de mesure difficiles affectent surtout le taux de comptage, problème en partie résolu par de multiples passages de l'outil, ou par une vitesse de remontée de l'outil plus réduite.



### 1.5.3 Mesure de l'indice d'hydrogène apparent ou porosité neutron

#### 1.5.3.1 Principe

La mesure utilise les réactions de diffusion élastique (voir Section 1.5.1.1) qui ralentissent les neutrons jusqu'à un état épithermique (moins de 10 eV) ou thermique (environ 0,025 eV) où ils sont détectés. La physique du transport des neutrons est décrite par l'équation de Boltzmann (Ellis, 1987). Elle donne le taux de perte de neutrons par absorption ou collision dans un volume considéré par rapport aux neutrons provenant de la source ou issus de collisions de neutrons de plus hautes énergies. Il est possible de simplifier analytiquement cette équation de façon à faire apparaître un coefficient de diffusion  $D$

$$D = \frac{1}{3(\Sigma_t - \bar{\mu}\Sigma_s)}$$

dépendant de la différence entre la section de capture totale  $\Sigma_t$  et la section de collision  $\Sigma_s$ , ainsi que la section d'absorption  $\Sigma_a$  ( $\bar{\mu}$  est une constante liée à l'angle d'incidence du neutron). On relie ces trois sections macroscopiques par  $\Sigma_t = \Sigma_a + \Sigma_s$ . En admettant seulement deux bandes d'énergie, épithermique et thermique, on peut écrire deux équations de diffusion couplées, décrivant chacune un état énergétique, et donnant les flux de neutrons épithermiques  $\Phi_{epith}(r)$  et thermiques  $\Phi_{th}(r)$  arrivant aux détecteurs. Ces flux dépendent de la longueur de ralentissement  $L_s$ , liée aux neutrons épithermiques, et à la longueur de diffusion thermique  $L_d$ , liée aux neutrons thermiques (Ellis, 1987).

Figure 1.11 donne un schéma de l'outil CNT, modèle G. A l'origine, il n'y avait qu'un détecteur des neutrons thermiques. Puis un second détecteur thermique a été ajouté de manière à augmenter le taux de comptage, ce qui permet de diminuer l'erreur statistique sur la porosité mesurée dans les formations très poreuses. De plus, le détecteur le plus proche de la source est utilisé pour compenser des effets de puits sur le détecteur placé plus loin, car on montre que le rapport entre les taux de comptage des deux détecteurs, bien que l'on détecte les neutrons thermiques, varie exponentiellement avec l'inverse de  $L_s$  (Ellis, 1987). Le CNT modèle G incorpore quant à lui deux détecteurs épithermiques en plus des deux détecteurs thermiques. Ces détecteurs sont entourés d'une gaine de cadmium (élément à très forte section de capture des neutrons thermiques) qui les isole des neutrons thermiques provenant de la formation. Il s'en suit une réduction notable du taux de comptage que l'on compense en plaçant les détecteurs épithermiques plus près de la source.

Dans l'espace, la distribution des neutrons thermiques est, nous l'avons vu plus haut, fonction à la fois de  $L_d$ , mais aussi de  $L_s$  qui est un facteur dépendant des neutrons épithermiques.  $L_s$  est très sensible à la quantité d'hydrogène présente dans la formation (Scott et al., 1982; Ellis, 1986) car c'est l'élément qui sera de loin le plus efficace à réduire l'énergie du neutron lors de la réaction élastique (jusqu'à 100% de l'énergie du neutron peut être perdue; voir Section 1.5.1.1). L'information concernant la porosité de la formation se trouve donc d'abord dans la distribution spatiale des neutrons épithermiques, mesurée par les deux détecteurs épithermiques (Davis et al., 1981), et donnée par (Scott et al., 1982)

$$\Phi_{epith}(r) = \frac{Q}{4\pi D r} \cdot e^{-\frac{r}{L_s}}$$

où  $\Phi_{epith}(r)$  est le flux de neutrons épithermiques,  $Q$  est l'intensité de la source en neutrons par seconde,  $D$  est le coefficient de diffusion des neutrons épithermiques (voir plus haut), et  $r$  est la distance à la source. Le rapport des taux de comptage des deux détecteurs épithermiques (voir

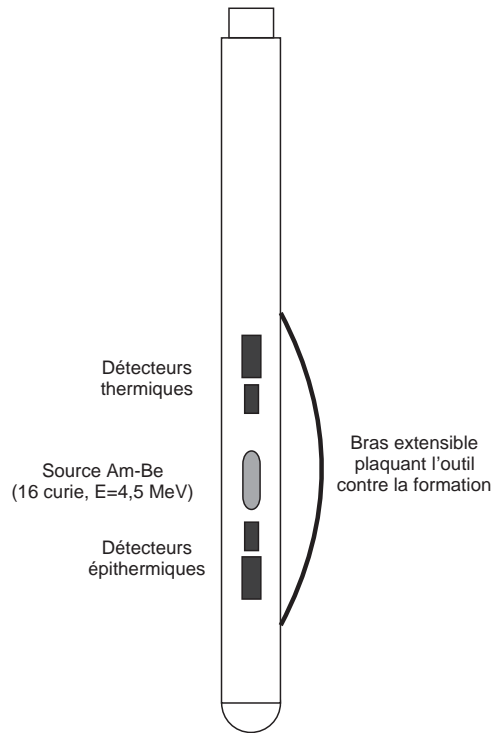


FIG. 1.11 – Schéma de l'outil CNT utilisé pour mesurer l'indice d'hydrogène apparent ou porosité neutron. La présence de deux détecteurs (plus ou moins près de la source) pour chaque type de neutron détecté (thermique et épithermique) permet de réduire les effets environnementaux (dus aux variations de diamètre du puits, ou à la nature du fluide de forage).

équation 1.15) placés à  $r_1$  et  $r_2$  de la source donne une mesure de  $L_s$  (Scott et al., 1982).

$$R_{epith} = \frac{\Phi_{epith}(r_1)}{\Phi_{epith}(r_2)} = \frac{r_1}{r_2} e^{\frac{-(r_1-r_2)}{L_s}} \quad (1.15)$$

En dessous de l'énergie épithermique, et avant qu'ils ne soient capturés, la diffusion des neutrons est caractérisée par le paramètre  $L_d$ , exprimé par

$$L_d = \frac{1}{3[\Sigma_a + \Sigma_s(1 - \bar{\mu})]}$$

Pour caractériser la mesure globale du CNT, qui combine les 4 détecteurs, on introduit le troisième paramètre  $L_m$ , défini à la Section 1.5.1. Le rapport des taux de comptage entre les deux détecteurs thermiques sera une mesure directe de  $L_m$ . Elle est fortement affectée par la présence d'absorbeurs de neutrons thermiques, autres que l'hydrogène, dans le fluide de forage ou la formation, comme le bore, le chlore ou le lithium.

La mesure épithermique donne une mesure de l'indice d'hydrogène indépendante de toute perturbation. La comparaison entre mesures thermique et épithermique permet le calcul des caractéristiques de capture thermique (Scott et al., 1982). La porosité apparente ainsi mesurée, acquise en coups par seconde, est exprimée en % à partir d'une calibration préalable de la sonde pour une lithologie donnée (courbes d'étalonnage).

**Le détecteur** Ils sont composés d'un matériau cible, et d'un détecteur conventionnel (scintillation). Le matériau cible doit avoir une grande section de capture des neutrons et l'énergie produite par la réaction doit être suffisante pour pouvoir être détectée. On utilise généralement le  $^{10}\text{B}$  (réaction  $n,\alpha$ ) ou le  $^3\text{He}$  (réaction  $n,p$ ), l'hélium étant le plus couramment utilisé en diagraphe sous forme de compteur proportionnel à flux gazeux. Les neutrons sont détectés en deux étapes : le neutron réagit avec le matériau cible, et les particules ainsi chargées sont détectées grâce à leurs propriétés ionisantes (Ellis, 1987).

**Résolution verticale et profondeur d'investigation** L'outil CNT effectue une mesure tous les demi-pied (15,24 cm), mais sa résolution est d'environ 50 cm. La profondeur d'investigation est inversement proportionnelle à la porosité de la formation mesurée, et fonction aussi de la géométrie de l'outil (espacement détecteur source) et des conditions de mesure (excentrement de l'outil dans le puits, présence de mud-cake, diamètre du puits). On la considère comme étant d'environ 20 à 30 cm selon la porosité. Ellis (1987) a procédé à des simulations Monte-Carlo pour l'évaluer avec plus de précision. Il montre que, en considérant un modèle considérablement simplifié, comprenant un détecteur et une source espacés de 50 cm, 85% de l'information provient d'un rayon de 25 cm autour de la sonde. Plusieurs modélisations intégrant un paramètre tenant compte de l'isolant en Cd autour des détecteurs épithermiques, ou faisant varier l'espacement détecteur-source, ont été réalisées. Bien que la profondeur d'investigation varie, elle reste dans un intervalle de quelques dizaines de centimètres.

**Effets environnementaux** Un des artefacts les plus connus est l'effet des formations contenant des minéraux argileux en quantité notable (les "shales"). La présence de radicaux hydroxylés (eau liée) dans la plupart des minéraux argileux va induire une porosité neutron apparente plus importante qu'elle n'est en réalité. Plus la quantité d'hydroxyles sera grande, plus la différence entre la porosité neutron et une estimation de la porosité déduite de la mesure de la densité totale (censée être plus proche de la porosité réelle) sera grande. De plus, la présence d'éléments absorbeurs de neutrons thermiques (B, Li, Fe) dans la structure des minéraux argileux va encore augmenter la porosité apparente. La différence se fera par la mesure épithermique qui sera affectée par la présence d'eau liée mais pas par les éléments absorbeurs dans la formation (Ellis, 1986).

Les effets environnementaux principaux affectant la mesure épithermique sont la perte de contact de l'outil avec la mur du puits et le diamètre du puits. Les variations de ce dernier peuvent être corrigées par l'intégration de la mesure du caliper au traitement. Une méthode de processing des taux de comptage a été développée pour tenir compte à la fois de la perte de contact de l'outil et de l'effet du fluide de forage. Scott et al. (1982) montre que cette méthode donne une valeur de la porosité peu sensible à ces effets environnementaux, et permet de corriger de la température, de la salinité et de la pression présentes dans le puits durant la mesure.

## Chapitre 2

# La diagraphie électrique

### 2.1 Principes

#### 2.1.1 Résistance et résistivité

La *résistivité* est une grandeur physique qui indique la capacité pour un matériau à laisser circuler le courant électrique. C'est une notion différente de la *résistance* car elle inclut une dimension géométrique du matériau considéré. La loi d'Ohm, qui s'écrit  $U = R \cdot I$ , ou dans sa forme plus générale  $\vec{J} = \sigma \vec{E}$ , où  $\vec{J}$  est la densité de courant,  $\vec{E}$  le champ électrique, et  $\sigma$  la conductivité, permet de définir la relation entre résistivité et résistance (Ellis, 1987). La résistance d'un matériau de surface  $A$  et de profondeur  $l$  est ainsi donnée par

$$R = \frac{U}{I} = \frac{1}{n_i \mu q^2} \frac{l}{A} \quad (2.1)$$

Le second membre,  $l/A$ , représente le paramètre géométrique décrivant la zone d'influence étudiée. Le premier membre,  $1/(n_i \mu q^2)$ , est dépendant du matériau considéré, et est appelé *résistivité*. Elle est exprimée en *ohm.m* ( $\Omega.m$ ). Dans le système international, un  $\Omega.m$  est égal à la résistance électrique d'un volume unitaire de  $1 \text{ m}^3$ .

La conduction présentant un intérêt en diagraphie est la conduction *électrolytique*, faisant intervenir les sels dissouts dans un liquide qui est la plupart du temps de l'eau. Dans les formations géologiques, la conductivité dépendra en premier lieu de la résistivité et de la quantité d'eau présente dans les pores. Elle dépendra aussi, mais dans une moindre mesure, de la lithologie (présence d'argiles, texture de la roche et distribution des pores). Elle est aussi très fortement affectée par la température. Pour comprendre ceci, il faut introduire un paramètre supplémentaire dans la définition de la résistivité, la viscosité  $\eta$ . La dépendance de la viscosité à la température est exprimée par

$$\eta = \eta_0 \cdot e^{\frac{C}{T}}$$

où  $C$  est caractéristique du liquide. La loi de Stoke, associée à la relation 2.1 permet d'écrire la résistivité en fonction de la viscosité du fluide impliqué, en faisant l'hypothèse que l'électrolyte est présent sous la forme de sphères de diamètre  $a$  et de charge  $q$  :

$$R = \frac{6\pi\eta a}{nq^2} \frac{l}{A} \quad (2.2)$$

Pour une concentration en électrolytes donnée, la résistivité d'une solution décroît fortement lorsque la température augmente.

L'inverse de la résistivité est la conductivité  $\sigma$ . Elle s'exprime en milliSiemens par m (mS/m).

### 2.1.2 Le poids de l'empirisme

Toute l'interprétation de la résistivité dans les formations géologiques est basée sur les relations empiriques développées grâce aux travaux de Leverett (1938) et Archie (1942).

C'est en faisant des expériences dans le but de trouver un lien entre résistivité et perméabilité qu'Archie (1942) a proposé plusieurs paramètres rendant compte des propriétés électriques de la formation. Il a fait ses expériences dans des formations exemptes d'argiles, car l'eau liée contenue dans ces minéraux influe énormément sur la mesure. Il a ainsi remarqué que, quelle que soit la résistivité d'un fluide saturant l'espace poreux d'un échantillon, la résistivité totale de l'échantillon est toujours proportionnelle à la résistivité du fluide par un facteur constant, qu'il a appelé *Facteur de Formation*, noté  $FF$  :

$$R_{\text{échant}} \equiv R_0 = FF \cdot R_{\text{eau}}$$

Archie (1942) a aussi expérimentalement établi une relation entre  $FF$  et la porosité de la formation, sous la forme :

$$FF \approx \frac{1}{\phi^m}$$

où l'on voit apparaître  $m$ , le coefficient de cimentation, caractérisant la texture et la tortuosité (géométrie) du chemin poreux. De la même façon, la saturation en eau (par rapport aux hydrocarbures) dans la formation est reliée à la résistivité par une relation du type

$$S_w = \left[ \frac{R_t}{R_0} \right]^{-\frac{1}{n}}$$

où  $n$  est le coefficient de saturation, qui est aussi très proche de 2 pour les formations non argileuses. La relation d'Archie généralisée permet de calculer la résistivité totale de la formation  $R_t$  en fonction de leur porosité effective  $\phi_e$ , de la saturation en fluide conducteur  $S_w$  et de la résistivité du fluide  $R_w$ . Elle s'écrit sous la forme :

$$R_t = \frac{a \cdot R_w}{\phi_e^m \cdot S_w^n}$$

$a$  représente le type de porosité (dépendant de la consolidation de la roche,  $a$  varie de 0,81 dans les sables à 1 dans les roches indurées).  $n$  varie de 1,7 à 2,2. Pour les formations sans argiles, on obtient une bonne approximation en prenant  $m = n = 2$ . Néanmoins, il est préférable de les déterminer expérimentalement dans des formations mélangées. Dans de nombreux cas, on peut simplifier les relations d'Archie :

$$FF \approx \frac{1}{\phi^2} \quad S_w = \sqrt{\frac{R_t}{R_0}} \quad R_t = \frac{R_w}{\phi_e^2 \cdot S_w^2} \quad (2.3)$$

Pour décrire la résistivité des formations poreuses contenant une proportion substantielle d'argiles, il est nécessaire d'utiliser des relations plus complexes qui font intervenir plusieurs types de conducteurs, le chemin poreux d'une part, avec sa tortuosité, et les argiles d'autre part.

## 2.2 La diagraphie de résistivité : les dispositifs focalisés

Les outils de mesure de la résistivité se basent sur la loi d'Ohm appliquée aux milieux homogènes :

$$\vec{E} = -\vec{\nabla}V = R \cdot \vec{i}$$

où  $\vec{E}$  est le champ électrique,  $\vec{\nabla}V$  le gradient de potentiel,  $R$  la résistivité et  $\vec{i}$  la densité de courant. Ils sont destinés soit à la mesure de la résistivité de la zone non-envahie, appelée  $R_t$ , soit à la mesure de la résistivité de la zone immédiatement en contact avec le puits de forage, ou  $R_{xo}$ .

### 2.2.1 Le Laterolog™

Les premiers outils mesurant la résistivité étaient basés sur le principe de l'étude de la distribution d'un courant continu autour d'une électrode d'émission ponctuelle. Dans un milieu très conducteur, comme l'est en général un forage, les lignes de courant ont tendance à se concentrer dans la boue ou les formations conductrices (argiles), et évitent les formations résistantes. La profondeur d'investigation de ces outils était très faible, et le seul moyen de l'améliorer, donc d'accéder à la mesure de  $R_t$ , était d'augmenter la distance émetteur-récepteur. Mais ceci ne peut se faire sans une perte d'information due à une baisse drastique de la résolution verticale. C'est pour cela que les outils dits *focalisés* ont été mis au point.

La Figure 2.1 montre le principe de fonctionnement du Laterolog™. Cet outil se base non plus sur un courant continu, mais sur un courant alternatif de basse fréquence. Le croquis de gauche montre la différence entre les lignes de courant dans un dispositif non focalisé (gauche) et un dispositif focalisé (droite). La présence d'un niveau très résistif ( $R_t$ ) perturbe les lignes de champ du dispositif non focalisé, qui ont tendance à passer autour, plutôt à travers les épontes plus conductrices ( $R_s$ ). La focalisation oblige au contraire les lignes de courant à traverser le niveau très résistif qui sera ainsi correctement mesuré.

Le principe de la focalisation est d'utiliser non plus une, mais trois électrodes, dont le potentiel est maintenu constant à la même valeur. Comme il n'existe de courant qu'en cas de différence de potentiel, il n'y a pas de lignes de courant entre l'électrode centrale  $A_0$  et les deux électrodes de focalisation  $A_1$  et  $A'_1$  situées de part et d'autre de  $A_0$ . Le courant est forcé à l'horizontale dans la formation. Une autre configuration utilise 7 électrodes, les trois électrodes précédemment décrites, auxquelles on ajoute 4 électrodes de contrôle fonctionnant par paire placées entre  $A_1/A'_1$  et  $A_0$ . On fait varier le courant parcourant  $A_1$  et  $A'_1$  pour que le potentiel entre chaque paire d'électrodes de contrôle soit nul, ce qui a pour effet de focaliser le courant dans la formation.

### 2.2.2 La focalisation sphérique

Elle permet de s'affranchir des effets induits sur la mesure par l'existence même du puits. En effet, dans un milieu homogène, les lignes de courant sont également réparties autour de l'électrode, et les surfaces équipotentielles sont des sphères. Dans un milieu comme le puits de forage, le courant a tendance à aller vers les milieux plus conductifs, comme la boue de forage, et les équipotentielles deviennent des ellipsoïdes. La focalisation sphérique a pour but de forcer ces équipotentielles à redevenir des sphères. La configuration des électrodes ainsi que les lignes de courant obtenues sont montrées sur la Figure 2.1, schéma de droite. Les paires  $M_1 - M'_1$  et  $M_2 - M'_2$  sont utilisées pour contrôler le courant de focalisation, mais cette fois leur différence de potentiel n'est pas nulle mais maintenue constante.  $A_0$  fournit le courant de focalisation retournant aux électrodes  $A_1$  et  $A'_1$ , ainsi que le courant de mesure, retournant à une électrode de mesure disposée en surface. Le volume compris entre les deux demi-cercles pointillés, représentant les surfaces équipotentielles, est mesuré par la sonde, à l'exception de la zone à l'interface puits-formation, où les courants de focalisation bloquent toute mesure. Ce volume dépendra de la distance entre les électrodes de mesure  $A$ .

Ces trois configurations sont utilisées alternativement par l'outil Dual Laterolog™ de Schlumberger (voir Figure 2.2). En alternant rapidement le rôle des divers courants de mesure et de

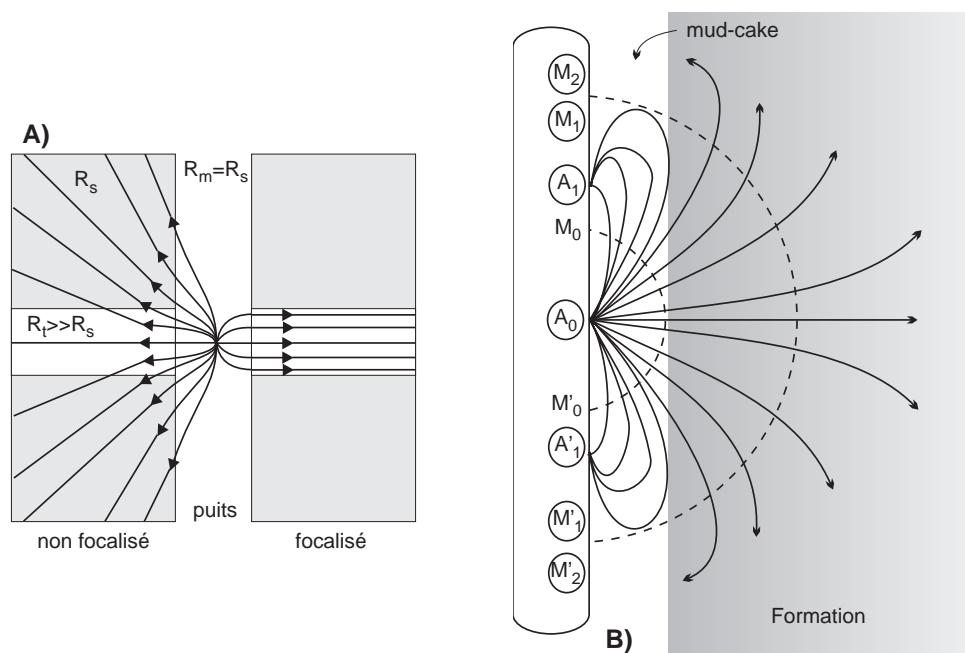


FIG. 2.1 – **A.** Différence entre un outil focalisé et un outil non focalisé : La résistivité de la formation  $R_s$  est sensiblement égale à celle du fluide de forage  $R_m$ . Le courant a tendance à éviter le banc central de résistivité  $R_t$  très supérieure à  $R_s$ . La focalisation oblige le courant à suivre l'horizontale quelle que soit la résistivité rencontrée. **B.** La focalisation sphérique permet de rétablir des équipotentielles sphériques pour ainsi s'affranchir de l'effet du puits et pouvoir mesurer la zone proche de l'interface puits–formation (d'après (Ellis, 1987)).

focalisation, on obtient une mesure de la résistivité à différentes profondeurs d'investigation : profonde (“deep”, ou LLD), moyenne (“shallow” ou LLS), et faible (utilisant le principe de la focalisation sphérique SFLU). Ceci permet d'avoir une sensibilité différente à  $R_t$  et  $R_{xo}$ . La résolution verticale sera fonction de la distance entre les électrodes de contrôle  $M - M'$ .

## 2.3 Les sondes à induction

Le paramètre de mesure requis pour l'usage du Laterolog<sup>TM</sup> est une boue de forage relativement conductrice. La précision sur  $R_t$  sera d'autant meilleure que la différence entre  $R_{xo}$  et  $R_t$  sera grande. Pour s'affranchir de ces limitations, les outils utilisant les phénomènes d'induction ont été développés.

### 2.3.1 La diagraphie de conductivité

Les sondes à induction utilisent des bobines parcourues par un courant alternatif  $I_t$  qui induit un champ magnétique alternatif dit primaire  $B_t$  dans la formation. Le principe physique est schématisé Figure 2.3. C'est la composante verticale  $(B_t)_z$  qui sert à la mesure, et on l'écrit :

$$(B_t)_z \propto I_0 e^{-i\omega t}$$

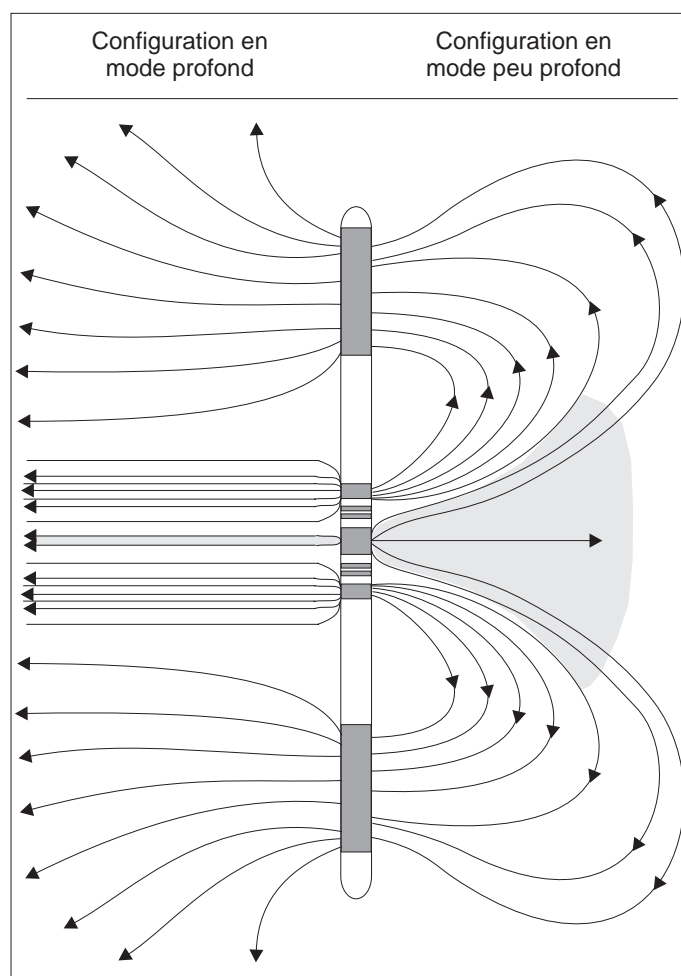


FIG. 2.2 – Schéma de l'outil Dual Laterolog ou DLL (d'après Chemali et al. (1983)). Il utilise plusieurs configurations pour mesurer la résistivité à différentes profondeurs d'investigation : LLD pour la zone non envahie et relativement loin de l'électrode (configuration de gauche), LLS pour la zone plus proche et potentiellement envahie par le fluide de forage (configuration de droite), et enfin utilise la focalisation sphérique (Figure 2.1) pour mesurer la résistivité de la zone très proche de l'interface puits-formation.

Ce champ magnétique variable dans le temps crée à son tour un champ électrique  $E$  dont l'amplitude est proportionnelle à la dérivée du flux magnétique, selon la loi de Faraday :

$$E \propto -\frac{\partial(B_t)_z}{\partial t} \propto i\omega I_0 e^{-i\omega t}$$

Ce champ électrique induit une densité de courant dans une portion de la formation que l'on peut représenter sous la forme d'un anneau dont l'axe est aligné avec l'axe de l'outil (voir Figure 2.3). La densité de courant  $J$  est proportionnelle à la conductivité et à la perméabilité magnétique des formations traversées :

$$J \propto \sigma E \propto i\omega\sigma I_0 e^{-i\omega t}$$

Ces courants circulaires secondaires sont appelés *courants de Foucault* ou “eddy currents” dans les pays anglo-saxons. Ils se comportent comme la bobine émettrice et créent à leur tour leur propre champ magnétique secondaire sinusoïdal,  $B_2$ , dont la composante verticale s'écrit, comme



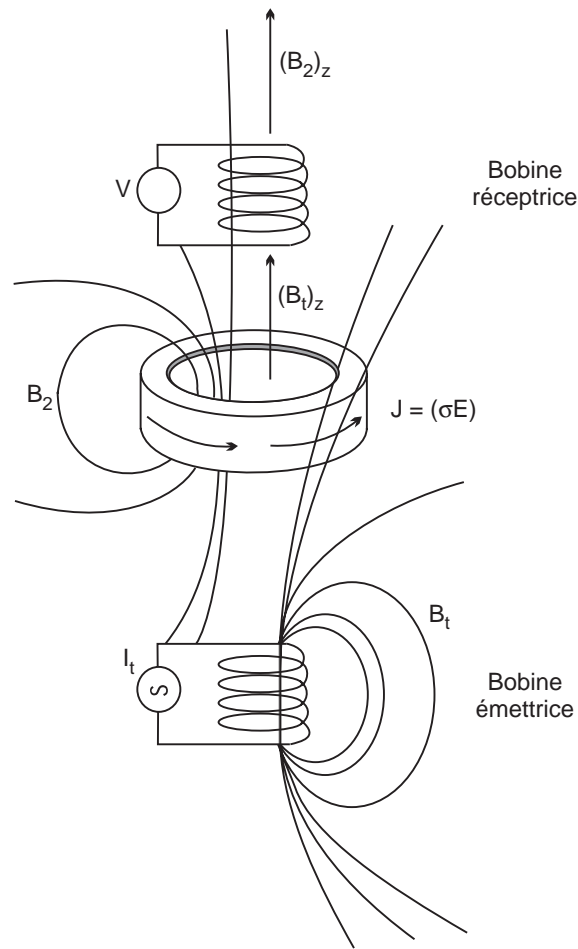


FIG. 2.3 – Principe des sondes à induction (d'après Ellis (1987)).

$(B_t)_z$

$$(B_2)_z = i\omega\sigma I_0 e^{-i\omega t}$$

et induit un voltage  $V$  dans la bobine réceptrice tel que

$$V \propto -\frac{\partial(B_2)_z}{\partial t} \propto -\omega^2 \sigma I_0 e^{-i\omega t}$$

La bobine réceptrice permet de mesurer l'induction totale (primaire et secondaire) qui s'exprime sous forme d'une partie réelle et d'une partie imaginaire, en prenant le courant de la bobine émettrice comme référence de base.

Un paramètre d'importance est le *facteur géométrique* de la sonde. La dépendance géométrique du signal détecté est exprimée par

$$g(z, r) \propto \frac{r}{\rho_t^3} \frac{r^2}{\rho_r^3}$$

Elle donne la contribution d'une tranche de formation située à une profondeur  $z$  et d'un rayon  $r$  au résultat délivré par la bobine réceptrice. On définit pareillement deux facteurs géométriques

radial,  $g(r)$ , et vertical,  $g(z)$ , pris en intégrant sur l'infini.  $g(r)$  donne l'importance relative de chaque cylindre de rayon  $r$  d'axe confondu avec l'axe de l'outil dans la réponse finale :

$$g(r) = \int_{-\infty}^{\infty} g(r, z) dz$$

On montre (Ellis, 1987) que  $g(r)$  est maximum pour des rayons légèrement inférieurs à la distance entre les deux bobines réceptrice et émettrice.

$g(z)$  donne l'importance d'une tranche de formation située à une profondeur  $z$  dans la réponse finale :

$$g(z) = \int_0^{\infty} g(r, z) dr$$

On montre (Ellis, 1987) que la réponse est fidèle pour la partie de la formation située entre les deux bobines, mais qu'au delà, la diminution du facteur géométrique vertical induit des distorsions au niveau de la reconnaissance des limites de bancs de conductivités différentes.

Pour corriger de ces effets, on utilise des rangées de bobines émettrices et réceptrices dans le même outil (c'est le cas par exemple du Dual Induction Tool ou DIT de Schlumberger). De plus les courants émis sont focalisés. Cette configuration permet de s'affranchir des effets environnementaux induits par la partie de la formation proche de l'interface puits-formation, et donc d'obtenir une profondeur d'investigation améliorée et surtout une mesure de  $R_t$  précise et reflétant la résistivité réelle de la formation. D'autres améliorations permettent encore d'améliorer la réponse de l'outil aux limites de bancs et la profondeur d'investigation, et sont expliquées en détail dans Barber (1986a) et Barber (1986b).

### 2.3.2 Induction ou électrode ?

Les sondes à induction sont d'autant plus performantes que la boue de forage est résistante, et que les formations étudiées sont conductrices, et elles sont donc en cela complémentaires des Latéologs. Les sondes à induction bénéficient d'une profondeur d'investigation supérieure dans de bonnes conditions de mesure, et donnent une mesure précise de la résistivité de la zone non-entahie  $R_t$ . La résolution des sondes à induction est aussi meilleure de l'ordre de 30 à 60 cm. Par contre, dans des formations très résistives, le latéolog peut être utilisé avec succès.

### 2.3.3 La diagraphie de susceptibilité magnétique

Elle correspond à une application particulière des dispositifs à induction (sonde dipôle-dipôle), montré Figure 2.4. Un courant alternatif de basse fréquence (200 Hz) est envoyé dans la formation par la bobine émettrice, et la bobine réceptrice enregistre la réponse de la formation. On rajoute une troisième bobine qui sert à compenser le flux direct entre les deux bobines. Le flux détecté par la bobine réceptrice se décompose en un signal en phase avec le flux direct et un signal en quadrature (Thibal, 1995). Le signal en quadrature est produit par les courants de Foucault (voir Section 2.3.1) et sert à mesurer la conductivité de la formation. La réponse en phase est d'intérêt ici car elle est liée linéairement à la susceptibilité magnétique de la formation (Thibal, 1995). Augustin (1990) a défini la distance émetteur-récepteur la plus appropriée pour réduire l'effet de la boue de forage et ainsi privilégier la réponse de la formation. Le phénomène physique étudié est classique et concerne la polarisation magnétique des minéraux ferromagnétiques, appelés "dipôles élémentaires" et présents dans la formation, sous l'effet d'un champ magnétique. On définit un vecteur d'aimantation  $\vec{J}$  et un vecteur d'induction magnétique  $\vec{B}$  :

$$\vec{J} = \chi \vec{H}$$

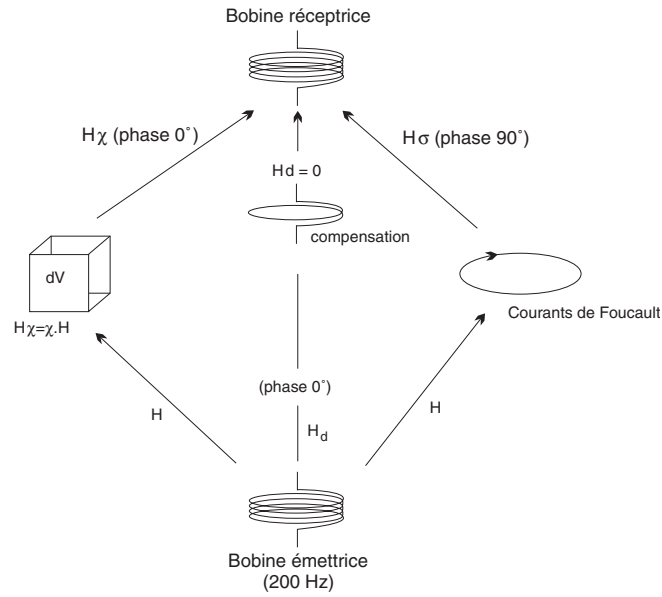


FIG. 2.4 – Principes de mesure utilisés dans la sonde dipôle-dipôle servant à la mesure de la susceptibilité magnétique de la formation (d'après Thibal (1995)).  $H_d$  est le flux direct,  $H_\sigma$  est le signal en quadrature produit par les courants de Foucault et donnant une information sur la conductivité, et  $H_\chi$  est le signal en phase lié à la susceptibilité magnétique de la formation.

$$\vec{B} = \mu \vec{H} = \mu_0(1 + \chi) \vec{H}$$

où  $\chi$  est la *susceptibilité magnétique*,  $\mu$  la perméabilité magnétique et  $\vec{H}$  le champ magnétique terrestre. La susceptibilité magnétique  $\chi$  est donnée en ppm de flux direct, sachant que 1 ppm vaut  $10^{-6}$  SI pour un diamètre de puits standard (Thibal, 1995). Elle dépend évidemment de la composition minéralogique de la formation, de la nature des minéraux magnétiques, de la taille des grains, et de leur concentration. La perméabilité magnétique du vide  $\mu_0$  vaut  $4\pi \cdot 10^{-7}$  Henry/m (H/m). Pour une formation géologique, elle s'exprime sous la forme :

$$\mu = \mu_r \cdot \mu_0$$

où  $\mu_r$  est la perméabilité relative, égale à  $1 + \chi$ . La valeur de  $\mu_r$  dépend de la teneur en minéraux ferromagnétiques de la formation. Pour mémoire, les minéraux paramagnétiques incluent les minéraux argileux (en particulier la chlorite, la smectite, la glauconie), les silicates ferromagnésiens, les carbonates de fer (sidérite), les disulfures de fer (pyrite), et les minéraux authigènes et oxyhydroxydes ferriques regroupés sous le terme de limonite. Leur susceptibilité magnétique est très faible et de l'ordre de  $10^{-8}$  SI (Dunlop and Özdemir, 1997). Les minéraux ferrimagnétiques incluent les oxydes fer-titane (titanomagnétite et titanomaghémite) dont les susceptibilités varient de  $10^{-2}$  à 1, et les monosulfures de fer (pyrrhotite, greigite), entre  $10^{-5}$  et  $10^{-2}$ . D'autres minéraux comme l'hématite ou la goethite sont ferrimagnétiques mais avec des susceptibilités magnétiques plus faibles (entre  $10^{-6}$  et  $10^{-4}$ ). Dans les formations géologiques exemptes de sulfures ou d'oxydes de fer, la susceptibilité magnétique est de l'ordre  $10^{-6}$  à  $10^{-2}$ , et on pose  $\mu = \mu_0$ .

La mesure de la susceptibilité magnétique, en dehors de l'estimation qualitative de la teneur en minéraux ferriques, sert aussi à retrouver l'aimantation rémanente  $J_r$  fossilisée par les grains ferromagnétiques (voir Chapitre 3).

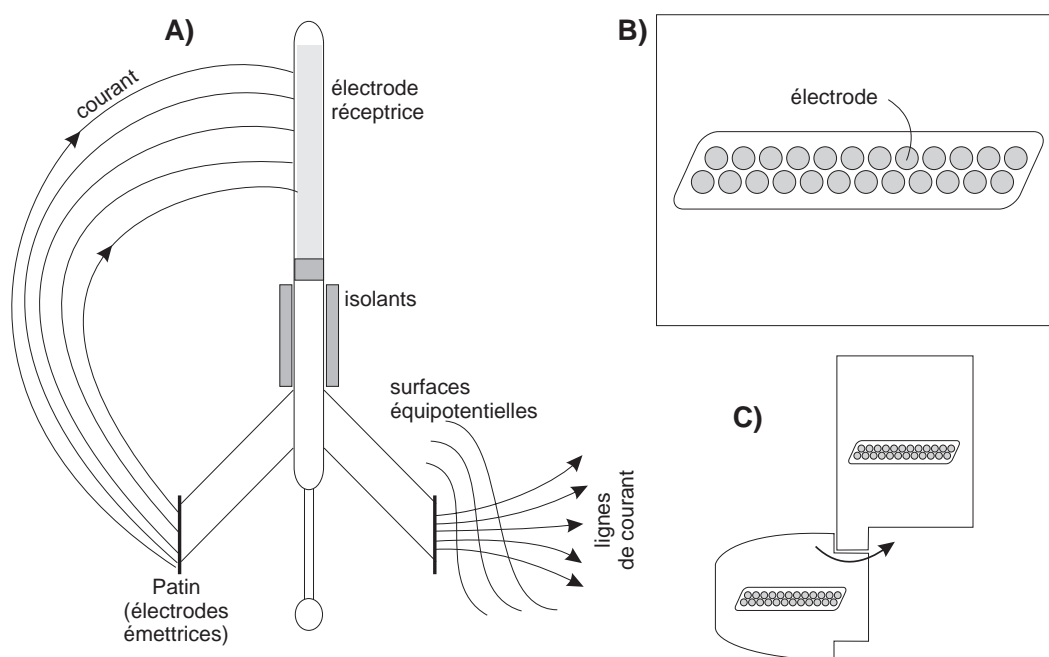


FIG. 2.5 – **A.** Principes des outils d'imagerie de micro-résistivité (Dipmètre, Formation Micro Scanner FMS et Fullbore Micro Imager FMI) : un courant électrique est envoyé dans la formation par les électrodes montées sur les patins et une réponse est enregistrée par l'électrode située au sommet de l'outil. **B.** Disposition des 24 électrodes sur un patin du FMI. Le diamètre de chaque électrode est de 5 mm. Les deux rangées sont disposées à 5 mm de distance, et la deuxième rangée est latéralement déplacée de 2,5 mm par rapport à la première (quinconce). Ceci permet d'obtenir une mesure tous les 2,5 mm. **C.** Le FMI possède en plus de 4 patins, 4 patins supplémentaires comportant eux aussi 24 électrodes chacun, qui se déploient pour couvrir près de 76% de la circonférence du puits de forage.

## 2.4 Imagerie de la paroi du trou

### 2.4.1 Principe

Des outils utilisant des micro-électrodes ont été inventés pour mesurer la résistivité  $R_{xo}$  de la zone envahie. La profondeur d'investigation de tels outils a été volontairement limitée à quelques centimètres. Le diamètre des électrodes est de l'ordre du centimètre. Elles sont montées sur des *patins* (Figure 2.5) qui par pression les gardent en contact permanent avec la formation (murs du puits).

Chaque électrode envoie un courant alternatif dans la formation, courant enregistré par une électrode réceptrice située dans la partie supérieure de l'outil, sous forme d'une série de courbes représentant des variations relatives dans la micro-résistivité des bords du puits. Ces variations sont dues par exemple à des changements de conduction électrolytique, fonction de la géométrie du chemin poreux, ou bien à des variations dans la capacité à échanger des cations (CEC) des minéraux argileux ou conducteurs. La variation de la micro-résistivité des murs du puits répondra donc principalement à des changements de porosité, de taille de grain, de minéralogie, de cimentation et de type de fluide. Durant l'acquisition, un accéléromètre tri-axial permet de déterminer la position de l'outil en profondeur, et trois magnétomètres donnent son orientation azimutale en temps réel. Le courant atteignant l'électrode réceptrice est caractérisé par trois composantes :

- une composante haute-fréquence, qui répond aux variations de micro-résistivité en face de chaque électrode, et qui permet une excellente résolution verticale et azimutale, mais réduit la profondeur d’investigation de l’outil. C’est cette composante qui est liée aux variations lithologiques, pétrophysiques et minéralogiques de la formation.
- une composante basse-fréquence, qui répond à la résistivité de la formation mais sur une profondeur latérale plus étendue (similaire à celle d’un latérolog réglé sur une faible profondeur d’investigation, ou LLs). Elle calibre l’image par rapport à une donnée plus quantitative, comme LLs, mais ayant une profondeur d’investigation du même ordre de grandeur. Ceci permet d’utiliser l’image de micro-résistivité d’une manière plus quantitative pour l’analyse texturale ou stratigraphique.
- des courants continus créés par la friction de l’outil sur les murs du puits, ou correspondant à des phénomènes de potentiel spontané. Ces courants sont éliminés lors du traitement du signal pour produire les images.

Le Dipmètre, premier outil développé par Schlumberger utilisant le principe de micro-électrodes montées sur patin, est utilisé pour mesurer l’orientation des couches sédimentaires. Il comprend deux ou quatre patins comportant chacun une ou deux électrodes. Les mesures ne sont pas nécessairement exploitées en terme de micro-résistivité, mais la succession verticale d’anomalies de micro-résistivité permet de déterminer l’orientation en trois dimensions des strates que traverse le puits de forage. Par exemple, pour un puits vertical traversant des couches horizontales, les anomalies de micro-résistivité marquant les limites entre deux strates doivent être mesurées par les patins à une profondeur identique. Si les couches sont affectées d’un pendage non nul, les anomalies seront perçues à une profondeur différente par chaque patin. La correction nécessaire pour aligner les anomalies dans le plan horizontal sera fonction de la taille du trou et du pendage des couches.

Le Stratigraphic High Resolution Dipmeter (SHDT) de Schlumberger mesure la microrésistivité à partir de deux patins comportant chacun de 27 électrodes. Le Formation Micro Scanner (FMS) dispose de 4 patins de 16 électrodes chacun. La couverture des murs du puits est évidemment meilleure que pour le SHDT et atteint environ 40%. Un outil récemment développé, le Fullbore Micro Imager (FMI), dispose de 4 patins, comportant chacun 24 électrodes ainsi qu’un patin plus petit et rabattable, comportant lui aussi 24 électrodes, ce qui donne en tout 8 patins de 24 électrodes chacun, couvrant environ 76% de la circonférence du puits. Les électrodes sont placées de telle façon que le pas d’échantillonnage est de 2,5 mm (Figure 2.5). Chaque électrode mesure 5 mm de diamètre. La résolution théorique de l’outil est de 2,5 cm, mais il peut détecter des objets très petits, dans le cas où le contraste de résistivité entre l’objet et la matrice est grand, et que de tels objets sont séparés par au moins 1 cm. Dans ce cas, la résolution peut atteindre 5 mm.

#### 2.4.2 Traitement des données et production de l’image

Quatre étapes sont nécessaires pour corriger des artefacts de mesure afin de produire des images utilisables.

1. Correction de vitesse : les données de micro-résistivité sont filtrées grâce à un filtre en deux étapes utilisant la mesure de l’accéléromètre pour compenser des mouvements irréguliers de l’outil durant l’acquisition, corrigeant aussi la mesure lorsque l’outil est resté coincé, et réduisant les effets “dents-de-scie”, qui sont un artefact de mesure dû à la présence de contraste de résistivité trop fort en face des électrodes (le problème a été partiellement résolu en réduisant la distance entre les électrodes sur les patins du FMS).
2. Correction de voltage : l’outil fait varier dynamiquement le courant qu’il émet pour garder

son intensité dans la fenêtre optimale de mesure. En clair, il augmente le courant en face d'une formation résistive, pour qu'une partie de ce courant pénètre dans la formation et que l'on puisse obtenir une mesure. Dans le cas d'une formation conductrice, l'outil diminue le courant émis pour éviter de saturer. Les corrections de voltage tiennent compte de ces ajustements permanents et permettent d'éliminer les distortions dues à des formations présentant des résistivités très différentes.

3. Egalisation du signal pour toutes les électrodes à un même niveau. Les mesures peuvent être affectées par une dérive électronique, ou par un mauvais contact d'un ou plusieurs patins contre la formation. Le principal objectif de l'égalisation est de modifier statistiquement les données de chaque électrode de manière à ce que chacune ait la même réponse à la résistivité mesurée que les autres. L'égalisation compense les différences de gain et le décalage entre les enregistrements ("offset") entre les électrodes sur un même niveau de mesure, en remplaçant pour chaque électrode, son gain et son offset par le gain et l'offset moyen de toutes les électrodes calculés sur une fenêtre glissante (d'environ 5 m). De plus on élimine de la statistique les parties extrêmes de l'histogramme des données, qui proviennent souvent de mauvaises mesures (puits détérioré). Seules les données provenant de mesures sur la roche saine sont utilisées dans l'égalisation. On utilise aussi des méthodes statistiques pour corriger des déficiences localisées d'une ou plusieurs électrodes lors de la mesure ("dead buttons").
4. Conversion de l'intensité du courant mesuré en tons de couleur. Le nombre de tons de couleur dépend de la capacité informatique dont l'on dispose. Il sera de 16 ou 48 sur les sorties papier, et atteint 256 niveaux de gris sur ordinateur. Par défaut, on établit une distribution cumulée des intensités de courant et l'intervalle total peut être ainsi partitionné en 48 classes, chacune comportant le même nombre de données. Les teintes s'échelonnent depuis le blanc (plus résistif), en passant par l'orange jusqu'au noir (plus conducteur).

La méthode la plus commune est d'utiliser l'ensemble des données pour établir l'intervalle des conductivités et définir la limite des classes. C'est la *normalisation statique*. Une même teinte indiquera une résistivité identique à n'importe quelle profondeur du puits. Cette représentation est utile pour décrire des variations de résistivité à grande échelle ou des contrastes entre plusieurs unités lithologiques. Par contre, les variations de résistivité à plus petite échelle, telle que fractures, figures de dissolution (stylolithes) ou pores, ne seront pas visible dans cette représentation. Il faut alors utiliser la *normalisation dynamique*, qui définit l'intervalle des conductivités sur une étroite fenêtre verticale (généralement 1 m). La totalité des teintes est utilisée dans cette fenêtre, ce qui permet d'obtenir le maximum de contraste pour la fenêtre donnée. Par contre, on ne peut pas avec cette méthode comparer directement des objets ayant la même teinte à deux profondeurs distinctes. Ce sont les variations relatives de micro-résistivité à un endroit donné qui seront étudiées.

Après ces étapes, les images sont prêtes à être affichées et analysées.



## Chapitre 3

# La diagraphie magnétique

### 3.1 La résonance magnétique nucléaire

#### 3.1.1 Principes

Les outils de mesure par résonance magnétique sont basés sur un principe physique simple, la réponse d'un noyau atomique à un champ magnétique (Bloch, 1946; Bloch et al., 1946; Feynman et al., 1965). Tous les noyaux des éléments chimiques ont un *moment angulaire*  $J$ , représentant le mouvement du noyau tournant rapidement sur lui-même. Mais seuls les noyaux à nombre impair de protons et de neutrons peuvent présenter un magnétisme détectable. Le plus simple de tous est l'hydrogène dont le noyau ne comporte qu'un proton. C'est une chance car c'est le composant majeur de l'eau, avec l'oxygène, mais celui-ci ne présente pas de propriétés magnétiques.

Le noyau de l'hydrogène (= proton) est animé d'un mouvement spontané de rotation rapide sur lui-même appelé *spin*, représenté par le vecteur de moment angulaire. La charge électrique positive du proton induit un champ magnétique microscopique représenté par un vecteur de *moment magnétique*  $\mu$ , dont la longueur et la direction correspondent respectivement à son intensité et à son orientation. Dans le cas du proton, le moment magnétique est donné par

$$\mu = \gamma \cdot J \quad (3.1)$$

où  $\gamma$  est la constante gyromagnétique caractérisant un type de noyau. Pour le proton,  $\gamma$  vaut  $4,2576 \cdot 10^3$  Hz/Gauss ( $4,2576 \cdot 10^7$  Hz/Tesla). A l'état naturel, le champ magnétique terrestre  $H_t$  polarise les noyaux, mais trop faiblement pour que cette polarisation soit perceptible. Néanmoins, les premiers outils de diagraphie par résonance magnétique essayaient d'en tirer partie en utilisant de grandes bobines pour produire un champ magnétique dont le but était de polariser les protons présents dans l'eau libre et les hydrocarbures. Ce champ était interrompu et le champ terrestre était utilisé pour produire le signal de décroissance représentant la porosité du fluide libre. Les problèmes de ces outils étaient nombreux : le fluide de diagraphie devait être dopé avec de la magnétite pour éliminer le signal provenant du puits, les courants provenant des bobines saturaient les récepteurs à chaque mesure, rendant l'outil sensible seulement aux protons relaxant lentement (représentant l'eau libre). Le développement des outils dits à "impulsion-écho" a permis de s'affranchir de ces problèmes.

A l'état naturel, les protons ont des moments magnétiques individuels  $\mu$  orientés au hasard, et leur somme correspond alors à une aimantation totale  $M$  nulle. Pour obtenir une aimantation mesurable, on applique, pendant tout le cycle de mesure, un champ magnétique statique externe  $H_0$  plus intense (entre 500 et 5000 Gauss) au champ magnétique terrestre (qui est environ de 0,5 Gauss ou  $0,5 \cdot 10^{-4}$  Tesla) de façon que celui-ci soit considéré comme négligeable (voir Figure 3.1).



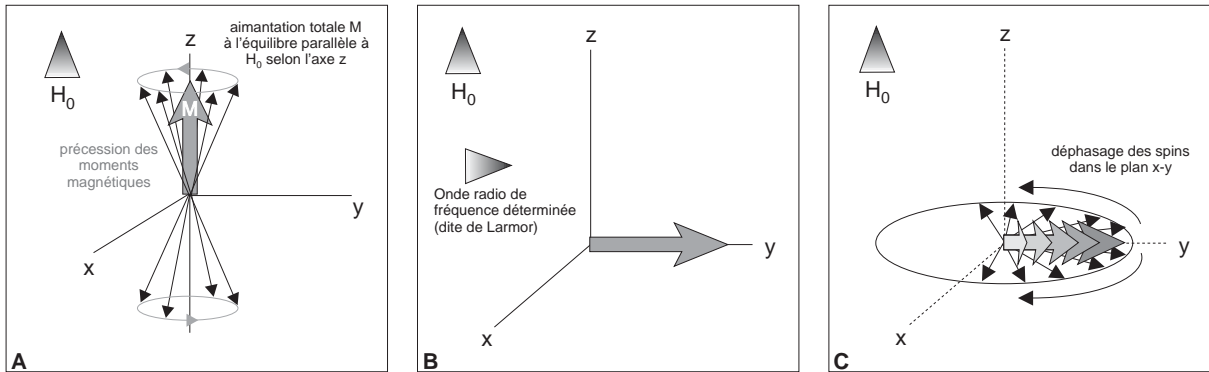


FIG. 3.1 – Les étapes d’une mesure par résonance magnétique nucléaire effectuée par l’outil CMR. **A.** les spins des noyaux d’hydrogène (protons) sont alignés par application d’un champ magnétique constant  $H_0$ . La composante longitudinale  $M$  suivant l’axe de  $H_0$  est non nulle, due aux deux états d’énergie des moments magnétiques des protons (parallèle et antiparallèle). **B.** Basculement des spins à 90 degrés de  $H_0$  par application d’un rayonnement électromagnétique  $H_1$  oscillant à la fréquence de Larmor. **C.** les protons sont alors affectés d’un mouvement de précession dans le plan x-y, générant un champ magnétique oscillant à la fréquence de Larmor, qui donne un signal détecté par l’antenne de l’outil. A la coupure de  $H_1$ , les spins se déphasent et ce signal, dit d’induction libre, décroît selon une constante  $T_2^*$ , indépendante de la formation. Une quatrième étape non indiquée ici consiste en plusieurs cycles resynchronisation des spins/relaxation servant à mesurer la constante  $T_2$  (d’après Kenyon et al. (1995)).

$H_0$  est produit par un aimant permanent, et sa direction est perpendiculaire à l’axe du puits de forage. Les spins s’alignent alors dans la direction de  $H_0$ , mais pas de façon strictement parallèle à son axe. Ils tournent autour de lui avec une vitesse angulaire et une inclinaison fixe proportionnelle à la puissance de  $H_0$ . Ce mouvement s’appelle la *précession* (voir Figure 3.2-A). Les moments magnétiques tournent autour de  $H_0$  sur deux cônes de précession, mais selon leur états d’énergie, ils se répartissent en deux groupes : ceux dont les spins sont “parallèles” à  $H_0$ , et ceux qui lui sont antiparallèles (voir Figure 3.1-A). L’aimantation nucléaire totale  $M$  est alors définie à partir de deux composantes : l’aimantation longitudinale  $M_L$  est la projection de  $M$  suivant un axe  $O_z$  parallèle à  $H_0$ , et l’aimantation transversale  $M_T$  correspond à la projection de  $M$  sur le plan  $xy$  perpendiculaire à  $H_0$ . La courbe de la Figure 3.2-B représente la mesure de la magnétisation induite en fonction du temps, décrivant le changement d’état du proton au fur et à mesure qu’il absorbe de l’énergie provenant des mouvements des molécules voisines, appelées généralement le “réseau”. On appelle  $T_1$  le temps de relaxation *spin-réseau*, ou temps de relaxation longitudinal, qui caractérise le temps nécessaire à l’orientation des protons dans le champ  $H_0$ .  $T_1$  n’est mesuré qu’en laboratoire. A l’équilibre, l’orientation quelconque des moments magnétiques sur les cônes de précession annule  $M_T$ , mais pas  $M_L$ . Pour pouvoir détecter l’aimantation longitudinale, il est nécessaire de placer le système hors de sa position d’équilibre.

Pour cela, l’étape suivante consiste à aligner les spins des protons dans une direction perpendiculaire à  $H_0$ . Pour cela, on applique un rayonnement électromagnétique (onde radio)  $H_1$  produit par une bobine. Pour que le transfert d’énergie de l’onde radio aux noyaux d’hydrogène puisse s’accomplir,  $H_1$  doit osciller à une fréquence  $f_L$  appelée *fréquence de Larmor*, telle que son énergie soit égale à la différence entre les deux états d’énergie des noyaux d’hydrogène. On dit alors que le système entre en *résonance*. La vitesse angulaire de précession des spins étant donnée par

$$\omega_0 = \gamma \cdot H_0$$

la fréquence de précession est la fréquence de Larmor, proportionnelle à  $H_1$  et au moment magnétique induit du proton. Elle s'écrit sous la forme

$$f_L = \frac{\omega_0}{2\pi} = 0,76gH_t$$

et s'exprime en Hz.  $g$  est un facteur caractéristique de l'atome considéré. Pour l'hydrogène,  $g$  vaut environ 5. Pour un champ artificiel  $H_0 \approx 550$  Gauss, la fréquence de Larmor est de 2,3 MHz.

Après une impulsion de  $H_1$ , le mouvement de rotation de l'aimantation transversale crée dans l'antenne de mesure un courant va décroître sous la forme d'une rotation dans le plan perpendiculaire à la direction de  $H_0$  (voir Figure 3.1-C). Ce mouvement de précession induit un faible champ magnétique oscillant à la fréquence de Larmor, et détecté par l'outil après amplification. A l'arrêt de  $H_1$ , plusieurs phénomènes vont accélérer le déphasage des spins et provoquer la décroissance exponentielle du signal détecté (voir Section 3.1.2). Ce signal, dont un exemple est donné par la Figure 3.2-C, est caractérisé par son temps de décroissance, constant et appelé  $T_2$ , ou temps de relaxation *spin-spin* (cette relaxation étant due aux interactions magnétiques des spins entre eux), son amplitude en début de précession et sa fréquence, qui est la fréquence de Larmor.  $T_2$  est aussi fonction des inhomogénéités du champ magnétique dans les roches dont l'aimantation est inhomogène.

Le déphasage causé par l'inhomogénéité de  $H_0$  peut être compensé par une nouvelle impulsion de  $H_1$  mais de 180 degrés cette fois (contre 90 degrés pour celle qui sert à aligner les protons). De la même manière, la relaxation des protons dans le champ  $H_0$  produit un signal écho, qui décroît aussi. Ces impulsions à 180 degrés sont appliquées plusieurs fois après la première pulsation à 90 degrés, à intervalle régulier, l'ensemble constituant une séquence de mesure CPMG (du nom de ses inventeurs, Carr, Purcell, Meiboom et Gill (Carr and Purcell, 1954)). A chaque impulsion, l'amplitude de l'écho diminue, le taux de décroissance donnant le temps de relaxation  $T_2$ . Le déroulement d'une mesure RMN classique est donnée Figure 3.1.

### 3.1.2 Paramètres influençant la mesure

La mesure de résonance magnétique nucléaire répond d'abord à la quantité de fluide présent dans les pores de la formation, dont les atomes d'hydrogène sont libres de s'orienter selon le champ magnétique appliqué. En effet, les hydrogènes de l'eau proche des surfaces des grains ("bound water"), ou liés dans le réseau cristallin des minéraux argileux vont être exclus de la mesure car il est difficile (= long) de les orienter selon la direction du champ magnétique et car leur temps de relaxation sera très court donc difficilement mesurable. La mesure de diagraphie par résonance magnétique nucléaire permet de calculer l'*index de fluide libre* (FFI ou "Free Fluid Index") qui est le volume de fluide libre rapporté à un volume de roche, que l'on note  $\phi_f$ .

L'outil CMR de Schlumberger (Combinable Magnetic Resonance Tool) mesure  $\phi_f$ , ainsi que les temps de relaxation  $T_1$  et  $T_2$ , grâce à une antenne orientable réceptrice du signal de précession, qui est placée entre deux aimants permanents produisant  $H_0$ . L'utilisation de cette configuration permet de focaliser la mesure dans une zone d'environ 15 cm de hauteur (cette épaisseur étant similaire à la résolution des autres outils habituels de mesure de paramètres pétrophysiques essentiels comme densité ou porosité).  $\phi_f$  est calculé à partir de l'enveloppe de l'onde de précession (voir Figure 3.2-C). Les temps de relaxation  $T_1$  et  $T_2$  sont gouvernés par trois mécanismes : la relaxation due à la présence des grains, la relaxation par diffusion moléculaire due aux gradients du champ magnétique, et la relaxation dans l'eau libre.

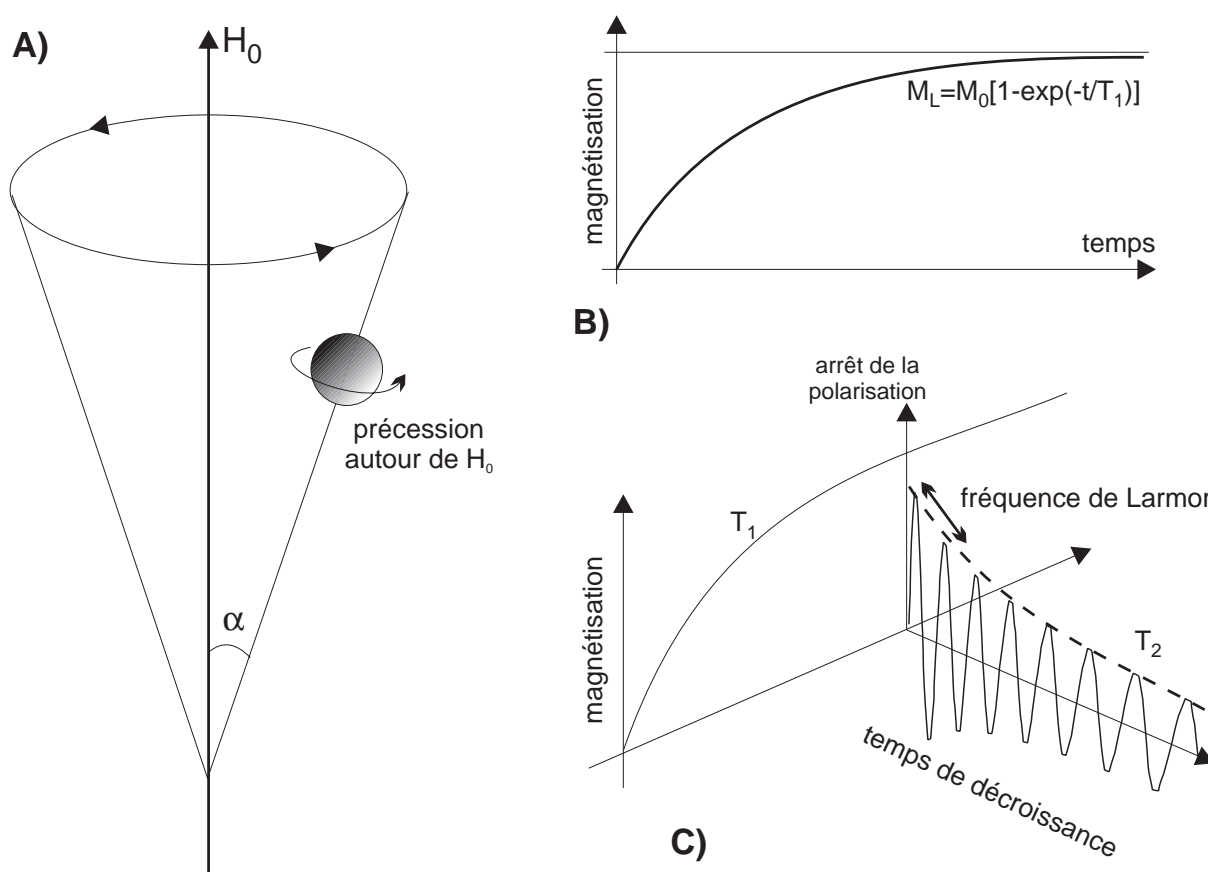


FIG. 3.2 – Principes de la résonance magnétique nucléaire **A**. Lorsque les spins de protons sont alignés selon un champ magnétique  $H_0$ , ils sont affectés d'un mouvement de précession ayant pour axe de rotation l'axe de  $H_0$  et dont l'angle et la vitesse de rotation sont proportionnelle à la puissance de  $H_0$ . **B**. Illustration du temps de relaxation longitudinal (ou spin-réseau)  $T_1$  caractérisant le taux de magnétisation d'une particule placée dans un champ magnétique constant. **C**. Après le temps  $T_1$  qui est nécessaire pour aligner les spins, le champ est coupé et les protons relaxent, produisant un champ magnétique oscillant à la fréquence de Larmor, et donc le signal décroît selon une constante de temps  $T_2$  appelée temps de relaxation transversal (ou spin-spin). d'après Ellis (1987).

**Les grains** Les molécules d'un fluide étant soumises au mouvement brownien, elles diffusent dans l'espace poreux et ricochent sans arrêt sur les surfaces des grains le délimitant. Pendant une mesure de résonance magnétique nucléaire, les noyaux d'hydrogène transfèrent leur énergie de moment cinétique à la surface du grain, ce qui permet aux protons soit de s'aligner selon le champ  $H_0$  produit par l'outil de diagraphie et influencer le temps de relaxation spin-réseau  $T_1$ , soit de se déphaser plus vite lorsque l'on coupe  $H_1$ , contribuant ainsi à la relaxation spin-spin  $T_2$ . Dans la plupart des roches poreuses, la relaxation due à la surface des grains est le paramètre qui a le plus d'influence sur  $T_1$  et  $T_2$ . La capacité des surfaces de grain à relaxer les protons est appelée *relaxivité de surface* et notée  $\rho$ .

La nature des grains détermine l'efficacité de leur surface à relaxer des protons. Par exemple, les grains de quartz d'un grès sont deux à trois fois plus efficaces que les grains d'une roche carbonatée. Une concentration élevée en minéraux magnétisables augmente la valeur de  $\rho$ , et donc baisse les temps de relaxation des protons. La taille des pores est aussi un facteur important, qui s'exprime sous la forme du *rapport surface-volume*, et qui va conditionner la fréquence des

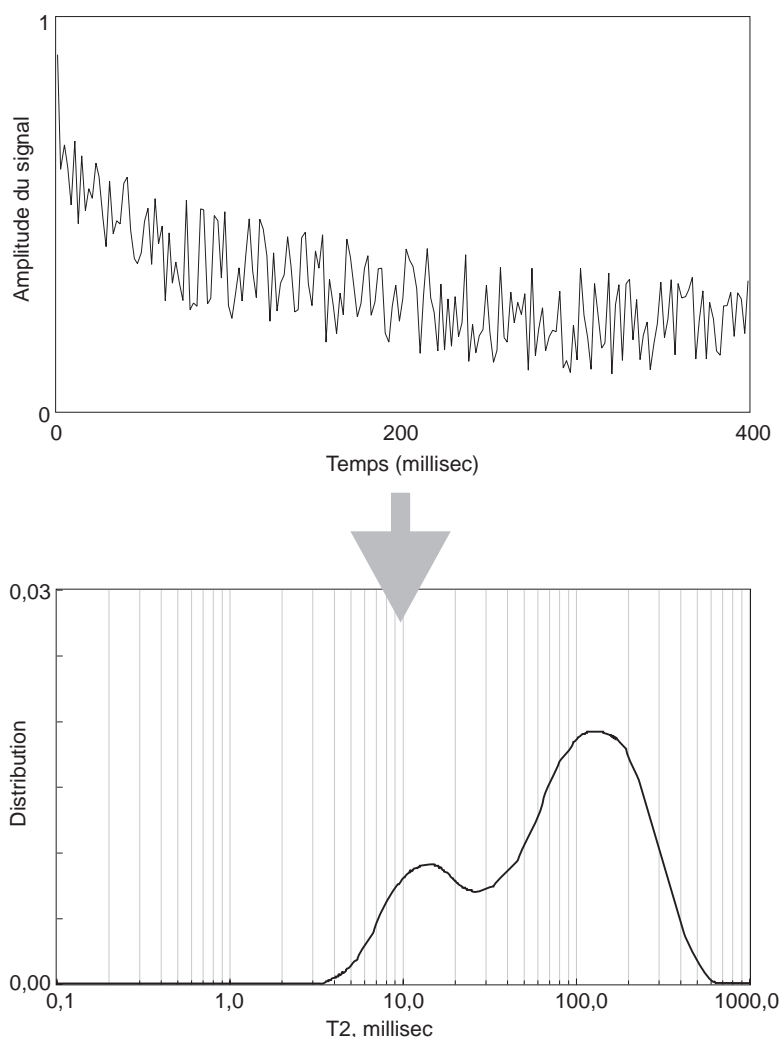


FIG. 3.3 – L'outil CMR mesure l'amplitude du signal de précession au fur et à mesure de la décroissance, qui est la somme de tous les signaux générés par les protons relaxant dans le volume mesuré. Une distribution des amplitudes des signaux est obtenue, qui correspond grossièrement à une distribution des tailles de pores, l'aire sous la courbe étant assimilable à  $\phi_f$ , la porosité du fluide libre et de l'eau liée par capillarité aux surfaces de grains.

collisions des protons avec la surface des grains (voir Figure 3.3) : des pores très grands ont un faible rapport surface–volume, donc les collisions sont peu fréquentes, donc les temps de relaxation sont relativement long, et inversement pour les pores étroits (Sen et al., 1990). Si l'on considère un pore, la magnétisation d'un proton décroît exponentiellement, et l'amplitude du signal en fonction du temps décroît elle aussi selon une constante caractéristique, et l'on écrit pour  $T_2$  par exemple :

$$\frac{1}{T_2} = \rho_2 \cdot \frac{S}{V}$$

On peut écrire la même chose pour  $T_1$ , avec  $\rho_1$  qui est la relaxivité de surface influant sur le temps de relaxation spin-réseau. La formation est composée de pores ayant chacun un rapport surface-volume propre. La magnétisation totale est la somme du signal provenant de l'ensemble des pores, et est donc proportionnelle à la porosité, et la décroissance du signal global reflète la

distribution de la taille des pores. L'amplitude de tous les signaux  $T_2$  générés par les protons relaxant dans le volume poreux est mesurée. Une inversion mathématique (voir Freedman et al. (1997)) distribue les temps  $T_2$  dans des classes pour donner la distribution  $T_2$  (Figure 3.3). Dans les milieux poreux simples (grains équants, un seul type de pore), il est possible de relier directement la distribution  $T_2$  à la distribution de la taille des pores, l'aire sous la courbe étant assimilable à la porosité  $\phi_f$ . Dans les milieux poreux complexes (grains de tailles variées, porosité complexe), comme par exemple les carbonates, on verra que la distribution  $T_2$  doit être analysée plus finement et que d'autres paramètres interviennent et influencent le temps de relaxation  $T_2$  (voir Chapitre 8).

**La diffusion moléculaire** Des gradients dans le champ  $H_0$  peuvent apparaître, dus soit à la configuration de l'aimant dans l'outil, soit au contraste de susceptibilité magnétique entre la matrice et le fluide. Dans ce cas, le mouvement des molécules augmente le déphasage des protons et influence la relaxation  $T_2$ . Les fréquences de résonance des différents éléments de volume de l'échantillon mesuré ne seront pas tout à fait identiques. La décroissance du signal est alors caractérisée par le paramètre  $T_2^*$ , plus petit que  $T_2$ . Ce paramètre  $T_2^*$  tient compte des interactions spin-spin (donnant le temps de relaxation  $T_2$ ), ainsi que d'une composante due aux inhomogénéités du champ  $H_0$ .  $T_1$  n'est pas affecté. La configuration actuelle des outils rend négligeable cet effet, et le temps de relaxation  $T_2^*$  mesuré est donc égal à  $T_2$ .

**La relaxation dans l'eau libre** Lorsque les pores sont grands, c'est à dire qu'aucune surface de grain n'est présente, les protons vont relaxer dans l'eau libre. Ceci se produit par exemple dans le cas des vacuoles de dissolution (aussi connues sous le nom générique de "vugs"), grands pores présents dans les formations carbonatées (voir Chapitre 8).

### 3.1.3 Une mesure de la porosité indépendante de la lithologie

Relier le temps de relaxation transverse  $T_2$  et la distribution  $T_2$  à des paramètres pétrophysiques est relativement aisé dans le cas de roches poreuses ayant un réseau poreux simple (grains équants, un seul type de pore). La distribution  $T_2$  donne une information sur la taille des pores, l'aire sous la courbe de distribution donne  $\phi_f$ , et il est possible de définir des seuils de coupure sur la courbe séparant la porosité correspondant au fluide libre, de celle correspondant à l'eau liée par capillarité aux surfaces des grains (l'eau liée aux argiles ou présente dans la matrice n'est pas vue par l'outil). Il est aussi possible de donner une estimation en continue de la perméabilité de la formation, dérivant néanmoins de relations empiriques développées en laboratoire entre la porosité obtenue par mesure de résonance magnétique et les temps de relaxation  $T_2$ . Par exemple, on utilise communément

$$k_{RMN} = C \cdot \phi_{RMN}^4 \cdot T_{2,log}^2$$

où  $k_{RMN}$  est la perméabilité estimée,  $\phi_{RMN}^4$  est la porosité déduite de la mesure RMN (résonance magnétique nucléaire),  $T_{2,log}^2$  est la moyenne logarithmique de la distribution  $T_2$ , et  $C$  est une constante dépendante de la formation étudiée.

## 3.2 La mesure de l'induction magnétique totale

Je ne fais ici que rappeler les principes de base ainsi que les applications potentielles des mesures magnétiques en forage. A l'origine de ces développements se trouve le projet GHMT, pour *Geological High sensitivity Magnetic Tool*, conjointement mené par Total, le CEA-LETI

de Grenoble, et l'équipe de Magnétisme et Propriétés Physiques des Roches de l'Ecole Normale Supérieure, depuis 1988, auxquels s'est ajoutée en 1992 la société Schlumberger. Pozzi et al. (1988) ont examiné le problème spécifique de la magnétostratigraphie en forage dans les sédiments. Augustin (1990), Lim (1991), et Bouisset and Augustin (1993) ont montré qu'il était possible de mettre en évidence des séquences d'inversion par mesure directe en forage et de les corrélérer aux échelles standards de polarités. Enfin Thibal (1995) a abordé dans une thèse récente le problème de la caractérisation des variations d'intensité du champ magnétique terrestre, ainsi que les applications de la mesure du GHMT pour des études paléoenvironnementales et paléoclimatiques.

### 3.2.1 Principe

Le champ total ou *induction magnétique totale* est mesuré dans le puits par l'intermédiaire d'une sonde NMRS portant un magnétomètre à résonance magnétique nucléaire. Le principe de base de la mesure par RMN est expliqué Section 3.1.1. Ici, au contraire de la mesure classique par RMN, on excite les noyaux dans le but de mesurer le champ magnétique induit dans le capteur par la cohérence de phase des spins. La partie sensible du capteur est un solvant. Les spins nucléaires sont couplés aux spins des électrons libres des radicaux présents dans le solvant durant la mesure par un effet de polarisation électronique dynamique (effet Overhauser), qui amplifie le signal nucléaire. La mesure de l'induction magnétique totale se réduit donc à une mesure très précise d'une fréquence. La sensibilité du magnétomètre développé par le CEA-LETI pour le GHMT est de l'ordre de 0,01 nT, sur une mesure de l'induction total pouvant atteindre 70000 nT.

### 3.2.2 Application

L'outil GHMT est donc constitué d'une sonde de mesure de l'induction totale (NMRS) et d'une sonde de mesure de la susceptibilité magnétique (SUMS), dont le principe de fonctionnement est détaillé Section 2.3.3. L'induction magnétique terrestre totale  $B_T(P, t)$  est fonction du point de mesure  $P$  et de l'instant  $t$ . Elle peut se décomposer dans le puits en plusieurs contributions (Pozzi et al., 1988; Augustin, 1990) :

$$B_T(P, t) = B_0(P) + B_a(P) + B_t(P, t) + B_f(P)$$

où  $B_0(P)$  est la valeur du champ magnétique terrestre, constituant la partie la plus importante de la mesure (compris entre 21000 et 68000 nT selon la position géographique),  $B_a(P)$  représente l'induction liée aux anomalies magnétiques locales dues aux forts contrastes d'aimantation (présence d'hétérogénéités magnétiques dans la formation, présence du tubage, pollution de l'environnement par les opérations de forage/carottage),  $B_t(P, t)$  est dépendant du temps et représente l'induction liée aux variations transitoires du champ magnétique terrestre. La composante  $B_f$ , due à l'aimantation de la formation  $J_f$  mesurée en puits de forage, se décompose en  $B_{fi}$  due à l'aimantation induite  $J_{fi}$ , et  $B_{fr}$  due à l'aimantation rémanente  $J_{fr}$  des sédiments (Pozzi et al., 1988, 1993; Thibal, 1995). La mesure de la susceptibilité magnétique par la sonde SUMS permet d'obtenir l'intensité de  $J_{fi}$ , qui s'exprime comme  $\chi H_t$ , avec  $\chi$ , susceptibilité magnétique, et  $H_t$ , champ magnétique terrestre. Un enregistrement de référence placé en surface (point 0) durant les opérations de diaggraphie permet de déduire de la mesure effectuée en puits les valeurs de  $B_a(P)$ ,  $B_0(P)$  et  $B_t(P, t)$ , tandis que la composante résiduelle  $B_a(P) - B_a(0)$  peut être éliminée par un filtrage ou une modélisation dipolaire. Dans le cas des forages ODP, la couche d'eau salée conductrice filtre les fréquences les plus hautes, mais introduit une composante due aux déplacements des masses d'eau dans le champ magnétique terrestre, qu'il est possible d'éliminer par des mesures simultanées faites à terre (Thibal, 1995).

La mesure du GHMT permet donc à l'aide des mesures de la susceptibilité magnétique et du champ total, de retrouver la polarité magnétique des niveaux traversés (Thibal et al., 1995). D'autres applications font intervenir la susceptibilité magnétique comme variable privilégiée dans la reconstruction paléoclimatique et paléoenvironnementale (Thibal, 1995; Thouveny et al., 1994; deMenocal et al., 1991).

Deuxième partie

Le traitement statistique des données  
de diagraphies





## Chapitre 4

# Qualité des données et préparation à l'analyse

### 4.1 Introduction

La mesure de diagraphie s'effectue *in situ*, dans des conditions souvent difficiles ou au moins peu optimales. Il est difficile de déterminer certaines sources potentielles d'erreur, et encore plus difficile de les quantifier. Il est nécessaire avant toute utilisation des données, dans des modèles d'inversion minérale en particulier, de savoir si l'on est capable d'évaluer l'erreur systématique (due à l'outil), et "analytique" (provenant des effets de l'environnement de mesure, de l'utilisation de standards, de la calibration, du filtrage), et de pouvoir l'introduire dans le traitement des données.

### 4.2 La population et l'échantillon

Un outil de diagraphie mesure un paramètre  $p$  lors de sa remontée dans le puits de forage. Selon l'outil utilisé, la profondeur d'investigation en  $x$  et la résolution verticale en  $z$  varient, et déterminent le volume d'investigation, qui est le volume de roche mesuré et moyenné à chaque niveau de mesure. Le pas de mesure le plus courant est le demi-pied (15,64 cm), mais peut être, comme on l'a vu au chapitre précédent, beaucoup plus réduit dans le cas des outils d'imagerie (0,25 cm). Le tableau A.4 récapitule les valeurs indicatives pour ces paramètres pour certains des outils classiquement utilisés dans le cadre du programme ODP et qui ont servi à acquérir les données traitées dans le cadre de cette thèse.

Le principe de mesure, intégrant habituellement un grand volume de roche (sauf en ce qui concerne les outils d'imagerie), permet d'assimiler chaque niveau de mesure à un échantillon statistique représentatif du volume d'investigation pris en compte, qui représente ici la population. En effet, par exemple dans le cas des outils nucléaires, le signal mesuré obéit à une loi de Poisson. En définitive, chaque niveau de mesure est la moyenne de cet échantillonnage, et l'incertitude sur la mesure pourra être considérée comme la déviation standard.

Si la lithologie varie significativement dans l'intervalle (le plus souvent en  $z$ ), par exemple à cause de la présence de bancs minces, l'outil fait une moyenne plus ou moins pondérée du signal provenant du volume d'investigation. L'échantillon mesuré sera alors représentatif d'une "moyenne" de l'ensemble des lithologies présentes dans le volume d'investigation. Ceci n'est pas trivial, ne serait-ce que par rapport au procédé classique d'évaluation de la précision des mesures de diagraphie par la comparaison avec des données obtenues sur des échantillons prélevés sur

carotte. Le volume mesuré alors en laboratoire ne peut que rarement être comparé aux volumes typiquement mesurés par les outils. Un échantillonnage sur carotte destiné à être comparé à une mesure de diagraphie se devrait d'être du même ordre de grandeur que cette mesure, au moins en ce qui concerne la résolution verticale. Toute comparaison devra se borner à vérifier une tendance générale de l'enregistrement de diagraphie, en passant à côté de l'intérêt particulier de la mesure de diagraphie, c'est à dire son caractère continu.

## 4.3 La qualité des données

### 4.3.1 La mesure de diagraphie

Pour quantifier les incertitudes qui entachent une mesure de logging, il faut déterminer les sources d'erreur. Le cycle d'acquisition des données en diagraphie est donné Figure 4.1. A chaque étape est associée une erreur. Les conditions environnementales affectent en priorité la mesure, et sont plus ou moins difficiles à corriger selon le principe de l'outil utilisé. Comme simple exemple, on peut citer la mesure de densité (HLDT) ou l'imagerie de micro-résistivité (FMS) qui nécessite un bon contact entre l'outil et la paroi du puits, pour minimiser l'effet dû à la boue de forage. Pour cela, l'outil est équipé d'un bras télescopique ou d'un ressort le maintenant collé contre la paroi. Si le puits devient trop large, l'outil peut néanmoins perdre le contact. Dans le cas du HLDT (mesure de la densité par effet Compton), on utilise un second détecteur placé plus près de la source et qui subit plus les effets du "mud-cake" se trouvant entre l'outil et la formation, ou de la boue de forage si l'outil est vraiment trop éloigné. La mesure de ce second détecteur permet de corriger la mesure de la densité. Cette correction est présentée sous la forme d'un paramètre DRHO (voir Section 1.4.1), dont la différence avec zéro indique l'importance de la correction et donc l'incertitude attachée à la mesure. Les effets environnementaux sont aussi corrigés ou au moins pris en compte par une description et une estimation avant les opérations des conditions de mesure. Cette estimation est précisée, durant l'acquisition, par la mesure *in situ* de paramètres comme le diamètre du puits (grâce au caliper du HLDT, ou du FMS), ou la section macroscopique de capture du puits de forage  $\Sigma_{BH}$  (nécessaire dans le traitement des données des outils de spectrométrie gamma). La différence pouvant exister entre les conditions de mesure estimées et réelles est une source importante d'erreur. L'incertitude sur la mesure dépend aussi des paramètres outil. La vitesse, l'accélération (influe sur le calage en profondeur de la mesure), l'excentrement par rapport à l'axe du puits (influe sur le volume effectif de formation pris en compte, variable importante pour des outils comme le GHMT ou le NGT), la configuration, sont enregistrés durant l'acquisition et utilisés pour corriger la mesure. Enfin, la calibration (par rapport à des standards par exemple), le traitement du signal (filtrage, corrections, calibrations) finissent de déterminer les conditions générales de la mesure.

### 4.3.2 La précision

La précision d'une mesure est la valeur de la différence entre cette mesure et la valeur vraie. De part le caractère continu de l'enregistrement de diagraphie, et les méthodes de mesure utilisées, le biais dû à l'échantillonnage n'existe en principe pas. Néanmoins, pour les outils mesurant la résistivité, il peut arriver que malgré la focalisation du courant, une différence de conductivité entre deux bancs provoque à l'approche du banc le plus conducteur un passage préférentiel du courant dans ce banc, le faisant paraître plus épais qu'il ne l'est en réalité. Ce phénomène sera aussi important dans le cas de l'imagerie de micro-résistivité électrique, où le diamètre fini de l'électrode provoque une augmentation artificielle de l'ouverture des fractures ou des

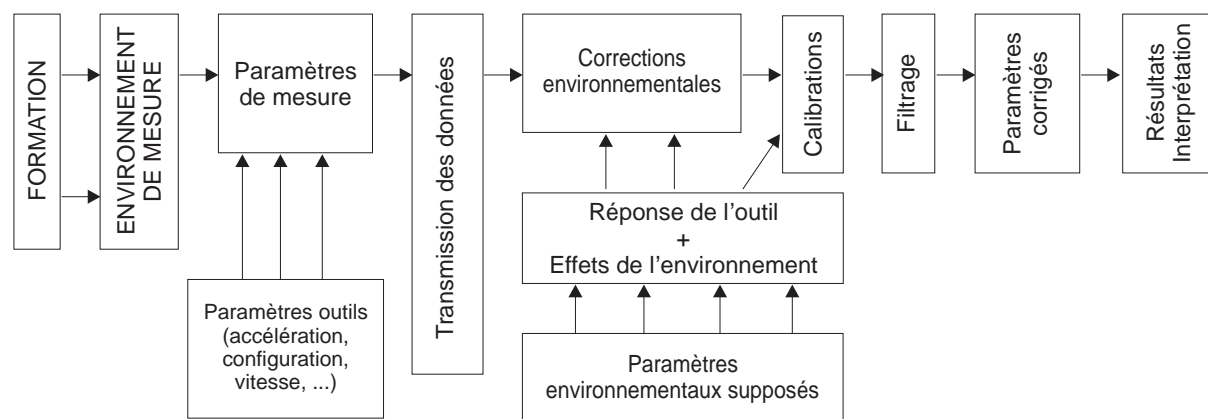


FIG. 4.1 – Cycle d'acquisition des données en diagraphie (adapté de Theys (1991)).

taille des pores de dissolution (voir Figure 4.2). Un biais peut aussi apparaître si on procède sans précaution à un “despiking”, (suppression des valeurs hors-échelle que l’on considère comme étant manifestement mauvaises—trou trop grand, outil défectueux, ...), ou bien lorsqu’on effectue un lissage (filtrage par fenêtre glissante) du signal souvent nécessaire pour séparer le bruit de variations réelles.

Dans le cas de la mesure de diagraphie, la précision dépend beaucoup de la minimisation de l’erreur instrumentale liée à chaque outil. Le biais instrumental peut exister, mais est pratiquement éliminé par une calibration effectuée avant chaque utilisation d’un outil de diagraphie. On utilise par exemple pour l’outil NGT mesurant la radioactivité naturelle gamma un standard de thorium de concentration connue, qui se présente sous la forme d’une gaine que l’on enroule autour du détecteur NaI(Tl). Le spectre obtenu après un temps de mesure fixé pour obtenir une bonne statistique sert à calibrer l’outil en terme de précision du positionnement du pic du thorium sur le spectre de radioactivité gamma, et à évaluer l’incertitude (pertes) liée au cristal et à sa sensibilité.

### 4.3.3 La reproductibilité

La *reproductibilité* d’une mesure sera bonne si de multiples duplicata effectués sur un même échantillon  $\lambda$  donnent des résultats similaires. Dans le cas idéal où un grand nombre de duplicata seraient obtenus, l’histogramme des valeurs obtenu serait une courbe de distribution normale (ou *gaussienne*), et sa variance donnerait une idée de la reproductibilité.

Pour estimer la reproductibilité d’une mesure de diagraphie, la meilleure façon est d’effectuer de multiples passages (Theys, 1994). Habituellement, dû à des contraintes de temps (ou financières pour les compagnies privées), on effectue, pour chaque chaîne d’outils, deux passages sur une partie ou la totalité du puits, le deuxième donnant la reproductibilité de la mesure. C’est encore peu pour vraiment en obtenir une estimation statistique, et on ne saurait trop encourager trois et même quatre passages pour chaque chaîne d’outils.

Un bon exemple de la bonne reproductibilité que l’on peut obtenir avec les outils de diagraphie est donné Figure 4.3. Les deux enregistrements de la susceptibilité magnétique effectués en continu par le GHMT dans ce puits montrent les mêmes variations. Un agrandissement d’une partie de l’intervalle prouve l’excellente répétabilité de cette mesure jusqu’à des variations minimes de l’ordre de la dizaine de ppm uSI.

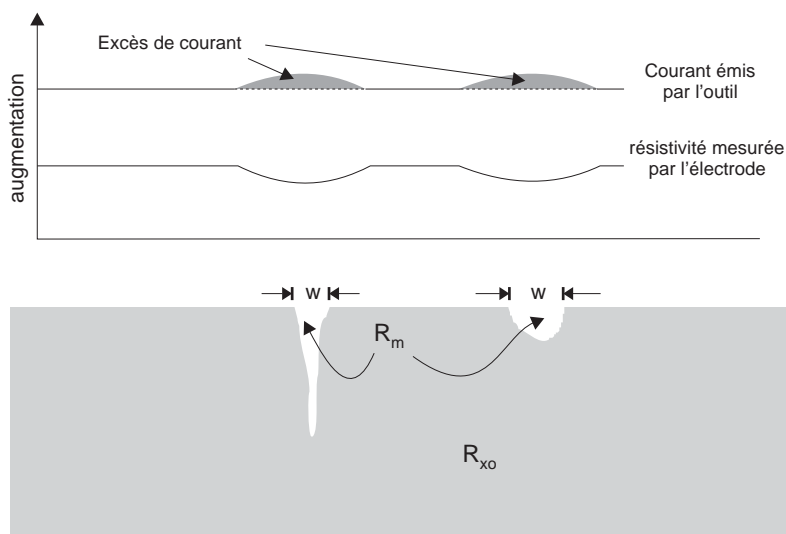


FIG. 4.2 – Illustration du biais de mesure en diagraphie à partir de la mesure de la micro-résistivité par l'outil FMI. Lorsque l'électrode s'approche d'une zone de résistivité significativement plus faible que la matrice (pore de dissolution, fracture), la densité de courant augmente à cause de la présence de cette anomalie, et restera élevée jusqu'à ce que l'électrode soit suffisamment éloignée pour ne plus être affectée. Pour cette raison, les anomalies de résistivité produites par des objets géologiques de faible dimension en donnent une image électrique erronée (d'après Schlumberger).

## 4.4 La représentation des données

### 4.4.1 Introduction

La technique de diagraphie permet, nous l'avons vu, la détermination en continu, de très nombreux paramètres géochimiques, pétrophysiques et même paléomagnétiques des formations recoupées par le forage. La première démarche, classiquement utilisée dans l'industrie du pétrole, consiste à reporter les mesures dans divers diagrammes X-Y standards, entre des paires de variables dont les variations conjointes s'interprètent directement en termes de variations lithologiques, texturales, ou pétrophysiques des formations traversées (voir Section 5.2). On peut aussi vouloir considérer toutes les variables simultanément. Un moyen de représenter les variations d'un grand nombre de variables est de projeter l'information dans un espace de dimension plus réduite. Pour cela, un certain nombre de méthodes statistiques regroupées sous le nom d'*analyse factorielle* ont été développées, qui ont pour objectif de rechercher une structure simple dans un jeu de données multivariées. Les méthodes d'analyse factorielle extraient les valeurs propres et les vecteurs propres d'une matrice carrée, produit de la matrice des données par sa transposée. Elles comprennent deux types de techniques. Celles cherchant les relations existant entre les variables, appelées techniques "*R-mode*", déterminent les valeurs et vecteurs propres à partir d'une matrice de covariance (dans le cas de variables de même ordre de grandeur) ou de corrélation (dans le cas de variables dont les intervalles de distribution sont d'échelles différentes). Ce sont des techniques statistiques. Les secondes, appelées techniques "*Q-mode*", se focalisent sur les similitudes pouvant exister entre les individus, et déterminent les valeurs et vecteurs propres d'une matrice des similitudes entre tous les objets pris deux à deux. Ces dernières ne sont pas considérées comme des techniques statistiques car elles considèrent indirectement le jeu de données comme la population entière disponible.

Il est nécessaire de revenir en détail sur la terminologie utilisée, ainsi que sur sa signification avant de poursuivre avec l'utilité de telles techniques pour traiter notre problème. Les termes français sont accompagnés de leurs équivalents en anglais, la traduction n'étant pas forcément évidente...

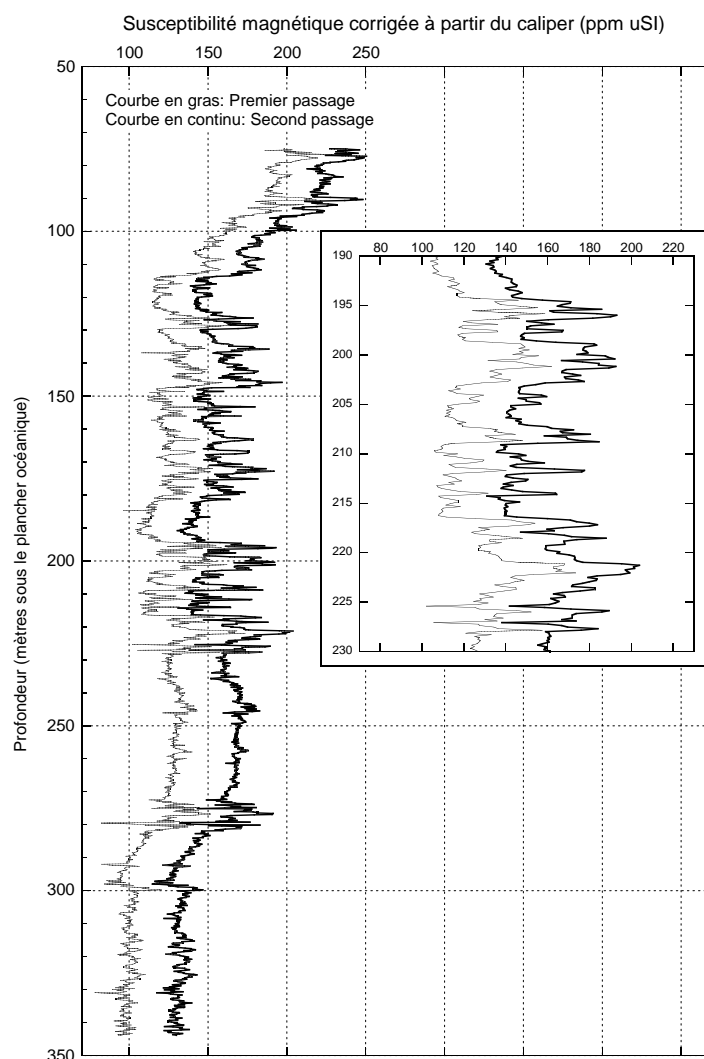


FIG. 4.3 – Mesure de la susceptibilité magnétique (outil GHMT) dans le puits ODP 966F (Leg 160). La courbe en gras correspond à la première passe, la courbe normale à la seconde passe. On remarque une excellente reproductibilité de la mesure, qui se retrouve jusqu'aux plus petites variations locales, de l'ordre de la dizaine de ppm SI. Le détail d'une portion de la courbe à droite confirme cette observation et affirme ces variations comme représentative des hétérogénéités présentes dans le sédiment. Le décalage observé entre le premier et le second passage est dû à la dérive thermique de la sonde de susceptibilité, provoquée par l'échauffement de la bobine émettrice au cours des opérations de diagraphies (Thibal, communication personnelle (1997)).

#### 4.4.2 La détermination des vecteurs propres

Il peut être judicieux avant toute manipulation d'un jeu de données, de standardiser et de centrer chaque variable. Cette transformation permet, pour toutes les variables du jeu de données, de ramener par la *standardisation* (ou réduction) leur variance (et leur écart-type) à 1, et, par

le *centrage*, leur origine à 0. Chaque individu pour la variable  $x$  se transforme selon la relation suivante en un individu de la variable  $z$  centrée-réduite :

$$z_i = \frac{x_i - \bar{x}}{s}$$

où  $\bar{x}$  est la moyenne sur la variable  $X$ , et  $s$  son écart-type. On verra que cette transformation des données, bien que très simple à mettre en oeuvre, permet d'éviter des biais de classification, provoqués par la différence des unités de mesure des variables utilisées dans la classification. Une autre transformation souvent utilisée est la transformation minimum-maximum, chaque individu  $x_i$  devient un individu  $z_i$  sans dimension suivant la relation

$$z_i = \frac{x_i - x_{min}}{x_{max} - x_{min}}$$

Pour trouver les vecteurs et valeurs propres, la première étape consiste à transformer la matrice  $X$  des  $n$  données et  $m$  variables en une matrice carrée symétrique. Pour cela on multiplie  $X$  par sa transposée  $X^T$ . Si l'on effectue  $X^T \cdot X$ , on obtient une matrice  $R$ , de dimension  $m \times m$ , dont les éléments sont les sommes des produits des  $m$  variables, soit

$$r_{jk} = \sum_{i=1}^n x_{ji} x_{ik}$$

où  $j$  et  $k$  sont deux colonnes de la matrice de données. Dans le cas où les variables ont été centrées et réduites, la matrice de corrélation entre les variables devient une matrice des covariances entre les variables. Si par contre on effectue  $X \cdot X^T$ , on obtient une matrice  $Q$ , de dimension  $n \times n$ , contenant les produits pour chaque paire d'individus, sommés pour chaque variables. On considèrera par la suite la matrice  $R$  des relations entre les variables.

Les vecteurs propres et les valeurs propres dans un espace vectoriel  $E$  sont associés de manière simple. Si  $f$  est une application linéaire de  $E$  dans lui-même, alors  $U$  est un *vecteur propre* de  $f$  si  $f(U) = \lambda U$ , où  $\lambda$  est une constante. On dit par ailleurs que  $U$  est associé à la *valeur propre*  $\lambda$ . Ramenons ceci à la matrice  $R$  :  $RU = \lambda U$ , et  $U$  est alors un vecteur propre de la matrice  $R$ , associé à la valeur propre  $\lambda$ . On appelle *spectre* de la matrice  $R$  l'ensemble de ses valeurs propres. A partir de celà, pour toute matrice carrée  $R$ , il existe un polynôme de degré  $m$   $D(\lambda) = 0$ , défini par le déterminant  $\det(R - \lambda I) = 0$ , appelé le *polynôme caractéristique* de  $R$ . Les racines de  $D(\lambda) = 0$  sont les valeurs propres de la matrice  $R$ .

Selon le théorème de Eckart-Young, il est possible de factoriser n'importe quelle matrice carrée  $A$  en lui combinant deux matrices orthogonales et orthonormales  $U$  et  $V$  (telles que  $U^T \cdot U = V^T \cdot V = I$ , avec  $I$ , matrice unité) pour que le produit  $V^T A U$  soit une matrice diagonale  $\Lambda$  dont tous les éléments sont positifs. On a diagonalisé la matrice  $A$ , et les éléments  $\lambda_i$  de la diagonale de la matrice  $\Lambda$  sont les valeurs propres ordonnées de  $A$ , tel que  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ . De même on appelle les vecteurs colonnes  $(u_1, \dots, u_n)$  de  $U$  les vecteurs propres de  $A$ . On note aussi que la somme des valeurs propres  $\lambda_i$  est toujours égale à la somme des valeurs de la diagonale de la matrice  $R$  des variances-covariances, appelée la *trace* de  $R$ . Ceci nous sera utile dans le cas de l'Analyse en Composantes Principales traitée dans la section suivante.

On peut introduire ici la méthode *SVD*, pour *Singular Value Decomposition* (Décomposition en Valeurs Singulières), qui est une généralisation du théorème de Eckart-Young. Elle s'applique à toute matrice  $A$  de dimension  $m \times n$ .  $A$  peut être écrite comme le produit  $U \Sigma V^T$ , où  $U$  est une matrice orthogonale  $m \times m$ ,  $V$  est une matrice orthogonale  $n \times n$ , et  $\Sigma$  est égal à  $\text{diag}(\sigma_1, \sigma_2, \dots, \sigma_p)$ , où  $p = \text{minimum}(m, n)$ , et  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p \geq 0$ . Les valeurs  $\sigma_i$  ( $i = 1, \dots, p$ ) sont appelées les *valeurs singulières* de  $A$ .

### 4.4.3 L'Analyse en Composantes Principales

Albarède (1995) donne une très claire définition du but recherché par l'Analyse en Composantes Principales (ACP) : "... une technique dont le but est d'expliquer la structure covariante d'un jeu de données multivariées par la contraction du jeu de données dans un espace de plus petite dimension dont les nouvelles variables sont indépendantes (non corrélées)". La base du système est orthogonale. On utilise les principes définis plus haut pour trouver les vecteurs propres de la matrice carrée  $R$  résultant de la multiplication de la matrice du jeu de données  $X$  par sa transposée  $X^T$ . Les vecteurs propres sont appelés *composantes principales*, et sont orthogonaux entre eux, leur covariance étant nulle. Géométriquement, cette transformation équivaut à une rotation des axes dans l'espace des  $m$  dimensions (variables d'origine) dans un nouvel espace à  $p$  dimensions (avec  $p = m$ ) dont les axes décrivent la variance maximum des données. L'origine des axes se trouve alors au centre du nuage de points dans l'espace des mesures. Le premier axe principal sera celui qui rendra compte de la variance la plus grande (les valeurs propres étant classées par ordre de valeurs d'écroissantes). La variance du nouveau jeu de données est fonction des valeurs des éléments de  $U$  ("loadings"), aussi une restriction leur est imposée : la somme des carrés des éléments de chaque vecteur propre sera égale à 1, soit par exemple pour le premier vecteur propre

$$u_{11}^2 + u_{12}^2 + u_{13}^2 + u_{14}^2 + \dots + u_{1m}^2 = 1$$

**Calcul du nouveau jeu de données** Pour obtenir les coordonnées des données dans le nouvel espace ("*principal components scores*"  $Y_{ij}$ ), on projette chaque vecteur individu  $(x_1, \dots, x_m)$  de  $X$  (avec  $i = 1, n$ , nombre d'individus, et  $j = 1, m$ , nombre de variables initiales) sur chaque composante principale en le multipliant par les éléments  $(u_j)$  de chaque vecteur propre  $U$  (composante principale). Considérons la projection d'un individu  $x_i$  sur l'axe principal 1 :

$$y_{i1} = \sum_{j=1}^m u_{1j} \cdot x_{ij}$$

ce qui donne d'une manière générale exprimée sous forme matricielle

$$X \cdot U = Y^R$$

où  $Y^R$  est la matrice résultante des données dans le nouvel espace,  $X$  la matrice  $n \times m$  des données de départ, et  $U$  la matrice carrée contenant les composants des vecteurs propres. On peut en donner la forme éclatée plus claire (pour moi...) :

$$\begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,m} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,m} \\ x_{3,1} & x_{3,2} & \cdots & x_{3,m} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,m} \end{bmatrix} \cdot \begin{bmatrix} u_{1,1} & \cdots & u_{1,m} \\ \vdots & \vdots & \vdots \\ u_{m,1} & \cdots & u_{m,m} \end{bmatrix} = \begin{bmatrix} y_{1,1} & y_{1,2} & \cdots & y_{1,m} \\ y_{2,1} & y_{2,2} & \cdots & y_{2,m} \\ y_{3,1} & y_{3,2} & \cdots & y_{3,m} \\ \vdots & \vdots & \vdots & \vdots \\ y_{n,1} & y_{n,2} & \cdots & y_{n,m} \end{bmatrix}$$

**Interprétation des composantes principales** On a vu à la section précédente que la somme des valeurs propres était égale à la trace de la matrice des variances-covariances  $R = X^T \cdot X$  de  $X$ . Les axes principaux contiennent de plus l'ensemble de la variance de  $X$  (c'est le but de l'ACP...). On va pouvoir retrouver la contribution relative de chaque composante principale simplement en divisant sa valeur propre correspondante  $\lambda_i$  par la trace de  $R$  (variance totale).



Bien que pour  $m$  variables d'origine,  $m$  composantes principales soient calculées à fin de l'ACP, il est souvent possible de ne retenir que les 3 ou 4 premières qui à elles seules recèlent la majeure partie de l'information du jeu de données. Combien de composantes principales retenir est subjectif et reste à la discrétion de l'utilisateur. Pour aider à la décision, on peut supprimer les composantes qui n'amènent que peu d'information supplémentaire à celles déjà retenues, où bien se fixer un pourcentage d'information à atteindre, 90% par exemple.

Il faut soulever ici le problème, particulièrement important en géochimie, des données exprimées selon des variables vérifiant une relation de fermeture. Lorsque la relation de fermeture est forcée, on ne considère plus ces variables comme statistiquement indépendantes, et la valeur d'une variable est toujours affectée des variations des autres. Des corrélations négatives ou positives très fortes, qui n'ont aucune signification géologique, vont apparaître entre les variables. L'ACP va évidemment les révéler, et la première composante principale en sera presque toujours uniquement le reflet. Ces corrélations, ne sont toutefois pas linéaires, et il est impossible de dire si les composantes principales suivantes en sont exemptes, et des auteurs (voir Chayes (1971)) ont montré que les deux sources possibles de corrélation (induite et naturelle) contribuaient chacune pour une part à la corrélation finale. Bien qu'il n'y ait pas pour l'instant de moyen vraiment satisfaisant pour tourner la difficulté, (Aitchison, 1984, 1986) a par exemple montré qu'une solution était de transformer les données selon  $x' = \ln(x/c)$ , où  $c$  est une des variables d'origine, utilisée comme dénominateur commun pour toutes les autres. Cette transformation débouche sur un jeu de données dont les variables sont indépendantes entre elles et peuvent alors se prêter à une analyse statistique.

#### 4.4.4 L'analyse factorielle

L'Analyse en Composantes Principales est une transformation linéaire de  $m$  variables originales vers  $m$  nouvelles variables, chacune de ces dernières étant une combinaison linéaire des premières. Elle fait partie des méthodes d'analyse factorielle, mais il est nécessaire de faire la différence entre l'ACP et l'analyse factorielle que je vais appeler *stricto sensu*. Le processus de l'ACP est effectué de telle manière que chaque composante principale successive décrive le plus de la variabilité totale. Il n'y a aucune probabilité ou hypothèse qui découle du principe, car l'ACP n'est pas considérée comme une méthode statistique, mais plutôt géométrique.

L'analyse factorielle *stricto sensu* est quant à elle une procédure statistique car elle s'appuie sur un certain nombre d'hypothèses *a priori* sur le jeu de données initial. Le principe, similaire à celui de l'ACP, est de chercher à représenter la variance totale d'un jeu de données par  $p$  facteurs géologiques, qui caractérisent les orientations préférentielles des données dans l'hyperespace des variables initiales. En ACP,  $p$  est égal au nombre de variables initiales  $m$ , alors que pour l'analyse factorielle,  $p$  est inférieur à  $m$ . Il est alors démontré que la variance dans ces  $p$  directions peut être augmentée par une *rotation* des axes, qui ne sont donc plus colinéaires aux vecteurs propres de la matrice des variances-covariances. On peut effectuer la rotation en gardant l'orthogonalité entre les nouveaux axes, ou bien en éliminant cette contrainte (certains axes principaux deviennent alors corrélés, cela s'expliquant par le fait que de vrais facteurs géologiques de variation peuvent présenter un certain degré de corrélation entre eux). Le problème de cette méthode est son degré important d'interactivité (comparée à l'ACP), le nombre initial de facteurs que la rotation des axes étant laissés au choix de l'utilisateur, ce qui introduit le paramètre subjectif dans l'analyse.

## Chapitre 5

# La classification de données multivariées

### 5.1 Introduction

L'utilisation de méthodes de classification de données multivariées semble évidente lorsque l'on veut débrouiller et ordonner la masse importante d'informations contenue dans un jeu de données de diagraphie provenant d'un simple puits. Le premier problème, partiellement résolu avec l'Analyse en Composantes Principales, dont les principes ont été expliqués à la fin du chapitre précédent, est celui de la représentation graphique dans un espace interprétable par l'homme, c'est à dire à 2 ou 3 dimensions. L'ACP trouve les axes de plus grande variabilité dans un jeu de données. Elle peut donc servir de traitement préalable à une classification, puisqu'elle orthogonalise et concentre l'information, et réduit la dimension de l'espace, rendant plus facile une classification quelle que soit la méthode mise en oeuvre. Pour autant, ceci requiert que le jeu de données soit le reflet d'une population relativement homogène (ellipsoïde dans l'hyperespace des variables). Si l'on considère un puits traversant plusieurs unités lithologiques très différentes (par rapport aux paramètres mesurés par la diagraphie), l'ACP travaille sur la variabilité de l'ensemble, alors que les données sont dispersées selon plusieurs ellipsoïdes et les composantes principales trouvées peuvent ne pas rendre compte tout à fait de la variance maximale des données, ou masquer une variabilité moindre, mais tout aussi importante.

Plusieurs méthodes d'analyse multivariée sont décrites dans les sections suivantes. Toutes permettent de partitionner les données en groupes plus ou moins homogènes selon les affinités pour les variables utilisées dans l'analyse.

### 5.2 Le diagramme X-Y

La méthode la plus simple pour discerner divers pôles de regroupement dans un jeu de données est encore le diagramme X-Y. Il est extensivement utilisé en diagraphie, et les pétroliers, pour des raisons de rapidité de traitement dans le but de séparer des groupes lithologiques simples (calcaires, grès, dolomies), ont établi de nombreux abaques où les données reportées se séparent en groupes plus facilement interprétables. Evidemment, si la lithologie est complexe, ou que les unités ne sont pas clairement définies mais montrent des gradients entre elles, ou que plusieurs subdivisions sont possible à plusieurs échelles, l'opération devient très vite malaisée. Mais cela reste une bonne approche pour commencer à visualiser les données et leur organisation par rapport aux variables mesurées.

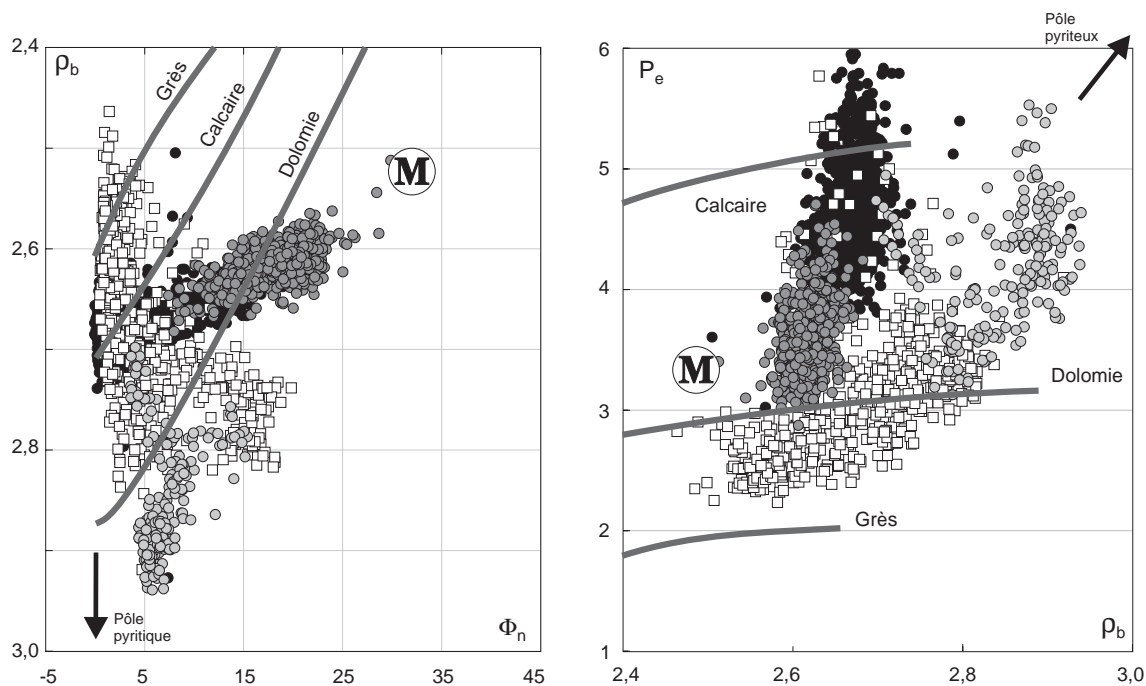


FIG. 5.1 – **Diagramme de gauche** : Densité totale ( $\rho_b$ ) reportée en fonction de la porosité neutron, pour quatre intervalles décrits comme relativement homogènes du forage Balazuc-1, Programme *Géologie Profonde de la France*. **Diagramme de droite** : Mêmes groupes mais cette fois on a porté le facteur photoélectrique ( $P_e$ ) en fonction de  $\rho_b$ . **Légende** : carrés vides : faciès fluviatile, grès+conglomérat+grès dolomitisés+argilite; cercles noirs : calcaires argileux; cercles gris foncé : alternance marnes et calcaires argileux; cercles gris clair : calcaires dolomitisés. Les trois lignes épaisses correspondent à la variation conjointe des deux paramètres reportés  $x$  et  $y$  pour la lithologie indiquée. L'influence de l'argilosité (pôle marneux) est représentée par le **M**. La flèche indique la direction du pôle pyriteux, situé en dehors des diagrammes (les valeurs pour la pyrite sont  $\rho_b \approx 5 \text{ g/cm}^3$ ,  $\phi_n \approx 30 \text{ p.u.}$ ,  $P_e \approx 17 \text{ barns/e}^-$ ).

Figure 5.1 montre un exemple de deux diagrammes classiquement utilisés en diagraphie pétrolière. Sur celui de gauche est reporté la densité totale  $\rho_b$  en fonction de la porosité neutron, sur celui de droite l'effet photoélectrique en fonction de  $\rho_b$ . Les données proviennent de quatre intervalles sélectionnés dans le forage Balazuc-1, réalisé dans le cadre du programme *Géologie Profonde de la France*. La lithologie présente dans chaque intervalle est considérée comme assez homogène et se distribuant entre des pôles comme calcaire, dolomie, grès. Les trois lignes épaisses indiquent l'évolution des deux paramètres reportés pour la lithologie pure concernée. Le **M** représente le pôle marneux (influence de l'argilosité; la position est choisie pour une argile illitique), et la flèche le pôle pyriteux. Les carrés vides correspondent à des grès de faciès fluviatile, localement peu à fortement dolomitisés, et mélangés avec des argilites et des conglomérats (Documents du BRGM, 1990). Dans le diagramme  $P_e = f(\rho_b)$ , ils tombent entre grès et dolomie, montrant l'importance de la dolomitisation. Certains points montrent une porosité neutron importante (jusqu'à 15 p.u.; diagramme  $\rho_b = f(\phi_n)$ ), indiquant probablement la présence d'argiles. Les calcaires (points noirs) sont décrits comme des calcaires argileux et micritiques. Ils tombent bien sur la ligne du calcaire dans les deux diagrammes, mais avec une tendance à se déplacer

vers la ligne dolomitique, montrant une dolomitisation occasionnelle. La densité quasi-constante de ce groupe et l'augmentation de son  $P_e$  au delà de la valeur de la calcite ( $P_e = 5.1$ ) peut être due au remplacement de la dolomite par de l'ankérite ( $\text{Ca}(\text{MgFe})(\text{CO}_3)_2$ ), ou par la présence de pyrite, ces deux minéraux étant indiqués dans l'analyse minéralogique ponctuelle effectuée sur carottes (Documents du BRGM, 1990). Les cercles gris clair, représentant les calcaires dolomités, montrent un profil intéressant en 'V', signant un mélange entre les trois pôles calcaire, dolomie, et pyrite. Les cercles gris foncé sont décrits comme des alternances marno-calcaire et calcaire argileux. L'influence du pôle marneux est évidente sur le diagramme  $\rho_b = f(\phi_n)$ , où les points de ce groupe traversent la ligne de la dolomie, mais aussi sur le diagramme  $P_e = f(\rho_b)$ , de part la densité plutôt faible du groupe ( $P_e=2,6$  à  $2,65$ , alors que la densité de la calcite est  $2,71$  et celle de la dolomie  $2,88$ ).

Un autre exemple de diagramme pouvant servir à l'identification qualitative des minéraux argileux présents dans un sédiments est le diagramme  $Th = f\left(\frac{[\text{Al}_2\text{O}_3]}{[\text{Al}_2\text{O}_3]+[\text{SiO}_2]}\right)$ . La figure 5.2 montre des données acquises sur les carottes du puits 948C, foré à travers la zone de décollement du prisme de la Barbade durant le Leg 156.  $\text{Al}_2\text{O}_3$  a été mesuré par fluorescence X, et la concentration en thorium par la méthode ICP-MS. Les cercles pleins correspondent à des échantillons prélevés dans l'unité lithologique sus-jacente au décollement, constituée des sédiments accumulés dans le prisme d'accrétion, et provenant pour partie du démantèlement de l'arc volcanique des Petites Antilles. Ils tombent plutôt dans le champ des smectite-montmorillonite. Ils montrent une meilleure homogénéité que les échantillons de sédiments prélevés dans l'unité sous-jacente au décollement et recouvrant la plaque subduite (cercles vides), qui se placent plutôt dans le champ des illites-muscovites. La dispersion des points pour ce groupe semble démontrer une proportion d'argile dans la roche totale plus variable.

## 5.3 Les méthodes de classification

La classification des données ("*clustering*") est une opération qui permet de partitionner un jeu de données en plusieurs groupes plus ou moins homogènes réunissant les données qui sont représentées par des propriétés (dans notre cas géologiques, pétrophysiques, ou géochimiques) similaires dans l'hyperespace défini par les grandeurs mesurées. Il existe de nombreuses méthodes utilisant divers algorithmes de classification, et il n'est pas obligatoire qu'elles mènent au même résultat à partir du même jeu de données. J'ai plutôt utilisé dans mes traitements de données la **classification itérative non-hiérarchique**. Je vais détailler cette méthode de classification, en la comparant avec l'analyse hiérarchique plus couramment utilisée.

### 5.3.1 La classification hiérarchique

Dans l'espace des variables d'entrée, la **classification hiérarchique**, par le biais d'une matrice de ressemblance (similarité), joint deux à deux les données les plus similaires entre elles, formant des regroupements composés de paires d'individus. La matrice de ressemblance est ensuite recalculée en prenant chaque groupe comme une donnée, et en calculant les similarités (liens) entre chaque groupe. Ceci forme d'autres groupes. Ce processus est répété jusqu'à ce que la matrice de ressemblance soit  $2 \times 2$ . Ce processus permet de construire un *dendrogramme* (voir Figure 5.3).

Comment construit-on la matrice de ressemblance entre chaque paire d'échantillons, chacun étant décrit par plusieurs variables? Il y a deux façon de mesurer la ressemblance entre deux objets :

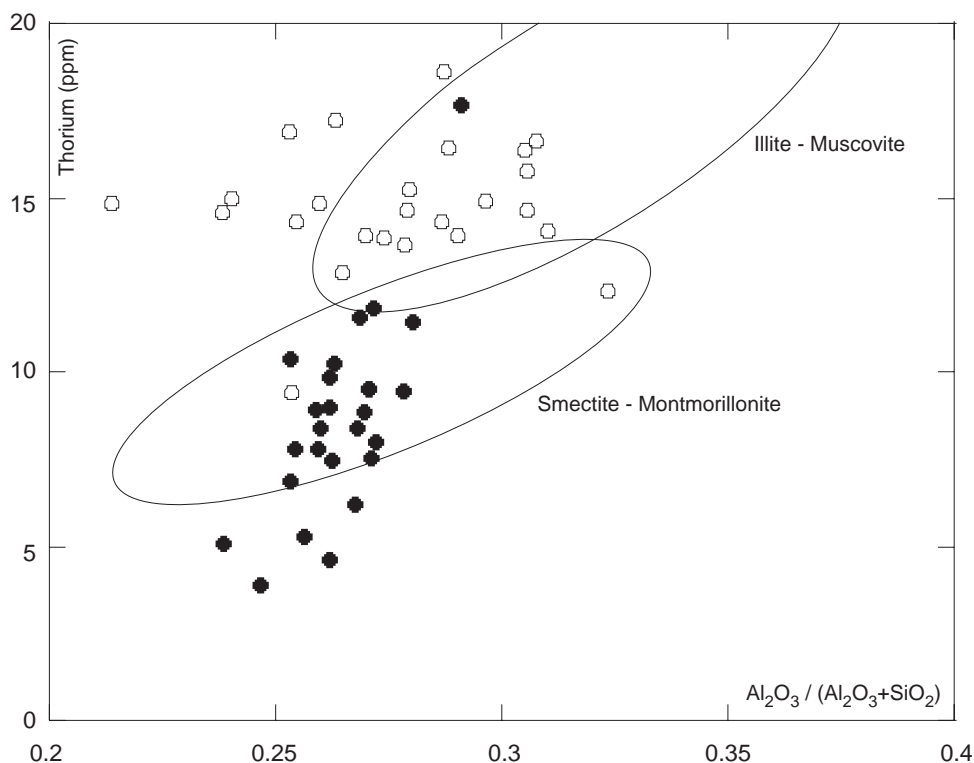


FIG. 5.2 – Détermination du type de minéral argileux prédominant en fonction de la concentration en thorium et aluminium. Les données ont été mesurées par fluorescence X et ICP-MS sur des échantillons du puits 948C foré à travers la zone de décollement du prisme de la Barbade. Cercles pleins : prisme, zone au-dessus du décollement. Cercles vides : section subduite.

- par la *distance*  $d_{AB}$  entre deux individus dans l'hyperplan. C'est le moyen que j'ai utilisé au cours de ce travail dans toutes les études faisant appel à la classification par groupe. C'est aussi le paramètre le plus simple à concevoir et à calculer. Nous standardisons toujours les variables avant la classification, pour éviter les disparités d'unités (par exemple une variable exprimée en ppm pourrait dominer les distances entre les points si les autres variables étaient exprimées en pourcentages). On peut calculer la distance entre deux points dans l'espace de deux manières différentes : d'abord la distance *euclydienne*, qui est la plus directe et se présente comme une ligne droite tracée entre les deux points. Dans le cas d'un hyperespace, la distance entre deux points  $A$  et  $B$  est

$$d_{AB} = \sqrt{\sum_{i=1}^m (x_{iA} - x_{iB})^2}$$

ou bien, si l'on ne veut pas que cette mesure augmente avec l'augmentation du nombre de variables :

$$d_{AB} = \sqrt{\frac{\sum_{i=1}^m (x_{iA} - x_{iB})^2}{m}}$$

Nous avons choisi cette distance dans nos analyses, par souci de simplicité.

- par le *coefficient de corrélation*  $r_{AB}$  entre chaque paire d'individus dans l'hyperespace. Le coefficient de corrélation est utilisé à l'inverse d'une régression par exemple : ici les axes

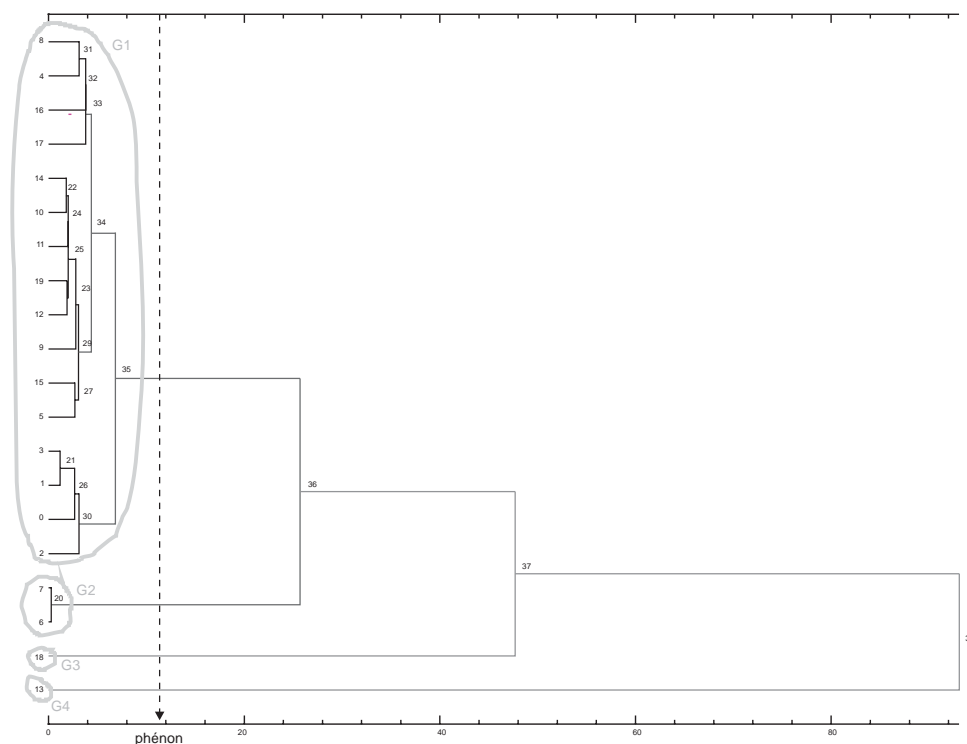


FIG. 5.3 – Exemple de dendrogramme obtenu par analyse hiérarchique par groupe. L'axe gradué au bas du diagramme montre la distance entre chaque groupe/individu. Les numéros à l'extrême gauche sont les numéros de chaque individu dans le jeu de données. Les numéros intermédiaires donnent l'ordre des connexions entre groupes/individus effectuées par l'algorithme. La distance utilisée est euclidienne, et la méthode de groupement est de choisir le meilleur coefficient de ressemblance entre un individu et le barycentre d'un groupe. La ligne pointillée verticale est le phénon ("phenon line") qui, placé à une certaine valeur du coefficient de similarité, coupe le dendrogramme et permet d'isoler des groupes plus ou moins homogènes (cercles gris clair).

sont les individus, et les points sont les variables. Le coefficient de corrélation est défini par

$$r_{AB} = \frac{\sum_{i=1}^m (x_{iA} - \bar{x}_A)(x_{iB} - \bar{x}_B)}{s_A s_B (n - 1)}$$

où les moyennes ( $\bar{x}_i$ ) et écart-types ( $s_i$ ) sont calculés pour toutes les variables. Plus  $r_{AB}$  sera proche de 1, plus les échantillons  $A$  et  $B$  seront semblables.

Maintenant que la matrice de ressemblance est construite, il reste à définir un critère à mettre en oeuvre pour grouper les individus deux à deux. Plusieurs méthodes existent, assez similaires dans leurs principes mais pouvant donner des résultats suffisamment différents pour devoir être détaillées.

Les premiers individus à être reliés sont simplement ceux qui ont le plus fort coefficient de corrélation. Le jeu de données comprend après cette étape des groupes de deux individus, et des individus seuls dont le coefficient de corrélation avec d'autres individus n'était pas assez fort pour permettre le lien. Il faut maintenant lier ces individus isolés aux groupes déjà formés. Pour cela, plusieurs méthodes sont utilisées :

- le regroupement par **plus proche voisin** ("Nearest-neighbour linkage" or "single linkage"). Le coefficient de ressemblance entre un individu isolé et un groupe est égal à celui existant entre cet individu et le *plus* ressemblant dans le groupe ( $d_{AB}$  minimum, ou  $r_{AB}$  le

plus proche de 1). Cette méthode, bien que simple et évidente dans son principe, favorise évidemment une structure hiérarchique peu organisée. En effet, elle favorise le lien d'un individu à une extrémité d'un groupe, car plus similaire à l'individu composant cette extrémité, et donc élimine la possibilité de former d'autres groupes lorsque les deux ou trois premiers groupes de départ ont été formés.

- le regroupement par **plus lointain voisin** (“*Furthest-neighbour linkage*” or “complete linkage”). C’est l’opposé de la méthode précédente : Le coefficient de ressemblance entre un individu isolé et un groupe est égal à celui existant entre cet individu et le *moins* ressemblant dans le groupe. Comme les distances entre les groupes sont augmentées artificiellement de cette façon, cette méthode donne une structure hiérarchique beaucoup plus organisée et des groupes très homogènes.
- on peut aussi choisir la solution intermédiaire, c’est à dire calculer le coefficient de ressemblance entre un individu et un groupe comme étant celui entre cet individu et la *moyenne* des coefficients de ressemblance entre cet individu et tous les individus composant le groupe. En clair, c’est la ressemblance entre l’individu à classer et le *barycentre* (“*centroid*”) du groupe. Cette méthode produit une structure hiérarchique organisée et est un bon compromis entre les deux extrêmes précédemment cités. Dans le cas où l’on doit calculer le coefficient de ressemblance entre deux groupes composés d’un nombre différent d’individus, il est possible de pondérer ce calcul en tenant compte de l’importance relative d’un des groupes par rapport à l’autre.

Lorsque le dendrogramme est complet, le positionnement arbitraire d’une ligne de “*phénon*” (“*phenon line*”) à une valeur du coefficient de ressemblance permet d’isoler des groupes (voir Figure 5.3). Le choix d’une certaine valeur de ressemblance peut être totalement arbitraire, ou basé sur l’allure du dendrogramme (certains groupes peuvent être très bien isolés par rapport au lien suivant), ou bien sur des connaissances *a priori* sur le jeu de données.

### 5.3.2 La classification itérative non-hiérarchique ou classification en nuées dynamiques

En diagraphie, le nombre  $n$  de données à classer est souvent très élevé (plusieurs milliers à dizaines de milliers). La classification hiérarchique oblige à manipuler une imposante matrice des similarités entre les individus, de dimension  $n \times n$ . La procédure appelée “*k-means*”, développée par McQueen (1967), permet une **classification itérative non-hiérarchique** des individus en différentes entités plus ou moins homogènes, les nuées ou groupes. Le principe consiste à choisir  $k$  points dans le jeu de données, qui sont les germes d’agrégation de départ (“*seeds*”). Ces points sont choisis au hasard (lire Ball and Hall (1967) pour une discussion sur les différentes façons de choisir les points d’agrégation de départ) ou positionnés par l’utilisateur selon des critères externes. Une matrice des ressemblances entre les  $k$  barycentres et les  $n$  individus est calculée, et un premier groupement est effectué entre les données les plus proches (similaires) et le barycentre le plus proche. On obtient  $k$  groupes, dont on recalcule le barycentre après accumulation des individus. On reforme alors la matrice  $k \times n$  des similarités entre les nouveaux  $k$  barycentres et les  $n$  individus, et on ré-affecte les données les plus proches au barycentre le plus proche. Le processus est réitéré jusqu’à ce que plus aucun changement important n’intervienne dans la composition des groupes. Les avantages de cette méthode sont sa convergence plus rapide vers une solution (la matrice est de dimension  $k \times n$  contre  $n \times n$  pour les méthodes hiérarchiques), et le fait que les barycentres des groupes individualisés ne sont pas fixes et évoluent en fonction de la géométrie du jeu de données dans l’hyperplan des variables. Ceci permet d’isoler des groupes plus homogènes tout en utilisant des critères de groupement simples (distance euclidienne).

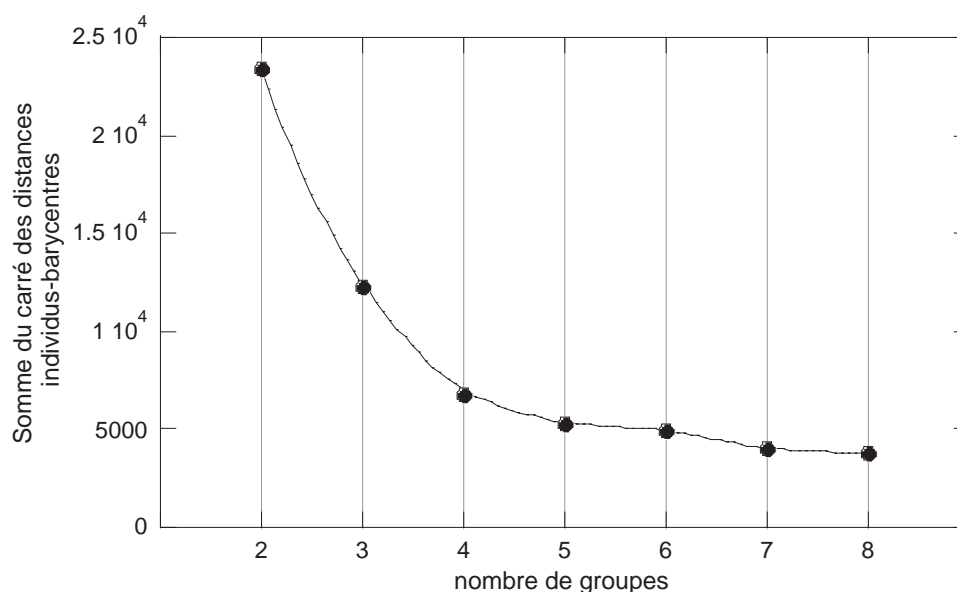


FIG. 5.4 – Méthode utilisée pour un choix optimum du nombre de barycentres servant au partitionnement d'un jeu de données. Nous utilisons un pseudo-test de Fisher, où le critère est la somme  $\Sigma$ , cumulée pour tous les groupes, des distances entre chaque barycentre et les individus du groupe.  $\Sigma$  est portée en fonction du nombre de groupes. Le nombre optimum est celui indiqué au point d'inflexion de la courbe : moins de groupes ne décriraient pas suffisamment la structure multivariée du jeu de données, alors qu'un nombre supérieur compliqueraient inutilement l'analyse.

La faiblesse potentielle reste le choix du *nombre initial* de barycentres de départ. On désire la plupart du temps éviter l'aide d'une quelconque information *a priori* sur le jeu de données, qui peut représenter un biais potentiel à une classification basée sur de purs critères mathématiques. Une méthode intéressante pour choisir le nombre optimal de barycentres est illustrée Figure 5.4. Elle consiste à effectuer plusieurs classifications complètes successives, en incrémentant à chaque fois le nombre de groupes, de façon à encadrer le nombre optimum de groupes décrivant au mieux la structure du jeu de données. On calcule ensuite, pour chaque nombre de groupes, la somme  $\Sigma$  des coefficients de similarités entre les individus et le barycentre qui leur est le plus proche. Par exemple si le critère de ressemblance entre un individu, décrit par  $m$  variables, et le barycentre du groupe auquel il est associé est la distance euclidienne les séparant

$$\Sigma = \sum_{h=1}^k \sum_{j=1}^n \sqrt{\sum_{i=1}^m (x_{i\lambda_j} - x_{i\Lambda_h})^2}$$

où  $h$  est le numéro du barycentre (variant de 1 à  $k$ ),  $j$ , le nombre d'individus,  $\lambda$ , un individu, et  $\Lambda$ , le barycentre du groupe auquel il est associé.

### 5.3.3 Introduction des probabilités d'appartenance dans le classement en nuées dynamiques

La classification itérative en nuées dynamiques (*k-means*) est très pratique pour traiter des systèmes simples, c'est à dire dont les données définissent plusieurs ( $k$ ) groupes bien individualisés. En cela, la classification satisfait trois conditions



1. tout groupe  $i$  ( $2 \leq i \leq k$ ) contient au moins un échantillon :  $X_i \neq \emptyset$ , où  $X_i$  est le groupe  $i$  formé par l'accumulation d'individus de  $X$ , le jeu de données original.
2. un individu appartenant au groupe  $i_1$  ne peut appartenir au groupe  $i_2$  :  $X_{i_1} \cap X_{i_2} = \emptyset$
3. la réunion de tous les individus de tous les groupes forme exactement le jeu de données de départ :  $\bigcup_{i=1}^k X_i = X$

En géologie, où la présence de faciès résultant d'un mélange entre deux pôles extrêmes est fréquente, il n'est pas rare de se trouver confronté à un jeu de données décrivant plutôt des tendances, sous formes de nuages de points allongés et déformés dans l'hyperespace. Il est alors assez difficile de définir des groupes pour lesquels le critère d'appartenance d'un individu serait purement binaire : on en fait partie, ou pas. Le cas d'individus "inclassables", car à la limite entre deux classes, ou étant des données aberrantes, est aussi problématique. Ces individus vont être forcément classés dans un des groupes (conditions 1 et 3), modifiant alors sa géométrie, ainsi que la localisation de son barycentre. Bezdek et al. (1984) ont développé un algorithme de classification basé sur la *logique floue*, dont le principe consiste à donner une probabilité d'appartenance d'un individu à chacun des groupes. Cette probabilité est comprise entre les deux cas extrêmes,  $< 1 >$ , pour appartenance exclusive, ou  $< 0 >$ , pour non appartenance exclusive, qui nous ramène à la classification en nuées "classique". Une matrice  $U$ , de dimension  $k \times n$ , est déterminée. Elle est la représentation d'un groupe  $X_i$  de  $X$  lorsque

$$u_i(x_j) = u_{ij} = \begin{cases} 1 & \text{si } x_j \in X_i \\ 0 & \text{sinon} \end{cases}$$

$U$  est la matrice qui contient pour chaque individu/vecteur  $x_j$  ( $j = 1, n$ ) de  $X$ ,  $k$  nombres compris entre 1 et 0 et qui sont les probabilités d'appartenance des individus  $x_j$  aux  $k$  groupes "flous"  $u_i$  de  $X$ .

Le critère utilisé par Bezdek et al. (1984) pour partitionner au mieux un jeu de données en utilisant la logique floue est une solution des moindres carrés pondérés. Elle s'écrit sous la forme

$$J_w(U, \Lambda) = \sum_{j=1}^m \sum_{i=1}^k (u_{ij})^w \|x_j - \Lambda_i\|_A^2$$

$w$  est le coefficient de pondération,  $\Lambda_i$  le barycentre du groupe  $i$ , et  $\Lambda$  le vecteur contenant tous les barycentres des groupes.  $A$  est une matrice des pondérations sur le calcul de la distance entre un individu et un barycentre, contrôlant donc la forme des groupes finaux dans l'hyperplan des  $m$  variables. Nous avons pris  $A$  comme étant la matrice identité. Le carré de la distance entre un individu  $x_j$  et un barycentre  $\Lambda_i$  est donc

$$d_{ij}^2 = \|x_j - \Lambda_i\|_A^2 = (x_j - \Lambda_i)^T A (x_j - \Lambda_i)$$

$(u_{ij})^w$  est le poids sur chaque  $d^2$ , correspondant à la probabilité d'appartenance d'un individu  $x_j$  au groupe  $i$  ( $u_{ij}$ ), élevé à la puissance  $w$ .  $w$  donne de bons résultats lorsqu'il est pris entre 1,5 et 3 (Bezdek et al., 1984). Le but auquel doit tendre l'algorithme est  $J_w$  minimum, ce qui se traduit par trouver de façon itérative les paires  $(\hat{U}, \hat{\Lambda})$  qui minimisent  $J_w$ .

Pour contrôler la présence de minima locaux, qui permettent la convergence, mais dont on ne sait pas s'ils décrivent au mieux le jeu de données, deux paramètres sont calculés pour chaque matrice  $\hat{U}$  produite à chaque itération, le *coefficient de partitionnement*  $F_k$ , et l'*entropie du système*  $H_k$ , représentant son état de désordre vis-à-vis de la classification, et exprimés comme

$$F_k(\hat{U}) = \sum_{j=1}^m \sum_{i=1}^k \frac{(\hat{u}_{ij})^2}{m}$$

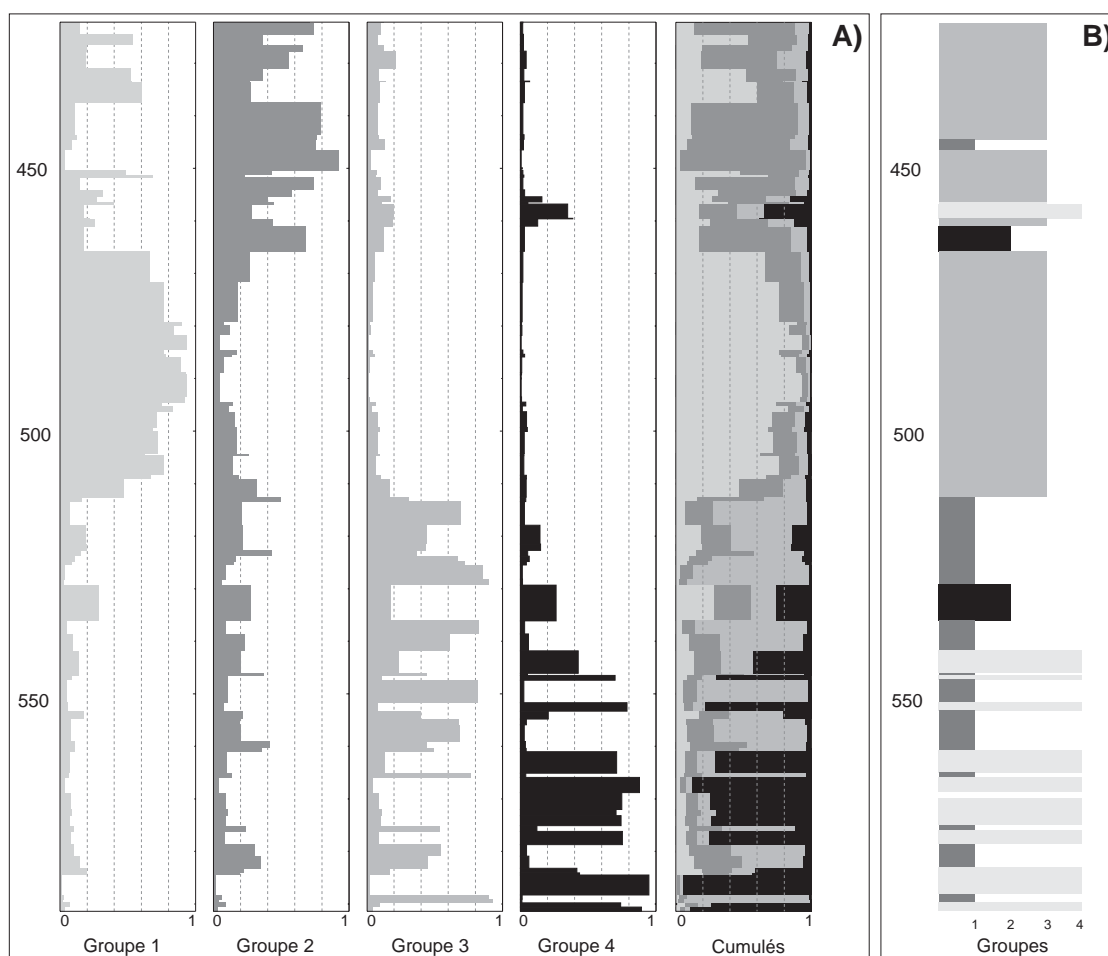


FIG. 5.5 – Comparaison entre un traitement en nuées dynamiques “floues” (A) et la classification en groupes “classique” (B). La première donne des informations sur la compositions des groupes et leur forme dans l’espace, donc leur homogénéité et la manière dont s’opèrent les transitions entre eux. Les données sont des pourcentages d’oxydes d’éléments majeurs mesurés par fluorescence X sur 82 échantillons provenant du puits 948C, carotté durant le Leg ODP 156.

$$H_k(\hat{U}) = - \sum_{j=1}^m \sum_{i=1}^k \frac{\hat{u}_{ij} \cdot \ln(\hat{u}_{ij})}{m}$$

Pour une classification en nuées dynamiques “classique”,  $F_k = 1$  et  $H_k = 0$ . Dans le cas de nuées “floues”, la validité de la classification finale est soumise aux conditions suivantes :

$$\frac{1}{k} \leq F_k \leq 1 \quad 0 \leq H_k \leq \ln(k)$$

Sur la figure 5.5-A., on a porté le résultat de la classification “floue”, effectuée pour  $k = 4$  groupes, sur les données géochimiques (majeurs) obtenues par fluorescence X sur des échantillons prélevés dans le puits 948C foré durant le Leg ODP 156. La norme utilisée est la norme euclidienne, et  $w$  est égal à 2. L’algorithme a convergé en 39 itérations. L’entropie de la classification finale  $H_k$  est égale à 0,87 alors que  $F_k$  est de 0,53, valeurs qui valident la classification comme étant une bonne approximation de l’organisation des individus dans l’espace des  $m$  variables.

Le cadre **B.** montre le résultat de la classification en nuées dynamiques “classique”. Si, d’une façon générale, ces deux résultats sont similaires, la classification floue apporte une information supplémentaire importante sur la dynamique des groupes dans l’espace. Les groupes définis par la classification classique (**B.**) seront dénotés  $G_i$ , et les nuées floues (**A.**)  $\hat{G}_i$ . Une première constatation concerne  $G_3$ , qui, s’il est relativement homogène de 470 à 515 mètres, l’est en fait très peu dans la partie supérieure. On voit que ses individus sont en fait assez partagés entre le  $\hat{G}_1$ ,  $\hat{G}_2$ , et même  $\hat{G}_3$ . On remarque les transitions dans les parties supérieures et inférieure de la partie homogène (470 à 515 mètres), où l’on voit l’apparition d’une probabilité d’appartenance à  $\hat{G}_2$ . Dans la partie inférieure de la colonne, même si l’on a une succession d’unités de faible épaisseur, ces unités semblent relativement homogènes (probabilité très élevées d’appartenance à  $\hat{G}_4$ ). Une autre observation intéressante est relative aux deux échantillons composant  $G_2$ . La classification floue trouve une quasi équiprobabilité d’appartenance aux 4 groupes  $\hat{G}_i$ . Cette situation se retrouve pour l’échantillon isolé appartenant à  $G_4$  vers 455 mètres. On verra dans une étude (Rabaute et al. soumis, voir Section 7.2.1, page 125) que ces échantillons seront très difficiles à caractériser par une inversion minérale, et la classification floue révèle déjà ce cas particulier.

## Chapitre 6

# Donner un sens aux classes individualisées

### 6.1 Introduction

Dans les chapitres 1 à 5, nous avons passé en revue toutes les étapes depuis l'acquisition et la nature de la mesure, jusqu'aux méthodes de représentation et de transformation permettant d'appréhender non seulement tous les paramètres simultanément, mais également les relations existant entre eux. Il s'agit maintenant d'utiliser toutes ces informations pour retrouver les propriétés géochimiques, pétrophysiques, hydrodynamiques, tectoniques, ou texturales des formations géologiques traversées par le forage. Pour cela, diverses méthodes plus ou moins compliquées à mettre en oeuvre existent.

### 6.2 Les informations directes

Comme première approche, les informations directes données par l'Analyse en Composantes Principales, ou la Classification Itérative Non-hiérarchique (classique ou "floue") sont déjà très utiles pour lier les groupes à une minéralogie, un lithofaciès, ou une lithologie. Les deux exemples suivants montrent une première analyse de ces informations directes.

#### 6.2.1 Les valeurs et vecteurs propres de l'ACP

La figure 6.1 donne les résultats de l'Analyse en Composantes principales effectuée sur un jeu de données d'exemple provenant du forage BALAZUC-1 (Programme *Géologie Profonde de la France*), entre 1500 et 1600 mètres de profondeur. En **A.**, les variables sont reportées dans le cercle de corrélation dans le plan défini par les deux premières composantes principales (qui portent à elles seules 94% de la variance du jeu de données initial; voir Tableau 6.1). On remarque la corrélation entre Ca, S,  $P_e$  et  $\rho_b$ , et leur anti-corrélation avec Si. La variable Fe, qui est presque orthogonal à Ca et Si, donc qui leur est non corrélée, trace un comportement géochimique différent (qui est sans doute caractéristique d'un faciès lithologique particulier). La partie **B.** de la figure 6.1 montre les individus projetés dans le système des nouveaux axes (vecteurs propres). On remarque l'individualisation de trois populations distinctes, que les corrélations variables-facteurs permettent d'associer à une lithologie : un premier groupe *G1* caractérisé exclusivement par le facteur 1, et dont les fortes contributions de  $P_e$  et  $\rho_b$ , ainsi que de Ca et S, permettent d'associer à de l'anhydrite. Par contre, *G2* montre une forte composante F1 mais aussi F2. C'est

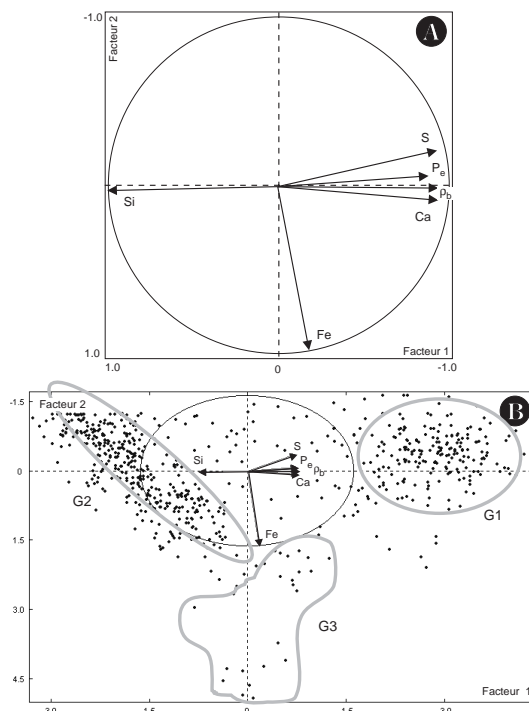


FIG. 6.1 – **A.** Cercle de corrélation entre les variables, portées en fonction des deux premières composantes principales. **B.** même diagramme, mais avec les individus recalculés (“scores”) pour les nouveaux axes (vecteurs propres), montrant trois populations distinctes. Les matrices de corrélation entre variables ainsi que les valeurs propres sont présentées Tableau 6.2.

une formation gréseuse, caractérisée par une teneur en argile variable (montrée par la composante F2 importante, qui porte la variance du fer).

### 6.2.2 La composition des barycentres

Reprenons nos données (oxydes d’éléments majeurs) de fluorescence X obtenues sur les échantillons des carottes du puits 948C (Leg ODP 156). La Figure 6.2 montre la classification obtenue à gauche. Les coordonnées dans l’hyperespace des variables des barycentres des 4 groupes indi-

Histogramme des 6 CP				Matrice des corrélations						
Numéro	Valeur propre	%	Cumulé	Si	Ca	Fe	S	$P_e$	$\rho_b$	
1	4,64	77,45	77,45	Si	1,00					
2	1,00	16,67	94,12	Ca	-0,96	1,00				
3	0,18	3,10	97,22	Fe	-0,21	0,25	1,00			
4	0,09	1,59	98,81	S	-0,96	0,89	0,01	1,00		
5	0,06	1,10	99,90	$P_e$	-0,90	0,84	0,16	0,88	1,00	
6	0,00	0,10	100,0	$\rho_b$	-0,91	0,88	0,21	0,88	0,91	1,00

TAB. 6.1 – Résultats d’une Analyse en Composantes Principales. Les deux premières composantes principales décrivent 94% de l’information du jeu de données initial.

Oxyde	Composition du barycentre			
	Groupe 1	Groupe 2	Groupe 3	Groupe 4
SiO <sub>2</sub>	63.0	47.3	59.9	44.2
Al <sub>2</sub> O <sub>3</sub>	21.2	17.6	20.7	16.4
Fe <sub>2</sub> O <sub>3</sub>	<b>7.7</b>	<b>9.9</b>	<b>9.4</b>	<b>6.3</b>
MgO	2.3	3.1	3.5	1.9
MnO	0.1	<b>7.5</b>	0.3	0.6
CaO	1.7	<b>9.1</b>	2.9	<b>25.1</b>
Na <sub>2</sub> O	0.2	0.2	1.1	0.1
K <sub>2</sub> O	<b>2.7</b>	1.7	<b>2.1</b>	1.0
TiO <sub>2</sub>	0.9	0.7	0.9	0.6
P <sub>2</sub> O <sub>5</sub>	0.1	0.3	0.2	0.3
# d'échantillons	26	2	38	16

TAB. 6.2 – Compositions des barycentres des groupes individualisés par la classification en nuées dynamiques des données géochimiques (pourcentages d'oxydes d'éléments majeurs) sur 82 échantillons du puits 948C, Leg ODP 156. Les valeurs en gras sont discriminantes entre les groupes ou permettent d'avoir une idée de la minéralogie.

vidualisés par la classification en nuées dynamiques sont données dans le tableau 6.2. A partir de ces compositions, on peut dégager une lithologie grossière. Les groupes 1 et 3 sont riches en Si et Al, donc composés probablement d'argiles. Le minéral argileux majoritaire est soit l'illite, soit la smectite, avec peut-être une présence un peu plus marquée de l'illite pour la section inférieure, due à la teneur légèrement plus élevée en potassium, alors que des teneurs plus élevées en Na signe probablement une légère prédominance de la smectite dans la partie supérieure de la section (Na est un cation échangeable pour la smectite). La kaolinite est sans doute présente en faibles concentrations car Fe montre des teneurs élevées. Le groupe 2 regroupe deux échantillons anomaux, très riches en Mn, et relativement en Ca. On peut penser à la présence d'oxydes de Mn et Fe (relativement concentré aussi), mais aussi à la présence de rhodocrosite (MnCO<sub>3</sub>), minéral présent en veines dans les sédiments marins. Le groupe 4 est très enrichi en Ca, ce qui signale la présence de calcite en forte concentration, associée avec la matrice argileuse présente mais réduite. On verra plus loin comment ces informations sont importantes pour choisir une solution *a priori* pour l'inversion minérale (voir Section 6.3.3 et Chapitre 7).

### 6.3 Les modèles déterministes

Parmi les nombreuses grandeurs physiques et chimiques mesurées en continu dans les puits de forage par les différentes méthodes de diagraphie, certaines satisfont à des lois déterministes, d'autres pas. Parmi les premières figurent les concentrations chimiques en éléments majeurs, exprimées sous forme de pourcentages d'oxyde (SiO<sub>2</sub>, Al<sub>2</sub>O<sub>3</sub>, Fe<sub>2</sub>O<sub>3</sub>, MgO, CaO, K<sub>2</sub>O, Na<sub>2</sub>O, TiO<sub>2</sub>, ...) dont la somme est généralement proche de 100%. On a vu que certains outils de diagraphie, comme le GLT, permettait d'obtenir les concentrations en 6 éléments majeurs, Si, Ca, Fe, Al, S, Ti, et une trace, le Gd. Pour le pétro-géochimiste, vouloir interpréter cette composition chimique en termes d'assemblages de minéraux est évidemment l'étape suivante naturelle. Beaucoup de méthodes ont été développées qui essaient de trouver par *inversion minérale* de compositions chimiques en éléments majeurs, une *norme* (assemblage minéralogique proposé) qui soit aussi proche que possible des proportions modales (le *mode*) de minéraux dans

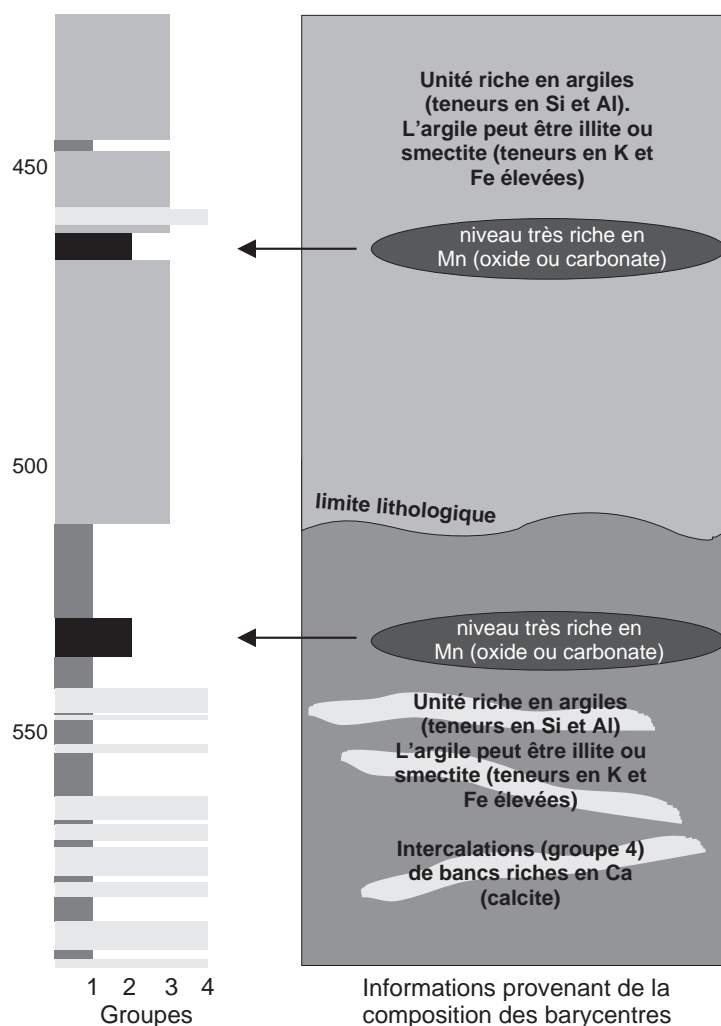


FIG. 6.2 – Première interprétation de l'information directe donnée par la classification en nuées dynamiques : à gauche est portée la classification en 4 groupes, à droite les informations que l'on peut déduire des compositions des barycentres des groupes, reportées dans le tableau 6.2.

la formation.

**Les schémas soustractifs** Ils ont été introduits il y a très longtemps pour les roches magmatiques, puis étendus aux roches métamorphiques. Pour les roches magmatiques, parce que la majorité des minéraux qui les composent ont souvent le temps de se cristalliser et sont de composition chimique simple, on utilise encore la norme CIPW (Cross et al., 1902; Kelsey, 1965). Bien qu'elle ait été développée dans un souci de classification, elle permet de transformer des données chimiques en une norme. L'analyse chimique est supposée précise et correcte, et les minéraux (un assemblage fixe) sont complétés les uns après les autres dans un ordre précis, avec à chaque étape, un des éléments chimiques agissant comme facteur limitant. Les minéraux de la norme ont bien sûr des compositions "idéales".

**Les modèles linéaires exacts** Les grandeurs mesurées sont régies par une loi élémentaire de mélange : la concentration en un élément chimique d'une formation géologique est égale au produit scalaire entre les concentrations en cet élément dans chacun des constituants minéraux et leurs proportions respectives  $x$ . De plus, les proportions des phases minérales dans la roche totale satisfont à une relation de fermeture : leur somme doit être égale à l'unité. Prenons par exemple une roche constituée de trois minéraux, pour laquelle on dispose d'une analyse de sa composition chimique exprimée sous la forme des pourcentages d'oxydes de deux éléments majeurs. Le système est *totalelement déterminé* (autant d'équations que d'inconnues) et s'écrit

$$x_1 \cdot c_{11} + x_2 \cdot c_{12} + x_3 \cdot c_{13} = y_1 \quad (6.1)$$

$$x_1 \cdot c_{21} + x_2 \cdot c_{22} + x_3 \cdot c_{23} = y_2$$

auxquelles on rajoute la contrainte de fermeture :

$$x_1 + x_2 + x_3 = 1$$

Le système se résume sous forme matricielle à  $C \cdot X = Y$ , où  $X$  est le vecteur inconnu. Une solution directe peut être obtenue par une résolution du système d'équations 6.1, qui prend dans ce cas la forme matricielle simple suivante :

$$X = C^{-1} \cdot Y$$

Dans ces deux cas, les erreurs ne sont pas intégrées au calcul, mais associées directement aux proportions  $x_i$  obtenues. Néanmoins, la rigidité de cette méthode (paragenèse et analyse chimique exactes, système totalement déterminé) a poussé à rechercher d'autres techniques d'inversion minérale plus souples.

### 6.3.1 La méthode des moindres carrés ("*least-squares*")

Souvent, le nombre de phases minérales  $j = 1, m$  dans une roche est inférieur au nombre d'éléments chimiques mesurés  $i = 1, n$ . On se trouve alors en présence d'un système sur-déterminé (plus d'équations que d'inconnues). Les méthodes de *minimisation de l'erreur* sont utilisées pour résoudre le problème (Harvey et al., 1990; Harvey and Lovell, 1992; Lofts, 1993). On se trouve à nouveau devant un système d'équations linéaires du type

$$y_i = \sum c_{ij}x_j \quad \text{soit} \quad Y = CX$$

Une solution consiste à résoudre le système de façon à trouver un vecteur  $X$ , puis à recalculer une estimation de la composition chimique de la roche noté  $\hat{Y}$ , qui est comparé au  $Y$  d'origine. Le vecteur  $X$  trouvé par l'inversion sera acceptable si la différence entre  $\hat{Y}$  et  $Y$  est petite. L'approche la plus commune est d'estimer  $X$  de telle façon que l'erreur  $E = (\hat{Y} - Y)^T(\hat{Y} - Y)$  soit minimum, où  $E$  est la somme du carré des différences entre  $\hat{Y}$  et  $Y$ . Le meilleur ajustement possible est donné par la *solution des moindres carrés*, exprimée

$$X = (C^T C)^{-1} C^T Y \quad (6.2)$$

Cette solution ne prend pas en compte les erreurs sur les mesures  $Y$ . Une amélioration consiste à intégrer une matrice de pondération  $M$ , dont les éléments seront les erreurs affectant les compositions mesurées  $Y$ . A partir de l'équation 6.2, on écrit la solution des moindres carrés pondérés :

$$X = (C^T M C)^{-1} C^T M Y$$



où  $M$  est par exemple la matrice des covariances des erreurs de mesure. Si l'on considère que ces erreurs sont indépendantes et suivent une distribution normale centrée sur zéro, cette matrice est diagonale de la forme :

$$M = \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{pmatrix}$$

où  $\sigma_i^2$  est la variance relative à  $y_i$ . On ajoute la contrainte de fermeture  $\sum_{i=1}^m x_i = 1$ , que l'on peut écrire sous forme matricielle  $x^T J_n = J_n^T x = 1$ , où  $J_n$  est le vecteur  $(1, 1, \dots, 1)^T$ . Pour que cette relation de fermeture soit rigoureusement satisfaite, la solution consiste en mettre en oeuvre la méthode des multiplicateurs de Lagrange (Albarède, 1995). Dans ce cas, la forme quadratique  $f^2$  à minimiser s'écrit

$$f^2 = (\hat{y} - C\hat{x})^T (\hat{y} - C\hat{x}) - 2\lambda(\hat{x}^T J - 1)$$

le dernier terme,  $\hat{x}^T J - 1$ , représentant la relation de fermeture, étant nul. Albarède (1995) rappelle que la solution  $\hat{x}$  est donnée par

$$\hat{x} = (C^T C)^{-1} C^T \hat{y} + \lambda (C^T C)^{-1} J = \hat{x}_0 + \lambda (C^T C)^{-1} J$$

où  $\hat{x}_0$  est la solution au sens des moindres carrés, soit  $\hat{x}_0 = (C^T C)^{-1} C^T \hat{y}$ . On calcule alors  $\lambda$ , que l'on insère dans le système.

De la même façon que précédemment, on peut tenir compte des erreurs de mesure par l'intermédiaire d'une pondération. Par contre, cette approche par la méthode des moindres carrés ne permet pas de prendre en compte facilement l'indétermination des paramètres du modèle linéaire, c'est à dire les incertitudes affectant la composition chimique des phases minérales. A l'instar de ce qui se pratique en géochronologie pour le traitement des données radiométriques (York, 1969; Brooks et al., 1972; Minster et al., 1982; Briquieu and de la Boisse, 1990), on peut tenir compte de ces incertitudes en complexifiant les coefficients de pondération :

$$w_i = \frac{1}{\sigma_{y_i}^2 - \sum_{j=1}^n x_j^2 \sigma_{c_{ij}}^2}$$

Introduits sous cette forme, les poids dépendent de la solution  $X$  elle-même ce qui conduit à mettre en oeuvre une méthode itérative tout en restant dans le cadre des moindres carrés.

### 6.3.2 L'évaluation quantitative de l'erreur

Harvey et al. (1990) et Lofts (1993) ont discuté des différents moyens d'estimer l'erreur sur la solution des systèmes d'équations sur-, sous- ou totalement déterminés, ainsi que de leur validité et de leur utilité à améliorer la solution. Doveton (1986) a proposé deux mesures basées sur les proportions minérales trouvées  $x_i$ . Il s'agit de la *variance proportionnelle* PVAR, et de la *somme absolue des erreurs* ASE. PVAR s'apparente à l'entropie maximale (variance) du système et est définie comme

$$PVAR = \frac{m}{m-1} \cdot \left( 1 - \sum_{j=1}^m x_j^2 \right)$$

Une valeur importante de PVAR indique une grande variance, qui atteint 1 lorsque la variance est maximale (proportions des minéraux égales). Doveton (1986) montre que ce paramètre est utile pour les systèmes sous-déterminés, et moins pour les systèmes sur-déterminés. ASE est

utile dans le cas des systèmes totalement déterminés. Elle indique la présence de proportions négatives. Elle est définie comme

$$ASE = \left( \sum_{j=1}^m |x_j| \right) - 1$$

Une forte ASE indique par exemple un assemblage minéral incorrect en entrée.

Dans le cas des systèmes sur-déterminés, la solution des moindres carrés permet une évaluation complète de l'erreur. Harvey et al. (1990) discute de l'utilité de deux mesures de l'erreur à estimer la validité d'une solution. Il s'agit de l'*erreur standard de l'estimation* SE et de la *déviatoin moyenne absolue* MAD. Elles sont calculées à partir de la définition de  $E$  (voir plus haut) :

$$SE = \sqrt{\frac{\sum_{i=1}^n e_i^2}{(n - m - 1)}}$$

SE varie de 0 pour une régression sans erreur, et augmente avec la dispersion des points autour de la droite de régression. Cette grandeur à la structure d'un  $\chi^2$  à  $n - m - 1$  degré de liberté, et l'on peut se reporter aux tables de  $\chi^2$  pour évaluer la pertinence de la solution. Lofts (1993) fait une remarque très juste, applicable à toutes les méthodes d'inversion minérale : une solution montrant une erreur acceptable, ou même nulle, n'implique pas forcément une solution correcte, mais simplement une bonne régression montrant peu d'erreur.

$$MAD = \frac{\sum_{i=1}^n |e_j|}{n}$$

MAD est similaire à SE, mais peut être calculé aussi pour les systèmes totalement déterminés, ce qui permet une comparaison entre les deux méthodes.

### 6.3.3 Modèles non linéaires : l'algorithme de l'inversion totale

Une approche alternative consiste à écrire le système sous la forme de  $n+1$  équations implicites dont on cherche les racines grâce à l'algorithme de l'inversion totale proposé par Tarantola and Valette (1982). Le système d'équations prend la forme suivante :

$$\begin{aligned} y_i - \sum_{j=1}^m c_{ij} x_j &= 0 & \text{avec } i = 1, n \\ \sum_{j=1}^m x_j - 1 &= 0 \end{aligned}$$

que l'on peut écrire  $G(X) = 0$  en considérant que données et paramètres du modèle sont les composantes d'un vecteur unique  $X$  auquel est affectée une matrice de covariance unique  $M$ .

Dès lors, le principe de l'algorithme consiste optimiser tous les paramètres du système en minimisant la distance quadratique

$$q^2 = (X - X_0)^T M_0^{-1} (X - X_0)$$

$X_0$  étant une solution *a priori* acceptable, sous les  $n + 1$  contraintes  $G(X) = 0$ . La méthode des multiplicateurs de Lagrange permet d'écrire  $q^2$  sous la forme (Albarède, 1995)

$$q^2 = (X - X_0)^T (X - X_0) - 2 \sum_{i=1}^{n+1} \lambda_i g_i(X)$$

soit en différenciant :

$$dq^2 = dx^T [X - X_0 - \sum_{i=1}^{n+1} \lambda_i \text{grad}(g_i(X))]$$

Le minimum de  $q^2$  est tel que  $X - X_0 = H\lambda$ ,  $H$  étant le Hessien de  $G(X)$ . Le vecteur  $\lambda$  des multiplicateurs de Lagrange est

$$\lambda = (H^T H)^{-1} H^T (X - X_0)$$

La solution proposée par Tarantola and Valette (1982) prend aussi en compte les incertitudes sur les paramètres par le biais de la matrice des covariances  $M_0$  *a priori*. Elle est donnée par

$$X_{k+1} = X_0 + M_0 \cdot H_k^T \cdot (H_k \cdot M_0 \cdot H_k^T)^{-1} \cdot [H_k \cdot (X - X_0) - G(X)]$$

Une bonne approximation de la matrice des covariances *a posteriori* est donnée par Tarantola and Valette (1982) :

$$M = M_0 - M_0 \cdot H^T \cdot (H \cdot M_0 \cdot H^T)^{-1} \cdot H \cdot M_0$$

Dans la mesure où l'on dispose d'une information *a priori* sur les paramètres du modèle descriptif, l'algorithme proposé de l'inversion totale est une approche bayésienne de l'estimation de ces paramètres, dont la particularité est d'être basée sur l'algorithme du point fixe.

On met ainsi en évidence l'importance de l'information *a priori* sur la qualité de l'estimation des paramètres du modèle et de la solution *a posteriori*. Les gains d'information sont d'autant plus importants que la solution *a priori* est adaptée au problème minéralogique traité. Nous voyons ici apparaître l'utilité d'une classification des données avant le traitement numérique, classification qui va permettre de choisir, pour chaque intervalle reconnu, une solution *a priori* d'autant plus proche du maximum de vraisemblance.

## 6.4 Les modèles statistiques : les réseaux de neurones

### 6.4.1 Introduction

Malgré un engouement relativement récent, l'intérêt scientifique pour les réseaux de neurones ne date pas d'hier. En effet, dès 1943 Warren McCulloch et Walter Pitts, deux neurobiologistes américains, proposèrent un modèle du cerveau où les cellules nerveuses (neurones) pouvaient être comparées à des fonctions logiques. Une théorie qu'implémente Frank Rosenblatt en 1958 avec son *Perceptron*, qui est en fait une modélisation de la perception visuelle. Le Perceptron se présente sous la forme d'une couche d'entrée jouant le rôle de la rétine, recevant l'image d'une lettre, une couche intermédiaire interprétant les signaux transmis via des connexions depuis la couche d'entrée, et une couche de sortie où chaque neurone est associé à une lettre, et doit reconnaître la lettre présentée en entrée. L'apprentissage se fait par une méthode proposée par Rosenblatt, appelée PCP (Procédure de convergence du Perceptron), consistant à modifier les poids affectés aux connexions, de façon à ce que l'activité des neurones auxquels arrivent ces connexions soit corrigée pour donner un résultat correct. Cette valeur ajoutée ou retranchée est une constante. Malgré cela, Rosenblatt note que certains problèmes sont efficacement traités, alors que pour d'autres cette machine neuronale est en échec..

Widrow et Hoff proposent en 1960 un modèle du neurone de McCulloch et Pitts, l'*Adaline* (pour ADaptive LInear NEuron). La règle d'apprentissage de ce type de réseau est appelée *règle de Widrow-Hoff* ou *règle Delta*. Cette règle obéit à une dynamique différente de la PCP, en ce

sens que l'amplitude de la correction est fonction à la fois de l'importance de l'erreur et de la valeur d'entrée. On montre que la règle Delta effectue une *descente de gradient* sur une mesure d'erreur quadratique, que l'on peut écrire simplement par

$$E = \frac{1}{2} \sum_{o \in \text{Sortie}} (d_o - y_o)^2$$

où  $d_o$  et  $y_o$  sont respectivement la sortie désirée et l'activation réelle du neurone de sortie  $o$ . La technique de descente de gradient consiste à modifier les poids du réseau dans le sens inverse du gradient de l'erreur.

En 1969, la publication d'une analyse théorique du perceptron à deux couches par M. Minsky et S. Papert, montrant les limites des possibilités de calcul d'une telle structure, engendre un désintérêt progressif pour les techniques neuronales de la plupart des scientifiques. Seuls certains irréductibles, à l'origine d'ailleurs d'avancées notables dans ce champ d'étude dans les années 70 et 80, continuent leurs efforts, comme Teuvo Kohonen, Stephen Grossberg, James Anderson, ou Kunihiko Fukushima.

Au début des années 80, la découverte de deux techniques capables d'entraîner des réseaux à neurones cachés relance l'intérêt pour les méthodes statistiques multidimensionnelles : la *rétro-propagation de l'erreur* ("backpropagation") et l'*apprentissage de la machine de Boltzmann*. Ces deux techniques sont à l'origine de deux modèles très différents de réseaux :

1. les *réseaux à couches*, représentés par le Perceptron Multi-Couches, auquel nous nous sommes plus particulièrement intéressés
2. les *réseaux récurrents*, comprenant les méthodes de réseau à compétition, dont la *carte topologique* ou *carte auto-organisatrice* proposée par T. Kohonen, est un exemple.

Aujourd'hui, les techniques neuronales sont utilisées dans des domaines aussi variés que la robotique, la météorologie (Thiria et al., 1997), la médecine, ou la biologie.

### 6.4.2 Présentation

Quel que soit leur type, les réseaux de neurones sont des structures composées de *noeuds* ou *neurones*, abritant une fonction de calcul non linéaire souvent simple, interconnectés entre eux par des poids. Ces poids sont adaptés en parallèle à partir d'un schéma d'apprentissage de manière à s'approcher du problème traité (approximation de fonction) (Lippmann, 1987). Les éléments du réseau, ces fameux *neurones*, contiennent une *fonction d'activation* qui anime le neurone en déterminant son *activation* (valeur scalaire). Une fois déterminée, l'activation du neurone est propagée à travers le réseau en utilisant les connexions ( $\equiv$  à des liens synaptiques). Cette propagation constitue le mode de calcul principal du réseau de neurones, qui présente alors un comportement collectif fort complexe à interpréter, à l'origine des grandes capacités de calcul de ces modèles.

**La connectivité** Le comportement du réseau est lié à sa connectivité. Chaque connexion est affectée d'un poids  $w_{ij}$  (poids entre un neurone de la couche  $i$  et un neurone de la couche  $j$ ). La *phase d'apprentissage* consiste à déterminer ces poids synaptiques, ainsi que d'autres paramètres dont on parlera plus loin. L'apprentissage est un processus itératif, qui suit une règle d'apprentissage. On distingue plusieurs types d'apprentissage :

1. apprentissage *supervisé*, lorsque l'utilisateur présente au réseau un ensemble (classification) de référence, à partir duquel sont calculées les modifications des poids synaptiques de façon à minimiser l'erreur entre son évaluation de la solution et la référence.

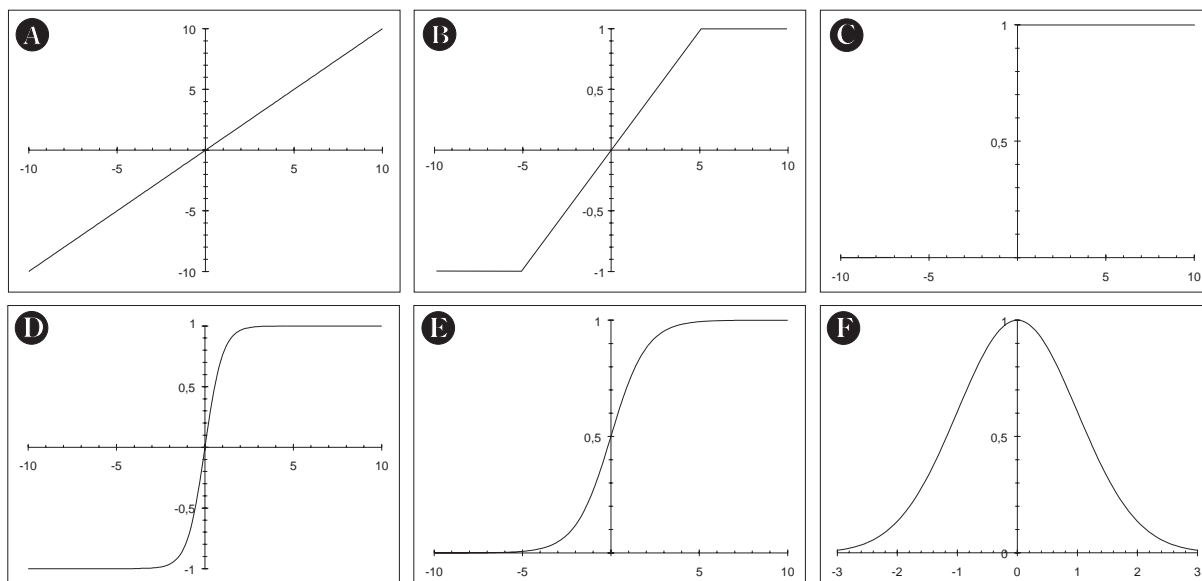


FIG. 6.3 – Exemples de fonctions d'activation utilisées dans les réseaux de neurones (d'après Jodouin (1994a)). **A.** Fonction linéaire :  $y = \lambda \cdot x$ . **B.** Fonction linéaire bornée : compromis entre la f. linéaire et la f. seuil, elle augmente le nombre de réponses possibles. **C.** Fonction seuil : elle est non-linéaire, très simple mais limitée. **D.** Fonction sigmoïde exponentielle : équivalent continu de la f. linéaire bornée, elle est dérivable, et est la f. généralement employée dans les Perceptron Multi-Couches. **E.** Fonction sigmoïde tangentielle (tanh) : très largement répandue également, permet souvent une convergence plus rapide vers la solution lors de l'apprentissage. **F.** Fonction gaussienne : ce n'est pas une f. monotone croissante, elle permet d'introduire une pondération dans l'activation du neurone. Elle est utilisée en particulier pour l'apprentissage des cartes auto-organisatrices.

2. apprentissage *semi-supervisé*, lorsque l'utilisateur n'a que des indications générales à transmettre au réseau (vrai ou faux par exemple) pour l'aider dans sa phase d'apprentissage.
3. apprentissage *non supervisé* : c'est le cas des cartes auto-organisatrices dont nous parlerons plus loin. Les poids du réseau sont modifiés selon des critères internes comme par exemple la coactivation des neurones.

**Types de neurones** Les neurones composant le réseau se classent en trois types :

- les neurones d'*entrée*, qui reçoivent les données d'entrée sous la forme de vecteurs de valeurs scalaires.
- les neurones de *sortie*, contenant la solution et le vecteur de sortie calculé. Les poids sont modifiés en fonction de l'erreur entre cette référence et la sortie calculée.
- les neurones des couches cachées, dont l'efficacité est dépendante du type de fonction d'activation utilisé. Ces neurones vont augmenter la puissance de calcul du réseau et lui permettre d'aborder des problèmes plus complexes.

**La fonction d'activation** La fonction d'activation joue un rôle prépondérant dans la capacité du réseau à s'approcher du résultat. Beaucoup de fonctions ont été proposées, et si, en théorie, n'importe quelle fonction scalaire peut être utilisée, la figure 6.3 donne les plus communément employées. Elles se classent selon trois axes :

1. la fonction est soit à valeurs discrètes soit continue. Les premières ont pour exemple la fonction seuil (1, 0) (Figure 6.3-C). Elles se prêtent à des problèmes simples et en permettent

une analyse théorique facile. De plus elles sont plus adaptées à une implémentation matérielle du réseau. Les secondes sont beaucoup plus utilisées car elles permettent de traiter plus d'informations à la fois dans un même neurone. On distingue

les fonctions linéaires, bornée ou non (Figure 6.3-A et B) sont simples à utiliser mais très limitées dans leurs applications.

les fonctions non linéaires continues : ce sont les plus employées dans les réseaux à couches (Perceptron Multi-Couches) du fait de leurs nombreux avantages, entre autres celui d'être dérivables, et aussi d'être seuillées, c'est à dire de posséder deux valeurs en deçà et au dessus desquelles la réponse de la fonction d'activation est respectivement négligeable et fixe. Cela permet une certaine résistance au bruit et évite de propager de grandes valeurs dans le réseau. Dans nos analyses, nous avons utilisé la plus employée, la fonction *sigmoïde exponentielle* (Figure 6.3-D), d'équation  $y = \frac{1}{1+e^{-x_i}}$  dont la dérivée est  $y' = y_i(1 - y_i)$ . Une autre fonction sigmoïde répandue est la *sigmoïde tangentielle* (Figure 6.3-E), d'équation  $y = \tanh(x_i)$ , dont la dérivée est aussi simple à calculer  $dy = 1 - y_i^2$ . Cette dernière est connue pour permettre une convergence plus rapide vers une erreur minimum lors de l'apprentissage des réseaux multi-couches (Jodouin, 1994a). On verra que la dérivabilité de ces fonctions de transfert est très utile.

- la fonction est soit déterministe (voir point 1), soit stochastique : ces dernières sont des fonctions dont la réponse est en partie aléatoire, et peut pour une même entrée, répondre de plusieurs façons différentes selon une distribution de probabilité donnée. Le neurone biologique en est un exemple. On peut citer aussi la machine de Boltzmann qui utilise une fonction de transfert stochastique

$$p(a_i = 1) = \frac{1}{1 + e^{(-\frac{1}{T}x_i)}}$$

où  $p(a_i = 1)$  exprime la probabilité que l'activation du neurone  $i$  soit à 1, et  $T$  un paramètre nommé "température",  $T \geq 0$ .

- La fonction est à mémoire ou sans mémoire. Dans le premier cas, on tient compte, dans le calcul de l'activité du neurone, des valeurs calculées aux itérations précédentes.

**La fonction d'erreur** Revenons un peu en détail sur la fonction d'erreur. La convergence du réseau se fait par la minimisation de l'erreur entre une solution connue *a priori*  $d_o$  et le résultat calculé  $y_o$ . Si l'erreur quadratique est le plus souvent employée, elle présente un gros désavantage bien connu : l'apprentissage converge très lentement vers la fin du processus (Jodouin, 1994b). Une solution consiste à mettre un seuil à la valeur de l'erreur prise en compte, évitant les minima locaux et diminuant considérablement le temps d'apprentissage, mais évidemment empêchant le réseau de s'approcher trop près de la solution. Une seconde solution est de prendre une autre fonction d'erreur. Certains auteurs ont proposé la *fonction de corrélation croisée* ("cross-correlation") qui s'écrit

$$E = - \sum_{o \in \text{Sortie}} \left[ d_o \ln(y_o) + (1 - d_o) \ln(1 - y_o) \right]$$

dont la dérivée est

$$\frac{\partial E}{\partial y_o} = \frac{y_o - d_o}{y_o(1 - y_o)}$$

Cette fonction partage avec la fonction quadratique la propriété de minimiser l'espérance de l'erreur (Jodouin, 1994b). Mais surtout, si l'on remplace l'erreur quadratique par l'erreur de

corrélation dans la dérivation de la rétropropagation (voir plus loin pour le détail de la règle de rétropropagation de l'erreur), l'erreur est simplifiée à une expression linéaire sur  $y_o$ , qui réagira mieux aux valeurs proches de 1 ou de 0 (la dérivée de la fonction quadratique est non linéaire pour  $y_o$  et 0 et 1 sont deux de ses racines).

Si en théorie, la descente de gradient tel que la pratique la règle Delta est un processus continu le long de la surface d'erreur, qui s'effectue par des mouvements infinitésimaux, en pratique on introduit dans la modification des poids un terme appelé *pas d'apprentissage* noté  $\gamma$ . Le processus d'apprentissage se fait alors de façon discrète, à pas réguliers ou non, selon que  $\gamma$  varie pendant l'entraînement. Lorsque  $\gamma$  est choisi trop grand, il peut arriver que le réseau "oscille" autour du minimum, sans jamais l'atteindre. Un moyen d'éviter cela est d'ajouter à la modification des poids un terme d'*inertie* ou *moment*. La modification des poids est alors fonction de la modification précédente. Ceci fait deux paramètres de plus à choisir, mais en pratique, les auteurs (voir par exemple Jodouin (1994a)) recommandent de choisir une inertie proche de zéro au début, et de faire varier à la fois le terme d'inertie et le pas d'apprentissage au cours de l'entraînement.

### 6.4.3 Le réseau à rétropropagation du gradient ou Perceptron Multi-Couches

Nous avons utilisé plus particulièrement le Perceptron Multi-Couches, à cause surtout de la littérature abondante le concernant (qui nous a permis de mieux en saisir les rouages), mais aussi des nombreux algorithmes existant déjà (et pour la plupart disponibles gratuitement sur Internet) et directement adaptables à nos problèmes.

La figure 6.4 donne le schéma en détail d'un Perceptron Multi-Couches. Comme nous l'avons déjà dit, un PMC se compose d'une couche d'entrée, d'une couche de sortie, et d'une ou plusieurs couches cachées. Le fonctionnement, et en particulier la méthode de rétropropagation, sont assez complexes, aussi me semble-t-il nécessaire de les expliciter. Plus de détails sur la méthode de la rétropropagation de l'erreur, ainsi que sur les différences introduites par les différentes fonctions d'activation ou d'erreur dans le calcul peuvent être trouvés dans Kröse and van der Smagt (1996) ou Jodouin (1994b). La fonction sigmoïde utilisée est la *sigmoïde exponentielle*, et la fonction d'erreur est l'erreur quadratique. Nous ajoutons un terme d'inertie et faisons intervenir un pas d'apprentissage variables dans la procédure d'apprentissage. De nombreux auteurs ont montré que la plupart des problèmes peut être traitée avec un réseau de configuration simple, aussi le perceptron présenté ne possède qu'une couche cachée.

#### Notation

- Chaque échantillon sera noté  $x_{ij}$ , où l'indice  $i$  sera le numéro de l'échantillon considéré et variera de 1 à  $N_i$ , et  $j$  sera la variable considérée et variera de 1 au nombre de variables  $N_v$ .
- Définition de la notation des différentes couches du réseau
  - La couche d'entrée : le numéro du neurone considéré sera notée  $j$  en indice.  $j$  variera de 1 au nombre de variables choisi  $N_v$ .
  - La couche cachée : le numéro du neurone considéré dans la couche cachée sera noté  $k$  en indice.  $k$  variera de 1 à  $N_h$ .
  - La couche de sortie : le numéro du neurone considéré sera noté  $o$  en indice.  $o$  variera de 1 à  $N_o$ .  $N_o$  sera le nombre de faciès/groupes/paragenèses considérés.
- Définition des poids
  - $w_{jk}$  seront les poids entre la couche d'entrée et la première couche cachée.
  - $w_{ko}$  seront les poids entre la dernière couche cachée et la couche de sortie.

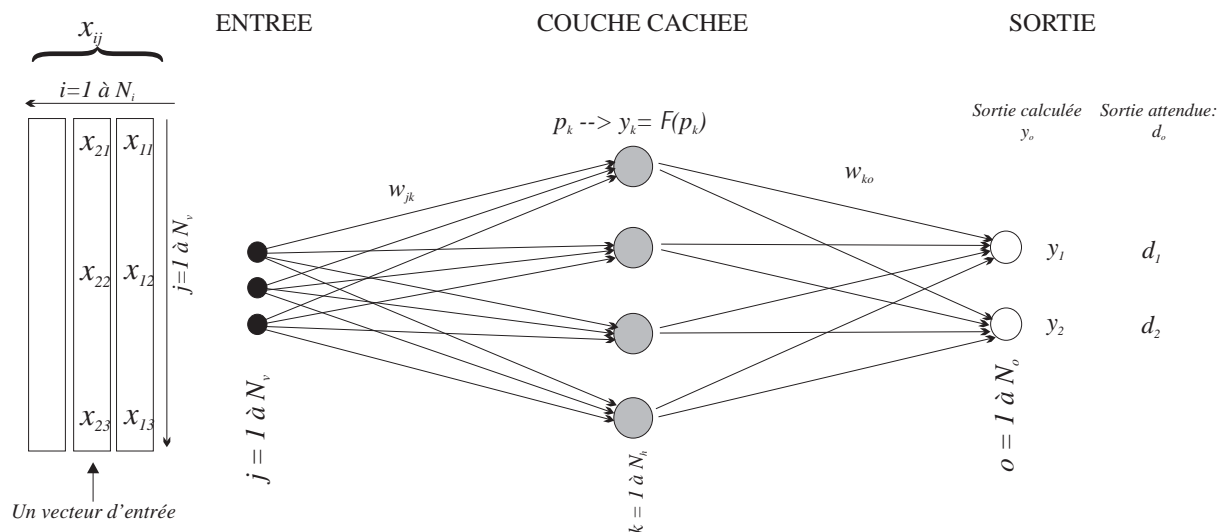


FIG. 6.4 – Schéma général d'un réseau neuronal de type Perceptron Multi-Couches.

- $p_k$  sera la somme des produits scalaires des poids avec les valeurs d'entrée pour la couche cachée.
- La fonction d'activation choisie est une fonction sigmoïde exponentielle notée  $\mathcal{F}$ .
- Notation des entrées pour chaque couche

Lorsque l'on passe de la couche d'entrée à la couche cachée,  $y_k$  est le résultat de la transformation de  $p_k$  par  $\mathcal{F}$  pour la couche cachée.

$y_o$  sera le résultat, pour un neurone de la couche de sortie, de la somme des produits scalaires des  $y_k$  par les poids entre la couche cachée et la couche de sortie, soit  $w_{ko}$ . Les sorties désirées seront notées  $d_o$ .

- $Q_o$  sera la forme quadratique correspondant aux contributions des neurones de la couche cachée pour le neurone  $o$  de la couche de sortie. On appellera *erreur quadratique totale*  $Q^i$  pour un échantillon  $x_i$  la somme des  $Q_o$  pour chaque neurone de la couche de sortie.

### Fonctionnement

- Lors de la phase d'initialisation, les poids des connections entre couche d'entrée, couche cachée et couche de sortie, sont tirés aléatoirement.
- Calcul de  $p_k$  qui correspond, pour chaque neurone de la couche cachée, à la somme des produits scalaires des poids avec les valeurs des échantillons en entrée. Par rapport à la structure de la figure 6.4, nous considèrerons toujours un seul échantillon à la fois :

$$p_k = \sum_{j=1}^{N_v} x_{ij} w_{jk} \quad (6.3)$$

- Le résultat de la fonction d'activation  $\mathcal{F}$  sur les  $p_k$  est

$$\text{pour la couche cachée} \quad (6.4)$$

$$y_k = \mathcal{F}(p_k) = \frac{1}{1 - e^{-p_k}} = \frac{1}{1 - e^{-\sum_{j=1}^{N_v} x_{ij} w_{jk}}} \quad (6.5)$$



- Pour arriver à la forme quadratique, on calcule  $y_o$  :

$$y_o = \sum_{k=1}^{N_h} w_{ko} y_k \quad (6.6)$$

L'erreur quadratique pour un neurone de la couche de sortie  $Q_o$  est :

$$\begin{aligned} Q_o &= \frac{1}{2} (d_o - y_o)^2 \\ &= \frac{1}{2} \left( d_o - \sum_{k=1}^{N_h} y_k w_{ko} \right)^2 \\ &= \frac{1}{2} \left( d_o - \sum_{k=1}^{N_h} \frac{w_{ko}}{1 + e^{-\sum_{k=1}^{N_h} p_k w_{jk}}} \right)^2 \end{aligned} \quad (6.7)$$

On définit l'*erreur quadratique totale* pour un échantillon  $x_i$  par la forme quadratique  $Q^i$  :

$$\begin{aligned} Q^i &= \sum_{o=1}^{N_o} Q_o \\ &= \frac{1}{2} \sum_{o=1}^{N_o} (d_o - y_o)^2 \\ &= \frac{1}{2} \sum_{o=1}^{N_o} \left( d_o - \sum_{k=1}^{N_h} \frac{w_{ko}}{1 + e^{-\sum_{k=1}^{N_h} p_k w_{kc}}} \right)^2 \end{aligned} \quad (6.8)$$

- les poids  $w_{jk}$  et  $w_{ko}$  sont optimisés de façon à minimiser la forme quadratique  $Q^i$ . Le principe utilisé dans la règle de rétropropagation du gradient mis en oeuvre dans le perceptron multicouche consiste à modifier les poids proportionnellement à la *dérivée partielle de l'erreur quadratique totale*, à un facteur  $\gamma$  près, pour un échantillon  $x_i$  par rapport à chaque poids :

$$\Delta^i w = -\gamma \frac{\partial Q^i}{\partial w} \quad (6.9)$$

où  $\gamma$  est le *pas d'apprentissage* du réseau, choisi constant au début, l'utilisateur pouvant le faire varier ou non au cours de l'entraînement. On peut écrire différemment le gradient de l'erreur par rapport à chaque poids du réseau :

pour les poids  $w_{jk}$

$$\frac{\partial Q^i}{\partial w_{jk}} = \frac{\partial Q^i}{\partial p_k} \frac{\partial p_k}{\partial w_{jk}} \quad (6.10)$$

pour les poids  $w_{ko}$

$$\frac{\partial Q^i}{\partial w_{ko}} = \frac{\partial Q^i}{\partial y_o} \frac{\partial y_o}{\partial w_{ko}} \quad (6.11)$$

- On définit le *signal d'erreur*  $\delta_k$  pour chaque neurone d'une couche cachée, pour chaque échantillon en entrée, comme étant respectivement :

$$\delta_k = -\frac{\partial Q^i}{\partial p_k} \quad (6.12)$$

$$\delta_o = -\frac{\partial Q^i}{\partial y_o} \quad (6.13)$$

$\delta$  permet de calculer la modification à appliquer aux poids initiaux  $w_{jk}$  et  $w_{ko}$ . Le poids d'une connection est ajusté par une quantité proportionnelle au produit du signal d'erreur  $\delta$ , soit :

$$\Delta^i w_{jk} = \gamma \delta_k x_j \quad (6.14)$$

pour l'échantillon  $i$  et pour la couche d'entrée où  $j$  varie de 1 à  $N_v$  et la couche cachée où  $k$  varie de 1 à  $N_h$ . Si l'on considère  $\delta_k$ , on peut écrire à partir de l'équation 6.12

$$\begin{aligned} \delta_k &= -\frac{\partial Q^i}{\partial p_k} \\ &= -\frac{\partial Q^i}{\partial y_k} \frac{\partial y_k}{\partial p_k} \end{aligned} \quad (6.15)$$

Pour trouver le second facteur, d'après eq. 6.4

$$\begin{aligned} \frac{\partial y_k}{\partial p_k} &= \mathcal{F}'(p_k) \\ &= \frac{\partial}{\partial p_k} \frac{1}{1 + e^{-p_k}} \\ &= \frac{-e^{-p_k}}{(1 + e^{-p_k})^2} \\ &= -\frac{1}{1 + e^{-p_k}} \frac{e^{-p_k}}{1 + e^{-p_k}} \\ &= y_k(1 - y_k) \end{aligned} \quad (6.16)$$

Pour un neurone de la couche de sortie, on a (eq. 6.11)

$$\frac{\partial Q^i}{\partial y_o} = -(d_o - y_o) \quad (6.17)$$

On obtient donc le signal d'erreur pour un neurone de la couche de sortie  $\delta_o$

$$\begin{aligned} \delta_o &= (d_o - y_o) \mathcal{F}'(y_k) \\ &= (d_o - y_o) y_o (1 - y_o) \end{aligned} \quad (6.18)$$

**Introduction du terme d'inertie** L'entraînement du réseau modifie les poids selon  $\frac{\partial Q^i}{\partial w}$ . Augmenter la valeur de  $\gamma$  peut améliorer la vitesse de convergence, mais des valeurs trop élevées induisent des oscillations autour de la solution, ralentissant alors le processus d'apprentissage. Il est possible d'introduire un *terme d'inertie* (ou *moment*), qui rend la modification des poids dépendante de la valeur des poids calculés lors de l'itération précédente :

$$\Delta^i w_{jk}(t+1) = \gamma \delta_k^i x_j^i + \boxed{\alpha \Delta^i w_{jk}(t)}$$

où  $t$  représentation le numéro du cycle de présentation de l'échantillon  $i$  au réseau de neurones, et  $\alpha$  la *constante d'inertie* qui indique l'importance à donner aux corrections précédentes. Les avantages de l'ajout du moment d'inertie sur l'apprentissage est schématisé Figure 6.5 : en **a.** et **b.**, la convergence est soit lente, soit compromise par des oscillations provoquées par un pas d'apprentissage trop grand. L'ajout d'un terme de d'inertie permet de supprimer les oscillations tout en convergeant relativement vite grâce au pas d'apprentissage élevé (**c.**).

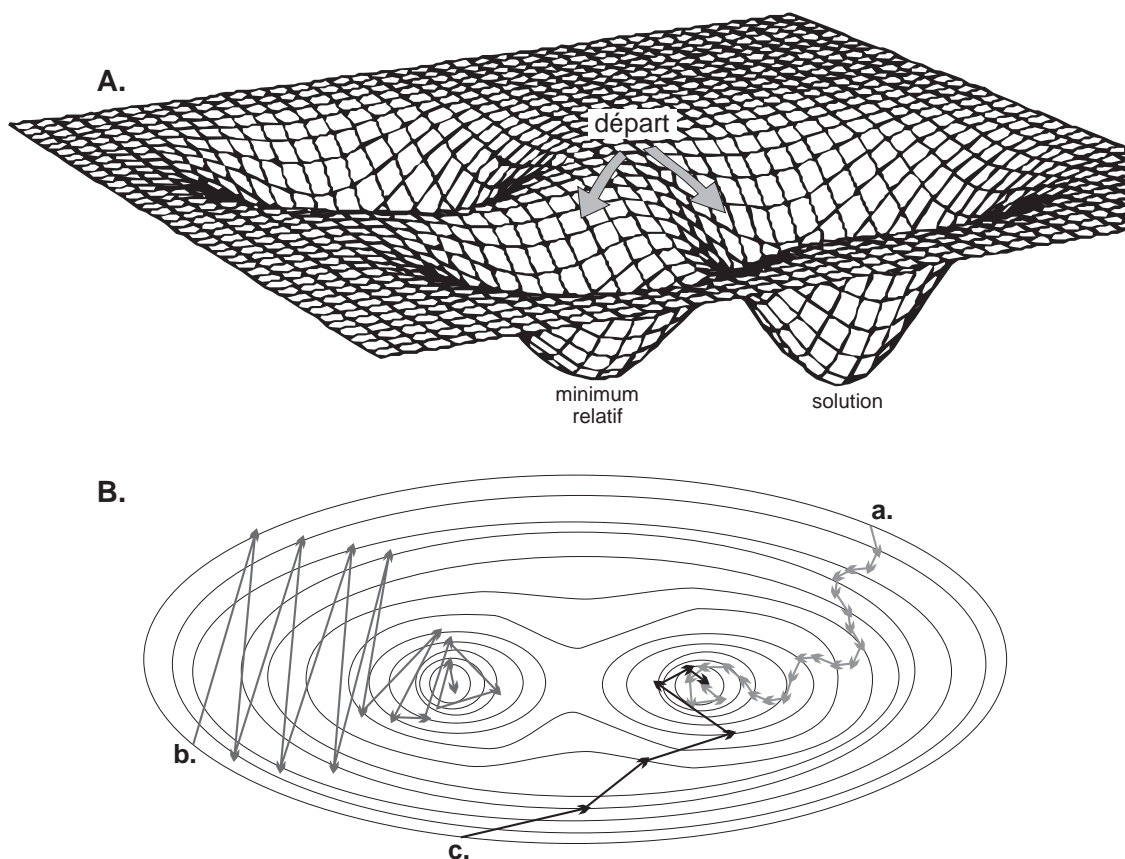


FIG. 6.5 – **A.** Représentation de la surface d'erreur par rapport aux poids synaptiques, montrant la possibilité pour des choix de poids initiaux très proches de converger vers des minima différents. La modification de la valeur du pas d'apprentissage durant l'entraînement permet d'éviter de rester coincé dans un minimum relatif (d'après Jodouin (1994b)). **B.** La descente de gradient est représentée pour plusieurs cas : **a.**  $\gamma$  est petit, pas de terme d'inertie, la convergence est très lente. **b.**  $\gamma$  est grand, pas de terme d'inertie, des oscillations préjudiciables à la convergence apparaissent. **c.**  $\gamma$  est grand mais un terme d'inertie est ajouté, les oscillations sont supprimées et la convergence est rapide (d'après Kröse and van der Smagt (1996)).

**De la théorie à la pratique** Mettre en oeuvre un perceptron n'est donc pas une chose aisée. L'utilisateur a besoin de connaître les principes de la règle d'apprentissage, et d'avoir réfléchi à son problème de façon à bien choisir la configuration de son réseau (qu'il est recommandé de choisir assez simple au début), ainsi que les paramètres de départ. Il lui faut construire un jeu de données d'apprentissage, avec lequel le réseau va apprendre et calculer le jeu de poids synaptiques, un jeu de données de test, contenant des réponses connues de l'utilisateur, permettant de valider le système (le réseau applique simplement les poids calculés lors de l'apprentissage aux vecteurs d'entrées, et les affecte à une classe de sortie) avant de l'utiliser sur un jeu de données à traiter.

Un problème important des réseaux de neurones est le *surapprentissage*. Le principe d'un réseau de neurones est d'approximer une fonction en utilisant une règle d'apprentissage basée sur la minimisation d'une erreur. En théorie, on peut arriver à une erreur nulle, si on laisse le réseau s'entraîner suffisamment longtemps, ou bien si on le complexifie (par exemple en ajoutant des neurones dans sa couche cachée) à outrance. Mais on s'aperçoit dans l'exemple donné par la figure 6.6 que le réseau, en cherchant à approximer au maximum les échantillons du jeu de

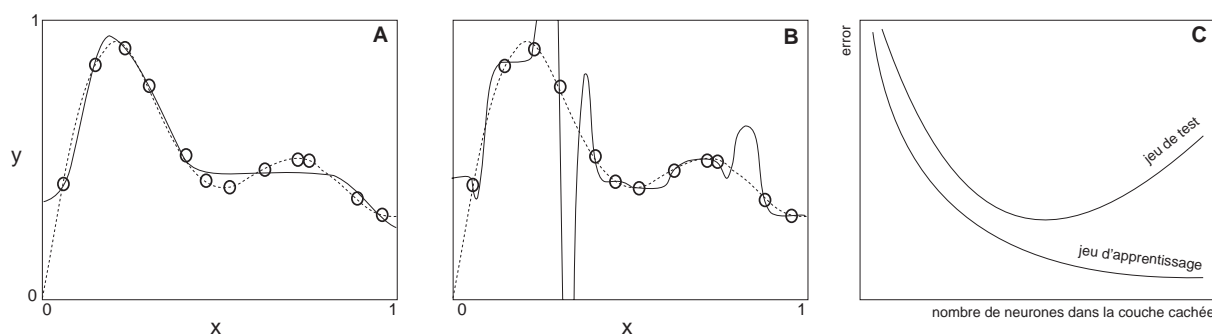


FIG. 6.6 – Le surapprentissage d'un réseau de neurones. Effet du nombre de neurones dans la couche cachée sur les performances : Les points noirs sont les échantillons disponibles dans le jeu de données d'apprentissage, la ligne pointillée donne la fonction désirée, la courbe continue donne l'approximation trouvée par le réseau. La couche cachée est composée de **A.** 5 neurones **B.** 20 neurones. **C.** Evolution de l'erreur d'apprentissage et de l'erreur de test en fonction du nombre de neurones dans la couche cachée.

données d'apprentissage, trouve une fonction qui est très éloignée de la fonction désirée. L'exemple donné concerne un nombre variable de neurones dans la couche cachée, mais le résultat serait le même avec un temps d'apprentissage trop long. Le problème se pose particulièrement lorsque l'on fournit des données entachées d'un bruit au réseau : au lieu de fournir une bonne approximation, le réseau va chercher à intégrer *tous* les points et finit par trouver une fonction n'ayant plus rien à voir avec la fonction désirée. Pour trouver une architecture optimale, une solution consiste à faire croître le nombre de neurones de la couche cachée, et de regarder l'évolution du coût de prédiction (erreur sur les données test) (Figure 6.6-C). On observe que le coût de prédiction (trait plein) passe par un minimum, qui correspond au meilleur compromis entre le biais et la variance : avant le minimum, l'approximation n'est pas assez bonne et le biais est trop important (faible précision), après le minimum, le biais est faible mais la variance est trop importante (faible stabilité  $\Rightarrow$  les approximations calculées ne correspondent plus à la fonction désirée).

#### 6.4.4 Les réseaux à compétition (récurrents)

##### 6.4.4.1 La notion de compétition

Dans le cas du Perceptron Multi-Couches, nous avons vu que la technique classique d'apprentissage consistait à minimiser l'erreur entre une solution calculée par le réseau, et une solution de référence donnée par l'utilisateur. Il existe toutefois des cas où cette base d'apprentissage n'est pas disponible. Il faut alors que le réseau trouve tout seul les différents motifs présents dans le jeu de données. La procédure d'apprentissage est alors *non supervisée*, et la propagation d'activation parmi les neurones se fait selon le principe de *compétition*. Pour cela, il faut s'assurer que les neurones répondent de façon différente aux entrées présentées, contrainte respectée par le tirage aléatoire des poids de départ, que les fonctions d'activation des neurones soient bornées, et que les neurones soient interconnectés entre eux pour permettre le mécanisme de concurrence.

Dans un cas simple, si un vecteur d'entrée est présenté au réseau, un seul neurone va réagir et être activé (c'est le neurone "gagnant"). Cet apprentissage compétitif va rapprocher les poids d'entrée du neurone ayant gagné la compétition du vecteur d'entrée responsable de la victoire. Si un vecteur similaire est présenté, ce neurone aura beaucoup plus de chances de gagner que les autres. En contrepartie, il aura moins de chances de gagner si l'entrée est différente.

Pour sélectionner le ou les gagnants, il existe plusieurs méthodes. Si par exemple un vecteur  $x$  est présenté, l'activation pondérée qui converge sur les neurones de sortie est

$$y_o = \sum_{c \in \text{Entrée}} w_o y_c = \sum_{c \in \text{Entrée}} w_o x_c$$

Cette valeur peut être considérée comme une mesure de l'angle  $\alpha$  entre le vecteur  $x$  et le vecteur des poids :

$$y_o = \sum_i w_{io} x_i = w_{io}^T x = |x| \cdot \|w_{io}\| \cdot \cos \alpha$$

Si ces grandeurs sont normalisées, on voit que l'activité pondérée d'un neurone se réduit à  $\cos \alpha$ . Elle sera d'autant plus grande que le vecteur des poids est parallèle à  $x$ , donc aura le maximum de chances de faire gagner ce neurone. La règle d'apprentissage redistribue alors les poids des neurones vainqueurs, de manière à réduire l'angle entre le vecteur des poids et le vecteur d'entrée  $x$  (voir Figure 6.7), selon

$$w_v(t+1) = w_v(t) + \gamma(x(t) - w_v(t))$$

où  $v$  représente les neurones gagnants, et  $t$  le cycle, et garde les poids des autres neurones intacts  $\Delta w_o = 0$ , où  $i$  représente les neurones perdants. Les vecteurs de poids des neurones gagnants sont, après l'apprentissage, particulièrement réglés sur différents domaines du jeu de données d'entrée. Dans le cas où les variables ne sont pas normalisées, au lieu d'utiliser le produit scalaire entre  $x$  et le vecteur poids, on choisit plutôt la *distance euclidienne* entre les deux vecteurs exprimée par  $d_o = \sum_{i=0}^{n-1} (x_i(t) - w_{io}(t))^2$ , où  $x_i(t)$  est l'entrée présente dans le neurone d'entrée  $i$ . La distance minimum détermine le neurone gagnant (Lippmann, 1987; Kohonen, 1990).

Il existe toutefois un problème à cette méthode d'apprentissage : les laissés-pour-compte ! Les neurones gagnants bénéficient de tout l'entraînement, aux dépens de certains autres qui se trouvent isolés dans l'espace, et ne codent pas d'information du jeu de données. Un second problème est celui du recodage : les données ne déterminent pas de groupes bien individualisés dans l'espace, et les neurones n'arrivent pas à se stabiliser sur le barycentre d'un groupe. Pour résoudre le problème de l'isolement, certains auteurs ont proposé une nouvelle règle d'apprentissage appelée *apprentissage à fuite* ("leaky-learning"). Avec cette règle, le neurone gagnant apprend le plus, mais tous les autres neurones apprennent un peu aussi à chaque correction, ce qui leur permet de se rapprocher de certains groupes avec lesquels ils ont des "affinités" géométriques. Une autre méthode, appelée "*frequency sensitive competitive learning*", enregistre pour chaque neurone le nombre de fois qu'il est gagnant. Au cours de l'apprentissage, la sensibilité du neurone à la compétition va décroître proportionnellement au nombre de fois qu'il a gagné. Inversement, un neurone qui perd souvent, va se retrouver de plus en plus sensible à la compétition. Malgré ces problèmes inhérents à de tels modèles statistiques, les réseaux à compétition sont simples et efficaces pour résoudre des problèmes quelquefois complexes de classification automatique.

#### 6.4.4.2 La carte topologique (ou auto-organisatrice) de Kohonen ("*Self-Organizing Map*")

**La structuration de la carte** Si les réseaux à compétition ne tiennent pas compte de la position de leurs neurones dans l'espace, la carte topologique ("*Self-Organizing Map*") lie les neurones qui la composent en une structure élastique (Jodouin, 1994b). La forme de cette structure évolue lors de l'apprentissage, et les neurones d'un voisinage sont liés entre eux et leur comportement final varie graduellement d'une région de la carte à l'autre, plutôt qu'arbitrairement comme c'était le cas dans l'exemple précédent. Les vecteurs de poids sont ainsi organisés pour que des noeuds

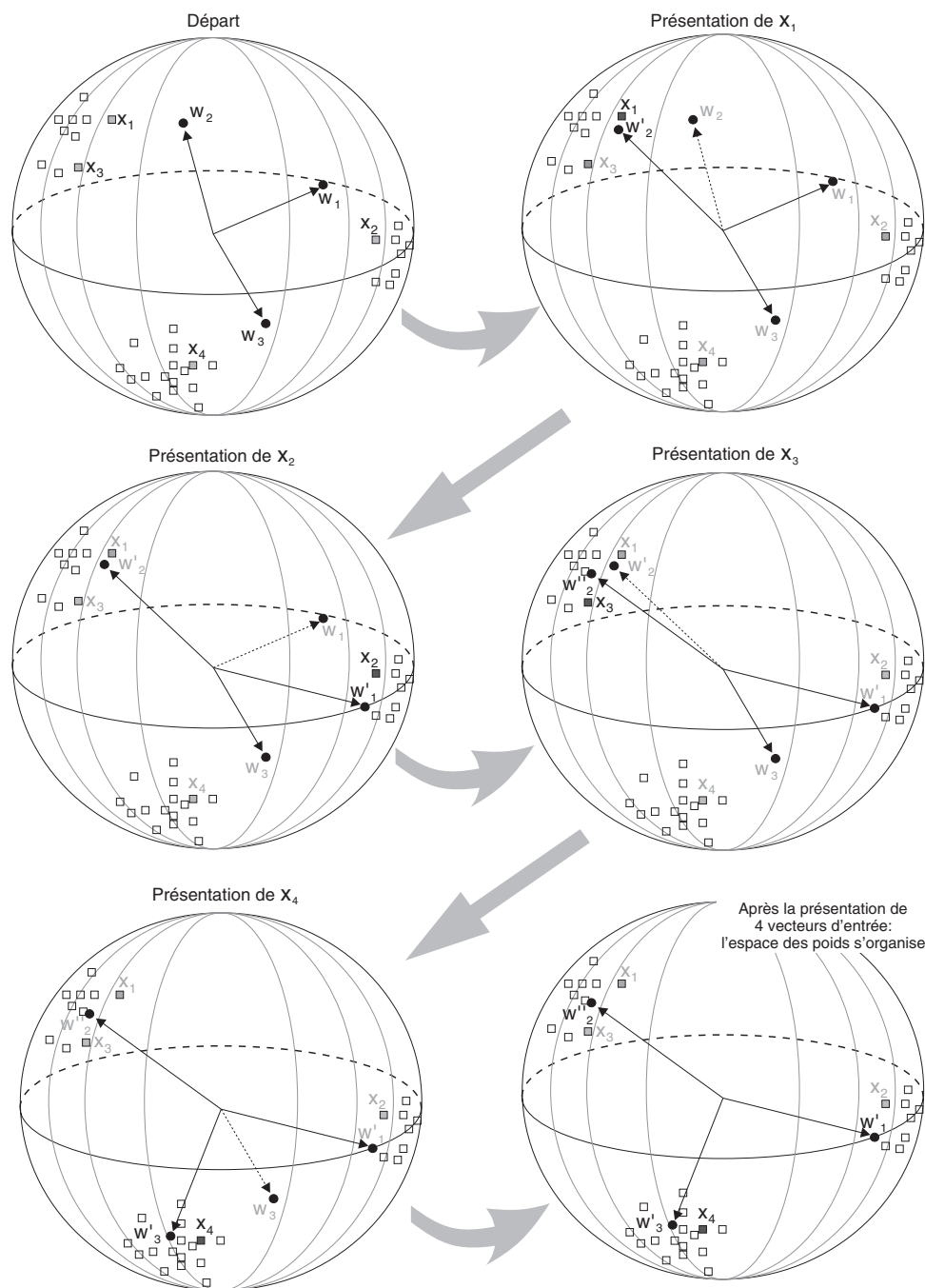


FIG. 6.7 – Schématisation de l'apprentissage par compétition entre neurones. Au départ, on considère 4 vecteurs d'entrée  $(x_1, \dots, x_4)$ , et trois poids tirés aléatoirement  $(w_1, \dots, w_3)$ . Lors de la présentation du vecteur  $x_1$ , le vecteur poids le plus proche,  $w_2$ , gagne la compétition, et est modifié de façon à se rapprocher de  $x_1$ . Les autres vecteurs poids sont inchangés. Les vecteurs poids  $x_2$  et  $x_4$  font gagner  $w_1$  et  $w_3$  de la même manière. On remarque que la présentation de  $x_3$  conforte le vecteur  $w_2$ , qui se rapproche du centre de la nuée. A la fin de l'apprentissage, les vecteurs poids seront organisés de manière à coder l'espace entier, avec un neurone codant pour une nuée (adapté de Kröse and van der Smagt (1996)). Bien sûr, ceci est le cas idéal où les nuées sont bien distinctes et le nombre de vecteurs poids restreint de manière à éviter les problèmes de neurones isolés ou de recodage cyclique.

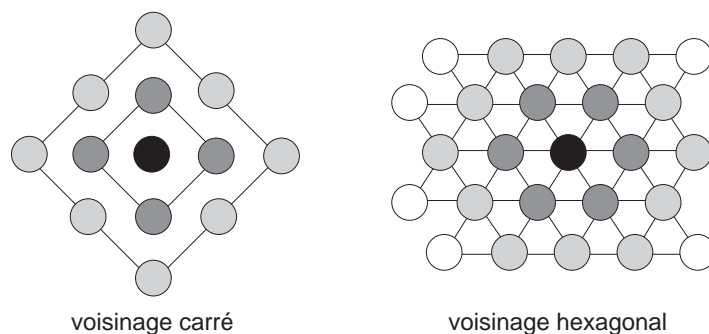


FIG. 6.8 – Les deux topologies les plus communément utilisées pour la carte auto-organisatrice de Kohonen : voisinage carré, ou voisinage hexagonal. Le neurone noir est le neurone vainqueur, les neurones gris foncé représentent le voisinage immédiat, tandis que les neurones gris clair représentent le voisinage étendu.

du réseau proche dans l'espace soient sensibles à des vecteurs d'entrées similaires (Lippmann, 1987; Kohonen, 1990).

La carte topologique fonctionne sur le même principe que les autres réseaux à compétition : une couche de neurones d'entrée reçoit chaque vecteur d'entrée les uns après les autres, et les présente à la grille de neurones interconnectés qui forment l'ossature de la carte. Les poids de chaque neurone étant au départ tirés aléatoirement, ils sont complètement désorganisés. L'algorithme d'auto-organisation définit un *voisinage* autour de chaque neurone, qui est simplement une mesure de la distance euclidienne entre les neurones de la carte. On définit un *voisinage immédiat*  $V_{v,1}$  autour de chaque neurone comme étant l'ensemble des neurones à au plus une distance unité, puis un *voisinage étendu*  $V_{v,E}$ , pour l'ensemble des neurones se trouvant à  $n$  unités du neurone. Cette notion de voisinage prendra tout son sens lors de l'apprentissage de la carte. Il faut aussi définir la topologie de la carte, c'est à dire la façon dont les neurones sont arrangés. Ceci détermine évidemment le nombre de neurones se trouvant dans le voisinage immédiat d'un noeud. La figure 6.8 montre deux topologies couramment utilisées, à voisinage carré et hexagonal. Bien évidemment, cette structure existe uniquement dans la fonction de voisinage lorsque l'apprentissage commence, mais se développe au fur et à mesure de l'organisation des vecteurs de poids stimulée par les vecteurs d'entrée qui leur sont présenté.

**L'apprentissage** La carte topologique s'organise grâce à une règle d'apprentissage non supervisée et compétitive, avec en plus l'intégration d'une variante de l'apprentissage à fuite (voir plus haut). Tous les neurones *dans le voisinage* du neurone gagnant ont droit à un apprentissage eux-aussi. La modification des poids se fait selon

$$\Delta w_v = \gamma(v, o, t) \cdot (x_i - w_v) \quad (6.19)$$

où  $w_v$  représente le neurone gagnant. Celui-ci profite complètement de cette modification, tandis que les neurones présents dans son voisinage voient les poids de leurs connections avec le vecteur entrée plus ou moins modifiés selon la fonction choisie : La figure 6.9 montre trois fonctions de modification du pas d'apprentissage couramment utilisées : la fonction gaussienne, où le pas décroît proportionnellement à la distance avec le neurone vainqueur (gaussienne). Le pas d'apprentissage est alors aussi fonction de la distance

$$\gamma(o, i, t) = \gamma(t) \cdot e^{-\frac{d(o,v)^2}{2\sigma(t)^2}}$$

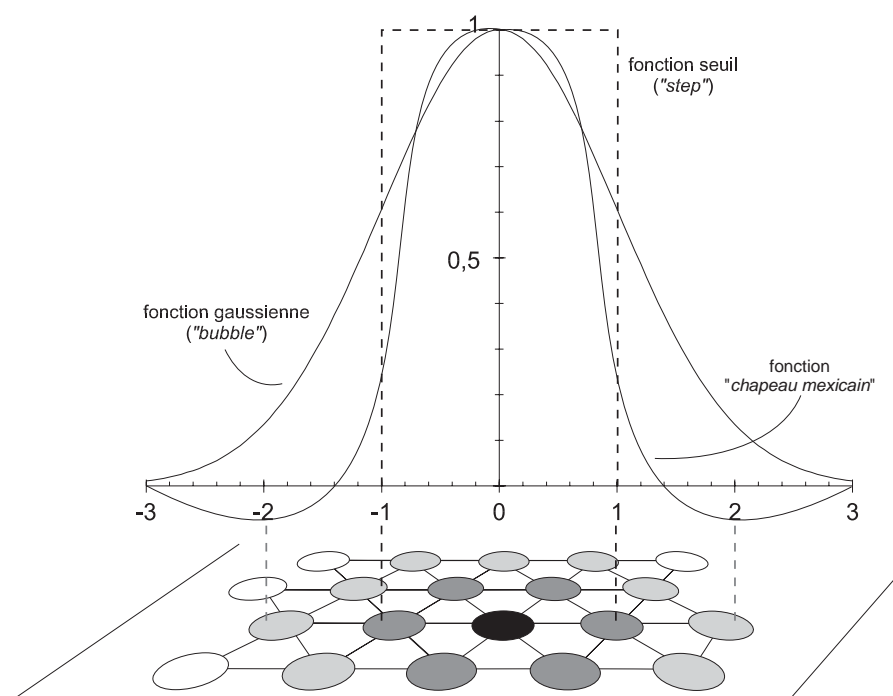


FIG. 6.9 – Trois formes de la fonction caractérisant le pas d'apprentissage au voisinage du neurone vainqueur : la fonction gaussienne permet une décroissance continue de l'apprentissage autour du neurone central gagnant. La fonction seuil permet une modification identique et maximale pour tous les neurones se trouvant dans le voisinage immédiat du neurone gagnant ( $V_{v,1}$ ), et ne modifie pas les autres neurones de la carte. La fonction "chapeau mexicain" permet un apprentissage pour les neurones dans le voisinage du neurone gagnant comme avec la fonction gaussienne, mais *désorganise* les neurones plus éloignés. On a représenté une carte simplifiée à voisinage hexagonal pour une meilleure visualisation de cette évolution du pas d'apprentissage.

où  $\sigma(t)$  est le rayon du voisinage étendu (écart-type de la gaussienne), et décroît avec le temps (Jodouin, 1994b; Kohonen, 1995). Cette manière d'entraîner les neurones voisins crée des une bulle centrée autour du vainqueur (d'où le nom de "bubble" pour la fonction gaussienne du pas d'apprentissage). Une autre possibilité est d'utiliser une fonction seuil ("step") : tous les neurones dans le voisinage immédiat du neurone vainqueur sont entraînés de la même façon que celui-ci, alors que les autres ne subissent aucune modification. Une troisième possibilité est d'utiliser la fonction "chapeau mexicain" : les neurones du voisinage apprennent de la même façon qu'avec la fonction gaussienne, mais les neurones plus éloignés sont cette fois désorganisés. Kohonen (1990) montre également qu'il suffit de réduire au cours de l'apprentissage les voisinages immédiat et étendu autour des neurones vainqueurs pour, au début, organiser grossièrement la carte, puis bien affiner le comportement des neurones et les stabiliser. Ces types d'apprentissage sont responsables des *gradations* dans les spécialisations observables sur la carte, très utiles car certains individus du jeu de données initial, représentant des transitions entre des nuées mieux individualisées, pourront alors être codés par un ou plusieurs neurones.

Le résultat de l'apprentissage est la structuration de la carte, où tous les neurones présents dans un voisinage codent des vecteurs d'entrée sensiblement identiques. Sammon (1969) a proposé un algorithme permettant de représenter sous la forme d'une projection bidimensionnelle un ensemble de vecteurs multidimensionnels. Cette projection a la particularité intéressante de préserver la notion de proximité entre les vecteurs. Cette représentation est classiquement utilisée



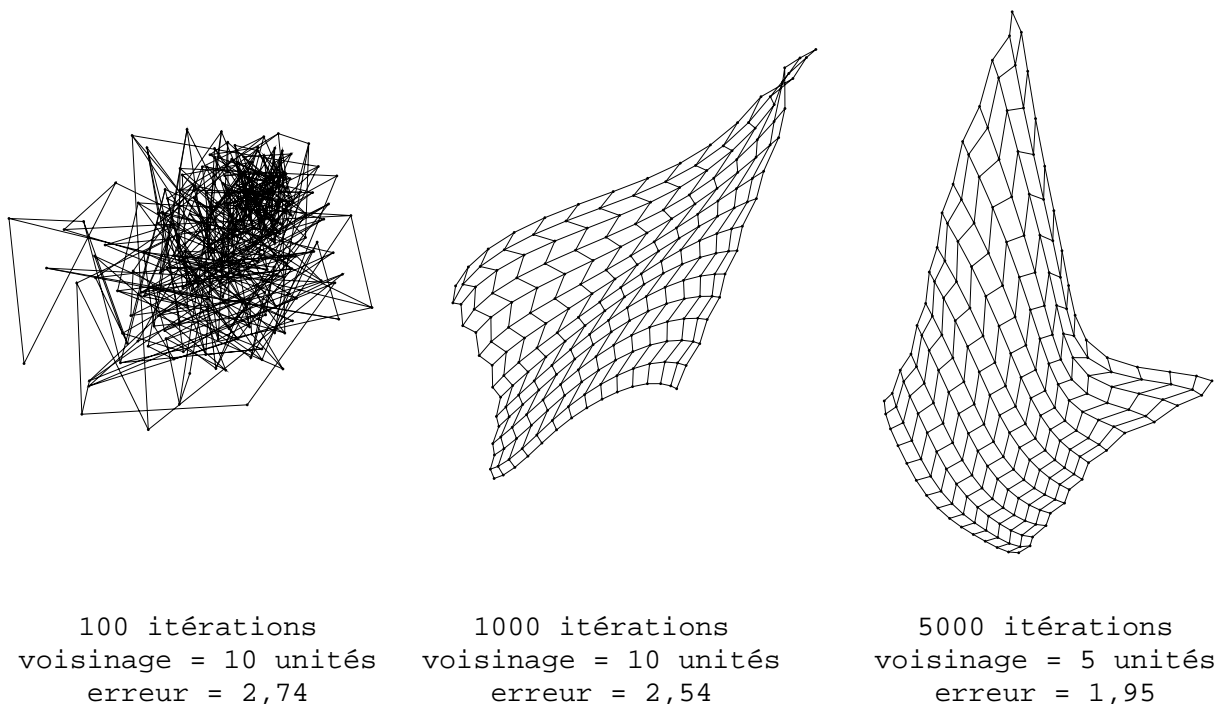


FIG. 6.10 – Structuration de la carte auto-organisatrice au cours de la phase d’apprentissage. La carte est de dimension  $15 \times 15$ , avec une topologie hexagonale et une fonction gaussienne pour l’évolution du pas d’apprentissage autour des neurones gagnants. Le voisinage autour du neurone gagnant a été diminué au cours de l’apprentissage, de façon à structurer la carte grossièrement, puis plus finement.

pour projeter en 2D les vecteurs associés aux neurones de la carte topologique, et ainsi de pouvoir visualiser le déploiement de la carte au cours de l’auto-organisation (Figure 6.10). De plus, la représentation 2D montre éventuellement les différents “agrégats” de neurones qui se sont formés lors de la phase d’entraînement, ce qui donne déjà une idée sur la structure du jeu de données initial.

#### 6.4.4.3 Les techniques de “*Vector Quantization*” et “*Learning Vector Quantization*”

De même que la carte auto-organisatrice, la technique de “*Vector Quantization*” est une technique de classification des données. Mais, contrairement à la classification multivariée dont nous avons parlé au chapitre précédent, on ne recherche pas à individualiser des groupes de données présentant des caractéristiques similaires, mais plutôt à quantifier l’espace. L’algorithme recherche les densités de probabilité les plus grandes dans le jeu de données d’entrée (Kohonen, 1990; Jodouin, 1994b; Kröse and van der Smagt, 1996). La densité des neurones sera plus grande dans les régions de l’espace où les données ont plus de chance d’apparaître. Cette méthode est très utilisée en compression de données (image, signal), car elle garde très précisément jusqu’au plus petites caractéristiques d’un signal (car celui-ci sera décrit par un grand nombre de neurones), tout en étant performante car utilisant une topologie “discontinue”, donc moins de neurones qu’une grille courante.

**Les techniques de “*Learning Vector Quantization*”** La méthode dite des “*Learning Vector Quantization*” ou LVQ, introduite par Teuvo Kohonen au milieu des années 80, est en fait une

technique *supervisée*, car elle prend des décisions basées sur les caractéristiques du jeu de données de départ. Si les techniques de “*Vector Quantization*” étaient surtout destinés à décrire les différentes composantes de l’espace des données, les méthodes de LVQ recherchent les *frontières* entre les classes, car c’est de leur détermination correcte que dépend la minimisation de l’erreur de classification. En ce sens, les LVQ sont une manière d’améliorer encore le résultat d’une classification (qu’elle soit multivariée classique ou effectuée par une carte topologique).

Les frontières entre deux classes est définie en statistiques par une *surface de décision bayésienne*, qui est simplement la surface de partage où la probabilité d’appartenir à une classe ou à une autre est égale. Les techniques de LVQ vont approximer ces surfaces entre les groupes, en utilisant comme point de départ (elles sont donc supervisées) les vecteurs trouvés par une analyse non-supervisée telle que la carte topologique de Kohonen ou plus généralement une technique de “*Vector Quantization*” (Kohonen, 1990; Jodouin, 1994b). Bien que beaucoup de techniques légèrement différentes de LVQ soient disponibles, le principe général de l’analyse est le suivant :

1. considérons qu’un label  $y_o$  soit associé à chaque neurone de sortie  $o$ .
2. on présente plusieurs échantillons consistant chacun en un vecteur d’entrée  $x$  associé à son label  $y_o$  correctement déterminé par la classification.
3. les distances entre les vecteurs poids  $w$  et les vecteurs d’entrée  $x$  sont calculées, et les vecteurs poids sont marqués et classés en plusieurs groupes selon leur proximité avec les vecteurs de référence représentés par les vecteurs échantillons. La modification des poids se fait de la même façon que pour la première classification (voir carte topologique, eq. 6.19), mais *sélectivement* selon que le vecteur poids a *bien répondu* (son label est le même que le label des vecteur d’entrée) ou *mal répondu* (son label est différent) (Kohonen, 1990) :

si le label  $y_k$  du neurone testé  $k$ , est différent du label des  $x$  :

$$w_k(t+1) = w_k(t) - \gamma(x - w_k(t))$$

si le label  $y_k$  du neurone testé  $k$  est le même que celui des  $x$  :

$$w_k(t+1) = w_k(t) + \gamma(x - w_k(t))$$

Cela revient en clair à déplacer itérativement les vecteurs poids de manière à les *éloigner* le plus possible des surfaces de décision bayésienne entre les classes. Les vecteurs poids viennent alors approximer au maximum les *limites des classes*, et laissent vides les zones inter-classes. Comme précédemment,  $\gamma$  est le pas d’apprentissage représenté par une fonction monotone décroissante sur  $t$ . Puisque les LVQ servent à préciser une structure ébauchée par les VQ, il est bon de choisir un pas d’apprentissage plus petit que celui pris lors de la classification, de manière à tenir compte de l’entraînement préalable qu’ont subi les vecteurs poids (Kohonen, 1990).

## 6.5 Remarque générale sur les réseaux de neurones

On se rend compte en utilisant un réseau de neurones, qu’il soit à entraînement non supervisé comme la carte topologique de Kohonen, ou supervisé comme le Perceptron Multi-Couches, combien la phase d’entraînement est importante. Du point de vue théorique, les techniques de classification classiques et les techniques neuronales sont extrêmement similaires. Du point de vue de l’utilisateur, la grande différence est l’*apprentissage*. Le même jeu de données traité par une carte topologique ne produira pas le même résultat selon la façon dont l’utilisateur aura mené l’apprentissage. De la même façon, un mauvais jeu d’indexation fourni au Perceptron ne lui permettra pas de converger vers la bonne solution. Ceci semble évident, mais ce qui l’est

moins, c'est que malgré un bon jeu d'indexation, un pas d'apprentissage trop important ou un mauvais choix de fonction d'activation rendra la convergence difficile, voire impossible. De plus, les problèmes géologiques étudiés définissent peu souvent des comportements bien individualisés, mais sont plutôt une somme (ou une moyenne) de plusieurs phénomènes qui se combinent. Ceci rend la tâche d'un réseau de neurones difficile, malgré les quantités de techniques développées. Pour obtenir une bonne réponse, il est souvent nécessaire de réfléchir à la paramétrisation du réseau et au problème demandé, et surtout de ne pas prendre des valeurs par défaut aveuglément, avant de commencer l'apprentissage sous peine de se voir confronté à des réponses peu discriminantes qui ne rendraient pas compte de la puissance du réseau, et dans notre cas, de la réalité géologique.

Ces difficultés de mise en oeuvre sont à l'origine de l'incrédulité, du scepticisme et du découragement qui entourent encore les techniques neuronales. Pourtant de nombreux exemples, y compris dans le domaine des sciences de la terre (Thiria et al., 1997; Chang et al., 1997; Rogers et al., 1992) montrent combien ces outils peuvent être efficaces pour résoudre des problèmes complexes pour lesquels on ne possède pas un modèle descriptif explicite.

## Troisième partie

# Retrouver l'information géologique à partir des diagraphies



## Chapitre 7

# Reconstruire un log minéralogique

- Le bassin d'Aoba, arc volcanique des Vanuatu, Leg ODP 134
- Le prisme d'accrétion Nord-Barbade, Leg ODP 156
- Le mont Eratosthènes, Méditerranée orientale, Leg ODP 160

7.1 Puits 833B, arc volcanique des Vanuatu, Leg ODP 134

**Article 2 :**

*Scientific Drilling*, Volume 6, Number 2, 1997, pp. 91–101

Characterisation of oceanic sediments and basalts from Hole 833B, New  
Hebrides Island arc (Vanuatu) from downhole measurements

**Alain Rabaute**, Louis Briqueu, Peter K. Harvey, and Henri Mercadier

## 7.2 Sites 947 et 948, prisme d'accrétion de la Barbade, Leg ODP 156

### 7.2.1 Inversion minéralogique de données géochimiques obtenues par analyse en fluorescence X sur des échantillons carottés

**Article 3 :**

soumis à *Geophysical Research Letters*,

Combination of cluster analysis and total inversion of geochemical data as an advantageous way of inferring mineralogy : Example from the North-Barbados accretionary prism, Leg ODP 156

**Alain Rabaute**, Louis Briquet, and Mohamed Ramadan



To appear in the *Geophysical Research Letters*, 1998.

## **Combination of cluster analysis and total inversion of geochemical data as a advantageous way of inferring mineralogy: Example from the North-Barbados accretionary prism, Leg ODP 156.**

Alain Rabaute, Louis Briqueu, Mohamed Ramadan

### **Abstract**

In order to infer mineralogy, the total inversion algorithm is used on major oxides weight percent data measured on samples from cores taken in ODP Hole 948C, North-Barbados accretionary prism. We show that using at first a simple cluster analysis on the data set helps the total inversion algorithm to account for varying mineralogy. Each cluster's centroid provides an average geochemical composition which can drive the choice for an appropriate *a priori* mineral assemblage. The advantages of the total inversion over a classical least-square fitting is the taking into consideration of the original uncertainties on the input parameters and data. The calculated error on the *a posteriori* solution is used to monitor the convergence process in order to track down the source of potential error.

## Introduction

Laboratory and field devices are capable nowadays to give precise major elements concentrations from bulk rock samples. One way for these data to be of any use for the geologist is to be represented in terms of an assemblage of mineral and fluid phases using various mineral inversion techniques. In case of sedimentary rocks, the problem becomes complex because most encountered minerals have different habits and origins, such as detrital or secondary authigenic, and can have very heterogeneous geochemical compositions, a good example being clay minerals.

The principal objective of Ocean Drilling Program Leg 156 was to evaluate and monitor the effects, rates and episodicity of fluid flow in an accretionary prism environment. Knowing precisely the mineralogy of the sediments is an important parameter to understand these effects on the formation. The classical X-ray diffraction method can give qualitative estimates of mineral abundances, but semiquantitative analysis is known to be difficult. *Fisher and Underwood* [1995] proposed a mathematical technique using matrix singular value decomposition that allows accurate conversion of XRD data to relative mineral abundances. Despite extremely good results, the authors acknowledge some limitations to this approach. The normalization factors calculated are valid only for the ranges of abundances used in the standard mixtures chosen. Varying chemical compositions could create mismatches in the model. The system needs to be fully-determined in order to be solved. In a tentative to overcome these problems, we propose here to use the total inversion algorithm [*Tarantola and Valette* 1982] as a way to calculate mineral abundances from geochemical data.

## Core sampling and sediment geochemistry

During Leg 156, 170 metres of marine sediments were cored in Hole 948C across the décollement zone of the North-Barbados accretionary prism. The cored section could be divided into two distinct lithological units, composed of sandy and silty claystones. The clay minerals that form the background sediment are mostly smectite (Unit I) and illite (Unit II), with some kaolinite. 82 samples, 10 cc in volume, were carefully taken from the cores, each from a single lithologic variation. X-ray fluorescence analysis was performed to obtain the concentrations of the 10 common major oxides CaO, SiO<sub>2</sub>, Al<sub>2</sub>O<sub>3</sub>, Fe<sub>2</sub>O<sub>3</sub>, K<sub>2</sub>O, TiO<sub>2</sub>, MgO,

MnO, Na<sub>2</sub>O, P<sub>2</sub>O<sub>5</sub>. The precision on these measurements are given for each element in Table 1.

## Mineral transform

A lot of approaches to the calculation of an accurate mode have been investigated over the last thirty years, either in the field of algebra such as solving for sets of simultaneous equations by matrix inversion [*Albarède and Provost*, 1977]. All of the latter are based on the assumption that a linear relationship exists between the bulk geochemical composition of a sample and the mineral and fluid phases present in that sample, that can be expressed by

$$y_i = \sum c_{ij} \cdot x_j \quad \text{that is} \quad Y = C \cdot X$$

where  $c_{ij}$  is the weight percentage of the  $i^{\text{th}}$  oxide in the  $j^{\text{th}}$  constituent phase,  $x_j$  the proportion of the  $j^{\text{th}}$  phase present in the mode, and  $y_i$  the weight percentage of the  $i^{\text{th}}$  oxide in the rock.  $i = 1, n$ , where  $n$  is the number of oxides analysed and  $j = 1, m$ , where  $m$  is the number of phases present. Moreover, the sum of minerals equals one, as do the sums of the oxide weight percentages in each mineral and in the rock.

The basic limitation becomes now the number of oxides (input responses) available for modelling. *Harvey et al.* [1990] have addressed the problem of under-, over- or fully-determined systems, the latter two cases describing classical geological systems. Another requirement is to take into account the uncertainties. Both systematic and random errors on the geochemical measurements have to be propagated through to the modal analysis and final solution. The classical least squares solution can take into account the uncertainties on the measurements  $y_i$  through a covariance matrix  $M$ . Assuming that these errors are independent and follow a normal law centered on zero,  $M$  is diagonal and is composed of the variances on the input parameters  $\sigma_1^2, \dots, \sigma_n^2$ . The least squares solution, expressed as  $X = (C^T M^{-1} C)^{-1} C^T M^{-1} Y$ , assumes fixed chemical composition of the mineral matrix  $C$ , which is clearly over-simplifying the problem when confronted to clay minerals.

## The total inversion algorithm

A different approach consists in solving for  $n + 1$  implicit equations using the total inversion algorithm proposed by *Tarantola and Valette* [1982]. The sys-

tem can be written as

$$y_i - \sum_{j=1; i=1}^{m; n} c_{ij} x_j = 0 \text{ and } \sum_{j=1}^m x_j - 1 = 0$$

the latter being  $G(X) = 0$  by stating that both inputs and parameters form parts of a single vector  $X$ , affected of a unique covariance matrix  $M$ . The problem is solved by minimizing the quadratic form

$$q^2 = (X - X_0)^T M_0^{-1} (X - X_0)$$

where  $X_0$  is an *a priori* acceptable solution. Using Lagrange multipliers, it is possible to write  $q^2$  such that it is minimum when  $X - X_0 = H\lambda$ , where  $H$  is the Jacobian of  $G(X)$  and  $\lambda$  is the vector of Langrange multipliers, written as  $\lambda = (H^T H)^{-1} H^T (X - X_0)$  [Albarède, 1995].

The global proposed solution includes the uncertainties on all the input parameters [Tarantola and Valette 1982] through an *a priori* covariance matrix  $M_0$  and is given by

$$X_{k+1} = X_0 + M_0 H_k^T (H_k M_0 H_k^T)^{-1} [H_k (X - X_0) - G(X)]$$

These authors give a good approximation of the uncertainties on the solution, expressed as an *a posteriori* covariance matrix  $M$ :

$$M = M_0 - M_0 H^T (H M_0 H^T)^{-1} H M_0$$

The total inversion algorithm is a bayesian approach to the mineral inversion problem based on the fixed point method, since an *a priori* information is needed. Obviously, the information gain will increase with the *a priori* information being more and more suitable to the studied case.

## The clustering process

To facilitate the choice of the *a priori* parameters, we choose to process the data set ( $y_i$ 's), prior to the mineral inversion, through an iterative non-hierarchical cluster analysis. The best number of groups was chosen according to the method of *Ball and Hall*, [1967]. The data set is thus partitioned into 4 groups having similar geochemical properties. The clustering algorithm first picks randomly seed points amongst the data, and clusters the remaining points around them according to their minimum euclidian distance to one seed. It then calculates for each cluster, its centroid and the sum of distances between the points and the centroid. It reallocates the points and move the centroids until the sum of the euclidian

distances within each cluster is minimum. The clustering algorithm gives the composition of each cluster centroid, which help in choosing an *a priori* solution  $X_0$ , the composition of the input minerals and their uncertainties.

## Results and discussion

### Clustering

Results of the iterative non-hierarchical cluster analysis are shown in Figure 1-B. The data set is clearly partitioned into 2 lithological units at about 513 meters below the sea floor (mbsf). Unit 1 is almost exclusively composed of Group 3, with only one occurrence of the other groups. Unit 2 is composed of Group 1 and Group 4, with one occurrence of Group 2. Table 2 gives the composition of the centroids of the 4 clusters. Knowing that the background mineral phase is clay, one can identify the dominant clay mineral in Group 1 to be illite, because of the higher potassium content, whereas it is probably smectite in Group 3, because of the high sodium and magnesium content (that are ones of the exchangable cations in the smectite crystal structure). Group 2 is characterized by a high Mn, Fe and Ca contents, which can be explained by the presence of Mn and Fe oxides and calcite. Group 4 is Ca-rich while very depleted in the other elements, signing the occurrence of calcite as the dominant mineral.

This important qualitative chemical and mineralogical information of the different units individualized in the studied section helps in choosing the right input mineral matrices and a better *a priori* solution for the mineral inversion.

### Mineral inversion

At this stage, we decided to remove  $TiO_2$  and  $P_2O_5$  from the data set because, only present at a minor level in the sediment, they proved to generate great instabilities in the subsequent mineral inversion without bringing much to the solution. However, the two samples constituting of Group 2 present high Mn content (461 and 525 mbsf), and this will have to be included in the discussion when dealing with the results of the mineral inversion for these samples later on. The concentrations of the 7 remaining oxides were recalculated to reach the original measured closure number. The total inversion algorithm was used separately on each individualized cluster with a particular input mineral matrix and *a priori* solution, according to the information provided by the cluster analysis.

To build the input mineral matrices, we averaged several compositions of smectites, illites, and kaolinites [Deer, Howie and Zussman, 1967], corresponding to the environment studied here, and eventually separated them into common, Fe-rich, or Mg-rich types. Table 3 shows the averaged mineral compositions used in the mineral inversion of each cluster that give the most stable results. A slight percentage of Mg in the input calcite was necessary to account for remaining Mg. Despite the fact that the XRD showed no Fe-rich phase such as magnetite, we introduced some in the analysis to take into account the somewhat higher Fe content of the top unit. This was supported by the higher magnetic susceptibility measured on the cores [Shipley, Ogawa, Blum *et al.*, 1995]. The plagioclase is arbitrarily chosen closer to the albite composition (Na-rich).

All the input minerals have uncertainties on each of their oxide content. They are calculated as one standard deviation of all the compositions for this type of mineral found in the literature. The model needs an *a priori* solution  $X_0$ , which is estimated at first, for each sample, from the compositions of the corresponding cluster. An interval of uncertainty around each *a priori* mineral proportion is also set according to the level of confidence of our prior knowledge. At last, the model integrate the uncertainties on the input data, which can be taken from Table 1, or calculated over the studied interval.

Figure 1-D, 1-E and 1-G show the way to track possible sources of uncertainties and errors during the mineral inversion process. At each step of the converging process, the cumulative difference, within each cluster and for each input variable (here oxide weight percents), between  $X$  and  $X_0$  is recorded (Figure 1-G). If the *a priori* solution is inappropriate, it is easier to see which element induces the greatest uncertainty. After the algorithm has reached a minimum (step 10), it calculates the *a posteriori* covariance matrix  $M$  giving the uncertainties on all the input parameters. Figure 1-D gives an example of the uncertainties on the final mineral abundances, for illite and smectite. Figure 1-E gives an example of the uncertainties on the elemental abundances (here  $\text{Al}_2\text{O}_3$ ,  $\text{Fe}_2\text{O}_3$  and  $\text{MgO}$ ) in the illite found to be the best match of the input parameters, the darker the uncertainty, the farther the composition of this illite is from the original illite as far as this particular element is concerned. As an example, around 465 and 530 mbsf, there is a rather large uncertainty on the estimate of  $\text{MgO}$  in the illite. These two particular

intervals are characterized by a high Mn and Fe contents (see Table 2). The analysis was performed again on these two discrete samples with a particular paragenesis that includes a Mn-oxide, needed to account for the high Mn content (Table 2).

Figure 1-C shows the results of the mineral inversion, carried on with the parameters detailed above, on the left. The mineralogy inferred from X-ray diffraction [Fisher *et al.*, 1995] is put on the right as a reference basis (with keeping in mind that it is only semi-quantitative). The resolution is higher in XRD mineralogy column because of the number of samples used, 183 for Fisher and Underwood [1995], and 82 on our study. We can see immediately that the dynamics along the hole is preserved, a top homogeneous unit, with the predominance of smectite, and a bottom perturbed unit [Shipley, Ogawa, Blum *et al.*, 1995], which is the result of the intercalation of several beds of different mineralogy. The boundary between the two lithologic units at 514 mbsf is clearly seen. The overall bulk mineralogy found by both methods is also comparable. The total inversion algorithm sees the occurrence of calcite around 450 mbsf and in the bottom unit (from 525 mbsf). It also finds no more plagioclase in the bottom unit, as the X-ray diffraction does. However, some discrepancies have to be noticed. The mineral inversion finds a higher illite content in the top unit, while the quartz content is lower. We know that volcanogenic rocks that compose the islands of the Lesser Antilles Arc produce abundant smectite by weathering [Capet *et al.*, 1990], and the proximity of this volcanic arc is probably the origin for the presence of plagioclase and magnetite in the top unit.

## Summary

Coupling statistics, through an Iterative Non-hierarchical Clustering, and probability, through the Total Inversion algorithm proved to be a good way to determine mineral abundances, even in sedimentary environments. Despite some discrepancies with the X-ray diffraction results, it is possible to have a good estimate of the total clay content from total inversion of elemental abundances (Figure 1-F). This algorithm provides uncertainties on all the input parameters, which makes it easier to monitor the behaviour of the latter during the converging process. It becomes easier to account for element varying concentrations by use of the classification scheme before performing any mineral inversion. Furthermore, clustering the data

set prior to the mineral inversion allow to deal with a large number of data, such as when using logging techniques. Logging can provide continuous measurement of a wide variety of parameters, from density and porosity to elemental abundances. Being able to calculate a mineralogy on a continuous basis can help quantifying the effects of the fluid circulations in the décollement zone of the North-Barbados accretionary prism, materialized by the two negative density peaks in Figure 1-A.

**Acknowledgments.** We are thankful to J. Vernières for his useful advices and big help with the fundamentals of the total inversion algorithm. Also acknowledged is the help and contribution of H. Mercadier, and Pierre Gaudon at Ecole des Mines d’Alès, for the X-ray fluorescence technique.

## References

- Albarède, F., *Introduction to Geochemical Modeling*, Cambridge University Press, 1995.
- Albarède, F., and A. Provost, Petrological and geochemical mass-balance equations: an algorithm for least-squares fitting and general error analysis, *Computers and Geosciences*, **3**, 309–326, 1977.
- Ball, G.H., and D.J. Hall, A clustering technique for summarizing multivariate data, *Behavioral Science*, **12**, 153–155, 1967.
- Capet, X., H. Chamley, C. Beck, and T. Holtzapffel, Clay mineralogy of Sites 671 and 672, Barbados Ridge accretionary complex and atlantic abyssal plain: paleoenvironmental and diagenetic implications, *In Moore, J.C., Mascle, A., et al., Proc. ODP, Sci. Results*, **110**, 85–96, 1990.
- Deer, W.A., R.A. Howie, and M.A. Zussman, *Rock-forming minerals*, vol 1–5, Longmans, Green and Coltd (Eds.), 1967.
- Fisher, A.T., M.B. Underwood, Calibration of an X-ray diffraction method to determine relative mineral abundances in bulk powders using matrix singular value decomposition: a test from the Barbados accretionary complex, *In Shipley, T.H., Ogawa, Y., Blum, P. and J.M. Bahr, Proc. ODP, Init. Repts*, **156**, 29–37, 1995.
- Harvey, P.K., J.F. Bristow, and M.A. Lovell, Mineral transforms and downhole geochemical measurements, *Scientific Drilling*, **1**, 163–176, 1990.
- Kohonen, T., *Self-organising maps*, Springer-Verlag, New-York, 1995.
- Lofts, Jeremy C., Intergrated geochemical–geophysical studies of sedimentary reservoir rocks, PhD Thesis, 156 pp., University of Leicester, UK, February 1993.
- Shipboard Scientific Party, Site 948. *In Shipley, T.H., Ogawa, Y., Blum, P. and J.M. Bahr, Proc. ODP, Init. Repts*, **156**, 87–192, 1995.
- Tarantola, A., and B. Valette, Generalized nonlinear inverse problems solved using the least squares criterion, *Rev. Geophys. and Space Phys.*, **20**, 2, 219–232, 1982.
- A. Rabaute, L. Briquieu, M. Ramadan, UMR 5567–CNRS, ISTEEM, cc 066, Université Montpellier II, 34095 Montpellier Cedex 5, France. (e-mail: rabaute@dstu.univ-montp2.fr)

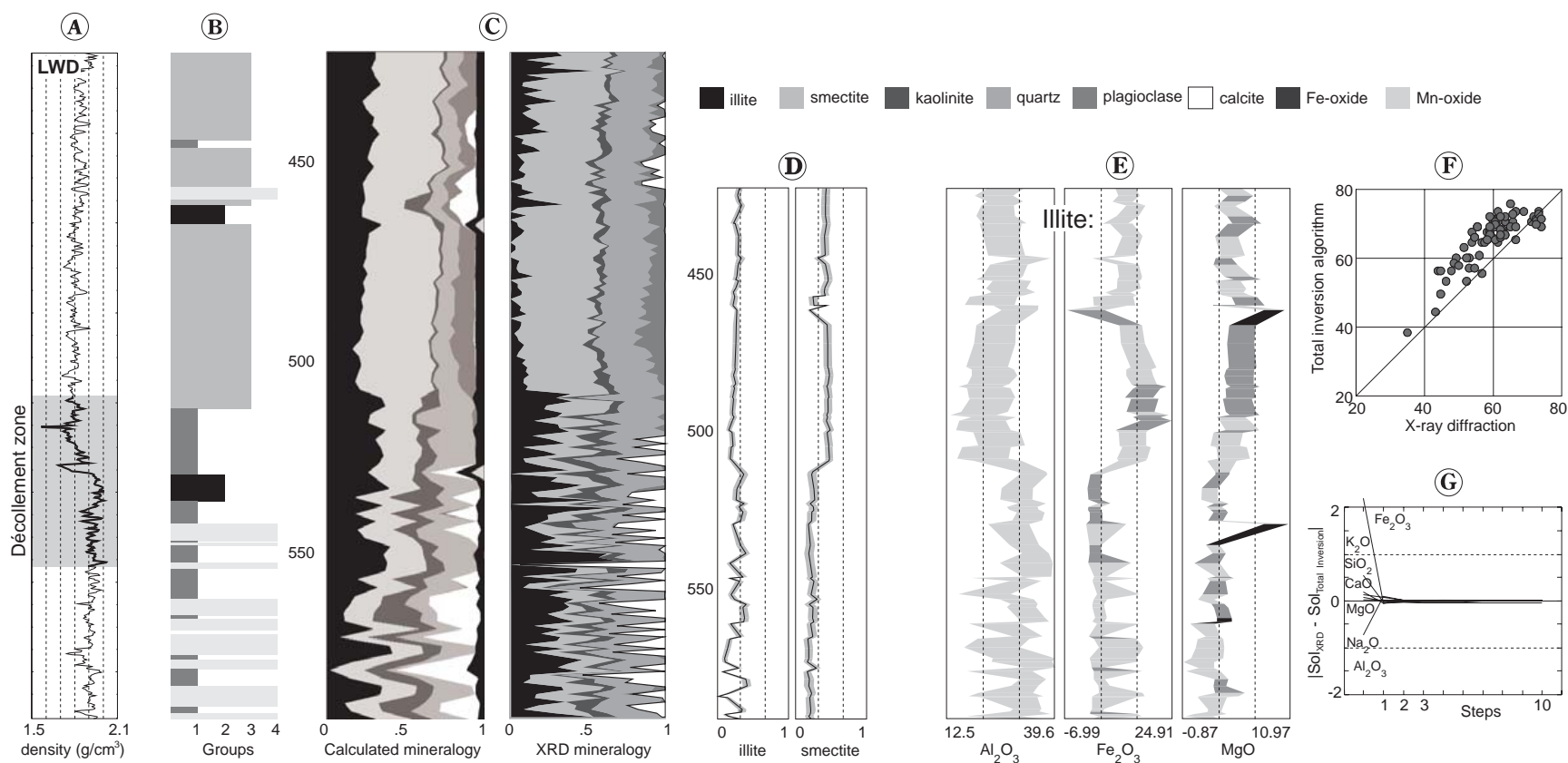


Figure 1: **A.** Example of Logging-While-Drilling (LWD) density measured downhole through the décollement zone of the North-Barbados accretionary prism. The two negative peaks around 525 mbsf are related to fluid circulations. **B.** Results of the Iterative Non-hierarchical Cluster Analysis. **C.** Results of the mineral inversion using the Total Inversion algorithm (left). Mineralogy from quantitative X-ray diffraction performed during Leg 156 (right). **D.** Example of uncertainties calculated on the output mineral abundances. **E.** Example of uncertainties calculated on the output mineral compositions. **F.** Comparison of the total clay volume calculated using mineral inversion and the one found by the quantitative X-ray diffraction method. **G.** Monitoring of the discrepancy between calculated and input solutions throughout the total inversion converging process.

**Table 1.** Precision Obtained With X-ray Fluorescence Analysis for the 10 Common Major Elements.

Oxide	Precision	Oxide	Precision
SiO <sub>2</sub>	0.5%	Al <sub>2</sub> O <sub>3</sub>	0.5%
Fe <sub>2</sub> O <sub>3</sub>	0.5%	MgO	1.0%
MnO	1.0%	CaO	0.6%
Na <sub>2</sub> O	5.0%	K <sub>2</sub> O	0.5%
TiO <sub>2</sub>	1.0%	P <sub>2</sub> O <sub>5</sub>	6.0%

**Table 2.** Composition of the Clusters Centroids after the Iterative Non-hierarchical Cluster Analysis.

Oxide	Centroid composition			
	Group 1	Group 2	Group 3	Group 4
SiO <sub>2</sub>	63.0	47.3	59.9	44.2
Al <sub>2</sub> O <sub>3</sub>	21.2	17.6	20.7	16.4
Fe <sub>2</sub> O <sub>3</sub>	7.7	9.9	9.4	6.3
MgO	2.3	3.1	3.5	1.9
MnO	0.1	7.5	0.3	0.6
CaO	1.7	9.1	2.9	25.1
Na <sub>2</sub> O	0.2	0.2	1.1	0.1
K <sub>2</sub> O	2.7	1.7	2.1	1.0
TiO <sub>2</sub>	0.9	0.7	0.9	0.6
P <sub>2</sub> O <sub>5</sub>	0.1	0.3	0.2	0.3
# samples	26	2	38	16



**Table 3.** Average mineral compositions chosen as input mineral matrices for total inversion algorithm. Numbers between parenthesis are the uncertainty on the element and are calculated as one standard deviation around the average composition of all the minerals found in the literature that belongs to the same type. Magnetite is 100%  $\text{Fe}_2\text{O}_3$ , and Mn-oxide is 100% Mn, both with the same uncertainties as quartz.

Elt	Smectites		Illites		Kaolinite	Quartz	Plagio	Calcite
	Group 3	Groups 1,2,4	Group 1	Groups 2,3,4				
$\text{SiO}_2$	66.7(2.6)	65.4(2.5)	53.6(4.8)	53.2(4.5)	53.3(0.5)	100(0.001)	63.5(0.01)	0(0.001)
$\text{Al}_2\text{O}_3$	25.8(2.7)	23.2(2.7)	30.9(5.3)	27.1(9.6)	45.4(0.7)	0(0.001)	22.4(0.01)	0(0.001)
$\text{Fe}_2\text{O}_3$	1.1(0.5)	3.4(2.7)	4.7(2.2)	8.3(8.0)	0.4(0.4)	0(0.001)	0.5(0.01)	0(0.001)
MgO	4.2(1.1)	4.1(1.1)	1.8(1.1)	2.9(2.2)	0.1(0.1)	0(0.001)	0.1(0.001)	3.0(0.5)
CaO	2.1(0.9)	2.9(0.9)	0.7(0.7)	0.7(0.6)	0.2(0.1)	0(0.001)	3.3(0.5)	97.0(0.5)
$\text{Na}_2\text{O}$	0.1(0.1)	0.3(0.3)	0.6(0.5)	0.5(0.5)	0.1(0.1)	0(0.001)	9.1(1)	0(0.001)
$\text{K}_2\text{O}$	0.05(0.01)	0.7(0.7)	7.6(1.8)	7.1(1.9)	0.3(0.2)	0(0.001)	0.9(0.05)	0(0.001)

7.2.2 Utilisation d'un Perceptron Multi-Couches pour reconstruire une minéralogie en continu à partir de mesures LWD : exemple du puits 948A

**Article 4 :**

en préparation

Inferring mineralogy with neural network classifier : Example using  
Logging-While-Drilling data from Hole 948A, ODP Leg 156

**Alain Rabaute**

# Inferring mineralogy with neural network classifier: Example using Logging-While-Drilling data from Hole 948A, ODP Leg 156

Alain Rabaute\*

en préparation

## Abstract

A neural network classifier was used on Logging-While-Drilling data obtained in Hole 948A during Ocean Drilling Program Leg 156. This improved logging technique, used for the first time in ODP, gives better data because of better hole and measurements conditions. The network is a classical backpropagation neural network, with an exponential sigmoid activation function. The training was done separately on two data sets, made with corresponding LWD intervals chosen after fuzzy *c-means* classification of 1) X-ray diffraction data and 2) physical properties measurements (grain density, magnetic susceptibility and natural gamma radiation), in order to find the weights corresponding to a minimal quadratic error between calculated and desired outputs. These data were measured on the cores from Hole 948C, drilled at the same location as Hole 948A. Better results are obtained when the network is trained after X-ray diffraction data. A continuous mineralogy is provided, showing a good agreement with the mineral concentrations calculated with a semi-quantitative x-ray diffraction method.

## 1 Introduction

**Logging-While-Drilling technique** ODP Leg 156 principal objective was to assess the existence of pore-water overpressure in décollement zone of the Northern Barbados accretionary prism. In such an active environment, using wireline logging techniques is quite illusive, and for the first time in ODP, the improved technique of Logging-While-Drilling was carried out. The LWD logging string is composed of two logging tools:

- the CDR or Compensated Dual Resistivity tool, which is an electromagnetic propagation and spectral gamma-ray tool built into a drill collar. It is in many ways similar to dual induction tools, responding to conductivity rather than resistivity, and providing two depths of investigation. But whereas it has a better vertical resolution than dual induction tools, it has a shallower depth of investigation. A 2-MHz electromagnetic wave is sent into the formation, and the tool measures the phase shift (PSR) and the attenuation of the wave (ATR) between two receivers. These quantities are then converted into two independent

---

<sup>0</sup>cc066 UMR 5567-CNRS, GGP, ISTEEM, Université Montpellier II, Place E. Bataillon, 34095 Montpellier Cedex 5, France

resistivities that provide the two depths of investigation. The phase shift PSR is a shallow resistivity, the attenuation ATR is the deep resistivity. An electrical hole diameter is computed from the CDR data and is used as an input to hole size corrections.

- the CDN or Compensated Density Neutron tool is similar in its principles of measurement to those of corresponding wireline tools. For the neutron porosity measurement (NPHI), fast neutrons are emitted from a 7.5-curie (Ci) americium-beryllium (Am-Be) source. The density section of the tool uses a 1.7-Ci  $^{137}\text{Cs}$  cesium (Cs) gamma-ray source in conjunction with two scintillation detectors to provide a borehole compensated density measurement.

Because the shaly composition of the considered geological formation, the neutron porosity was not believed to give accurate results and was not used in the analysis.

**X-ray diffraction and fluorescence data** During ODP Leg 156, Fisher and Underwood (1995) implemented a new method using matrix Singular Value Decomposition and spectral analysis using Fourier transform for estimating relative mineral abundances in bulk powders. They determined calcite, quartz, total clay, and plagioclase content, as well as were able to measure the major clay minerals (smectite, illite, kaolinite) contents from 176 samples picked at regular intervals from the cores in Hole 948C. The samples were carefully extracted in zones where no major variation in color, grain size, or texture was visible.

X-ray fluorescence measurements were carried out at the Université de Montpellier II on the same samples used by Fisher and Underwood (1995). Concentrations in Si, Ca, Al, Fe, Mg, Mn, P, and Na were measured on fusion beads with an average uncertainty of about 1%. Only concentrations in Al are used in this study.

**Multi-Sensor Track data** Routinely onboard the JOIDES *Resolution*, a number of physical measurements are made on the cores immediately after recovery. The petrophysical properties GRAPE (grain density), NGR (natural gamma radiation) and MAGSUS (magnetic susceptibility) are measured on a continuous basis on the whole core using a special chain of devices called the Multi-Sensor Track, composed of a natural gamma radiation detector (scintillation crystal), a susceptometer, a grain density measurement device using a  $^{137}\text{Cs}$  gamma-ray activation source, and a P-wave velocity measurement device. We used the three first ones and discarded the P-wave velocity data because of some flaky results.

**Validity of the total clay concentration data** In a recent paper, Herron and Herron (1996) prove that Al is an excellent estimator of the total clay content. In Figure 1,  $\text{Al}_2\text{O}_3$  concentration from X-ray fluorescence analysis is plotted against the total clay content after Fisher and Underwood (1995). The good agreement indicates that X-ray diffraction results can be used with confidence as a reference basis in this work.

## 2 Composition of the training data sets

We decided to prepare two training data sets from what we will call the *Physical properties*, that is, **NGR**, **MAGSUS**, and **GRAPE**, measured on the whole cores onboard the Joides Resolution, and from the mineral concentrations calculated from X-ray diffraction results on core samples by Fisher and Underwood (1995).

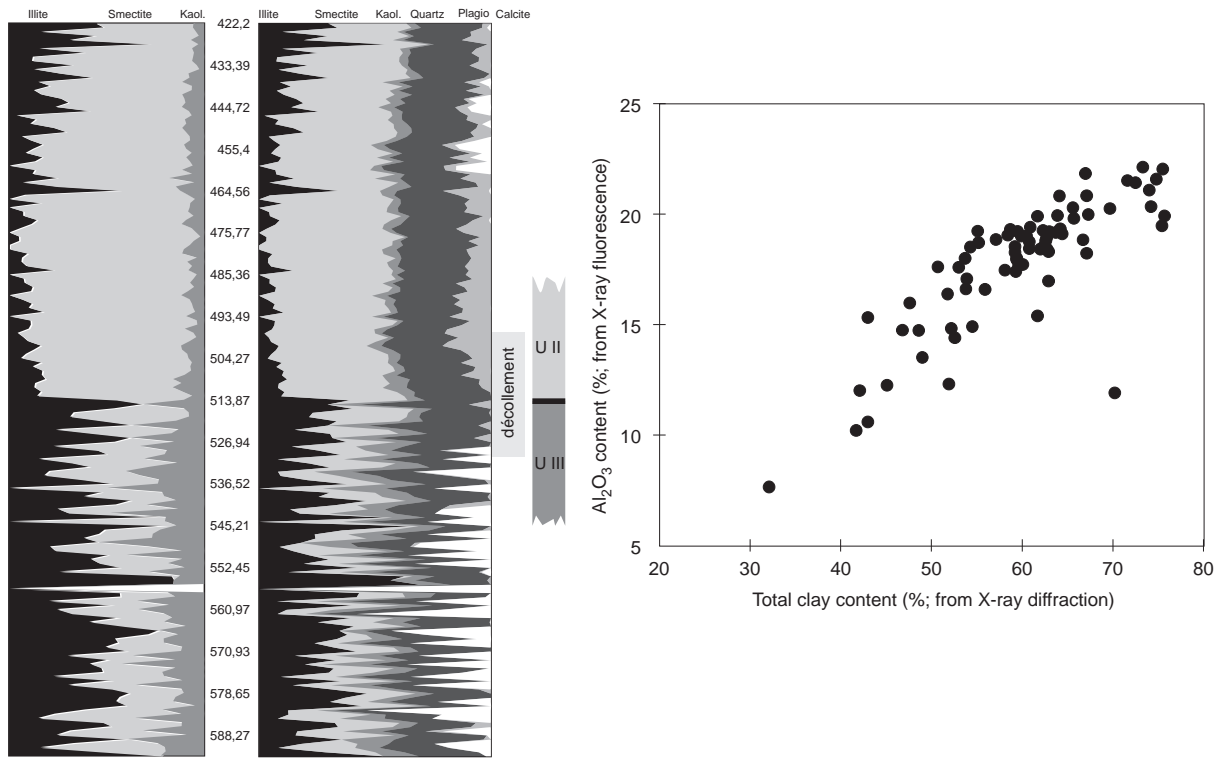


Figure 1: Left: Clay mineralogy and total mineralogy of the cored section in Hole 948C. Right: Comparison between total clay content and aluminium content around the décollement zone.

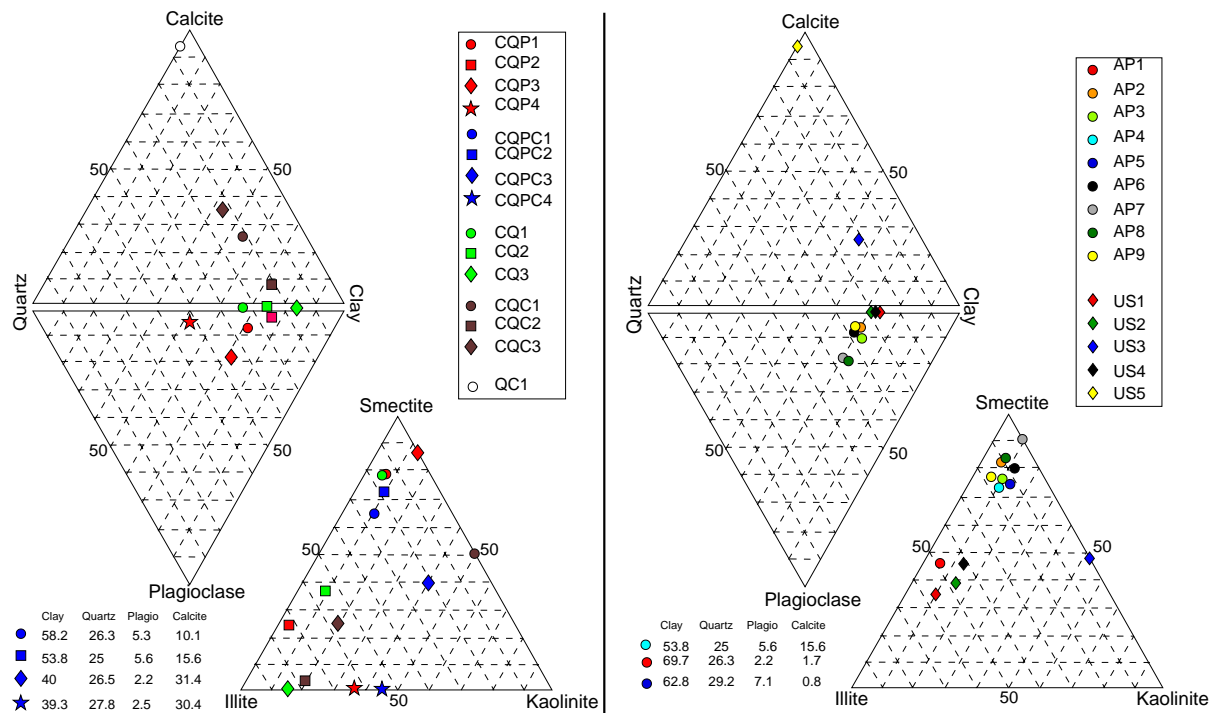


Figure 2: Compositions of the 15 groups made after the X-ray diffraction results and of the 14 groups made after the fuzzy clustering of LWD physical properties (resistivity, density, photoelectric effect, total gamma ray).

Code	CQP1	CQP2	CQP3	CQP4	CQPC1	CQPC2	CQPC3	CQPC4
section	2X-6	2X-4	9X-1	12X-3	4X-4	5X-4	13X-6	15X-2
interval (cm)	52–54	116–118	94–96	107–109	12–14	47–49	19–21	104–106
depth (mbsf)	428.82	426.46	489.34	521.47	444.72	454.77	534.49	547.94
smectite	50.7	18.5	48.7	0.1	37.6	39.0	15.7	0.1
illite	10.0	54.6	0.0	29.7	13.8	9.2	8.2	18.0
kaolinite	4.0	2.3	6.5	17.1	6.7	5.6	16.0	21.3
total clay	64.7	75.4	55.2	46.8	58.2	53.8	40.0	39.3
quartz	28.9	23.1	28.1	47.7	26.3	25.0	26.5	27.8
plagio	6.4	1.5	16.7	5.5	5.3	5.6	2.2	2.5
calcite	0.0	0.0	0.0	0.0	10.1	15.6	31.4	30.4
Code	CQ1	CQ2	CQ3	CQC1	CQC2	CQC3	QC1	
section	11X-4	12X-1	16X-2	14X-1	14X-5	17X-4	16X-2	
interval (cm)	68–70	79–81	24–26	146–148	61–63	21–23	91–93	
depth (mbsf)	512.88	518.19	556.54	537.56	542.71	568.81	557.21	
smectite	53.1	26.6	0.0	27.5	2.4	10.1	0.1	
illite	9.6	40.3	69.5	0.0	55.0	23.9	0.0	
kaolinite	4.2	7.8	13.1	27.0	14.1	8.9	0.0	
total clay	66.9	74.6	82.6	54.5	71.5	43.0	0.0	
quartz	33.1	25.4	7.4	21.0	20.9	22.2	5.9	
plagio	0.1	0.1	0.1	0.0	0.1	0.0	0.0	
calcite	0.0	0.0	0.0	24.5	7.4	34.9	94.1	

Table 1: Mineral composition of the 15 groups chosen after the X-ray diffraction results.

## 2.1 XRD mineral concentrations

Figure 1 shows the mineralogy (total clay + quartz + calcite + plagioclase) along the borehole in the cored section (420 to 590meters below sea floor; mbsf), and the relative amounts of clay minerals (smectite + illite + kaolinite) (Fisher and Underwood 1995). The shaded rectangle on the right side represents the dcollement zone as described on cores, and the boundary between Units II and III, defined by sedimentological studies onboard is showed. 15 intervals ( $\equiv$  X-ray diffraction sample locations) were picked on this log, as representative of the diverse mineral assemblages present in the entire cored interval. Each depth was reported in the LWD data set, where over a total interval of 75 cm (X-ray diffraction sample depth +/- 30 cm), LWD variables were averaged to form the first training data set,  $TS_1$ . The mineral composition of these averaged intervals are seen in table 1, and plotted on ternary and diamond diagrams in Figure 2. As a rule, the letters ‘‘CQPC’’ indicate the presence of *clay*, *quartz*, *plagioclase* and *calcite* in the assemblage. If clay is present, the presence of *smectite*, *illite* or *kaolinite* is indicated by the letters S, I or K (see Figure 4). An idea of the proportion of the clay mineral in the total clay content relative to the other clay minerals present in the assemblage is given by a ‘+’ or ‘-’ sign (see Figure 4).

## 2.2 Physical properties from Multi-Sensor Track (MST)

Figure 3 shows the results of a cluster analysis performed on the MST variables chosen, GRAPE, NGR, and MAGSUS. I used the *fuzzy c-means* algorithm, which yields to a probability for a sample to belong to a particular group. This is particularly useful in the mixed clay environment present in Hole 948A, which makes it very difficult to clearly classify a sample on a binary way. On Figure 3, 4 units can be clearly distinguish, according to the physical properties from MST:

- the first one, ending around 460 mbsf, is characterized by Groups 2 and 6.
- the second one, between 460 and 500 mbsf, is represented by Groups 1 and 3.

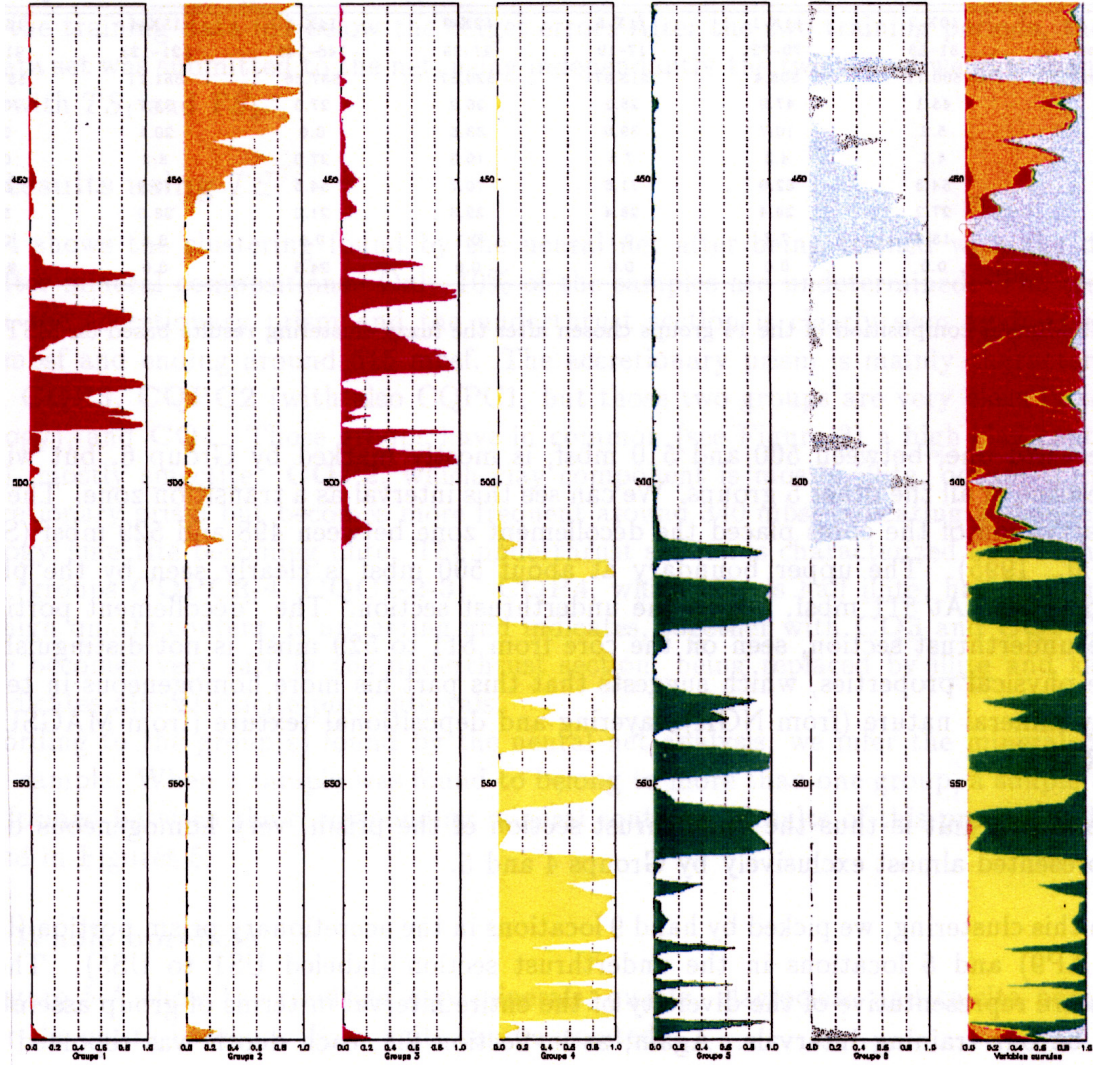


Figure 3: Fuzzy clustering on the MST petrophysical properties, GRAPE, NGR and MAGSUS.

Code	AP1	AP2	AP3	AP4	AP5	AP6	AP7
section	2X-1	3X-2	5X-1	5X-4	7X-4	8X-2	9X-4
interval (cm)	146–148	16–18	63–65	47–49	10–12	19–21	59–61
depth (mbsf)	422.26	432.16	450.43	454.77	473.6	480.39	493.49
smectite	32.0	52.8	47.1	39.0	48.0	48.0	48.7
illite	33.7	7.5	8.4	9.2	8.9	5.7	0
kaolinite	4.1	4.0	6.1	5.6	5.9	6.9	5.0
total clay	69.7	64.4	61.7	53.8	62.8	60.6	53.7
quartz	26.3	30.3	27.7	25.0	29.2	29.7	28.4
plagio	2.2	5.3	10.6	5.6	7.1	9.7	17.9
calcite	1.7	0.0	0.0	15.6	0.8	0.0	0.0
Code	AP8	AP9	US1	US2	US3	US4	US5
section	10X-2	11X-1	11X-5	12X-3	14X-1	15X-4	16X-2
interval (cm)	51–53	70–72	17–19	27–29	146–148	121–123	91–93
depth (mbsf)	500.11	508.4	513.87	520.67	537.56	551.11	557.21
smectite	45.1	47.9	25.1	26.9	27.5	33.7	0.1
illite	5.2	10.7	39.0	33.5	0.0	20.1	0.0
kaolinite	4.1	4.3	7.5	10.3	27.0	8.2	0.0
total clay	54.3	62.9	71.6	70.7	54.5	72.0	0.0
quartz	27.2	29.4	28.4	29.3	21.0	28.0	5.9
plagio	18.2	7.7	0.1	0.1	0.0	0.1	0.0
calcite	0.0	0.0	0.0	0.0	24.5	0.0	94.1

Table 2: Mineral composition of the 14 groups chosen after the fuzzy clustering results based on MST data.

- the third one, between 500 and 510 mbsf, is mostly marked by Group 6, but with the presence of all the other 5 groups. We can see this interval as a transition zone. The visual observation of the cores placed the décollement zone between 498 and 529 mbsf (Shipley et al. 1995). The upper boundary at about 500 mbsf is clearly seen by the physical properties. At 511 mbsf, begins the underthrust section. The “décollement portion” of the underthrust section, seen on the core from 511 to 529 mbsf, is not distinguished by the physical properties, which suggests that this part is more homogeneous in terms of clay mineral nature (from NGR), layering and depositional texture (from MAGSUS and GRAPE).
- the fourth unit is thus the underthrust section of the prism, very homogeneous because represented almost exclusively by Groups 4 and 5.

From this clustering, we picked by hand 9 locations in the accretionary prism portion (labeled AP1 to AP9) and 5 locations in the underthrust section (labeled US1 to US5). These 14 locations are representative of the diversity of the entire interval in terms of group assemblages, and will be our training intervals. Again, as in Section 2.1, each depth was reported in the LWD data set, and the LWD variables were averaged over a 75 cm-interval (MST depth +/- 30 cm) to compose the second training data set,  $TS_2$ . In table 2 are reported the corresponding mineralogical compositions for each location (according to the closest XRD analyzed sample, see Fisher and Underwood (1995)). In Figure 2, those compositions are plotted in ternary and diamond diagrams. One can notice that the training intervals AP1 to US5 are far less scattered than the training set CQP1 . . . QC1 chosen from XRD results. The cored interval at Site 948 can therefore be considered more complex in terms of mineralogy and chemistry than it is in terms of physical properties, because the various group assemblages distinguished after the cluster analysis in Figure 3 show small differences in mineralogy (see Figure 2).



### 3 Neural Network analysis

We used a classical backpropagation neural network with a sigmoid transfer function. Before using the network on the entire LWD data set, one has to train it on the corresponding training data set (see Section 2). We took care of not over-training the network, thus 100% of the training data below the target error was not required for convergence. The learning rate was set to 0.4 and the momentum to 0.6, and be allowed to vary during the training process, in order to avoid local minima. Convergence was attained when average error was below 0.05 (5% relative uncertainty in attributing all training data samples to their corresponding group), and about 85% of the training data fell below the target error. After the two training periods, the entire LWD data set was submitted to the net, using independently the two sets of weights found when trained with  $TS_1$  and  $TS_2$ .

#### 3.1 Results using $TS_1$

Figure 4 shows the clustering found by the neural net after being trained with the data set made after mineral compositions. Only 10% of the samples are undetermined. The transition between the accretionary prism and the underthrust section is clearly seen again, beginning at 490 mbsf and ending around 515 mbsf. The accretionary prism is mainly characterized by **CQP1**, **CQP3**, **CQPC2** (with also **CQPC1**, but those two groups are very close in terms of mineralogy), and **CQ1**. Those groups have in common (see Figure 3) a high clay content (54 to 66%), mostly smectite. **CQP2**, which clay component is mostly illite, occurs throughout the accretionary prism but becomes more frequent around 340 mbsf indicating a change in clay mineralogy, smectite becoming illite. The underthrust section is characterized by a high calcite content (groups **CQPC3-4**, **CQC1-2-3**). **CQP4**, which clay is half illite, half kaolinite, and with a high quartz content, is appearing and indicates, together with **CQ3** and **CQPC4**, that smectite becomes very rare in the underthrust section, being replaced by illite and kaolinite. **QC1** indicates almost calcite-pure intervals.

According to the group(s) found by the neural net analysis, we infer the mineral contents for each sample. When a sample was found to belong to more than one group, a simple average was performed between the corresponding mineral contents. Results of this processing for  $TS_1$  is showed in Figures 5.

##### 3.1.1 Whole borehole

- Figure 5-A show the resulting clay minerals, quartz, plagioclase and calcite percentages for the entire borehole. Four units can be isolated:

from 0 to 320 mbsf, *Unit I* is quite stable and is composed on average by 60% clay, 25% quartz, 5% plagioclase and some calcite. Some alternance between smectite-rich units and illite-rich units can be noted.

then a increase in clay content and a decrease in plagioclase occur between 320 and 400 mbsf (*Unit II*). Meanwhile, the quartz content shows higher fluctuations, as clay content does also.

*Unit III* ends at 480 mbsf and is characterized by 60% total clay content, an increase in plagioclase and quartz contents.

*Unit IV* runs between 480 until the end of the log. It can be further subdivided in three parts: *subunit IV-1* from 480 to 510 mbsf, with no calcite and a slight increase in clay and quartz content. *subunit IV-2*, which shows a lot of fluctuations in mineral contents, a

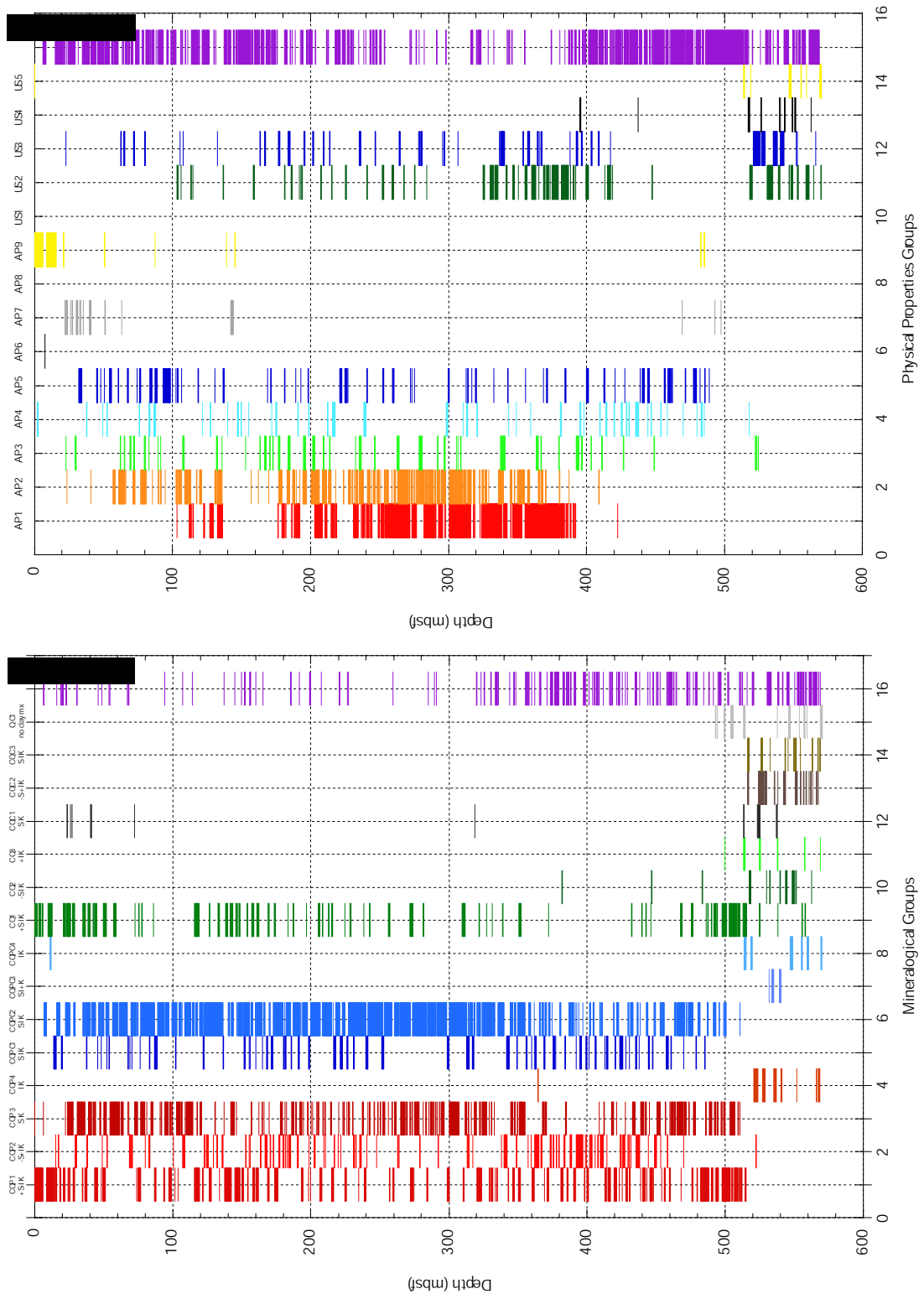


Figure 4: Results given by the neural network for the two training data sets. The right-hand columns (purple) indicate the samples that the neural network was unable to allocate.

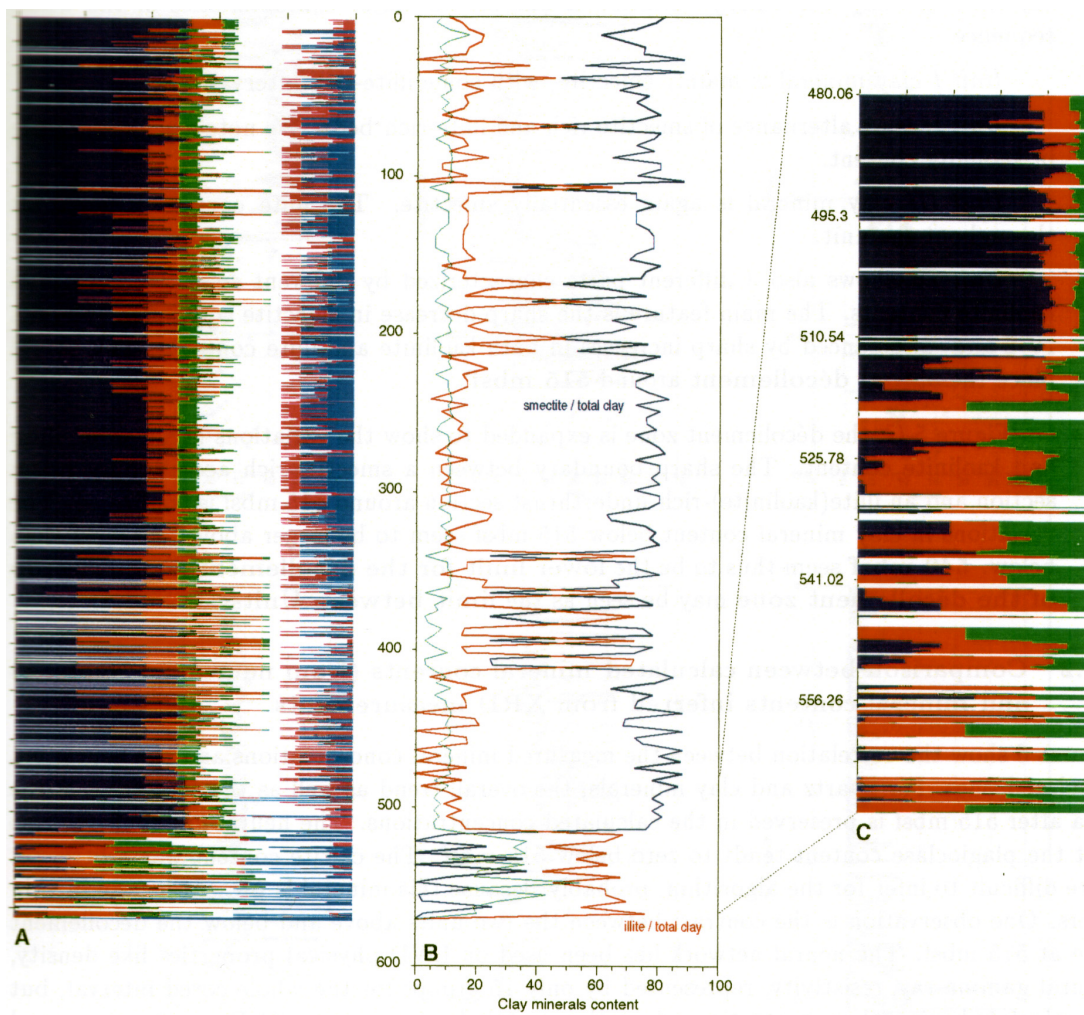


Figure 5: Mineralogy inferred from the neural network results for the first training data set  $TS_1$ . **A.** cumulative mineral contents. **B.** Clay minerals contents with respect to total clay content, for the whole borehole. **C.** Clay minerals contents with respect to total clay content, focused on the dcollement zone (cored section).

decrease in clay content and runs until 540 mbsf. *subunit IV-3*, with almost no plagioclase and an increase in calcite content.

- Figure 5-B brings some more information by displaying the relative amount of clay minerals with respect to total clay. The data have been interpolated to give a curve easier to interpret. The four units described above are easily seen on this figure, and one can even say that the clay mineralogy variation is one of the major characteristics of the whole sequence.

*Unit I* clay mineral is mainly smectite, with some illite-rich intervals.

*Unit II* is an alternance of smectite-rich and illite-rich beds. We note a slight decrease in kaolinite content.

*Unit III* clay mineral is again essentially smectite. The illite content varies a lot throughout the unit.

*Unit IV* shows also 3 different parts characterized by different smectite, illite, and kaolinite contents. The main feature is the sharp decrease in smectite content in subunits IV-2 and -3, balanced by sharp increases in both kaolinite and illite contents. This would place the **actual décollement** around **515 mbsf**.

- On Figure 5-C, the décollement zone is expanded to show the variations in smectite, illite and kaolinite content. The sharp boundary between a smectite-rich accretionary prism section and an illite(kaolinite)-rich underthrust section around 515 mbsf is confirmed. The variations in clay mineral content below 515 mbsf seem to be larger above 540 mbsf than below. **540 mbsf** seem thus to be the **lower limit for the décollement zone**. The **top of the décollement zone** may be seen as the **limit between Units III and IV**.

### 3.1.2 Comparison between calculated mineral contents (from neural net analysis) and mineral contents inferred from XRD measurements

Figures 6 show the correlation between the measured mineral concentrations and the neural net calculated ones. For quartz and clay minerals, the overall trend as well as the scattering of the data after 515 mbsf is preserved in the calculated concentrations. The neural net confirms also that the plagioclase content tends to zero below 515 mbsf. The calcite content was a little bit more difficult to infer for the algorithm, probably because this mineral is present locally in thin layers. One observation is the contrast between the two units above and below the décollement zone at 515 mbsf. The neural network has been used on LWD physical properties like density, natural gamma-ray, resistivity, represented by only 15 groups for the whole cored interval, but the calculated mineral concentrations show the same behaviour as the XRD measured mineral contents (see Figure 7). This indicates that mineralogical and chemical variations influence such physical properties, that can of course be locally the result of tectonic features (thrust at 130 mbsf, compaction trend visible on the resistivity profile).

## 3.2 Results using $TS_2$

The same analysis was carried out after the network has been trained with the second training data set, composed with LWD intervals found after MST data cluster analysis (see Section 2.2). Figure 4 shows how the whole logged interval has been partitioned into the 14 facies. We notice a very high level of uncertainty as the neural net could not allocate 41% of the data. Another observation is that two groups (AP8 and US1) have not been used at all. We can explain that

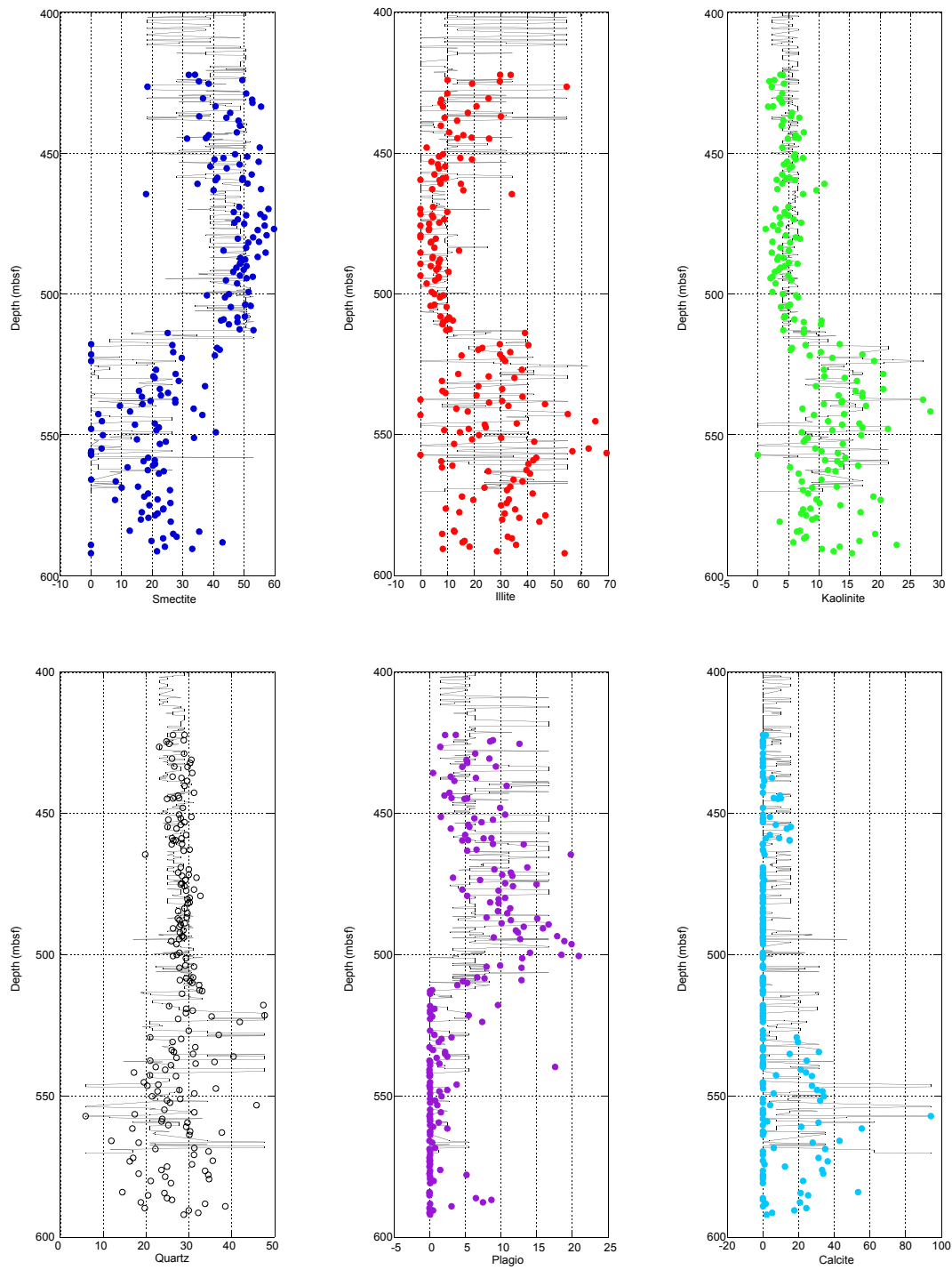


Figure 6: Comparison between the mineral contents found by the neural network and the X-ray diffraction data..

by the difference in groups mineralogical composition. The 14 facies individualized after the MST physical properties span a larger range than those found after the X-ray diffraction data (Figure 2). The variations in mineralogy that is visible on the XRD results does not show up in the physical properties measured on the cores, at least around the décollement zone. Thus the groups from  $TS_2$  are not representative of the entire logged interval.

### 3.2.1 Whole borehole

Being made after physical properties, the training set has strongly influenced the mineralogy output: the compaction trend is visible, with a sharp decrease in clay content and increases in plagioclase and calcite contents that mark the thrust at 130 mbsf. The limit at 400 mbsf is still visible, as well as the top of the décollement zone near 490 mbsf. Figure 7 shows the same separation into different units as in Figure 5, simply the limits between those units are fuzzier, mostly because of high undetermination. Around the décollement zone, the clay minerals contents show the same pattern as previously: sharp limit at 515 mbsf between the smectite-rich accretionary prism section above and the illite-rich underthrust section below. The décollement portion of this section ends at 540 mbsf with decreasing clay minerals variations.

## 4 Fuzzy Clustering of the entire LWD data set

Figure 8 shows the results of a fuzzy clustering of the entire LWD data set for Hole 948A. The limit between the accretionary prism section and the underthrust section is very sharp at 515 mbsf, marked by strong proportion of Group 1, and disappearance of Groups 3, 4 and 5. Group 1 seems to characterize the tectonic events (thrust at 130 mbsf), so we can guess that something is happening around 380 and 420 mbsf. The rest of the accretionary prism section shows repetitive alternance of the same units, presumably having the same physical properties.

## References

- Fisher, A. and M. Underwood (1995). Calibration of an x-ray diffraction method to determine relative mineral abundances in bulk powders using matrix singular value decomposition: a test from the barbados accretionary complex. In *Proc. of ODP, Initial Reports*, Volume 156, pp. 29–37.
- Herron, S. and M. Herron (1996, 16-19 June). Quantitative lithology: an application for open and cased hole spectroscopy. In *Transactions of SPWLA Thirty-seventh Annual Logging Symposium*, Number Paper E, New Orleans.
- Shipley, T., Y. Ogawa, and P. Blum (1995). Site 948, Hole 948C, Sedimentology chapter. In *Proc. ODP, Initial Reports*, Volume 156.

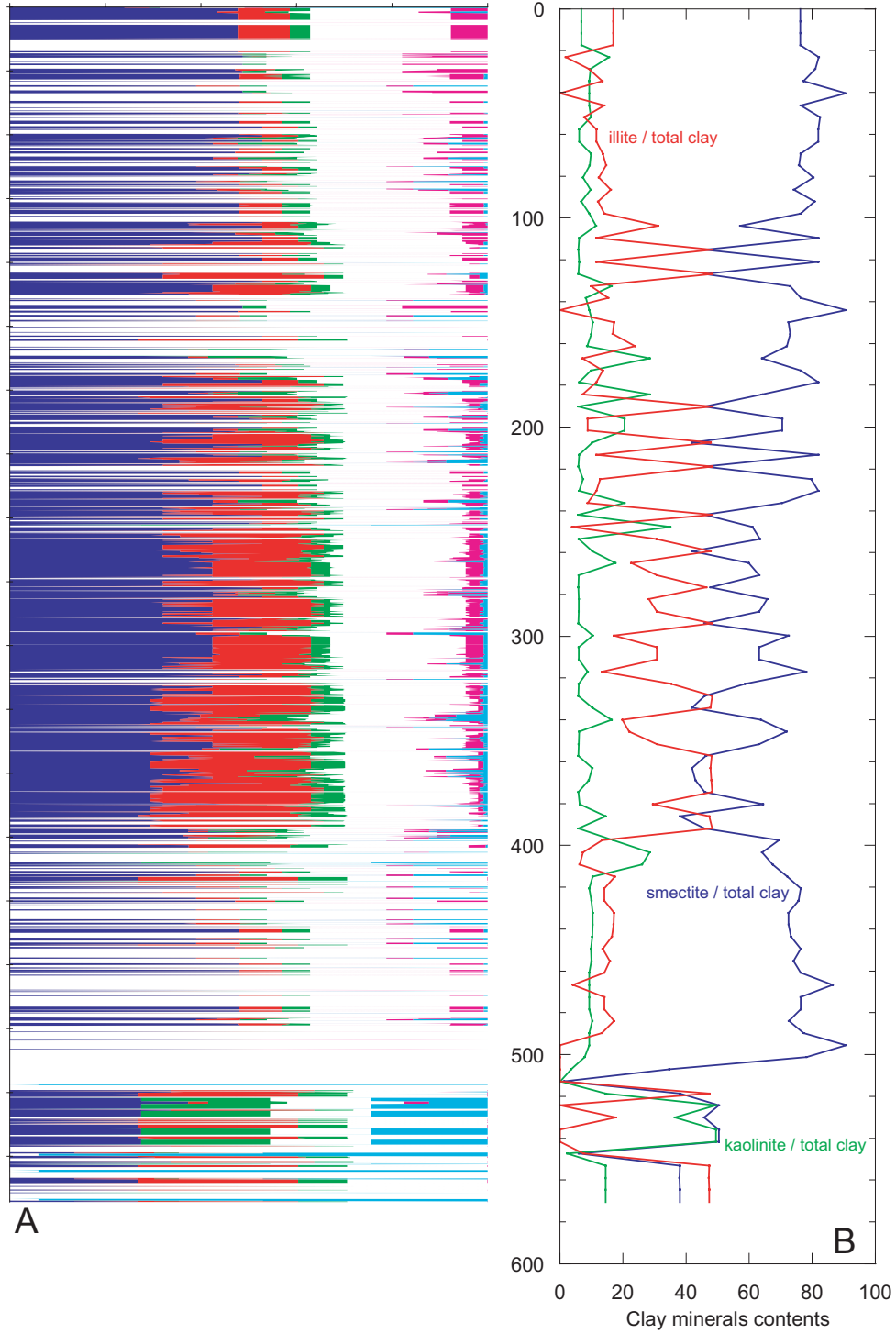


Figure 7: Mineralogy inferred from the neural network results for the second training data set  $TS_2$ . **A.** cumulative mineral contents. White gaps spanning the entire width are undetermined data. **B.** Clay minerals contents with respect to total clay content, for the whole borehole.

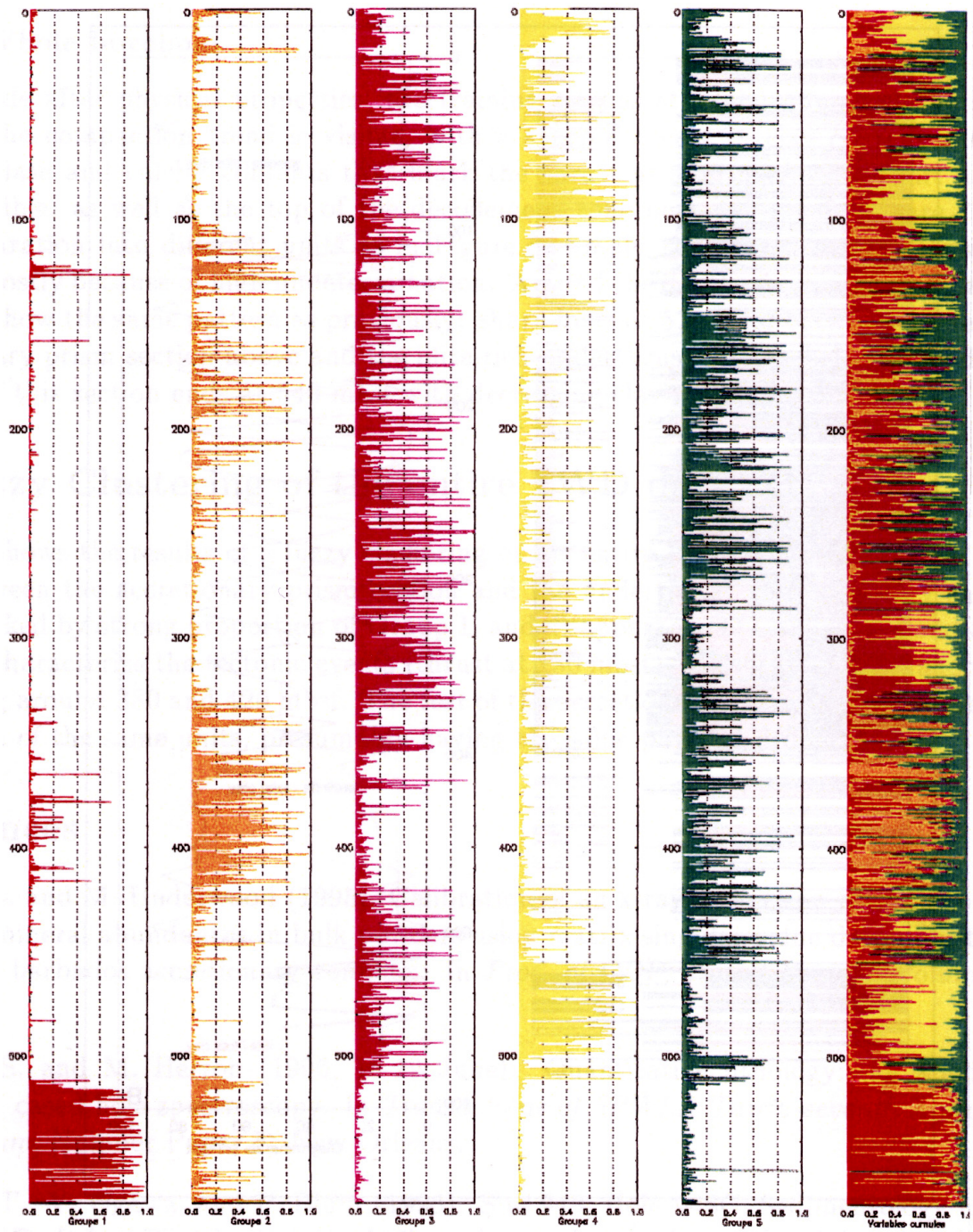


Figure 8: Fuzzy clustering on the LWD variables, for the whole borehole.



## 7.3 Le puits 967E, Mont sous-marin Erathosthenes, Leg ODP 160

### 7.3.1 Introduction

Le Leg 160 d'Ocean Drilling Program s'est déroulé du 12 Mars au 3 Mai 1995 en mer Méditerranée. Un de ses objectifs était de tester l'hypothèse controversée selon laquelle le Mont sous-marin Eratosthènes est un morceau de croûte continentale qui entre en collision avec la plaque eurasiennne, représentée par la marge active de Chypre (Robertson, 1990; Kempler, 1993). Cette collision aurait été le moteur du soulèvement de l'île de Chypre, entraînant celui de l'ophiolite de Troodos. Les données sismiques de reconnaissance collectées en vue du leg 160 durant la campagne UNESCO-ESF du N/O Gelendzhik (Limonov and Woodside, 1994) ont fourni la base d'une telle interprétation que le forage de 4 puits le long du Mont Eratosthènes permettait de tester en situant dans le temps le début de la collision.

### 7.3.2 Géologie du site

Le mont Eratosthènes est un haut fond océanique localisé dans le Bassin Levantin en Méditerranée orientale entre Chypre et la marge continentale de l'Égypte. 4 sites 965, 966, 967 et 968, (Figure 7.1) atteignant respectivement une profondeur totale de 250, 297, 600 et 303 mètres sous le plancher océanique ont été forés à travers le mont sous-marin.

Dans sa partie supérieure, le forage du site 965, localisé sur le haut de la pente nord du mont Eratosthènes, a rencontré une couche de 30 m d'épaisseur composée de boues à nannofossiles d'âge plio-quadernaire, contenant quelques rares niveaux de sapropèles. Cet unité montre des preuves d'instabilité tectonique (failles, hiatus de sédimentation). Au dessous, une unité plus mince (environ 3 m) argileuse à clastes carbonatés contient aragonite et dolomite dont le dépôt a eu lieu dans des environnements caractérisés par une salinité faible à élevée. Elle est supposée d'âge Messinien. L'interprétation des données de diagraphies couplée avec l'observation des carottes montre ensuite un intervalle de 200 m composé de lits de calcaire relativement fins intercalés avec des sédiments micritiques plus abondants. La composition du calcaire indique un environnement de dépôt en eau peu profonde (lagon), adjacent à des zones récifales.

Le site 966 a été choisi au sommet du mont Eratosthènes (Figure 7.1). Les relevés sismiques indiquent une zone assez plane, coupée par plusieurs linéaments interprétés comme étant autant de failles normales actives (Robertson et al., 1994a,b; Limonov and Woodside, 1994). On relève aussi des indications de structures en compression. Les sédiments Plio-quadernaire sur le site sont d'épaisseur quasi-constante et les carottages effectuées durant la campagne du R/V Gelendzhik ont montré qu'ils étaient composés de boues à nannofossiles avec intercalation de niveaux de sapropèles (S1 et S5) et d'un niveau de cendres (Y5) (Lucchi et al., 1994).

Le site 967 se trouve au pied de la pente nord du mont Eratosthenes, sur le flanc d'une ride sous-marine de faible altitude (Figures 1 et 2). On retrouve la séquence Plio-quadernaire de boues argileuses, turbiditiques et à nannofossiles, avec intercalation de sapropèles, d'une épaisseur de 100 m environ. Ensuite, 15 à 25 m de sédiments rougeâtres à grain fins bien consolidés et à clastes carbonatés se sont déposés rapidement probablement dans un environnement de subsidence en eau profonde. Les données géochimiques et de diagraphies indiquent la présence de gypse, d'âge présumé Messinien, vers 120–130 mètres sous le plancher océanique. La section continue avec 310 m de craies à nannofossiles d'âge Eocène moyen à Crétacé supérieur (Santonien supérieur), puis avec 95 m environ de dépôts carbonatés d'eau peu profonde, fracturés et n'ayant pas donné jusqu'à présent de fossiles identifiables. L'étude des vitesses des ondes acoustiques et des résultats de la FMS pour cette dernière unité donne des indications d'intense activité tectonique et de brèches formées dans des environnements extensifs. Les données géochimiques

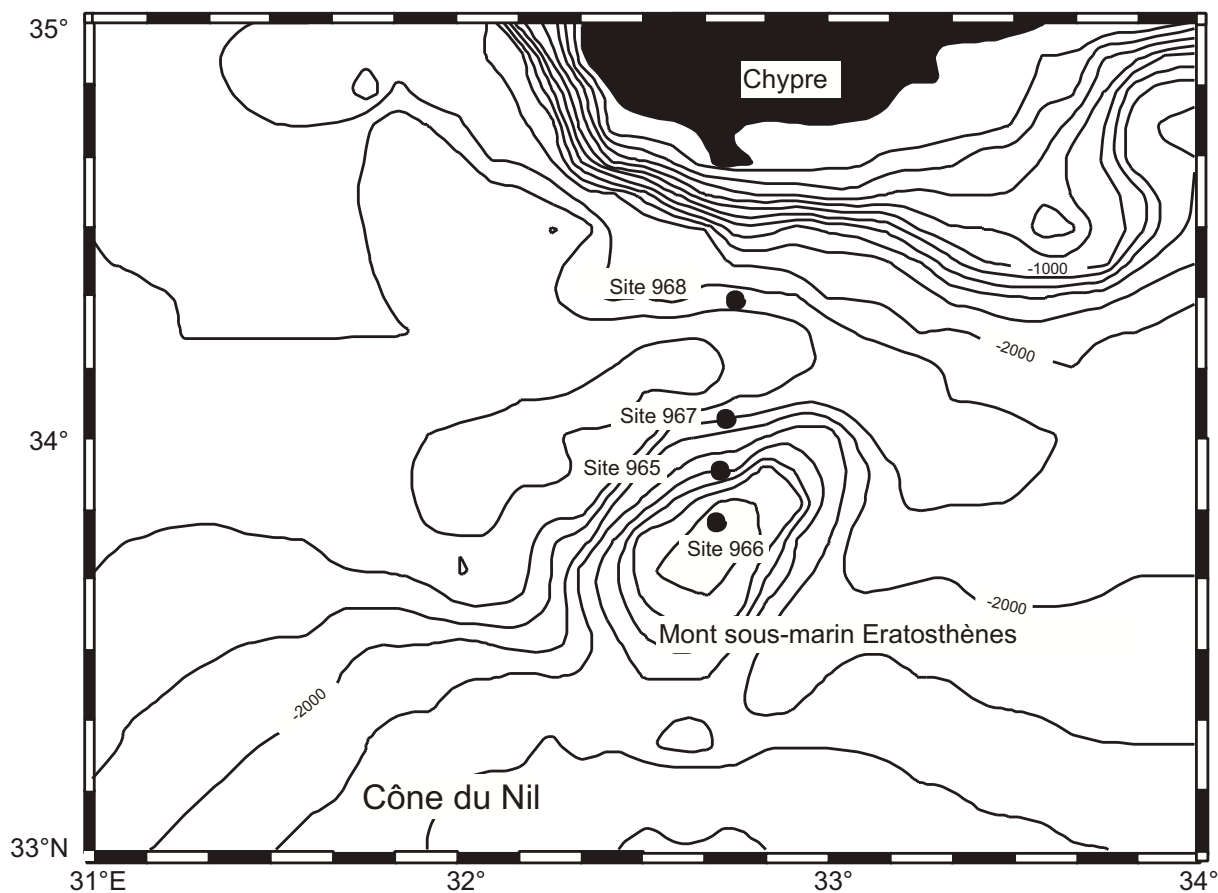


FIG. 7.1 – Localisation des 4 sites forés sur le mont sous-marin Eratosthènes durant le Leg ODP 160.

montrent la présence de solutions hypersalines en profondeur pouvant indiquer les évaporites messiniennes.

Le site 968 correspond au dernier site foré dans la zone du mont Eratosthènes et est localisé sur la crête d'un haut-fond en bas de pente de la marge active de Chypre (Figure 7.1). Il a pour but d'aider à la compréhension de l'évolution tectonique et sédimentaire de la marge de Chypre en relation avec les sites forés plus au sud sur le mont Eratosthènes. La partie supérieure de la section carottée sur le site 968 se compose de 143 m de boues argileuses à nannofossiles présentant un certain nombre de fines couches de cendres. Plus de 80 sapropèles interstratifiés avec des niveaux argilo-silteux et turbiditiques ont été récupérés. 23 m environ de sédiments à grains fins beaucoup plus argileux (kaolinite et smectite principalement) et 34m environ de turbidites calcaires interstratifiées avec des niveaux pélitiques et argileux (ayant probablement les dépôts du Nil comme origine). Plusieurs couches de gypse détritique ont été relevées. Les âges varient de l'Holocène à la limite entre le Pliocène inférieur et moyen. Un hiatus est présent entre le Pliocène supérieur et le Pléistocène inférieur et peut indiquer une période de fort soulèvement du sud de Chypre. Les données de géochimie des eaux interstitielles indiquent un niveau à halite sous la section carottée au dessous de 600 mètres sous le plancher océanique.

### 7.3.3 Histoire tectonique

L'histoire tectonique et sédimentaire du mont sous-marin Eratosthènes commence avec l'accumulation à la fin du Mésozoïque-début du Tertiaire, de dépôts carbonatés en eau peu profonde récupérés sur le site 967, et supposés d'âge Crétacé supérieur. Cette unité surmonte des carbonates d'origine plus profonde, dont seuls quelques fragments ont été récupérés, et dont l'âge précis reste pour l'instant inconnu. Entre le Crétacé supérieur et l'Eocène moyen, le mont Eratosthènes est interprété comme ayant fait partie de la marge passive africaine, peut-être comme un fragment crustal marginal. La nature du socle sous le mont sous-marin est inconnue mais représente probablement une croûte continentale transitionnelle, pénétrée par des intrusions de roches mafiques, au vu de la forte anomalie magnétique présente dans cette zone (Woodside, 1972, 1977). Par contre, plus au nord, d'importants mouvements tectoniques régionaux incluant la genèse et la dislocation tectonique des ophiolites de Troodos ont marqué le Crétacé supérieur dans la Téthys. Mais aucune preuve évidente de ces événements ne subsiste dans la séquence sédimentaire récupérée sur le mont Eratosthènes, peut-être à cause de sa position au sud des zones impliquées dans ces processus de collision et de marge active. Après l'Eocène moyen, on note un hiatus, et plus tard un soulèvement d'au moins 1 km du mont sous-marin suivi du dépôt de carbonates d'eau peu profonde. Ces dépôts calcaires sont les mêmes pour les sites les plus au sud (965, 966) tandis que pour le site 967 ils sont datés Miocène inférieur. Les données de logging indiquent des énergies de dépôt variables (environnement lagunaire à récifal), les carottes du site 966 ayant montré une grande abondance de corail relativement au site 965 amène à considérer le premier comme le plus proche des zones récifales. En ce qui concerne le site 967, le calcaire pélagique d'âge Crétacé supérieur à Eocène moyen est directement surmonté par des sédiments à grain fins d'environnement profond datés du début du Pliocène. La série Miocène est donc absente, mais il existe des indices de dépôts de cette période à la base du Pliocène sous forme de microfossiles indiquant des âges Eocène, Oligocène, Miocène et Pliocène (Emeis et al., 1996).

Pendant le Messinien, la mer Méditerranée connut sa crise de salinité, événement confirmé durant les Legs 13 et 42 du Deep Sea Drilling Project (DSDP) (Ryan and Hsü, 1973). Des niveaux d'évaporites sont connus pour exister sous le bassin Méditerranéen, provoquant le réflecteur très énergétique "M" sur les profils sismiques de la région (Ryan and Hsü, 1973; Ross and Uchupi, 1977). Mais ce réflecteur s'est révélé très mince aux abords du mont Eratosthènes (Limonov and Woodside, 1994), indiquant une position bien au-dessus des bassins profonds de la Méditerranée durant le Messinien. Le Leg 160 a confirmé la non-existence de niveaux épais à évaporites sous le mont Eratosthènes (site 966) ou sous ses abords nord (sites 965 et 967). Bien que le Messinien soit absent au niveau du mont sous-marin, plusieurs mètres de niveaux argileux rougeâtres ont été récupérés sur le site 965, contenant des faunes d'ostracodes pouvant être d'âge Messinien à Pliocène inférieur. De plus, sur le site 967, les données de diagraphies (voir plus loin) ont indiqué un niveau de gypse (non récupéré lors du carottage) de plusieurs mètres d'épaisseur vers 120 mètres sous le plancher océanique. Les données de la géochimie des eaux interstitielles ainsi que l'outil de diagraphie géochimique confirment une importante augmentation à ce niveau de la teneur en sulfates solubles. Durant le Messinien, le mont Eratosthènes existait en tant que haut-fond émergé, marqué par une érosion et/ou accumulation d'évaporites et de boues ferrugineuses, dans une configuration de lagons ou lacs marginaux. La présence de gypse sélénite et d'albâtre détritiques dans les sédiments argileux, ainsi que d'ostracodes et de foraminifères benthiques typiques indique un environnement probablement lacustre, conditions les plus largement répandues à la fin du Messinien, surtout sur le pourtour de la mer Noire (Steininger and Rogl, 1984). En résumé, le Messinien est très faible ou absent sur le mont Eratosthènes (sites 965, 966 et 967). Sur le site 968 (marge active de Chypre) par contre, le niveau à halite indiquée par la géochimie

des eaux interstitielles peut être considéré comme un dépôt messinien (faciès "Lago Mare") ou au moins tardi-Miocène.

A la fin de la crise messinienne, tous les sites forés enregistrent un regain de sédimentation, bien que les conditions varient à chaque site. Le Site 965 montre un intervalle d'épaisseur réduite (6 mètres) dont les nannofossiles indiquent des âges compris entre 3,5 et 6 Ma (fin du Miocène - début du Pliocène). La base du Pliocène est marquée par un réflecteur très énergétique, qui correspond au fort contraste lithologique entre le calcaire miocène dessous et les sédiments faiblement consolidés du Pliocène au dessus. Ce contraste se retrouve aussi au Site 967 plus au nord, incluant là des clastes calcaires et des foraminifères remaniés d'âge Eocène et Oligocène. Dans la zone forée par le Site 966, cette inconformité à la base du Pliocène est moins bien marquée que sur les parties adjacentes du mont sous-marin. On trouve ici environ 65 mètres de boues calcaires à nannofossiles. On peut y discerner une stratigraphie primaire, composée généralement d'abord d'un intervalle riche en clastes, devenant ensuite plus argileux, avec des intercalations riches en clastes détritiques. Ces clastes sont principalement calcaires, d'eau peu profonde, identiques à ceux de l'unité sous-jacente, et contiennent des microfossiles d'âge miocène, preuve que l'unité de brèches consolidées datée du pliocène s'est formée par érosion d'une source contenant des calcaires miocène. L'accumulation s'est totalement faite dans un environnement marin, et assez profond, d'après la faune fossile pélagique omniprésente. L'érosion des calcaires a pu être soit subaérienne, soit immergée, soit les deux. Il existe des preuves de transport sédimentaire - clastes roulés - qui dans certains cas indiquent des environnements de haute énergie (plages, eaux peu profondes). Parmi les clastes se trouvent des boues à nannofossiles, indiquant aussi probablement que l'instabilité tectonique locale a favorisé la remontée de sédiments marins d'origine profonde, inclus ensuite dans le remaniement. En résumé, le Site 966, et dans une moindre mesure le Site 965, ont montré qu'une tectonique active de faille a pris place avant 4,5 Ma au niveau du mont Eratosthènes, avec érosion des calcaires puis dépôts des clastes ainsi formés dans un environnement marin profond. Le Site 968, foré à la base de la pente de la marge de Chypre, a montré une histoire très différente : Considérant que l'unité plus riche en argiles à la base est d'âge messinien, la transition vers le Pliocène inférieur voit un apport de niveaux turbiditiques depuis la marge de Chypre, mais sans montrer de limite nette. Ce site est interprété comme ayant déjà été un site sous-marin au Messinien, et a donc connu une sédimentation en eau profonde au Pliocène inférieur.

Au Pliocène-Pléistocène, tous les sites forés montrent une sédimentation en eau profonde durant cette période. Le mont Eratosthènes se trouvait alors à une profondeur allant de 700 à 2900 mètres. La subsidence a été très rapide (les fossiles indiquant le Pliocène inférieur sont déjà d'environnement marin profond) au début du Pliocène, et a repris après la fin du Pliocène, comme le montre le carottage du Site 967 (fossiles indiquant des environnements encore plus profonds). L'activité tectonique entre la fin du Pliocène et le début du Pléistocène est à l'origine soit d'un hiatus de sédimentation (Sites 965 et 967), ou d'un taux extrêmement faible de sédimentation (Site 966), pouvant refléter des déformations du mont sous-marin (indiquées par les discordances et les plis observés sur les profils sismiques hautes résolution (Limonov and Woodside, 1994)). Ce hiatus n'a pas été clairement mis en évidence sur le Site 968.

En résumé, le mont sous-marin Eratosthènes fait partie intégrante de la plaque Afrique, et après un soulèvement et le dépôt en eau peu profonde d'une plateforme carbonatée au Miocène, la crise messinienne voit son émergence totale et le dépôt de niveaux évaporitiques. A la fin du Miocène, et jusqu'au début du Pliocène, la collision du mont Eratosthènes avec la marge active de Chypre entraîne une rapide subsidence et une sédimentation contrôlées par une tectonique extrêmement active, et sera à l'origine du soulèvement intense de Chypre au début du Pléistocène.

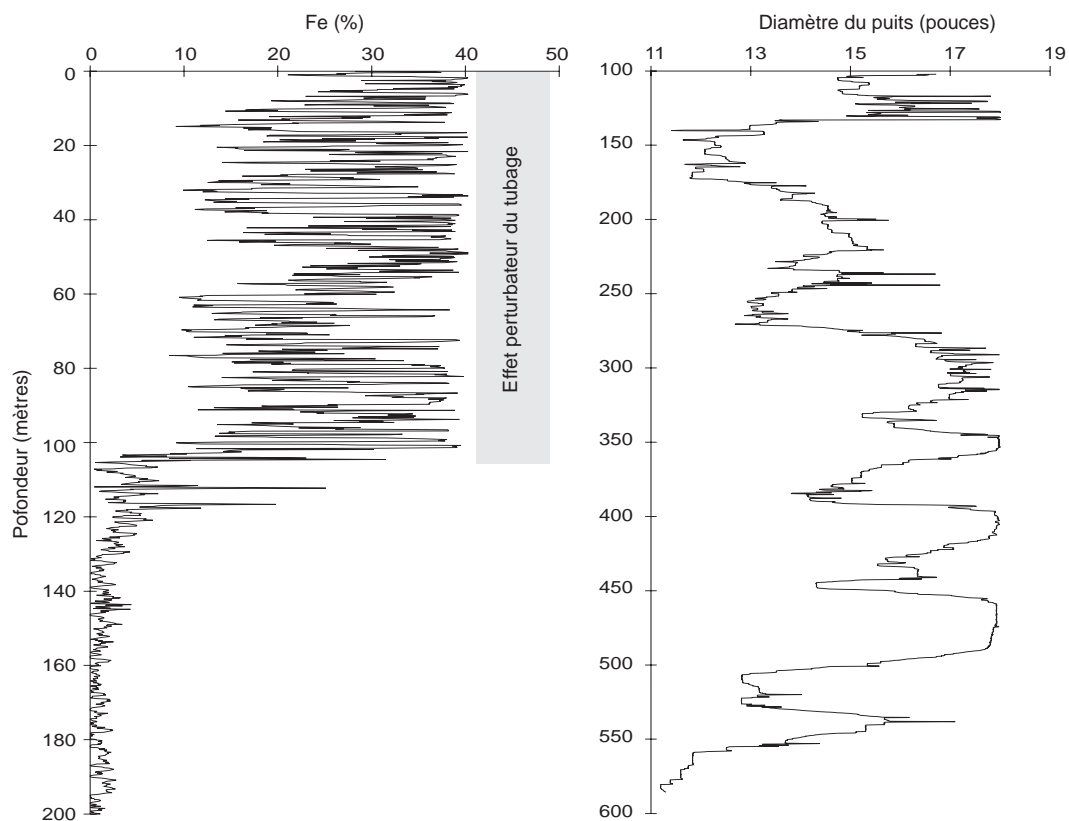


FIG. 7.2 – **Gauche** : Effet perturbateur du tubage sur la mesure du fer par l'outil GLT. Il est à noter que le tubage affecte la mesure de tous les éléments à des degrés différents, Al et Fe étant les plus fortement perturbés. **Droite** : Diamètre du puits 967E mesuré (en pouces) par l'outil HLDT pendant les opérations de logging par l'outil HLDT. Au delà de 13 pouces, l'erreur sur la mesure est souvent trop grande pour des outils comme le NGT, le CNT, ou le HLDT. Au delà de 18 pouces, le diamètre ne peut plus être mesuré, le bras du HLDT n'étant pas assez extensible.

### 7.3.4 Interprétation des diagraphie géochimiques et géophysiques, Site 967, puits 967E

Il est utile de préciser en préambule que les diagraphies réalisées dans ce puits sont entachées d'une incertitude assez importante. Néanmoins, il est possible dans certains intervalles d'interpréter qualitativement les variations de paramètres tels que le facteur photoélectrique PEF, la porosité neutron NPHI, ou bien encore les pourcentages d'éléments majeurs et traces (Si, Ca, S, Fe, Al, Th, K, Ti). Une inversion minérale a été tentée sans succès sur les pourcentages d'oxydes mesurés par l'outil GLT. L'incertitude est trop grande pour permettre une inversion correcte, ou du moins tout à fait objective.

La figure 7.2 montre l'effet perturbateur du tubage sur la mesure d'éléments comme le fer. Le tubage perturbe la mesure car il atténue le rayonnement gamma de capture provenant de la formation en réponse aux neutrons rapides provenant du GST. De plus, sa proximité, dans l'intervalle 100 à 120 mètres sous le plancher océanique (le tubage est à 80 mètres sous le plancher océanique durant les opérations de diagraphie), amène à douter de la précision des concentrations des éléments majeurs mesurées par le GLT. De plus, l'état du puits est typique des puits de forage ODP. Avant les opérations de diagraphie, le puits 967E a été carotté durant plusieurs dizaines

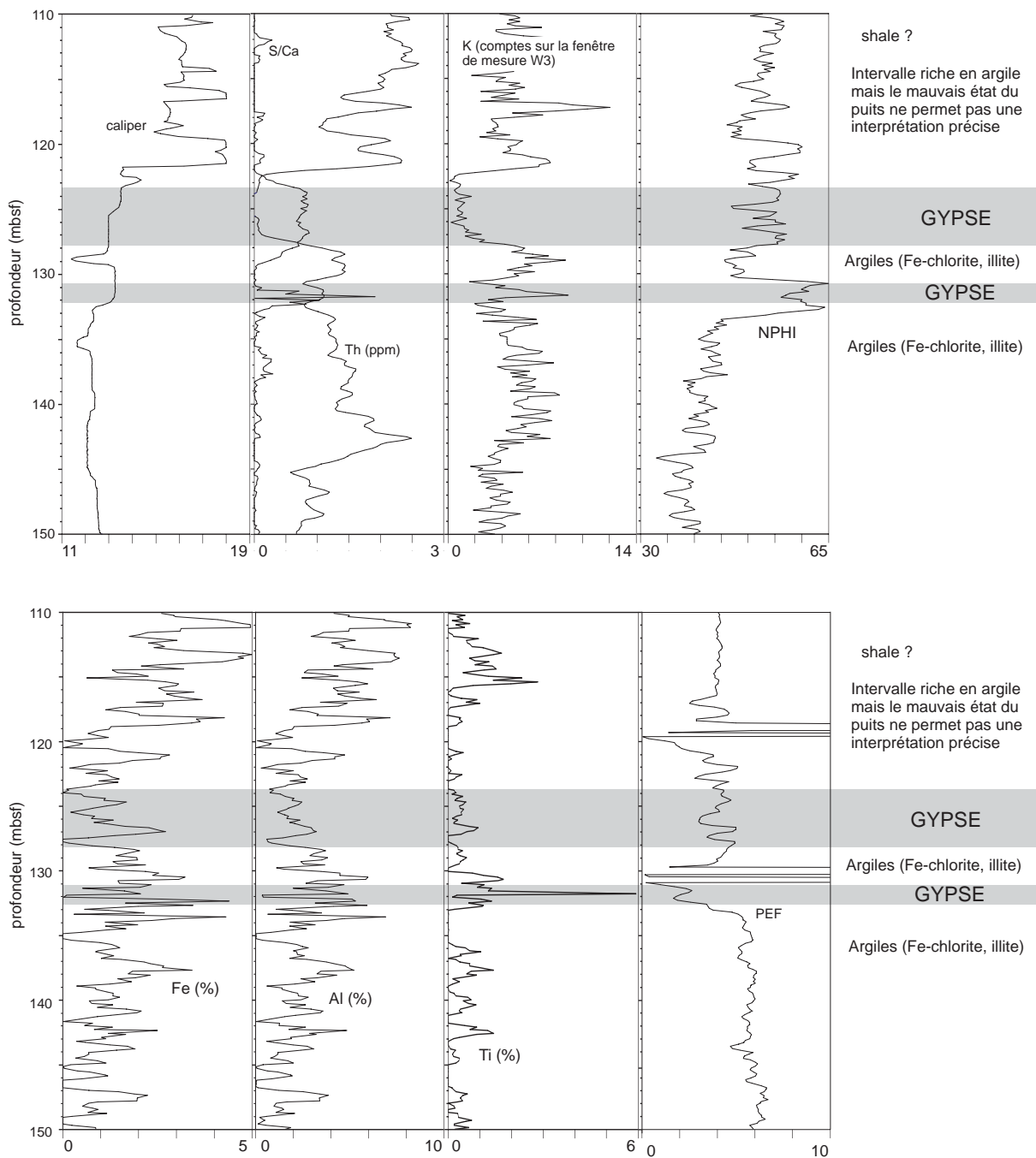


FIG. 7.3 – Interprétation de l'intervalle caractéristique correspondant aux évaporites messiniennes.

d'heures. Les parois du puits ont tendance à naturellement se déformer, des cavités se forment aux endroits où le sédiment est moins consolidé (sable, boue), rendant la mesure de diagraphie très aléatoire.

La figure 7.3 montre le comportement de plusieurs paramètres géochimiques et pétrographiques autour de 120 mètres sous le plancher océanique. Le caliper, stable autour de 12–13 pouces, indique des données de qualité moyenne. Le rapport S/Ca très proche de 1 caractérise sans ambiguïté la couche d'évaporites messiniennes. La porosité montre une valeur de 60 p.u., indiquant qu'il s'agit de gypse. On peut remarquer que l'outil CNT mesurant la porosité neutron donne étonnamment une valeur standard (voir Tableau D.1) pour le gypse. De part et d'autre de la couche de gypse, les valeurs en Th, K et Al, indiquent la présence possible de niveaux riches en argiles (Fe-chlorite, illite). Un autre pic du rapport S/Ca vers 132 mètres indique encore la présence d'évaporites, mais cette fois moins massivement que précédemment, et probablement intercalées avec des argilites et des boues calcaires. Les épontes de la couche d'évaporites sont des boues calcaires riches en argiles (glauconite, chlorite).

Le caliper (Figure 7.4, à gauche) mesure le diamètre du puits pendant les opérations de diagraphies. Au delà de 13 pouces, la mesure de la radiation gamma naturelle (NGT) est sujette à caution, tandis que au delà de 17 pouces, ce sont pratiquement toutes les mesures qui sont entachées d'une incertitude trop importante pour permettre une interprétation correcte. Les traces de Si et Ca permettent de séparer la section en quatre unités : la boue calcaire riche en argiles sert d'éponte à la couche de gypse. Vers 175 mètres, on passe à une unité plus calcaire, avec toutefois des passées argileuses (correspondant aux pics maximums de Si) et la présence de pyrite (correspondant aux pics maximums de S, les autres faisant parti du bruit très important dû à l'état du puits). Vers 400 mètres (la limite peut être placée à 395 ou 420 mètres), on trouve une lithologie très calcaire et plus homogène que les précédentes unités. Les pics de soufre peuvent aussi indiquer la présence localisée de nodules/veines de pyrite. La quatrième unité est plus hétérogène, correspond toujours à un calcaire, mais de texture plus grossière, avec des passées de chert au sommet.

Il ressort de ce bref descriptif que les conditions de mesure (état du puits, bonne utilisation des outils, bonne calibration) influencent extrêmement la qualité des données, et donc celle de leur interprétation, tant directe par une description des traces, qu'indirecte, par l'utilisation de méthodes d'inversion minérale, qui n'ont pas donné de résultats significatifs ici. L'étude des diagraphies effectuées dans le forage Balazuc-1, qui est l'objet de la section suivante, montre au contraire que des données de bonne qualité apportent de nombreuses informations pour la connaissance de la lithologie et de la minéralogie des formations géologiques traversées.

## 7.4 Le forage BALAZUC-1, Programme *Géologie Profonde de la France*

### 7.4.1 Introduction et cadre géologique

Le forage BALAZUC-1 (BA-1) a été réalisé dans le cadre du programme *Géologie Profonde de la France*, programme placé sous l'autorité d'un Comité Directeur composé de trois organismes à l'origine du projet : Le Bureau de Recherches Géologiques et Minières, Le Centre National de la Recherche Scientifique, et la Direction de la Recherche du Ministère chargé de l'Éducation Nationale. Le forage BALAZUC-1 a été réalisé entre le 23 Juillet et le 25 Septembre 1990, et constitue le premier ouvrage du programme Ardèche dont le thème est l'étude des interactions fluides-roches sur une paléomarge distensive. La zone étudiée est une portion de la marge continentale européenne appartenant à la bordure nord du paléo-océan Téthys ligur (Boillot et al.,

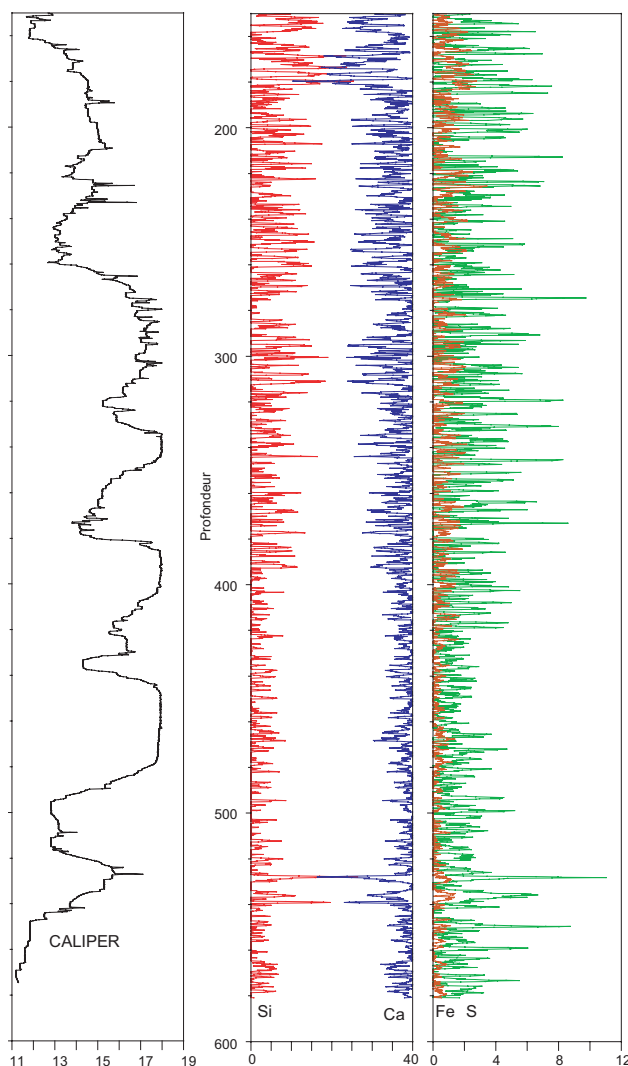


FIG. 7.4 – Données de diagraphie dans le puits 967E. Le caliper donne une idée de l'état du puits. Si et Ca sont caractéristiques de la lithologie dominante.

1984). La durée de vie de cette structure géologique en tant que marge passive couvre une large partie du Mésozoïque (Trias-Crétacé inférieur). Une coupe géologique visible Figure 7.5 a été réalisée au cours du forage grâce à l'étude pétrographique et micropaléontologique des déblais de forage, lorsque celui-ci était réalisé en destructif, ou des carottes, lorsque des sections carottées étaient réalisées.

L'approche sismique du site, précédant le forage, a permis de découvrir une faille normale dite d'Uzer, cachetée par le Jurassique moyen et supérieur et affectant le socle anté-triasique. Le forage a été localisé de façon à ne pas recouper cette faille, permettant ainsi de traverser toute la série mésozoïque (Trias moyen à Jurassique supérieur). Les données préliminaires (Documents du BRGM, 1990) montrent que les modifications diagénétiques sont très intenses, et se manifestent par une cimentation-compaction poussée responsable de valeurs très élevées de la densité des roches et de vitesses sismiques proches de 6 km/sec., ou par la concentration de minéraux tels que barytine, dolomite, anhydrite sur pratiquement toute la série carottée en continu (Trias-



Hettangien).

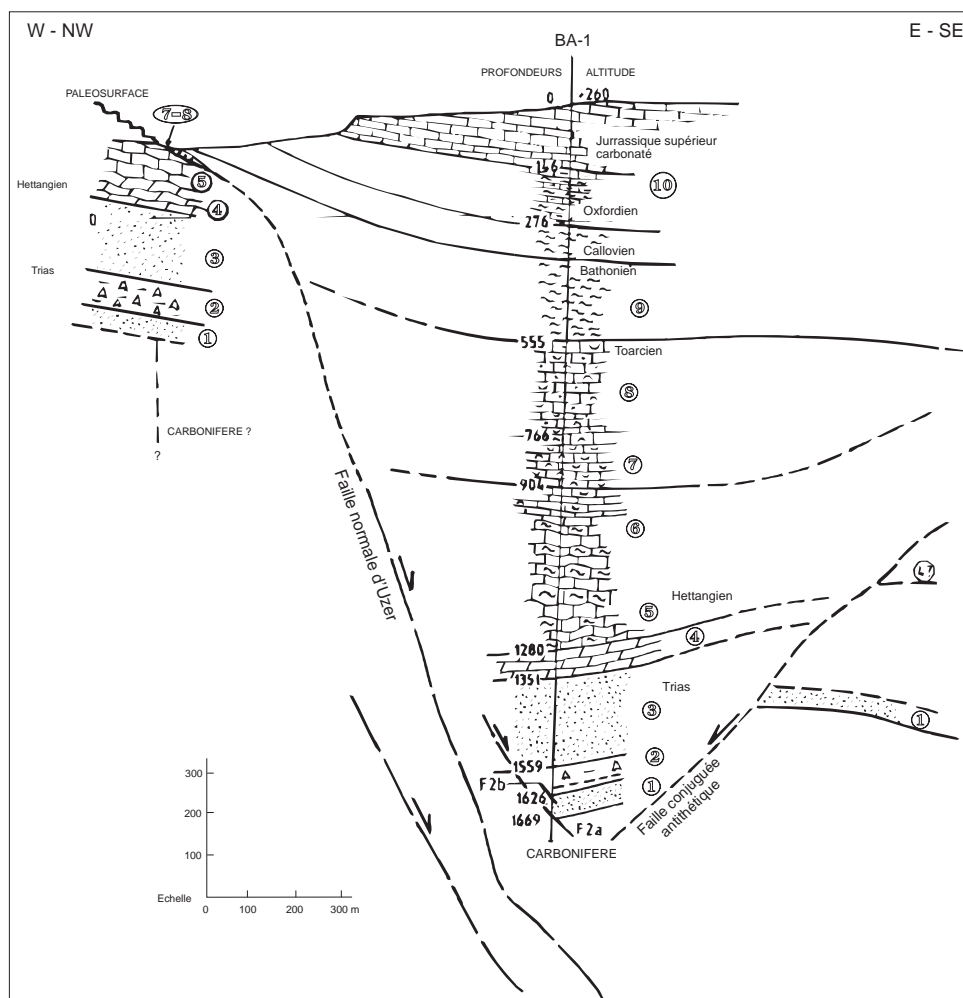


FIG. 7.5 – Coupe géologique réalisée lors du forage BA-1. 1. Ensemble gréseux inférieur. 2. Ensemble argilo-sulfaté. 3. Ensemble gréseux supérieur, marin au sommet, continental, puis lagunaire à la base. 4. Ensemble dolomitisé, fortement au sommet, partiellement à la base. 5. Ensemble calcaréo-argileux, alternant, ou 6. à faciès noduleux. 7. Biocalcarénites. 8. Biocalcarénites quartzueuses partiellement silicifiées. 9. Ensemble marneux. 10. Ensemble carbonaté (d'après Documents du BRGM (1990)).

#### 7.4.2 L'inversion minérale

**Choix des données** La tentative de reconstruction d'une minéralogie en continu a été effectuée entre 100 et 1700 mètres de profondeur pour plusieurs raisons : dans cet intervalle, on constate la plus grande variabilité lithologique (comparé au sommet de la section forée), allant du calcaire à un complexe argilo-sulfaté. Ceci nous permet de tester la souplesse d'adaptation du modèle employé. De plus, la zone a fait l'objet de trois carottages : entre 1115 et 1124 mètres, entre 1160 et 1169 mètres, et entre 1199,5 et 1729,48 mètres (Documents du BRGM, 1990). Les résultats obtenus pourront être comparés à la lithologie reconnus sur carottes.

Les variables choisies sont les pourcentages d'oxydes d'éléments majeurs mesurés en continu par l'outil GLT (voir Section 1.5.2), soit  $\text{SiO}_2$ ,  $\text{Al}_2\text{O}_3$ ,  $\text{K}_2\text{O}$ ,  $\text{CaO}$ ,  $\text{Fe}_2\text{O}_3$ ,  $\text{K}_2\text{O}$ ,  $\text{SO}_3$ ,  $\text{TiO}_2$ . Trois

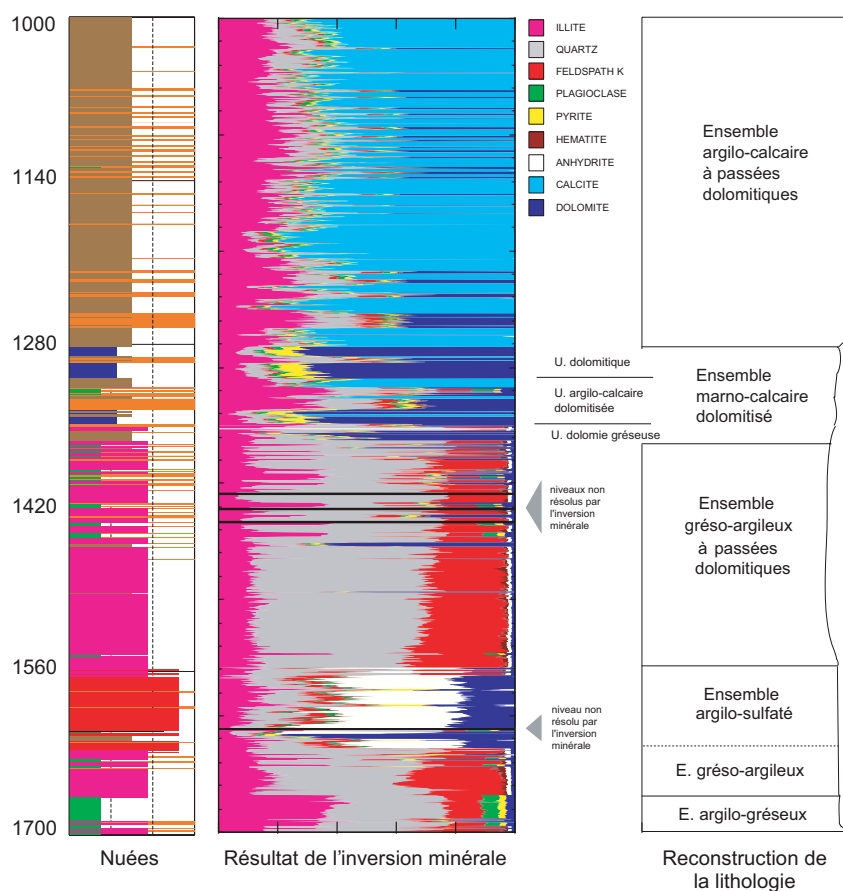


FIG. 7.6 – **Gauche** : résultats de l'analyse en nuées dynamiques. 7 groupes ont été individualisés dans la section étudiée. **Milieu** : résultats de l'inversion minérale. L'algorithme d'inversion totale a trouvé une solution satisfaisante pour la quasi-totalité de l'intervalle. Quelques niveaux particuliers, vers 1420 et 1620 mètres n'ont pas été résolus. **Droite** : Il est possible de diviser la section en plusieurs ensembles lithologique d'après leur minéralogie dominante.

autres variables représentant des propriétés plus pétrophysiques, ont aussi été retenues : l'index photoélectrique PEF, la densité totale RHOB, et la porosité neutron NPHI. Les pourcentages d'oxydes ont été fournis par Serge Scheerens, après recalcul tenant compte dans le modèle de fermeture de la possibilité d'avoir plusieurs oxydes associés au même élément majeur dans la formation (voir Scheerens (1995) pour les détails de ce traitement).

**Schéma de mise en oeuvre** L'algorithme utilisé pour l'inversion minérale est l'algorithme de l'inversion totale, proposé par Tarantola and Valette (1982), et dont le principe est explicité Section 6.3.3. La distance quadratique à minimiser s'écrit

$$q^2 = (X - X_0)^T M_0^{-1} (X - X_0)$$

où l'on voit apparaître  $X_0$ , qui est une solution *a priori* acceptable, utilisée comme point de départ à la descente du gradient. Evidemment, les gains d'informations seront d'autant plus importants, du moins en théorie, que la solution *a priori* sera proche du minimum global de la surface d'erreur. D'un autre côté, une solution *a priori* trop précise ne laissera que peu de liberté à

	PEF	NPHI	RHOB	SiO <sub>2</sub>	Al <sub>2</sub> O <sub>3</sub>	K <sub>2</sub> O	CaO	Fe <sub>2</sub> O <sub>3</sub>	SO <sub>3</sub>	TiO <sub>2</sub>
Groupe 1	3,32	15,94	2,66	50,64	18,71	7,72	15,80	6,28	0,69	0,16
Groupe 2	4,17	6,50	2,87	18,20	3,56	0,87	62,92	13,68	0,61	0,17
Groupe 3	4,49	5,33	2,73	18,23	6,71	2,43	69,37	2,60	0,65	0,02
Groupe 4	15,43	2,79	2,67	70,65	7,36	6,06	12,45	1,68	1,79	0,01
Groupe 5	1,97	15,16	1,92	8,99	22,31	0,79	5,64	9,67	47,13	5,48
Groupe 6	4,46	3,71	2,87	20,41	5,15	2,24	35,78	1,57	34,85	0,01
Groupe 7	3,87	11,78	2,73	39,04	10,73	4,47	40,49	4,35	0,87	0,04

TAB. 7.1 – Compositions des barycentres des groupes individualisés par une analyse en nuées dynamiques sur les données de diagraphie acquises dans le forage BA-1.

Groupe 1	Illite, Quartz, Orthose, Albite, Anhydrite, Calcite
Groupe 2	Illite, Quartz, Orthose, Albite, Pyrite, Calcite
Groupe 3	Illite, Quartz, Orthose, Pyrite, Anhydrite, Calcite
Groupe 4	Illite, Quartz, Orthose, Pyrite, Hématite, Dolomite
Groupe 5	groupe anomal
Groupe 6	Illite, Quartz, Orthose, Albite, Anhydrite, Dolomite
Groupe 7	Illite, Quartz, Orthose, Albite, Hématite, Pyrite, Dolomite

TAB. 7.2 – Paragenèses de départ affectées aux groupes individualisés par l'analyse en nuées dynamiques. Le Groupe 5 présente des concentrations anormales en éléments majeurs, et ne comprend que peu d'échantillons (9). Il sera traité à part.

l'algorithme pour trouver un minimum géologiquement correct. La méthode de l'inversion totale considère données et paramètres du modèle comme des inconnues, qui sont les composantes d'un vecteur unique  $X$ , auquel est affectée une matrice de covariance  $M$ . L'algorithme optimise tous ces paramètres pour résoudre le système  $G(X) = 0$ .

Pour avoir une idée de la solution *a priori*, les 8 oxydes mesurés par le GLT, ainsi que l'index photoélectrique, la densité totale et la porosité neutron ont servi de base à une classification des données par la méthode des nuées dynamiques (Section 5.3.2). 7 groupes ont été testés avec un pseudo-test de Fisher, ayant pour critère la somme de carrés des distances inter-nuées, comme étant la meilleure représentation du jeu de données. Les compositions des barycentres de ces groupes ont été utilisés pour choisir une paragenèse de départ correspondant à chaque groupe, devant servir comme paramètres à l'inversion minérale. Le Tableau 7.1 présente les compositions des barycentres et le Tableau 7.2 les paragenèses retenues. Les observations pétrographiques effectuées sur les 63 échantillons prélevés sur les carottes entre 904 et 1730 mètres ont montré (Documents du BRGM, 1990) que l'illite était le seul minéral argileux présent dans une concentration significative en dessous de 800 mètres de profondeur, et que le feldspath majoritaire était le feldspath potassique (orthose), par rapport au plagioclase. Les compositions des minéraux, avec les erreurs associées, utilisés pour l'inversion minérale, sont reportées dans le tableau 7.3 et sont tirés de Scheerens (1995).

Seuls les pourcentages d'oxydes mesurés par le GLT ont été utilisés dans l'inversion minérale. Les erreurs sur ces données, introduites dans l'analyse, sont celles typiquement reportées par Schlumberger (voir Sections 1.5.2.2 et 1.5.2.3). Les trois variables physiques précitées, PEF, RHOB et NPHI sont utilisées qualitativement dans les cas où l'inversion minérale ne parvient pas à converger vers une solution satisfaisante.

	Ill.	Qtz	Orth.	Albi.	Dolom.	Calc.	Pyrite	Hématite	Anhydr.
SiO <sub>2</sub>	56,14 (2,8)	100 (0,001)	63,50 (3,17)	71,27 (2,24)	0 (0,001)	0 (0,001)	0 (0,001)	0 (0,001)	0 (0,001)
Al <sub>2</sub> O <sub>3</sub>	29,65 (1,5)	0 (0,001)	18,30 (0,91)	24,39 (1,78)	0 (0,001)	0 (0,001)	0 (0,001)	0 (0,001)	0 (0,001)
K <sub>2</sub> O	7,89 (0,39)	0 (0,001)	18,20 (0,91)	0,34 (0,02)	0 (0,001)	0 (0,001)	0 (0,001)	0 (0,001)	0 (0,001)
CaO	0,34 (0,02)	0 (0,001)	0 (0,001)	4,15 (0,95)	100 (0,001)	100 (0,001)	0 (0,001)	0 (0,001)	41,20 (0,001)
Fe <sub>2</sub> O <sub>3</sub>	5,52 (0,27)	0 (0,001)	0 (0,001)	0 (0,001)	0 (0,001)	0 (0,001)	69,00 (0,001)	100 (0,001)	0 (0,001)
SO <sub>3</sub>	0 (0,001)	0 (0,001)	0 (0,001)	0 (0,001)	0 (0,001)	0 (0,001)	31,00 (0,001)	0 (0,001)	58,80 (0,001)
TiO <sub>2</sub>	0,45 (0,02)	0 (0,001)	0 (0,001)	0 (0,001)	0 (0,001)	0 (0,001)	0 (0,001)	0 (0,001)	0 (0,001)

TAB. 7.3 – Compositions des minéraux utilisés pour l’inversion minérale sur le forage BA-1 (d’après Scheerens (1995)). Les nombres entre parenthèses correspondent aux erreurs utilisées pour l’inversion totale.

### 7.4.3 Résultats

Sur presque toute la section étudiée (1000–1700 mètres), l’analyse a permis de reconstituer une lithologie satisfaisante. Il est nécessaire d’explicitier ce terme. En effet, on a vu qu’en théorie, les algorithmes d’inversion peuvent converger vers une solution correcte du point de vue mathématique, mais sans aucune réalité géologique. Il est difficile, sans avoir de base de comparaison, tel que des observations sur échantillons, de vérifier que la solution trouvée reflète effectivement la lithologie. Néanmoins, mis à part l’illite et le feldspath potassique, les autres minéraux rencontrés dans les formations traversées par le forage BA-1 montrent des compositions chimiques relativement stables, ce qui laisse présager de l’existence d’une solution quasi-unique, aux marges d’erreur près, correspondant à un minimum de l’erreur quadratique relativement isolé. Un test de contrôle a été effectué, en prenant, pour les 7 groupes individualisés par l’analyse en nuées dynamiques, une solution *a priori* identique. Le résultat est très mauvais, même si l’analyse réussit à trouver une solution mathématiquement correcte pour quelques niveaux de mesure.

La figure 7.6 montre le résultat de l’analyse en nuées dynamiques, ainsi que celui de l’inversion minérale. Le log stratigraphique à droite montrent les grandes unités qui ont été isolées d’après les variations minéralogiques visibles. Deux grands ensembles lithologiques, le premier à dominante argilo-carbonatée, entre 1000 et 1365 mètres, le second à dominante grés-argileuse, interrompu par un ensemble évaporitique, entre 1365 et 1700 mètres.

**L’ensemble argilo-carbonaté** Cet ensemble se subdivise en deux grandes unités selon la fréquence des passées dolomitiques. Sur la colonne minéralogique, le motif bleu foncé, appelé “dolomite”, correspond en fait à un faciès où le minéral carbonaté principal est la dolomite. Le motif bleu clair, quant à lui, correspond aux endroits où la calcite est le carbonate de calcium principal est la calcite. L’outil GLT ne mesurant pas la magnésium, nous utilisons l’index photoélectrique (PEF) pour estimer l’importance relative de la dolomite et de la calcite : un PEF proche de 3 montrera une présence de dolomite plus importante, tandis que plus proche de 5, le minéral carbonaté dominant sera la calcite. Les échantillons du Groupe 3 (couleur marron) montrent un PEF moyen supérieur à celui des échantillons des Groupes 2 et 7 (couleurs bleu foncé et orange), et l’on peut grossièrement en déduire que la phase carbonatée majeure pour

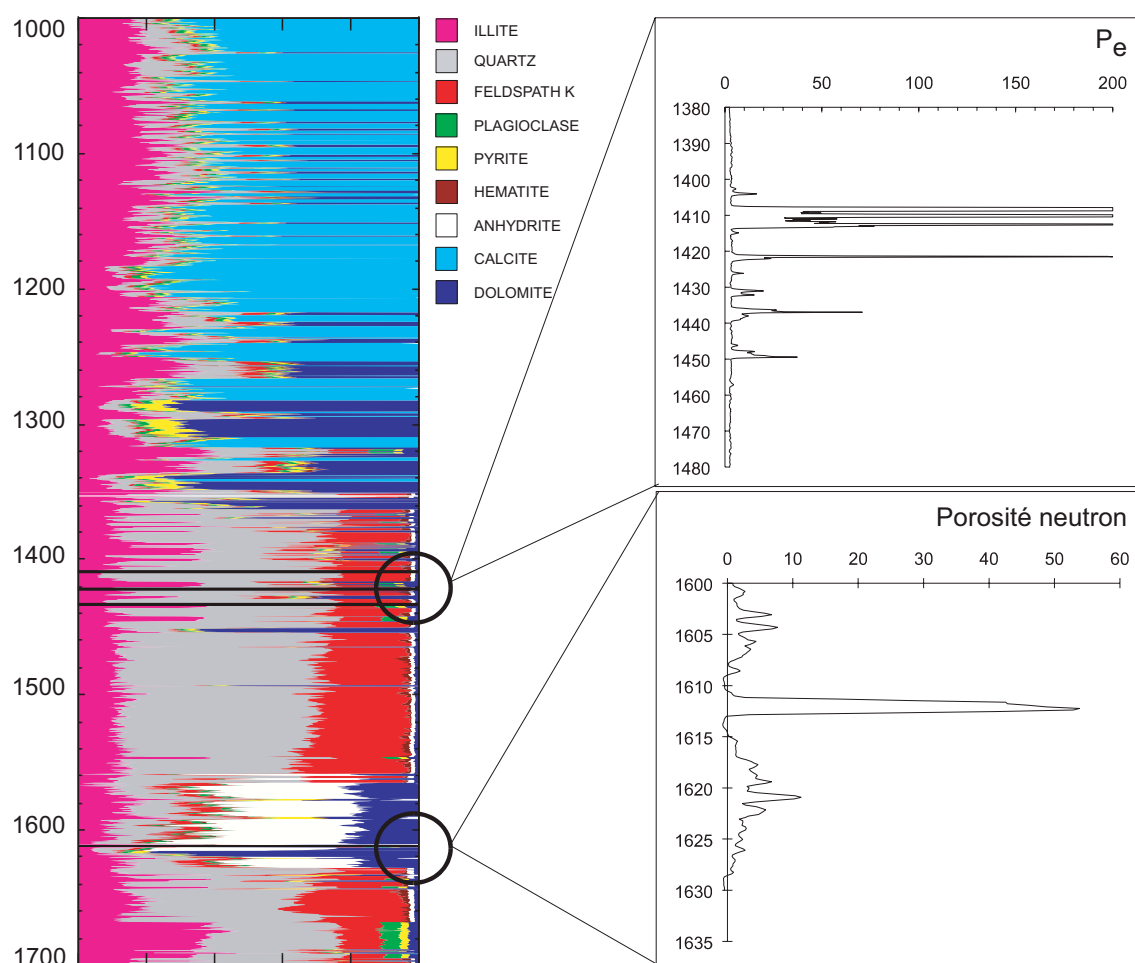


FIG. 7.7 – Utilisation des propriétés pétrophysiques mesurées en diagraphie, en l'occurrence l'index photo-électrique et la porosité neutron, pour confirmer ou interpréter des niveaux dont il est difficile de retrouver la minéralogie.

le Groupe 3 est la calcite, tandis que celle des Groupes 2 et 7 est la dolomite. Les bancs où la dolomitisation semble plus importante (Groupe 7) deviennent logiquement plus épais vers le bas de l'unité. Il semble aussi pertinent de considérer que les intervalles où la pyrite est présente en quantité non négligeable sont plutôt dolomitisés, bien qu'à 1320 mètres, le PEF indique plutôt la calcite (mais cette augmentation du PEF peut aussi être due à l'influence du PEF de la pyrite, qui est de 17 barns/e<sup>-</sup>). On remarque une proportion d'illite non négligeable, pouvant atteindre 50% dans les parties plus dolomitiques. Ceci indique s'il est besoin une diagénèse (circulations de fluides?) très importante dans la série carbonatée, diagénèse d'intensité toutefois variable en profondeur (variations de la concentration en argiles).

La première unité (1000 à 1282 mètres) est caractérisée par un PEF relativement élevé (entre 4 et 5), donc par des teneurs élevées en calcite, bien que la dolomite soit de plus en plus présente vers la base de l'unité. La deuxième unité (1282 à 1363 mètres) est caractérisée par une variabilité minéralogique plus grande et l'apparition de la pyrite. Elle peut être subdivisée en 3 sous-unités sur la base des assemblages minéraux et de l'abondance en dolomite. Une première sous-unité, entre 1282 et 1310 mètres, est un marno-calcaire assez fortement dolomitisé, où la

pyrite est abondante (10 à 15%). La deuxième sous-unité (1310 à 1363 mètres) est argilo-calcaire dolomitisé, où l'on remarque la présence du quartz et de la pyrite (phase gréseuse). La troisième sous-unité est une dolomie gréseuse.

**L'ensemble grés-argileux** Cet ensemble est caractérisé par une grande unité grés-argileuse plutôt homogène, entre 1363 et 1668 mètres, interrompue entre 1558 et 1629 mètres par une unité argilo-sulfatée. Au sommet de l'unité grés-argileuse, on trouve de fréquente passées dolomitiques, ainsi que la présence localisée de pyrite et d'évaporites en faible proportion. A trois reprises, quelques niveaux ne correspondant pas du tout à l'assemblage minéralogique proposé n'ont pu être résolus par l'inversion minérale. On verra plus loin comment les interpréter. Entre 1450 et 1558 mètres, la base de l'unité grés-argileuse, tant au-dessus qu'au-dessous de l'unité sulfatée, montre une très bonne homogénéité, avec présence fréquente d'hématite.

L'unité sulfatée est caractérisée par des teneurs en soufre et en carbonate très importantes (entre 25 et 40%), alors que le fer est peu abondant, tout ceci confirmant la présence d'anhydrite, mélangée à des argiles, et présentant localement des faciès à pyrite-hématite. On note la présence d'une barre carbonatée (1610–1620 mètres) au milieu de la séquence évaporitique. Ici aussi, l'inversion minérale n'a pas donné de résultats satisfaisants pour un niveau particulier, vers 1610 mètres.

**L'ensemble argilo-gréseux** Cet ensemble qui termine la section étudiée est caractérisé par une augmentation de la teneur en argile (illite), et montre des teneurs élevées en feldspath potassique et plagioclase. Il s'agit probablement d'un shale, dont l'âge est indiqué comme carbonifère (Documents du BRGM, 1990).

**Les niveaux non reconnus** A plusieurs reprises, l'inversion minérale n'est pas arrivé à une solution satisfaisante. Dans le premier cas, on remarque (Figure 7.7) que ces niveaux sont caractérisés par des valeurs du PEF hors échelle (plus de 200 barns/e<sup>-</sup>). Ces niveaux montrant aussi des quantités non négligeables de soufre (jusqu'à 7%), une interprétation logique est de les associer à du ciment barytique. Etant donné la porosité relativement faible de l'unité gréseuse (quelques %), il s'agit d'un colmatage matriciel, la barytine ayant rempli les vides entre les grains de quartz, diminuant ainsi drastiquement la porosité.

Le second niveau non résolu est caractérisé par une porosité neutron très élevée (plus de 50%; voir figure 7.7) ainsi qu'une densité et un index photoélectrique faibles (respectivement 1,3 g/cm<sup>3</sup> et 2 barns/e<sup>-</sup>). Il pourrait s'agir d'un niveau riche en matière organique (charbon) mais la monotonie du log de l'uranium à cet endroit ne valide pas cette hypothèse. On peut penser à une couche massive de halite (NaCl), coiffant la barre carbonatée à 1610–1620 mètres, indétectable par le GLT qui ne mesure ni Na, ni Cl. La présence de halite a été confirmée par les études préliminaires (Documents du BRGM, 1990), et citée par Scheerens and Steinberg (1997) et Aquilina et al. (1997).

**Validation** Figure 7.8 montre la comparaison entre la densité calculée à partir des proportions minérales trouvées par l'inversion minérale et la densité RHOB mesurée par diagraphie (diagramme de gauche), pour l'unité argilo-gréseuse triasique, entre 1350 et 1700 mètres. Les valeurs standard de densité pour chaque minéral sont exprimées dans le tableau de l'Annexe D. On remarque un bon accord entre les deux, les valeurs calculées étant néanmoins légèrement supérieures. Ceci est dû au fait que la porosité n'est pas prise en compte dans le calcul. La différence  $\frac{|\rho_{calc}-\rho_b|}{\rho_b}$  est comprise entre 0 et 7%, avec une moyenne autour de 3%, correspondant bien à la porosité moyenne dans cet intervalle, mesurée par diagraphie (NPHI) et sur carottes.

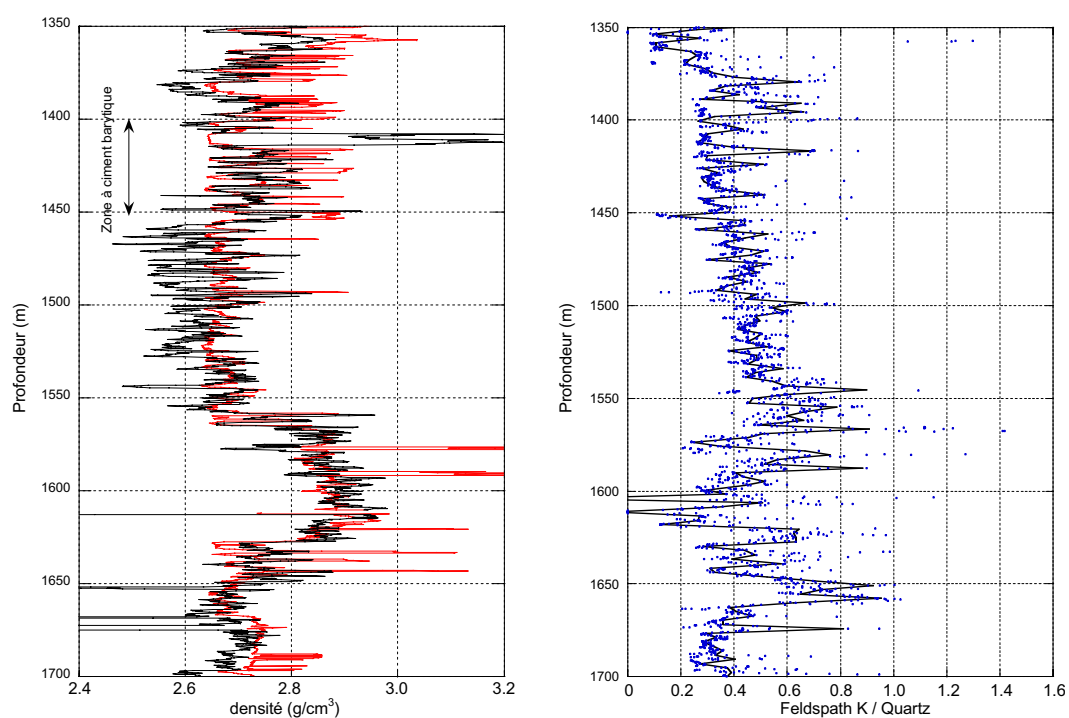


FIG. 7.8 – Validation des résultats de l'inversion minérale. **Gauche** : Comparaison entre la densité calculée à partir de la minéralogie en continu (courbe rouge) et celle mesurée par diagraphie (RHOB, courbe noire) pour la série gréseuse du Trias. **Droite** : Variations du rapport Feldspath K / Quartz pour la série gréseuse du Trias; la courbe est le résultat d'une interpolation par les points.

Le diagramme de droite de la figure 7.8 montre l'évolution du rapport Orthose/Quartz dans la série gréseuse triasique. Ce rapport est souvent un bon indicateur pour la maturité du sédiment. Il varie suivant les unités décrites plus haut : 1) le rapport OR/Q est faible pour les dolomies du sommet de la section, puis augmente avec la profondeur dans l'unité gréseuse supérieure (1350 à 1560 mètres); 2) dans la séquence évaporitique (1560 à 1630 mètres), le rapport Or/Q est décroît pour arriver à des valeurs très faibles, dues à la disparition du feldspath potassique; 3) le rapport augmente à nouveau jusqu'à des valeurs proche de 1 dans la partie inférieure de l'unité gréso-argileuse; 4) le rapport Or/Q est faible dans les shales carbonifères. Cette évolution de la maturité des apports détritiques correspond bien avec les séquences sédimentologiques de second ordre décrites par Courel et al. (1997) et Aquilina et al. (1997), et que les concentrations en quartz et feldspath potassique reflètent l'intensité des apports détritiques lors de la phase de sédimentation triasique.





## Chapitre 8

# Retrouver les paramètres pétrophysiques des formations traversées

- Le prisme d'accrétion Nord-Barbade, Leg ODP 156
- Les formations carbonatées

## 8.1 Dynamique du prisme d'accrétion nord-Barbade, Leg ODP 156

### Article 5 :

*C. R. Acad. Sci. Paris*, Tome **320**, Série IIa, 1995, pp. 977–984

Circulation et surpression de l'eau interstitielle dans le prisme d'accrétion nord-Barbade : résultats du Leg ODP 156

Pierre Labaume, Pierre Henry **Alain Rabaute**, et l'équipe scientifique du Leg ODP 156

**Article 6 :**

soumis à *Water Resources Research*

*In Situ* determination of permeability from downhole measurements  
analysis

André Revil, Delphine Fitzenz, **Alain Rabaute**, Mohamed Ramadan, Louis Briquet

## 8.2 Structure du réseau poreux dans les formations carbonatées

### Article 7 :

La confidentialité, toujours active sur les données ayant servi de base de travail pour cet article, ne me permet pas de citer les noms des puits étudiés (Well 1 et Well 2), ni de montrer les résultats obtenus. Ma participation du 12 au 19 Juillet 1998 au SPE Forum Series Meeting, qui doit se tenir à Breckenridge, CO, USA, devrait voir la levée de cette confidentialité et la publication de ce travail dans un journal du type *Journal of Sedimentary Research*.

Pore space components and their partitioning in carbonate rocks

**Alain Rabaute**, T.S. Ramakrishnan, Stefan Luthi, Marina Polyakov

# Pore space components and their partitioning in carbonate rocks

Alain Rabaute, T.S. Ramakrishnan, Stefan Luthi, Marina Polyakov

December 12, 2006

## **Abstract**

Pore space partitioning into different porosity types was carried out on numerous carbonate samples selected from two wells with diverse sedimentological, petrophysical and petrographic range. The vuggy fraction of porosity was evaluated from core and plug slabs using a novel but simple method, while intergranular (“macro”) and intragranular (“micro”) fraction of porosity were estimated by judiciously combining optical microscope (OM) and environmental scanning electron microscope (ESEM) views of thin sections. Total porosity estimates with the three pore types compare reasonably well with the conventional He intrusion measurements. In addition to pore type partitioning, pore network characteristics in terms of lithofacies are also identified. The implication of classification to petrophysical data interpretation and transport calculation is briefly discussed.

To be submitted to *Journal of Sedimentary Petrology*

## Introduction

The petrophysical response and the flow characteristics of carbonate rocks is governed to a large extent by their sedimentation and diagenetic history. While the depositional environment is responsible for initial variations in texture, post depositional processes such as dissolution, cementation, and replacement add to the tremendous variability in the pore structure of carbonates (Choquette and Pray 1970; Lucia 1983; Nurmi 1984; Lucia 1995). For this reason, Dunham (1962) classified carbonates into two groups: (i) where the depositional texture is recognizable and (ii) where it is not (see Figure 1).

Depositional features of carbonate rocks differ from sandstones since (i) many carbonate sediments are characterized by a bimodal grain distribution (Folk 1959); (ii) organisms play an important role in producing the sediment both in terms of fossils and fecal pellets; and (iii) much of the carbonate sediments deposit in shallow water.

Initially, a carbonate deposit may have void space ranging from 60 to 80% of the total volume, often as interparticle porosity (Choquette and Pray 1970; Lucia 1983). Unlike the siliciclastics, the physico-chemical processes following deposition have a tremendous influence in altering the mineralogy and the texture, and in the formation of the carbonate rock. The shape and the size of the final open network is likely to be heterogeneous and difficult to predict. Pore sizes may vary from sub-microns to meters, sometimes only a few feet apart. Pore shapes may be altered due to dissolution and/or recrystallization processes.

It is clear that a petrography study of carbonate rocks cannot cover all aspects and features that are likely to be encountered. This is especially true with regards to quantitative partitioning of porosity and pore-types. Our attempt has therefore been to simplify the petrography classes to the essential features that are relevant to transport properties, be it electrical resistivity or permeability.

## 1 Geological background

The petrography studies are based on two wells. The cores from the two wells were found to cover a broad range of pore types, porosity and permeability.

### 1.1 Well 1

The lithology consists of mainly crystalline dolomite. The original deposit was as a shelf-margin carbonate, but extensively dolomitized, and mixed with nodules or intercalations of anhydrite. We were provided with 39 samples out of a cored interval of about 90 m. The depositional texture is no longer recognizable..

A tight arrangement of euhedral or semi-euhedral crystals of 30–100  $\mu\text{m}$  is evident from the thin section analysis. The pore space is dominated by intercrystalline and vuggy porosity with minor amounts of microporosity, present as microvugs. Much of the interval has a porosity less than 5%, and the permeability is mostly within 0.1 mD (100  $\text{nm}^2$ ). The lower limit of the permeameter that we used is 0.01 mD (10  $\text{nm}^2$ ).

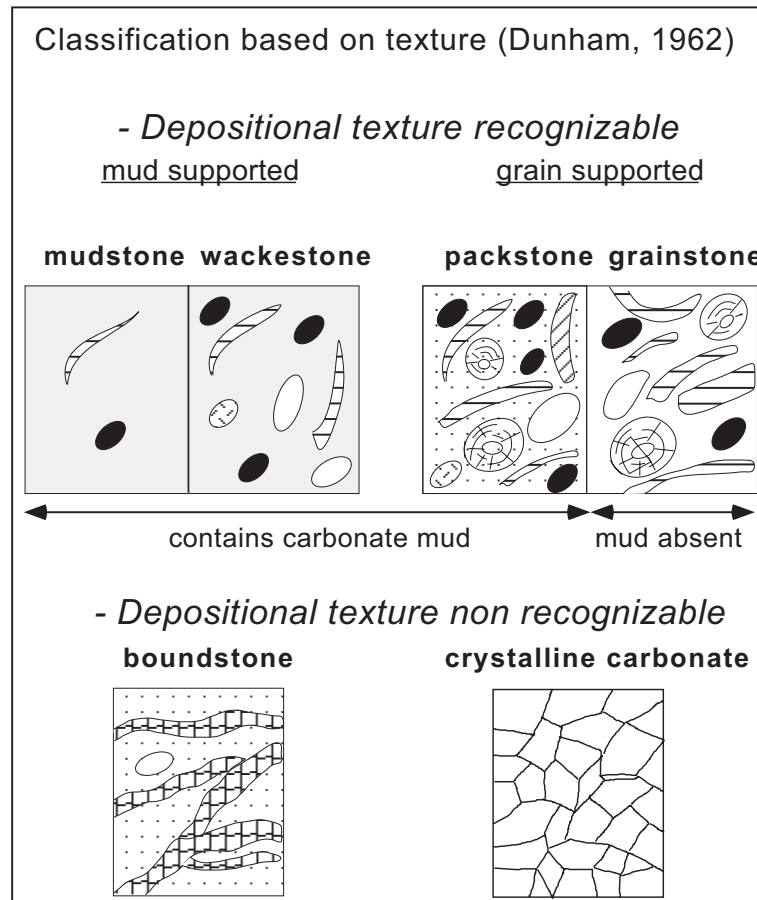


Figure 1: The Dunham classification of carbonate rocks

## 1.2 Well 2

More than a hundred samples of 2 cm diameter from three depth interval groups were collected from whole cores. The samples were cut in vertical and horizontal directions at each depth to determine orientation effects if any. The samples spanned a shoaling-upward sequence of mudstones and wackestones to open-marine packstones and grainstones. The cores had suffered little diagenesis.<sup>1</sup> The grains are either pellets or bioclasts (algae, foraminifera, globigerinidae), and range from 50–200  $\mu\text{m}$  in size. Recrystallisation occurs sporadically and is mostly inside the bioclasts, and occasionally between the grains.

In this sequence, permeability generally decreases with depth while the porosity shows a slight increase with depth, consistent with the observed trend in facies of grainstone to mudstone. Local sharp decreases in permeability correspond probably to sea-level fluctuations between the deposition of different sedimentary sub-units inside the sequence.

<sup>1</sup>In this article, the term *diagenesis* refers to the natural changes that occur after the deposition of a sediment and are caused by changes in temperature, pressure, and/or weathering; see Murray and Pray (1965) for exhaustive definition.

## 2 Pore space partitioning

Quantitative petrography was one of the steps in a sequence of several petrophysical measurements. To ensure that different types of data were obtained on samples of close proximity, we used a multistep sampling procedure. From the whole core, a 3 cm diameter plug was cut orthogonal to the wellbore. The plug was then slabbed for petrography work, and the butt portion was used to core a 2 cm cylinder, about 3.75 cm long. The remaining length was used to make thin sections and chips for mercury porosimetry (see Figure 2).

Given that carbonate rocks are extremely complex, any attempt at pore space partitioning is likely to be inexact. Our purpose here was to identify the main types of porosity contributors to enable us to compare them with our petrophysical interpretation and measurements. Since many of the petrophysical data are sensitive only to gross features, the main emphasis in petrography was to quantify the three main contributors to porosity. These were namely the intergranular ( $f_m$ ), intragranular ( $f_\mu$ ), and vuggy fractions ( $f_v$ ). The precise definitions of these quantities is elaborated further below. Subtle variations from this gross picture were not taken into account. As illustrated in Figure 2, the three fractions of porosity ( $f_v$ ,  $f_m$ , and  $f_\mu$ ; see Section 2.2 for a definition of these quantities) are from the slabs and the thin sections, while permeability, capillary pressures, NMR relaxation, and He porosity are on the plugs. The 2 cm cylindrical core was used for measurements such as permeability, resistivity and NMR relaxation. For acoustic velocities, additional larger size plugs were cut, whenever it was feasible to do so .

Choquette and Pray (1970) stated the fact that the wide range of sedimentary carbonate particles is one of the reason for the complexity of the carbonates pore space. They proposed a very precise classification by focusing on the concept of *fabric selectivity*, emphasizing the geological differences between carbonates and siliciclastic rocks porosity origin. They refers to the porosity as *fabric selective* when a dependent relation can be found between porosity and the fabric elements (*i.e.*, the *grains*) and as *not fabric selective* when such a relation cannot be established easily. Inside these two broad classes, they distinguished 7 fabric selective pore types (interparticle, intraparticle, intercrystalline, moldic, fenestral, shelter, and growth-framework) and 4 not fabric selective pore types (fracture, channel, vug, and cavern). On top of this classification, they also described modifications processes like *solution*, *cementation*, or *vadose circulation*, that can affect widely the porosity and other petrophysical parameters of carbonates such as permeability or acoustic properties.

Lucia (1983) proposed a simpler classification based more on the petrophysical aspects of the carbonate pore space components. Three classes of pore types were identified as mainly contributing to the variations observed in the porosity, permeability, capillarity and Archie's  $m$  values: *interparticle porosity*, defined as pore space between the particles—regardless of the nature of the particle, pellets, bioclast, or crystal—and *all other pore space*, called *vuggy porosity*. He divided the latter into *touching* vugs, which tend to improve the overall connectivity of the pore space, and *separate* vugs, which on the contrary, reduce the permeability while increasing the total porosity.

In this study, the pore space component classification is made in a slightly different way. Our goal is primarily to be able to account for the variations of such different measurements as acoustics, T2 distribution and permeability, without changing the petrological classification. Further on, we will refer to '*grain*' as the elementary particle in a carbonate rock, and will specify as necessary if it is a pellet, a bioclast or a crystal.



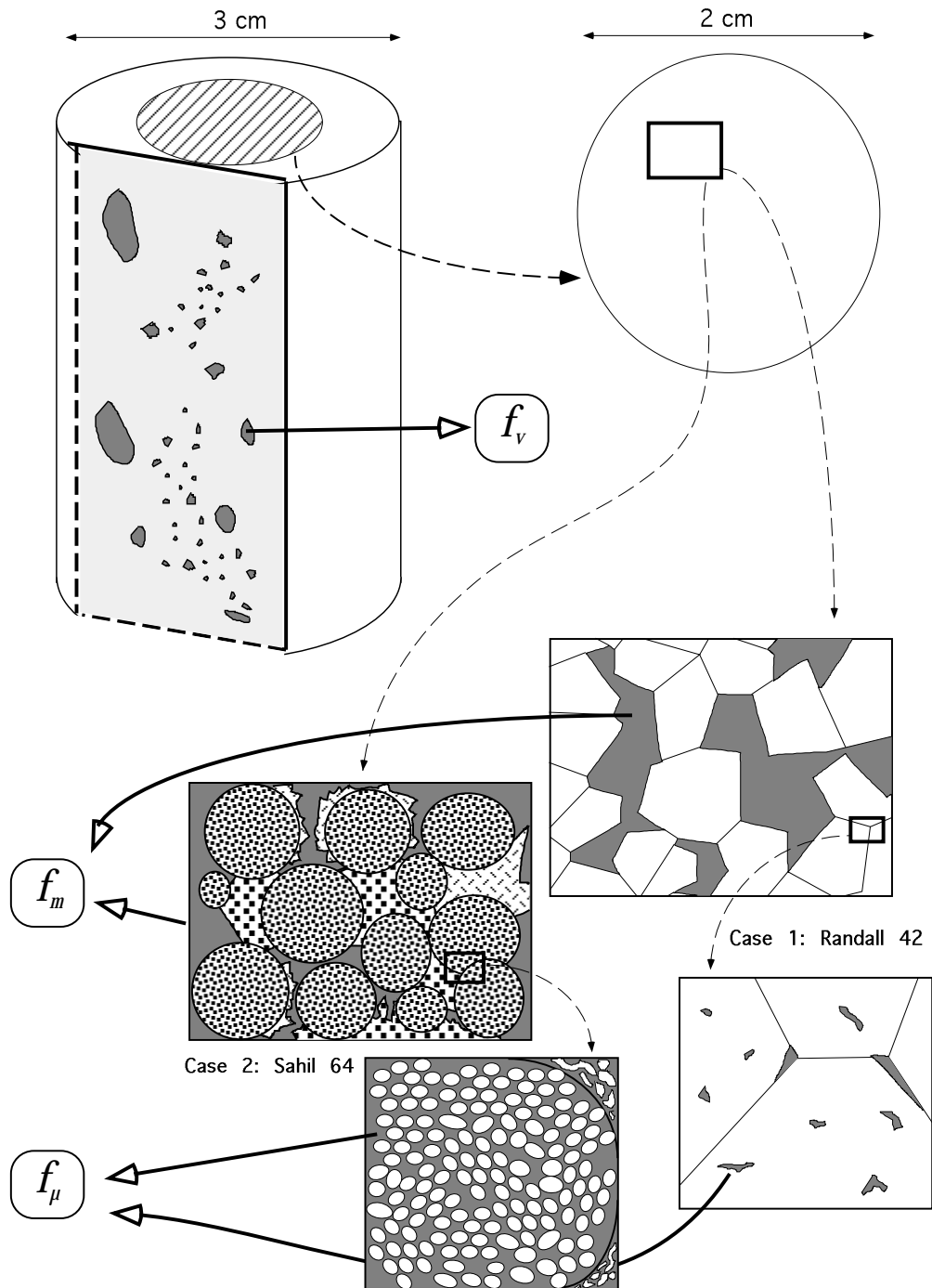


Figure 2: Method of extraction of the three fraction of porosity —  $f_v$ ,  $f_m$ , and  $f_\mu$  — from plugs, slabs and thin sections taken out from whole cores. For each sample, the thin section was made from the bottom or the top of the plug, thus allowing direct comparison between petrophysical measurements from plugs and quantities extracted from image analysis on thin section views.

## 2.1 The concept of *fraction of porosity*

For simplification purpose, we have broken the total volume of the rock into three *volume fractions*:

- the *vuggy* fraction, noted  $f_v$ , which is the volume occupied by the vugs. We are not yet discerning between touching or non-touching vugs, and we will call *vug* a pore that we can easily see with the naked eye.
- the *macro* fraction, noted  $f_m$ , which can be assimilated to the volume between the grains.
- the *grain* fraction, noted  $f_g$ , which is the volume occupied by the grains.

These three quantities verify the simple relation:

$$f_v + f_m + f_g = 1$$

Let  $\phi_v$  and  $\phi_m$  be the intrinsic porosities of respectively a vug and a macropore. We can write that:

$$\phi_v = \phi_m = 1$$

A simple formula can link these three quantities to the total porosity  $\phi_T$ :

$$\phi_T = f_v + f_m + f_g \phi_\mu$$

where  $\phi_\mu$  is the *intrinsic porosity* of the grain, or *intrinsic microporosity*. The fraction of microporosity  $f_\mu$  is then:

$$f_\mu = f_g \phi_\mu$$

Further on, the raw fractions determined from Petrographic Image Analysis will be noted as  $\hat{f}_v$ ,  $\hat{f}_m$  and  $\hat{f}_\mu$  to emphasize that they are *non corrected* fractions. In Section 3.2.4, we explain how to correct the raw numbers in order to come back to the general notation defined above.

## 2.2 Definitions

As we are using Image Analysis to quantify the fractions of porosity, the critical parameter will be the *size* of the pore. The following *size thresholds* are used:

- *vuggy* fraction: all pores larger than  $200\mu\text{m}$  in size, that can easily be seen with the naked eye.
- *macro* fraction: all pores between  $200\mu\text{m}$  and  $7\mu\text{m}$  in size, seen on thin section using an optical microscope.
- *micro* fraction: all pores smaller than  $7\mu\text{m}$  in size.

As the only criteria is the size, there is a need to define carefully the different pore types that we are using, not only on a geological aspect, but also on a petrophysical aspect.

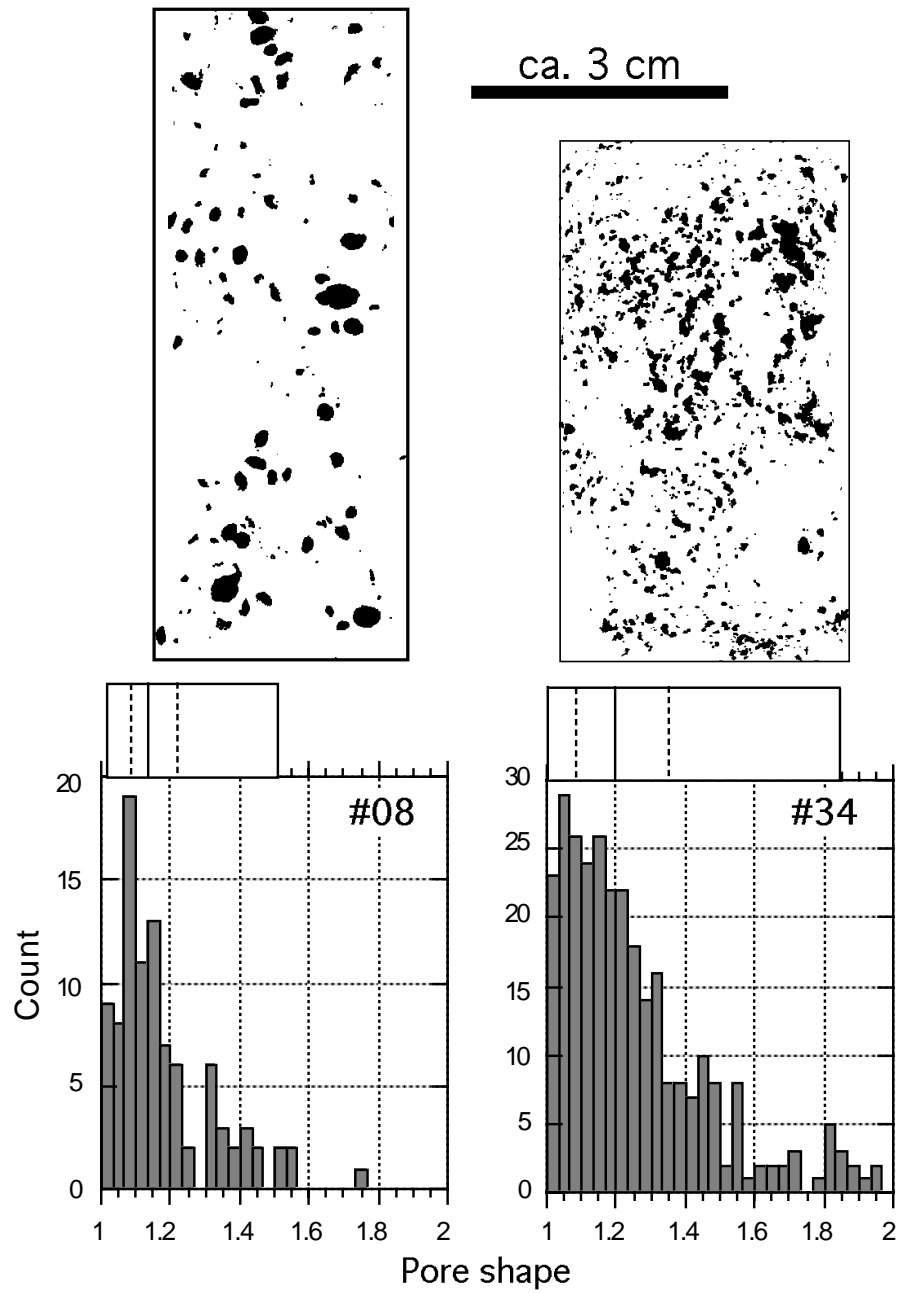


Figure 3: Example of different vug shape. Both samples are taken from the Well 1 core set. Sample 08 has rounded separate vugs, most of them having a pore shape between 1.1 and 1.2 (dashed lines are 25% and 75% of the data). Sample 34 shows very connected vugs, most of them having a complicated shape of 1.1 to 1.35, but reaching almost 1.9.

### 2.2.1 Vuggy pore space

Choquette and Pray (1970) proposed a quite restrictive definition of a vug, as it is a pore “(1) somewhat equant, or not markedly elongate, (2) large enough to be visible with the unaided eye, and (3) does not specifically conform in position, shape or boundaries to particular fabric elements of the host rock (*i.e.*, is *not* fabric selective)”. They distinguished it from a mold, which is the removal of a former individual constituent from the host rock, and is thus considered as being fabric selective. In fact, a solution-enlarged mold, as long as its precursor’s identity is lost, may be called a vug.

Lucia (1983) assimilated all “pore space larger than or within the particles of rock and commonly present as leached particles, fractures, and large irregular cavities” to a vug. These pore types, completely different in a geological sense, are indeed playing a similar role in terms of flow properties of the rock. Lucia (1983) then distinguished between *touching* vugs and *separate* vugs. He stated that the effect of adding separate vugs to a non vuggy rock is to increase the porosity, with little or no increase in permeability. On the other hand, adding touching vugs will give a higher permeability than expected if all the porosity were interparticle.

Figure 3 shows two samples from Well 1. Sample 08 has a vuggy fraction  $f_v$  equal to 5%, while sample 34 vuggy fraction is 10%. The difference stands in the shape of the vugs. The *pore shape*  $\gamma$  is a parameter defined by Anselmetti et al. (1996) as:

$$\gamma = \frac{Perimeter}{2\sqrt{\pi Area}}$$

normalized to give a convenient value of 1 for a circle. It has a value of 1.9 for an interparticle pore between spherical grains, and cracks or fractures may have values of  $\gamma > 5$ .

It is obvious in Figure 3 that most of the vugs in sample 08 are *separate* vugs, while they are more connected in sample 34. This is confirmed by the distribution of the pore shape for the two samples, showed below each slab picture. Most vugs in sample 08 have a shape between 1.1 and 1.2, which is those of an elongated circle. The shape of most vugs from sample 34 is between 1.1 and 1.4, but can be 1.8, which is those of an interparticle pore (see above). This difference in the shape of the pore will matter in acoustics because this petrophysical measurement is *sensitive to the shape*, acoustics, and we will have to discuss the results qualitatively. In the case of T2 distribution however, the *volume of the pore* is the important criteria, and thus the way we measure the vuggy fraction of porosity will be valid.

### 2.2.2 Interparticle pore space

As the criteria in also the *size*, interparticle porosity will be the pore space *between the grains*. The nature of the grain (pellet, bioclast, crystal) will not be taken into account although it can be of some importance in acoustics or NMR, and will then be considered as a qualitative variable. In the case where the grains are micritized, and the interparticle pore space is filled with micritized matrix, all the pore space will be accounted as *microporosity* (see below).

### 2.2.3 Intraparticle pore space

Intraparticle porosity will be the pore space inside the grains. Again, describing the shape and nature of these pores will result in a qualitative parameter. Samples from Well 1 presents isolated

rounded solution-induced micropores that we can define as *microvugs* (see Choquette and Pray (1970) for a first use of this term), while in samples from Well 2, the microporosity forms a complicated network due to the micritization of the grains. The microporous network is connected through the interparticle pore space if present (in the case of a grainstone), or forms a one-scale network if this space is filled with micritized matrix (case of mudstone).

### 3 Pore space component measurements

#### 3.1 Data acquisition

We used the standard approach followed by Anselmetti et al. (1996), with some changes and improvements, and introduced the determination of the vuggy fraction  $f_v$  from rock slabs.

As explained in Figure 2, from the original 2" diameter plug drilled from the whole core, a slab face is cut and a 1" diameter plug sub-cored. The latter is devoted to the measurements of petrophysical properties, while the slab is used to measure  $f_v$ . 30  $\mu\text{m}$ -thick thin sections were prepared from the end pieces of the 1" diameter plugs, after impregnating them with blue-dyed epoxy (see also Ehrlich et al. (1991), Gerard et al. (1992)).

##### 3.1.1 Determination of $f_v$

**Well 1** The slab face is polished on a grinding wheel in order to obtain a uniform color on its surface. Then we use *red Play-Doh paste* to fill the vugs. The filling is made in such a way that we have at the end a flat surface with the vugs in red, and the polished rock in a contrasted uniform color. We will refer to this technique throughout this article as the "*silly putty*" technique. The crystalline dolomite from allows polishing and filling without modifying shape and size of the vugs. We then digitize the surface in a 24 bits color image file using a regular flatbed scanner, at a resolution of 300 dpi.

**Well 2** Vugs are present during a few feet at the top of the sequence. The loose texture of these carbonates make it very difficult to use the slab method because polishing and filling operations can result in enlarging the vug, destroying together its original shape and size. We decided to use as a rough approximation the CT Scan images.

##### 3.1.2 Determination of the interparticle and intraparticle fractions of porosity

**Well 1** Interparticle and intraparticle pore networks are very well individualized both in size, location and shape in samples from Well 1. We used the Optical Microscope (equipped with a Toshiba CCD camera) to characterize the macroporosity. In order to avoid bias, we took 6 views randomly distributed across each thin section, using a x5 magnification objective. In case of a thin section presenting a big heterogeneity in the size and location of the macropores, we took 5 additional views at lower magnification (x1). A good contrast between pore space and solid phase is provided by the blue-dye epoxy filling the pore space.

We used the Environmental Scanning Electron Microscope (ESEM) to characterize the microporosity, by taking 10 "pseudo-random" image views at a magnification of x2000 (the term "pseudo-random" means that the original locations of the views were randomly taken, but then adjusted to avoid the presence of macropore in the views). The contrast between pore space and

solid phase is mainly a density contrast between the two phases (the epoxy appearing black while the rock is grayish). The pore space present in the image is  $\hat{f}_\mu$ .

**Well 2** The two types of pore space seen in the thin section *i.e.*, inter (macro) and intragranular (micro) are not easily distinguished, especially based on size alone. In a grainstone, the distinction is made depending on whether the pores exist between grains or within grains. In a wackestone/packstone, one has to clearly identify the grain first, and then make a judgement on whether the pore is within the grain or outside. Normally when micrite particles partially fill the interparticle pore space, we assign the porosity among the micrite particles to be intragranular. This method of partitioning was chosen not from a geological point, but from a petrophysical and transport calculation view. The quantification presented here maps easily to a simple, but consistent pore structure model that enables us to predict resistivity and permeability (see chapter xxx). By no means is this a rigid and complete framework for predicting flow behavior of these rocks. Occasionally deviations between predicted and observed petrophysical properties are explained through qualitative description of features that are not quantifiable through the classification employed here.

Because of the micritization of the grains and of the matrix, the measurement of the macrofraction of porosity from Optical Microscope views of thin sections will suffer a big uncertainty. In fact, the optical method uses transmitted light, thus penetrating the entire 30  $\mu\text{m}$  of the thin section, and the small micrite particles (4 to 10  $\mu\text{m}$ ) are diffusing the incoming light, greatly lowering the contrast between the blue-dye epoxy and the rest of the thin section. In order to avoid bias we acquired 6 randomly distributed views (24 bits RGB color) at x5 magnification.

The similarities in size and shape of the interparticle and the intraparticle pore networks prevent to use the ESEM to simply count the microporous space. We chose to use the ESEM at a lower magnification (x300) and to measure the *total image porosity*, which can be symbolized by  $\phi_I$ . The image view at that magnification is a square of 320x320  $\mu\text{m}$  and thus includes both interparticle and intraparticle networks. 10 randomly distributed views were taken across the thin section, and their results averaged to give a value of  $\phi_I$ .

In all cases, the final value for the fraction of porosity was obtained by averaging the results from the several views acquired across the thin section.

## 3.2 Petrographic Image Analysis

We used a standard image processing package (such as Adobe Photoshop + NIH Image). The acquired digital images are then *binarized* into a pore space and a solid/matrix phase, using a simple *threshold*. The fidelity of the thresholded image is verified by visual comparison with the original image. Each particle (= pore) is individualized, and occasionally slight noise removal and smoothing was applied. Several parameters are then computed for each particle, such as *area*, *perimeter*, *major* and *minor elongation axis*. The sum of areas of all the particles in an image represents an estimate of the pore space and is noted  $\hat{A}_{pores}$ .

### 3.2.1 Slab face scanning image processing

Red paste was chosen to fill the vug because this color is absent in the range of color of carbonate rocks. In a 24 bits RGB color image, each pixel has a value from 0 to 255 for Red, Green and Blue.

In the case of rock images, the RGB values for the background are quite similar (assuming that it has been color smoothed by polishing), while the vugs have a high red value and small green and blue values. Subtracting the green value out of the red for each pixel in the image will result in a black background and light grey vugs. Then a simple *threshold* will give the *binarized image*, that is used to compute  $\hat{f}_v$ :

$$\hat{f}_v = \frac{\hat{A}_{pores}}{A_{total}}$$

where  $A_{total}$  is the total area of the binarized image. We used a threshold of  $200\mu\text{m}$  for the minimum size of detected particle.

### 3.2.2 Optical Microscope view processing

As the pore space is filled with blue-dye epoxy, the same technique explained above applies. Here, subtracting the red value out of the blue for each pixel in the image will result in a black background and light grey pores. The thresholded image is binarized and all the particles larger than  $7\mu\text{m}$  are counted.  $\hat{f}_m$  is computed as

$$\hat{f}_m = \frac{\hat{A}_{pores}}{A_{total}}$$

### 3.2.3 Environmental Scanning Electron Microscope view processing

The overall good contrast between the pore space (seen in black) and the matrix (seen in grayish colors) in the ESEM images allows direct threshold and binarization to compute:

- $\hat{f}_\mu$  in case of Well 1, expressed as

$$\hat{f}_\mu = \frac{\hat{A}_{pores}}{A_{total}}$$

- $\phi_I$  in case of Well 2. The microfraction of porosity is then derived by doing

$$\hat{f}_\mu = \phi_I - \hat{f}_m$$

### 3.2.4 Corrections applied

Some corrections have to be applied on these raw numbers, depending on the way the data acquisition is accomplished and also on the way the three fractions of porosity are computed. The basic relationship

$$f_v + f_m + f_g = 1$$

is verified in all cases.

In all cases also, the measured  $\hat{f}_v$  will be equivalent to the true vuggy fraction and we can write that

$$\hat{f}_v = f_v$$

**Well 1** We measure the interparticle porosity in the thin section, assuming that there is *no* vug present in the image. Thus the true value of the macrofraction of porosity is

$$f_m = \hat{f}_m(1 - f_v)$$

In the same way, by avoiding interparticle porosity, the measure fraction  $\hat{f}_\mu$  is the *intrinsic microporosity* of the grain/crystal, noted  $\phi_\mu$  (see Section 2.1). Then, the true microporosity is:

$$f_\mu = \hat{f}_\mu(1 - f_m - f_v) = \phi_\mu(1 - f_m - f_v)$$

The *corrected* total image porosity  $\phi_{TPIA}$ , comparable to the total Helium porosity  $\phi_{He}$  measured on plugs, will be

$$\phi_{TPIA} = \underbrace{\hat{f}_v}_{f_v} + \underbrace{\hat{f}_m(1 - \hat{f}_v)}_{f_m} + \underbrace{(1 - \hat{f}_m(1 - \hat{f}_v) - \hat{f}_v)}_{f_g} \hat{f}_\mu$$

**Well 2** The interparticle fraction of porosity  $\hat{f}_m$  is measured on thin section, assuming that there is *no* vugs present. As in the case of Well 1,  $f_m$  will be

$$f_m = \hat{f}_m(1 - f_v)$$

We measure the *total image porosity* noted  $\phi_I$ , from the ESEM view at magnification x300. We can say that

$$\phi_I = f_m + f_g \phi_\mu$$

The true *microfraction* of porosity noted  $f_\mu$  is then:

$$f_\mu = f_g \phi_\mu = \phi_I - \hat{f}_m(1 - f_v)$$

The *corrected* total image porosity  $\phi_{TPIA}$ , directly comparable to the total Helium porosity  $\phi_{He}$ , is

$$\phi_{TPIA} = \phi_I + f_v$$

## 4 Results and discussion

### 4.1 Well 1

**Comparison between  $\phi_{TPIA}$  and  $\phi_{He}$ .**  $\phi_{TPIA}$  obtained from Petrographic Image Analysis show a good agreement with the conventional method (Helium volumetry), which validates the partitioning of the porosity into three volume fractions. Most of the samples where  $f_v$  represents more than 25% of  $\phi_{TPIA}$  fall away from the 1:1 line, this discrepancy possibly being the result of bad statistics during the measurement of  $f_v$  (we have only one measurement of  $f_v$  from one slab face). Another source of error is the assumption of an isotropy in the third dimension by plotting the fractions measured on 2D sections with volumes (Helium porosity).



**Correlation between total porosity and permeability.** Solving the differential equations allows to find the effective permeability of a medium with a specified  $f_v$ , that is

$$k = \frac{k_m}{1 - 3f_v}$$

As demonstrated in Chapter XXX,  $k_m$  can be expressed as follows:

$$k_m = K \cdot \phi_{interparticle}^m \cdot r^2$$

Where  $K$  is a constant,  $m$  is the cementation exponent, and  $r$  is the average size of the interparticle pores. As  $\phi_{interparticle} = \phi_m = \phi_{THe} - f_v$ , we can write

$$k_m = K \cdot (\phi_{THe} - f_v)^m \cdot r^2$$

In order to simplify,  $r$  is considered constant for a given lithotype, and  $m$  is roughly equal to 2.  $k$  is plotted as

$$k(1 - 3f_v) = f(\phi_{THe} - f_v)^2$$

The permeability  $k$  plotted against  $\phi_{He}$  correlate fairly well, showing ideally that estimating  $f_v$  allows to consider the medium as a single pore type network, where the behaviour of the petro-physical properties is easier to model.

## 4.2 Well 2

**Comparison between  $\phi_{TPIA}$  and  $\phi_{He}$ .**  $\phi_{TPIA}$  is in good agreement with  $\phi_{He}$ . To evaluate  $f_v$ , and because the results given by the “silly putty” technique were suffering from too much uncertainty, we used CT Scan images that we analyzed exactly in the same way as our Optical Microscope images. However, these images were taken at a quite large interval (0.5 to 1 foot), and at different depth than the samples’. Considering the common heterogeneity of this pore type, inferring a possible vuggy fraction for the depth locations of our samples is difficult and can reveal itself inaccurate. That leads to the uncertainty affecting the samples which present vugginess.

Considering the evolution of the three fractions of porosity according to depth in Well 2, the microfraction of porosity  $f_\mu$  is increasing with depth, as the intergranular fraction  $f_m$  is decreasing, as expected in a shoaling-upward sequence.  $f_v$  is a rough estimate and only present in the grainstone at the top of the sequence. The agreement between  $\phi_{He}$  and the sum of the three fractions of porosity is better at the bottom of the sequence (mud-supported lithotypes). This is due to the uncertainty on the measurement of  $f_m$ , which is much larger than those of  $f_\mu$  (due to the technique of measurement; see Section 3.1.2). The grain-supported lithotypes having by definition a larger  $f_m$ , the total uncertainty on the measurement of  $\phi_{TPIA}$  is larger than for the mud-supported lithotypes.

**Correlation between total porosity and permeability.** Despite a larger scattering of the Petrographic Image Analysis data, it is obvious that there is *strong textural dependence of permeability*. The porosity of the mud-supported lithotypes (wackestones and mudstones) is mainly supported by a single pore-type network. Since all intergranular pore space has been filled by micritic cement (with microporosity) or sparite cement (without microporosity), the resulting pore network is intergranular, where the grain is here a *micrite particle*, leading to a direct correlation between  $k$  and the total porosity. The grain-supported lithotypes (grainstones and packstones) have a two-pore types network, *inter-* and *intragranular*.  $k$  is then not directly dependent on the total porosity.

**Fractions of porosity and lithotypes** Observation and description of the thin sections with both the Optical and the Environmental Scanning Electron microscope lead to the classification of the 53 samples into the 4 basic carbonates lithotypes defined by Dunham in 1962 (See Figure 1). The scattering of the data around the 1:1 line, illustrating the fact that the way we measure  $f_m$  is quite inaccurate, is however less when going from the grainstone lithofacies, to the mudstone lithofacies, that is when the part of  $f_m$  in the total porosity is decreasing.

### 4.3 Implications for NMR relaxation in carbonates

Two common assumptions are made in sandstones for interpretation of NMR relaxation. The first is that magnetization relaxation in each pore is controlled by surface relaxivity (Sen et al. 1990; Kleinberg et al. 1994), i.e.  $\frac{V_s \rho}{\mathcal{D}} \ll 1$ , where  $\mathcal{D}$  is the molecular diffusion coefficient,  $\rho$  the surface relaxivity and  $V_s$  the pore volume-to-surface ratio. The characteristic transverse relaxation constant for the pore then becomes

$$\frac{1}{T_2} = \frac{1}{T_{2b}} + \frac{\rho}{V_s}$$

where  $T_{2b}$  is the bulk relaxation of the fluid. Its effect is normally ignored, since it can be easily accounted for (Wilkinson et al. 1991). The second assumption is that in multipore systems, each pore act independently, and that the measured magnetization is an integral response of all the pores.

In carbonates, especially in grainstones and packstones, these assumptions fail due to significant diffusion of magnetic moments between inter- and intra-granular pore space, whose we have seen can see their characteristic sizes may vary by an order of magnitude or more. Thus the direct link between  $T_2$  and pore size is lost. In mudstones, only one pore type is present (intergranular, between the micrite crystals), and the problem can roughly be solved as for sandstones.

The petrography results in grainstones and mudstones, presented in the previous sections, are indicative of a bimodal pore size distribution, absent in the  $T - 2$  spectra. This is explained by the concept of enhanced relaxivity (Ramakrishnan et al. 1998). The diffusion of magnetic moments from a fraction  $f_m$  into the intragranular pore space leads to an enhanced relaxation of magnetization than if two types of pores were isolated from each other. Solving for the quasisteady equation gives an apparent relaxivity  $\rho_a$  for intergranular pores

$$\rho_a = \rho(1 - \phi_\mu) + \left[ \sqrt{\frac{\phi_\mu \mathcal{D}}{F_\mu T_{2\mu}}} \coth\left(\sqrt{\frac{\phi_\mu \mathcal{D}}{\mathcal{D} T_{2\mu}}} R\right) - \frac{\mathcal{D}}{F_\mu R} \right]$$

where  $F_\mu$  is the formation factor of the grain,  $T_{2\mu}$  the relaxation time constant within the grain, and  $R$  the grain radius. Substituting typical numbers suggests for a typical grainstone an order of magnitude increase in relaxivity from intrinsic  $\rho$  values. Given the above equation, it is possible to write an approximate expression for relaxation in carbonates with micro and macro porosity:

$$M(t) = (\phi - f_m) \exp\left(-\frac{t}{T_{2\mu}}\right) + f_m \exp\left(-\frac{\rho_a t}{V_{sm}}\right)$$

where  $V_{sm}$  is the ratio of the volume of the macri pore to the external surface area of the grain in the continuum representation (Ramakrishnan et al. 1998). This is the basis of the two-exponential fit to the relaxation data in carbonates. However, because  $\rho_a$  is an order of magnitude larger than  $\rho$ ,

the time scale for macro pore relaxation is not significantly different from  $T_{2\mu}$ . This approximation then introduces errors. Furthermore, the boundary conditions are not satisfied and our estimates based on the two exponential approximation to laboratory data are larger than the petrography data, as confirmed by numerical simulations (“*random walk*”). These numerical simulations can correct for that by estimating an overlap volume to be subtracted from the measured value of  $f_m$  to obtain the corrected one. Because the intergranular pore may have a complicated geometry, the quasistatic approximation may be extended to include transient terms within the grain.

In conclusion, once the facies identification step is carried out, whenever it is recognized that a grainstone/packstone is present, a model based inversion for  $f_m$ ,  $T_{2\mu}$  and  $V_{sm}$  may be implemented. These can be readily incorporated into transport calculation. If the lithotypes belongs to a wackestone or a mudstone, Ramakrishnan et al. (1998) find that a two exponential fit is sufficient to deduce the pore sizes. For a transport calculation, they use an average of the two sizes.

## References

- Anselmetti, F., S. Luthi, and G. Eberli (1996). A study of carbonate “end-member” rocks—Part 1: Petrographic Image Analyses. Research Note ISD-001-96-05, SDR. Confidential.
- Choquette, P. and L. Pray (1970). Geologic nomenclature and classification of porosity in sedimentary carbonates. *AAPG Bulletin* 54, 207–250.
- Dunham, R. (1962). Classification of carbonate rocks according to depositional texture. In W. Ham (Ed.), *Classification of carbonate rocks—a symposium*, Volume 1, pp. 108–121. Am. Assoc. Petroleum Geologists Mem.
- Ehrlich, R., S. Crabtree, K. Horkowitz, and J. Horkowitz (1991). Petrography and Reservoir Physics III: Objective classification of reservoir porosity. *AAPG Bulletin* 75, 1547–1562.
- Folk, R. (1959). Practical classification of limestones. *AAPG Bulletin* 43, 1–38.
- Gerard, R., C. Philipson, F. Manni, and D. Marshall (1992). Petrographic Image Analysis: An alternate method for determining petrophysical properties. In P. I. and S. S.K. (Eds.), *Automated pattern analysis in petroleum exploration*, pp. 249–263. Springer-Verlag.
- Kleinberg, R. L., W. E. Kenyon, and P. P. Mitra (1994). On the mechanism of NMR relaxation of fluids in rocks. *Journal of Magnetic Resonance* 108A(2), 206–214.
- Lucia, F. (1983, March). Petrophysical parameters estimated from visual description of carbonate rocks: a field classification of carbonate pore space. *Journal of Petroleum Technology* 35, 626–637.
- Lucia, F. (1995). Rock-fabric/petrophysical classification of carbonate pore space for reservoir characterization. *AAPG Bulletin* 79(9), 1275–1300.
- Murray, R. and L. Pray (1965). Dolomitization and limestone diagenesis, an introduction. In L. Pray and R. Murray (Eds.), *Dolomitization and limestone diagenesis—a symposium*, Volume 13, pp. 1–2. Soc. Econ. Paleontologists and Mineralogists Spec. Pub.
- Nurmi, R. (1984). Pore structure in carbonate rocks. *Schlumberger Technical Review* 32(1), 14–23.

Ramakrishnan, T. S., L. M. Schwartz, E. Fordham, W. E. Kenyon, and D. J. Wilkinson (1998). Forward Models for Nuclear Magnetic Resonance in Carbonate Rocks. In *Transactions of the SPWLA Annual Meeting*.

Sen, P. N., C. Straley, W. E. Kenyon, and M. S. Whittingham (1990). Surface-to-volume ratio, charge density, nuclear magnetic relaxation and permeability in clay-bearing sandstones. *Geophysics* 55(1), 61–69.

Wilkinson, D. J., D. L. Johnson, and L. M. Schwartz (1991). Nuclear magnetic relaxation in porous media: the role of the mean lifetime. *Phys. Rev. B* 44, 4960–4973.



# Synthèse et perspectives

Une conclusion (temporaire bien évidemment !) à l'issue de ce travail pourrait être : *“Et pourtant, ça marche !”*. En effet, après toutes les difficultés d'acquisition, les erreurs de mesure, de traitement, les incertitudes inhérentes aux méthodes d'inversion et aux méthodes de classification des données, il semble bien que retrouver une minéralogie, et partant de là, une lithologie, en continu, à partir de mesures de diagraphie, soit du domaine du possible. Ca n'est pas nouveau, d'autres personnes (bien que peu nombreuses) se sont penchées sur le sujet. Je pense en particulier à l'équipe du Professeur Peter Harvey, du Borehole Research Group de l'Université de Leicester (thèse de Jeremy Lofts en 1993, thèse de Carlos Gonçalves en 1996), ainsi qu'à Christine Lauer, du Laboratoire de Mesures en Forage de l'Université Aix-Marseille III, CEREGE, qui présente dans sa thèse une tentative plutôt réussie pour retrouver la minéralogie de formations argileuses à partir de propriétés physiques mesurées par diagraphie.

Les nouveaux développements dont ce travail fait état apportent des perspectives intéressantes, qui sont toutefois tempérées par d'inévitables problèmes qu'il est urgent de résoudre avant d'aller plus loin.

- Reconstruire un log minéralogique cohérent s'est révélé possible dans plusieurs cas (puits 948C, données sur carottes ; puis 948A, données de diagraphie en cours de forage ; puits Balazuc-1, données de diagraphie géochimique). Des méthodes différentes ont été utilisées, allant de l'inversion totale au réseau de neurones.
- La précision du résultat est fonction de la qualité des données dont on dispose. Dans le cas des données de diagraphie, la qualité réelle est assez difficile à estimer, car si la variation du diamètre du puits en est un bon indicateur, elle seule n'est pas suffisante. La vitesse de l'outil durant l'acquisition est aussi un facteur important pour tous les paramètres dont la précision dépend du temps de comptage (radioactivité naturelle ou induite). Il est malheureusement souvent passé sous silence ou bien oublié, en particulier dans le cas des opérations effectuées dans le cadre du programme ODP. De même, la régularité de l'opération de diagraphie, perturbée par les mouvements du bateau dus à la houle en mer, ou bien par les accidents géologiques le long du puits, sont un facteur d'incertitude. La technique de diagraphie en cours de forage (*“Logging- While-Drilling”, LWD*), expérimentée lors du Leg ODP 156 (auquel j'ai participé), a fourni des données de très bonne qualité. Elles permettent d'envisager sérieusement la possibilité de retrouver une colonne minéralogique dans un environnement aussi instable qu'un prisme d'accrétion, dans lequel les techniques de diagraphie au câble sont d'ordinaire presque inutiles, lorsque, comme c'est le cas pour les campagnes ODP, le puits loggé subit au préalable un carottage de plusieurs dizaines d'heure, avec tous les effets prévisibles sur l'état du puits. De plus la mesure en cours de

forage permet de “capturer” des événements qui ne sont visibles que sous les conditions *in situ*, comme des circulations de fluides. En 1997, le Leg ODP 171A a eu pour objectif l’exploration du prisme d’accrétion de la Barbade à l’aide exclusivement de la technologie LWD, ce qui révèle l’enthousiasme suscité par les données de diagraphie acquises en cours de forage lors de la précédente campagne 156 trois ans plus tôt.

- En ce qui concerne l’inversion minérale, la méthode de l’inversion totale a été utilisée avec succès dans plusieurs cas. Cette méthode permet de prendre en compte dans le traitement les incertitudes à la fois sur les paramètres du modèle, et sur les mesures, ce qui n’est pas possible avec des méthodes classiques comme les moindres carrés. Dans la problématique géologique, ceci est extrêmement précieux, car les formations étudiées sont souvent des mélanges de phases plus ou moins homogènes, d’où la nécessité de donner au modèle des paramètres souples, modifiables dans un intervalle fixé selon des critères précis (qui peuvent être statistiques, géologiques, . . . ). Un des problèmes inhérents à la méthode reste le choix de la solution *a priori*. Une tentative, qui a donné de bons résultats, a été d’effectuer une classification de type nuées dynamiques, et d’utiliser la composition moyenne de chaque groupe ainsi individualisé pour guider le choix de la solution *a priori*. Une comparaison récente entre deux méthodes différentes (inversion totale et carte auto-organisatrice) utilisées sur les données du puits Balazuc-1 (voir Annexe E) a montré qu’une telle approche était pertinente. De plus, la méthode de classification utilisée est une classification dite “dure”, où chaque individu appartient à un groupe et un seul. Disposer de probabilités d’appartenance permettraient de définir des proximités de composition entre les groupes, et rendraient la classification plus réaliste. Mais il faudrait alors trouver le moyen d’intégrer ces probabilités d’appartenance dans la méthode d’inversion, peut être sous la forme d’une pondération.
- Le réseau de neurones est sans aucun doute un très bon outil de traitement de données. Son avantage sur des méthodes d’inversion linéaire ou non linéaire est la possibilité de prendre en compte des variables non additives dans le traitement. Par exemple, l’outil géochimique ne mesurant pas le magnésium, la discrimination entre calcite et dolomite ne peut se faire sur la seule base chimique. L’index photoélectrique peut lever l’indétermination, car la valeur de la calcite (5,0) est très différente de celle de la dolomite (3,2), mais ceci n’est valable pratiquement que lorsque le carbonate est la phase majeure. Dans le cas où des phases à fort  $P_e$  sont présentes (comme la pyrite,  $P_e = 17$  barnes/e<sup>-</sup>, ou la barytine,  $P_e = 266$  barnes/e<sup>-</sup>), il faut évaluer leur influence sur le  $P_e$  total, ce qui est possible en utilisant le paramètre  $U$ , en fait  $P_e \times \rho_b$ , donnant une expression volumique du  $P_e$ . Mais une application originale du réseau de neurones pourrait être sa capacité à combler les intervalles, par exemple dans le cas d’une corrélation puits à puits. Dans une application récente (voir Annexe E, présentation à la 17<sup>ème</sup> Réunion des Sciences de la Terre, avril 1998), nous avons utilisé une carte auto-organisatrice de Kohonen pour retrouver la lithologie dans le puits Balazuc-1, à partir de l’entraînement sur des intervalles discrets reconnus sur carottes. Toutes les variables mesurées par diagraphie (depuis les pourcentages d’oxydes jusqu’à la densité et au facteur photoélectrique) ont été utilisées. Mis à part le fait qu’une carte simple (5 × 5) permet une reconnaissance quasi-parfaite de la lithologie, et donne un résultat identique à l’inversion totale, nous avons tenté avec moins de succès d’utiliser une carte plus complexe (30 × 30), dont l’avantage est en théorie de pouvoir affecter un nombre plus grand de neurones à un même faciès, montrant ainsi la variabilité des compositions des individus composant le faciès. Nous n’avons utilisé qu’un algorithme d’apprentissage simple, et donc de nombreux neurones se sont retrouvés isolés et non indexés. Il faudrait maintenant utiliser des méthodes plus probabilistes, capables d’affecter des faciès intermé-

---

diaires aux neurones isolés, par une interpolation entre des faciès plus homogènes reconnus à partir d'observations discrètes (carottes). C'est alors la capacité de prédiction du réseau de neurone, autrement plus intéressante que la classification pure et simple, qui serait ainsi mise à contribution.

- Un problème récurrent à toute inversion minérale est la *colinéarité de composition*, déjà citée dans de nombreux articles par P. Harvey, J. Lofts, et leurs collègues. Elle rend compte du fait que plusieurs minéraux correspondant à la même composition chimique peuvent être présents simultanément dans la formation, ce qui rend bien évidemment l'inversion minérale très difficile, et parfois impossible. Il est par exemple difficile d'estimer la proportion relative de plusieurs minéraux argileux, sur la seule base de composition en éléments majeurs (voir la discussion sur le prisme de la Barbade, section 7.2.1). Le choix de variables discriminantes est crucial. Le réseau de neurones peut être aussi un bon outil car il peut prendre en compte des paramètres qualitatifs, comme la résistivité, la densité ou la radioactivité naturelle, qui peuvent éliminer le problème de la colinéarité de composition.
- Enfin, un effort doit être entrepris pour améliorer la chaîne de traitement des données du GLT, depuis la régression linéaire utilisant les moindres carrés pondérés entre les spectres de capture et les spectres standard, jusqu'au modèle de fermeture souffrant de nombreuses imperfections, comme par exemple la prise en compte d'un seul oxyde par élément chimique (obligeant par exemple à faire un choix entre allouer le soufre soit à la pyrite, soit à l'anhydrite). Il faudrait mettre au point un modèle tenant compte *a priori* des informations données par des paramètres comme l'index photoélectrique, la densité, la radioactivité naturelle. Ce pourrait être une autre application des techniques neuronales.



## Annexe A

# Outils de diaggraphie et paramètres mesurés

Pour le fonctionnement et les principes physiques relatifs aux outils suivants, se référer aux Chapitres 1 à 3. Les acronymes de la colonne 2 dans les Tables suivantes A.1-A.3 sont ceux couramment utilisés par la société *Schlumberger*.

Diagraphie nucléaire			
<i>Mesure du rayonnement gamma naturel :</i> Natural Gamma-ray Tool (NGT)	Rayonnement gamma naturel total :	SGR	API <sup>1</sup>
	<i>Spectral Gamma-ray</i>		
	Rayonnement gamma naturel calculé :	CGR	API
	<i>Computed Gamma-ray</i>		
	Concentration en Potassium	POTA	poids%
	Concentration en Uranium	URAN	ppm
	Concentration en Thorium	THOR	ppm
<i>Les sondes gamma-gamma :</i> Hostile-environment Litho-Density Tool (HLDT)	Mesure de la masse volumique de la formation (appelée abusivement densité totale)	RHOB	g/cm <sup>3</sup>
	Mesure du facteur photoélectrique	PEF	barnes/e <sup>-</sup>
	Mesure du diamètre du trou (caliper)	CALI	pouces (in)
<i>Les sondes neutron-neutron :</i> Compensated Neutron Tool (CNT-G)	Mesure de la porosité neutron	NPHI	unité porosité (p.u.) ou %
	<i>Les sondes neutron-gamma :</i> Geochemical Logging Tool (GLT : NGT / CNT-G / AACT-Aluminium Activation Clay Tool / GST-Geochemical Spectrometry Tool)	Mesure de la concentration en Ca, Si, Fe, Ti, S, Al, Mg, H, Cl, Gd	— poids % d'oxydes ou ppm selon l'élément

TAB. A.1 – Outils et principaux paramètres mesurés par diagraphie nucléaire.

Diagraphie électrique			
<i>Les dispositifs à induction :</i> Induction Tool (DIT-E)	Dual	Mesure de la résistivité de la zone non-envahie par les fluides de forage : Induction à grande profondeur d'investigation	ILD $\Omega.m$ ou IDPH
		Mesure de la résistivité de la zone intermédiaire : Induction à moyenne profondeur d'investigation	ILM $\Omega.m$ ou IMPH
		Mesure de la susceptibilité magnétique	SUSC ppm SI
<i>Les dispositifs à focalisation sphérique :</i> Micro Spherically Focused Log (MSFL)		Mesure de la résistivité de la zone envahie $R_{xo}$ : faible profondeur d'investigation	SFLU $\Omega.m$

TAB. A.2 – Outils et principaux paramètres mesurés par diagraphie électrique.

Diagraphie magnétique			
<i>Magnétomètre à résonance magnétique nucléaire :</i> Resonance Sonde (NMRS)		Mesure de l'induction magnétique totale	TOTF nT

TAB. A.3 – Outils et principaux paramètres mesurés par diagraphie magnétique.

Mesure	Résolution intrinsèque de l'outil	pas d'échantillonnage	Résolution optimale réelle reportée
Formation Micro Scanner FMS	5 mm	3 mm	$\leq 5$ mm
Spherically Focused log SFLU	76 cm	15 cm	76 cm
Phasor Induction log DIT			
grande profondeur	2 m	15 cm	2 m
moyenne profondeur	1,5 m	15 cm	1,5 m
LithoDensity log HLDT			
densité	38 cm	15 cm	46 cm
$P_e$	5 cm	15 cm	46 cm
Compensated Neutron log CNT	38 cm	15 cm	46 cm
Gamma Ray log NGT, GST	20-50 cm	15 cm	46 cm

TAB. A.4 – Résolution verticale et pas d'échantillonnage de quelques outils de diagraphie..

## Annexe B

# Glossaire de quelques termes particuliers utilisés en diagraphie

**API** abbréviation de *American Petroleum Institute*

**unité API** unité utilisée pour standardiser la mesure des outils nucléaires de mesure du rayonnement gamma naturel et de la porosité neutron. Pour le rayonnement gamma naturel : la différence entre deux zones calibrées dans le puits API (situé à Houston, TX) est 200 unités API. 1/200<sup>ème</sup> de cette différence est l'unité API. Pour la porosité neutron : la différence entre le zéro électrique (porosité zéro) et la mesure d'une lithologie de référence (calcaire de l'Indiana, USA) à 19% de porosité, représente 1000 unités API. 1/1000<sup>ème</sup> de cette différence est l'unité API.

**barn**  $10^{-24}$  cm<sup>2</sup>, c'est l'unité de mesure de la section de capture.

**darcy** unité de mesure de la perméabilité. Un darcy correspond à la perméabilité d'un milieu qui laisse passer un fluide d'une viscosité égale à un centipoise à travers un cm<sup>2</sup> sous un gradient de pression de une atmosphère à la vitesse de 1 mm par seconde. L'unité commune utilisée est le millidarcy (mD).

**effet environnemental du puits** influence sur la mesure de diagraphie des conditions de mesure dues principalement à la configuration du puits : sa rugosité, son diamètre, sa forme, le type de fluide de mesure employé, ou la présence de "mud-cake".

**densité électronique** population d'électrons par unité de volume.

**électron-volt (eV)** unité d'énergie équivalente à l'énergie cinétique acquise par un électron sous l'effet d'une différence de potentiel de 1 volt.  $1 \text{ eV} = 1.6 \times 10^{-19}$  volt.

**facteur de formation** mesure du rapport entre la résistivité d'une roche saturée d'eau à 100% avec la résistivité du fluide qui la sature. On mesure plus souvent le facteur de formation apparent, qui est fonction de la porosité, de la salinité du fluide, de la géométrie des pores, et de la concentration en argiles.

**facteur géométrique** pondération, dépendant de la géométrie de la mesure, déterminant comment les conductivités des différents milieux présents au voisinage d'un outil dans un puits affectent la mesure de la conductivité apparente.

**facteur pseudo-géométrique** coefficient utilisé pour estimer la résistivité de la zone non-envahie par le fluide de forage à partir de la mesure d'un Laterolog.

“**mud-cake**” résidu déposé sur les murs du puits de forage au fur et à mesure que le filtrat de la boue de forage remplit les pores de la formation.

**numéro atomique** le nombre de protons dans un noyau atomique, ou le nombre d'électrons orbitaux dans un atome neutre.

**poids atomique** le poids relatif d'un atome en prenant comme base relative l'atome de carbone égal à 12. Est égal au nombre de neutrons+protons du noyau de l'atome.

**porosité** fraction du volume occupé par les pores sur le volume total. La porosité *primaire* est la porosité d'un sédiment après compaction mais avant l'action de la diagenèse. La porosité *secondaire* est la porosité créée par l'action chimique des circulations de fluides par exemple (dissolution), la tectonique (fractures, ou la diagénèse (dolomitisation). La porosité *effective* concerne le volume des pores interconnectés pouvant contenir des fluides.

**propre (formation)** en langage pétrolier, se dit d'une formation qui ne contient pas de quantité d'argile appréciable

**profondeur d'investigation (grande)** la mesure des propriétés d'une formation géologique effectuée suffisamment loin du puits pour que les effets de la zone envahie soit minima.

**section de capture** 1. la section de capture nucléaire pour les neutrons est la sphère limite autour du noyau dans laquelle le neutron doit rentrer pour être capturé par le noyau. C'est une valeur probabiliste dépendante de la nature et de l'énergie de la particule incidente, ainsi que de la nature du noyau considéré. Elle est mesurée en barns. 2. la section *macroscopique* de capture est l'aire effective pour la capture des neutrons par unité de volume. Elle dépend donc du nombre d'atomes présents ainsi que de leur section de capture propre. Elle est mesurée en  $\text{cm}^{-1}$ , mais aussi en unités de capture, tel que  $1 \text{ c.u.} = 10^{-3} \text{ cm}^{-1}$ .

**zone envahie** la portion de la formation autour du puits dont la partie poreuse a été envahie par le fluide de forage.

## Annexe C

# Tableaux récapitulatifs (éléments majeurs et traces)

**Data Report :**

*Non publié*

Trace and major elements concentrations from ODP Leg 156 deep sea  
cores : a data report

**Alain Rabaute**, Peter Blum, Pierre Gaudon, James F. Allan

# Trace and major elements concentrations from ODP Leg 156 deep sea cores: a data report.

Alain Rabaute<sup>1</sup>, Peter Blum<sup>2</sup>, Pierre Gaudon<sup>3</sup>, James F. Allan<sup>2</sup>

<sup>1</sup> ISTEEM, CC 066, Laboratoire de Géochimie Isotopique, UMR 5567, Université Montpellier II, Place E. Bataillon, 34095 Montpellier Cedex 5, France.

<sup>2</sup> Ocean Drilling Program, 1000 Discovery Drive, Texas A&M University, College Station, Texas, U.S.A.

<sup>3</sup> Ecole Nationale Supérieure des Techniques Industrielles et des Mines d'Alès, Laboratoire P3MG, 6 Avenue de Clavières, F-30319 Alès Cedex, France.

## Abstract.

Geochemical trace and major elements concentrations measured on samples taken from sediment cores are presented here from two holes (948C and 949B) drilled in the North-Barbados accretionary prism during ODP Leg 156. These data include X-ray Fluorescence (XRF), Inductively Coupled Plasma Mass Spectrometry (ICP-MS), and Instrumental Neutron Activation (INA) analyses of the samples, providing accurate elemental concentrations in 10 major and 20 trace elements. The data set includes: 79 samples, each of them cut along a 20 cm interval in the section, in order to match corresponding gamma-ray spectra (see full study in Blum et al., this volume); and 113 samples from physical properties measurements, considered as representative of a particular lithology in the section. The sampling interval spans the decollement zone with a very good resolution in Hole 948C (almost one sample per meter), thus providing useful information on the effect of fluid flow on chemical and mineralogical nature of the formation.

## Introduction.

Leg 156 of the Ocean Drilling Program (ODP) cored two sites on the Northern Barbados ridge accretionary prism. At the opportunity of an investigation involving gamma-ray measurements on sediments from ODP Leg 156 (See Blum et al., this volume), 79 samples were taken from the cores recovered in Holes 948C and 949B. 113 discrete samples were also selected from the physical properties measurements sample set. Various geochemical measurements were carried out in order to obtain the elemental concentrations in K, U and Th, that were used as calibration data set (See Blum et al., this volume). As a standard output of these analyses, concentrations in 10 major elements, including K, and 20 trace elements, including U and Th are obtained. This data report presents the results of this chemical analysis of bulk sediments. A similar study was carried out during ODP Leg 110 (Wang et al., 1990) but we focused particularly our interest in the decollement zone.

In Hole 948C, the sampling interval ranges from 421.55 meters below sea floor (mbsf) to 591.43 mbsf, which corresponds to cores 948C-2X-01 to 948C-19X-06. The average core recovery was 94.9% and allowed encountered formations to be divided into two units. The first one (cores 2X-01 through 11X-05, 17 cm) consists mainly in claystone and silty claystone of late Miocene to middle Miocene age, occasionally crossed by local thin beds of altered tuffaceous siltstone and altered volcanish ash, and is characterized by changes in color used to describe lithologic variations. The second one (cores 11X-05, 17 cm, through 19X) consists in variegated claystone of early to middle Miocene age at the top, and below 523.78 mbsf, appears as an intercalation of carbonate-rich beds (nanofossil chalk with clay to silty claystone with nanofossils), siliciclastic-rich beds (silty claystone to clayey siltstone), and claystone, which forms the background sediment (see Shipboard Scientific Party, 1995; Site 948).

Cores 949B-2X-02 to 949B-25H-02 were sampled in Hole 949B, with a depth ranging from 255.48 mbsf to 460.53 mbsf. Despite the very low recovery (39.8%), the same two lithologic units as in Hole 948C were recognized. The first unit (Cores 2X-02 to 22X-03, 52 cm) is characterized by mottled nanofossil-bearing claystone, with thin layers of volcanic ash occurring sporadically. When going deeper, the claystone becomes barren of nanofossils, and volcanic ash is rarer. The second unit consists in a mottled variegated claystone, crossed by thin beds of sandy siltstone, siltstone, clayey siltstone, nanofossil claystone, and micrite (see Shipboard Scientific Party, 1995; Site 949).

## Sampling protocols.

All samples in this study were collected by P. Blum and A. Rabaute during ODP Leg 156, using two different sampling protocols.

A first set of 79 samples were first taken from Holes 948C and 949B. Each of them consists in a 20 cm long thin slice of material, centered at the core depth within the area of influence of the MST-NGR apparatus, during the gamma-ray measurements carried out during ODP Leg 156 (see Blum et al., this volume). Each sample weights about 250 g. This particular sampling mode was used to prevent potential sampling error due to lithologic variations. These samples were freeze-dried onboard the JOIDES Resolution, crushed in a Frysche mortar grinder Pulverisette, and splitted 7 times at the Laboratoire de Geochimie Isotopique in order to obtain 2 g-splits that can be considered as representative of the original sample. We will refer to this set of samples as 'GR (gamma-ray) samples' in this report.

The second set of samples comprises 113 samples chosen from the physical properties measurements residues. These samples were 10 cc in volume, and each of them was carefully taken out from a single lithologic variation, which reflects mostly a redox-associated change in color. This second data set was intended as a calibration for the possible geochemical variations that could happen along with the lithological changes, and therefore should represent geochemical "end-members". These samples were freeze-dried and crushed onboard the JOIDES Resolution. We will refer to this set of samples as 'Index Properties samples' later in this report.

## Methods of study.

X-ray fluorescence spectrometry was carried out at the Ecole Nationale Supérieure des Techniques Industrielles et des Mines d'Ales, France. The X-ray laboratory is equipped with a Philips PW-1400 spectrometer with a Philips PW-1775 70 port automatic sample changer, driven by a Philips X-40 software. Machine settings for all the samples analyzed were as follows:

	SiO <sub>2</sub>	Al <sub>2</sub> O <sub>3</sub>	Fe <sub>2</sub> O <sub>3</sub>	MgO	MnO	CaO	Na <sub>2</sub> O	K <sub>2</sub> O	TiO <sub>2</sub>	P <sub>2</sub> O <sub>5</sub>
Collimator	coarse	coarse	fine	coarse	fine	fine	coarse	fine	fine	coarse
Crystal	PeT	PeT	LiF100	PX1	LiF100	LiF100	PX1	LiF100	LiF 100	PeT
Detector	F	F	F+S	F	F+S	F	F	F	F	F
Peak time	60	60	10	80	40	40	80	40	40	40

**Table 1.** X-ray fluorescence measuring conditions for each element. PeT = Pentaerythritol; LiF100 = lithium fluoride; PX1 = artificial monodimensional crystal made with coating layers of Si and W; F = gas flow detector; S = scintillation detector; coarse collimator = 550 mm; fine collimateur = 150 mm; Peak time is the counting time at the element peak for each sample, in seconds.

generator=50 kV and 40 Ma; medium=vacuum; tube anode = Rh, with no beam filter; wavelength = Cu-K . Table 1 gives the measuring conditions for each chemical element. Fusion beads were prepared by mixing a sample split with lithium tetraborate + lanthanum oxide (LiBO<sub>4</sub> + La<sub>2</sub>O<sub>5</sub>) flux, with 1 for 8 relative respective proportions (0.750 g of sample per 6 g of flux). The method depends on the quantitative measurements of specific X-rays emitted during electronic transitions within atoms excited by irradiation with primary X-rays generated by a tube anode with a metal target. X-ray fluorescence analysis gives the concentrations of the 10 common major oxides SiO<sub>2</sub>, Al<sub>2</sub>O<sub>3</sub>, Fe<sub>2</sub>O<sub>3</sub>, MgO, MnO, CaO, Na<sub>2</sub>O, K<sub>2</sub>O, TiO<sub>2</sub>, and P<sub>2</sub>O<sub>5</sub>. Table 2 summarizes the results of a repeat analysis carried out on 10 samples from Hole 948C and on 21 samples from Hole 949B. It shows that the relative precision for silicon, calcium, iron, potassium, magnesium, titanium, aluminium, and manganese data is within 0.5 to 1%. P<sub>2</sub>O<sub>5</sub>, due to low concentrations, and Na<sub>2</sub>O, due to the masking effects of the lanthanum oxide on the Na peak, are considered precise within respectively 7% and 9%. Replicate analysis of rock standards show excellent to very good correlation for all elements except Na, which leads to an accuracy of 1 to 3 % for silica, calcium, iron, manganese, titanium, aluminium and potassium, and about 11% for MgO and P<sub>2</sub>O<sub>5</sub> (see Table 3). Na is not considered accurate because of its huge uncertainty and is thus indicative here.



The Inductively Coupled Plasma Mass Spectrometer analysis (ICP-MS) was carried out at the Université de Montpellier II, France. The basic principle is to use a solution of the sample which is passed, as an aerosol from a nebulizer, into an electrodeless argon plasma. The plasma is sustained by inductive coupling to radio frequency magnetic fields, and is connected with a multi-channel spectrometer. We determined only U and Th concentrations from the first set of 79 samples. The sample preparation involves the dissolution of 100 mg of powdered sample in a 15 ml

	SiO <sub>2</sub>	Al <sub>2</sub> O <sub>3</sub>	Fe <sub>2</sub> O <sub>3</sub>	MgO	MnO	CaO	Na <sub>2</sub> O	K <sub>2</sub> O	TiO <sub>2</sub>	P <sub>2</sub> O <sub>5</sub>
Absolute Average Deviation	0.149	0.100	0.029	0.036	0.005	0.016	0.071	0.010	0.007	0.011
Relative Average Deviation	0.27%	0.49%	0.35%	1.00%	0.95%	0.61%	8.96%	0.48%	0.86%	6.89%

**Table 2** . Evaluation of the precision of the X-ray fluorescence measurements given by absolute and relative average deviations from duplicate XRF measurements of each of 30 samples taken from the Index Properties samples of Holes 948C and 949B.

	SiO <sub>2</sub>	Al <sub>2</sub> O <sub>3</sub>	Fe <sub>2</sub> O <sub>3</sub>	MgO	MnO	CaO	Na <sub>2</sub> O	K <sub>2</sub> O	TiO <sub>2</sub>	P <sub>2</sub> O <sub>5</sub>
Root Mean Square ~ absolute average deviation (wt %)	0.705	0.266	0.216	0.305	0.007	0.111	1.482	0.062	0.025	0.016
Relative average deviation	1.32%	1.39%	2.80%	10.74%	2.12%	3.23%	N/A	3.07%	3.25%	11.43 %

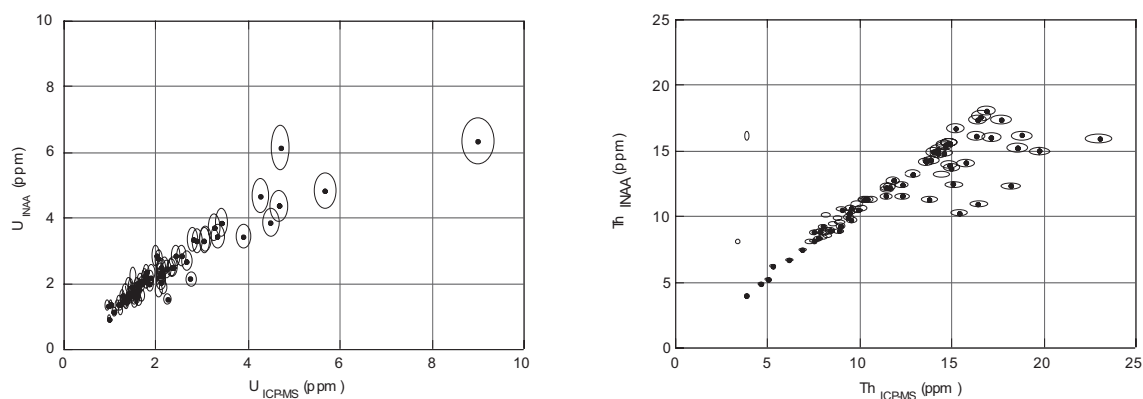
**Table 3** . Evaluation of the accuracy of the X-ray fluorescence measurements given by the square root of the sum of the squares of the absolute deviations from the linear regression between 11 Geostandard concentrations and their corresponding measured values. Input geostandard values during calibration were recommended values from Govindaraju (1989).

screw-top Teflon® bomb (Saville®) using HF 40% and HClO<sub>4</sub> 70%, at a temperature of 100°C. After evaporation of the HF-HClO<sub>4</sub> mixture, this dissolution/evaporation operation is repeated in order to get a complete dissolution of the sample. After evaporation to dryness, the sample is then taken into solution in HNO<sub>3</sub> 65%, transferred in a 20.6 g bottle, and diluted with water until the bottle is filled. For the analysis, 10 ml of this solution is taken, and, with the addition of 10 ppb of a In-Bi internal calibration standard, flushed into the plasma field. We used routinely doubly distilled reagents to minimize the blank. The use of screw-top bombs generally ensures more complete dissolution of resistant silicate and oxide minerals. The presence of a known amount of In-Bi standard in each sample preparation leads to the estimation of accuracy and precision. U is precise within 4% and accurate within 3%, while Th has a precision of 3% but an accuracy of 7%.

Instrumental Neutron Activation (INA) analyses were conducted at the Center for Chemical Characterization at Texas A&M University, using a TRIGA reactor. Fifty milligrams of sample were irradiated for 14 hours and counted for six hours at 9 to 12 and 40 to 43 days using germanium detectors. Spectral analysis was made using a Nuclear Data program and was followed by manual U- and Th-series overlap and interference corrections. 12 samples of the international AGV-1 standard were irradiated at multiple can levels during three runs and counted. Table 4 shows the results of error calculations from these repeated measurements of AGV-1 for all the trace elements. All elements, except Cr, Ni, Zr, Yb and U, show an overall good precision ranging from 1 to 8%, and an accuracy of 1 to 9%. Cr, Yb, and U are considered meaningful because of their good accuracy, despite their common 11% of precision. Ni and Zr are not considered accurate or precise and their values are thus only indicative here.

## Results.

Tables 5 to 10 present the major and trace element concentrations obtained on the two sets of samples from Holes 948C and 949B. In Hole 948C, because coring during Leg 156 was concentrated around the decollement zone, our sampling covers a 170 meters-zone, between 421.35 and 591.4 mbsf. With a total of 135 samples taken from this hole, the resolution is 1.25 meters. In Hole 949B, the sampling was sparse due to poor and uncontinuous core recovery. The 57 samples range from 255.49 to 460.53 mbsf. The GR sample set is considered to be representative of a global geochemical behaviour because of the technique of sampling used. It should be usefully compared directly to the logging data that were acquired while drilling in Hole 948A (close to Hole 948C), or open-hole in Hole 948C after coring operations. The Index Properties sample set can be used as discrete measurements of trace and major element concentrations around the decollement zone, and as geochemical nature of particular sediment layers, since each Index Properties sample was taken within a single-colored layer, which was interpreted as lithologic variation. All trace and major elements concentrations were obtained for the GR sample set. On the Index Properties sample set, only major elements concentrations were



**Figure 1.** Comparison between ICP-MS and INAA for U and Th. Plain circles are data from Hole 948C; open circles are data from Hole 949B. The ellipse around each data point is the precision of the measurement (see text and Table 4).

measured using X-ray fluorescence spectrometry.

Figure 1 compares the Th and U values obtained with the two different methods - ICP-MS and INAA. There is an overall good correlation, which is excellent for Hole 948C (plain circles), although data from Hole 949B (open circles) show a scattering, especially for the Thorium. Data from Hole 949B must be use with great care, in first hand because of these method-related problems, and also because low recovery in this hole may have induced problems of depth matching.

## References.

- Govindaraju, K. (1989) 1989 compilation of working values and sample description for 272 geostandards. *Geostand. Newsl.*, 13: 1-114.
- Shipboard Scientific Party, 1995. Site 948. In Shipley, T.H., Ogawa, Y., Blum, P., et al., *Proc. ODP, Init. Repts.*, 156: College Station, TX (Ocean Drilling Program), 87-192.
- Shipboard Scientific Party, 1995. Site 949. In Shipley, T.H., Ogawa, Y., Blum, P., et al., *Proc. ODP, Init. Repts.*, 156: College Station, TX (Ocean Drilling Program), 193-296.
- Wang, Y-C, Gieskes, J.M., and Musoke, L. (1990) Bulk chemical analysis of sediments - Hole 671B. In Moore, J.C., Mascle, A., et al., *Proc. ODP, Sci. Results*, 110: College Station, TX (Ocean Drilling Program), 179-188.

AGV-1	Sc	Cr	Co	Ni	Zn	Rb	Sr	Zr	Ba	La	Ce	Sm	Eu	Tb	Yb	Lu	Hf	Ta	Th	U
Run 1	12,00	11,50	15,80	13,60	92,50	63,40	760,0	143,0	1190	39,10	70,50	5,88	1,63	0,606	1,78	0,237	5,46	0,931	6,34	1,95
	12,00	8,52	15,80	26,70	94,30	65,40	715,0	187,0	1178	39,00	71,50	6,03	1,66	0,681	1,60	0,247	5,66	0,917	6,55	2,09
	12,14	11,85	16,18	<22,15	91,90	75,25	743,9	155,7	1215	38,87	71,80	5,93	1,67	0,532	1,64	0,260	5,31	0,880	6,44	2,01
	11,87	11,38	15,66	<18,94	88,38	75,00	773,0	134,0	1143	37,67	70,44	5,73	1,64	0,637	1,64	0,245	5,35	0,972	6,51	2,13
	12,09	10,73	15,90	<19,75	83,95	62,51	653,2	128,1	1214	39,16	71,42	6,07	1,65	0,588	1,62	0,209	5,42	0,894	6,34	1,99
to	12,00	10,80	15,93	22,62	88,03	67,15	672,4	157,5	1196	39,10	72,61	6,12	1,65	0,620	1,58	0,238	5,57	0,923	6,54	2,11
	11,88	8,89	15,42	<17,55	95,34	71,20	740,8	131,2	1214	38,58	69,55	5,88	1,58	0,624	1,46	0,224	5,39	0,871	6,16	2,00
	11,75	10,52	15,22	<20,11	86,43	68,69	698,5	122,9	1224	39,39	69,24	6,02	1,56	0,566	1,90	0,229	5,38	0,856	6,40	2,10
	11,93	9,47	15,82	<17,94	85,61	69,78	730,8	175,9	1201	38,94	70,71	6,02	1,56	0,717	1,65	0,218	5,49	0,889	6,54	1,68
	11,78	13,25	15,48	<16,58	85,92	71,63	768,1	178,8	1185	39,62	69,51	6,14	1,58	0,607	1,91	0,232	5,41	0,913	6,38	2,71
	12,32	11,17	16,02	<16,43	88,48	66,38	756,3	196,8	1181	39,09	73,00	6,06	1,68	0,612	1,46	0,216	5,71	0,931	6,58	2,19
Run 12	11,94	10,53	15,53	<24,41	85,84	71,10	701,9	204,9	1214	39,12	70,59	6,08	1,63	0,694	1,28	0,264	5,37	0,910	6,27	2,19
Mean	11,98	10,72	15,73	20,97	88,89	68,96	726,2	159,6	1196	38,97	70,91	6,00	1,62	0,624	1,63	0,23	5,46	0,91	6,42	2,10
Standard Deviation	0,15	1,25	0,26	5,47	3,58	3,96	36,80	27,26	21,58	0,46	1,15	0,11	0,04	0,05	0,17	0,02	0,12	0,03	0,12	0,23
<b>Precision (%)</b>	<b>1,26</b>	<b>11,67</b>	<b>1,68</b>	<b>26,10</b>	<b>4,02</b>	<b>5,75</b>	<b>5,07</b>	<b>17,08</b>	<b>1,80</b>	<b>1,19</b>	<b>1,62</b>	<b>1,91</b>	<b>2,53</b>	<b>8,13</b>	<b>10,62</b>	<b>6,97</b>	<b>2,20</b>	<b>3,32</b>	<b>1,93</b>	<b>10,83</b>
AGV-1*	12,20	10,10	15,30	16,00	88,00	67,30	667,0	227,0	1226	38,00	67,00	5,90	1,64	0,700	1,72	0,22	5,1	0,9	6,5	1,92
<b>Accuracy (%)</b>	<b>1,84</b>	<b>6,11</b>	<b>2,81</b>	<b>31,08</b>	<b>1,01</b>	<b>2,46</b>	<b>8,87</b>	<b>29,67</b>	<b>2,43</b>	<b>2,55</b>	<b>5,83</b>	<b>1,64</b>	<b>0,97</b>	<b>10,90</b>	<b>5,43</b>	<b>6,78</b>	<b>7,06</b>	<b>0,80</b>	<b>1,22</b>	<b>9,16</b>

**Table 4.** Precision and accuracy of INAA measurements from AGV-1 NIST geostandard (AGV-1\* values are recommended values from Govindaraju (1989))

Core/Sec	Interval (cm)	Depth (mbsf)	LOI (%)	SiO <sub>2</sub> (wt%)	Al <sub>2</sub> O <sub>3</sub> (wt%)	Fe <sub>2</sub> O <sub>3</sub> (wt%)	MgO (wt%)	MnO (wt%)	CaO (wt%)	Na <sub>2</sub> O (wt%)	K <sub>2</sub> O (wt%)	TiO <sub>2</sub> (wt%)	P <sub>2</sub> O <sub>5</sub> (wt%)	Total	Th <sub>(ICP-MS)</sub> (ppm)	U <sub>(ICP-MS)</sub> (ppm)	Th <sub>(INAA)</sub> (ppm)	U <sub>(INAA)</sub> (ppm)
2X 01	65-85	421.35	8.1	53.99	19.8	8.96	3.15	0.1	0.8	1.22	2.87	0.87	0.12	100.03	11.61	1.87	12.21	1.99
2X 03	65-85	424.35	8	55.19	19.3	8.23	3.35	0.11	1.08	1.59	2.2	0.84	0.1	100.05	7.79	1.58	8.29	1.59
2X 05	65-85	427.35	7.57	53.69	19.99	9.29	3.04	0.4	1.39	0.69	2.84	0.79	0.3	100	11.85	2.06	12.69	2.74
3X 01	65-85	431.25	7.72	55.97	19.82	8.07	3.37	0.09	0.99	1.33	2.38	0.84	0.08	100.67	9.02	1.71	10.52	1.97
3X 03	65-85	434.25	6.98	56.51	19.72	8.13	3.32	0.07	1.05	1.24	2.4	0.94	0.08	100.45	8.97	1.89	9.23	2.16
3X 05	65-85	437.25	12.8	46.09	17.59	7.99	2.69	3.28	5.32	0.88	2.19	0.71	0.46	100.02	no data	no data	10.74	2.41
4X 01	65-85	440.85	7.92	54.55	20.66	8.3	3.06	0.76	1.14	0.77	2.49	0.75	0.24	100.64	no data	no data	12.18	1.96
4X 03	65-85	443.85	12.16	49.04	18.2	6.98	2.88	0.25	7.02	0.89	2.11	0.64	0.12	100.29	9.56	1.28	9.70	1.63
4X 05	65-85	446.85	8.05	54.78	21.32	8	3.15	0.31	0.96	0.82	2.34	0.74	0.14	100.62	11.45	1.54	12.21	1.72
5X 01	65-85	450.55	7.8	55.6	20.62	8.24	3.48	0.15	0.99	1.4	2.2	0.85	0.1	101.44	9.53	1.49	10.60	1.62
5X 03	65-85	453.55	9.72	52.6	19.38	7.29	3.4	0.24	3.91	1.58	1.86	0.72	0.14	100.84	8.86	1.29	8.88	1.46
5X 05	65-85	456.55	13.4	47.85	17.77	6.49	2.92	0.27	7.81	1.07	1.62	0.63	0.14	100.01	7.55	1.35	8.11	1.43
6X 01	65-85	460.15	7.92	55.11	20.16	7.77	3.25	0.29	0.99	0.88	1.97	0.76	0.12	99.23	8.4	1.44	8.92	1.69
6X 03	65-85	463.15	8.2	54.27	20.89	7.94	3.06	0.46	1.07	1.1	2	0.85	0.15	100	9.48	1.6	10.30	1.93
7X 01	65-85	469.75	7.6	55.12	20.12	7.75	3.68	0.32	1.38	1.47	1.63	0.79	0.14	100.05	6.2	1.03	6.66	1.36
7X 03	65-85	472.75	7.62	53.75	19.09	8.49	3.7	0.46	1.22	1.47	2	0.85	0.17	98.82	7.5	1.18	8.76	1.34
7X 05	65-85	475.75	8.4	53.98	18.58	9.04	3.83	0.23	1.51	1.81	1.55	0.94	0.13	100.07	5.29	0.95	6.17	1.33
8X 01	65-85	479.45	7.58	55.98	19.09	9.49	3.46	0.3	1.38	1.45	1.74	0.93	0.13	101.53	7.83	1.38	8.81	1.47
8X 03	65-85	482.45	8	53.85	19.08	9.16	3.5	0.15	2	1.93	1.15	1.06	0.13	100.03	4.65	1.1	4.84	1.12
8X 05	65-85	485.45	7.27	55.89	18.94	9.65	3.51	0.22	1.7	1.33	1.53	0.94	0.13	101.09	6.91	1.22	7.43	1.34
9X 01	65-85	489.15	5.81	57.17	18.69	8.66	3.48	0.18	2.92	2.16	0.81	1.06	0.08	101.02	3.88	1	4.00	0.92
9X 03	65-85	492.15	6.22	58.64	18.35	8.41	3.49	0.26	2.34	1.88	1.13	1.13	0.11	101.98	5.07	1.68	5.14	2.06
9X 05	65-85	495.15	8.3	54.74	20.43	8.28	2.84	0.17	1.56	1.29	1.39	0.84	0.17	100.04	8	1.46	9.20	1.90
10X 01	65-85	498.85	6.9	56.23	19.74	8.38	2.89	0.22	1.77	1.57	1.33	0.82	0.16	100.09	8.41	1.61	8.87	1.81
10X 03	65-85	501.85	7.92	56.45	19.11	7.17	2.67	0.57	1.11	0.87	1.55	0.8	0.12	98.33	10.4	1.8	11.25	2.33
10X 05	5-25	504.25	7.9	56.86	20.26	7.1	2.63	0.45	1.19	1.28	1.47	0.7	0.17	100.04	10.28	1.76	11.30	2.19
11X 01	65-85	508.45	7.9	57.56	20.41	6.55	2.77	0.38	1.1	1.27	1.23	0.64	0.2	100.03	9.9	1.39	10.43	1.53
11X 03	65-85	511.45	9.3	53.96	22.11	8.23	1.91	0.1	0.65	1	1.8	0.79	0.15	100.07	17.71	2.24	17.37	2.42
11X 05	93-113	514.73	7.9	56.99	22.04	6.94	1.93	0.04	0.48	0.71	2.12	0.73	0.13	100.02	14.64	2.14	15.37	2.22
12X 01	65-85	518.15	8.33	62.12	19.44	6.89	1.82	0.04	0.46	0.38	2.37	0.78	0.06	102.7	14.6	2.01	14.80	2.84
12X 03	65-85	521.15	6.25	65.83	17.89	5.98	1.85	0.1	0.65	0.45	2.2	0.82	0.11	102.13	14.87	2.67	13.89	2.67
12X 05	65-85	524.15	6.5	58.76	21.15	7.83	2.22	0.05	0.54	0.41	2.74	0.76	0.09	101.05	12.88	4.72	13.20	6.13
13X 01	65-85	527.55	7.67	56.35	20.86	6.2	1.88	0.03	0.53	0.42	2.59	0.78	0.08	97.39	13.89	2.89	14.28	3.28
13X 03	95-115	530.55	18.4	38.19	14.43	9.35	2.02	2.95	11.67	0.1	1.46	0.54	0.3	99.4	13.82	5.69	11.26	4.83
13X 05	65-85	533.85	11.77	49.15	20.1	10.96	2.32	1.68	1.18	0.17	2.5	0.67	0.11	100.6	13.89	3.12	no data	no data
14X 01	65-85	536.85	8.8	54.8	23.07	6.71	1.8	0.1	0.51	0.69	2.66	0.75	0.11	100	14.9	2.56	15.60	2.83

**Table 5.** Chemical compositions of bulk sediments: Major elements concentrations, U and Th contents. Hole 948C. GR sample set..

Core/Sec	Interval (cm)	Depth (mbsf)	LOI (%)	SiO <sub>2</sub> (wt%)	Al <sub>2</sub> O <sub>3</sub> (wt%)	Fe <sub>2</sub> O <sub>3</sub> (wt%)	MgO (wt%)	MnO (wt%)	CaO (wt%)	Na <sub>2</sub> O (wt%)	K <sub>2</sub> O (wt%)	TiO <sub>2</sub> (wt%)	P <sub>2</sub> O <sub>5</sub> (wt%)	Total	Th (ICP-MS) (ppm)	U (ICP-MS) (ppm)	Th (NAA) (ppm)	U (NAA) (ppm)
14X 03	65-85	539,85	8,9	52,98	23,57	7,97	1,79	0,02	0,38	0,61	2,88	0,79	0,11	100,04	16,62	2,12	17,62	2,47
14X 05	65-85	542,85	10	54,1	21,73	6,47	1,81	0,08	1,93	0,58	2,46	0,73	0,11	100	14,28	3,43	15,10	3,84
15X 01	65-85	546,15	10	52,88	23,28	7,04	2	0,49	1,21	0,47	1,66	0,78	0,19	100,06	14,63	1,79	15,48	2,06
15X 03	65-85	549,15	12,9	49,45	21,77	5,77	1,69	0,03	5,31	0,37	1,93	0,68	0,11	100	15,78	3,35	14,04	3,42
15X 05	65-85	552,15	7,56	56,38	19,27	7,18	1,85	0,04	1,51	0,38	2,54	0,77	0,08	97,56	14,29	1,47	14,76	1,71
16X 01	65-85	555,55	8,9	55,74	21,61	6,79	1,86	0,03	1,4	0,47	2,3	0,76	0,12	100	15,23	2,1	16,70	2,22
16X 03	65-85	558,55	10,6	51,85	22,78	7,15	1,93	0,03	2,38	0,46	1,88	0,78	0,16	100,09	16,34	1,59	16,13	1,81
16X 05	65-85	561,55	22,4	34,22	11,62	4,35	1,57	0,35	22,95	0	1,17	0,46	0,26	99,35	9,39	3,27	9,89	3,70
17X 01	65-85	564,85	6,95	59,1	20,01	7,43	1,89	0,03	0,46	0,64	3,24	0,91	0,07	100,74	16,91	2,33	18,01	2,46
17X 03	65-85	567,85	8,24	56,51	19,84	8,2	1,97	0,04	2,54	0,41	2,98	0,84	0,09	101,67	14,85	2,11	15,60	2,35
17X 05	65-85	570,85	11,85	52,2	16,51	5,71	1,62	0,07	8,84	0,11	2,49	0,76	0,11	100,28	15	3,06	13,67	3,31
18X 01	65-85	574,05	15,03	46,13	18,57	5,48	1,49	0,11	10,78	0,81	1,94	0,74	0,23	101,3	18,63	4,69	15,20	4,35
18X 03	65-85	577,05	16	43,05	20,6	5,46	1,56	0,08	10,66	0,38	1,49	0,56	0,15	100	12,31	2,12	12,38	2,03
18X 05	65-85	580,05	8	55,96	22,7	6,78	1,72	0,03	0,72	0,47	2,71	0,83	0,11	100,01	16,45	2,45	17,36	2,86
19X 01	65-85	583,55	13,8	46,06	20,74	4,87	1,65	0,14	8,03	0,39	1,96	0,64	0,14	98,41	14,01	1,63	14,88	1,80
19X 03	65-85	586,55	7,75	58,78	21,03	7,62	1,89	0,08	0,82	0,26	3,02	0,92	0,08	102,26	17,19	2,81	16,01	3,34
19X 05	65-85	589,55	10,6	53,04	20,47	7,49	1,8	0,17	3,1	0,45	2,03	0,7	0,15	100	13,62	2,36	14,17	2,50

**Table 5 (continued).** Chemical compositions of bulk sediments: Major elements concentrations, U and Th contents. Hole 948C. GR sample set.

Core/Sec	Interval (cm)	Depth (mbsf)	Sc	Cr	Co	Ni	Zn	Rb	Sr	Zr	Ba	La	Ce	Sm	Eu	Tb	Yb	Lu	Hf	Ta	
2X 01	65-85	421,35	21,28	85,36	20,64	54,60	131,9	113,4	92,2	90,2	346,5	38,08	80,50	7,71	1,64	1,080	3,50	0,478	3,57	0,877	
2X 03	65-85	424,35	24,01	54,18	23,96	<28,81	108,4	83,8	<111,3	<69,20	178,4	25,14	57,28	5,82	1,48	0,933	2,69	0,403	3,35	0,618	
2X 05	65-85	427,35	20,30	82,64	20,09	47,34	120,7	128,3	<116,9	68,1	266,4	36,88	78,99	7,08	1,42	0,823	2,75	0,380	3,72	0,956	
3X 01	65-85	431,25	23,20	66,27	22,17	<28,41	125,7	99,8	<108,3	<67,49	269,2	30,49	62,12	6,41	1,58	0,874	2,99	0,419	3,48	0,770	
3X 03	65-85	434,25	25,36	62,91	23,92	43,05	125,2	88,6	88,3	91,9	194,8	28,05	69,19	5,87	1,40	0,868	2,58	0,345	3,50	0,692	
3X 05	65-85	437,25	18,82	73,62	21,38	47,34	101,7	102,9	189,0	<62,69	264,4	35,96	80,28	6,68	1,42	0,894	3,13	0,445	2,96	0,738	
4X 01	65-85	440,85	19,79	84,25	22,47	48,76	120,5	119,4	95,0	99,7	306,7	38,39	84,54	7,50	1,62	1,090	3,31	0,445	3,50	0,802	
4X 03	65-85	443,85	17,80	68,50	21,50	46,20	115,0	82,4	211,0	78,8	239,0	31,20	66,90	6,30	1,38	0,908	3,15	0,437	3,34	0,713	
4X 05	65-85	446,85	20,53	90,97	25,02	50,13	120,1	103,1	58,7	109,8	257,0	36,84	82,26	7,29	1,56	0,965	3,29	0,429	3,77	0,830	
5X 01	65-85	450,55	23,40	77,50	25,90	48,30	121,0	82,0	<77,0	94,7	185,0	33,20	73,10	6,83	1,52	0,867	2,95	0,413	3,91	0,886	
5X 03	65-85	453,55	20,67	64,14	23,00	43,67	118,3	70,2	212,4	88,0	264,0	28,29	60,71	6,12	1,41	0,914	3,04	0,433	3,70	0,639	
5X 05	65-85	456,55	19,20	56,90	20,40	31,70	109,0	62,0	336,0	79,7	238,0	26,50	58,30	5,52	1,26	0,788	2,87	0,400	3,34	0,632	
6X 01	65-85	460,15	21,88	70,97	21,89	37,07	117,6	79,0	<68,21	59,3	205,3	30,24	64,66	6,39	1,41	0,908	3,00	0,485	3,49	0,676	
6X 03	65-85	463,15	20,90	77,90	22,20	46,20	110,0	79,6	<75,7	87,6	242,0	35,70	76,00	6,98	1,53	0,993	3,51	0,476	4,08	0,919	
7X 01	65-85	469,75	23,43	52,82	33,28	47,77	136,9	60,0	<115,6	<76,92	209,4	21,29	49,45	5,31	1,26	0,751	2,59	0,378	3,43	0,488	
7X 03	65-85	472,75	23,10	70,30	32,50	62,70	120,1	65,0	<76,9	<53,3	244,0	30,70	67,00	6,67	1,48	1,000	3,24	0,434	3,45	0,717	
7X 05	65-85	475,75	30,82	52,03	35,35	51,31	106,5	47,2	<133,9	88,4	206,1	20,43	44,38	5,11	1,33	0,803	2,57	0,414	2,48	0,422	
8X 01	65-85	479,45	29,60	69,50	30,40	60,10	127,0	62,2	<84,9	92,4	262,0	28,50	65,50	6,22	1,48	0,893	3,01	0,439	3,20	0,726	
8X 03	65-85	482,45	34,06	53,16	36,81	51,74	119,6	47,8	<137,5	<89,79	136,0	18,40	40,93	5,04	1,35	0,861	2,82	0,429	2,75	0,402	
8X 05	65-85	485,45	29,20	62,50	30,80	47,60	132,0	69,7	<84,9	<58,6	265,0	25,20	55,60	5,79	1,41	0,890	2,86	0,405	3,04	0,576	
9X 01	65-85	489,15	35,10	52,33	44,27	53,65	139,5	26,6	<125,4	<78,90	93,0	11,06	26,02	2,99	803,70	0,394	1,65	0,239	2,93	0,333	
9X 03	65-85	492,15	35,65	41,56	43,68	<34,26	94,0	43,1	<132,6	<81,34	250,6	19,16	42,31	5,37	1,43	0,680	2,83	0,417	2,78	0,404	
9X 05	65-85	495,15	24,74	69,36	34,92	62,31	139,2	76,3	129,3	69,9	482,6	31,40	69,56	6,69	1,57	1,030	3,27	0,481	3,48	0,721	
10X 01	65-85	498,85	25,40	72,30	36,40	85,00	144,0	70,0	96,5	<57,3	182,0	28,20	64,40	6,29	1,47	0,783	3,32	0,465	3,14	0,637	
10X 03	65-85	501,85	20,45	91,18	30,50	94,56	123,5	78,7	76,7	125,1	328,6	36,96	85,50	7,12	1,53	0,979	3,24	0,450	4,15	1,020	
10X 05	5-25	504,25	20,20	89,50	29,80	86,50	151,0	88,3	201,0	66,2	2349,0	35,30	86,00	6,84	1,54	0,990	3,26	0,472	3,21	0,786	
11X 01	65-85	508,45	19,62	77,83	36,49	72,27	127,8	75,4	110,1	80,3	198,6	31,54	75,69	6,69	1,48	0,914	2,77	0,394	3,15	0,714	
11X 03	65-85	511,45	19,04	97,25	19,15	81,18	131,4	110,8	164,1	124,6	479,3	41,56	104,60	7,34	1,51	0,924	2,77	0,402	3,68	1,080	
11X 05	93-113	514,73	18,11	92,24	72,53	108,30	185,7	119,0	<102,1	91,4	319,3	40,91	90,99	7,01	1,39	0,866	2,76	0,392	4,05	1,160	
12X 01	65-85	518,15	18,00	85,80	83,30	134,00	112,0	135,0	<76,0	68,0	289,0	42,90	85,90	7,36	1,34	0,792	2,76	0,385	3,81	1,040	
12X 03	65-85	521,15	17,29	97,56	17,97	53,54	128,4	126,4	85,2	119,4	336,1	37,51	78,39	6,36	1,17	0,627	2,58	0,365	4,17	1,040	
12X 05	65-85	524,15	15,90	94,20	26,90	77,20	204,0	106,0	<65,7	168,0	435,0	35,60	75,80	6,73	1,34	0,984	3,00	0,468	6,42	1,050	
13X 01	65-85	527,55	18,97	101,90	31,77	82,62	195,4	128,8	142,6	138,1	397,7	38,69	80,85	6,49	1,25	0,732	2,44	0,372	4,16	1,030	
13X 03	95-115	530,55	11,94	77,07	12,89	32,66	88,2	78,5	368,0	125,6	348,4	32,06	68,13	5,49	1,17	0,702	2,50	0,400	4,76	0,734	
13X 05	65-85	533,85																			

no data

**Table 6.** Chemical compositions of bulk sediments: Trace elements concentrations. Hole 948C. GR sample set.

Core/Sec	Interval (cm)	Depth (mbsf)	Sc	Cr	Co	Ni	Zn	Rb	Sr	Zr	Ba	La	Ce	Sm	Eu	Tb	Yb	Lu	Hf	Ta
14X 01	65-85	536,85	20,20	105,00	21,10	47,10	183,0	154,0	92,6	94,6	369,0	39,50	85,30	6,58	1,32	0,835	2,58	0,369	4,54	1,090
14X 03	65-85	539,85	20,01	99,45	15,92	51,57	108,8	168,6	<101,1	137,9	356,0	47,06	103,81	8,16	1,63	0,921	2,73	0,398	3,97	1,210
14X 05	65-85	542,85	18,54	101,20	21,23	51,51	151,2	144,1	163,6	139,7	379,6	40,17	82,57	6,78	1,37	0,731	2,68	0,405	4,26	1,020
15X 01	65-85	546,15	16,14	115,10	14,25	71,50	114,8	82,0	108,6	150,0	553,5	46,05	95,18	7,08	1,46	0,922	2,95	0,404	4,90	1,150
15X 03	65-85	549,15	15,41	95,43	20,78	76,39	134,2	15,4	307,4	110,5	571,4	37,69	77,09	5,85	1,17	0,737	2,29	0,345	4,30	0,994
15X 05	65-85	552,15	18,88	99,84	10,83	<27,69	121,9	145,9	<94,41	<75,40	362,3	41,16	88,40	7,09	1,44	0,959	2,62	0,382	3,45	0,990
16X 01	65-85	555,55	16,30	97,40	17,00	53,30	101,0	117,0	129,0	157,0	417,0	42,40	89,20	6,95	1,36	0,817	2,65	0,395	5,02	1,190
16X 03	65-85	558,55	16,32	108,30	14,78	52,37	142,9	91,3	213,6	137,9	546,0	43,68	93,10	6,61	1,38	0,784	2,51	313,8	4,54	1,160
16X 05	65-85	561,55	10,50	69,20	10,30	34,30	109,0	56,9	777,0	115,0	543,0	34,80	72,20	5,32	1,14	0,669	2,28	0,347	3,60	0,752
17X 01	65-85	564,85	17,96	88,82	12,51	<21,52	149,2	159,7	<106,9	155,9	377,1	48,27	100,58	8,07	1,43	0,861	2,88	0,418	4,62	1,470
17X 03	65-85	567,85	17,20	90,40	13,60	48,30	123,0	134,0	249,0	94,9	411,0	45,20	91,30	7,81	1,48	0,897	2,90	0,386	4,08	1,170
17X 05	65-85	570,85	17,34	108,15	17,74	61,505	133	113,04	136,1	144,85	466,55	43,11	88,875	6,93	1,415	0,8265	2,815	0,4045	4,58	1,085
18X 01	65-85	574,05	15,60	97,70	14,40	47,40	110,0	98,2	503,0	131,0	418,0	49,90	109,00	9,29	1,92	1,300	3,33	0,489	4,29	1,140
18X 03	65-85	577,05	15,82	91,81	13,69	71,53	100,8	84,6	618,4	165,1	672,2	37,49	81,65	6,48	1,33	0,777	2,40	0,373	3,53	0,923
18X 05	65-85	580,05	18,04	89,60	14,60	42,08	133,5	156,5	<109,1	<71,11	366,6	47,07	97,42	7,73	1,41	0,890	2,65	0,384	4,42	1,370
19X 01	65-85	583,55	15,87	100,70	13,33	54,15	105,6	116,4	396,8	73,9	697,9	41,78	91,10	7,04	1,42	0,763	2,30	0,336	3,73	1,080
19X 03	65-85	586,55	16,37	79,68	18,34	34,79	130,9	131,1	120,0	133,2	334,8	45,43	86,06	7,74	1,35	0,840	2,96	0,43214	4,53985	1,254
19X 05	65-85	589,55	15,34	99,66	14,32	65,00	131,8	102,6	166,8	147,7	409,9	36,69	74,91	6,05	1,17	0,712	2,59	0,376	5,28	1,100

**Table 6 (continued).** Chemical compositions of bulk sediments: Trace elements concentrations. Hole 948C. GR sample set.

Core/Sec	Interval (cm)	Depth (mbsf)	LOI (%)	SiO <sub>2</sub> (wt%)	Al <sub>2</sub> O <sub>3</sub> (wt%)	Fe <sub>2</sub> O <sub>3</sub> (wt%)	MgO (wt%)	MnO (wt%)	CaO (wt%)	Na <sub>2</sub> O (wt%)	K <sub>2</sub> O (wt%)	TiO <sub>2</sub> (wt%)	P <sub>2</sub> O <sub>5</sub> (wt%)	Total	Th (ICP-MS) (ppm)	U (ICP-MS) (ppm)	Th (INAA) (ppm)	U (INAA) (ppm)
2X 04	65- 85	259,05	7,95	53,91	21,11	9,66	2,97	0,17	0,81	0,84	2,57	0,84	0,24	101,05	11,48	1,58	12,08	1,71
3X 01	65- 85	264,25	7,93	55,91	19,99	7,26	3,67	0,14	1,02	1,62	1,82	0,86	0,09	100,31	3,41	1,51	8,03	2,24
3X 03	65- 85	267,25	8,78	55,95	20,36	7,56	3,57	0,2	0,92	1,49	1,98	0,78	0,14	101,72	8,83	1,42	9,85	1,51
3X 06	75- 95	271,85	7,53	56,6	19,25	8,19	3,72	0,18	1	1,53	1,99	0,85	0,12	100,95	8,01	1,35	8,38	1,79
4X 01	55- 75	275,75	7,78	56,89	19,6	7,16	3,87	0,18	0,9	1,49	1,99	0,74	0,1	100,72	8,23	1,36	8,51	1,37
5X 01	65- 85	283,55	7,82	56,75	18,61	9,08	3,58	0,23	1,03	1,62	1,91	0,87	0,13	101,66	7,31	1,19	8,09	1,46
5X 03	65- 85	286,55	7,94	56,76	19,48	7,59	3,29	0,58	0,93	1,35	1,91	0,89	0,15	100,86	10,69	1,59	11,32	1,59
5X 05	65- 85	289,55	8,58	56,24	17,55	8,03	3,52	1,28	1,25	2,41	1,79	0,78	0,16	101,59	7,82	1,24	8,86	1,19
7X 01	65- 85	302,85	17,83	40,95	15,37	6,43	2,66	0,29	15,49	0,54	1,65	0,56	0,16	101,92	16,53	2,26	10,89	1,51
7X 03	65- 85	305,85	12,21	47,42	18,19	7,56	2,72	0,27	8,18	0,81	2,48	0,72	0,13	100,68	15,14	2,07	12,39	1,79
7X 05	65- 85	308,85	15,99	42,57	16,43	6,44	2,54	0,43	13,26	0,57	1,93	0,61	0,13	100,9	12,34	1,7	11,49	1,63
13X 02	65- 85	352,45	8,78	53,45	20,45	8,24	3,24	0,21	0,88	1,18	2,35	0,81	0,13	99,71	11,46	1,64	11,47	1,48
14X 02	90-110	357,4	8,74	53,92	20,24	8,77	3,14	0,21	0,8	1,26	2,18	0,89	0,12	100,27	15,41	2,17	10,26	1,86
14X 03	70- 90	358,7	7,89	54,97	20,15	7,9	3,42	0,3	0,96	1,6	1,96	0,68	0,09	99,93	8,19	3,08	8,76	3,36
14X 05	60- 80	361,6	7,94	54,91	20,68	8,26	3,2	0,14	0,82	1,12	2,21	0,82	0,14	100,25	10,09	1,58	10,58	1,80
15X 02	55- 75	362,05	7,97	55,07	20,64	7,95	3,32	0,26	0,87	1,39	2,24	0,76	0,15	100,62	10,27	1,55	11,21	1,50
15X 04	65- 85	365,15	9	53,16	19,46	8,69	3,24	0,42	1,39	1,4	2,28	0,79	0,2	100,04	8,16	1,53	10,09	1,42
15X 06	80-100	368,3	8,01	54,24	19,16	8,63	3,98	0,5	1,4	1,3	2,3	0,81	0,17	100,5	8,56	1,42	9,40	1,93
19X 01	100-120	399,9	8,1	58,71	18,38	7,41	2,74	0,23	1,1	1,24	1,55	0,73	0,16	100,35	9,85	1,59	10,88	1,68
19X 04	15- 35	403,55	7,64	60,27	17,87	7,04	3,23	0,28	1,2	1,54	1,44	0,7	0,13	101,34	8,94	2,2	9,50	2,53
22X 01	65- 85	428,15	8,17	55,29	21,6	9,42	2,46	0,09	0,74	0,78	2,17	0,83	0,12	101,66	3,89	1,92	16,08	2,32
22X 03	55- 75	431,05	8,14	56,19	20,8	8,63	2,55	0,06	0,58	0,73	2,22	0,83	0,11	100,84	23	4,5	15,87	3,87
22X 05	65- 85	434,15	-2,77	69,2	19,22	7,93	2,76	0,05	0,59	0,86	2,39	0,89	0,1	101,24	14,49	4,28	13,19	4,68
25H 01	56- 76	459,01	8,21	54,16	23,13	9,12	2,23	0,04	0,54	0,55	3,12	0,74	0,08	101,92	19,79	2,76	14,89	2,11
25H 01*	82-102	459,32	8,4	55,41	23,09	7,56	2,18	0,1	0,5	0,39	2,98	0,75	0,08	101,44	18,85	3,91	16,12	3,43
25H 02	53- 73	460,53	27	33,01	14,21	5,24	1,4	0,57	16	0,03	0,73	0,52	0,16	98,87	18,24	9,01	12,25	6,35

**Table 7.** Chemical compositions of bulk sediments: Major elements concentrations, U and Th contents. Hole 949B. GR sample set.



Core/Sec	Interval (cm)	Depth (mbsf)	Sc	Cr	Co	Ni	Zn	Rb	Sr	Zr	Ba	La	Ce	Sm	Eu	Tb	Yb	Lu	Hf	Ta
2X 04	65-85	259,05	20,06	87,97	17,41	59,46	129,1	95,9	<75,27	105,5	221,3	40,08	85,61	7,18	1,48	0,943	3,05	0,458	3,33	0,859
3X 01	65-85	264,25	24,32	56,18	29,32	29,07	116,0	60,0	<94,29	83,1	145,5	25,54	57,99	6,02	1,32	0,758	2,93	0,428	4,42	0,670
3X 03	65-85	267,25	22,32	74,50	31,16	47,52	143,2	85,0	<83,87	98,2	263,3	32,41	73,92	6,98	1,60	1,060	3,58	0,486	3,78	0,748
3X 06	75-95	271,85	23,13	66,59	28,76	44,51	119,2	75,4	<83,67	102,1	228,8	28,38	63,19	6,62	1,60	1,060	3,30	0,495	3,36	0,616
4X 01	55-75	275,75	20,91	66,90	36,52	66,34	133,1	74,2	55,8	123,7	267,3	27,87	67,41	6,52	1,45	0,934	2,90	0,427	4,14	0,704
5X 01	65-85	283,55	26,57	65,76	24,93	46,30	106,5	64,1	120,3	<69,65	169,5	28,57	63,80	6,14	1,43	0,834	2,81	0,394	2,86	0,662
5X 03	65-85	286,55	21,66	84,41	31,37	76,75	131,4	82,4	142,9	120,3	403,1	37,75	87,57	7,23	1,60	0,901	3,57	0,499	4,50	1,170
5X 05	65-85	289,55	22,09	104,80	35,34	68,88	124,0	74,0	115,5	88,0	515,9	28,99	66,87	6,43	1,42	0,984	3,20	0,436	3,27	0,699
7X 01	65-85	302,85	13,35	63,05	15,52	34,34	86,2	99,1	618,9	<53,50	263,7	35,57	74,06	6,43	1,28	0,786	2,33	0,321	2,77	0,812
7X 03	65-85	305,85	15,77	69,54	20,00	51,38	98,0	101,3	340,3	100,6	256,3	40,09	83,82	7,42	1,51	0,987	2,94	0,426	3,39	0,912
7X 05	65-85	308,85	14,69	63,99	18,39	47,74	86,2	92,2	532,8	129,1	261,5	36,38	75,66	6,75	1,35	0,779	2,71	0,360	3,23	0,882
13X 02	65-85	352,45	20,68	86,05	20,48	50,80	122,6	87,6	103,0	65,7	231,6	36,11	77,53	6,91	1,44	0,876	3,07	0,437	3,69	0,844
14X 02	90-110	357,4	21,76	70,80	21,66	32,90	119,2	80,0	<81,69	91,8	251,9	35,98	73,87	7,19	1,48	1,060	3,53	0,499	3,79	0,853
14X 03	70-90	358,7	21,45	57,40	24,99	45,84	114,2	60,7	182,9	149,7	1882,1	28,10	62,29	6,43	1,34	0,878	3,00	0,438	5,42	0,791
14X 05	60-80	361,6	21,37	79,17	20,48	51,90	137,3	86,5	<102,5	96,8	194,3	35,86	78,46	7,24	1,61	1,020	4,21	0,535	3,53	0,746
15X 02	55-75	362,05	20,36	83,79	21,72	57,32	140,1	101,5	76,3	82,4	200,2	35,32	79,81	7,17	1,51	1,030	3,51	0,504	4,11	0,887
15X 04	65-85	365,15	22,91	79,53	16,42	42,44	113,9	97,9	<84,23	74,6	309,9	31,80	76,81	6,59	1,49	0,960	3,25	0,474	3,45	0,728
15X 06	80-100	368,3	22,70	74,50	21,07	44,62	126,5	89,8	92,9	<59,25	237,1	30,36	69,62	6,22	1,42	0,967	2,80	0,420	3,49	0,766
19X 01	100-120	399,9	18,77	91,28	25,58	94,59	136,8	78,2	<72,25	73,8	420,0	34,60	81,37	6,49	1,37	0,865	3,10	0,422	3,20	0,808
19X 04	15-35	403,55	19,67	86,78	60,19	104,90	127,9	72,7	68,6	<51,96	273,5	29,18	73,06	6,27	1,48	0,934	2,95	0,436	3,11	0,664
22X 01	65-85	428,15	19,69	104,40	20,83	81,11	162,5	114,7	88,2	74,3	275,8	44,89	105,30	7,33	1,48	0,903	2,87	0,393	4,02	1,150
22X 03	55-75	431,05	18,66	103,90	18,80	64,01	127,3	112,0	107,1	119,5	343,5	41,85	96,25	7,10	1,37	0,747	2,59	0,412	3,97	1,110
22X 05	65-85	434,15	16,53	93,92	102,40	131,30	116,1	125,7	<100,0	105,0	295,8	34,92	78,37	5,97	1,23	0,633	2,76	0,394	4,50	1,130
25H 01	56-76	459,01	19,69	106,60	17,21	41,69	143,0	141,9	<100,6	98,7	351,1	42,49	90,93	7,30	1,42	0,944	2,47	0,367	3,47	1,020
25H 01*	82-102	459,32	19,97	107,70	22,42	58,37	175,5	149,6	<99,16	127,0	343,3	40,08	85,46	6,35	1,27	0,764	2,40	0,356	4,05	1,080
25H 02	53-73	460,53	12,14	93,89	13,51	56,06	138,6	61,5	716,2	108,8	566,9	35,37	75,06	5,96	1,19	0,715	2,19	0,363	4,07	0,925

**Table 8.** Chemical compositions of bulk sediments: Trace elements concentrations. Hole 949B. GR sample set.

Core/Sec	Interval (cm)	Depth (mbsf)	LOI (%)	SiO <sub>2</sub> (wt%)	Al <sub>2</sub> O <sub>3</sub> (wt%)	Fe <sub>2</sub> O <sub>3</sub> (wt%)	MgO (wt%)	MnO (wt%)	CaO (wt%)	Na <sub>2</sub> O (wt%)	K <sub>2</sub> O (wt%)	TiO <sub>2</sub> (wt%)	P <sub>2</sub> O <sub>5</sub> (wt%)	Total
2X 01	140-142	422,2	7,01	54,83	19,99	8,15	3,02	0,11	1,03	0,74	2,67	0,77	0,12	98,45
2X 01	146-148	422,3	6,99	55,61	20,26	8,39	3,11	0,12	1,04	0,75	2,63	0,79	0,14	99,84
2X 03	72- 74	424,5	8,55	54,94	18,84	8,35	3,4	0,12	1,15	1,07	2,31	0,86	0,08	99,68
2X 04	116-118	426,5	8,54	54,89	20,04	8	3,19	0,11	1,03	0,84	2,63	0,77	0,12	100,17
2X 05	95- 97	427,7	7,2	54,77	19,48	9,87	2,92	0,48	0,89	0,45	2,97	0,76	0,2	100
3X 01	55- 57	431	7,08	55,77	19,34	8,3	3,4	0,11	1,17	1,14	2,35	0,89	0,07	99,63
3X 02	139-141	433,4	6,6	57,21	20,82	6,91	3,7	0,09	1,27	1,86	1,84	0,64	0,09	101,03
3X 03	1- 3	433,5	7,03	56,49	19,81	7,49	3,49	0,09	1,31	1,29	2,05	0,73	0,09	99,88
3X 05	95- 97	437,5	11,69	48,96	17,73	8,49	2,78	0,34	5,03	0,58	2,38	0,87	0,11	98,98
4X 03	55-57	443,6	11,4	49,41	18,45	7,86	3,19	0,32	7,05	0,51	2,34	0,68	0,12	101,33
4X 03	142-144	444,5	10,69	50,53	19,21	6,89	2,86	0,31	5,24	0,5	2,27	0,66	0,15	99,31
4X 05	42- 44	446,5	10,54	49,99	18,82	8,24	2,83	0,31	5,11	0,54	2,33	0,69	0,13	99,55
5X 01	63- 65	450,4	7,62	54,82	19,91	8,58	3,44	0,14	1,07	0,88	2,09	0,88	0,13	99,55
5X 01	146-148	451,3	7,56	55,7	19,94	8,45	3,96	0,18	1,65	1,08	1,94	0,81	0,18	101,45
5X 02	43- 45	451,7	9,51	53,22	20,29	8,41	3,01	0,14	0,88	0,56	2,26	0,78	0,13	99,18
5X 03	127-129	454,1	13	47,93	18	7,84	3,03	0,3	7,01	0,84	1,97	0,71	0,15	100,79
5X 04	110-112	455,4	15,58	43,81	16,62	6,4	2,84	0,7	12,19	0,67	1,68	0,6	0,15	101,25
5X 05	63- 65	456,4	12,39	48,54	17,37	7,29	3,16	0,29	8,39	0,83	1,81	0,7	0,17	100,93
5X 05	99-101	456,8	16,79	41,27	15,4	5,76	2,62	0,57	14,17	0,25	1,54	0,5	0,12	98,99
5X CC	25- 27	459,5	17,32	41,57	14,41	5,22	2,8	0,58	15,38	0,56	1,17	0,52	0,16	99,68
6X 01	22-24	459,6	14,52	45,76	17,08	6,19	2,81	0,46	10,25	0,56	1,61	0,6	0,12	99,96
6X 02	6- 8	461	11,81	47,02	18,44	8,02	3,03	5,37	1,96	0,36	2,13	0,7	0,19	99,05
6X CC	15- 17	465,8	7,65	54,96	19,22	8,48	3,6	0,42	1,37	1	2,07	0,84	0,16	99,76
7X 02	123-125	471,7	7,47	55,29	19,01	8,59	3,5	0,24	1,18	0,99	1,97	0,81	0,15	99,21
8X 01	52- 54	479,2	7,26	55,44	18,43	9,44	3,59	0,25	1,44	1,08	1,71	0,86	0,15	99,65
8X 01	122-124	479,9	7,93	54,49	18,54	8,95	3,49	1,03	1,5	0,89	1,75	0,91	0,13	99,6
8X 03	9- 11	481,8	7,42	55,95	19,02	8,94	3,51	0,36	1,44	1	1,56	0,96	0,14	100,3
8X 04	141-143	484,6	7,71	54,93	18,31	9,59	3,51	0,33	1,33	0,86	1,78	0,92	0,15	99,43
8X 05	66- 68	485,4	6,85	55,47	17,95	9,97	3,6	0,17	1,99	1,41	1,4	0,98	0,13	99,9
8X 05	115-117	485,9	7,24	55,18	18,54	9,7	3,42	0,24	1,7	1,14	1,61	0,93	0,15	99,85
9X 01	52- 54	488,9	7,29	55,85	18,75	8,68	3,52	0,2	1,64	1,22	1,48	0,88	0,17	99,7
9X 01	94- 96	489,3	7,07	55,42	18,72	9,37	3,31	0,22	1,79	1,17	1,54	0,94	0,17	99,72
9X 04	59- 61	493,5	6,91	57,23	18	8,77	3,47	0,26	1,56	1,22	1,48	1	0,12	100,03
9X 05	7- 9	494,5	6,7	58,51	17,69	10,23	3,68	0,18	2,24	1,46	1,33	0,93	0,12	103,07
9X 05	81- 83	495,2	7,06	56,76	19,24	9,84	3,08	0,22	1,86	1,14	1,63	0,94	0,16	101,92
9X 06	43- 45	496,3	6,73	55,91	17,6	9,99	3,04	0,19	2,25	1,43	1,23	0,94	0,15	99,45
10X 01	124-126	499,3	6,88	57,95	17,47	7,38	3,61	0,63	1,8	1,42	1,14	0,72	0,11	99,11
10X 02	51- 53	500,1	7	57,06	18,51	7,95	2,81	0,32	1,85	1,23	1,55	0,79	0,16	99,22
10X 05	17- 19	504,3	7,84	58,3	19,42	8,39	2,86	0,23	1,39	0,92	1,62	0,73	0,17	101,87
10X 05	59- 61	504,7	7,16	58,34	17,41	8,13	3,02	0,29	1,56	1,01	1,42	0,77	0,13	99,24
11X 01	70- 72	508,4	7,75	59,78	16,97	7,55	3,12	0,54	1,22	1,05	1,31	0,66	0,15	100,1
11X 01	137-139	509,1	7,44	58,35	18,27	8,01	2,62	0,23	1,46	0,76	1,55	0,82	0,15	99,66
11X 04	37- 39	512,6	7,84	55,63	20,84	8,84	2,34	0,1	0,72	0,37	2,11	0,78	0,16	99,73
11X 05	17-19	513,4	6,92	57,97	21,52	6,74	2,33	0,04	0,55	0,26	2,67	0,76	0,07	99,83
12X 01	47- 49	517,9	4,54	69,96	15,32	4,69	1,64	0,04	0,47	0,52	2,15	0,76	0,07	100,15
12X 03	107-109	521,5	4,52	68,95	14,75	4,83	1,63	0,04	0,49	0,41	1,99	0,82	0,09	98,52
12X 04	83- 85	522,7	7,85	53,95	21,42	9,29	2,14	0,08	0,65	0,25	2,44	0,75	0,16	98,99
12X 05	45-47	523,8	6,17	64,79	17,61	5,45	1,87	0,04	0,6	0,27	2,33	0,85	0,08	100,06
12X 05	145-147	524,8	6,67	63,32	18,18	5,51	2,19	0,04	0,53	0,33	2,67	0,81	0,06	100,31

**Table 9.** Chemical compositions of bulk sediments: Major elements concentrations. Hole 948C. Index Properties sample set.

Core/Sec	Interval (cm)	Depth (mbsf)	LOI (%)	SiO <sub>2</sub> (wt%)	Al <sub>2</sub> O <sub>3</sub> (wt%)	Fe <sub>2</sub> O <sub>3</sub> (wt%)	MgO (wt%)	MnO (wt%)	CaO (wt%)	Na <sub>2</sub> O (wt%)	K <sub>2</sub> O (wt%)	TiO <sub>2</sub> (wt%)	P <sub>2</sub> O <sub>5</sub> (wt%)	Total
12X 06	59- 61	525,5	7,68	58,97	20,49	6,46	2,2	0,18	0,58	0,25	2,66	0,82	0,09	100,39
13X 01	142-144	528,2	7,47	60,58	19,26	6,09	1,85	0,09	0,6	0,19	2,39	0,85	0,08	99,45
13X 02	101-103	529,3	22,05	31,53	11,91	9,78	2,6	6,77	12,68	0	0,85	0,54	0,28	98,97
13X 07	28- 30	536,1	7,08	62,35	18,85	6,02	1,64	0,08	0,47	0,17	2,46	0,83	0,07	100,04
14X 02	100-102	538,6	8	55,08	21,84	6,43	1,96	0,03	0,47	0	2,93	0,77	0,07	97,57
14X 04	116-118	541,8	18,51	39,2	19,32	4,92	1,34	0,08	13,94	0	0,98	0,69	0,1	99,09
15X 01	66- 68	546,1	8,49	55,02	22,14	7,81	2,33	0,03	0,68	0,17	2,18	0,89	0,18	99,93
15X 01	106-108	546,5	18,61	40,32	14,83	5,1	1,74	0,24	16,91	0,48	1,2	0,56	0,11	100,13
15X 02	51- 53	547,4	7,41	60,61	19,18	7,1	2,11	0,04	0,78	0,07	2,15	0,88	0,2	100,54
15X 05	28- 30	551,7	19,43	39,61	12,01	4,63	1,68	0,63	18,61	0	1,24	0,72	0,14	98,71
15X 06	42- 44	553,3	10,81	59	13,52	6,51	1,54	0,05	5,79	0	1,71	0,75	0,14	99,82
15X 07	44- 46	554,8	6,88	58,81	19,92	7,66	2	0,04	0,48	0,01	3,3	0,86	0,08	100,04
16X 01	106-108	555,9	6,21	61,5	18,23	7,42	2,38	0,06	0,67	0,36	2,8	0,77	0,11	100,5
16X 03	126-128	559,1	8,05	56,08	21,1	8,4	2,52	0,03	0,7	0,14	2,64	0,79	0,16	100,61
16X 04	109-111	560,4	8,18	56,02	21,59	7,36	2,59	0,03	0,73	0,12	2,41	0,96	0,17	100,14
16X 05	17- 19	561	17,93	42,97	14,74	5,32	1,29	0,07	14,85	0	1,2	0,62	0,11	99,1
17X 01	104-106	565,1	9,88	57,04	18,02	5,71	1,87	0,05	4,5	0,41	2,63	0,87	0,08	101,06
17X 02	33- 35	565,9	23,1	31,23	12,26	4,66	1,5	0,23	23,12	0	1,08	0,39	0,19	97,75
17X 04	21- 23	568,8	22,37	32,67	10,59	6,45	1,43	1,5	21,74	0	0,84	0,42	0,68	98,69
17X 06	38- 40	572	23,02	32,62	16,39	3,93	1,28	0,18	20,51	0	0,55	0,55	0,13	99,16
17X 07	12- 14	573,2	21,98	33,68	15,98	4,05	1,1	0,12	20,19	0	0,58	0,54	0,19	98,42
18X 02	28- 30	575,1	14,6	46,99	18,95	5,9	1,66	0,06	8,49	0,09	2,07	0,7	0,09	99,6
18X 02	145-147	576,2	24,63	31,4	10,21	4,93	1,21	0,5	25,64	0	0,47	0,41	0,3	99,7
18X 04	85- 87	578,7	7,72	56,01	22,05	7,98	2,18	0,04	0,61	0,2	3,01	0,79	0,08	100,67
18X 06	11- 13	580,9	7,22	55,99	20,34	9,51	2,1	0,06	0,59	0,14	3,34	0,72	0,08	100,09
19X 01	43- 45	583,2	22,04	32,03	14,06	9,43	1,9	1,75	17,25	0	0,8	0,49	0,24	99,99
19X 01	134-136	584,1	29,48	19,3	7,65	9,06	1,73	1,28	29,64	0	0,11	0,25	0,47	98,97
19X 02	15- 17	584,5	20,79	35,96	14,91	4,85	1,52	0,1	20,2	0,06	0,85	0,48	0,14	99,86
19X 04	97- 99	588,3	9,24	56,25	19,11	7,05	1,84	0,22	2,81	0,15	2,73	0,8	0,11	100,32
19X 05	33- 35	589,1	7,79	60,52	19,05	6,52	1,79	0,17	1,43	0,17	2,65	0,85	0,1	101,03
19X 05	97- 99	589,8	21,37	34,95	16,59	4,57	1,39	0,44	19,04	0	0,6	0,56	0,13	99,63
19X 06	30- 32	590,6	20,5	37,79	12,31	5,56	1,19	0,09	19,18	0	0,9	0,49	0,14	98,15
19X 06	113-115	591,4	10,48	54,01	18,81	7,17	1,68	0,09	3,92	0,01	2,45	0,77	0,09	99,47

**Table 9** (continued). Chemical compositions of bulk sediments: Major elements concentrations. Hole 948C. Index Properties sample set.

Core/Sec	Interval (cm)	Depth (mbsf)	LOI (%)	SiO2 (wt%)	Al2O3 (wt%)	Fe2O3 (wt%)	MgO (wt%)	MnO (wt%)	CaO (wt%)	Na2O (wt%)	K2O (wt%)	TiO2 (wt%)	P2O5 (wt%)	Total
2X 02	18- 20	255,49	8,19	55,19	19,56	7,5	3,7	0,28	1,46	1,39	1,95	0,87	0,16	100,24
2X 02	106-108	256,37	7,97	55,4	19,82	7,42	3,76	0,33	1,1	1,15	1,99	0,83	0,13	99,89
2X 03	65- 67	257,46	8,15	54,09	21,19	8,55	3,24	0,17	0,78	0,84	2,36	0,8	0,16	100,33
2X 04	28- 30	258,59	8,01	53,41	21,01	8,91	2,99	0,13	0,74	0,64	2,41	0,81	0,17	99,22
2X 05	31- 33	260,12	8,07	55,22	20,84	7,48	3,28	0,15	0,83	1,13	2,09	0,8	0,11	99,99
2X 06	140-142	262,71	9,9	51,19	19,41	8,16	3,47	2,98	1,33	1,02	2,18	0,75	0,21	100,59
3X 02	140-142	266,41	7,98	54,01	20,49	8,61	3,35	0,17	0,89	1,18	2,35	0,87	0,15	100,06
3X 03	50- 52	267,01	7,98	55,01	20,24	8,03	3,45	0,19	0,86	1,05	2,17	0,89	0,13	100,02
3X 06	14- 16	271,15	7,86	55,33	19,38	8,17	3,89	0,32	1,04	1,18	2,11	0,82	0,11	100,21
3X 06	63- 65	271,64	7,7	55,86	19,29	8,36	3,99	0,2	0,96	1,23	2,1	0,81	0,12	100,62
5X 01	35- 37	283,16	7,7	56,44	17,91	8,18	4,24	0,45	1,18	1,67	1,64	0,82	0,14	100,38
5X 06	23- 25	290,54	6,64	57,06	17,8	8,24	3,72	0,22	1,81	1,95	1,49	0,88	0,16	99,98
5X CC	18- 20	292,47	6,56	58,32	17,33	8,41	4,15	0,25	1,82	1,7	1,35	0,85	0,14	100,87
7X 01	86- 88	302,97	18,86	38,49	14,42	5,97	2,01	0,29	17,98	0,32	1,45	0,5	0,21	100,51
7X 02	76- 78	304,37	14,71	43,8	16,45	6,51	2,52	0,28	12,7	0,8	1,83	0,67	0,13	100,41
7X 03	127-129	306,38	9,79	51,55	19,69	6,87	3	0,19	4,8	0,99	2,65	0,77	0,13	100,43
7X 07	11-13	311,22	12,41	47,05	17,73	7,63	2,83	0,3	8,51	0,81	2,38	0,68	0,13	100,46
13X 02	27- 29	351,98	7,75	56,24	19,61	7,26	3,9	0,21	1,55	1,53	1,85	0,76	0,13	100,79
14X 03	48- 50	358,39	7,86	56,05	20,38	8,08	3,73	0,25	0,93	1,3	1,95	0,78	0,11	101,42
15X 01	144-146	361,35	7,68	55,6	19,97	7,99	3,68	0,22	0,93	1,18	2,02	0,9	0,12	100,3
15X 02	94- 96	362,35	7,68	54,94	20,67	8,46	3,38	0,21	1,02	1	2,31	0,83	0,22	100,73
15X 04	133-135	365,74	7,68	56,05	19,14	8,01	4,08	0,3	1,12	1,34	1,99	0,79	0,14	100,63
15X 05	95- 97	366,86	7,68	54,18	19,28	9,44	3,48	0,31	1,17	0,92	2,49	0,77	0,21	99,95
19X 01	135-137	400,16	8,11	57,95	19,19	7,42	2,69	0,18	1,01	0,85	1,58	0,89	0,15	100,02
19X 03	8-10	401,39	8,01	58,11	19,07	7,77	3,04	0,24	1,14	0,92	1,48	0,76	0,14	100,68
19X 04	48- 50	403,29	7,39	58,81	18,02	7,9	2,71	0,16	1,7	1,39	1,37	0,83	0,14	100,42
22X 01	102-104	428,43	7,88	55,44	20,95	9,21	2,36	0,1	0,64	0,92	2,19	0,85	0,13	100,67
22X 05	91- 93	434,32	7,33	57,1	20,62	8,2	2,5	0,06	0,61	0,46	2,38	0,84	0,12	100,23
22X 06	73- 75	435,64	7,2	56,79	20,29	9,33	2,74	0,07	0,58	0,46	2,56	0,78	0,09	100,89
25H 01	96- 98	459,37	8,4	55,24	22,81	7,05	2,46	0,13	0,51	0,17	3	0,73	0,08	100,58
25H 02	13- 15	460,04	7,43	55,48	22,06	8,75	2,55	0,04	0,5	0,3	3,35	0,74	0,09	101,29

**Table 10.** Chemical compositions of bulk sediments: Major elements concentrations. Hole 949B. Index Properties sample set.

## Annexe D

# Caractéristiques diagraphiques des minéraux



TAB. D.1 – Caractéristiques diagraphiques des principaux minéraux des formations sédimentaires (d'après Schlumberger).

# Annexe E

## Présentations et posters

### Géosciences marines

Réunion spécialisée de la Société Géologique de France  
Paris, 16 et 17 décembre 1993

**Analyse statistique et comparaison avec les mesures sur carottes des données du G.L.T. du puits 833B, Bassin Nord d'Aoba, Arc insulaire des Vanuatu (Leg 134 d'ODP).**

**A. Rabaute**, L. Briqueu, P.K. Harvey, C. Laporte et H. Mercadier

Dans sa partie centrale, l'arc volcanique insulaire des Vanuatu est composé de trois chaînes d'îles parallèles. Cette particularité structurale résulte de la collision d'un haut-fond océanique : la Zone d'Entrecasteaux, dont le Leg 134 d'ODP a permis de préciser la nature. Cette collision provoque la surrection de la chaîne Ouest et la formation du bassin Nord d'Aoba qui a enregistré l'histoire volcanique de l'arc depuis le Miocène. Le puits 833B, foré dans ce bassin, traverse 1000m de roches sédimentaires, volcano-sédimentaires et volcaniques. Des mesures en Fluorescence X effectuées au laboratoire sur des échantillons prélevés dans les carottes permettent d'évaluer la précision des mesures obtenues par diagraphie et montrent que celles-ci peuvent être entachées d'une forte erreur analytique. De ce fait, une utilisation de ces données pour la reconstitution de la minéralogie à l'aide des méthodes d'inversion ne fournira pas de résultats très fiables. Néanmoins, ces données restent statistiquement significatives. En effet, les 9 éléments mesurés en continu par le Geochemical Logging Tool de Schlumberger sur 500m environ de profondeur dans ce puits ont permis par des techniques d'analyses statistiques multivariées de visualiser les différents groupes qui composent la séquence lithostratigraphique : la transition sédiments-sill basaltique vers 830 mbsf est notamment très bien marquée. De même, les traitements individualisent bien les sédiments carbonatés des sédiments plus siliceux. Afin de mieux contraindre le dépouillement des données GLT, nous développons un protocole d'analyse rapide des échantillons par fluorescence X sur pastilles de poudres pressées. Nous l'avons testé avec succès pour 3 éléments majeurs (K, Ti, Fe). Il pourrait permettre,

à bord du Joides Resolution, l'acquisition rapide d'un nombre important de mesures sur carottes permettant de contraindre le dépouillement des données G.L.T.

---

### 1er Colloque des Doctorants Géochimie Grand-Sud

Observatoire Midi-Pyrénées, Toulouse  
19 et 20 Octobre 1994

**Etude et interprétation des données de diagraphie géochimique réalisée dans les puits de forage ODP. Contribution à la compréhension des subductions intra-océaniques.**

**A. Rabaute**

Ce travail a été commencé lors de mon DEA par l'évaluation de données des diagraphies géochimiques réalisées dans le puits 833B pendant le Leg ODP 134, dont l'objet était l'étude de l'arc volcanique des Vanuatu. Ce puits se trouve dans la partie nord du bassin d'Aoba, dont la formation résulte de la collision de la Zone d'Entrecasteaux, ancienne ride océanique, avec la partie centrale de l'arc. Ce bassin sédimentaire, situé à l'aplomb de la ligne volcanique active, a enregistré toute l'histoire volcanique de l'arc.

Les données de diagraphies géochimiques ont été réalisées grâce à l'outil GLT de Schlumberger, et couvrent environ 1000 mètres de formation comprenant toute la partie volcano-sédimentaire et sédimentaire du bassin, ainsi que le sill basaltique rencontré à 830 mètres sous le plancher océanique. L'outil GLT permet d'avoir les abondances en poids pourcent d'oxydes de 6 éléments majeurs (Ca, Si, Fe, Al, K et Ti) ainsi que les teneurs en ppm de 3 éléments traces (U, Th, et Gd). Les traitements statistiques effectués, soit à l'aide de méthodes classiques comme l'Analyse en Composantes Principales, ou bien d'analyses multivariées (INCA, Itérative Non-hierarchical Cluster Analysis) permettant de classer les données en groupes de mêmes affinités géochimiques, ont permis de différencier la partie sédiment du sill basaltique, et de séparer les bancs plus siliceux (volcano-sédimentaires) des bancs plus carbonatés (sédiments pélagiques). Néanmoins, l'analyse n'a pas

fait ressortir la brèche basaltique, décrite de 360 à 560 mètres sous le plancher océanique, de la partie volcano-sédimentaire. La qualité des données de diagraphie est fortement dépendante des conditions d'acquisition (en particulier l'état du trou). On arrive par l'analyse en Composantes Principales à isoler un groupe parfaitement corrélé avec les anomalies du puits de forage, enregistrées sur la trace du caliper. Des mesures en fluorescence X effectuées sur des échantillons prélevés sur les carottes du puits 833B ont permis d'évaluer l'incertitude des données de diagraphie. Ces résultats pourront être utilisés par la suite pour calibrer l'outil de diagraphie et améliorer la qualité des données de logging.

Ma participation au Leg 156 pendant les mois de Juin-Juillet 94 m'a permis de voir la mise en oeuvre d'une nouvelle technique de logging expérimentée lors de ce leg, le logging en cours de forage (Logging-While-Drilling, LWD). Elle a fourni des mesures de résistivité, porosité, densité et rayonnement gamma naturel en continu le long de deux puits de forage (947A et 948A), d'une qualité très supérieure à celle couramment obtenue en diagraphie classique (en trou ouvert ou "au câble") dans ce type d'environnement très instable que sont les prismes d'accrétion. Les données de fluorescence X obtenues sur des échantillons prélevés de façon systématique sur les carottes remontées du puits 948C vont permettre, de part leur fréquence et leur nombre, de calibrer les outils du LWD et d'évaluer la précision de la mesure de diagraphie ainsi effectuée ainsi que l'impact des conditions environnementales. Le logging en cours de forage a permis lors du Leg 156 de confirmer la présence au niveau de la zone de décollement du prisme d'accrétion de la Barbade, de niveaux à forte porosité. En effet, la diminution conjointe, en plusieurs parties du log, de la résistivité et de la densité mesurées en continu dans les deux puits, suggère des profils de compaction différents de ceux couramment envisagés dans les prismes d'accrétion. Il a en outre été possible, à partir des données du LWD, de différencier plusieurs unités se corrélant assez bien avec les unités définies sur carottes, de faire apparaître les intervalles plus ou moins riches en minéraux argileux, et de distinguer avec une très bonne précision la transition lithologique reconnue sur carotte à 514 mètres sous le plancher océanique. Les données de K, Th et U obtenues en fluorescence X sur les échantillons du puits 948C, couplées avec les données de logging, devraient permettre d'individualiser des niveaux ayant des propriétés physiques et chimiques directement corrélables avec les études sédimentologique et minéralogique.

## 2ème Colloque des Doctorants Géochimie Grand-Sud

Institut des Sciences de la Terre, de l'Eau et de l'Espace de Montpellier  
Université Montpellier II  
18 et 19 Septembre 1995

## La structure du mont Eratosthènes (Méditerranée orientale) et son implication dans la collision Afrique du Nord - Eurasie. Interprétations préliminaires des diagraphies géochimiques réalisées lors de la campagne ODP 160.

A. Rabaute

Le mont Eratosthènes est un haut fond océanique localisé dans le Bassin Levantin en Méditerranée orientale entre Chypre et la marge continentale de l'Égypte. L'un des objectifs de la campagne ODP 160, qui a eu lieu du 12 Mars au 3 Mai 1995, était de mieux comprendre la structure et l'âge du mont Eratosthènes, ainsi que la dynamique du processus de collision engagé entre la plaque Afrique et l'Eurasie au niveau de Chypre. Cette collision aurait été le moteur du soulèvement de l'île de Chypre, entraînant celui de l'ophiolite de Troodos. Les données sismiques de reconnaissance collectées en vue du Leg 160 durant la campagne UNESCO-ESF du N/O Gelendzhik (Limonov et al., 1994) ont fourni la base d'une telle interprétation que le forage de 4 puits le long du mont Eratosthènes permettait de tester en situ dans le temps le début de la collision.

Les résultats préliminaires de ces quatre forages effectués sur la zone permettent de contraindre une histoire géologique pour cette structure :

Le site 966, foré au sommet de ce haut-fond, a traversé des calcaires micritiques d'âge crétacé supérieur à éocène moyen (85 à 45 Ma), intercalés localement de niveaux de chert et de matière organique. Le mont Eratosthènes était alors un fragment crustal détaché de la marge passive africaine, alors subsidente. S'il n'a pas souffert de la tectonique active qui s'est développée plus au nord de la Thétys vers 70-75 Ma, le processus de subduction va entraîner son soulèvement de plus d'un kilomètre, ce qui est montré par les calcaires corraliens d'eau peu profonde rencontrés dans les forages 966 et 965, implantés sur le flanc nord du mont sous-marin. Vers la fin du Miocène (5-7 Ma), le mont Eratosthènes forme un relief émergé, le niveau de la Méditerranée étant plus bas à cette époque que de nos jours. Cette période est présente dans les carottes du puits 965 sous forme de niveaux dolomitiques et de paléosols.

Les données des diagraphies géochimique et géophysique, ainsi que du Formation Micro-Scanner (FMS), obtenues sur le site 967 montrent très clairement la présence d'un niveau d'évaporites déposé lors de la crise messinienne. Une période de subsidence commence ensuite au début du Pliocène, due à une charge crustale de plus en plus élevée à cet endroit, ainsi qu'à une tectonique active, résultant de la flexuration de la plaque induite par la subduction au nord. Les données de terrain sur Chypre ont montré que le maximum du soulèvement de l'île datait de la fin du Pliocène, soit environ 3 Ma après le début de la subsidence du mont Eratosthènes.

Les données de logging géophysique et FMS, recueillis sur les quatre sites forés, ainsi que les données de



diagraphie géochimique obtenues sur le site 967, profond de 600 mètres, sont particulièrement utiles dans cet environnement car leur analyse permet de combler les nombreux vides laissés par un carottage très difficile et incomplet (en moyenne 30% de récupération). Leur analyse préliminaire, faite rapidement à bord, a permis d'identifier différentes unités lithologiques et de les corréliser avec les observations sur carottes. Une analyse plus poussée utilisant des méthodes statistiques (analyse discriminante, par nuées dynamiques, suivies d'une inversion minérale) pourra donner des indications précises et utiles sur les formations présentes dans les intervalles non récupérés par le carottage. En particulier, l'utilisation de l'outil géochimique dans le puits 967 a déjà permis de confirmer la présence du niveau de gypse correspondant à l'émergence de la structure lors de la crise messinienne. Ces traitements fourniront des indications sur la minéralogie et la composition chimique des formations traversées, éléments importants pour l'étude des flux de chaleur, des circulations de fluides et des conditions paléo-océanographiques au moment de leur dépôts.

---

#### Géosciences marines

Réunion spécialisée de la Société Géologique de France  
Les forages océaniques : Nouveaux enjeux  
Paris, 27 et 28 janvier 1998

#### Structure minéralogique du prisme d'accrétion de la Barbade (Leg ODP 156) : inversion linéaire des données géochimiques et non linéaire des données diagraphiques LWD.

Alain Rabaute, Louis Briquet, Mohamed Ramadan

Le but du Leg 156 d'Ocean Drilling Program, qui a eu lieu en Juin-Juillet 1994 sur la partie nord du prisme d'accrétion de la Barbade, était d'étudier la dynamique de la zone de décollement située entre le prisme et la plaque subduite, et notamment les circulations de fluides, leurs effets sur les formations géologiques, ainsi que la tectonique qui y est attachée. Pour la première fois au cours d'une campagne ODP, une technique particulière de diagraphie, le "Logging-While-Drilling, LWD", ou diagraphie en cours de forage, a été mise en oeuvre dans deux puits, dont l'un, le 948A, a traversé la zone de décollement (située entre 498 et 580 mspo), qui a ensuite été carottée en continu dans le puits 948C.

La technique de diagraphie en cours de forage est très performante comparée à la technique classique de diagraphie au câble car elle permet de réduire les effets environnementaux (instabilité du puits, changement des conditions in-situ) qui peuvent dégrader la qualité des mesures durant l'acquisition. Les paramètres mesurés par LWD sont la résistivité à grande profondeur d'investigation par induction, la résistivité de la zone envahie par un dispositif focalisé, la porosité neutron, le rayonnement gamma naturel, et la masse volumique de

la formation. En parallèle, 150 échantillons ont été prélevés à intervalle régulier sur les carottes du puits 948C et ont fait l'objet de la mesure par fluorescence X des concentrations en 9 éléments chimiques (Si, Ca, Fe, Al, Mg, Mn, K, Ti, Na), ainsi que de la détermination de la paragenèse minérale par diffraction X. La densité de grain, la susceptibilité magnétique et le rayonnement gamma naturel ont été mesurés en continu à bord sur les carottes du puits 948C par l'outil Multi Sensor Track (MST).

Nous avons mis en oeuvre en parallèle deux méthodes déterministes d'inversion minérale, l'algorithme des moindres carrés, et celui de Tarantola et Valette, et une méthode statistique, le réseau de neurone à rétropropagation du gradient, dont nous comparons ici les résultats respectifs. Dans l'algorithme des moindres carrés, nous avons introduit la prise en compte des erreurs de mesure ainsi que l'indétermination sur les composants de la matrice des compositions minérales. L'algorithme de Tarantola et Valette permet une inversion moins rigide car il considère les compositions des minéraux de la matrice des paramètres comme des variables et non comme des constantes. Il intègre de plus les indéterminations sur toutes les variables. Grâce à l'inversion minérale sur les données de diagraphie LWD par l'utilisation d'un réseau de neurones, nous obtenons une colonne lithologique en continu, donnant les proportions des différentes phases minérales présentes, dont l'estimation des proportions des différents minéraux argileux. Cette inversion se base sur les résultats obtenus par inversion des données géochimiques obtenues par fluorescence et diffraction X, ainsi que les données MST. L'information lithologique en continu ainsi obtenue est nécessaire pour établir un modèle de la dynamique du prisme qui permettra de recalculer une porosité "hors-argiles", de tracer le chemin poreux et ainsi de comprendre les circulations de fluides. *Présentation sous forme de poster, voir ci-contre.*

---

#### 17ème Réunion des Sciences de la Terre Brest, 31 Mars au 3 Avril 1998

#### Détermination en continu de la minéralogie à partir de mesures de diagraphie : exemple du forage profond de l'Ardèche BA-1, Programme BRGM Géologie Profonde de la France.

Alain Rabaute, Louis Briquet, Mohamed Ramadan

Les données de la diagraphie géochimique acquises dans le puits Balazuc-1, ont été utilisées pour recalculer les compositions minéralogiques des formations traversées en utilisant des modèles à la fois déterministe (modèle de mélange traité en utilisant l'algorithme proposé par Tarantola et Valette qui permet de prendre en compte les incertitudes sur tous les paramètres) et statistique (réseau de neurones). La première méthode permet d'arriver à un résultat très satisfaisant sous réserve d'un bon choix de la solution *a priori* et de son

incertitude. Pour permettre de mieux définir cette information *a priori*, nous faisons précéder l'inversion minérale d'un partitionnement du jeu de données en groupes de composition homogène grâce à une classification automatique en nuées dynamiques. On procède alors à l'inversion minérale sur chaque groupe en utilisant des paramètres de départ différents. L'inversion minérale étant un modèle de mélange, elle ne s'applique qu'à des variables additives (pourcentages d'oxydes provenant de l'outil géochimique de diagraphie, masse volumique par exemple). Par contre l'utilisation d'un réseau de neurones, du type carte auto-organisatrice de Kohonen, va permettre de prendre en compte dans le traitement des variables mesurées aussi en continu par diagraphie mais ne satisfaisant pas à une loi déterministe comme la résistivité, la vitesse des ondes acoustiques, la porosité. Pourtant, ces grandeurs rendent compte de propriétés importantes de la formation. La carte auto-organisatrice est un outil non linéaire de classification automatique indépendant de tout modèle *a priori*. L'algorithme tire aléatoirement dans le jeu de données les informations dont il a besoin pour son auto-organisation. Les résultats de l'inversion minéralogique peuvent être utilisés *a posteriori* pour identifier les différents faciès individualisés par la carte auto-organisatrice et reconstruire en continu la lithologie des formations traversées.

---

**XXXèmes Journées de Statistique**  
Campus de Ker Lann, Bruz, Mai 1998

**Mise en oeuvre complémentaire de différentes approches statistiques (classifications, moindres carrés, perceptron) et probabilistes (inversion totale) pour l'interprétation des diagraphies mesurées dans le forage 948 (Campagne 156 d'ODP). Comparaison et validation.**

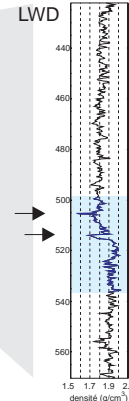
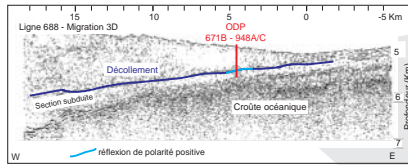
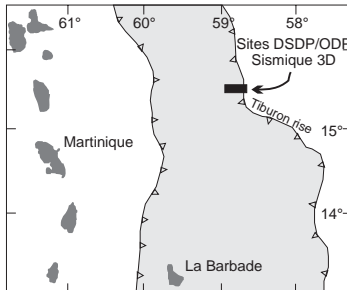
**Alain Rabaute**, Louis Briquet, Mohamed Ramadan,  
Olivier Voutay

**Principe** Parmi les nombreuses grandeurs physiques et chimiques mesurées en continu dans les puits de forage par les différentes méthodes de diagraphie, certaines satisfont à des lois déterministes, d'autres pas. Parmi les premières figurent les concentrations chimiques en éléments majeurs, exprimées sous forme de pourcentages d'oxyde ( $\text{SiO}_2$ ,  $\text{Al}_2\text{O}_3$ ,  $\text{Fe}_2\text{O}_3$ ,  $\text{MgO}$ ,  $\text{CaO}$ ,  $\text{K}_2\text{O}$ ,  $\text{Na}_2\text{O}$ ,  $\text{TiO}_2$ , ...) dont la somme est généralement proche de 100%. Ces grandeurs sont régies par une loi élémentaire de mélange : la concentration en un élément chimique d'une formation géologique sera égale au produit scalaire entre concentrations en cet élément dans

# Structure minéralogique du prisme d'accrétion de la Barbade (Leg ODP 156): inversion linéaire de données géochimiques

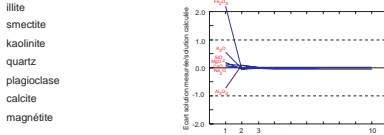
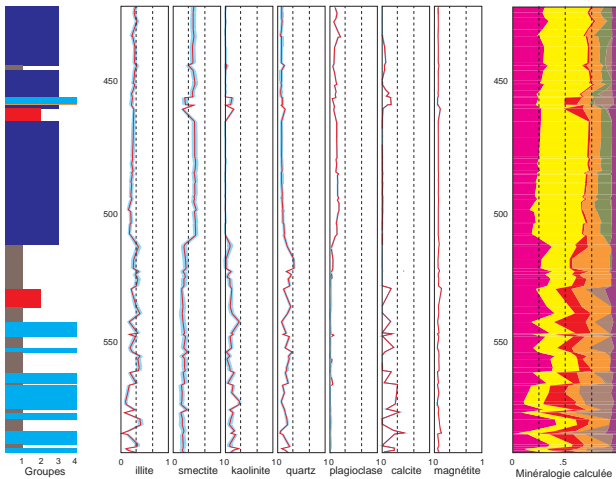
Alain RABAUTE, Louis BRIQUEU, Mohamed RAMADAN

Coordonnées : UMR 5567-CNRS, ISTEEM, Université de Montpellier II  
rabaute@dstu.univ-montp2.fr

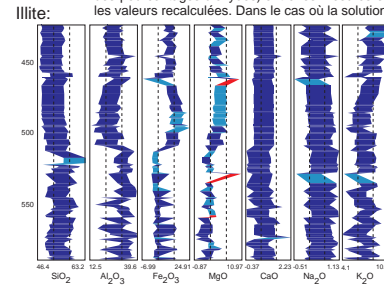


La campagne 156 d'ODP s'est déroulée dans la partie frontale du prisme nord-Barbade et avait pour but d'étudier divers paramètres (perméabilité, débit, épisodicité, surpression) liés à la dynamique des fluides qui circulent dans la zone de décollement. Une campagne de sismique réflexion 3D effectuée avant le leg 156 montra que le décollement est caractérisé par une réflexion très intense de polarité soit négative, due à la compaction, soit positive qui serait due à une forte porosité liée à une pression de l'eau interstitielle anormalement élevée à cet endroit. La méthode de diaggraphie en cours de forage ou Logging-While-Drilling a été utilisée dans le puits 948A car elle permet de minimiser les problèmes d'instabilité de puits qui rendent la diaggraphie en puits ouverts très difficile dans les environnements instables comme un prisme d'accrétion actif. Des enregistrements en continu de la résistivité, du rayonnement gamma naturel, de la densité totale ainsi que de la porosité neutron ont été obtenus jusqu'à 50 m en dessous du décollement dans le puits 948A. Au niveau de la zone de décollement, le log de densité montre deux pics (flèches) à faible densité. Ces porosités anormalement

L'approche



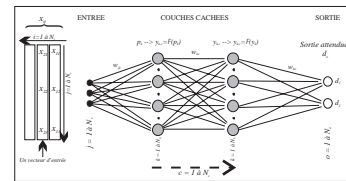
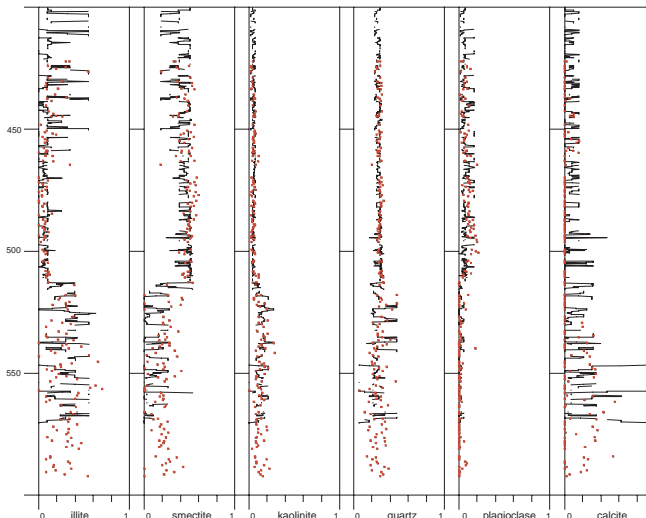
Il est possible d'enregistrer à chaque itération l'écart cumulé pour chaque variable d'entrée (ici, des pourcentages d'oxydes) entre les mesures et les valeurs recalculées. Dans le cas où la solution



L'algorithme de l'inversion totale proposé par Tarantola et Valette (1982) a été utilisé sur les pourcentages d'oxydes mesurés par fluorescence X sur des échantillons du puits 948C. Cette méthode considère données et paramètres du problème comme les composantes d'un vecteur unique X affecté d'une matrice de covariance M représentant les incertitudes. On cherche à minimiser la forme quadratique entre X et une solution *a priori* acceptable X0. Une classification automatique partitionne au préalable le jeu de données (à gauche) en 4 groupes, ce

De la même façon, pour chaque phase minérale déterminée, il est possible de déterminer l'écart entre les pourcentages d'oxydes calculés et les paramètres de départ, en tenant compte de leurs incertitudes et de l'erreur calculée sur la solution *a posteriori*. Par exemple, pour l'illite (ci-dessus), le bleu foncé signale que la solution est dans l'intervalle de

L'approche



Le principe du perceptron, ou réseau de neurone à rétropropagation du gradient, est illustré ci-dessus. Il se compose de plusieurs couches successives de neurones dont chaque élément est le produit scalaire entre une transformation de type  $y = F(x_1, \dots, x_n; W)$  où la fonction F est une fonction de transfert sigmoïde, et W la matrice

Le réseau de neurones a été utilisé sur les paramètres physiques mesurés par LWD dans le puits 948A, soit résistivités à faible et moyenne profondeurs d'investigation, densité totale, porosité neutron et rayonnement gamma naturel total. Le jeu d'entraînement fourni au perceptron pour son apprentissage a été composé à partir d'intervalles homogènes LWD individualisés par classification automatique sur des données densité, susceptibilité magnétique et rayonnement gamma naturel acquises sur carottes. Le perceptron donne pour chaque niveau de mesure LWD une probabilité d'appartenance à un ou plusieurs des groupes de départ, donc une minéralogie (courbe noire ci-contre). Les points rouges donnent la minéralogie

chacun des constituants minéraux et leurs proportions respectives  $x_i$ , soit par exemple pour le potassium :

$$[K_2O]_{Total} = \sum_{i=1}^n [K_2O]_i \cdot x_i \quad (E.1)$$

De plus, les proportions des phases minérales dans la roche totale satisfont à une relation de fermeture : leur somme doit être égale à l'unité.

**Les moindres carrés** Les concentrations en éléments chimiques majeurs sont mesurées soit *in situ* dans le forage tous les 15,24 cm, soit sur des échantillons prélevés sur carottes et analysés au laboratoire par quelque méthode analytique que ce soit. On cherchera à partir des concentrations chimiques, à déterminer les proportions des différentes phases minérales de façon à reconstituer la lithologie des différentes couches géologiques. Le problème posé consiste à résoudre un système élémentaire d'équations linéaires du type

$$y_i = \sum c_{ij} x_j \text{ soit } Y = CX \quad (E.2)$$

où  $i$  représente le nombre d'éléments chimiques considérés, et  $j$  le nombre de phases minérales. La relation de fermeture est traitée par la méthode des multiplicateurs de Lagrange (Albarède 1995).

Même si les concentrations chimiques pour certains éléments ( $K_2O$ ,  $TiO_2$ ,  $Fe_2O_3$ ) sont mesurées plus précisément que d'autres ( $SiO_2$ ,  $Al_2O_3$ , ...), toutes ces mesures réalisées *in situ* ou en laboratoire sont entachées d'erreur. Leur prise en compte par le truchement d'une matrice de pondération ne pose aucune difficulté particulière lors de la mise en oeuvre des moindres carrés quel que soit l'algorithme utilisé (SVD ou pas) et s'exprime sous la forme :

$$X = (C^T M^{-1} C)^{-1} C^T M^{-1} Y \quad (E.3)$$

$M$  étant la matrice des covariances des erreurs de mesure. Si l'on considère que ces erreurs sont indépendantes et suivent une distribution normale centrée sur zéro, cette matrice est diagonale de la forme :

$$M = \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{pmatrix} \quad (E.4)$$

où  $\sigma_i^2$  est la variance relative à  $y_i$ . Par contre une difficulté importante résulte du fait que les paramètres du modèle sont eux-mêmes entachés d'incertitude. En effet tous les feldspaths n'ont pas la même composition chimique, et il en est *a fortiori* de même pour les différents minéraux argileux (illite, smectite, kaolinite). La matrice des paramètres est définie soit à partir de valeurs moyennes déduites de compilations prises dans des ouvrages spécialisés (Deer et al. 1967), soit à partir de mesures ponctuelles à la microsonde électronique sur des échantillons de roche. L'incertitude sur ces paramètres peut être relativement bien quantifiée mais reste difficile à prendre en compte dans la méthode des moindres

carrés. A l'instar de ce qui se pratique de longue date en géologie pour le traitement des données radiométriques (York 1969, Brook 1972, Minster et al. 1982, Briquieu et de la Boisse 1990), nous prenons en compte ces incertitudes en complexifiant les coefficients de pondération :

$$w_i = \frac{1}{\sigma_{y_i}^2 - \sum_{j=1}^n x_j^2 \sigma_{c_{ij}}^2} \quad (E.5)$$

Introduits sous cette forme, les poids dépendent de la solution  $X$  elle-même ce qui conduit à mettre oeuvre une méthode itérative tout en restant dans le cadre des moindres carrés.

**L'inversion totale** Une approche alternative consiste à écrire le système sous la forme de  $n+1$  équations implicites dont on cherche les racines grâce à l'algorithme de l'inversion totale proposé par Tarantola et Valette (1982). Le système d'équations prend la forme suivante :

$$\begin{aligned} y_i - \sum_{j=1}^m c_{ij} x_j &= 0 & \text{avec } i = 1, n \\ \sum_{j=1}^m x_j - 1 &= 0 \end{aligned} \quad (E.6)$$

que l'on peut écrire  $G(X) = 0$  en considérant que données et paramètres du modèle sont les composantes d'un vecteur unique  $X$  auquel est affectée une matrice de covariance unique  $M$ .

Dès lors, le principe de l'algorithme consiste à minimiser la distance quadratique

$$q^2 = (X - X_0)^T M_0^{-1} (X - X_0) \quad (E.7)$$

$X_0$  étant une solution *a priori* acceptable, sous les  $n+1$  contraintes  $G(X) = 0$ . La méthode des multiplicateurs de Lagrange permet d'écrire  $q^2$  sous la forme

$$q^2 = (X - X_0)^T (X - X_0) - 2 \sum_{i=1}^{n+1} \lambda_i g_i(X) \quad (E.8)$$

soit en différenciant :

$$dq^2 = dx^T [X - X_0 - \sum_{i=1}^{n+1} \lambda_i \text{grad}(g_i(X))] \quad (E.9)$$

Le minimum de  $q^2$  est tel que  $X - X_0 = H\lambda$ ,  $H$  étant le Hessian de  $G(X)$ . Le vecteur  $\lambda$  des multiplicateurs de Lagrange est

$$\lambda = (H^T H)^{-1} H^T (X - X_0) \quad (E.10)$$

La solution proposée par Tarantola et Valette (1982) prend aussi en compte les incertitudes sur les paramètres par le biais de la matrice des covariances  $M_0$  *a priori*. Elle est donnée par

$$X_{k+1} = X_0 + M_0 \cdot H_k^T (H_k \cdot M_0 \cdot H_k^T)^{-1} [H_k \cdot (X - X_0) - G(X)]$$

Une bonne approximation de la matrice des covariances *a posteriori* est donnée par Tarantola et Valette (1982) :

$$M = M_0 - M_0 \cdot H^T \cdot (H \cdot M_0 \cdot H^T)^{-1} \cdot H \cdot M_0 \quad (E.11)$$

Dans la mesure où l'on dispose d'une information *a priori* sur le système des paramètres, l'algorithme proposé de l'inversion totale est une approche bayésienne de l'estimation de ces paramètres, dont la particularité est

d'être basée sur l'algorithme du point fixe. On met ainsi en évidence l'importance de l'information *a priori* sur la qualité de l'estimation des paramètres du modèle et de la solution *a posteriori*. Les gains d'information sont d'autant plus importants que la solution *a priori* est adaptée au problème minéralogique traité. Nous choisissons de partitionner, préalablement au traitement numérique, le jeu de données à l'aide d'une classification automatique, du type nuées dynamiques (ou de sa généralisation : la carte auto-organisatrice ; voir Kohonen 1994, Anouar et al 1997, Lechevallier et Le Corre 1997), qui va permettre de choisir, pour chaque intervalle reconnu, une solution *a priori* d'autant plus proche du maximum de vraisemblance.

**Le réseau de neurones** Les méthodes citées précédemment, moindres carrés ou inversion totale, sont soumises à une contrainte importante : elles ne s'appliquent qu'à des grandeurs qui satisfont à des fonctionnelles établies. C'est le cas des compositions chimiques, à partir desquelles on peut faire une inversion minérale. Par contre, des variables non-additives, comme la résistivité, la vitesse des ondes acoustiques, la porosité dans une moindre mesure, ne pourront pas être utilisées dans de tels modèles. Or ces variables apportent des renseignements importants sur les propriétés physiques des roches. La résistivité électrique, par exemple, renseigne sur les fluides présents dans la formation, et donne des indications sur son degré de consolidation ou sa texture. Une façon de prendre en compte ces grandeurs peut être la mise en oeuvre de réseaux de neurones, que ce soit un perceptron, ou une carte auto-organisatrice. Ce type de méthode nécessite un apprentissage ou une indexation à l'aide d'un jeu de données d'entraînement. Pour le constituer, on peut exploiter les résultats de l'inversion minérale obtenus lors de l'étape précédente et définir plusieurs intervalles lithologiques relativement homogènes, et le plus représentatifs possible de la complexité de la géologie. En supposant raisonnablement que les paramètres physiques mesurés en continu par diagraphie décrivent ces groupes avec la même cohérence que les mesures géochimiques réalisées sur les échantillons, on entraîne le réseau à reconnaître ces intervalles en utilisant les paramètres physiques, d'abord sur un jeu de données de test connu, puis sur le jeu de données de diagraphie complet.

**Application** L'exemple traité concerne le forage 948 localisé sur le prisme d'accrétion de la Barbade, dans les Petites Antilles, et utilise à la fois les données de diagraphie et des mesures géochimiques obtenues en laboratoire. Le but de la campagne océanographique était de mieux comprendre la zone de décollement située entre le prisme sédimentaire proprement dit et la plaque lithosphérique plongeante sous-jacente. Du fait de la convergence de plaques, cette zone est le lieu de circulations de fluides en surpression dont les effets sur les formations adjacentes sont mal connus. Dans notre exemple, 50 à 70% de la roche est composée d'assem-

blages de minéraux argileux, en majorité illite, smectite, et kaolinite, de minéraux détritiques (tels que quartz, plagioclase) et marins (calcite). En laboratoire, il est difficile de déterminer la proportion des minéraux argileux car ils sont souvent mal cristallisés, ou de petite taille. Ils se prêtent mal à des études en microscopie optique ou en diffraction X. Pourtant, cette détermination est cruciale pour calculer la porosité de la formation, car par exemple, la smectite est 10 à 20 fois plus riche en eau que l'illite. Or, la porosité mesurée par diagraphie est déduite de la mesure de la quantité d'atomes d'hydrogène présents dans la formation, qu'ils soient libres ou liés aux argiles. Connaître la proportion de smectite va permettre de corriger cette mesure pour calculer la quantité de fluides (eau ou hydrocarbure) effectivement disponible. Nous avons vu que la méthode de l'inversion totale des mesures géochimiques sur échantillons de roche, sous réserve d'un bon choix des solutions *a priori*, pouvait aider à reconstituer la séquence minéralogique de la zone carottée autour de la zone de décollement. L'information apportée par échantillonnage reste toutefois discrète (le pas d'échantillonnage est d'environ 1 m), donc incomplète. Le couplage des informations données par les classifications (perceptron ou carte auto-organisatrice) utilisant les paramètres physiques mesurés en continu par diagraphie, avec les différents groupes minéralogiques individualisés par l'inversion totale, permet d'obtenir un log lithologique continu de la zone de décollement dont la résolution atteint maintenant une quinzaine de centimètres.

### Bibliographie

- Albarède, F.** (1995) Introduction to geochemical modeling. Cambridge University Press.
- Anouar, F., Badran, F. et Thiria, S** (1997) Cartes topologiques et nuées dynamiques. *Actes des 29e Journées de Statistique*, Carcassonne, 26-29 Mai 1997, 1-4.
- Briqueu, L. and de la Boisse, H.** (1990) U-Pb geochronology : systematic development of mixing equation and application of Monte-Carlo numerical simulation to the error propagation in the Concordia diagram. *Chemical Geology*, **88**, 69-83.
- Brooks, C., Hart, S.R. and Wendt I.** (1972) Realistic Use of two-error regression treatments as applied to Rubidium-Strontium data. *Review of Geophysics*, **10**, 551-577.
- Deer, W.A., Howie, R.A. and Zussman, M.A.** (1967) Rock-Forming Minerals, vol 1-5. Longmans, Green and Coltd *Eds.*
- Kohonen, T** (1994) Self-organizing maps. Springer Verlag.
- Le Chevallier, Y. et Le Corre F.** (1997) Classification des profils hydrologiques par une carte auto-organisatrice. *Actes des 29e Journées de Statistique*, Carcassonne, 26-29 Mai 1997, 527-530.
- Minster, J.F., Richard, L.P. and Allègre, C.J.** (1982)  $^{87}\text{Rb}/^{86}\text{Sr}$  chronology of enstatite meteorites. *Earth Planet. Sci. Lett.*, **44**, 420-440.

- Tarantola, A.** and Valette, B. (1982) Generalized non linear inverse problem solved using the least square criterion. *Rev. Geophysics and Space Physics*, **20**, 219-232.
- York, D.** (1969) Least squares fitting of a straight line with correlated errors. *Earth Planet. Sci. Lett.*, **5**, 320-324.

# Bibliographie

- (1984). *Spectrométrie du rayonnement gamma naturel. Principes essentiels d'interprétation des diagraphies NGS*. Schlumberger, afm-marketing edition.
- Adams, J. A. and Weaver, C. E. (1958). Thorium-to-uranium ratios as indicators of sedimentary processes : example of concept of geochemical facies. *Am. Assoc. Pet. Geol. Bull.*, 42(2) :387–430.
- Aitchison, J. (1984). The statistical analysis of geochemical compositions. *Mathematical Geology*, 16(6) :531–564.
- Aitchison, J. (1986). *The statistical analysis of compositional data*. John Wiley, New-York.
- Albarède, F. (1995). *Introduction to geochemical modeling*. Cambridge University Press.
- Aquilina, L., Scheerens, S., Vinchon, C., Dromart, G., and Le Strat, P. (1997). Cored section of Balazuc well (South East Basin of France). Geological and geochemical approach (Deep Geology of France Program). *Mar. Petrol. Geol.*, 13(6) :725–735.
- Archie, G. E. (1942). The electrical resistivity log as an aid in determining some reservoir characteristics. *Journal of Petroleum Technology*, 5(TP 1422) :310–314.
- Augustin, A. (1990). *Diagraphie d'induction magnétique et de susceptibilité magnétique en forage ; cas des terrains sédimentaires, interprétation vers des applications en magnétostratigraphie*. PhD thesis, Université de Grenoble I.
- Ball, G. H. and Hall, D. J. (1967). A clustering technique for summarizing multivariate data. *Behavioral Science*, 12 :153–155.
- Barber, T. D. (1986a). Invasion profiling with the phasor induction tool. In *Transactions of the SPWLA*.
- Barber, T. D. (1986b). Real-time environmental corrections for the DIT-E digital dual induction tool. In *Transactions of the SPWLA*.
- Bertozzi, W., Ellis, D., and Wahl, J. S. (1981). The physical foundations of formation lithology logging with gamma rays. *Geophysics*, 46(10) :1439–1455.
- Bevington, P. R. (1969). *Data reduction and error analysis for the physical sciences*, pages 164–176. McGraw-Hill, New-York.
- Bezdek, J. C., Ehrlich, R., and Full, W. (1984). FCM : The Fuzzy *c*-Means Clustering algorithm. *Computers and Geosciences*, 10(2–3) :191–203.
- Bloch, F. (1946). Nuclear induction. *The Physical Review*, 70.
- Bloch, F., Hansen, W. W., and Packard, M. (1946). The nuclear induction experiment. *The Physical Review*, 70.

- Blum, P., Rabaute, A., Gaudon, P., and Allan, J. F. (1997). Analysis of natural gamma-ray spectra obtained from sediment cores with the shipboard scintillation detector of the Ocean Drilling Program : Example from Leg 156. In Shipley, T. H., Ogawa, Y., Blum, P., and Bahr, J. M., editors, *Proceeding of the Ocean Drilling Program, Scientific Results*, volume 156, pages 183–195.
- Boillot, G., Montadert, L., Lemoine, M., and Biju Duval, B. (1984). *Les marges continentales actuelles et fossiles autour de la France*. Masson.
- Bouisset, P. and Augustin, A. (1993). Borehole magnetostratigraphy, absolute age dating and correlations of sedimentary rocks, with examples from the Paris basin (France). *Amer. Assoc. Petr. Geol.*
- Briqueu, L. and de la Boisse, H. (1990). U-Pb geochronology : systematic development of mixing equation and application of Monte-Carlo numerical simulation to the error propagation in the Concordia diagram. *Chemical Geology*, (88) :69–83.
- Brooks, C., Hart, S. R., and Wendt, I. (1972). Realistic use of two-error regression treatment as applied to rubidium-strontium data. *Review of Geophysics*, (10) :551–577.
- Carr, H. Y. and Purcell, E. M. (1954). Effects of diffusion on free precession in nuclear magnetic resonance experiments. *The Physical Review*, 94(3) :630–638.
- Chang, H. C., Chen, H. C., and Fang, J. H. (1997). Lithology determination from well-logs with fuzzy associate memory neural network. In *IEEE Transactions on Geosciences and Remote Sensing*, volume 35, pages 773–780.
- Chayes, F. (1971). *Ratio correlation*. University of Chicago Press.
- Chemali, R., Gianzero, S., Strickland, R., and Tijani, S. M. (1983). The shoulder bed effect on the dual laterolog and its variation with the resistivity of the borehole fluid. In *Transactions of the SPWLA*, number UU.
- Courel, L., Poli, E., Le Strat, P., Jacquin, T., Baud, A., and Vannier, F. (1997). Sequence stratigraphy along a Triassic transect on the western Peritethyan margin in Ardèche (Basin Sud-Est France), correlations with Subalpine and Germanic realms. In Graciansky, P. C., Hardenbol, J., Jacquin, T., Farley, B., and Vail, P. R., editors, *Mésozoic and Cenozoic sequence stratigraphy of European Basins*, *SEPM Spec. Pub.*
- Cross, W., Iddings, J. P., Pirsson, . V., and Washington, H. S. (1902). *Quantitative classification of igneous rocks*. University of Chicago Press, Chicago.
- Davis, R. R., Hall, J. E., Boutemy, Y. L., and Flaum, C. (1981). A dual porosity CNL logging system. In *56th Annual Fall Technical Conference and Exhibition of the Society of Petroleum Engineers of AIME*, number 10296, San Antonio, TX.
- Deer, W. A., Howie, R. A., and Zussman, J. (1978). *Rock forming minerals*. Halsted Press, 2nd edition. 6 volumes.
- deMenocal, P., Bloemendal, J., and King, J. (1991). A rock magnetic record of monsoonal dust deposition to the arabian sea : evidence for a shift in the mode of deposition at 2.4 Ma. In *Proceeding of the Ocean Drilling Program, Scientific Results*, volume 117, pages 389–407.



- Desbrandes, R. (1968). *Théorie et interprétation des diagraphies*. Technip, Paris.
- Documents du BRGM, editor (1990). *Forages Scientifiques de l'Ardèche, les interactions fluides-roches sur une paléomarge distensive ; rapports d'exécution et données préliminaires sur le forage Balazuc-1, Programme Géologie Profonde de la France*, number 200.
- Doveton, J. (1986). *Log analysis of subsurface geology. Concepts and computer methods*. John Wiley & Sons, New-York.
- Dunlop, D. J. and Özdemir, O. (1997). *Rock magnetism : Fundamentals and frontiers (Cambridge Studies in Magnetism, 3)*. Cambridge University Press.
- Ellis, D., Flaum, C., Roulet, C., Marienbach, E., and Seeman, B. (1983). The Litho-density Tool calibration. In *58th Annual Technical Conference and Exhibition of the Society of Petroleum Engineers of AIME*, number 12048, San Francisco, CA.
- Ellis, D. V. (1982). Correction of NGT logs for the presence of KCl and barite mud. In *Transaction of the SPWLA 33rd Annual Logging Symposium*, number O.
- Ellis, D. V. (1986). Neutron porosity devices—what do they measure? *First Break*, 4(3) :11–17.
- Ellis, D. V. (1987). *Well Logging for Earth Scientists*. Elsevier Science Publishing Company, New York.
- Ellis, D. V. (1988). *SPE Petroleum Production Handbook*. Society of Petroleum Engineers, h. bradley, spe, dallas edition. Chapitre 50 : *Nuclear Logging Techniques*.
- Ellis, D. V., Grau, J., Schweitzer, J., and Hertzog, R. (1987). Basics of nuclear logging. *The Technical Review*, 35(3) :4–10.
- Emeis, K.-C., Robertson, A. H. F., Richter, C., and Shipboard Scientific Party (1996). Sites 965, 966, 967 and 968. In Emeis, K.-C., Robertson, A. H. F., and Richter, C., editors, *Proceeding of the Ocean Drilling Program, Initial Reports*, volume 160.
- Feynman, R. P., Leighton, R. B., and Sands, M. L. (1965). *The Feynman lectures on physics*, volume 2. Addison-Wesley.
- Freedman, R., Boyd, A., Gubelin, G., Morriss, C. E., and Flaum, C. (1997). Measurement of total NMR porosity adds new value to NMR logging. In *Transaction of the SPWLA 38th Annual Logging Symposium*, number 00, Houston, TX.
- Grau, J. A., Antkiw, S., Hertzog, R. C., Manente, R. A., and Schweitzer, J. S. (1984). *In situ* neutron capture spectroscopy of geological formations. In *Proceedings of the 5th International Symposium on Capture Gamma-ray Spectroscopy*.
- Grau, J. A. and Herron, M. M. (1987). Absolute elemental concentrations estimated from geochemical well logging using neutron-induced gamma-ray spectroscopy and a geological model. *Am. Assoc. Pet. Geol. Bull.*, 71 :561–.
- Grau, J. A. and Schweitzer, J. S. (1987). Prompt  $\gamma$ -ray spectral analysis of well data obtained with nai(tl) and 14 mev neutrons. *Nucl. Geophys.*, 1(2) :157–165.
- Grau, J. A. and Schweitzer, J. S. (1989). Elemental concentrations from thermal neutron capture gamma-ray spectra in geological formations. *Nucl. Geophys.*, 3(1) :157–165.

- Grau, J. A., Schweitzer, J. S., Ellis, D. V., and Hertzog, R. C. (1989). A geological model for gamma-ray spectroscopy logging measurements. *Nucl. Geophys.*, 3(4) :351–359.
- Grau, J. A., Schweitzer, J. S., and Hertzog, R. C. (1980). From spectra to elemental yields in 256 easy steps. *The Technical Review*, 28(3) :43–50.
- Grau, J. A., Schweitzer, J. S., and Hertzog, R. C. (1990). Statistical uncertainties of elemental concentrations extracted from neutron-induced gamma-ray measurements. *IEEE Transactions on Nuclear Science*, 37 :2175–2178.
- Grim, R. E. (1968). *Clay mineralogy*. McGraw-Hill, 2nd edition.
- Harvey, P. K., Bristow, J. F., and Lovell, M. A. (1990). Mineral transforms and downhole geochemical measurements. *Scientific Drilling*, 1 :163–176.
- Harvey, P. K. and Lovell, M. A. (1992). Downhole mineralogy logs : mineral inversion methods and the problem of compositional colinearity. In Hurst, A., Griffiths, C. M., and Worthington, P. F., editors, *Geological Applications of Wireline Logs II*, number 65, pages 361–368. Geological Society Special Publication.
- Hassan, M., Hossin, A., and Combaz, A. (1976). Fundamentals of the differential gamma ray log interpretation technique. In *Transaction of the SPWLA 17th Annual Logging Symposium*, number H, Denver, CO.
- Herron, S. L. and Herron, M. M. (1996). Quantitative lithology : an application for open and cased hole spectroscopy. In *Transactions of the SPWLA 37th Annual Logging Symposium*, number E, New Orleans, LA.
- Hertzog, R. (1978). Laboratory and field evaluation of an inelastic–neutron–scattering and capture gamma-ray spectroscopy tool. In *53rd Annual Fall Technical Conference and Exhibition of the Society of Petroleum Engineers*, number 7430, Houston, TX.
- Hertzog, R., Colson, L., Seeman, B., O'Brien, M., Scott, H., McKeon, D., Wraight, P., Grau, J., Ellis, D., Schweitzer, J., and Herron, M. (1989). Geochemical logging with spectrometry tools. *SPE Formation Evaluation*, pages 153–162.
- Hertzog, R. and Plasek, R. (1979). Neutron-excited gamma ray spectrometry for well-logging. *IEEE Transactions on Nuclear Science*, NS-26.
- Jodouin, J.-F. (1994a). *Les réseaux de neurones*. Hermès, Paris.
- Jodouin, J.-F. (1994b). *Les réseaux neuromimétiques*. Hermès, Paris.
- Kelsey, C. H. (1965). Calculation of the C.I.P.W. norm. *Mineralogical Magazine*, (34) :276–282.
- Kempler, D. (1993). *Tectonic patterns in the Eastern Mediterranean*. PhD thesis, Hebrew University of Jerusalem.
- Kenyon, W. E., Kleinberg, R., Straley, C., Gubelin, G., and Morriss, C. (1995). Nuclear magnetic resonance imaging–technology for the 21st century. *OilField Review*, 7(3).
- Kohonen, T. (1990). The self-organizing map. In *Proceedings of the IEEE*, volume 78, pages 1464–1480.

- Kohonen, T. (1995). *Self-Organizing Maps*. Springer.
- Kröse, B. and van der Smagt, P. (1996). An introduction to neural networks. Copyright University of Amsterdam. Cet ouvrage est disponible sur Internet à l'adresse suivante <http://www.robotic.dlr.de/Smagt/books>.
- Leverett, M. C. (1938). Flow of oil-water mixtures through unconsolidated sands. In *Transactions of AIME*, volume 132.
- Lim, M. (1991). *Diagraphies magnétiques : conséquences des mesures de laboratoire pour l'interprétation*. PhD thesis, Université Paris VI.
- Limonov, A. and Woodside, J. (1994). 4.2. Results 4.2.1a. Seismic reflection profiling. In Limonov, A. F. and Woodside, J. M., editors, *Mud volcanism in the Mediterranean and Black Seas and shallow structure of the Eratosthenes Seamount. Initial Results of the geological and geophysical investigations during the Third UNESCO-ESF "Training-through-Research Cruise of R.V. Gelendzhik (June-July 1993). UNESCO Reports in Marine Science*, volume 64, pages 120–127.
- Lippmann, R. P. (1987). An introduction to computing with neural nets. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 2(4) :4–22.
- Lofts, J. C. (1993). *Integrated geochemical-geophysical studies of sedimentary reservoir rocks*. PhD thesis, University of Leicester, UK.
- Lucchi, R., Pott, M., Ivanova, E., Lottotskaya, A., Glover, C., Flecker, R., Basov, E., Kidd, R., and Robertson, A. H. F. (1994). 4.2d. Bottom sampling. In Limonov, A. F. and Woodside, J. M., editors, *Mud volcanism in the Mediterranean and Black Seas and shallow structure of the Eratosthenes Seamount. Initial Results of the geological and geophysical investigations during the Third UNESCO-ESF "Training-through-Research Cruise of R.V. Gelendzhik (June-July 1993). UNESCO Reports in Marine Science*, volume 64, pages 134–135.
- McQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *5th Berkeley Symposium on Mathematics, Statistics, and Probability*, volume 1, pages 281–298.
- Minster, J.-F., Richard, L. P., and Allègre, C. J. (1982).  $^{87}\text{Rb}/^{86}\text{Sr}$  chronology of enstatite meteorites. *Earth Planet. Sci. Lett.*, (44) :420–440.
- Pettijohn, F. J. (1975). *Sedimentary rocks*. Harper and Row, 3rd edition.
- Pozzi, J.-P., Barthès, V., Thibal, J., Pocachard, J., Lim, M., Thomas, T., and Pagès, G. (1993). Downhole magnetostratigraphy in sediments : comparison with the paleomagnetism of a core. *Journal of Geophysical Research*, 98(B5) :7939–7957.
- Pozzi, J.-P., Martin, J.-P., Pocachard, J., Feinberg, H., and Galdeano, A. (1988). In-situ magnetostratigraphy : interpretation of magnetic logging in sediments. *Earth Planet. Sci. Lett.*, 88 :357–373.
- Robertson, A. H. F. (1990). Tectonic evolution of Cyprus. In Malpas, J., Moores, E. M., Panayiotou, A., and Xenophontos, C., editors, *Ophiolites and Oceanic Lithosphere, Proc. Internat. Symp. "Troodos 1987"*, pages 235–252, Nicosia, Cyprus. Geol. Surv. Dept.

- Robertson, A. H. F., Woodside, J., and Kidd, R. (1994a). 4. Eratosthenes Seamount (Study area 3), 4.1. Geological setting. In Limonov, A. F. and Woodside, J. M., editors, *Mud volcanism in the Mediterranean and Black Seas and shallow structure of the Eratosthenes Seamount. Initial Results of the geological and geophysical investigations during the Third UNESCO-ESF "Training-through-Research Cruise of R.V. Gelendzhik (June-July 1993). UNESCO Reports in Marine Science*, volume 64, pages 115–119.
- Robertson, A. H. F., Woodside, J., Limonov, A., and Kidd, R. (1994b). 4.3 General interpretation. In Limonov, A. F. and Woodside, J. M., editors, *Mud volcanism in the Mediterranean and Black Seas and shallow structure of the Eratosthenes Seamount. Initial Results of the geological and geophysical investigations during the Third UNESCO-ESF "Training-through-Research Cruise of R.V. Gelendzhik (June-July 1993). UNESCO Reports in Marine Science*, volume 64, pages 148–158.
- Rogers, S. J., Fang, J. H., Karr, C. L., and Stanley, D. A. (1992). Determination of lithology from well logs using a neural network. *Am. Assoc. Pet. Geol. Bull.*, 76 :731–739.
- Ross, D. A. and Uchupi, E. (1977). Structure and sedimentary history of southeastern Mediterranean sea—Nile cone. *AAPG Bull.*, (61) :872–902.
- Ryan, W. B. F. and Hsü, K. J., editors (1973). *Initial Reports, DSDP*, volume 13. U.S. Government Printing Office.
- Sammon, Jr., J. W. (1969). A nonlinear mapping for data structure analysis. *IEEE transactions on computers*, C-18(5) :401–409.
- Scheerens, S. (1995). *Reconstitution de la lithologie et de la minéralogie à partir de mesures réalisées à l'aide de l'outil diagaphique GLT (Geochemical Logging Tool). Application au forage scientifique de Balazuc (Ardèche, Programme Géologie Profonde de la France)*. PhD thesis, Université Paris-Sud, Orsay.
- Scheerens, S. and Steinberg, M. (1997). Mineralogy from geochemical data, a case study : Balazuc-1 borehole (Ardèche, Programme Géologie Profonde de la France). *Bull. Soc. Géol. France*, 168(2) :207–217.
- Schweitzer, J. S., Ellis, D. V., Grau, J. A., and Hertzog, R. C. (1988). Elemental concentrations from gamma-ray spectroscopy tools. *Nucl. Geophys.*, 2(3) :175–181.
- Schweitzer, J. S., Grau, J. A., and Hertzog, R. C. (1987). Precision and accuracy of short-lived activation measurements for *in situ* geological analyses. In *2nd Int. Workshop on Activation Analysis With Short-Lived Nuclides*, Vienna.
- Schweitzer, J. S., Manente, R. A., and Hertzog, R. C. (1982). Gamma-ray spectroscopy tool environmental effects. In *57th Annual Fall Technical Conference and Exhibition of the Society of Petroleum Engineers of AIME*, number 11144, New Orleans, LA.
- Scott, H. D., Flaum, C., and Sherman, H. (1982). Dual porosity CNL count rate processing. In *57th Annual Fall Technical Conference and Exhibition of the Society of Petroleum Engineers of AIME*, number 11146, New Orleans, LA.
- Scott, H. D. and Smith, M. P. (1973). The Aluminium Activation Log. In *Transactions of the SPWLA 14th Annual Logging Symposium*, Lafayette, LA.

- Sen, P. N., Straley, C., Kenyon, W. E., and Whittingham, M. S. (1990). Surface-to-volume ratio, charge density, nuclear magnetic relaxation and permeability in clay-bearing sandstones. *Geophysics*, 55(1) :61–69.
- Serra, O. (1984). *The acquisition of logging data*, chapter 1, pages 1–26. Elsevier.
- Serra, O., Baldwin, J., and Quirein, J. (1980). Theory, interpretation and practical applications of natural gamma-ray spectroscopy. In *Transactions of the SPWLA 21st Annual Logging Symposium*, Lafayette, LA.
- Steininger, F. F. and Rogl, F. (1984). Paleogeography and palinspastic reconstruction of the Neogene of the Mediterranean and Paratethys. In Dixon, J. E. and Robertson, A. H. F., editors, *The Geological Evolution of the Eastern Mediterranean, Spec. Publ. Geol. Soc. Lond.*, volume 17, pages 659–668. Blackwell Scientific Publ., Oxford.
- Tarantola, A. and Valette, B. (1982). Generalized non-linear inverse problem solved by using the least square criterion. *Rev. Geophysics and Space Physics*, (20) :219–232.
- Theys, P. P. (1991). *Log data acquisition and quality control*. Editions Technip, Paris.
- Theys, P. P. (1994). A serious look at repeat sections. In *Transaction of the SPWLA 35th Annual Logging Symposium*, number F.
- Thibaut, J. (1995). *Analyse de l'aimantation des sédiments par diagraphies magnétiques. Magnétostratigraphie, cyclicités climatiques et intensité du champ géomagnétique*. PhD thesis, Université Paris-Sud, Orsay.
- Thibaut, J., Pozzi, J.-P., Barthès, V., and Dubuisson, G. (1995). Continuous record of geomagnetic field intensity between 4.7 and 2.7 Ma from downhole measurements. *Earth Planet. Sci. Lett.*, 136 :541–550.
- Thiria, S., Meija, C., Badran, F., and Crepon, M. (1997). *Statistiques et méthodes neuronales*, chapter 5, pages 75–89. Dunod.
- Thouveny, N., de Beaulieu, J.-L., Bonifay, E., Creer, K. M., Gulot, J., Icole, M., Johnsen, S., Jouzel, J., Reille, M., Williams, T., and Williamson, D. (1994). Climate variations in Europe over the past 140 kyr deduced from rock magnetism. *Nature*, 371 :503–506.
- van den Oord, R. J. (1991). Evaluation of geochemical logging. *The Log Analyst*, 32(1) :1–12.
- Wahl, J. S. (1983). Gamma-ray logging. *Geophysics*, 48(11) :1536–1550.
- Westaway, P., Hertzog, R., and Plasek, R. E. (1980). The gamma-spectrometer tool inelastic and capture gamma-ray spectroscopy for reservoir analysis. In *55th Annual Fall Technical Conference and Exhibition of the Society of Petroleum Engineers of AIME*, number 9461, Dallas, TX.
- Woodside, J. M. (1972). Disruption of the African plate margin in the Eastern Mediterranean. In Salem, M. J., editor, *The Geology of Libya*, volume 6, pages 2319–2329. Elsevier.
- Woodside, J. M. (1977). Tectonic elements and crust of the eastern Mediterranean Sea. *Marine Geophys. Res.*, pages 317–354.
- York, D. (1969). Least squares fitting of a straight line with correlated errors. *Earth Planet. Sci. Lett.*, (5) :320–324.