



Fonds National
des Thèses

Reproduction & Diffusion interdite

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITE M'HAMED BOUGARA BOUMERDES
Faculté des Hydrocarbures et de la Chimie
Laboratoire de Physique de la Terre



MEMOIRE

Présenté par

MIHOUBI ABDELHAFID

En vue de l'obtention du diplôme de

MAGISTER

Spécialité : **TRAITEMENT DU SIGNAL**
Option: **GÉOPHYSIQUE et TRAITEMENT DU SIGNAL**

Classification Lithologique des Attributs Sismiques
par les Réseaux de Neurones Artificiels
Comparaison entre
les Cartes Auto-organisatrices de Kohonen et le Perceptron
Multicouches

Date de soutenance 26/04/2008

Mrs. K. BADDARI	(Professeur, UMBB)	Président
M. DJEDDI	(Professeur, UMBB)	Rapporteur
A. BOUGUERNE	(Maître de conférences, UMBB)	Examineur
M.A. AITOUCHE	(Chargé de cours, UMBB)	Examineur
Mme Z. BENAÏSSA	(Maître de conférences, USTHB)	Examineur

Année universitaire 2007/2008

REMERCIEMENTS

Je tiens à exprimer ma gratitude au professeur DJEDDI M, d'avoir accepté ce thème comme sujet de magister et d'en avoir supervisé la concrétisation.

Mes remerciements s'adressent aussi aux membres du jury, pour leur disponibilité à évaluer le présent mémoire et leur participation au jury de soutenance.

Mes reconnaissances et mes plus vifs remerciements s'adressent au Docteur AITOUCHE MA, chargé de cours à l'Université de Boumerdes, pour avoir accepté la lecture et la correction de ce mémoire.

Je remercie également Monsieur KOUADRI A, enseignant à l'Université de Boumerdes, pour ses remarques et ses corrections relatives au développement théorique des réseaux de neurones artificiels.

Je remercie chaleureusement le Docteur TANER M T, chef géophysicien chercheur au Rock Solid Images USA, pour son aide en mettant à ma disposition ses connaissances et ses documents.

Je remercie Monsieur BELAHCEN A, Ingénieur géophysicien à la Sonatrach Division Exploration direction études et synthèse, pour son aide dans la partie interprétation et préparation des données.

Je tiens à remercier Monsieur BOUGUERN A, Maître de conférences, chef de département Géophysique à l'Université de Boumerdes, pour la confiance et le soutien qu'il m'a accordés.

Je voudrais également remercier Monsieur DJARFOUR, enseignant à l'Université de Boumerdes, pour son aide, sa disponibilité et son soutien moral.

En fin j'exprime ma gratitude à tous ceux qui ont contribué de près ou de loin à l'élaboration de ce mémoire.

Abstract

Lithological classification of Seismic attributes using artificial neural networks

The aim of this work is to introduce the artificial neural networks (ANN) as classification tools in petroleum exploration. In this work we try to exploit the capacities of artificial neural networks training to realize a lithological classification of the seismic attributes derived from seismic data. The present report is composed of five chapters. In the part annexes we developed the different programmers of amplitude corrections in preserved amplitudes processing sequence.

We first introduce in chapter one seismic attributes definitions and determination. The characterization of reservoirs requires integration of data and knowledge from various types. The relationships between seismic attributes and lithofacies may provide important information for understanding subsurface. Seismic attributes offer a good input data set for lithological classification. Classification is the division of a given set of objects into different groups based on their discriminating features. Discriminating features are those that maximize the differences between each group. In seismic event classification the discriminating features are the seismic attributes.

The second chapter refers to elementary definitions and general informations of the artificial neural networks. Artificial neural networks can be used effectively to identify lithology from seismic data set. A neural network consists of an interconnection of a number of neurons. This operation is made up of two principal stages, training and classification. During the phase of training, the network is formed to draw up relationship between seismic attributes and lithology of the geological formations. Only the seismic data around wells (seismic attributes) are employed to learn the network. Then, all the seismic dataset will be classified by this network (qualified by training). In cases where well logs are available, Artificial Neural Networks (ANN) may be trained in a supervised manner. In cases where no well information is available, we use an unsupervised approach to classify the data.

The third chapter describes the first technique of ANN applied in our project. Kohonen's self-organizing maps (K-SOM) are simple analogs of the human brain's ability to organize information in a logical manner. A self-organizing maps tries to establish rules for the characterization of the physical problem. Kohonen's self-organizing maps consist of a single layer of neurons organized in one, two or multi-dimensional arrays. Each neuron has as many input connections as there are attributes to be used in the classification. The training procedure consists of finding the neuron with weights that are closest to the input data vector and declaring that neuron as the winning neuron. Then the weights of all of the neurons in the vicinity of the winning neuron are adjusted by an amount inversely proportional to the distance. The radius of the accepted vicinity is reduced as the iterations increase. The training process is terminated if the errors of all of the inputs are reduced to an acceptable level or a prescribed number of iterations is reached.

The second technique of ANN applied is Feed forward Fully Connected Multi-Layer Perceptrons (MLP). the chapter four present this method. the Multi-Layer Perceptrons are one of the first successfully utilized artificial neural networks. I use only one type of network proposed by Rumelhart and others and used by Taner in seismic classification, which is called the "Semilinear Feed-Forward". In this network the data flows forward to the output continuously without any feed-back. It is also called the Perceptron. It consists of input section, where input data is weighted and summed. A bias is added to the sum and input to the activation function.

These weights and the bias represent the knowledge. Result from activation function represents the output of this neuron.

As an application for methods described above and in order to present the performance of these techniques, real data are used. It is the contents of the final chapter. This application is a study of detail in an area situated in the South-west of Algeria. This region is characterized by gas production since almost the totality of wells highlighted an accumulation of gas.

The information obtained by the two applications is of great interest compared to the rough seismic section. The number of classes was not defined before starting the classification by K-SOM method and the result has five colors which indicate five lithological classes. This result is confirmed by the result of the method MLP.

An interesting possibility might be to use the output of self-organizing maps (K-SOM) as the input for multi-layers perceptron (MLP).

ملخص البحث

التصنيف الليتولوجي للمعطيات السيسمية بواسطة الشبكة العصبية الاصطناعية

موضوع بحثنا يتناول تطبيقات الشبكة العصبية الاصطناعية (Réseaux de neurones artificiels) في ميدان الاستكشاف البترولي ويخص بالذات استعمال هذه الشبكة في التصنيف الليتولوجي (Classification lithologique) للمقاطع السيسمية (Sections sismiques) ومستخلصاتها (Attributs sismiques).

البحث يتكون من خمس محاور كبرى إضافة إلى ملحق يشمل بعض التفاصيل الدقيقة و التي لم يتسن لنا وضعها داخل المحاور تجنباً للتفرع و الخروج عن لب الموضوع رغم علاقتها بالبحث.

المحور الأول يتضمن التعاريف الخاصة والأسس النظرية للإشارة السيسمية (Signal sismique) ومستخلصاتها.

المحور الثاني خصصناه للشبكة العصبية الاصطناعية، تعريفها، عناصرها، أنواعها، أسسها النظرية و استعمالاتها.

المحور الثالث قدمنا فيه خريطة كوهونان الذاتية التنظيم (les cartes auto-organisatrices de Kohonen)

(K-SOM) و التي تمثل النوع الأول من الشبكة العصبية الاصطناعية المطبق في بحثنا هذا تضمن هذا المحور تعريف مفصل لهذه التقنية و استعمالاتها و عرضنا خوارزمية استعمالها في ميدان التصنيف الليتولوجي للمعطيات السيسمية.

المحور الرابع عرضنا فيه تقنية البارسابترون المتعدد الطبقات (Perceptron multicouches, PMC) و شملنا

بعض التفاصيل مع وضع خوارزمية استعمالها في الترتيب والتصنيف الليتولوجي.

المحور الخامس جعلناه تطبيقاً عملياً لكلتا التقنيتين المتقدمتين في المحور الثالث والرابع على معطيات حقيقية وتم

مقارنة النتائج المحصل عليها من كلتيهما. التقنيتان (K-SOM) و (PMC) تختلفان عن بعضيهما البعض اختلافاً جوهرياً من

حيث طريقة التلقين (Apprentissage) والتدريب لكل نوع، فالأول يمتاز بالتلقين الذاتي من دون توجيه

(Apprentissage non supervisé) أي دون إدخال أي معلومة خارجية عن المعطيات المقدمة للتصنيف. والثاني يتطلب

تلقين موجه (Apprentissage supervisé) وهنا نستعمل معطيات خارجية عن المادة المقدمة. في هذه الحالة المادة المقدمة

للتصنيف هي معطيات سيسمية ومستخلصاتها، والمادة المستعملة في تلقين الشبكة العصبية الاصطناعية هي عبارة عن قياسات بئرية لبعض الخصائص الفيزيائية للصخور المكونة للطبقات الجيولوجية.

عموماً النتائج التي توصلنا إليها تبين انه لكل تقنية خصوصية ايجابية. فالتقنية الذاتية التلقين من دون توجيه توفر لنا

حلاً في غياب المقاييس البئرية والتقنية الموجهة التلقين تعطينا تصنيفاً دقيقاً دون الابتعاد عن النموذج الجيولوجي الخاص

بالمنطقة المستكشفة. كما أننا توصلنا إلى انه يمكننا استعمال نتائج التقنية الذاتية التلقين كمدخل للتقنية الموجهة التلقين.

LISTE DES FIGURES

- Figure : I.3.1. Porosité et Perméabilité
Figure : I.3.2. Mode de propagation des ondes P et S.
Figure : I.4.1.a. Trace sismique réelle.
Figure : I.4.1.b. Réponse fréquentielle du filtre de quadrature.
Figure : I.4.1.c. Trace en quadrature de trace sismique réelle.
Figure : I.4.3.a. Représentation de la trace complexe.
Figure : I.4.3.b. Schéma illustratif d'une trace complexe.
Figure : I.6.2. Principe de calcul des attributs de l'ondelette
Figure : II.3.1.a. Principe de fonctionnement du système nerveux.
Figure : II.3.1.b. Schéma synoptique simplifié d'un neurone.
Figure : II.4.1.1. Modèle d'un neurone formel.
Figure : II.4.1.2. Représentation graphique d'un neurone.
Figure : II.4.1.3. Comparaison entre le neurone biologique et le neurone artificiel.
Figure : II.4.2.1.a. Exemple d'un réseau à connexions locales.
Figure : II.4.2.1.b. Exemple d'un réseau à connexions totales.
Figure : II.4.2.3.a. Exemple d'un Réseau à une couche.
Figure : II.4.2.3.b. Exemple d'un réseau multicouches ``feedforward``.
Figure : II.4.2.3.c. Exemple d'un réseau récurrent.
Figure : II.4.2.3.d. Réseaux en treillis.
Figure : II.5.1. Architecture du Perceptron.
Figure : II.5.2. Architecture du Perceptron Multicouche.
Figure : II.5.3. Architecture d'un modèle de Rétropropagation.
Figure : II.5.4. Architecture d'un modèle de Hopfield.
Figure : II.5.5. Architecture d'un modèle de Kohonen.
Figure : III.4. Une carte auto-organisatrice (réseau de Kohonen) à 2-dimensions et une couche d'entrée à 5 composantes.
Figure : III.5.a. Connexion des données d'entrée au nœud neural.
Figure : III.5.b. Distance Euclidienne séparant deux vecteurs.
Figure : III.6.a. Forme Topologique Unidimensionnelle.
Figure : III.6.b. Forme Topologique Bidimensionnelle.
Figure : III.9. Organigramme du Programme S.O.M Appliqué.
Figure : IV.2.a. Représentation graphique d'un Perceptron.
Figure : IV.2.b. Fonction d'activation de type sigmoïde.
Figure : IV.4. Réseaux de neurones semi linéaire
Figure : IV. 5. Pointé des premières arrivées.
Figure : IV. 6. a. Adaptation des Couches de sortie.
Figure : IV. 6. b. Adaptation des Couches Cachées.
Figure : V.4. Document Sonatrach/Exploration.
Figure : V.5.2.a. Séquence de Traitement
Figure : V.5.2.b. Contrôle de qualité des différentes étapes de traitement.
Planche : V.5.2.c. Section sismique avec Migration et Filtrages
Figure : V. 6. Analyse des attributs Sismiques avant classification.
Figure : V.7.1.1. Paramètres du réseau (PMC).
Figure : V.7.1.2. Cross Plot, outil de choix des attributs sismiques.
Figure : V.7.1.3.a. Définition des classes pour l'apprentissage.
Figure : V.7.1.3.b. Types de classes d'amplitudes définies pour l'apprentissage.
Tableau V.7.1.4. Statistique et QC de l'analyse par PCM.
Figure V.7.1.5. Coupe lithologique après analyse par PCM.

Figure V.7.2.1. Définition des paramètres pour la classification par K-SOM.
Figure V.7.2.2. Choix des attributs à utiliser par K-SOM.
Figure V.7.2.3.a. Evolution du réseau K-SOM.
Tableau V.7.2.3.b. Statistiques et QC de l'analyse par K-SOM.
Figure V.7.2.3.c. Coupe lithologique selon K-SOM.
Figure : V.7.3.a. Comparaison entre les résultats du PMC et la K-SOM.
Planche : V.7.3.b. Classification Lithologique par le PMC, Résultat après
l'itération N°100
Planche : V.7.3.c. Classification Lithologique par le PMC, Résultat après
l'itération N°15000
Planche : V.7.3.d. Classification Lithologique par K-SOM, Résultat après
l'itération N°2600

NOTATIONS ET ABREVIATIONS

NMO : Normal moveout (correction dynamique)
CS : Correction statique
S/B : Rapport signal sur bruit
Stack : Sommation des traces sismiques
RNA : Réseaux de neurones artificiels
ANN : Artificial neural networks
PMC : Perceptron multicouches
MLP : Multi layers perceptron
K-SOM : Kohonen Self Organizing Map
AVO : Amplitude versus Offset
TH : Transformation de Hilbert
TF : Transformation de Fourier
CDP : Common depth point
CMP : Common mid point
Thb : Thin bed (Couche mince)
 F_i : Fréquence Instantanée
Dip : Pendage
Sembl(t) : La fonction de semblance
Net : La somme pondérée des entrées (l'entrée d'un neurone)
RMS : Erreurs quadratiques moyennes,
 W_i : Well,(puits)
 C_i : Classe lithologique
QC : Quality Control (contrôle de qualité)

SOMMAIRE

INTRODUCTION GENERALE

Chapitre I. SIGNAL ANALYTIQUE ET ATTRIBUTS SISMIQUES

Aspect Théorique et Développement Mathématique

I.1. Historique	1
I.2. Introduction	
I.3. Elasticité et Ondes élastiques	
I.3.1. Paramètres d'élasticité	
I.3.2. Ondes Elastiques	
I.4. Théorie du signal analytique	
I.4.1. Filtre de quadrature	
I.4.2. Transformée de Hilbert directe et inverse	
I.4.3. Signal analytique	
I.4.4. Procédures du calcul du signal analytique	
I.5. Classification des attributs sismiques	
I.6. Calcul des attributs de base	
I.6.1. Attributs instantanés	
I.6.2. Attributs de l'ondelette	
I.7. Calcul des attributs sismiques en tenant compte de l'information spatiale	
I.7.2. Calcul des attributs géométriques	
I.7.3. Calcul des attributs Hybrides	

Chapitre II. DES RESEAUX DE NEURONES ARTIFICIELS

Bases Théoriques

II.1. Historique	23
II.2. Introduction	
II.3. Eléments de neurophysiologie	
II.3.1. Le neurone	
II.3.2. Caractéristiques principales du neurone physiologique	
II.3.3. Circuits neuronaux	
II.3.4. Activation d'un neurone	
II.4. Les neurones Artificiels	
II.4.1. Le neurone formel	
II.4.1.1. Définition	
II.4.1.2. Modélisation d'un neurone	
II.4.1.3. Comparaison entre le Neurone biologique et neurone artificiel	
II.4.1.4. Fonction d'activation	
II.4.1.5. Types des Fonctions d'activation	

- II.4.2. Le réseau de neurones artificiels
 - II.4.2.1. Connectivité
 - II.4.2.2. Apprentissage
 - II.4.2.2.a. Types d'apprentissage
 - II.4.2.2.b. Application des réseaux de neurones selon le mode d'apprentissage
 - II.4.2.3. Topologie
- II.5. Les types de réseaux
 - II.5.1. Le Perceptron
 - II.5.2. Perceptron multicouches
 - II.5.3. Perceptron multicouches avec l'algorithme de Retro-propagation
 - II.5.4. Réseaux Hopfield
 - II.5.5. Réseau de Kohonen
- II.6. Mise en œuvre et Application des réseaux de neurones artificiels

Chapitre III. LES CARTES AUTO-ORGANISATRICES DE KOHONEN (K-SOM)

Outil de Classification à Apprentissage non Supervisé

- III.1. Introduction 39
- III.2. Principe des cartes auto-organisatrices
- III.3. Classification et cartes auto-organisatrices
- III.4. Architecture Neurale des cartes auto-organisatrices
- III.5. Méthodologie de la technique S.O.M
- III.6. Différentes organisations topologiques
- III.7. L'application des cartes auto-organisatrices dans l'interprétation sismique
- III.8. Procédé de calcul des S.O.M
- III.9. Organigramme Du Programme S.O.M Appliqué
- III.10. Conclusion

Chapitre IV. Perceptron Multicouches (PMC)

Technique de Classification et Apprentissage Supervisé

- IV.1. Introduction 50
- IV.2. Neurones
- IV.3. Discrimination linéaire par un neurone
- IV.4. Réseau Neurologique
- IV.5. Exemple d'application des réseaux de neurones
- IV.5.1. Choix des données pour l'apprentissage
- IV.5.2. Choix des paramètres d'apprentissage
- IV.5.3. Exécution du processus d'apprentissage et contrôle des résultats

IV.6. Algorithme de retropropagation du gradient de l'erreur et règle d'adaptation des poids	
IV.6.1. Règle d'adaptation des poids et des seuils de la couche de sortie	
IV.6.2. Règle d'adaptation des poids et des seuils des couches cachées	
IV.6.3. Stratégies des adaptations des poids et des seuils	
IV.7. Conditions d'une bonne exécution	
IV.7.1. Configuration du réseau	
IV.7.2. Convergence de l'algorithme	
IV.7.3. Choix d'entrée pour l'apprentissage	
IV.8. Critère d'un choix optimal de la configuration du réseau	
IV.9. Critères d'arrêt de l'apprentissage	
IV.10. Séquence de la phase d'apprentissage	
IV.11. Conclusion	

Chapitre V. APPLICATION PMC et K-SOM, Outils de classification lithologique

V.1. L'objet de l'application.....	64
V.2. Données Géologiques	
V.2.1. Roches mères	
V.2.2. Roches couvertures	
V.2.3. Réservoirs	
V.3. Analyses des résultats des trois puits	
V.3.1. GF-1	
V.3.2. GF-2	
V.3.3. GF-3	
V.4. Description structurale de la Structure 'Garet El Guefoul'	
V.5. Traitement des données sismiques	
V.5.1. Paramètres d'acquisition	
V.5.2. Séquence de traitement en amplitude Préservée et Analyse des résultats	
V.6. Calcul des attributs sismiques	
V.7. Classification Lithologique	
V.7.1. Classification par les Réseaux de Neurones Artificiel (ANN)	
V.7.1.1. Choix du réseau	
V.7.1.2. Choix des attributs pour l'apprentissage	
V.7.1.3. Définition des classes	
V.7.1.4. Critères d'arrêt de l'apprentissage	
V.7.1.5. Analyse des résultats	

V.7.2. Classification par les Cartes Auto Organisatrice de Kohonen
(K-SOM)

V.7.2.1. Définition des Paramètres

V.7.2.2. Choix des attributs (Kohonen Attributes Planes)

V.7.2.3. Contrôle de processus d'apprentissage et analyse
des résultats

V.7.3. Comparaison ANN-SOM

CONCLUSION ET RECOMMANDATIONS

ANEXES

INTRODUCTION GENERALE

Le progrès et l'amélioration des techniques d'acquisition de traitement et d'interprétation en prospection sismique, permettent de fournir aujourd'hui une image, aussi proche que possible de l'image réelle de la subsurface.

Les événements sur une section sismique apparaissent souvent consistants d'une trace à une autre. Généralement ils sont utilisés pour construire des modèles géophysiques et géologiques. Les explorateurs ne s'intéressent pas uniquement aux amplitudes des réflexions mais aussi à l'évaluation de la morphologie et au caractère des réflexions (la phase de l'ondelette).

L'amplitude d'une réflexion sismique a longtemps été la moins prise en compte par la majorité des programmes de traitement. Ces derniers s'attachent en général à améliorer le rapport signal sur bruit « S/B » (filtrage fréquentiel, atténuation des bruits, stack...) et à repositionner les événements d'une section en temps et en espace (Corrections statiques, corrections dynamiques « NMO », Migration).

Ces dernières années ont vu l'importance du paramètre amplitude s'accroître sensiblement, notamment en raison de l'émergence de la sismique stratigraphique et des méthodes d'inversion qui essaient de relier quantitativement les données sismiques avec les paramètres d'élasticité et pétrophysiques des milieux de propagation.

L'étude et l'interprétation des attributs sismiques donnent des informations qualitatives sur les paramètres physiques et sur la configuration géométrique de la subsurface. Il est à noter que le contenu en amplitudes des données sismiques est le facteur principal pour la détermination des paramètres physiques, comme l'impédance acoustique, les coefficients de réflexion, les vitesses, l'absorption, etc....

Les attributs calculés à partir de la trace complexe, telle que l'enveloppe, la phase instantanée..., correspondent aux diverses mesures effectuées sur le mode de propagation du front d'ondes. On les appellera Attributs physiques (*Physical Attributes*). Ceux calculés à partir de la configuration et de la continuité des réflexions, seront appelés Attributs géométriques (*Geometrical Attributes*).

Les attributs peuvent être calculés avant sommation ou après sommation des données, avant ou après migration. Ils peuvent être classés de différentes manières. Plusieurs auteurs ont donné leur propre classification. Ici nous illustrons la classification de Taner. C'est une classification basée sur les caractéristiques des attributs. Taner avait établi une classification générale des attributs basée sur leurs données d'entrée et leur utilisation.

Beaucoup d'anomalies d'amplitudes (par exemple le Bright spot) sont identifiées et conduisent souvent à des résultats positifs après forage des structures. Cette application a constitué la première utilisation des attributs sismiques à savoir l'enveloppe, la fréquence instantanée et la polarité apparente. La recherche dans cet axe a permis de développer de nouveaux attributs. Durant les dernières années, les explorateurs ne se contentaient pas uniquement de l'exploitation des attributs séparément mais aussi de leur combinaison. La combinaison des attributs avec d'autres techniques comme l'analyse AVO et l'inversion sismique a joué un rôle très important dans la caractérisation des réservoirs.

La caractérisation des réservoirs exige l'intégration des données de divers types et de diverses propriétés physiques. Parmi ces derniers, la distribution des litho-faciès est d'une importance fondamentale. Un problème important dans la prévision de la distribution spatiale de cette propriété physique est que les données de puits sont souvent limitées au voisinage de ce dernier.

Afin de surmonter ce problème, les données sismiques au voisinage du puits sont parfois utilisées pour exploiter les rapports entre la réponse sismique et les propriétés physiques tirées des données de puits. Il est alors difficile de prédire ces rapports avec des

méthodes simples parce qu'ils sont compliqués et non linéaires. Parmi les solutions proposées récemment, l'utilisation des réseaux de neurones artificiels. En se basant sur le caractère et la forme de l'ondelette sismique on peut construire une image lithologique par le biais des réseaux de neurones artificiels.

Dans ce travail nous avons tenté d'exploiter les capacités d'apprentissage des réseaux de neurones artificiels pour faire une classification lithologique des attributs sismiques. Cette opération est composée de deux étapes principales, l'apprentissage et la classification. Pendant la phase d'apprentissage, le réseau est formé pour établir des rapports entre les attributs sismiques et la lithologie des formations géologiques. Seules, les données sismiques (attributs) dans les zones au voisinage des puits sont employées dans la phase d'apprentissage. Ensuite, toutes les données sismiques de la région seront classifiées par ce réseau (après apprentissage). La sortie est une coupe 2D (ou un cube 3D) formée des différentes classes lithologiques appropriées.

L'efficacité de l'utilisation des " attributs sismiques " dans l'interprétation dépend des capacités de discrimination de l'ensemble des attributs choisis. Comme dans beaucoup de problèmes mathématiques, quelques attributs peuvent être suffisants. Ceci peut être déterminé par l'expérimentation avec des combinaisons logiques de divers types d'attributs. Dans le cas où les données de puits ou les colonnes stratigraphiques sont disponibles, des réseaux de neurones artificiels (RNA) peuvent être formés d'une façon supervisée. Le perceptron multicouches (PMC) est un exemple. Dans les cas où aucune information n'est disponible, nous employons l'approche non supervisée pour classifier les données.

Les cartes auto-organisatrices de Kohonen (K-SOM) est un exemple des méthodes à apprentissage non supervisé qui analyse les données sismiques et extrait l'information à partir de ces dernières à l'aide d'un algorithme de classification adapté. C'est une nouvelle technique qui utilise uniquement les données sismiques ou des attributs dérivés, et n'exige pas une estimation a priori sur les données d'entrée.

Cette étude s'inscrit dans le contexte de la caractérisation de réservoir. Généralement ce domaine est lié directement aux données de puits mais il n'est pas facile de disposer du nombre nécessaire de puits. Cette situation a poussé les géophysiciens à explorer de nouvelles méthodes pour remédier à ce problème. L'exploitation de l'amplitude sismique et l'extraction des informations lithologiques directement de cette amplitude ont fait l'objet de la plus grande partie des recherches dans le domaine de la caractérisation de réservoir. Cette étude repose sur l'exploitation directe des attributs extraits de l'amplitude sismique et leurs combinaisons en utilisant les réseaux de neurones artificiels.

Le premier but du présent travail est de présenter les attributs sismiques sous leurs différentes formes avec une nouvelle vision. Le premier chapitre est consacré au développement des attributs avec une nouvelle classification. L'aspect de la caractérisation de réservoir dans notre travail se situe dans la combinaison des attributs pour effectuer une analyse lithologique. La méthode utilisée pour faire cette analyse est les réseaux de neurones artificiels. Le développement de deux techniques à savoir le perceptron multicouches (PMC) et les cartes auto-organisatrices de Kohonen (K-SOM) et la comparaison entre ces deux méthodes forment le deuxième but de ce travail. Les cartes auto-organisatrices de Kohonen (Kohonen Self Organizing Map « K-SOM ») et le perceptron « PMC » multicouche font l'objet des chapitres trois et quatre. Le cinquième chapitre sera consacré à l'application sur des données réelles d'une analyse lithologique par la combinaison des attributs sismiques en utilisant les réseaux de neurones artificiels.

I. Signal analytique et attributs sismiques, Aspect théorique et développement Mathématique

I.1. Historique

Le calcul et l'utilisation des attributs sismiques remontent aux origines des méthodes de la prospection sismique dans l'exploration pétrolière. Le temps d'arrivée et le pendage des événements sismiques ont été employés dans l'évaluation géologique des structures. L'introduction de l'auto-corrélogramme (Anstey et Newman) a mené à une bonne estimation des multiples et à une autre utilisation de la déconvolution. Les analyses de vitesses pour effectuer la correction dynamique « NMO » ont donné de bonnes évaluations de la vitesse d'intervalles et par conséquent une géométrie plus précise de la subsurface. L'interprétation des anomalies « Bright-spot » a mené à la détection de gaz ; elle a été améliorée par l'introduction de la technique AVO. Chacun de ces développements a réduit les incertitudes. Cependant, l'un des principaux échecs de chacune de ces différentes techniques était la dépendance implicite à la technique elle-même et chaque méthode nécessite une autre complémentarité par une ou plusieurs autres. Pour remédier à ce problème la combinaison d'un ensemble de données est devenue une nécessité. Une méthode très puissante, basée sur l'utilisation combinée des attributs sismiques, a été développée.

Les premiers attributs sismiques de la trace complexe ont été présentés en 1970 comme des représentations utiles pour aider à interpréter qualitativement les données sismiques. Le premier article a été publié par Walsh de Marathon dans la revue Geophysics en 1971 sous le titre de " Color Sonograms ". Dans le même temps Nigel Anstey de Seiscom-Delta avait publié " Seiscom 1971 " et introduit l'enveloppe et la fréquence moyenne. L'enveloppe était le résultat d'un filtrage passe bas, c'est une trace sismique rectifiée. Les couleurs ont révélé plus d'information que celle montrée sur les sections sismiques bicolores. Taner, Koehler et Anstey s'étaient intéressés par la propagation de l'onde et au mouvement harmonique simple. Ceci a conduit à l'identification du signal enregistré en tant que représentation de la partie cinétique du flux de l'énergie. A la base de ce modèle, Koehler a développé une méthode pour calculer la composante potentielle à partir de sa partie cinétique. Un autre auteur, Neidell a suggéré l'utilisation de la transformation de Hilbert. Koehler a poursuivi les travaux par le développement de la fréquence et du programme de la transformée de Hilbert dans le domaine temporel. Dans les années 70, trois principaux attributs étaient assez bien développés. Au cours des dernières années, d'autres ont été ajoutés.

Premiers travaux:

- A.H Balch (Marathon), 1971. « Color Sonograms » Geophysics
- Research Team at Seiscom-Delta
 - Nigel Anstey (Seiscom 71 & 72)
 - Fulton Koehler (Energy Equation)
 - Norman Neidell (Hilbert Transform)
 - Tury Taner (Phase, Polarity)
 - Robert Sheriff (AAPG Memoirs 26)

Premiers Attributs:

- Event time and Dips (Rieber Sonograms), 1940's
- Event time, Dips and Velocities (Migration), 1950's
- NMO Velocities and Semblance, 1960's
- Bright spot, True relative amplitudes, Acoustic Impedance Inversion, 1970's
- Complex trace attributes, 1970-1980's
- AVO-AVA, Elastic Impedance, 1980-1990's

I.2. Introduction

Un attribut est une qualité attribuée à toute personne ou à toute chose. Les attributs sismiques sont toute information obtenue à partir des données sismiques, par des mesures directes ou à la base d'un raisonnement logique indirect.

L'étude et l'interprétation des attributs sismiques donnent des informations qualitatives sur les paramètres physiques et sur la configuration géométrique de la subsurface.

L'interprétation de chaque attribut, fournit une information sur un seul paramètre physique. La complexité des formations géologiques a poussé les explorateurs à la combinaison des attributs afin d'exploiter plusieurs paramètres physiques à la fois, permettant de faire une évaluation lithologique du modèle à étudier. Cette utilisation combinée nécessite d'abord une classification et un regroupement des attributs dans des classes selon leur sens physique, leur développement mathématique et leurs types d'application.

La fiabilité des attributs sismiques dépend directement de la qualité des données sismiques (données Input). Il est à noter que le contenu en amplitudes des données sismiques est le facteur principal pour la détermination des paramètres physiques, comme l'impédance acoustique, les coefficients de réflexion, les vitesses, l'absorption, ...etc.

Les attributs calculés à partir de la trace complexe, telle que l'enveloppe, la phase instantanée..., correspondent aux diverses mesures de la propagation du front d'ondes. On les appellera « attributs physiques » (*Physical Attributes*). Ceux calculés à partir de la configuration et de la continuité des réflexions, seront appelés « attributs géométriques » (*Geometrical Attributes*).

Le développement des attributs sismiques a connu deux grandes phases. Les premiers travaux ont été concentrés sur la décomposition de la trace complexe et à l'interprétation de chaque attribut d'une façon singulière. Plus tard, les explorateurs ont inclus la combinaison des attributs pour expliquer certaines anomalies (Exemples : Bright spot et AVO). Actuellement les attributs représentent l'axe principal dans la caractérisation de réservoir.

Plusieurs procédés de calcul ont été utilisés pour extraire les attributs. Ce calcul peut être effectué à la base d'une théorie fondamentale comme l'étude du signal complexe où par l'utilisation de méthodes statistiques.

I.3. Théorie de l'élasticité et Ondes élastiques

I.3.1. Paramètres d'élasticité

La réponse sismique dépend de plusieurs paramètres, les plus importants étant : la matrice, la porosité et le type de fluide de remplissage (Figure : I.3.1).

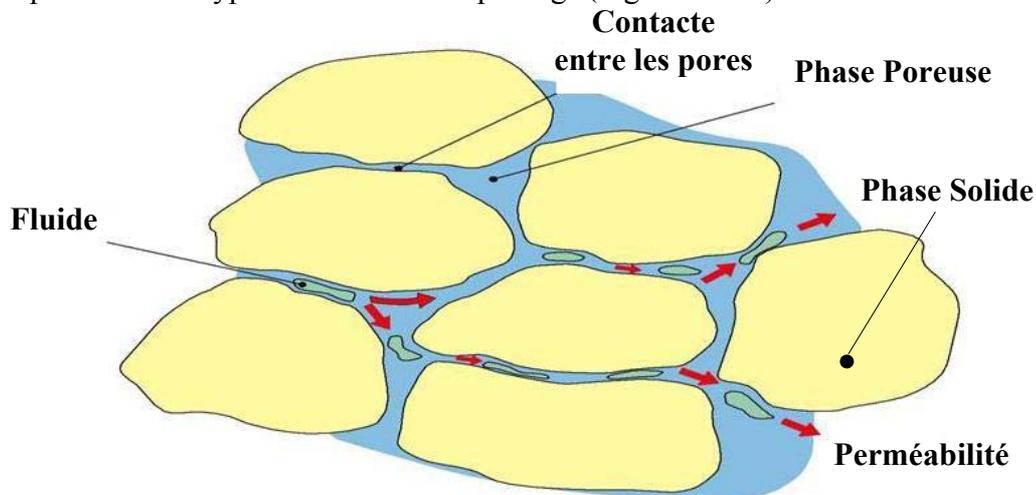


Figure:I.3.1. Porosité et Perméabilité

Le passage d'une onde sismique dans une roche induit de petites déformations qui disparaissent après le passage de celle-ci. Ceci est connu en tant que comportement élastique et est commandé par les propriétés élastiques du matériau. Les propriétés élastiques des roches régissent le passage des ondes sismiques.

Pour de petites déformations dans un matériau élastique, la loi de Hooke déclare que la déformation est proportionnelle à la contrainte.

Les propriétés des matériaux qui obéissent à la loi de Hooke sont caractérisés par les modules d'élasticité suivants:

- Module de d'incompressibilité (de Bulk), k
- Module de cisaillement, μ
- Module de Young, E
- Constante de Lamé, λ
- Le coefficient de Poisson, σ

Le Module d'incompressibilité, k

Le module d'incompressibilité k est le rapport contrainte-déformation sous une pression hydrostatique simple[4].

$$k = \frac{\Delta P}{\Delta V/V} \quad (I.1)$$

où : ΔP = Variation de pression
 V = Volume
 ΔV = le changement de volume
 $1/k$, désigne parfois la compressibilité

Le Module de cisaillement, μ

Le module de cisaillement, μ (Module de rigidité), est le rapport contrainte-déformation pour un cisaillement simple[4].

$$\mu = \frac{\Delta F/A}{\Delta x/x} \quad (I.2)$$

Où: ΔF = Force tangentielle.
 A = Surface de cisaillement
 x = Distance entre les surfaces de cisaillement.
 Δx = Déplacement de cisaillement.

Les fluides n'ont aucune résistance au cisaillement, donc $\mu = 0$.

Le Module de Young, E

Le module de Young, E , est le rapport contrainte-déformation quand une tige est étirée ou comprimée[4].

$$E = \frac{\Delta F/A}{\Delta L/L} = \frac{9k\mu}{3k + \mu} \quad (I.3)$$

Où : ΔF = Force par unité de surface,
 A = Section,
 L = Longueur initiale,
 ΔL = Variation de la longueur L .

Le coefficient de Poisson, σ

Le coefficient de Poisson σ , est le rapport de la contrainte transversale à la contrainte longitudinale quand une tige de longueur L est tirée et allongée de ΔL et sa largeur W se contractant d'une quantité ΔW [4].

$$\sigma = \frac{\Delta W/W}{\Delta L/L} = \frac{3k - 2\mu}{6k + 2\mu} = \frac{E}{2\mu} - 1 \quad (I.4)$$

Le coefficient de Poisson σ , varie de 0 pour les matériaux rigides à 0,5 pour les liquides.

La constante de Lamé λ

La constante de Lamé λ , est une constante d'élasticité dans les matrices des contraintes-déformations[4].

$$\lambda = k - \frac{2\mu}{3} = \frac{\sigma E}{(1 + \sigma)(1 - 2\sigma)} \quad (I.5)$$

Cette constante désigne parfois l'incompressibilité des fluides, pour les fluides non visqueux, $\mu = 0$ et par conséquent $\lambda = k$

I.3.2. Vitesse sismiques et propriétés des roches

Les modules d'élasticité d'une roche et du fluide qu'elle contient sont affectés par un certain nombre de facteurs tels que: la lithologie et contenu d'argile, la porosité, la forme des pores, le type de fluide, la saturation, la pression et la température. Par conséquent, chacune de ces variables affecte la vitesse des ondes **P** et **S** dans la roche.

Les relations entre V_p , V_s , et σ seront résumées par les formules suivantes :

$$V_s = V_p \sqrt{\frac{0.5 - \sigma}{1 - \sigma}} \quad (I.6)$$

avec : $0 < V_s < 0.707 V_p$

$$\sigma = \frac{V_p^2 - 2 V_s^2}{2 \left(V_p^2 - V_s^2 \right)} \quad (I.7)$$

avec : $0 < \sigma < 0.5$

Les vitesses des ondes P et S dans un milieu isotrope peuvent être exprimé en fonction des modules d'élasticité par :

$$V_p = \sqrt{\frac{k + \frac{4}{3}\mu}{\rho}} \quad \text{Et} \quad V_s = \sqrt{\frac{\mu}{\rho}} \quad (I.8)$$

ρ est la densité de la formation.

I.3.3. Ondes élastiques

Pendant le passage des ondes **P**, les particules de la formation se déplacent dans la même direction de propagation de l'onde. D' autre part, les ondes **S** déplacent ses particules perpendiculairement à la direction dans laquelle elles se déplacent. Ces deux modes de propagation des ondes peuvent être illustrés par la figure (figure:I.3.3.).

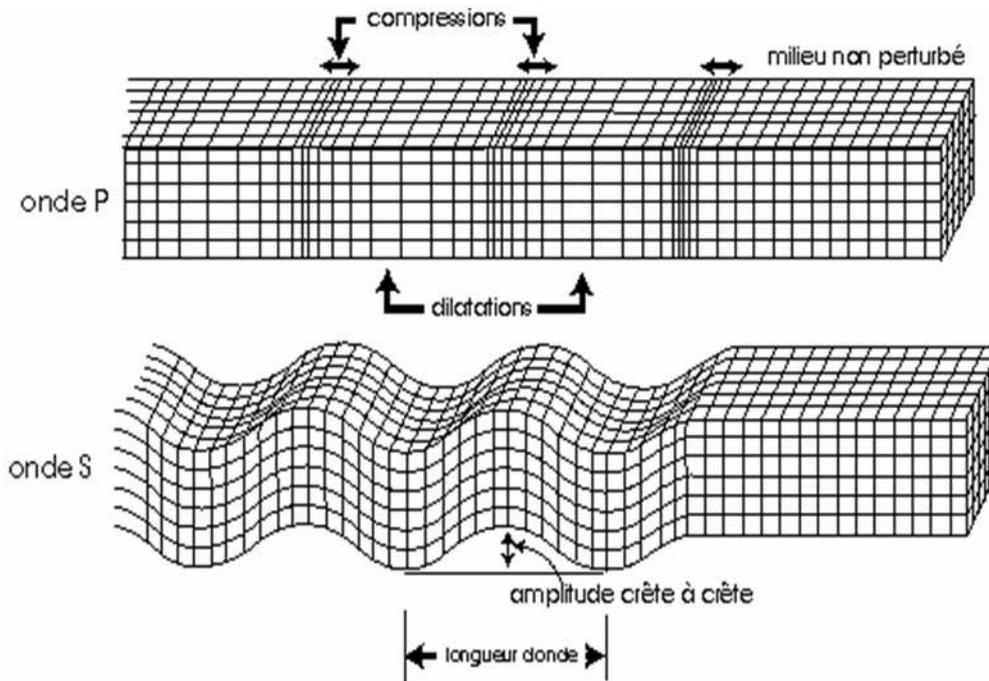


Figure:I.3.3. Mode de propagation des ondes P et S

I.4. Théorie du signal analytique

La théorie de la trace complexe est une simple transposition à la sismique de la théorie du signal analytique. Le concept de signal analytique est employé en télécommunication et dans toute la science de l'ingénierie. Le développement de ce type de signal a fait l'objet de toute une littérature de l'analyse du signal (Gabor,1946 ; Bracewell,1965 ;Cramer,1967 ; Openheim et Schaer,1975). L'application de la trace complexe a été introduite en la sismique par Anstey (1972,1973), Farnbasch (1975), Taner et Sheriff (1977)

I.4.1. Filtre de quadrature

Soit le signal sismique $s(t)$ suivant.

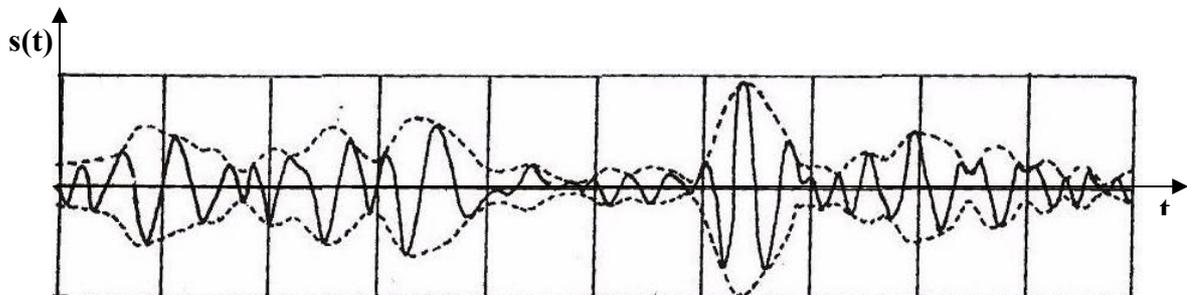


Figure:I.4.1.a. Trace sismique réelle $s(t)$, (Taner et Sheriff 1979) [59]

Le filtre qui fait la conversion de la fonction sinus en une fonction cosinus est appelé un filtre déphaseur de 90° . Il est appelé aussi filtre en quadrature (CLAERBOUT, Fundamental of Geophysical data processing).

La réponse fréquentielle $Q(\omega)$ de ce filtre est donnée par la formule suivante [15] :

$$Q(\omega) = \frac{-i\omega}{|\omega|} \tag{I.9}$$

$$iQ(\omega) = \begin{cases} 1 & \text{si } \omega > 0 \\ -1 & \text{si } \omega < 0 \end{cases}$$

$$iQ(\omega) = \text{sgn}(\omega) \tag{I.10}$$

On peut représenté ce filtre par le graphe suivant :

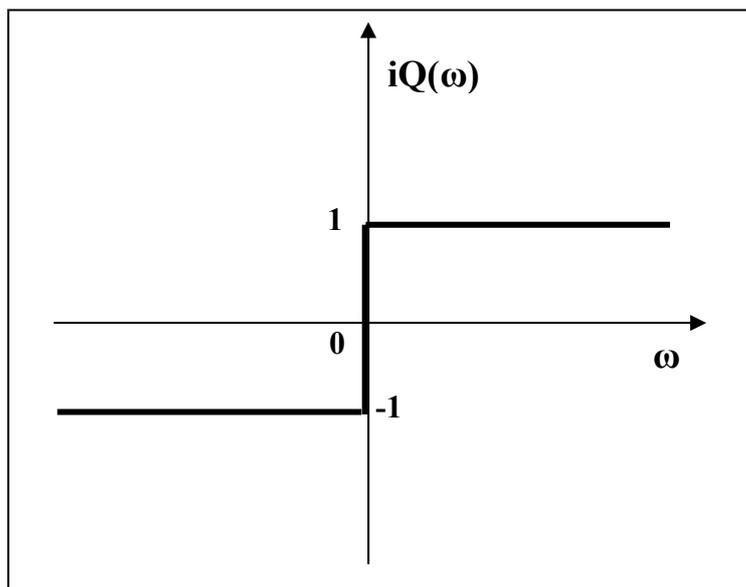


Figure:I.4.1.b. Réponse fréquentielle du filtre de quadrature

La trace en quadrature du signal $s(t)$ est présentée par la figure suivante.

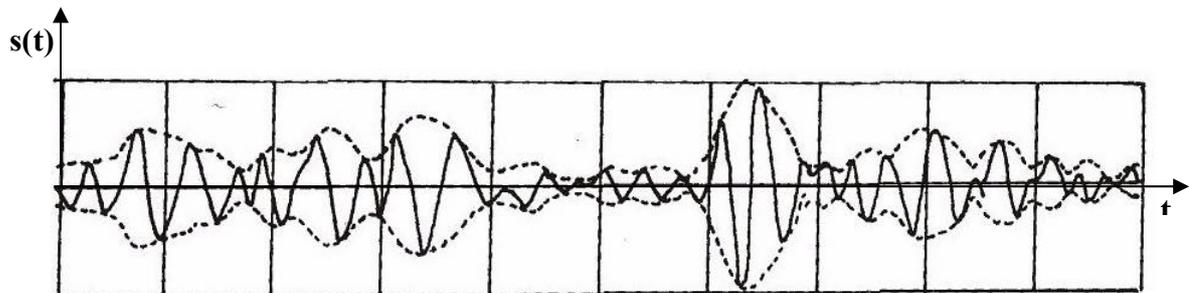


Figure:I.4.1.c Trace en quadrature de $s(t)$, (Taner et Sheriff 1979) [59]

Ce signal s'obtient en retardant le signal d'entrée de 90 degrés à toute fréquence (positive). Le signal $s(t)$ et son signal en quadrature sont présentés respectivement par les deux figures (Figure:I.4.1.a) et (Figure:I.4.1.c).

I.4.2. Transformée de Hilbert

La convolution d'un signal causal $s(t)$ par les coefficients du filtre de quadrature $q(t)$ est appelée la transformation de HILBERT de ce signal.

La transformation de Hilbert est comparable a une opération de filtrage linéaire de réponse impulsionnelle $q(t)$:

$$q(t) = \frac{1}{\pi} \text{V.P.} \left[\frac{1}{t} \right] \tag{I.11}$$

où : V_p désigne la valeur principale au sens de Cauchy

Dans le domaine temporel la transformation de Hilbert d'un signal $\mathbf{s}(t)$ sera formulée comme suit:

$$\text{TH}[\mathbf{s}(t)] = \frac{1}{\pi} V_p \left[\frac{1}{t} \right] * \mathbf{s}(t) \quad (\text{I.12})$$

$$\hat{\mathbf{s}}(t) = \text{TH}[\mathbf{s}(t)] = \frac{1}{\pi} V_p \left[\frac{1}{t} \right] * \mathbf{s}(t) \quad (\text{I.13})$$

La transformée de Fourier du signal en quadrature :

$$\text{TF}[\hat{\mathbf{s}}(t)] = \hat{\mathbf{S}}(f) = \text{TF} \left(\left[\frac{1}{\pi} V_p \frac{1}{t} \right] * \mathbf{s}(t) \right) \quad (\text{I.14})$$

$$\hat{\mathbf{S}}(f) = \text{TF}[\mathbf{s}(t)] \text{TF} \left[\frac{1}{\pi} V_p \frac{1}{t} \right]$$

$$\hat{\mathbf{S}}(f) = -j \text{sgn}(f) \mathbf{S}(f) \quad (\text{I.15})$$

Le filtre de Hilbert est un filtre linéaire dont la réponse impulsionnelle est $1/\pi t$, il n'est donc pas causal. La fonction de transfert de ce filtre est $-j \text{sgn}(f)$

Un signal $\mathbf{s}(t)$ et sa $\text{TH}\{\mathbf{s}(t)\} = \hat{\mathbf{s}}(t)$ ont la même caractéristique d'amplitude.

On remarque que le signal en quadrature et celui réel ont le même module.

$$|\hat{\mathbf{S}}(f)| = |\mathbf{S}(f)| \quad (\text{I.16})$$

et la phase de $\hat{\mathbf{S}}(f)$ est altérée de $\pm \frac{\pi}{2}$, selon le signe de f , de celle de $\mathbf{S}(f)$. Donc le signal en quadrature ne diffère de celui réel que par sa phase. Le déphasage est de $+\frac{\pi}{2}$ pour les fréquences positives, et de $-\frac{\pi}{2}$ pour les fréquences négatives.

La transformée de Hilbert d'un signal réel $\mathbf{s}(t)$ est équivalente à une opération de filtrage pour laquelle les amplitudes des composantes spectrales du signal d'entrée $\mathbf{s}(t)$ sont conservées, tandis que leurs phases sont altérées de $\pm \frac{\pi}{2}$ selon le signe des fréquences f . Cette dernière opération est appelée **quadrature**. C'est la raison pour laquelle le filtre de Hilbert est dit : **filtre de quadrature**.

La transformée de Hilbert est une transformation réversible et sa transformée inverse a pour réponse impulsionnelle $-\frac{1}{\pi} V_p \frac{1}{t}$ et comme fonction de transfert $j \text{sgn}(f)$

I.4.3. Signal analytique

Un signal $\mathbf{s}(t)$ et sa transformée de Hilbert $\hat{\mathbf{s}}(t)$ peuvent être associés pour former un signal complexe $\mathbf{z}(t)$ dont le module est appelé enveloppe.

$$\mathbf{z}(t) = \mathbf{s}(t) + j\hat{\mathbf{s}}(t) \quad (\text{I.17})$$

Le signal $\mathbf{z(t)}$ ainsi obtenu est appelé **signal analytique** associé au signal réel $\mathbf{s(t)}$. $\hat{\mathbf{s(t)}}$, transformée de Hilbert du signal $\mathbf{s(t)}$, est appelé **signal en quadrature** associé au signal $\mathbf{s(t)}$.

à partir de la définition du signal analytique $\mathbf{z(t)}$, on déduit :

$$\mathbf{s(t)} = \text{Re}\{\mathbf{z(t)}\} \tag{I.18}$$

donc $\mathbf{z(t)}$ contient la même information que $\mathbf{s(t)}$. Il s'agit tout simplement d'une autre représentation du signal $\mathbf{s(t)}$. L'importance de cette représentation apparaît dans le domaine de la fréquence. Calculons la TF du signal analytique $\mathbf{z(t)}$.

$$\mathbf{Z(f)} = \text{TF}\{\mathbf{z(t)}\} = \text{TF}\{\mathbf{s(t)}\} + j\text{TF}\{\text{TH}\{\mathbf{s(t)}\}\} \tag{I.19}$$

Puisque

$$\text{TF}\{\text{TH}\{\mathbf{s(t)}\}\} = -j\text{sgn}(f)\mathbf{S(f)} \tag{I.20}$$

on déduit

$$\mathbf{Z(f)} = \mathbf{S(f)} + j(-j\text{sgn}(f)\mathbf{S(f)}) = (1 + \text{sgn}(f))\mathbf{S(f)} \tag{I.21}$$

Pour les fréquences négatives, $\mathbf{Z(f)} = \mathbf{0}$ donc le signal analytique n'a pas de composantes à fréquences négatives! D'autre part pour les fréquences positives $\mathbf{Z(f)} = 2\mathbf{S(f)}$ et l'énergie totale de $\mathbf{z(t)}$ est la même que celle de $\mathbf{s(t)}$.

Le signal analytique lui-même est obtenu en effectuant la transformée de Fourier inverse de cette expression (I.21). Mais il faut connaître la transformée de la fonction **sign(t)** qui vaut **-1** pour $t < 0$ et **+1** pour $t > 0$

La transformation de Hilbert donne lieu à de nombreuses applications en sismique, par exemple pour :

- introduire un déphasage,
- mesurer une enveloppe,
- mesurer la phase instantanée,
- mesurer la fréquence instantanée,
- transformer un signal en un signal à phase minimale.

La trace sismique conventionnelle peut être vue comme la partie réelle d'une trace complexe qui s'obtient à partir de la transformée de Hilbert. Cette analyse des traces sismiques, comme un signal analytique à une dimension, est appelée analyse complexe des traces. Elle peut être aussi développée en deux. Le signal analytique bidimensionnel résulte de la transformée de Hilbert 2D permet de calculer les attributs sismiques 2D.

Dans l'espace, la représentation de la trace complexe associée à la trace réelle et la trace sismique en quadrature est donnée par la figure (Figure : I.4.3.a.).

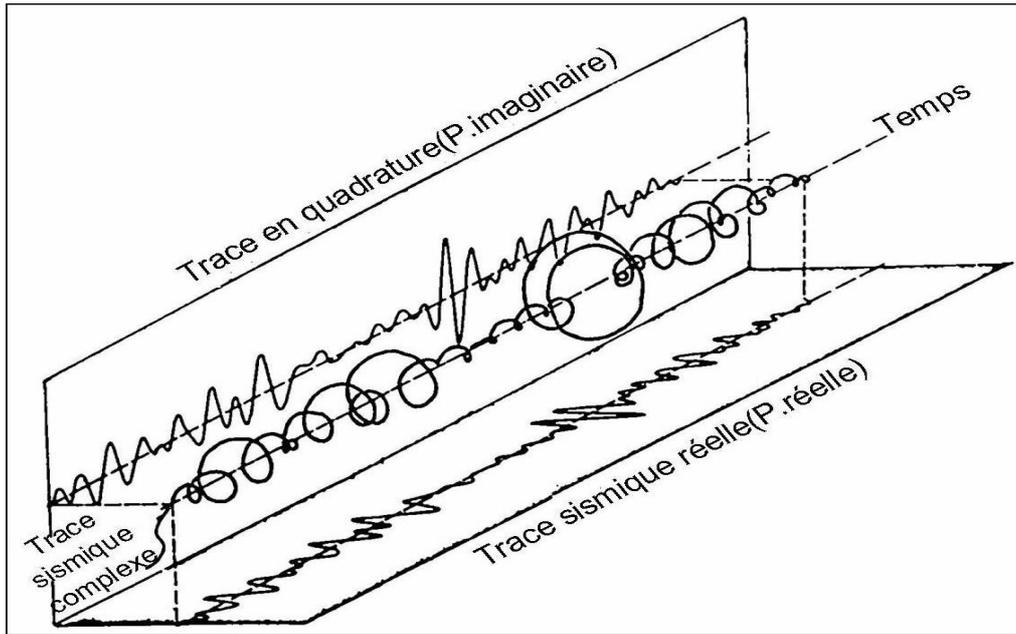


Figure:I.4.3.a. Représentation de la trace complexe,(Taner et Sheriff 1979) [59]

Une rotation constante de phase de la forme d'une onde sismique aura comme conséquence une autre forme différente que celle avant la rotation. Ainsi, une onde sismique peut être décomposée en une partie réelle et une partie imaginaire. Une rotation de phase de cette onde représente la partie imaginaire. Ce concept est illustré en utilisant le modèle physique simple ci-dessous (Figure:I.4.3.b), dans lequel une longueur de fil raide se développe en spirales autour d'une tige terminée par une poignée pour la tourner. Le fil représente une forme d'onde complexe, la tige représente l'axe du temps et la poignée représente le rotateur de phase. Le mouvement vertical du fil à un point quelconque est connu comme **amplitude réelle**, tandis que le mouvement horizontal au même point est connu comme **quadrature**. Mathématiquement, ceci peut être dérivé de la trace sismique réelle en utilisant la transformation de HILBERT.

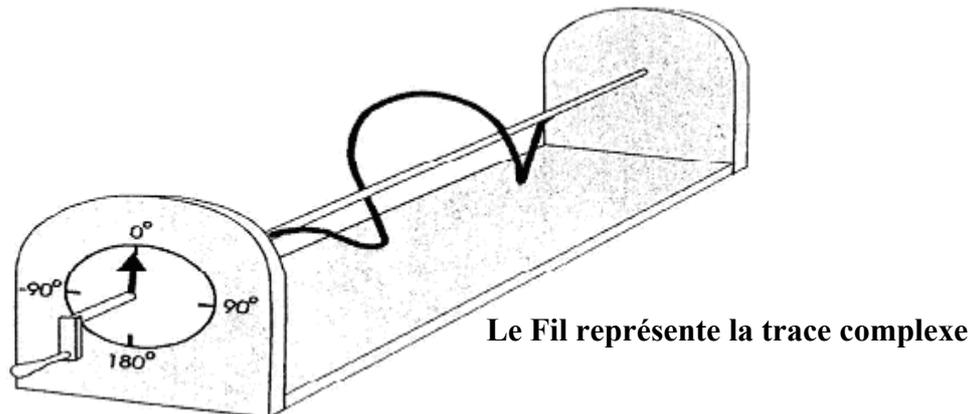


Figure:I.4.3.b. schéma illustratif d'une trace complexe

La longueur de la ligne séparant un point sur le fil au plus proche point sur la tige représente l'**amplitude instantanée**. C'est égale à la racine carrée de la somme des carrés de l'amplitude réelle et sa quadrature à ce point. La poignée étant placé à 0° , l'angle entre l'amplitude instantanée et la verticale représente la phase instantanée correspondante.

La phase instantanée est la rotation de phase qui donne l'amplitude maximale à chaque instant.

L'enveloppe enferme une suite de formes d'onde qui représentent la même forme d'onde après différentes rotations de phase, s'étalant de 0° à 180° , avec un incréments de 45° .

I.5. Classification des attributs sismiques

L'objectif initial des explorateurs était de développer autant d'attributs physiques afin de définir les paramètres lithologiques du réservoir.

Les attributs peuvent être calculés avant sommation ou après sommation des données, avant ou après migration. Les attributs peuvent être classifiés de différentes manières. Plusieurs auteurs ont donné leurs propres classifications. Ici nous illustrons la classification de Taner[54]. C'est une classification basée sur les caractéristiques suivantes :

- Paramètres géométriques. (Offset et Azimut)
- Mode de calcul. (Instantanée et Ondelette)
- Relation des attributs avec la géologie. (Physique et Géométriques)
- Type de propagation de l'onde. (Transmission et Réflexion)

En tenant compte des informations de l'Offset et de l'Azimut on peut classer les attributs comme suit (Prestack – Poststack Attributs).

I.5.1. Attributs avant sommation (Pre-Stack)

Les données d'entrée sont des traces arrangées sous forme de CDP gather (common depth point). Elles auront l'information relative à l'azimut et à l'offset. Les attributs calculés dans ce cas donnent des renseignements sur l'azimut et sur l'offset ainsi que sur le caractère du signal des traces élémentaires constituant chaque CDP. Les calculs produisent des quantités énormes de données; par conséquent ils ne sont pas pratiques pour des études préliminaires.

I.5.2. Attributs après sommation (Post-Stack)

Après avoir effectué la sommation, on perd toute information relative à l'azimut et à l'offset. Les données d'entrée sont des CDP sommés ou sommés et migrés. Notons que pour la migration en temps, les données maintiennent leurs relations de temps. Par conséquent les variables temporelles (telle que la fréquence) maintiendront leurs dimensions physiques. Pour les sections migrées en profondeur, la fréquence est remplacée par le nombre d'onde, qui est une fonction de la vitesse de propagation et de la fréquence.

Les attributs post-stack sont meilleurs pour observer de grandes quantités de données pour des investigations initiales. Pour des études détaillées, les attributs pré-stack peuvent être utilisés.

par rapport à l'information contenue, les attributs sont divisés en deux catégories générales (Instantaneous -Wavelet Attributs).

I.5.3. Attributs Instantanés

Les attributs instantanés calculés échantillon par échantillon, représentent des variations instantanées de divers paramètres. Des attributs instantanés tels que l'enveloppe de la trace, ses dérivés, la fréquence et la phase peuvent être déterminées à partir des traces complexes.

I.5.4. Attributs De l'ondelette

Les attributs instantanés calculés à la crête de l'enveloppe de la trace ont une relation directe avec la transformée de Fourier de l'ondelette à proximité de la crête de cette enveloppe. Par exemple, la fréquence instantanée à la crête de l'enveloppe est égale à la fréquence moyenne du spectre d'amplitude de l'ondelette. La phase instantanée correspond à l'intercept de la phase de l'ondelette.

La troisième classification est basée sur la relation des attributs à la géologie (Physical – Geometrical Attributes).

I.5.5. Attributs Physiques

Les attributs physiques sont liés aux qualités physiques et aux quantités des propriétés des roches. La magnitude de l'enveloppe de la trace est proportionnelle au contraste d'impédance acoustique, les fréquences sont reliées à l'épaisseur des couches minces, à la dispersion des ondes et à l'absorption. Les vitesses instantanées et moyennes sont reliées directement aux propriétés des roches. En conséquence, ces attributs sont dans la plupart du temps employés pour la classification lithologique et la caractérisation de réservoir.

I.5.6. Attributs Géométriques

Les attributs géométriques décrivent le rapport spatial et temporel de tous les autres attributs. La continuité latérale mesurée par semblance est un bon indicateur de la similitude des couches aussi bien que de la discontinuité. Le pendage des couches et les courbures fournissent des informations sur les dépôts. Le but de l'utilisation des attributs géométriques est l'interprétation stratigraphique. Cependant, l'expérience a prouvé que les attributs géométriques définissant les caractéristiques des événements et leurs relations spatiales et mesurent les modes de dépôts qui aident directement à l'identification de ces modes de dépôt, et leurs relations lithologiques.

On admet que la majeure partie des attributs, instantanés ou d'ondelette, étudie les caractéristiques de l'ondelette sismique réfléchie, c'est-à-dire, en rapport avec les interfaces séparant deux couches. Cependant, la vitesse et l'absorption sont mesurées comme quantités se produisant entre deux interfaces, ou dans une couche. Par conséquent, nous pouvons diviser les attributs en deux catégories basées sur leurs origines.

I.5.7. Attributs de Réflectivité

Ce sont des attributs qui correspondent aux caractéristiques des interfaces. Tous les attributs instantanés et d'ondelette peuvent être inclus dans cette catégorie. Les attributs Pré-stack comme l'AVO sont également des attributs de réflectivité, puisque l'AVO étudie la dépendance de la réponse réfléchie d'une interface avec l'angle d'incidence.

I.5.8. Attributs De Transmissivité

Les attributs de transmissivité sont reliés aux caractéristiques des strates entre deux interfaces, l'intervalle, les vitesses moyennes, le facteur de qualité Q, l'absorption et la dispersion faisant partie de cette catégorie.

I.6. Calcul des attributs de base

I.6.1. Attributs instantanés

La formulation initiale des attributs pour l'interprétation des données sismiques a été donnée par Taner dans un article publié en 1979. Cet article a couvert cinq attributs, L'enveloppe, la phase instantanée, la fréquence instantanée, la fréquence moyenne et la polarité apparente. Plus tard leur application a été discutée par Robertson et Nogami (1984) pour l'analyse des couches minces, et Robertson et Fisher (1988) pour l'interprétation générale. Le développement le plus récent des attributs sismiques a été fait par Taner en 1992 et révisé en 2000[54].

L'Enveloppe

Soit la trace analytique donnée par :

$$Z(t) = s(t) + ig(t) \quad (I.22)$$

Où:

$s(t)$: est la partie réelle correspondant aux données sismiques enregistrées.

$g(t)$: partie imaginaire de la trace complexe, transformée de Hilbert de $s(t)$.

Alors l'enveloppe est le module de la fonction complexe;

$$E(t) = \sqrt{[s^2(t) + g^2(t)]} \quad (I.23)$$

$E^2(t)$ représente l'énergie instantanée. Sa magnitude (grandeur) est du même ordre que celle des traces d'entrée. Elle change approximativement entre 0 et l'amplitude maximale de la trace sismique. Comme indiqué sur l'équation (I.23), l'enveloppe est indépendante de la phase et elle est reliée directement aux contrastes d'impédance acoustique. Elle peut représenter le contraste d'interface individuel ou, plus probable, la réponse combinée de plusieurs interfaces.

L'enveloppe de la trace est un attribut physique et elle peut être employée comme discriminateur efficace pour les caractéristiques suivantes :

- . Représentation du contraste d'impédance acoustique. (D'où la réflectivité)
- . Bright spots
- . Possibilité d'accumulation de gaz
- . Limites des séquences
- . Tuning effects (couches minces)
- . Discordances
- . Grandes variations lithologiques
- . Indication des failles par variation latérale

Première dérivée de l'enveloppe (taux de variation de l'enveloppe)

Le taux de changement de l'enveloppe en fonction du temps montre la variation de l'énergie des événements réfléchis. La dérivée de l'enveloppe $\partial[E(t)]/\partial t$ des événements montre l'effet d'absorption. Une variation plus lente indique un grand effet d'absorption. Taner utilise un filtre dans le domaine temporel particulièrement conçu pour performer la différentiation. Un taux de variation de l'enveloppe relativement brusque implique un effet d'absorption faible. Cet attribut est également un attribut physique et il peut être employé pour détecter la possibilité des fractures et les effets d'absorption.

- . La variation de l'enveloppe Elle est caractérisée par une relation inversement proportionnelle à l'absorption
- . Une variation latérale indique la présence de discontinuités.
- . Elle est utilisée dans le calcul de la direction de propagation de groupe. En comparaison avec la direction de propagation de phase, elle peut indiquer les ondes dispersives.

Deuxième dérivée d'enveloppe

La deuxième dérivée de l'enveloppe donne une mesure du changement de la crête de l'enveloppe. Avec une représentation en noir et blanc, elle montre toutes les crêtes de l'enveloppe, ceci correspond à toutes les interfaces des réflecteurs visibles dans une séquence de couches. Elle montre aussi les limites des événements et indique les changements lithologiques brusques.

Phase instantanée

L'argument de la trace analytique est la phase instantanée :

$$\varphi_i(t) = \arctan\left[\frac{g(t)}{s(t)}\right] \quad (I.25)$$

L'information de phase est indépendante des amplitudes des traces et elle est reliée à la phase de propagation du front d'ondes sismiques. Puisque le front d'onde est défini comme des lignes de phase constante (généralement), l'attribut de phase est également un attribut physique et peut être efficacement utilisé comme discriminateur pour des classifications des formes géométriques :

- . Bon indicateur de la continuité latérale.
- . Utilisé dans le calcul de la vitesse de phase.
- . Indicateur des événements de faibles amplitudes.
- . Montre les discontinuités, (Failles et changement lithologique)
- . Visualisation des limites des séquences,
- . Visualisation des détails dans une séquence d'événements.
- . Appliqué dans le calcul de la fréquence instantanée et accélération

Amplitude Normalisée

L'interprétation des sections caractérisées par un mauvais rapport signal sur bruit sera très difficile. L'application d'un AGC (Automatic Gain Control) avec une petite fenêtre change les caractéristiques des événements sismiques. La représentation d'une section sans changement de ces caractéristiques de base est nécessaire pour ce genre de données. Le cosinus de la phase instantanée présente une solution parfaite. Cet affichage a tous les détails d'une section phase instantanée et montre la continuité latérale désirée. Cet attribut n'a aucune information d'amplitude et son enveloppe est une constante égale à l'unité.

Fréquence Instantanée

Le taux de changement de la phase par rapport au temps est la fréquence instantanée:

$$F_i(t) = \frac{1}{2\pi} \frac{\partial \varphi_i(t)}{\partial(t)} \quad (I.26)$$

$$F_i(t) = \frac{1}{2\pi} \frac{\partial \arctan[g(t)/s(t)]}{\partial(t)} = \frac{1}{2\pi} \frac{[s(t).dg/dt - g(t).ds/dt]}{[s^2(t) + g^2(t)]} \quad (I.27)$$

La fréquence instantanée est montrée par une table de couleur commencent par le rouge comme la plus basse fréquence et elle change graduellement en jaune, en vert et finalement en bleue pour les fréquences les plus élevées. L'unité de la fréquence est le cycle par seconde. La phase instantanée représente la phase du vecteur résultant des différents mouvements harmoniques simples. Quoique les différents vecteurs ont un mouvement dans le sens des aiguilles d'une montre, leur vecteur résultant peut, parfois, former un modèle de cardioïde et sembler tourner dans la direction opposée. Taner a interprété ce phénomène comme effet de l'interférence de deux ondelettes arrivées presque en même temps. Ceci peut

également être provoqué par l'interférence des bruits dans des zones caractérisées par des faibles amplitudes. En raison de ces inversions, la fréquence instantanée aura des fluctuations et des grandeurs anormales. Puisque les fréquences instantanées sont influencées par l'épaisseur des couches, nous voudrions les observer sans d'interférence. Pour cela Taner a utilisé plusieurs traces adjacentes pour former un résultat cohérent.

Il a montré que la fréquence instantanée, calculée comme la dérivée de la phase instantanée par rapport au temps, est reliée au centre du spectre de puissance de l'ondelette sismique.

Les fréquences instantanées relient la propagation des ondes à l'environnement des dépôts. Cet attribut est un attribut physique et il correspond à la fréquence moyenne (centroïde) du spectre de puissance de l'ondelette sismique,

- . Corrélateur du caractère sismique dans la direction latérale,
- . Indique les bords des bancs minces de faibles impédances,
- . Indicateur d'hydrocarbures par anomalie de basse fréquence. Cet effet est quelque fois accentué par des sables non consolidés dus à la teneur en huile dans les pores.
- . Indicateur des zones fracturées, Ces zones sont généralement caractérisées par de basses fréquences.
- . Indicateur de l'épaisseur des bancs. Les hautes fréquences indiquent les limites des interfaces et les bancs minces argileux. Les basses fréquences indiquent les bancs de sable.
- . Indicateur du rapport sable/argile dans un environnement clastique.

Indicateur de couches Minces

La présence des sauts de fréquences et des fréquences négatives présentent des anomalies de fréquences instantanées. Nous pouvons extraire une information très importante de ces anomalies. Les endroits où les fréquences instantanées montrent des variations brusques ou changent de signe sont représentatifs des ondelettes réfléchies étroitement proches de point du vue temps d'arrivée. Par conséquent, la dérivée par rapport au temps de la fonction phase contiendra les indicateurs des bancs minces (des interfaces très proches), sous formes de grandes variations de la fréquence instantanée. Cette variation lissée, est liée aux caractéristiques de stratification. L'indicateur de couches minces est donc calculé comme la différence entre les fréquences instantanées et la fréquence moyenne par rapport au temps :

$$Thb(t) = F(t) - \bar{F}(t) \quad (I.28)$$

Thb :Thin bed(Couche mince)

F (t) : Fréquence

$\bar{F}(t)$: Fréquence moyenne pondérée

La fréquence moyenne par rapport au temps a été définie par Taner comme fréquence pondérée. Cette méthode utilise l'enveloppe, ou le carré de l'enveloppe (toute la puissance) comme fonction de pondération.

$$\bar{F}(t) = \frac{\sum_T [E(t) \cdot F(t)]}{\sum_T [E(t)]} \quad (I.29)$$

Où : **T** : est la fenêtre de lissage (en temps).

E(t) : Enveloppe

F (t) : Fréquence

Cet attribut de fréquence est moins influencé par des effets de courtes longueurs d'onde. Les facteurs de grandes longueur d'onde, tels que l'absorption due aux bancs épais ou aux corps de sable massifs, changeront les caractéristiques de la propagation de l'ondelette qui peuvent être observées sur l'attribut fréquence moyenne. C'est un attribut physique, indiquant les variations des grandes longueurs d'onde.

Impédance Acoustique Relative

Nous supposons que les données sismiques ont été traitées pour filtrer les bruits et atténuer les multiples. Partant de cette hypothèse, la trace sismique représente la série de réflectivité par bande limitée, qui peut être exprimée comme :

$$s(t) = \frac{1}{2} \frac{\Delta(\rho v)}{(\rho v)} \quad (I.30)$$

ρ :Densité

v :Vitesse

De façon équivalente:

$$s(t) = \frac{1}{2} \Delta \ln(\rho v) \quad (I.31)$$

Par conséquent, en intégrant la trace zéro phase, nous obtiendrons l'estimation du log d'impédance acoustique dans une bande limitée (signal borné). Puisque c'est une bande limitée, l'impédance n'est pas absolue et la section stack est souvent une estimation de la réflectivité zero offset; par conséquent elle est appelée l'impédance acoustique relative.

$$\ln(\rho v) = 2 \cdot \int_{t=0}^{t=T} s(t)dt \quad (I.32)$$

Dans la pratique, et en raison du bruit et du contenu imparfait des données sismiques, l'impédance acoustique relative calculée par intégration développera d'une façon arbitraire une tendance vers les grandes longueurs d'onde. Comme les données sismiques ne contiennent aucune information fiable aux basses fréquences, ces tendances vers les grandes longueurs d'onde ne peuvent pas être utilisées. Afin de remédier à ce problème on applique un filtre coupe bas.

Le calcul est une intégration simple suivie d'un filtrage coupe bas. Il reflète le contraste des propriétés physiques, par conséquent c'est un attribut physique efficacement utilisé dans beaucoup de procédures de calibrage.

- . L'impédance acoustique relative démontre le contraste d'impédance acoustique apparente dans une bande limitée.
- . Un fort contraste indique les limites possibles des séquences.
- . L'impédance indique les surfaces de discordance.
- . L'impédance détecte les discontinuités.

I.6.2. Attributs de l'ondelette

Les attributs de l'ondelette sont, par définition (Taner), les attributs instantanés calculés aux points maxima de l'enveloppe de la trace sismique. La crête de l'enveloppe représente une position où la plus grande partie de l'énergie des différentes bandes de fréquences est en phase. Par conséquent, les attributs calculés à cette position auront des implications statistiques directement liés aux caractéristiques spectrales de l'ondelette. Beaucoup d'auteurs (Bodine, 1984, Barnes, 1991) ont publié des études statistiques du signal analytique et du rapport entre la transformée de Fourier et les attributs instantanés. Ces études

prouvent que les attributs calculés à la crête de l'enveloppe sont directement liés aux divers moments de la densité spectrale de puissance. Par conséquent le calcul et l'affichage de ces attributs fourniront des informations sur les caractéristiques de l'ondelette réfléchie. Bodine (1984) appelle ces attributs « 'response' attributes ». Taner a préféré 'attributs de l'ondelette'. Puisque une seule valeur va représenter toute ondelette, nous devons déterminer et définir la longueur de cette ondelette.

Définition : Nous supposons qu'une ondelette sismique, simple ou composée, occupe la période entre deux minimums adjacents de l'enveloppe. En réalité, les ondelettes se prolongeront au delà de ces minimums. Cependant, pour l'affichage on considère les minimums comme frontière. En outre la partie de l'enveloppe située entre les minimums représente une partie significative de cette dernière.

Par conséquent, on recherche les positions 'minimums' et 'maximums' de l'enveloppe. Toutes les positions 'maximum' de l'enveloppe sont employées pour calculer les attributs instantanés et ces valeurs seront distribuées sur la fenêtre de temps entre les deux minimums de chaque côté de la position du calcul (pic maximum de l'enveloppe). Pour plus de précision, on utilise une interpolation quadratique (figure : I.6.2).

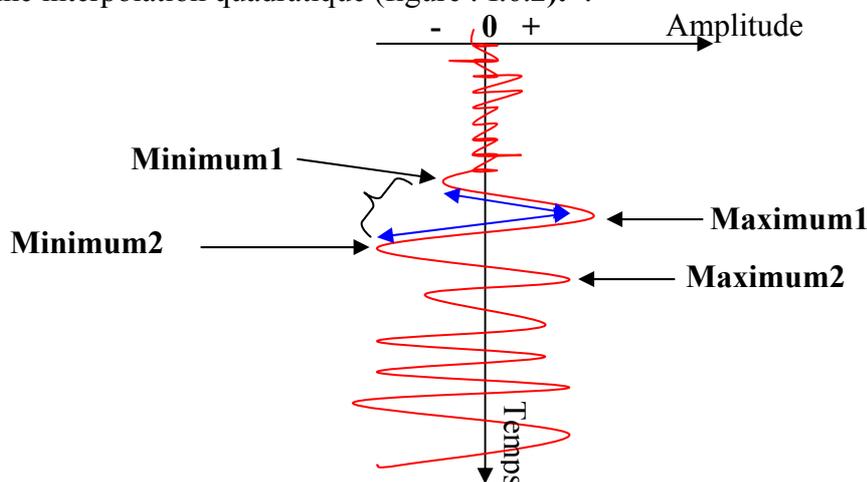


Figure : I.6.2 Principe de calcul des attributs de l'ondelette

Les attributs de l'ondelette les plus significatifs sont :

- L'Enveloppe de l'ondelette ,
- La dérivée du maximum positif de l'enveloppe de l'ondelette
- La dérivée seconde de l'enveloppe de l'ondelette,
- La phase de l'ondelette, La polarité apparente,
- La fréquence instantanée de l'ondelette,
- L'accélération de l'ondelette,

Ces attributs donnent des indications sur les épaisseurs des couches, la continuité des dépôts, la continuité latérale et sur les limites des séquences.

I.7. Calcul des attributs sismiques dans la direction l'information spatiale.

Quoique la mesure de la phase et de sa dérivée par rapport au temps donnent une information directe sur la variation de l'énergie dans la direction temporelle, cette mesure est étendue pour inclure des informations dans la direction spatiale. Elle porte des informations sur le nombre d'onde et la direction de propagation des ondes. Il est bien connu que, l'onde se propage dans deux modes séparés, phase et groupe. Dans un milieu dispersif, ces deux modes

auront des vitesses de propagation différentes. Ici, on suppose un mode non dispersif et on ne considère que la composante de phase.

Un autre avantage important, vient de l'information contenue le long du front d'onde. Cette information aide à réduire les effets des bruits. La fréquence instantanée souffre de l'influence des bruits qui affectent la continuité spatiale. Ceci apparaît ensuite sur la représentation de la fréquence moyenne, comme une bande large de couleur. O'Doherty (1992) [43] a prouvé qu'en incluant l'information spatiale dans le processus de calcul de la fréquence instantanée et du pendage, les événements peuvent être suivis avec plus de rigueur.

L'objectif de ces attributs est de développer les relations entre les traces dans les directions spatiales et temporelles. Ce type d'attributs destiné beaucoup plus à l'interprétation de la sismique 3D, est appelé aussi attributs volumiques.

Le calcul des attributs dans cette classe est basé sur la relation entre chaque échantillon et les échantillons adjacents. Plusieurs auteurs utilisent le terme anglais 'scanning' (balayage des échantillons). Les expressions statistiques représentent la première méthode de calcul. La fonction de corrélation présente aussi une méthode de calcul très puissante. Cette fonction est appliquée sur une fenêtre de calcul en temps et en espace, en déplaçant cette fenêtre avec un recouvrement jusqu'à la fin de la zone d'intérêt. La longueur de la fenêtre dépend du pendage. Les attributs les plus utilisés sont la cohérence, la continuité latérale, l'azimut, le dip(pendage) et le dipAzimut.

Soit $s(t, x)$ une trace sismique enregistrée, cette même fonction représente la partie réelle de la trace complexe. $g(t, x)$ est calculé par la transformée de Hilbert de $s(t, x)$. Les traces complexes bidimensionnelles sont définies comme:

$$Z(x,t) = s(x,t) + i.g(x,t) \quad (I.33)$$

L'amplitude instantanée (Enveloppe) est indiquée par:

$$E(x,t) = \sqrt{[s^2(x,t) + g^2(x,t)]} \quad (I.34)$$

La phase instantanée est indiquée par:

$$\varphi_i(x,t) = \arctan[g(x,t) / s(x,t)] \quad (I.35)$$

La fréquence instantanée est indiquée par :

$$F_i(x,t) = \frac{1}{2\pi} \partial \{ \varphi_i(x,t) \} / \partial t \quad (I.36)$$

I.7.1. Calcul des attributs géométriques

Il y a plusieurs méthodes de calcul des attributs géométriques. Nous avons cité la méthode basée sur l'idée de lissage dans des directions spatiales. Une deuxième méthode basée sur le calcul des différences entre les traces sismiques dans la direction spatiale, échantillon par échantillon, donne des résultats positifs. Si les pendages se prolongent au delà d'un échantillon par trace cette méthode donnera des résultats aliasés. Une troisième méthode de balayage de pendage en utilisant plus de deux échantillons peut maîtriser cette difficulté.

I.7.1.1. Continuité des événements

C'est un résultat intermédiaire d'un calcul hybride des attributs. L'objectif de cet attribut est de développer la continuité latérale des crêtes [pics] et des creux, et de classer le type de discontinuité. On affecte la valeur +1 à toutes les crêtes positives de la trace réelle et la valeur -1 à toutes les crêtes négatives, tout le reste sera nul (zéro). L'affichage de cet attribut montrera seulement les crêtes et les creux, et avec la même grandeur.

I.7.1.2. Dip instantané

La phase instantanée pour trois traces adjacentes est calculée. Une courbe parabolique est adaptée aux valeurs de la phase dans la direction latérale, et le pendage (Dip) à la trace centrale est calculé. Ce pendage est alors ajusté selon l'inter-trace. Pour les données 3D, des dips de phase sont calculées dans les deux directions, in-line et cross-line, à partir desquelles la direction du maximum pendage et son azimuth sont obtenus. Soient dt/dx et dt/dy les pendages (Dips) dans les directions in line et cross line respectivement:

L'expression :

$$\Delta T = \sqrt{(dt / dx)^2 + (dt / dy)^2} \quad (I.37)$$

Donnera le pendage maximum, et sa direction (**dip azimuth instantané**) sera:

$$\phi = \arctan(dt / dy , dt / dx) \quad (I.38)$$

Comme mentionné ci-dessus, ce calcul est valide pour des pendages ayant plus de 180 degrés comme différences de phase. Les pendages réels au delà de cette limite produiront des résultats aliasés, par conséquent ils seront erronées. Dans ce cas le balayage [scanning] de pendages avec un plus grand nombre de traces devient nécessaire.

I.7.1.3. Semblance latérale maximale et la cohérence maximale

La semblance est une mesure de la puissance cohérente entre un certain nombre de traces (groupe de traces, Gather) en fonction de la puissance totale de toutes les traces, elle est calculée comme suit[54];

$$\text{Semb}(t) = \frac{\sum_{\tau=-N/2}^{\tau=N/2} \left\{ \sum_{m=1}^M s_m(t + \tau) \right\}^2 - \sum_{\tau=-N/2}^{\tau=N/2} \sum_{m=1}^M s_m^2(t + \tau)}{\sum_{\tau=-N/2}^{\tau=N/2} \sum_{m=1}^M s_m^2(t + \tau)} \quad (I.39)$$

Où ; $s_m(t)$:est le m^{eme} trace du groupe (gather)

N : est la longueur de la fenêtre de calcul.

M : Nombre total de traces choisies par groupe (gather)

L'indication de "cohérence" latérale est calculée comme la semblance déterminé en considérant un nombre choisi de traces. Le nombre de traces dépend des éléments suivants :

1. Le rapport signal sur bruit (S/B). On augmente le nombre de traces pour des faibles (S/B).
2. La courbure des bancs (plus de courbure, moins de traces),
3. La résolution latérale la plus élevée (une meilleure résolution, moins de traces).

La semblance est généralement calculée par fenêtre de temps, de 40 a 80 millisecondes (des fenêtres de temps plus courtes sont appropriées pour une bande large de données). Pour des données de bon rapport S/B, trois traces adjacentes pour le balayage [scanning] donnent une résolution plus élevée. Cinq ou un nombres de traces plus élevé tendront à perdre des discontinuités latérales.

I.7.1.4. Variance du maximum cohérence, variation instantanée

Cet attribut est la différence entre la semblance latérale maximale instantané et la semblance latérale maximale lissée. Il montre les anomalies locales de dépôts. Il peut également montrer les zones de bruit.

I.7.2. Attributs Hybrides

Les attributs présentés précédemment mesurent fondamentalement des conditions particulières physiques ou géométriques et par fois statistiques. Ce sont des attributs de base, mais ils peuvent également être utilisés par une combinaison logique pour détecter des conditions plus compliquées. Cette combinaison logique peut être une expérience ou une connaissance de base. Par exemple, le facteur de fluide calculé par l'analyse d'AVO est un attribut hybride basé sur l'expérience. Dans cette partie nous présentons des descriptions de plusieurs attributs hybrides (développés par Taner). Leur première introduction était présentée comme outil de discrimination. Le développement de ce type d'attributs nécessite l'incorporation d'autres connaissances des propriétés physiques des roches de la géologie et de la caractérisation de réservoir.

I.7.2.1. Indicateur de couches parallèles

C'est l'écart type des pendages (dips) instantanés (déviations standard du pendage instantané) calculés dans une fenêtre de temps. Après avoir calculé la variance par la formule suivante (I.40), on calcule l'écart type qui est la racine carrée de la variance.

$$\sigma^2(t) = \frac{1}{N} \sum_{i=-N/2}^{i=N/2} [\mathbf{d}(t+i) - \mathbf{d}(t)]^2 \quad (\text{I.40})$$

N : nombre d'échantillons dans la fenêtre de calcul

d(t) : le pendage moyen

Les zones avec des couches parallèles auront des valeurs voisines de zéro (déviations standard). Les valeurs croissantes de la déviation standard indiqueront une variation de pendage dans une fenêtre de calcul. Cet attribut, calculé comme pendage statistique, fournit une indication directe de la configuration géométrique des bancs et des interfaces.

I.7.2.2. Indicateur des zones chaotiques

Les zones chaotiques sont définies comme zones avec un degré élevé de la continuité latérale et avec une variation aléatoire du pendage des couches. Cet attribut est calculé pour détecter les zones caractérisées par une variation aléatoire du pendage des événements sismiques organisés. Les zones chaotiques auront une cohérence latérale très élevée mais leurs pendages changeront rapidement dans le temps et dans l'espace. Les zones sourdes (caractérisées par une absence de réflecteurs) ayant un écart type de pendage très élevée auront également une semblance moyenne latérale faible. L'attribut est calculé en tant que l'écart type des pendages pondérés par la semblance latérale.

$$\mathbf{Chao}(t) = \sigma(t) \cdot \mathbf{Semb}(t) \quad (\text{I.41})$$

σ(t) : L'écart type

Semb(t) : La fonction semblance

Cet attribut a des grandeurs élevées dans les zones des réflexions caractérisées par une distribution aléatoire des pendages, et il présente de faibles grandeurs dans les zones carbonatées.

I.7.2.3. Révélateur de zone de discordances

C'est un attribut hybride utilisé pour détecter les surfaces de discordance. Le programme détecte d'abord les crêtes de l'enveloppe de la trace sismique. De point de vue sismique, la position de la crête de l'enveloppe représente l'interface visible des terminaisons des couches (possibilité d'une surface de discordance). Sa grandeur est relativement proche de celle du contraste d'impédance acoustique. Ensuite on calcule d'autres attributs pour ajouter d'autres contrastes. Trois échantillons de pendage, de la semblance latérale maximale, de la fréquence instantanée, de la semblance latérale et de la variance de pendage sont sélectionnés des deux côtés de l'interface. On calcule le produit scalaire entre l'ensemble des attributs des deux côtés. Après normalisation de ce vecteur (dont ces composantes sont des attributs), toutes ces valeurs seront entre +1 et -1. La valeur +1 représente une similitude très importante, elle indique qu'il n'y a aucune différence entre les deux côtés de la crête de l'enveloppe de point de vue pendage, fréquence et continuité latérale. Les plus petites valeurs indiqueront un certain degré de différence. La soustraction du produit scalaire normalisé de 1 donnera un facteur de dissimilitude, qui est alors multiplié par la valeur de l'amplitude à la crête l'enveloppe. Le résultat est alors considéré comme indicateur des zones de discordance. Les résultats caractérisés par des grandes valeurs représentent un contraste très élevé.

I.7.2.4. Indicateur d'argile

C'est un attribut hybride qui combine plusieurs attributs de base pour détecter les zones d'argiles réparties dans un environnement clastique. Les bancs d'argile sont identifiés par leur configuration géométrique comme couches minces parallèles avec une continuité latérale élevée. Cet attribut utilise les hautes fréquences instantanées comme indicateur de couches minces, des valeurs faibles de l'écart type de pendage comme indicateur de couches parallèles, et enfin de la semblance et de sa variance comme indicateur de la continuité latérale. Une valeur très élevée du résultat de la combinaison de ces attributs indique une forte possibilité de présence d'argile. Des petites valeurs indiquent la possibilité d'une lithologie autre que l'argile, telle que des bancs de sable ou de carbonates.

II. Les Bases théoriques des réseaux de neurones artificiels,

II.1. Historique

Les réseaux de neurones artificiels, nés il y a environ une cinquantaine d'années, sont toujours en cours de développement. Cet axe de recherche a tiré l'attention de beaucoup de chercheurs de différentes disciplines. Ces réseaux de neurones artificiels (RNA) sont issus de la combinaison entre la neurobiologie comme idée de base, la physique comme champs d'application et les mathématiques avec l'informatique en tant que moyens de réalisation. Nous présentons dans ce qui suit un bref historique des réseaux de neurone artificiels.

1943: J. McCulloch & W. Pitts

- . Proposent un modèle simple de neurone capable de reproduire la machine de Turing,
- . Démontrent qu'un assemblage synchrone de tels neurones est une machine universelle de calcul (c'est-à-dire que toute fonction logique peut être représentée par des unités à seuil),

A logical Calculus of the Ideas Immanent in Nervous Activity,
Bull. of Math. Biophysics 5, 115 (1943).

1948: D. Hebb

- . Propose une règle d'apprentissage pour des réseaux de neurones.
- The Organization of Behaviour,
Wiley, New York 1949.

1958: F. Rosenblatt

- . Propose le modèle du perceptron et démontre son théorème de convergence.

1969: M. Minsky & S. Papert

- . Démontrent les limitations du modèle du perceptron.
- Perceptrons,
the MIT Press, Cambridge 1969.

1972 :T. Kohonen

- . Présente ses travaux sur les mémoires associatives.
- . Propose des applications à la reconnaissance de formes.

1982: J.J. Hopfield

- . Présente une théorie de fonctionnement et des possibilités des réseaux de neurones.
- Neural Networks and Physical Systems with Emergent Collective Computational Abilities, Proceedings of the National Academy of Sciences, USA, 1982, pp. 2,554.

1986: D.E. Rumelhart et al

- . Apprentissage par rétro-propagation pour les réseaux multi-couches.
- Learning representations by back-propagating errors, Nature, vol. 323 (1986).

II.2. Introduction

Le neurone biologique est une cellule vivante spécialisée dans le traitement des signaux électriques. Les neurones sont reliés entre eux par des liaisons appelées axones. Ces axones vont eux mêmes jouer un rôle important dans le comportement logique de l'ensemble. Ils transmettent les signaux électriques de la sortie d'un neurone vers l'entrée (synapse) d'un autre neurone. Les neurones font une sommation des signaux reçus en entrée et en fonction du résultat obtenu, vont fournir un courant en sortie.

Les réseaux de neurones biologiques réalisent un certain nombre d'applications telles que la reconnaissance de formes, le traitement du signal, l'apprentissage par l'exemple et la mémorisation. Ces applications sont, malgré tous les efforts déployés en algorithmique et en intelligence artificielle, à la limite des possibilités actuelles. C'est à partir de l'hypothèse que le comportement intelligent émerge de la structure et du comportement des éléments de base du cerveau que les réseaux de neurones artificiels, désigné dans ce travail par RNA, se sont développés. Il est nécessaire de mettre en œuvre des réseaux de neurones pour toute application nécessitant de trouver, par apprentissage, une relation non linéaire entre des données numériques. Le réseau de neurones artificiels est basé sur un modèle simplifié du neurone biologique. Les RNA sont donc des modèles et à ce titre, ils peuvent être décrits par leurs composantes, leurs variables descriptives et les interactions des composantes.

II.3. Eléments de neurophysiologie

II.3.1. Le neurone

Le neurone est une cellule composée d'un corps cellulaire et d'un noyau. Le corps cellulaire se ramifie pour former ce que l'on nomme les dendrites. C'est par les dendrites que l'information est acheminée de l'extérieur vers le soma, corps du neurone. L'information traitée par le neurone chemine ensuite le long de l'axone pour être transmise aux autres neurones. La transmission entre deux neurones n'est pas directe. Il existe un espace intercellulaire de quelques dizaines d'Angström (10^{-9} m) entre l'axone du neurone afférent et les dendrites du neurone efférent. La jonction entre deux neurones est appelée la synapse (Figure. II.3.1.b.).

Les réseaux de neurones sont utilisés pour résoudre des cas complexes en optimisation non linéaire. Les techniques utilisées sont basées sur le fonctionnement des neurones du cerveau. Le système nerveux est un système à trois étages (Figure: II.3.1.a).

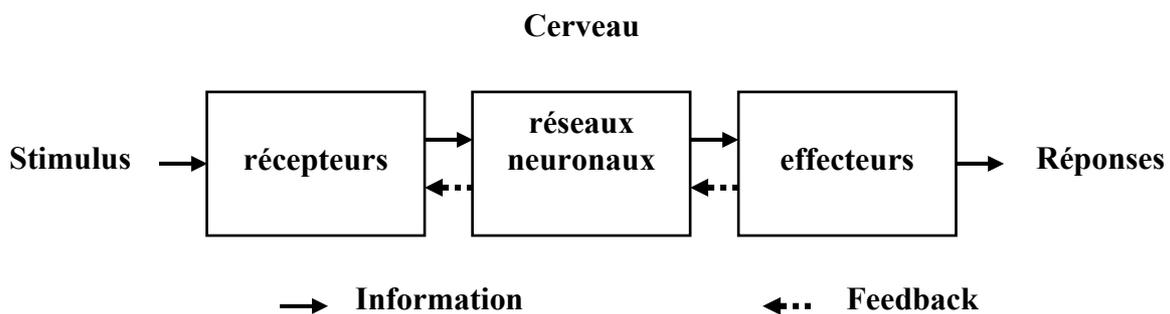


Figure II.3.1.a. Principe de fonctionnement du système nerveux

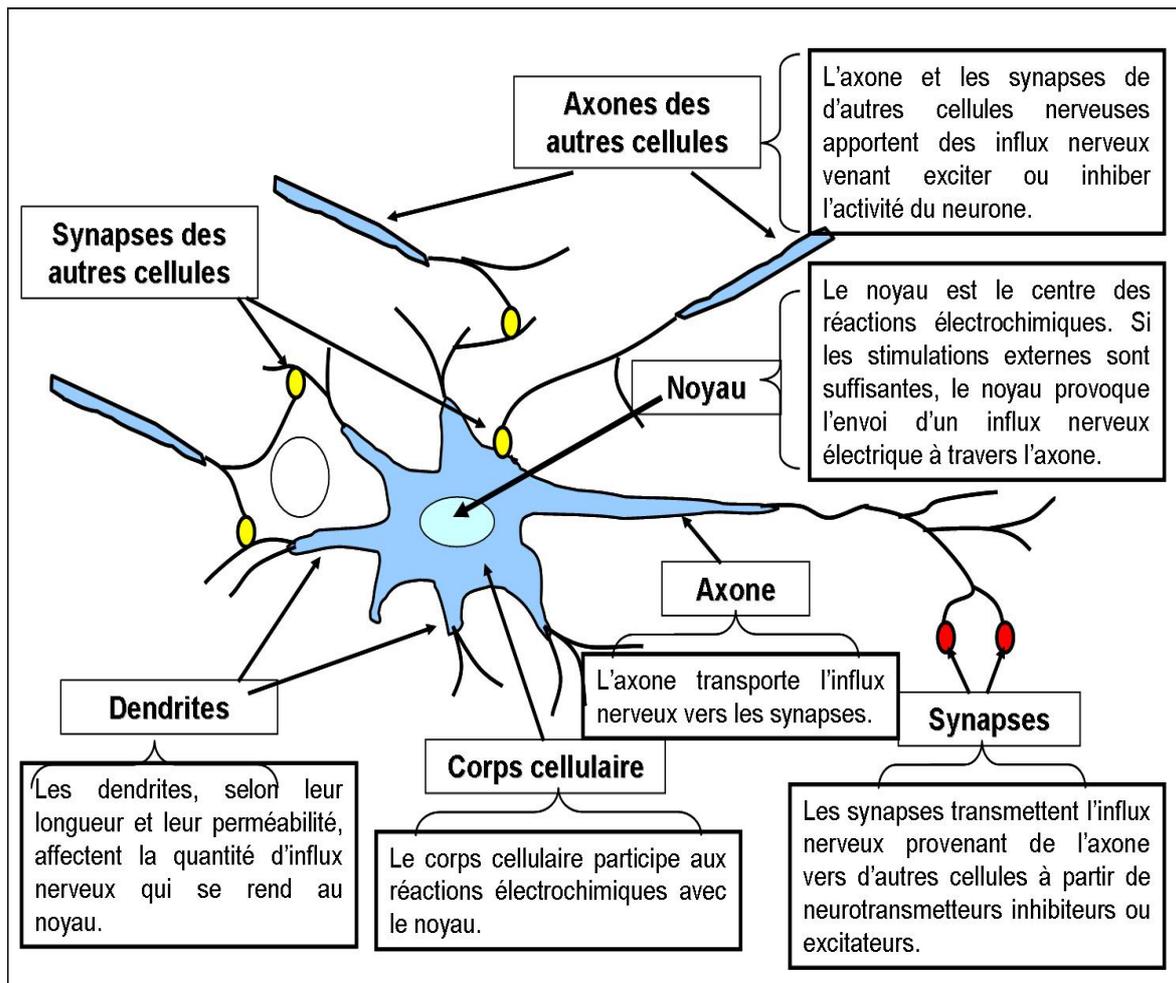


Figure II.3.1.b. Schéma synoptique simplifié d'un neurone biologique

II.3.2. Circuits neuronaux

Nous avons vu que chaque neurone est une unité autonome au sein du cerveau. Le neurone reçoit en continu des entrées. Le corps cellulaire du neurone, centre de contrôle, concentre les informations reçues pour ensuite être interprétées. La réponse, unique, à ces signaux est envoyée au travers de l'axone. L'axone transmet le message aux autres neurones. Le signal transmis peut avoir un effet excitateur ou inhibiteur. Le traitement très simple réalisé par chaque neurone indique que l'information n'est pas stockée dans les neurones, mais est plutôt le résultat du comportement de toute la structure interconnectée. L'information est, principalement, dans l'architecture des connexions et dans la force de ces connexions.

II.3.3. Activation d'un neurone

L'ensemble des signaux d'entrée (excités et inhibés) est moyenné. Si cette moyenne est suffisamment importante sur un laps de temps très court, la cellule délivre un signal électrique aux cellules suivantes par l'intermédiaire de son axone. Les impulsions délivrées (ou potentiel d'action) sont du type tout ou rien. Cela signifie que les neurones communiquent entre eux en langage binaire. Les réseaux neuronaux sont basés sur cette logique.

II.4. Les neurones artificiels

On désigne par "réseau de neurones" (neural network), ou réseau neuromimétique, un réseau de neurones artificiels basé sur un modèle simplifié du neurone. Ce modèle permet de réaliser certaines fonctions du cerveau. Parmi ces fonctions la mémorisation associative, l'apprentissage par l'exemple, le travail en parallèle. Cependant le neurone formel ne possède pas toutes les capacités des neurones biologiques (comme le partage de synapses, l'activation membranaire ou la structuration prénatale des neurones). Par conséquent les réseaux de neurones artificiels actuels sont loin d'intégrer toutes les possibilités du cerveau).

II.4.1. Le neurone formel

II.4.1.1. Définition

Le neurone artificiel s'inspire largement du neurone naturel. Il est donc constitué d'entrées et d'une sortie qui peut être dédoublée. La valeur de sortie est une fonction pondérée des valeurs d'entrée qui est modulée par une autre fonction généralement sigmoïde. On obtient donc un intégrateur dont la sortie se déclenche si la somme des entrées est supérieure à un seuil (Figure : II.4.1.1).

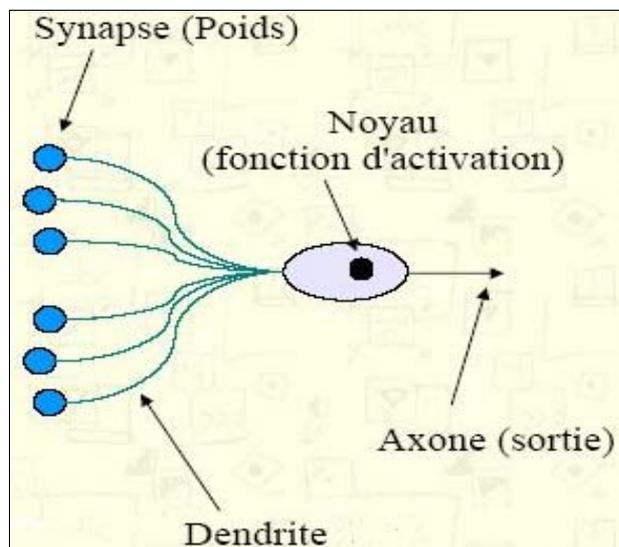


Figure : II.4.1.1. Modèle d'un neurone formel

II.4.1.2. Modélisation d'un neurone

Un neurone se compose généralement d'une entrée formée des variables sur lesquelles opère ce neurone et une sortie représentant la valeur de la fonction réalisée (fonction d'activation). La sortie du neurone est une fonction non linéaire d'une combinaison des entrées x_i (signaux d'entrées) pondérées par les paramètres w_i (poids synaptiques). Graphiquement le neurone est présenté sous la forme indiquée sur la figure (figure:II.4.1.2).

Le neurone est modélisé par deux opérateurs :

- Un opérateur de sommation qui élabore un potentiel égale à la somme pondérée des entrées de la cellule.
- Un opérateur calculant l'état de la sortie du neurone en fonction de son potentiel (fonction d'activation).

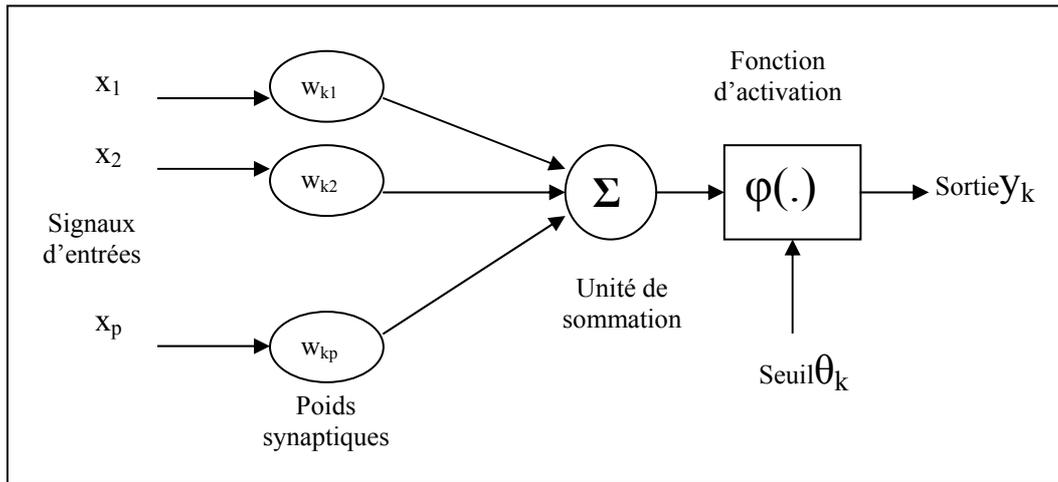


Figure:II.4.1.2. Représentation graphique d'un neurone

Mathématiquement le neurone est une fonction algébrique non linéaire, paramétrée, à valeurs bornées[22]. La sortie du neurone est donnée par l'expression suivante :

$$y_k = \varphi \left(\sum_{j=1}^p w_{kj} \cdot x_j - \theta_k \right) \quad (II.1)$$

- où : $x_1, x_2, x_3, \dots, x_p$: entrées,
- $w_{k1}, w_{k2}, w_{k3}, \dots, w_{kp}$: Poids synaptiques du neurone k ,
- y_k : est la sortie d'activation,
- θ_k : Le Seuil
- $\varphi(.)$: La fonction d'activation,

II.4.1.3. Comparaison entre le neurone biologique et le neurone artificiel

Chaque neurone artificiel est un processeur élémentaire. Il reçoit un nombre variable d'entrées en provenance de neurones en amont ou des capteurs composant la machine dont il fait partie. A chacune de ces entrées est associée un poids représentatif de la force de la connexion. Chaque processeur élémentaire est doté d'une sortie unique, qui se ramifie ensuite pour alimenter un nombre variable de neurones en aval. A chaque connexion est associé un poids. Tout d'abord on additionne les entrées avec leurs poids respectifs, puis une fonction de transfert calcule la valeur de l'état du neurone (Figure : II.4.1.3)..

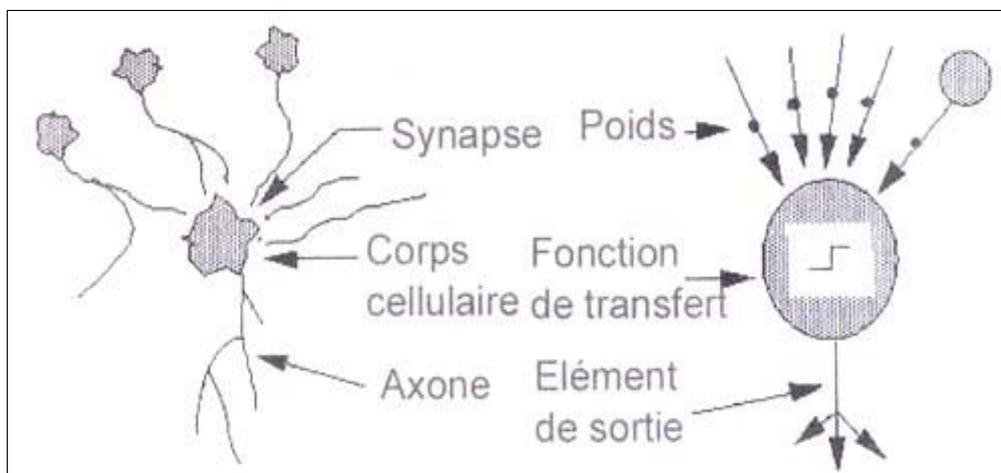


Figure : II.4.1.3. Comparaison entre le neurone biologique et le neurone artificiel

II.4.1.4. La fonction d'activation

La fonction d'activation définit la valeur de la sortie d'un neurone en termes des niveaux d'activité de ses entrées.

II.4.1.5. Types de fonctions d'activations :

La fonction d'activation considère différentes formes selon l'utilisation du réseau et selon sa nature, continue ou discrète. La plupart des neurones formels utilisés actuellement sont des variantes du neurone de McCulloch et Pitts dans lesquels la fonction de Heaviside est remplacée par une autre fonction d'activation. Les fonctions les plus utilisées sont :

- la fonction sigmoïde ;
- la fonction tangente hyperbolique ;
- fonction à seuil (fonction échelon) ;
- fonctions linéaires par parties.

Ces choix sont motivés par des considérations théoriques et pratiques. Les propriétés de la fonction d'activation influent en effet sur celles du neurone formel et il est donc important de bien choisir celle-ci pour obtenir un modèle utile en pratique.

Quand les neurones sont combinés en un réseau de neurones formels, il est important par exemple que la fonction d'activation de certains d'entre eux ne soit pas un polynôme sous réserve de limiter la puissance de calcul du réseau obtenu. Un cas caricatural de puissance limitée correspond à l'utilisation d'une fonction d'activation linéaire, comme la fonction identité : dans une telle situation le calcul global réalisé par le réseau est lui aussi linéaire et il est donc parfaitement inutile d'utiliser plusieurs neurones, un seul donnant des résultats strictement équivalents.

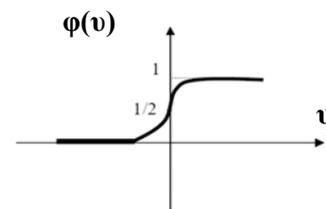
Cependant, les fonctions de type sigmoïde sont généralement bornées. Dans certaines applications, il est important que les sorties du réseau de neurones ne soient pas limitées a priori : certains neurones du réseau doivent alors utiliser une fonction d'activation non bornée. On choisit généralement la fonction identité.

Il est aussi utile en pratique que la fonction d'activation présente une certaine forme de régularité. Pour calculer le gradient de l'erreur commise par un réseau de neurones, lors de son apprentissage, il faut que la fonction d'activation soit dérivable. Pour calculer la matrice de l'erreur, ce qui est utile pour certaines analyses d'erreur, il faut que la fonction d'activation soit dérivable deux fois. Comme elles comportent généralement des points singuliers, les fonctions linéaires par morceaux sont relativement peu utilisées en pratique.

La fonction sigmoïde

La fonction sigmoïde, définie par la formule (II.2) et représentée par le graphe (Graphe : II.1) possède les propriétés importantes évoquées précédemment (elle n'est pas polynomiale et est indéfiniment continûment dérivable).

$$\varphi(v) = \frac{1}{1 + e^{-v}} \quad (\text{II.2})$$



Graphe: II.1

En outre, une propriété simple permet d'accélérer le calcul de sa dérivée, ce qui réduit le temps calcul nécessaire à l'apprentissage d'un réseau de neurones. On a en effet

$$\frac{d\varphi(v)}{dv} = \varphi(v)(1 - \varphi(v)) \quad (\text{II.3})$$

On peut donc calculer la dérivée de cette fonction en un point de façon très efficace à partir de sa valeur en ce point. De plus, la fonction sigmoïde est à valeurs dans l'intervalle $[0,1]$, ce qui permet d'interpréter la sortie du neurone comme une probabilité. Elle est aussi liée au modèle de régression logistique et apparaît naturellement quand on considère le problème de la séparation optimale de deux classes de distributions gaussiennes avec la même matrice de covariance.

La fonction tangente hyperbolique

La fonction tangente hyperbolique, définie par :

$$\varphi(v) = \frac{e^v - e^{-v}}{e^v + e^{-v}} \quad (\text{II.4})$$

est aussi très utilisée en pratique, car elle partage avec la fonction sigmoïde certaines caractéristiques pratiques :

- non polynomiale
- indéfiniment continûment dérivable
- calcul rapide de la dérivée par la formule

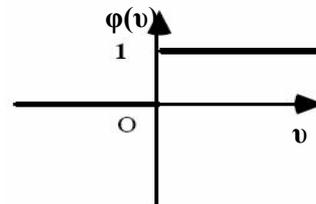
$$\frac{d\varphi(v)}{dv} = 1 - (\varphi(v))^2 \quad (\text{II.5})$$

On ne peut pas cependant lui donner une interprétation probabiliste assez claire.

La fonction à seuil

Certains auteurs utilisent les neurones à sortie binaire. La fonction d'activation est alors la fonction signe (-1 ou +1) ou la fonction échelon (0 ou +1) (Graphe: II.2).

$$\varphi(v_k) = \begin{cases} 1 & \text{si } v_k \geq 0 \\ 0 & \text{si } v_k < 0 \end{cases} \quad (\text{II.6})$$

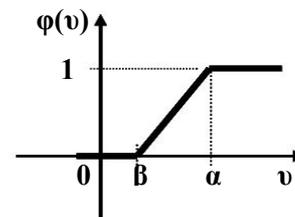


Graphe: II.2

La fonction linéaire par parties

Ses modèles peuvent être décrits par une partie linéaire à pente constante et une valeur de saturation (Graphe : II.3)

$$\varphi(v) = \begin{cases} 1 & \text{si } v \geq \alpha \\ v & \text{si } \alpha > v > \beta \\ 0 & \text{si } v \leq \beta \end{cases} \quad (\text{II.7})$$



Graphe: II.3

II.4.2. Le réseau de neurones artificiels

Un neurone réalise une fonction non linéaire. L'intérêt des neurones réside dans les propriétés qui résultent de leurs associations en réseaux. c'est-à-dire de la composition des fonctions non linéaires réalisées par chacun des neurones constitutifs.

La plupart des modèles de réseaux de neurones artificiels (RNA) ne retirent du fonctionnement réel que les principes suivants :

- ceux caractérisés par des interconnexions denses entre des unités de traitement simples agissant en parallèle.

- Ceux qu'à chaque connexion est associé un poids qui détermine l'influence réciproque des deux unités connectées.
 - Ceux pour lesquels les poids de connexions sont modifiables et c'est cette plasticité qui donne lieu aux facultés d'adaptation et d'apprentissage.
- Les neurones sont donc fortement connectés entre eux. Mais ces connexions ne sont pas aléatoires. Elles correspondent à des réseaux dont les architectures sont assez connues.

II.4.2.1. Connectivité

La connectivité des réseaux, c'est à dire la manière dont ils sont reliés, peut être totale (tous les neurones connectés entre eux) ou par couche (les neurones d'une couche ne sont connectés qu'aux couches suivantes) - il existe des réseaux monocouches ou multicouches (*multi layer networks*). Les connexions entre les neurones qui composent le réseau décrivent la topologie du modèle.

Selon le type d'interconnexion on peut classer les réseaux comme réseaux à connexion locale (partielle) ou totale (complète).

a. Réseaux à connexions locales (figure : II.4.2.1.a.) : Chaque neurone entretient des relations avec un nombre réduit et localisé de neurones de la couche en aval.

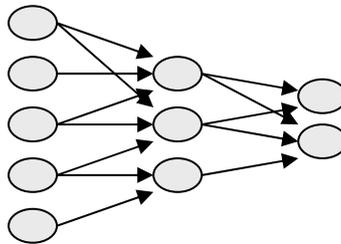


Figure: II.4.2.1.a. Réseau à connexions locales

b. Réseaux à connexions totales (figure : II.4.2.1.b.) : Chaque neurone est connecté à tous les neurones du réseau.

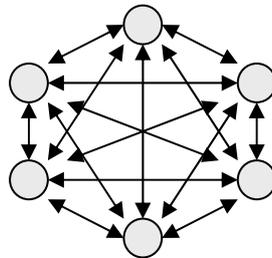


Figure: II.4.2.1.b. Réseau à connexions totales

II.4.2.2. Apprentissage

Une caractéristique des réseaux de neurones est leur capacité à apprendre (par exemple à reconnaître une lettre, un son...). Mais cette connaissance n'est pas acquise dès le départ. La plupart des réseaux de neurones apprennent par l'exemple.

On appelle « phase d'apprentissage » des réseaux de neurones la procédure qui consiste à estimer les paramètres des neurones du réseau. afin que celui-ci remplisse au mieux la tâche qui lui est affectée. La phase de l'apprentissage est une phase du développement d'un réseau de neurones durant laquelle le comportement du réseau est modifié jusqu'à l'obtention du comportement désiré.

II.4.2.2.a. Types d'apprentissage

Selon le critère d'apprentissage les réseaux de neurones se subdivisent en deux classes, les réseaux à apprentissage supervisé (Supervised learning) et les réseaux à apprentissage non supervisé (Unsupervised learning).

- Pour les réseaux à apprentissage supervisé, on présente au réseau des entrées et au même temps les sorties que l'on désirerait pour cette entrée. Le réseau doit alors se reconfigurer. C'est-à-dire calculer ses poids afin que la sortie qu'il donne corresponde bien à la sortie désirée.

- Pour les réseaux à apprentissage non supervisé, on présente une entrée au réseau et on le laisse évoluer librement jusqu'à ce qu'il se stabilise.

Il existe aussi des réseaux à apprentissage dit semi supervisé (reinforcement learning) qui ne tiennent compte que d'une évaluation partielle ou qualitative des sorties.

II.4.2.2.b. Application des réseaux de neurones selon le mode d'apprentissage

Les réseaux à apprentissage supervisé servent à :

- La modélisation statique des processus non linéaires décrits par des équations algébriques.
- La modélisation dynamique des processus décrits par des équations récurrentes.
- La commande de processus. Ils servent aussi à imposer une dynamique de réponse à une sollicitation.

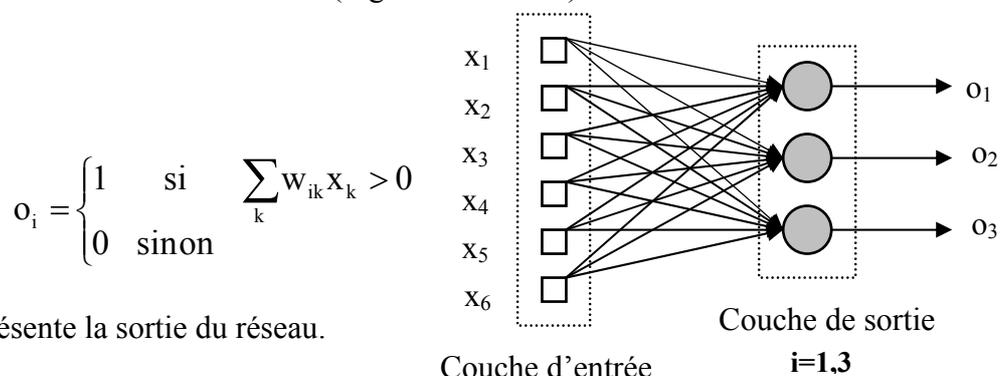
Les réseaux à apprentissage non supervisé permettent l'analyse, la classification et la visualisation des données.

II.4.2.3. Topologie

Si un neurone isolé a un comportement prévisible, la complexité d'un réseau naît de ses nombreuses possibilités d'interconnexion. L'adaptation d'un réseau à un problème donné passe par un choix de la topologie et des poids de liaison entre les neurones.

La topologie des réseaux de neurones peut être très variée. On peut concevoir plusieurs types de réseaux seulement en modifiant les règles de connexion[22].

a. Réseau à une couche (Figure : II.4.2.3.a) :



o_i représente la sortie du réseau.

Figure:II.4.2.3.a. Exemple d'un réseau à une couche

b. Réseaux multicouches ``feedforward`` (Figure:II.4.2.3.b.):

Notons les couches cachées par la lettre 'c' et les couches de sorties par la lettre 'o'. On appelle :

- Couche d'entrée : l'ensemble des neurones d'entrée,
- Couche de sortie : l'ensemble des neurones de sortie.
- Couches cachées : l'ensemble des couches intermédiaires, elles n'ont aucun contact avec l'extérieur.

Couches cachées :

$$c_j = \begin{cases} 1 & \text{si } \sum_k w_{jk} X_k > 0 \\ 0 & \text{sinon} \end{cases}$$

avec : $j=1, \dots, n_{ne}$

Couches de sortie :

$$o_i = \begin{cases} 1 & \text{si } \sum_k w_{ik} c_k > 0 \\ 0 & \text{sinon} \end{cases}$$

avec : $i=1, \dots, n_{ns}$

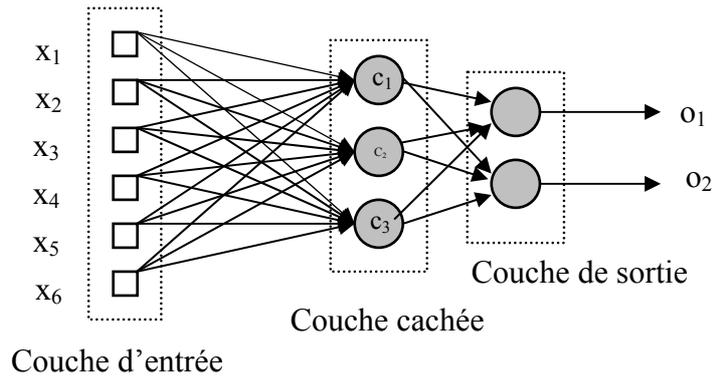


Figure:II.4.2.3.b. Exemple d'un réseau multicouches "feedforward"

(**ne**) indique les neurones de la couche d'entrée et (**ns**) indique les neurones de la couche de sortie

c. Réseaux récurrents (Figure : II.4.2.3.c.) : L'idée des connexions récurrentes est que le réseau est capable de « se rappeler » des valeurs des états précédents par l'intermédiaire de leurs poids synaptiques.

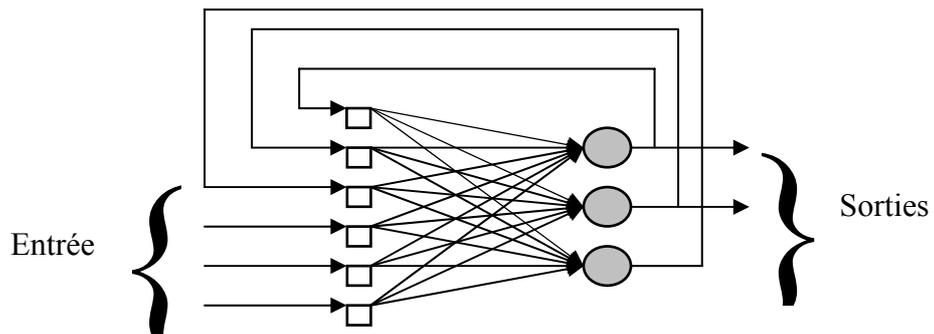


Figure:II.4.2.3.c. Exemple d'un réseau récurrent

La sortie du réseau à un instant donné est une fonction non linéaire de l'entrée externe et de la sortie des neurones cachées à l'instant précédent. Chaque unité i est connectée à chaque autre unité j par un poids w_{ij} . Les poids sont supposés symétriques : $w_{ij}=w_{ji}$

d. Réseaux en treillis : (Figure:II.4.2.3.d) Réseaux en treillis 3X3 bidimensionnel.

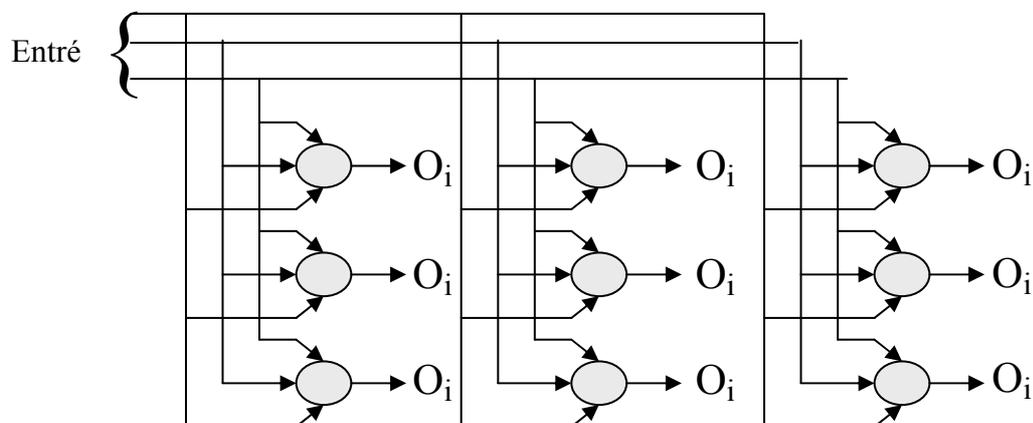


Figure:II.4.2.3.d. Exemple d'un Réseau en treillis

II.5. Les types de réseaux.

Par définition, un réseau de neurones est constitué d'un assemblage d'éléments, d'unités ou de noeuds processeurs pour lequel un sous-groupe effectue un traitement indépendant et transmet le résultat à un deuxième sous-groupe et ainsi de suite (cas d'un réseau à couches multiples). Les capacités de traitement du réseau dépendent des poids w_i auxquels sont affectées des valeurs produisant un filtre affectant la capacité d'apprentissage du réseau.

Dans un réseau de neurones, les neurones sont regroupés en couches. Habituellement, chaque neurone dans une couche est connecté à tous les neurones dans la couche précédente et la couche suivante (excepté dans la couche d'entrée et celle de sortie du réseau).

L'information donnée à un réseau de neurones est propagée couche par couche de la couche d'entrée à la couche de sortie en passant par une ou plusieurs couches intermédiaires (couches cachées).

II.5.1. Le perceptron

Le perceptron est un réseau de neurones très simple. Il est linéaire et monocouche, acceptant uniquement des valeurs d'entrées et de sorties binaires (0 ou 1). Il est inventé par le psychologue Frank Rosenblatt et inspiré du système visuel. Le procédé d'apprentissage est supervisé et le réseau est capable de résoudre des opérations logiques simples comme les opérations "ET logique" ou "OU logique". Ces réseaux sont efficaces pourvu que les données traitées soient des vecteurs orthogonaux ou linéairement indépendants. Ce que l'on entend par méthode d'apprentissage supervisé est que les sorties du réseau sont comparées avec des sorties cibles. Ce type de réseau est aussi utilisé pour la classification. Les opérations logiques plus compliquées comme un problème "XOR" (OU Exclusif) ne peuvent pas être résolues par un perceptron. L'algorithme d'apprentissage est la règle d'apprentissage de Hebb qui affecte le changement des poids en multipliant l'entrée d'un neurone par sa sortie et le taux d'apprentissage du réseau[22].

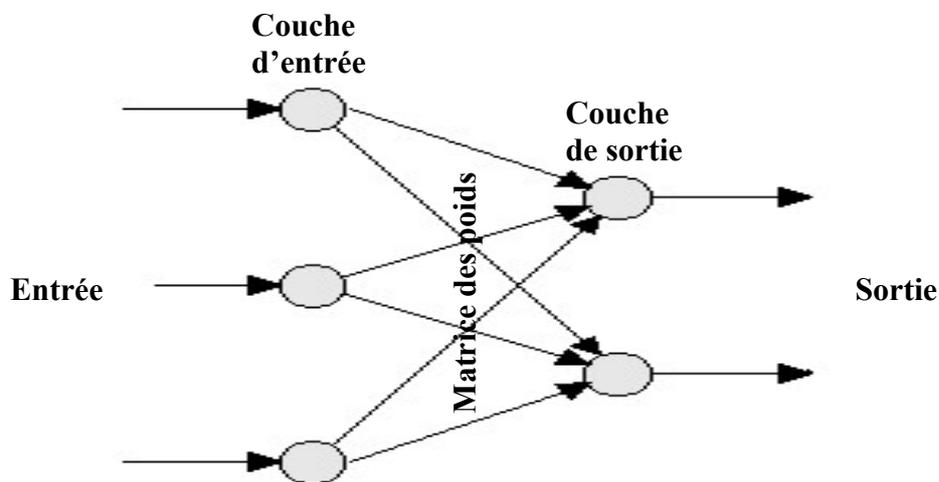


Figure:II.5.1. Architecture d'un perceptron

Un réseau "feedforward" à une seule couche (perceptron) peut uniquement représenter des fonctions linéairement séparables. C'est-à-dire celles pour lesquelles la surface de décision séparant les cas positifs des cas négatifs est un (hyper-)plan.

La sortie du perceptron dépend de la somme des composantes x_i du vecteur d'entrée, pondérées par des poids réels w_i . Conformément aux notations utilisées dans ce travail, nous appellerons un potentiel, la somme pondérée notée par 'net', suivante.

$$\text{net}_k = \sum_{i=0}^n w_{ki} X_i \quad (\text{II.8})$$

La sortie du réseau sera désignée par 'O'. Elle sera de la forme suivante :

$$O_k = \begin{cases} 1 & \text{si } \text{net}_k \geq \theta_k \\ 0 & \text{si } \text{net}_k < \theta_k \end{cases} \quad (\text{II.9})$$

w_{ji} : Coefficients synaptiques de j vers i, x_i : sortie(ou état) de j , et θ_k : seuil

II.5.2. Perceptron multicouches

Les réseaux multicouches sont actuellement les plus employés. Plusieurs couches de traitement leur permettent de réaliser des associations non linéaires entre l'entrée et la sortie.

Le perceptron à niveaux multiples (perceptron multi couche'PMC' ou 'MLP' en anglais) est une extension du perceptron qui a une ou plusieurs sous-couches cachées entre ses couches d'entrée et de sortie. A cause de sa structure étendue, un perceptron à niveaux multiples est capable de résoudre toute opération logique incluant le problème XOR. La méthode d'apprentissage est supervisée. L'algorithme d'apprentissage est la règle d'apprentissage delta (Δ) qui affecte le changement des poids en multipliant l'entrée d'un neurone par la différence entre la sortie réelle et la sortie désirée et le taux d'apprentissage du réseau.

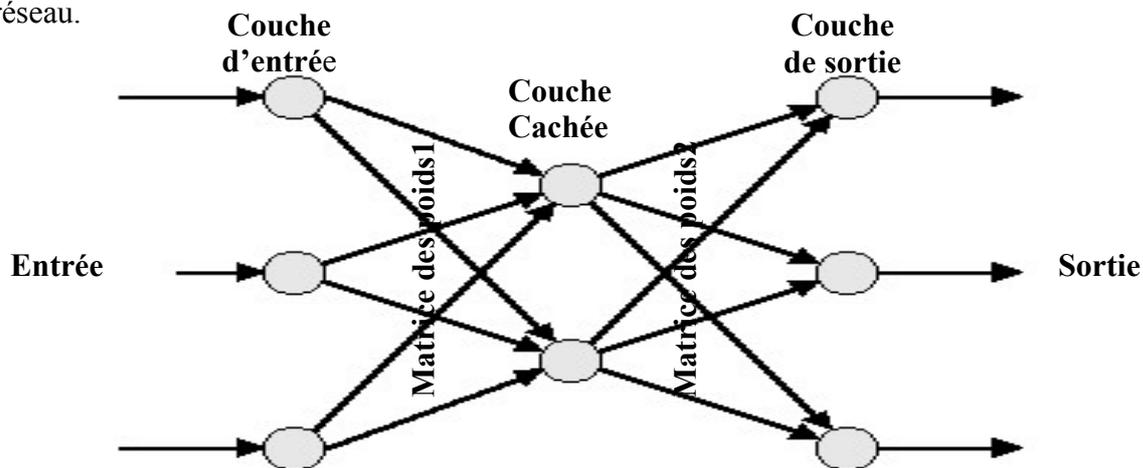


Figure:II.5.2. Architecture du perceptron Multicouche

Le perceptron a les caractéristiques suivantes :

- Le perceptron multicouches est une suite de couches contenant chacune des neurones formels.
- Les entrées de la couche n sont les sorties de la couche n-1.
- Les neurones de la première couche sont reliés à l'extérieur et reçoivent tous la même entrée (vectorielle).
- La sortie du perceptron multicouches est un vecteur constitué des sorties des neurones de la dernière couche.

Apprentissage du perceptron multicouches

Pour la couche de sortie, on peut appliquer l'apprentissage du perceptron, mais comment modifier les poids pour les connexions qui ne sont pas en relation avec un neurone de sortie? Le problème est l'obtention d'une estimation de la valeur désirée pour chaque neurone de la couche cachée. Ils utilisent pour modifier leurs poids, un algorithme de rétropropagation du gradient (backpropagation). Il s'agit toujours de minimiser l'erreur quadratique.

II.5.3. le perceptron multicouche par l'algorithme de rétro-propagation

Ce type de réseau est l'un des plus puissants réseaux de neurones (Figure:II.5.3). Il a la même structure que le perceptron à niveaux multiples et utilise l'algorithme de rétro-propagation (back propagation). La méthode d'apprentissage est sous contrôle supervisé. On utilise fréquemment des neurones avec une caractéristique entrée-sortie de type sigmoïde.

Les perceptrons multicouche agissent comme un séparateur non linéaire et peuvent être utilisés pour la classification, le traitement de l'image ou l'aide à la décision.

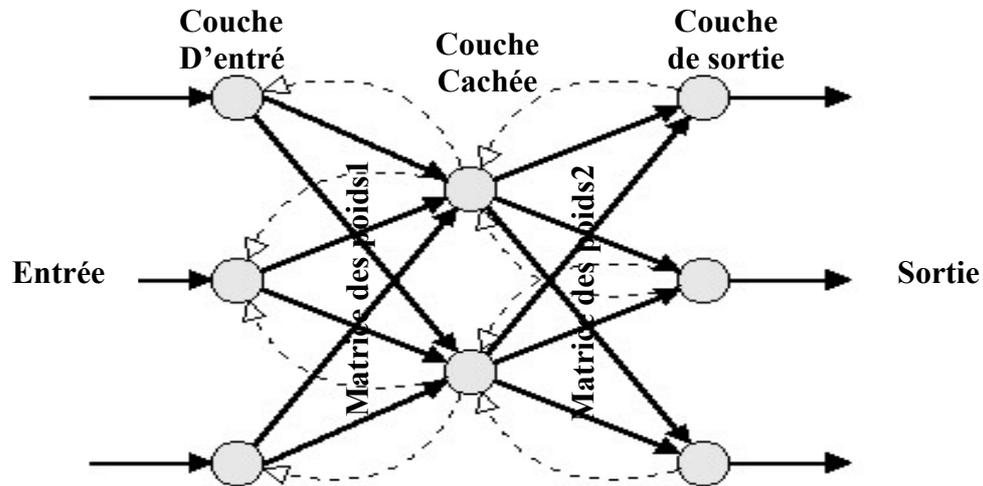


Figure:II.5.3. Architecture d'un modèle de Rétropropagation

Le développement de la rétropropagation et son utilisation pour la classification des données sismiques font l'objet du chapitre IV.

II.5.4. Le réseau de Hopfield

Pour présenter les spécificités de ces réseaux, il est intéressant de comparer les divers aspects du formalisme et de la terminologie entre le traitement de signal et la physique. Là où le traitement de signal cherche à minimiser les fonctions de corrélation, la physique cherche à minimiser les fonctions d'énergie. Si le traitement du signal analyse la convergence d'un système, la physique étudie le phénomène de relaxation. Dans les deux cas, les matrices ont les mêmes valeurs propres et les mêmes vecteurs propres. La différence est plus subtile : en plus de sa terminologie propre, la physique a apporté toute une classe de problèmes liés à l'optimisation, ainsi que des méthodes d'approche de ces problèmes [18].

Ce modèle a été développé en 1982 selon une analogie formelle avec des réseaux de particules orientées et en interaction. Sa propriété fondamentale réside dans le fait qu'il existe une grandeur homologue à une énergie qui décroît au cours des itérations.

Ce type de réseau est un ensemble de neurones, où chacun de ces derniers est connecté à chaque autre neurone. Il n'y a aucune différenciation entre les neurones d'entrées et de sorties.

La loi d'apprentissage est la règle de Hebb, qui stipule qu'une synapse améliore son activité si et seulement si l'activité de ses deux neurones est corrélée d'une autre façon cela signifie que le poids w_{ij} d'une connexion entre un neurone i et un neurone j augmente quand les deux neurones sont activés en même temps (Figure:II.5.4.).

L'application principale d'un réseau Hopfield est la collecte des données, la reconnaissance des formes et l'optimisation des images.

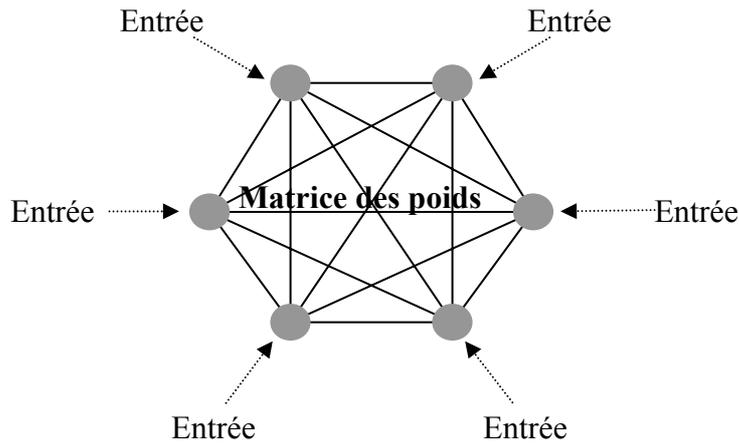


Figure:II.5.4. Architecture d'un model de Hopfield

II.5.5. Le réseau de Kohonen

Les phénomènes d'auto-organisation sont bien connus dans le système nerveux des vertébrés, particulièrement au niveau cortical. Citons a titre d'exemple le cortex visuel qui est constitué de feuilletts dits d'orientation, c'est-à-dire des régions qui répondent préférentiellement à des stimulations visuelles (spots lumineux) d'orientation différentes. De plus, l'activité des régions corticales voisines est associée à des orientations voisines des stimulations rétiniennes (rétinotopie). un phénomène similaire est observé au niveau du cortex auditif : deux stimulations de fréquences voisines activent des régions corticales voisines. L'organisation corticale est non supervisée[18]. Elle est pilotée par la structure des stimulations extérieures.

Des modèles d'auto-organisation, inspirés de l'organisation corticale des vertébrés, ont été proposés des les années 70. Au début ces modèles ont pour objectif de représenter des données complexes dans un espace discret dont la topologie est limitée à un, deux, voire trois dimensions. Il s'agit donc de modèles de quantification vectorielle qui seront dotés de propriétés topologiques particulières. Le type le plus connu des cartes auto-organisatrice est celui de Kohonen. C'est probablement le type de réseau de neurones le plus utilisé quant à sa ressemblance simulé du processus d'apprentissage du cerveau humain. Le coeur de ce type de réseau est une organisation applicative où les neurones s'organisent eux-mêmes par rapport à certaines données se trouvant à l'entrée (Figure : II.5.5.). Le chapitre III sera consacré au développement de ce type de réseau. Il fait l'objet de l'une des applications des réseaux de neurones dans l'exploration pétrolière et exactement dans le domaine du traitement des données sismiques.

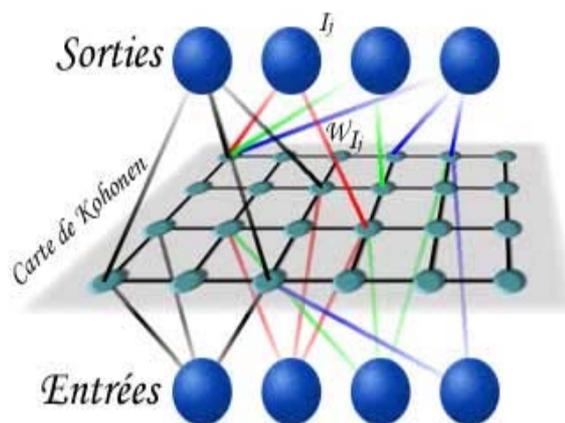


Figure:II.5.5. Architecture d'un model de Kohonen

II.6. Mise en œuvre et Application des réseaux de neurones artificiels

Les réseaux de neurones réalisent des fonctions non linéaires paramétrées; ils peuvent approcher toute fonction de régression si l'on trouve un réseau de neurones qui rend minimum la fonction de coût (somme des carrés des différences entre les valeurs calculées par le réseau de neurones et les valeurs mesurées).

La mise en œuvre d'un réseau de neurones nécessite :

- La détermination des entrées pertinentes: les grandeurs qui ont une influence significative sur le phénomène que l'on cherche à modéliser.
- La collection des données nécessaires à l'apprentissage et à l'évaluation des performances du réseau de neurones.
- La détermination du nombre de couches cachées nécessaire pour obtenir une approximation satisfaisante.
- L'estimation des valeurs des paramètres correspondant à un minimum de la fonction de coût, c'est-à-dire effectuer un apprentissage.
- L'évaluation des performances du réseau de neurones à l'issue de l'apprentissage.

À l'intersection des différents domaines (informatique, électronique, science de la nature et neurobiologie, etc...), l'étude des réseaux de neurones est une voie prometteuse de l'intelligence artificielle. Cet outil a des applications dans plusieurs domaines.

Les réseaux de neurones ont été développés pour résoudre des problèmes de contrôle, de reconnaissance de formes, de classification et de décision. Les caractéristiques et les performances des réseaux de neurones artificiels ont fait de cette technique une solution idéale pour beaucoup de problèmes.

Généralement, on peut classer les domaines d'application des réseaux de neurones comme suit.

a. Outil de classification et de reconnaissance.

Reconnaissance de formes, classification des signaux, vérification des signaux, reconnaissance des objets et contrôle de qualité des documents.

b. Traitement de signal.

En se basant sur la performance des réseaux de neurones dans l'amélioration significative des systèmes de traitement des signaux, on peut les utiliser pour le filtrage des bruits, dans la classification des signaux et pour la compression des images.

c. Modélisation prédictive.

A titre d'exemple on peut citer deux champs d'application : l'Analyse financière des bourses et la prédiction du climat.

d. Analyse des données.

C'est un outil très puissant pour la classification et la généralisation des informations dans un volume de données important (exp: les services des assurances, les services des renseignements et la recherche de nouvelles caractéristiques et relations entre des différents événements).

e. Application dans le secteur pétrolier.

Cette technique est implantée dans plusieurs disciplines, les plus connues :

- Le traitement des données sismiques (Veezhinatan 1990, McCormack 1991, Musuma 1992).
- La caractérisation des roches (Derek 1990, Braunschweig 1990).
- Interprétation des logs de diagraphie (Baldwin 1989-1993, Avelino 1991, Riva, 1992).
- Diagnostique de l'outil de forage (Arehart 1989).
- Control des processus (Lambert 1991).

III. Les cartes auto-organisatrices de Kohonen (K-SOM)

III.1. Introduction

Dans ce chapitre nous développons le principe des cartes auto-organisatrices de Kohonen, leurs fondements théoriques et leur méthodologie d'application dans le domaine de traitement des données sismiques. Notre application a comme objet une classification lithologique en utilisant les attributs sismiques comme données d'entrée et les cartes de Kohonen comme outil de classification.

L'efficacité de l'utilisation des " attributs sismiques " dans l'interprétation dépend des capacités de discrimination de l'ensemble des attributs choisis. Comme dans beaucoup de problèmes mathématiques, quelques attributs peuvent être nécessaires et suffisants. Ceci peut être déterminé par expérience avec des combinaisons logiques des divers types d'attributs. Dans les cas où les données de puits sont disponibles, des réseaux de neurones artificiels (ANN) peuvent être formés d'une façon supervisée (apprentissage supervisé) pour réaliser une telle combinaison. Dans les cas où aucune information n'est disponible, nous utilisons l'approche non supervisée (apprentissage non supervisé).

Dans le cas où les données sont représentées par une suite d'attributs, l'utilisateur en choisit un ensemble à utiliser dans la classification. Il existe plusieurs méthodes non supervisées pour faire ce travail. Une des approches la plus simple, regroupe les données séquentiellement. Par exemple, si nous supposons que nous avons un ensemble de traces sismiques, chaque échantillon est représenté par un vecteur d'attributs. Ceci peut être observé comme des données représentées dans un espace à N dimensions, où chaque attribut représente l'un des axes. Nous calculons d'abord une moyenne simple de tous les attributs; puis nous assemblons (regroupement) tous les échantillons de données sur une certaine distance de cette valeur moyenne. Après avoir contrôlé tous les échantillons, nous pouvons identifier les échantillons sur une distance bien définie comme des membres du premier faisceau (groupe) et les séparer de tout autre calcul. Ensuite on calcule la moyenne des échantillons rejetés, qui représente la moyenne du deuxième ensemble et nous continuons de grouper de cette manière jusqu'à ce qu'il ne reste aucun échantillon significatif.

Le problème avec ce procédé, bien qu'il puisse être rapide, est que tous les faisceaux (groupe ou classe) sont arbitraires et qu'ils ne présentent pas des relations raisonnables entre les groupes. Cette caractéristique représente un aspect très important dans la méthode de Kohonen. Les classes, produites par le réseau de Kohonen, sont formées de telle manière qu'elles soient reliées entre elles (l'une par rapport à l'autre) d'une façon organisée. Par exemple, les classes du sable, du sable-argileux, et des argiles peuvent occuper des régions adjacentes sur la carte. C'est une analogie au cerveau humain (mémoire associative).

Nous pouvons expliquer cette méthode non linéaire par l'analogie avec une autre méthode linéaire. C'est une autre approche plus classique, et elle est basée sur le développement des statistiques des données, en formant la matrice de covariance et en calculant les composantes principales. Ceci fournit des valeurs propres et des vecteurs propres classés des plus importants vers les moins significatives. Le vecteur propre correspondant à la valeur propre la plus importante représente la combinaison des attributs qui est la plus significative dans l'ensemble des données. La prochaine valeur propre représentera la prochaine classe la plus significative et ainsi de suite. Après cette étape on aborde la classification des données en calculant la distance entre l'ensemble des données et tous les vecteurs propres et en choisissant le vecteur propre qui a la distance euclidienne minimale. Cette méthode formera des classes basées sur la population et ne peut avoir aucun rapport avec des valeurs et des vecteurs adjacents.

III.2. Principe des cartes auto-organisatrices de Kohonen

Les cartes auto-organisatrices de Kohonen (Kohonen Self Organizing Map, K-SOM) sont tirées par analogie de la capacité dont le cerveau humain organise l'information d'une façon logique. La recherche a démontré que le cortex cérébral du cerveau humain est divisé en subdivisions fonctionnelles. Le cortex contient des milliards de neurones et plusieurs de milliards de raccords (synapses) entre eux. La subdivision du cortex est ordonnée. Les composants les plus importants sont : le cortex sensoriel, le cortex visuel et le cortex auditif. Le cortex auditif est également subdivisé en de nombreuses cellules, chacune fonctionne comme partie spéciale pour chaque signal auditif. Certaines de ces cellules ont un apprentissage supervisé, tandis que d'autres sont développées d'une façon non supervisée. Selon leur Topologie, les secteurs adjacents exécutent des fonctions relativement reliées. La méthode de Kohonen utilise un apprentissage non supervisé d'une façon élégante et extrêmement simple.

Les cartes auto-organisatrices présentent un exemple des méthodes à apprentissage non supervisé. Elles analysent les données sismiques et extraient l'information à partir de ces dernières en utilisant un algorithme de classification sophistiqué. Cette technique utilise uniquement les données sismiques ou des attributs dérivés de ces données, et n'exige pas une estimation a priori sur les données d'entrée.

Les cartes auto-organisatrices ont été introduites pour la première fois par Teuvo Kohonen en 1981. C'est un réseau de neurones artificiel dont l'organisation topologique de ses neurones présente sa principale caractéristique.

Les applications visées par les cartes auto-organisatrices de Kohonen devaient pouvoir concerner de très grands ensembles de données, pour lesquelles les observations traitées pouvaient atteindre de grandes dimensions. La visualisation par cartes topologiques proposées par Kohonen cherche, par apprentissage à partir des données, à partitionner l'ensemble des observations disponibles en groupement similaires. Les groupements proposés possèdent la particularité caractéristique d'avoir une structure de voisinage qui peut être matérialisée à l'aide d'un espace discret que l'on appelle carte topologique.

La carte auto-organisatrice est habituellement composée d'une grille de neurones à deux dimensions. Parfois des cartes à une dimension peuvent être employées pour résoudre certains problèmes. Cependant, des cartes multidimensionnelles peuvent être aussi bien mises en application. Une carte auto-organisatrice offre la possibilité d'effectuer une réduction de dimensions, car elle permet de projeter des données multidimensionnelles sur une carte à deux dimensions. Chacun des neurones composants cette carte sera relié à chaque neurone d'entrée (il y a autant de neurones d'entrée que de vecteurs composant les données d'entrée).

III.3. Classification et cartes auto-organisatrices

Les tâches de classifications sont les plus importantes parmi les différentes applications réalisées par les cartes auto-organisatrices. Le résultat d'un apprentissage non supervisé permet d'affecter une observation à un sous-ensemble d'une partition, indépendamment de toute notion de classe. On suppose que l'on dispose d'un très grand nombre de données bruitées (observations) dont on ne connaît pas exactement la classe. L'idée générale de ce type d'application est de faire apparaître des groupements cohérents (classes) en relation avec les données traitées. Ces groupements sont appelés généralement « Classes ». Dans notre cas, on les appelle classes lithologiques car chacune de ces classes représente une formation lithologique à part. Pour une trace sismique qui coïncide avec un puits, il sera facile de reporter la lithologie correspondante. Mais au delà de cette trace on ne peut rien dire sur toute la section. Connaissant la lithologie sur une portion d'une section et en utilisant les différentes réponses des attributs sismiques de l'ensemble de la section, on peut

faire une classification lithologique de la totalité du profil sismique 2D ou cube sismique 3D par le biais des cartes auto-organisatrices.

III.4. Architecture neurale des cartes auto-organisatrice

Ces cartes sont des réseaux de neurones constitués de deux couches (Figure:III.4.) :

La couche d'entrée sert à la présentation des observations (données) à classer. Cette couche contient alors exactement N neurones (N étant la dimension de l'espace des observations)

La couche d'adaptation est formée du treillis des neurones qui forme la carte. La structure du réseau employé peut être soit fixée a priori, soit évoluer lors de l'apprentissage.

Les neurones utilisés à ce niveau sont de simples neurones, chacun d'entre eux étant connecté à tous les éléments de la couche d'entrée. Le vecteur référent W_k associé à un neurone K de la carte n'est autre que le vecteur de poids. Puisque le réseau est totalement connecté à la couche d'entrée, chaque neurone possède N connexion (poids)

La figure (Figure :III.4) montre un exemple schématique des cartes auto-organisatrices. C'est un réseau (réseau de Kohonen) à deux dimensions et une couche d'entrée de cinq composantes. Chaque neurone d'entrée est relié à chaque neurone de la carte

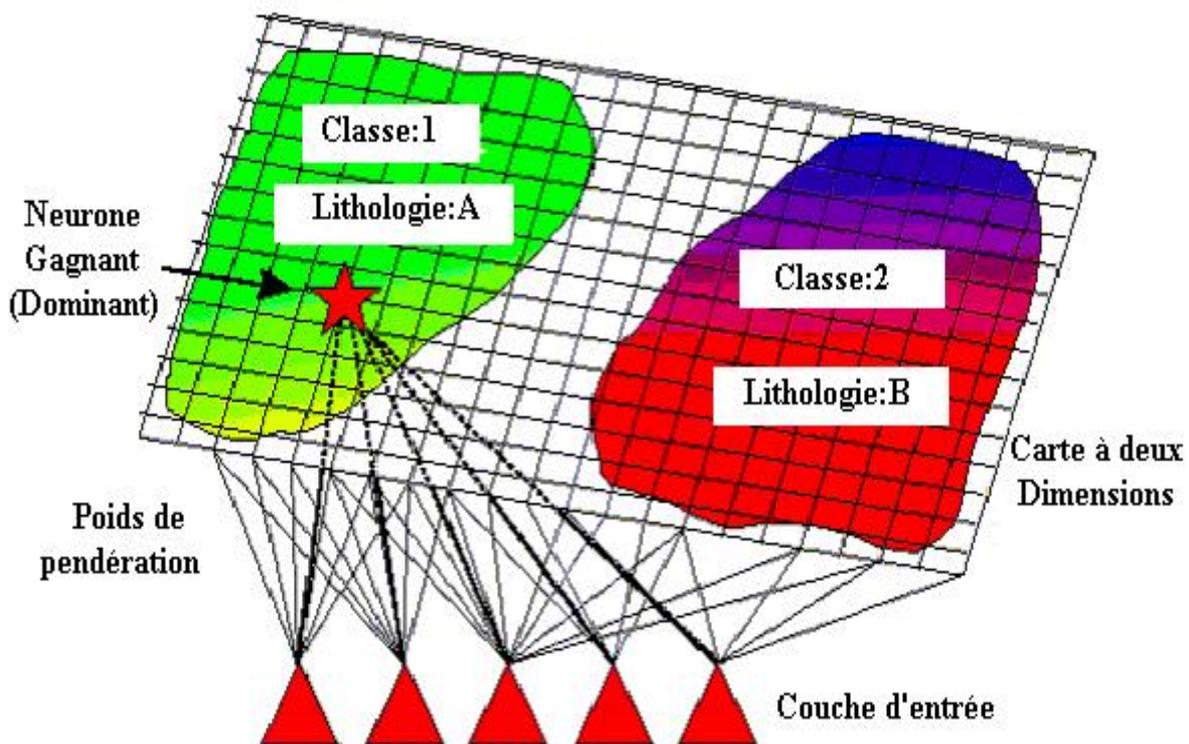


Figure:III.4. Une carte auto-organisatrice (réseau de Kohonen) à 2-dimensions et une couche d'entrée à 5 composantes.

III.5. Méthodologie de la technique

De la même façon que les études de la covariance, les entrées des cartes auto-organisatrices doivent être normalisées de telle sorte que leurs valeurs des erreurs quadratiques moyennes (RMS) seront égales à l'unité[55]. Ceci empêchera n'importe quel attribut de dominer arbitrairement la classification.

Les cartes auto-organisatrices de Kohonen se composent d'une seule couche de neurones organisés dans une dimension, deux dimensions ou multidimensionnelles (plus de deux dimensions). Chaque neurone a un nombre de connexions d'entrée égale au nombre d'attributs utilisés dans la classification comme représenté sur la Figure: III.5.a.

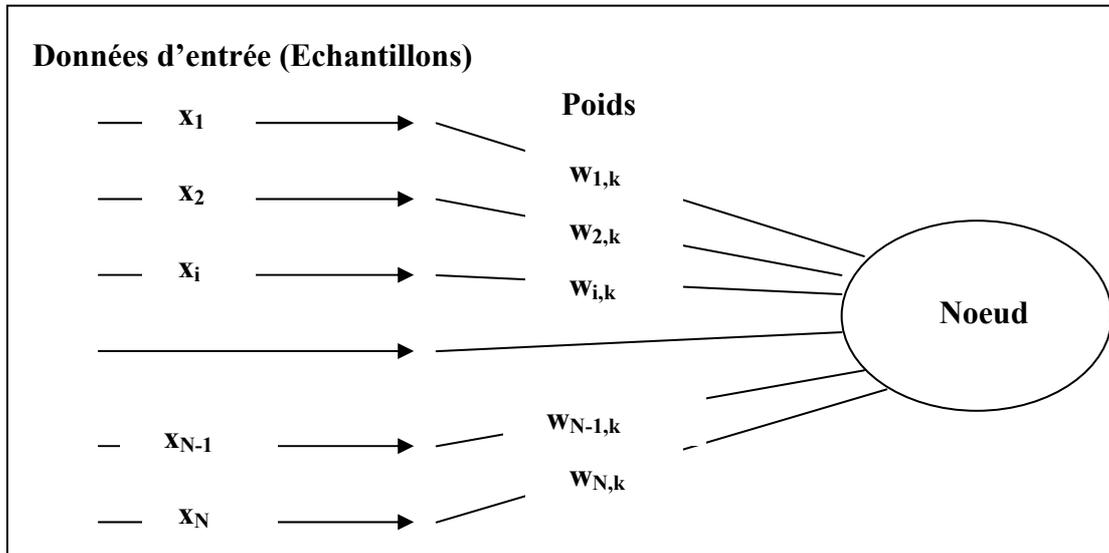


Figure:III.5.a. Connexion des données d'entrée au nœud neural

La procédure d'apprentissage consiste à trouver le neurone ayant le poids le plus proche du vecteur de données d'entrée et le déclarer comme neurone gagnant. Les poids de tous les neurones à proximité du neurone gagnant sont ajustés par une quantité inversement proportionnelle à leur distance. Au fur et à mesure que le nombre d'itérations augmente le rayon du voisinage admis sera réduit. Le processus d'apprentissage se termine si les erreurs RMS de toutes les entrées sont réduites à un niveau acceptable ou après un nombre d'itérations prédéfini.

Il y a deux méthodes pour déterminer la similitude des réponses de données :

Dans la première, chaque entrée est pondérée par le vecteur poids du neurone correspondant et les résultats sont additionnés. Ceci représente l'entrée net du neurone. Soit k l'indice du $k^{\text{ème}}$ neurone avec N attributs employés, alors l'entrée net sera (en termes de produit scalaire);

$$\text{net}_k = \sum_{i=1}^N [x(i) \cdot w(i,k)] \tag{III.1}$$

Le produit scalaire (III.1) donnera la projection d'un vecteur sur un autre. Une entrée **net** égale à **1** représenterait deux vecteurs colinéaires, qui sont orientés dans la même direction et leurs paramètres sont semblables. Une valeur **net** égale **0** signifiera que les deux vecteurs sont perpendiculaires, alors ils ne sont pas semblables.

La deuxième méthode pour mesurer la similitude des deux vecteurs est de calculer la distance euclidienne les séparant (Figure : III.5.b) comme;

$$\|\text{net}_k\| = \sqrt{\sum_{i=1}^N [x(i) - w(i,k)]^2} \tag{III.2}$$

Dans ce cas un résultat net égale à zéro signifiera que les deux vecteurs sont identiques. Une valeur de 2 signifiera qu'ils sont exactement opposés, donc ne sont pas semblables. On applique cette opération à tous les neurones et le neurone qui admet le résultat le plus grand du produit scalaire ou la distance euclidienne minimale sera choisi en tant que gagnant.

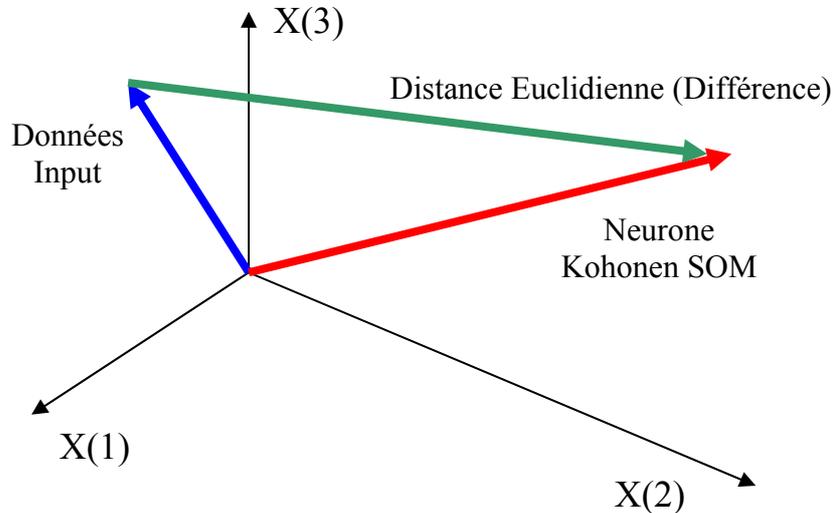


Figure : III.5.b. Distance Euclidienne séparant deux vecteurs

Afin de maintenir la similitude des neurones les plus proches de point de vue topologie, les poids de tous les neurones dans un rayon choisi doivent être ajustés. Le neurone gagnant est le plus ajusté par un processus modifiant les poids du neurone gagnant de telle sorte qu'ils soit plus près des valeurs des données d'entrée. Tous les autres neurones sont ajustés par de petites quantités inversement proportionnelles à leur distance du neurone gagnant. Avec la progression des itérations, les erreurs RMS se réduisent.

À la fin de l'apprentissage, les poids de chaque neurone représentent une référence et pourraient être utilisés pour la classification. Si l'entrée comporte des attributs sismiques, alors les poids représentent les attributs de référence de chaque classification. Le processus de classification est le calcul de la distance de chaque ensemble (ensemble de données d'entrée) à chaque neurone et l'attribution des données d'entrée à la classe représentée par le neurone gagnant.

On peut résumer cette procédure sous forme d'un algorithme réalisant les tâches suivantes :

- . Choix du nombre des neurones de la topologie du réseau et du nombres d'itérations.
- . Initialisation des poids (souvent d'une façon aléatoire).
- . Perturbation des poids.
- . Minimisation (la distance euclidienne minimale) ou (plus grand produit scalaire)
- . Fixation des poids référents.
- . Classification

III.6. Différentes organisations topologiques

Les neurones peuvent être organisés sous plusieurs formes topologiques. La figure (Figure:III.6.a.) montre une organisation unidimensionnelle. Ceci permet de grouper ou de classifier avec un rapport (relation) topologique unidimensionnel, c'est-à-dire, chaque neurone aura de petites différences par rapport aux deux neurones adjacents. La différence augmentera graduellement avec l'augmentation de la distance. Chaque neurone est relié aux données d'entrée par leurs propres poids. Dans le cas de la S.O.M, ces poids seront équivalents aux

attributs réels représentant les valeurs de l'attribut moyen de chaque groupe (Taner1997) [55]. Ce dispositif du S.O.M est très différent des réseaux neurones artificiels ANN, où les poids ne sont pas directement reliés aux données d'entrée.

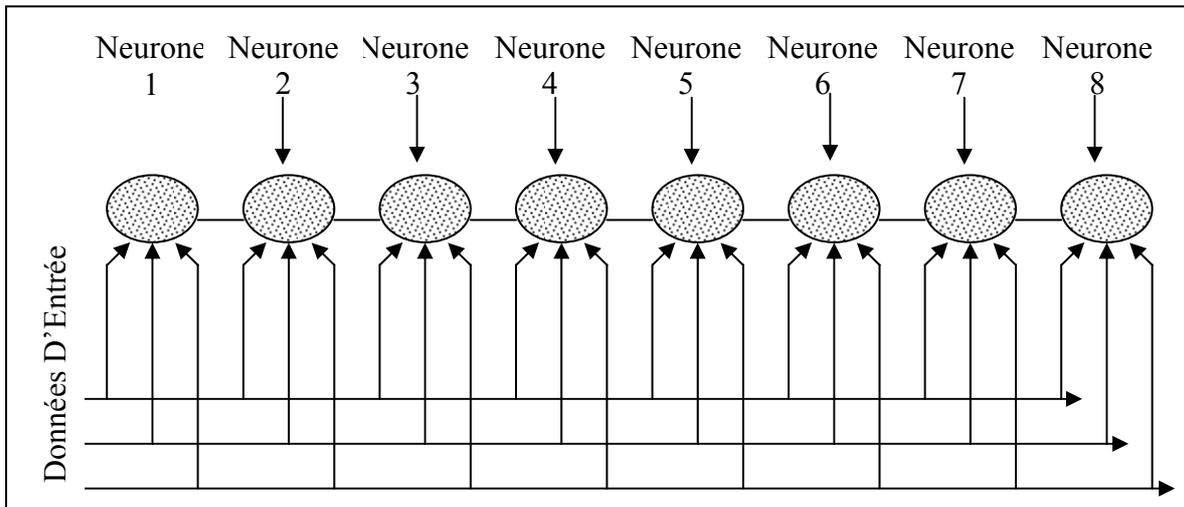


Figure:III.6.a. Forme Topologique Unidimensionnelle

La figure (Figure:III.6.b) montre une forme topologique bidimensionnelle. Cette fois nous avons quatre neurones adjacents pour chaque neurone. Les différences seront inversement proportionnelles à la distance topologique entre les neurones. Nous pouvons utiliser, de la même manière, une forme tridimensionnelle. Dans la pratique, les cartes mono ou bidimensionnelle sont habituellement utilisées.

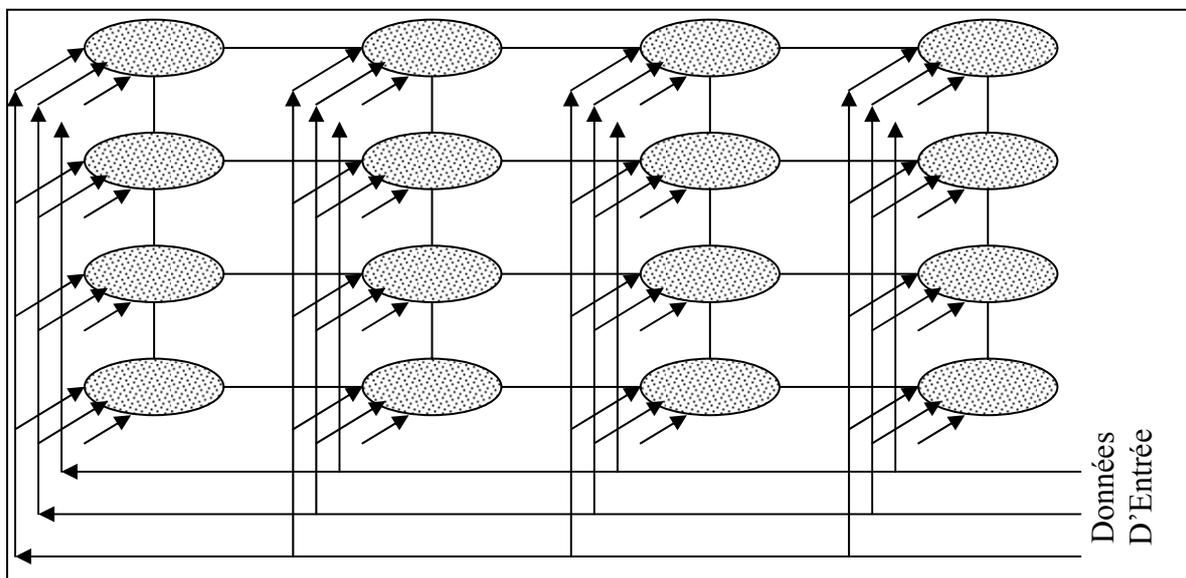


Figure:III.6.b. Forme Topologique Bidimensionnelle

III.7. L'application des cartes auto-organisatrices dans l'interprétation sismique

Les cartes auto-organisatrices de Kohonen peuvent être utilisées dans l'interprétation sismique pour deux tâches principales : une comme générateur de groupes de données (Attributs sismiques, logs de diagraphie ou données géologiques), l'autre comme outil de classification. Les données d'entrée sont généralement à deux ou trois dimensions, segment de trace sismiques ou leurs attributs. Chaque échantillon des données d'entrée et leurs attributs correspondant représentent un point dans un espace à N dimensions qui est utilisé pour former

un ensemble de poids pour chaque noeud neural selon la topologie du réseau, ces poids s'appellent mémoires associatives, parce qu'ils représentent des attributs de référence de chaque faisceau (classe), ils ont une certaine association organisée avec leurs voisins. Ce processus comprend un grand nombre d'itérations, de l'ordre de dizaines de milliers. La mesure de la réduction (minimisation) des erreurs RMS après chaque itération peut contrôler la convergence. Une fois que la convergence atteint un minimum satisfaisant, alors les poids sont enregistrés pour la classification. Puisque chaque vecteur de poids représente une référence de l'attribut moyen d'un faisceau (une classe) particulier, on applique la classification en calculant la distance euclidienne entre chaque échantillon des données d'entrée et chaque noeud, et en choisissant le noeud ayant la distance minimale. Les cartes auto-organisatrices de Kohonen S.O.M est un algorithme caractérisé par un apprentissage non supervisé. Nous devons rattacher la signification lithologique par l'utilisation des informations des puits après que la classification soit accomplie.

III.8. Procédé de calcul des S.O.M

L'application des cartes auto-organisatrices de Kohonen (S.O.M) pour une classification lithologie peut être résumer sous forme d'un algorithme descriptif organisé par la suite dans un organigramme et transcrit à la fin en un programme exécutable[55].

III.8.1. Initialisation De Données

Une fois que le groupe des attributs à utiliser dans la classification est sélectionné, chaque attribut doit être normalisé ($RMS=1$). Ceci empêche n'importe quel attribut de maîtriser le reste du groupe pendant les étapes de la classification. Si, cependant, nous considérons quelques attributs plus diagnostiques que les autres, nous pouvons ajuster leurs amplitudes RMS qui rendront la distance de cet attribut particulier plus importante que le reste, par conséquent la minimisation sera affectée par cet attribut.

III.8.2. Initialisation des poids neuraux :

Chaque neurone aura un poids pour chaque attribut. Au départ nous affectons de petits nombres aléatoires à chacun des poids et nous initialisons également le rayon de voisinage. Ce rayon sera grand au début pour couvrir un plus grand nombre de noeuds adjacents. Avec l'augmentation des itérations, ce rayon peut être systématiquement réduit à une distance égale à l'unité. Nous initialisons également un ensemble de compteurs qui indiqueront le nombre de fois consécutives qu'un neurone est déclaré gagnant. Dans l'algorithme de "conscience", si un neurone gagne excessivement il permettra à d'autres neurones de gagner.

III.8.3. Choix des données d'entrée : Les données d'entrée sont choisies au hasard à partir de l'ensemble des traces sismiques et de leurs attributs. Ceci empêche les attributs voisins semblables d'influencer l'arrangement et l'adaptation des poids. Tous les échantillons des données sont entrés à chaque cycle d'itération.

III.8.4. Calcul de la distance euclidienne ou du produit scalaire et détermination du neurone gagnant :

La distance euclidienne ou le produit scalaire sont calculés entre l'ensemble des données d'entrée (vecteur) et chacun des noeuds neuraux. Le noeud muni de la distance euclidienne minimale ou du produit scalaire maximum est déclaré gagnant. Ceci détermine le noeud avec les poids les plus semblables aux attributs d'entrée. Avant que le noeud gagnant ne soit déterminé, son nombre de "victoires" consécutives est vérifié. Si le noeud dépasse une limite définie, alors les résultats sont réduits proportionnellement pour permettre à un autre

noeud de "gagner". Les compteurs des noeuds perdants sont remis à zéro et le compteur du noeud gagnant est incrémenté d'une unité.

III.8.5. l'adaptation des poids du noeud gagnant et de ses voisins :

Les poids des noeuds gagnants sont modifiés comme suit;

$$w_{(n+1)}(k,j) = w_n(k,j) + \eta(n)[x(j) - w_n(k,j)] \quad (\text{III.3})$$

où : n : le nombre d'itérations, k : le noeud gagnant, j : le j^{eme} l'attribut.

$\eta(n)$: le poids modifié sous forme d'une fonction décroissante du nombre d'itérations appelé aussi taux d'apprentissage.

Tous les noeuds associés à ce même noeud gagnant seront également modifiés, par une quantité inversement proportionnelle à la distance sur la carte. Si nous attribuons une dimension de l'ordre de un, de deux ou d'une dimension plus élevée pour chaque noeud, nous pouvons calculer la distance relative du noeud gagnant par rapport à n'importe quel autre noeud et calculer facilement les modifications.

III.8.6. Adaptation des noeuds voisins selon leurs distances par rapport au noeud gagnant :

L'adaptation des noeuds voisins du noeud gagnant selon leurs distances par rapport à ce dernier sera faite par la formule suivante :

$$w_{(n+1)}(m,j) = w_n(m,j) + d(m,k,n) \cdot \eta(n)[x(j) - w_n(m,j)] \quad (\text{III.4})$$

où : $d(m, k, n)$ est une fonction du nombre d'itérations n et de la distance entre le noeud gagnant k et son noeud voisin m . Tous les autres noeuds en dehors du rayon associé au noeud gagnant n'auront aucune correction.

w_n : Le poids du n^{ieme} noeud

III.8.7. Adaptation du rayon associé à chaque noeud corrigé :

Généralement, on considère un rayon initial assez long pour couvrir la plupart des noeuds neuraux, ce rayon diminue en fonction de l'augmentation du nombre d'itérations. La fonction $d(m,k,n)$ peut rendre n'importe quelle fonction inversement proportionnelle au nombre d'itérations n , et à la distance entre le noeud gagnant k et le noeud à ajuster m .

III.8.8. Contrôle de la convergence pour continuer ou pour interrompre l'itération :

On calcule l'erreur RMS (la racine carrée de la somme des carrés des erreurs. C'est-à-dire le carré de la distance entre les données d'entrée et les poids des neurones gagnants sur le nombre des données d'entrées) et le contrôler s'il est au-dessous d'un certain niveau prédéfini. Si ce test est vrai, alors l'objectif est atteint (les poids désirés), ces poids peuvent être utilisés pour la classification. Si non, on poursuit avec une nouvelle série d'itérations (nous reprenons de nouveau la boucle à partir de l'étape III.8.3).

III.8.9. Contrôle de qualité :

Une des caractéristiques principales des cartes de Kohonen est la détermination des neurones avec les poids associés. Le premier Contrôle à faire est de voir si les distances euclidiennes calculées entre les noeuds neuraux sont proportionnelles aux distances topologiques de la carte désirée. Si cette condition n'est pas satisfaite, une série d'itérations additionnelle sera nécessaire.

III.8.10. Enregistrement des poids calculés.

Les poids des noeuds neuraux calculés sont des fonctions mémorisées qui seront employées pour la classification. Ces poids peuvent avoir de petites différences entre les neurones adjacents.

III.8.11. Classification de l'ensemble des données d'entrée :

Dans l'étape de classification pour chaque entrée, nous calculons la distance euclidienne de chaque noeud neural et classifions l'entrée en tant que membre de la classe représentée par le noeud gagnant.

III.8.12. Contrôle de validité des résultats :

Après chaque itération une nouvelle carte de densité de population des différentes classes sera présentée. La forme de cette carte change en fonction de l'évolution du processus de la classification. Dans le cas idéal tous les nœuds convergent vers un seul point d'où la notion du neurone gagnon. En plus du résultat graphique, d'autres paramètres sont également imprimés pour le contrôle et la surveillance de l'évolution du réseau. Le pas d'apprentissage, la correction du rayon entre deux neurones voisins, la largeur de la base de la carte SOM et l'erreur RMS de chaque passe(itération) sont les principaux paramètres de contrôle utilisés.

III.9. Organigramme du programme SOM

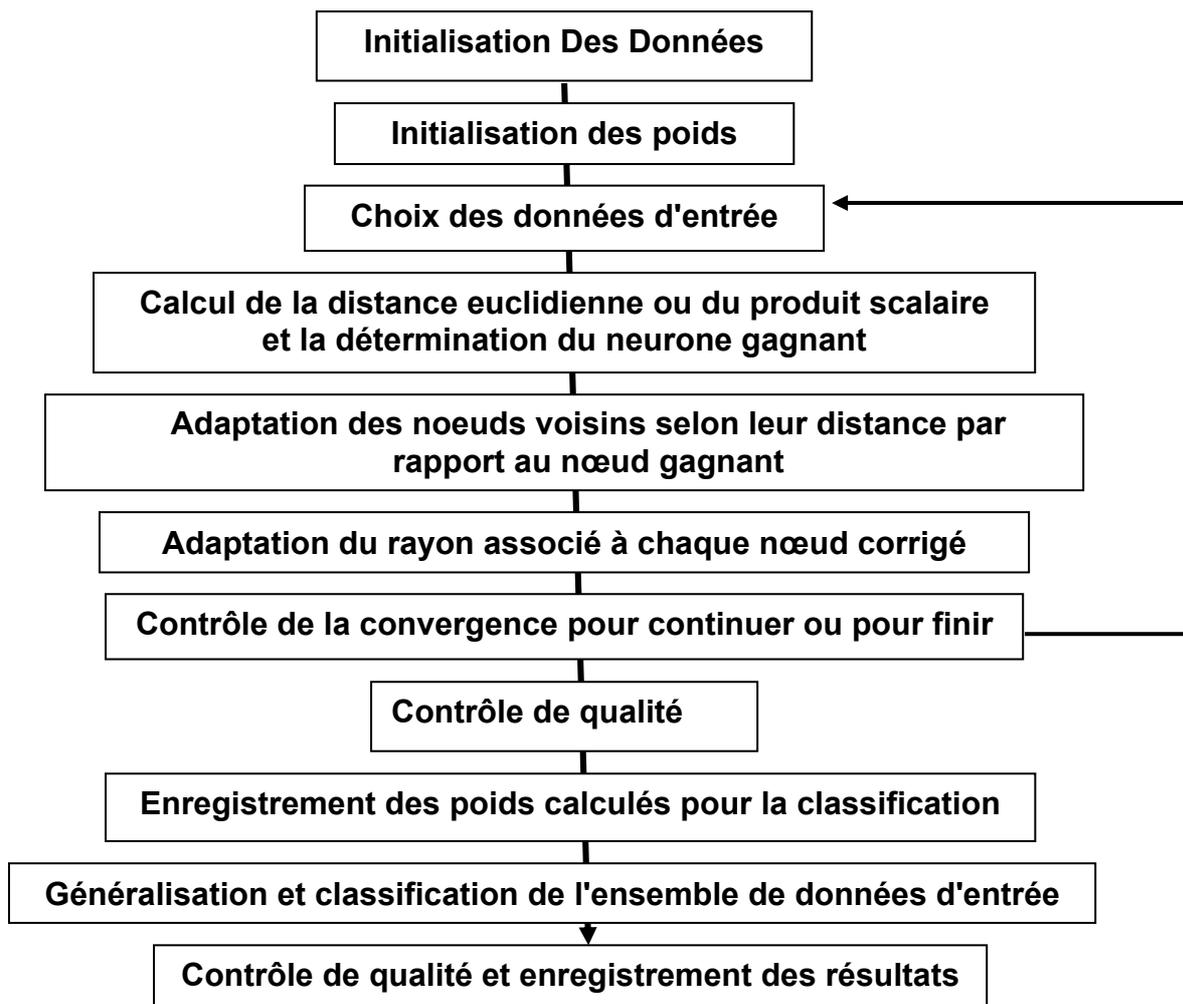


Figure:III.9. Organigramme Du Programme SOM Appliqué

IV. Le perceptron Multicouches (PMC), Technique de classification et apprentissage supervisé

IV.1. Introduction

On entend par classification la tâche qui consiste à séparer des données en classes distinctes, à partir de leurs caractéristiques. La classification, en d'autres termes, n'est autre que la détermination des frontières entre les classes ou surfaces discriminantes. Les statisticiens appellent aussi « classification » la tâche qui consiste à regrouper des données qui se ressemblent dans des classes qui ne sont pas définies à l'avance. Nous avons vu (dans le chapitre III) que les réseaux de neurones à apprentissage non supervisé peuvent réaliser ce genre de tâches.

Nous savons qu'un réseau non bouclé réalise une relation entre ses entrées et ses sorties. On peut donc affecter à un réseau la tâche qui consiste à réaliser une fonction algébrique non linéaire ; celle-ci peut être :

- Connue analytiquement : le réseau réalise alors une tâche d'approximation de fonction.

- Inconnue analytiquement : mais pour laquelle on dispose de valeurs, en nombre fini, qui sont entachées de bruit ;

C'est essentiellement pour ce dernier type d'applications que sont mis en œuvre les réseaux de neurones à apprentissage supervisé qui sont alors des approximateurs universels.

Toute fonction bornée suffisamment régulière peut être approchée uniformément, avec une précision arbitraire, dans un domaine fini de l'espace de ces variables, par un réseau de neurones comportant une couche de neurones cachés en nombre fini, possédant tous la même fonction d'activation, et un neurone de sortie linéaire [Hornik et al.1989], [Hornik et al.1990], [Hornik 1991]. Cette propriété mathématique donne une base solide à l'application du perceptron multicouches comme outil de classification des données sismiques avec un apprentissage supervisé pour donner à la fin (sortie du réseau) une coupe lithologique (classification lithologique). Pour les réseaux à apprentissage supervisé, on présente au réseau des entrées, et au même temps les sorties que l'on désire pour cette entrée. Le réseau doit alors se configurer, c'est-à-dire calculer ses poids afin que la sortie qu'il donne corresponde bien à la sortie désirée.

Dans cette partie nous développerons des équations pour déterminer les poids et des seuils optimaux qui minimisent les erreurs entre les sorties désirées et les sorties réelles d'un réseau de neurones (la phase d'apprentissage). Le procédé informatique est la règle généralisée d'adaptation des poids et des seuils appelée en littérature anglo-saxonne (the Generalized Delta Rule) avec l'utilisation de l'algorithme de rétropropagation du gradient (Back Propagation Approach) [58].

Dans cette application, les données d'entrée sont des attributs sismiques. La sortie désirée sera formée des amplitudes sismiques choisies à partir des zones dont la lithologie est connue préalablement (information tirée d'un profil passant par un puits)

IV.2. Neurones

Un réseau de neurones à N entrées, N_c neurones cachés et N_o neurones de sortie réalise N_o fonctions non linéaires de ces N variables d'entrée par composition des N_c fonctions réalisées par ses neurones cachés. Les réseaux de neurones non bouclés dont les couches cachées ont une fonction d'activation sigmoïde, sont souvent appelés perceptron multicouches. La figure (Figure:IV.2.a) illustre le modèle d'un perceptron. Ce neurone se compose d'un ensemble d'entrées $X=[x_1, x_2, \dots, x_N]$ et d'une seule sortie. Fonctionnellement, chaque entrée x_i sera multipliée par un poids w_i (pondération), tous les produits $(x_i \cdot w_i)$ seront additionnés. Le résultat de la sommation, noté net du neurone (le potentiel du neurone), sera calculé par la formule suivante :

$$\text{net} = \sum_{i=1}^N [x(i).w(i)] \quad (IV.1)$$

Le neurone représenté par la figure (Figure : IV.2.a) a été développé par Widrow (1962) et est appelé le Perceptron linéaire.

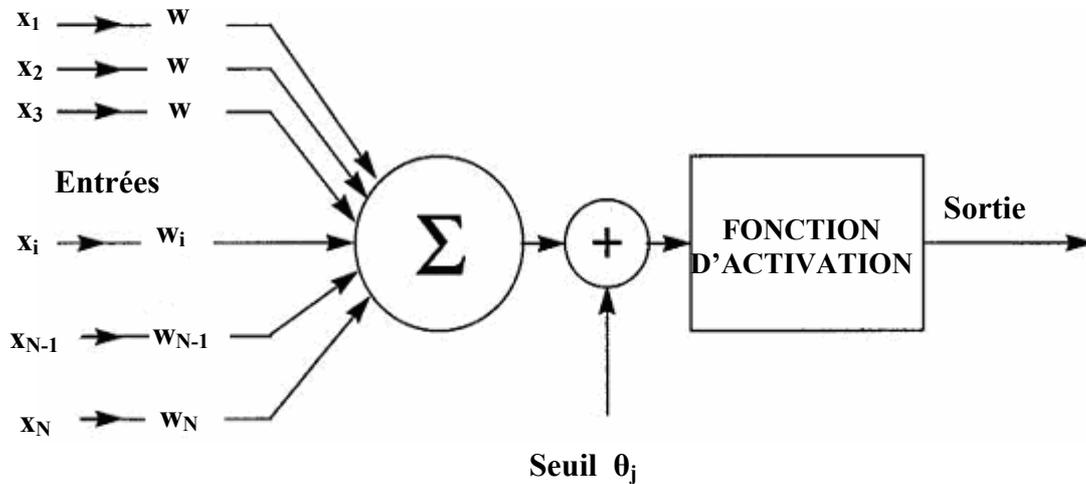


Figure:IV.2.a. Représentation graphique d'un perceptron

L'entrée de la fonction d'activation est toujours formée du net (potentiel) et du seuil θ_j . Cette fonction d'activation peut être une fonction linéaire par partie, arctangente ou une fonction sigmoïdale. La fonction sigmoïdale est la plus utilisée (calculs plus faciles). Avant que la fonction d'activation ne soit calculée, on ajoute une valeur comme seuil à la fonction d'activation. La sortie du neurone a comme expression;

$$b(j) = 1 / \{1 + \exp[-\frac{\text{net}(j) + \theta(j)}{\theta(0)}]\} \quad (IV.2)$$

$\theta(0)$ est un facteur de la forme sigmoïde.

Pour un nombre élevé de $\theta(0)$, la fonction sigmoïde change doucement. Pour des valeurs plus basses de $\theta(0)$ la fonction sigmoïde prend une forme plus raide. Ceci est illustré sur la Figure (Figure:IV. 2.b).

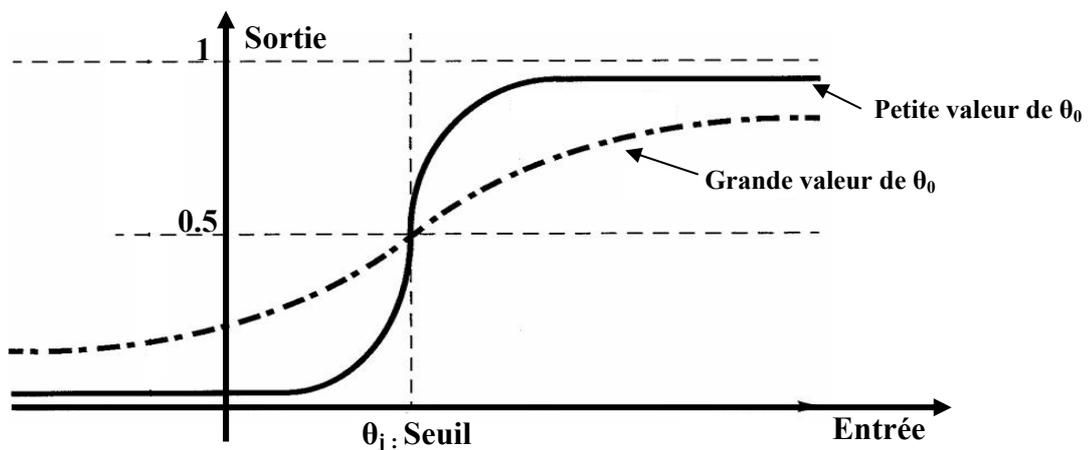


Figure:IV. 2.b. Fonction d'activation de type sigmoïde

L'apprentissage d'un neurone signifie le calcul des valeurs $w(i)$ (poids) et $\theta(j)$ (seuil) pour classifier correctement l'ensemble de données.

$$\mathbf{X} \cdot \mathbf{w} = \begin{cases} 1 & \text{si } \mathbf{X} \text{ appartient à la classe A} \\ 0 & \text{si } \mathbf{X} \text{ appartient à la classe B} \end{cases}$$

Si les données d'entrée sont très bruitées, qui est généralement le cas, nous n'obtiendrons jamais le 0 et le 1 comme résultats, elles seront dans cet intervalle (0,1). C'est pourquoi un seuil subjectif doit être ajouté pour décider la fin de la classification de A et le début de la classification de B. après apprentissage, un simple neurone peut réaliser la classification comme discriminateur linéaire.

IV.3. Discrimination linéaire par un neurone

Si $\mathbf{X} = \{x_1, x_2, \dots, x_N\}$, représente le modèle d'entrée pour l'apprentissage sous forme de vecteur colonne, le but est de chercher le vecteur colonne \mathbf{w} de M éléments tel que;

$$\mathbf{X} \cdot \mathbf{w} = \mathbf{b} \quad (\text{IV.3})$$

Les éléments de \mathbf{b} sont les valeurs de sortie spécifiées par l'apprentissage. Dans la classification binaire, $b=1$ indique que l'ensemble \mathbf{X} appartient à la classe A et $b=0$ indique l'ensemble \mathbf{X} appartient à la classe B. Puisque la matrice \mathbf{X} est composée de N colonnes et M lignes et $M < N$, alors nous pouvons résoudre ce système d'équations par la méthode classique au sens des moindres carrés. Nous pouvons obtenir une matrice carrée en multipliant les deux côtés du système par la matrice transposée de \mathbf{X} ;

$$\mathbf{X}^T \cdot \mathbf{X} \cdot \mathbf{w} = \mathbf{X}^T \mathbf{b} \quad (\text{IV.4})$$

et calculant \mathbf{w} ,

$$\mathbf{w} = (\mathbf{X}^T \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^T \mathbf{b} \quad (\text{IV.5})$$

Cette expression peut être simplifiée et devenir;

$$\mathbf{w} = \hat{\mathbf{X}} \cdot \mathbf{b} \quad (\text{IV.6})$$

où $\hat{\mathbf{X}}$ est le pseudo-inverse de la matrice rectangulaire $(\mathbf{X}^T \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^T$,

$$\hat{\mathbf{X}} = (\mathbf{X}^T \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^T \quad (\text{IV.7})$$

Théoriquement on peut calculer cet inverse directement. Cependant nous pouvons obtenir quelques valeurs absurdes qui ne représentent pas un cas réel. Ceci est calculé par un procédé itératif appelé l'algorithme linéaire du perceptron.

Nous souhaitons calculer un seul ensemble de poids \mathbf{w} pour réaliser à la sortie un ensemble correct \mathbf{b} pour tous les modèles d'entrée $\mathbf{x}(\mathbf{p})$. Nous commençons par un ensemble de valeurs arbitraire de $\mathbf{w}^1(\mathbf{i})$, après on passe à l'adaptation par la règle suivante :

$$w^{(k+1)}(j) = w^k(j) + \rho [b(j) - w^k(j)x(j)] \cdot x(j) \quad (\text{IV.8})$$

Cette adaptation se poursuit jusqu'à une classification correcte du modèle, Alors l'expression $[b(j) - w^k(j)x(j)]$ sera nulle ou devient très petite. Dans la pratique ceci ne peut pas être atteint dans la plupart des cas, par conséquent l'itération s'arrêtera quand la somme des carrés des erreurs atteint un certain seuil préalablement défini.

Nous pouvons écrire l'expression (IV.8) sous la forme;

$$\Delta(\mathbf{W}) = \eta \cdot \delta \cdot \mathbf{X} \quad (\text{IV.9})$$

Où δ est la différence entre la sortie désirée et la sortie réelle calculée par le perceptron.

L'expression (IV.9) est appelée l'expression de la règle delta (delta-rule expression), qui déclare que le changement du vecteur de poids devrait être proportionnel au delta (l'erreur) et au modèle d'entrée.

IV.5. Exemple d'application des réseaux de neurones dans la sismique

IV.5.1. Choix des données pour l'apprentissage

Il est nécessaire d'identifier les objets par leurs particularités qui les distinguent des autres. Ces particularités sont généralement appelées des caractéristiques (features). Pour la reconnaissance des formes (identification des modèles géologiques) nous devons définir chaque objet par ses caractéristiques sous forme d'une série de valeurs présentées comme vecteurs caractéristiques. Par exemple, si nous voulons faire apprendre un programme de pointer automatiquement les premières arrivées, nous devons d'abord définir certaines caractéristiques des premières arrivées, ce qui les rend différentes de tous les autres événements sismiques ou de toutes formes de bruit.

D'une façon générale, les premières arrivées sont distinguées par un caractère très énergétique après une zone de bruit caractérisée par des amplitudes plus faible. Par conséquent la différence entre le niveau de l'énergie de l'amplitude moyenne du bruit et le niveau de l'énergie de l'amplitude moyenne des premières arrivées est une caractéristique intéressante à utiliser. Il est connu qu'après les arrivées premières, les amplitudes restent presque semblables pendant quelques temps. Par conséquent on ne peut pas utiliser les mêmes critères pour différencier entre les premières arrivées et les événements qui arrivent après (postérieurs). Dans ce cas nous employons la dérivée de l'amplitude moyenne. Dans la zone des premières arrivées le niveau de l'énergie de l'amplitude moyenne croit du niveau moyen de l'énergie du bruit jusqu'au niveau moyen de l'énergie des réflexions. Par conséquent le taux de changement est beaucoup plus grand que le changement du niveau de l'énergie prévu dans les zones des ondes réfléchies. Ceci est habituellement mesuré comme le rapport de puissance qui atteint un maximum autour de la première arrivée.

Généralement on prend un pic positif ou un pic négatif de la trace sismique (Figure:IV. 5.) De cette façon nous choisissons seulement un endroit sur la trace sismique où la première dérivée égale à zéro. Si nous sélectionnons un pic positif, alors nous réduisons les points pointés seulement aux positions ayant les valeurs positives et leurs premières dérivées égales à zéro. Puisque nous supposons généralement que les ondes réfléchies ou réfractées ont une énergie plus élevée que la zone des bruits, nous pouvons définir un niveau de seuil. Au dessous de ce seuil les traces seront classifiées comme zone de bruits et au dessus duquel les traces seront classifiées comme zone de réflexions ou réfractions. Par conséquent l'amplitude de l'enveloppe de la trace devient l'une des caractéristiques. Si nous considérons la pente de l'enveloppe comme une autre caractéristique nous pouvons voir que dans la partie gauche de la réflexion (Figure:IV.5.) cette pente sera positive et négative dans la partie droite. À la crête de l'enveloppe elle sera nulle. Nous pouvons donc définir l'endroit où on doit pointer tout en donnant une partie de la pente de l'enveloppe. Si nous considérons des valeurs positives alors nous sommes entrain de pointer sur la partie gauche de l'ondelette (première arrivée). Nous pouvons voir aussi que divers attributs sismiques comme les temps d'arrivées et les vitesses peuvent être utilisés pour décrire le temps des premières arrivées.

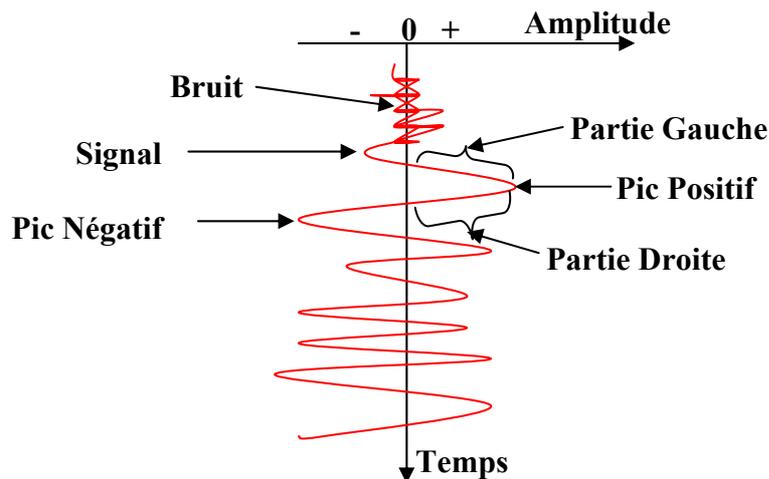


Figure:IV. 5. Pointé des premières arrivées

IV.5.2.Choix des paramètres d'apprentissage

Pour débiter la phase d'apprentissage, on doit choisir quelques points de tirs (Shots) dans un profil sismique. Deuxièmement il faut sélectionner une zone de quelques centaines de millisecondes autour de la première arrivée. L'utilisateur donnera le pointé en temps de la première arrivée pour une gamme d'offset. Le programme détermine, d'abord, les temps de toutes les crêtes avec la même polarité selon la demande de l'utilisateur, puis il calcule tous les attributs et les temps pour toutes les crêtes. Il marquera le pointé donné par l'utilisateur comme pointé correct et tous les autres comme pointés erronés, pré first-break ou post first-break. Dans ce cas nous aurons trois différentes classes. Si nous avons huit attributs pour décrire une ondelette, alors notre réseau de neurones aura huit noeuds d'entrés et trois noeuds de sortie. Nous devons avoir autant de groupe de données d'entrée pour l'apprentissage que de nombre de traces pointées. Nous pouvons définir nos pointés en donnant une valeur (prévue) pour chaque groupe d'apprentissage. Par exemple nous assignons le premier noeud de sortie pour représenter la classification des pointés de la partie gauche des premières arrivées (près First-break, Figure:IV. 5.), le second aux pointés réels (pic des premières arrivées) et le troisième noeud aux pointés de la partie droite des premières arrivées (post First-break, Figure:IV. 5.). Si l'ensemble d'entrée appartient aux événements du pré First-break, alors nous prévoyons que le sortie sera (1.0, 0.0, 0.0), dans le cas des pointés réels la sortie sera (0.0, 1.0, 0.0) et la sortie pour les pointés du post First-break sera (0.0, 0.0, 1.0). Les données d'entrée et leurs correspondantes valeurs de sorties prévues sont entrées dans le réseau pour l'apprentissage.

IV.5.3. Exécution du processus d'apprentissage et contrôle des résultats

Le réseau déterminera tous les coefficients par **Retro-propagation** des erreurs tout en essayant de minimiser la somme des carrés de la différence (erreurs) entre la sortie prévue et la sortie calculée (réelle). Ceci prend habituellement des centaines ou des milliers d'itérations. La vitesse de la convergence est plus rapide dans les premières itérations et devient plus lente au fur et à mesure que le nombre des itérations augmente.

Parfois, les itérations atteignent un seuil, qui peut être soit un minimum local soit un minimum global. Dans le cas où il n'y a pas d'amélioration significative des résultats, alors le calcul devra être arrêté et les résultats peuvent être exploités. En comparant les résultats calculés aux valeurs prévues, nous pouvons décider si un degré suffisant de convergence a été atteint ou non. Si nous atteignons un degré de convergence raisonnable, alors nous considérons que l'apprentissage est terminé et le réseau est maintenant prêt à être employé

pour pointer les premières arrivées du reste des données voisines. Si l'apprentissage n'est pas suffisant, nous pouvons utiliser une deuxième série de données pour un autre apprentissage. Un ensemble important de données pour l'apprentissage devrait avoir comme conséquence un meilleur degré de convergence.

IV.6. Algorithme de retro-propagation du gradient de l'erreur et règle d'adaptation des poids

La plupart des algorithmes d'apprentissages permettent d'obtenir les poids par minimisation d'une fonction de coût dérivable. la méthode la plus simple pour procéder à une minimisation est la méthode du gradient.

On considère un réseau constitué de « n » neurones recevant des vecteurs à « p » composantes. Les entrées x_k du réseau sont distribuées sur tous les neurones.

A un vecteur d'entrée, on veut associer un vecteur de sortie désiré. Si les poids ont des valeurs quelconques, le vecteur de sortie observé est a priori différent de celui désiré. on peut associer à cette différence l'erreur quadratique. Nous allons montrer que l'on peut modifier les poids du réseau en minimisant cette erreur quadratique.

La mise à jour de l'ensemble des poids d'une couche nécessite la connaissance des erreurs associées à chaque neurone de la couche suivante. On appliquera l'algorithme d'ajustement des poids en partant de la dernière couche (pour laquelle les erreurs sont connues) vers la première, d'où l'appellation de cet algorithme : Algorithme de retropropagation du gradient de l'erreur, ou simplement « backpropagation Algorithme » dans la désignation anglo-saxonne.

IV.6.1. Règle d'adaptation de la couche de sortie

Dans ce chapitre nous utiliserons la notation de Taner.

Soit $X = [x_1, x_2, x_3, \dots, x_N]$ un vecteur de données d'entrée (Input). On utilise p comme index d'ordre du vecteur des données d'entrée. Soit X un vecteur de données d'entrée appartenant à la classe b. Cette classe est définie par les valeurs prévues à atteindre $T[t_1, t_2, t_3]$. Ce qui signifie que lorsqu'on considère comme entrée du réseau un vecteur X nous obtenons T comme résultat à la sortie du réseau. (Voir la Figure:IV.6.a) [58];

$$\mathbf{X} \cdot \mathbf{w} = \mathbf{T} \quad (\text{IV.10})$$

La sortie (résultat) réelle est cependant $O[o_1, o_2, o_3]$, le vecteur d'erreur sera :

$$\mathbf{e} = (\mathbf{T} - \mathbf{O}) \quad (\text{IV.11})$$

En utilisant la méthode de calcul au sens des moindres carrés, on calcule les opérateurs W qui minimisent la somme des carrés de la fonction d'erreur E.

$$E = \frac{1}{2} \sum_{k=1}^K [t(k) - O(k)]^2 \quad (\text{IV.12})$$

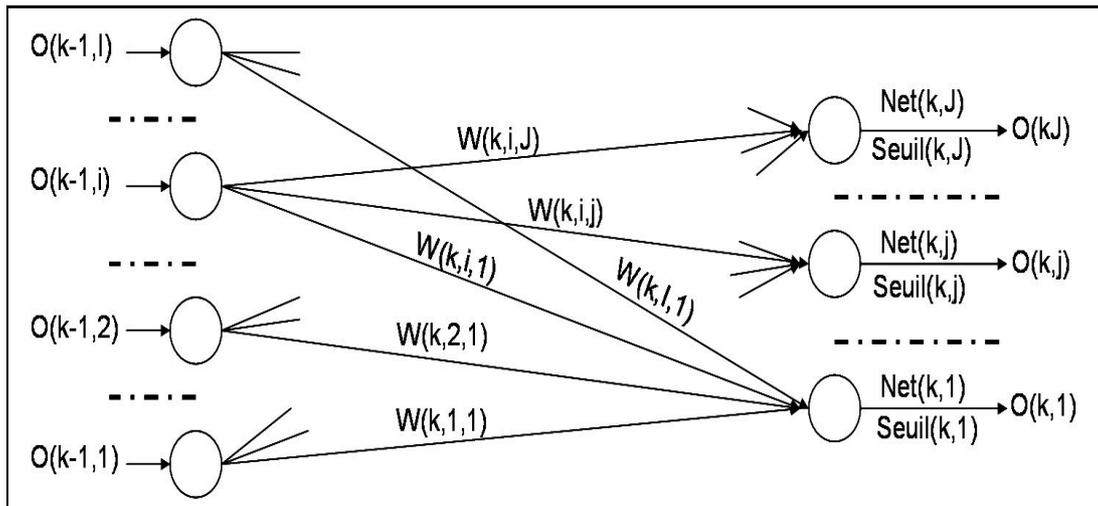


Figure:IV.6.a. Adaptation des couches de sortie

Le gradient E d'une surface d'erreur de n dimensions est donné par le vecteur défini par les dérivées partielles de la fonction d'erreur en respectant les valeurs des poids et des seuils, ces deux derniers sont des inconnus. Lorsqu'on atteint un minimum, cette fonction $(-\frac{\partial E}{\partial w(k,i,j)})$ sera égale à zéro pour toutes les valeurs de $w(i,j)$. Pour la raison que les dérivées des fonctions quadratiques sont des équations linéaires et que nous disposons de nombreux algorithmes pour résoudre un système linéaire. Nous avons annulé les dérivées de la fonction E pour une solution directe des fonctions de l'erreur quadratiques. Malheureusement la fonction d'erreur pour les réseaux de neurones n'est pas une fonction quadratique et les dérivés résultant ne sont pas des équations linéaires.

Puisque le gradient est donné par les dérivées partielles, si nous utilisons comme variation des poids $\Delta w(i,j)$ proportionnelle aux dérivées, nous obtenons l'erreur correspondante;

$$\Delta w(k,i,j) = -\eta \frac{\partial E}{\partial w(k,i,j)} \tag{IV.13}$$

où : k représente l'indice de la couche de sortie, j représente le $j^{\text{ème}}$ noeud de la couche de sortie et i représente le $i^{\text{ème}}$ noeud de la couche avant la couche de sortie.

Selon la définition du réseau, l'erreur E est exprimée en termes de résultats désirés et réels. Le résultat réel (sortie réelle) est une sortie non linéaire du noeud k :

$$O(k,j) = f[\text{net}(k,j)] \tag{IV.14}$$

Où : $\text{net}(k)$: est la somme pondérée de toutes les sorties de la couche précédente,

$$\text{net}(k,j) = \sum_{i=1}^I [w(k,i,j).O(k-1,i)] \tag{IV.15}$$

Nous ne pouvons pas calculer directement la dérivée partielle de E en respectant ces poids, mais nous pouvons les évaluer indirectement;

$$\frac{\partial E}{\partial w(k,i,j)} = \left[\frac{\partial E}{\partial \text{net}(k,j)} \right] \cdot \left[\frac{\partial \text{net}(k,j)}{\partial w(k,i,j)} \right] \tag{IV.16}$$

Nous savons par définition qu'au $j^{\text{ème}}$ noeud de la couche de sortie, le terme $\text{net}(k, j)$ est la somme pondérée des sorties réelles de la couche précédente. Alors la dérivée demandée est;

$$\frac{\partial \text{net}(k, j)}{\partial w(k, i, j)} = \frac{\partial \sum_{i=1}^I \{w(k, i, j) \cdot O(k-1, i) + \theta(k, j)\}}{\partial w(k, i, j)} = O(k-1, i) \quad (\text{IV.17})$$

Ce qui correspond à la sortie calculée du $i^{\text{ème}}$ noeud de la dernière couche cachée ($(k-1)^{\text{ème}}$). A ce moment, nous introduisons la fonction δ :

$$\delta(k, j) = -\frac{\partial E}{\partial \text{net}(k, j)} \quad (\text{IV.18})$$

Par conséquent nous pouvons écrire;

$$\Delta w(k, i, j) = \eta \cdot \delta(k, j) \cdot O(k-1, i) \quad (\text{IV.19})$$

Cette expression est connue sous le nom « Generalized Delta rule [58] » Règle généralisée de delta.

Pour calculer la deuxième composante de la partie droite de la formule (IV.16), on calcule la dérivée partielle $\frac{\partial E}{\partial \text{net}(k, j)}$.

$$-\frac{\partial E}{\partial \text{net}(k, j)} = \left[-\frac{\partial E}{\partial O(k, j)} \right] \cdot \left[\frac{\partial O(k, j)}{\partial \text{net}(k, j)} \right] \quad (\text{IV.20})$$

de l'expression (VI.12);

$$\frac{\partial E}{\partial O(k, j)} = -[t(j) - O(k, j)] \quad (\text{IV.21})$$

et;

$$\frac{\partial O(k, j)}{\partial \text{net}(k, j)} = O(k, j) \cdot [1.0 - O(k, j)] \quad (\text{IV.22})$$

Si nous utilisons la fonction sigmoïde $[1.0 + \exp(-O(k, j) + \theta)]^{-1}$ comme fonction d'activation, avec le seuil θ , dans ce cas la fonction δ sera;

$$\frac{\partial E}{\partial \text{net}(k, j)} = - [t(j) - O(k, j)] \cdot O(k, j) \cdot [1.0 - O(k, j)] \quad (\text{IV.23})$$

de l'expression (IV.18), on peut écrire ;

$$\delta(k, j) = [t(j) - O(k, j)] \cdot O(k, j) \cdot [1.0 - O(k, j)] \quad (\text{IV.24})$$

Par conséquent la valeur adaptée du poids pour la couche de sortie sera;

$$\Delta w(k, i, j) = \eta \cdot O(k, j) [1 - O(k, j)] \cdot [t(j) - O(k, j)] \cdot O(k-1, i) \quad (\text{IV.25})$$

Pour mettre à jour la valeur du seuil nous utiliserons une approche similaire avec l'exception qu'au lieu d'une sortie (résultat) arbitraire de la couche précédente nous aurons la valeur de 1. Le reste des développements sera identique. Le seul changement est :

$$\frac{\partial \text{net}(k, j)}{\partial \theta(k, j)} = \frac{\partial \sum_{i=1}^I \{w(k, i, j) \cdot O(k-1, i) + \theta(k, j)\}}{\partial \theta(k, j)} = 1 \quad (\text{IV.26})$$

Par conséquent l'adaptation du seuil au $j^{\text{ème}}$ noeud de couche de sortie sera :

$$\Delta \theta(k, j) = \eta \cdot O(k, j) [1 - O(k, j)] \cdot [t(j) - O(k, j)] \quad (\text{IV.27})$$

IV.6.2. Règle d'adaptation des couches cachées

Les mises à jour des poids des couches cachées et des seuils suivent le même développement[58]. La Figure (Figure : IV.6.b) montre les indices et les relations entrée-sortie des couches cachées.

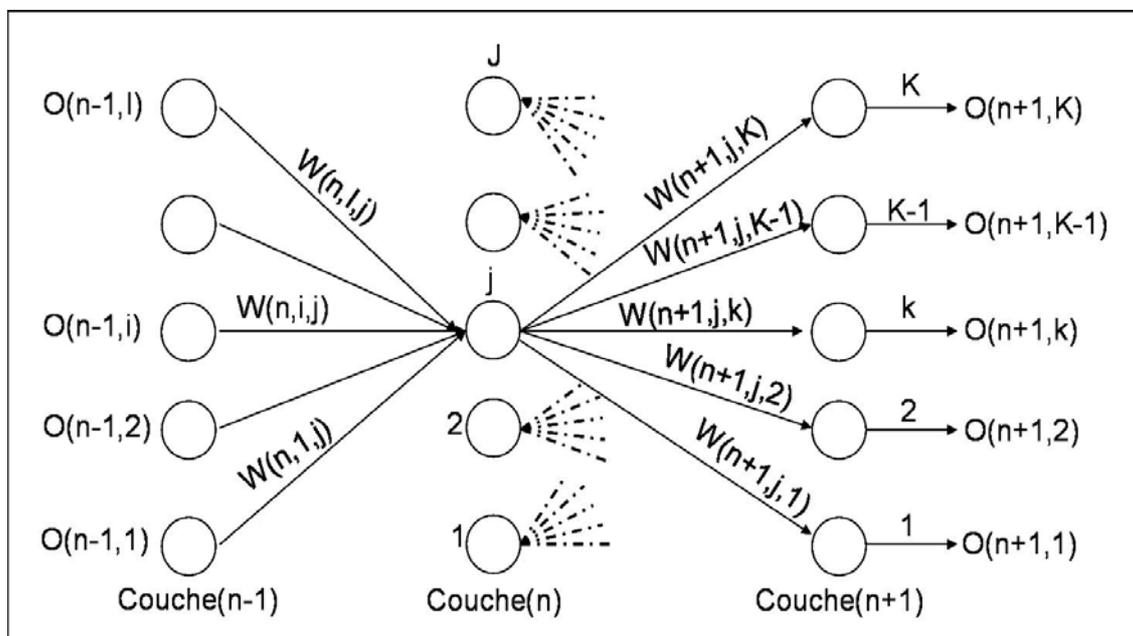


Figure:IV.6.b. Adaptation des Couches Cachées

Le concept principal est que nous commençons de la couche de sortie et nous propageons par rétro-propagation l'influence de l'erreur vers la couche d'entrée en passant par les couches cachées, d'où le nom « Back-Propagation ». Nous calculons les corrections des poids et des seuils connectés à la couche de sortie dans la première étape. Ensuite nous calculerons les poids et les seuils de la dernière couche cachée, puis la couche cachée suivante et ainsi de suite.

Supposons que nous avons calculé la $(n+1)^{\text{ème}}$ couche et nous souhaitons calculer les poids de la $(n)^{\text{ème}}$ couche. Nous utiliserons la lettre j comme indice du nœud de la $(n)^{\text{ème}}$ couche, la lettre i comme indice du nœud de la $(n-1)^{\text{ème}}$ couche et k pour la $(n+1)^{\text{ème}}$ couche, sachant qu'il peut y avoir différent nombre de noeuds dans différentes couches.

Le gradient est donné par les dérivées partielles $-\frac{\partial E}{\partial w(n, i, j)}$

Nous utiliserons encore la formule (IV.16) avec d'autres indices pour évaluer les dérivées :

$$-\frac{\partial E}{\partial w(n, i, j)} = \left[-\frac{\partial E}{\partial \text{net}(n, j)} \right] \cdot \left[\frac{\partial \text{net}(n, j)}{\partial w(n, i, j)} \right] \quad (\text{IV.28})$$

Comme nous avons procédé ci-dessus, la deuxième dérivée partielle est :

$$\frac{\partial \text{net}(n, j)}{\partial w(n, i, j)} = \frac{\partial}{\partial w(n, i, j)} \left\{ \sum_{i=1}^I [w(n, i, j) \cdot O(n-1, i) + \theta(n, i)] \right\} = O(n-1, i) \quad (\text{IV.29})$$

$O(n-1, i)$ est la sortie réelle du $i^{\text{ème}}$ noeud de la $(n-1)^{\text{ème}}$ couche et $\theta(n, i)$ est un seuil .

Si $O(n-1, i)$ est la couche d'entrée nous devons prendre l'entrée réelle $x(i)$ au lieu de $O(n-1, i)$.

Cette dérivée (expression:IV.29) sera aussi égale à 1 pour le cas de la dérivée partielle en respectant les seuils.

La première dérivée partielle de l'expression (IV.28) est donnée par :

$$\frac{\partial E}{\partial \text{net}(n, j)} = \left[\frac{\partial E}{\partial O(n, j)} \right] \cdot \left[\frac{\partial O(n, j)}{\partial \text{net}(n, j)} \right] \quad (\text{IV.30})$$

La première partie de l'expression (IV.30) sera égale à la somme des dérivées partielles à chaque noeud de la $(n+1)^{\text{ème}}$ couche, qui est la sortie (output) du $j^{\text{ème}}$ noeud de la $n^{\text{ème}}$ couche;

$$-\frac{\partial E}{\partial O(n, j)} = -\sum_{k=1}^K \left[\frac{\partial E}{\partial \text{net}(n+1, k)} \right] \cdot \left[\frac{\partial \text{net}(n+1, k)}{\partial O(n, j)} \right] \quad (\text{IV.31})$$

La deuxième partie de l'expression ci-dessus est la dérivée partielle des sorties (résultats) de la $(n+1)^{\text{ème}}$ couche par rapport aux sorties de la $n^{\text{ème}}$ couche. Puisque les potentiels (désigné par net) sont les sommes pondérées des sorties de la $n^{\text{ème}}$ couche, alors :

$$\text{net}(n+1, k) = \sum_{j=1}^J \{w(n+1, j, k) \cdot O(n, j)\} \quad (\text{IV.32})$$

et dont la dérivée partielle sera égale à;

$$-\frac{\partial \text{net}(n+1, k)}{\partial O(n, j)} = w(n+1, j, k) \quad (\text{IV.33})$$

Nous avons déjà calculé la première partie de l'expression (IV.31) (toutes les étapes entre l'équation (IV.21) et l'équation (IV.24)), ainsi:

$$-\frac{\partial E}{\partial O(n, j)} = \sum_{k=1}^K \{ \delta(n+1, k) \cdot w(n+1, j, k) \} \quad (\text{IV.34})$$

Selon la règle delta généralisée, nous aurons la nouvelle fonction δ :

$$\delta(n, j) = O(n, j) [1 - O(n, j)] \sum_{k=1}^K \{ \delta(n+1, k) \cdot w(n+1, j, k) \} \quad (\text{IV.35})$$

Puisque la mise à jour des poids selon la règle d'adaptation est :

$$\Delta w(n, i, j) = \eta \cdot \delta(n, j) \cdot O(n-1, i) \quad (\text{IV.36})$$

alors, la règle de mise à jour des seuils dans les couches cachées est :

$$\Delta \theta(n, j) = \eta \cdot \delta(n, j) \quad (\text{IV.37})$$

Ceci est répété pour tout le chemin inverse jusqu'à la couche d'entrée, à laquelle les valeurs de la sortie $O(n-1, i)$ de la couche précédente sont remplacées par les valeurs de l'entrée réelle $x(i)$.

Comme nous l'avons vu dans la séquence de calcul, les valeurs $\delta(n, j)$ dans chaque couche, cachée ou autre couche, et pour chaque nœud, peuvent être calculées indépendamment des poids à partir de la couche de sortie calculée récursivement jusqu'à la couche d'entrée. C'est la partie la plus significative du calcul qui a rendu l'optimisation du réseau possible. Une fois que les valeurs $\delta(n, j)$ sont calculées, la mise à jour des valeurs des poids sera facile. Il y a d'autres algorithmes d'optimisation tels que les algorithmes génétiques et la méthode du gradient conjugué.

IV.6.3. Stratégie de l'adaptation des poids et des seuils (mise à jour des poids et des seuils)

Nous calculons les poids $\Delta w(n, i, j)$ pour toutes les données entrées dans la phase d'apprentissage. Ceux-ci seront cumulés et une mise à jour sera faite après que tous les modèles d'apprentissage soient utilisés. Alors la valeur de la mise à jour sera :

$$\Delta w(n, i, j) = \sum_{p=1}^P \Delta w_p(n, i, j) \quad (IV.38)$$

On commence le processus d'apprentissage par des poids et des seuils aléatoires. Initialement les erreurs seront grandes, et les valeurs Δw calculées indiqueront seulement la direction pour une correction appropriée, mais pas les valeurs réelles des poids. Donc nous approchons le minimum avec précaution.

Soient $w_m(n, i, j)$, les poids à la $m^{\text{ième}}$ itération. $\Delta w_m(n, i, j)$ est l'ensemble des mises à jour (updates). Nous pouvons appliquer une contrainte dans chaque étape de mise à jour en ajoutant un terme, utilisant une partie de la mise à jour précédente et l'inclure dans la mise à jour suivante :

$$\Delta w^{m+1}(n, i, j) = \eta \delta(k, j) \cdot O(k-1, i) + \alpha \Delta w^m(n, i, j) \quad (IV.39)$$

α est un facteur. Si α est grand nous serons fortement influencés par les dernières mises à jour. Par conséquent nous aurons une certaine inertie pour suivre une telle direction particulière avec un petit taux de changement. La valeur du δ doit être petite ce qui nous donnera un petit rapprochement au minimum. Si δ est grand, nous pouvons tester de grandes variations dans l'espace des poids. D'autre part, de plus petites valeurs de δ exigent un plus grand nombre d'itérations pour atteindre la solution optimale qui peut être comprise entre deux minimums locaux.

IV.7. Conditions d'une bonne exécution

La modélisation et le training d'un réseau de neurones dépendent de plusieurs paramètres. Principalement, la détermination de sa configuration optimale (sa structure), la distribution optimale des poids d'un réseau spécifique et la garantie d'une bonne exécution sont les plus éminentes. Dans des applications de classification de réservoir, le choix des attributs est un élément important à ajouter aux trois premiers points. Les problèmes rencontrés généralement sont :

- . Une configuration inadéquate du réseau.
- . L'algorithme converge vers un minimum local.
- . Un apprentissage non souhaité.

IV.7.1. Configuration du réseau

Le nombre de neurones d'entrée et de sortie est lié au choix des attributs d'entrée et des classes de la sortie. Il n'y a aucune règle pour déterminer le nombre exact de neurones dans la couche cachée. Le principe de base est semblable aux problèmes linéaires des moindres carrés, c'est-à-dire nous devons avoir 10 fois plus l'apprentissage des échantillons que le nombre de poids et du seuil à calculer. Si nous disposons d'un grand nombre de neurones, le réseau sera sur-adapté. Il atteint la convergence, mais il n'identifiera aucune autre donnée. Il est préférable de commencer à tester avec un petit nombre de neurones.

IV.7.2. Convergence de l'algorithme

La méthode du gradient essaye de diriger la solution vers la minimisation de la fonction d'erreur. Dans les cas non linéaires (exemple : des réseaux avec plus de neurones) il y aura des minimums locaux. L'objectif est d'atteindre le minimum global. Si la fonction d'erreurs est extrêmement complexe, alors il y a une forte probabilité que le réseau converge vers un minimum local. La solution de ne pas atteindre des minimums locaux, est que le gradient ne change pas sa direction rapidement. Cependant, on laisse l'exécution se dérouler plusieurs fois avec différentes valeurs initiales aléatoires et comparer les résultats. C'est un problème commun pour tous les systèmes non linéaires d'optimisation. L'algorithme génétique et la méthode de Monté Carlo sont quelques alternatives possibles. La solution la plus pratique est de réexécuter le réseau de 5 à 10 fois. Des poids initiaux sont générés aléatoirement. Ceci assure que la solution débute à des endroits différents dans l'espace d'erreurs et ira dans différentes directions. Le réseau peut converger vers des différents minimums locaux ou atteindre le minimum global. L'utilisateur choisira le réseau qui a la meilleure convergence, ou la plus petite erreur de RMS.

IV.7.3. Choix d'entrée pour l'apprentissage

Dans notre cas il s'agit du choix des attributs. Les deux problèmes les plus fréquents sont : un excès d'apprentissage ou un choix de données non représentatif pour l'apprentissage. Un excès d'apprentissage se produit quand nous avons trop de poids à calculer (trop de noeuds) et peu d'échantillons de données d'apprentissage. Le principe de base est de 10 pour 1. Nous devrions avoir dix échantillons à utiliser dans l'apprentissage pour chaque poids (un seul poids) ou seuil (un seul seuil). Dans le cas de mauvaises données, nous ne pouvons pas représenter toutes les classes, par conséquent le réseau ne va pas identifier cette classe particulière, ou la combinaison des attributs n'est pas assez significative pour que le réseau puisse identifier les différences.

IV.8. Critère d'un choix optimal de la configuration de réseau

Ce problème est résolu par test, mais ceci prend beaucoup de temps. La première étape est de réduire le nombre de dimensions des données d'entrée. Ceci peut se faire par une analyse en composante principale (ACP). Le rapport de la plus grande valeur propre à la somme des valeurs propres rapporte une évaluation précise de l'information contenue dans les données projetées. Au lieu des données originales, les dimensions de l'entrée peuvent être réduites en employant les données projetées sur quelques axes des composantes principales. Si, cependant, beaucoup de classes différentes existent toujours, alors cette projection peut ne pas être efficace.

L'ACP peut être appliquée entre les poids de chaque couche. Ceci indiquera la signification de la contribution de chaque neurone. Si deux neurones présentent une certaine semblance, alors on peut éliminer l'un des deux. Ceci réduira le nombre d'inconnues d'une manière significative. En outre, si un neurone contribue d'une façon remarquable moins que les autres, il peut également être éliminé et le réseau pourrait être relancé. On observe,

cependant, que pour éviter les minimums locaux il faut réduire le nombre de neurones. En outre, peu de nombre de noeuds signifie une meilleure généralisation de la classification des réseaux de neurones

IV.9. Critères d'arrêt de l'apprentissage

Taner a proposé une nouvelle méthode comme test d'arrêt de l'apprentissage. Il s'est basé sur l'analyse statistique des résidus de l'erreur. Il déclare que l'apprentissage est fait d'une façon optimale si les résidus (différence entre la classe d'apprentissage et la classe prévue) et le bruit ont les mêmes caractéristiques. Si le bruit est supposé blanc, l'apprentissage s'arrêtera quand les résidus auront des moyennes nulles et ne présente aucune corrélation.

IV.10. Séquence de la phase d'apprentissage

1. Introduction des données d'apprentissage et calcul de la sortie du réseau
2. Comparaison du résultat réel (sortie du neurone) avec le résultat désiré (données réelles)
3. Détermination de la direction du changement des poids afin de réduire l'erreur
4. Détermination de la variation des poids (la variation de chaque poids)
5. L'application de la correction de chaque poids
6. Répétition de toute la séquence de 1 à 5 avec les vecteurs de données d'apprentissage, jusqu'à la réduction de l'erreur à une valeur acceptable.

V. Application du PMC et K-SOM comme outils de classification lithologique sur des données réelles

V.1. L'objet de l'application

Cette étude a comme objet l'application d'une classification lithologique dans une région du Sud-Ouest Algérien. Le piège en question est de type structural en présence d'un réseau de failles très complexe. Les objectifs pétroliers dans cette région sont le GIVETIEN et l'ORDOVICIEN.

Notre objectif est de faire une classification lithologique à partir de données sismiques, par une combinaison des attributs sismiques et pour obtenir une coupe lithologique. Cette combinaison est faite par les réseaux de neurones artificiels. C'est une étude de détail dans une région à gaz, puisque la presque la totalité des puits ont mis en évidence une accumulation de gaz.

Nous avons choisi un profil sismique qui passe par trois puits dont deux positifs (GAZ). Ces trois puits ont été forés sur la même structure (GARET EL GUEFOUL).

Le profil utilisé datant des années 70 est de bonne qualité sismique. La résolution sismique est limitée mais les objectifs sismiques ciblés sont clairs.

L'acquisition a été faite par une technique conventionnelle avec un laboratoire de type DFS IV. La couverture est de 1200% et l'inter trace de 100m.

V.2. Données Géologiques.

Le bassin de L'AHNET est caractérisé par une structuration intense. L'image actuelle a principalement été acquise au paléozoïque supérieur (orogénèse hercynienne). D'autres phases de plissements (phase autrichienne fin-jurassique) sont bien marquées par des failles inverses et des flexures importantes, particulièrement dans la partie Nord du bassin avec de sensibles déformations au niveau de la discordance hercynienne, à CARET EL BEFNAT, CARET EL GUEFOUL, HASSI MOUMENE et DJOUA.

En général ces phases post-hercyniennes ont surtout provoqué des mouvements en décrochement avec des plis d'entraînement le long des axes préférentiels.

L'analyse des cartes structurales laisse apparaître trois grands ensembles structurellement bien différenciés, limités par des accidents majeurs subméridiens :

- L'AHNET ORIENTAL.
- L'AHNET OCCIDENTAL
- L'AHNET CENTRAL auquel appartient le permis IN SALAH.

V.2.1. Roches mères

Les principales roches mères sont :

- La Série argileuse du Carbonifère.
- La Série argileuse du Dévonien Supérieur.
- La Série argileuse Carbonatées du Dévonien Moyen et Inférieur.
- La Série argileuse du Silurien.

V.2.2. Roches couvertures

La couverture des réservoirs Tournaisien-Strunien est assurée par les argiles Tournaisiennes et Viséennes. Celle des réservoirs du Dévonien Inférieur par les argiles du Dévonien supérieur, les argiles du Couvinien, de l'Emsien et les alternances argileuses du Gédinnien. Alors que celle du Cambro-Ordovicien est assurée par les argiles de l'Ordovicien et essentiellement l'épaisse série argileuse du Silurien.

V.2.3. Réservoirs Cambro-Ordovicien

En général les réservoirs du Cambro-Ordovicien sont représentés par des grès quartzitiques compacts; les qualités pétrophysiques primaires sont médiocres. Ce réservoir dans l'ensemble de l'Ahnet ne produit que par fracturation.

Dévonien Inférieur

- Gedinnien (zone des alternances) : Cette série est représentée par des alternances de grès et d'argile. Le Gédinnien a produit du gaz sec à DJOUA.
- Siegenien : Représenté par des grès quartzitiques compacts le Siégénien a de mauvaises qualités pétrophysiques primaires. Comme pour le Cambro-Ordovicien cette barre peut produire par fracturation.
- Emsien : L'Emsien gréseux, évolue très rapidement; son épaisseur varie de 2 à 20m. Les qualités matricielles de ce réservoir sont bonnes. L'Emsien a produit du gaz sec à IN SALAH.

Dévonien Moyen

- Givetien : Les calcaires bioclastiques fracturés du GIVETIEN peuvent constituer de bons réservoirs par fracturation.

Dévonien Supérieur-Carbonifère

- Tournaisien-Strunien : En général les qualités pétrophysiques primaires sont médiocres. Les séries gréseuses du Tournaisien- Strunien donnent de faibles débits de gaz sec.

V.3. Analyses des résultats des puits

V.3.1. Le puits GF-1

Grès de l'Emsien

A GF-1, les caractéristiques réservoir de ces grès sont généralement médiocres. Toute fois, la porosité s'améliore localement et atteint. Dans sa partie basale, ce réservoir s'est révélé sec.

Barre Quartzitique du Siégénien

Vu les mauvais paramètres pétrophysiques primaires, les possibilités en hydrocarbures dans ce réservoir sont directement liées à sa fracturation éventuelle. Au puits GF-1, les porosités et perméabilités « carotte » sont très faibles; toutefois, il est à noter l'existence de quelques fissures subverticales, Néanmoins le test effectué dans la partie sommitale de ce réservoir s'est montré sec.

Gédinnien (Zone des Alternances)

Au puits GF-1, les porosités et perméabilités « carottes » sont faibles, un test a produit du gaz en très faible débit et un autre, effectué plus bas, a mis en évidence un faible débit d'eau salée.

Cambro-Ordovicien

Les porosités et perméabilités primaires sont très faibles dans la région; de ce fait, le critère de fracturation s'avère indispensable quant au potentiel en hydrocarbures de ce réservoir. L'évaluation diagraphique a cependant montré des niveaux gréseux imprégnés d'hydrocarbures.

V.3.2. Le puits GF-2

En guise de conclusion, nous essayerons de récapituler les résultats les plus probants apportés par le sondage GF-2. Du point de vue stratigraphique, il y a lieu de signaler la similitude du lithofaciés de la couverture sédimentaire avec les puits de la région. Nous mentionnerons aussi, le creusement probable d'une vallée glaciaire durant la période fin-Ordovicienne. La confirmation de cette hypothèse définira une direction de recherche appropriée, suivant l'axe de cette paléovallée qui serait orientée Nord-Ouest- Sud-Est, car le

changement latéral de faciés en plus des ravinements, fréquents dans ce type de dépôt, sont des facteurs favorables à une possibilité de piégeage des hydrocarbures par variation latérale de perméabilité.

Le puits GF-2, a aussi confirmé sa position régionale haute, qui en a fait le lieu d'élection des hydrocarbures gazeux générés à travers les temps géologiques. Dans le domaine pétrolier, nous retiendrons que malgré la médiocrité des paramètres pétrophysiques, les résultats prouvés par l'objectif Ordovicien demeurent satisfaisants, pour peu qu'un réseau de fissures ouvertes y soit développé[64].

V.3.3. Le puits GF-3

La réalisation du sondage GF-3 a permis une augmentation des réserves prouvées en hydrocarbures dans la structure de Garet El Guefoul influençant positivement sur le futur projet de développement en gaz dans la région de IN-SALAH[63].

V.4. Description structurale de la Structure 'Garet El Guefoul'

Au niveau du toit de l'ordovicien, la structure de Garet El Guefoul, se présente comme une grande voûte dont la configuration est compliquée par l'existence d'un important réseau de failles normales et inverses. Sensiblement allongée Nord-Sud, malgré l'existence d'un axe pratiquement Est-Ouest dans sa terminaison Sud. La structure présente deux culminations dans sa partie centrale : l'une au SW du puits GF-1 culminant à 1805 m et l'autre juste à l'Ouest du puits GF-2 culminant à moins de 1683 m. Le flanc occidental est le plus abrupt, avec une dénivelée moyenne de 140 m par 1 km (Figure : V.4).

Si l'on considère l'isobathe fermée de 2800m, les dimensions de la structure sont de 41/24 km en moyenne avec une amplitude de 1117m, soit une superficie de 966 km². Dans le cas où l'on considère l'isobathe fermée de 2600 m, les dimensions sont de 37/20 km, soit 684 km² avec une amplitude de 917 m.

Le sommet de la structure est divisé en trois blocs :

- Les culminations I et II, séparées par un léger ensellement III. La première culmination I au Sud est affectée par des failles dirigées NW-SE, notamment la faille inverse F1 dont le rejet varie de 50 à 200m. Elle est divisée par ce réseau de failles en quatre (04) petits blocs (I_A, I_B, I_C, I_D), étroits (1 à 2 km), orientés NW-SE. Le bloc central I_A est le plus soulevé; le puits GF-2 a été foré dans les limites de ce bloc et a atteint l'ordovicien à 1683 m de profondeur. Le bloc I_C est à 100m plus bas que I_A. Et I_D, tout au Sud, à 50-60m plus bas que I_C. Le bloc I_B au Nord est le plus bas d'entre eux.

- La culmination II au Nord est délimitée par la faille inverse F2 et la faille normale F4, toutes deux orientées NNW-SSE (Figure : V.4). Le rejet des failles à ce niveau est de 50 à 80m. La faille inverse 2 s'amortit vers le Sud ; son rejet n'est que de 20m au niveau de la ligne sismique 81-TEG- 40.

- L'élément structural III qui se présente comme un sillon séparant les deux culminations précédentes est caractérisé par une configuration plus complexe. Cet élément est réellement constitué par plusieurs petits blocs, subdivisés par de nombreuses petites failles en compartiments soulevés et abaissés, notamment dans sa partie Sud. Trois anomalies positives sont observées sur les profils sismiques 71-TEG-05 (PT 110-111) ,74-TEG-36 (PT132) et 80-GOU-46bis (PT137). La partie Sud de cet élément se présente comme un monoclinale plongeant vers le NW (Figure : V.4).

Concernant le réseau de failles, il faut noter le rôle prédominant des failles inverses notées F1, F2, F2a et F3 (cf. carte en isobathes) dans la configuration de la structure. Ces failles affectent toute la série Paléozoïque, ainsi que les terrains situés au-dessus. L'amplitude des rejets au niveau de la discordance au toit du paléozoïque est égale ou inférieure parfois à

celle des rejets au Dévonien et à l'Ordovicien. Il en est de même pour les failles normales F6 et F7 au Sud et F4 au Nord qui semblent associées aux failles inverses et qui affectent les bordures des compartiments chevauchants. (Figure : V.4) [3]

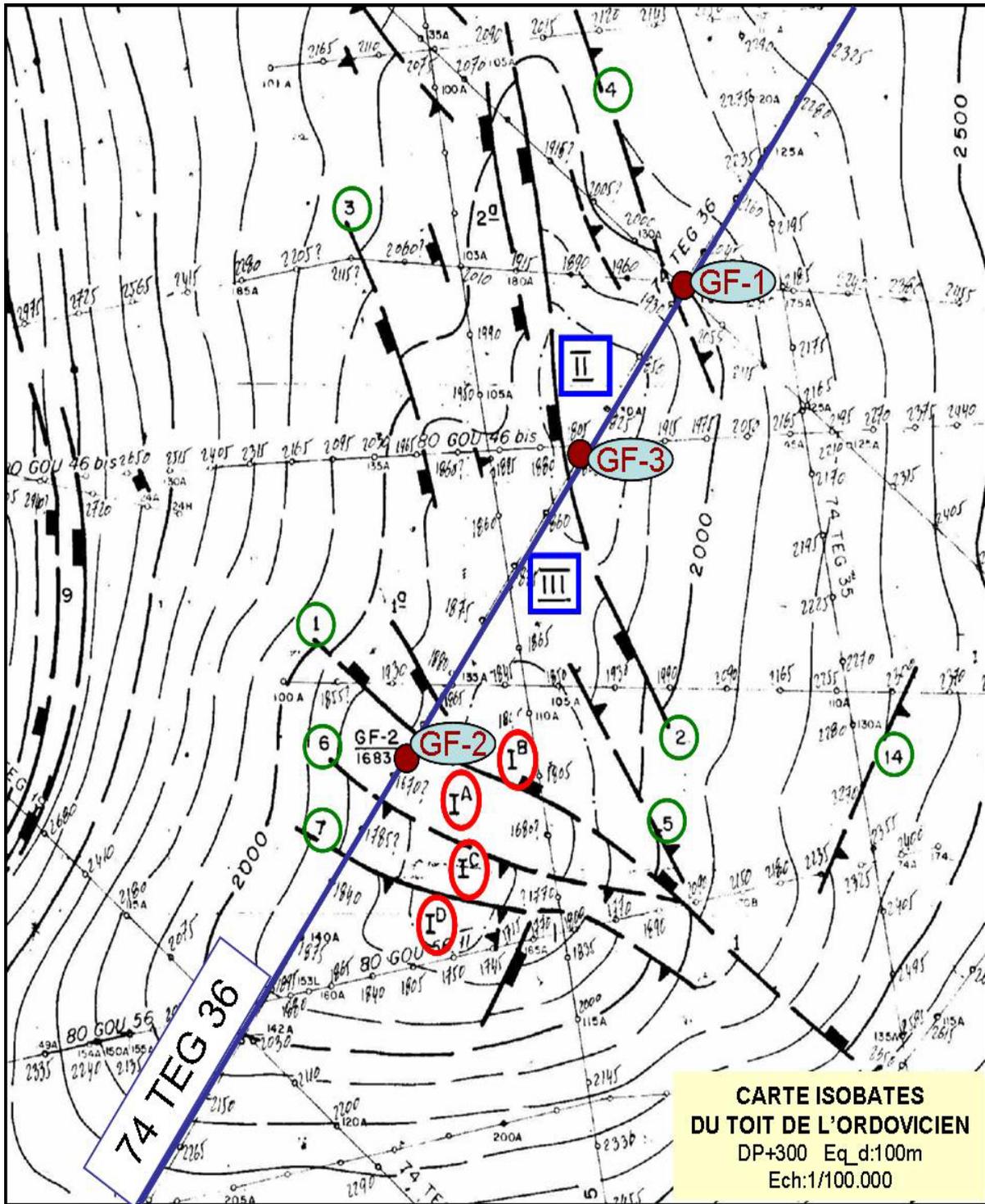


Figure : V.4 Document Sonatrach/Exploration

V.5. Traitement des données sismiques

Le profil utilisé date des années 70. Les données sismiques sont de bonne qualité. La région est connue comme une région à gaz. Le but de notre travail est d'exploiter les données sismiques en terme d'amplitude et de tirer des informations sur la lithologie directement du caractère de ces amplitudes.

Le profil sismique 74 TEG 36 joignant trois puits a été choisi (Figure : V.4). Les différents résultats entre le premier puits d'exploration et le deuxième de délimitation ont fait l'objet d'un troisième puits intercalé entre le premier et le deuxième. Le profil sismique en question a été retraité. La séquence de traitement est composée d'une série de programmes choisis essentiellement pour remédier aux problèmes posés.

Le traitement a été réalisé dans le centre de traitement sismique de la division Exploration par le logiciel GeoVecteur Plus de la Compagnie Générale de Géophysique et Le logiciel RSI du Rock Solid Images.

Nous avons commencé par la résolution du problème des corrections statiques. La technique utilisée se base sur la modélisation de la WZ en utilisant les carottages sismiques verticaux. Le résultat obtenu est satisfaisant.

Le deuxième problème rencontré est celui des analyses de vitesses. Cette région a une particularité de point de vue des vitesses sismiques. Généralement on peut obtenir un bon stack avec une seule loi de vitesses ou conserver la même loi de vitesses pour tout le profil. Pour confirmer cette étape de traitement, nous avons appliqué une migration (en temps) avant sommation.

Une autre phase très importante a été traitée. Elle concerne les bruits aléatoires et les bruits linéaires. Deux programmes d'atténuation des bruits ont été appliqués. Le premier avant sommation et le second après sommation. Le rapport signal sur bruit est très appréciable et le résultat final de bonne qualité.

Nous avons appliqué un traitement en amplitudes préservées car ces données font l'objet d'une classification lithologique qui se base directement sur la qualité de l'amplitude sismique.

Les principaux objectifs pétroliers dans cette région sont le Dévonien Inférieur et l'Ordovicien. La corrélation des horizons est très facile vu le caractère énergétique très fort; la continuité des événements est assurée.

V.5.1. Paramètres d'acquisition

- 1974 Sismique conventionnelle.
 - Couverture 1200 %
 - Inter trace 100m.
 - Labo : DFS IV.
 - Dispositif : 2400m ← 100m ↗ PT ↘ 100m → 2400m
- DP=300m, VC=2200m/s



V.5.2. Séquence de traitement en amplitudes préservées et analyse des résultats

Nous avons résumé toute la séquence de traitement sous forme d'un algorithme tout en prenant en considération l'ordre d'application des différents programmes appliqués (Figure : V.5.2.a.). Les détails des programmes de récupération des pertes d'amplitudes seront présentés dans l'annexe. La planche (Figure : V.5.2.b.) présente quelques illustrations des différentes étapes du QC (Qualité Contrôle).

Le traitement en amplitudes préservées a permis d'obtenir un modèle bien défini sur le plan structural. Nous constatons aussi une meilleure définition du système de failles, sachant que l'un des grands problèmes dans cette région est le système de failles très dense et très compliqué. L'horizon objectif dans ce projet (Ordovicien) est très bien défini et sa continuité est assurée. Selon le caractère des amplitudes, on constate qu'il y a une variation remarquable spatialement. Planche (Planche : V.5.2.c)

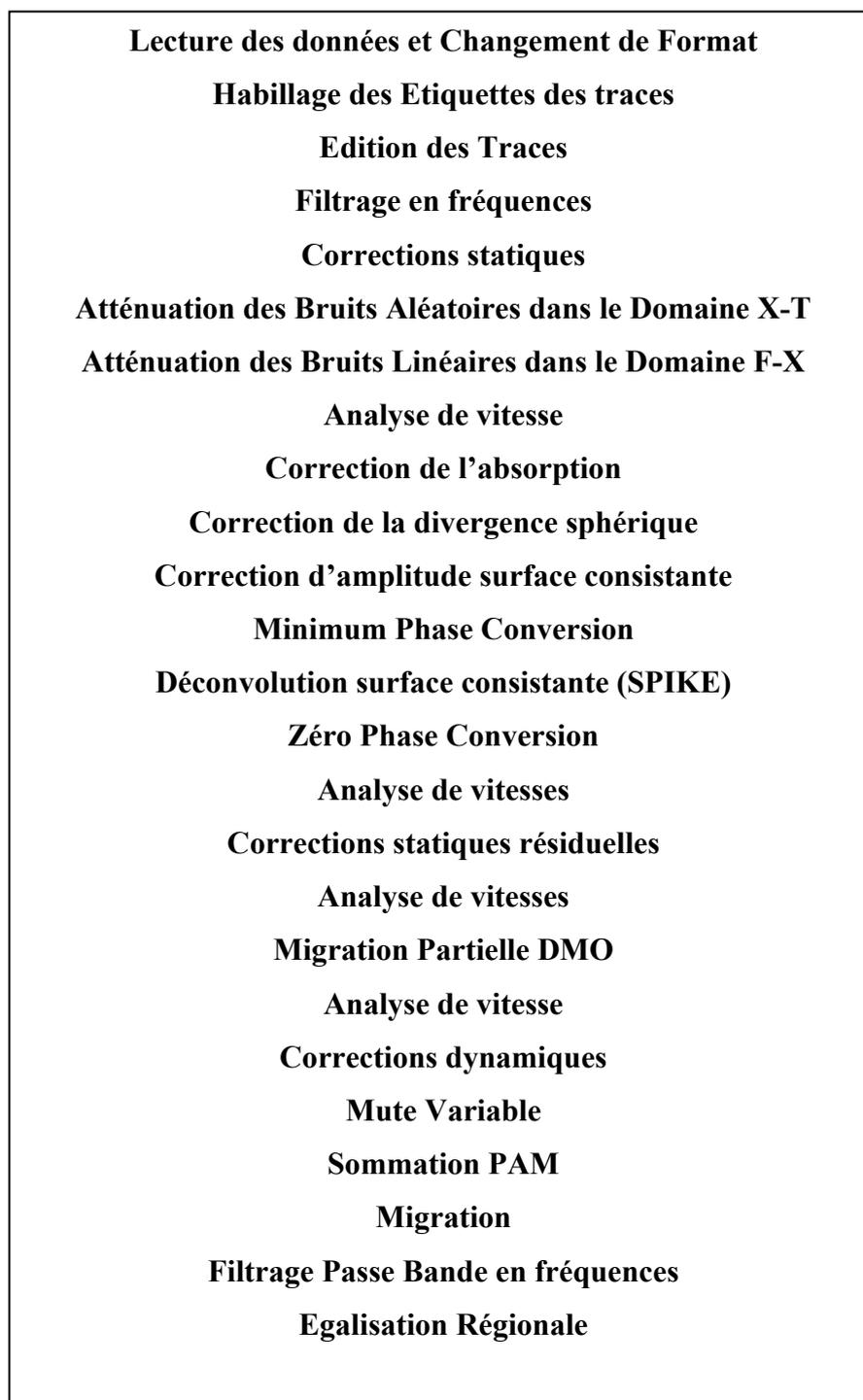
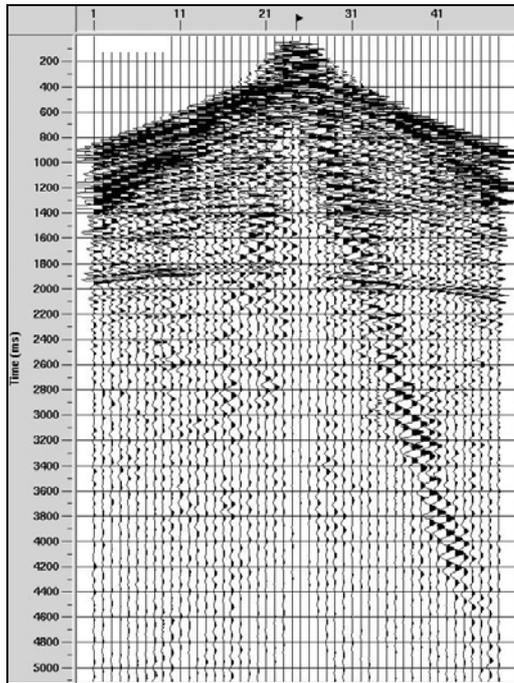
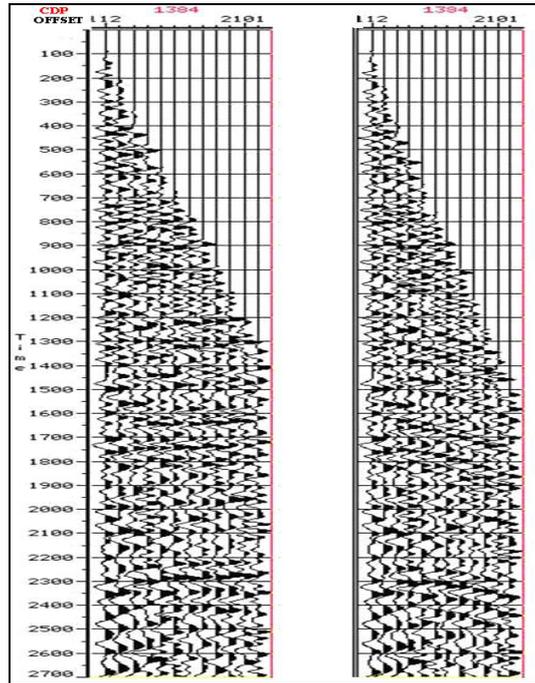


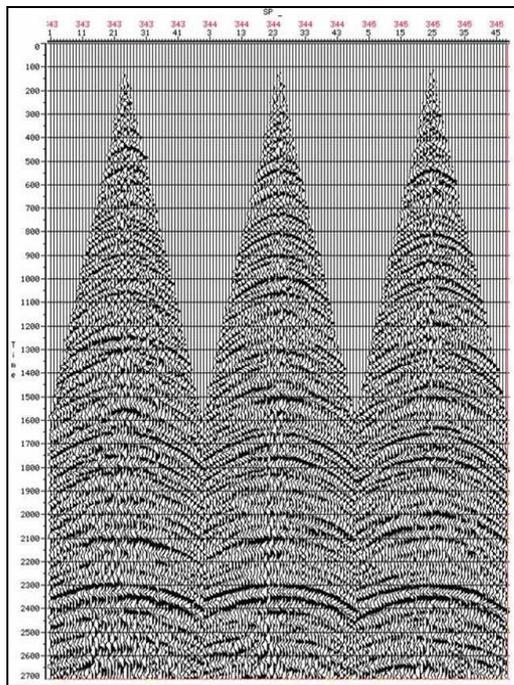
Figure : V.5.2.a. Séquence de Traitement



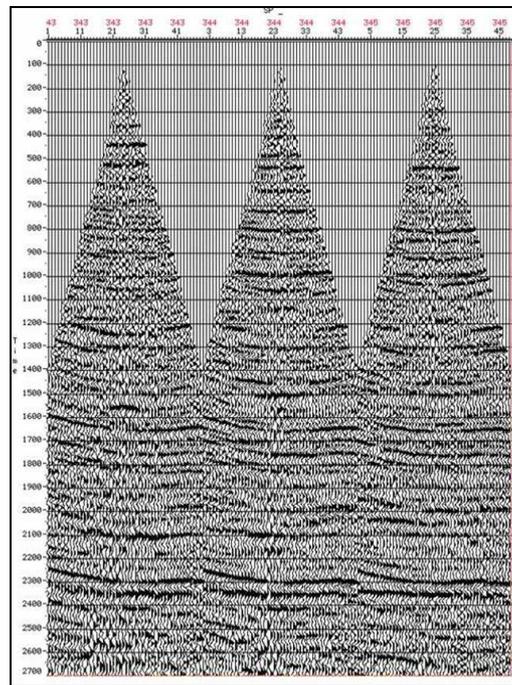
Un point de tir sans aucun traitement



Un CDP après et avant NMO



Trois Points de tirs après filtrage des Bruits et Mute



Trois Points de tirs avec NMO

Figure : V.5.2.b. Contrôle de qualités des différentes étapes de traitement

V.6. Calcul des attributs sismiques

Le but du calcul des attributs sismiques dans ce projet n'est pas une finalité en elle même, mais ils font l'objet des données d'entrée (input) à utiliser dans une autre tâche encore plus compliquée et plus détaillée. Toute classification des attributs sismiques nécessite d'un côté une bonne maîtrise de l'utilité de chaque attribut et d'un autre de chercher l'aspect discriminatoire entre les différents attributs pour réaliser une telle combinaison. Avant toute analyse il faut définir la zone d'étude en temps et en espace. Pour délimiter cette zone et déterminer les différents paramètres de la classification nous avons utilisé en plus de la section en amplitudes préservées, l'amplitude instantanée, la phase instantanée, l'amplitude normalisée, la fréquence instantanée et l'impédance acoustique relative.(Figure :V.6)

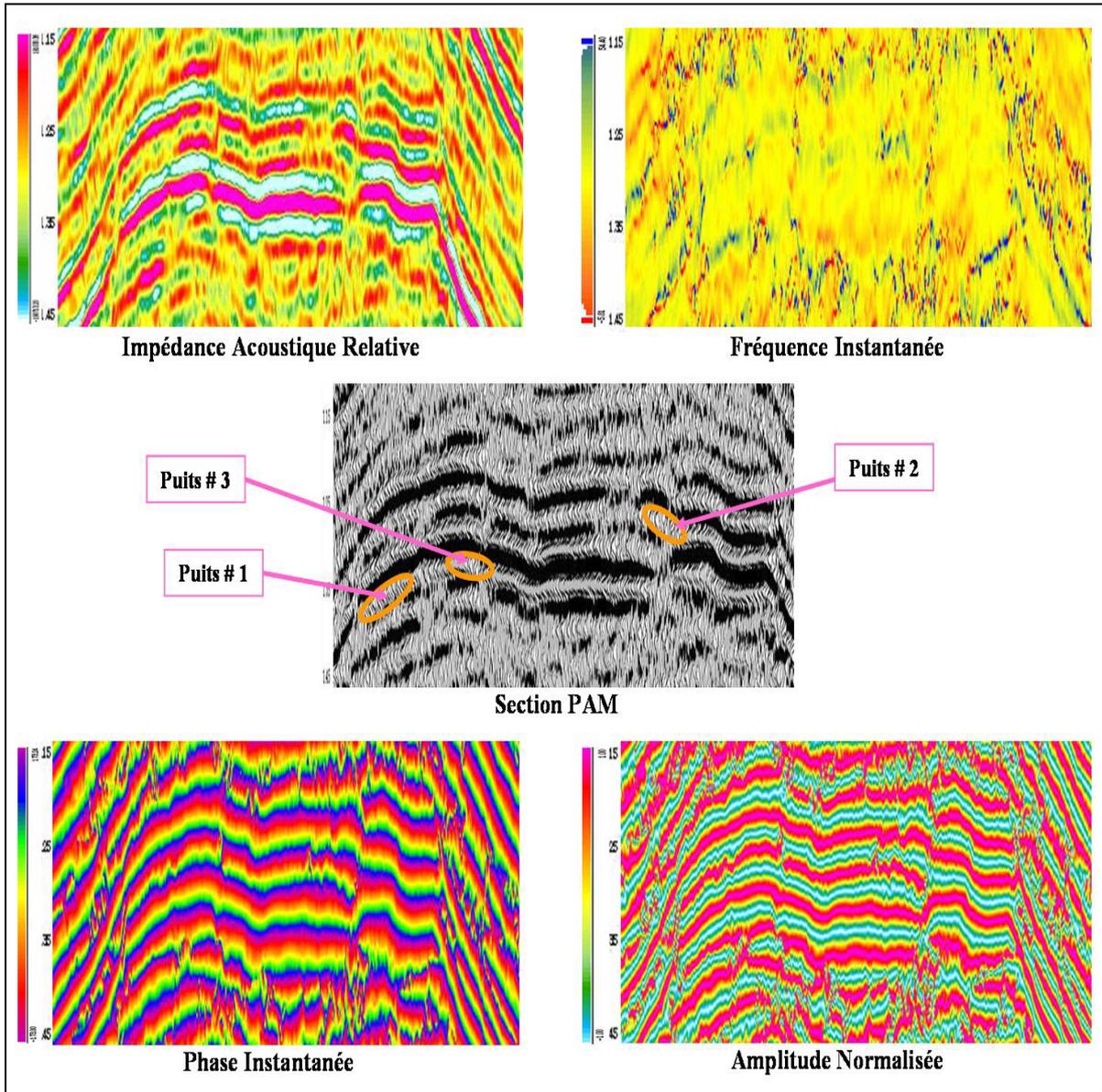


Figure : V. 6. Analyse des attributs sismiques avant classification

V.7. Classification Lithologique

Notre application consiste en une analyse lithologique en utilisant les attributs sismiques comme information brute et les réseaux de neurones artificiels comme outil de classification. Selon le type d'apprentissage, nous avons utilisé le perceptron multicouche (PMC) comme méthode d'apprentissage supervisé et les cartes auto-organisatrices de Kohonen (K-SOM) comme méthode à apprentissage non supervisé.

V.7.1. Classification par le perceptron multicouches (PMC)

V.7.1.1. Choix du réseau

Nous discutons seulement les réseaux statiques, multicouches directs et à apprentissage supervisé. Le modèle statique implique que les poids, une fois déterminés, demeurent fixes. Multicouche direct indique que le résultat (la sortie) ne rétroagit pas, il progresse de l'entrée vers la sortie à sens unique.

La fonction d'activation utilisée est le sigmoïde. Les réseaux neurologiques avec la fonction d'activation de type sigmoïde s'appellent Multi-Layer Perceptron (MLP).

Le gradient permet de diriger la solution vers une minimisation de la fonction erreur. Dans les cas non linéaires (des réseaux avec beaucoup de neurones) il y aura des minimums locaux. L'objectif est d'atteindre le minimum global; dans ce cas la fonction erreur est extrêmement complexe ; il y a alors un risque pour que le réseau converge vers un minimum local. Pour contrôler le gradient et éviter des variations brusques, l'expérience a montré qu'il faut faire quelques tests (5 à 10 fois) au début et prendre la solution qui présente une variation stable du gradient. Les poids initiaux sont aléatoires. Ceci assure que la solution débute en des endroits différents dans l'espace d'erreurs et ira dans des différentes directions. Un réseau qualifié est un réseau qui présente la meilleure convergence, ou la plus petite erreur RMS. Les paramètres du réseau utilisé dans notre projet sont les suivants :

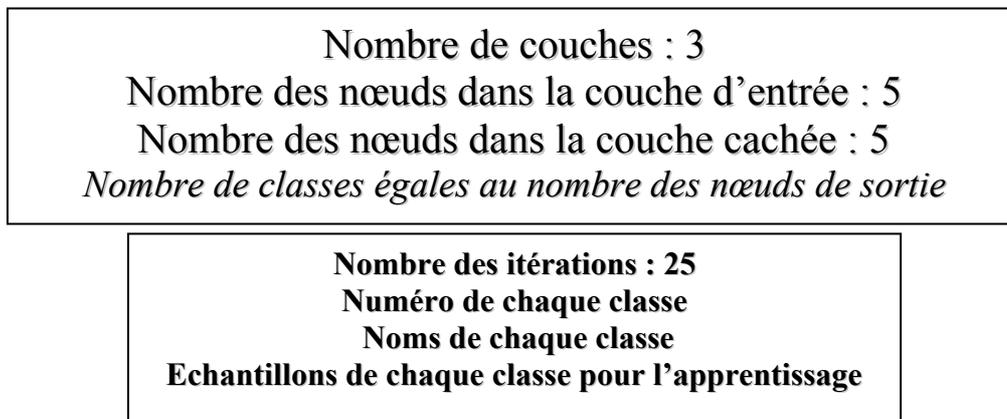


Figure : V.7.1.1. Paramètres du réseau de neurones artificiel (PMC)

V.7.1.2. Choix des attributs sismiques pour l'apprentissage

Dans cette étape les deux problèmes les plus fréquents sont : le premier porte sur l'apprentissage qui est mauvais et le second sur l'ensemble d'attributs non représentatifs dans la phase d'apprentissage. Celui concernant l'apprentissage se produit si nous avons trop de poids à calculer (trop de nœuds) et peu d'échantillons de données pour l'apprentissage. Le principe de base est de 10 pour 1. Nous devrions avoir dix échantillons de données pour chaque poids. Dans le cas où les données sismiques sont de mauvaise qualité, nous ne pouvons pas représenter toutes les classes. Par conséquent, la combinaison des attributs n'est pas assez significative pour que le réseau puisse identifier les différences.

Outil de choix des attributs (Cross-Plot)

C'est un outil de sélection des attributs pour l'apprentissage. Les couples d'attributs qui ne présentent aucune discrimination et qui montrent toutes les données de la même façon ne seront pas utilisés. Le principe de l'utilisation des « Cross Plot » comme outil de sélection des attributs est le suivant :

- Préparation des Cross Plot :
 - Définitions des deux classes à étudier.
 - Sélection de la zone d'étude.
 - Choix des deux attributs à tester.
- Analyse de la figure La figure (V.7.1.2.) :
 - Le symbole carré noir (surface pleine) est la valeur moyenne de la classe une.
 - Le symbole cercle (cercle plein) est la valeur moyenne de la classe deux.
 - Une droite séparant les deux classes a mi-distance entre les valeurs moyennes des deux classes d'une manière discriminante (entre les deux classes).
- Interprétation :
 - Si la distance entre les valeurs moyennes est grande, alors les deux attributs peuvent séparer les deux classes (figure : V.7.1.2.a).
 - Si la distance entre les deux valeurs moyennes est petite, alors ces deux attributs n'ont pas le caractère de discrimination et ne peuvent séparer les deux classes. (figure : V.7.1.2.b)

La figure (V.7.1.2.) montre que pour séparer les deux classes présentées dans ce cas, il faut utiliser la fréquence instantanée et l'Enveloppe. Alors que la Fréquence et l'Impédance acoustique relative ne présentent pas un bon couple discriminatoire.

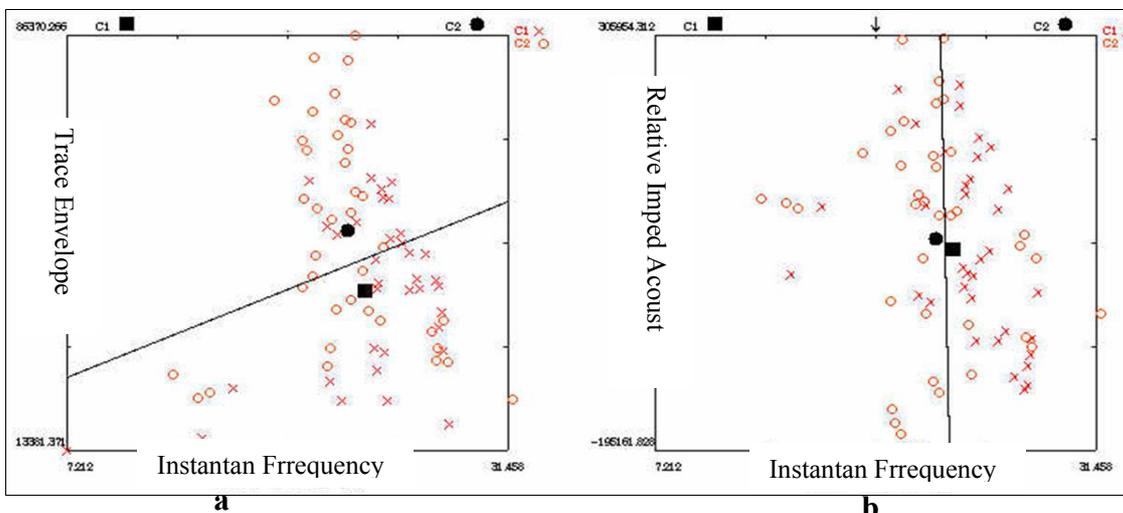


Figure : V.7.1.2. Cross Plot, outil de choix des attributs sismiques

Les attributs choisis pour réaliser cette classification sont les suivants :

- L'Enveloppe de la trace sismique
- La phase instantanée
- La fréquence instantanée
- L'indicateur des couches minces
- L'amplitude normalisée
- L'impédance acoustique relative
- La polarité apparente de l'ondelette

V.7.1.3. Définition des classes

Cinq classes ont été prédéfinies pour classifier le profil en question. Dans cette opération on définit trois paramètres à savoir : le numéro de la classe, le nom de la classe (la formation géologique) et la sélection d'une zone représentative de chaque classe.

La première classe C_1 représente des échantillons dans une partie du réservoir traversé par le premier puits désigné par W_1 . La deuxième classe C_2 représente des échantillons dans une partie du réservoir traversé par le deuxième puits désigné par W_3 . La troisième classe C_3 pour représenter le troisième puits désigné par W_2 . La quatrième classe C_4 caractérise une zone mixte et la cinquième C_5 représente un réflecteur caractérisé par une forte énergie (Figure : V.7.1.3.a).

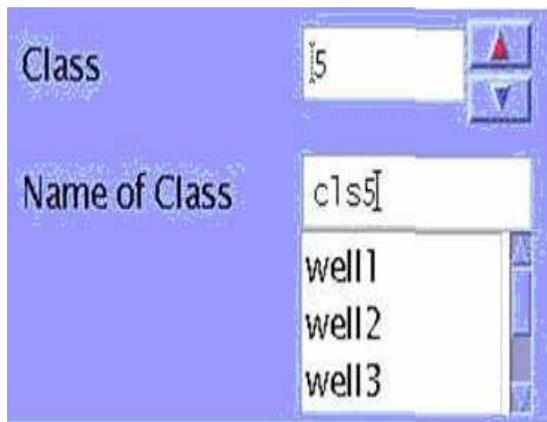


Figure : V.7.1.3.a. Définition des classes pour l'apprentissage

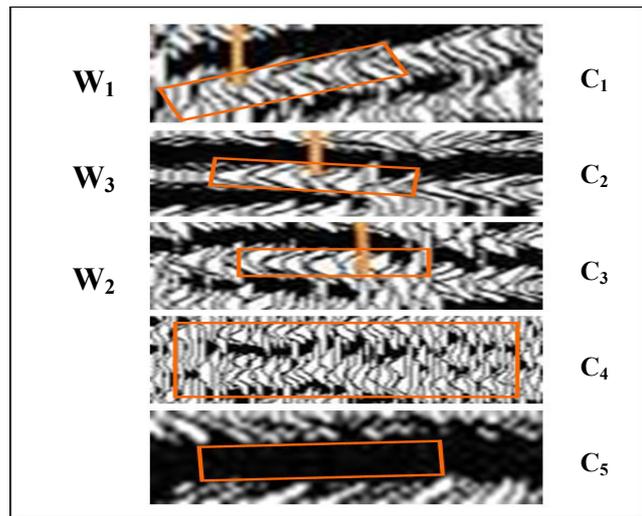


Figure : V.7.1.3.b. Types de classes d'amplitudes définies pour l'apprentissage

V.7.1.4. Critères d'arrêt de l'apprentissage

La méthode appliquée est simple. Elle est basée sur l'analyse statistique des erreurs résiduelles. On déclare qu'un réseau qualifié a été obtenu si les erreurs résiduelles et le bruit ont les mêmes caractéristiques. Les erreurs résiduelles sont définies par la différence entre la classe de l'apprentissage et la classe prévue, le bruit est supposé blanc. On arrête l'apprentissage quand les résiduels auront une moyenne nulle et ne presteront aucune corrélation. Pendant le calcul, le programme liste l'erreur RMS et l'erreur maximale pour chaque passe. (Figure : V.7.1.4.)

RMS ERROR	Iteration	1	-	0.276689.
MAX ERROR	Iteration	1	-	0.968991.
RMS ERROR	Iteration	2	-	0.219753.
MAX ERROR	Iteration	2	-	0.986719.
RMS ERROR	Iteration	3	-	0.212667.
RMS ERROR	Iteration	99	-	0.150060.
MAX ERROR	Iteration	99	-	0.999989.
RMS ERROR	Iteration	100	-	0.144772.
MAX ERROR	Iteration	100	-	0.999999.
RMS ERROR	Iteration	999	-	0.137565.
MAX ERROR	Iteration	999	-	1.000000.
RMS ERROR	Iteration	4999	-	0.133478.
MAX ERROR	Iteration	4999	-	1.000000.
RMS ERROR	Iteration	5000	-	0.140296.
MAX ERROR	Iteration	5000	-	1.000000.
RMS ERROR	Iteration	9999	-	0.134390.
MAX ERROR	Iteration	9999	-	1.000000.
RMS ERROR	Iteration	10000	-	0.136458.
MAX ERROR	Iteration	10000	-	1.000000.
RMS ERROR	Iteration	14999	-	0.133737.
MAX ERROR	Iteration	14999	-	1.000000.
RMS ERROR	Iteration	15000	-	0.134449.
MAX ERROR	Iteration	15000	-	1.000000.
RMS ERROR	Iteration	64999	-	0.130930.
MAX ERROR	Iteration	64999	-	1.000000.

Tableau V.7.1.4. Statistiques et QC de l'analyse par PMC

V.7.1.5. Analyse des résultats de la classification par le « PMC »

Nous partons d'une image qui présente uniquement deux caractères d'amplitude vers une coupe de cinq classes parfaitement séparées (Tableau V.7.1.5.). Cette classification rend l'interprétation facile et précise. La convergence du processus est acceptable car le nombre de classes est le même par rapport aux prévisions. Dans le cas contraire, le résultat de l'analyse peut présenter des classes en plus ou en moins par rapport à celui qui est fixé par les paramètres du réseau. La coupe PMC montre une nette continuité lithologique entre les deux puits W_1 et W_3 . Le contact entre les zones des puits W_1 et W_3 d'un côté et la zone du puits W_2 n'est pas clair structuralement. Le résultat obtenu par cette classification présente une définition détaillée des différentes unités de l'ordovicien (objectif ciblé par cette étude). Dans la zone faillée et en fonction de la distribution des couleurs, nous pouvons facilement faire la corrélation entre les deux compartiments de part et d'autre de la faille (Figure V.7.1.5.).

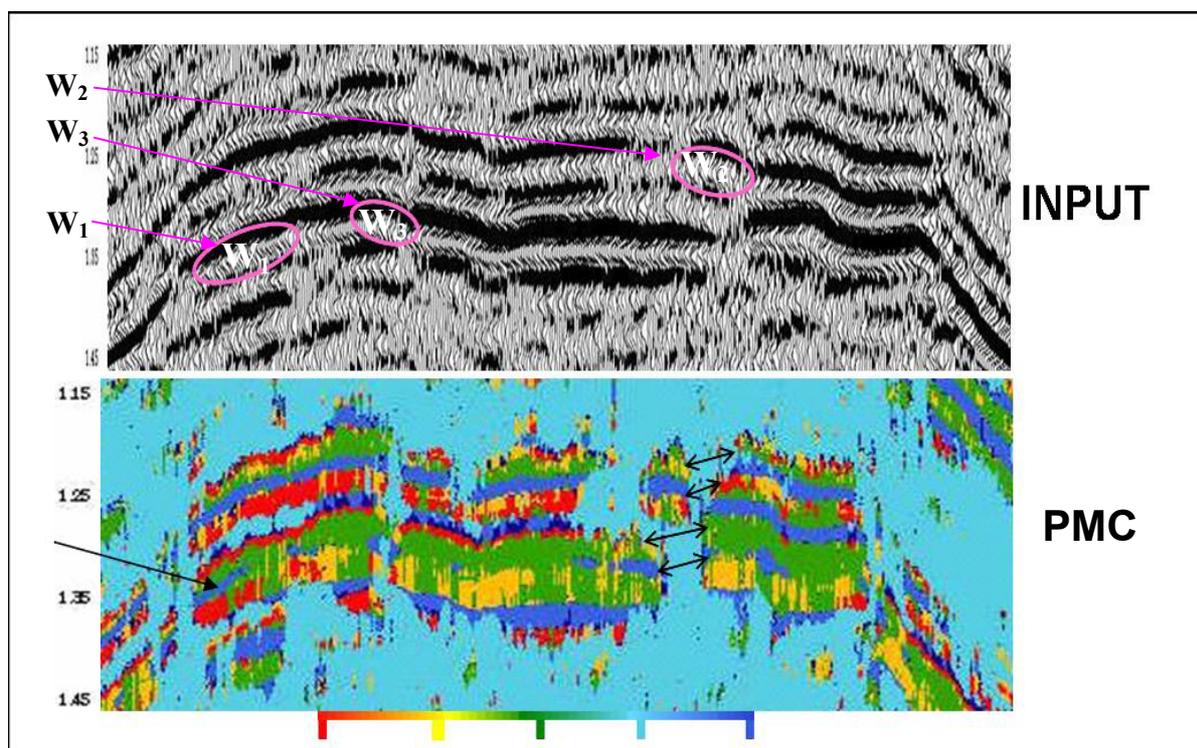


Figure V.7.1.5. Coupe lithologique après analyse par PMC

V.7.2. Classification par les Cartes Auto Organisatrice de Kohonen (K-SOM)

Les Cartes Auto Organisatrices de Kohonen (K-SOM, Kohonen's Self Organizing Maps) se composent d'un ensemble de neurones, placés côte à côte d'une façon organisée. Chaque neurone a des poids associés. Ces poids sont employés pour calculer une certaine forme de similitude entre le vecteur d'entrée et chacun des neurones. Une grande valeur du produit scalaire entre le vecteur de données à classifier et le vecteur des poids associé à un neurone donné ou une petite valeur de la distance euclidienne entre ces deux vecteurs indique le degré de similitude.

V.7.2.1. Définition des Paramètres

Une bonne classification par la méthode SOM nécessite quelques conditions que nous résumons comme suit (Figure V.7.2.1) :

- Les données utilisées pour l'apprentissage doivent contenir toutes les différentes classes qui peuvent exister dans les données à traiter. (Sélection de la zone d'apprentissage)
- Les attributs utilisés pour réaliser la classification doivent avoir des possibilités de

discrimination pour identifier les différentes classes. (Choix des attributs)

- Le nombre de neurones et la configuration de carte dépendront du type de problème posé. Si la configuration lithologique est simple, une structure linéaire du réseau (unidimensionnelle, 1-D) peut suffire. Il a été montré que, dans des cas plus compliqués une carte bidimensionnelle (2-D) est préférable. (Définition des dimensions de la carte)

- Le nombre de neurones représente la dynamique du réseau pour définir les diverses classes. Devant un cas caractérisé par un petit nombre de classes, un petit nombre de neurones sera suffisant. Dans la plupart des cas de classification des réservoirs, l'intérêt principal est toujours lié aux réservoirs producteurs. Chaque réservoir représente un horizon et donc la zone à classifier représentera une petite partie de la section totale. Ceci suggère que nous devons avoir un plus grand nombre de neurones. Ainsi, des classes avec une petite densité seront séparés correctement. (Définition du nombre de Neurones)

Comme indiqué dans la liste des paramètres, plusieurs paramètres doivent être choisis correctement pour obtenir la convergence désirée. Comme pour la plupart des autres techniques des réseaux de neurones, la technique SOM est une science d'une part et un art de l'autre. Tous les paramètres dépendront du type de problème réel et doivent être choisis par expérience.

La classification par les cartes de Kohonen (K-SOM) a deux phases importantes; La première phase de l'apprentissage s'appelle " phase de commande ". Dans cette phase on définit la topologie et tous les paramètres du réseau. La phase suivante appelée la " phase de convergence ", est la phase de classification proprement dite, au cours de laquelle on observe les différents résultats jusqu'à la convergence voulue.

Type	2-D		
X-Rows	8		
Y-Columns	8		
Number of Passes	25		
Starting Radius	6.0		
Radius Decrement	0.98		
Y-Columns	8		
Number of Passes	25		
Starting Radius	6.0		
Radius Decrement	0.98		
Minimum Radius	1.5		
Start Learning Rate	0.01		
Weights Scale	1.0		
Confidence Limit %	50		
Conscience Limit (Percent)	2		
Gaussian Curve			
Bias Scalar	5		
<input type="checkbox"/> Use PCA Factors (2D)			
<input checked="" type="checkbox"/> Save Segy Files			
First Cdp	Last Cdp	Starting Time	Ending Time
400	1400	250	2300

Figure V.7.2.1. Définition des paramètres pour la classification par K-SOM

V.7.2.2. Choix des attributs (Kohonen Attributes Planes)

C'est un moyen de contrôle permettant la sélection et le filtrage des attributs à utiliser dans la classification. Deux attributs qui présentent la même distribution des différentes classes ne peuvent être utilisés à la fois. On ne considère qu'un seul discriminateur parmi ces deux attributs. La figure (V.7.2.2.) présente le résultat de cette distribution. Après analyse nous avons observé que la fréquence instantanée et l'indicateur des couches minces (thin-bed-indicateur) présentent la même distribution des classes, par conséquent nous avons décidé d'éliminer le deuxième attribut.

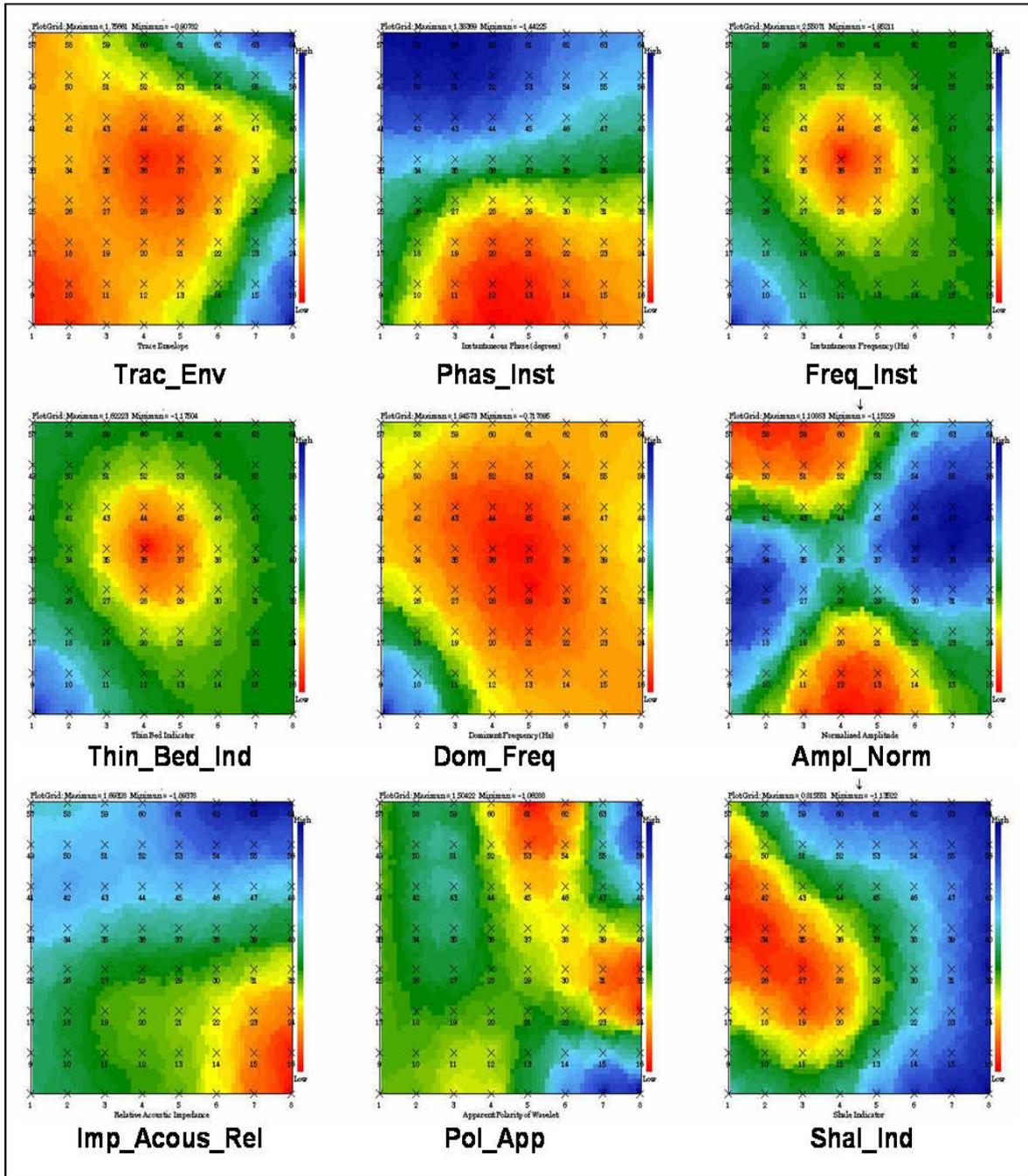


Figure V.7.2.2. Choix des attributs à utiliser par K-SOM

V.7.2.3. Contrôle du processus d'apprentissage et analyse des résultats.

Après chaque itération une nouvelle carte de densité de population des différentes classes sera présentée. La distribution de population est un bon indicateur pour suivre le progrès de la classification. La figure (V.7.2.3.a) montre quatre cartes en deux dimensions. La première représente l'état de la carte de Kohonen au début du processus après la cinquième itération. Au début cette carte prend la forme d'un rectangle. Elle est composée de huit lignes et huit colonnes en fonction de la définition des paramètres du réseau. La forme de cette carte change en fonction de l'évolution du processus de la classification. Dans le cas idéal tous les nœuds convergent vers un seul point d'où la notion de neurone gagnant.

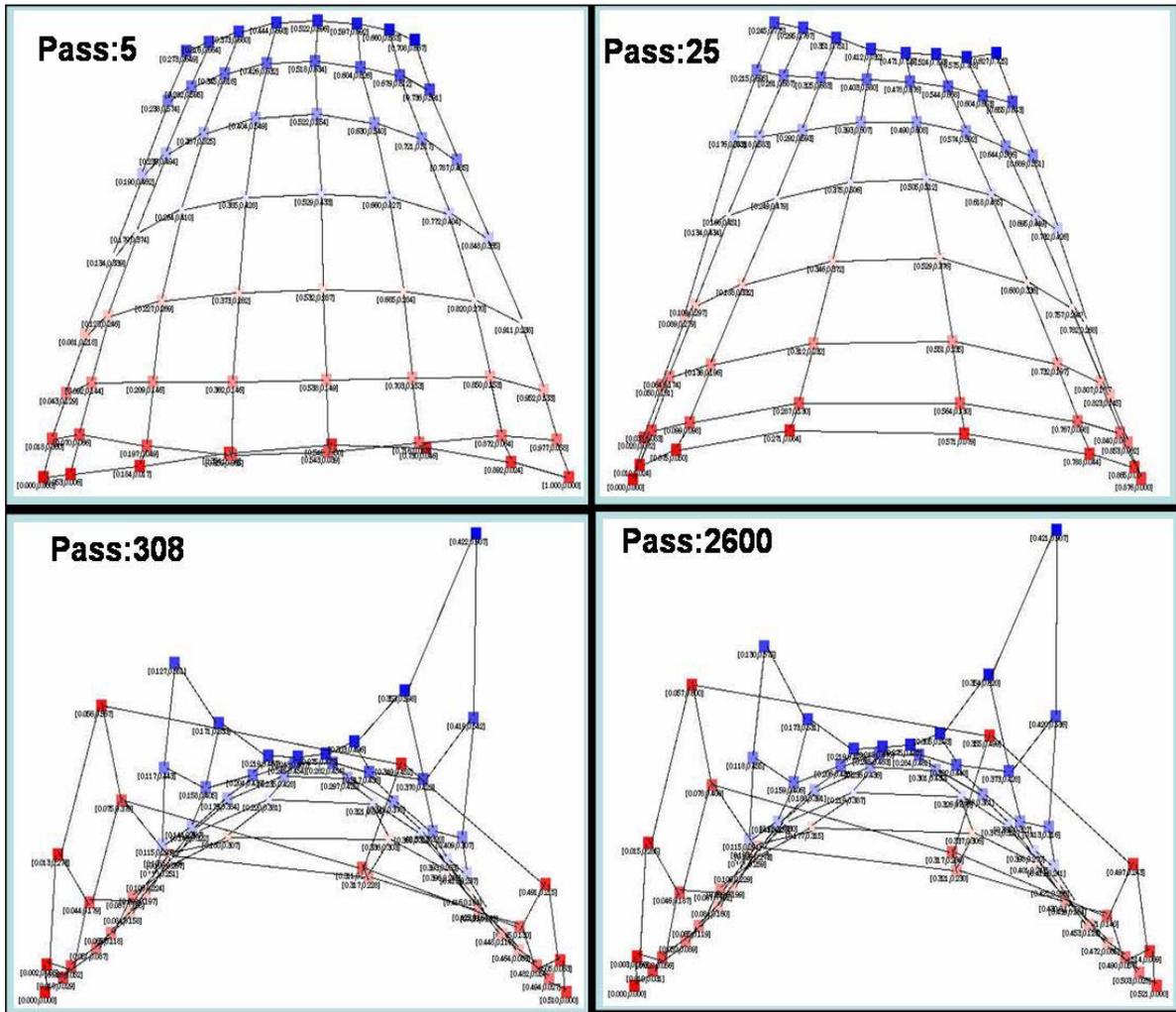


Figure V.7.2.3.a. Evolution du réseau K-SOM

En plus du résultat graphique, d'autres paramètres sont également imprimés pour le contrôle et la surveillance de l'évolution du réseau. Le pas d'apprentissage, la correction du rayon entre deux neurones voisins, la largeur de la base de la carte SOM et l'erreur RMS de chaque itération sont les principaux paramètres de contrôle utilisés (Figure V.7.2.3.b).

Learning Rate - 0.010000.
RADIUS ---- 6.000000. Conscience ON.
BASE ----- 8695.961914.
RMS ERROR Pass 1 - 788329.562500.
Learning Rate - 0.009531.
RADIUS ---- 3.694683. Conscience ON.
BASE ----- 330361.656250.
RMS ERROR Pass 25 - 510494.156250.
Learning Rate - 0.005409.
RADIUS ---- 1.500000. Conscience ON
BASE ----- 3.125302.
RMS ERROR Pass 308 - 1.522260.
Learning Rate - 0.001000.
RADIUS ---- 1.500000. Conscience ON
BASE ----- 3.131869.
RMS ERROR Pass 2600 - 1.521138.

Tableau : V.7.2.3.b. Statistiques et QC de l'analyse par K-SOM

Le nombre de couleurs utilisées indique le nombre de classes trouvées dans une coupe lithologique. Une coupe lithologique présentée par cinq couleurs montre que la section sismique a été classifiée en cinq classes distinctes. L'interprétation de la coupe obtenue par la K-SOM (Figure V.7.2.3.c.) est plus facile que celle de section sismique brute. Pour la classification par K-SOM, le nombre de classes n'est pas défini préalablement. Le résultat obtenu pour notre cas d'étude est jugé appréciable, car le nombre de classes est de l'ordre de cinq. Ce même nombre a été défini pour l'apprentissage du réseau PCM.

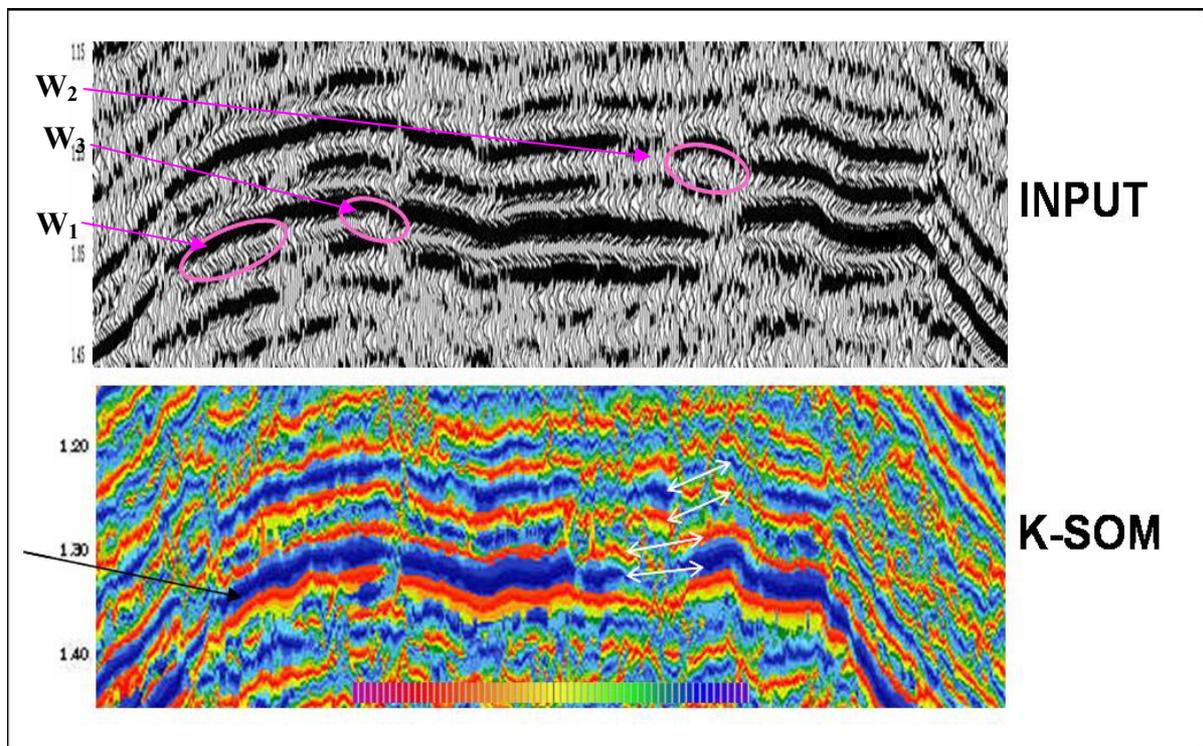


Figure V.7.2.3.c. Coupe lithologique selon K-SOM

V.7.3. Comparaison ente les résultats des deux méthodes PMC et K-SOM

Les informations obtenues par les deux applications PMC (perceptron multi couches) et K-SOM présentent un grand intérêt par rapport à la section sismique brute.

Dans les zones caractérisées par un faible rapport signal sur bruit (mauvaise continuité, perte de l'énergie...), la technique K-SOM donne de bons résultats par rapport au PMC. Nous pouvons observer cette différence à gauche de la position des flèches sur la figure(V.7.3.a)

D'un point de vue résolution et détails des classes, les résultats du PMC sont plus précis par rapport à ceux de K-SOM. La zone indiquée par une flèche (figure :V.7.3.a) représente un banc de faibles dimensions sur la coupe obtenue par le PMC alors que la technique K-SOM ne montre pas cette particularité.

Dans la zone faillée à droite de la section, le PMC ne montre aucun contact entre les deux compartiments soulevé et affaissé. La K-SOM assure le lien de part et d'autre de la faille.

Quoique le nombre de classes n'a pas été défini préalablement pour la méthode K-SOM, le résultat présente cinq couleurs qui indiquent cinq classes lithologiques. Ce résultat est confirmé par le résultat de la méthode PMC.

La distribution des couleurs au toit de la formation ciblée par les forages montre des intercalations fines avec des dispositions sous forme d'un nuage de points qui peut représenter des formations argileuses. Ce résultat est mieux indiqué par le PMC que par la K-SOM.

D'un point de vue lithologique, les trois zones forées (même formation) pressentent le même comportement. Mais la présence de la faille a rendu l'interprétation structurale de la sismique très difficile. Dans ce genre de situation ce type d'analyse et de classification devient nécessaire et indispensable pour minimiser le risque et rendre la décision d'implantation facile.(voir planches ; Planche : V.7.3.b , Planche : V.7.3.c et Planche : V.7.3.d)

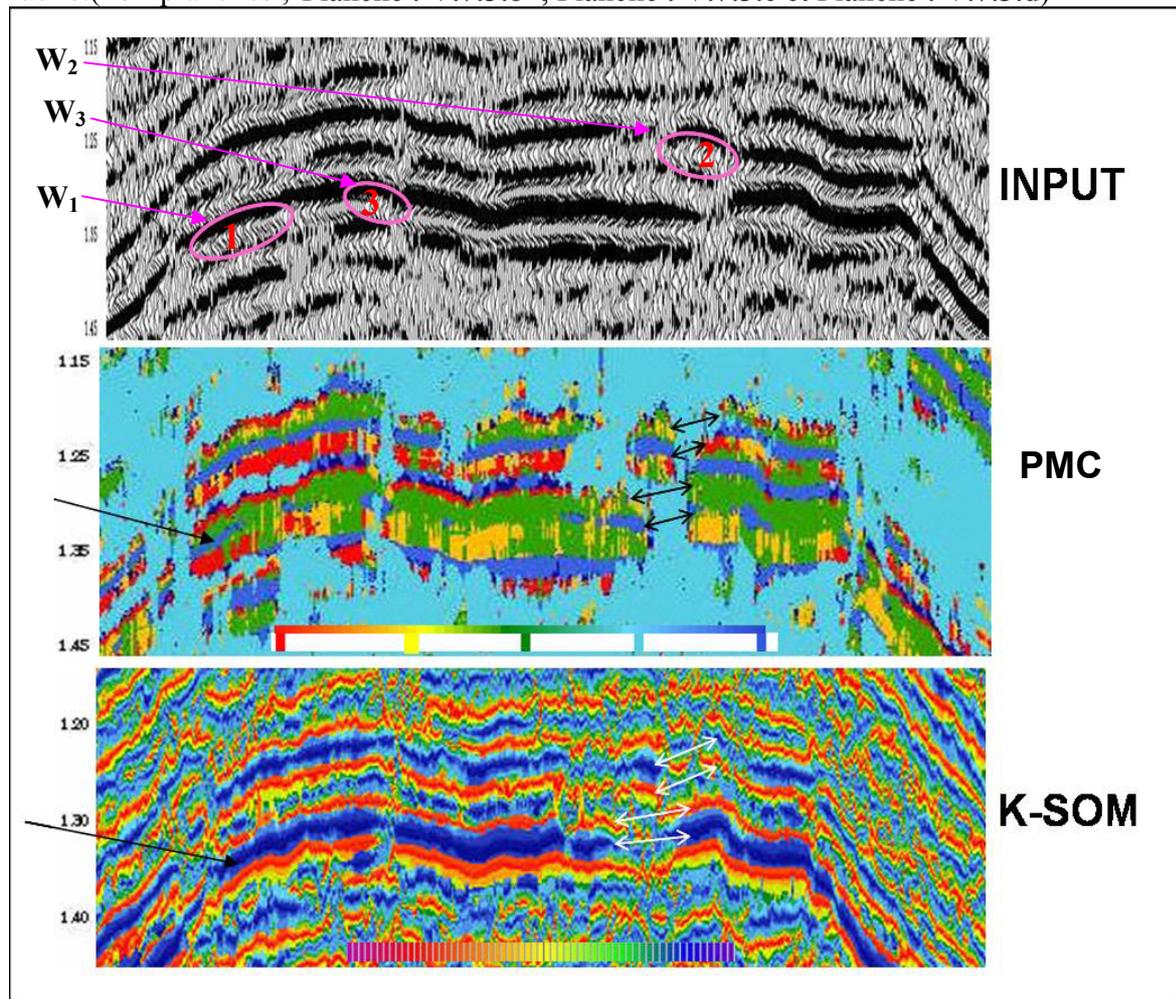


Figure : V.7.3.a. Comparaison entre les résultats du PMC et de la K-SOM

CONCLUSION GENERALE

L'intégration des données de puits dans la caractérisation des réservoirs en utilisant les réseaux de neurones artificiels exige une méthode à apprentissage supervisé.

Le rapport entre les attributs sismiques et la lithologie est compliqué et non linéaire, et il est difficile de prédire ce rapport avec des méthodes simples. Le perceptron multicouches est la technique la plus recommandée pour ce genre de problème.

La plupart des algorithmes d'apprentissages permettent d'obtenir les poids par minimisation d'une fonction de coût dérivable. La méthode la plus simple pour procéder à une minimisation est la méthode du gradient.

La mise à jour de l'ensemble des poids d'une couche nécessite la connaissance des erreurs associées à chaque neurone de la couche suivante. On appliquera l'algorithme d'ajustement des poids en partant de la dernière couche (pour laquelle les erreurs sont connues) vers la première, d'où l'appellation de cet algorithme : Algorithme de rétropropagation du gradient de l'erreur « backpropagation Algorithme ».

Un bon fonctionnement du perceptron multicouches avec l'algorithme de rétropropagation comme classifieur des données sismiques nécessite :

- Une configuration adéquate du réseau.
- L'algorithme doit converger vers un minimum global.
- Un bon choix des données d'apprentissage.
- Un apprentissage souhaité.
- Un bon choix des attributs sismiques utilisés dans la classification.

Les cartes de Kohonen permettent d'avoir une coupe lithologique en exploitant les différentes réponses des attributs sismiques extraits des données sismiques brutes (Profils 2D ou Cube 3D).

L'absence de données de puits dans certaines régions oblige les explorateurs à utiliser des techniques à apprentissage non supervisé. Dans ce cas, on fait souvent appel aux cartes de Kohonen.

Par rapport aux algorithmes de classification classiques, la propriété essentielle des cartes auto-organisatrices est la conservation de la notion de voisinage. Les données voisines qui se ressemblent, sont référées par des neurones voisins dans le réseau.

La classification par les cartes de Kohonen présente l'avantage de fournir une représentation réduite des données :

- Le nombre d'échantillons est réduit au nombre de neurones choisi a priori.
- Les données en dimension N (généralement assez grand) sont représentées dans un réseau dont la topologie est généralement à une ou à deux dimensions. La sélection de la zone d'apprentissage, le choix des attributs, la définition des dimensions de la carte et la définition du nombre de neurones sont les paramètres essentiels dans une classification par les cartes auto-organisatrices de Kohonen

Toute classification des attributs sismiques nécessite d'une part une bonne maîtrise de l'utilité de chaque attribut et de l'autre de chercher l'aspect discriminatoire entre les différents attributs pour réaliser une telle combinaison. Avant toute analyse il faut définir la zone d'étude en temps et en espace.

Les informations obtenues par les deux applications PMC et K-SOM présentent un grand intérêt par rapport à la sismique brute.

Dans les zones caractérisées par un faible rapport signal sur bruit (mauvaise continuité, perte d'énergie ...), les cartes de Kohonen donnent de bons résultats par rapport au perceptron.

Du point de vue résolution, les résultats du perceptron sont plus précis par rapport à celle des cartes de Kohonen.

Dans la zone faillée le PMC ne montre aucun contact entre les deux compartiments. Et la K-SOM assure le lien de part et d'autre de la faille.

Quoique le nombre de classes n'a pas été défini préalablement dans la méthode K-SOM, le résultat présente cinq couleurs qui représentent cinq classes lithologiques. Ce résultat est confirmé par le résultat de la méthode PMC.

La distribution des couleurs au toit de la formation ciblée par les forages montre des intercalations fines avec des dispositions sous forme d'un nuage de points qui peut représenter des formations argileuses. Ce résultat est mieux indiqué par le perceptron que par les cartes de Kohonen.

Du point de vue lithologique, les trois zones forées (même formation) présentent le même comportement. Mais la présence de la faille a rendu l'interprétation structurale très difficile. Dans une telle situation ce type d'analyse et de classification devient nécessaire et indispensable pour rendre la décision d'implantation facile.

Dans la phase d'apprentissage on a constaté qu'un sur-apprentissage donne de mauvais résultats, de même que l'utilisation des attributs non représentatifs du modèle étudié affecte la classification.

Annexe I. Préservation des l'amplitudes dans un traitement sismique.

Le traitement en amplitude préservé à pour but de donner des informations sur les caractéristiques lithologique des formations géologiques. il consiste à compenser l'effet de certains facteur qui affectent les amplitude de l'onde sismique, à atténuer certains bruits et à améliorer la résolution.

Une séquence PAM (preserved amplitudes) ne comporte généralement pas d'égalisation, sauf au stade final de représentation où l'on utilisera une égalisation régionale .cependant, pour certains modules, il est préférable de travailler avec des traces égalisées afin que le niveau d'amplitude moyen soit constant sur toute la trace. C'est le cas pour les programmes d'analyse de vitesses et les programmes de calcul de corrections statiques résiduelles (pour optimiser le ponté).On emploiera donc ces programmes dans des séquences intermédiaires pour définir les corrections statiques résiduelles et pour bien optimiser les vitesse de sommation.

A.I.1. Programmes de correction d'amplitude

Les programmes de traitement des données sismiques s'attachent en général à améliorer le rapport signal sur bruit S/B (filtrages, anti-bruit, sommation) et à repositionner les événements d'une section en temps et en espace (correction statiques, NMO, migration).

Les programmes de traitement des amplitudes peuvent être classés en trois grandes catégories en fonction de leurs objectifs.

Programmes destinés à améliorer le rapport signal sur bruit.

Programmes destinés à compenser les variations d'amplitude qui ne sont pas directement liées aux variations des coefficients de réflexion.

Programmes destinés à rendre compatible la très grande dynamique des amplitudes sismiques avec celle des plotters.

A.I.1.2. Correction de l'effet de la divergence sphérique

Cet effet est l'expression du principe de conservation de l'énergie appliquée à l'énergie dissipée sur un front d'onde ; cette invariance de la densité d'énergie par unité de surface impose une décroissance de l'amplitude proportionnelle à la longueur du rai parcouru. Dans un milieu homogène (à vitesse constante, Figure : A.I.1.2.a), les fronts d'ondes sont des portions de sphères et la décroissance d'amplitude est exprimée par $1/VT$. Si l'amplitude à un point quelconque est indiquée par A , alors donc A est proportionnelle à $1/VT$. Le produit VT est la distance parcourue. Avec V comme vitesse RMS et T le temps de parcours.

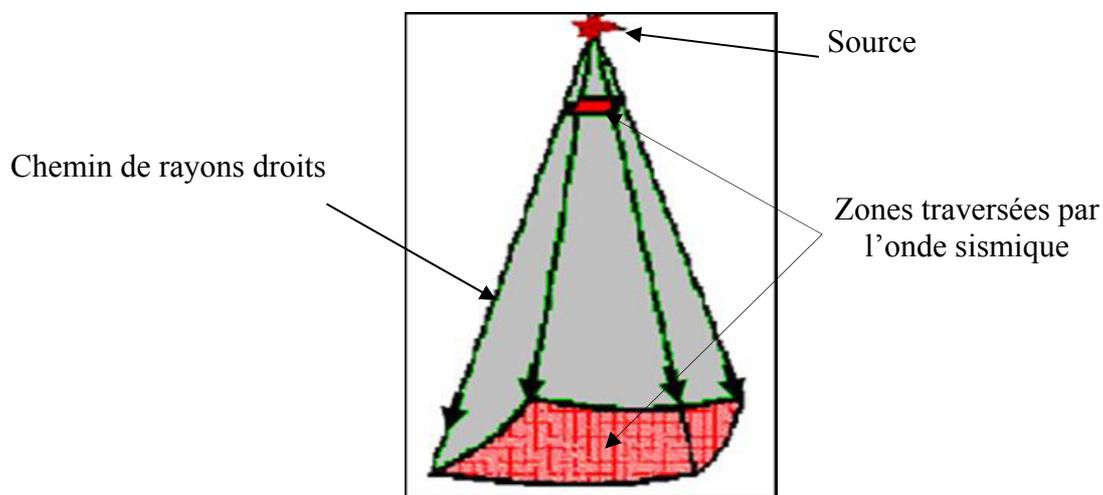


Figure : A.I.1.2.a Milieu à vitesse constante

Pour les milieux à vitesse varie avec la profondeur, les fronts d'ondes sont des surfaces quelconques et il n'existe pas d'expression analytique exacte pour la divergence sphérique (Figure : A.I.1.2.b) Newmann (1973) a établie que dans un modèle stratifié plan et à offset nul, la décroissance d'amplitude n'était en $1/V^2T$ ou V est la vitesse RMS et T est le temps de parcours.

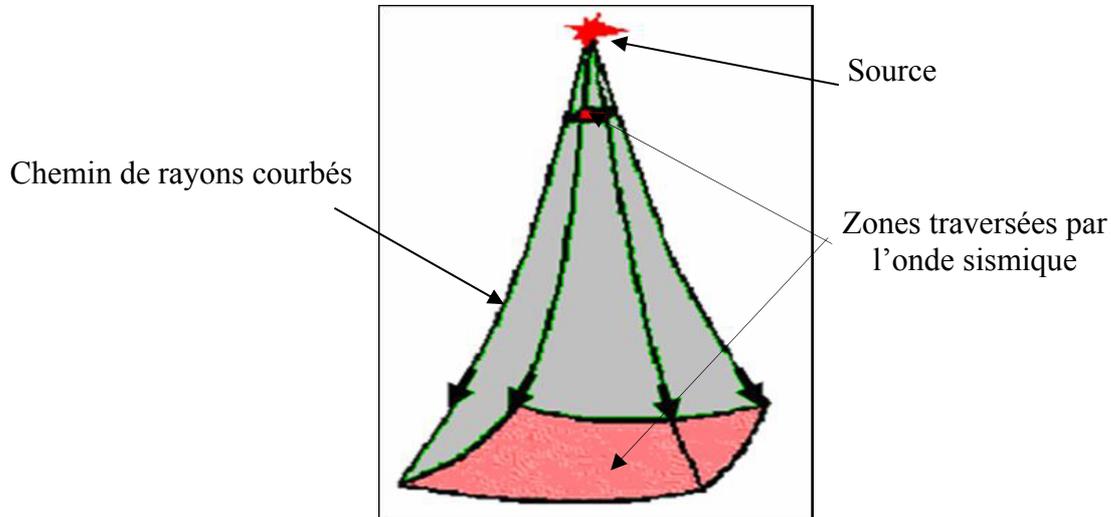


Figure : A.I.1.2.b. Milieu où la vitesse varie avec la profondeur

Le programme appliqué pour corriger l'effet de la divergence sphérique est basé sur l'approximation de Newmann et nécessite la connaissance de la loi de vitesse de sommation. Cette connaissance n'étant généralement pas acquise en début de traitement. Dans la pratique on applique une correction temporaire au début du traitement et on applique la correction définitive après une bonne estimation des vitesses de sommation.

La principe de Newmann utilisé pour compenser les effets de la divergence sphérique se base sur la multiplication de chaque échantillon de la trace sismique par :

$$T.(V^2T_0) \quad (A.I.1)$$

Avec ; T : est le temps de trajet Source-Géophone.

T_0 : est le temps de trajet à offset nul.

V : vitesse de sommation pour le CDP concerné.

Avant l'application de la correction NMO, T : le temps de trajet t

T_0 : le temps t corrigé de NMO

Après l'application de la correction NMO, T : le temps avant NMO

T_0 : le temps t

A.I.1.2. Correction de l'effet de couplage

En sismique terrestre, de nombreux effets de surface sont susceptibles de modifier les conditions d'acquisition. La variation de l'épaisseur de la WZ, la différence de puissance de la source d'un point à un autre, la variation des conditions de couplages source-sol et géophone-sol, sensibilité des géophones, sont autant de facteurs qui peuvent altérer sensiblement les amplitudes. Lorsqu'ils sont surface-consistante (effets liés directement à la surface du sol), ces effets peuvent bénéficier de la redondance d'information apportée par la couverture multiple.

La correction appliquée dans ce cas est appelée correction d'amplitude en mode surface-consistante. Son principe se base sur le regroupement des traces suivant différentes collections (offsets, ponts de tir, géophones, CDP) et en analysant leur amplitude moyenne. Le

principe du programme appliqué pour faire cette correction est identique à celui des programmes des corrections statiques résiduelles

Le programme permet de calculer des corrections de gain pour chaque point de tir et chaque géophone afin de corriger les conditions de couplage à l'émission et à la réception. Des coefficients correctifs en fonction de l'offset sont aussi calculés (figure : A.I.1.2).

Les différentes étapes suivies pour faire cette correction sont :

- Calcul de l'amplitude moyenne de chaque offset (courbe moyenne).
- Correction des amplitudes moyennes à un terme constant pour toutes les offsets.
- Enregistrement des corrections (différence entre l'amplitude moyenne de chaque offset et le terme constant utilisé pour toutes les offsets).
- Après le calcul de la correction en terme d'offsets en fait un lissage de toutes les corrections offset dans un CDP.
- Calcul d'une correction en terme de point de tir après avoir appliqué les correction en modes offset et CDP.
- Calcul d'une correction en terme géophone après avoir appliqué les correction en modes offset et CDP.

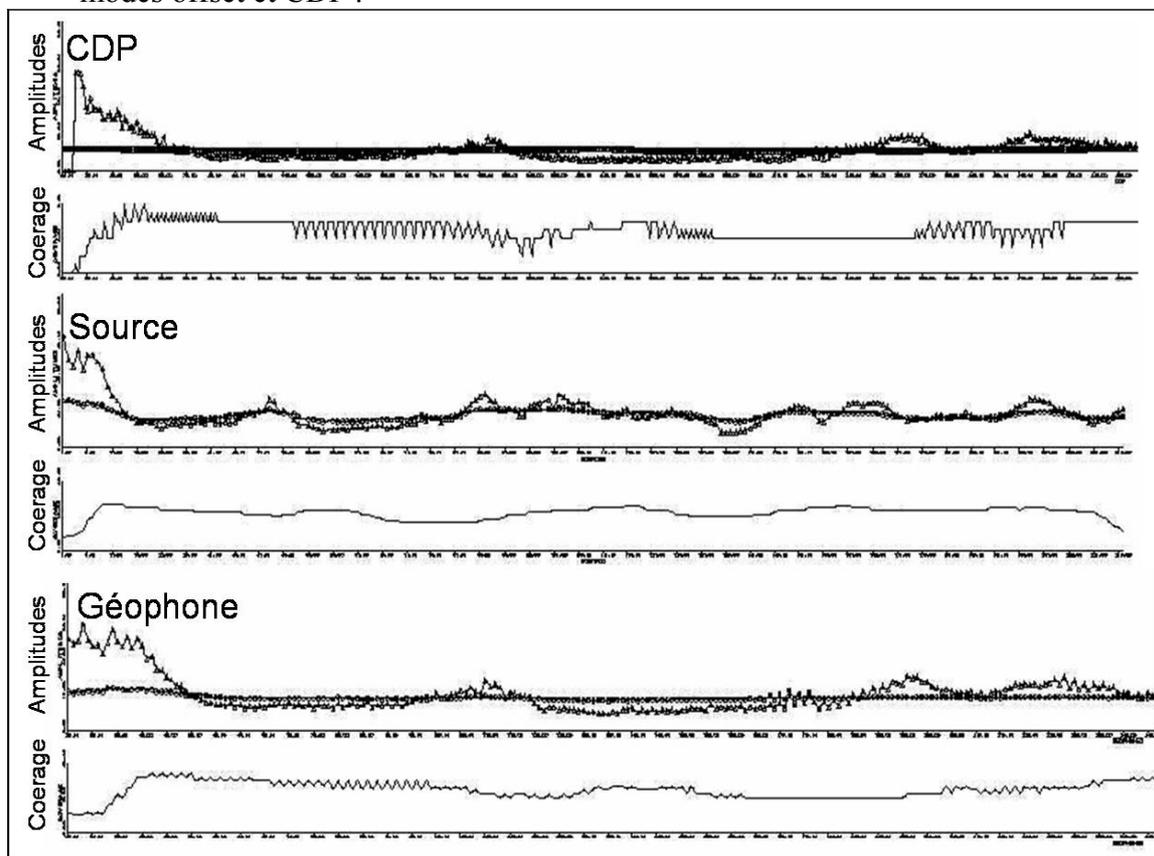


figure : A.I.1.2 Correction de l'effet de couplage

A.I.1.3. Correction de l'effet de l'absorption (atténuation non élastique)

Lors de la propagation de l'onde sismique, une partie de l'énergie du front d'onde est dissipée en chaleur. Cette dissipation est très dépendante des paramètres pétro-physiques des formations traversées, ainsi que des fluides qu'elles contiennent. Par analogie avec les grandeurs électrique, on peut affecter à chaque couche (formation) un « facteur de qualité » Q caractéristique de ses capacités d'absorption. Q est généralement compris entre 25 et 300 (les milieux très absorbants ayant les plus petites valeurs de Q ; Grès:20-130, Calcaire :40-700, Granite :200-300).

Pour une distance donnée, les hautes fréquences parcourent un nombre de longueur d'onde plus importante que les basses fréquences (Figure :A.I.1.3.a); le phénomène d'absorption atténue donc plus l'amplitude des hautes que des basses fréquences (figure : A.I.1.3.b) (O'DOHERTY et ANSTY,1971)

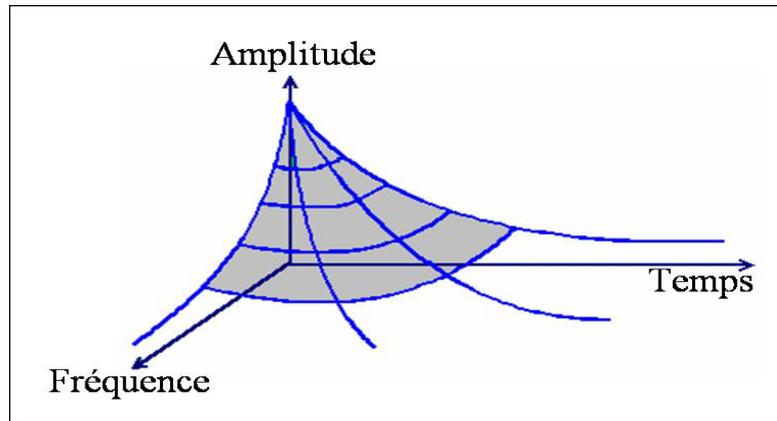


Figure : A.I.1.3.a Effet de l'absorption

L'expérience a montré que la perte d'amplitude est une fonction du temps parcouru et de la fréquence de l'onde sismique. On constate approximativement que A est proportionnel à $e^{-\alpha t}$.

Avec : A : l'amplitude.

α : coefficient d'absorption (appelé aussi coefficient d'atténuation)

t : le temps parcouru.

Le coefficient d'absorption est donné par : $\alpha = \frac{\pi f}{QV}$

Et la décroissance de l'amplitude peut s'exprimer sous la forme suivante :

$$A(t) = A(0) \cdot e^{-\frac{\pi f t}{Q}} \quad (A.I.2)$$

où Q : facteur de qualité.

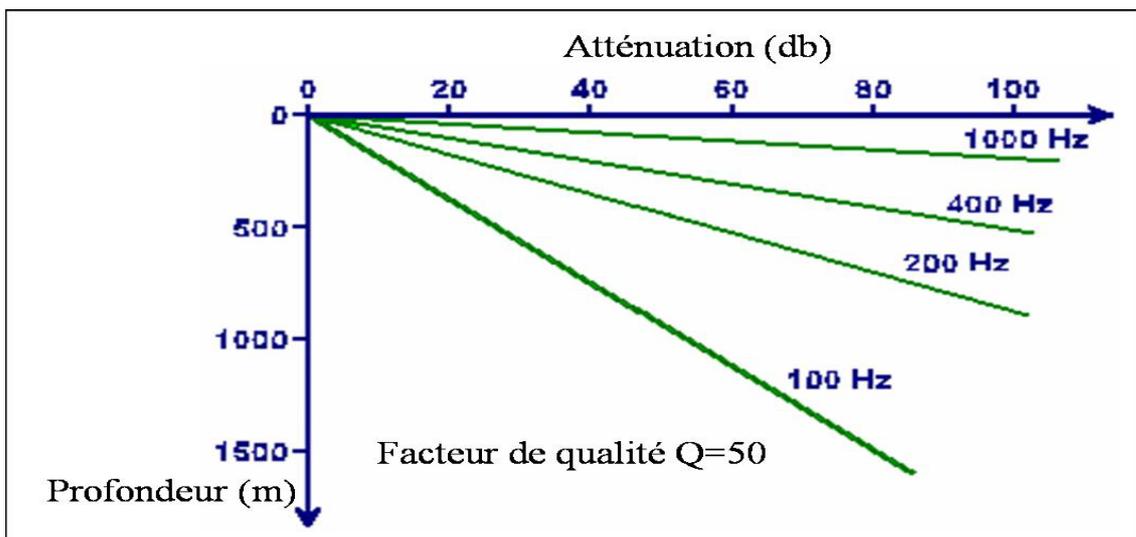


Figure : A.I.1.3.b . Effet de l'absorption sur fréquences

La figure (Figure : A.I.1.3.c.) présente l'effet de l'absorption sur une impulsion unitaire pour un coefficient d'atténuation de 0,2 db/cycle (O'DOHERTY et ANSTY,1971)

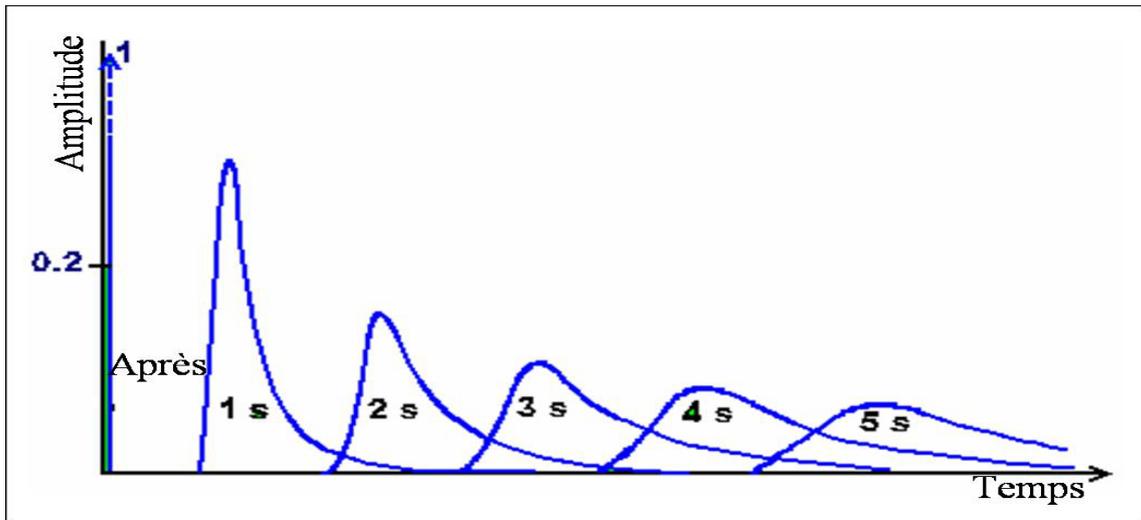


Figure : A.I.1.3.c. Effet de l'absorption sur l'amplitude

A.I.1.4. Correction de l'effet de transmission

Dans cette phase on parle de l'influence du coefficient de transmission Aller-Retour sur l'amplitude. L'amplitude réfléchié à l'interface est donc égale au coefficient de réflexion de cette interface multiplié par les coefficient de transmission Aller-Retour de toutes les couches qui sont situées au-dessus de cette interface(Figure : A.I.1.4).

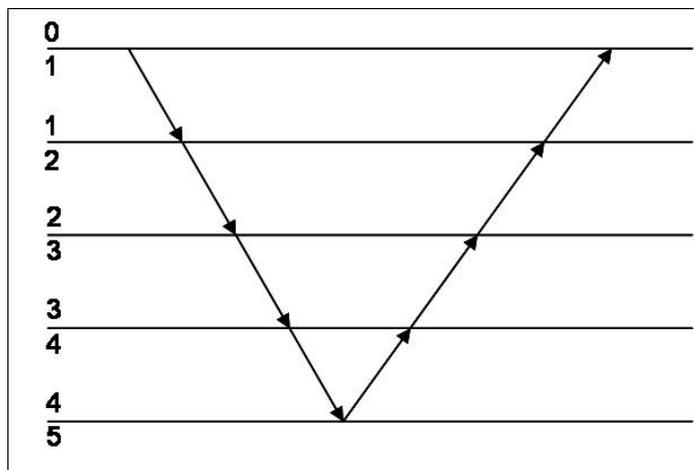


Figure : A.I.1.4. Modele géologique

Considérons par exemple quatre couches sédimentaires dont impédances acoustiques Z_1, Z_2, Z_3, Z_4 et dont les coefficients de réflexions sont respectivement $R_{1,2}, R_{2,3}, R_{3,4}, R_{4,5}$.

Le coefficient de transmission Aller-Retour (désigné par T) en incidence normale pour la première couche est :

$$T = 1 - R_{1,2}^2 \quad (A.I.1.4.1)$$

L'amplitude du signal réfléchié à la 2^{ème} interface sera :

$$A_2 = R_{2,3} (1 - R_{1,2}^2) \quad (A.I.1.4.2)$$

L'amplitude du signal réfléchi à la 3^{ième} interface est :

$$A_3 = R_{3,4} \left(1 - R_{1,2}^2\right) \left(1 - R_{2,3}^2\right) \quad (\text{A.I.1.4.3})$$

.....

L'amplitude du signal réfléchi au n^{ième} interface sera donc égale à :

$$A_n = R_{n,n+1} \left(1 - R_{1,2}^2\right) \left(1 - R_{2,3}^2\right) \dots \left(1 - R_{n-1,n}^2\right) \quad (\text{A.I.1.4.4})$$

au fur et à mesure que le front d'onde traverse les différents réflecteurs, l'énergie descendante disponible diminue en fonction des coefficients de réflexion des couches supérieures. Si l'on considère une incidence nulle, ces pertes de transmissions peuvent s'écrire pour un trajet aller-retour :

$$T_0 = \left(1 - R_{1,2}^2\right) \left(1 - R_{2,3}^2\right) \dots \left(1 - R_{n-1,n}^2\right) \quad (\text{A.I.1.4.5})$$

L'expérience a montré que lorsque l'impédance acoustique croît progressivement la transmission est meilleure alors qu'un brusque changement de l'impédance acoustique engendre une transmission moins bonne.

Pour composer ces pertes de transmission, on procède généralement par une analyse statistique des traces sismiques en fonction du temps et de l'offset afin de définir un coefficient de gain (β). En suite un programme permet l'application de gain de la forme $e^{\beta t}$. Le programme peut appliquer plusieurs coefficients β le long d'une même trace, interpoler ces coefficients en fonction de l'offset ou les déplacer en fonction de la profondeur d'eau.

A.I.2 Programmes d'Egalisation pour les représentations graphiques

A.I.21. Egalisation régionale

Avec cette option, le programme ne calcule qu'un seul coefficient pour toutes les traces entrées. Les amplitudes relatives sont donc conservées, à la fois, en temps et en espace. C'est l'option recommandée pour les traitements en amplitudes préservées (PAM).

Planche : V.5.2.c . Section sismique avec Filtrages et Migration

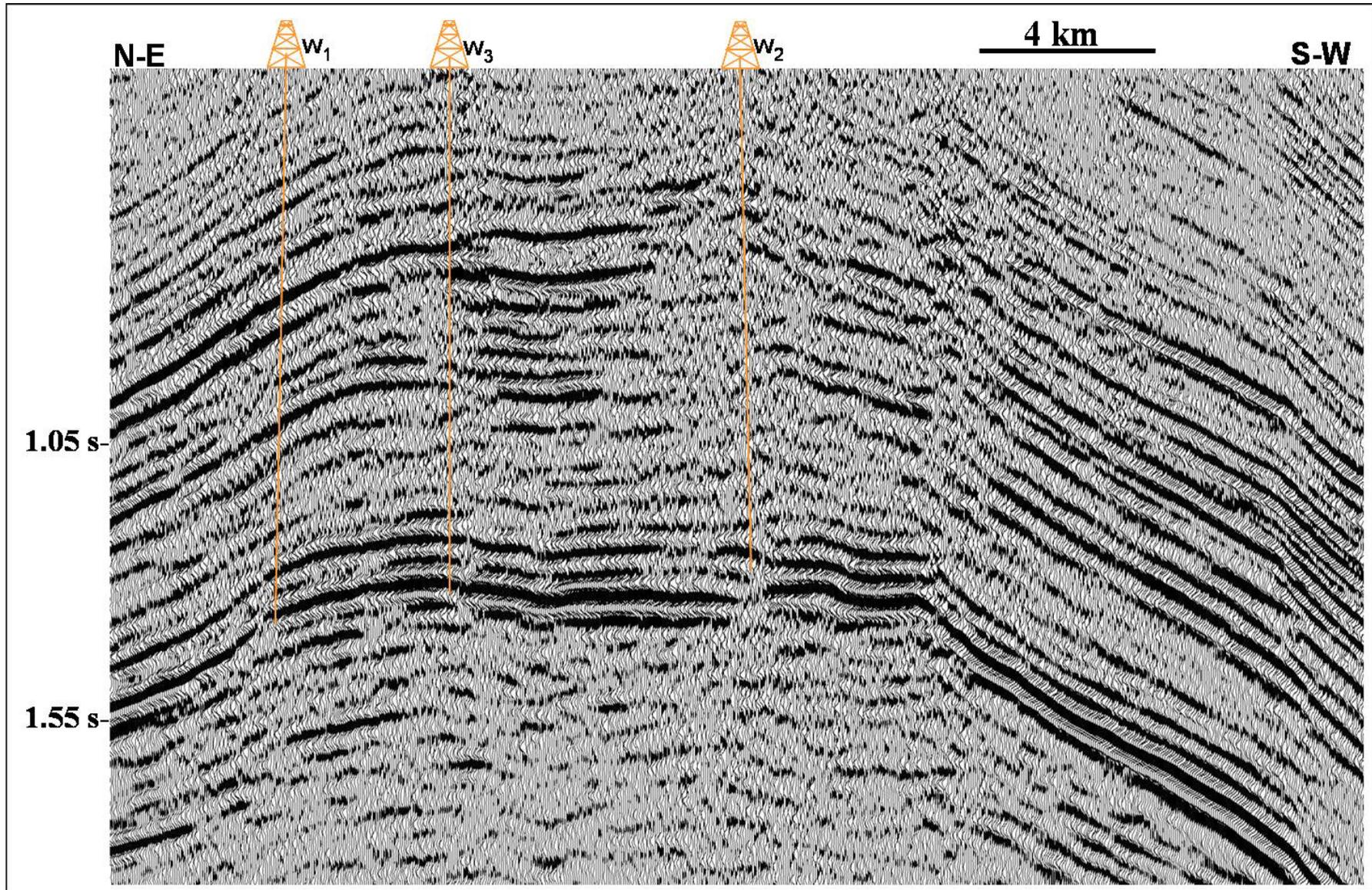


Planche : V.7.3.b Classification Lithologique par le PMC, Résultat après l'itération N°100

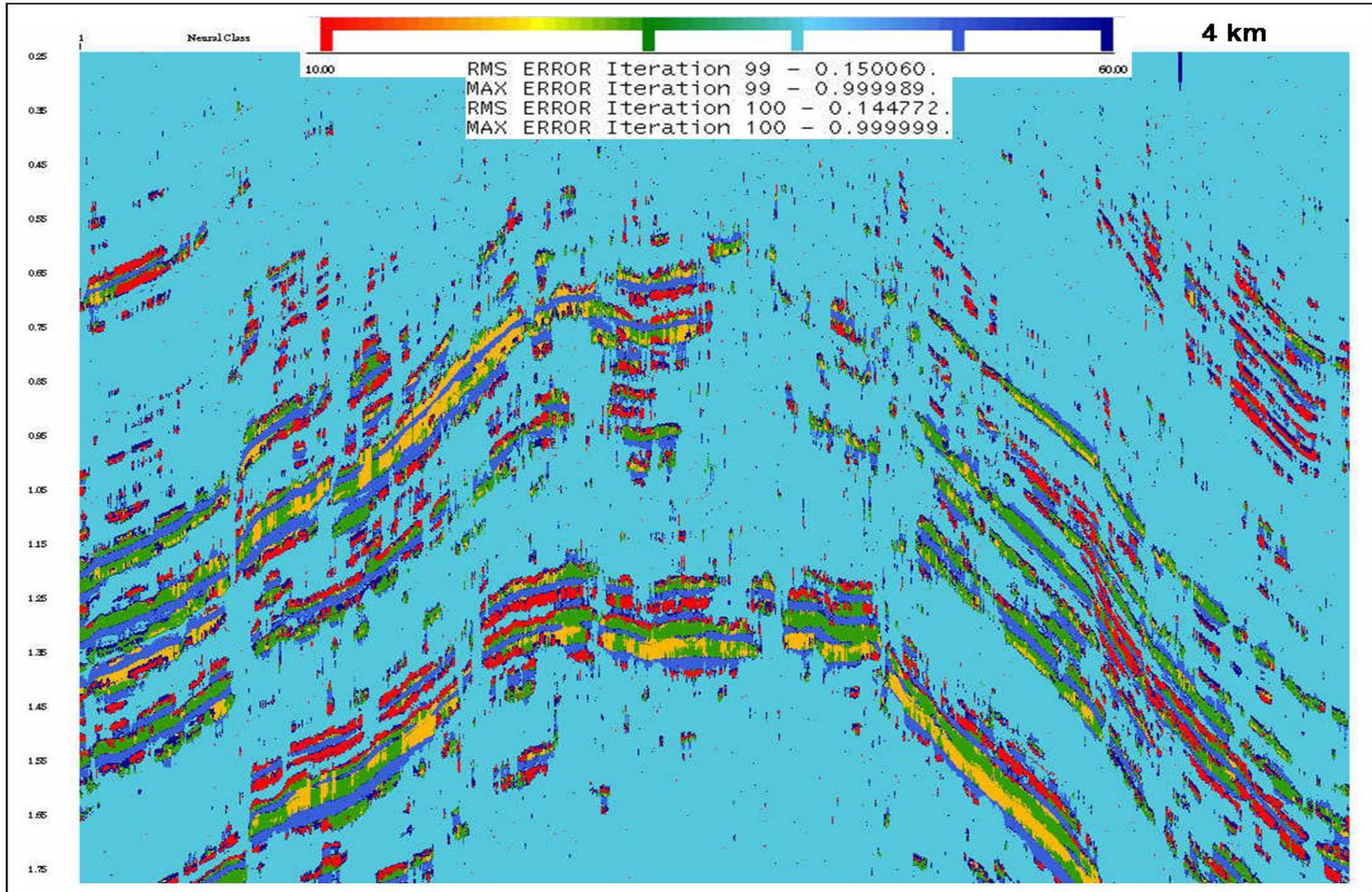
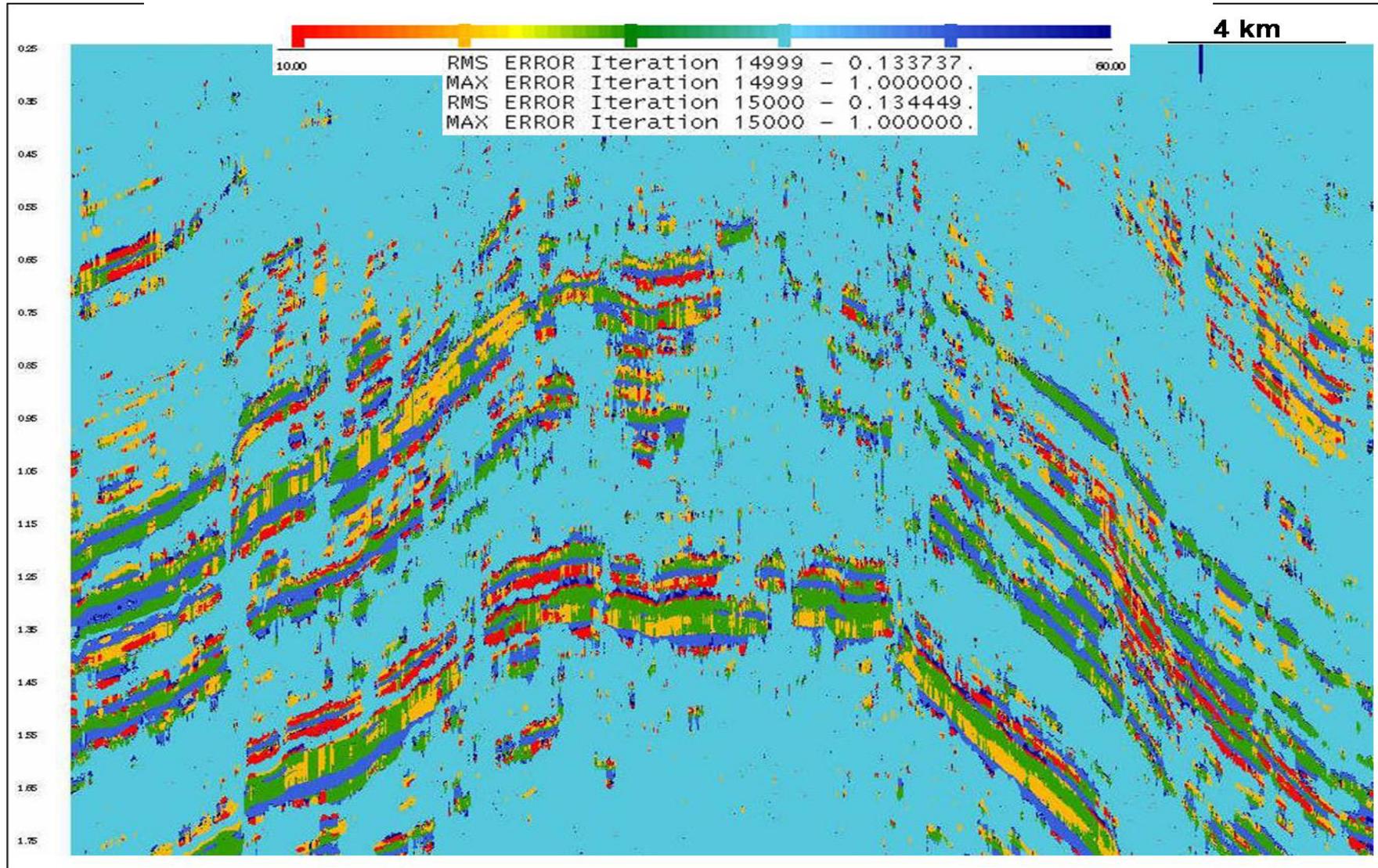
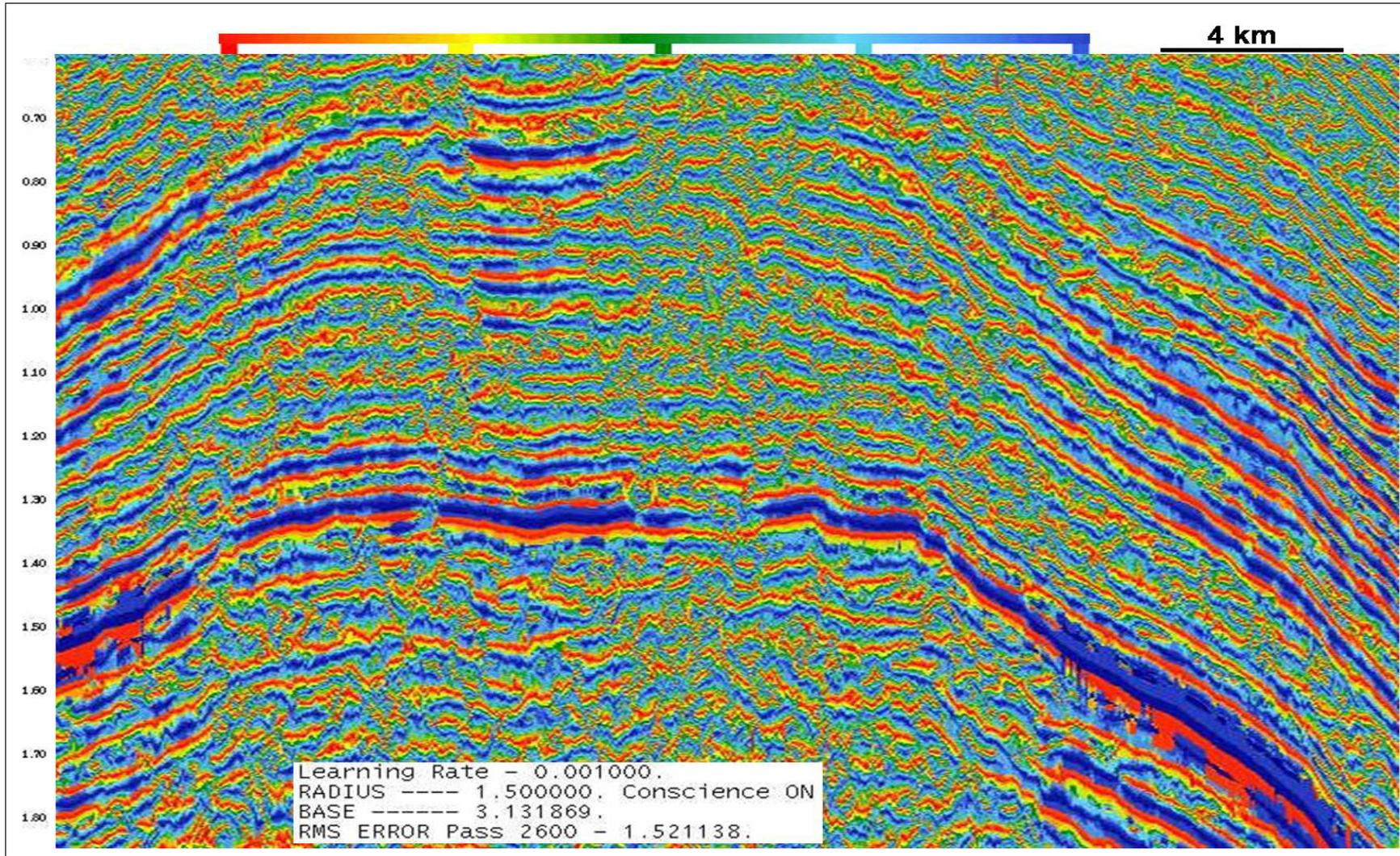


Planche : V.7.3.c Classification Lithologique par le PMC, Résultat après l'Itération N°15000



0.55 s-

Planche : V.7.3.d Classification Lithologique par K-SOM, Résultat après l'itération N°2600



document téléchargé depuis www.pnst.cerist.dz CERIST

BIBLIOGRAPHIE

- [1] Alam A., P. Caragounis, S. Matsumoto, & C. Hurst, 1995,
``Reservoir Classification with Seismic Attributes``: 57th Mtg. Eur. Assoc. Expl
Geophys., Extended Abstracts, 95, Session: A037.
- [2] Anstey N. A., & P. Newman, 1966,
``The sectional auto-correlogram and the sectional retro-correlogram``
Geophys. Prosp., Eur. Assn. Geosci. Eng., 14, 389-42.
- [3] Aouine S., 1991,
``Etude structurale par sismique réflexion...``
Document Interne, Sonatrach, Exploration.
- [4] Badari K., M. Jdeddi, 1994,
``Eléments de sismologie``
OPU-Alger.
- [5] Bahorich M. S., & S. L. Farmer, 1995,
``3-D seismic discontinuity for faults and stratigraphic features, The coherence cube``
Expl. Geophys., 95, 93-96.
- [6] Balch A. H., 1971,
``Color sonagrams - A new dimension in seismic data interpretation``
Geophysics, Soc. Of Expl. Geophys., 36, 1074-1098.
- [7] Barnes A. E., 1994,
``Theory of two-dimensional complex seismic trace analysis``
Soc. Expl. Geophys., p1580-1583.
- [8] Barnes A.E., 1992,
``Instantaneous spectral bandwidth and dominant frequency
with applications to seismic reflection data`` Geophysics.
- [9] Barnes A.E., 1991,
``Instantaneous frequency and amplitude at the envelope peak of a constant-phase
wavelet``
Geophysics, 56, 1058-1060.
- [10] Braunschweig B., R. Day, 1995,
``Artificial intelligence in the petroleum industry – Symbolic and computational
applications``, Tome1. Editions technip – Paris.
- [11] Bodine J.H., 1984,
``Waveform analysis with seismic attributes``
S.E.G, Atlanta, Georgia.
- [12] Braunschweig B., B. Bremdal, 1996,
``Artificial intelligence in the petroleum industry – Symbolic and computational
applications``, Tome2. Editions technip, Paris.

- [13] Carrillat A. et al, 2002,
``Seismic stratigraphic mapping of carbonate mounds using 3D texture attributes``
EAGE 64th Conference&Exhibition-Florence,Italy.
- [14] Chen Q., & S. Sidney, 1997,
``Seismic attribute technology for reservoir forecasting and monitoring``
The Leading Edge, 16, no. 05, 445,447-448, 450,453-456.
- [15] Clearbout J.F., 1976,
``Fundamentals of geophysical data processing With applications to petroleum prospecting``
McGraw-Hill Book Company.
- [16] Cohen L., & C. Lee, 1990,
``Instantaneous bandwidth for signals and spectrogram``
Proc. IEEE ICASSP-90, 2451-2454.
- [17] Cooke D., 1999,
``Techniques for improving seismic attribute versus well-log property cross plots with synthetic and real data examples from Prudhoe Bay Field``
Society Of Exploration Geophysicists,p 69-72.
- [18] Dreyfus G., et al, 2004,
``Réseaux de neurones, Méthodologie et applications``
Editions Eyrolles.
- [19] Freeman J., D. Skapura, 1991,
``Neural networks – Algorithms, Applications and Programming Techniques``
Addison-Wesley Publishing Cmpany.
- [20] Hebb D., 1949,
``The Organization of Behaviour``,
Wiley, New York.
- [21] Hengchang Dai, et al, 1999,
``Inversion of fracture parameters by using the artificial neural network``
IC-AI'99 International Conference.
- [22] Herault J., C. Jutten, 1994.
``Réseaux neuronaux et traitement du signal``
Editions Hermés, Paris,
- [23] Hilterman F., 2001,
``Seismic amplitude interpretation``
SEG & EAGE.
- [24] Hoon Sung and D.S.Lee, 2001,
``Neuro-Fuzzy recognition system for detecting wave patterns using wavelet coefficients``, IEICE TRANS.INF &SYST, vol.84-D, no.8.

- [25] Hopfield J.J., 1982,
``Neural Networks and Physical Systems with Emergent Collective Computational Abilities``
Proceedings of the National Academy of Sciences, USA, pp. 2,554.
- [26] Johnston D. H., 1993,
``Seismic attribute calibration using neural networks``
Soc. Expl. Geophys., Expanded Abstracts, 93, p250-253.
- [27] Koehler F., & M. T. Taner, 1996,
``The use of the conjugate-gradient algorithm in the computation of predictive deconvolution operators``
Society Of Exploration Geophysicists, 358-365.
- [28] Kohonen T., 1990,
``The Self-Organizing Map``
Proceedings of IEEE, n° 78, p. 1464-1480..
- [29] Kohonen T., 1982,
``Self-Organized Formation of Topologically Correct Feature Maps``, Biological Cybernetics, n° 43, p. 53-69.
- [30] Kohonen T., 1982,
``Clustering, Taxonomy, and Topological Maps of Patterns``
Proceedings of the 6th International Conference on Patterns Recognition, Munich, Germany, p. 114-128.
- [31] Kileiko E., 1991,
``Interprétation sismique de la structure Garet EL Guefoul``
Document Interne, Sonatrach, Exploration, Mai.
- [32] Kouadri A., 2002,
``Modelisation et commande d'un système Boiler-Turbine par les réseaux de neurones et les algorithmes génétiques``
Mémoire de magister,
- [33] Lefeuvre F. E., K. H. Wrolstad, K. S. Zou, L. J. Smith, J-P. Maret, & Nyein, 1995,
``Sand-shale ratio and sandy reservoir properties estimation from seismic attributes, An integrated study``
Expl. Geophys., Expanded Abstracts, 95, p108-110.
- [34] Lewis C. , 1995.
``Seismic attributes for reservoir monitoring ; A feasibility study using forward modeling``
Soc. Expl. Geophys., 95, 309-312.
- [35] Li X., & T. J. Ulrych, 1996,
``Multi-scale attribute analysis and trace decomposition``
Society Of Exploration Geophysicists, 1634-1637.

- [36] Lynn H. B., D.Campagna, K. M. Simon, & W. E. Beckham, 1999,
 ``Relationship of P-wave seismic attributes, azimuthal anisotropy, and commercial gas
 pay in 3-D P-wave multiazimuth data `` Rulison Field, Piceance Basin, Colorado:
 Geophysics, 64, no. 4, 1312-1346. ,.
- [37] Matteucci G., 1996,
 ``Seismic attribute analysis and calibration, A general procedure and a case study``
 Annual Meeting Abstracts, Society Of Exploration
 Geophysicists, 373-376.
- [38] Mazzotti A., 1991,
 ``Amplitude, phase and frequency versus offset applications``
 Geophys. Prosp., 39, 863-886.
- [39] McCulloch J. & W. Pitts, 1439,
 ``A logical Calculus of the Ideas Immanent in Nervous Activity``
 Bull. of Math. Biophysics 5, 115.
- [40] Minsky M. & S. Papert, 1969,
 ``The Perceptrons``,
 The MIT Press, Cambridge.
- [41] Morlet J., G.Arens, E. Fourgeau, & D. Giard, 1982,
 `` Wave propagation and Sampling theory- Part II : Sampling theory and complex
 waves``
 Geophysics, 47, 222-236.
- [42] Neidell N. S., & Taner, M. T., 1971,
 ``Semblance and other coherency measures for multichannel data``
 Geophysics, 36, no. 03, p482-497.
- [43] O'Doherty R. F., & M. T. Taner, 1992,
 ``A method of computing instantaneous frequency and dip``
 European Association Of Geophysical Exploration, 180-181.
- [44] Pathak R., S. Badola, M. Chandra, & S. Singh, 1999,
 `` Role of dip and azimuth attributes in delineation of subtle faults``
 Society Of Exploration Geophysicists, 965-968.
- [45] Renders J-M, 1995,
 `` Algorithmes génétiques et réseaux de neurones``
 Editions Hermés, Paris.
- [46] Rihaczek A.W., 1968,
 ``Signal energy distribution in time and frequency``
 IEEE,Info. Theory, IT-14, 369.
- [47] Rijks E. J. H., & J. C. E. M. Jauffred, 1991,
 ``Attribute extraction, An important application in any detailed 3-D interpretation
 study`` , The Leading Edge, 10, no. 9, 11-19.

- [48] Robertson J.D., & D. A. Fisher, 1988,
``Seismic interpretation - Complex seismic trace attributes``
The Leading Edge, 7, no. 6, 22-26.
- [49] Robertson J.D., & H.H. Nogami, 1984,
``Complex seismic trace analysis of thin beds``
Geophysics, 49, 344-352.
- [50] Rumelhart D.E. et al, 1986,
``Learning representations by back-propagating errors``,
Nature, vol. 323.
- [51] Schmitt D. R., 1999,
``Seismic attributes for monitoring of a shallow heated heavy oil reservoir : A case study``.
Geophysics, 64, no. 2, 368-377.
- [52] Srivastava A. K., M. Chandra, D. J. Rao, D. V. Singh, & S. Singh, 1999,
``Role of seismic attributes in finding new reserves``: Cambay Basin, India``
Society Of Exploration Geophysicists,p 919-922.
- [53] Steeghs P., 1999,
``Fast computation of the sliding-window radon transform applied to 3-D seismic attribute extraction``
Society Of Exploration Geophysicists, 1146-1149.
- [54] Taner T., 2000,
``RSIATTRIB3D, Seismic trace attributes computation & Display program``
Rock Solid Images.
- [55] Taner T., 2000,
``Kohonen's Self Organizing Networks With 'Conscience'``
Rock Solid Images.
- [56] Taner M. T., & L. Lu, 1996,
``Wavelet estimation by exponential decay``, in Robinson, E. A., Ed., Seismic source signature estimation and measurement:
Society Of Exploration Geophysicists, p204-233.
- [57] Taner M. T., J. S.Schuelke, R. O'Doherty, & E. Baysal, 1994,
``Seismic attributes revisited``
Society Of Exploration Geophysicists, 1104-1106.
- [58] Taner T., 1991,
``Neural Networks and computation of Neural Network Weights and Biases by the Generalized Delta Rule and Back-propagation Errors``
Rock Solid Images,.

[59] Taner M. T., F. Koehler, & R. E. Sheriff, 1978,
``Complex seismic trace analysis``
Geophysics,44, no. 06,p 1041-1063.

[60] Taner M. T. & R. E. Sheriff, 1977,
``Application of Amplitude, Frequency, and other attributes to Stratigraphic and
hydrocarbon Determination``
Geophysics.

[61] Walls J., N. Derzhi, D. Dumas, T. Guidish, , M. T. Taner, & G. Taylor, 1999,
``North Sea reservoir characterization using rock physics, seismic attributes, and
neural networks ``A case history
Society Of Exploration Geophysicists, 1572-1575.

[62] Yilmaz O., & M. T. Taner, 1997,
``Semi-elastic plane-wave depth migration and velocity analysis`` ,Society Of
Exploration Geophysicists, 1579-1582.

[63] Youbi L., 1991,
``Rapport de fin de sondage Garet El Guefoul 3``
Document Interne, Sonatrach, Exploration.

[64] Youbi L., 1991,
``Rapport de fin de sondage Garet El Guefoul 2``
Document Interne, Sonatrach, Exploration,Aout,.