



UNIVERSITE M'HAMED BOUGARA DE BOUMERDES

Faculté des Hydrocarbures et de la Chimie

Laboratoire de Physique de la Terre

THÈSE DE DOCTORAT

Présentée par

DJARFOUR Noureddine

En vue de l'obtention du diplôme de Docteur en géophysique

Option : Sismique, sismologie et méthode de potentiels

Thème

***Inversion des données sismiques par les réseaux de neurones
artificiels***

***Application du Réseau à Fonction de Base Radiale et du Réseau
Récurrent d'Elman.***

Devant le Jury composé de :

M. DJEDDI	Professeur	Université de Boumerdès	Président
K.BADDARI	Professeur	Université de Boumerdès	Rapporteur
T.AIFA	M.C.H.D.R	Université de Rennes1-France	Examineur
Mme Z. BENAÏSSA	M.C	Université sciences et technologie Houari Boumediène	Examineur
A.BOUGUERN	M.C	Université de Boumerdès	Examineur
A.LACHOURI	M.C	Université de Skikda	Examineur

Table des Matières



Table des Matières

Résumé en arabe	I
Résumé en anglais	II
Résumé en français	III
Nomenclatures	IV
Principales abréviations utilisées	V
Liste des figures	V

Introduction générale

1	Problématique	01
2	Objectifs de la recherche	02
3	Présentation de la thèse	03

Chapitre I Etat de l'art des problèmes inverses

I-1	Introduction	04
I-2	Aperçu historique	05
I-3	Classification des approches du problème inverse	07
I-4	Formulation mathématique du problème inverse	08
I-4-1	Paramétrisation du problème	08
I-4-2	Problème direct et inverse	09
I-4-3	Difficulté et résolution du problème inverse	11
I-4-3-1	Problème bien ou mal posé	11
I-4-3-2	Conditionnement du problème	12
I-4-3-3	Régularisation et solution du problème inverse	14
I-4-3-3-a	Information a priori	14
I-4-4-3-b	Régularisation par troncature	15
I-4-4-3-b	Régularisation par optimisation	15
I-4-4-3-d	Critère d'arrêt du processus	17
I-5	Etude de cas et exemple d'application	17
I-5-1	La tomographie sismique	17
I-5-2	Tomographie linéaire ou non linéaire	18
I-5-3	Les méthodes algébriques de reconstruction tomographique	20
I-5-3-a	La méthode technique de reconstruction algébrique (ART)	20
I-5-3-b	La méthode de reconstruction itérative simultanée (SIRT)	20
I-5-3-c	La méthode du gradient conjugué (CG)	21
I-6	Deconvolution stratigraphique	21
I-7	Conclusion	23

Chapitre II Les réseaux de neurones

II-1	Introduction	25
II-2	Historique de la technique	25
II-3	Caractéristiques des techniques d'intelligence artificielles	26
II-4	Le neurone artificiel	27
II-5	Comportement d'un neurone artificiel	28
II-6	Les réseaux de neurones	28
II-7	Structure d'interconnexion	28
II-8	Réseau récurrent d'Elman	29
II-9	Apprentissage	30
II-9-a	Principe de rétro propagation	32
II-9-b	Adaptation des poids	33
II-10	Réseaux de neurones à fonction de base radiale	34
II-11	Conclusion	37

Chapitre III Application et résultats de la tomographie

III-1	Application et résultats de la tomographie	38
III-2-a	Modèle 1	40
III-2-b	Modèle 2	44
III-2-c	Modèle 3	49
III-3	Robustesse au bruit sur les données de temps de parcours	53
III-4	Conclusion	57

Chapitre IV Application et résultats de l'inversion stratigraphique

IV-1	Application et résultats de l'inversion stratigraphique	58
IV-2-a	Modèle 1	58
IV-2-b	Modèle 2	62
IV-2-c	Modèle 3	65
IV-3	Robustesse au bruit sur les données sismiques synthétiques	69
IV-4	Application des RNA sur un exemple d'un prospect Sud algérien	74
IV-5	Conclusion	81
	<i>Conclusions, nouveautés scientifiques et perspectives</i>	82
	<i>Références bibliographiques</i>	89
	<i>Annexe A</i>	95
	<i>Annexe B</i>	103

RESUME



ملخص

لا تزال المسألة العكسية تطرح تحديا كبيرا في الجيوفيزياء و ذلك راجع أساسا لأن تفسير المعطيات الجيوفيزيائية المسجلة عند سطح الأرض يهدف أساسا إلى تقدير المعايير الصوتية، التي يمكن اعتمادها أساسا لتفسير و ذلك بغية البحث المباشر عن الهيدروكربونات.

تنصب هذه الأطروحة أساسا على معالجة المسائل العكسية في الجيوفيزياء و على وجه التحديد لا و وجودية الحل و لا وحدانيته و لا استقراره.

خلال هذا العمل قمنا بالتطرق إلى الإطار النظري مع التحليل الرياضي للمسألة المباشرة و العكسية. ثم مختلف الطرق المعتمدة أساسا للمعالجة كل جانب من هذه المسألة، ضمن هذا الإطار تم تحديد مسألتين إحداها خطية و الأخرى لا خطية كنموذج للتطبيق و المقارنة و يتعلق الأمر بالطورمغرافيا و العكسية الستراتيغرافية كمساهمة أساسية لهذه الأطروحة قمنا باقتراح مجموعة من الطرق تعتمد على تقنيات الذكاء الاصطناعي وفقا لاتجاهين مختلفين الشبكات العصبية الاصطناعي.

تندرج هذه الطرق ضمن الإطار العام لأنظمة المعالجة المقتبسة من البيولوجيا والهندسة الوراثية ذات الخصائص المميزة: التعاون ، التأقلم التعلم بأمثلة الخ

النتائج المحصل عليها ضمن تطبيقات الطوموغرافيا اعطت تفوقا جديا معتبرا للشبكات العصبية الاصطناعية من نوع الدوال ذات الأساس القطري متميزة بقدرة فائقة على التعلم و التعميم مع دقة معتبرة في تحديد الصور التحت سطحية لنماذج حيولوجية معقدة مقترحة ، متبوعة بالطرق الكلاسيكية المعتمدة غيران الشبكة العصبية الاصطناعية التراجعية اظهرت صعوبة في التعلم و التعميم و ذلك راجع إلى صعوبة التقارب المؤدي إلى تصحيح معاملات الشبكة. نفس النتائج تم الحصول عليها عبر مختلف الاختبارات عن وجود تشويش على المعطيات.

أما فيما يخص العكسية الإستراتيغرافية تم تطبيق علي عدة نماذج حيث عملية التعلم و التعميم تمت بنجاح نسبة للشبكتين ، أيضا اختبار المعالجة عند وجود تشويش على المعطيات بنسبة 5% و 10% أعطى تقدما ملحوظا للشبكات الاصطناعية التراجعية على الدوال ذات الأساس القطري و التي تراجعت تراجع ملحوظ، و يرجع تفوق شبكات المن إلى وجود البعد الزمني لهاته الشبكات أثناء المعالجة.

بعد اختبار الطرق المقترحة على معطيات نموذجية قمنا بتطبيقها على معطيات حقيقية مستخرجة من الاستكشاف الجيوفيزياء منجز بإحدى مناطق الجنوب الجزائري. نصف المعطيات استخدام للتعلم و النصف الآخر للتعميم مختلف اختبارات التعلم و التعميم تمت بنجاح بالنسبة للشبكتين مع تقارب جد ملحوظ في الدقة بالنسبة للشبكتين.

ABSTRACT

The inverse problem theory, to the largest sense, has been developed initially by researchers working with the geophysical data; the reason is that geophysicist tries to extract information about the Earth structure from data of observation collected at the surface of the earth.

Otherwise, several techniques have been elaborated for the resolution of the inverse problem, most of them tries to treat its complexity and its non linear characterization, that pose serious problems of stability and convergence. The non-existence, non-uniqueness and instability are the most important properties that we deal with in the solution of inverse problem, due to noisy and/or insufficient measurement data, an accurate solution may not exist.

However, we considered in this work, the seismic tomography and stratigraphic inversion as examples of application.

In this work the neural networks of Elman type and Radial Basis Function (RBF) were addressed to inverse seismic data. The results show that networks could establish relationship between acoustic parameters and seismic data, based on their ability of approximation and adaptation. The training of the network can be considered as a means to synthesize automatically a function (control mapping), generally a non-linear one, which plays a role of filter.

In our case the learning ANN algorithm used a steepest descent technique which is based on straight downhill in the weight space till the first valley is reached. This makes the choice of initial starting point in the multi-dimensional weight space critical. To overcome this limitation, the training process was repeated a number of times with different starting weights. The excellent results indicate that training of the proposed networks was done successfully, and the Elman and RBF network have the ability to obtain strict convergence for a complicated set of data in case of inversion of seismic data. Moreover, the method is able to reproduce features that were not included in its training set. In fact, based on the results of the ability of generalization, the application of the ANN technique in a same seismic exploration area produced reasonable reconstruction of acoustic impedance properties from seismic data.

Several tests were performed with synthetic data even in the presence of random noise. The results show that the used ANN structure (i.e. Elman and RBF type) was able to recognize and reconstruct subtitle details of geological model. The same results were also obtained on real seismic data.

We think that Elman feedback and RBF neural networks, with their ability to discover input-output relationships will increasingly be used in engineering applications, especially in petroleum engineering usually associated with intrinsic complexity. The results obtained demonstrate that the neural network can detect several shapes of geologic layers in complex data.

RESUME

L'appréhension de la structure interne de la Terre à partir de données collectées en surface demeure sans doute le souci constant de la plupart des géophysiciens. C'est dans ce contexte typique que s'est initialement développée la théorie du problème inverse, visant à interpréter, voire caractériser au mieux les problèmes géophysiques ainsi posés.

Au cours de ces dernières années, plusieurs techniques ont été élaborées pour la résolution du problème inverse. La plupart de ces techniques cherchent à traiter son caractère mal posé qui pose en lui-même de sérieux problèmes de stabilité et de convergence.

Dans cette thèse, notre contribution s'inscrit justement au cœur de cette problématique. En effet, nous ne prétendons pas apporter des solutions clefs à tous les problèmes posés, mais nous avons examiné l'application des réseaux de neurones artificiels (RNA) à la résolution du problème inverse. Sont considérées le long de ce travail les deux principales classes suivantes : la tomographie sismique et l'inversion stratigraphique.

Soulignons tout d'abord que les RNA font partie des systèmes adaptatifs non linéaires, qui sont constitués d'un grand nombre d'unités élémentaires (appelées aussi cellules ou agents). Ils traitent l'information d'une manière parallèle et distribuée. De telles propriétés leur confèrent bien la qualité de robustesse vis-à-vis des données incomplètes, incertaines et bruitées. Chose qui motive leur utilisation dans le cadre du problème inverse.

En vue de justifier l'applicabilité des RNA, une étude comparative a été réalisée. Pour ce faire, les réseaux de neurones à Base Radiale (FBR) et de type Elman sont utilisés comme outils d'inversion et comparés à des techniques classiques (ART, SIRT, et GC) sur la base d'exemples synthétiques, puis sur des données réelles.

Dans le cadre de la tomographie, les résultats obtenus ont montré que les RNA de type FBR sont les plus performants en terme d'amélioration de la reconstruction d'images sismiques très complexes, viennent ensuite les méthodes GC et ART avec un niveau de reconstruction assez acceptable. La méthode SIRT, quant à elle, s'est révélée moins performante. D'autre part, il est important de noter que le réseau de type Elman s'avère assez conservatif en raison des difficultés liées à la procédure d'apprentissage. Le test de reconstruction en présence de bruit montre la capacité de la structure FBR à retrouver avec une précision largement acceptable la répartition de vitesse à partir du temps de trajet, tandis que les autres méthodes ont donné des images altérées présentant des artefacts. Ceci est principalement dû à la présence de bruits et à l'instabilité de la solution obtenue.

Quant à l'inversion stratigraphique, nous avons pu obtenir un apprentissage simple et facile pour les deux structures proposées (FBR et Elman) à l'aide de l'algorithme de rétro-propagation du gradient conjugué. Notons qu'en présence de bruits, le RNA de type Elman conduit bien à des résultats intéressants en termes d'inversion, alors que la performance préalablement discutée du réseau FBR se dégrade de plus en plus au cours du processus d'inversion. Une autre constatation qu'on puisse également faire réside dans le fait que l'utilisation d'une seule couche cachée simplifie davantage la structure des réseaux et permet, par voie de conséquence, de rendre aisée l'estimation des paramètres acoustiques à partir des données sismiques.

Sur la base de l'ensemble des considérations examinées le long de cette thèse, il ne deviendra guère difficile d'affirmer que les RNA offrent bien les outils nécessaires à l'amélioration de la caractérisation de la structure interne de la Terre, ce qui se traduit par un net allègement de la tâche de l'interprétation.

NOMENCLATURES



NOMENCLATURES

- M : espace de paramètres.
- D : espace des données.
- G : opérateur de modélisation.
- d : vecteur de données.
- m : vecteur de paramètres.
- η : bruit.
- α : parameter d'échelle.
- λ : valeur singulière.
- $g(s,x)$: noyau de convolution.
- $E(m)$: distance Euclidienne.
- $J(m)$: fonction objective.
- T : vecteur temps de trajet.
- S : lenteur (inverse de la vitesse).
- P : trajet du rayon sismique.
- l_{ij} : distance traversée par le rayon i dans la cellule j .
- t_i : temps de trajet du rayon i .
- gr : gradient
- β : pas de descente.
- r : coefficient de réflexion.
- I : impédance acoustique.
- C : vitesse des ondes sismiques.
- u : potentiel de déplacement.
- W : champ adjoint.
- f : densité de force.
- h : pas de discrétisation.
- x : entrée de neurone.
- w_{ij} : force de la connexion simulant les poids synaptiques des neurones
- O_i : fonction de stimulation de neurone i .
- f_i : fonction d'activation de neurone.
- ρ_i : sortie de neurone.
- e_i : erreur de neurone i à la sortie du réseau de neurones.
- ΔW_{ij} : mécanisme d'adaptation

- y_i : sortie réelle du neurone i à la sortie du réseau de neurone.
- t_i : sortie désiré du neurone i à la sortie du réseau de neurone.
- Q_j : matrice des rayons gaussiens,
- c_j : vecteur du centre du nœud j

PRINCIPALES ABREVIATIONS UTILISEES

- IFK: intégrale de Fredholm de première espèce (Integral of Fredholm First Kind).
- SVD : décomposition en Valeur Singulière (Singular Value Decomposition).
- ART: technique de reconstruction algébrique (Algebraic Reconstruction Technique).
- SIRT : reconstruction itérative simultanée (Simultaneous Iterative Reconstruction Technique).
- GC : Gradient Conjugué.
- RNA : Réseau de Neurones Artificiel.
- RNA-FBR : Réseau de Neurones Artificiel de Fonction à Base Radiale.

LISTE DES FIGURES

Figure I-01 problème direct

Figure I-02 problème inverse

Figure I.03 trajet d'un rai

Figure I-04 étapes du problème direct

Figure I-05 étapes de l'inversion stratigraphique

Figure II-01 neurone formel

Figure II.02 structure de connexion du Perceptron multicouches

Figure II-03 connexion directe (a) et récurrentes (b)

Figure II.04 structure de connexion du réseau d'Elman

Figure II-05 schéma de fonctionnement de l'apprentissage

Figure II-06 réseau de neurones à fonction de base

Figure II-07 quelques fonctions radials

Figure III-01 couverture et le design de l'acquisition des données de sismique entre puits

Figure III -02 exemple de paramétrisation, tracé du rayon

Figure III -03 densité des rayons

Figure III -04 résultats de la tomographie sur le Modèle 1 : a- Modèle proposé, b- résultat de la méthode GC, c- méthode SIRT, d- la méthode ART. e- résultat de RNA de Elman, f)- résultat de FBR, g)- Courbes des résidus de différentes méthodes.

Figure III -05: résultats de la tomographie sur le Modèle 2 : a- modèle proposé, b- résultat de la GC, c- méthode SIRT, d- méthode ART. e- résultat du RNA de Elman, f- résultat du FBR, g- Courbes des résidus des différentes méthodes.

Figure III -06. résultats de la tomographie sur le Modèle 3 : a- modèle proposé, b- résultat de la GC, c- méthode SIRT, d- méthode ART. e- résultat de RNA de Elman, f)- résultat de FBR, g)- Courbes des résidus des différentes méthodes.

Figure III -07 résultats de la tomographie sur le Modèle 3 : a- résultat de la GC, b- méthode SIRT, c- méthode ART. d- résultat du RNA Elman, e- résultat du RNA FBR, f- Courbes des résidus des différentes méthodes.

Figure IV-01: résultats de l'inversion stratigraphique sur le Modèle 1 : a- section sismique (entrée du RNA), b- modèle à retrouver (sortie désirée), c- résultat du RNA-Elman cas de 90

neurones dans la couche cachée. e- résultat du RNA-FBR cas de 90 neurones dans la couche cachée. f- les courbes d'apprentissage.

Figure IV-02 résultats de l'inversion stratigraphique sur le Modèle 2 : a- section sismique (entrée du RNA), b- modèle à retrouver (sortie désirée), c- résultat du RNA-Elman cas de 90 neurones dans la couche cachée. e- Résultat du RNA-FBR cas de 90 neurones dans la couche cachée. e- courbes d'apprentissage.

Figure IV-03: résultats de l'inversion stratigraphique sur le Modèle 1 : a- section sismique (entrée du RNA), b- modèle à retrouver (sortie désirée), c- résultat du RNA-Elman cas de 90 neurones dans la couche cachée. e- résultat du RNA-FBR cas de 90 neurones dans la couche cachée.

Figure IV-04 résultats de l'inversion stratigraphique sur le Modèle 3 : a- section sismique avec 10% de bruits aléatoires (entrée du RNA), b- résultat du RNA-Elman cas de 90 neurones dans la couche cachée. e- résultat du RNA-FBR cas de 90 neurones dans la couche cachée.

Figure IV-05 résultats de l'inversion stratigraphique sur le Modèle 3 : a- section sismique avec 20% de bruits aléatoires (entrée du RNA), b- le résultat du RNA-Elman cas de 90 neurones dans la couche cachée. e- résultat du RNA-FBR cas de 90 neurones dans la couche cachée.

Figure IV-06 a) Section sismique sommée traitée en amplitude préservée (b) Impédance acoustique (section désirée)

Figure IV-07 a) Résultat du RNA Elman avec 5 neurones (b) Résultat du RNA Elman à 90 neurones

Figure IV-08 a) Résultat de la FBR à 10 neurones (b) Résultat de la FBR à 50 neurones

Figure IV-09. Courbes d'apprentissage des deux structures neuronales proposées.

LISTE DES TABLEAUX

Tableau 01 Récapitulatif de la paramétrisation, modélisation

Tableau 02 Récapitulatif des caractéristiques de reconstruction sur le modèle 01

Tableau 03 Récapitulatif des caractéristiques de reconstruction sur le modèle 02

Tableau 04 Récapitulatif des caractéristiques de reconstruction sur le modèle 03

Tableau 05 Récapitulatif des caractéristiques de reconstruction en présence de bruit sur le modèle 02

Tableau 06 Tableau récapitulatif de l'apprentissage sur le modèle 1

Tableau 07 Tableau récapitulatif de l'apprentissage sur le modèle 2

Tableau 08 Tableau récapitulatif de l'apprentissage sur le modèle 3

Tableau 09 Tableau récapitulatif de l'apprentissage sur les données réelles

INTRODUCTION GENERALE



1- PROBLEMATIQUE

Le plus grand défi du géophysicien est d'extraire à partir de données sismiques mesurées en surface, les paramètres acoustiques de la structure interne de la Terre. Ainsi un modèle direct décrivant l'interaction entre le signal sismique émis et les couches de la Terre est développé. Ce modèle s'exprime sous la forme d'un produit de convolution ou d'équations aux dérivées partielles. Connaissant les données mesurées, la résolution du problème inverse permet de remonter aux paramètres recherchés [Backus 1967, Tarantola 1984a, 1984b, 1987, Symes 1994, Carlos 2000, Moraes 2000, Adde 2005]. Ce problème, est en général mal posé, intervient dans toutes les méthodes géophysiques. En effet, l'estimation des paramètres physiques (vitesse, densité, impédance ...) à partir des données collectées à la surface de la Terre, représente l'objectif principal des études géophysiques [Duijndam 1988, Stewart 1991, Zhou 2001].

Les difficultés de résolution du problème inverse viennent du fait que :

- Les données ont une origine expérimentale, ce qui implique l'existence d'erreurs de mesure [Tarantola 1984a, 1984b, 1987].
- Les mesures sont collectées en nombre fini, même si, dans le modèle mathématique, elles sont décrites par des fonctions.
- L'algorithme d'inversion peut parfois créer une altération des données à savoir l'interpolation requise pour la discrétisation d'un modèle initialement continu par exemple [Boulianger 2004, Lascaux 2004].
- La modélisation procède d'une idéalisation de la réalité physique et repose sur des hypothèses simplificatrices, elle est donc également une source d'incertitude [Lascaux 2004].

La résolution du problème inverse passe nécessairement par les étapes suivantes : la paramétrisation, la modélisation (problème direct), et l'inversion. Une bonne estimation des propriétés acoustiques s'effectue correctement à condition de bien cerner les trois étapes précédentes [Cooke 1983, Hirahara 1993, Andrés 2002].

L'inexistence, la non-unicité et l'instabilité de la solution sont en effet des problèmes inévitables lors de la résolution du problème inverse. La prise en compte des incertitudes et des incomplétudes sur les données expérimentales, et la non-linéarité de la relation données-paramètres sont des questions fondamentales. Ces trois problèmes sont résolus conjointement en introduisant dans le processus d'inversion de l'information *a priori* [Levy 1980, Pan 1988, Kormendi 1989, Snieder 1999, Kern 2003].

Le travail effectué dans le cadre de cette thèse porte justement sur cette problématique : le traitement de l'existence, de l'unicité et de la stabilité de la solution du problème inverse en sismique.

2- OBJECTIFS DE LA RECHERCHE

A la lumière de ce qui a été dit, il semble que la recherche d'un algorithme d'inversion ne nécessitant pas d'hypothèses restrictives sur la solution, un modèle initial et capable de traiter le problème dans son contexte réel, est une alternative intéressante pour résoudre le problème d'estimation des paramètres. Cette alternative est offerte par les méthodes d'intelligence artificielle. Cette nouvelle formulation permet de s'affranchir des problèmes complexes. De plus, la technique est intéressante, car elle peut être programmée facilement sur machine.

Depuis des décennies, les réseaux de neurones artificiels, et plus particulièrement les *perceptrons multicouches*, se sont montrés très efficaces dans le domaine de la reconnaissance de formes, de l'étude de systèmes complexes. Ces systèmes sont composés d'un ensemble structuré en unités de traitement, appelées *neurones*, qui fonctionnent en parallèle et qui sont fortement interconnectés. Ces neurones réalisent chacun une fonction non-linéaire à partir de leurs entrées. Cette fonction qui est déterminée par un ensemble de paramètres dont les valeurs sont établies à la suite d'un apprentissage par l'exemple. Les non-linéarités présentes dans ces systèmes connexionnistes, ainsi que l'apprentissage discriminant qu'ils subissent, les rendent particulièrement bien adaptés aux tâches d'inversion. Cependant, leurs performances sont fortement affectées par la complexité des données à traiter. Vu les caractéristiques du problème inverse en sismique (non linéarité, dimensionnalité élevée, présence de bruits .etc.), l'apprentissage des perceptrons multicouches peut devenir délicat. Les perceptrons ne peuvent en outre être performants que si les caractéristiques qui leur sont fournies, sont pertinentes. Enfin, des réseaux de neurones de grande taille sont aussi très exigeants en puissance de calcul, ce qui restreint fortement leurs possibilités d'application.

3- PRESENTATION DE LA THESE

Dans le premier chapitre, après avoir mis en évidence les difficultés de résolution des problèmes inverses en géophysique de par leur caractère mal-posé, on procédera à la mise en équation du problème inverse et la présentation des méthodes de régularisation classiques les plus utilisées. La suite du chapitre est consacrée à la présentation des applications très particulières, d'une part l'inversion tomographique, qui représente l'un des problèmes inverses linéaires, où trois algorithmes à savoir : ART et SIRT, et le GC seront présentés. D'autre part, l'inversion stratigraphique.

Le deuxième chapitre est consacré à la présentation des réseaux de neurones artificiels. Leur principe et leur fondement mathématique sont exposés en détail. Ces méthodes sont intéressantes dans le sens où elles sont à la fois non-linéaires et plus simples à mettre en œuvre dans le cas d'apprentissage supervisé. Il s'agit alors de trouver une fonction de transfert non-linéaire, qui permette le passage des données observées aux paramètres acoustiques. Nous avons développé pour cela deux structures, l'une de type connexion directe et l'autre de type feedback. L'apprentissage est assuré par l'algorithme de rétro-propagation du gradient conjugué.

L'entraînement des réseaux de neurones, souvent considéré comme délicat, fera l'objet de plusieurs simulations aux troisième et quatrième chapitres. Dans le cadre de l'inversion des données sismiques, nous avons implémenté deux structures neuronales qui peuvent être placées dans deux catégories différentes: les architectures statiques qui ne tiennent compte que de l'organisation spatiale mais qui négligent la "composante" temps, et les architectures temporelles qui incorporent la dimension temps. Dans la catégorie des architectures statiques, nous avons implémenté le réseau à Fonctions de Base Radiale (RBF: Radial-Basis Functions network). En ce qui concerne la catégorie des architectures temporelles, nous avons expérimenté le réseau récurrent d'Elman. Les performances des diverses méthodes classiques présentées au chapitre 1 sont ensuite comparées aux structures neuronales proposées sur la base des modèles synthétiques de complexité variable.

Finalement, nous terminerons le présent manuscrit par une conclusion générale et des perspectives qui en découlent.

Etat de l'art des problèmes
inverses

CHAPITRE I



Etat de l'art des problèmes inverses

I-1 INTRODUCTION

La résolution du problème inverse permet d'estimer les caractéristiques d'un système connaissant sa réponse à une excitation donnée. L'inversion des données sismiques cherche à résoudre le problème fondamental de la prospection géophysique, qui est la détermination de la répartition spatiale des propriétés des roches à partir des observations de surface.

Le problème inverse est un sujet trop vaste pour que nous puissions en donner un exposé exhaustif dans le cadre de cette thèse. C'est aussi un sujet en évolution constante et ses applications se développent de plus en plus. Dans ce chapitre, nous essayons de présenter le cadre général du problème inverse et d'établir les principales notions que nous utiliserons par la suite. L'accent sera exclusivement mis sur le caractère mal posé de certains problèmes inverses.

Le formalisme général du problème inverse est basé sur l'ajustement de l'accord entre les mesures obtenues et les simulations du modèle direct tout en tenant compte des différentes sources de contradictions [Backus 1967, Lines 1977]. L'estimation de la valeur des paramètres que l'on veut restituer ainsi que l'évaluation de l'incertitude découlant de ces estimations est possible si le problème inverse est correctement posé.

Dans le cadre particulier de ce travail, il s'agit de déterminer une série de paramètres décrivant le plus précisément possible les propriétés acoustiques du milieu géologique à partir de mesures effectuées à la surface de la Terre. D'après certains auteurs [Tarantola 1984a, Snieder 1999, Andrés 2002, Pierrick 2003, Boulanger 2004, Adde 2005], la résolution d'un tel problème peut se décomposer en trois étapes, qui nécessitent d'être étudiées de manière coordonnée à savoir :

1- Une acquisition des données qui consiste à recueillir des données les plus informatives possibles afin de pouvoir restituer des paramètres décrivant le milieu de façon pertinente.

2- Un problème direct où l'on établit, connaissant l'objet et le mécanisme d'observation, une description mathématique des données observées (figure I-01). Ce modèle doit être assez précis pour fournir une description correcte du phénomène physique observé, et assez simple cependant pour se prêter à un traitement numérique ultérieur.

3- Un problème inverse où l'on doit restituer les propriétés caractéristiques du milieu à partir du modèle et des données précédentes. Une méthode d'inversion permettra de restituer la

distribution des propriétés acoustiques de la subsurface (figure I-02).

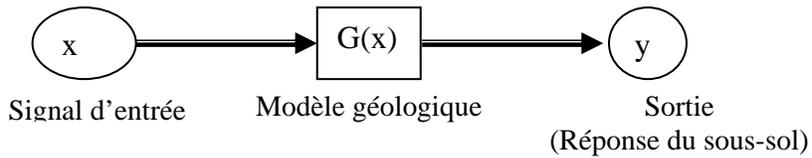


Figure I-01 Organigramme d'un problème direct

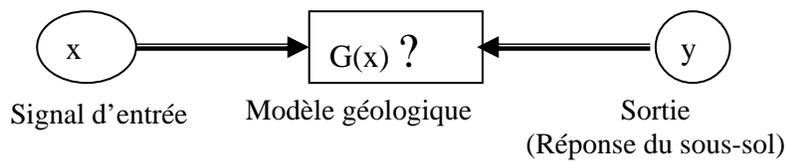


Figure I-02 Organigramme d'un problème inverse

I-2 APERÇU HISTORIQUE

On propose en premier lieu une revue de la littérature sur les méthodes de l'inversion pour pouvoir définir par la suite nos objectifs et notre méthodologie.

Au cours des trente dernières années, les techniques d'inversion ont été intensément utilisées en géophysique dans les études d'exploration, d'exploitation et de développement des gisements d'hydrocarbures, de même pour l'étude des séismes et de la localisation des sources, et d'autres champs de l'application de la géophysique. Les méthodes analytiques ont fourni des solutions précises du problème exprimé explicitement à l'aide de fonctions mathématiques connues. Les chercheurs se sont cependant vite rendus compte que la résolution analytique des problèmes géophysiques complexes était limitée par des hypothèses bien souvent trop simplistes. Un grand nombre d'entre eux se sont donc naturellement tournés vers les méthodes numériques. Ces méthodes, bien que demandant des efforts de calculs parfois plus importants, fournissent des outils puissants d'analyse et de simulation, qui permettent de résoudre des problèmes d'hétérogénéité et de variations latérales des propriétés physiques du sous-sol, dans des structures bi- ou tri-dimensionnelles, fortement hétérogènes et à géométrie complexe [Gjoystdal 1981, Mora 1987, Moraes 2000].

La théorie du problème inverse date de l'époque de Laplace et Gauss, qui ont travaillé sur les

problèmes sur-déterminés en utilisant les critères des moindres absolues (norme l_1) et les critères des moindres carrés. Le concept d'un problème bien posé (mal posé), a été introduit par Hadamard (1900), avec une première application de la théorie du problème inverse en sismologie. Geiger (1910) a formulé l'estimation de la localisation de l'hypocentre, comme problème inverse [Mosegaard, 2002]. Cette méthode est devenue pratique avec l'événement de l'apparition des ordinateurs, qui a aussi donné un important développement de la théorie du problème inverse.

Au cours des années soixante-dix, [Rietsch 1977] a introduit le concept de distribution *a priori* non-informative qui décrit l'état de référence de l'information Backus (1967, 1968) et Jackson (1979) ont proposé l'incorporation de l'information *a priori* dans la formulation du problème inverse linéaire, qui a été généralisée plus tard, pour les problèmes inverses non-linéaires par Tarantola et Valette (1984a, 1984b), introduisant la notion de la conjonction de l'état d'information et l'approche Gaussienne pour la résolution du problème d'estimation non-linéaire des paramètres.

L'ouvrage de *Tarantola* [Tarantola 1987] est la plus importante monographie dans le champ de la théorie du problème inverse. Bien que cet ouvrage soit écrit du point de vue géophysique, les concepts exposés sont de caractère général, qui peut être appliqué dans d'autres champs physiques. Le travail de Tarantola est focalisé sur la description de la formulation probabiliste des moindres carrés pour l'estimation des paramètres.

Ces dernières années, plusieurs méthodes d'inversion à 1D et à 2D ont été échafaudées. Pan et al, (1988) ont inversé les données sismiques dans le domaine (τ, p) pour déterminer la vitesse et la densité en fonction de la profondeur. Leur modèle est constitué d'un grand nombre de couches homogènes. Ils ont généré les données de référence et les inversé directement dans le domaine (τ, p) . Amundsen et Ursin (1991) ont décrit une procédure d'inversion pour obtenir les vitesses, les densités et les épaisseurs des couches à partir des sismogrammes de sismique réflexion multi-offsets, générés par une excitation d'un point source dans un milieu acoustique à stratification horizontale. Par décomposition du sismogramme de référence en ondes sphériques, l'inversion est réalisée dans le domaine du nombre d'onde.

Kormendi et Dietrich (1991) ont proposé une méthode pour déterminer les paramètres élastiques du milieu à stratification horizontale à partir de la réflectivité des ondes-plans. Leur méthode utilise la propriété de la convergence rapide de l'algorithme du gradient conjugué.

Bleistein (1987) a traité une méthode d'inversion multidimensionnelle pour les données prestack et poststack à partir d'un seul profil, cette méthode est basée sur la théorie d'inversion linéaire, en supposant que la vitesse est approximée à partir des données des puits (log sonic) et par perturbations successives jusqu'à l'obtention d'un modèle optimal.

Une méthode d'inversion de vitesse par la technique de DSO (Differential Semblance Optimisation) a été élaborée par Symes et all (1991). Cette technique convertie le problème d'inversion de vitesse en un problème d'optimisation, avec une fonction objective convexe. Par conséquent, le problème de la difficulté de la réalisation de la convergence, liée à la non-convexité de la fonction objective, a été surmonté par une stratégie basée sur la séparation des paramètres. Ces paramètres peuvent être classés en paramètres de contrôle (la composante à grande longueur d'onde de vitesse) et en paramètres d'état (la série de réflectivité). La minimisation se réalise en deux étapes : la première consiste à linéariser la fonction objective par rapport à la variable d'état et ajustement des résidus par rapport à la variable de contrôle.

La méthode d'inversion non-linéaire peut être classée dans les méthodes dites du gradient ou non du gradient, selon qu'elle exploite l'information contenue ou non dans le gradient de la fonction objective. Citons quelques exemples des techniques du gradient, les méthodes de descente, de Gauss Newton, de quasi-Newton, de décomposition en valeurs singulières, du gradient conjugué [Gill 1993]. Tandis que les méthodes du non gradient regroupent les méthodes du recuit stimulé, d'algorithme génétique, des réseaux de neurones artificiels [Carlos 1997,2000, Djarfour 2007].

I-3 CLASSIFICATION DES APPROCHES DU PROBLEME INVERSE

Les méthodes d'inversion peuvent être globales ou partielles, suivant qu'elles prennent en compte toutes les informations sismiques enregistrées ou seulement une partie. Par exemple, à partir des traces sismiques on obtient un profil d'impédance acoustique par un programme d'inversion globale [Bamberger 1982]. En revanche, à partir du temps de parcours extrait des traces, on peut obtenir une répartition de vitesses dans une série de cellules par inversion partielle (tomographie de vitesses) [Bishop 1985, Fara 1987]. Une autre distinction faite sur les nombreuses méthodes d'inversion réside dans la formulation des équations reliant le champ des données observées à la distribution des paramètres à estimer. Cette distinction permet de scinder les méthodes inverses en deux catégories :

a) **Inversion linéaire** qui se ramène à la résolution d'une équation intégrale. Dans ce cas, le recours à l'algèbre linéaire et à l'analyse fonctionnelle permet d'obtenir des résultats précis et des algorithmes efficaces. L'outil fondamental consiste en la décomposition en valeurs singulières (SVD) de l'opérateur considéré et les méthodes algébriques [Levy 1980, Day-Lewis 2002].

b) **Inversion non-linéaire** qui est le plus souvent une estimation des paramètres dans des équations aux dérivées partielles. Notons que la non-linéarité dont il est question fait référence au problème inverse lui-même, et non pas au problème direct (en considérant connus les paramètres). Les méthodes **non-linéaires** utilisent une procédure itérative pour optimiser ces paramètres [Berkhout 1984, Tarantola 1984a, 1984b, Bleistein 1991, Patrick 1996, Snieder 1998].

I-4 FORMULATION MATHEMATIQUE DU PROBLEME INVERSE

I-4-1 PARAMETRISATION DU PROBLEME

Dans cette étape, l'objectif est la détermination d'un nombre minimal de paramètres, appelés paramètres principaux [Tarantola 1987]. Ce nombre doit être capable de décrire le comportement du système selon le phénomène observé. Ces paramètres choisis seront estimés par l'approche inverse. Plus précisément il s'agit tout d'abord du choix des paramètres et de l'optimisation de ce choix. Cependant des difficultés sont inhérentes à cette démarche car ce choix a une influence sur les conditions imposées par le problème de minimisation associé, et sur l'unimodalité de la fonction coût. C'est souvent un problème non trivial que de choisir la représentation discrète des paramètres à estimer, selon le type du phénomène physique observé (réflexion ; réfraction, absorption, temps de parcours, etc.) Plusieurs paramétrisations sont possibles pour le même modèle [Tarantola 1987, Snieder 1999].

Après le choix des paramètres, vient le mode de la représentation de ces paramètres. Les méthodes numériques discrétisent le domaine spatial en engendrant un très grand volume d'éléments. Chacun de ces éléments peut se voir affecté d'une ou de plusieurs valeurs des paramètres physiques inconnus [Gualtierio 1999]. Le problème inverse possède dès lors un nombre appréciable d'inconnues (ou de degrés de liberté) considérées comme les **paramètres** ajustables de la subsurface. Il est clair qu'un nombre d'inconnues trop élevé est souvent la situation la plus rencontrée.

I-4-2 PROBLEME DIRECT ET INVERSE

On peut expliciter et reformuler les notions du problème direct et du problème inverse à l'aide de l'artifice de la présentation assez commode utilisant la notion de système. Les scientifiques cherchent souvent à relier les paramètres qui caractérisent un modèle \mathbf{m} avec les données observées \mathbf{d} . On suppose souvent que la physique fondamentale est bien maîtrisée et un opérateur noté \mathbf{G} peut définir une certaine relation entre \mathbf{m} et \mathbf{d} :

$$\mathbf{G}(\mathbf{m})=\mathbf{d}, \quad \text{I-01}$$

En pratique, \mathbf{d} peut être une fonction du temps et/ou de l'espace, ou encore une collection discrète des observations. En réalité, aucun dispositif expérimental n'est complètement affranchi d'une incertitude dont l'origine la plus simple est la précision finie des mesures. De même, aucun modèle mathématique ne peut refléter exactement la réalité. Il est donc plus réaliste de considérer que l'objet recherché et les mesures sont reliées par une équation de la forme :

$$\begin{aligned} \mathbf{d} &= \mathbf{G}(\mathbf{m}_{\text{exac}}) + \boldsymbol{\eta} \\ &= \mathbf{d}_{\text{exac}} + \boldsymbol{\eta}, \end{aligned} \quad \text{I-02}$$

Où $\boldsymbol{\eta}$ est appelé communément bruit, représente l'ensemble des erreurs et \mathbf{d}_{exac} satisfait l'équation (I.01) pour \mathbf{m} égal au modèle exact, \mathbf{m}_{exac} .

Dans le cas où \mathbf{m} et \mathbf{d} sont des fonctions, alors \mathbf{G} est considéré comme un opérateur. D'autre part, \mathbf{G} représentera une fonction quand \mathbf{m} et \mathbf{d} sont des vecteurs. L'opérateur \mathbf{G} peut prendre plusieurs formes. Dans certain cas, il s'agit d'une équation différentielle ou bien d'une équation aux dérivées partielles. Dans d'autre cas, \mathbf{G} peut être un système algébrique d'équations linéaires ou non-linéaires.

le problème direct consiste à trouver \mathbf{d} connaissant \mathbf{m} . Le calcul de $\mathbf{G}(\mathbf{m})$ peut se faire par la résolution d'une équation différentielle, ou d'une équation aux dérivées partielles, l'évaluation d'une intégrale, ou par l'application d'un algorithme approprié dans le cas d'absence d'une formule analytique explicitant $\mathbf{G}(\mathbf{m})$. Le deuxième problème est le problème inverse qui consiste à calculer \mathbf{m} sachant \mathbf{d} . Le troisième est l'identification du modèle par la détermination de \mathbf{G} sachant \mathbf{m} et \mathbf{d} .

Dans plusieurs cas, on cherche à déterminer un nombre fini de paramètres pour définir un modèle. Les paramètres peuvent décrire une entité physique (densité, vitesse), des coefficients ou toutes autres constantes d'une fonctionnelle représentant un processus physique. Dans ce cas, on peut exprimer les paramètres du modèle par un vecteur de n éléments. De même, les données

sont arrangées en un vecteur de **m** éléments. Ce type de problème est appelé **problème inverse discret** ou **problème d'estimation de paramètres**.

Dans d'autres cas, quand le modèle et les données sont fonction du temps et de l'espace, la tâche associée d'estimation de **m** à partir de **d** est appelée **problème inverse continu**. Ce problème est aussi souvent bien approximé par le problème inverse discret.

On désignera par problème d'estimation des paramètres, un problème avec un petit nombre de paramètres.

Une classe de modèles mathématiques pour lesquels des résultats intéressants existent, est le système linéaire, qui est tout système obéissant au principe de superposition

$$\mathbf{G}(\mathbf{m1} + \mathbf{m2}) = \mathbf{G}(\mathbf{m1}) + \mathbf{G}(\mathbf{m2}) \tag{I-03}$$

et d'échelle $\mathbf{G}(\alpha\mathbf{m}) = \alpha\mathbf{G}(\mathbf{m}), \tag{I-04}$

Dans le cas du problème inverse discret, l'équation (I-01) peut être écrite sous forme d'un système linéaire d'équation algébrique

$$\mathbf{G}(\mathbf{m}) = \mathbf{Gm} = \mathbf{d}. \tag{I-05}$$

Dans le cas d'un problème inverse linéaire continu, **G** s'exprime par un opérateur linéaire intégral, qui prend la forme suivante :

$$\int_a^b g(s, x)m(x)dx = d(s) \tag{I-06}$$

Où $g(s, x)$ est une fonction appelée noyau. Notons que la linéarité est facile à faire apparaître étant donné que :

$$\int_a^b g(s, x)(m1(x) + m2(x))dx = \int_a^b g(s, x)m1(x)dx + \int_a^b g(s, x)m2(x)dx \tag{I-07}$$

et

$$\int_a^b g(s, x)\alpha m(x)dx = \alpha \int_a^b g(s, x)m(x)dx \tag{I-08}$$

L'équation (I.06) où $\mathbf{m(x)}$ est l'inconnu, est appelée équation intégrale de Fredholm de première espèce (IFK) qui couvre un large spectre du problème inverse.

Dans plusieurs cas, le noyau dans l'équation (I.06) peut être écrit de telle façon qu'il dépend explicitement de $(s-x)$, et produit ainsi l'équation de convolution suivante :

$$\int_a^b g(s - x)m(x)dx = d(s) \tag{I-09}$$

Quand le problème direct admet la forme de l'équation (I.09), le fait de trouver $\mathbf{d(s)}$ à

partir de $\mathbf{m}(\mathbf{x})$ est appelé problème de convolution, alors que le problème inverse de détermination de $\mathbf{m}(\mathbf{x})$ à partir de $\mathbf{d}(\mathbf{s})$ est appelé déconvolution, ou convolution inverse.

Rappelons que les problèmes inverses définis par une équation intégrale ne sont jamais bien posés. L'équation intégrale définit un exemple générique de problèmes inverses, dont on trouve de multiples exemples d'application, comme le problème de la déconvolution. Ce qui nous permet de mettre en évidence sur ce cas de figure, les caractéristiques en continu et en discret des problèmes mal posés et d'établir les bases qui serviront au développement des techniques classiques de régularisation.

I-4-3 DIFFICULTE ET RESOLUTION DU PROBLEME INVERSE

I-4-3-1 PROBLEMES BIEN OU MAL POSES

Avant de montrer que les problèmes d'inversion sont parfois mal posés, rappelons ce qu'est un problème bien posé au sens de Hadamard [Snieder 1999, Kern 2003]. Il s'agit d'un problème dont :

1. la solution existe ;
2. la solution est unique ;
3. la solution dépend continûment des données.

D'après Hadamard (1923), seul un problème bien posé pouvait modéliser correctement un phénomène physique. Un problème qui n'est pas bien posé au sens de la définition ci-dessus est dit « mal posé » (ill-posed en anglais).

Un **Problème mal posé** est un problème dont la solution n'existe pas, ou si elle existe, elle n'est pas unique ou ne dépend pas de façon continue des données, autrement dit, qu'il satisfait l'une des conditions **d'inexistence**, de **non-unicité** ou **d'instabilité** de la solution.

On parle **d'inexistence** lorsque aucune fonction satisfaisant les équations du système n'existe, de **non-unicité** lorsque plus d'une solution existe, c'est-à-dire lorsque différentes combinaisons de valeurs des paramètres permettent un ajustement optimum sur les observations, et **d'instabilité** lorsque de faibles variations des données d'entrées modifient de manière importante les résultats du modèle. On parle également d'insensibilité lorsque les observations ne contiennent pas assez d'information pour assurer l'estimation des paramètres.

L'existence et **l'unicité** sont les deux conditions classiques pour l'inversion. **L'existence** peut souvent être assurée en élargissant l'ensemble des solutions (par exemple dans le cas d'un

système linéaire discret surdéterminé, en redéfinissant le problème avec moins d'équations ou en cherchant la solution au sens des moindres carrés). L'**unicité** peut être obtenue en rajoutant des informations sur le modèle pour réduire l'ensemble des solutions (introduction de contraintes sur la solution). La **stabilité** est une condition liée à la physique : elle signifie qu'une petite perturbation sur les données observées n'entraîne qu'une petite erreur sur la solution. Malheureusement cette condition n'est jamais vérifiée pour tous les problèmes inverses, ce qui implique des méthodes dites de régularisation.

I-4-3-2 CONDITIONNEMENT DU PROBLEME INVERSE

La résolution de l'équation de Fredholm de première espèce (I-06) est un problème inverse mal posé car il ne satisfait pas les conditions de l'unicité, l'existence et la stabilité de la solution vis-à-vis des incertitudes de mesures à cause au bruit. La résolution numérique de cette intégrale passe par une étape de discrétisation (paramétrisation déjà discutée) qui conduit à :

$$\mathbf{d} = \mathbf{G}\mathbf{m} + \boldsymbol{\eta},$$

où \mathbf{d} désigne le vecteur des données, \mathbf{G} la matrice de transformation, \mathbf{m} le vecteur des paramètres à estimer et $\boldsymbol{\eta}$ le vecteur contenant les bruits de mesures dues aux erreurs de différentes sources.

Dans ce cas, l'opérateur \mathbf{G} est une matrice de modélisation de dimension $n \times m$.

- 1 Le critère d'existence de la solution est rempli si le **rang(G)**= n (rang maximal) ;
- 2 Le critère d'unicité de la solution est rempli si le **rang(G)**=m ;

Pour assurer les deux critères d'existence et d'unicité de la solution simultanément, il faut que la matrice \mathbf{G} soit une matrice carrée, et de rang plein. Cette condition est aussi suffisante pour assurer l'existence et l'unicité de la solution du problème inverse.

- 3 La stabilité est mesurée par le conditionnement de la matrice \mathbf{G} . Si \mathbf{G} est régulière, son conditionnement est mesuré par :

$$\text{Cond}(G) = \|G\| \|G^T\| \tag{I-10}$$

Où $\|G\|$ est la norme matricielle induite d'une norme vectorielle sur \mathbb{R}^N (ici $N=2$).

On a toujours $\text{Cond}(G) \geq 1$. Une matrice bien conditionnée (donnant lieu à un problème bien stable) est telle que $\text{Cond}(G)$ soit proche de la valeur 1. Notons que $\text{Cond}(I)=1$ où I est la matrice identité. Avec la norme euclidienne sur \mathbb{R}^N on montre que :

$$\text{cond}(G) = \frac{\lambda_{\max}}{\lambda_{\min}} \tag{I-11}$$

Où λ_{\max} et λ_{\min} sont les valeurs singulières de \mathbf{G} . Le conditionnement dépend donc de l'étalement du spectre de \mathbf{G} .

Le conditionnement est relié au problème de sensibilité par rapport au bruit car si l'on considère la relation matrice/vecteur à inverser $d=Gm$ et que δd est une perturbation sur les observations, alors la perturbation δm sur la solution vérifie [Jupp 1975, Lascaux 2000, 2004] :

$$\frac{\|\delta m\|}{\|m\|} \leq \text{cond}(G) \cdot \frac{\|\delta d\|}{\|d\|} \tag{I-12}$$

Ainsi, si G est bien conditionnée ($\text{Cond}(G) \approx 1$), alors une petite perturbation $\|\delta d\|$ relative aux données $\|d\|$ engendre une petite perturbation $\|\delta m\|$ relative aux paramètres $\|m\|$. Une grande valeur de $\text{cond}(G)$ est l'indice d'une grande sensibilité du problème aux perturbations possibles des données. Dans ce cas, $\frac{\|\delta m\|}{\|m\|}$ ne peut être borné par une petite constante même si $\frac{\|\delta d\|}{\|d\|}$ l'est.

Ainsi, les solutions basées sur l'inversion (inverse de G ou solution des moindres carrés) $\min \|d - Gm\|$ ne donnent en général pas une solution correcte mais un résultat dominé par le bruit.

Les causes d'incertitude sont nombreuses :

- Les données ont une origine expérimentale, ce qui implique l'existence d'erreurs de mesure.
- Elles sont collectées en nombre fini, même si, dans le modèle mathématique, elles sont décrites par des fonctions.
- L'algorithme d'inversion lui-même peut parfois créer une altération des données : interpolation requise pour la discrétisation d'un modèle initialement continu par exemple.
- Le modèle lui-même procède d'une idéalisation de la réalité physique et repose sur des hypothèses simplificatrices, il est donc également une source d'incertitudes.

L'étude du problème inverse serait donc incomplète sans l'examen de la stabilité des solutions vis-à-vis de perturbations du second membre.

Dans le cas général (matrice G rectangulaire quelconque), on est confronté à un problème double :

- Définir une solution approchée à (I-05), stable par rapport aux erreurs sur \mathbf{d} .

– Faire en sorte que le conditionnement de la méthode d'inversion retenue ait une valeur inférieure à $\text{cond}(G)=\lambda_{\max}/\lambda_{\min}$,

Pour cela, il faudra « régulariser » l'inversion du problème.

I-4-3-3 REGULARISATION ET SOLUTION DU PROBLEME INVERSE

Durant les années 60, le mathématicien Russe Tikhonov a mis en place les bases théoriques des méthodes modernes d'inversion en introduisant le concept de solutions régularisées [Kern 2003]. Ces solutions sont un compromis entre la fidélité aux données et la régularité. Ce compromis permet d'obtenir des problèmes reformulés donc "bien posés". Cette méthode est nommée principe de régularisation, elle permet de stabiliser les problèmes d'inversion mal posés.

La régularisation d'un problème mal posé consiste à définir une solution unique pour toutes les mesures possibles dans l'espace d'observation, et à assurer la stabilité de cette solution vis-à-vis des erreurs sur ces mesures.

L'objectif de toute méthode de « régularisation » d'un problème inverse est double. D'une part, on définit un cadre qui permet d'incorporer toutes les informations supplémentaires dont on dispose, et qui proviennent nécessairement de considérations physiques ou qualitatives extérieures aux mathématiques (les mathématiques ne remplacent pas l'absence d'information, elles peuvent seulement utiliser au mieux celle dont on dispose). D'autre part, on fait en sorte que la solution obtenue, compte tenu de ces informations supplémentaires, soit stable par rapport aux données [Geoffroy 2005]

La plupart des méthodes qui ont été proposées, depuis les années 70 pour résoudre et stabiliser les problèmes mal posés, sont classées en deux grandes familles : celles qui procèdent par troncature et celles qui procèdent par optimisation.

A)- INFORMATIONS A PRIORI

On appelle information *a priori* tout type d'information que l'on peut obtenir indépendamment des données de mesure. Ces informations *a priori* correspondent à une information directe sur la variable inconnue du problème inverse. Dans le cas des modèles géologiques, elle consiste en des valeurs "locales" des propriétés acoustiques, le plus souvent obtenues par interprétation précédente ou bien disponibles en un nombre limité de points au sein des puits. Après une éventuelle "mise à l'échelle", ces informations complémentaires peuvent être prises en compte dans la procédure d'inversion à l'aide d'un "critère de plausibilité", qui mesure

la déviation entre les paramètres optimisés et les estimations *a priori* de ces paramètres. De cette manière, non seulement un maximum d'information est introduit dans le processus d'inversion, mais surtout le processus d'inversion est stabilisé et le problème est rendu non-unique [Jiang 1989].

B)- REGULARISATION PAR TRONCATURE

Il existe de nombreuses techniques permettant de conditionner la matrice \mathbf{G} , en particulier les techniques de décomposition [Lines 1984] par exemple la décomposition en valeurs singulières (SVD).

Un exemple typique de méthode de la première famille est la décomposition tronquée en valeurs singulières pour supprimer le mauvais conditionnement du problème. Il suffit de tronquer le développement en retenant les composantes correspondant à des valeurs singulières suffisamment grandes pour que les termes d'erreur restent faibles. Elle est très efficace pour assurer la stabilité numérique.

C)- REGULARISATION PAR OPTIMISATION

Les méthodes que nous avons vues précédemment sont qualifiées de directes, parce qu'elles donnent (dans le cas de la dimension finie) la solution exacte du problème régularisé (aux erreurs d'arrondis près) en un nombre fini d'opérations. Pour des problèmes de taille modérée, ces méthodes sont les plus utilisées. Toutefois, pour des problèmes de grande taille, où les matrices obtenues après discrétisation sont souvent creuses, ces méthodes s'avèrent inadaptées car, d'une part, le nombre d'opérations devient trop grand, et d'autre part, ces méthodes ne respectent pas la structure des matrices. Une alternative consiste alors à se tourner vers des méthodes itératives en construisant une suite de solutions approchées qui (dans le cas non bruité) convergent vers la solution désirée, dans le cas où le rapport signal/bruit est faible.

Dans cette classe des méthodes, la régularisation est exprimée par la minimisation d'une fonction objective des moindres carrés, le problème revient à rechercher la solution \mathbf{m} minimisant la norme L_2 de l'espace D^*M dans lesquels D et M décrivent respectivement les espaces de données et de paramètres. Dans ce cas, le problème inverse peut être formulé ainsi : Chercher \mathbf{m} de façon à minimiser les distances suivantes [Tarantola 1987] :

$$E(\mathbf{m}) = \left\| d^{obs} - g(\mathbf{m}) \right\|_D^2 - \left\| \mathbf{m} - \mathbf{m}_0 \right\|_M^2 \quad \text{I-13}$$

Les normes $\| \cdot \|_D$ et $\| \cdot \|_M$ peuvent être décrites à l'aide de deux matrices symétriques définies positives, appelées en général matrices de covariance [Tarantola 1984a, 1984b, Duijndam 1988, Shang 1994]. Par conséquent, la fonction objective s'écrit comme suit :

$$2J(m) = \left(d^{obs} - g(m) \right)^t C_p^{-1} \left(d^{obs} - g(m) \right) + \left(m - m_0 \right)^t C_m^{-1} \left(m - m_0 \right) \quad \text{I-14}$$

Où C_p et C_m sont les matrices de covariance des données et du modèle. La pondération permet d'augmenter l'importance de certaines observations par rapport à d'autres au sein du processus d'optimisation. Les poids accordés aux observations sont inversement proportionnels à leur écart-types de ceux-ci, les plus grands poids étant attribués aux observations les plus "sûres". La pondération est fortement conseillée lorsqu'on travaille avec des observations de différents types (par exemple sismique de puits et sismique de surface) afin que les observations ayant des ordres de grandeurs plus élevés ne dominent le processus d'estimation établie [Tarantola 1984a, 1984b].

Le problème de minimisation d'une fonctionnelle se résoud le plus souvent par une méthode de descente. On détermine, pour une approximation donnée, un sens de décroissance de la fonctionnelle J en se déplaçant dans ce sens. Si le pas du déplacement n'est pas très grand, la valeur de la fonctionnelle diminue nécessairement.

Comme la fonction coût associée au problème inverse est quadratique, l'algorithme de minimisation choisi est le gradient conjugué. La différentiation de $J(\mathbf{m})$ par rapport à \mathbf{m} donne la fonction du gradient de celle-ci [Tarantola 1984a] ;

A partir de l'équation (1-14), la détermination du minimum de J revient à le rechercher dans la direction du gradient [Tarantola 1984a, 1984b, Faurre 1988]:

$$\mathbf{m}_{k+1} = \mathbf{m}_k - \alpha (\text{grad}(J))_k \quad \text{I-15}$$

Où α est une constante arbitraire ayant une petite valeur assurant la convergence de l'algorithme. Cet algorithme améliore au cours des itérations les valeurs des paramètres recherchés. Bien entendu, ce n'est pas l'unique algorithme pour résoudre ces problèmes d'optimisation.

d) CRITERE D'ARRET DU PROCESSUS D'INVERSION

Plusieurs critères d'arrêt de l'ensemble du processus d'estimation, basés sur l'analyse de l'évolution de la fonction objective et sur le comportement des paramètres ajustables, existent. Le processus d'estimation s'arrête lorsque:

1. la valeur de la fonction objective atteint zéro (cette condition est très rare voire jamais rencontrée) ;
2. la fonction objective ne diminue pas ou seulement très peu après un nombre donné d'itérations d'optimisations successives ;
3. les paramètres ajustables ne sont que très faiblement modifiés après un nombre donné d'itérations d'optimisations successives ;
4. le nombre maximum d'itérations du processus d'estimation est atteint, indépendamment de la valeur de la fonction objectif et de l'estimation des paramètres ajustables.

Dans les trois premiers cas, l'arrêt du processus signifie que les paramètres ajustables ont convergé vers leurs valeurs minima, mais ne signifie pas que l'optimum global est atteint. En effet, le processus peut s'arrêter à un minimum local puisque de faibles variations des paramètres ne permettent pas de s'en éloigner. Il est dès lors toujours intéressant d'effectuer un test.

I-5 ETUDE DE CAS ET EXEMPLE D'APPLICATION**I-5-1 LA TOMOGRAPHIE SISMIQUE**

La tomographie appartient à la classe des problèmes inverses, portant sur des mesures de l'énergie qui se propage dans un milieu. En utilisant le caractère de cette énergie à la réception (amplitude, temps de parcours), les caractéristiques de l'intérieur de ce milieu seront explorées. Principalement, la description de cette énergie peut être effectuée par l'intégrale d'un certain nombre de paramètres. La détermination de ces paramètres à partir de leur intégrale constitue le problème de la tomographie. Bien qu'il existe des méthodes qui se basent sur la résolution complète des équations d'ondes, les méthodes les plus utilisées sont celles basées sur l'approximation de l'onde par un rai [Landa 1988, Hirahara 1993, Zhou 2001].

Pour un trajet arbitraire L , le temps de parcours est donné par [Peterson 1985, Krajewski 1989, Asakawa 1993, Liao 1997, Erwan 2004] :

$$T = \int_L s(x) dl \quad \text{I-16}$$

où dl est une distance infinitésimale le long du trajet P

Le problème direct est non-linéaire, du fait que le trajet lui-même dépend des paramètres du milieu comme la distribution de vitesses. Il convient de rappeler ici que les problèmes inverses définis par cette équation intégrale ne sont jamais bien posés.

Si nous avons un ensemble d'observations $t=t_1, \dots, t_{nt}$, pour nt couples émetteur-récepteur dans un domaine de lenteur $s(x)$ (inverse de la vitesse), les données expérimentales disponibles sont presque toujours constituées d'un nombre nécessairement fini de points du domaine d'observations. Elles sont donc naturellement discrètes, et si le domaine est discrétisé en np cellules, alors l_{ij} désigne la longueur du $i^{ème}$ rai traversant la $j^{ème}$ cellule. Le temps de trajet total pour un rai sismique s'exprime par [Krajewski 1989, Stewart 1991, Zhang 1994, Hanyga 1995, Stefani 1995, Hyndman 1996, Ming 2003, Erwan 2004]:

$$\sum_{j=1}^{np} l_{ij} s_j = t_i \quad \text{Avec } i=1, \dots, nt \tag{I-17}$$

Nous pouvons noter que la plupart des cellules ne sont pas visitées pour un rai donné. La figure.I.03 illustre le trajet d'un rai dans un milieu aux propriétés acoustiques non homogènes. Les inflexions du rai suivent le principe de Fermat et répondent aux lois de Snell-Descarte.

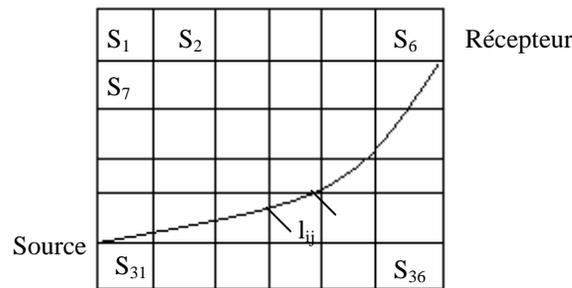


Figure.I.03 Le trajet d'un rai

I-5-2 TOMOGRAPHIE LINEAIRE OU NON LINEAIRE

Dans un problème direct, il s'agit de déterminer L et t en supposons connu S . Cela revient à calculer le trajet des rais entre chaque couple émetteur-récepteur puis de calculer l'intégrale des temps de parcours le long de chaque parcours.

En inversion non-linéaire, seule t est donné. Le but est de trouver S et L . Il n'existe cependant pas de méthodes directes assurant une combinaison entre L et S sachant t . Il faut utiliser un processus itératif.

En inversion linéaire, L et t sont donnés. L'objectif est de déterminer S . Les trajets sont supposés connus et sont par conséquent indépendants du champ des lenteurs. Dans ce cas, le trajet est supposé une ligne droite reliant la position de la source à celle du récepteur (principe du trajet minimal).

Après la discrétisation du milieu à imager (paramétrisation discuté précédemment), il est possible de réécrire l'équation I-17 sous la forme matricielle suivante :

$$S = \begin{pmatrix} s_1 \\ s_2 \\ \cdot \\ s_n \end{pmatrix}, \quad t = \begin{pmatrix} t_1 \\ t_2 \\ \cdot \\ t_m \end{pmatrix}, \quad G = \begin{pmatrix} l_{11} & l_{12} & \cdot & l_{1n} \\ l_{21} & l_{22} & \cdot & l_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ l_m & l_{m2} & \cdot & l_{nm} \end{pmatrix}$$

L'équation devient l'équation de base de la modélisation directe par la méthode des rais [Bregman 1989, Nemeth 1997, Erwan 2004].

$$\mathbf{GS}=\mathbf{t}$$

I-18

La matrice G dépend des paramètres du modèle S , le problème est donc non-linéaire.

D'autre part, plus la discrétisation est fine (le pas est petit), plus les lignes de la matrice sont semblables et donc proche d'une matrice singulière. Cependant, la discrétisation peut contribuer à rendre un problème inverse mal posé en un problème bien posé, rappelons que les problèmes inverses liés à la physique sont en général mal posés fondamentalement.

La linéarisation de ce problème suppose que :

1. Le trajet du rayon connu *a priori*, est justifié sous l'approximation linéaire, qui néglige la dépendance entre le trajet du rayon et la distribution de la lenteur.
2. Le trajet est supposé une ligne droite reliant la position de la source à celle du récepteur (principe du trajet minimal).

Cette tomographie linéaire est initialement appliquée en médecine, puis progressivement introduite en géophysique.

I-5-3 LES METHODES ALGEBRIQUES DE RECONSTRUCTION TOMOGRAPHIQUE**I-5-3 –a LA METHODE ART**

La matrice \mathbf{G} caractérise la géométrie d'acquisition des données. Chaque ligne de cette matrice représente l'équation d'un rayon de projection. L'ART est un algorithme itératif de résolution du système d'équations faisant agir une ligne de la matrice à la fois, dans la version de base, l'idée est de corriger la solution à chaque itération de manière à la rendre consistante avec l'équation considérée lors de cette itération. Cette modification est donnée par la formule suivante [Mason 1981, Peterson 1985, Stewart 1991] :

$$s_j^{k+1} = s_j^k + \frac{t_n^{k+1} l_{nj}}{\sum_{m=1}^M l_{nm}^2} \quad \text{I-19}$$

Où $t_n^{k+1} = d_n - t_n^k$ est la différence entre le résidu de la première itération et de la $k^{\text{ième}}$ itération, s_j^k est l'approximation $j^{\text{ème}}$ des paramètres du modèle à la $k^{\text{ème}}$ itération, Notons que $t_n^1 = 0$.

L'image est initialisée à $s^{(0)}$. A partir de cette image on calcule le temps de parcours qui lui est associé. L'erreur est la différence entre le temps calculé et le temps observé. A chaque itération k on ajoute à l'image $s^{(k)}$ un terme additif ne dépendant que de l'équation j . Ensuite on passe à la deuxième ligne de la matrice et ainsi de suite. Le processus est continu pour les équations, et on peut retourner à la première équation et résoudre la séquence de toutes les autres équations une autre fois. Cette procédure est similaire à celle qui est connue sous le nom de la méthode de Kaczmarz [Stewart 1991] pour la résolution d'un système d'équations. L'algorithme ART est instable en présence de données bruitées.

I-5-3 –b LA METHODE SIRT

Cette méthode dresse le problème de la convergence de la méthode ART. C'est une méthode itérative comme l'ART, mais au lieu de procéder rayon par rayon, elle procède pixel par pixel [Krajewski 1989, Stewart 1991].

A chaque pixel, toutes les mesures de projection correspondante à tous les rayons passant par ce pixel sont calculées et sommées ; un facteur de correction est appliqué à la densité des mesures de ce pixel [Krajewski 1989].

$$s_j^{k+1} = s_j^k + \frac{1}{N} \sum_{n=1}^N \frac{t_n^{k+1} l_{nj}}{\sum_{m=1}^M l_{nm}^2} \quad \text{I-20}$$

où N est le nombre de rayons passant par la cellule.

I-5-3-c LE GRADIENT CONJUGUE (GC)

La méthode du GC est très utilisée pour la résolution des systèmes de matrices symétriques, positives et définies.

Un des intérêts du GC est qu'il ne nécessite pas de calculer explicitement l'inverse $G^T G$. L'algorithme démarre avec une solution initiale *a priori* et converge vers la solution optimale en prenant comme direction la plus grande pente. Cette direction est l'opposée du gradient de la fonctionnelle quadratique $(t - Gs)^T (t - Gs)$ [Scales 1987].

Une caractéristique intéressante du GC pour la résolution des problèmes mal posés est que lorsque l'algorithme est appliqué pour résoudre les équations normales non régularisées ($G^T G S = t$) alors, les composantes de basses fréquences de la solution tendent à converger plus rapidement que les composantes hautes fréquences. Le GC possède donc un effet « régularisateur ». Il se résume aux étapes suivantes telle qu'il a été présenté [Faurre 1988, Lascaux 2000, 2004]:

Etant donnés $G, t, k_{\max}, s^{(0)}, \varepsilon$

Initialisation : $k=0$ nombre d'itération

- $r^{(0)} = Gs^{(0)} - t$
- $d^{(0)} = r^{(0)}$

tant que ($k < k_{\max}$) ou $\left(\frac{\|r\|}{\|b\|} < \varepsilon \right)$

- $k = k + 1$
- $\alpha^{(k)} = \frac{d^{(k)T} r^{(k)}}{d^{(k)T} G d^{(k)}}$
- $s^{(k+1)} = s^{(k)} + \alpha^{(k)} d^{(k)}$
- $r^{(k+1)} = t - Gs^{(k+1)}$
- $\beta^{(k)} = \frac{r^{(k+1)T} r^{(k+1)}}{r^{(k)T} r^{(k)}}$
- $d^{(k+1)} = r^{(k+1)} + \beta^{(k)} d^{(k)}$

Fin tant que

I-6 DECONVOLUTION STRATIGRAPHIQUE

Après corrections, filtrages et migration, on peut considérer que l'essentiel des problèmes de types géométriques ont été résolus sur la section sismique, dans la mesure où les traitements sont correctement appliqués. Si dans certains cas, les formes géométriques peuvent aider à l'identification des faciès, la méthode la plus rigoureuse est le rattachement à un puits, sachant que les mesures sismiques spéciales sont nécessaires pour en assurer la validité. En dehors de ce cas ou bien loin du puits, il faut utiliser toutes les informations qui permettent d'aboutir à cette identification.

Pour caractériser le faciès on utilise un certain nombre de paramètres telles que la vitesse d'intervalle qui est déterminée par l'analyse de vitesses et qui sera interprétée en terme de faciès global ou de leurs variations. Un autre paramètre qui est l'impédance acoustique, utilisée pour l'étude détaillée de faciès, est calculé par la technique d'estimation de pseudo-log ou par inversion stratigraphique.

Afin d'extraire des données mesurées sismiques les grandeurs caractéristiques de la structure du sous-sol (réflectivité, impédance acoustique), un modèle direct décrivant l'interaction entre le signal sismique et le sous-sol est développé. La propagation des ondes élastiques dans un milieu est gouvernée par l'équation d'onde. Pour calculer la réponse sismique d'un tel modèle, on fait appel à la méthode d'approximation par le modèle de convolution [Sheriff 1984, Tarantola, 1984, Boyer 1994, Buxton 2000, Boschetti 2000, Perrodon 2002, Glangeaud 2004]. Ce modèle est dérivé de l'approximation acoustique de l'équation d'onde élastique. Il s'exprime sous la forme d'un produit de convolution donné par l'équation suivante :

$$T(t) = W(t)*R(t) + B(t) \quad \text{I-21}$$

où (*) représente le produit de convolution, T(t) désigne le signal mesuré, W(t) le signal émis, R(t) la réflectivité et B(t) le bruit sismique.

Connaissant les données mesurées, la résolution du problème inverse permet de remonter aux grandeurs recherchées (réflectivité puis impédance acoustique). Compte tenu de la définition du modèle direct, le problème inverse spécifique traité dans cette partie est celui de la déconvolution. La résolution de cette classe de problèmes se heurte à deux difficultés liées d'une part à la présence de bruit et d'autre part à la perte d'informations due à la convolution.

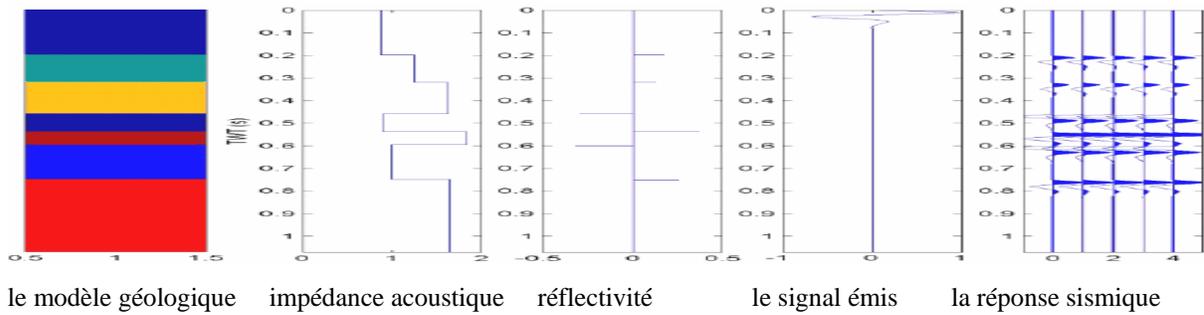


Figure I-04 les étapes du problème direct

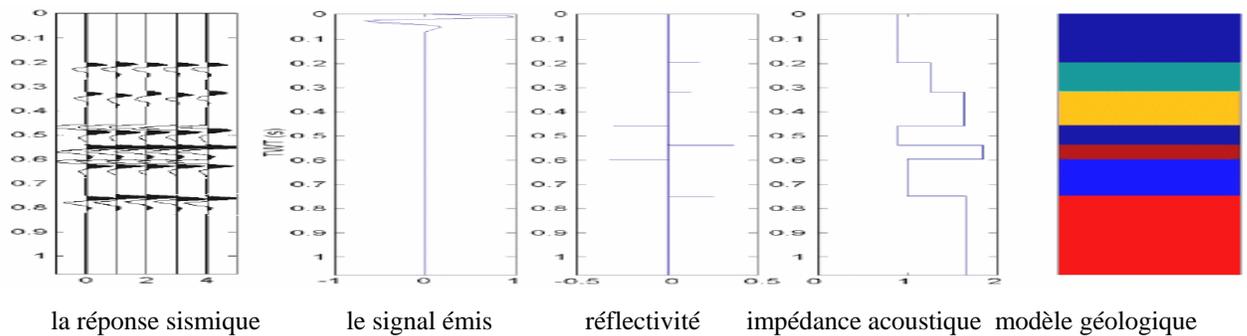


Figure I-05 les étapes de l'inversion stratigraphique

L'Impédance Acoustique (IA) est le produit de la vitesse par la densité de roches ; ce qui signifie que l'IA est une propriété de la roche mais non pas de l'interface. Cette distinction est la caractéristique la plus importante de l'IA.

La dernière étape consiste à passer des coefficients de réflexion aux impédances acoustiques qui traduisent beaucoup mieux les caractéristiques des roches. Ce passage est donné par l'expression suivante [Bilgeri 1981, Cooke 1983, Henry 1997, Glangeaud 2004] :

$$r_m = \frac{I_{m+1} - I_m}{I_{m+1} + I_m} \tag{I-22}$$

La suite des coefficients de réflexion est en quelque sorte une dérivée (dérivée logarithmique) du log d'impédance acoustique. A partir des coefficients de réflexion, il est évident que l'on peut revenir au log d'impédance acoustique par une opération d'intégration **inverse**. En pratique, le plus souvent est d'inverser la formule suivante [Mari 1998] :

$$I_{m+1} = I_m \frac{1+r_m}{1-r_m}$$

I-23

Cependant, un calage sera nécessaire tant pour déterminer la constante de l'équation ci-dessus que pour tenir compte des basses fréquences non présentes sur la trace sismique.

I-7 CONCLUSION

Dans ce chapitre, nous avons présenté les principes de base de la théorie du problème inverse, et décrit en particulier leur caractère mal posé. Il en ressort que le problème inverse comprend trois étapes essentielles à savoir la paramétrisation, la modélisation, et l'inversion. Si le problème direct est toujours bien posé, alors le problème inverse est toujours mal posé au sens de Hadamard. Il devient donc nécessaire de développer des méthodologies d'inversion qui visent le traitement de ce caractère mal posé.

Une difficulté pratique de l'étude des problèmes inverses est qu'elle demande souvent une bonne connaissance du problème direct, ce qui se traduit par le recours à une grande variété de notions tant physiques que mathématiques. Le succès dans la résolution d'un problème inverse repose en général sur l'intégration de trois types d'information, à savoir, les données de modélisation, les données d'observation et l'information *a priori*.

Nous avons exposé deux types du problème inverse. L'inversion partielles qui utilise le temps de trajet pour passer au modèle géologique paramétré par la vitesse, le deuxième type est l'inversion globale qui part de toute la trace sismique pour retrouver l'impédance acoustique le paramètre qui caractérise le sous sol. Nous avons présenté et analysé les équations qui régissent les applications considérées.

CHAPITRE II

Les Réseaux de Neurones Artificiels



II LES RESEAUX DE NEURONES ARTIFICIELS

II-1 INTRODUCTION :

Les réseaux de neurones artificiels (RNA) ont fait l'objet d'un intérêt soutenu et de nombreux travaux depuis plus d'une vingtaine d'années à cause de leur capacité à résoudre des problèmes non linéaires par apprentissage. Ce sont donc des méthodes numériques qui appartiennent au domaine des méthodes non paramétriques, en ce sens qu'ils ne reposent pas sur un modèle explicite des données. Il a été également démontré qu'un réseau de neurones ayant une couche cachée et une fonction sigmoïde est un « approximateur universel » [Hertz 1991, Jean 1994, Gondran 1997] en ce sens qu'il peut approximer toute fonction multivariable non linéaire et ses dérivées avec une précision fixée, pourvu qu'il soit soumis à un apprentissage efficace. Ce potentiel d'approximation est cependant difficile à exploiter, car il n'existe à ce jour aucune méthodologie précise permettant de réaliser avec certitude l'approximation désirée pour un système donné. C'est pourquoi le choix et l'application d'un modèle neuronal demeure un domaine de recherche très actif contrairement aux modèles stochastiques classiques.

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 se sont développés [Hertz 1991, Gérard 2004]. Les réseaux de neurones artificiels sont des modèles, à ce titre ils peuvent être décrits par leurs composants, leurs variables descriptives et les interactions des composants.

II-2 HISTORIQUE DE LA TECHNIQUE

Sous le terme de réseau de neurones, on regroupe aujourd'hui un certain nombre de modèles, qui ont pour objectif d'imiter certaines fonctions du cerveau humain. Cette imitation repose sur la reproduction de quelques unes de ses structures de base [Jean 1994, Hertz 1991, Gondran 1997].

Historiquement, les origines de cette discipline sont diverses. En 1943, McCulloch et Pitts ont étudié un ensemble de neurones formels interconnectés, et ont montré leurs capacités à calculer certaines fonctions logiques.

En 1949, Hebb dans une perspective psycho-physiologique [Gérard 2004, Jean-Michel 1994, Hertz 1991] souligna l'importance du couplage synaptique dans le processus d'apprentissage. Les premiers réseaux de neurones appartiennent à un type appelé perceptron, terme proposé par le chercheur américain Frank Rosenblatt (1962) [Jean 1994]. Ils sont

conçus à partir, des connaissances de l'époque, des systèmes de visions biologiques. Ces machines sont présentées comme des dispositifs de reconnaissance des formes.

Les recherches sur ce genre de systèmes se sont poursuivies jusqu'à 1969. À cette date, les chercheurs américains Marvin Minsky et Seymour Papert publient leurs travaux dans lesquels ils décrivent une série de problèmes de reconnaissance des formes qu'un perceptron ne peut pas résoudre. Ces problèmes sont appelés *problèmes d'apprentissage difficiles*, raison pour laquelle les travaux sur la technologie des réseaux de neurones ont été interrompus [Hertz 1991].

Cependant quelques études se sont poursuivies dont certaines ont donné lieu au système WISARD, réalisé en 1980. Le nom du système dérive de celui de ses inventeurs britanniques, Wilkie, Stonham, et Aleksander WISARD. Le problème d'*apprentissage difficile* est résolu en incorporant plus d'une couche dans le réseau. Les perceptrons à couches multiples peuvent accomplir les tâches que Minsky et Papert avaient démontré comme impossibles à résoudre par un perceptron à couche unique [Jean-Michel 1994, Hertz 1991].

Le regain d'intérêt pour les réseaux de neurones est dû à des contributions originales, comme celle de Hopfield (1982) qui a montré l'analogie des réseaux de neurones avec certains systèmes physiques permettant d'appliquer un formalisme riche et bien maîtrisé. Plus récemment, en 1985 des nouveaux modèles mathématiques avaient dépassé les limites du perceptron [Gérard 2004, Hertz 1991].

Aujourd'hui, les premières applications des réseaux de neurones commencent à voir le jour, et cette discipline va concerner un large public d'étudiants, de chercheurs et d'universitaires. En terme d'application en géophysique on peut citer les travaux de [Djarfour 2007, carlos 2000, Devilee 1999, Groof 1992] qui ont utilisé les réseaux de neurones de type perceptron multicouche pour la résolution du problème inverse. Les travaux de [Djarfour 2008, Mousset 1992] pour le filtrage des données sismiques, ainsi que la compression des images sismiques [Epping 1992].

II-3 CARACTERISTIQUES DES TECHNIQUES D'INTELLIGENCE ARTIFICIELLE

Les RNA font partie des réseaux adaptatifs non linéaires qui sont des systèmes constitués d'un grand nombre d'unités élémentaires (appelées aussi cellules ou agents) telles que :

- Ces unités forment un réseau qui traite l'information d'une manière parallèle et distribuée ;

- Ces unités interagissent d'une façon non linéaire et compétitive dans un certain environnement, le plus souvent en l'absence d'autorité centrale ;
- Des opérateurs modifient l'état des unités et leurs interactions de telle façon que le système global (population d'agents) s'adapte progressivement à l'environnement, qui n'est pas forcément fixe.

Les RNA font apparaître en général des propriétés d'auto-organisation très particulières. Ces propriétés découlent naturellement du caractère non-linéaire des interactions tel que le comportement de la somme des agents ne peut pas se déduire directement de la somme des comportements, car le principe de superposition n'est évidemment plus valide. Il faut remarquer que, du caractère non-linéaire des interactions découle aussi l'existence du groupe d'agents cohabités, c'est-à-dire pour qu'un bon agent soit véritablement efficace au niveau de la somme d'agents, il doit nécessairement agir en collaboration avec d'autres bons agents. Cette propriété est bien connue en génétique et est appelée "épistasie", elle réfère dans ce cas à des groupes cohabités de gènes [Jean 1994].

Le caractère distribué et parallèle du traitement par les RNA offre des avantages de robustesse par rapport à des données incomplètes, contradictoires et même par rapport à des données bruitées.

L'auto-organisation de ces agents est fondamentalement reliée au concept d'apprentissage, autrement dit aux mécanismes permettant l'amélioration des performances futures, sur la base d'une connaissance acquise au fur et à mesure. Bien sûr, la nécessité d'un tel apprentissage, et donc de l'emploi des RNA pour l'étude d'un système, ne se justifie que dans le cas où il serait difficile, voire impossible de concevoir «classiquement» un système basé sur la connaissance disponible *a priori*, capable d'analyser le système d'une façon satisfaisante. Ces difficultés de conception découlent généralement de la complexité du système, dont il faut assurer la conduite : non-linéarité, dimensionnalité élevée, erreur de modélisation, bruit de mesure [Gérard 2004].

Etant donné que le problème inverse est un problème mal posé au sens d'Hadamard (1923), nous nous proposons d'analyser l'apport des méthodes d'intelligence artificielle (RNA) dans l'amélioration de la solution du problème inverse.

Il nous apparaît nécessaire de commencer par présenter ces méthodes puis d'expliquer les manières les plus simples de les mettre en œuvre.

II-4 LE NEURONE ARTIFICIEL :

Un neurone artificiel est un processeur élémentaire. Il reçoit un nombre variable d'entrées en provenance des neurones amonts. A chacune de ces entrées est associé un poids w représentatif de la force de connexion. Chaque processeur élémentaire est doté d'une sortie unique, qui se ramifie ensuite pour alimenter un nombre variable de neurones avals (Figure II.01).

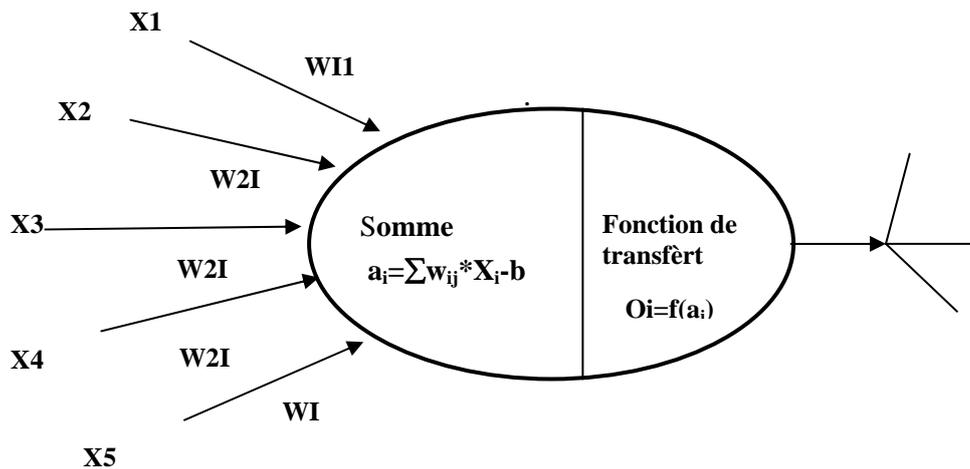


Figure II-01 : Neurone formel.

II-5 COMPORTEMENT D'UN NEURONE ARTIFICIEL :

On distingue deux phases. La première est habituellement le calcul de la somme pondérée des entrées selon l'expression suivante :

$$o_i = \sum_{j=1}^n w_{ij} x_j \quad (\text{II-01})$$

A partir de cette valeur, une fonction de transfert calcule la valeur de l'état du neurone.

Dans les premiers travaux fondateurs de McCulloch et Pitts sur le Perceptron la fonction d'activation était une simple fonction de seuillage brutale de type $sign(x)$. Dès lors, la sortie du réseau ne pouvait prendre que des valeurs binaires (-1 ou +1). Cela était peut-être utile pour des tâches de classification. Cependant, de nos jours, il est plus usuel d'utiliser une fonction continue de forme sigmoïdale. C'est de loin la fonction la plus utilisée dans la littérature sur les réseaux de neurones artificiels (RNA). Elle a les avantages d'être continue, dérivable :

$$p_i = f_i(o_i) = \frac{1 - e^{-o_i}}{1 + e^{+o_i}} \quad (\text{II-02})$$

C'est cette valeur qui sera transmise aux neurones avals.

II-6 LES RESEAUX DE NEURONES :

Les réseaux de neurones sont constitués de neurones élémentaires connectés entre eux par l'intermédiaire de poids qui jouent le rôle des synapses. L'information est portée par la valeur de ces poids, tandis que la structure du réseau de neurones ne sert qu'à traiter cette information et à l'acheminer vers la sortie.

Un RNA est généralement constitué de trois couches (Figure II.02). La première est celle de l'entrée. La deuxième, dite couche cachée, constitue le cœur du réseau ; ses fonctions d'activation sont du type sigmoïde ou linéaire. La troisième est la couche de sortie.

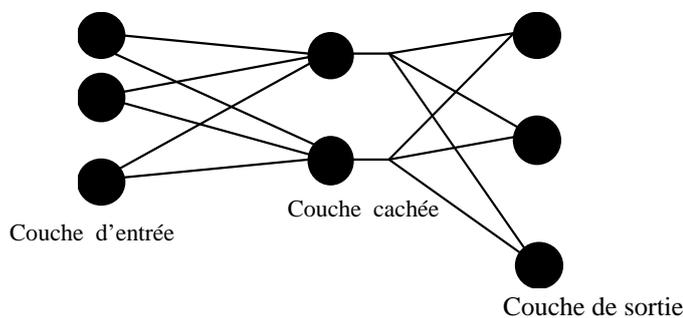


Figure II.02 structure de connexion du Perceptron multicouche

II-7 STRUCTURE D'INTERCONNEXION :

Les connexions entre les neurones qui composent le réseau décrivent la topologie du modèle. Cette topologie peut être quelconque, mais le plus souvent il est possible de distinguer une certaine régularité.

Réseau multicouche : les neurones sont arrangés par couches. Il n'y a pas de connexion entre neurones d'une même couche et les connexions ne se font qu'avec les neurones des couches avales (Figure II.03 a).

Réseau à connexions récurrentes : les connexions récurrentes ramènent l'information en arrière par rapport au sens de propagation défini dans un réseau multicouche. Ces connexions sont le plus souvent locales (Figure II.03 b).

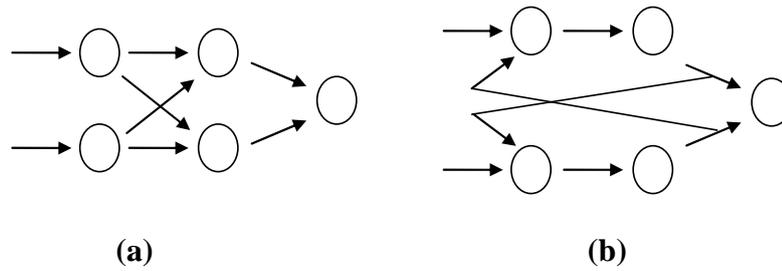


Figure II-03 : Connexion directe (a) et récurrente (b)

II-8 RESEAU RECURRENT D'ELMAN

Elman [Elman 1989, 1990] a proposé en 1990 une architecture de réseau de neurones récurrente pour la tâche de prédiction de phonèmes. Le réseau d'Elman (figure II.04) comprend des connexions récurrentes de neurones cachés vers une couche d'unités contextuelles constituées de délais unitaires. Ces unités contextuelles emmagasinent les sorties des neurones cachés pour un pas de temps donné et en alimentent les neurones d'entrée. Les neurones cachés alimentent aussi la couche de sortie.

Le rôle des unités contextuelles est de fournir au réseau une mémoire dynamique utile pour une tâche de prédiction. Selon Hertz [Hertz 1991] et Haykin [Haykin 1995] et grâce à la présence de rétroaction, les neurones cachés continuent à recycler l'information dans le réseau à travers de multiples pas de temps et peuvent dès lors découvrir des représentations abstraites de l'information temporelle. Ce réseau récurrent ne fait donc pas une copie conforme simpliste des données du passé mais possède la faculté d'effectuer un traitement non-trivial.

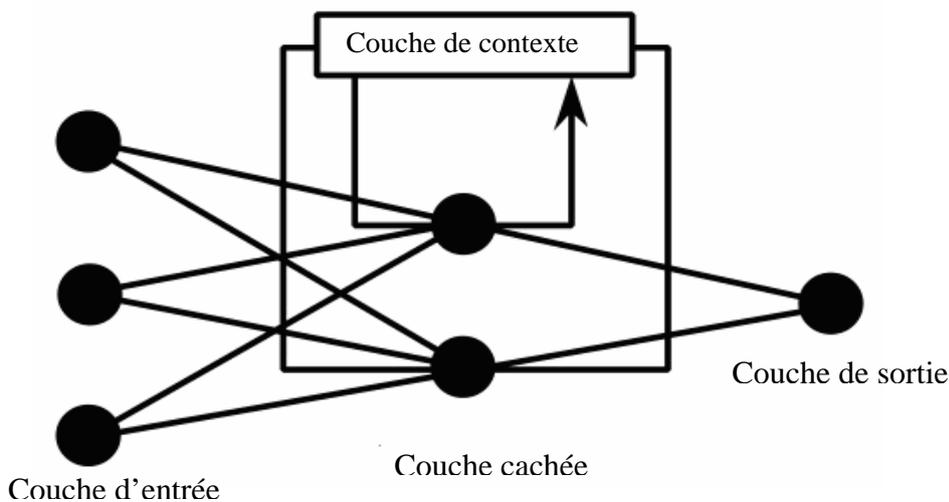


Figure II.04 Structure de connexion du réseau d'Elman

II-9 APPRENTISSAGE :

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. Les variables modifiées pendant l'apprentissage sont les poids des connexions. L'apprentissage est la modification des poids du réseau dans l'optique d'accorder la réponse du réseau aux exemples et à l'expérience. Il est souvent impossible de décider *a priori* des valeurs des poids des connexions d'un réseau pour une application donnée. A l'issue de l'apprentissage, les poids sont fixés : c'est alors la phase d'utilisation.

Une fois la structure du réseau est fixée, il faut passer par le processus d'apprentissage, par lequel les poids synaptiques vont être ajustés de manière à satisfaire un critère d'optimisation. On peut, en effet définir l'apprentissage comme l'acquisition des propriétés associatives stables ou «toute modification du comportement relativement durable, que l'on peut attribuer à l'expérience sensorielle passée de l'organisme» [Gondran 1997], et la mémorisation comme l'enregistrement et le rapport de ces propriétés ou modification.

Historiquement, la première règle d'apprentissage a été formulée par le psychologue Hebb (1949) «quand une cellule **A** excite par son axone une cellule **B** et que, de manière répétée et persistante, elle participe à la genèse d'une impulsion dans **B**, un processus de croissance ou un changement métabolique a lieu dans une ou deux cellules, de telle façon que l'efficacité de **A**, à déclencher une impulsion dans **B** est, parmi les autres cellules qui ont cet effet, accrue» [Gondran 1997].

L'apprentissage est basé sur le même principe que la règle de Widrow–Hoff, on dispose d'un ensemble d'exemples qui sont des couples, entrée-sortie désirée (Figure II.04). A chaque étape, un exemple est présenté en entrée du réseau et une sortie réelle est calculée. Ce calcul est effectué de proche en proche de la couche d'entrée à la couche de sortie, cette phase est appelée propagation avant ou encore, relaxation du réseau. Ensuite, l'erreur (somme quadratique des erreurs sur chaque cellule de sortie) est calculée et rétropropagée dans le réseau, donnant lieu à une modification de chaque poids. Ce processus est répété, en présentant successivement chaque exemple. Si pour tous les exemples, l'erreur quadratique est inférieure à un seuil choisi, cela implique que le réseau a convergé.

Toute la difficulté pour effectuer cette descente dans un réseau multicouche est de pouvoir calculer la dérivée de l'erreur quadratique par rapport à un poids donné. L'utilisation de neurones à fonction d'activation dérivable peut résoudre ce problème simple

Pour la plupart des réseaux, l'apprentissage requiert essentiellement trois caractéristiques :

- 1) Une structure permettant de mettre en œuvre une association ou une fonction «mapping» correspondante à une fonction de transfert.
- 2) Une capacité d'ajustement des poids de connexion entre unités et une loi d'ajustement.
- 3) Une mesure de performance et sa rétroaction.

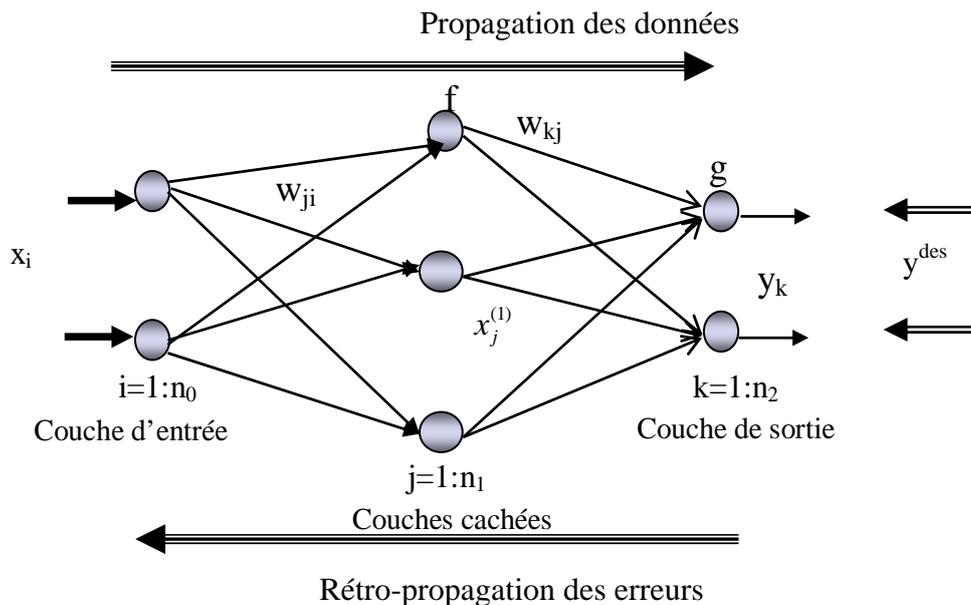


Figure II-05 Schéma de fonctionnement de l'apprentissage

a)-Principe de rétropropagation:

La technique de **rétropropagation du gradient** (*Back-propagation*) est une méthode qui permet de calculer le gradient de l'erreur pour chaque neurone du réseau, de la dernière couche vers la première. De façon abusive, on appelle souvent technique de rétropropagation du gradient l'algorithme classique de correction des erreurs basé sur le calcul du gradient grâce à la rétropropagation que nous allons présenter ici. En vérité, la correction des erreurs peut se faire selon d'autres méthodes, en particulier le calcul de la dérivée seconde. Cette technique modifie de façon itérative les poids pour adapter la réponse obtenue à la réponse désirée. Il s'agit en fait de modifier les poids lorsqu'il y a erreur seulement.

Les poids dans le réseau de neurones sont au préalable initialisés avec des valeurs aléatoires. On considère ensuite un ensemble de données qui vont servir à l'apprentissage. Chaque

échantillon possède ses valeurs cibles qui sont celles que le réseau de neurones doit à terme prédire lorsqu'on lui présente le même échantillon.

Soit $W_1=[w_{ji}]$ les poids de la 1^{ère} couche

Soit $W_2=[w_{kj}]$ les poids de la 2^{nde} couche (figure II-05)

$$a_j^{(1)} = \sum_{i=0}^{n_0} w_{ji} x_i \quad (\text{II-03})$$

La sortie de la couche cachée est donnée par l'expression suivante

$$x_j^{(1)} = f(a_j^{(1)}) \quad (\text{II-04})$$

La minimisation de l'erreur se fait par l'approximation d'une descente de gradient comme dans la règle de Widrow Hoff (règle delta). Toute la difficulté pour effectuer cette descente dans un réseau multi couches était de pouvoir calculer la dérivée de l'erreur quadratique par rapport à un poids donné. Généralement on utilise une fonction d'activation de type sigmoïde telle que $f'(x) = f(x)(1 - f(x))$

La sortie de la couche de sortie est donnée par l'expression suivante

$$a_k^{(2)} = \sum_{j=0}^{n_1} w_{kj} x_j^{(1)} \quad (\text{II-05})$$

$$y_k = g(a_k^{(2)}) \quad (\text{II-06})$$

Cette phase, dite de propagation, permet de calculer la sortie du réseau en fonction de l'entrée.

On présente un exemple $x=[x_1... x_{n0}]$ (avec y^{des} sortie désirée)

– on calcule la sortie correspondante $y = [y_1... y_{n2}]$

$$\text{– erreur : } e_k = y_k^{des} - y_k \quad (\text{II-07})$$

$$\text{– coût associé à l'exemple : } J_{(exemple)} = \frac{1}{2} \sum_{k=1}^{n_2} e_k^2 \quad (\text{II-08})$$

$$\text{– coût global : } J = \frac{1}{2} \sum_{l=1}^n J_{(exemple l)} \quad (\text{II-09})$$

Mise à jour de w_{ji} et w_{kj} selon une règle delta:

$$w(t+1) = w(t) + \Delta w \quad (\text{II-10})$$

$$\text{avec } \Delta w = -\eta \frac{\partial J}{\partial w} \quad \text{et } \eta \text{ facteur d'apprentissage} \quad (\text{II-11})$$

L'algorithme de rétropropagation consiste à effectuer une descente de gradient sur le critère 'J'. On distingue alors deux cas, suivant que le neurone d'indice i est un neurone de sortie ou non. Le gradient de J est calculé pour tous les poids de la manière suivante :

Calcul de $\frac{\partial J}{\partial w_{kj}}$ pour les neurones de la couche de sortie, pour un exemple fixé

$$\frac{\partial J}{\partial w_{kj}} = \frac{\partial J}{\partial y_k} \frac{\partial y_k}{\partial a_k^{(2)}} \frac{\partial a_k^{(2)}}{\partial w_{kj}} = -(y_k^{des} - y_k) g'(a_k^{(2)}) x_j^{(1)} \quad (II-12)$$

Calcul de $\frac{\partial J}{\partial w_{ji}}$ pour les neurones de la couche cachée, pour un exemple fixé

$$\frac{\partial J}{\partial w_{ji}} = \frac{\partial J}{\partial x_j^{(1)}} \frac{\partial x_j^{(1)}}{\partial a_j^{(1)}} \frac{\partial a_j^{(1)}}{\partial w_{ji}} = \left(\sum_{k=0}^{n_2} -(y_k^{des} - y_k) g'(a_k^{(2)}) w_{kj} \right) f'(a_j^{(1)}) x_i \quad (II-13)$$

L'apprentissage peut se voir comme un moyen de synthétiser automatiquement une fonction de transfert. (Normal mapping) en libérant l'homme au moins en principe, du processus interactif de conception sous forme d'essais-erreurs ce sont surtout les facultés d'optimisation des performances et d'adaptations à un environnement incertain qu'il faut souligner [Carlos 2000].

II-10 RESEAUX DE NEURONES A FONCTION DE BASE RADIALE :

Les réseaux de fonctions à base radiale sont utilisés dans la classification, l'approximation et la reconnaissance de parole. Leur but est d'approximer un comportement désiré par une collection de fonctions appelées noyaux. Un noyau est caractérisé par un centre et des champs récepteurs [Seghouane 2003, Angel 2007].

La structure du réseau est constituée d'une couche d'entrée, d'une couche de sortie et d'une couche cachée de n unités caractérisées par une fonction d'activation de type gaussien (Figure II.06)

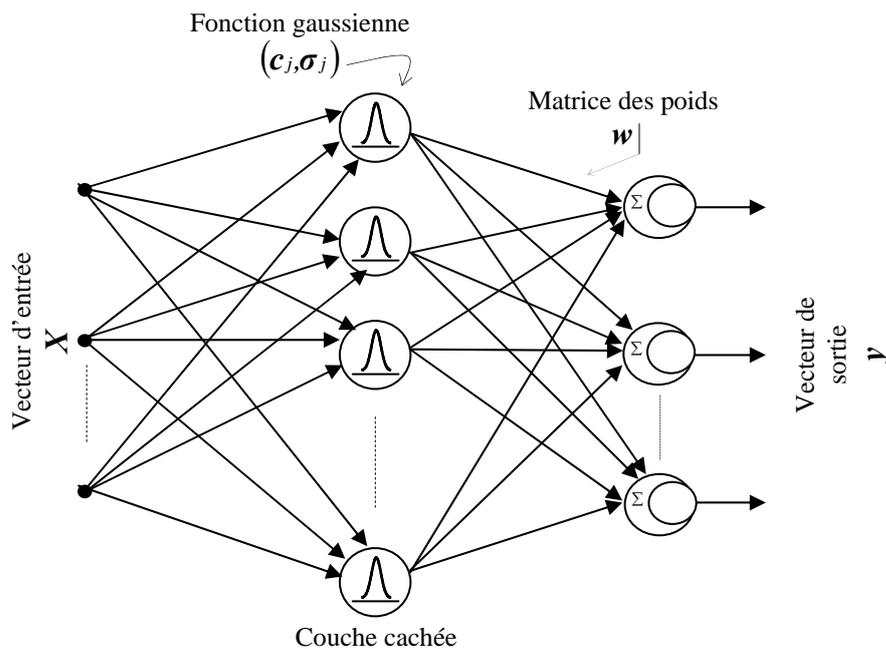


Figure II-06 Réseau de neurones à fonction de base radiale

La valeur à la sortie est donnée par :

$$y_i = \frac{\sum_{j=1}^{n_n} w_{ij} \rho_j(x)}{\sum_{j=1}^{n_n} \rho_j(x)} \quad \left(\sum_{j=1}^{n_n} \rho_j(x) \text{ est le terme de normalisation} \right) \quad (\text{II-14})$$

tel que:

$$\rho_j = \exp\left(-\frac{1}{2}(X - \mathbf{c}_j)^T \mathbf{Q}_j^{-1}(X - \mathbf{c}_j)\right) \quad (\text{II-15a})$$

où c_j représente le vecteur du centre du nœud j , \mathbf{Q}_j est la matrice des rayons gaussiennes, exprimée par l'équation suivante :

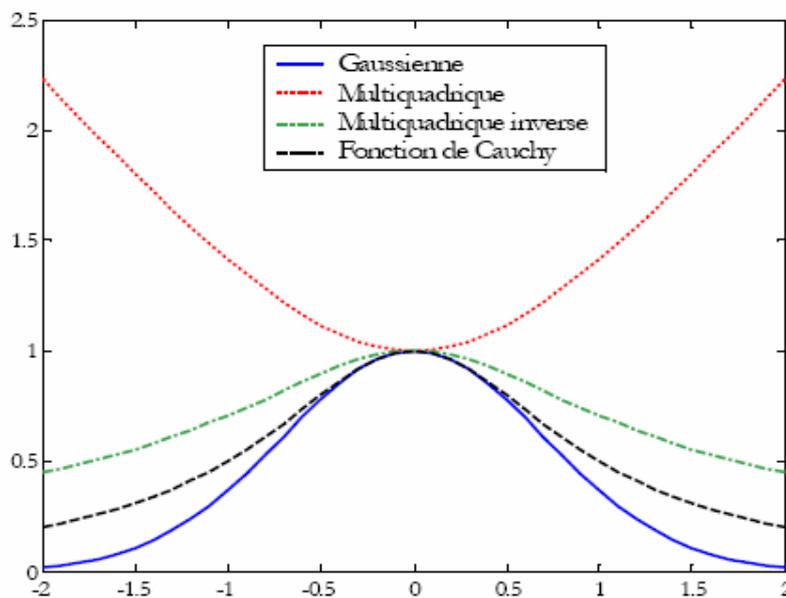


Figure II-07 Quelques fonctions radiales

$$\mathbf{Q}_j = \left(\text{diag}[\sigma_{1,j}, \sigma_{2,j}, \sigma_{3,j}, \dots, \sigma_{n_e,j}] \right)^2 \quad (\text{II-15b})$$

et X est le vecteur d'entrée.

D'après les formules précédentes, on voit que la réponse des noyaux dépend de la distance entre l'entrée et un paramètre interne appelé centre, modulée par un paramètre d'échelle. Ce type de réseaux normalisés possède la propriété où le neurone requiert uniquement des données locales pour le calcul de son activation. Divers types de fonctions peuvent être utilisés comme noyaux ou fonctions de base (Figure II.07) la fonction gaussienne reste cependant la plus utilisée. Des investigations théoriques ont montré qu'un paramètre d'échelle uniforme pour chaque unité cachée est suffisant pour l'approximation non-linéaire [Seghouane 2003, Angel 2007]. Les performances d'un tel réseau en approximation

dépendent donc du nombre de centres et par conséquent du nombre de noyaux, de leurs positions, de la valeur des paramètres d'échelle et de la méthode utilisée pour l'apprentissage de la relation entrée-sortie recherchée.

La plupart des applications utilisant des fonctions à base radiale disposent de centres et de rayons fixés une fois pour toutes. Dans ce cas la sortie du réseau est linéaire en ces paramètres, il est tout à fait concevable d'utiliser les techniques d'optimisation linéaire classique pour obtenir un apprentissage rapide et précis. En outre, la disposition des centres a priori s'avère quelques fois impraticable. En effet, le nombre de centre requis est parfois trop élevé.

Les performances d'un tel réseau en approximation dépendent donc du nombre de centres et par conséquent du nombre de noyaux, de leurs positions, de la valeur des paramètres d'échelle et de la méthode utilisée pour l'apprentissage de la relation entrée-sortie recherchée.

Le recours aux réseaux à base radiale modifiés appelés aussi réseaux à base radiale généralisés s'avère nécessaire. Ils permettent non seulement l'apprentissage des poids mais aussi celui des centres et des rayons.

Par conséquent, la sortie n'est plus linéaire en ces paramètres, cette caractéristique rend l'utilisation des algorithmes d'optimisation linéaire inutile. L'ajustement effectué par une descente de gradient sur les poids, les positions des centres ainsi que les valeurs des rayons associés, risque de tomber dans des minima locaux. L'introduction des connaissances a priori sur la distribution initiale des centres, des rayons et des poids permet de disposer de valeurs très éloignées de celles des valeurs optimales.

En utilisant un réseau à base radiale normalisé et en se référant aux équations (II-14), et (II-15), les valeurs des gradients s'expriment par les équations suivantes :

$$\frac{\partial J}{\partial w_{ij}} = (y_i - \hat{y}_i) \frac{\rho_j(x)}{R(x)} \quad (\text{II-16})$$

$$\frac{\partial J}{\partial c_{j,k}} = \frac{\partial J}{\partial \rho_j} \frac{\partial \rho_j}{\partial c_{j,k}} \quad (\text{II-17})$$

$$\frac{\partial J}{\partial \sigma_{j,k}} = \frac{\partial J}{\partial \rho_j} \frac{\partial \rho_j}{\partial \sigma_{j,k}}$$

$$\frac{\partial J}{\partial \rho_j} = (y_i - y_{id}) \left[\frac{w_{ij} R(x) - \sum_{i=1}^n w_i \rho_i(x)}{R(x)^2} \right] \quad (\text{II-18})$$

$$\frac{\partial \rho_j}{\partial c_{j,k}} = \rho_j \frac{x_k - c_{j,k}}{\sigma_{j,k}^2} \quad (\text{II-19})$$

$$\frac{\partial \rho_j}{\partial \sigma_{j,k}} = \rho_j \frac{(x_k - c_{j,k})^2}{\sigma_{j,k}^3} \quad (\text{II-20})$$

Les corrections s'obtiennent par :

$$\Delta w_{ij} = -\eta_w \cdot \frac{\partial J}{\partial w_{ij}} \quad (\text{II-21})$$

$$\Delta c_{j,k} = -\eta_c \cdot \frac{\partial J}{\partial c_{j,k}} \quad (\text{II-22})$$

$$\Delta \sigma_{j,k} = -\eta_\sigma \cdot \frac{\partial J}{\partial \sigma_{j,k}} \quad (\text{II-23})$$

On constate qu'il y a une augmentation du nombre de paramètres à ajuster à l'ordre $(1 + 2n_e)n_{cc}$ au lieu de n_{cc} valeurs dans le cas précédent (où n_e est le nombre d'entrée et n_{cc} le nombre de neurones dans la couche cachée).

II-11 CONCLUSION :

Dans ce chapitre, nous avons, en premier lieu, procédé à une présentation des propriétés des nouvelles méthodes du traitement de l'information, qui sont classées sous le vocable de système de traitement parallèle et distribué. Les réseaux de neurones à partir des définitions et de la propriété fondamentale d'approximation universelle parcimonieuse à partir du moment où cette propriété est établie, la nature essentiellement non-linéaire des réseaux de neurones apparaît très clairement. Bien que ce ne fût pas là l'essentiel de notre propos, nous avons tenu à exposer brièvement le principe de fonctionnement des réseaux de neurones. **Nous avons également** insisté sur le fait que les réseaux de neurones rentrent dans la catégorie des meilleurs approximateurs.

L'objectif essentiel de ce chapitre était de montrer que les réseaux de neurones reposent à présent sur des bases mathématiques qui permettent d'envisager des applications très diversifiées en géophysique, notamment dans le domaine de l'inversion des données sismiques. **Sachant que** la résolution du problème inverse est toujours difficile et nécessitant encore très longtemps beaucoup de travail et un éventail étendu de connaissances en optimisation, traitement du signal, etc., il n'est pas douteux que les réseaux de neurones peuvent alléger considérablement la tâche des ingénieurs en permettant une approche efficace et générique des problèmes non linéaires.

CHAPITRE III

La Tomographie sismique :
Application et résultats



III-1 APPLICATION ET RESULTATS DE LA TOMOGRAPHIE

Dans cette phase de simulation, nous décrivons l'application des différents algorithmes de reconstruction tomographique. Ces algorithmes permettent d'estimer des répartitions de vitesses à partir des vecteurs de temps de parcours calculés sur des modèles géologiques synthétiques des milieux non homogènes. Ces modèles sont proposés pour tester l'efficacité des différents algorithmes. Nous comparons les résultats des méthodes classiques ART, SIRT, GC avec des méthodes proposées (RNA-Elman et RNA-FBR) à partir de données synthétiques.

La figure III-01 montre la couverture et le design de l'acquisition des données de sismique entre puits. Cette couverture est la même pour tous les modèles proposés. Les puits sont espacés de 1000 m et font 1000 m de profondeur. La modélisation des temps d'arrivée a été effectuée avec un algorithme de rais rectiligne. La discrétisation de la grille de modélisation est de 20 m 20 m. Les émetteurs (Sx) sont situés dans le puits de droite et les récepteurs (Rx) dans le puits de gauche. Les récepteurs et les émetteurs sont espacés de 20 m et sont situés entre 10 m et 990 m de profondeur. La modélisation des temps de parcours est effectuée pour toutes les combinaisons d'émetteur-récepteurs, sur les 2401 combinaisons possibles.

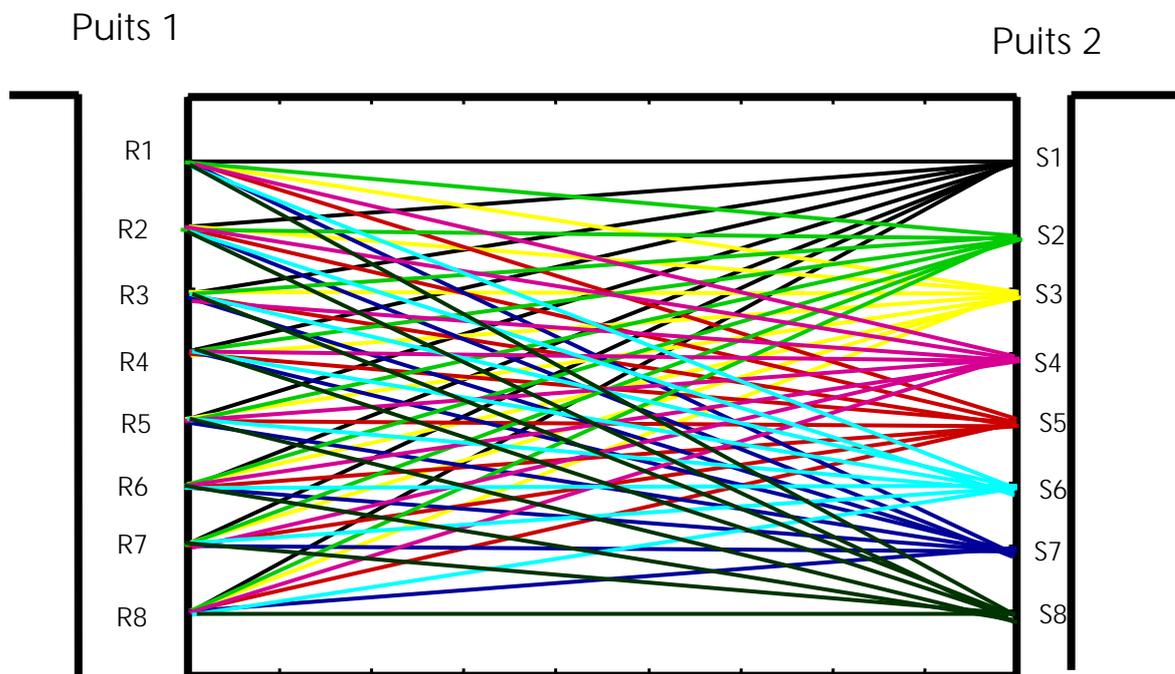


Figure III-01 Couverture et design de l'acquisition des données de sismique entre puits

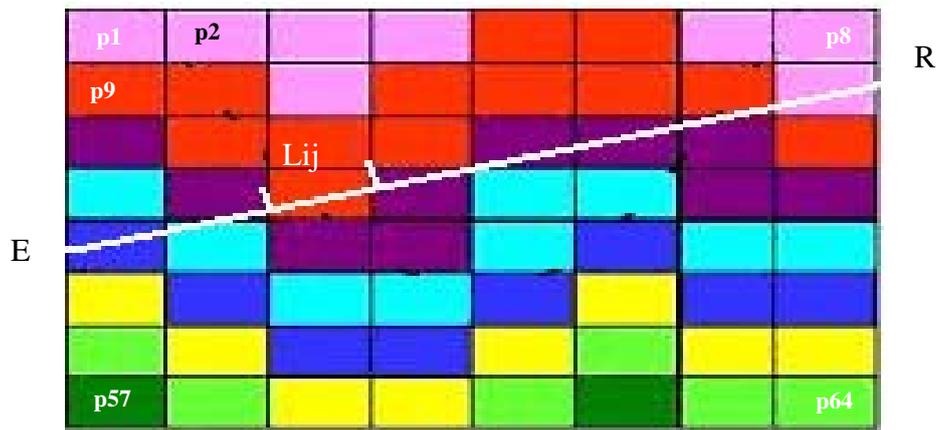


Figure III-02 Exemple de paramétrisation, tracé du rayon

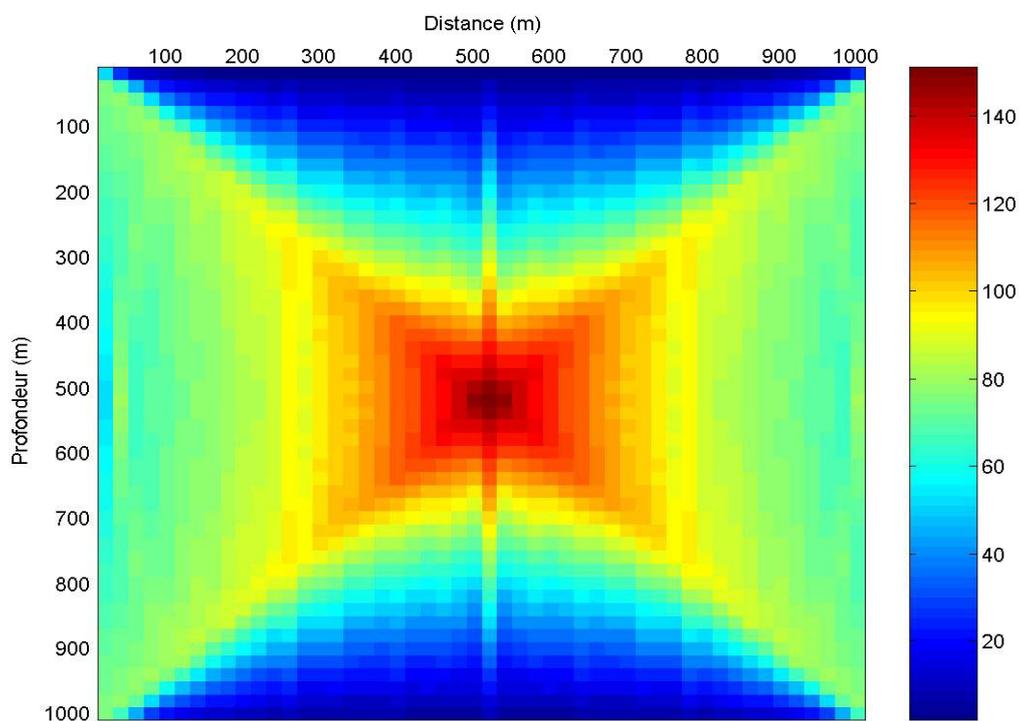


Figure III-03 Densité des rayons

La grille de tomographie est discrétisée en 2401 cellules carrées de 20 m de côté (49 le long de la verticale (Z) et 49 horizontalement (X)) (figure III-02). La figure III-03 montre la fonction de sensibilité qui représente le nombre de rayons par cellule. Dans la partie centrale du modèle réside une forte concentration (140 rayons par cellule) par rapport à la partie la plus profonde et superficielle où cette distribution est relativement faible (10 à 20 rayons par cellule). **Ce qui** signifie que les cellules constituant cette zone de centre ont une bonne distribution des rayons.

De plus, dans tout le processus de tomographie simulé, le calcul du temps des trajets est le plus coûteux en terme de temps calcul. Pour chaque itération, nous avons calculé la norme L1 de chacune des simulations. Le temps résiduel est calculé en faisant la différence entre les temps réels et ceux prédits pour chaque couple émetteur-récepteur. Ce résidu sera utilisé dans le processus d'inversion.

Avant de chercher à tester l'efficacité de différents algorithmes de reconstruction tomographiques, il est naturel de se demander si le problème de la reconstruction à partir du temps de trajet est bien ou mal posé.

Le tableau ci-après, issu de nos calculs, montre clairement que notre problème est mal posé pour les deux raisons suivantes :

- le rang de la matrice de modélisation est inférieur au nombre de lignes et colonnes de la matrice de modélisation, cela indique que la solution peut ne pas exister comme il peu y avoir une infinité de solutions.
- Le nombre de conditionnement est très grand par rapport à 1, cela indique une très grande sensibilité à l'erreur sur les données, et prouve que le problème de reconstruction est mal conditionné.

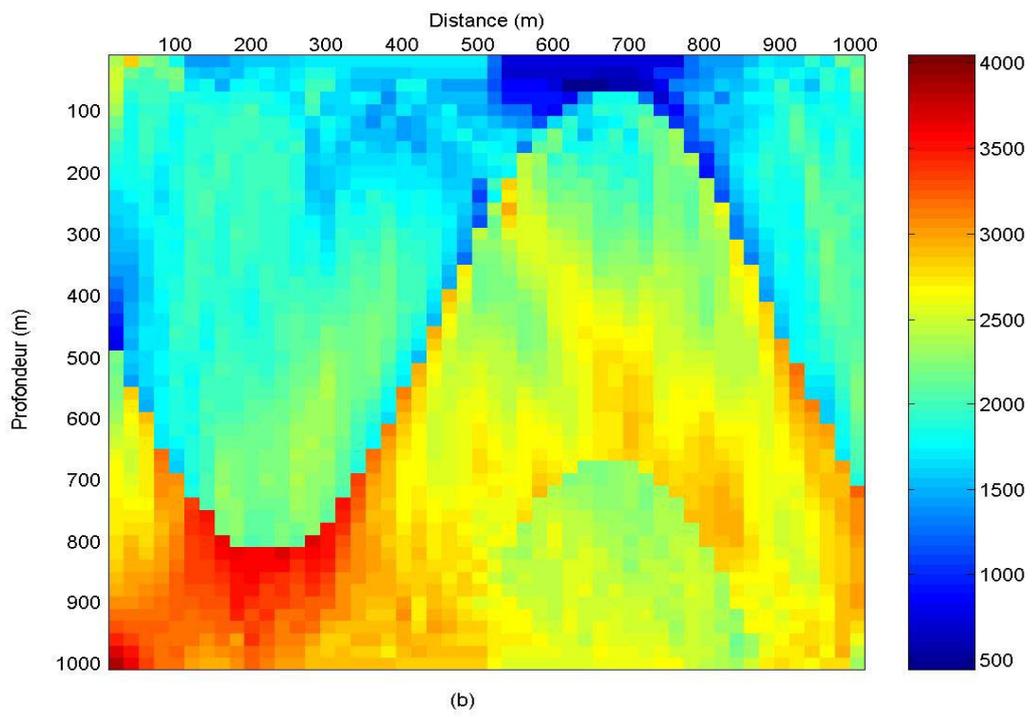
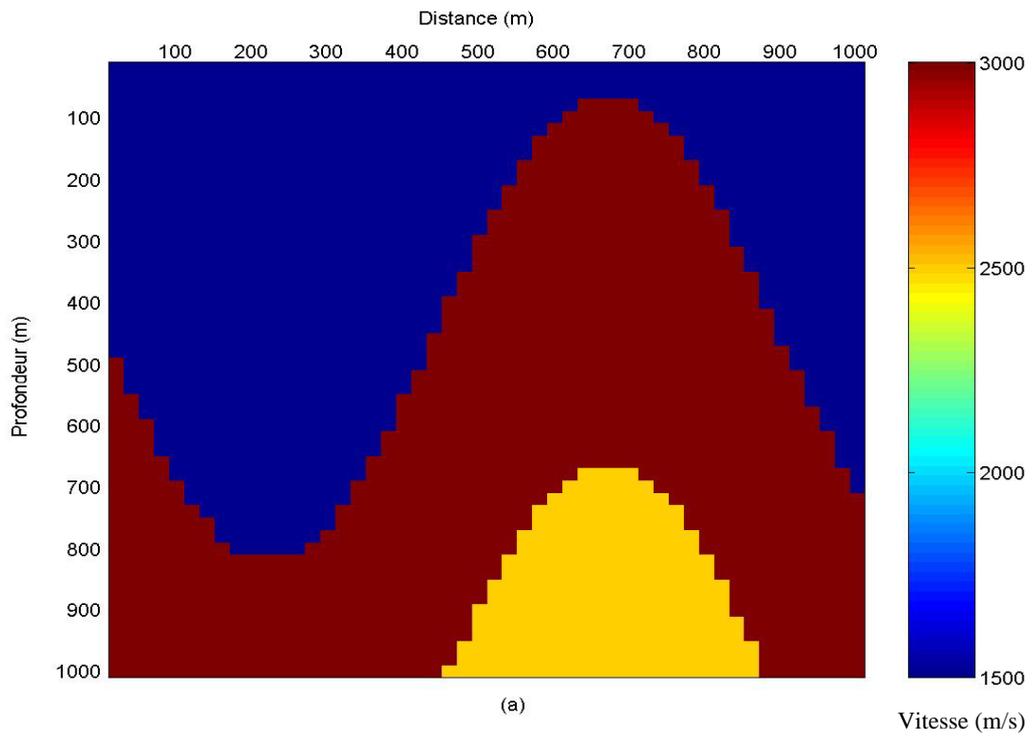
Nombre d'observations	Nombre de paramètres	Rang de la matrice de modélisation	Nombre du conditionnement
2401	2401	2135	6.411 e+19

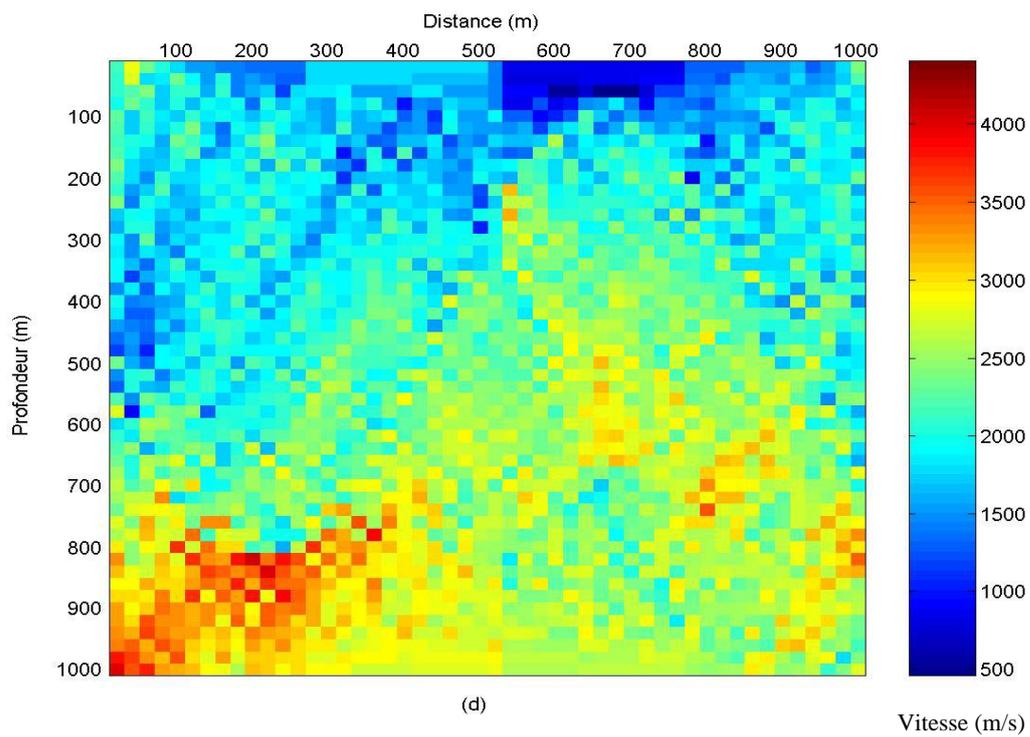
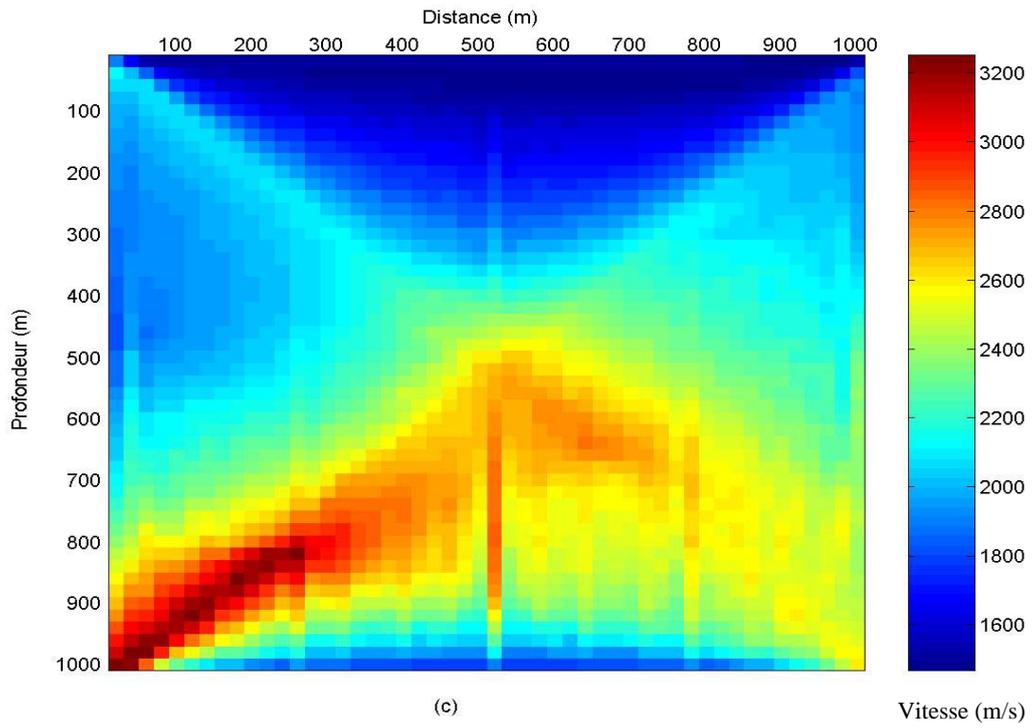
Tableau 01 Récapitulatif de la paramétrisation, modélisation

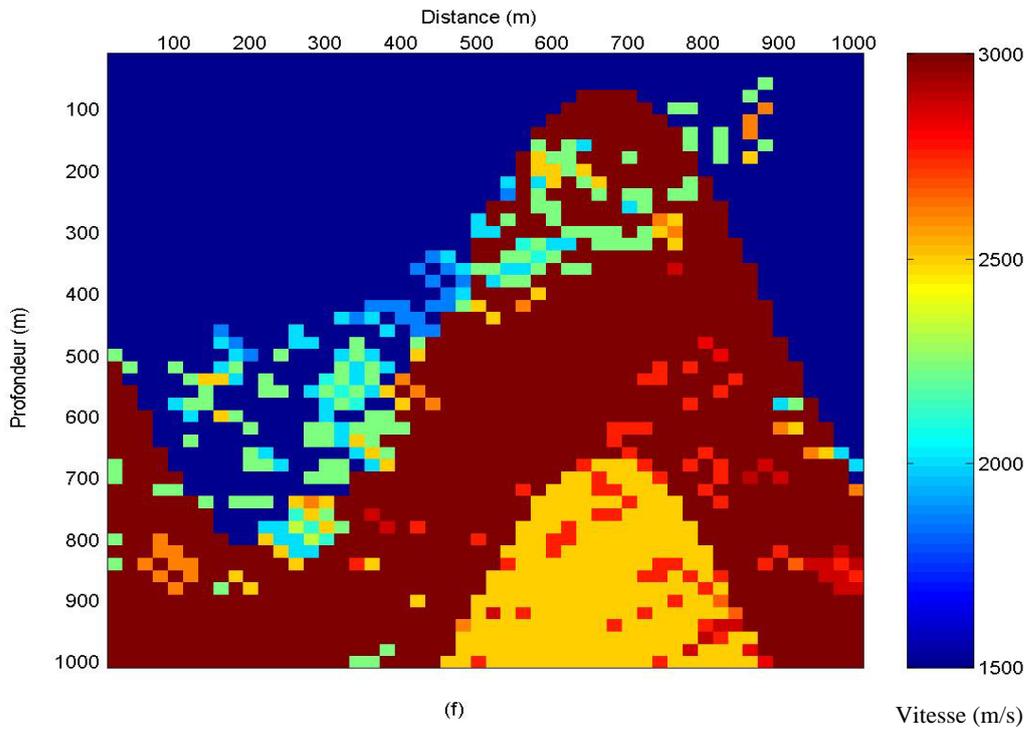
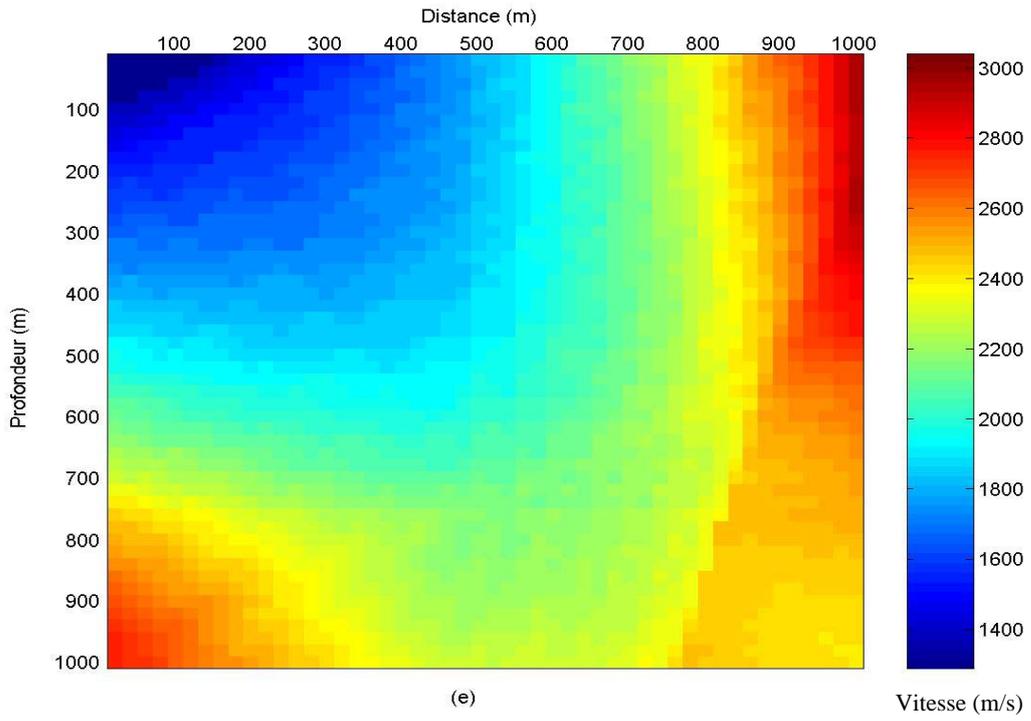
Mais il est important de comprendre que ce mauvais conditionnement est lui-même la conséquence d'une difficulté fondamentale de la résolution des problèmes inverses, difficulté déjà discutée dans le premier chapitre.

III-2-a Modèle 1

La figure III-04-a présente le premier modèle proposé pour cette simulation. Ce modèle consiste en une anomalie de vitesse négative, d'une forme anticlinale entourée par deux couches différentes. La vitesse de l'encaissant est 1500 m.s^{-1} et 2500 m.s^{-1} , tandis que la vitesse à l'intérieur de la couche anomale est de 3000 m.s^{-1} .







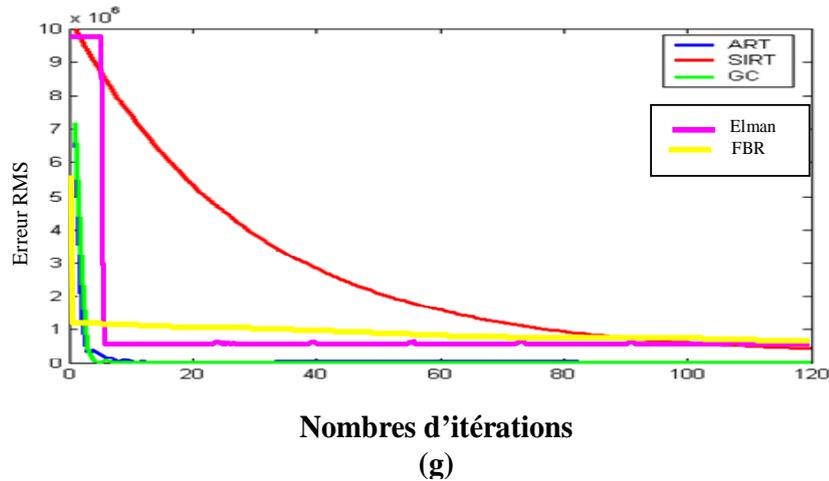


Figure III-04: a- modèle proposé, les résultats de tomographie des méthodes b-GC, c-SIRT, d-ART. e-RNA-Elman, f- RNA-FBR, g- Courbes des résidus des différentes méthodes.

Pour toutes les méthodes itératives (GC, SIRT, ART), nous avons initialisé le processus de reconstruction par une distribution uniforme d'une valeur de vitesse constante de 1500m.s^{-1} . La figure III-04-b représente le résultat de la simulation de la tomographie par l'algorithme du GC, alors que les figures III-04-c et III-04-d présentent les résultats de la méthode SIRT et la méthode ART, respectivement. Nous remarquons que ces deux dernières génèrent de nombreux artefacts sur l'image reconstruite. Ceci est inhérent à ce genre de méthode si le nombre de rais n'est pas très supérieur au nombre de cellules et si les trajets ne sont pas très proches des trajets réels. La méthode du SIRT est beaucoup plus lourde en temps calcul que la GC et l'ART. Les méthodes GC et ART retrouvent bien les grandes lignes du modèle synthétique, cependant, la résolution spatiale semble meilleure avec la méthode du gradient conjugué. En revanche, il apparaît clairement que la méthode SIRT déforme la structure du modèle.

La figure III-05-e représente le résultat de la reconstruction par le RNA de type Elman. On remarque une très grande difficulté d'apprentissage de la structure neuronale proposée puisque l'image obtenue diffère entièrement de l'image à retrouver. La deuxième structure proposée (FBR) a donné le résultat illustré à la figure III-04-f. Le comportement de cette structure s'est traduit par un apprentissage simple et donne une nette amélioration par rapport à toutes les méthodes préalablement examinées, malgré les quelques cellules mal reconstruites. La forme et la position de l'anomalie sont bien retrouvée par le réseau de type FBR.

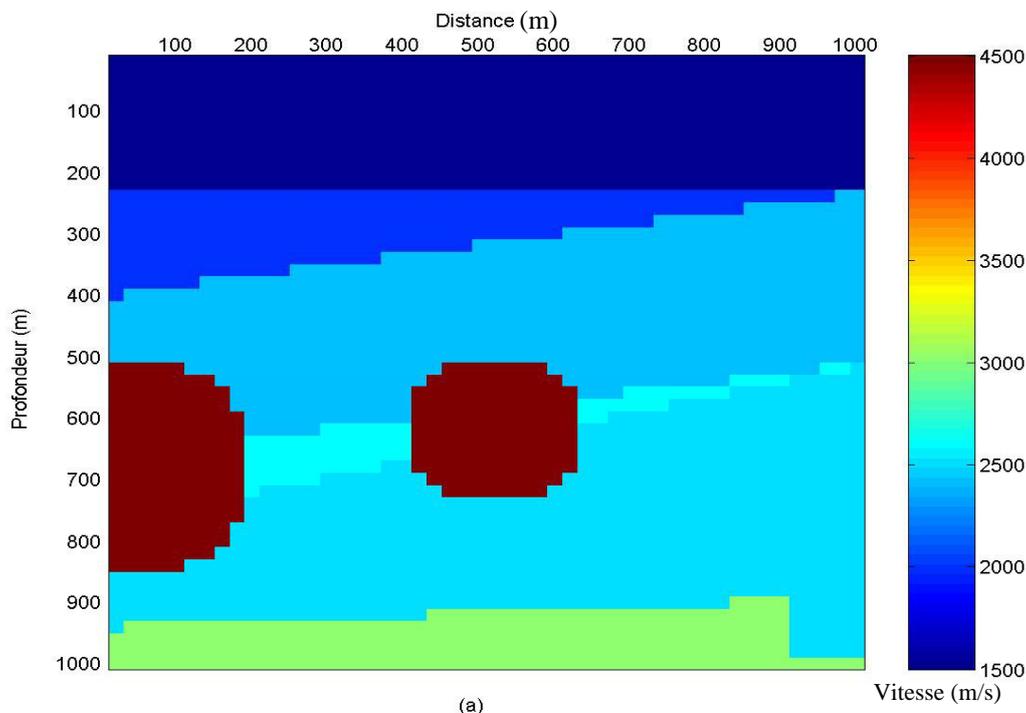
Afin de tester la bonne convergence de la solution, il est indispensable de suivre l'évolution de l'erreur quadratique moyenne en fonction du nombre d'itérations. Dans le cas de la méthode CG, le *RMS* décrit une courbe strictement décroissante telle que montré à la figure (III-04-g) Afin de quantifier la qualité de chaque image obtenue par toutes ces méthodes, nous avons calculé le coefficient de corrélation entre l'image exacte à retrouver et l'image obtenue par chacune des méthodes. Les coefficients résultants sont indiqués dans le tableau 01 qui contient également une estimation du temps de calcul réalisé à l'aide d'un processeur 1.47 GHz Cet indicateur de performance laisse bien prévaloir la méthode RNA-FBR par rapport à toutes les autres méthodes utilisées.

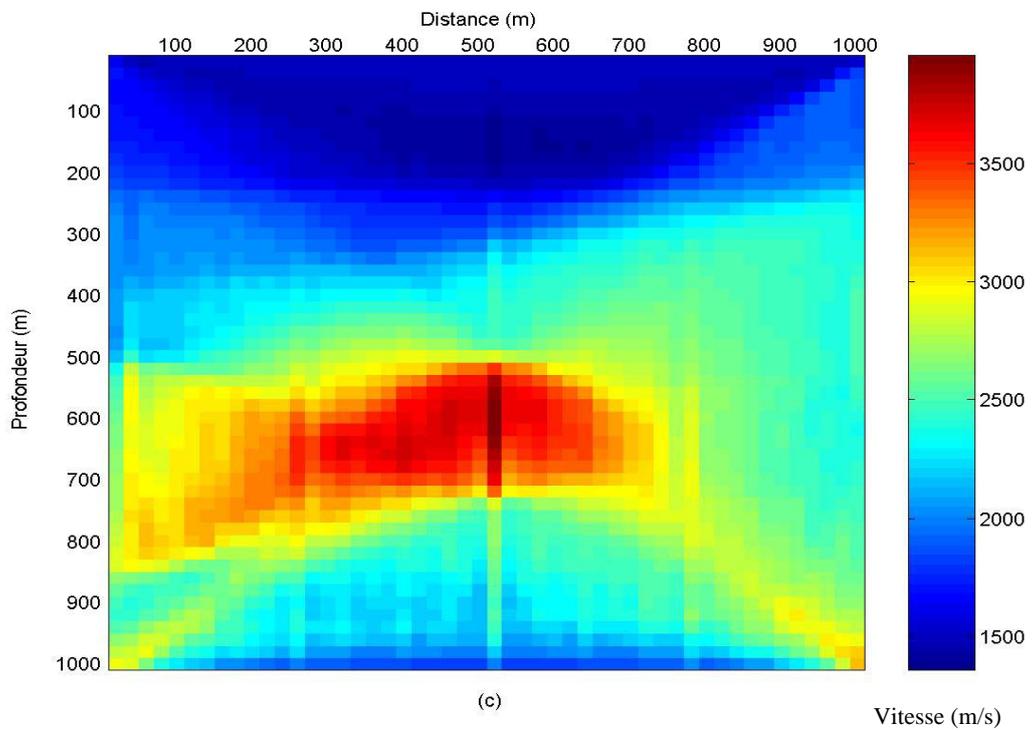
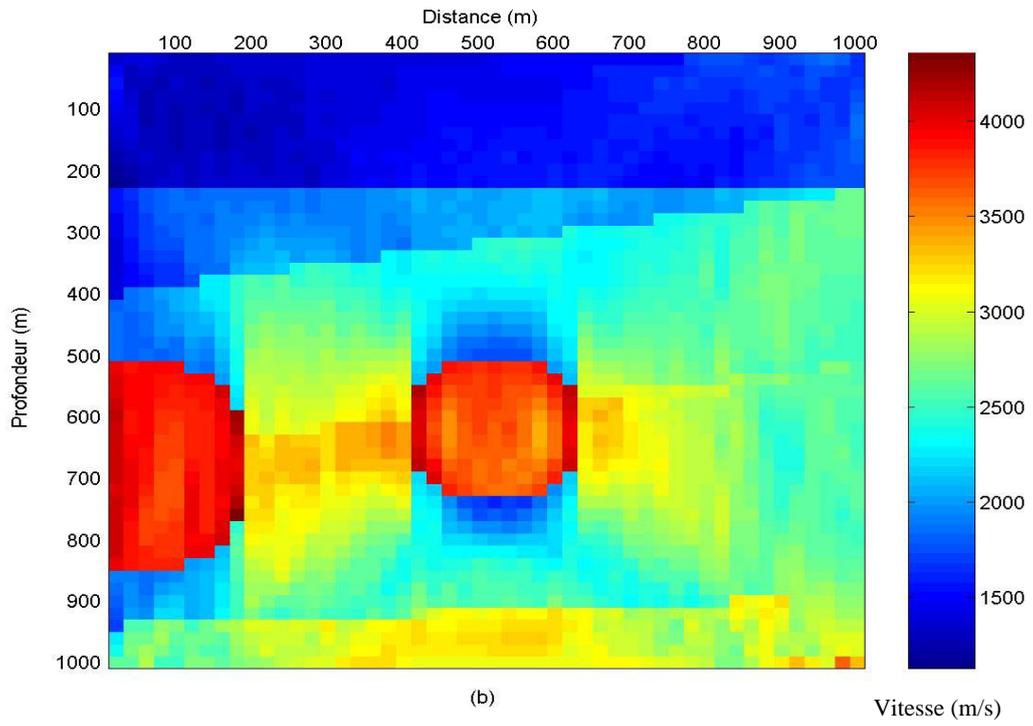
Méthode	RNA –FBR	RNA-Elman	GC	ART	SIRT
Coefficient de corrélation	0.9473	0.3487	0.7850	0.6573	0.5446
Temps de calcul (secondes)	86	30	58	129	209

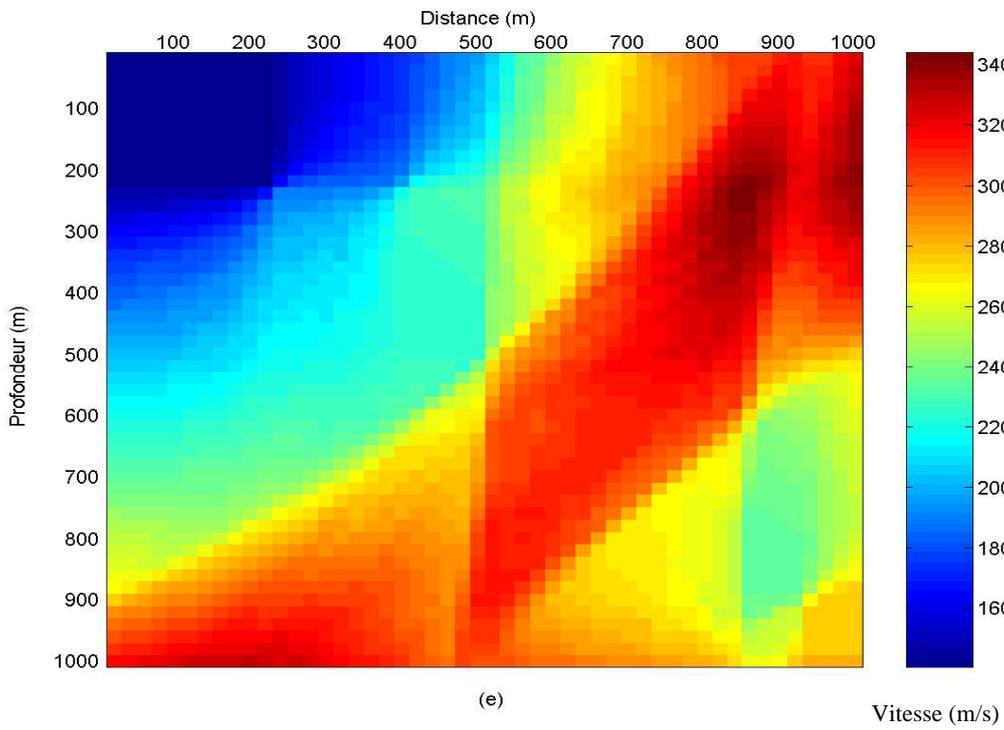
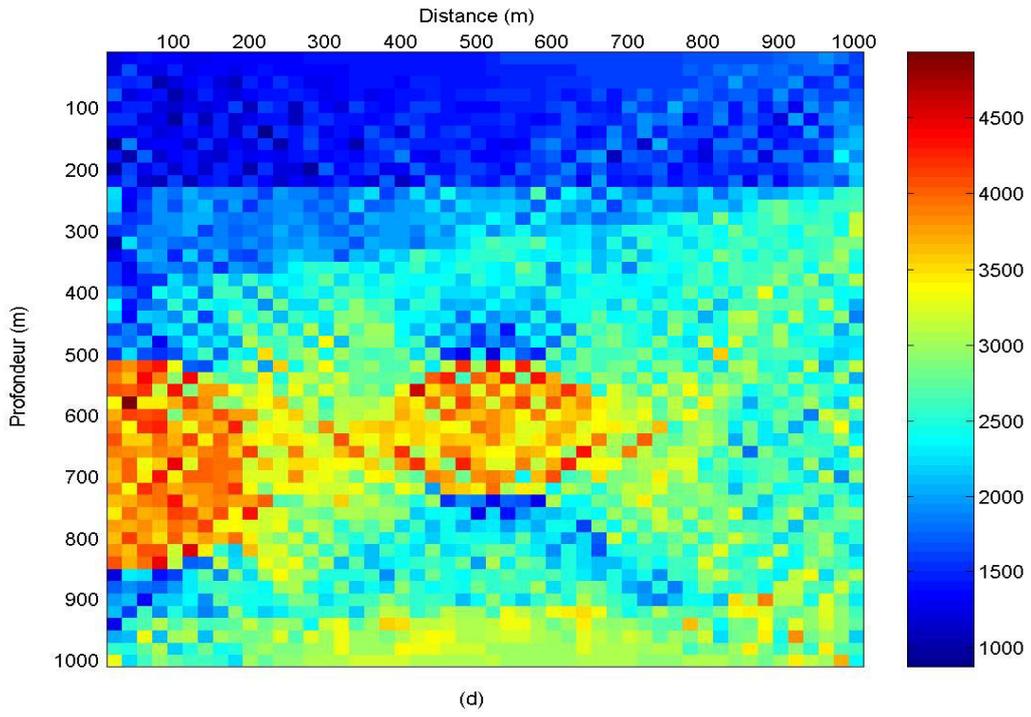
Tableau 02 Récapitulatif des caractéristiques de reconstruction sur le modèle 01

III-2-b Modèle 2

La figure III-05-a illustre le deuxième modèle proposé pour la simulation. Ce modèle consiste en deux anomalies de vitesses négatives, d'une forme sphérique et d'épaisseurs variables. La vitesse de l'encaissant est de 1500 m.s^{-1} et la vitesse à l'intérieur de la couche anormale est de 4500 m.s^{-1} .







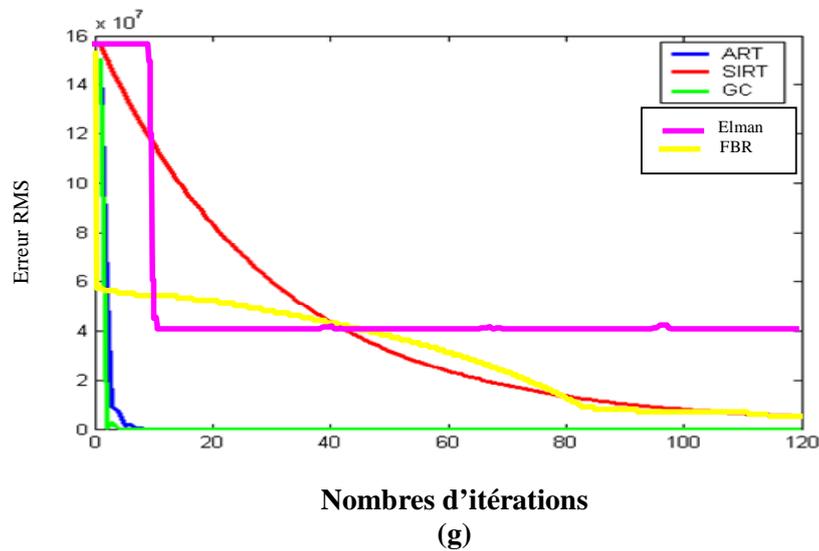
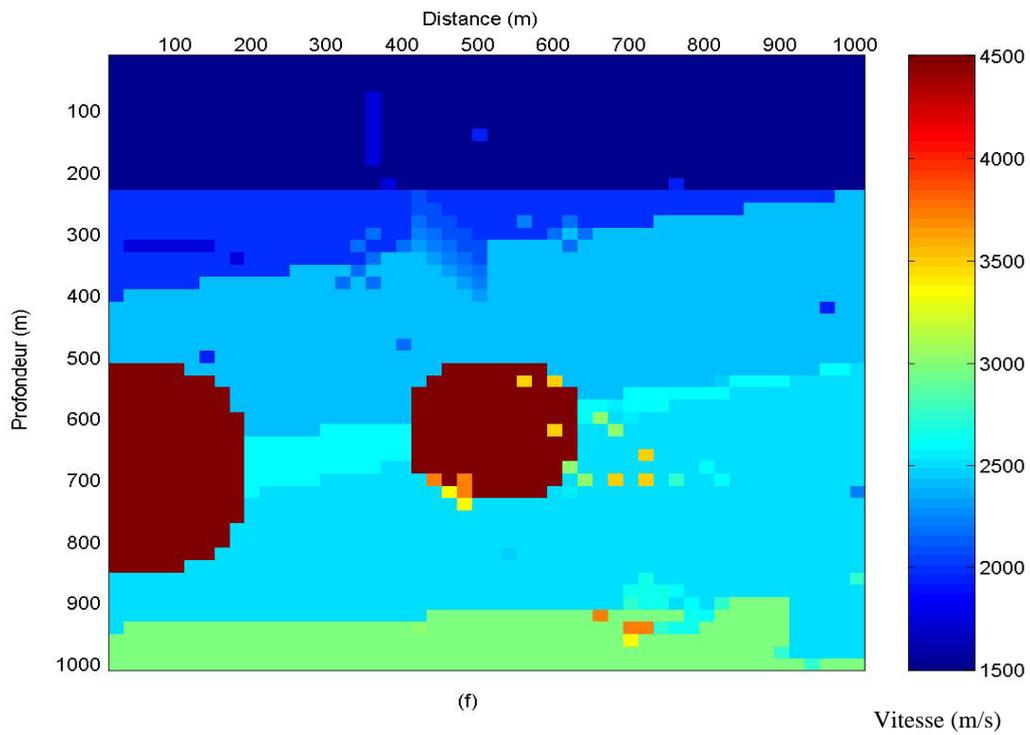


Figure III-05: a- modèle proposé, les résultats de tomographie des méthodes b- GC, c- SIRT, d- ART, e- RNA-Elman, f- RNA-FBR, g- Courbes des résidus des différentes méthodes.

En ce qui concerne les anomalies des couches fortement inclinées, nous remarquons que les deux méthodes (ART et SIRT) donnent des résultats fortement altérés ; les couches minces

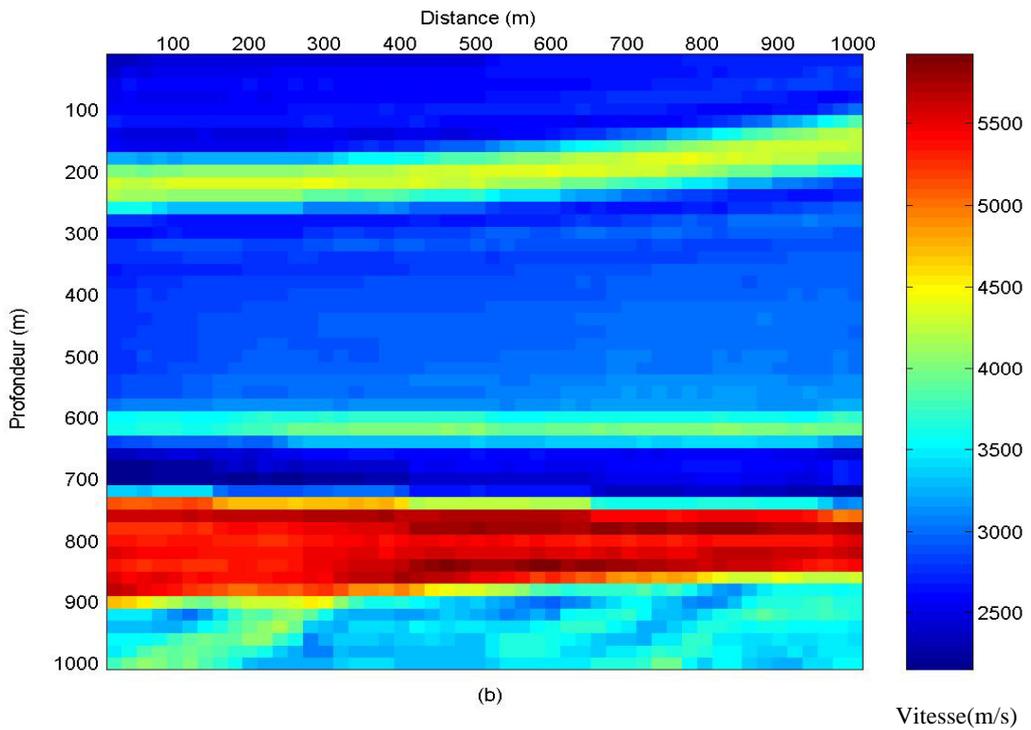
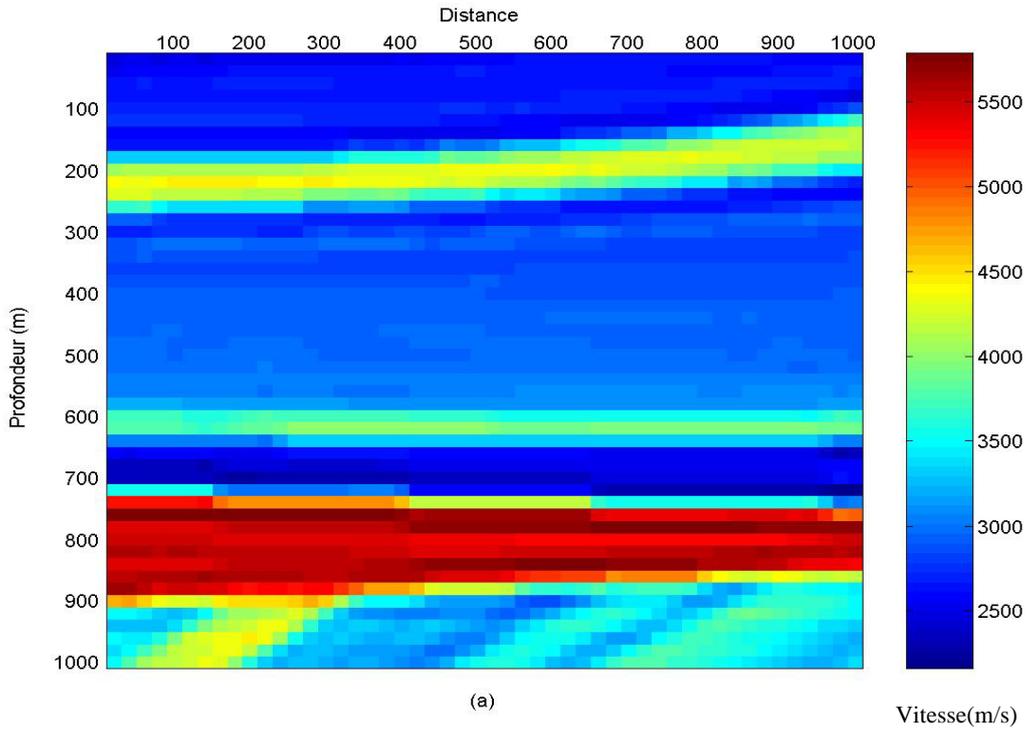
sont indétectables par ces algorithmes. La raison est que nous approchons de la limite de ces méthodes. Tandis que l’algorithme du GC donne un résultat plus ou moins satisfaisant. Si nous examinons ce modèle synthétique, nous remarquons qu’il y a deux anomalies de forte vitesse de forme sphériques. La méthode GC localise ces anomalies en bas à gauche alors que les ART et SIRT les déforment. La figure III-05-e représente le résultat de la reconstruction par RNA de type Elman, une très grande difficulté d’apprentissage de la structure proposée est rencontrée, et l’image obtenue est complètement différente de l’image à retrouver, tandis que la deuxième structure proposée de RNA-FBR figure (III-05-f), a présenté un apprentissage simple, et montre une net amélioration par rapport à toutes les méthodes précédentes. Elle permet, avec une précision, de bien localiser la position de cette anomalie. On voit que la FBR identifie clairement les deux anomalies en terme de forme et de position.

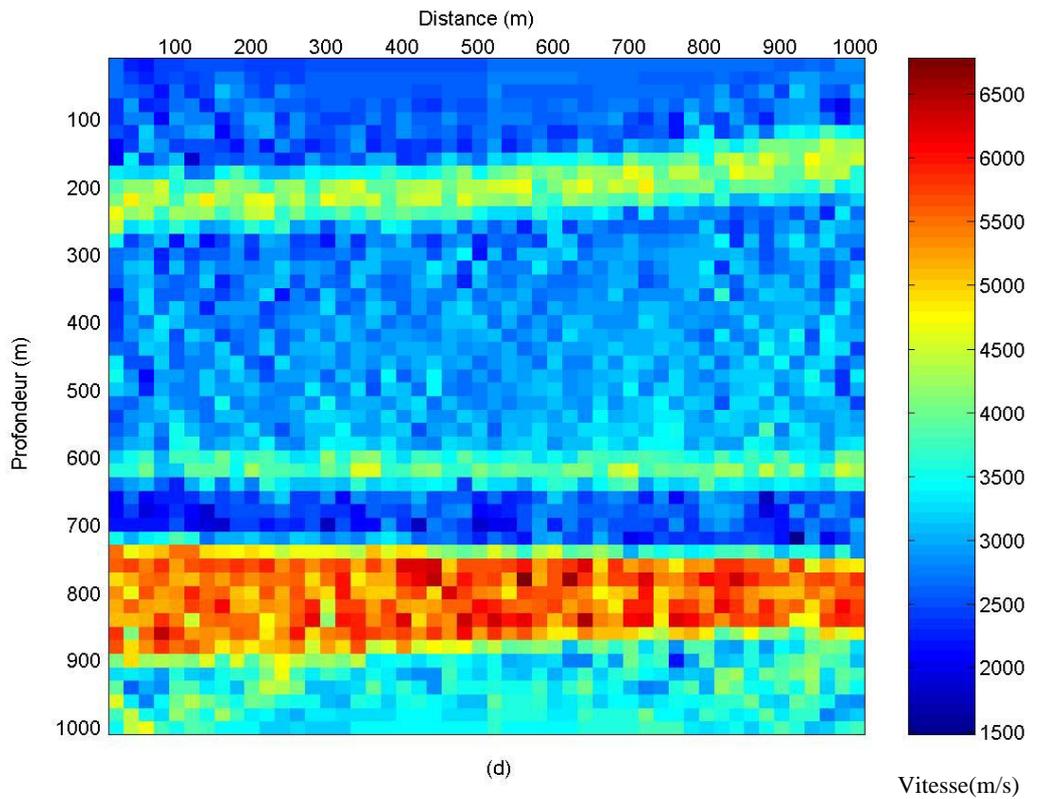
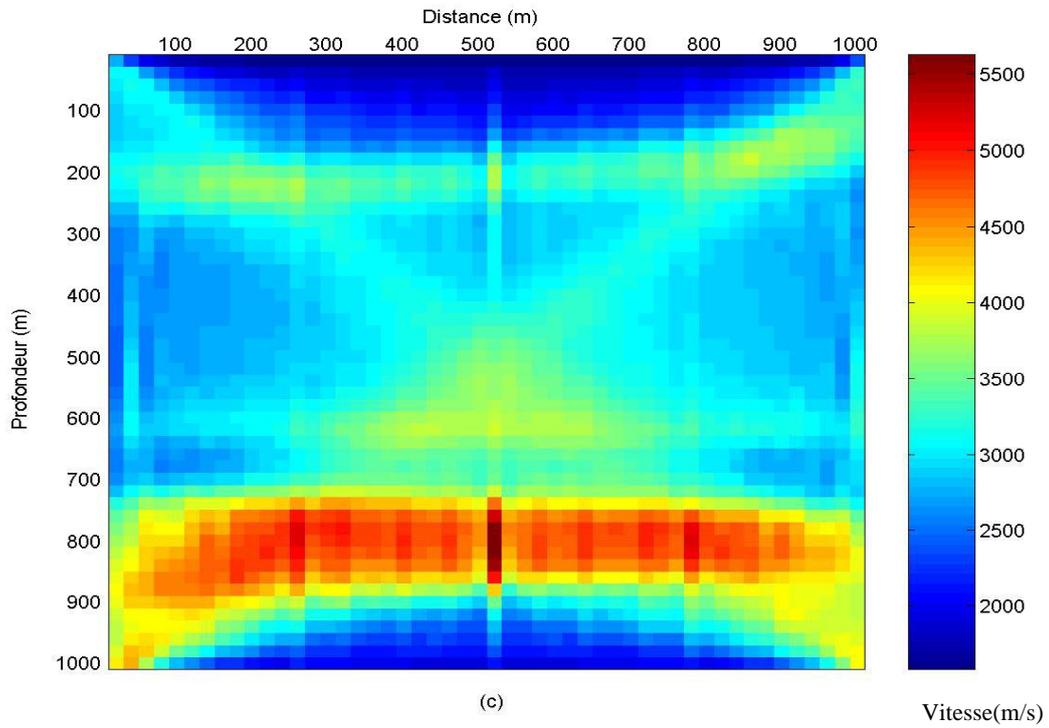
Méthode	RNA -FBR	RNA-Elman	GC	ART	SIRT
Coefficient de corrélation	0.9944	0.2340	0.89864	0.82684	0.7186
Temps de calcul (secondes)	86	30	58	129	209

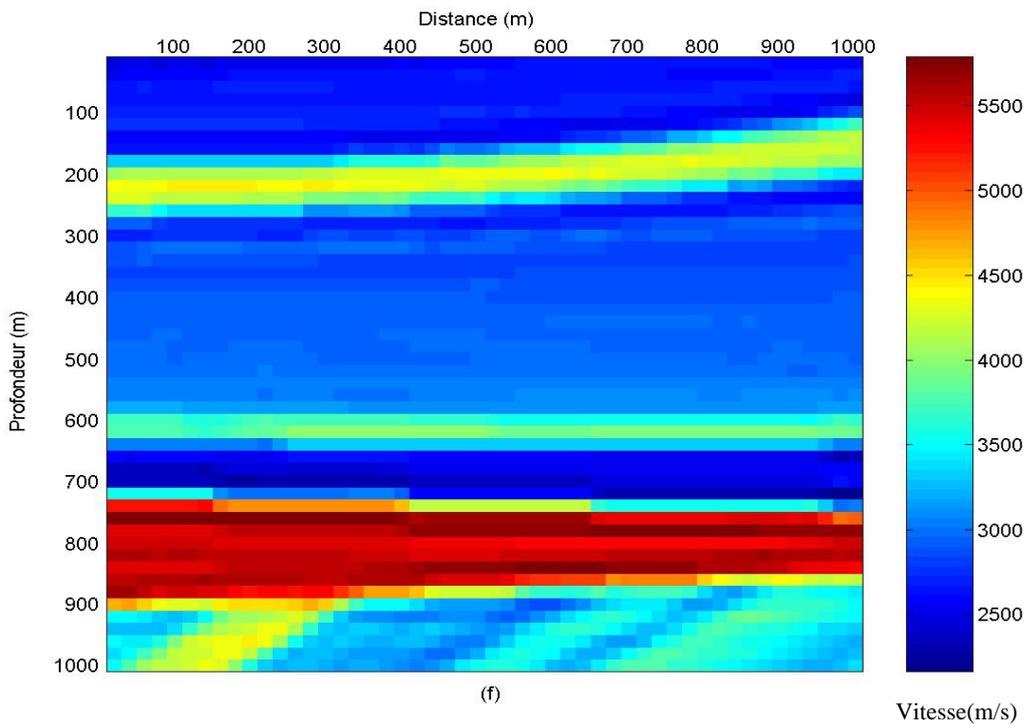
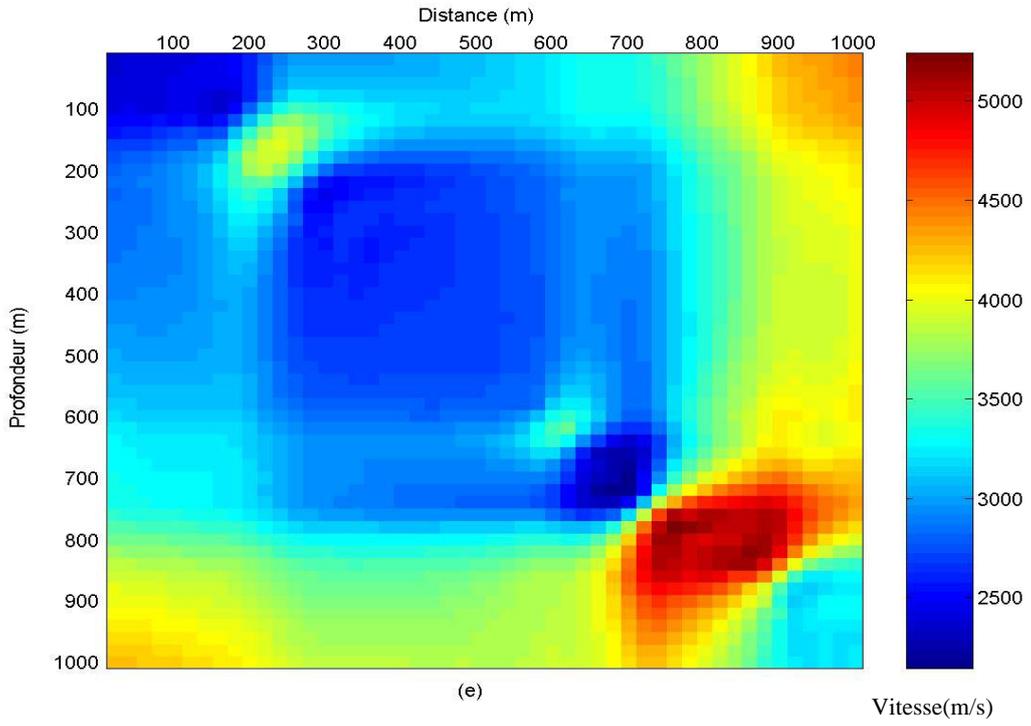
Tableau 03. Récapitulatif des caractéristiques de reconstruction sur le modèle 02

III-2-c Modèle 3

La figure III-06-a illustre le troisième modèle stochastique proposé. Ce modèle est constitué de plusieurs couches de formes et de vitesses variables. Ce modèle représente une complexité relativement importante en comparaison avec les modèles précédents. La figure III-06-b montre les résultats de la méthode GC, alors que les figures III-06-c et III-06-d représente les résultats du SIRT et de l’ART respectivement, tandis que les figures III-06-e et III-06-f représentent les résultats du RNA Elman et du RNA-FBR.







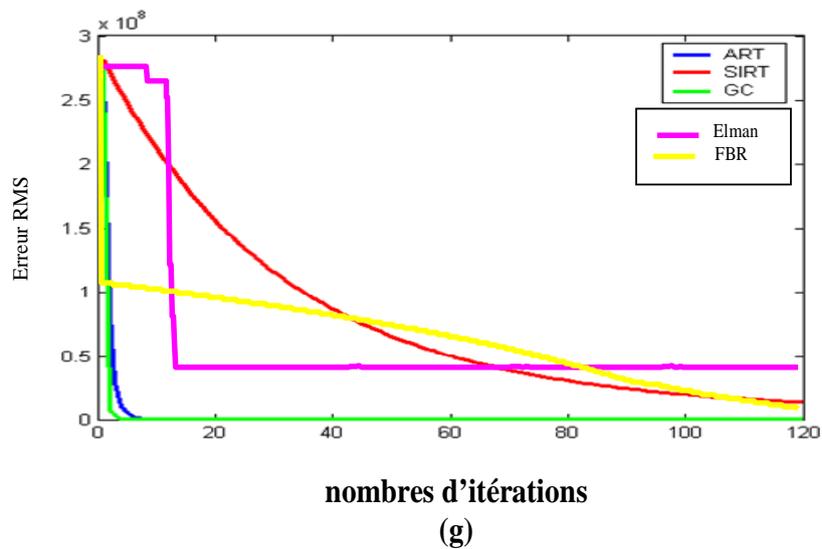


Figure III-06. : a- modèle proposé, les résultats de tomographie des méthodes b- GC, c-SIRT, d- ART. e- RNA-Elman, f- RNA-FBR, g- Courbes des résidus des différentes méthodes.

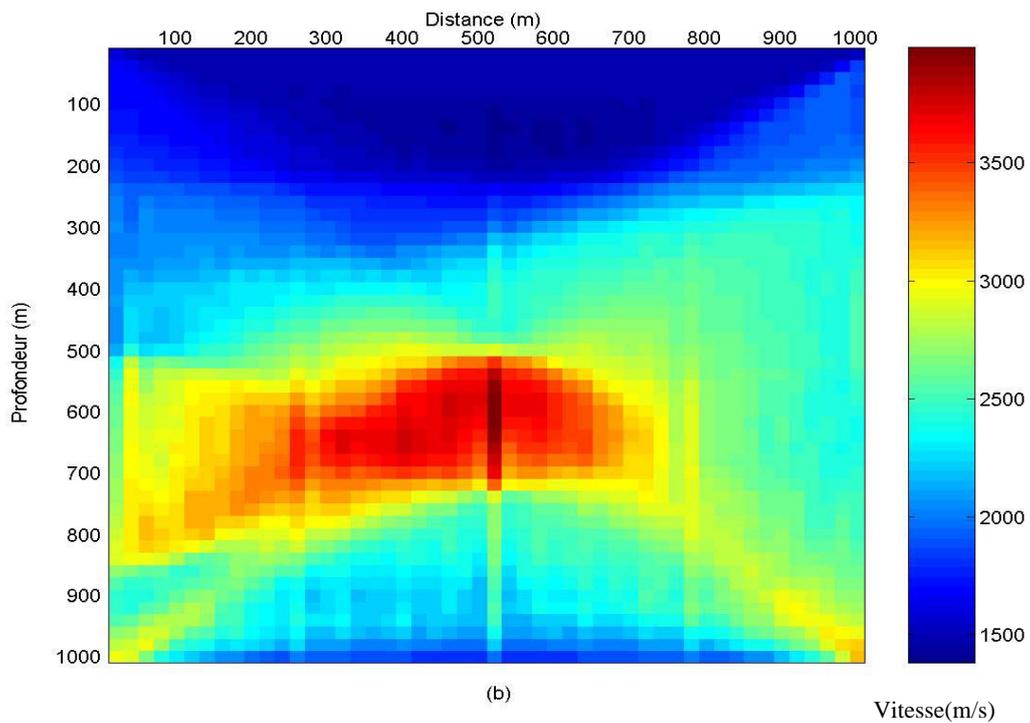
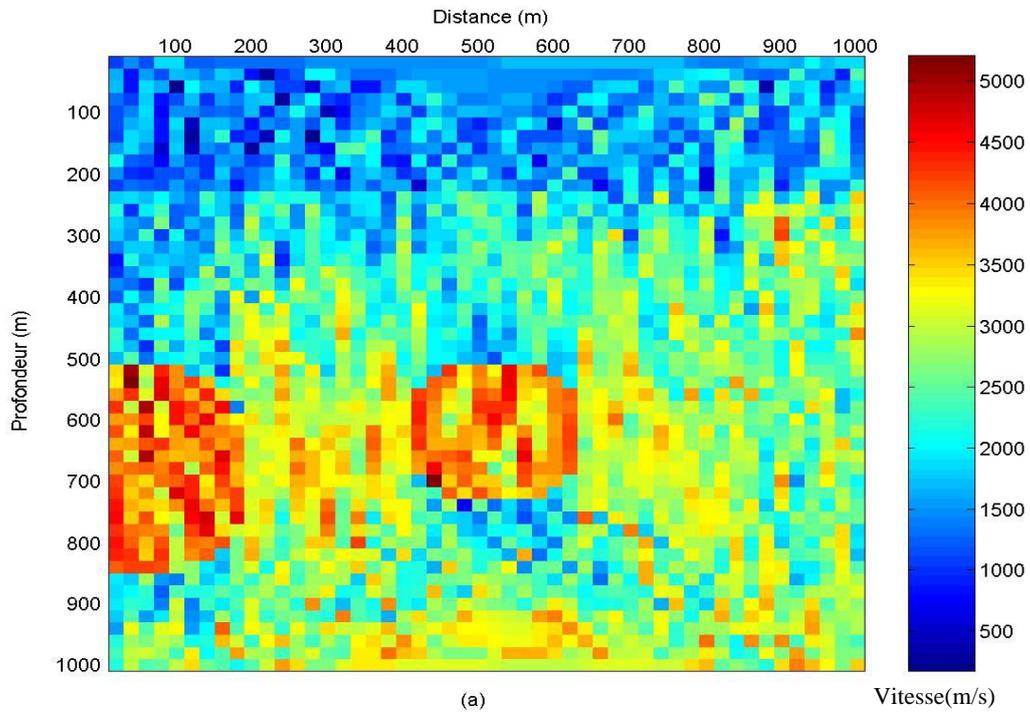
Les différentes méthodes retrouvent bien les grandes lignes du modèle synthétique. Cependant, là encore, la localisation spatiale des anomalies est meilleure avec les méthodes RNA-FBR et GC et moins importante avec l'ART et la SIRT, par contre on trouve que le RNA-Elman est incapable de réaliser cette reconstruction. La figure III-6-g représente les courbes d'erreur de reconstruction par les différentes méthodes. Comme pour le modèle synthétique précédent, nous avons calculé la corrélation entre les images issues des tomographies et le modèle synthétique. Le tableau suivant résume les caractéristiques de reconstruction sur le modèle 03.

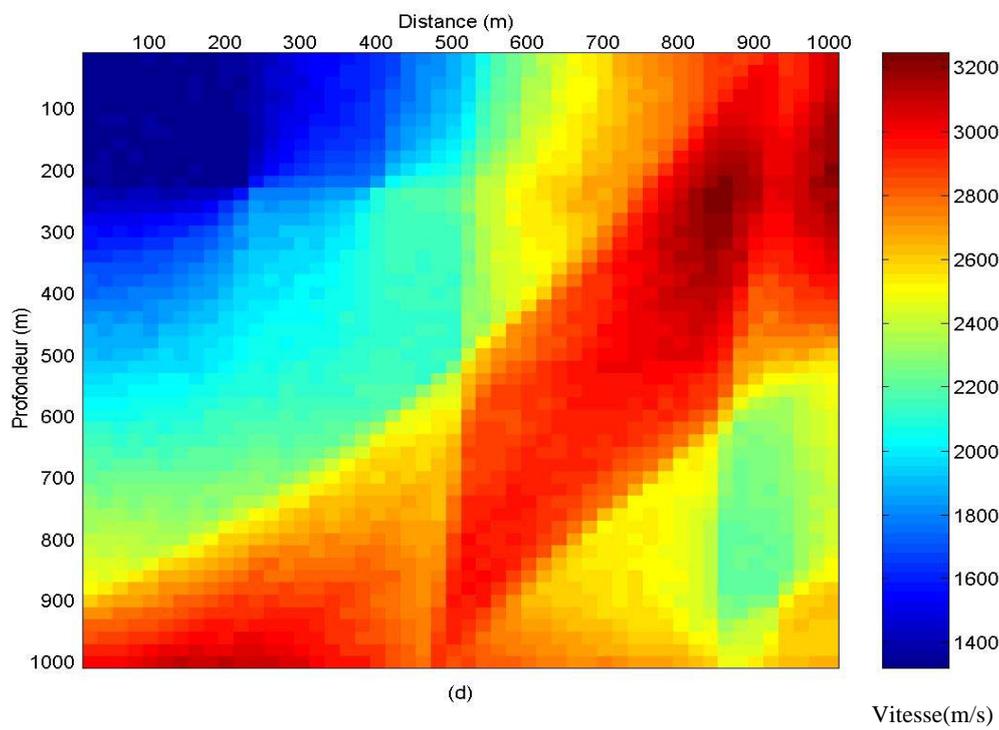
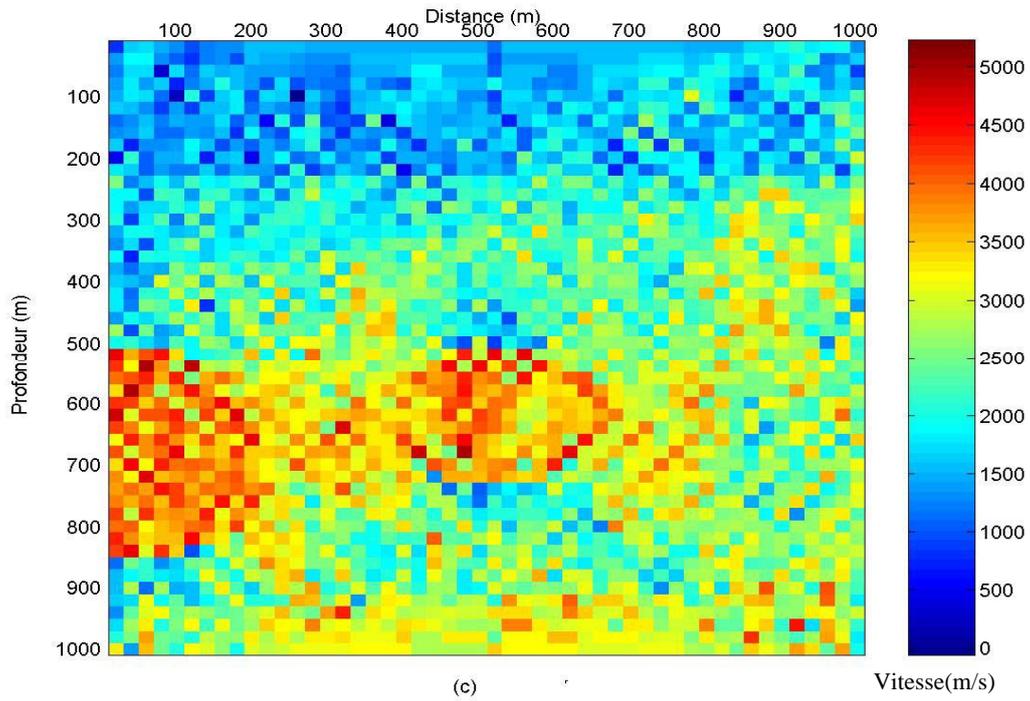
Méthode	RNA -FBR	RNA-Elman	GC	ART	SIRT
Coefficient de corrélation	0.9977	0.3110	0.9435	0.9229	0.74185
Temps de calcul (secondes)	86	30	58	129	209

Tableau 04. Récapitulatif des caractéristiques de reconstruction sur le modèle 03

III-3 Robustesse au bruit sur les données de temps de parcours

Toutes les données précédemment utilisées sont dénuées de bruit. Dans cette partie, nous avons ajouté un bruit blanc, de moyenne nul et d'écart type 1, aux temps de parcours dans le Modèle 2.





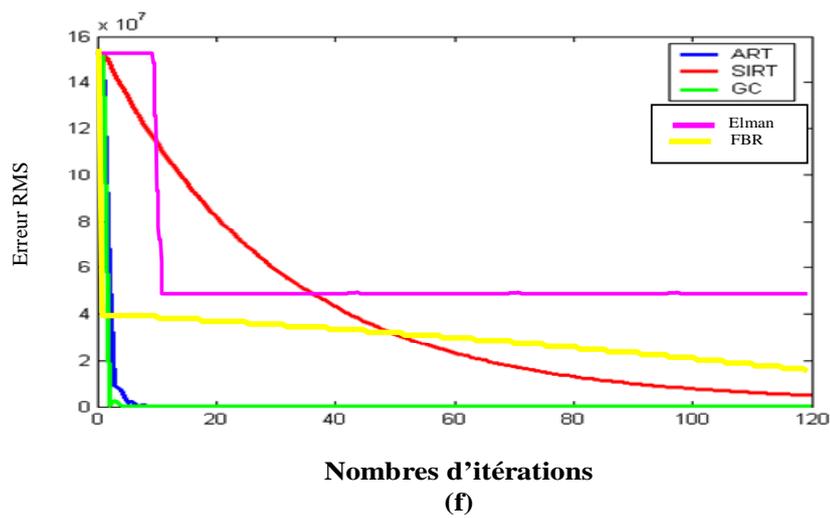
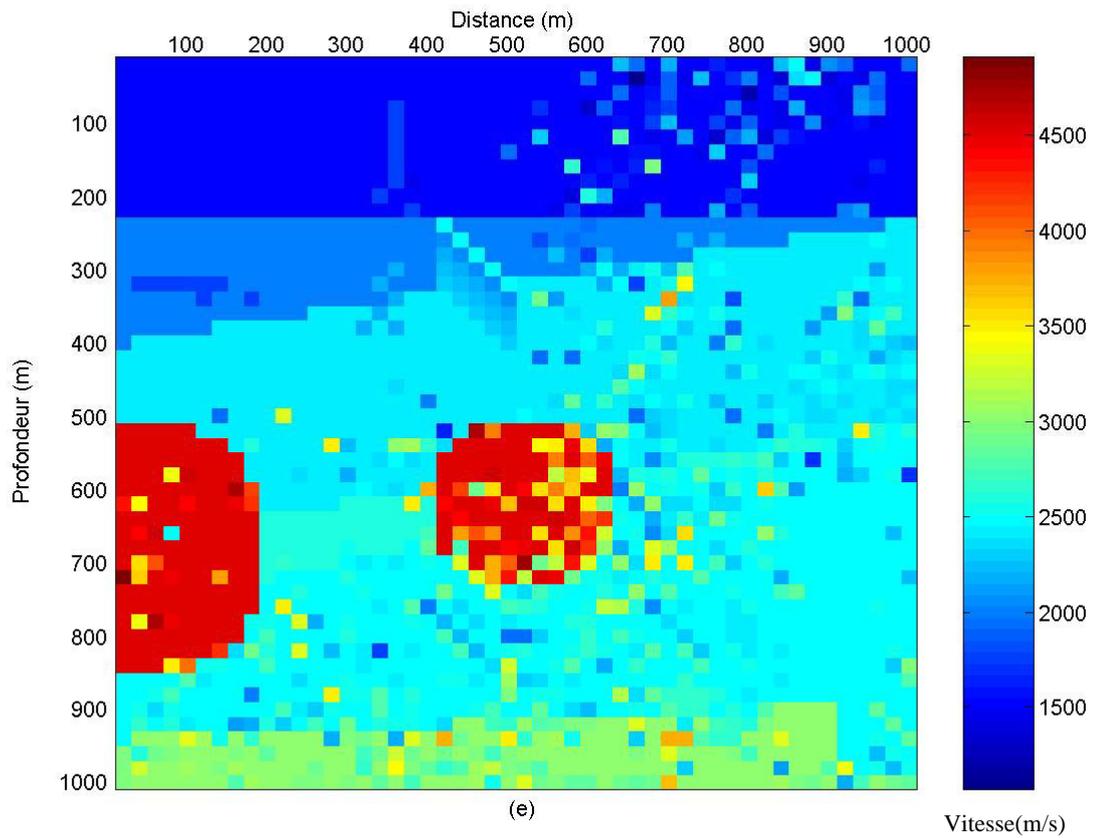


Figure III-07. a- GC, b- SIRT, c- ART. d- RNA-Elman, e- RNA-FBR, f- Courbes des résidus des différentes méthodes.

La figure III-07 illustre le résultat de la tomographie sur le Modèle 3. Bien que les grandes structures du Modèle 2 soient retrouvées par les RNA FBR (figure III-07-e), GC (figure III-

07-a) et ART (figure III-07-c) avec une résolution spatiale faible, avec la présence de plusieurs artefacts pour ces deux dernières méthodes, les méthode SIRT et RNA-Elman (figures III-07-b et d) ont complètement déformé ces structures. Cependant, les méthodes SIRT et RNA-Elman peuvent atteindre un palier au-delà duquel l'algorithme aura fini de traiter l'information utile contenue dans le vecteur d'observation, et commencera à se stabiliser autour d'une valeur constante. De même, on remarque la diminution de l'erreur en fonction du nombre d'itérations et ceci pour toutes les méthodes (figure III-07-f). Il est à signaler également que la résolution est meilleure dans le cas du RNA FBR comparée aux autres méthodes.

Méthode	RNA -FBR	RNA-Elman	GC	ART	SIRT
Coefficient de corrélation	0.93	0.2343	0,89	0,45	0,3
Temps de calcul (secondes)	86	30	58	129	209

Tableau 05. Récapitulatif des caractéristiques de reconstruction en présence de bruit sur le modèle 02

III-4 CONCLUSION

Dans ce chapitre, des réseaux de neurones, de types FBR et Elman, ont été entraînés pour l'estimation des répartitions de vitesse à partir des temps de trajets dans une expérience de tomographique **sismique**.

Avec les techniques de reconstruction effectuées sur des données synthétiques on a pu démontrer l'efficacité et les limites des méthodes classiques, la méthode du **GC** a donné des résultats acceptables, suivie par la méthode ART, **en revanche** la SIRT a échouée dans des cas complexes.

Les résultats obtenus montrent que :

1. Les réseaux ont pu établir une relation entre les temps de trajet et la vitesse, grâce à leurs capacités d'approximation et d'adaptation.
2. l'utilisation des réseaux de neurones comme outil d'inversion à un avantage notable de par rapport à des techniques classiques plus élémentaires telles que la ART, SIRT, GC, **même si le modèle est complexe**.

CHAPITRE IV

L'inversion stratigraphique : Application et résultats



IV-1 L'INVERSION STRATIGRAPHIQUE : APPLICATION ET RESULTATS

Étant donné que l'objectif de cette phase de simulation est l'estimation de la répartition des impédances acoustiques à partir des données sismiques par RNA (Elman et FBR), trois modèles de subsurface, avec des complexités variables, ont été utilisés. Le sismogramme utilisé pour l'apprentissage du RNA ne simule que l'onde de compression P calculée à incidence normale. Sur la base de ces modèles, des sections sismiques ont été calculées par convolution avec une impulsion de Ricker de fréquence centrale 35 Hz ; les trajets sont calculés suivant le principe de Snell–Descartes [Henry 1997]. Les propriétés acoustiques sont utilisées pour générer des données synthétiques qui seront utilisées pour l'apprentissage et le test de l'efficacité du réseau.

Dans nos simulations, le réseau ne possède qu'une seule couche cachée. Pour les deux premières couches (couche d'entrée, couche cachée), la fonction de transfert utilisée est linéaire et sigmoïde ; alors que les neurones de sortie utilisaient une fonction linéaire. **Ce qui** constitue le choix le plus approprié pour une tâche d'approximation par opposition à une tâche de classification et/ou de catégorisation.

En accord avec la méthodologie des réseaux de neurones, chaque base a été divisée en un ensemble d'apprentissage (collection de traces) et un ensemble de validation. On a utilisé 50% des traces pour l'entraînement et les traces restantes pour la validation. Pour chaque architecture, nous avons varié les paramètres de conception pour observer l'effet sur la performance de l'inversion. Nous avons étudié l'impact du nombre de neurones dans la couche cachée. C'est le paramètre par excellence qui permet de sélectionner à volonté la complexité intrinsèque du modèle implémenté par une architecture de réseaux de neurones donnée.

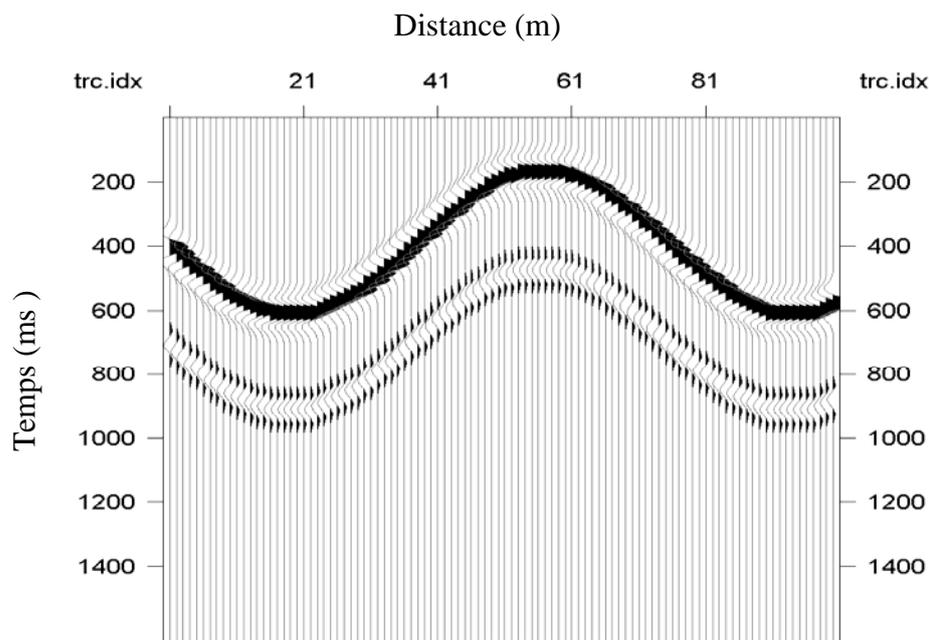
Suivant la variation de la performance pour chacune des architectures expérimentées, telles que décrite précédemment, le nombre de neurones **peut varier** entre 10 et 90 neurones. Les résultats sont reportés dans des tableaux pour chaque modèle. Pour le cas de 90 neurones, nous présentons les résultats pour chaque architecture individuellement.

V-2-a Modèle 1

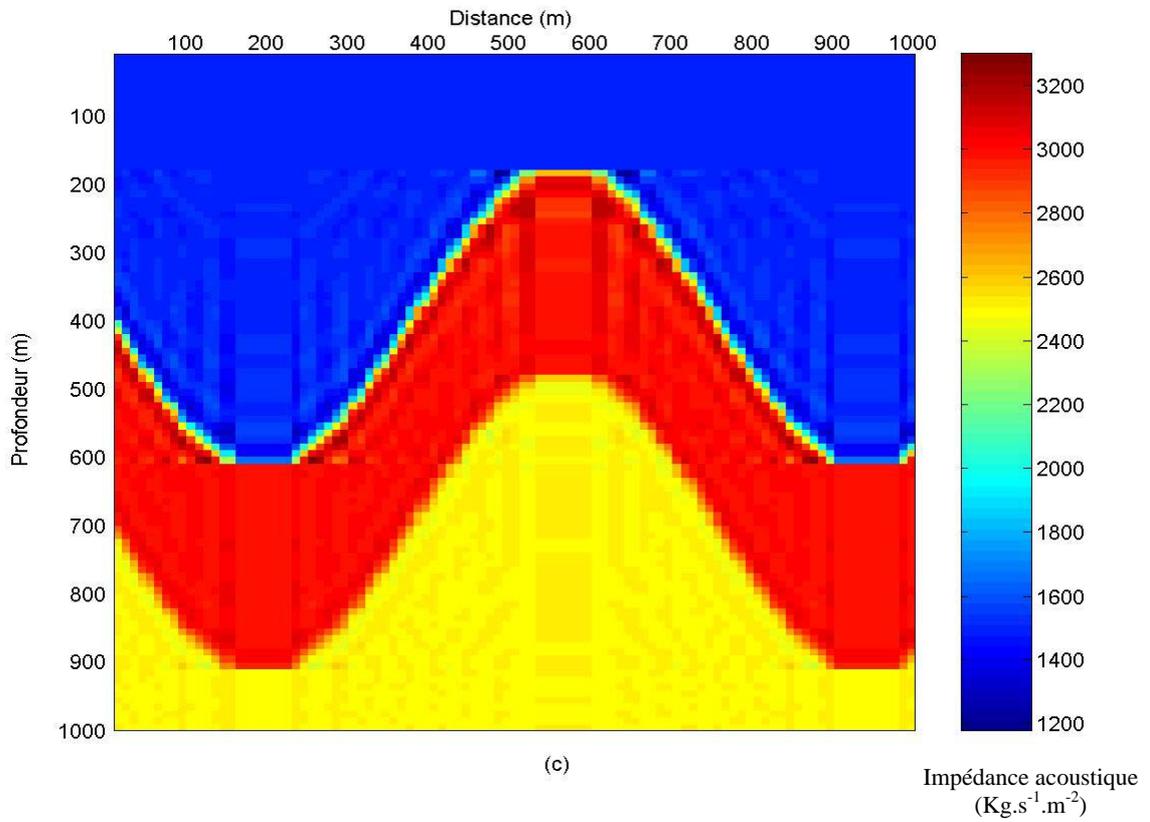
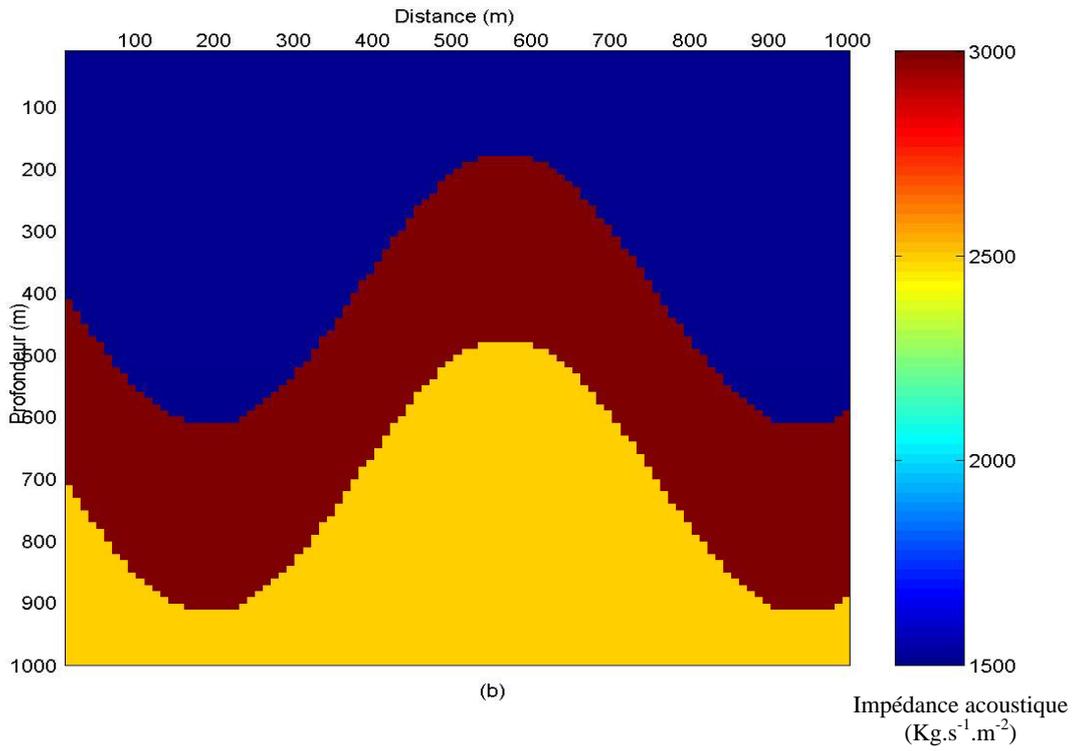
La figure IV-01-a représente la section sismique calculée sur le modèle de répartition des impédances acoustiques (figure IV-01-b). Ce modèle consiste en une anomalie négative de

forme anticlinale. Ce modèle est à retrouver par inversion à base des RNA de types Elman et FBR.

Plusieurs tests d'apprentissage on été réalisés avec un nombre variable de neurones dans la couche cachée. Sur le même modèle, les performances obtenues sont reportées dans le tableau 06 et montre que les performances obtenues sont concluantes. Dans la figure IV-01-c la sortie de réseau Elman dans le cas de 80 neurones dans la couche cachée, nous avons utilisé toutes les traces (50% utilisées en apprentissage, 50% restantes), ce que constitue le test d'apprentissage et de généralisation. En comparant la sortie du réseau après apprentissage avec le modèle utilisé pour le calcul de la section sismique, on constate que le réseau a parfaitement établi la tâche d'inversion. La figure IV-01-d montre le résultat du RNA-FBR (80 neurones) selon la même procédure **donnant ainsi** une certaine indication de l'utilité des modèles neuronaux par rapport aux techniques linéaires simplistes. Il semble que les architectures apprennent avec une performance comparable pour la phase d'apprentissage et de généralisation. Les méthodes Elman et FBR (figure IV-01-c et d) présentent des images similaires, cependant, l'impédance acoustique est meilleure pour la FBR. Les courbes d'apprentissage (figure IV-01-e) montrent la simplicité de l'apprentissage par rapport aux deux structures neuronales proposées. De même les RNA-FBR apprennent plus facilement que le modèle Elman.



(a)



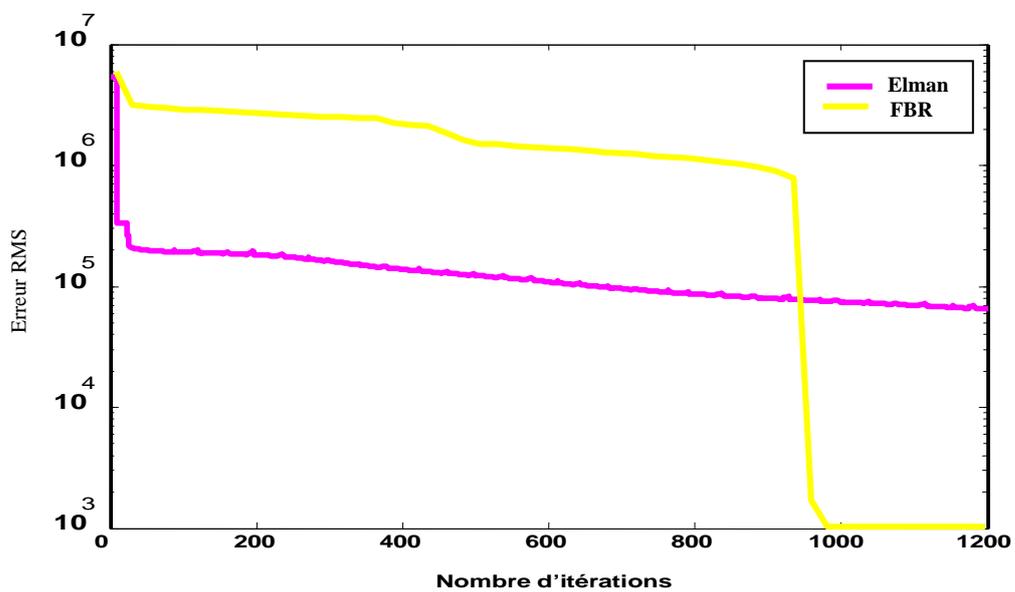
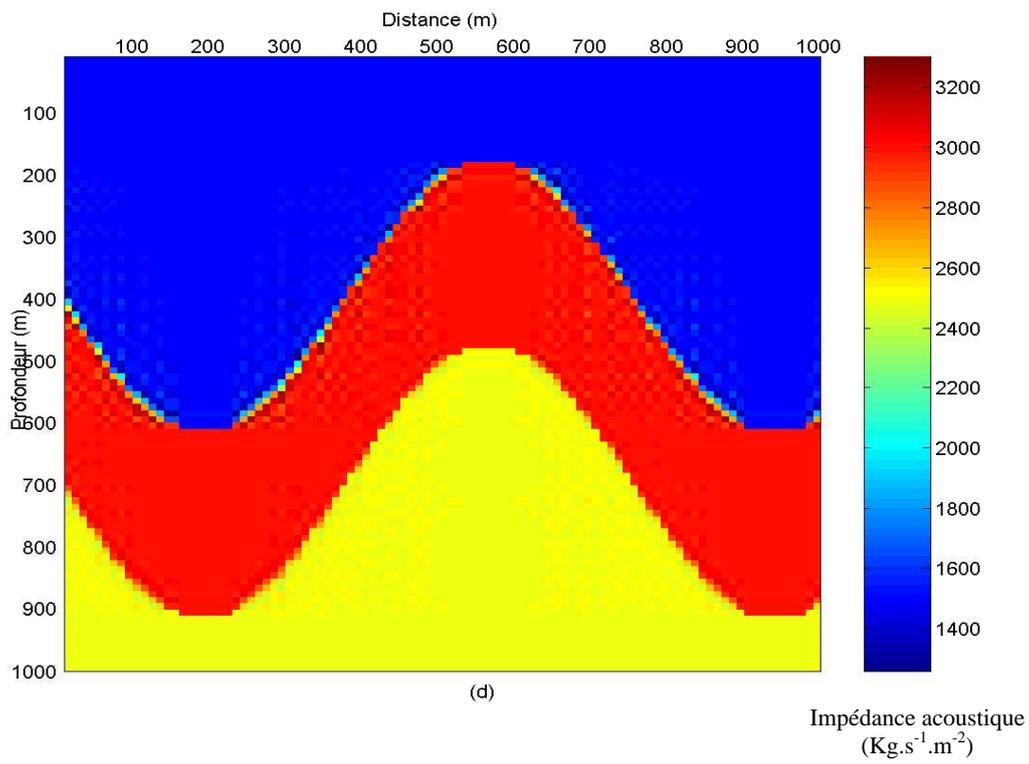


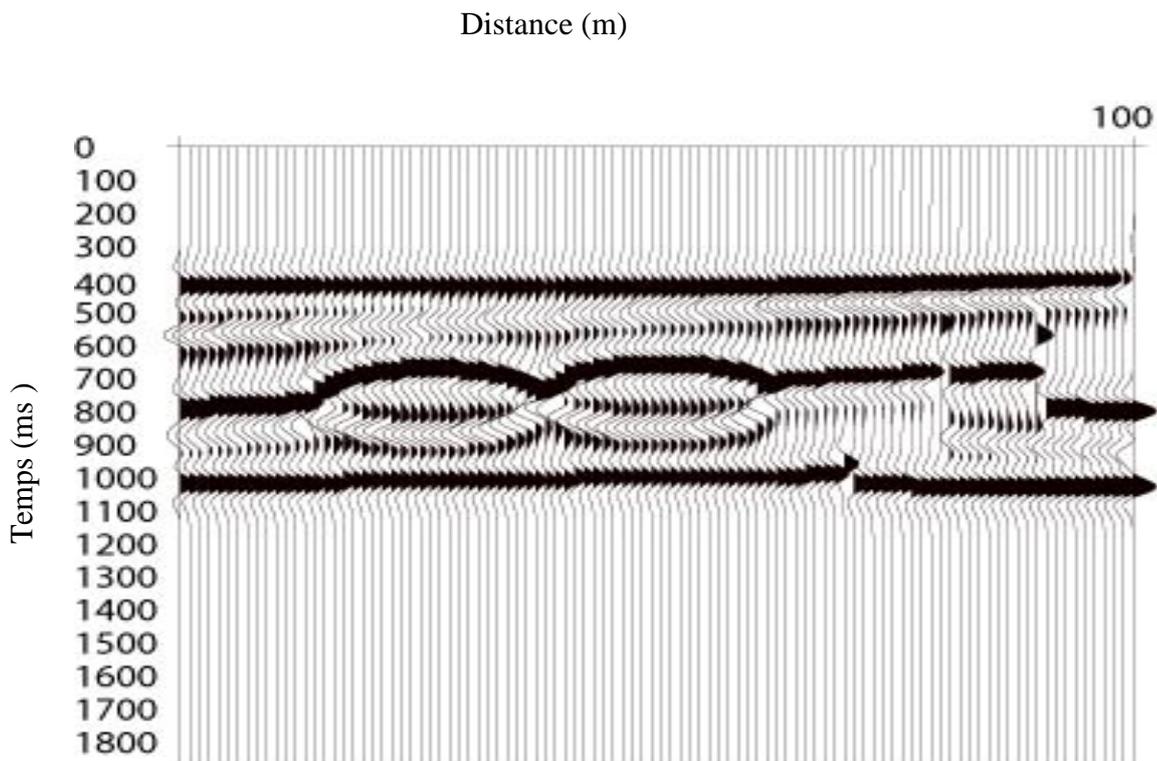
Figure IV-01 a- section sismique (entrée du RNA), b- modèle à retrouver (sortie désirée), les résultats de l'inversion des méthodes : c- RNA-Elman cas de 90 neurones dans la couche cachée. d- RNA-FBR cas de 90 neurones dans la couche cachée. f- Courbes des résidus de différentes méthodes.

RNA	Nombre de neurones dans la couche cachée	Nombre d'itérations	Performance (erreur RMS)
Elman	10	1200	9639.23
	50	1200	7657.45
	80	1200	46645.9
RBF	10	1200	3.40594e+008
	50	1200	2.87807e+007
	80	1200	871304

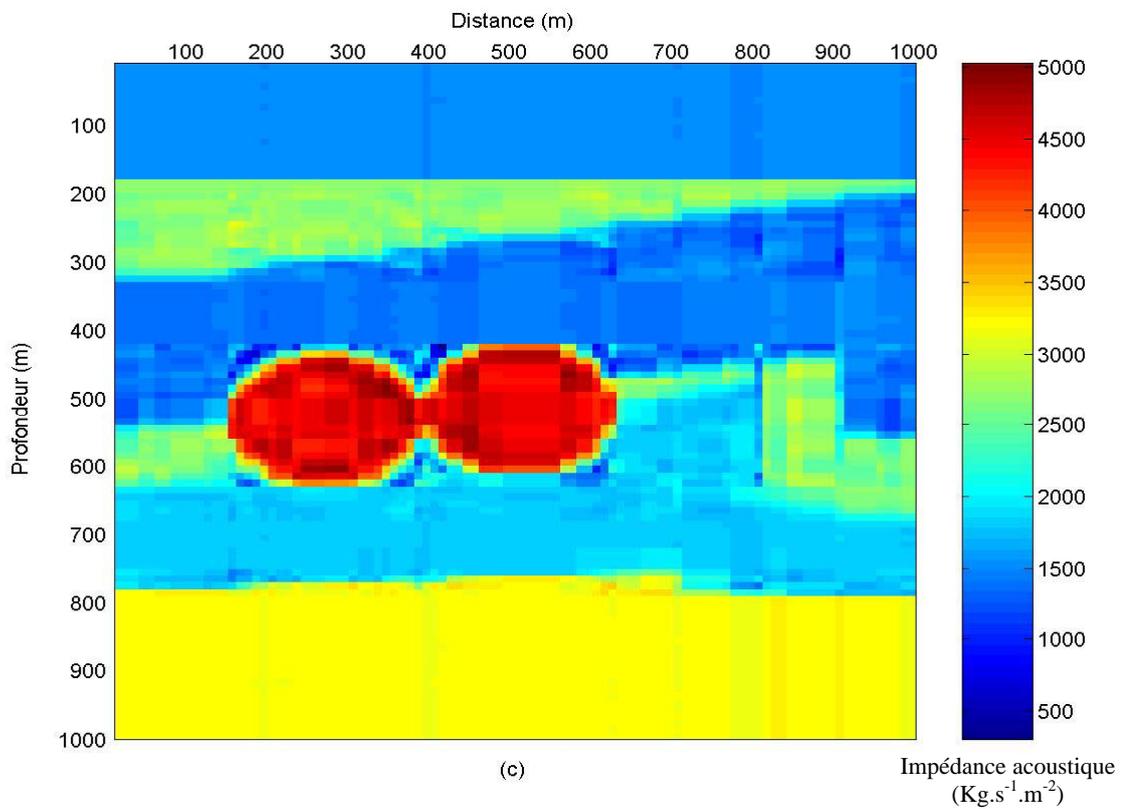
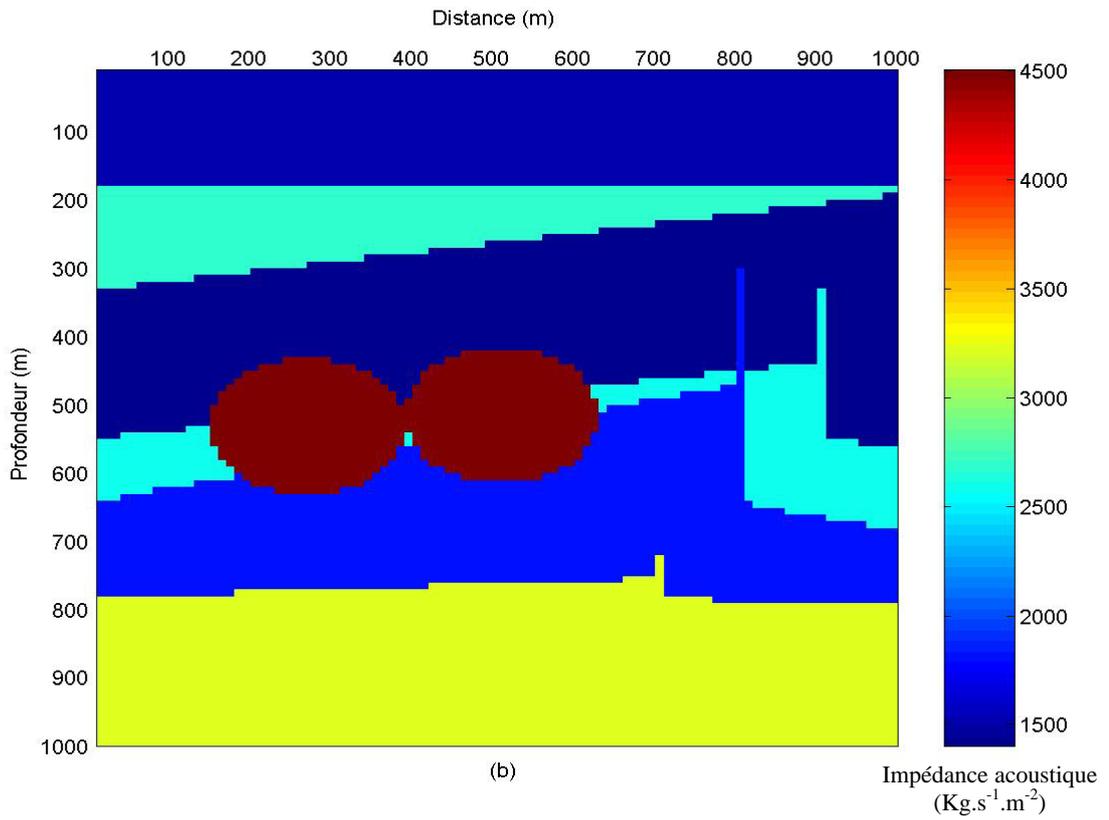
Tableau N° 06. Tableau récapitulatif de l'apprentissage du modèle 1

IV-2-b **Modèle 2**

La figure IV-02 présente le deuxième modèle proposé pour cette simulation. Ce modèle consiste en deux anomalies négatives, de forme sphérique..



(a)



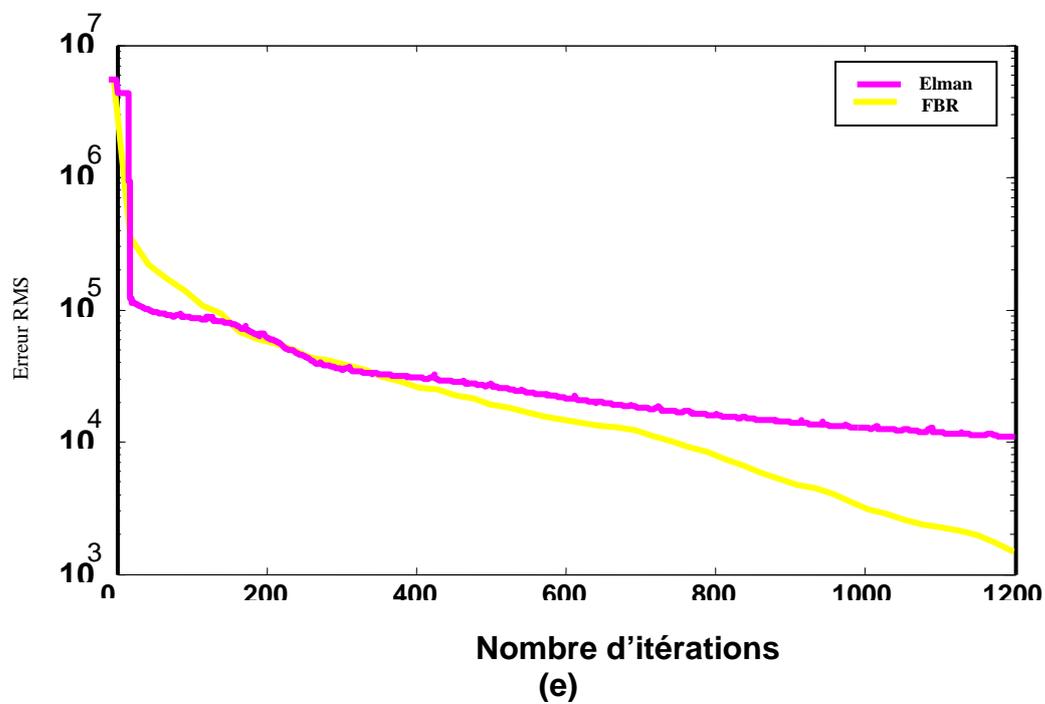
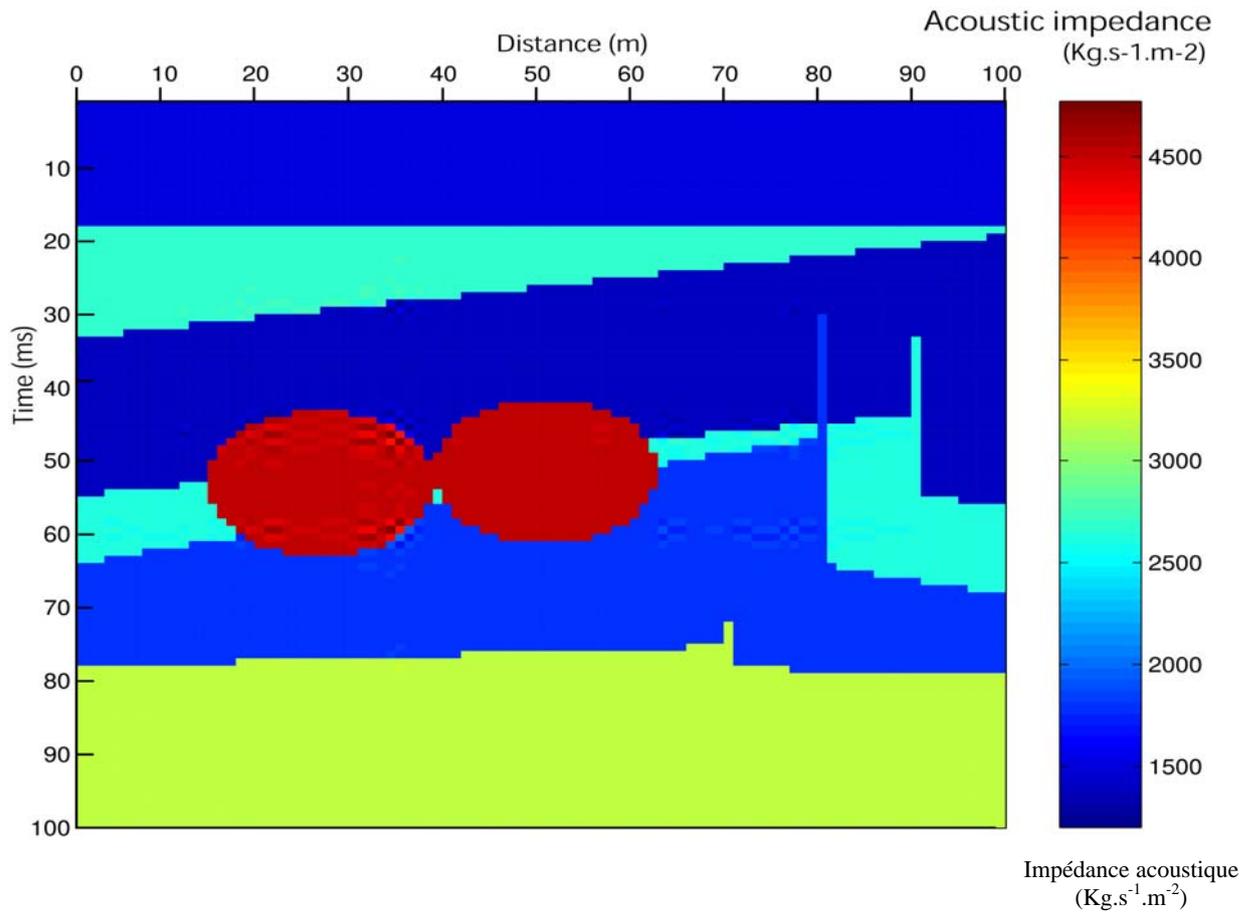


Figure IV-02 a- section sismique (entrée du RNA), b- modèle à retrouver (sortie désirée), les résultats de l'inversion des méthodes : c- RNA-Elman cas de 90 neurones dans la couche cachée. d- RNA-FBR cas de 90 neurones dans la couche cachée. f- Courbes des résidus de différentes méthodes.

RNA	Nombre de neurones dans la couche cachée	Nombre d'itérations	Performance (Erreur RMS)
Elman	10	1200	29806.6
	50	1200	36963.8
	80	1200	20437.9
RBF	10	1200	2.99237e+008
	50	1200	298351
	80	1200	184801

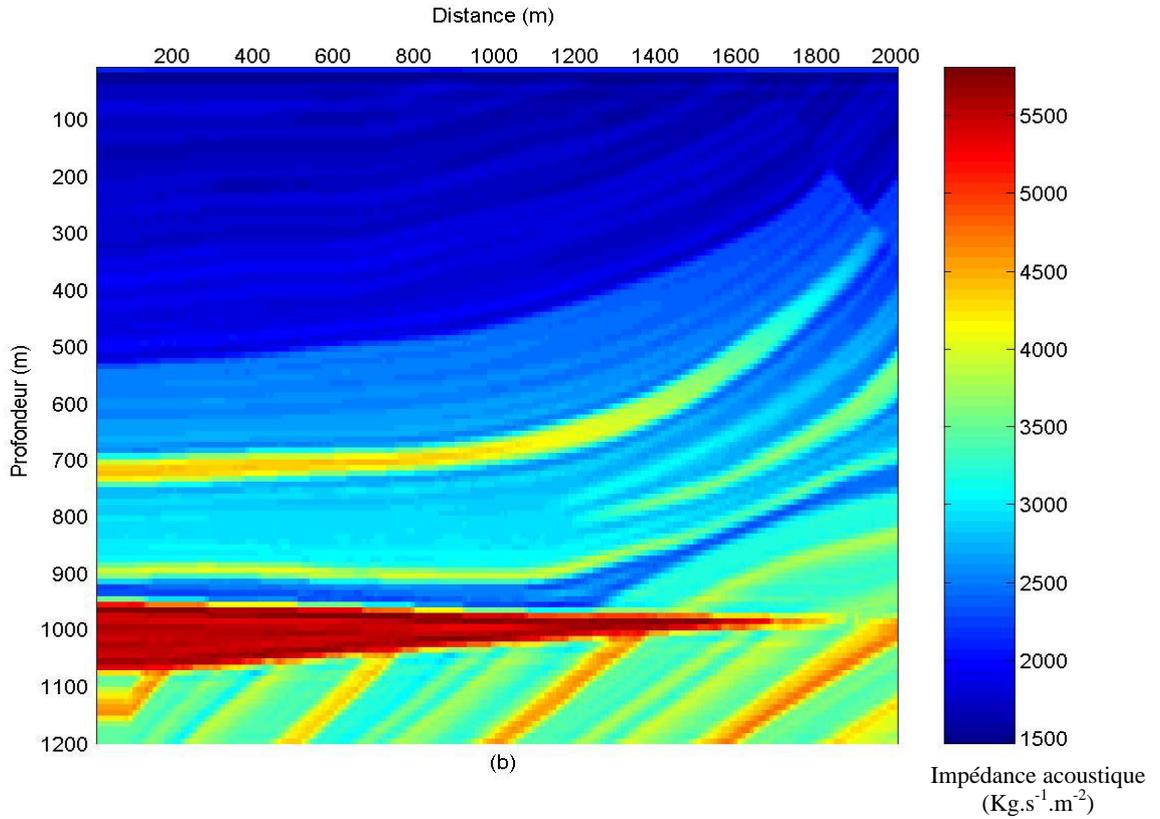
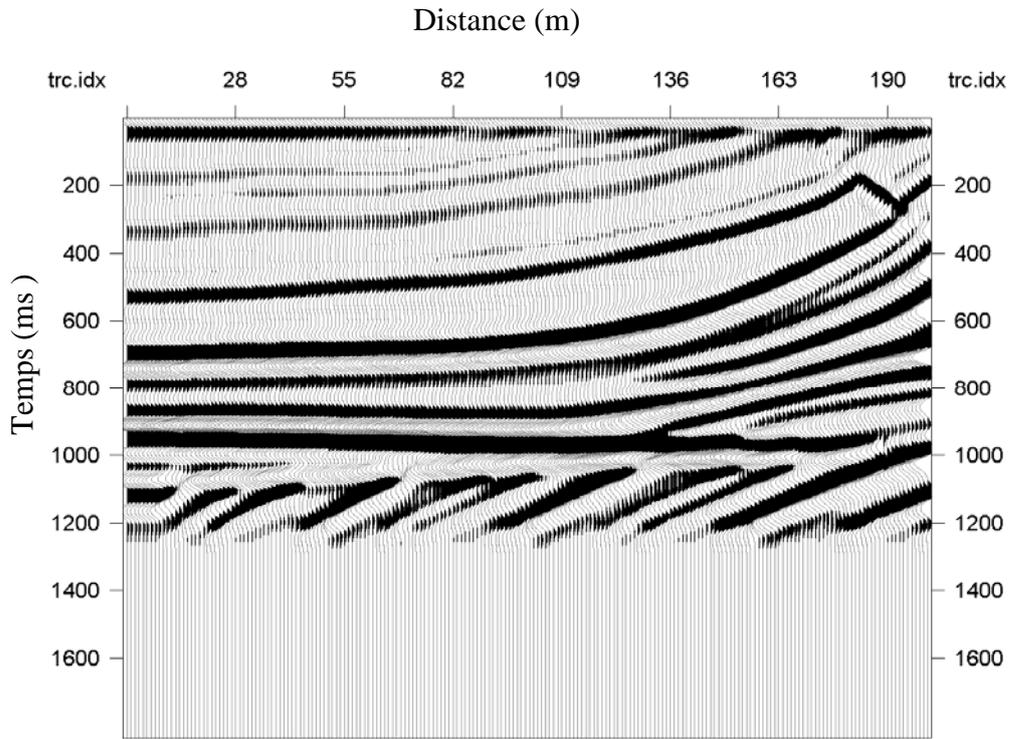
Tableau 07. Tableau récapitulatif de l'apprentissage du modèle 2

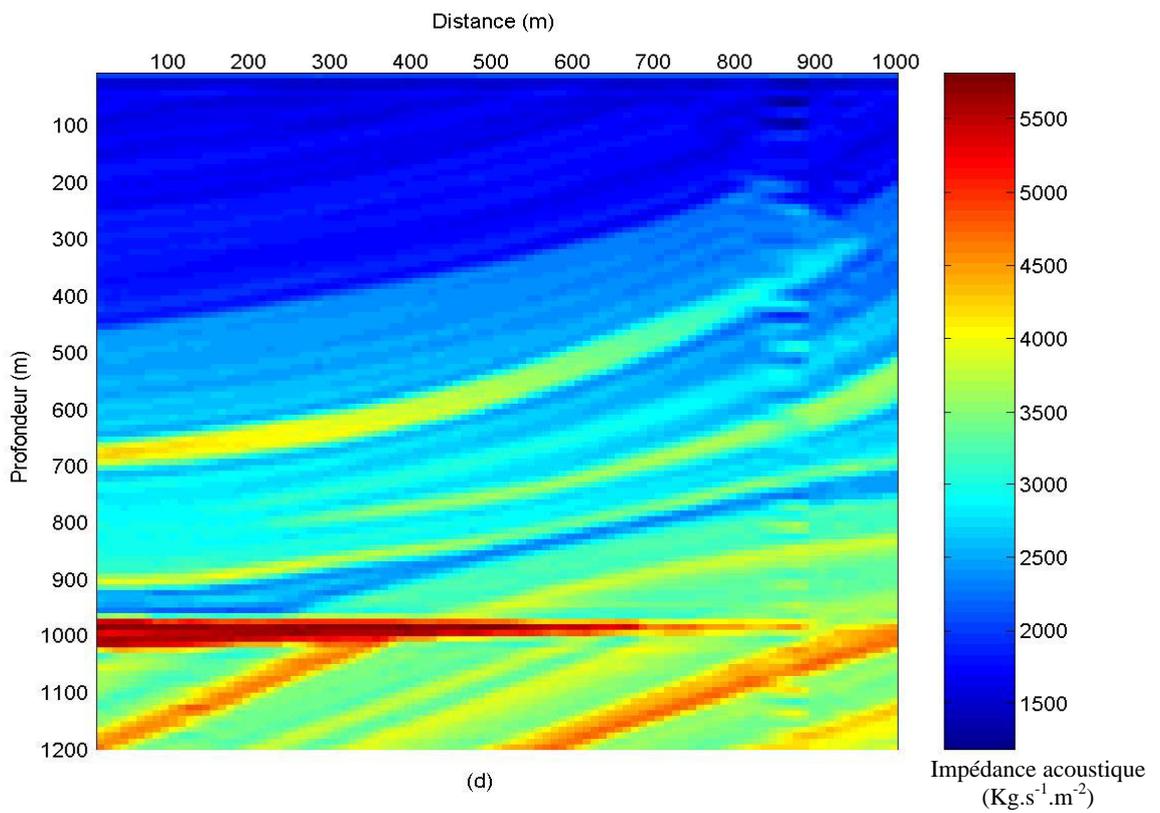
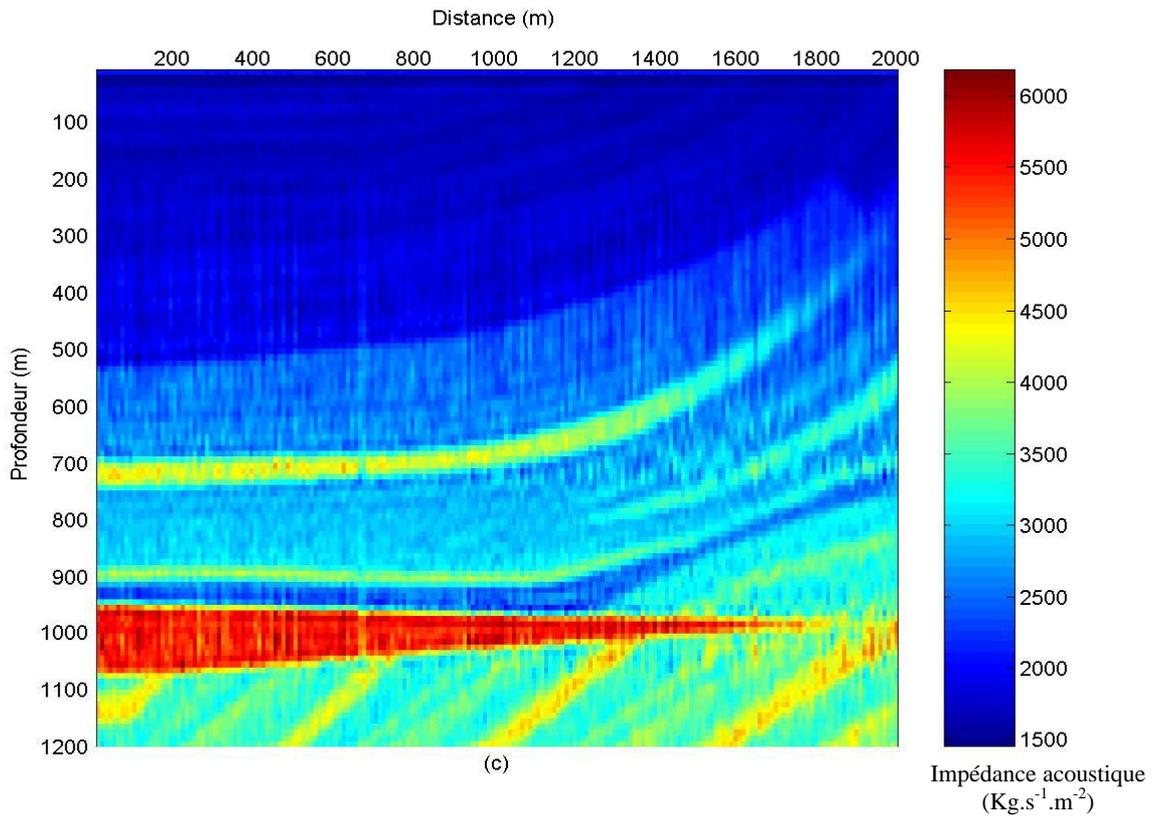
Dans le cas de la figure IV-02 relative aux tests sur le modèle 2, le réseau FBR démontre une performance globale acceptable mesurée sur l'ensemble de validation (figure IV-02-d). Il est suivi par le réseau d'Elman (figure IV-02-c). Notons aussi, que la différence entre ces deux modèles neuronaux est palpable dans la précision de l'inversion. La résolution est meilleure avec les RNA-FBR, où une nette distinction des interfaces entre toutes les couches est à signaler avec la reconstitution des deux anomalies sphériques. Les performances obtenues sont reportées dans le tableau 06.

IV-2-c Modèle 3

La figure IV-03 présente le troisième modèle proposé pour cette phase de simulation. Ce modèle est constitué de plusieurs couches avec des formes et d'impédances variables. Il représente une complexité relativement importante en comparaison avec les modèles de la figure IV-01 et IV -02. Les figures IV-03-c et IV-03-d présentent les résultats du RNA Elman et RNA-FBR.

La figure IV-03-c représente la sortie du réseau Elman avec 80 neurones dans la couche cachée, tandis que la figure IV -03-d la sortie de RNA-FBR avec 80 neurones. Les performances pour l'ensemble des tests sur le même modèle sont reportées dans le tableau 07, et montrent une performance globale comparable entre les RNA-Elman et RNA-FBR qui est mesurée sur les courbes d'évolution d'erreurs à la sortie de chaque réseau de neurones (figure IV -03-e).





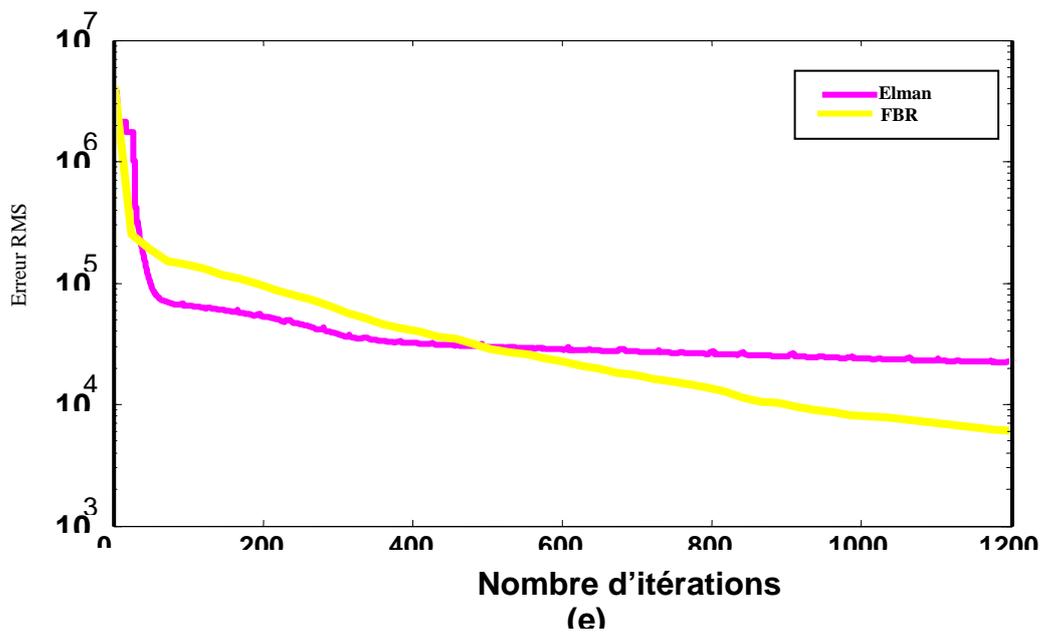


Figure IV-03 a- section sismique (entrée du RNA), b- modèle à retrouver (sortie désirée), les résultats de l'inversion des méthodes : c- RNA-Elman cas de 90 neurones dans la couche cachée. d- RNA-FBR cas de 90 neurones dans la couche cachée. f- Courbes des résidus de différentes méthodes.

RNA	Nombre de neurones dans la couche cachée	Nombre d'itérations	Performance (erreur RMS)
Elman	10	1200	15912.4
	50	1200	19881.1
	80	1200	9507.82
FBR	10	1200	4.35818e+007
	50	1200	444152
	80	1200	6.61776e-009

Tableau 08. Tableau récapitulatif de l'apprentissage du modèle 3

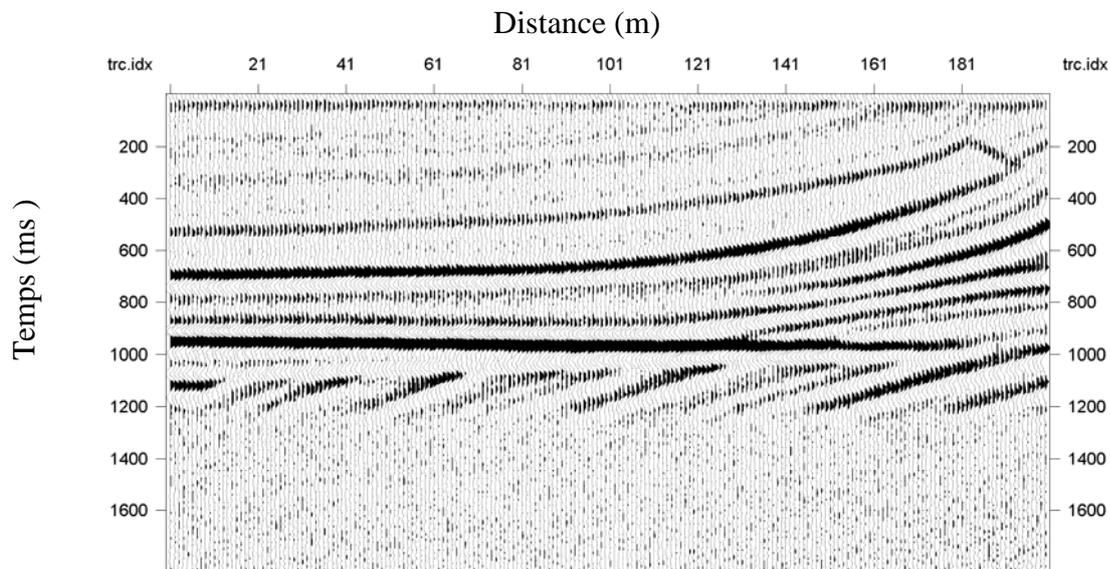
De façon générale, on observe dans ces différents résultats que la qualité de l'inversion est liée au nombre de neurones dans la couche cachée pour chaque structure proposée.

Nous constatons que la précision des estimations du réseau FBR est meilleure par rapport à celle du RNA-Elman. De même, nous avons constaté qu'à partir de 50 neurones dans la couche cachée, la qualité de l'inversion réalisée par les deux structures est largement acceptable (tableau 08). La stratégie que nous avons adoptée pour le test d'apprentissage et de

généralisation, consiste à utiliser 50% de traces, prises d'une façon aléatoire, pour l'apprentissage. Une fois que l'apprentissage est achevé, la structure entraînée est utilisée ensuite pour l'inversion de la totalité des traces et afin de passer au modèle géologique recherché, principe qui constitue le test d'apprentissage et de généralisation en même temps.

IV-3 Robustesse au bruit sur les données sismiques synthétiques

Les données réelles sont toujours bruitées, avec du bruit d'origines diverses (vent, activité sismique, vagues, activité humaine . . .). Au contraire, les données synthétiques ne sont généralement pas bruitées. Dans l'optique d'inverser des données réelles, il est donc intéressant de tester la stabilité de l'inversion par rapport au bruit. Pour cela, le sismogramme synthétique du modèle 3 est perturbé par du bruit aléatoire à 10% et 20% des amplitudes des traces simulées non bruités.



(a)

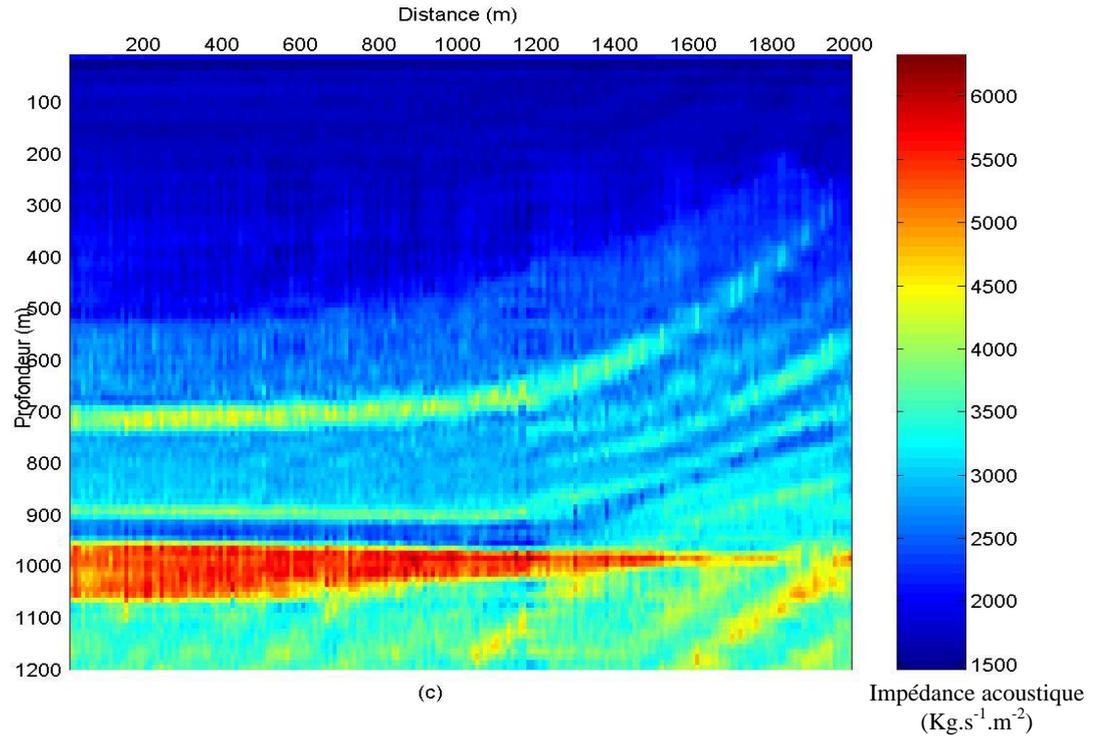
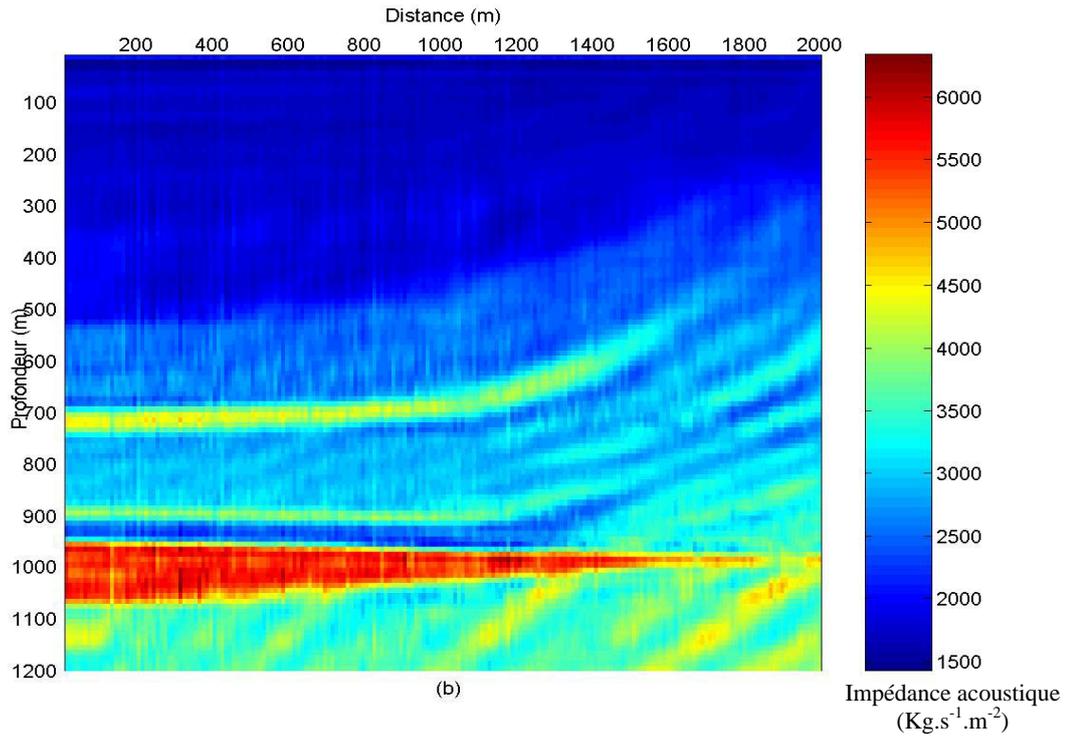
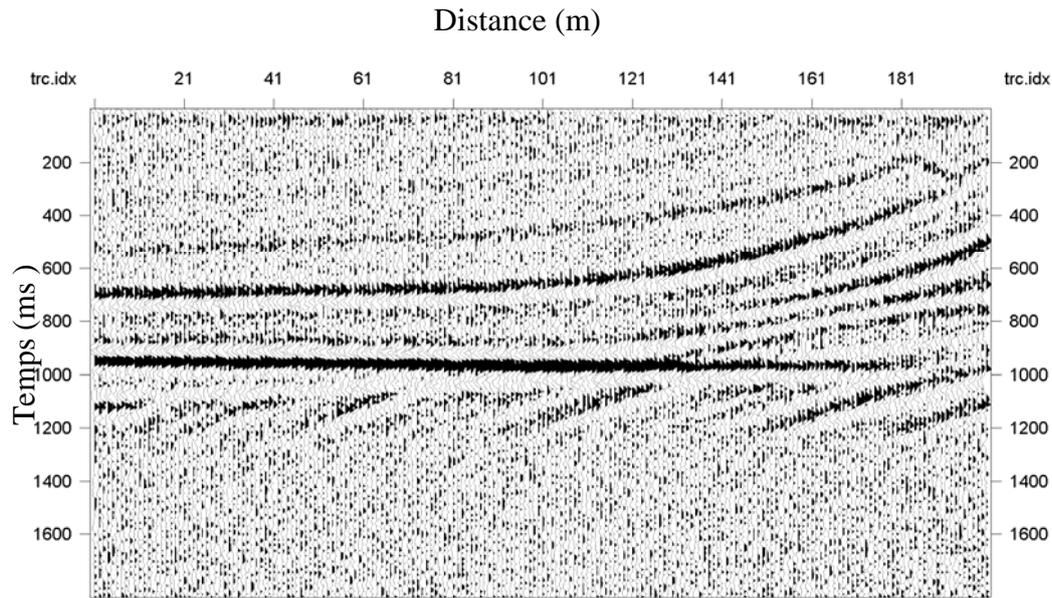


Figure IV-04. a- section sismique avec 10% de bruits aléatoires (entrée du RNA), b- résultat du RNA-Elman cas de 90 neurones dans la couche cachée. c- résultat du RNA-FBR cas de 90 neurones dans la couche cachée.

Dans le cas de la figure IV-04 relative aux tests sur le modèle 3, le réseau RNA-Elman (figure IV-04-b) démontre une performance acceptable en présence de 10% de bruits comparés au modèle initial (figure IV-03-3). En revanche, le réseau RNA-FBR (figure IV-04-c) rencontre une difficulté notable d'apprentissage, la résolution est meilleure avec Elman, dont on observe une nette distinction des interfaces entre toutes les couches qui constituent ce modèle.



(a)

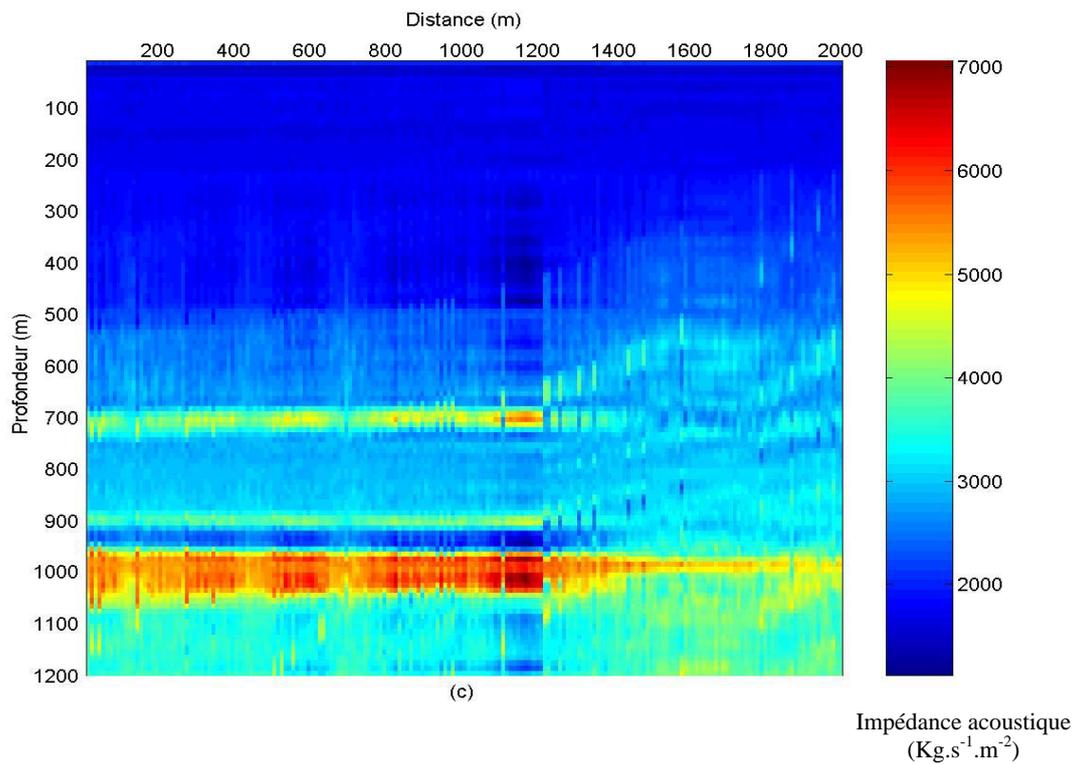
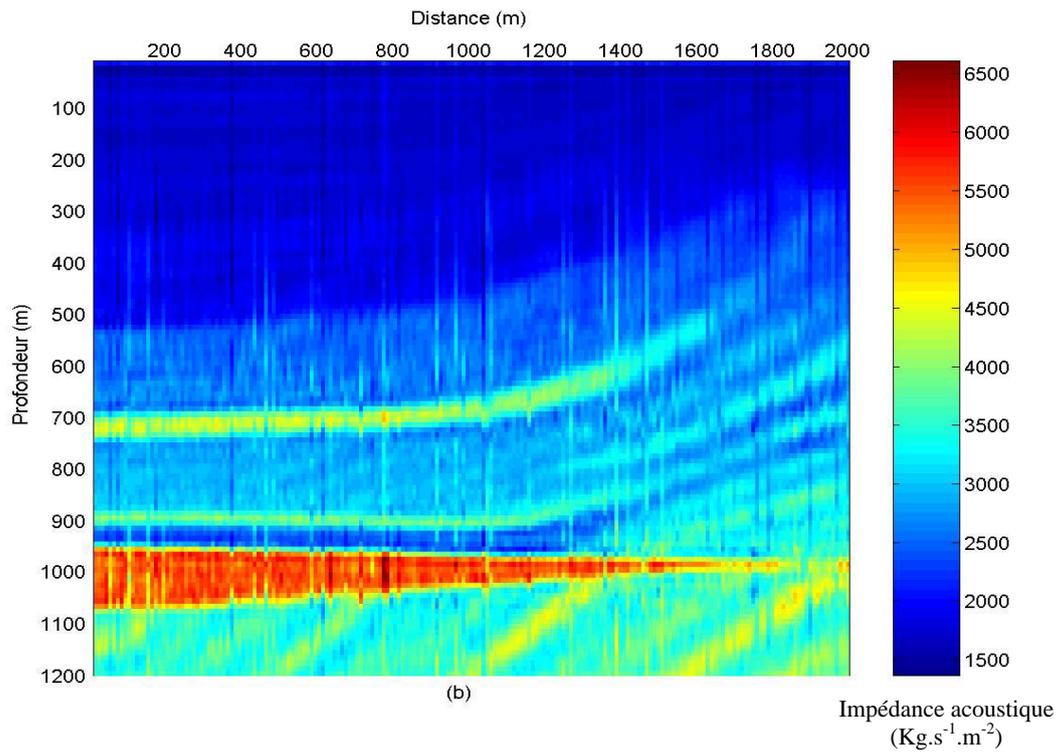


Figure IV-05 a- section sismique avec 20% de bruits aléatoires (entrée du RNA), b- le résultat du RNA-Elman cas de 90 neurones dans la couche cachée. c- résultat du RNA-FBR cas de 90 neurones dans la couche cachée.

Dans le cas de la figure IV-05 relative aux tests sur le modèle 3, le réseau Elman (figure IV-04-b) continue à donner une performance acceptable même en augmentant le pourcentage

du bruit à 20%, les traits qui constituent les différentes structure de ce modèle sont encore perceptibles. L'efficacité du réseau FBR (figure IV-04-c) se dégrade de plus en plus avec l'augmentation du pourcentage de bruit. On observe une déformation de la structure comme une conséquence de l'augmentation de pourcentage de bruit.

IV-4 APPLICATION DE L'INVERSION SUR UN PROSPECT SUD OUEST ALGERIEN.

Dans cette partie, nous appliquons les techniques de RNA dans l'inversion stratigraphique sur des données réelles, celles qui consistent à déconvoluer et à intégrer les traces sismiques. On transforme par intégration une trace sismique en une pseudo-diagraphie d'impédance acoustique dans le but de pouvoir différencier les séries lithologiques et suivre les variations de faciès et l'extension latérale de caractéristiques identifiées sur la section sismique.

L'exemple traité est extrait d'une étude réalisée sur un prospect au Sud algérien. Un traitement complet de ce profil, à partir de la mise en collection CDP, a été réalisé afin de produire des traces représentatives de la réflectivité, et donc directement utilisables pour réaliser l'inversion stratigraphique après sommation. Le traitement en amplitude préservée a permis d'éliminer les distorsions du signal, provoquées en surface par les variations de couplage des émetteurs et des récepteurs, par les interférences avec les bruits sismiques et par les atténuations différentielles lors de la traversée de la zone d'altération superficielle.

Lors du traitement de ce profil, les corrections des effets de surface ont été réalisées par une égalisation des amplitudes dans le domaine temporel (fenêtre glissante), mais qui a pour inconvénient un lissage latéral de la réflectivité du sous-sol. Pour ne compenser que les artefacts de surface, nous avons utilisé une décomposition en surface consistante qui tient compte des effets liés à chaque émission et à chaque réception, et nous avons corrigé ces anomalies dans le domaine temporel.

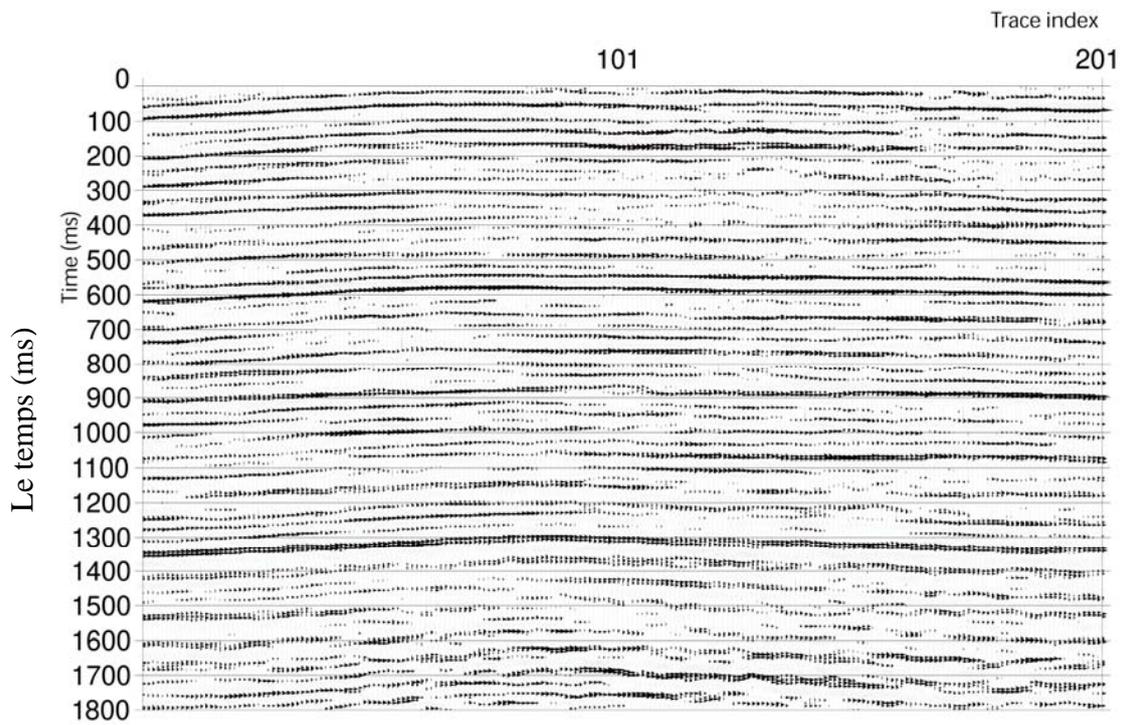
- Le profil M 100 a été exploité avec un système d'enregistrement à 48 traces en couverture 24. La source sismique consistait en des charges d'explosif.
- Des essais de filtrage FK sous NMO sur les collections points de tirs et de réceptions n'ont pas été concluants.
- Des essais d'interpolation de traces ont aussi été effectués pour réduire l'effet de l'aliasing par le processus du «Beam Steering Interpolation » qui ne furent pas non plus convaincants.
- Le traitement est mené de sorte à obtenir un document de bonne qualité pour servir de guide aux opérations nécessaires à une sortie de l'information en amplitude préservées.

Les étapes de base du traitement réalisées sont :

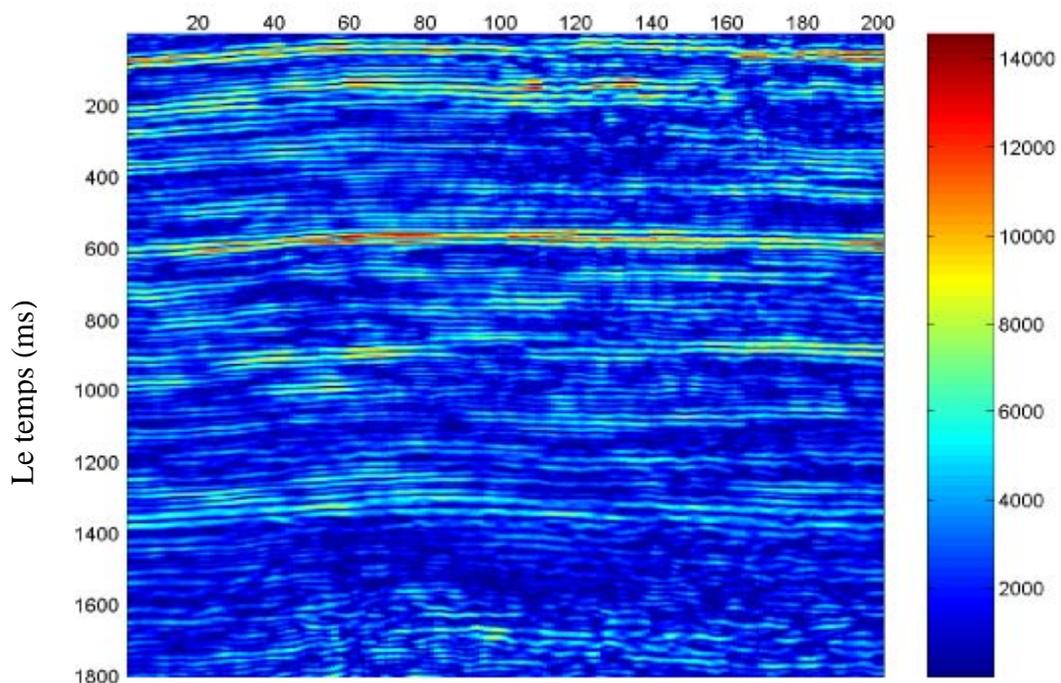
- Compensation des pertes par transmission et par divergence sphérique.
- Déconvolution réduisant la distorsion des formes de l'ondelette.
- Egalisation **dynamique** et **corrections** statiques résiduelles basées sur la compatibilité surface pour l'élimination des anomalies superficielles et du relief.
- Analyse de vitesse pour aligner les signaux et application des corrections dynamiques avec une fonction mute adéquate.

Le résultat du traitement est donné dans la figure IV-06a. Sur la base de cette section nous avons appliqué la technique de l'inversion stratigraphique sur une fenêtre de 400-2000 ms. Le résultat de l'inversion est donné en figure IV-06b. Sur la section en impédance acoustique (figure. 06-b) on observe une zone d'anomalie négative d'impédance acoustique et qui est souvent observée sous les réservoirs à gaz où les réflexions ont une faible amplitude. Cette anomalie constitue un indicateur direct de gaz. De plus, on observe sur la même section de la figure IV-06 b des variations latérales de plusieurs niveaux, dont les valeurs absolues sont probablement inexactes, mais dont l'évolution relative est sans doute significative. Sur cette section on observe plusieurs niveaux de variation d'impédance acoustique avec la continuité de plusieurs horizons et discontinuité des autres interfaces. De même sur cette section il y a un changement de contraste d'impédance qui se manifeste dans certaines régions de la section. Ce résultat montre un exemple représentatif de données sismiques qu'un RNA doit apprendre à identifier et à inverser.

Un prétraitement a été appliqué sur les données d'entrée et de sortie désiré. Il s'agit d'une transformation classique et universellement utilisée pour les réseaux de neurones qui permettent d'obtenir une convergence plus rapide et une meilleure adaptation. Cette normalisation assure que les variables d'entrée et de sortie soient de l'ordre de l'unité, auquel cas, on s'attend à ce que les poids synaptiques du réseau soient également de l'ordre de l'unité. Les poids peuvent alors recevoir une initialisation aléatoire appropriée avant de commencer l'apprentissage. Sans ce changement d'échelle linéaire, on se trouverait dans l'obligation de trouver une solution pour les poids dans laquelle certains poids auraient des valeurs dramatiquement différentes des autres. Dans cette phase nous avons utilisé la fonction PRESTD (PREprocesses STandard Deviations) qui permet d'avoir des données avec une moyenne 0 et une déviation standard égale à 1.



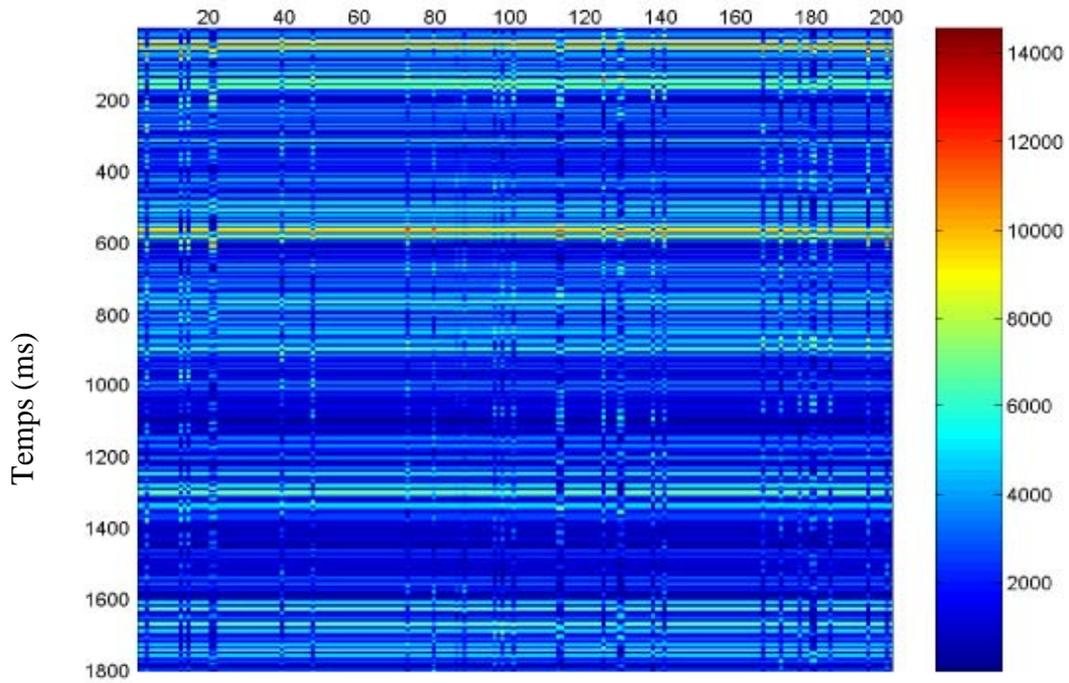
(a)



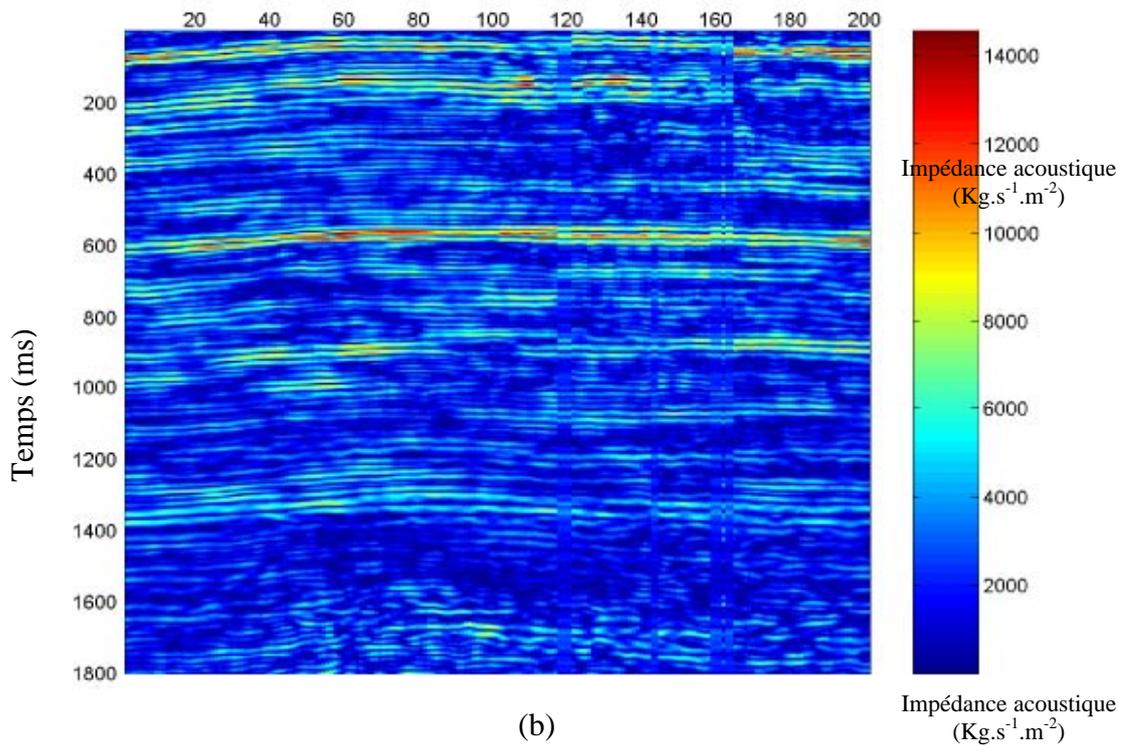
(b)

Impédance acoustique
($\text{Kg.s}^{-1}.\text{m}^{-2}$)

Figure IV-06 (a) Section sismique sommée traitée en amplitude préservée
(b) Impédance acoustique (section désirée)



(a)



(b)

Figure V-07 (a) Résultat du RNA Elman avec 5 neurones
(b) Résultat du RNA Elman à 90 neurones

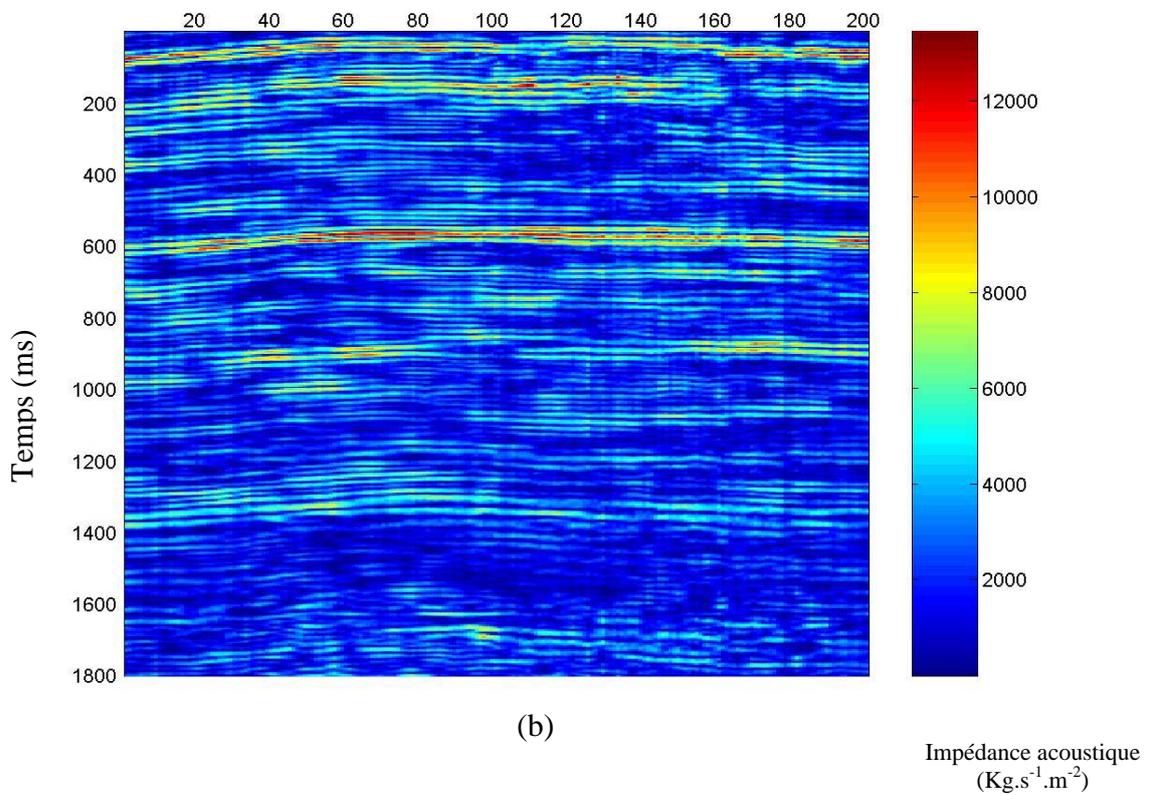
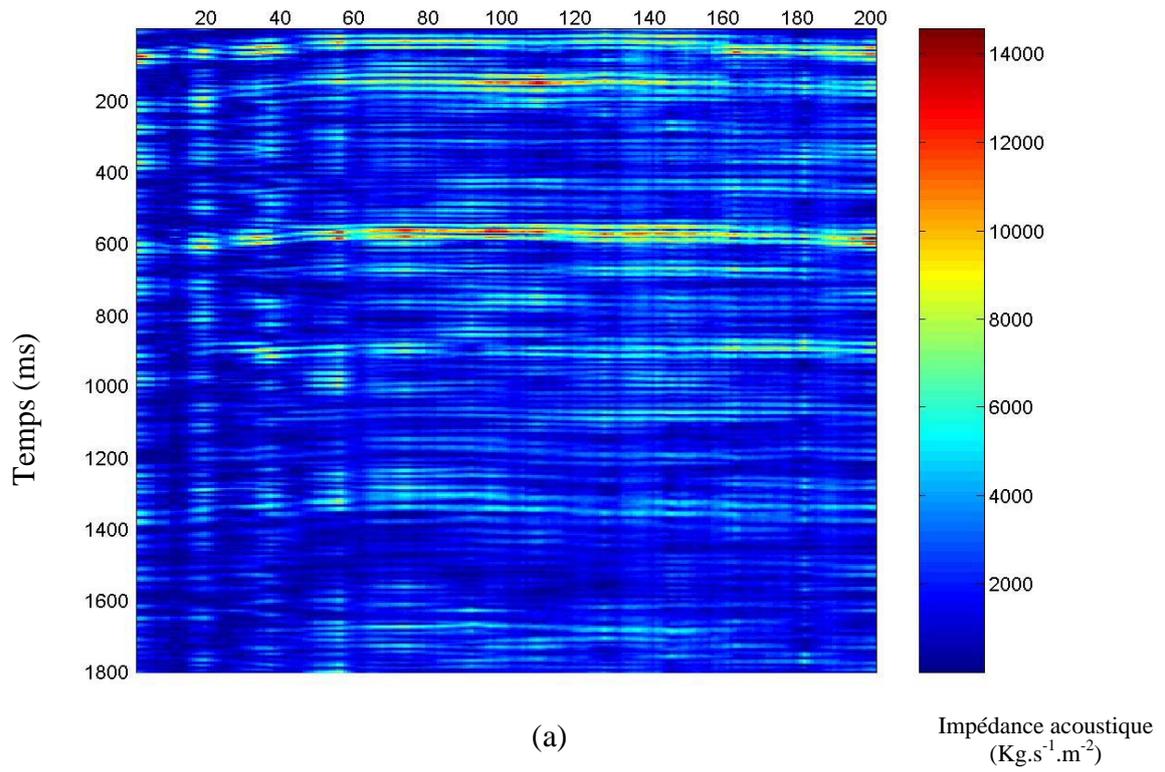


Figure V-08 (a) Résultat de la FBR à 10 neurones
 (b) Résultat de la FBR à 50 neurones

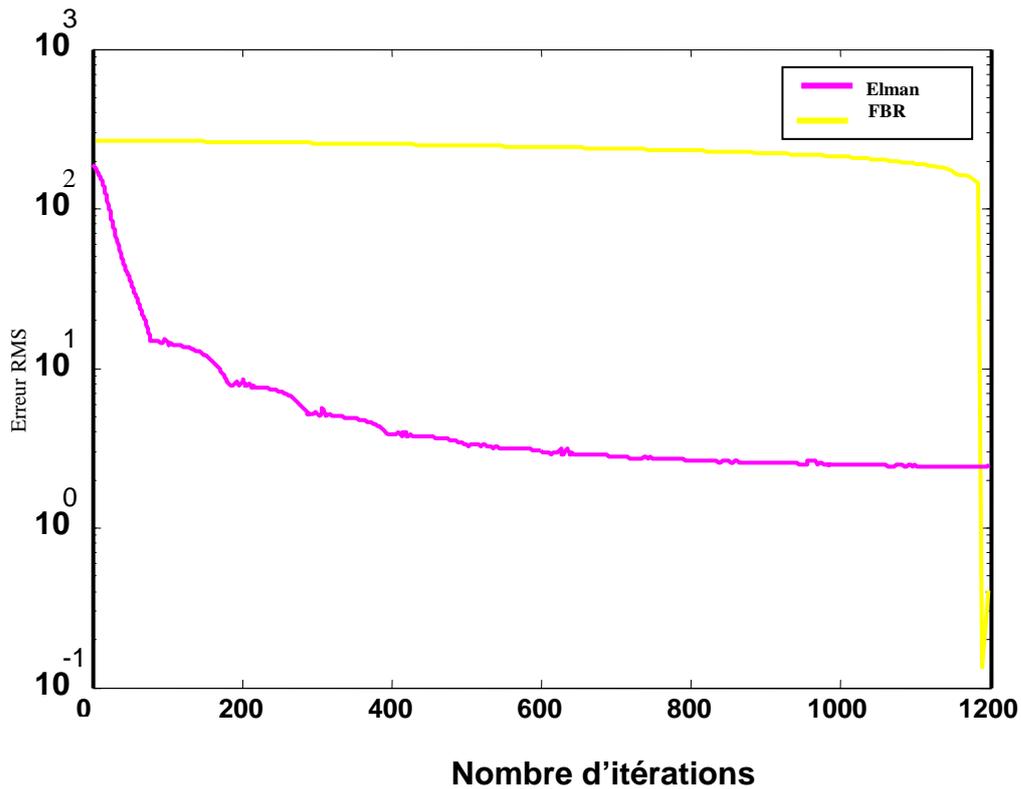


Figure V-09. Courbes d'apprentissage des deux structures neuronales proposées.

Type de RNA	Nombre de neurones dans la couche cachée	Nombre d'itération	Performance (Erreur RMS)	Correlation
Elman	5	1200	1.08654	0.3198
	15	1200	1.13275	0.3218
	55	1200	0.952139	0.4913
	75	1200	0.775327	0.5925
	90	1200	0.718436	0.6304
FBR	23	1200	509.069	0.1837
	25	1200	610.27	0.2310
	34	1200	2194.92	0.3067
	67	1200	4199.89	0.4623
	190	1200	3515.09	0.8336

Tableau 10. Tableau récapitulatif de l'apprentissage des données réelles

Plusieurs essais d'apprentissage ont été réalisés sur les données de la figure IV-06 par variation du nombre de neurones dans la couche cachée de chaque structure proposée. Le nombre d'itérations a été fixé à 1200. Les performances obtenues sont reportées dans le tableau 09.

La Figure IV-07 illustre le résultat du RNA- Elman avec 5 neurones dans la couche cachée (Figure IV-07-a). Dans ce cas, le test d'apprentissage et de généralisation, montre l'incapacité de cette structure à réaliser le passage de la trace sismique à l'impédance acoustique, tout en conservant la configuration de la répartition désirée (Figure IV-06-b). Le comportement du RNA dans ce cas se manifeste par une tendance générale à horizontaliser tous les réflecteurs. Afin d'améliorer la capacité d'inversion par RNA-Elman, nous avons augmenté le nombre de neurones dans la couche cachée jusqu'à 90 neurones, et nous avons obtenu la répartition d'impédance acoustique illustrée par la Figure IV-07-b. Sur cette figure, on observe une nette amélioration de la réponse de la structure neuronale Elman, qui a pu trouver la majeure partie des variations latérales et verticales de l'impédance acoustique estimée.

La Figure IV-08 illustre le résultat du RNA-FBR avec 23 neurones dans la couche cachée (Figure IV-08 a). On remarque un bon ajustement des traces utilisées en apprentissage, mais dans les régions des traces non utilisées dans l'apprentissage on observe le même comportement d'horizontalisation des réflecteurs obtenu par RNA-Elman avec 5 neurones dans la couche cachée. En fait, l'apprentissage est réussi mais la capacité de généralisation est limitée. La Figure IV-08-b est obtenue avec un RNA-FBR à 190 neurones dans la couche cachée. Concernant les autres cas intermédiaires, nous avons reporté les performances obtenues dans le tableau 08.

On peut aussi constater, dans certains cas, une dégradation de la performance avec l'augmentation du nombre de neurones cachés, c'est le cas du passage de 25 neurones à 34 et à 67 neurones dans la couche cachée pour le RNA- FBR. Bien que difficile à expliquer, on peut postuler que l'utilisation de modèle d'ordre élevé causerait un problème de surentraînement. Effectivement on peut bien l'observer sur la courbe d'apprentissage du RNA-FBR avec 1200 itérations, tandis que celle du RNA-Elman continue à diminuer bien que lentement (figure IV-09). Le coefficient de corrélation indique bien le rapprochement de la distribution de l'impédance acoustique recherchée. Une augmentation du nombre de neurones améliore donc la corrélation.

On peut immédiatement constater que la répartition de l'impédance acoustique trouvée par le RNA-FBR est beaucoup plus proche de la répartition recherchée que celle calculée par le RNA-Elman.

Il faut remarquer, néanmoins, que l'augmentation du nombre de neurones dans la couche cachée du RNA-FBR apporte une amélioration sensible au niveau la cohérence la répartition

d'impédance acoustique. Ce résultat s'harmonise parfaitement avec les valeurs recherchées. En revanche, le progrès essentiel vient de la rapidité de l'algorithme et la récupération de la distribution d'impédance acoustique en amplitude et en position.

On constate donc sur cet exemple que l'amélioration de l'approche proposée (RNA) sur l'inversion sismique est significative, malgré le peu d'information ajouté. Le progrès se fait soit au niveau:

- de la rapidité des calculs, avec un apprentissage rapide et facile.
- des valeurs des impédances retrouvées après inversion, valeurs plus proches de celles de la répartition recherchée.

V-5 CONCLUSION

Dans ce chapitre, des réseaux de neurones de types FBR et Elman ont été entraînés pour l'estimation des paramètres acoustiques à partir de données sismiques.

Les résultats obtenus montrent que ces réseaux ont pu établir une relation entre les données sismiques et les paramètres acoustiques, grâce à leurs capacités d'approximation et d'adaptation. L'apprentissage du réseau peut être considéré comme un moyen de synthèse automatique d'une fonction de transfert généralement non-linéaire. Les réseaux de neurones semblent être une approche prometteuse pour la tâche définie par la problématique.

Toutefois, dans la présente étude, les données réelles disponibles étaient malheureusement limitées influençant par conséquent nos conclusions. Malgré tout, les résultats obtenus sur ces données réelles limitées, indiquent que les réseaux de neurones peuvent être un outil très efficace pour l'estimation des paramètres acoustiques à partir des données sismiques.

Conclusion, nouveauté scientifique et perspectives



Conclusion, nouveauté scientifique et perspectives

Au terme de ce travail, nous avons présenté les principes de base de la théorie du problème inverse. Il en ressort que la **stratégie de résolution** repose sur trois étapes fondamentales : la paramétrisation, modélisation, et inversion. Si le problème direct est toujours supposé bien posé, le problème inverse est toujours mal posé au sens de Hadamard. Il devient alors nécessaire de développer des méthodologies d'inversion qui visent plus précisément le traitement de ce caractère mal posé. Après avoir mis en évidence les difficultés inhérentes à la résolution des problèmes inverses, un bref rappel des méthodes classiques de régularisation a été présenté. Les différentes stratégies d'incorporation de l'information a priori servant à stabiliser la solution obtenue y ont été examinées.

Dans cette thèse, nous avons présenté un certain nombre de méthodes mathématiques et d'algorithmes pour l'inversion des données sismiques. Loin d'être exhaustif, nous avons passé en revue de nombreux travaux dans le domaine de la tomographie et de l'inversion stratigraphique. Nous avons essayé de présenter à la fois l'état de la théorie du problème inverse ainsi que les résultats mathématiques existants, et les applications pratiques sur des problèmes concrets des applications géophysiques courantes.

A propos des techniques de l'intelligence artificielle sur lesquelles sont porte les contributions essentielles de **ce travail**, nous avons procédé **par** une présentation des propriétés des ces méthodes de traitement de l'information (réseaux de neurones) classées sous le vocable de système de traitement parallèle et distribué.

Cette recherche visait à identifier les architectures les mieux adaptées pour l'inversion tomographique et l'inversion stratigraphique. Les résultats obtenus montrent que les réseaux de neurones de types FBR et Elman offrent l'avantage d'un apprentissage simple et rapide par l'application d'un algorithme de rétropropagation du gradient conjugué de l'erreur.

Après toute une série d'essais avec un réseau de type perceptron multicouches (**feedforward neural network**), de piètres résultats ont été obtenus puisque l'erreur se stabilise à 10^7 , signifiant que la fonction objective converge vers un minimum local, malgré **l'existence** d'un certain nombre de théorèmes garantissant la capacité d'approximation universelle de ce type de structure. Cependant, ces théorèmes ne fournissent pas de méthodes systématiques constructives guidant le

choix de la topologie à adopter pour une application particulière. En outre, ce type d'architecture entraîne des difficultés d'apprentissage dont les plus importantes sont :

- la présence de minima locaux dans la fonction de coût s'explique par le fait que les paramètres (poids des connexions) paraissent non linéaires dans l'expression de la sortie du réseau. La permutation indifféremment des nœuds ainsi que leurs poids correspondants dans les couches cachées tout en gardant le même comportement de sortie augmente le nombre d'optima. Les méthodes d'ajustement de type « gradient » peuvent donc conduire à un apprentissage sous-optimal, tandis que les méthodes à caractère stochastique s'avèrent plus fiables mais plus pesantes en terme temps calcul et espace mémoire.
- la possibilité de paralysie durant l'apprentissage à cause de la saturation introduite par les sigmoïdes.
- La difficulté d'introduire l'information *a priori*, due au caractère distribué et non local de la représentation.

Certains auteurs Carlos, Calderón-Macias et al [Carlos 2000] ont étudié l'apprentissage de ce type de réseau par des algorithmes hybridés.

Les réseaux de neurones que nous avons implémentés pour l'inversion des données sismiques, peuvent être placés globalement dans deux catégories différentes: les architectures statiques qui ne tiennent compte que de l'organisation spatiale tout en négligeant la "composante" temps, et les architectures temporelles qui incorporent la dimension temporelle. Dans la catégorie des architectures statiques, un réseau à fonctions de base radiales (Radial-Basis Functions network) a été implémenté. Dans la catégorie des architectures temporelles, nous avons expérimenté le réseau récurrent d'Elman. Ces RNA constituent deux classes de modèles paramétriques largement utilisées en identification de systèmes non-linéaires. En effet, ces réseaux avec une seule couche cachée peuvent approximer n'importe quelle fonction continue ayant un nombre fini de discontinuités sur tout intervalle compact.

En accord avec la méthodologie des réseaux de neurones, chaque base a été divisée en un ensemble d'apprentissage et un ensemble de validation. On a utilisé 50% des exemples pour faire l'apprentissage et les exemples restants pour la validation. Pour chaque architecture, nous avons

varié les paramètres de conception pour observer l'effet sur la performance de l'inversion. Dans le cadre de ce travail, nous avons étudié l'impact du nombre de neurones dans la couche cachée. C'est le paramètre par excellence qui permet de sélectionner à volonté la complexité intrinsèque du modèle implémenté pour une architecture de réseaux de neurones donnée.

Différents tests d'apprentissage ont été réalisés sur ces types de réseaux en utilisant des données de modèles caractérisés par un nombre de couche variable et par des formes géométriques complexes. De meilleures performances sont par voie de conséquence obtenues avec un nombre d'itérations réduit.

Les tests que nous avons effectués attestent d'une bonne réussite du passage de la réponse sismique aux paramètres acoustiques, même en présence de bruit **POURCENTAGE**. En effet, l'apprentissage de ces réseaux peut être considéré comme un moyen de synthèse automatique d'une fonction de transfert généralement non-linéaire qui permet le passage des données sismiques aux paramètres acoustiques.

Afin d'apporter une justification pratique quant à l'applicabilité de l'approche considérée, nous avons comparé nos résultats de simulation avec les méthodes classiques GC, ART et SIRT. Ces résultats indiquent que les réseaux de neurones de type FBR constituent un outil d'approximation potentiellement avantageux, ayant une puissance de reconstruction supérieure aux techniques classiques, viennent ensuite les méthodes GC et ART avec un niveau de reconstruction assez acceptable. La méthode SIRT, quant à elle, s'est révélée moins performante. D'autre part, il est important de noter que le réseau de type Elman s'avère assez conservatif en raison des difficultés liées à la procédure d'apprentissage.

Les performances des modèles proposés avec les données bruitées sont également testées. Le test de reconstruction en présence de bruit montre la capacité de la structure **RNA-FBR** à retrouver avec une précision acceptable la répartition de vitesses à partir du temps de trajet, tandis que les autres méthodes ont donné des images largement altérées présentant plusieurs artefacts. Ceci est principalement dû à la présence **de bruits et à l'instabilité** de la solution obtenue par ces méthodes.

Quant à l'inversion stratigraphique, nous avons pu obtenir un apprentissage simple et facile pour les deux structures proposées (**RNA-FBR** et **RNA-Elman**). Notons qu'en présence de bruits, le RNA de type Elman conduit bien à des résultats intéressants en terme d'inversion, alors que la

performance préalablement discutée du réseau RNA-FBR se dégrade de plus en plus au cours du processus d'inversion.

Il ressort de notre recherche, que la richesse de la base d'apprentissage est nécessaire si nous envisageons la généralisation de l'utilisation des RNA. Concernant le réseau lui-même, l'utilisation d'une couche cachée simplifie davantage la structure du réseau et s'avère suffisante pour assurer la tâche de l'estimation des paramètres acoustiques à partir des données sismiques.

En comparant l'approche neuronale avec celle de l'inversion généralisée, il en ressort les remarques suivantes :

- La technique d'inversion est caractérisée par une structure algorithmique complexe dont la formulation mathématique est très lourde à mettre en oeuvre, ce qui n'est pas le cas de l'approche neuronale, puisqu'elle est basée sur l'apprentissage.
- La technique des réseaux de neurones départit des hypothèses restrictives, qui se répercutent généralement sur la qualité des résultats obtenus.

L'intérêt porté aux réseaux de neurones se justifie par leurs propriétés et devraient permettre de dépasser les limites de l'inversion basée sur les algorithmes d'optimisation classiques tant au niveau programmation qu'au niveau calcul. Parmi ces propriétés, on mentionne le parallélisme, la capacité d'adaptation et la mémoire distribuée, provenant directement de la modélisation des propriétés du système nerveux de l'homme.

Nouveautés scientifiques et contributions :

Dans ce travail nous avons développé et implémenté une approche nouvelle dans l'inversion sismique basée sur la technique des réseaux de neurones artificiels (RNA) pour l'estimation des paramètres acoustiques à partir des données sismiques.

Le caractère non-linéaire de ces réseaux rend difficile la maîtrise du choix des paramètres intrinsèques du réseau. Cependant, les excellentes capacités d'apprentissage et de généralisation observées lors de l'inversion, permettent d'améliorer la résolution du problème inverse, quelque soit l'ordre de complexité du modèle contrairement aux méthodes classiques d'inversion. L'originalité de notre approche réside dans l'utilisation des RNA de type Elman et les RNA à Fonctions de Base Radiale (FBR) appliqués aux inversions tomographique et stratigraphique.

Contributions

Les divers algorithmes d'inversion les plus utilisés **ont été explorés** de façon à mettre en valeur les améliorations apportées par l'utilisation des réseaux de neurones proposés dans l'inversion sismique. Les principales contributions engendrées se résument **comme suit** :

Les réseaux de neurones artificiels de type FBR appliqués aux inversions tomographiques ont donné une **meilleure** amélioration de la qualité de reconstruction en comparaison avec les méthodes classiques les plus utilisées (ART, SIRT, GC).

1. Concernant l'inversion stratigraphique, nous avons pu obtenir un apprentissage simple et plus rapide pour les deux structures proposées (**RNA-FBR** et **RNA-Elman**). Notons qu'en présence de bruits, les RNA de type Elman conduisent à des résultats significatifs en termes d'inversion, alors que les performances de **type FBR** préalablement discutées se dégradent de plus en plus au cours du processus d'inversion.
2. De nombreux résultats expérimentaux et de simulation indiquent clairement l'efficacité observée dans l'utilisation de réseaux de neurones de type feedback et FBR pour la résolution du problème inverse en sismique.
3. En comparaison avec les travaux de recherche les plus **récents**, notamment ceux de Calderón-Macias et al. [Carlos 2000], les résultats que nous avons obtenus sont à notre sens très probants et présentent certains avantages à savoir :
 - Une meilleure performance avec un nombre d'itérations réduit.
 - Une meilleure capacité d'apprentissage et de généralisation de la tâche d'inversion.
 - L'utilisation d'un algorithme d'apprentissage unique d'où un gain en temps de calcul et en espace mémoire.

Perspectives

Les directions de recherche sur le problème inverse sont évidemment très nombreuses et je me bornerai à indiquer seulement quelques pistes...

Une des extensions naturelles des travaux précédents est de considérer la résolution du problème inverse dans le cadre de l'approximation acoustique, dont la paramétrisation est établie sur la base de deux paramètres acoustiques simultanément.

L'objectif est de pouvoir traiter des problèmes inverses les plus réalistes possibles. Or dans bon nombre de situations les milieux de propagation sont considérés comme élastiques ou

acoustiques, et sans hypothèses restrictives sur le type des incidences, soit encore qu'on ait une géométrie particulière ou une singularité à prendre en compte localement... Une des voies d'avenir pour repousser les limitations actuelles est d'être capable d'utiliser les réseaux de neurones pour inverser des données calculées à base d'équation d'ondes. Pour pouvoir comparer l'apport des RNA avec les méthodes d'optimisation classique utilisées dans le cadre d'inversion, nous avons terminé le code de modélisation par la méthode des différences finies la base théorique est donnée en annexe A. Reste maintenant modifier le même code pour pouvoir l'utiliser afin de surmonter la difficulté de calcul du gradient de la fonction objective lors de l'implémentation d'un algorithme de type gradient. Vu le volume de calcul important lié à la taille des grilles de discrétisation, nous proposons le passage à la programmation en parallèle.

Le code à paralléliser effectue deux tâches principales :

a. La première est la simulation numérique de la propagation des ondes dans un modèle du sous-sol pour chaque source de l'expérience sismique à l'aide d'une méthode aux différences finies implémentée dans le domaine temporel.

b. La deuxième est le calcul du gradient de la fonction coût nécessitant la modification du programme de différences finies pour le réadapter au calcul du gradient selon la formulation du modèle adjoint.

Il existe également plusieurs pistes concernant les RNA et leurs applications dans le cadre du problème inverse. Nous allons donc mener d'autres expériences avec des bases de données beaucoup plus importantes. L'avantage serait de faire une étude du comportement des modèles sur les données de puits avec les données sismiques de terrain. Il est également envisageable que d'autres architectures de réseaux de neurones puissent être expérimentées. En particulier, il serait souhaitable de chercher des architectures plus spécialement adaptées pour traiter les données sismiques avec les données de puits. L'application sur des données de simulation à base d'équation d'onde (modélisation par différences finies) n'a pu être terminée à cause de volume de calcul demandé.

Enfin, une autre direction dans laquelle la recherche peut se poursuivre concerne l'applicabilité des algorithmes génétiques pour le choix des paramètres de RNA. Etant données qu'il n'existe pas une règle systématique pour le choix du nombre de neurone, le type de connexion, la fonction d'activation (dérivable ou non) et les valeurs initiales des poids

synaptiques en début d'apprentissage, il semble que les algorithmes génétiques offrent une solution potentielle à ces problèmes **qu'il est souhaitable de tester.**

LISTE DES TRAVAUX SCIENTIFIQUES

Cette thèse a fait l'objet des travaux scientifiques suivants :

1. PUBLICATIONS INTERNATIONALES :

- ✓ Tomographic Velocity Images by Artificial Neural Networks. Journal of Engineering and Applied Sciences 2(4):775-782, 2007.
- ✓ Application of feedback connection artificial neural network to seismic data filtering. Compte rendu of Geoscience. 2008, vol. 340, p. 335-344. Published by Elsevier Masson SAS on behalf of Académie des sciences. Paris-France.
- ✓ Application of Radial Basis Function Artificial Neural Network to Seismic Data Inversion. (Submitted to Computer & Geosciences (09.08) Published by Elsevier)(en cours).

3. COMMUNICATIONS INTERNATIONALES :

- ✓ Estimation des paramètres acoustiques par des réseaux de neurones artificiels Sixth International conference on the Geology of Arab World (GAW-6), 11-15 /02/2002, Cairo, Egypt
- ✓ Tomographic Velocity Images by Artificial Neural Networks. 2nd international symposium on hydrocarbons & chemistry. GHARDAÏA-ALGERIA March 21-23, 2004.
- ✓ Application of Elman artificial neural networks to filter seismic data. 4nd international symposium on hydrocarbons & chemistry. GHARDAÏA-ALGERIA March 21-23, 2008.

4. COMMUNICATIONS NATIONALES

- ✓ Seismic data filtering by artificial neural networks. Les 6èmes Journées Scientifiques et Techniques de Sonatrach (JST 6) Hôtel Hilton, Alger, 20 – 22 Juin 2004.

REFERENCES BIBLIOGRAPHIQUES



REFERENCES BIBLIOGRAPHIQUES

1. Adde, G, Méthodes de Traitement d'Image Appliquées au Problème Inverse en Magnéto-Electro-Encéphalographie. Thèse de doctorat, Informatique, Ecole nationale des Ponts et chaussée, France, 2005, 228 p.
2. Amundsen, L, Ursin, B, Frequency-wavenumber inversion of acoustic data *Geophysics*, 1991, Vol. 56, p. 1027-1039.
3. Andrés, E, M. Predicting petrophysical properties by simultaneous inversion of seismic and reservoir engineering data. Thèse de doctorat, Géophysique, université de Stanford, 2002, 220 p.
4. Angel, C.A.M. Valentin, B. Application of radial basis neural networks for area forecast, *International Journal "Information Theories & Applications"*, 2007, vol. 14, p. 218-222.
5. Asakawa, E. Kawanaka, T. Seismic ray using linear traveltime interpolation, *Geophysical prospecting*, 1993, vol. 41, p. 99-111.
6. Backus, G.E. Gilbert, J.F. Numerical application of a formalism for geophysical inverse problems, *Journal of the Royal Astronomical Society*, 1967, vol. 13, p. 247-276.
7. Backus, G.E. Gilbert, J.F. *The resolving power of gross earth data*, *Geophysical Journal of the Royal Astronomical Society*, 1968, vol. 16, p. 169-205.
8. Bamberger, A.; Chavent, G. Hemon, C. Lailly, P. Inversion of normal incidence seismograms, *Geophysics*, 1982, vol. 47, p. 757-770.
9. Berkhout, A.J. Multidimensional linearized inversion and seismic migration, *Geophysics*, 1984, vol. 49, no. 11, p. 1881-1895.
10. Bilgeri, D. Carlini, A. Non-linear estimation of reflection coefficients from seismic data, *Geophysical Prospecting*, 1981, vol. 29, p. 672-686.
11. Bishop, T.N. Bube, K.P; Cutler, R.T; Langan, R.T; Love, P.L; Resnick, J.R; Shuey, J.T; Spindler, D.A; Wyld, H.W. Tomographic determination of velocity and depth in laterally-varying media, *Geophysics*, 1985, vol. 50, p. 903-923.
12. Bleistein, N, On the imaging of reflectors in the earth: *Geophysics*, 1987, Vol. 52, no. 7, 931-942.
13. Bleistein, N; Gray, S.H, An extension of the born inversion: method to a depth dependent reference profile, *Geophysical prospecting*, 1985, vol. 33, p. 999-1022.
14. Boschetti, F; Moresi, L. Interactive inversion in geosciences, *Geophysics*, 2001, vol. 66: p. 1226-1234.

15. Boulanger, O. Modélisation et inversion tri-dimensionnelles en prospection gravimétrique et électrique, Thèse de doctorat, Géophysique, université de Montréal, Canada, 2004, 262 p.
16. Boyer, S; Mari, J.L. Sismique et diagraphies, Université Paris VI, Editions Technip, Paris, 1994, 266 p.
17. Bregman, N.D; Bailey, R.C; Chapman, C.H. Crosshole seismic tomography, *Geophysics*, 1989, vol. 54, p. 200-215.
18. Buxton, L.R; Davison, R; Van R.P. An interpreter's guide to understanding and working with seismic-derived acoustic impedance data, *THE LEADING EDGE*, 2000, vol. 19, no. 3, p. 242-257.
19. Carlos, C.M; Sen M; Stoffa, P. Artificial neural networks for parameter estimation in geophysics, *Geophysical prospecting*, 2000, vol. 48, p.27-47.
20. Carlos, C.M; Sen M; Stoffa, P. Hopfield neural networks, and mean field annealing for seismic deconvolution and multiple attenuation: *Geophysics*, 1997, vol. 62, no. 3, p. 992–1002.
21. Cooke, D.A; Schneider, W.A. Generalized inversion of reflection seismic data, *Geophysics*, 1983, vol. 48, p. 665-676.
22. Day-Lewis, F.D; Harris, J.M; Gorelick, S.M. Time-lapse inversion of crosswell radar data, *Geophysics*, 2002, vol. 67, p. 1740-1752.
23. Devilee, R.J.R; Curtis, A; Roy-Chowdhury, K. An efficient, probabilistic neural network approach to solving inverse problems: Inverting surface wave velocities for Eurasian crustal thickness, *Journal of geophysical research*, 1999, vol. 28, p. 841–28.
24. Djarfour, N; Aifa, T; Baddari, K; Mihoubi, A; Ferahtia, J. Application of Feedback Connection Artificial Neural Network to Seismic Data Filtering. *Comptes Rendus Geoscience*, 2008, vol. 340, p. 335-344.
25. Djarfour, N; Aifa, T; Baddari, K; Mihoubi, A; Ferahtia, J. Ferahtia, J; Baddari, K. Tomographic Velocity Images by Artificial Neural Networks. *Journal of Engineering and Applied Sciences*. 2007, vol. 2, no. 4, p. 775-782.
26. Duijndam, A.J.W. Bayesian estimation in seismic inversion, part II: Uncertainty analysis, *Geophysical Prospecting*, 1988, vol. 36, p. 899-918.
27. Elman, J.L. Finding structure in time, *Cognitive Science*, 1990, vol. 14, p. 179–211.
28. Elman, J.L. Distributed Representations, Simple Recurrent Networks and Grammatical Structure, *Machine Learning*, 1991, vol. 7, no.2, 195-225.

29. Epping, W.J.M; L'ISTELLE, A.R. Data compression of seismic image by neural networks, *Revue de l'IFP*, 1992, vol. 47, no. 3, p. 423-430.
30. Gérard, D; Manuel, S; Jean-Marc, M; Mirta, B; Gordon, F.B; Sylvie T; Laurent, H. Réseaux de neurones Méthodologies et applications, Edition EYROLLES. 2004, 408 p.
31. Erwan, G. Reconstruction des images bidimensionnelles de levés de RADAR géologique de forage. Thèse de l'école polytechnique de Montréal (canada). 2004, 204 P.
32. Fara, V. Approche hamiltonienne de la théorie des rais : Inversion des temps de parcours en sismique réflexion, (thèse), Edition Technip, 1987, 215 p.
33. Faure, P. Analyse numérique notes d'optimisation, Ellipses, Ecole Polytechnique, Paris, 1988, 175 p.
34. Gjoystdal, H; Ursin, B. Inversion of reflection times in three dimensions, *Geophysics*, 1981, vol. 46, p. 972-983.
35. Glangeaud, F; Mari, J.L; Mars, J; Lacoume, J.L. Traitement de signal pour géologue et géophysicien, Edition Technip, 2004, 472 p.
36. Gondran, M; Muller, C. Réseaux de neurones et apprentissage, *Revue de l'électricité et de l'électronique (REE)*, 1997, vol. 3, p. 6-9.
37. Groof, P.F. Neural networks experiments on synthetic seismic data, *Artificial intelligence in the petroleum industry*, Edition Technip, 1992, p. 93-124.
38. Gualtiero, B; Aldo, L; Vesnaver. In quest of the grid, *Geophysics*, 1999, vol. 64, no. 4, p. 1116-1125.
39. Hanyga, A; Pajchel, J. Point-to-curve ray tracing in complicated geological model. *Geophysical prospecting*, 1995, vol. 43, p. 859-872.
40. Haykin, S; Liang, L. Nonlinear adaptive prediction of non stationary signals. *IEEE Transactions on Signal Processing*, 1995, vol. 43, no. 2, p. 526-535.
41. Henry, G., La sismique réflexion Principe et développement, Edition Technip, 1997, 172 p.
42. Hertz, J; Krogh, A; Palmer, R.G., Introduction to the theory of neural computation. Addison-Wesley, Reading, Massachusetts, 1991, 328 p.
43. Hirahara, K; Iyer, H.M. Seismic Tomography: Theory and Practice, Edition Kluwer Academic Publishers, 1993, 865 p.
44. Hyndman, D.W; Harris, J.M. Travelttime inversion for the geometry of aquifer lithologies, *Geophysics*, 1996, vol. 61, p. 1728-1737.

45. Jean-Michel, R. Algorithmes génétiques et réseaux de neurones, Edition Hermes Sciences Publicat, 1994, 334 p.
46. Jupp, D.L.B. et Vozoff, K. Stable iterative methods for the inversion of geophysical data, *Geophysical Journal of the Royal Astronomical Society*, 1975, vol. 42, p. 957-976.
47. Kern, M., Problèmes inverses, thèse de doctorat, Géophysique, Ecole supérieur d'ingénieurs Léonard de Vinci, 2003, 140 p.
48. Kormendi, F. & Dietrich, M. Non-linear waveform inversion of plane-wave seismograms in stratified elastic media. *Geophysics*, 1991, Vol. 56, p. 664–674.
49. Krajewski, C; Dresen, L; Gelbke, C; Ruter, H. Iterative tomographic methods to locate seismic low-velocity anomalies: a model study. *Geophysical Prospecting*, 1989, vol. 37, p. 717-75 1.
50. Landa, E; Kosloff, D. A method for determination of velocity and depth from seismic reflection data, *Geophysical prospecting*, 1988, vol. 36, p.223-243.
51. Lascaux, P; Theodor, R. Analyse numérique matricielle appliquée à l'art de l'ingénieur : méthodes directes, Tome 1, Edition Dunod, 2000, 326 p.
52. Lascaux, P; Theodor, R. Analyse numérique matricielle appliquée à l'art de l'ingénieur : méthodes itératives, Tome 2, Edition Dunod, 2004, 636 p.
53. Levy, S; Clowes, R.M. Debubbling. A generalised linear inversion approach. *Geophysical prospecting*, 1980, vol. 28, p. 840-858.
54. Liao, Q; et McMechan, G.A. Tomographic imaging of velocity and Q, with application to crosswell seismic data from the Gypsy Pilot Site, Oklahoma, *Geophysics*, 1997, vol. 62, p. 1804-1811.
55. Lines, L.R; Ulrych, T.J. The old and the new in seismic deconvolution and wavelet estimation, *Geophysical Prospecting*, 1977, vol. 25, p. 512-540.
56. Mari, J.L; Arens, G; Chapellier, D; Gaudiani, P. Géophysique de gisement et de génie civil, Edition Technip, 1998, 468 p.
57. Mason, I.M. Algebraic reconstruction of a 2-D velocity in homogeneity in the High Hazels Seam of Thoresby Colliery, *Geophysics*, 1981, vol. 46, p. 298–308.
58. Moraes, F.S; Scales, J.A. Local Bayesian inversion: theoretical developments, *Geophysical Journal International*, 2000, vol. 141, p. 713-723.
59. Mora, P.R. Nonlinear two-dimensional elastic inversion of multioffset seismic data; *Geophysics*, 1987, vol. 52, p. 121 1-1228.

60. Mousset, E. Spikes filtering with neural networks: a Two-Stage Detection System, *Revue de l'IFP*, 1992, 4 vol. 7, no. 3, p. 407-421.
61. Nemeth, T.E; Normark; Qin, F. Dynamic smoothing in crosswell travelttime tomography, *Geophysics*, 1997, vol. 62, p. 168-176.
62. Pan, G.S.; Phinney, R.A.; Odom, R.I. Full-waveform inversion of plaewave. seismograms in stratified acoustic media: Theory and feasibility. *Geophysics* 1988, vol. 53, p. 21-31.
63. Perrodon, A. Géodynamique pétrolière, Edition MASSON, 2002, 388 p.
64. Peterson, J.E; Paulsson, B.N.P; McEvelly, T.V. Applications of algebraic reconstruction techniques to crosshole seismic data, *Geophysics*, 1985, vol. 50, p. 1566-1580.
65. Pierrick, C. Inversion gravimétrique tridimensionnelle utilisant un modèle de covariance des paramètres, Thèse de école polytechnique de Montréal, 2003, 113p.
66. Scales, J.A. Tomographic inversion via the conjugate gradient method, *Geophysics*, 1987, vol. 52, p. 179-185.
67. Stewart, R. Exploration Seismic Tomography: Fundamentals, Society of Exploration Geophysicists, Course Notes Series, 1991, 140 p.
68. Seghouane, A; Fleury, G. Learning radial basis function neural networks with noisy input-output data set, *Traitement du Signal*, 2003, vol. 20, no. 4, p. 213-221.
69. Shang, G; Young, C.Y; Castagna, J.P. An integrated target-oriented prestack elastic waveform inversion: Sensitivity, calibration, and application, *Geophysics*, 1994, vol. 59, no. 9, p.1392-1404.
70. Sheriff, R.E, et Geldart, L.P. *Traité de prospection sismique*, Edition ERG, 1984, 371 p.
71. Snieder, R; Trampert, J. *Inverse Problems in Geophysics*, Samizdat Press, 1999, 200 p.
72. Snieder, R; Trampert, J. The role of nonlinearity in inverse problems, *Inverse Problems*, 1998, vol. 14, p. 387-404.
73. Snieder, R; Trampert, J. An extension of Backus-Gilbert theory to nonlinear inverse problems, *Inverse Problems*, 1991, vol. 7, p. 409-433.
74. Stefani, J.P. Turning-ray tomography, *Geophysics*, 1995, vol. 60, p. 1917-1929.
75. Stewart, R. R., *Exploration Seismic Tomography: Fundamentals*, Society of Exploration Geophysicists, Course Notes Series, 1991, 3, 140 p.

76. Symes, W; Kern, M. Inversion of reflection seismograms by differential semblance analysis: algorithm structure and synthetic examples. *Geophysical prospecting*, 1994, vol. 42, p. 565-614.
77. Tarantola, A. Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, 1984a, vol. 49, no. 8, p. 1259-1266.
78. Tarantola, A. Linearized inversion of seismic reflection data. *Geophysical prospecting*, 1984b, vol. 32, p. 998-1015.
79. Tarantola, A. Inverse problem theory: Methods for Data Fitting and Parameter Estimation, Amsterdam: Elsevier, 1987, 613 p.
80. Zhang, J; McMechan, G.A. 3-D transmission tomography using wide aperture data for velocity estimation of irregular salt bodies, *Geophysics*, 1994, vol. 59, p. 1620-1630.
81. Zhou, C.L; Lane, J.W. Simultaneous inversion for velocity and attenuation, *Journal of Applied Geophysics*, 2001, vol. 47, p. 199-215.

ANNEXE A



ANNEXE A

INVERSION NON LINEAIRE

Nous abordons maintenant les problèmes inverses non-linéaires, avec une concentration particulière sur l'identification de paramètres dans les équations aux dérivées partielles.

A -1 LE PROBLEME DIRECT

Il existe plusieurs modèles physiques pouvant rendre compte (avec des degrés divers d'approximation) de l'expérience de sismique réflexion. Nous nous limitons à étudier l'un des plus simples. Dans ce cas, on peut démontrer que la propagation des ondes est régie par l'équation des ondes acoustiques, et la quantité mesurée est un champ (scalaire) de pression donné par l'expression suivante [Shttevelman 1984] : *cette réf. n'existe pas dans la liste*

$$\frac{\partial^2 u(x, z, t)}{\partial t^2} = C^2(x, z) \nabla^2 u(x, z, t) + f(x, z, t) \quad \text{A-10}$$

$$u(x, z, 0) = \frac{\partial u}{\partial t}(x, z, 0) = 0 \quad \text{A-11}$$

$$\frac{\partial u}{\partial t}(x, 0, t) = 0 \quad \text{A-12}$$

avec

$(x, z) \in \Omega$ ouvert de \mathbb{R}^2

$t \in [0, T]$

$u = u(x, z, t)$ pression, inconnue du problème

$C = C(x, z)$ champs de vitesses de propagation d'onde acoustique

$f = f(x, z, t)$ densité de force externe.

Et ∇^2 est le Laplacien donné par la relation suivante : $\nabla^2 u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial z^2}$

Le équation A-10 décrit la propagation des ondes acoustiques, la seconde A-11 est l'état initial du système (ici supposé au repos). La troisième A-12 est une condition aux limites, ici de surface libre sur la surface du sol. L'ensemble de ces équations forme le problème direct : qui s'énonce comme suit : connaissant la répartition de la vitesse C ; le sismogramme $P(x, 0, t) = f(x, t)$ peut être calculé à partir des équations (A-10, A-11, A-12) en utilisant la méthode des différences finies.

Ce modèle mathématique correspond à la formulation en pression des équations de l'élastodynamique linéaire pour des matériaux élastiques linéaires isotropes à densité constante (vérifiant l'hypothèse des petites déformations).

Le milieu est alors excité par une source ponctuelle par exemple, une explosion à support limité dans le temps. Il s'agit de densité de force à support ponctuel que l'on modélise par

$$f(x,z,t)=sr(t)*\delta(x-x_s,z-z_s) \quad \text{A-13}$$

où $Sr(t)$ représente la signature temporelle de la source utilisée. Dans nos applications, une fonction Ricker, dérivée seconde d'une gaussienne, est employée pour modéliser la signature d'une explosion.

La méthode des différences finies consiste à approcher les dérivées partielles par des différences finies, ceci permet d'une part de remplacer une équation mathématique aux dérivées partielles par un schéma numérique qui réalise une approximation de l'équation, et d'autre part la résoudre numériquement. Cette méthode représente un outil efficace pour résoudre des équations aux dérivées partielles. Elle est particulièrement intéressante dans le cas où la solution analytique n'est pas encore connue.

A -1-a)- Schémas en espace

L'approximation s'effectue en deux étapes successives. La première consiste à discrétiser le domaine Ω ; pour cela, n et m étant deux nombres entiers strictement positifs donnés. On découpe le domaine Ω en $(n+1)$ et $(m+1)$ parties de longueurs h_x et h_z , telles que $h_x = \frac{x_{\max}}{n+1}$ et

$h_z = \frac{z_{\max}}{m+1}$; h_x et h_z sont les pas de discrétisation dans la directions x et z . On considère alors le

maillage par les $(n+2) \times (m+2)$ points, $x_i = ih_x$ $i=0,1,\dots,n,n+1$; $z_j = jh_z$ $j=0,1,\dots,m,m+1$; on cherche une approximation de $u(x,z)$ que l'on note $u(x_i,z_j) = p_{ij}$

La seconde étape consiste à discrétiser les opérateurs ; à ce stade, on remplace les dérivées par des quotients différentiels faisant intervenir les valeurs de la pression $p(x,z)$ inconnue aux points du maillage. On considère les développements en série de Taylor limités à l'ordre 2, et on cherche un schéma aux Différences Finies centré qui approche l'opérateur ∇^2 à l'ordre 2. Posons $\nabla^2 p_{ij}$ une valeur approchée à l'ordre 2 de $\nabla^2 p(x,z)$ au point (i,j) . En utilisant le développement de Taylor d'ordre 2, on trouve l'approximation de l'opérateur de Laplace suivante [Young et al. 1984] :

$$\nabla^2 u_{ij} = \frac{u_{i+1,j} - 2u_{ij} + u_{i-1,j}}{h_x^2} + \frac{u_{ij+1} - 2u_{ij} + u_{ij}}{h_z^2} \quad \text{A-14}$$

A -1-b)- Schémas en temps

Soit $t \in [1T]$ sur lequel on définit un maillage uniforme de pas h_t . On note :

$$t = kh_t \quad k \in \mathbb{Z}$$

$u(t) = u(kh_t)$ pour toute fonction p

Le schéma le plus naturel pour approcher l'opérateur en temps est le schéma centré d'ordre 2 qui s'écrit, comme en espace :

$$\frac{\partial^2 u}{\partial t^2} = \frac{u^{k+1} - 2u^k + u^{k-1}}{h_t^2} \quad \text{A-15}$$

En additionnant les deux expressions précédentes, on obtient finalement l'approximation en différences finies de l'équation d'onde:

$$u_{ij}^{k+1} = 2u_{ij}^k - u_{ij}^{k-1} + \frac{C_{ij}^2 h_t^2}{h^2} (u_{i+1j}^k + u_{i-1j}^k + u_{ij+1}^k + u_{ij-1}^k - 4u_{ij}^k) + s^k \quad \text{A-16}$$

Pour réduire le temps de calcul, on suppose $h_x = h_z = h$, le champs de vitesse discrétise $c(x,z)=c(ih,jh)=C_{ij}$ ainsi que la source est représenté par s^k indice de **tems. ??**

A -1-c)- Stabilité du schéma aux différences finies

Lors de l'approximation de l'équation d'onde par le schéma aux différences finies, la valeur du champ est déterminée à partir des deux étapes précédentes dans le temps. La vitesse de propagation de l'information dans le maillage est bien inférieure à la vitesse de discrétisation qui est égale h/h_t ,

Le pas h_t doit être alors limité, la condition de stabilité est donnée par la formule suivante ;

$$\frac{ch_t}{h} \leq \frac{1}{\sqrt{2}} \quad \text{A-17}$$

le pas maximum en temps est délimité par cette condition de stabilité

$$h_t \leq \frac{h}{c\sqrt{2}} \quad \text{A-18}$$

A -2 LE PROBLÈME INVERSE :

L'opérateur de modélisation (problème directe) est :

$$G : (C) \rightarrow u(x,z,t) \text{ pour } x \in X_g \text{ et } t \in [0,\pi].$$

Le problème inverse consiste à déterminer c , et f (qu'il n'est pas réaliste de supposer connu) à partir des mesures ; c'est-à-dire de la connaissance de p . Ce problème est non-linéaire, car la solution p dépend de façon non-linéaire de c même si l'équation aux dérivées partielles précédente est linéaire.

Dans la réalité, il y a un paramètre supplémentaire : l'acquisition des données est répétée en déplaçant le «dispositif» source–récepteur. L'ensemble de ces «tirs» fournit une immense quantité de données (couverture multiple). Ce problème inverse est surdéterminé.

On définit une fonction de coût J mesurant la distance entre la solution de (A-16) associée à M et les observations à D par :

$$J(C) = \sum_{s \in R_s} \|G(C) - u_s^{ob}\|^2 \quad \text{A-19}$$

Le problème inverse consiste à résoudre le problème de contrôle optimal suivant :

Trouver le modèle (C_{opt}) qui minimise la fonctionnelle :

$$J(C_{opt}) = \inf_{C \in R^N} J(C) \quad \text{A-20}$$

Une condition nécessaire pour que C_{opt} soit une solution est donnée par l'équation d'Euler-Lagrange :

$$\nabla J(C_{opt}) = 0 \quad \text{A-21}$$

L'inversion de données apparaît alors simplement comme un problème d'optimisation. On cherche C_{opt} réalisant le minimum de J .

Pour résoudre un tel problème d'optimisation on utilise une méthode numérique de descente : étant donnée une première approximation C^0 , on construit une suite définie par :

$$C^{n+1} = C^n + \lambda^n G^n \quad \text{A-22}$$

G^n est la direction de descente et λ^n le pas de la descente. En général, on utilise des méthodes de type gradient, gradient conjugué ou quasi-Newton pour lesquelles la détermination de G^n nécessite le calcul du gradient de la fonction du coût $\Delta J(C_n)$. Ce calcul du gradient est un point clef de ce type d'inversion de données sismiques. Les outils mathématiques et numériques essentiels permettant d'obtenir efficacement le gradient sont ce que l'on appelle le modèle adjoint[Kern 2003] et le code adjoint (certains utilisent même la terminologie de "méthode adjointe")

A-3 LA TECHNIQUE DE L'ETAT ADJOINT

La technique de l'état adjoint est couramment utilisée en optimisation pour calculer le gradient d'une fonctionnelle non linéaire. Cette technique consiste à calculer un champ auxiliaire (état adjoint) par la résolution de l'équation adjointe. Cet état adjoint représente directement le gradient de la fonction objective. Par exemple, considérons le problème des moindres carrés suivant :

Trouver le minimum c de la fonctionnelle

$$J(C) = \sum_{s \in R_s} \|G(C) - u_s^{ob}\|^2 \quad A-23$$

Où F est un opérateur non linéaire d'un espace d'Hilbert du modèle M vers l'espace d'Hilbert de données D :

$$\begin{aligned} G: M &\rightarrow D \\ C &\rightarrow G(C) = u \end{aligned} \quad A-24$$

Pour trouver le minimum C en utilisant les techniques d'optimisation locale, on a besoin de calculer le gradient de la fonctionnelle J . la dérivation de J au point C dans la direction δC est donnée par la relation suivante :

$$J'(C) \cdot \delta C = (G'(C) \cdot \delta C, G(C) - u^{obs})_D \quad A-25$$

Où (\cdot, \cdot) est le produit scalaire en D . La dérivé $G'(C)$ de la fonctionnelle $G(C)$ au point C est un opérateur linéaire de M à D . Si on suppose que $G'(C)$ est une application subjective, alors celle-ci admet un inverse $L = (G'(C))^{-1}$ défini comme suivant :

$$\begin{aligned} L: D &\rightarrow M \\ \delta u &\rightarrow L\delta u = \delta C \end{aligned} \quad A-26$$

Maintenant on introduit un nouveau champ W solution de l'équation adjointe :

$$L^*W = F(C) - u^{obs} \quad A-27$$

Le problème revient à faire apparaître la linéarité de $J(C)$ par rapport à C , afin d'obtenir une expression de $\Delta J(C)$. L'équation peut être réécrite comme suit:

$$\begin{aligned} J'(C) \cdot \delta C &= (G'(C) \cdot \delta C, G(C) - u^{obs})_D & A-28 \\ &= (G'(C) \cdot \delta C, L^*W)_D \\ &= (LG'(C) \cdot \delta C, W)_M \\ &= (\delta C, W)_M \end{aligned}$$

Où $(\cdot, \cdot)_M$ est le produit scalaire dans l'espace M . Puisque le gradient Gr de J est défini comme élément de M alors

$$J'(C) \delta C = (\delta C, Gr)_M \quad A-29$$

avec $Gr = W$.

El est claire, d'après l'équation que nous obtenons le gradient complet de J s'obtient par la résolution de la seule équation adjointe. Ceci implique que la méthode de l'état adjoint permet de calculer le gradient à un coût proportionnel à celui du calcul de la fonction elle-même.

Soit W est la solution de l'équation du modèle adjoint suivante [Tarantola 1987]:

$$\frac{\partial^2 W(x, z, t)}{\partial t^2} - C^2(x, z) \nabla^2 W(x, z, t) = G(C) - u^{obs} \quad \text{A-30}$$

$$W(x, z, T) = \frac{\partial W}{\partial t}(x, z, T) = 0 \quad \text{A-31}$$

$$\frac{\partial W}{\partial t}(x, Z_{\max}, t) = 0 \quad \text{A-32}$$

Ces équations décrivent la propagation du champ adjoint. Elles sont intégrées de manière rétrograde en temps.

Dans ce cas, les étapes de mise en œuvre de l'algorithme du gradient dans ce cas sont les suivantes :

Etant donné à l'itération k : la solution C_k , le sismogramme $G(C_k)$, la solution à l'itération $k-1$, la solution C_{k-1} , le sismogramme $G(C_{k-1})$, et les réels α_k et β_k

- Calculer le coût $J(C_k)$ à partir de $G(C_k)$.
- Calculer le gradient de $Gr(C_k)$ par la résolution de l'équation adjointe.
- Calculer de α_k et β_k
- Définir la solution à l'itération $k+1$
- Définition du sismogramme de la solution à l'itération $k+1$,
- Test d'arrêt

Différences

A-4 RESULTAS DE LA MODELISATION PAR LA METHODE DES DIFFERENCES FINIES :

Ce genre de modélisation de la propagation des ondes demande des ressources de calcul importantes du fait de la grande taille des grilles dans notre programme, nous avons utilisé un schéma d'ordre 2 en espace, et un maillage avec un pas de 1 mètres afin de minimiser la dispersion numérique du signal (pour une fréquence centrale de la source de 35 Hz). Les domaines typiques que l'on cherche à calculer ont pour taille 100 m 100m, soit un maillage de 10000 points (figure A-01). Deux types d'informations sont intéressants à observer :

1. Les snapshots : cela consiste à prendre une photographie instantanée du champ de pression dans le sous-sol et à observer son évolution au cours du temps (figures A-02, 03, 04, 05).
2. Les sismogrammes : c'est l'enregistrement du signal au cours du temps en un point donné de l'espace (par exemple à la surface du sol). Chaque signal forme ce que l'on appelle une trace sismique. L'ensemble des traces forme le sismogramme (figure A-06).

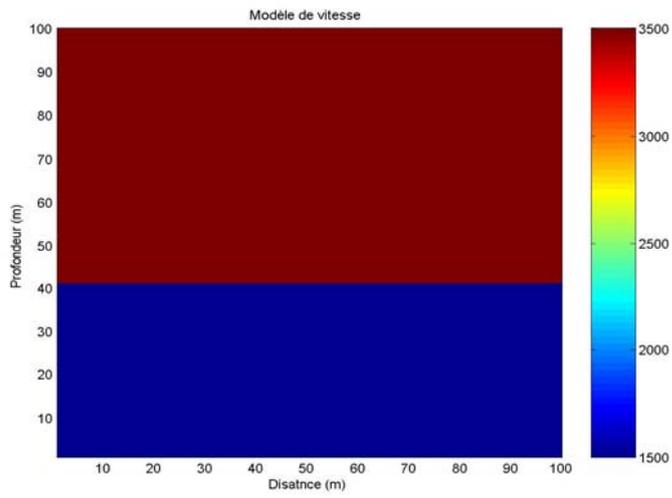


Figure A-01 Modèle de vitesse proposé pour la simulation

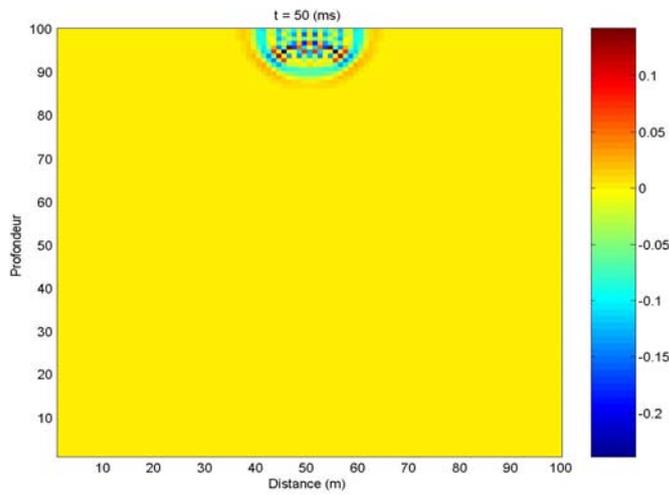


Figure A-02 Snapshot à l'instant 50ms.

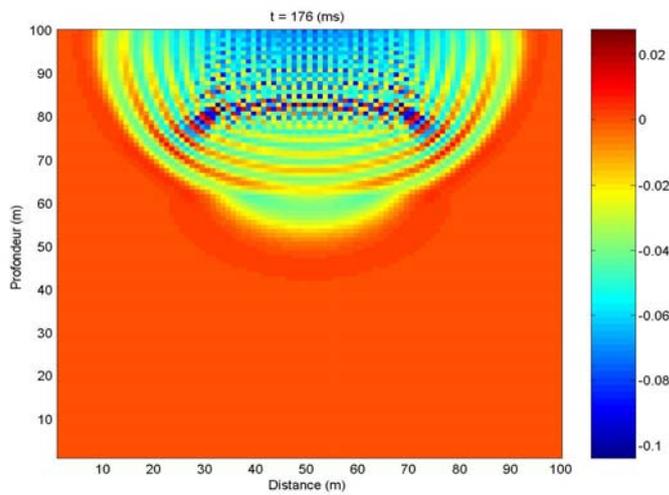


Figure A-03 Snapshot à l'instant 176 ms.

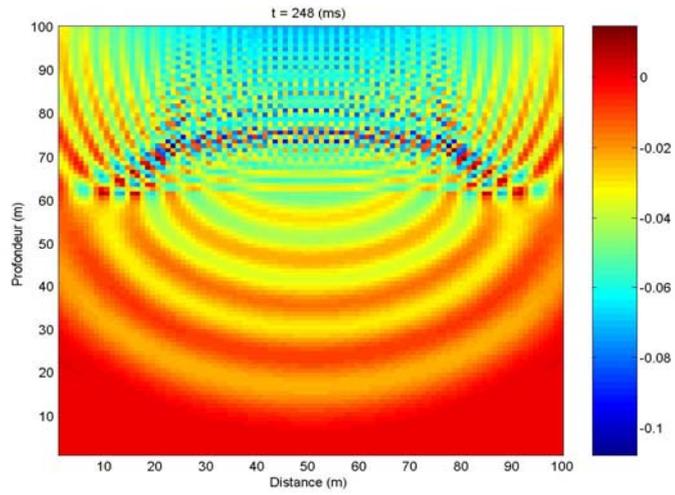


Figure A-04 Snapshot à l'instant 248ms.

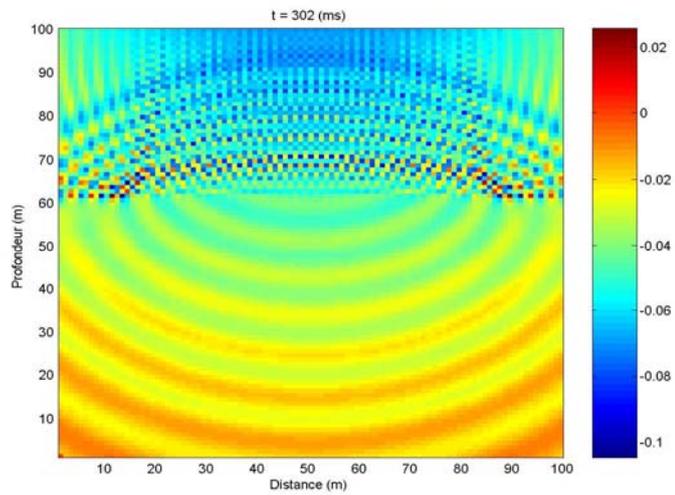


Figure A-05 Snapshot à l'instant 202ms.

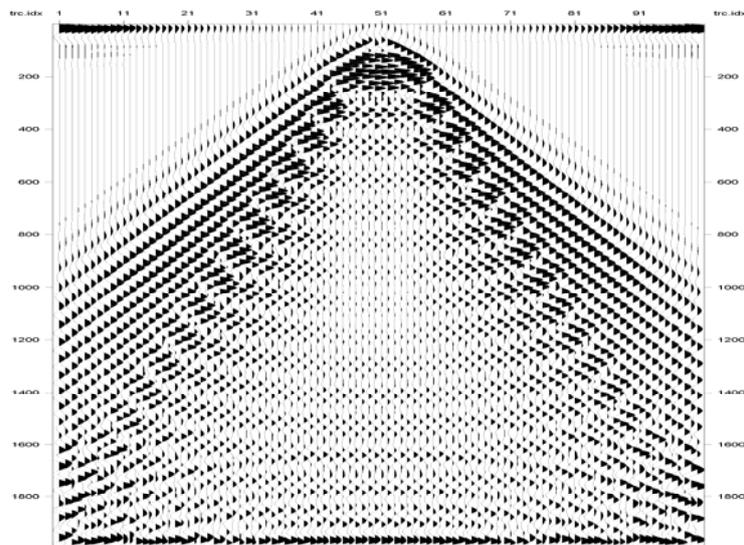


Figure A-06 Sismogramme synthétique.

ANNEXE B



ANNEXE B

LES PRINCIPAUX ALGORITHMES DEVELOPPES DANS LE CADRE DE CETTE THESE :

Les étapes de l'algorithme de rétropropagation:

Etape 01- Lecture des données d'apprentissage et de généralisation, (pour les données réelles conversion des données du format SEG Y au format MAT). généralement la taille des données d'apprentissage est supérieure à celle des données de tests. Ces données sont normalisées pour qu'elles deviennent dans l'ordre de la sortie d'une fonction de transfert de type sigmoïde.

Etape 02- On doit décider du nombre de couches intermédiaires ou cachées. Sans couche cachée, le réseau n'offre que de faibles possibilités d'adaptation; avec une couche cachée, il est capable, avec un nombre suffisant de neurones, d'approximer toute fonction continue.

Etape 03- Déterminer le nombre de neurones par couches cachées, le type de la fonction d'activation de chaque neurone.

Etape 04- Initialiser les poids de toutes les couches (cachée, contexte, sortie). On attribue à chaque poids une valeur aléatoire comprise entre -1 et +1, ce que détermine les points de départ dans la surface d'erreur pour les méthodes de descente du gradient, chose qui influe grandement sur la convergence de l'algorithme d'apprentissage.

Etape 05- Propager les signaux d'activation des entrées vers les sorties. On calcule les sorties de la première couche qui deviennent les entrées de la couche suivante jusqu'à la dernière couche.

Etape 06- Comparer les sorties du réseau avec les sorties désirées et calculer l'erreur associée à cet exemple, présenter un autre couple (entrée, sortie désirée), jusqu'à ce que tout l'ensemble des couples sont présenté,

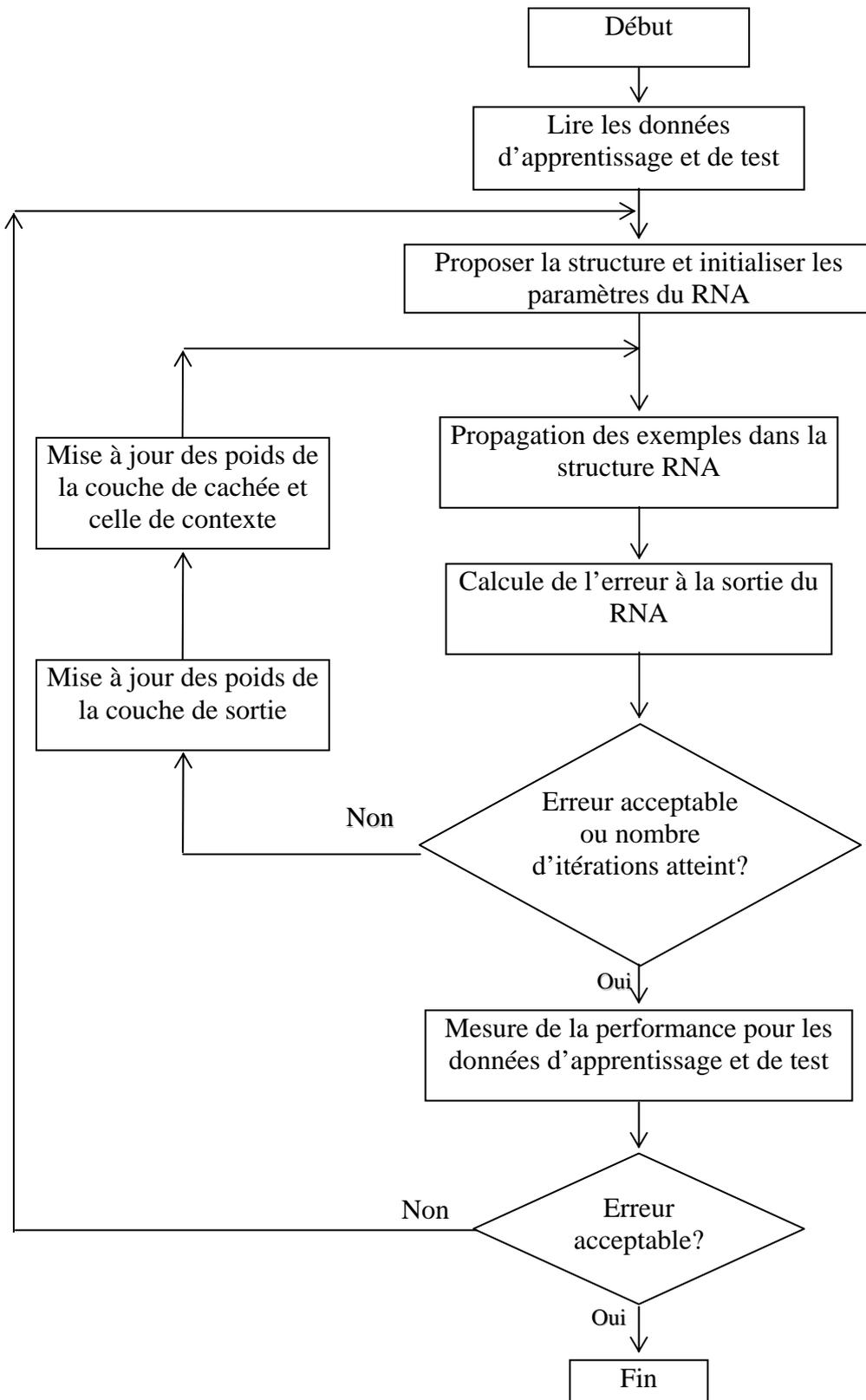
Etape 07- calculer le coût global

Si l'erreur globale est inférieure à un seuil donné ou le nombre d'itération est atteint
Allez à l'étape 08

Si non : On met à jour les poids dans toutes les couches, allez à l'étape 05

Etape 08- Mesure de la performance du RNA pour les données d'apprentissage et de test.

Re- entraîner si nécessaire



ORGANIGRAMME D'APPRENTISSAGE DU RNA-ELMAN

L'algorithme d'apprentissage du RNA-RBF:

Étape 1. Lecture des données d'apprentissage et de généralisation, (pour les données réelles conversion des données du format SEG Y au format MAT), généralement la taille des données d'apprentissage est supérieure à celle des données de tests. Ces données sont normalisées pour qu'elles deviennent dans l'ordre de la sortie d'une fonction de transfert de type sigmoïde.

Étape 2. Simulation et initialisation du RNA

➤ Pour la couche cachée

On définit les n centres C_j , chaque centre est défini comme le centre de gravité des points x_i

Détermination des rayons de la fonction d'activation Gaussienne, chaque centre est la racine carré de la distance maximum inter centres

➤ Pour la couche de sortie, on attribue à chaque poids une valeur aléatoire comprise entre -1 et +1.

Étape 3. Itérer jusqu'à la convergence

➤ Pour la couche cachée

- Un neurone à base radiale avec nombre de centres égale au nombre des données, jusqu'au nombre maximum de neurones fixé à l'avance

➤ Pour la couche de sortie

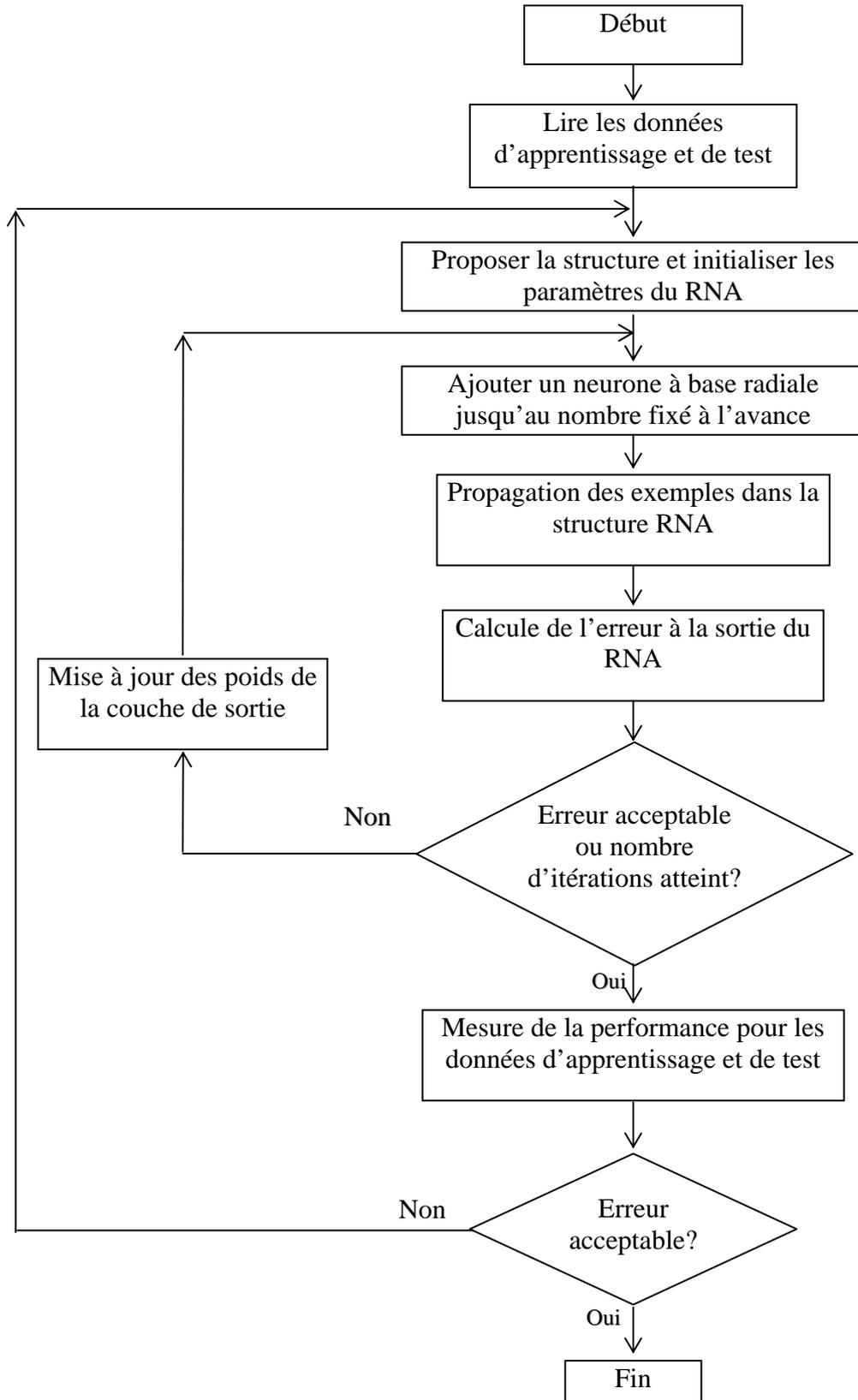
- La mise à jour des poids de connexion selon la descente de gradient.

Étape 4. Calcul de l'erreur entre la sortie du RNA et la sortie désirée :

- Si l'erreur à la sortie du RNA est supérieure à l'erreur tolérée ou le nombre d'itération n'est pas atteint, aller à l'étape 3.
- Si l'erreur à la sortie du RNA est inférieure à l'erreur tolérée ou le nombre d'itération est atteint, aller à l'étape 5.

Étape 5. Mesurer la performance du RNA par rapport aux données d'apprentissage et de généralisation.

- Re-entériner si nécessaire.



ORGANIGRAMME D'APPRENTISSAGE DU RNA-RBF