

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



UNIVERSITE M'HAMED BOUGARA DE BOUMERDES
FACULTÉ DES HYDROCARBURES ET DE LA CHIMIE

Laboratoire de Physique de la Terre (LABOPHYT)



THÈSE DE DOCTORAT

Présentée par

Jalal FERAHTIA

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

Option : Sismique, sismologie et méthodes de potentiels

Thème

Traitement de données géophysiques à base de filtres non- linéaires et adaptatifs.

Application des filtres : de diffusion anisotropique, trilatéral et SD-ROM à l'atténuation des bruits cohérents et aléatoires.

Devant le Jury composé de :

Mabrouk	DJEDDI	Professeur (Université M'hamed Bouguara de Boumerdès)	Président
Kamel	BADDARI	Professeur (Université M'hamed Bouguara de Boumerdès)	Rapporteur
Hocine	SHOUT	Professeur (Université de Constantine)	Examineur
Tahar	AIFA	Maître de conférences habilité à diriger des recherches (Université de Rennes 1-France)	Examineur
Abd rezak	LACHOURI	Maître de conférences Rang A (Université de Skikda)	Examineur



بسم الله الرحمن الرحيم

وقل ربّ زدني علما

صدق الله العظيم

Table des matières

Introduction générale.....	1
1. Etat de l'art dans le domaine du filtrage en géophysique	
1.1 Introduction	6
1.2 Les bruits en géophysique.....	6
1.2.1 Cas de la sismique.....	6
1.3 Méthodes d'atténuation des bruits en sismique.....	7
1.3.1 Méthodes conventionnelles.....	7
1.3.1.1 Limitations.....	8
1.3.2 Méthodes basées sur l'intelligence artificielle.....	8
1.3.2.1 Limitations des RNA.....	9
1.4 Cas de la méthode électrique à courant continu.....	10
1.4.1 Historique de l'imagerie électrique.....	10
1.4.2 Dispositifs utilisés.....	10
1.4.3 Représentation des données.....	10
1.4.4 Les bruits en prospection électrique.....	11
1.4.5 Techniques de filtrage en prospection électrique.....	11
1.5 Notions sur les images.....	13
1.5.1 Définition d'une image numérique.....	13
1.5.2 Qu'est ce qu'un traitement d'images ?.....	14
1.5.3 Thèmes du traitement de l'image.....	14
1.5.4 Opérations de base en traitement de l'image.....	14
1.5.4.1 Traitement de voisinage.....	15
1.5.5 Types de bruits en imagerie.....	16
1.5.6 Filtrage de bruits en imagerie.....	16
2. Théorie des filtres : Diffusion anisotropique, Trilatéral et SD-ROM	
2.1 Filtre de diffusion anisotropique.....	18
2.1.1 Bases théoriques	18
2.1.2 Filtrage par convolution.....	18
2.1.3 Analyse de l'EDP.....	19
2.1.3.1 Cas non linéaire.....	20
2.1.4 Approximation discrète par différence finie.....	21
2.1.5 Remarques sur la stabilité.....	22
2.1.6 Description de l'algorithme de filtrage par diffusion anisotropique	23
2.1.7 Algorithme du filtre anisotropique.....	24
2.2 Filtre Trilatéral	
2.2.1 Introduction.....	25
2.2.2 Arrière plan théorique.....	25
2.2.3 Principe du filtre bilatéral.....	25
2.2.4 Formulation mathématique du filtre bilatéral.....	26
2.2.5 Etude des paramètres seuils.....	27

2.2.6 Formulation du filtre bilatéral.....	27
2.2.7 Définition de la statistique d'ordre.....	28
2.2.8 Définition de la statistique ROAD.....	28
2.2.9 Le filtre trilatéral.....	29
2.2.10 Version modifiée du filtre trilatéral.....	30
2.2.11 Algorithme du filtre trilatéral.....	31
2.3 Filtre SD-ROM	
2.3.1 Introduction	32
2.3.2 Théorie du filtre SD-ROM.....	32
2.3.3 Description de l'action du filtre.....	32
2.3.4 Importance de la taille de la fenêtre.....	33
2.3.5 Algorithme du filtre SD-ROM.....	35
2.3.6 Organigramme du filtre SD-ROM.....	36
3. Atténuation des bruits par le filtre de diffusion anisotropique	
3.1 Implémentation du filtre de diffusion anisotropique.....	37
3.1.1 Données synthétiques.....	37
3.1.1.1 Description du modèle.....	37
3.1.1.2 Paramétrisation.....	38
3.1.1.3 Résultats de simulations et discussions.....	40
a) Etude qualitative.....	40
b) Comparaison des sections différences.....	41
c) Investigation dans le domaine spectrale.....	42
3.1.1.4 Influence du nombre d'itérations.....	44
3.1.1.5 Description des sections différences.....	49
3.1.1.6 Restitution de données perdues.....	51
3.1.2 Application sur des données sismiques réelles.....	54
3.1.2.1 Présentation des données.....	54
3.1.2.2 Paramétrisation.....	55
3.1.2.3 Analyses et résultats.....	55
3.2 Conclusion.....	57
4. Atténuation du bruit sismique par le filtre trilatéral	
4.1 Introduction.....	59
4.2 Simulations.....	59
4.2.1 Données synthétiques.....	59
4.2.2 Résultats et commentaires.....	61
4.2.3 Influence du nombre d'itérations.....	63
4.2.4 Analyse spectrale.....	67
4.2.5 Comparaison avec les filtres usuels.....	70
4.3 Application sur des données réelles.....	73
4.3.1 Stratégie proposée pour l'élimination des bruits.....	73
4.3.2 Description des données réelles	73
4.3.3 Résultats du filtrage sur les données réelles.....	74
4.4 Etude comparative avec le filtre anisotropique.....	77
4.4.1 Cas des données synthétiques	77
4.4.2 Cas des données réelles.....	84
4.5 Conclusion.....	86

5. Application du filtre SD-ROM aux données de prospection électrique	
5.1 Simulations.....	88
5.1.1. Description du modèle synthétique.....	88
5.1.2 Résultats du filtrage sur les résistivités apparentes calculées.....	90
5.1.3 Effets du filtrage sur le processus d'inversion.....	91
5.2 Application du filtre SD-ROM aux données réelles.....	93
5.2.1 Description des données et implémentations des filtres.....	93
5.2.2 Paramétrisation.....	95
5.2.3 Géologie de la région d'étude.....	96
5.3 Résultats et discussions.....	96
5.3.1 Cas du premier profil.....	96
5.3.2 Analyse des pseudosections différences.....	97
5.3.3 Evaluation des performances des filtres.....	99
5.3.4 Impact du filtrage sur l'inversion.....	100
5.3.5 Interprétation géologique.....	101
5.3.6 Cas du second profil.....	101
5.3.7 Examen des pseudosections différences.....	102
5.3.8 Evaluation quantitative des performances de filtres.....	103
5.3.9 Impact du filtrage sur l'inversion.....	103
5.3.10 Conclusion.....	105
Conclusion générale, nouveautés scientifiques et perspectives	106
Nouveautés scientifiques et contributions	109
Perspectives	109
Liste des travaux Scientifiques	110
Bibliographie	112
Annexe A	120
Annexe B	131

Table des figures

1.1.	Représentation d'une image numérique en deux dimensions	13
1.2.	Masque de filtrage appliqué à un pixel centré en x,y	15
2.1.	Schéma de discrétisation utilisé pour le calcul des gradients ($\nabla\Psi$) et des coefficients de diffusion (w)	22
2.2.	Bloc diagramme du filtre SD-ROM. $x(p)$ est l'échantillon central d'une fenêtre réarrangée de dimension 3x3, $Si(p)$ sont les valeurs triées de la fenêtre 3x3 à l'exception de l'échantillon $x(p)$, $di(p)$ est le vecteur différence entre $x(p)$ et $Si(p)$, Ti sont les valeurs seuils, $f(p)$ est la sortie du filtre et $Rm(p)$ est la ROM.	32
3.1.	Modèle 2D utilisé pour la génération des sections sismiques synthétiques.	38
3.2.	Section sismique, sans bruit, correspondant au modèle de la figure 3.1	38
3.3.	Section sismique synthétique avec 10% de bruit aléatoire ajouté	39
3.4.	Section sismique synthétique avec 50% de bruit aléatoire ajouté	39
3.5.	Résultats du filtrage de diffusion anisotropique : a) 10% et b) 50% de bruit aléatoire ajouté.	40
3.6.	Sections différences, a) 10% et b) 50 % de bruit aléatoire.	41
3.7.	Spectrogramme de la trace numéro 27 tirée de la section sismique sans bruit.	42
3.8.	Spectrogramme de la trace numéro 27 avec 10% de bruit aléatoire ajouté a) avant et b) après filtrage.	43
3.9.	Spectrogramme de la trace numéro 27 avec 50% de bruit aléatoire ajouté a) avant et b) après filtrage.	44
3.10.	Influence du nombre d'itérations sur l'atténuation du bruit, a) 2, b) 5, c) 10, d) 50 et e) 100 itérations.	45
3.11.	Graphes représentant l'évolution du RSB en dB et de l'EQM en fonction du nombre d'itérations. a) RSB en présence de 10 % de bruit aléatoire, b) EQM en présence de 10 % de bruit aléatoire, c) RSB en présence de 50 % de bruit aléatoire et d) EQM en présence de 50 % de bruit aléatoire.	48
3.12.	Sections différences montrant l'effet du nombre d'itérations sur le signal d'intérêt, a) 2, b) 5, c) 10, d) 50 et e) 100 itérations.	50
3.13.	Simulation de traces mortes (Flèches)	51
3.14.	Récupération de traces mortes	51
3.15.	Simulation de 5 traces mortes à différentes position sur la section	52
3.16.	Restitution d'un paquet de traces mortes (5 traces)	52
3.17.	Simulation de 7 traces mortes à différentes positions sur la section (cas limite).	53
3.18.	Limite de la restitution des traces mortes par le filtre anisotropique.	53
3.19.	Section sismique réelle sommée brute (Raw stack).	54
3.20.	Section stack après filtrage anisotropique	55
3.21.	Agrandissement dans une fenêtre de temps a) avant filtrage, b) après filtrage anisotropique	56
3.22.	Section réelle après filtrage dans le domaine f-k	57
4.1.	Shot gather synthétique (sans bruits) montrant quatre événements sismiques.	60
4.2.	gather contaminé par a) 5%, b) 10% et c) 50% de bruit aléatoire.	61
4.3.	Résultat du filtrage dans le cas de a) 5%, b) 10% et c) 50% de bruit aléatoire	62

	ajouté et d) résidu de filtrage du cas (c).	
4.4.	Influence du nombre d'itérations sur le résultat du filtrage après a) 2, b) 5, c) 10 et d) 50 itérations.	64
4.5.	Courbes du RSB et de l'EQM RSB en présence de 5% (a), 10% (c) et 50% (e) de bruit aléatoire, EQM en présence de 5% (b), 10% (d) et 50% (f) de bruit aléatoire.	66
4.6.	Spectrogramme de la trace numéro 29 tirée du shot gather de la Figure 4.1	68
4.7.	Spectrogrammes de la trace numéro 29 avec 5% (a), 10% (c) et 50% (e) de bruit et leurs filtrages respectifs (b,d,f).	69
4.8.	Comparaison entre données contaminées par a) 5% de bruit aléatoire, et résultats des filtres b) trilatéral, c) f-k et d) f-x.	70
4.9.	Comparaison entre données contaminées par a) 10% de bruit aléatoire, et résultats des filtres b) trilatéral, c) f-k et d) f-x.	71
4.10.	Comparaison entre données contaminées par a) 50% de bruit aléatoire, et résultats des filtres b) trilatéral, c) f-k et d) f-x.	72
4.11.	Stratégie proposée pour l'extraction puis la soustraction du bruit.	73
4.12.	Shot gather bruité issue d'une campagne sismique 3D.	74
4.13.	La composante bruitée isolée par le filtre trilatéral (a) et, le shot filtré(b).	75
4.14.	Agrandissement à l'intérieur d'une fenêtre du shot bruité (a) et filtré (b).	76
4.15.	Modèle 2D utilisé pour l'étude comparative.	77
4.16.	Section sismique synthétique correspondant au modèle de la figure 4.15.	78
4.17.	Données contaminées par a) 5%, b) 10% et c) 50% de bruit aléatoire.	79
4.18.	Résultats du filtre trilatéral en présence de a) 5%, b)10% et c) 50% de bruit aléatoire.	80
4.19.	Résultats du filtre anisotropique en présence de a) 5%, b) 10% et c) 50% de bruit aléatoire.	81
4.20.	Rapport RSB des données filtrées en fonction du nombre d'itérations pour les filtres trilatéral et anisotropique	83
4.21.	EQM relative aux données filtrées, en fonction du nombre d'itérations, pour les filtres trilatéral et anisotropique.	84
4.22.	Section sismique brute (a) et ses versions filtrées par les filtres anisotropique (b) et trilatéral (c).	85
4.23.	Vue élargie de la section sismique filtrée par les filtres anisotropique (a) et trilatéral (b).	86
5.1.	Modèle 2D utilisé pour générer les pseudosections de résistivités apparentes, en présence des inhomogénéités de surface (NSIs). L'espacement 'a' varie de 3m à 9m pour n=1,2...8.	89
5.2.	Pseudosection de résistivités apparentes calculées correspondant au modèle de la figure 5.1, en présence des NSIs.	89
5.3.	Pseudosection de résistivités apparentes calculées sans effets des NSIs.	90
5.4.	Pseudosections de résistivités apparentes calculées, après filtrage par filtre a) SD-ROM, b) médian (fenêtre 3) et c) médian (fenêtre 5).	91
5.5.	Résultats de l'inversion des données synthétiques (Les contours noirs reproduisent le modèle synthétique) : a) en l'absence des NSIs, b) en présence des NSIs, après filtrage par filtres c) SD-ROM, d) médian (fenêtre de longueur 3) et e) médian (fenêtre de longueur 5).	92
5.6.	Données réelles tirées du premier profil de direction Nord- Sud (Les flèches indiquent les valeurs anormales).	97
5.7.	Pseudosections différences après filtrage a) SD-ROM, b) médian (fenêtre 3) et	98

	c) médian (fenêtre 5).	
5.8.	Pseudosections inversées issues du premier profil après filtrage a) manuel, b) SD-ROM, c) médian (fenêtre 3) et d) médian (fenêtre 5).	100
5.9.	Données réelles issues du second profil de direction Est-Ouest (Les flèches indiquent des valeurs de résistivités anormalement élevées).	101
5.10.	Pseudosections différences après filtrage SD-ROM (a), médian (fenêtre 3) (b) et médian (fenêtre 5) (c).	102
5.11.	Pseudosections inversées issues du second profil après filtrage a) manuel, b) SD-ROM, c) médian (fenêtre 3) et d) médian (fenêtre 5).	104

Liste des tableaux

2.1.	Exemple de l'influence de la taille de la fenêtre sur le rapport RSB (d'après Chandra et al., 1998).	34
3.1.	Influence du nombre d'itérations sur le résultat du filtrage en termes du RSB et de l'EQM de données contaminées par 10% de bruit aléatoire. Le RSB et l'EQM des données bruitées sont 21.97 dB et 8.24E-4, respectivement.	46
3.2.	Influence du nombre d'itération sur le résultat du filtrage en termes du RSB et de l'EQM de données contaminées par 50% de bruit aléatoire. Le RSB et l'EQM des données bruitées sont -7.13 dB et 0.015, respectivement.	47
4.1	Résultat du filtrage en termes de RSB et de l'EQM en fonction du nombre d'itérations.	65
4.2	Résultats du filtre trilatéral et anisotropique en termes de rapport RSB et de l'EQM en fonction du nombre d'itérations.	82
5.1.	Profondeur d'investigation moyenne (Z_e) correspondant à différents dispositifs de longueur totale L (d'après Edwards, 1977).	94
5.2.	Etude comparative, en termes de RSB et déviation standard, des différents filtres du premier profil.	99
5.3.	Etude comparative, en termes de RSB et de déviation standard, des différents filtres du second profil.	103

Remerciements

Tout d'abord, j'exprime ma gratitude envers les membres de jury : Messieurs le professeur Hocine Shout, le Docteur Tahar Aifa et le Docteur Abd Rezak Lachouri qui m'ont fait l'honneur d'examiner mon travail et rapporté dessus et de faire partie de ce jury de thèse. Leurs remarques et suggestions m'ont permis d'apporter des améliorations à la qualité de ce travail.

Je remercie également Monsieur le Professeur Mabrouk Djeddi pour avoir présidé à ma soutenance de thèse.

Je remercie chaleureusement mon directeur de thèse Monsieur le Professeur Kamel Baddari, qui m'a fait confiance en acceptant de diriger ma thèse, pour la pertinence de ces conseils, sa rigueur scientifique et son soutien sans faille : qu'il trouve ici l'expression de ma profonde reconnaissance.

Je tient aussi à vivement témoigné ma reconnaissance à Monsieur le Professeur Roger Guérin, de l'université Pierre et Marie Curie-Paris 6, pour m'avoir accepté au sein de son laboratoire durant mon stage en France, et avec lequel j'ai eu des discussions constructives à plus d'une reprise.

Bien sûr je n'oublie pas mon frère, ami et collègue le Docteur Djarfour Nouredine avec lequel je partage mes quotidiens et avec lequel j'ai eu des discussions, toujours agréables, intéressantes et fertiles.

Merci à toutes les personnes que j'ai pu côtoyer pendant ces années. Merci à tous les doctorants et chercheurs, Nabil, Aissa, Hassen, Mouloud, Mme Alliouene. Ainsi, je tiens à remercier entre autres Ghanem Brahmi et Badr Ezamen Hedjazi pour leurs aides.

Enfin, et surtout, je remercie chaleureusement mon entourage proche (ma grande et petite famille) pour leurs encouragements et leurs soutien permanent et inconditionnel, que ce soit dans les bons ou mauvais moments. Qu'ils trouvent en ce témoignage toute l'affection que je leur porte.

Dédicaces

À ma raison d'être, ma mère,
À mon idéal, mon père,
À mes frères et sœurs,
À ma femme,
À toi mon ange Mohamed.

Introduction générale

Introduction

Le problème d'extraction des signaux à partir des données bruitées remonte à l'époque du mathématicien Allemand Carl Friedrich Gauss (18^{ème} siècle), qui a utilisé sa méthode des moindres carrées pour l'étude des mouvements des planètes (détermination de l'orbite de l'astéroïde Cérès) sur la base d'observations célestes. Plus récemment, le problème du filtrage a été abordé séparément au cours des années trente et quarante par **Kolmogorov (1939)** et **Wiener (1949)** qui ont traité l'estimation linéaire des processus stationnaires à l'aide des représentations spectrales, puis par **Kalman** dans les années soixante (**Kalman 1960**) avec l'apparition de son filtre révolutionnaire qui porte désormais son nom. Depuis, beaucoup de chemin a été parcouru dans le domaine du filtrage. Aujourd'hui, on assiste à l'avènement de méthodes de traitement non-classiques et complexes, appelées systèmes d'apprentissages ou procédés de calcul intelligent, parmi lesquelles on trouve les réseaux de neurones artificiels, la logique floue et les algorithmes génétiques, et la liste est loin d'être exhaustive.

Le thème du filtrage est identifié et caractérisé par plusieurs termes à savoir : filtrage, lissage, prédiction (**Haykin 1996**) et déconvolution (**Hayes 1996**). Il a toujours été considéré comme une pièce maîtresse dans l'échiquier du domaine du traitement numérique du signal, qui comporte aussi l'analyse du signal (Les lecteurs, désireux d'enrichir leurs connaissances dans le domaine du traitement numérique du signal, peuvent se référer à l'excellent ouvrage *Statistical and adaptive signal processing* de **Manolakis et al. 2005**).

Au sens le plus large, le filtrage est un processus permettant d'atténuer le bruit, présent dans une séquence de mesures donnée, tout en favorisant ou améliorant une information présentant un intérêt quelconque. Cette opération peut être définie soit dans le domaine spatio-temporel, incarnée par la convolution, soit dans le domaine fréquentiel, où la convolution est remplacée par un produit. Dans le cas particulier du traitement numérique des images, la convolution s'apparente à un filtrage spatial, où la luminosité d'un pixel est remplacée par une valeur qui a été déterminée grâce à son voisinage.

D'une manière générale, les filtres se décomposent en deux classes, linéaires et non-linéaires. La distinction se porte sur la relation qui existe entre la sortie du filtre et son entrée. Ainsi, un filtre est dit linéaire, si son entrée et sa sortie sont liées par une fonction linéaire, et non-linéaire dans le cas contraire.

Le filtrage linéaire stationnaire est probablement le plus commun et le plus important dans le traitement numérique du signal. Le caractère linéaire de ces filtres est propre à leur interprétabilité et leur utilisation aisée. Leur implémentation suppose que le signal est stationnaire et que l'information statistique *a priori* du signal et du bruit est connue. L'idée derrière ces filtres est de minimiser l'effet du bruit vis-à-vis du signal en satisfaisant un certain critère statistique. Ce qui s'applique parfaitement au filtre linéaire de Wiener, où l'erreur quadratique moyenne est minimisée entre la sortie désirée du signal et la réponse actuelle du filtre. Pour le cas des signaux non-stationnaires, c'est plutôt le filtre de Kalman qui est utilisé. Néanmoins, dans plusieurs cas (bruits dépendant du signal ou bruits impulsifs), les performances des filtres linéaires sont médiocres, et possèdent l'inconvénient de lisser beaucoup d'informations pertinentes telles que les contours des objets, en les déplaçant et les rendant flous.

Malgré leur mise en œuvre et leur conception ardue, les filtres non-linéaires sont la promesse de performances accrues. Cependant, les résultats obtenus sont très souvent de meilleure qualité que les résultats obtenus avec les filtres linéaires. Les filtres d'ordre, homomorphiques et morphologiques, sont les plus communs des filtres non-linéaires. Toutefois, une attention particulière est donnée lors de l'utilisation de ces filtres, car toute modification est irréversible. Dans le cas des filtres d'ordre, la sortie est définie comme une fonction des statistiques d'ordre. L'exemple du filtre médian est le plus connu. En réalité, les filtres d'ordre peuvent être considérés comme une modification des filtres linéaires discrets et non récurrents (fenêtrage suivi par une pondération linéaire) obtenue par l'insertion d'une opération de tri entre la phase de fenêtrage et celle de pondération.

En l'absence d'information statistique *a priori* sur le signal et le bruit, les méthodes (algorithmes) de traitement du signal ne sont plus valables. La solution vient donc par une utilisation récursive du filtre. Un filtre adaptatif n'est autre qu'un filtre récursif qui modifie automatiquement ces caractéristiques, généralement en temps réel, en optimisant ses paramètres internes. Ces modifications doivent suivre l'évolution des systèmes dans leurs

environnements aussi rapidement que possible. C'est pour cette raison qu'un filtre adaptatif est un système strictement non-linéaire.

Et la géophysique dans tout cela ?

En géophysique, le filtrage est un concept de base et occupe une place privilégiée dans le processus du traitement des données. En effet, les observations géophysiques sont souvent entachées de bruits issus de l'environnement d'enregistrement et de l'appareillage de mesure d'une part, et des perturbations dues au sous-sol, lui-même, d'autre part. Le bruit en géophysique est tout ce qui gêne l'interprétation. Le filtrage doit donc, non seulement améliorer le rapport signal/bruit (S/B) mais permettre aussi d'accéder à des grandeurs qui ne sont qu'implicitement contenues dans les données brutes.

Le bruit en géophysique est classé en i) aléatoire, ii) cohérent et/ou iii) impulsif. Les techniques d'atténuation du bruit ont toujours suivi l'évolution des méthodes du traitement numérique du signal. C'est pour cette raison qu'on retrouve une grande panoplie de filtres, appliqués dans d'autres domaines, en géophysique. En réalité, il n'existe pas de méthode miracle permettant de résoudre l'ensemble des problèmes qui se présentent, mais un ensemble de techniques appropriées (adaptées) à un type ou une catégorie de bruits.

Contexte et Objectifs

Des progrès considérables ont été accomplis dans le domaine du filtrage adaptatif et non-linéaire en traitement numérique du signal. Cependant, les applications issues de ce domaine, demeurent plutôt faibles en géophysique, voir inexistantes. En dépit des nombreux filtres étudiés, seuls quelques uns ont été examinés jusqu'à maintenant. Le développement de nouvelles applications, et l'adaptation de ces filtres suivant le type et l'intensité des bruits contaminants les données géophysiques, a été pour nous un vrai « challenge ». Notre principal objectif était donc de combler ces lacunes, et de dépasser les limitations des filtres conventionnels, en ouvrant la voie à de nouvelles possibilités dans le domaine du traitement des données géophysiques.

Dans le cadre de ce travail, nous nous sommes intéressés plus particulièrement à introduire les filtres adaptatifs et non-linéaires à base des statistiques d'ordre afin de mieux appréhender les problèmes rencontrés lors de la manipulation des données, particulièrement denses et redondantes, et qui sont issues d'un sous sol souvent complexe. L'originalité de ce travail réside principalement dans l'analogie dressée entre les représentations des données en

géophysiques (sections sismiques et pseudosections de résistivités apparentes en électrique) et l'imagerie. Analogie rendue possible grâce à la numérisation des images. Ainsi, une image numérisée est une matrice A en deux dimensions, dont chaque élément $A(x,y)$ est la mesure de l'intensité lumineuse en chaque point. Le couple (x,y) prend des valeurs entières qui désignent le numéro de la ligne et de la colonne du pixel correspondant. En géophysique, l'intensité du pixel est remplacée par la caractéristique physique propre à chaque méthode (coefficient de réflexion, résistivité apparente, etc.). Le choix des filtres non-linéaires et adaptatifs est argumenté par le besoin d'améliorer la résolution et le rapport S/B avec le moindre coût numérique, en plus d'une simple et rapide implémentation dans les différents environnements de traitement. Les questions essentielles pour lesquelles nous avons tenté d'apporter des solutions sont les suivantes :

Comment :

- identifier le bruit, avec le minimum d'intervention humaine?
- adapter le filtre au bruit ?
- garantir le bon fonctionnement du filtre ?

Les réponses à ces questions nous ont permis de sortir du cadre conventionnel du traitement des données géophysiques et de déboucher sur de nouveaux outils et algorithmes. C'est ainsi que nous avons introduit le filtre SD-ROM (Signal Dependent-Rank Order Mean) de **Chandra et al. (1998)** qui est basé sur les statistiques d'ordre (Order statistics), à la détection, et par la suite, à la suppression des bruits impulsifs présents dans les données de prospection électrique. A notre connaissance, son application est une première en géophysique. Nous avons aussi utilisé le filtre de diffusion anisotropique de **Perona et Malik (1990)** pour l'élimination des bruits aléatoires présents sur des données de sismique réflexion. Le dernier filtre (càd le filtre trilatéral de **Choudhury et Tumblin (2003)**) a aussi été utilisé pour l'atténuation des bruits de sismique réflexion. Le point commun de tous ces filtres est leur comportement non-linéaire et leur caractère adaptatif. En plus, dans tous les filtres, le bruit est atténué sans modification notable de l'amplitude du signal d'intérêt. Cette caractéristique est primordiale lors des études en amplitude préservée.

Organisation du document

Ce manuscrit est composé de 5 chapitres et s'organise de la façon suivante :

Le premier chapitre est consacré essentiellement à la présentation du problème de filtrage et les principales méthodes de filtrage appliquées en géophysique. Bien que cette présentation

soit obligatoirement trop succincte en regard des travaux menés dans le domaine du filtrage en géophysique. Ce chapitre est en quelque sorte l'état de l'art du domaine du filtrage.

Nous consacrons le **deuxième chapitre** à l'étude des fondements théoriques des différents filtres que nous avons appliqués, indispensables et nécessaires à la compréhension de cette thèse. Ensuite, nous nous intéressons à l'aspect algorithmique des filtres anisotropique, trilatéral et SD-ROM.

Le **troisième chapitre** traite de l'application du filtre par diffusion anisotropique au filtrage des données sismiques synthétiques, générées par Matlab, et validées par une application sur des données sismiques réelles. L'aspect algorithmique du filtre a aussi été développé. Les différents tests menés sur l'application du filtre anisotropique aux données sismiques bruitées ont montré qu'il existe une limite à définir entre l'atténuation du bruit et l'interprétabilité, en terme géologique, des résultats.

Le **quatrième chapitre** est dédié à la modélisation et à l'application réelle du filtre trilatéral. Les simulations ont été réalisées sur des données sismiques synthétiques générées par le programme libre Seismic Unix (SU). Les résultats ont été validés par une application sur une section sismique du Sud algérien. Pour des raisons de confidentialité, nous avons jugé de garder les informations portant sur la région d'étude et le nom du profil cachées.

Le **cinquième chapitre** présente une étude sur un modèle 2D pour la simulation d'une prospection électrique par courant continu en présence de bruits impulsifs. Les pseudosections de résistivité apparente calculées ainsi générées ont été inversées pour établir un modèle du sous-sol. Le comportement du filtre SD-ROM vis-à-vis du bruit a été investi. Le filtre a aussi été appliqué à des données réelles issues de deux profils électriques multi-électrodes acquis dans la région de Sétif. Les résultats ainsi obtenus ont montré la grande efficacité de la méthode proposée.

Une conclusion sur l'ensemble de ces méthodes ainsi que les perspectives d'avenir sont données à la fin de ce document.

Chapitre 1

Etat de l'art dans le domaine du filtrage en géophysique

Dans ce chapitre nous commençons par une présentation des principales techniques et algorithmes de filtrage existants dans le domaine de la géophysique, particulièrement en prospections sismique et électrique. Nous nous limiterons aux références bibliographiques des travaux les plus marquants et les plus importants dans ce domaine. En effet, regrouper toutes les techniques de filtrage en géophysique dans un seul chapitre, voire même dans un livre, relève de l'impossible. Car chaque progrès enregistré dans le domaine du filtrage trouve une application aussi prompte en géophysique. Dans la deuxième partie de ce chapitre, nous avons jugé profitable d'introduire quelques notions élémentaires sur les images numériques et le traitement numérique du signal.

1.1. Introduction

Le recours aux données de très hautes qualités est devenu une priorité en industrie pétrolière. Les méthodes d'interprétation modernes telles que l'analyse de la variation de l'amplitude en fonction de l'offset (AVO), l'analyse des attributs sismiques, la caractérisation des réservoirs, la classification des faciès ainsi que la tomographie électrique en sont les principaux bénéficiaires de ces données. La méthode 3D ainsi que la méthode 4D produisent des volumes de données importants, et sont de plus en plus utilisées en industrie pétrolière. Malheureusement, force est de constater que la nature n'a pas facilité la tâche aux géophysiciens, et que toutes ces données sont souvent contaminées par le bruit. Le terme *bruit* peut avoir plusieurs indications. En géophysique, nous considérons le bruit comme la partie du signal nuisible pour l'interprétation. La notion de bruit est étroitement liée à l'objectif de l'étude et aux traitements utilisés.

1.2. Les bruits en géophysique

1.2.1 Cas de la sismique

Le bruit en sismique peut apparaître comme :i) aléatoire, ii) cohérent et/ou impulsif. Le bruit aléatoire, comme son nom l'indique, représente la partie imprévisible du signal et non corrélable d'une trace à une autre. Ce type de bruit est le plus commun des bruits contaminants les données sismiques. Le bruit aléatoire est dû principalement à des

événements tels que le vent, les lignes électriques, le mouvement des véhicules, l'activité humaine, les géophones mal implantés (Yilmaz 1987). Les ondes directes, ondes réfractées ainsi que les ondes de surface sont des exemples des bruits cohérents (Spitzer et al. 2001). Ces ondes sont générées par la source (*source related*), et apparaissent avec des « move out » « corrections de temps » linéaires sur les sections sommées (Larner et al. 1983), et sont caractérisées par des niveaux d'énergies assez élevés comparées aux réflexions superficielles. Le ground roll est le plus gênant des bruits de surface. Il est caractérisé par de fortes amplitudes, comparées aux réflexions primaires, des fréquences basses ainsi que des pendages temporels assez marqués. Le problème des multiples, considérées comme du bruit cohérent, est très connu en sismique. En effet, l'énergie des multiples qui interfère avec celle des primaires, donne une représentation distordue du sous sol et conduit, souvent, à une interprétation erronée.

D'une manière générale, l'énergie des bruits cohérents est basse fréquence, tandis que celle des bruits aléatoires est plutôt haute fréquence. Le bruit impulsif quand à lui, est caractérisé par des pics isolés, ou par des traces isolées bruitées et de fortes amplitudes.

1.3. Méthodes d'atténuation des bruits en sismique

L'un des objectifs les plus importants du traitement du signal est, en plus de l'extraction du signal, l'atténuation des bruits tout en préservant le signal sismique.

Fondamentalement, deux méthodes d'atténuation du bruit existent i) le filtrage et ii) la modélisation et l'inversion. Cette dernière méthode tend à prédire puis à soustraire le bruit des données brutes (Tarantola 1987 ; Wegleim 1999 ; Guitton 2002 ; Voss et Hearn 2003).

La multiplication des sources de bruit a conduit à une multiplication des techniques d'atténuation de ces mêmes bruits. La plus simple de ces techniques est, sans doute, la fonction « mute » qui consiste à mettre à zéro une partie des données contaminées par le bruit. Néanmoins, cette méthode conduit à une perte significative du signal en plus du temps requis pour la réaliser.

1.3.1 Méthodes conventionnelles

Les méthodes de filtrages des bruits sont nombreuses en géophysique et la liste est loin d'être exhaustive. Les méthodes matricielles, basées sur la construction d'une matrice qui sert à décomposer l'espace d'observation en sous espaces, où l'on trouve : 1- la décomposition à valeurs singulières (DVS, SVD en anglais) (Ursin et Zheng 1985; Wenkai 2002), 2- l'analyse à composantes principales (PCA) (Hagen 1982), 3- la transformation de Karhunen-Loève (K-

L) (Jones et Levy 1987, Al-Yahia 1991). Les méthodes à gabarit opérants par changement de domaine tels les domaines temps-ralentissement (τ -p), fréquence- nombre d'onde (f-k), radial (R-T) (Claerbout 1975, 1983, 1985; Henley 2003) ou f-x (Canales 1984, Gulunay 1986), et où un gabarit est appliqué afin d'éliminer ou d'atténuer les parties bruitées. Ces méthodes s'appliquent souvent à des catégories d'arrivées particulières, par exemple les arrivées premières. D'autres méthodes telles la sommation (Mayne 1962, Naess et Bruland 1985), l'image propre (eigenimage en anglais) (Ulrych et al. 1999, Trickett 2003) ou la transformée en ondelette (Deighan et Watts 1997) sont aussi utilisées.

D'une manière générale, les méthodes d'atténuation des bruits cohérents sont mieux investies que celles destinées à l'atténuation des bruits aléatoires. Pour la plupart de ces méthodes, se sont les propriétés physiques qui différencient le bruit du signal, principalement la vitesse apparente ou le nombre d'ondes, qui sont utilisées lors du filtrage. En industrie pétrolière, les filtres linéaires, en particulier les filtres de Wiener (Meyerhoff 1966; Galbraith et Wiggins 1968; Cassano et Rocca 1973, 1974), f-k (March et Bailey 1983) et τ -p (Tatham et al 1982, Tatham 1984) sont les plus utilisées.

1.3.1.1. Limitations

Néanmoins, comme toute technique, ces méthodes possèdent aussi des inconvénients. Ainsi, les performances du filtre f-k sont assez médiocres lorsque les données ont été acquises avec un pas d'échantillonnage spatial non uniforme, ou bien lorsque le bruit cohérent est spatialement aliasé (Linville et Meek 1995), et peut parfois échouer de séparer le bruit du signal lorsqu'ils ont le même pendage (Karsli et Bayrak 2003; Haines et al. 2007). L'atténuation des bruits dans le domaine f-k, est particulièrement déconseillée lors des analyses en amplitudes préservées, car elle tend à détériorer l'amplitude du signal (Fatti et al. 1994). Tandis que la méthode de filtrage f-x souffre lorsque le rapport signal-sur-bruit (RSB) est faible, notamment aux hautes fréquences, (Harris et White 1997). Les autres méthodes, telles que le filtre τ -p ou la transformation K-L, peuvent causer de sérieuses distorsions aux données.

1.3.2. Méthodes basées sur l'intelligence artificielle

Plus récemment, de nouvelles techniques, relativement complexes, à l'exemple des réseaux de neurones artificielles (RNA) (Essenreiter 1999; Djarfour et al. 2008) et de logique floue (Hashemi et al. 2008), ont été employées pour le filtrage des données sismiques. Pour la

plupart de ces méthodes, l'intervention humaine est nécessaire, en particulier lors de la mise en place des règles, de la définition des ensembles ou lors de la définition des seuils. L'autre handicap est leurs paramétrisations difficiles, en plus de leurs implémentations assez compliquées. Nous nous focaliserons sur le cas des RNA :

1.3.2.1. Limitations des RNA

- **L'architecture optimale.** On mentionnera qu'il n'existe pas encore de théorie permettant de déterminer la structure optimale d'un réseau pour la réalisation d'une tâche donnée. En particulier, la détermination du nombre de couches cachées et du nombre de neurones sur chacune d'entre elle, qui relève encore largement de l'intuition de l'utilisateur, et de sa capacité à expérimenter plusieurs architectures afin de retenir celle qui donne les résultats les meilleurs.
- **Le pouvoir explicatif des réseaux.** Un réseau de neurones reste encore aujourd'hui en partie une « boîte noire ». Il est encore difficile d'extraire les relations pertinentes entre les variables. De ce point de vue, le système n'a donc qu'un pouvoir explicatif médiocre.
- **L'intervention humaine.** Comme cela a été signalé pour ce qui est du choix de l'architecture, ces systèmes font toujours trop largement appel à l'intuition de l'utilisateur. En particulier, l'apprentissage est guidé par des paramètres qu'il convient de régler manuellement. Un taux d'apprentissage (paramètre réglant la vitesse de convergence de l'algorithme) trop important peut aboutir à une oscillation du système alors qu'un taux trop faible se traduira par une convergence très lente, ce qui peut se révéler prohibitif en temps de calcul. Par ailleurs, il revient encore à l'utilisateur de choisir le point d'arrêt de l'apprentissage de façon à ce que le réseau conserve ses capacités de généralisation.

Les filtres dits « adaptatifs » constituent une classe à part des filtres usuels, et se distinguent par leurs facultés à modifier leurs paramètres au fur et à mesure du déroulement de l'opération de filtrage, sans aucune intervention de l'utilisateur. A l'inverse des méthodes de filtrage précédemment citées, les filtres adaptatifs sont faciles à implémenter et moins gourmands en calculs et peuvent donc s'intégrer facilement dans les différents environnements de traitements. Durant la dernière décennie, ces filtres ont été appliqués avec succès au domaine du traitement de l'image et de la parole. En géophysique, ces filtres ont aussi fait leurs

preuves, l'exemple le plus marquant de ces filtres est celui de la déconvolution (Robinson 1984). Ces applications, inspirées du domaine du traitement de l'image, ont ouvert de nouveaux horizons au traitement des données géophysiques (Bednar 1983 ; Hale 2001 ; Ristau et Moon 2001 ; Fehmers et Hocker 2003 ; Ferahtia et al. 2009).

1.4. Cas de la méthode électrique à courant continu

1.4.1. Historique de l'imagerie électrique

L'imagerie par résistivités électriques (ERI) a été à l'origine proposée par Henderson et Webster (1978) comme outil d'imagerie médicale, et introduite par la suite en géophysique par Lytle et Dines (1978). Ces dernières années, plusieurs auteurs ont établi les bases de la méthode ERI (Edwards 1977 ; Pelton et al. 1978 ; Dey et Morrison 1979). Néanmoins, elle reste relativement nouvelle en géophysique (Griffiths et Barker 1993). La méthode ERI trouve plusieurs domaines d'application, notamment dans le domaine de l'environnement et de la pollution (Daily et al. 1998), l'engineering (Chambers et al. 2006; Grellier et al. 2008), la contamination du sous-sol dans les carrières ou les décharges (Bernstone et al. 2000; Guérin et al. 2004; Chambers et al. 2005) ainsi que la caractérisation de la surface proche (Dahlin 1996).

La technique ERI mesure la résistivité du sous-sol en injectant un courant continu (CC) et en mesurant les différences de potentiels résultants. Ces dernières seront utilisées pour calculer les résistivités apparentes du sous-sol (mesures), puis seront représentées sous forme de cartes de résistivités vraies du sous-sol. Les résistivités apparentes sont mesurées par des dispositifs géoélectriques.

1.4.2. Dispositifs utilisés

En prospection électrique, plusieurs dispositifs sont utilisés, les plus communs sont les dispositifs Wenner, Schlumberger, pole-pole, pole-dipôle, dipôle- dipôle, et plus récemment d'autres dispositifs plus sophistiqués ont vu le jour tels ceux proposés par Wilkinson et al. (2006a, 2006b). Chaque dispositif possède ses avantages et ses inconvénients (Dahlin et Zhou 2004) vis-à-vis de la profondeur d'investigation, de la sensibilité aux changements verticaux et horizontaux des résistivités, ainsi qu'au niveau du bruit et de l'intensité du signal.

1.4.3. Représentation des données

Les mesures des résistivités apparentes sont représentées sous forme de pseudosection. La relation entre la résistivité apparente et la résistivité vraie est une relation complexe que l'on obtient par la résolution du problème inverse.

L'inversion des résistivités apparentes permet de produire des modèles du sous-sol en deux et en trois dimensions. L'objectif de l'inversion est de trouver un modèle de distribution des résistivités vraies capable de générer une réponse similaire aux valeurs des résistivités mesurées. L'inversion des données électriques s'est grandement améliorée par l'utilisation d'algorithmes robustes (Sasaki 1989, 1992 ; deGroot et Constable 1990 ; Loke et Barker 1996 ; Loke et al. 2003) et par le développement de l'outil informatique. Néanmoins, comme toutes mesures physiques, la présence des bruits est inévitable. L'étape de filtrage, ou suppression du bruit, est indispensable avant toute manipulation des données (LaBrecque et al. 1996; Ritz et al. 1999; Zhou et Dahlin 2003).

1.4.4. Les bruits en prospection électrique

Dans le domaine de la prospection électrique, les sources du bruit sont nombreuses : (i) bruits cohérents dus aux structures géologiques tels que les inhomogénéités de surface (IS) (ou *Near Surface Inhomogeneities NSI*, en anglais) (Ritz et al. 1999), les effets 3D des structures et les effets d'électrodes, (ii) bruits aléatoires et impulsifs tels que les courants telluriques, polarisation des électrodes (Cornacchiulo et Bagtzoglou 2004), ainsi que les bruits résultants de mesures avec un mauvais RSB. Rappelons que les NSI sont définies comme des corps géologiques, souvent de petites tailles localisés près de la surface, et caractérisés par des résistivités anormalement élevées, ou anormalement faibles, comparativement aux résistivités de l'arrière plan.

Bobachev et al. (1996) et Ritz et al. (1999) ont démontré que les NSI distordent les courbes des sondages électriques ainsi que les pseudosections, débouchant ainsi sur des interprétations géologiques erronées. L'inversion électrique est grandement affectée par la présence des NSI, en particulier lors de l'investigation de structures profondes.

1.4.5. Techniques de filtrages en prospection électrique

L'élimination du bruit avant toute inversion est donc nécessaire. Il existe plusieurs méthodes de filtrage du bruit, la plus simple est, sans doute, la méthode manuelle. Malheureusement, celle-ci requiert énormément de temps, en particulier lorsque le volume des données est

important (cas le plus fréquent). L'autre méthode, consiste à utiliser des filtres linéaires de type passe-bas. L'inconvénient de ce dernier filtre est qu'il tend à créer des zones de « non-valeurs », et qu'il est sensible aux valeurs anormales. L'alternative à ces méthodes peut être trouvée dans le domaine du traitement de l'image. Le filtre médian (MED) (Tukey 1974) est le plus répandu des filtres non-linéaires utilisés en traitement de l'image et de la parole (Gonzalez et Woods 2002), particulièrement lorsque le bruit est de type impulsif. Le principe du filtre MED est assez simple, mais son action face au bruit est très efficace. Ainsi, à l'intérieur d'une fenêtre glissante de longueur l , les données sont triées et l'échantillon central d'ordre p est remplacé par la valeur médiane. Néanmoins, lorsque le signal est affecté par une succession d'impulsions, alors il sera nécessaire d'utiliser une fenêtre plus large, ce qui produit un signal de plus basse fréquence. En traitement de l'image, ce phénomène est responsable de la dégradation des détails et de l'apparition du flou dans l'image (Chandra *et al.* 1998; Manglem *et al.* 2002). Gersztenkorn et Scales 1988 avaient combiné le filtre médian et le filtre moyenneur pour l'atténuation du bruit afin d'améliorer le processus d'inversion. Le filtre médian, en prospection électrique, est souvent utilisé lors du filtrage des anomalies dans les pseudosections de résistivités apparentes liées aux hétérogénéités de surface. Il a été décrit par Ritz *et al.* (1999) et implémenté par Bobachev et Robain dans le logiciel X2IPI.

1.5. Notions sur les images

Le matériel de cette partie, utilisant le traitement numérique des images est tiré, entre autres, des excellents ouvrages « *Digital Image Processing* » de Gonzales et Woods (2002) et « *Introduction to digital image processing* » d'Alasdair McAndrew (2004).

1.5.1. Définition d'une image numérique

Une image peut être considérée comme une fonction bidimensionnelle continue $f(x,y)$, d'amplitude f et de coordonnées cartésiennes x et y . f est appelée aussi *intensité* ou *niveau de gris* de l'image en ce point.

Une image numérique est une représentation matricielle en deux dimensions de la fonction $f(x,y)$, de valeurs numériques $A(x,y)$ aux coordonnées x, y (Figure 1.1), avec x et y qui représentent les numéros de ligne et de colonne, respectivement. Une image numérique est composée d'un nombre fini d'éléments, où chacun possède une position et une valeur bien définie. Chaque élément est appelé *pixel*, caractérisé par une luminosité ou une intensité. Les pixels aux alentours d'un pixel central s'appellent voisinages (*neighbourhoods* en anglais). Un voisinage est caractérisé par sa forme matricielle, on dit alors, voisinage 3x3, 5x5 etc.

Un pixel central q , possède ainsi quatre voisins, horizontaux et verticaux, de coordonnées $(x+1,y)$, $(x-1,y)$, $(x,y+1)$, $(x,y-1)$, et quatre voisins diagonaux de coordonnées $(x+1,y+1)$, $(x+1,y-1)$, $(x-1,y+1)$, $(x-1,y-1)$ (Figure 1.1). D'une manière générale, un voisinage possède un nombre impair de lignes M et de colonnes N , de cette manière le pixel central se trouve toujours au centre de la matrice.

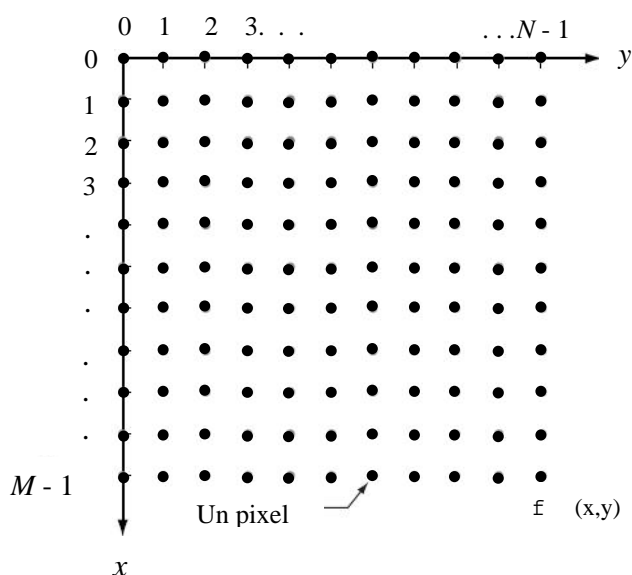


Figure 1.1. Représentation d'une image numérique en deux dimensions.

La représentation matricielle de l'image numérisée de la **Figure 1.1.** est donnée par :

$$f(x, y) = \begin{bmatrix} \Lambda(0,0) & \Lambda(0,1) & \dots & \Lambda(0, N-1) \\ \Lambda(1,0) & \Lambda(1,1) & \dots & \Lambda(1, N-1) \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \Lambda(M-1,0) & \Lambda(M-1,1) & \dots & \Lambda(M-1, N-1) \end{bmatrix} \quad (1.1)$$

Dans le cas où Δx et Δy , représentent les distances entre deux points consécutifs de la matrice (1.1), nous pouvons écrire :

$$\Lambda(m,n) = \Lambda(m\Delta x, n\Delta y), \text{ avec } m=0,1,\dots,M-1 \text{ et } n=0,1,\dots,N-1 \quad (1.2)$$

1.5.2. Qu'est ce qu'un traitement de l'image ?

Le mot traitement de l'image implique un changement de la nature de l'image, afin d'améliorer au mieux possible son aspect pour la vision humaine, ou bien pour l'adapter à un traitement automatisé.

1.5.3 Thèmes du traitement de l'image

En traitement de l'image il existe une multitude d'algorithmes et de méthodes qui manipulent les images numériques, chaque méthode est dédiée à une tâche donnée. Parmi les tâches les plus importantes citons : l'amélioration, le traitement et la segmentation de l'image.

1.5.4 Opérations de base en traitement de l'image

Les opérations en traitement de l'image ont pour but de changer la valeur d'un pixel. Selon l'information de base, nécessaire pour appliquer la transformation, on distingue trois opérations de base :

- Transformation : une transformation présente les valeurs des pixels sous une forme différente, mais équivalente, au pixel d'origine.
- Traitement de voisinage : afin de changer le niveau de gris d'un pixel, il suffit de connaître le niveau de gris d'un nombre fini de pixel de son voisinage.

- Traitement ponctuel : le changement dans la valeur du pixel se fait indépendamment de son voisinage.

Nous allons nous intéresser au traitement de voisinage, car dans nos applications, nous avons fait recours à ce type de traitement.

1.5.4.1 Traitement de voisinage

L'idée de base est de faire glisser un *masque* (*kernel, fenêtre ou filtre*) sur l'image à traiter. Le résultat est une nouvelle image, dont les valeurs de chaque pixel sont calculées à partir des valeurs à l'intérieur du masque. Chaque valeur dans un masque est appelée coefficient. Le masque est déplacé de proche en proche jusqu'à atteindre la fin de l'image (Figure 1.2).

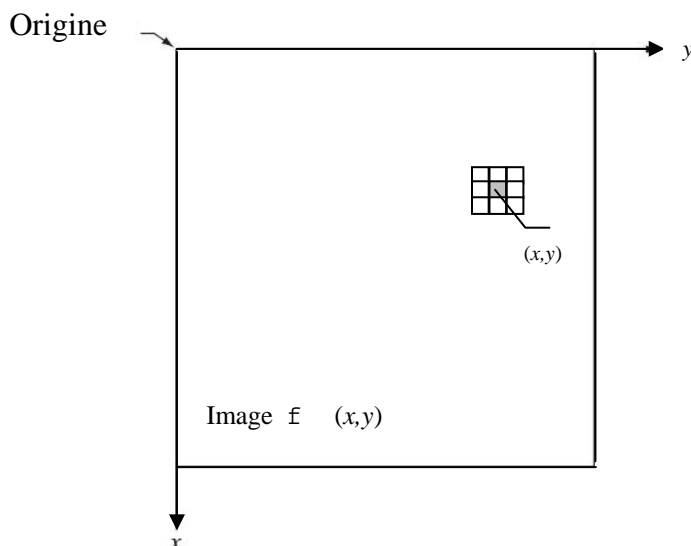


Figure 1.2. Masque de filtrage appliqué à un pixel centré en x, y .

Si la relation reliant le pixel de sortie aux pixels d'entrée est linéaire, alors on parle de filtrage linéaire. Cette opération est aussi assimilée au filtrage spatial. Dans ce cas, on dit aussi convolution spatiale, définie par :

$$g(m, n) = (h * f)(x, y) = \frac{1}{MN} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} h(m-k, n-l) f(k, l) \quad (1.3)$$

Avec : $g(\cdot)$ l'image de sortie, $h(\cdot)$ le masque de convolution, $f(\cdot)$ portion de l'image d'entrée (de même dimension que le masque $M \times N$). Cette opération de filtrage peut aussi être appliquée dans le domaine fréquentiel, où la convolution est remplacée par un produit :

$$G(u,v)=H(u,v).F(u,v) \quad (1.4)$$

Dans le cas où il n'est pas possible de mettre l'opération de filtrage sous forme de convolution, on parle alors de filtrage non-linéaire.

1.5.5. Types de bruits en imagerie

Les principaux types de bruits rencontrés en imagerie sont:

- Bruit impulsif : connu aussi sous le nom de poivre et sel (*salt and pepper*) : ce bruit apparaît comme une distribution aléatoire de pixels blancs ou noirs sur l'image. Ce bruit est obtenu en ajoutant aux pixels originaux des valeurs aberrantes.
- Bruit Gaussien : ce bruit est une représentation idéale d'un bruit blanc. Le bruit gaussien est un bruit blanc normalement distribué. Modélisable en ajoutant des valeurs aléatoires.
- Bruit moucheté (*speckle noise*) ou bruit multiplicatif : ce bruit est modélisable en multipliant une valeur aléatoire par la valeur actuelle du pixel.
- Bruit périodique : ce cas se produit lorsque l'image est sujette à un bruit périodique.

1.5.6. Filtrage de bruits en imagerie.

Le filtrage de bruit en imagerie peut se pratiquer dans deux domaines différents : spatial ou fréquentiel. Il peut aussi être linéaire ou non-linéaire. Les filtres linéaires les plus connus sont :

- Les filtres linéaires du domaine spatial :
 - Le filtre moyenneur : où la valeur du pixel de sortie est une moyenne des valeurs de son voisinage (équivalent au filtre passe-bas).
 - Le filtre gaussien
 - ..
- Les filtres opérants dans le domaine fréquentiel :
 - Passe-bas,
 - Passe-haut.
 - Orientation sélective
 - ..
- Filtres linéaires de restauration d'images

- Inverse, pseudo-inverse.
- Wenner
- ..

L'inconvénient majeur de ces filtres est qu'ils ont tendance à flouer l'image d'origine et à améliorer certaines fréquences tandis que d'autres sont atténuées. En plus, ils sont incapables de différencier, pour une seule composante fréquentielle, entre la partie désirable et la partie indésirable du signal.

Pour y remédier à ces problèmes, les filtres non-linéaires ont été introduits. Le filtre médian, et ses variantes (Weighted median filter WM, Centered weighted median filter CWM, Space variant median SVM...), est sans doute le plus connu des filtres non-linéaires. Ce filtre appartient à une famille de filtres plus grande appelée filtres à base de statistique d'ordre (OSF). On retrouve dans la littérature du traitement de l'image un grand nombre de filtre non-linéaires, nous citerons ici quelques exemples:

Erosion (Min), Dilatation (Max), Ouverture, Fermeture, Laplacien, Prewitt, Sobel, filtres basés sur la décision : SD-ROM, Tri-state median filter (TSM), pixel-wise Pad (PWMAD), Homomorphique,

Chapitre 2

Théorie des filtres : Diffusion anisotropique, Trilatéral et SD-ROM

Nous abordons avec ce chapitre l'aspect théorique et algorithmique des filtres de diffusion anisotropique, trilatéral et SD-ROM. Bien que cette étude soit concise, elle donne néanmoins les concepts de base de chaque méthode.

Ces filtres appartiennent à la famille des filtres adaptatifs non-linéaires. Leur utilisation en traitement numérique de l'image est devenue courante. Néanmoins, à notre connaissance, ces filtres n'ont jamais été utilisés en traitement des données géophysiques (cas des filtres trilatéral et SD-ROM).

2.1. Filtre de diffusion anisotropique

2.1.1 Bases théoriques

Premier volet des trois filtres qui seront explorés, le filtre de diffusion anisotropique non-linéaire a été introduit par [Perona et Malik \(1990\)](#) comme outil de filtrage des images bruitées, avec la particularité de préserver les détails et les contours.

La diffusion est un phénomène physique décrivant la compensation des différences de concentration entre deux solutions, sans perte ni création de matière. L'analogie avec le traitement de l'image peut être établie en remplaçant la concentration d'une solution par l'intensité d'un pixel d'une image donnée.

2.1.2 Filtrage par convolution

Soit $\Psi_0(x,y)$ une image bruitée aux coordonnées cartésiennes x,y . Considérons la fonction gaussienne

$$G_\sigma(x,y) = \frac{1}{2\pi\sigma} \exp\left(-\frac{x^2 + y^2}{2\sigma}\right) \quad (2.1)$$

σ est la variance de la gaussienne.

Le problème de filtrage du bruit peut se résumer à une convolution spatiale de l'image avec la fonction gaussienne :

$$\Psi(x,y,t)=(\Psi_0*G_\sigma)(x,y) \quad (2.2)$$

Avec $\Psi(x,y,t)$ qui représente l'image restaurée au temps t ($t \in [0,t]$), $\Psi_0(x,y,t)$ représente l'image bruitée et $(*)$ l'opérateur de convolution.

Quand t varie continuellement, on obtient une famille d'images $\Psi(.,t)$, qui ne sont autres que la solution de l'équation aux dérivées partielles (équation de la chaleur) :

$$\begin{cases} \frac{\partial \Psi(x,y,t)}{\partial t} = \frac{\partial^2 \Psi(x,y,t)}{\partial x^2} + \frac{\partial^2 \Psi(x,y,t)}{\partial y^2} \\ \Psi(x,y,0) = \Psi(x,y) \end{cases} \quad (2.3)$$

L'opération de lissage gaussien (équation 2.2) est similaire à un phénomène de conduction de la chaleur avec diffusivité constante.

Comme cité plus haut, l'équation de diffusion est une équation aux dérivées partielles (EDP), et elle peut se réécrire sous la forme :

$$\frac{\partial \Psi(x,y,t)}{\partial t} = \text{div}(w(x,y,t)\nabla \Psi(x,y,t)) = \nabla w \nabla \Psi + w(x,y,t)\Delta \Psi(x,y,t) \quad (2.4)$$

Où $\Psi(x,y,t)$ est l'intensité lumineuse (amplitude ou résistivité en géophysique) du pixel de l'image d'entrée (événement sismique) aux coordonnées x et y et au temps de diffusion t ; et $w(x,y,t)$ représentent les coefficients de diffusion (fonction de diffusivité), lesquels contrôlent le taux de lissage. ∇ représente le gradient de l'image, div la divergence tandis que Δ est l'opérateur Laplacien ($\Delta \Psi = \nabla^2 \Psi = \text{div}(\text{grad} \Psi)$). L'équation (2.4) est similaire à la deuxième loi de Fick's.

2.1.3 Analyse de l'EDP

On remarque que si les coefficients w sont constants, alors l'équation (2.4) est dite diffusion linéaire et est équivalente à l'équation de la chaleur (équation 2.3) :

$$\frac{\partial \Psi(x,y,t)}{\partial t} = w \nabla^2 \Psi(x,y,t) \quad (2.5)$$

- lorsque la fonction $w(x,y,t)$ est remplacée par une fonction de contrôle de diffusion aux niveaux des bords « ou edge-stopping function », la diffusion est dite non-linéaire.
- dans le cas où $\nabla\Psi$ et le flux de diffusion $j(j = -w\frac{\partial\Psi}{\partial x})$ sont parallèles, on parle de diffusion isotrope.
- Le cas d'anisotropie réfère à une diffusion isotrope non-linéaire, non homogène.

Le filtrage isotrope linéaire utilise l'équation (2.5) comme modèle, et sa solution est équivalente à la convolution avec un opérateur (masque) gaussien. Ce type de filtrage est utilisé pour l'atténuation des bruits aléatoires dans les zones homogènes (faibles gradients) des images 2D. Néanmoins, les images résultantes sont souvent flouées. La diffusion linéaire ne possède aucune direction privilégiée, atténuant ainsi bruits et contours sans distinction.

2.1.3.1. Cas non linéaire

La diffusion non linéaire aide à surmonter les handicaps de la diffusion linéaire isotrope en contrôlant la diffusion au niveau de chaque point, c'est ainsi que la diffusion peut être réduite le long des contours, il s'agit d'un lissage conditionnel. L'idée de base du filtre de diffusion anisotropique non-linéaire est justement de préserver les contours. Cette préservation se fait grâce à une orientation maîtrisée parallèlement aux contours. C'est ainsi que la diffusion est maximale dans les zones homogènes (faibles gradients), et minimale près des contours (forts gradients). Pour leurs filtres de diffusion anisotropique non-linéaire, [Perona et Malik \(1990\)](#) ont proposé une fonction de diffusivité w monotonement décroissante, en fonction du gradient de magnitude ($\nabla\Psi$), pour le contrôle de la diffusion. La stabilité impose que cette fonction (w) soit une fonction décroissante. Habituellement, les fonctions suivantes sont utilisées :

$$w_1(x, y, t) = e^{-\left(\frac{\nabla\Psi}{\eta}\right)^2} \quad (2.6a)$$

$$w_2(x, y, t) = \frac{1}{1 + \left(\frac{\nabla\Psi}{\eta}\right)^2} \quad (2.6b)$$

η est appelée constante de diffusion. La constante de diffusion permet une diffusion rétrograde lorsqu'elle est trop faible, comparée au gradient $\nabla\Psi$. Cette constante fait office de « seuil » permettant de distinguer les zones de faibles gradients de celles à forts gradients.

2.1.4. Approximation discrète par différence finie

La discrétisation des EDP est obtenue par remplacement de toutes les dérivées par des formules de différences. Les références relatives à la solution numérique des EDP sont nombreuses, nous nous limiterons aux suivantes : Ames (1992), Morton et Mayers (1994), Cooper (1998). Perona et Malik (1990) ont proposé la forme discrète suivante de l'équation (2.4) :

$$\Psi_q^{t+\Delta t} = \Psi_q^t + \Delta t \sum_{l \in n_q} w_{q,l}^t \nabla \Psi_{q,l}^t \quad (2.7)$$

Où n_q représente l'ensemble des pixels au voisinage du pixel central q dans les quatre directions (Nord, Sud, Est et Ouest), Δt est le pas espace-temps (ou bien l'incrément Δt), $\nabla \Psi_{q,l} = \Psi_{nq}^t - \Psi_q^t$ et $|n_q|$ sont les pixels au voisinage du pixel central à la position q .

Avec :

$$\left\{ \begin{array}{l} \nabla \Psi_{q,N}^t = \Psi_N^t - \Psi_q^t \\ \nabla \Psi_{q,S}^t = \Psi_S^t - \Psi_q^t \\ \nabla \Psi_{q,E}^t = \Psi_E^t - \Psi_q^t \\ \nabla \Psi_{q,O}^t = \Psi_O^t - \Psi_q^t \end{array} \right. \quad \text{et} \quad \left\{ \begin{array}{l} w_{q,N}^t = g_i(|\nabla \Psi_N^t|) \\ w_{q,S}^t = g_i(|\nabla \Psi_S^t|) \\ w_{q,E}^t = g_i(|\nabla \Psi_E^t|) \\ w_{q,O}^t = g_i(|\nabla \Psi_O^t|) \end{array} \right. \quad (2.8)$$

g_i représente la fonction de diffusion (équations 2.6a ou 2.6b).

Le calcul du gradient et des poids correspondants est schématisé par la Figure (2.1).

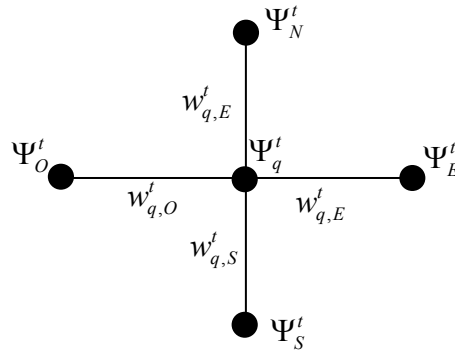


Figure 2.1. Schéma de discrétisation utilisé pour le calcul des gradients ($\nabla\Psi$) et des coefficients de diffusion (w).

Comme cité plus haut, les dérivées spatiales sont approximées par des différences entre les échantillons du voisinage et l'échantillon central, selon la formule suivante :

$$\frac{\partial\Psi}{\partial x} \approx \frac{\Psi_{n_q} - \Psi_q}{\Delta x} \quad (2.9)$$

Où l'indice « n_q » réfère aux voisinages de l'échantillon central Ψ_q , et Δx au pas d'échantillonnage spatial.

Tandis que les dérivées temporelles (∂t) sont remplacées par leurs approximations directes des différences. Les indices (N , S , E et O) indiquent respectivement le Nord, le Sud, l'Est et l'Ouest par rapport à l'échantillon central à la position q et au temps t .

2.15. Remarques sur la stabilité

Comme nous l'avons vu précédemment, l'équation aux dérivées partielles peut être approximée par une représentation discrète à l'aide de la méthode des différences finies. Le mot « discrète » veut dire que la solution n'est connue que pour un certain nombre de points. D'une manière générale, augmenter le nombre de point signifie l'amélioration de la résolution et l'exactitude de la solution numérique. L'incrément (Δt) dans l'équation (2.5) doit être dans l'intervalle $[0, P]$, P est un nombre réel positif. Il est clair que plus cet incrément est proche de zéro, plus l'approximation numérique est proche du cas continu. Néanmoins, pour la stabilité

de l'intégration numérique, l'incrément (Δt) est généralement pris inférieur à $\frac{1}{4}$. Pour plus de détail sur l'étude de stabilité se référer à [Gerig et al. \(1992\)](#) et [Deriche et Faugeras \(1996\)](#).

2.1.6. Description de l'algorithme de filtrage par diffusion anisotropique

Une fenêtre, d'habitude carrée, centrée à l'échantillon Ψ positionné en q est choisie sur l'image originale. On procède par itération, et à chaque itération les gradients ($\nabla\Psi$) sont calculés pour les quatre directions (N, S, E et O) ainsi que les coefficients de diffusion (w_i) correspondants. L'image restaurée est une somme des mises à jour de l'image originale. En effet, la partie droite de l'équation (2.5) représente le changement d'intensité de l'image pour une seule itération du filtre. La fenêtre est déplacée à la position suivante à la fin du nombre d'itérations préétabli (les critères concernant le choix du nombre d'itération seront donnés au chapitre suivant). Ce procédé, préserve les bords pour un certain nombre d'itérations, mais tend vers la moyenne de l'image lorsque le nombre d'itérations dépasse un certain seuil. Dans ce cas précis, il n'existe pas de critère universel d'arrêt. Il faudra donc décider du moment de l'arrêt du procédé d'itérations.

2.1.7. Algorithme du filtre anisotropique

Etape 1. Lecture des données

- Lire les données à partir du fichier des données sismiques sous format SEG Y
- Convertir les données en format Matlab.
- Identifier le nombre de lignes m et le nombre de colonnes n de la matrice des données.

Etape 2. Initialisation

- o Introduire le nombre d'itérations, introduire la constante de diffusion (η), le facteur de stabilité (pas espace-temps), proposer la fonction de diffusion (w_1 ou w_2).
- o Rembourrer les bords de la matrice (étape nécessaire afin d'éviter de perdre l'information aux bords).

Etape 3. Filtrage

Pour chaque ligne
Pour chaque colonne
Pour chaque itération

} Position initiale de la fenêtre

- 1- Calculer la magnitude du gradient ($\nabla\Psi$) pour chaque direction (N, S, E, O)
- 2- Calculer les coefficients de diffusion $w(x,y,t)$ associés à la fonction de diffusion (w_1 ou w_2) en utilisant l'équation 2.8 pour chaque itération.
- 3- Mettre à jour la valeur du pixel se trouvant à la position actuel, selon l'équation 2.7.

Prochaine itération
Prochaine colonne
Prochaine ligne

} Déplacement de la fenêtre

Etape 4. Réorganisation des données pour la sortie

Les données sont transformées du format Matlab au format SEG Y.

Algorithme 1: Le filtre de diffusion anisotropique

2.2. FILTRE TRILATERAL

2.2.1. Introduction

L'application du filtre trilatéral sur les données sismiques est rendue possible grâce à l'analogie dressée avec les images. La section sismique a été considérée comme une fonction Λ de deux variables x et t , avec x qui représente la distance émetteur-récepteur et t le temps vertical double. La magnitude de cette fonction peut être remplacée par l'intensité d'un pixel d'une image 2D. Partant de ce raisonnement, nous avons tenté d'appliquer les algorithmes, déjà existants en traitement d'images, au domaine de la sismique.

Le moyen le plus simple en traitement de l'image pour atténuer le bruit est d'utiliser les filtres spatiaux. Un filtre spatial agit par convolution d'un masque (appelée aussi kernel ou fenêtre, généralement carrée) au voisinage du pixel central. C'est ainsi que la nouvelle valeur du pixel filtré est calculée à partir du voisinage de ce pixel. Chaque élément du masque est un coefficient du filtre (poids). Les filtres spatiaux sont regroupés dans trois familles : linéaires, non linéaires et adaptatives. Le filtre trilatéral est un filtre adaptatif non-linéaire.

2.2.2. Arrière plan théorique

Le filtre trilatéral de [Choudhury et Tumblin \(2003\)](#) est une extension du filtre bilatéral de [Tomasi et Manduchi \(1998\)](#). Ce filtre est particulièrement robuste en présence de bruits aléatoires en plus de son aptitude à préserver les contours. Dans ce qui suit, nous présenterons le filtre trilatéral établi par [Garnett et al. \(2005\)](#) qui, lui-même, se base sur le filtre bilatéral mais diffère du filtre trilatéral de [Choudhury et Tumblin \(2003\)](#).

2.2.3. Principe du filtre bilatéral

Le filtre bilatéral est un filtre non-linéaire, dont le comportement est analogue à celui d'un filtre passe-bas à deux dimensions. L'idée de base de ce filtre est que chaque échantillon central ζ dans une fenêtre glissante est remplacé par une moyenne pondérée de son voisinage dans l'espace et la magnitude. Les poids les plus importants sont attribués aux échantillons spatialement « proches », noté filtrage du domaine (ou *domain filtering*) et de grandeurs analogues à l'échantillon central ζ , noté filtrage de portée (ou *range filtering*). Signalons que tous les poids doivent être normalisés afin de garder la moyenne locale.

2.2.4. Formulation mathématique du filtre bilatéral

Considérons un échantillon central $\zeta_{x,t}$ dans une fenêtre de taille $2n+1 \times 2n+1$, et $S_{x+p, t+q}$ ($-n \leq p, q \leq +n$) l'ensemble des échantillons au voisinage de ζ , où x représente l'offset, t le temps tandis que les indices p et q représentent la position spatiale de chaque échantillon du voisinage, ζ et S se réfèrent aux amplitudes des échantillons.

Chaque échantillon ($S_{x+p, t+q}$) est multiplié par un poids $\Gamma_{pq}(\zeta, S)$, lui-même résultat de la multiplication d'un poids de proximité $\Gamma_c(\zeta, S)$ et d'un poids de similitude $\Gamma_s(\zeta, S)$ tel que:

$$\Gamma_{pq}(\zeta, S) = \Gamma_c(\zeta, S) \cdot \Gamma_s(\zeta, S) \quad (2.10)$$

Le poids de proximité est donné par:

$$\Gamma_c(\zeta, S) = e^{-\frac{|d(\zeta, S)|^2}{2\sigma_d^2}} \quad (2.11)$$

avec :

$$d(\zeta, S) = d(\zeta - S) = \|\zeta - S\| \quad (2.12)$$

Où $\|\cdot\|$ est une distance géométrique entre l'échantillon central ζ et les échantillons du voisinage S , et $|\cdot|$ représente la valeur absolue.

Il est clair que plus la distance spatiale entre ζ et S est importante plus le poids de proximité $\Gamma_c(\zeta, S)$ est amoindri. Dans ce cas, on parle de proximité dans le domaine.

Le paramètre σ_d est défini par **Tomasi et Manduchi (1998)** comme une constante de diffusion.

Il est responsable du taux de filtrage passe-bas. Ce paramètre joue le rôle d'un seuil, il sert à identifier les échantillons spatialement proches.

Le poids de similarité est défini par :

$$\Gamma_s(\zeta, S) = e^{-\frac{\|\zeta - S\|^2}{2\sigma_s^2}} \quad (2.13)$$

Où $|\zeta|-|S|$ est une mesure du gradient ou de la “distance” séparant les deux magnitudes. Avec l’augmentation de la distance séparant les deux magnitudes (c.-à-d. les deux échantillons sont très différents), le poids de similarité diminue. La similarité renvoie à la proximité en termes de magnitude.

2.2.5. Etude des paramètres seuils

Le paramètre σ_s joue aussi le rôle de seuil, servant à identifier les échantillons de magnitudes proches (similaires). Les paramètres σ_s et σ_d contrôlent ainsi le comportement des poids. Ces paramètres définissent les points auxquelles les dérivées des fonctions gaussiennes de pondération prennent leurs valeurs maximales.

Notons que :

- 1) Dans le cas particulier où $\sigma_s \rightarrow \infty$ alors le comportement du filtre bilatéral approche celui d’un filtre gaussien de déviation standard σ_s .
- 2) Dans le cas où σ_s et $\sigma_d \rightarrow \infty$ le comportement du filtre bilatéral se rapproche de celui d’un filtre moyeneur.

2.2.6. Formulation du filtre bilatéral

Les deux poids $\Gamma_c(\zeta, S)$ et $\Gamma_s(\zeta, S)$ représentent des fonctions gaussiennes, inversement proportionnelles à la « distance ».

Notons que dans les équations (2.11 et 2.13), il est possible de remplacer la fonction gaussienne par n’importe quelle autre fonction décroissante.

Enfin, l’échantillon central ($\hat{\zeta}$) restauré est donné par la forme normalisée suivante :

$$\hat{\zeta}(\zeta, S) = \frac{\sum_{(p,q) \in [-n,+n]} (\Gamma_{pq}(\zeta, S) \cdot S_{x+p,t+q})}{\sum_{(p,q) \in [-n,+n]} \Gamma_{pq}(\zeta, S)} \quad (2.14)$$

Il s’ensuit, d’après la formule (2.14), que l’échantillon en question est remplacé par une moyenne pondérée d’échantillons qui lui sont spatialement proches et semblables en termes d’amplitude.

2.2.7 Définition de la statistique d'ordre

Considérons les réalisations $\{x_1, x_2, \dots, x_n\}$ de la variable aléatoire X . Le tri croissant de ces réalisations nous donne l'ensemble ordonné $\{x_{(1)}, x_{(2)}, \dots, x_{(n)}\}$.

On définit alors la statistique d'ordre (i) la valeur $x_{(i)}$. Cette valeur est la $i^{\text{ème}}$ valeur dans l'ensemble des valeurs ordonnées.

Les valeurs suivantes dans la série ordonnée sont particulièrement intéressantes :

- $x_{(1)}$ représente le minimum
- $x_{(n)}$ représente le maximum
- $x_{((n+1)/2)}$ représente la valeur médiane

Les combinaisons linéaires de ces statistiques d'ordre donnent naissance aux L-filtres, tandis que les combinaisons non-linéaires donnent naissance aux NL-filtres. Le filtre médian est le plus connu des L-filtres. Ce filtre sera utilisé plus tard dans ce document.

2.2.8. Définition de la statistique « ROAD »

Garnett et al. (2005) ont introduit une statistique appelée Rank-Ordered Absolute Differences (ROAD) afin de détecter les échantillons affectés par des bruits impulsifs.

Nous commencerons par définir les différences absolues $\delta_{p,q}(\zeta_{x,t})$ en magnitudes (intensités) entre les échantillons du voisinage $S_{x+p,t+q}$ et l'échantillon central $\zeta_{x,t}$:

$$\delta_{p,q}(\zeta_{x,t}) = |S_{x+p,t+q} - \zeta_{x,t}|, \quad (\text{avec } -n \leq p, q \leq +n) \quad (2.15)$$

Ensuite, les valeurs des différences $\delta_{p,q}(\zeta_{x,t})$ sont triées de façon croissante, l'élément r_j représente le $j^{\text{ème}}$ plus petit élément des différences (c'est une statistique d'ordre j).

La statistique ROAD est ainsi définie par :

$$\text{ROAD}_m(\zeta) = \sum r_j(\zeta_{x,t}) \quad (j=1, \dots, m) \quad (2.16)$$

avec $2 \leq m \leq (2n+1)^2 - 2$.

Pour $n=1$, nous aurons $2 \leq m \leq 7$. Dans la suite, nous considérons que le cas où $m=4$, donc $\text{ROAD}_m(\zeta) = \text{ROAD}_4(\zeta)$.

La statistique ROAD permet de mesurer le degré de proximité de l'échantillon central (ζ) de ces quatre échantillons du voisinage (S) les plus semblables. Les échantillons bruités devront avoir des valeurs ROAD élevées, tandis que ceux non-bruités devront avoir de faibles

valeurs. En d'autres termes, les échantillons bruités vont avoir des valeurs qui varient grandement de leur voisinage, tandis que ceux non-bruités doivent avoir, au minimum, la moitié des valeurs du voisinage ayant des magnitudes similaires. Ainsi, la statistique ROAD est utilisée comme un détecteur de bruit.

2.2.9. Le filtre trilatéral

Dans ce qui suit, nous introduisons le filtre trilatéral basé sur la statistique ROAD, selon les lignes directrices données par Garnett et al. (2005) :

Une troisième fonction de pondération a été introduite aux poids du filtre bilatéral existant, et faisant intervenir la statistique ROAD, appelée poids impulsif $\Gamma_I(\zeta)$:

$$\Gamma_I(\zeta) = e^{-\frac{ROAD(\zeta)^2}{2\sigma_I^2}} \quad (2.17)$$

où le paramètre σ_I joue le rôle de seuil au-delà duquel les valeurs ROAD élevées seront pénalisées.

En présence de bruit, le poids impulsif $\Gamma_I(\zeta)$ est incompatible avec le poids de similarité $\Gamma_s(\zeta, S)$, d'où la nécessité d'introduire une nouvelle fonction appelée « *joint impulsivity* », ou impulsivité jointe, afin de limiter l'utilisation du poids de similarité.

La fonction d'impulsivité « J » de S vis-à-vis à l'échantillon central ζ est donnée par :

$$J(\zeta, S) = \left(1 - e^{-\frac{\left(\frac{ROAD(\zeta) + ROAD(S)}{2}\right)^2}{2\sigma_J^2}} \right) \quad (2.18)$$

Cette fonction prend des valeurs dans l'intervalle [0,1] suivant les valeurs de ζ et S par rapport à σ_J . Si aucun des échantillons (ζ, S) n'est bruité, et ainsi aucun n'a de valeur ROAD élevée vis-à-vis de σ_J , alors $J(\zeta, S) \approx 0$, sinon $J(\zeta, S) \approx 1$. L'impulsivité jointe se comporte donc comme un commutateur.

2.2.10. Version modifiée du filtre trilatéral

Il est maintenant possible de définir le filtre trilatéral, comme présenté par Garnett et al. (2005) qui est une version différente du filtre trilatéral de Choudhury et Tumblin (2003):

$$\Gamma_{pq}(\zeta, S) = \Gamma_c(\zeta, S) \cdot \Gamma_s(\zeta, S)^{1-J(\zeta, S)} \cdot \Gamma_I(\zeta)^{J(\zeta, S)} \quad (2.19)$$

Remarquons que lorsque dans (2.19) $J(\zeta, S) \approx 1$ alors $1 - J(\zeta, S) \approx 0$. Le seuil σ_s dans ce cas devient trop grand ce qui provoque l'annulation de l'effet des différences de magnitudes, tandis que le poids d'impulsivité reste inchangé. Lorsque $J(\zeta, S) \approx 0$, c'est l'inverse qui se produit et seul le poids de similitude est utilisé pour la distinction entre les échantillons, le seuil du poids impulsif est alors trop grand. D'après Garnett et al. (2005), la combinaison des équations (2.14) et (2.19) produit le filtre trilatéral. En l'absence de bruits impulsifs, le filtre trilatéral se transforme en filtre bilatéral.

2.2.11. Algorithme du filtre trilatéral

Etape 1. Lecture des données

- Lire les données à partir du fichier des données sismiques sous format SEG Y
- Convertir les données en format Matlab.
- Identifier le nombre de lignes m et le nombre de colonnes n de la matrice des données.

Etape 2. Initialisation

Introduire les valeurs : σ_s ; σ_r ; σ_l ; σ_j

début

Pour chaque ligne

Pour chaque colonne

Etape 2. Partie Calculs

- Sélectionner la fenêtre centrée à l'échantillon $\zeta_{x,t}$
- Calculer le vecteur des différences $\delta_{p,q}(\zeta_{x,t})$ selon la formule (2.15).
- Trier le vecteur $\delta_{p,q}(\zeta_{x,t})$.
- Calculer la statistique ROAD, selon la formule (2.14).
- Calculer les poids : $\Gamma_c(\zeta,S)$; $\Gamma_s(\zeta,S)$; $\Gamma_l(\zeta)$; $J(\zeta,S)$; $\Gamma_{pq}(\zeta,S)$, selon les formules : (2.11, 2.13, 2.17, 2.18 et 2.19, respectivement).

Etape 3. Mise à jour

- Remplacer l'échantillon central $\zeta_{x,t}$ par la nouvelle valeur selon la formule (2.14)

Prochaine colonne

Prochaine ligne

Etape 4. Réorganisation des données pour la sortie

Les données sont transformées du format Matlab au format SEG Y.

Algorithme 2 : Le filtre trilatéral

2.3. Filtre SD-ROM

2.3.1. Introduction

Le filtre SD-ROM a été proposé initialement par [Abreu et al. \(1996\)](#), [Chandra et al. \(1998\)](#), [Moore et al. \(1999\)](#), et [Moore et Mitra \(2000\)](#) pour le traitement des bruits impulsifs contaminant les images en 2 dimensions. Nous avons étudié, analysé et utilisé ce filtre pour l'élimination des bruits impulsifs présents sur des données de résistivités électriques.

Le mot **SD-ROM** est une abréviation de **Signal Dependent Rank-Ordered Mean**.

2.3.2. Théorie du filtre SD-ROM

Le filtre SD-ROM appartient à la famille des filtres non-linéaires basés sur la décision, appelé aussi filtre adaptatifs à base de statistique d'ordre (Order Statistic Filters ou OSF). L'exemple le plus connu des OSF est le filtre médian standard, introduit par [Tukey \(1974\)](#), comme outil de lissage des signaux discrets. Durant cette dernière décennie, les recherches sur les OSF se sont multipliées, et leurs spectres d'utilisation se sont élargis.

2.3.3. Description de l'action du filtre

Le filtre SD-ROM procède en deux temps : en premier temps les échantillons suspects sont détectés par un détecteur d'impulsion, puis en second temps, l'échantillon bruité est remplacé par une estimation appropriée notée ROM. Le bloc diagramme de la Figure 2.2 illustre le comportement du filtre SD-ROM :

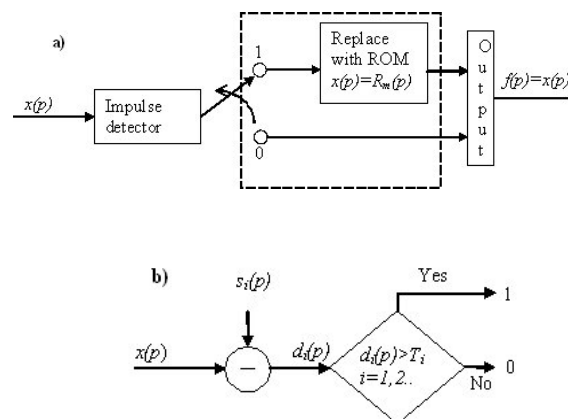


Figure 2.2. Bloc diagramme du filtre SD-ROM. $x(p)$ est l'échantillon central d'une fenêtre réarrangée de dimension 3×3 , $S_i(p)$ sont les valeurs triées de la fenêtre 3×3 à l'exception de l'échantillon $x(p)$, $d_i(p)$ est le vecteur différence entre $x(p)$ et $S_i(p)$, T_i sont les valeurs seuils, $f(p)$ est la sortie du filtre et $R_m(p)$ est la ROM.

Les données de résistivités électriques apparentes sont ici notées sous forme matricielle $\rho_{i,j}$ de telle façon que la valeur de chaque résistivité apparente, pour chaque point de mesure, se retrouve par ses coordonnées (x,y) correspondant aux coordonnées spatiales (dans la direction horizontale et verticale). A chaque position, une fenêtre de dimension 3×3 est choisie puis les données à l'intérieur de cette fenêtre seront transformées en un vecteur 1 dimension.

L'échantillon $x(p)$ est l'échantillon central à l'intérieur de la fenêtre. Deux cas de figures peuvent se présenter. Dans le premier cas, l'échantillon est identifié comme bruité et le filtre se met en position 1 où l'échantillon $x(p)$ est remplacé par la Rank Order Mean (ROM) R_m . L'identification est réalisée par comparaison entre le vecteur des différences $d_i(p)$ et les valeurs seuils T_i prédéterminées par l'utilisateur. Le vecteur des différences est le résultat de la différence entre les valeurs d'un vecteur $S_i(p)$, constitué d'éléments se trouvant à l'intérieur d'une fenêtre de dimension 3×3 centrée à l'échantillon $x(p)$ (excepté l'échantillon $x(p)$ lui-même), et l'échantillon central $x(p)$. Dans le second cas, l'échantillon est identifié comme non contaminé par le bruit et passe inchangé à travers le filtre (commutateur en position 0).

Les valeurs seuils peuvent être déterminées statistiquement ou par des tests empiriques (Moore et Mitra, 2000). Dans le cas d'une fenêtre de taille 9, quatre valeurs seuils seront nécessaires. Tandis que pour des fenêtres plus larges, l'augmentation du nombre des valeurs seuils est nécessaire.

Comme indiqué par Chandra et al. (1998), les performances du filtre peuvent être grandement améliorées par une utilisation itérative.

2.3.4. Importance de la taille de la fenêtre

Dans leurs études du filtre SD-ROM, Chandra et al. (1998) ont remarqué que la taille de la fenêtre est un élément très important dans le comportement du filtre. Ainsi, l'augmentation de la taille de la fenêtre n'est pas toujours synonyme de résultats meilleurs, mais elle tend à améliorer la performance, en particulier lorsqu'on trouve plus de deux impulsions. Néanmoins, l'augmentation de la taille de la fenêtre conduit à l'augmentation du nombre de valeurs seuils (T_i). Le tableau 2.1 montre l'effet de la taille de la fenêtre sur le résultat final, où la taille cinq produit les meilleurs résultats. Néanmoins, la taille de la fenêtre change en fonction des données.

Longueur de la fenêtre	S/B(dB)
5	35.87
7	34.41
9	32.77

Tableau 2.1. Exemple de l'influence de la taille de la fenêtre sur le rapport RSB (d'après Chandra et al., 1998).

2.3.5. Algorithme du filtre SD-ROM

1. Introduire les valeurs des résistivités apparentes mesurées $\rho_{i,j}$
2. Sélectionner une fenêtre T de taille $l \times l$ (dans notre cas une fenêtre de taille 3×3) centrée à l'échantillon de rang p noté $x(p)$.
3. Réarranger les valeurs de résistivité ($x_i(p)$) à l'intérieur de la fenêtre T en un vecteur 1 dimension que nous noterons $w_i(p)$, avec $p=(i-1)*l+j$ où $l=3$ et $i=j=l$, tel que:

$$w_i(p)=[x_{-4}(p),x_{-3}(p),x_{-2}(p),x_{-1}(p),x(p),x_1(p),x_2(p),x_3(p),x_4(p)]^T \quad (2.20)$$

4. Exclure l'échantillon central $x(p)$ du vecteur $w_i(p)$. Définir ainsi le vecteur résultant $\hat{w}_i(p)$ de taille huit :

$$\hat{w}_i(p)=[x_{-4}(p),x_{-3}(p),x_{-2}(p),x_{-1}(p),x_1(p),x_2(p),x_3(p),x_4(p)]^T \quad (2.21)$$

5. Réaliser un tri ascendant du vecteur $\hat{w}_i(p)$

$$\hat{w}_i(p)_{trié}=S_i(p)=[s_1(p),s_2(p),s_3(p),s_4(p),s_5(p),s_6(p),s_7(p),s_8(p)]^T \quad (2.22)$$

tel que $s_1(p) \leq s_2(p) \dots \leq s_8(p)$. Les $s_i(p)$ sont les statistiques d'ordre p de x_i

6. Définir les valeurs différences d'ordre p , $d_i(p)$, entre les éléments de $S_i(p)$ et l'échantillon central $x(p)$:

$$d_i(p) = \begin{cases} s_i(p) - x(p) & \text{si } x(p) \leq R_m \\ x(p) - s_{9-i}(p) & \text{sinon} \end{cases} \quad (2.23)$$

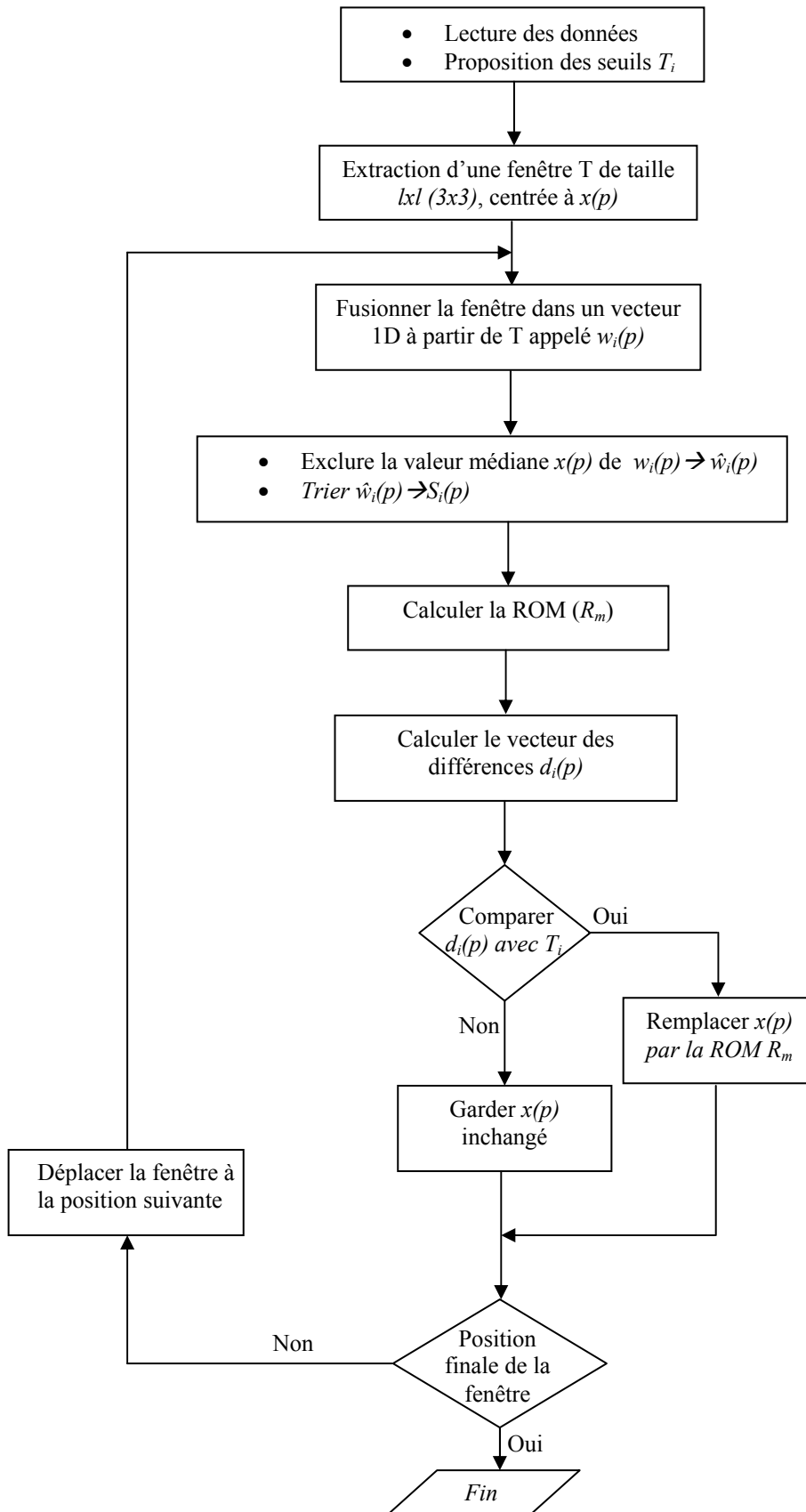
Pour ($i=1,\dots,4$), avec R_m la Rank-Ordered mean (ROM) définie par :

$$ROM=[s_4(p)+s_5(p)]/2 \quad (2.24)$$

7. Tester si l'une des valeurs $d_i(p)$ correspondantes est supérieure à l'une des valeurs seuils T_i , alors considérer l'échantillon comme bruit, sinon laisser la valeur de $x(p)$ inchangée.
8. Remplacer l'échantillon $x(p)$ contaminé par la ROM R_m (expression (2.24)), et déplacer la fenêtre vers la position prochaine, jusqu'à atteindre la position finale.

Algorithme 3 : Le filtre SD-ROM

2.3.6. Organigramme du filtre SD-ROM



Organigramme 1 : Filtre SD-ROM

Chapitre 3

Atténuation des bruits par le filtre de diffusion anisotropique

Dans le chapitre précédent, l'aspect théorique et algorithmique du filtre de diffusion anisotropique a été présenté. En outre il a été souligné le rôle fondamental joué par ce filtre dans l'élimination du bruit tout en conservant le signal d'intérêt. Ce filtre se base sur le phénomène physique de diffusion dans les fluides. Le but de ce chapitre est de présenter l'applicabilité de ce filtre au domaine du filtrage des données sismiques synthétiques et réelles.

3.1. Implémentation du filtre de diffusion anisotropique

3.1.1. Données synthétiques

Nous avons appliqué le filtre de diffusion anisotropique sur des données sismiques synthétiques et réelles. Des programmes sous Matlab© ont été développés afin de générer les données synthétiques, d'étudier les effets du bruit aléatoire sur les données ainsi que le filtrage. Nos programmes permettent à l'utilisateur de spécifier le niveau de bruit, de choisir son mode de distribution (normalement distribué ou uniformément distribué). En plus, il est possible de modifier les paramètres du filtre et de visionner les résultats en temps réel.

3.1.1.1. Description du modèle

Les données synthétiques ont été générées à partir d'un modèle géologique 2D composé essentiellement de 6 couches, dont deux couches avec variation latérale d'épaisseur, simulant ainsi la présence de biseaux ([Figure 3.1](#)). En plus, une discontinuité près de la trace numéros 88 a été générée afin d'imiter la présence de faille. La section sismique synthétique correspondant à la convolution avec une ondelette de Ricker de fréquence centrale égale à 25 Hz, est fournie en [Figure 3.2](#).

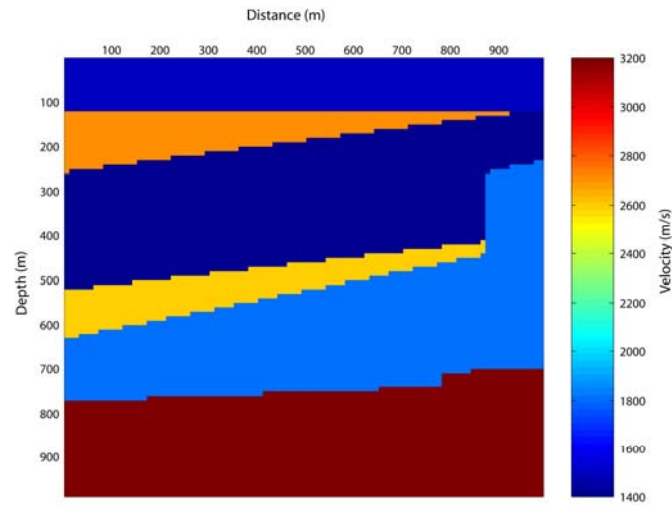


Figure 3.1. Modèle 2D utilisé pour la génération des sections sismiques synthétiques.

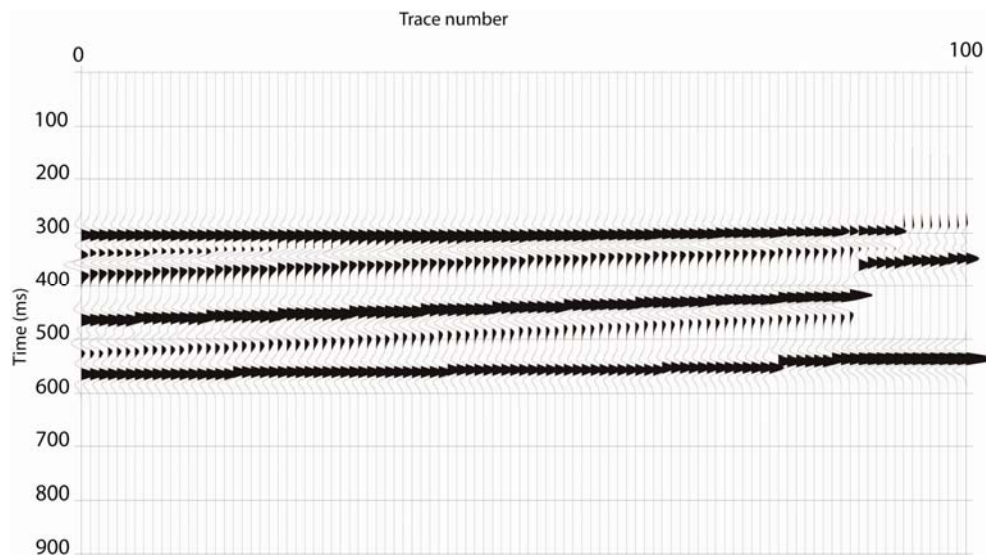


Figure 3.2. Section sismique sans bruit, correspondant au modèle de la figure 3.1.

3.1.1.2. Paramétrisation

Après plusieurs tests, nous avons pu déterminer les paramètres optimaux du filtre, et qui ont été fixés durant toutes les autres simulations. Dans la suite les paramètres utilisés seront les suivants :

- constante de diffusion = 50
- intervalle de temps = 0.10 (à cause de la stabilité)

- nombre maximum d'itérations = 3 et 8 pour le cas de 10 et 50% de bruit, respectivement.
- une fonction de diffusivité analogue à celle présentée par l'équation (2.6a).

En vue de tester l'applicabilité du filtre anisotropique, et afin de rapprocher les données synthétiques de la réalité géologique, nous avons ajouté aux données synthétiques 10 et 50% de bruits aléatoires et le résultat est présenté en [Figures 3.3 et 3.4](#), respectivement. Remarquons, que la faille déjà présente dans le modèle de la [Figure 3.1](#), est partiellement voilée par le bruit, en particulier en présence de 50% de bruit aléatoire ([Figure 3.4](#)).

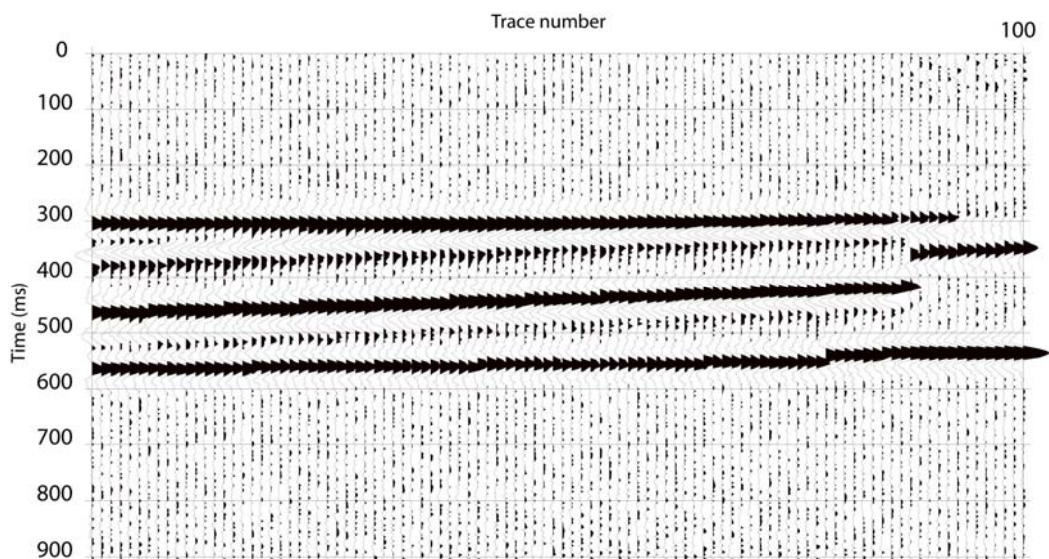


Figure 3.3. Section sismique synthétique avec 10% de bruit aléatoire ajouté

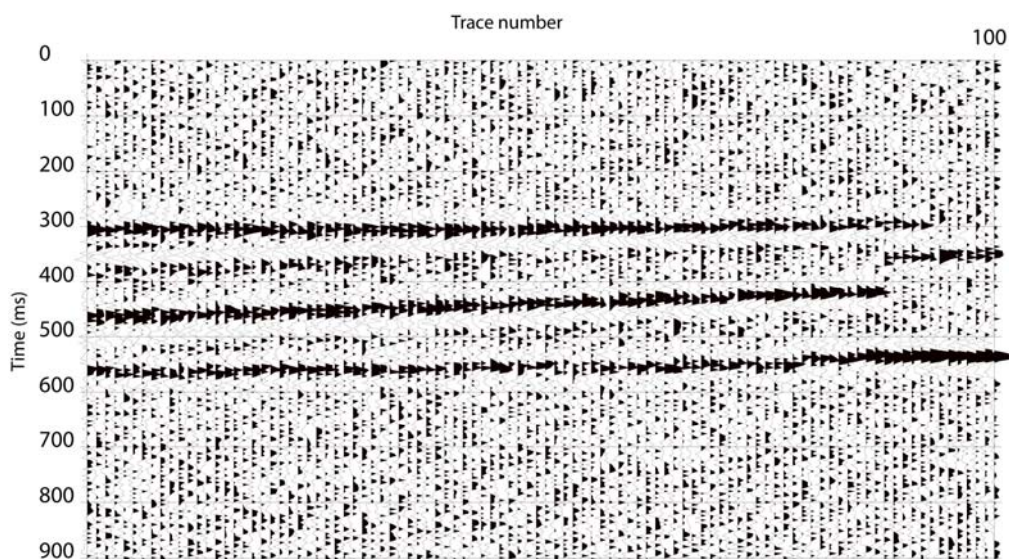


Figure 3.4. Section sismique synthétique avec 50% de bruit aléatoire ajouté

3.1.1.3. Résultats des simulations et discussions

a- Etude qualitative

Comme prévu, la majorité du bruit aléatoire ajouté à la section sismique synthétique a été atténué, avec une nette amélioration du rapport signal-sur-bruit (RSB) (Figures 3.5a et 3.5b). Soulignons, que l'amplitude et la continuité des réflecteurs sismiques ont été substantiellement préservées et améliorées. La faille simulée près de la trace numéro 88 a été parfaitement repérée sur les deux sections filtrées.

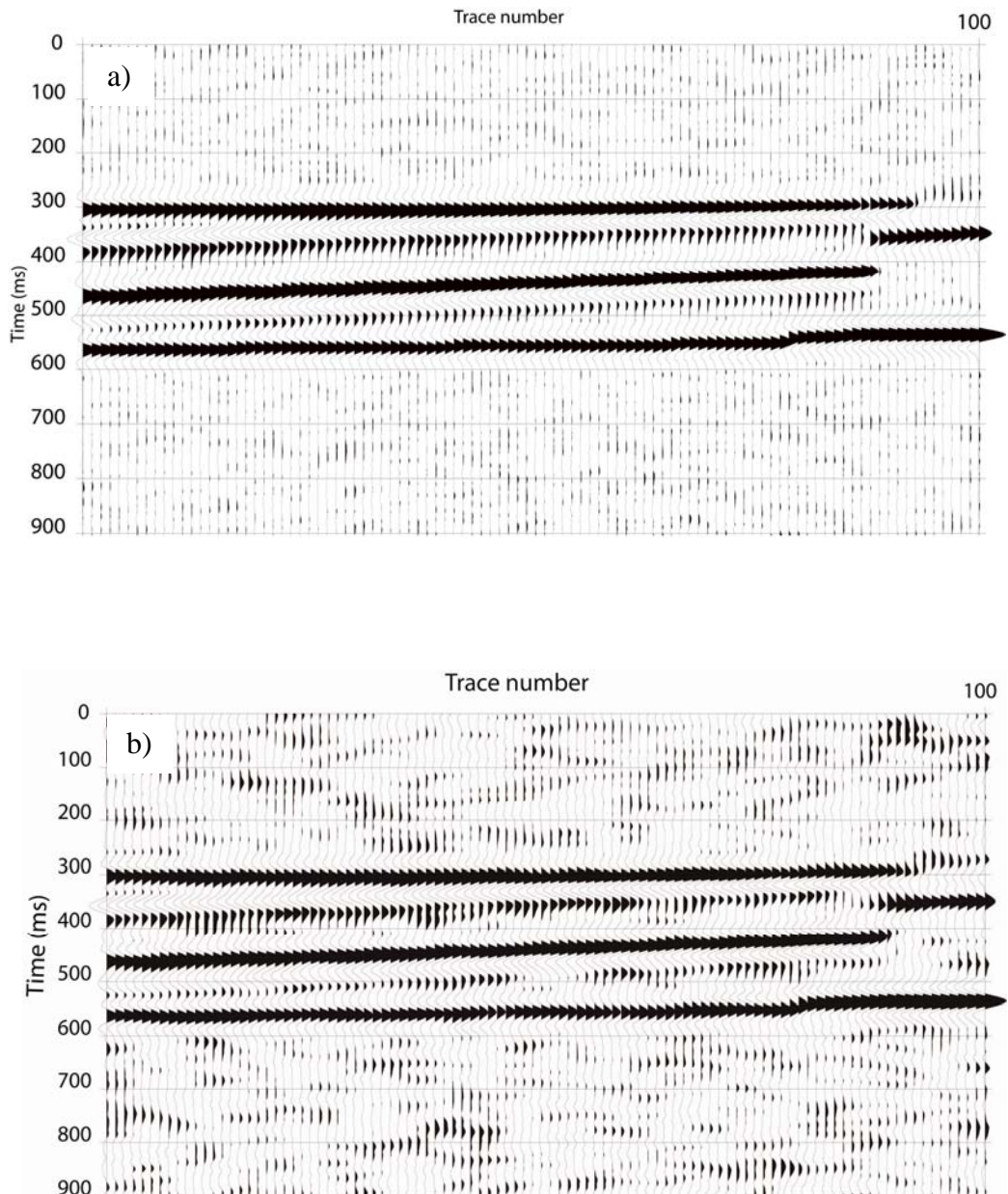


Figure 3.5. Résultats du filtrage de diffusion anisotropique : a) 10% et b) 50% de bruit aléatoire ajouté.

b- Comparaison des sections différences

Les sections différences découlent de la soustraction, échantillon par échantillon, des sections bruitées et de la section synthétique sans bruit, comme illustré par les [Figures 3.6a et 3.6b](#), pour le cas de 10 et 50% de bruit aléatoire, respectivement.

Les sections différences sont très utiles pour faire ressortir les différences entre deux images, le résultat de cette différence définit le changement.

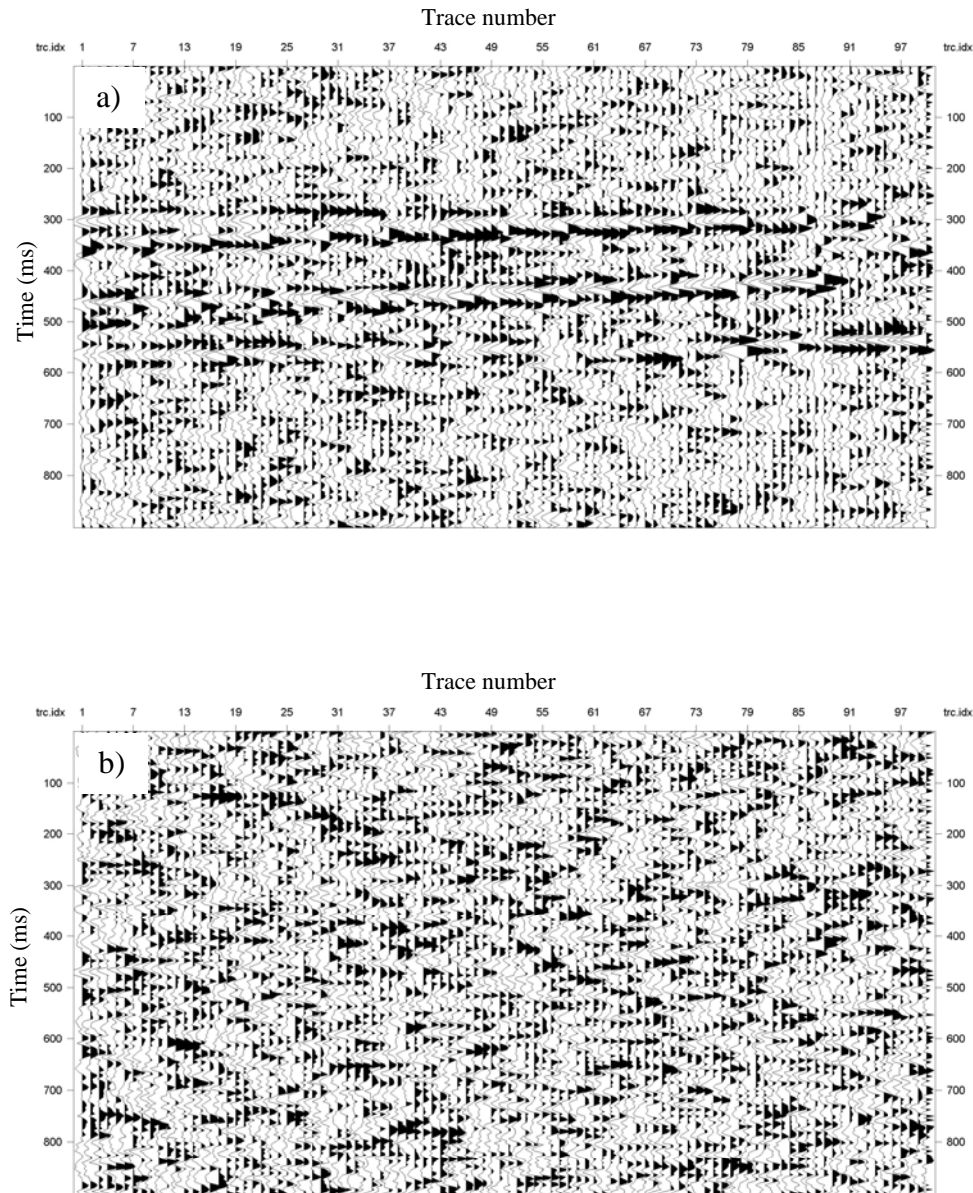


Figure 3.6 Sections différences, a) 10% et b) 50 % de bruit aléatoire.

Il est observé que le bruit a été parfaitement localisé et fortement atténué, sans changement notable du signal d'intérêt.

Indiquons toutefois, qu'une faible atténuation du signal d'intérêt est observée sur la [Figure 3.6a](#). Ce qui s'explique, en partie, par le fait que nous avons gardé les paramètres du filtre inchangés durant toutes les simulations afin d'éviter de biaiser le comportement de notre filtre.

c- Investigation dans le domaine spectrale

L'autre moyen d'analyse des données est le domaine spectral à travers l'analyse des spectrogrammes. Il convient de rappeler qu'un spectrogramme est une transformation de Fourier d'une séquence temporelle (appelée aussi *short-time Fourier transform STFT*). L'analyse par spectrogrammes est très bénéfique, car elle permet de visualiser simultanément la densité spectrale d'un signal donnée en fonction du temps et de la fréquence. Nous avons fait recours à la fonction Matlab « *spectrogram* » pour le calcul des spectrogrammes. La [Figure 3.7](#) représente le spectrogramme de la trace numéro 27 extraite de la section synthétique sans bruit de la [Figure 3.2](#).

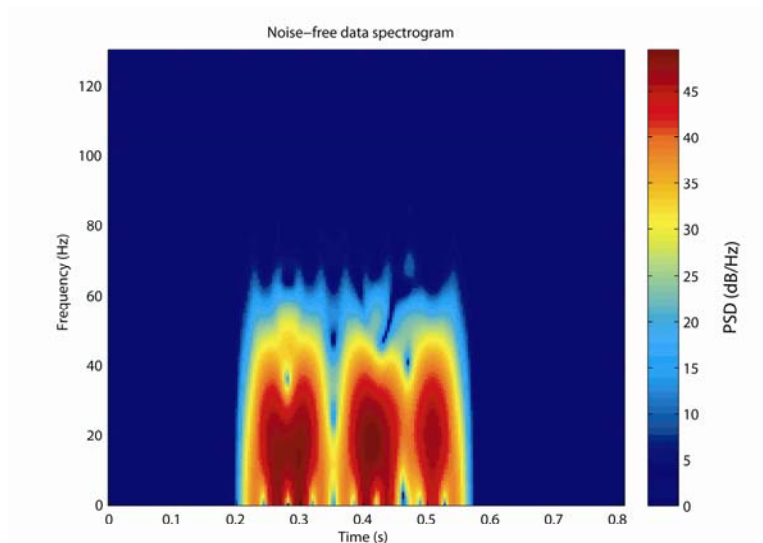


Figure 3.7. Spectrogramme de la trace numéro 27 tirée de la section sismique sans bruit.

Remarquons, que tous les événements (réflecteurs) ont été bien localisés et que la densité spectrale est concentrée autour de la position, en temps, des réflecteurs. La barre de couleurs indique la densité de l'énergie spectrale (Power spectral density PSD, en anglais) en dB/Hz. La [Figure 3.8a](#) montre le spectrogramme des données contaminées par 10 % de bruits et sa version filtrée ([Figure 3.8b](#)).

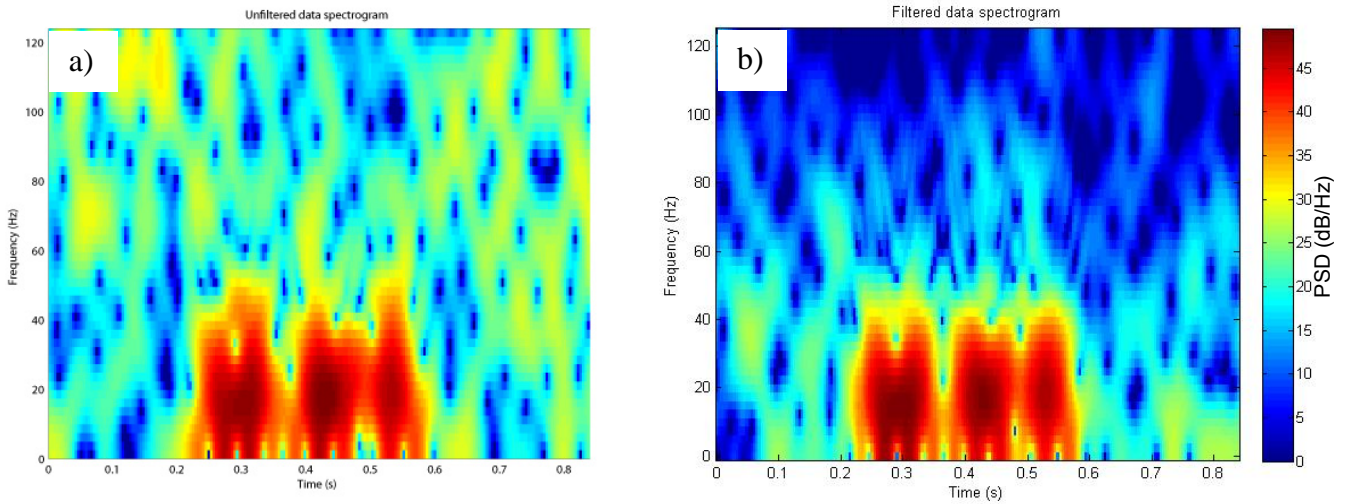


Figure 3.8. Spectrogramme de la trace numéro 27 avec 10% de bruit aléatoire ajouté a) avant et b) après filtrage.

Notons que dans le cas de données bruitées, la dispersion de l'énergie sur toutes les fréquences et tous les temps, en particulier vers les hautes fréquences, révèle la présence de bruits aléatoires, souvent identifiés par son caractère haut fréquent. De plus, la localisation des différents événements est assez subtile. En revanche, le spectrogramme des données filtrées (Figure 3.8b) montre une concentration de densité spectrale aux positionnements des différents réflecteurs sur la section sismique. Les Figures 3.9a et 3.9b représentent les spectrogrammes de données contaminées par 50% de bruits aléatoires et sa version filtrée respectivement. Les mêmes conclusions peuvent être tirées pour le cas précédent. Notons également que dans ce cas, la concentration d'énergie est moins localisée que pour le cas précédent.

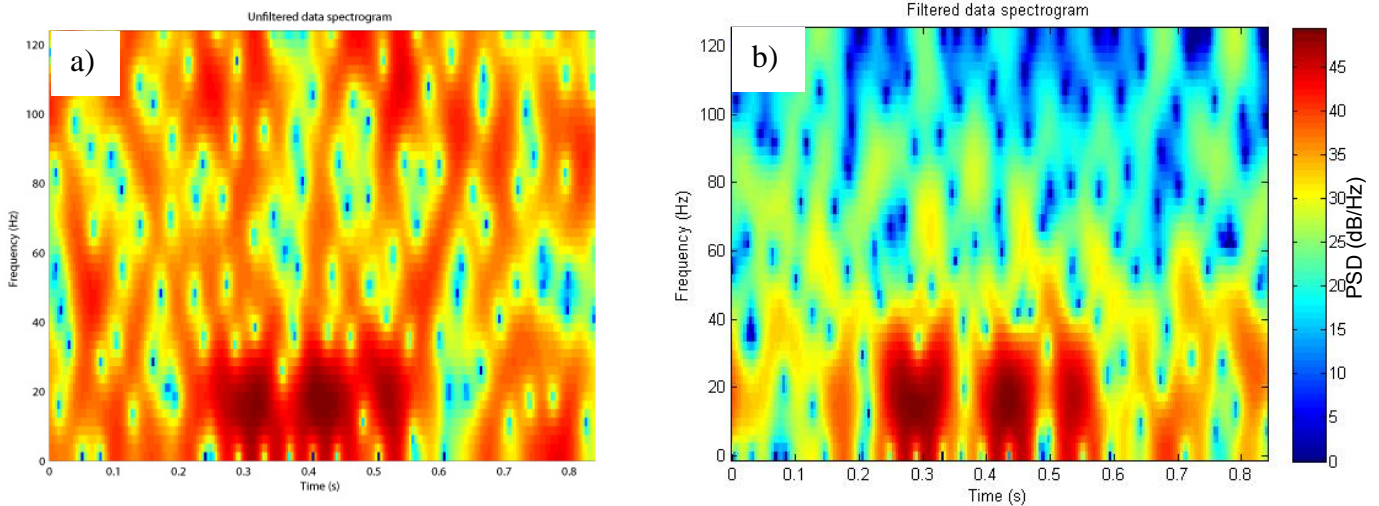


Figure 3.9. Spectrogramme de la trace numéro 27 avec 50% de bruit aléatoire ajouté a) avant et b) après filtrage.

Enfin, nous avons pu constater lors des simulations que le comportement du filtre de diffusion anisotropique est assez semblable, en termes de résultats, au filtre bilatéral de [Tomasi et Manduchi \(1998\)](#).

3.1.1.4. Influence du nombre d'itérations

L'itération est l'essence même du filtre anisotropique. L'expérience montre qu'une simplification extrême de l'image par une diffusion anisotropique importante (100 par exemple) n'améliore en rien les résultats. Afin d'évaluer l'effet du nombre d'itérations, nous avons donc réalisé le filtrage anisotropique, dans le cas d'ajout de 50% de bruit aléatoire, après 2, 5, 10, 50 et 100 itérations et les résultats sont illustrés sur les [Figures 3.10a, 10b, 10c et 10d](#) respectivement.

Après analyse des résultats, il s'avère que l'augmentation du nombre d'itération atténue les bruits en donnant une image plus lisse. Malheureusement, cela implique une distorsion du signal d'entrée, et conduit même, à une perte de certains caractères géologiques (cas de la faille sur les [Figures 3.10d et 10e](#)).

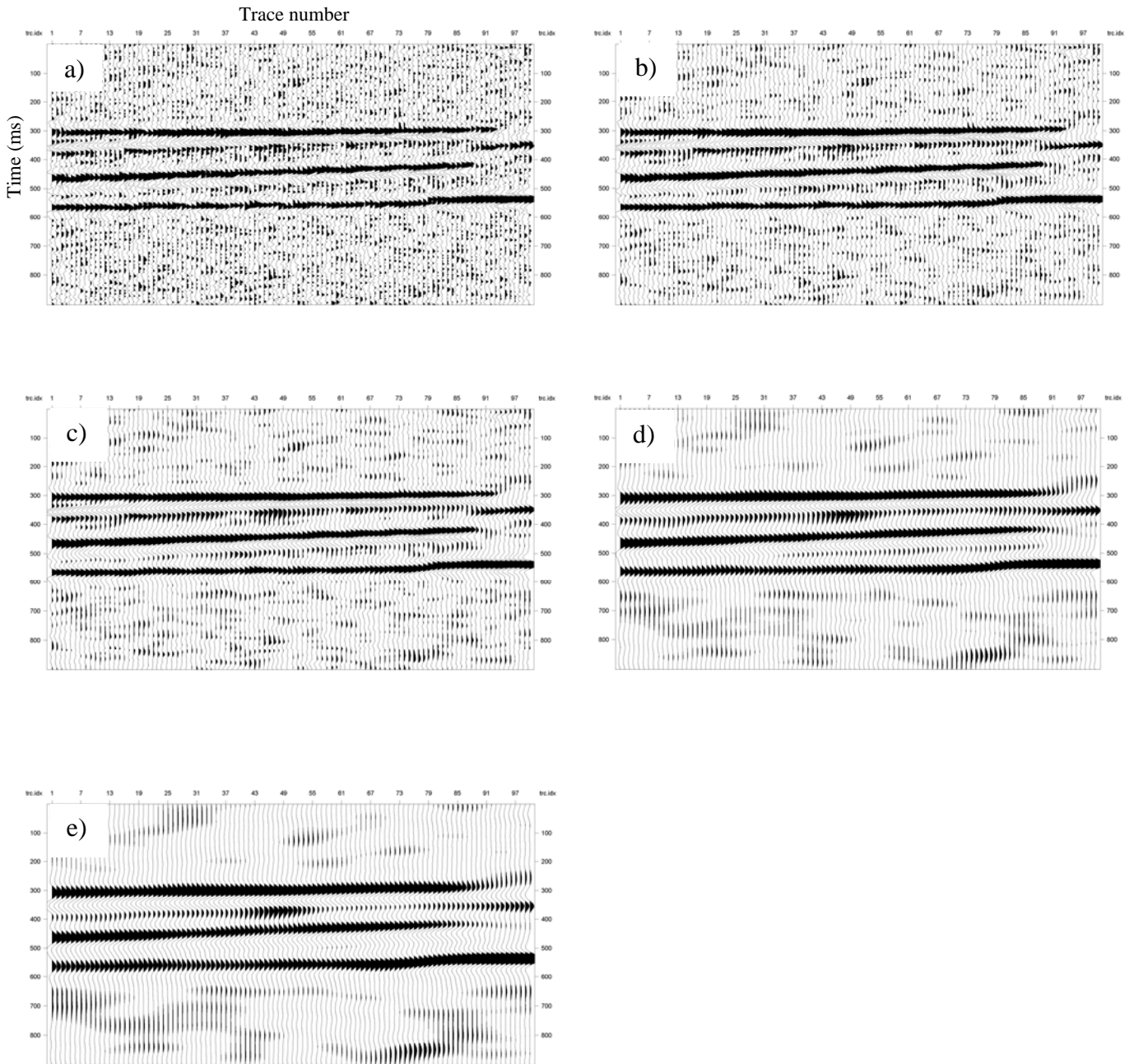


Figure 3.10. Influence du nombre d'itérations sur l'atténuation du bruit, a) 2, b) 5, c) 10, d) 50 et e) 100 itérations.

L'effet visuel du nombre d'itération constaté auparavant sur les données sismiques a été quantifié en calculant le RSB en décibels ainsi que l'erreur quadratique moyenne (EQM) ou (Mean squared error MSE, en anglais). Ainsi, nous avons utilisé les formules suivantes pour le calcul du RSB et de l'EQM :

$$RSB(dB) = 10 \log_{10} \left(\frac{\sum_i \sum_j U(i, j)^2}{\sum_i \sum_j (f(i, j) - U(i, j))^2} \right) \quad (3.1)$$

Où $U(i, j)$ représente le signal supposé sans bruit; $f(i, j)$ est le signal à la sortie du filtre anisotropique. ($i=j=1 : M \times N$, où M et N représentent les dimensions de la matrice d'entrée).

L'EQM est calculée selon la formule suivante:

$$EQM = \frac{1}{M \times N} \sum_i \sum_j (U(i) - f(i))^2 \quad (3.2)$$

Ces paramètres ont été utilisés pour guider notre choix sur le nombre d'itérations. Les [Tableaux 3.1 et 3.2](#) regroupent les résultats du filtrage en termes de RSB et d'EQM des données contaminées par 10 et 50% de bruit aléatoire en fonction du nombre d'itérations.

Nombre d'itérations	RSB en dB	EQM
1	30.76	3.42 E-4
2	35.17	2.20 E-4
3	36.4	1.94 E-4
4	35.79	2.06 E-4
5	34.39	2.38 E-4
6	32.74	2.80 E-4
7	31.08	3.31 E-4
8	29.50	3.88 E-4
9	28.03	4.49 E-4
10	26.68	5.14 E-4
11	25.44	5.82 E-4
12	24.29	6.53 E-4
13	23.24	7.26 E-4
14	22.26	8.00 E-4
15	21.36	8.76 E-4
20	17.68	0.0013
30	12.93	0.0020
40	9.97	0.0027
50	7.96	0.0033

Tableau 3.1. Influence du nombre d'itérations sur le résultat du filtrage en termes du RSB et de l'EQM de données contaminées par 10% de bruit aléatoire. Le RSB et l'EQM des données bruitées sont 21.97 dB et 8.24E-4, respectivement.

Nombre d'itérations	RSB en dB	EQM
1	1.97	0.0061
2	7.86	0.0034
3	11.87	0.0023
4	14.63	0.0017
5	16.53	0.0014
6	17.82	0.0012
7	18.65	0.0011
8	19.13	0.0011
9	19.35	0.0011
10	19.38	0.0011
11	19.25	0.0011
12	19.03	0.0011
13	18.27	0.0011
14	18.36	0.0012
15	17.97	0.0012
20	15.85	0.0015
30	12.18	0.0022
40	9.57	0.0028
50	7.71	0.0034

Tableau 3.2. Influence du nombre d'itération sur le résultat du filtrage en termes du RSB et de l'EQM de données contaminées par 50% de bruit aléatoire. Le RSB et l'EQM des données bruitées sont -7.13 dB et 0.015, respectivement.

On notera que plus le RSB est grand et que l'EQM est faible mieux sont les résultats du filtrage. Les résultats des [Tableaux 3.1 et 3.2](#) sont reportés sur les courbes de la [Figure 3.11](#).

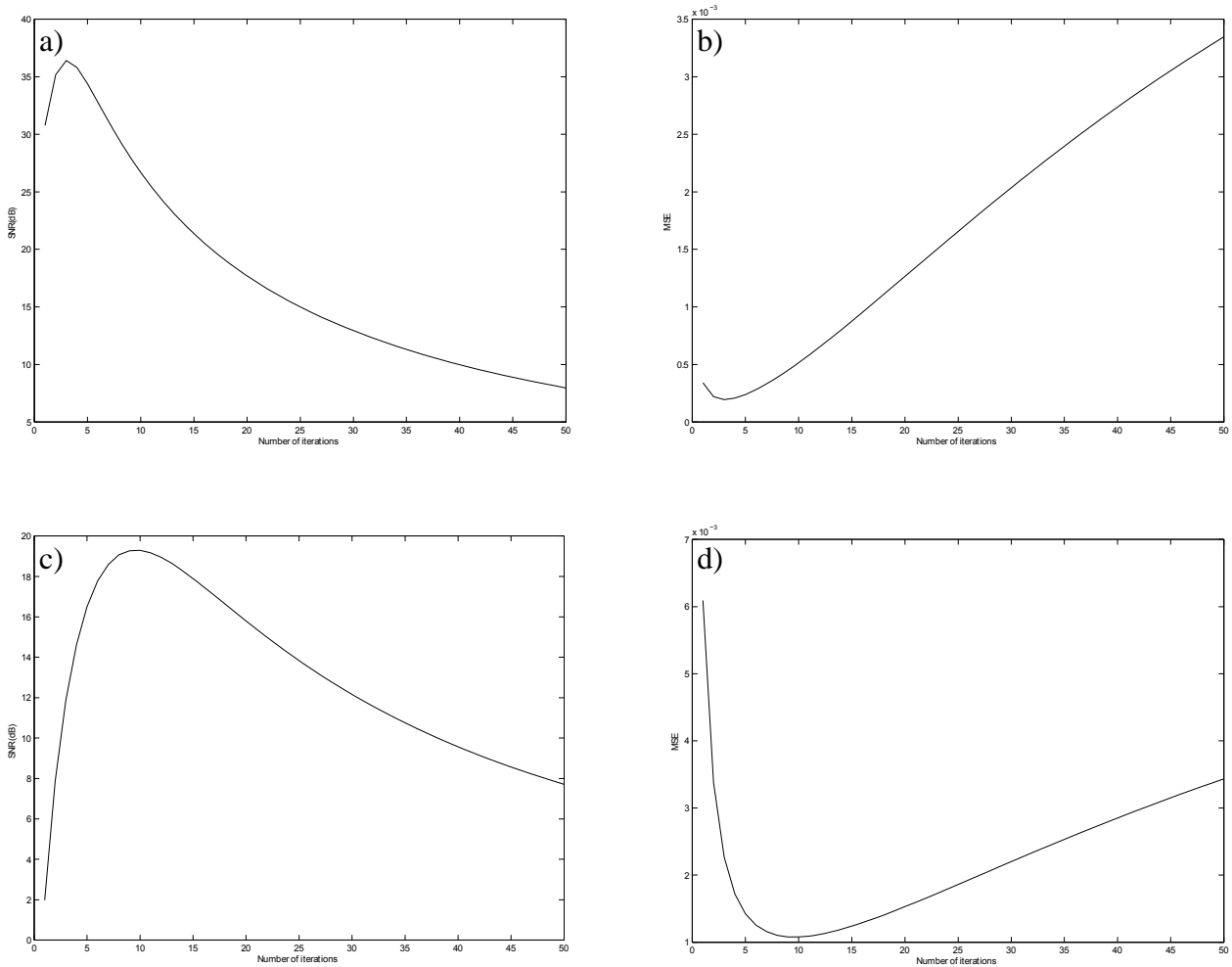


Figure 3.11. Graphes représentant l'évolution du RSB en dB et de l'EQM en fonction du nombre d'itérations. a) RSB en présence de 10 % de bruit aléatoire, b) EQM en présence de 10 % de bruit aléatoire, c) RSB en présence de 50 % de bruit aléatoire et d) EQM en présence de 50 % de bruit aléatoire.

La [Figure 3.11a](#) montre l'évolution du RSB (dB) en fonction du nombre d'itérations pour des données contaminées par 10% de bruit aléatoire. Le graphe montre une augmentation du RSB jusqu'à une valeur de 36.4 dB pour un nombre d'itération=3, ensuite une diminution jusqu'à une valeur minimale de 7.96dB lorsqu'il atteint un nombre d'itérations maximal =50. Tandis que l'EQM ([Figure 3.11b](#)) diminue jusqu'à une valeur minimale de 1.94 E-4, puis augmente jusqu'à atteindre une valeur maximale de 0.0033 lorsqu'elle atteint un nombre d'itérations maximal =50. A l'ombre de cela, nous avons fixé le nombre d'itérations à 3. De même, dans le cas où les données sont contaminées par 50% de bruit, le RSB ([Figure 3.11c](#)) augmente jusqu'à atteindre une valeur maximale de 19.31 dB après 10 itérations, puis diminue jusqu'à atteindre une valeur de 7.71 dB à un nombre d'itérations maximum=50. Tandis que l'EQM

(Figure 3.11d) diminue jusqu'à atteindre une valeur minimale de 0.11 après 7 itérations et se stabilise au niveau de cette valeur jusqu'à l'itération 13, puis augmente jusqu'à atteindre une valeur maximale de 0.0034 à l'itération 50. A la lumière de ces résultats, nous avons fixé le nombre maximum d'itérations, qui garantit un bon compromis entre atténuation du bruit et préservation du signal, pour le cas de 50% de bruit aléatoire, à 10 itérations.

3.1.1.5. Description des sections différences

Les sections différences (Figures 3.12) confirment le constat dressé par l'étude quantitative de la section précédente. Elle a montré que lorsque le nombre d'itérations dépasse un certain seuil, défini par le RSB et l'EQM, le signal d'intérêt sera gravement affecté.

Ainsi, à partir de 5 itérations on assiste à une atténuation du signal d'intérêt et le phénomène s'accroît avec l'augmentation du nombre d'itérations. Ce travail montre qu'il est nécessaire de garder un compromis entre lissage des données et préservation de la réalité géologique.

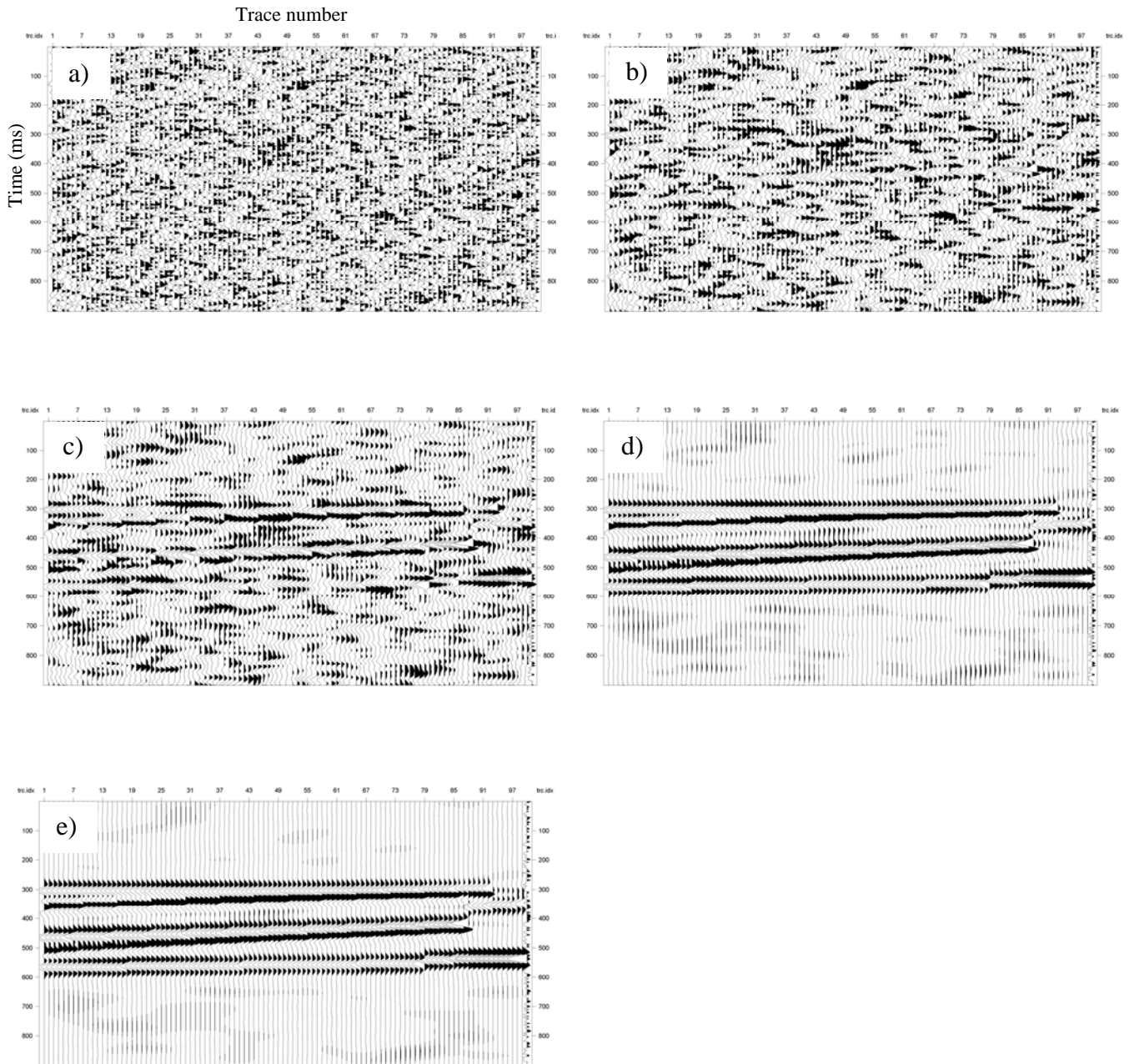


Figure 3.12. Sections différences montrant l'effet du nombre d'itérations sur le signal d'intérêt, a) 2, b) 5, c) 10, d) 50 et e) 100 itérations.

3.1.1.6 Restitution de données perdues

Une autre particularité du filtre anisotropique est testée dans l'exemple suivant. Nous avons simulé des traces mortes par la mise à zéro de l'information d'une trace dans une ou plusieurs parties de la section sismique (ce cas est fréquent en sismique, où les traces mortes sont causées par les problèmes électroniques ou par le géophone lui-même). La [Figure \(3.13\)](#) reprend la section synthétique de la [Figure 3.3](#) où certaines traces ont été mises à zéro afin de simuler des traces mortes (indiquées par des flèches). L'emplacement des traces mortes est choisi au début, au milieu et au niveau de la faille.

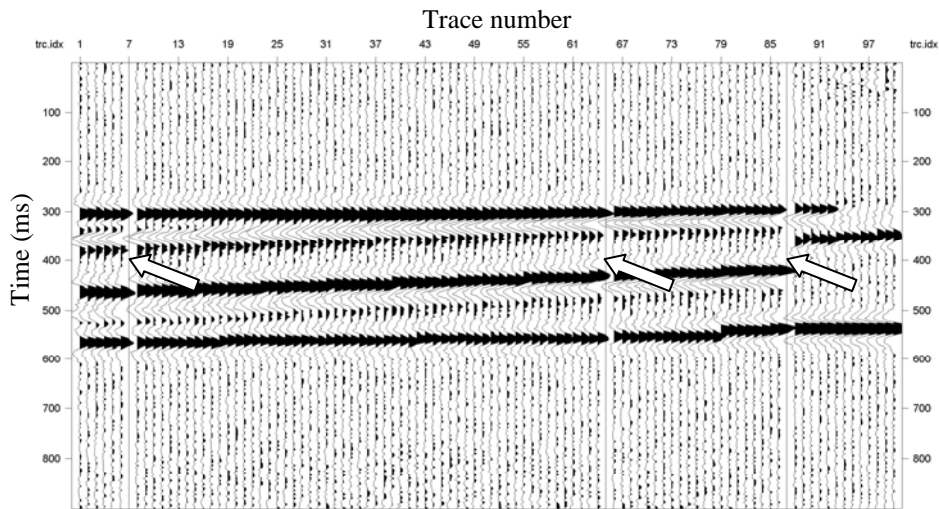


Figure 3.13. Simulation de traces mortes (Flèches)

Après passage du filtre anisotropique ([Figure 3.14](#)), nous constatons une très bonne restitution de ces traces, en termes d'amplitude et de position.

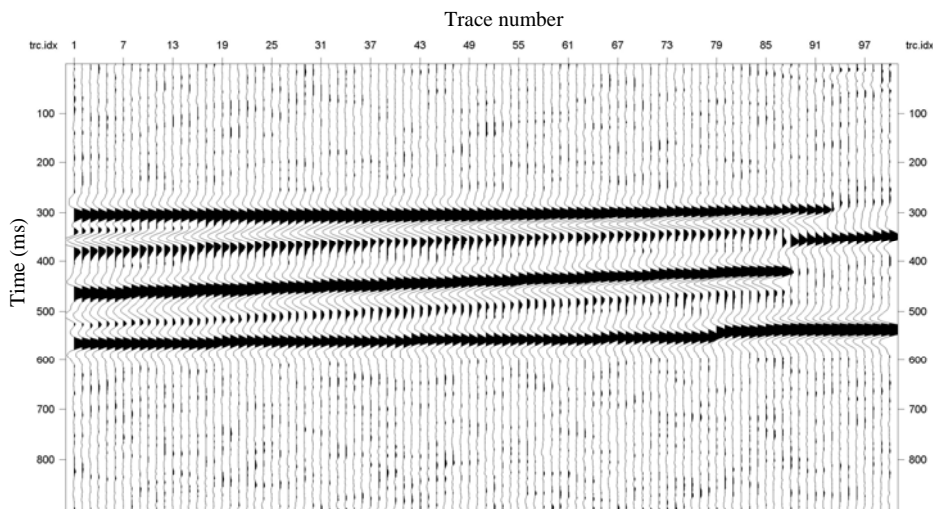


Figure 3.14. Récupération de traces mortes

L'efficacité d'une méthode est souvent testée lors des cas extrêmes. Nous avons choisi d'enlever un paquet de 5 traces (Figure 3.15), ce qui est déjà très important, et de voir le comportement du filtre en face d'un « déficit » aussi important.

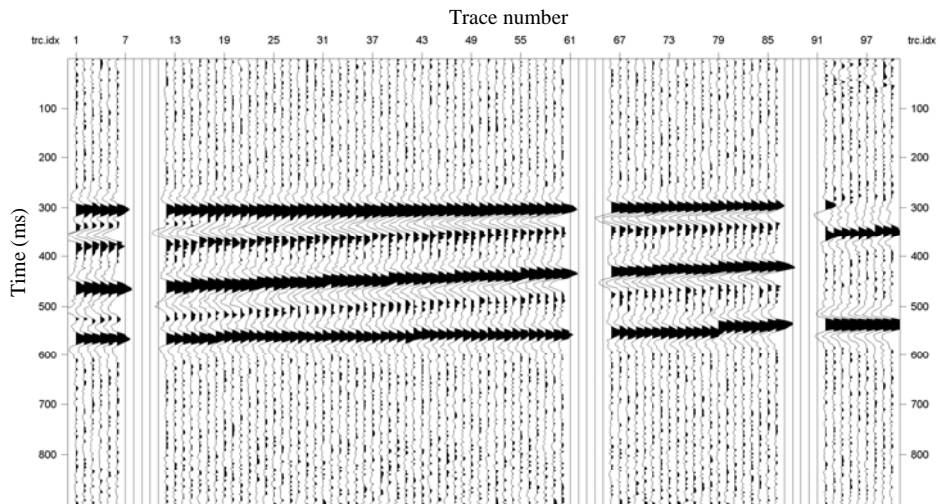


Figure 3.15. Simulation de 5 traces mortes à différentes position sur la section

Sur la Figure (3.16) le résultat est assez marquant, on remarque que la section initiale a été reproduite avec une très bonne fidélité.

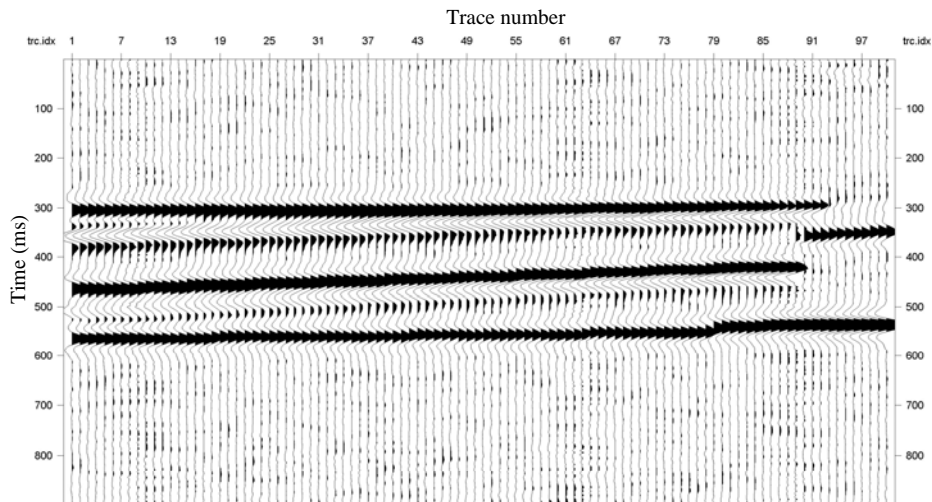


Figure 3.16. Restitution d'un paquet de traces mortes (5 traces)

A partir d'un déficit de 7 traces (Figure 3.17), le filtre anisotropique n'est plus capable de reconstruire fidèlement les traces perdues (Figure 3.18). Ceci s'explique en partie par le choix de la taille de fenêtre.

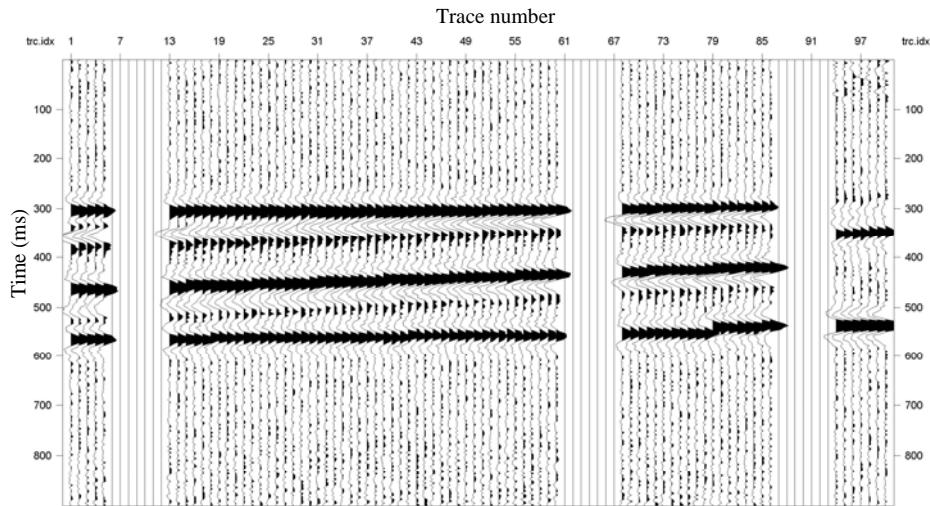


Figure 3.17. Simulation de 7 traces mortes à différentes positions sur la section (cas limite).

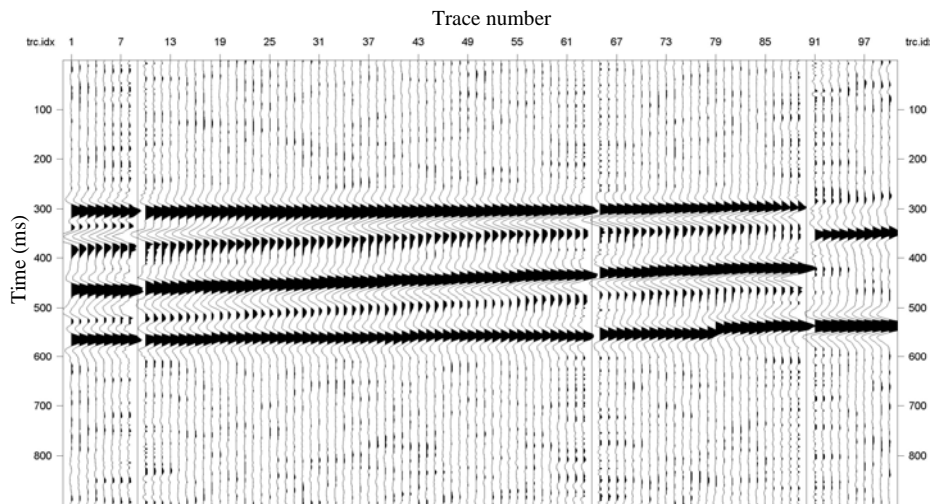


Figure 3.18. Limite de la restitution des traces mortes par le filtre anisotropique.

3.1.2. Application sur des données sismiques réelles

3.1.2.1 Présentation des données

Nous avons appliqué le filtre anisotropique, en plus des simulations sur les données synthétiques, sur une section sismique réelle 2D sommée, provenant du Sud algérien (Figure 3.19). Pour des raisons de confidentialité, nous avons évité de citer la région d'étude et le nom du profil (données encore non traitées). Nous nous contentons seulement de donner les paramètres d'acquisition : l'intertrace utilisé est de 25m, l'inter-tir est de 25m et le pas d'échantillonnage est égal à 2ms. La bande fréquentielle est de 8-72Hz. Les sources utilisées sont des vibrateurs.

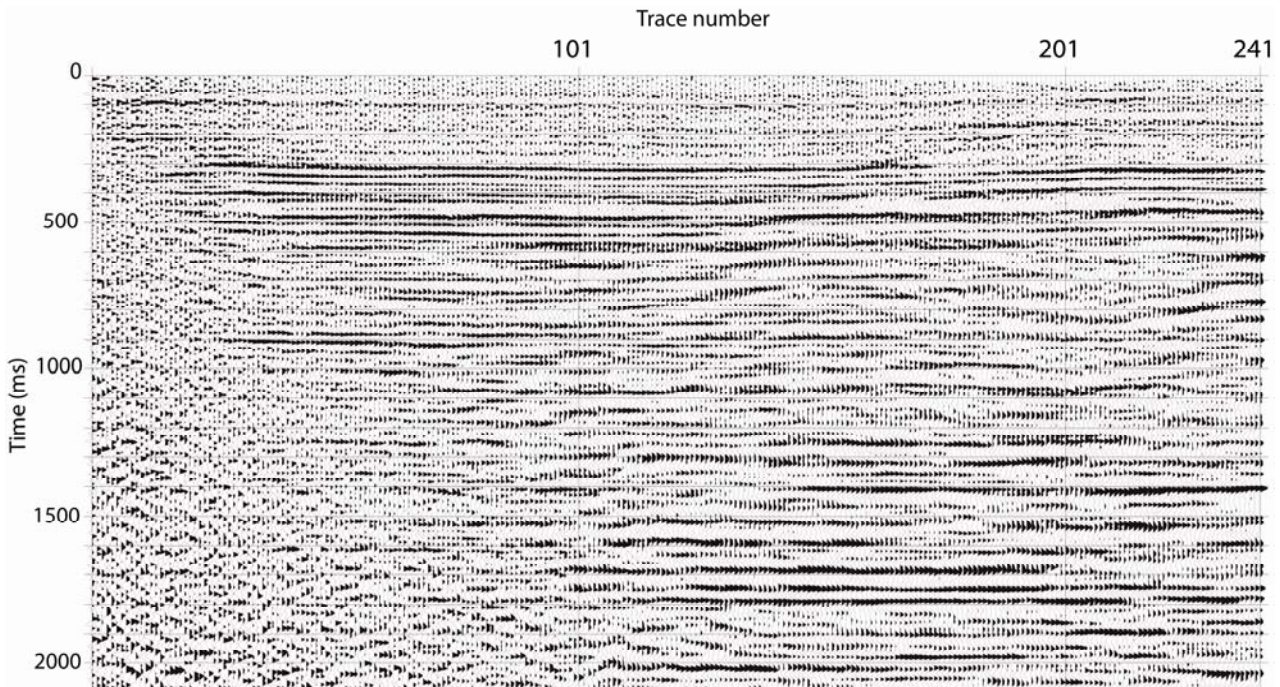


Figure 3.19. Section sismique réelle sommée brute (Raw stack).

Malgré la sommation, la section sismique de la Figure 3.19 contient encore un fort taux de bruit aléatoire. Nous avons toutefois décidé de n'utiliser aucune autre étape de traitement supplémentaire.

3.1.2.2. Paramétrisation

Dans un souci d'atténuer le bruit, tout en veillant à ne pas affecter les amplitudes des réflecteurs, nous avons choisi les paramètres définissant le filtre anisotropique après plusieurs tests comme suit :

- constante de diffusion = 50
- le pas espace-temps = 0.15
- fonction de diffusion = w_I
- nombre maximum d'itérations = 3

3.1.2.3. Analyse et résultats

A l'issue du filtrage, il s'est avéré qu'au bout de 3 itérations de très bons résultats ont été achevés (Figure 3.20).

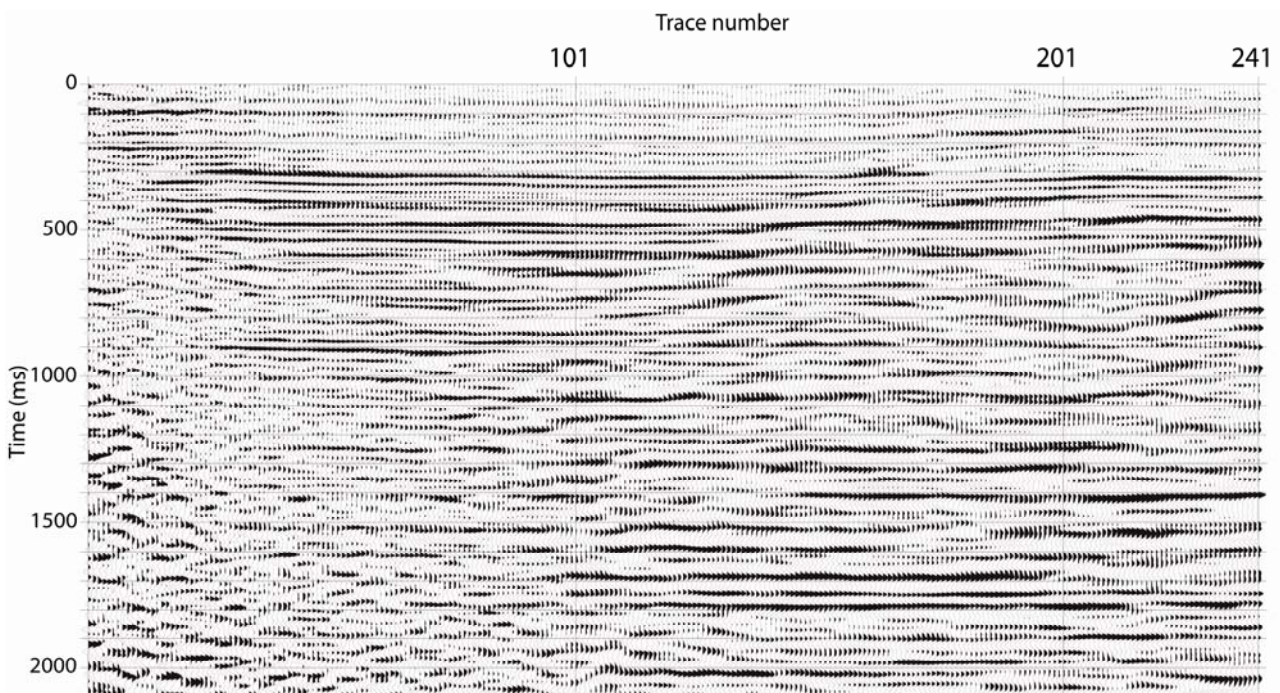


Figure 3.20. Section stack après filtrage anisotropique.

Les améliorations les plus marquées sont visibles sur les réflecteurs superficiels et profonds, là où la continuité des réflecteurs est faible. L'autre point qui mérite d'être soulevé est qu'aucun changement notable n'a été enregistré sur la forme du signal d'intérêt. Il est à noter que de meilleurs résultats peuvent être obtenus en augmentant le nombre d'itérations. Néanmoins, il est clair qu'il faudra garder un compromis entre atténuation du bruit et

déformation du signal. Un zoom sur une fenêtre entre les CDP 264-418 au temps double 0-500ms montre une nette amélioration, en terme de RSB, de la continuité des réflecteurs (Figure 3.21a), particulièrement au temps double 390 et 450ms, comparée à la section non filtré (Figure 3.21b).

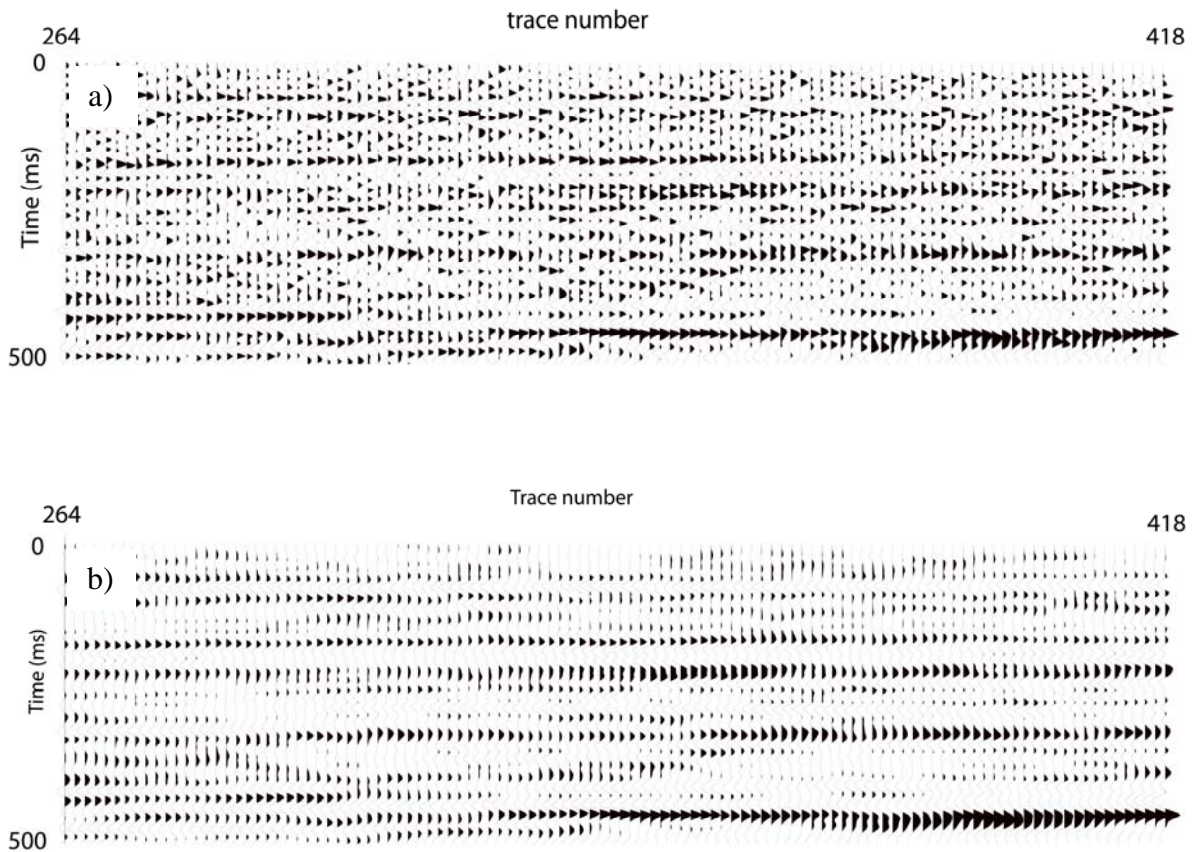


Figure 3.21. Agrandissement dans une fenêtre de temps a) avant et b) après filtrage anisotropique.

Enfin, nous avons comparé les résultats du filtre anisotropique avec ceux d'un filtre de type f-k (Figure 3.22). Il est clair que les performances de notre filtre sont supérieures au filtre f-k. Les bruits présents dans la partie supérieure de la section sont encore visibles sur la section traitée par le filtre f-k, ce qui n'est pas le cas avec la section traité par filtre anisotropique (Figure 3.20).

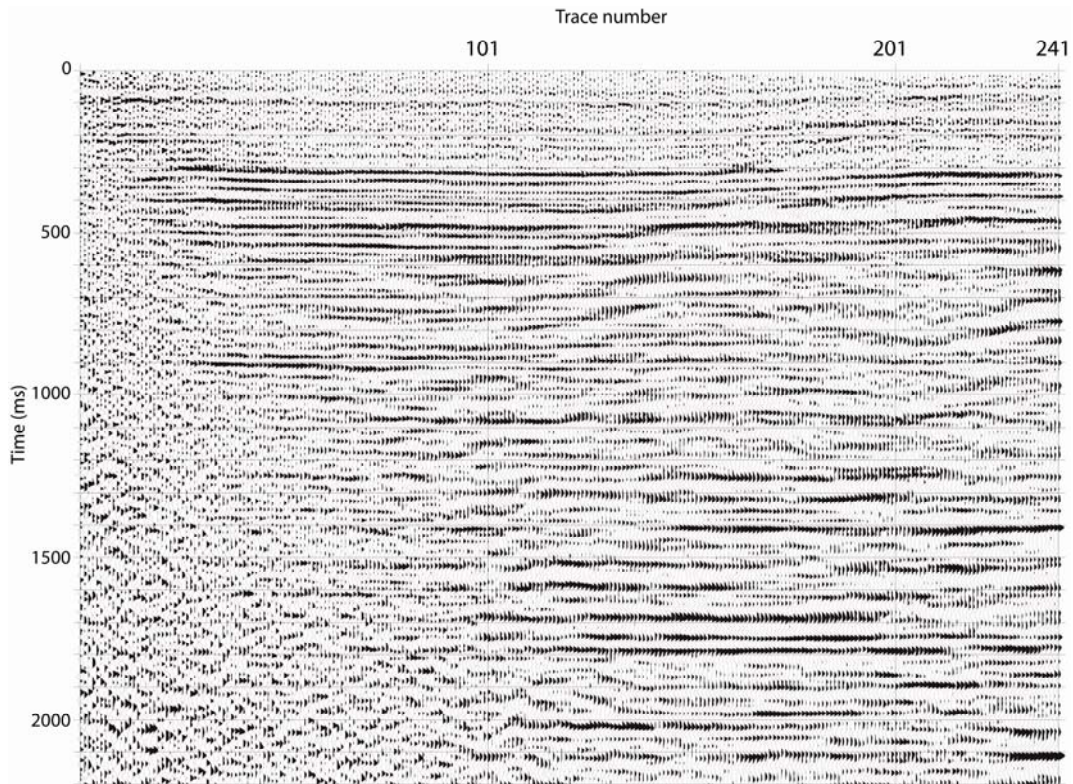


Figure 3.22. Section réelle après filtrage dans le domaine f-k

3.2. Conclusion

Le filtre de diffusion anisotropique est un outil puissant pour l'atténuation des bruits et la préservation des bords. Nous avons utilisé ce filtre pour l'atténuation des bruits contaminants des données sismiques synthétiques 2D et des données réelles. Les résultats obtenus durant ce chapitre avaient montré une nette amélioration du RSB, en plus des améliorations apportées à la continuité des réflecteurs et à la détection des phénomènes géologiques. L'analogie que nous avons dressée avec le traitement de l'image a montré l'applicabilité du filtre anisotropique au domaine du traitement des données sismiques. Nous avons aussi pu constater que le filtre était en mesure d'extraire le signal d'intérêt noyé dans du bruit, de même qu'il a permis la détection d'une faille, même en présence d'un fort taux de bruit. Les tests intensifs menés sur les données synthétiques, contaminées par des niveaux de bruits différents, et les calculs du RSB et de l'EQM, ont montré que le nombre d'itération est un facteur clé qu'il faudra bien paramétrer, faute de quoi, le filtrage risque de distordre les données, et ainsi fausser l'interprétation. Nous avons aussi utilisé le filtre comme moyen de récupérer des traces mortes. Nous avons pu constater qu'avec un paquet de 5 traces mortes, le filtre était encore en mesure de reconstituer la section sismique, avec une grande fidélité. L'application

Chapitre 3. Atténuation des bruits par le filtre de diffusion anisotropique

sur des données sismiques réelles, a montré que le filtre de diffusion anisotropique offre de meilleurs résultats comparativement au filtre f-k. Nous croyons que, potentiellement, ce filtre offrira une alternative aux méthodes existantes, souvent complexes et gourmandes en temps, et qu'il bénéficiera d'une attention particulière dans le futur.

Chapitre 4

Atténuation du bruit sismique par le filtre trilatéral

Deuxième volet des trois filtres que nous avons utilisés, le filtre trilatéral a été appliqué à des données sismiques synthétiques et réelles. L'objectif de l'utilisation de ce filtre est de supprimer le bruit, tout en gardant les bords inchangés. En sismique, il se traduit par une atténuation des bruits aléatoires et/ou cohérent, une meilleure résolution et une amélioration de la continuité des réflecteurs. Ce filtre se base sur une distribution préférentielle de poids. L'idée est que tous les échantillons proches dans l'espace, similaires en magnitude et possédant une certaine statistique (ROAD) se voient attribués de forts poids, et sont ainsi privilégiés. Le filtre trilatéral est un filtre de voisinage et se base sur le filtre bilatéral de Tomasi et Manduchi (1998).

4.1. Introduction

L'applicabilité du filtre trilatéral (TL) a été testée sur des données sismiques synthétiques et réelles. En traitement de l'image, ce filtre est devenu un outil de base. Malheureusement, en sismique son utilisation est quasiment nulle. Ce filtre est simple à implémenter, en revanche, ses performances sont excellentes en présence de bruits aléatoires et gaussiens indépendamment des fréquences, en plus de son aptitude à préserver les bords. Dans la suite de cette application, nous allons utiliser une version modifiée basée sur une statistique d'ordre appelée ROAD (Rank-Ordered Absolute Differences) d'après Garnett et al (2005). Les résultats après application du filtre sur des données sismiques synthétiques et réelles témoignent de l'efficacité du filtre trilatéral à réduire considérablement différents types de bruits.

4.2. Simulations

4.2.1. Données synthétiques

La validité du filtre trilatéral à atténuer le bruit, tout en gardant le signal d'intérêt inchangé, a été testée sur des données synthétiques générées par le programme libre Seismic unix (CMP/SU) développé au « Center of Wave phenomena » à la « Colorado School of Mines ». Les données synthétiques en question sont représentées sous forme d'un shot gather. Le shot

gather est composé de quatre événements situés aux temps doubles 310, 470, 530 et 700 ms (Figure 4.1).

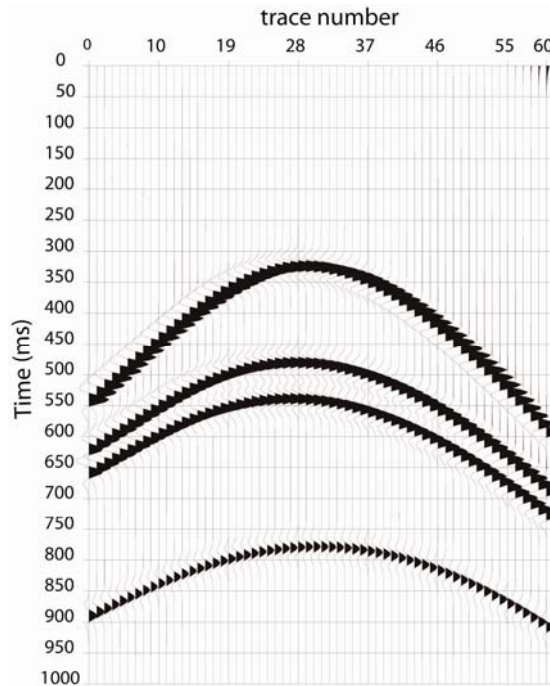


Figure 4.1. Shot gather synthétique (sans bruit) montrant quatre événements sismiques.

Le shot gather de la Figure 4.1 est composé de 60 traces obtenues par convolution de la série de réflectivité avec une ondelette de Ricker de fréquence centrale égale à 25 Hz. Chaque trace est composée de 251 échantillons, avec un pas d'échantillonnage de 4 ms. Nous avons aussi réalisé des codes Matlab pour les opérations de filtrage et la réalisation des spectrogrammes. Nous avons ajouté aux données de la Figure 4.1 du bruit aléatoire, normalement distribué, en utilisant la commande matlab « *noise* » (Margrave, 1991). Différents niveaux de bruit ont été ainsi ajoutés (5%, 10% et 50%) et les résultats sont illustrées sur les Figures 4.2a, 4.2b, et 4.2c respectivement. Notre objectif était de montrer l'efficacité de notre procédé de filtrage même en présence de bruits de niveau très élevé, comme c'est le cas de la Figure 4.2c.

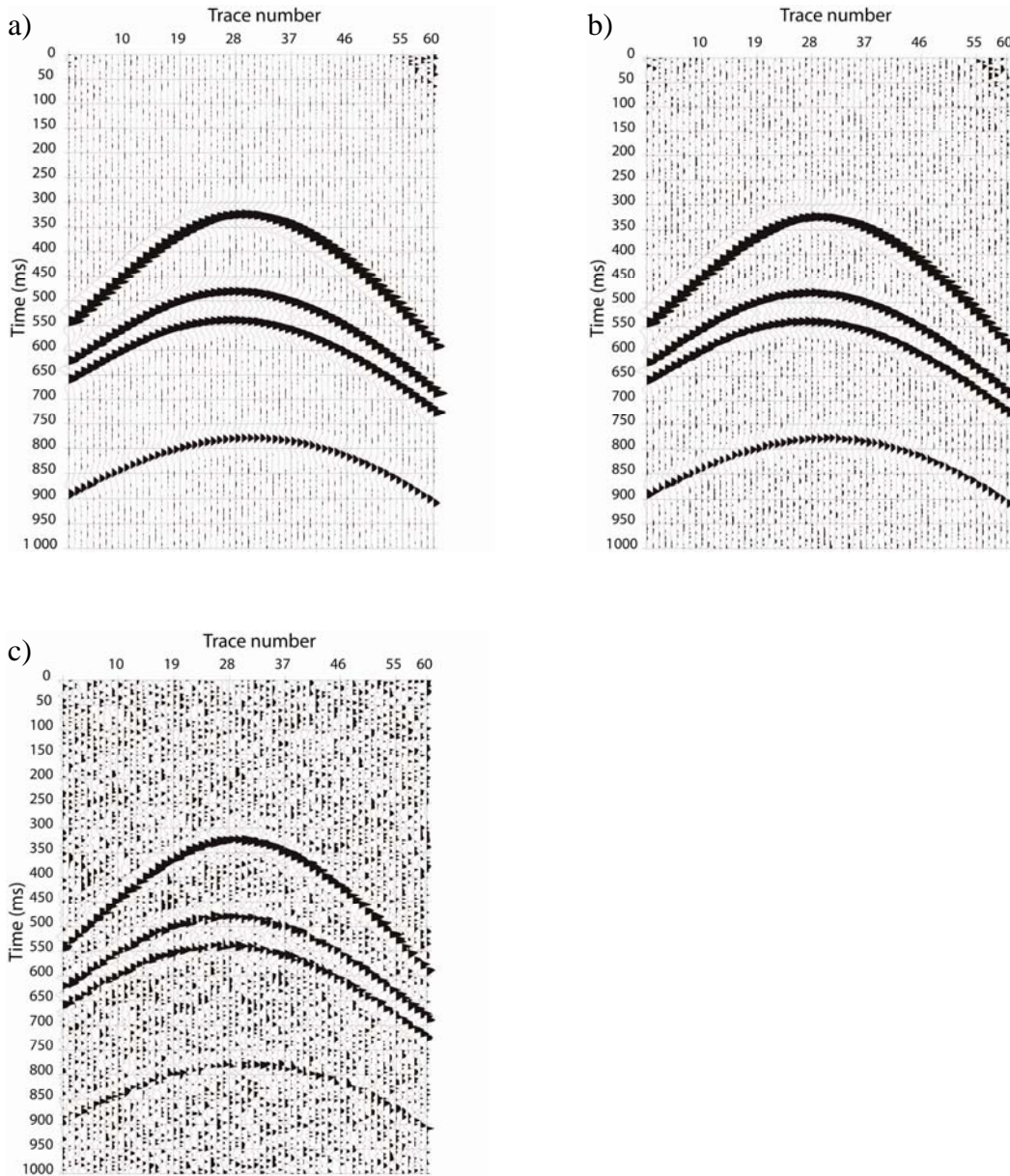


Figure 4.2. Shot gather contamin   par a) 5%, b) 10% et c) 50% de bruit al  atoire.

4.2.2. R  sultats et commentaires

Lors des simulations, nous avons utilis   une fen  tre glissante de taille 3 x 3. Notons qu'il n'existe pas de r  gle g  n  rale pour le choix de la taille de la fen  tre, car cela d  pend du r  sultat escompt  . D'une mani  re g  n  rale, on choisit une fen  tre carr  e de dimension impaire (3x3, 5x5, 7x7, etc.) afin de garantir une bonne sym  trie par rapport    l'  chantillon central ζ . Les diff  rents param  tres du filtre (σ_s , σ_d , σ_I et σ_J) ont   t   choisis apr  s plusieurs tests, et ont   t   modifi  s en fonction du niveau de bruit et des r  sultats escompt  s. Avant d'appliquer notre filtre directement sur les donn  es brutes, nous avons utilis   l'op  ration de

rembourrage «ou *padding*, en anglais ». Cette opération consiste à rajouter des valeurs nulles autour de la matrice d'entrée afin d'éviter que la dimension de la matrice de sortie ne soit inférieure à celle de la matrice d'entrée. Le « padding » est retiré à la fin du filtrage.

La Figure 4.3 illustre les résultats du filtrage. Comme attendu, après seulement une itération (cas de 5% et 10% de bruit aléatoire), le bruit présent sur les shots gathers a été fortement atténué. Cependant, en présence de 50% de bruit, nous avons augmenté le nombre d'itérations à 3 (par la suite nous discuterons des critères du choix du nombre d'itérations).

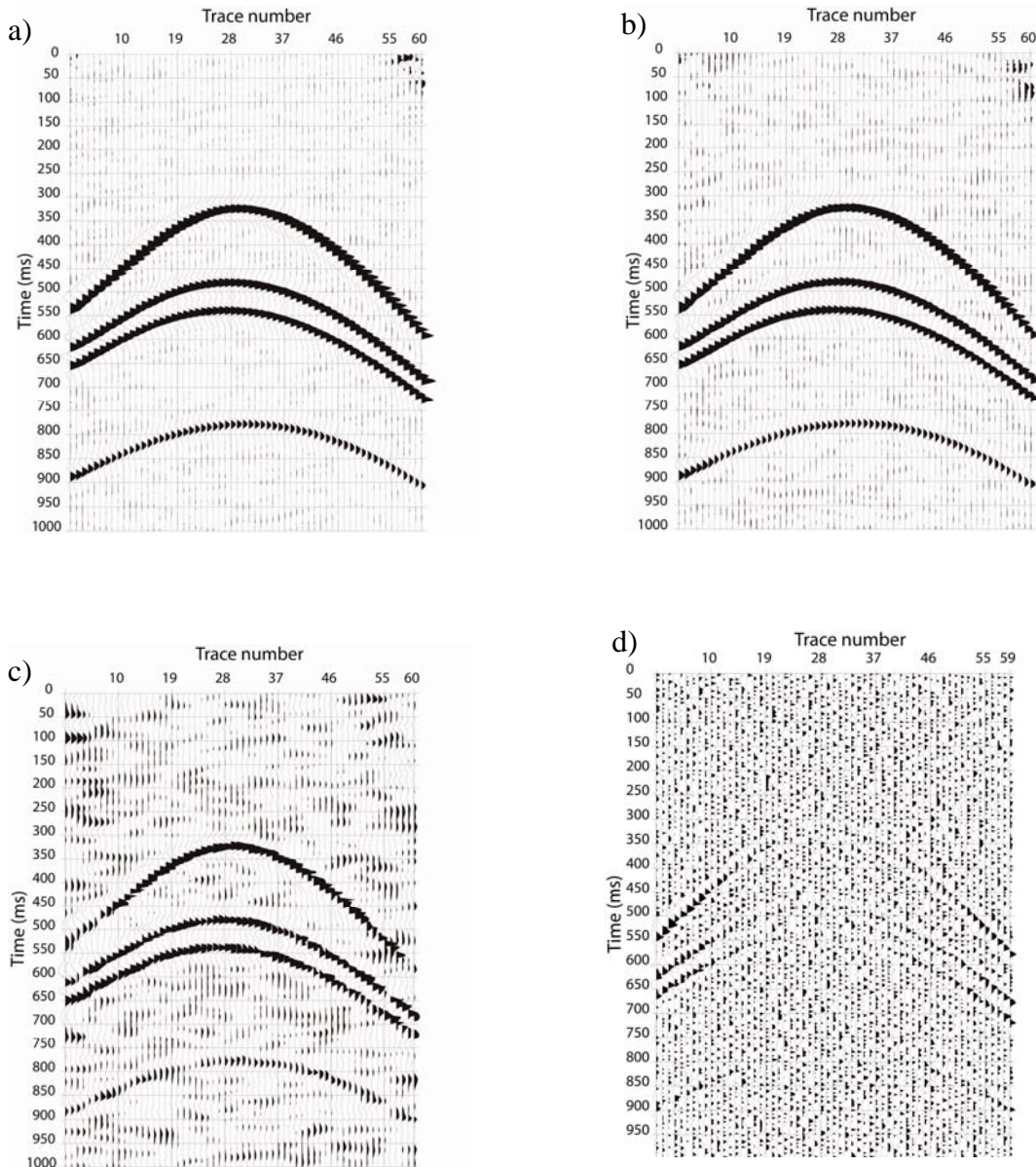


Figure 4.3. Résultat du filtrage dans le cas de a) 5%, b) 10% et c) 50% de bruit aléatoire ajouté et d) résidu de filtrage du cas (c).

Il est à remarquer la bonne continuité des réflecteurs avec une bonne préservation des amplitudes (Figure 4.3a et 4.3b). En revanche, une légère dégradation dans la continuité, au niveau du premier réflecteur, est visible près de la trace numéro 49 (Figure 4.3c). D'une manière générale, tous les réflecteurs montrent une bonne continuité avec une meilleure préservation des amplitudes. Cette dernière caractéristique est très importante lors des études en amplitudes préservées. En pratique, les réflecteurs profonds sont les plus vulnérables aux pertes de résolution et aux effets de la divergence sphérique et à l'absorption inélastique. Néanmoins, quelques artefacts de filtrage sont visibles au temps double 50ms du côté haut droit du shot gather (Figure 4.3a et 4.3b). Ces artefacts peuvent être aisément supprimer par l'application d'une fonction de mute. La Figure 4.3d est le résultat de la soustraction, échantillon par échantillon, des données bruitées de la Figure 4.2c de celles du shot filtrées de la Figure 4.3c. Le bruit aléatoire, dans l'ensemble, a été fortement atténué. Toutefois, on remarque une légère dégradation du signal, en particulier pour les réflecteurs superficiels.

4.2.3. Influence du nombre d'itérations

Comme il a été déjà mentionné, le filtre trilatéral est basé sur le filtre bilatéral. Néanmoins, les performances du filtre peuvent être substantiellement améliorées par l'utilisation itérative. Lors de nos simulations, tous les résultats ont été obtenus après une seule itération, à l'exception du cas où nous avons ajouté 50% de bruit aléatoire, où il a été nécessaire d'utiliser 3 itérations. Afin d'améliorer les performances du filtre, nous l'avons utilisé itérativement sur les données de la Figure 4.2a. L'utilisation itérative suppose que la sortie précédente du filtre redevient son entrée actuelle. Les résultats après 2, 5, 10 et 50 itérations sont illustrés sur la Figure 4.4. La Figure 4.4a visualise les effets après 2 itérations. Une amélioration substantielle en termes de RSB est visible. L'amélioration est accentuée en augmentant le nombre d'itérations à 5 (Figure 4.4b), en revanche, une légère dégradation dans la continuité des réflecteurs est observée. Après 10 itérations, on note que le signal d'intérêt commence à se déformer (Figure 4.4c), en particulier près des réflecteurs superficiels. Néanmoins, la résolution est toujours acceptable, et il est encore possible de bien positionner les différents réflecteurs. Après 50 itérations (Figure 4.4d), on perd complètement la résolution, et il devient impossible d'interpréter les données. Il en ressort de ces résultats, qu'il doit y avoir un compromis entre atténuation de bruit et lissage excessif des données.

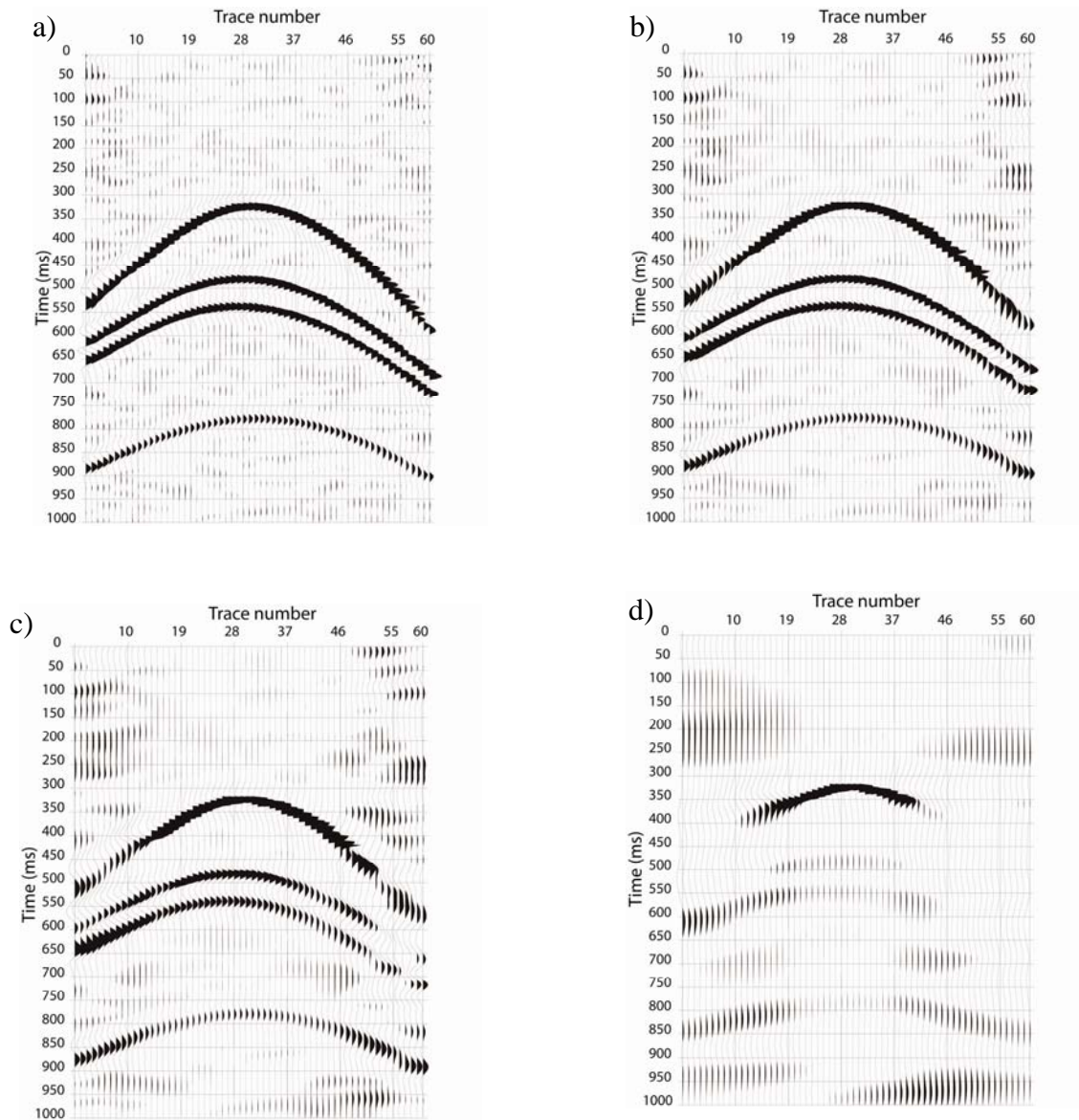


Figure 4.4. Influence du nombre d'itérations sur le résultat du filtrage après a) 2, b) 5, c) 10 et d) 50 itérations.

Une évaluation quantitative des performances du filtre trilatéral a été réalisée en calculant le rapport RSB en décibels ainsi que l'EQM par rapport aux données sans bruit, en utilisant les formules suivantes :

$$RSB(dB) = 10 \log_{10} \frac{\sum_{i=1}^N \sum_{j=1}^M U(i, j)^2}{\sum_{i=1}^N \sum_{j=1}^M (f(i, j) - U(i, j))^2} \quad (4.1)$$

$$EQM = \frac{1}{MN} \sum_{i=1}^N \sum_{j=1}^M (U(i, j) - f(i, j))^2 \quad (4.2)$$

Où $U(i, j)$ sont les données sans bruit (données de la [Figure 4.1](#)), $f(i, j)$ représente la sortie du filtre trilatéral, et M et N désignent la dimension de la matrice d'entrée.

Les résultats sont regroupés au [Tableau 4.1](#).

Nombre d'itérations	5% de bruit aléatoire ajouté		10% de bruit aléatoire ajouté		50% de bruit aléatoire ajouté	
	RSB	EQM	RSB	EQM	RSB	EQM
1	32.15	0.0029	29.11	0.0039	1.29	0.063
2	22.48	0.0076	21.62	0.0083	7.97	0.032
3	17.40	0.012	16.97	0.013	10.16	0.026
4	14.17	0.017	14.01	0.017	10.08	0.026
5	11.91	0.021	11.92	0.021	9.37	0.028
10	6.60	0.034	6.95	0.036	5.97	0.039
15	4.62	0.045	4.90	0.044	4.35	0.046
20	3.71	0.049	3.88	0.049	3.39	0.051
30	2.74	0.054	2.86	0.054	2.20	0.058
40	2.19	0.057	2.32	0.057	1.60	0.061
50	1.82	0.060	1.98	0.059	1.22	0.064
Données bruitées	35.83	0.0020	21.97	0.0080	-7.13	0.146

Tableau 4.1. Résultat du filtrage en termes de RSB et de l'EQM en fonction du nombre d'itérations.

Les résultats du [Tableau 4.1](#) sont fonction du nombre d'itérations. Nous considérons que plus le RSB est grand et que plus l'EQM est faible, meilleurs sont les résultats du filtrage. Il est évident qu'avec l'augmentation du nombre d'itérations les résultats sont encore plus lisses, donc visuellement plaisants bien qu'ils résultent d'une dégradation du signal. Cette dégradation se traduit par une diminution du RSB et d'une augmentation de l'EQM (voir [Tableau 4.1](#)). Nous jugeons donc utile d'utiliser la forme itérative du filtre trilatéral avec prudence lors de la manipulation de données traitées en amplitudes préservées.

La comparaison est plus facile en traçant les données du [Tableau 4.1](#) en fonction du nombre d'itérations, sous forme de courbes ([Figure 4.5](#)).

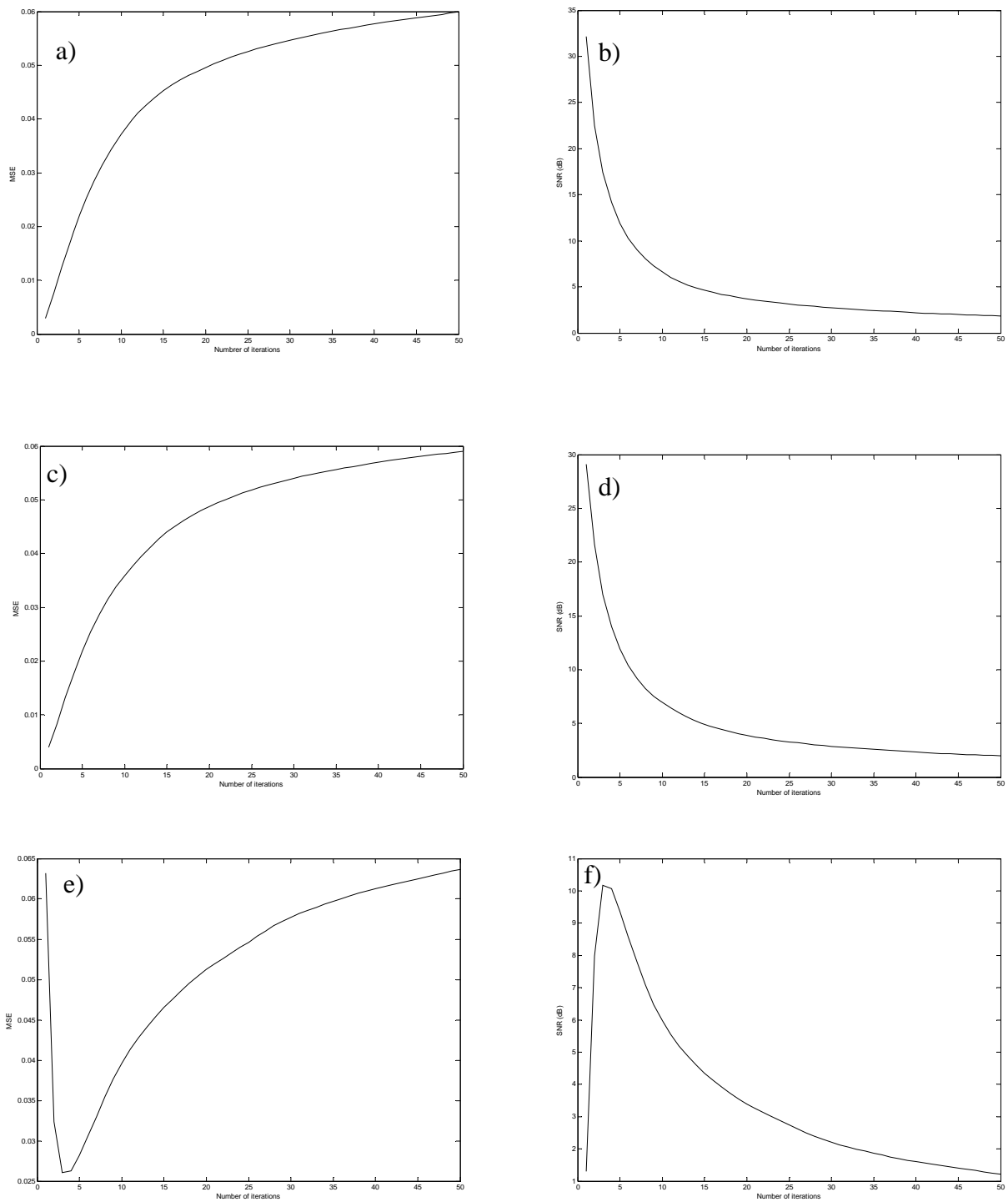


Figure 4.5. Courbes du RSB et de l'EQM RSB en présence de 5% (a), 10% (c) et 50% (e) de bruit aléatoire, EQM en présence de 5% (b), 10% (d) et 50% (f) de bruit aléatoire.

Comme nous pouvons le remarquer, les courbes du RSB et de l'EQM, correspondant aux données contaminées par 5% et 10% de bruit aléatoire, sont similaires (Figures 4.5a, 5b, 5c et 5d). Dans ce cas, nous observons que la courbe du RSB augmente monotonement, tandis que

celle de l'EQM diminue monotonement. Ce comportement nous a gênés lors du choix du nombre d'itérations. Nous avons donc opté pour une seule itération, afin de ne pas distordre les données. En présence de 50% de bruit aléatoire (Figure 4.5e et 4.5f) les courbes du RSB et de l'EQM sont plus explicites. Ainsi, la courbe de l'EQM (Figure 4.5e) diminue jusqu'à une valeur minimale de 0.026, ensuite augmente jusqu'à une valeur maximale de 0.064 lorsqu'elle atteint un nombre d'itérations=50. Tandis que le RSB (Figure 4.5f) augmente jusqu'à atteindre une valeur maximale de 10.16 dB à un nombre d'itérations=3, puis diminue jusqu'à une valeur minimale de 1.22dB à un nombre d'itérations=50. C'est ainsi que notre choix du nombre d'itérations a été fixé à 3 en présence de 50% de bruit aléatoire, car ce nombre garantit une EQM minimale et un RSB maximal.

4.2.4. Analyse spectrale

Un autre moyen d'analyser les résultats est de les étudier dans le domaine temps-fréquence. A cette fin, nous avons fait recours à la fonction matlab « *spectrogram* » laquelle calcule la transformation de Fourier (appelée aussi Short-Time Fourier Transform STFT) d'un signal à l'intérieur d'une fenêtre ($t_i-w/2, t_i+w/2$), où w est la longueur de la fenêtre. Un spectrogramme associe à chaque temps ' t_i ', d'un signal donné, son spectre de fréquence. Il décrit, en quelque sorte, l'évolution de l'énergie vis-à-vis des fréquences en fonction du temps.

La Figure 4.6 illustre le spectrogramme de la trace numéro 29, issue du shot gather sans bruit de la Figure 4.1, utilisant une fenêtre de longueur 520ms. Sur cette figure, l'axe des abscisses représente le temps, l'axe des ordonnées représente les fréquences et chaque pixel à l'intérieur du spectrogramme donne l'intensité pour une fréquence donnée en fonction du temps. La barre de couleurs indique la densité spectrale (PSD) en dB/Hz .

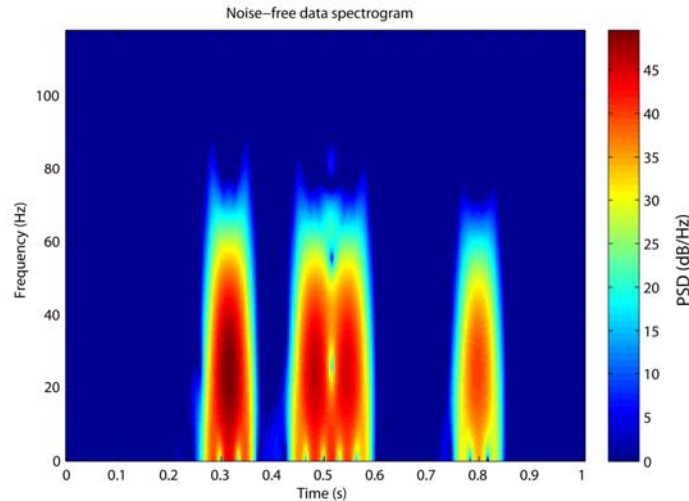


Figure 4.6. Spectrogramme de la trace numéro 29 tirée du shot gather de la Figure 4.1

Remarquons que tous les événements sismiques sont présents et qu'ils sont bien localisés en temps. Notons aussi que le maximum d'énergie se concentre autour de la fréquence 25 Hz. L'énergie décroît avec la fréquence et le temps, en particulier pour le réflecteur profond (temps double 770ms) qui s'explique, en partie, par le phénomène de divergence sphérique et de l'absorption.

La Figure 4.7 montre les spectrogrammes des données bruitées avec 5%, 10% et 50% de bruit aléatoire (Figures 4.7a, 4.7c et 4.7 e, respectivement). Dans ce cas, l'énergie est étendue sur toutes les fréquences et est présente sur tous les temps. Cette distribution éparse témoigne de la présence de bruit aléatoire, en particulier près des hautes fréquences. On note donc que les événements sismiques sont complètement noyés dans le bruit et qu'il est difficile de préciser l'endroit exact de la concentration de la densité spectrale. Dans le cas de la présence de 50% de bruit (Figure 4.7e), la localisation des événements sur le spectrogramme est très difficile, voire impossible (cas du réflecteur profond). L'application du filtre trilatéral sur ces données bruitées produit des spectrogrammes assez proches du spectrogramme des données sans bruit (Figures 4.7b, 4.7d et 4.7f). L'amélioration est assez notable. Dans ce cas, la concentration d'énergie correspond à la position des événements en temps et est autour de 25 Hz. Donc, le contenu fréquentiel des données filtrées n'a pas beaucoup changé comparativement aux données non bruitées.

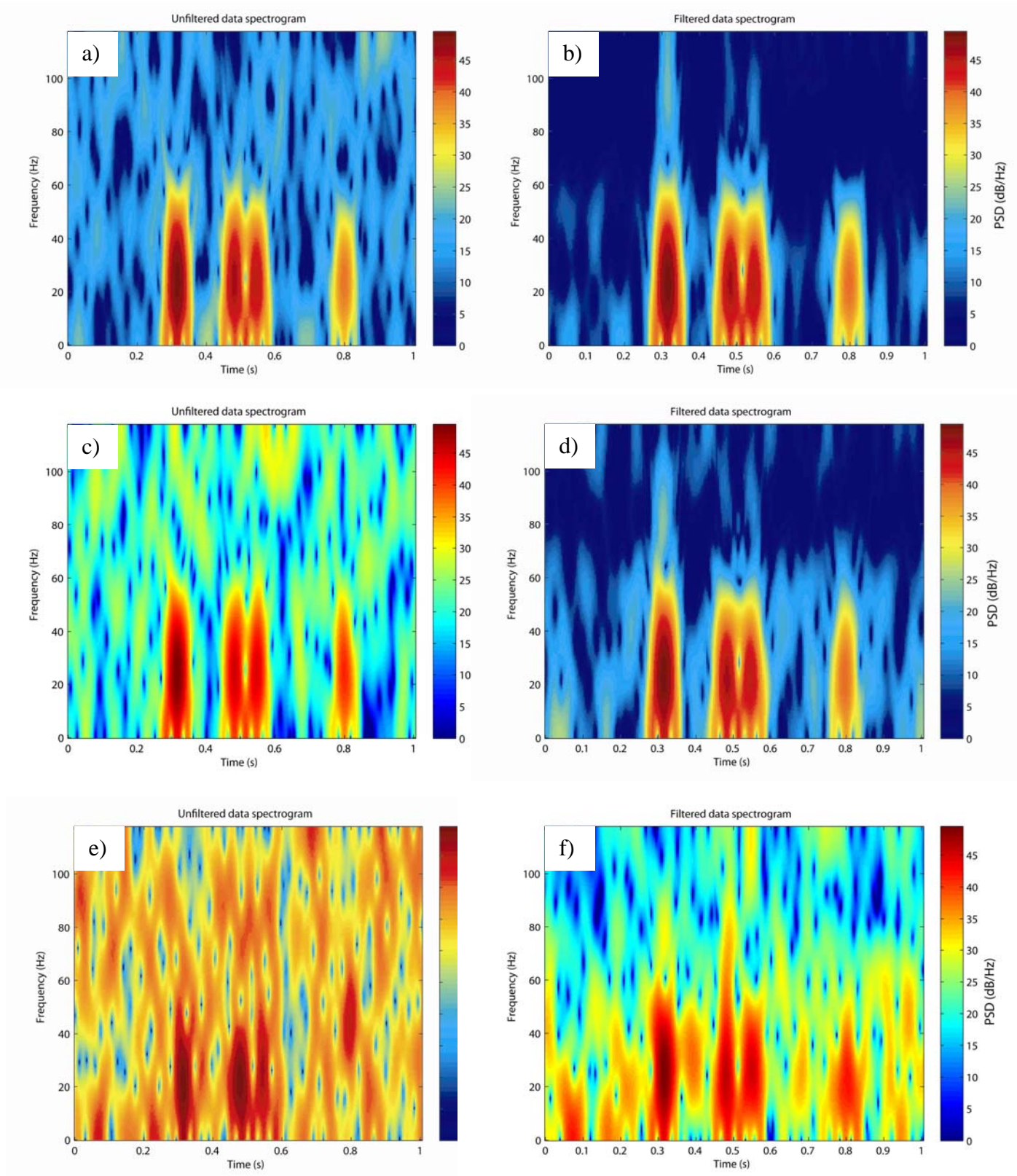


Figure 4.7. Spectrogrammes de la trace numéro 29 avec 5% (a), 10% (c) et 50% (e) de bruit et leurs filtrages respectifs (b,d,f).

4.2.5. Comparaison avec les filtres usuels

Nous avons procédé à une comparaison des résultats du filtre trilatéral avec ceux de deux filtres très populaires en traitement des données sismiques, à savoir le filtre $f-k$ et le filtre $f-x$. Ces filtres ont été appliqués aux données de la Figure 4.2.

La Figure 4.8 montre les résultats du filtrage correspondant à 5% de bruit aléatoire.

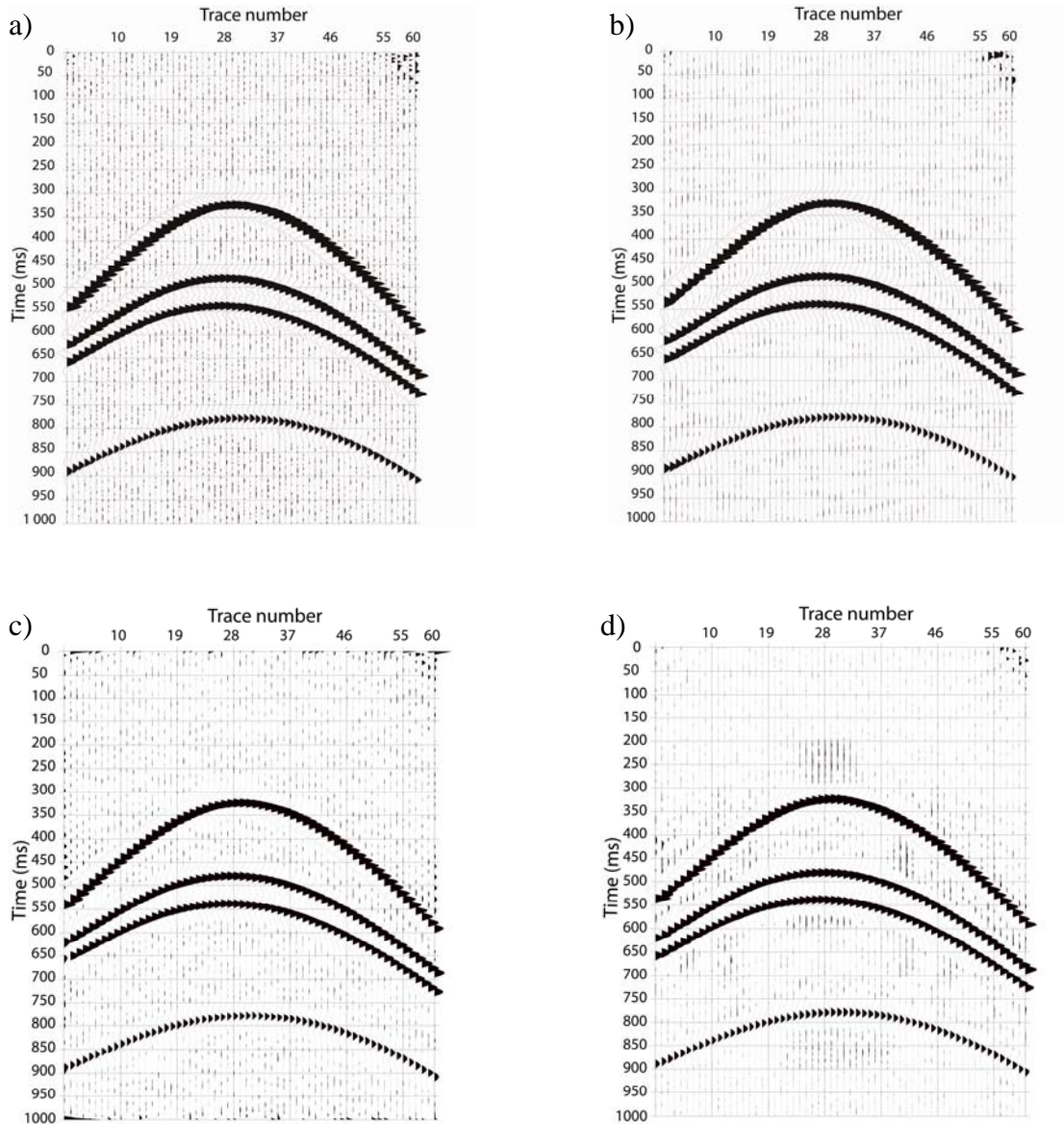


Figure 4.8. Comparaison entre données contaminées par a) 5% de bruit aléatoire, et résultats des filtres b) trilatéral, c) $f-k$ et d) $f-x$.

De la Figure 4.8b, il est clair que le filtre trilatéral produit la meilleure atténuation du bruit en comparaison avec le filtre $f-k$ (Figure 4.8c) qui tend à laisser beaucoup de bruit non traité, tandis que le filtre $f-x$ est meilleur que le filtre $f-k$ mais ses performances restent en-dessous

du filtre trilatéral à cause des artefacts générés au-dessus et en-dessous de chaque réflecteur. Notons que le résultat du filtrage du filtre trilatéral est donné seulement pour une seule itération. Nous avons pu avoir de meilleurs résultats en augmentant le nombre d'itérations, mais au risque de détériorer le signal d'intérêt. La [Figure 4.9](#) reproduit le même raisonnement de la [Figure 4.8](#) mais en présence de 10% de bruit aléatoire.

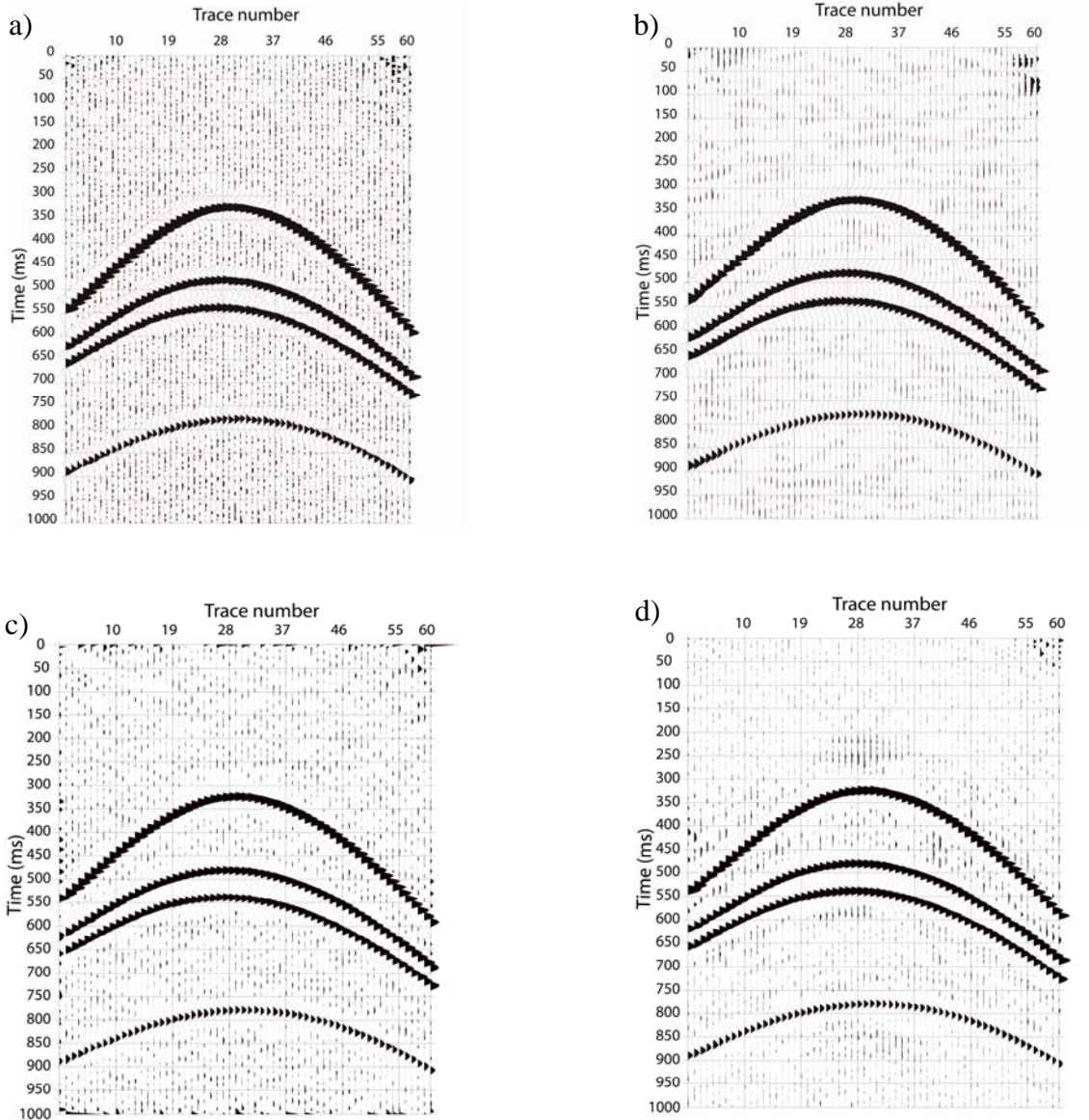


Figure 4.9. Comparaison entre données contaminées par a) 10% de bruit aléatoire, et résultats des filtres b) trilatéral, c) $f-k$ et d) $f-x$.

Les mêmes conclusions dressées pour le cas des données contaminées par 5% de bruit aléatoire sont mentionnées pour le cas de 10% de bruit aléatoire.

Le cas de la [Figure 4.10](#) est très important, car avec un taux de 50% de bruit aléatoire, nous avons remarqué que les performances des différents filtres se détériorent grandement.

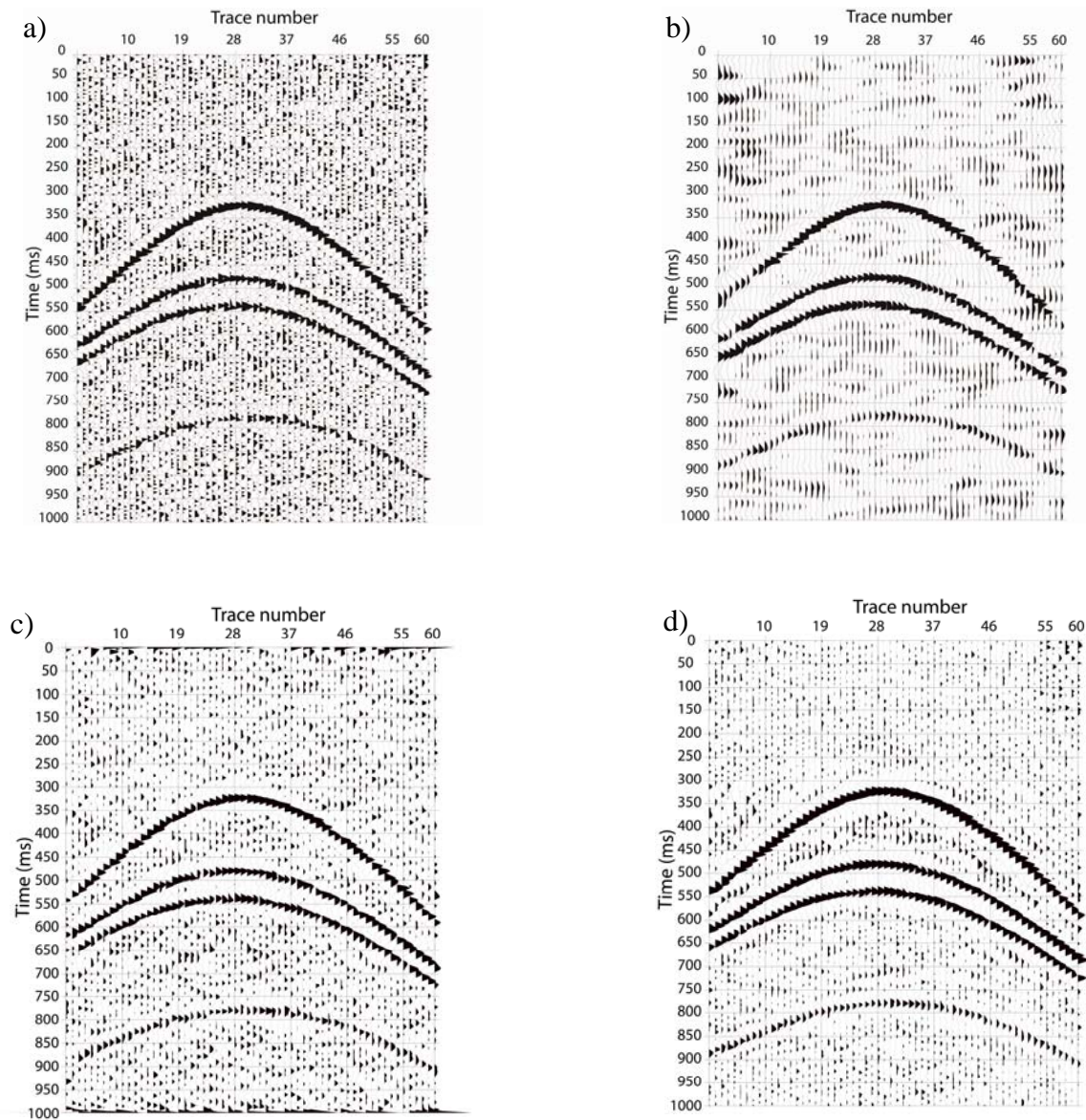


Figure 4.10. Comparaison entre données contaminées par a) 50% de bruit aléatoire, et résultats des filtres b) trilatéral, c) $f-k$ et d) $f-x$.

Il est clair qu'avec l'augmentation du taux du bruit aléatoire à 50% les trois filtres laissent de plus en plus de composantes du bruit non traités. En plus, nous remarquons une certaine détérioration du signal d'intérêt, en particulier au niveau du premier réflecteur, cette détérioration est accentuée dans le cas du filtre trilatéral ([Figure 4.10b](#)). La détérioration du signal d'intérêt est moins visible dans les résultats du filtre $f-x$, tandis que le filtre $f-k$ montre les performances les moins concluantes. Néanmoins, en termes d'atténuation de bruit, les

résultats du filtre trilatéral restent les plus concluants comparés au filtre $f-k$ (Figure 4.10c) et au filtre $f-x$ (Figure 4.10d).

4.3. Application sur des données réelles

Initialement, le filtre trilatéral, comme introduit par Garnett et al. (2005), était désigné pour le filtrage des images en 2 dimensions. Nous avons modifié cette particularité pour l'appliquer aux données 1 dimension. Ce filtre est appliqué à des données extraites d'une fenêtre glissante 3x3, puis transformées en un vecteur 1D. De plus, la statistique ROAD, présentée par Garnett et al. (2005), est la somme de m valeurs ascendantes représentant les échantillons non bruités. Nous avons modifié cette définition en calculant une valeur ROAD inverse (IROAD) en triant, puis ne sommant que les r_j différences dans un ordre décroissant. Ce qui nous a permis d'extraire la composante bruit du signal.

4.3.1. Stratégie proposée pour l'élimination des bruits

Nous avons suivi une stratégie, pour l'élimination du bruit, un peu différente de celle proposée par Garnett et al. (2005) (Figure 4.11). Il est assumé que le signal sismique (shot brut) est une combinaison d'une composante signal utile et d'une composante bruit. La composante signal utile est extraite via le filtre TL. La composante bruit est extraite par le biais du filtre TL. La composante du signal utile n'est autre que la différence entre la composante du bruit et le signal original.

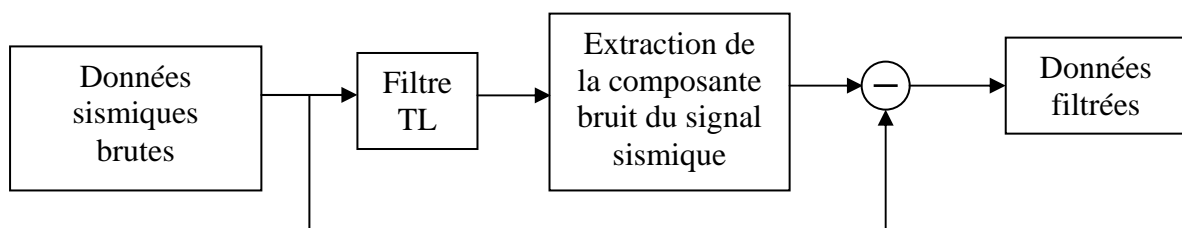


Figure 4.11. Stratégie proposée pour l'extraction puis la soustraction du bruit.

4.3.2. Description des données réelles

Les données réelles, utilisées dans notre étude, sont issues d'une campagne sismique 3D au Sud algérien. Le shot gather utilisé est composé de 240 traces (Figure 4.12). L'offset utilisé varie de 20 à 4800m avec un intervalle de groupe égale à 20m et un pas d'échantillonnage de 4ms. Les sources utilisées sont de type Vibroseis®. La Figure 4.12 montre les différentes

variétés de bruits linéaires. Les arrivées directes et réfractées obscurcissent les réflexions à 420ms, tandis que d'autres bruits cohérents prédominent à 950ms.

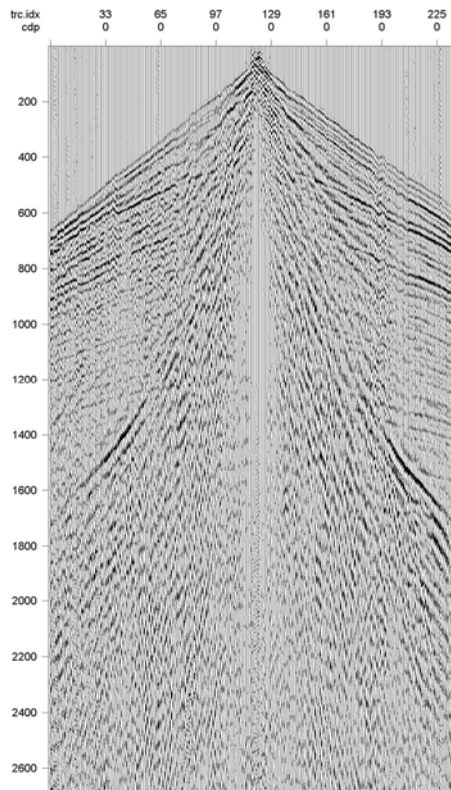


Figure 4.12. Shot gather bruité issu d'une campagne sismique 3D.

4.3.4. Résultats du filtrage sur les données réelles

Nous avons utilisé le filtre TL pour éliminer le bruit cohérent dans le shot gather. Nous avons implémenté le filtre en deux étapes :

- 1- Extraction de la composante bruit du shot gather : nous avons supposé que le ground roll et l'onde aérienne appartiennent à la composante bruit.
- 2- Élimination du bruit cohérent : une simple soustraction de la composante bruit (Figure 4-13a). Le bruit aléatoire peut être supprimé par simple mute La Figure 4-13b, illustre clairement que le bruit cohérent a été atténué. Un bon rapport S/B est obtenu entre 0 et 1400ms. Sur cette figure, on lit une bonne continuité des réflecteurs sur tous les

offsets. Ce qui produit un bon arrière plan pour des traitements ultérieurs. Néanmoins, quelques bruits résiduels sont visibles à environ 1450ms.

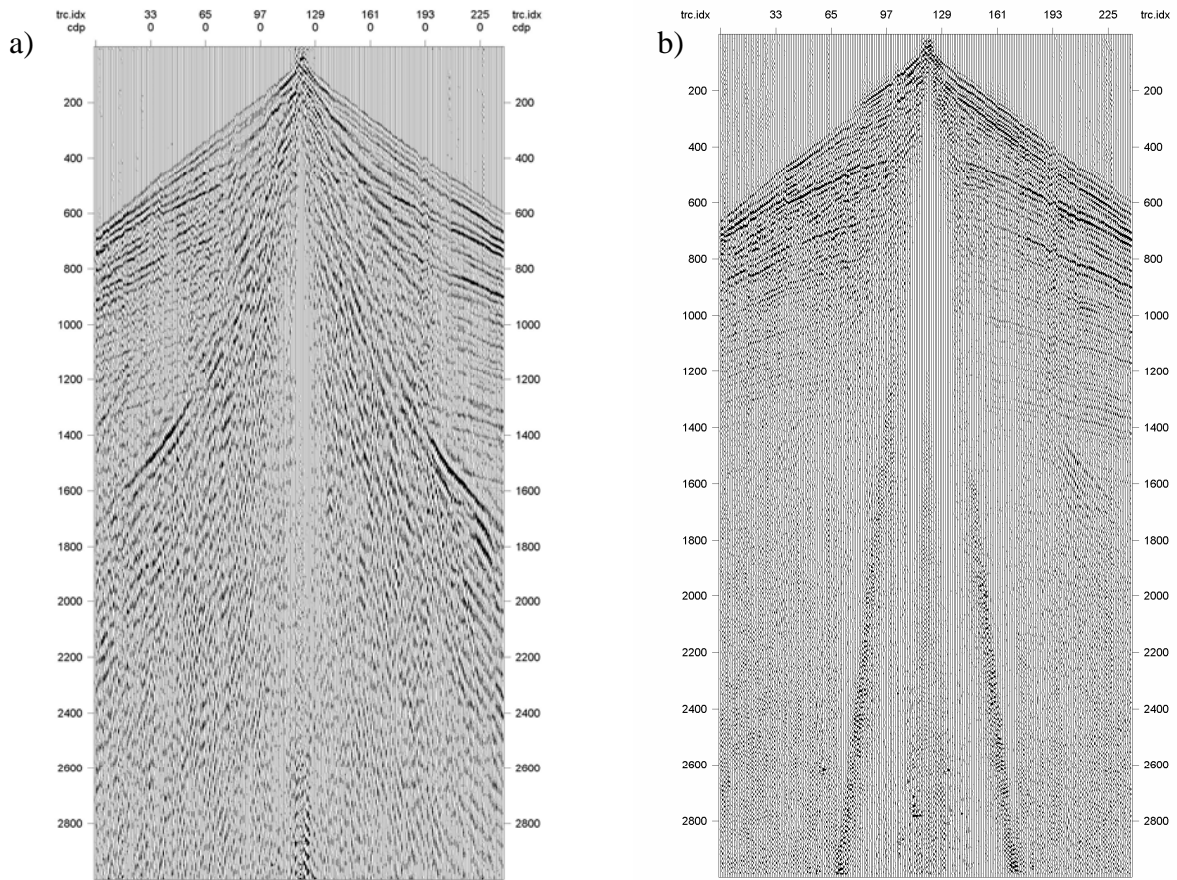


Figure 4.13. La composante bruitée isolée par le filtre trilatéral (a) et, le shot filtré(b).

Afin de mieux apprécier les effets du filtrage, nous avons réalisé une vue élargie (Figure 4.14) sur une fenêtre entre 320 et 650ms des Figures 4.13a et 4.13b autour de 500ms. Il est clair que les événements linéaires de la Figure 4.14a ont été considérablement atténués, après passage du filtre trilatéral (Figure 4.14b). Notons aussi l'amélioration de la continuité du réflecteur au temps double 450ms.

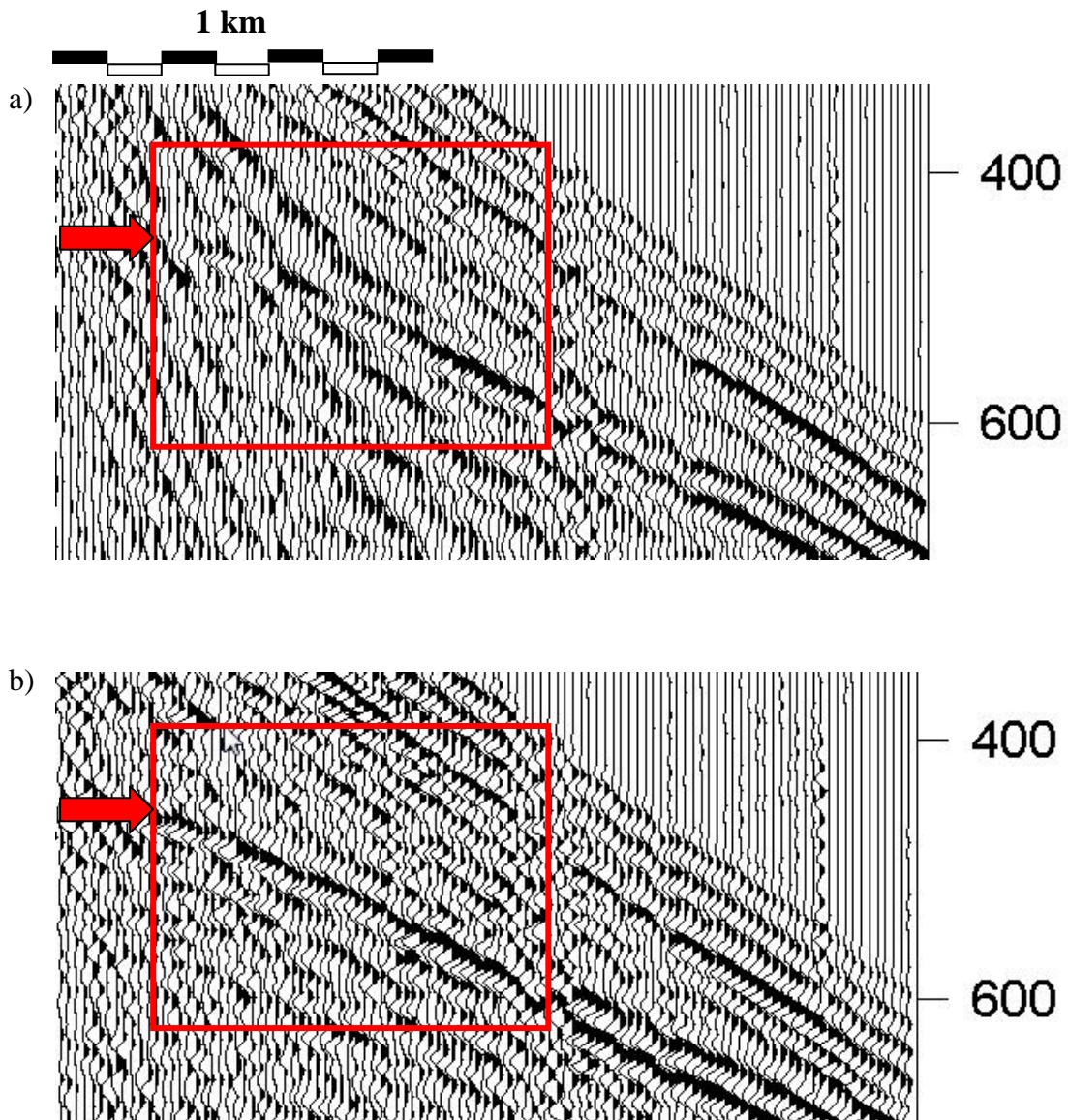


Figure 4.14. Agrandissement à l'intérieur d'une fenêtre du shot bruité (a) et filtré (b).

4.4 Etude comparative avec le filtre anisotropique

4.4.1 Cas des données synthétiques

A l'issue des chapitres précédents, nous avons constaté la robustesse des filtres trilatéral et de diffusion anisotropique vis-à-vis du bruit aléatoire. Néanmoins, il est indispensable de comparer le comportement des deux filtres en face des mêmes données. C'est ainsi que nous avons procédé à cette étude comparative. Nous avons repris le modèle de la [Figure 3.1](#) avec une légère modification. Ainsi, une couche à 550m de profondeur, simulant une lentille de sable, a été insérée et la faille précédemment présente a été supprimée ([Figure 4.16](#)).

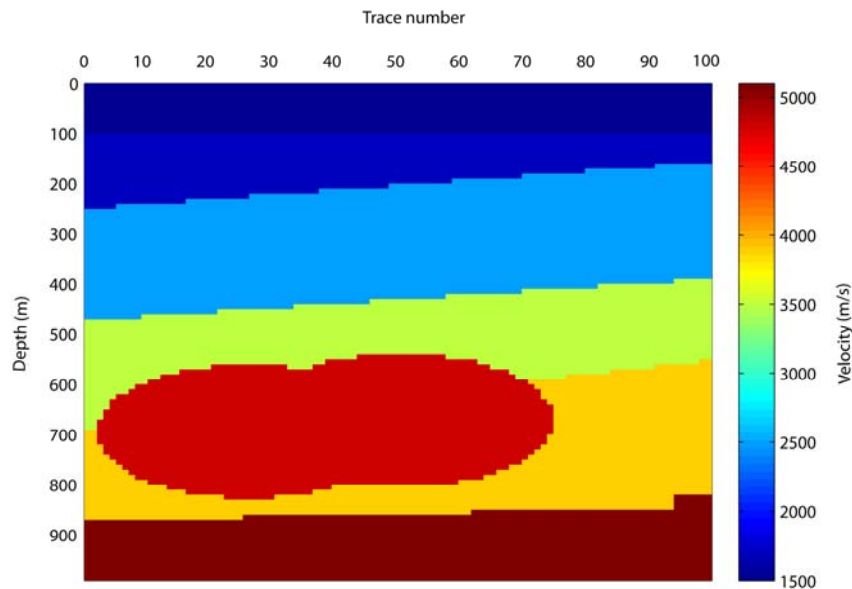


Figure 4.15. Modèle 2D utilisé pour l'étude comparative.

Ce modèle est constitué de cinq couches subhorizontales, avec des vitesses qui augmentent en fonction de la profondeur. La section sismique synthétique résultante de la convolution d'une ondelette de Ricker, de fréquence centrale 25 Hz, avec la série de réflectivité du modèle de la [Figure 4.15](#) est donnée en [Figure 4.16](#).

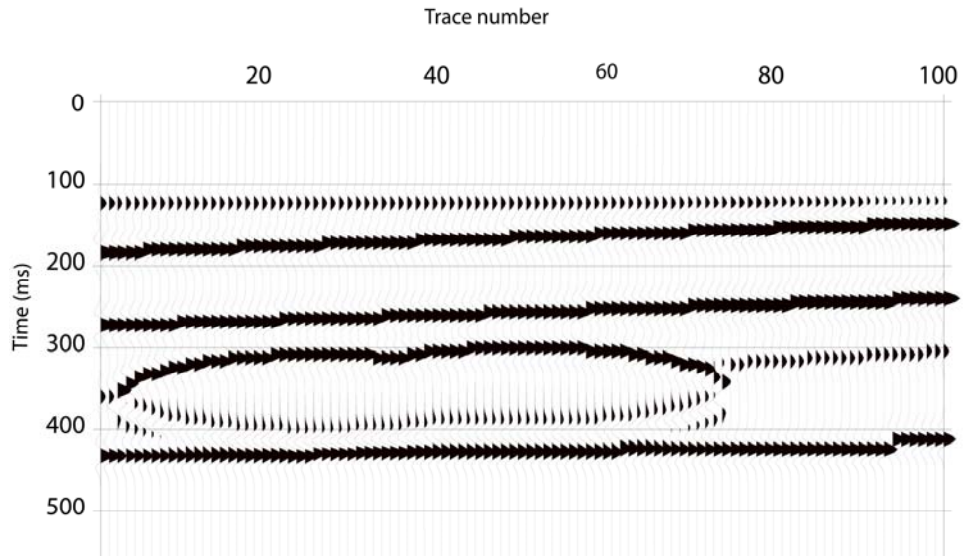


Figure 4.16. Section sismique synthétique correspondant au modèle de la figure 4.15.

Nous avons suivi la même procédure que celle utilisée pour tester le filtre trilatéral et anisotropique. Du bruit aléatoire à différents pourcentages (5%, 10% et 50%) a été ajouté aux données de la [figure 4.16](#) et le résultat est représenté en [Figure 4.17](#).

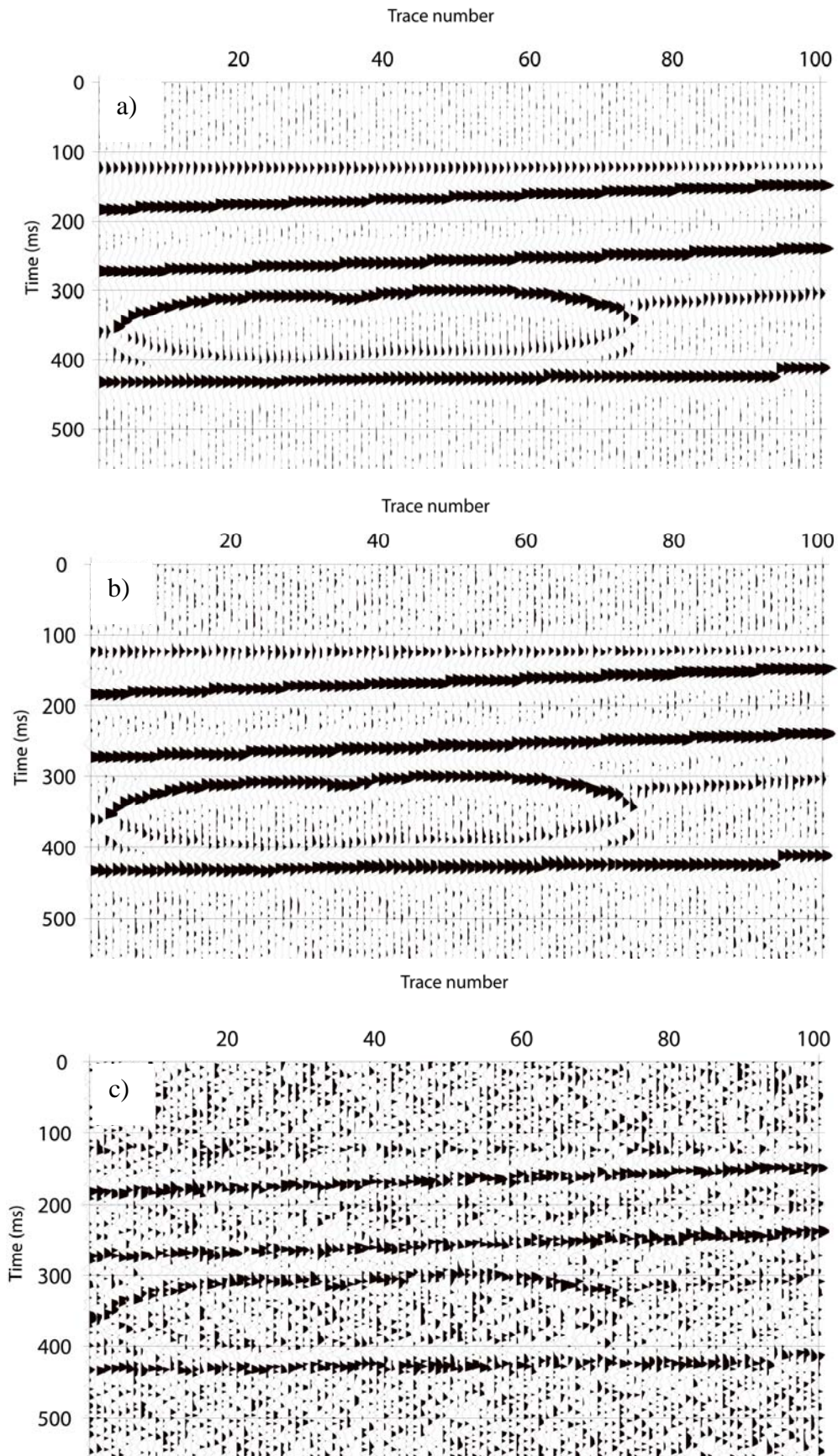


Figure 4.17. Données contaminées par a) 5%, b) 10% et c) 50% de bruit aléatoire.

Il est facile de constater la dégradation de la résolution latérale avec l'augmentation du niveau de bruit, en particulier lorsque le bruit atteint un taux de 50% (Figure 4.17c) où le toit de la quatrième couche a complètement disparu. Les Figures 4.18 et 4.19 montrent les résultats du filtrage par les filtres trilatéral et anisotropique respectivement.

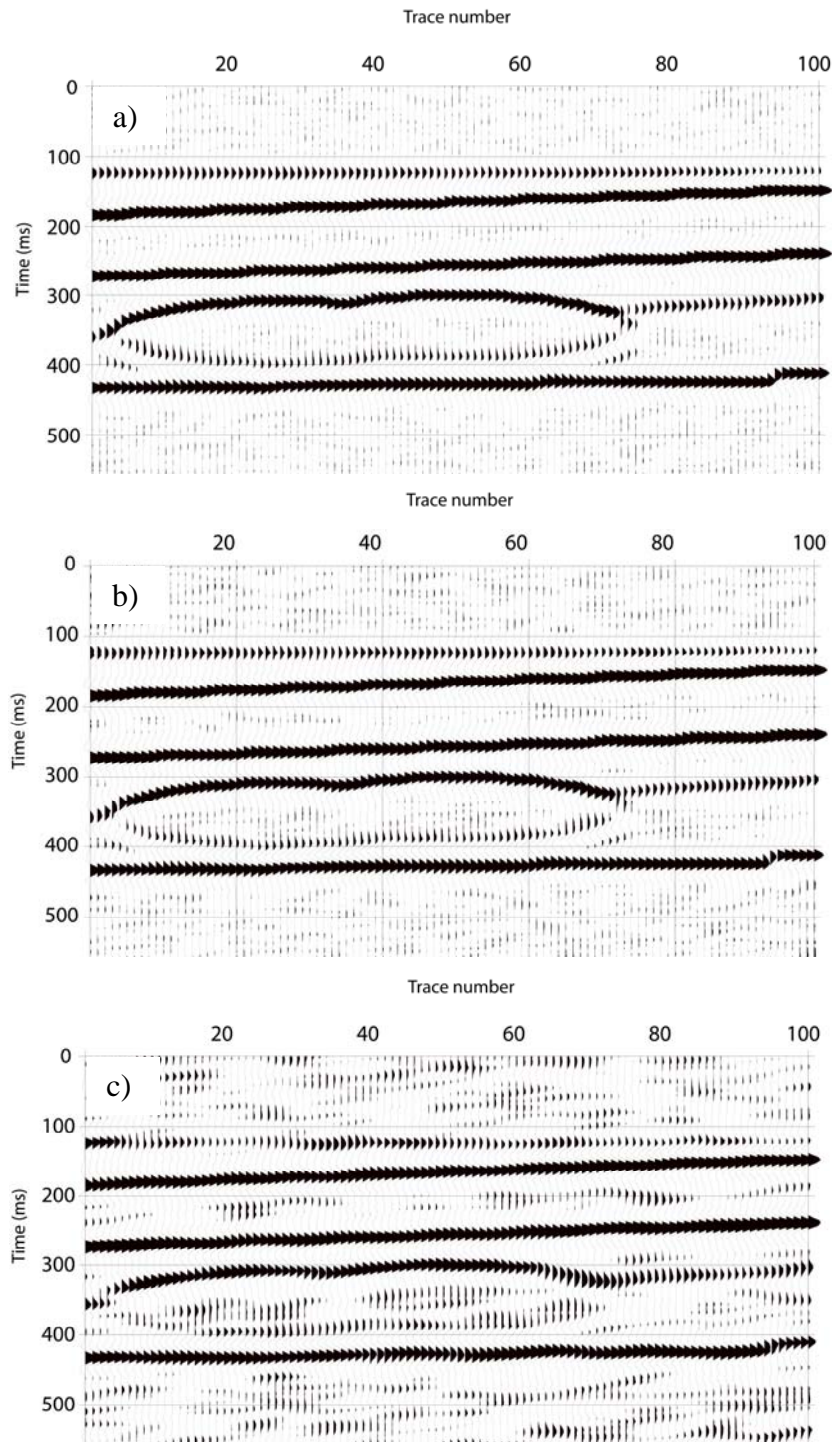


Figure 4.18. Résultats du filtre trilatéral en présence de a) 5%, b)10% et c) 50% de bruit aléatoire.

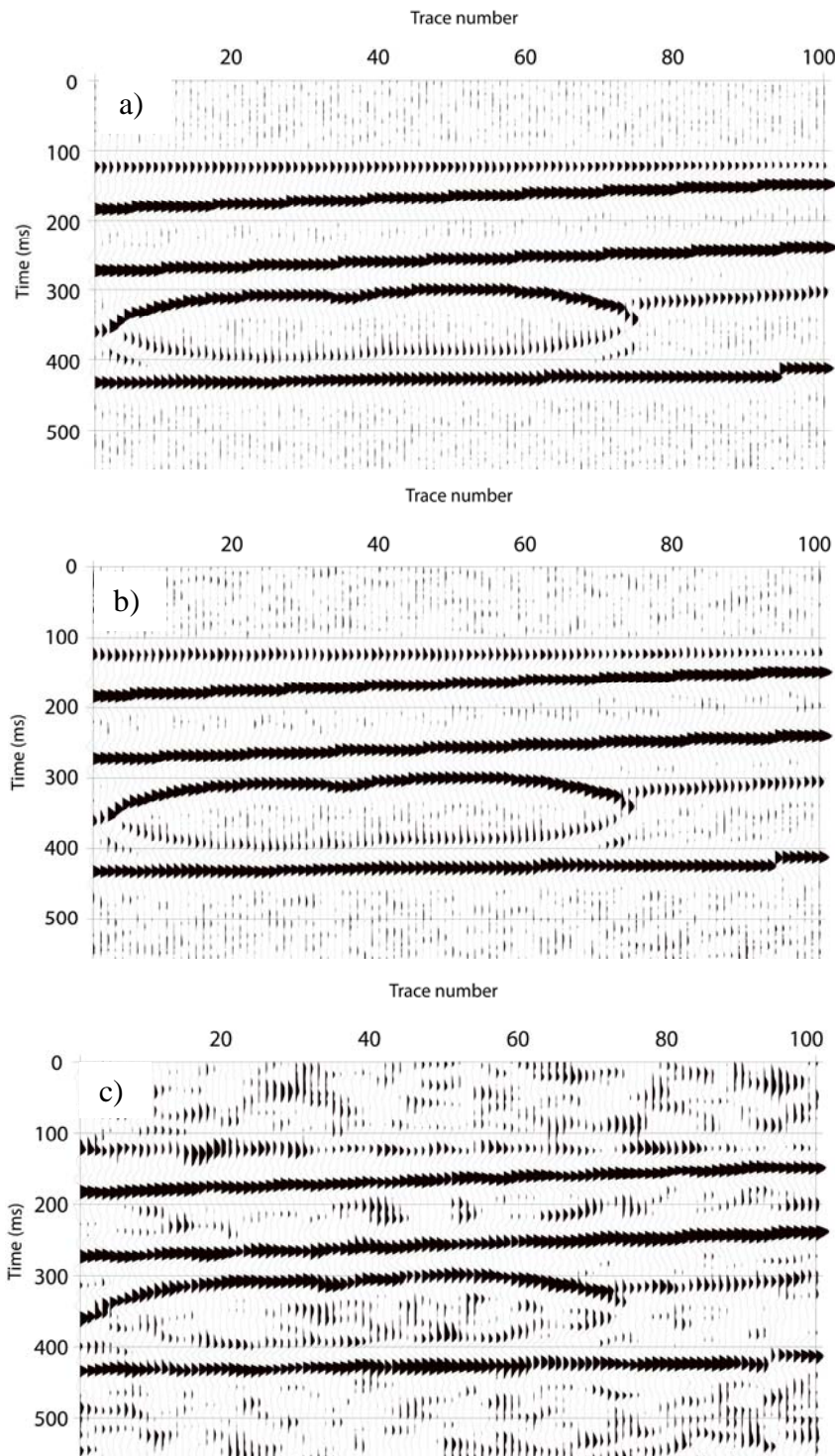


Figure 4.19 Résultats du filtre anisotropique en présence de a) 5%, b) 10% et c) 50% de bruit aléatoire.

Nous constatons que les deux filtres ont réussi à atténuer grandement le bruit aléatoire. Les meilleurs résultats sont visibles pour le cas où le bruit est à 50% du signal d'intérêt (Figures 4.18c et 4.19c). Dans les deux cas, du bruit résiduel est présent, en particulier lorsque le RSB

est faible. Cependant, la sortie du filtre trilatéral est visuellement plus lisse (Figure 4.18c) tandis que le filtre anisotropique est plus efficace pour l'atténuation du bruit. On remarque aussi que la continuité des différents réflecteurs a été préservée. Les paramètres des filtres ont été choisis par des tests et les valeurs optimales fixées lors des simulations. Nous avons choisi les paramètres suivants pour le filtre anisotropique : une valeur de 50 pour la constante de diffusion, des coefficients de diffusion comme indiqué par l'équation (2.6b) et un facteur η égale à 0.10. Le nombre d'itérations a été choisi sur la base du RSB et de l'EQM. Nous avons aussi utilisé les paramètres suivants pour le filtre trilatéral : $\sigma_s=1$, $\sigma_r=0.5$, $\sigma_l=0.3$ et $\sigma_j=200$.

Les résultats du filtrage sont regroupés au [tableau 4.2](#).

Algorithme	Filtre trilatéral						Filtre anisotropique					
	5%		10%		50%		5%		10%		50%	
	EQM	RSB	EQM	RSB	EQM	RSB	EQM	RSB	EQM	RSB	EQM	RSB
Nombre d'itérations												
1	5.33 ^{E-5}	38.26	1.09 ^{E-4}	31.06	0.0015	5.09	4.18^{E-5}	41.05	1.25 ^{E-4}	30.08	0.0021	2.07
2	6.93 ^{E-5}	35.64	1.02^{E-4}	31.76	8.98 ^{E-4}	10.02	6 ^{E-5}	37.41	1.04^{E-4}	31.88	0.0011	7.92
3	9.18 ^{E-5}	32.83	1.14 ^{E-4}	30.65	6.61 ^{E-4}	13.08	9.51 ^{E-4}	32.81	1.23 ^{E-4}	30.23	7.98 ^{E-4}	11.54
4	1.16 ^{E-4}	30.49	1.35 ^{E-4}	29.16	5.41 ^{E-4}	15.07	1.38 ^{E-4}	29.05	1.58 ^{E-4}	27.73	6.48 ^{E-4}	13.62
5	1.41 ^{E-4}	28.5	1.54 ^{E-4}	27.63	4.78 ^{E-4}	16.32	1.87 ^{E-4}	26.05	2.01 ^{E-4}	25.3	5.84 ^{E-4}	14.65
6	1.65 ^{E-4}	26.92	1.83 ^{E-4}	25.92	4.45 ^{E-4}	17.03	2.35 ^{E-4}	23.76	2.56 ^{E-4}	22.88	5.64^{E-4}	15
8	2.20 ^{E-4}	24.06	2.35 ^{E-4}	23.39	4.27^{E-4}	17.44	3.43 ^{E-4}	19.98	3.61 ^{E-4}	19.45	5.86 ^{E-4}	14.62
10	2.80 ^{E-4}	21.66	2.85 ^{E-4}	21.49	4.42 ^{E-4}	17.10	4.57 ^{E-4}	17.1	4.62 ^{E-4}	16.99	6.47 ^{E-4}	13.64
15	4.25 ^{E-4}	17.49	4.27 ^{E-4}	17.44	5.32 ^{E-4}	15.25	7.25 ^{E-4}	12.50	7.27 ^{E-4}	12.48	8.48 ^{E-4}	10.93
20	5.68 ^{E-4}	14.6	5.68 ^{E-4}	14.59	6.48 ^{E-4}	13.29	9.64 ^{E-4}	9.65	9.65 ^{E-4}	9.65	0.0011	8.76
30	8.31 ^{E-4}	10.79	8.3 ^{E-4}	10.81	8.84 ^{E-4}	10.18	0.0013	6.30	0.0013	6.31	0.0014	5.89
Données bruitées	2.75^{E-5}	45.19	7.0^{E-5}	35.88	0.001	8.83						

Tableau 4.2. Résultats du filtre trilatéral et anisotropique en termes de rapport RSB et de l'EQM en fonction du nombre d'itérations.

Il est évident qu'avec l'augmentation du nombre d'itérations les résultats sont plus lisses, mais les données sont distordues comme indiqué par la diminution du RSB et l'augmentation de l'EQM. Ce résultat montre que l'utilisation du filtre trilatéral et du filtre de diffusion

anisotropique doit être prudente lors de la manipulation de données traitées en amplitudes préservées. Nous avons donc choisi une itération en présence de 5% de bruit aléatoire, et deux itérations pour le cas de 10% de bruit aléatoire. Cependant, lorsque le bruit atteint 50%, nous avons augmenté le nombre d'itérations à six pour le filtre anisotropique et à huit pour le filtre trilatéral afin d'atténuer au mieux le bruit tout en préservant le signal d'intérêt. Les résultats du [tableau 4.2](#) sont représentés sur les [Figures 4.20 et 4.21](#) sous forme de courbes du RSB et de l'EQM en fonction du nombre d'itérations, respectivement.

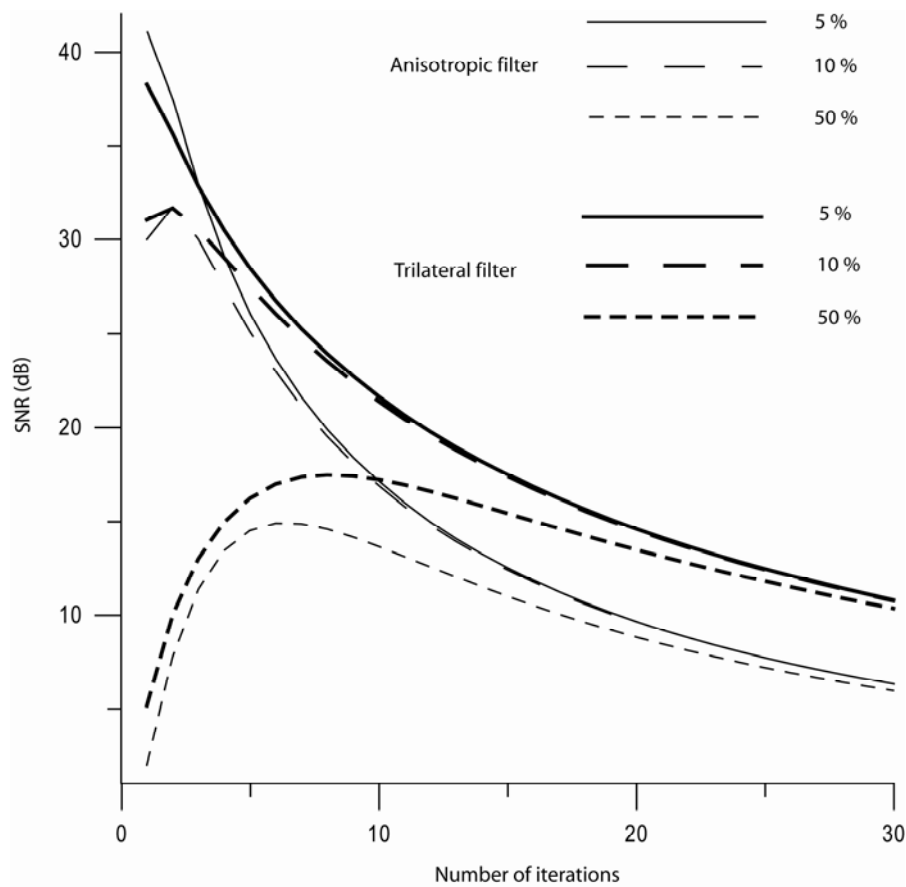


Figure 4.20. Rapport RSB des données filtrées en fonction du nombre d'itérations pour les filtres trilatéral et anisotropique.

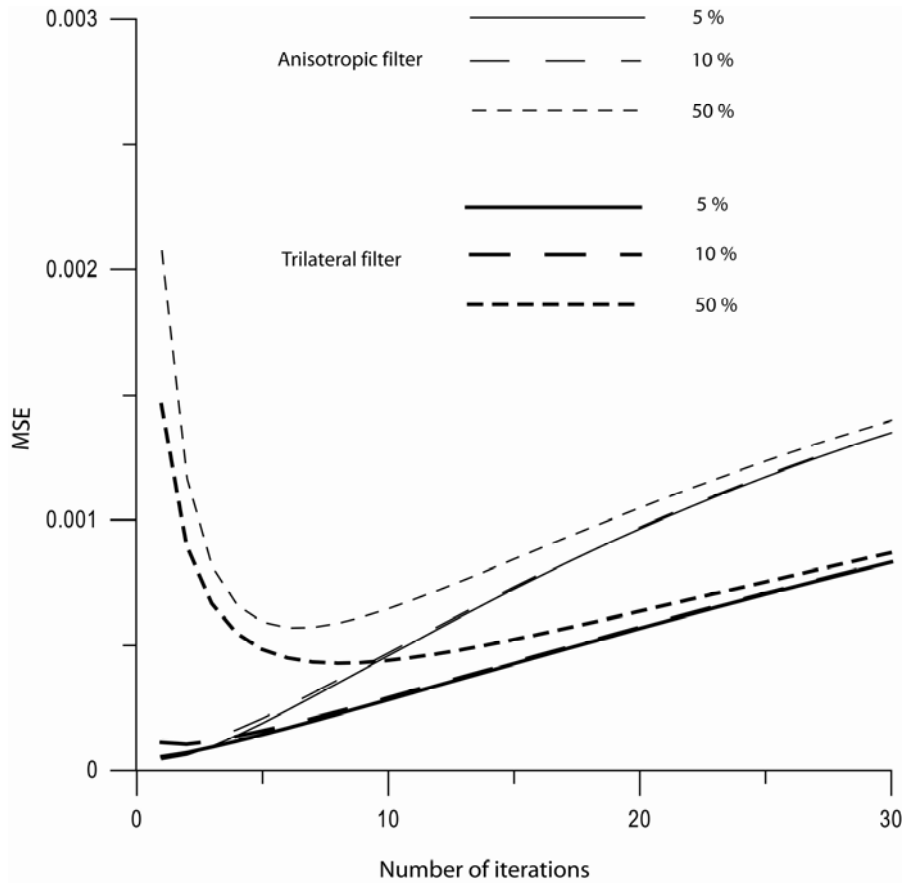


Figure 4.21. EQM relative aux données filtrées, en fonction du nombre d'itérations, pour les filtres trilatéral et anisotropique.

On remarque, d'après les figures 4.20 et 4.21, une bonne similitude dans le comportement des deux filtres, se traduisant par une concordance des courbes des deux filtres. Cette concordance est plus marquée lorsque le niveau du bruit est assez faible, et tend à diverger en augmentant le niveau du bruit.

4.4.2 Cas des données réelles

L'étude comparative s'est aussi portée sur des données réelles (Figure 4.22a). Ces données proviennent d'une campagne sismique au Sud algérien, l'intertrace est de 25m, l'inter-tir de 25m et le pas d'échantillonnage est de 2ms. La bande fréquentielle du signal est de 8-72 Hz, tandis que les sources sont de type vibreur. Malgré la sommation, la section de la figure 4.22a montre la présence d'un fort bruit, en particulier en surface.

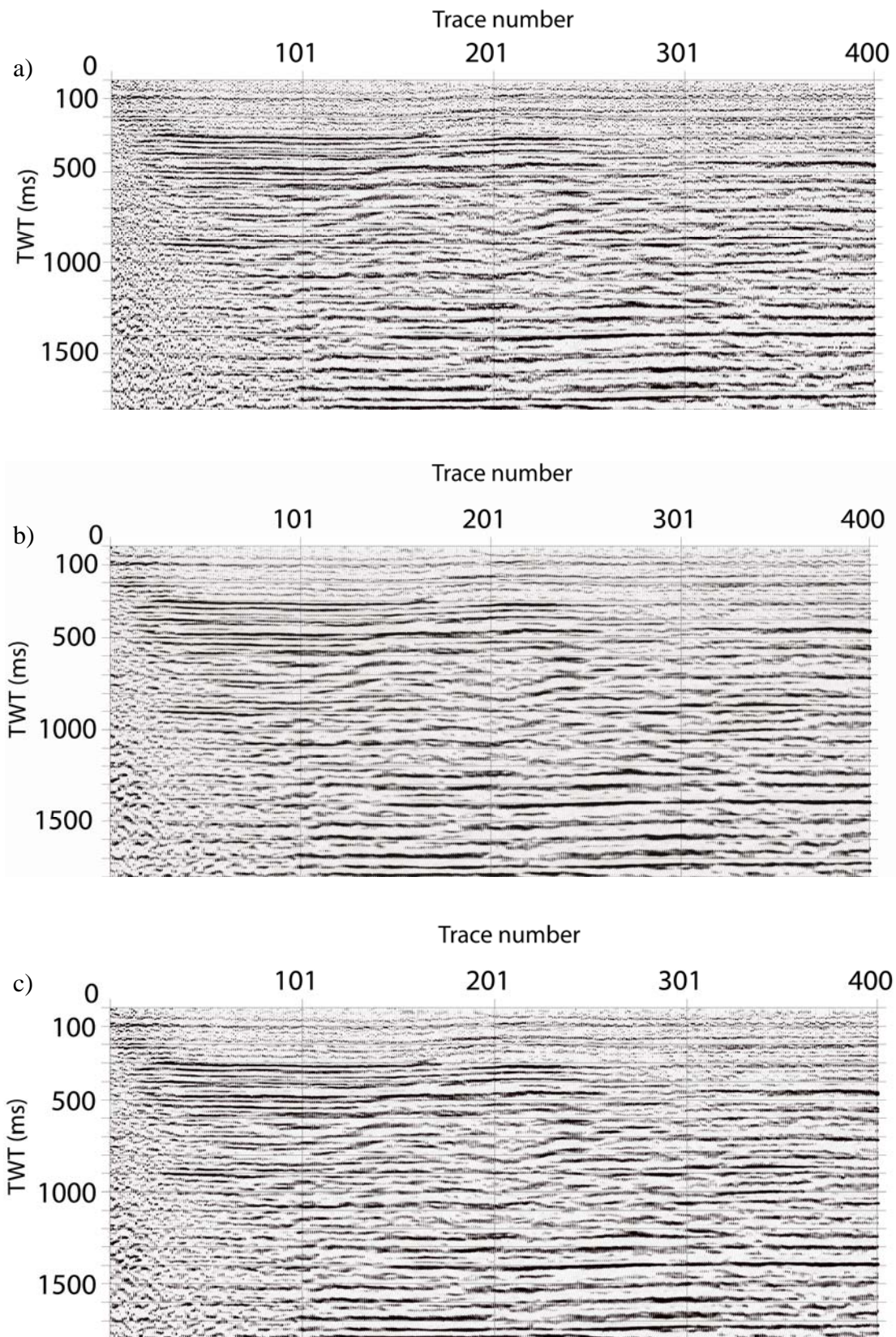


Figure 4.22. Section sismique brute (a) et ses versions filtrées par les filtres anisotropique (b) et trilatéral (c).

Les deux filtres ont été appliqués aux données de la section sommées, et les résultats correspondants aux filtres anisotropique et trilatéral sont montrés sur les figures 4.22b et 4.22c, respectivement. L'amélioration en termes de RSB est visible pour les deux filtres, avec une bonne continuité des réflecteurs et une préservation du signal d'intérêt. Un zoom sur une fenêtre permet de mieux différencier l'action des deux filtres (Figure 4.23).

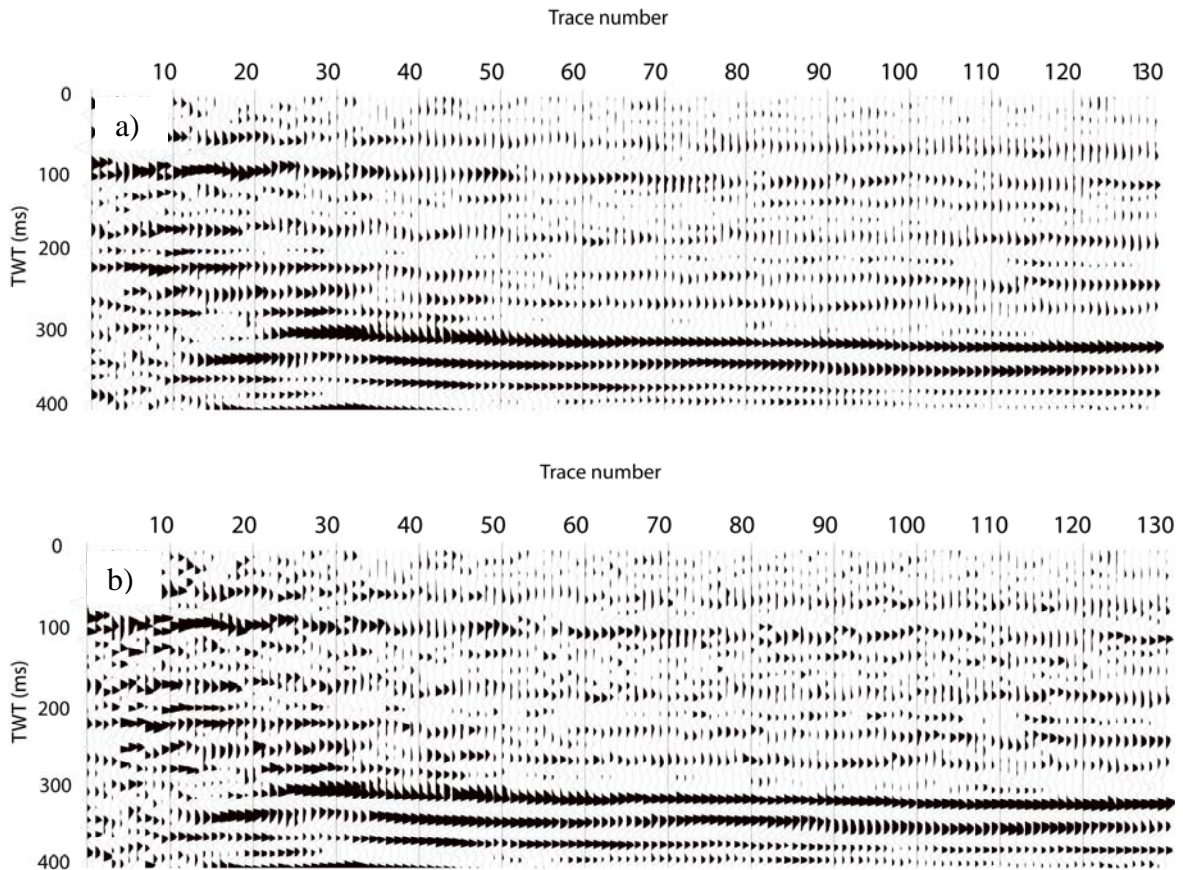


Figure 4.23. Vue élargie de la section sismique filtrée par les filtres anisotropique (a) et trilatéral (b).

Il est évident que les deux filtres ont réussi à atténuer le bruit ; néanmoins, l'action du filtre anisotropique (Figure 4.23a) est plus marquée que celle du filtre trilatéral (Figure 4.23b).

4.5. Conclusion

Nous avons utilisé le filtre trilatéral (TL) pour éliminer les bruits aléatoires et cohérents de données sismiques synthétiques et réelles.

Les résultats numériques obtenus sur les données synthétiques, ont montré que le filtre trilatéral est efficace dans l'élimination des bruits aléatoires et cohérents. En plus, nos tests ont montré qu'il est préférable de garder le nombre d'itérations au plus bas niveau lorsque le bruit entachant les données n'est pas trop important, et qu'il faudra l'augmenter lorsque le bruit devient plus important. Bien entendu il faut veiller à garder un certain compromis entre atténuation du bruit et lissage excessif des données. La comparaison avec les filtres $f-k$ et $f-x$ a montré que notre filtre est le plus performant en termes d'atténuation du bruit, tandis que le filtre $f-k$ est le moins performant des trois filtres. En présence d'un fort taux de bruit, l'atténuation du bruit est bien effectuée par le filtre trilatéral mais au détriment du signal d'intérêt, cependant c'est le filtre $f-x$ qui préserve le mieux le signal mais avec une atténuation du bruit assez médiocre. Nous avons modifié la structure du filtre TL en calculant une valeur ROAD inverse et en calculant le signal filtré par simple différence entre les données brutes et celles isolées par le filtre TL. L'application du filtre sur des données réelles a montré qu'il est très robuste en présence de bruit tels le Ground roll et autres bruits liés aux sources instrumentales ou anthropiques. L'implémentation de ce filtre est très facile en plus de sa robustesse.

Enfin, nous avons comparé les filtres trilatéral et anisotropique sur les données d'un même modèle synthétique, contaminées par différents niveaux de bruit, et les résultats ont été en accord avec nos prévisions. Ainsi, l'atténuation du bruit est appréciable pour les deux filtres. Le filtre anisotropique a donné les meilleurs résultats en termes d'atténuation du bruit, tandis que le filtre trilatéral a donné des sorties plus lissées. L'application des deux filtres sur les données réelles a aussi confirmé la supériorité du filtre anisotropique dans l'atténuation du bruit.

Chapitre 5

Application du filtre SD-ROM aux données de prospection électrique

Nous abordons dans ce chapitre l'application du filtre DS-ROM (Abreu et al., 1996; Chandra et al., 1998; Moore et al., 1999; Moore et Mitra, 2000) aux données électriques. Ce filtre se base sur les statistiques d'ordre. Il procède en deux temps i) identification, puis ii) estimation. Lors de la phase d'identification, l'échantillon est comparé à un nombre de valeurs seuils, qui décideront s'il est contaminé par le bruit ou non. L'échantillon passera inchangé à travers le filtre s'il est considéré comme « non contaminé », sinon il sera remplacé par une valeur (ROM) calculée (phase d'estimation) à partir de son voisinage. L'application de ce filtre sur des données électriques a montré sa robustesse et son efficacité en présence de bruits impulsifs et cohérents (cas des Inhomogénéités de surface : NSIs).

5.1 Simulations

Nous avons testé le filtre SD-ROM sur des données synthétiques de résistivité électrique à courant continu. Un modèle synthétique à deux dimensions (Figure 5.1) est généré par le programme libre Res2dmod (Loke et Barker 1996). Ce programme calcule les résistivités apparentes par la méthode des différences finies ou par la méthode des éléments finis, pour un modèle prédéfini. Le modèle ainsi utilisé divise le sous-sol en blocs rectangulaires et une valeur de potentiel est définie pour chaque nœud de la grille. Les résistivités sont affectées à chaque bloc par l'utilisateur.

5.1.1 Description du modèle synthétique

Le modèle 2D utilisé pour générer les pseudosections de résistivité apparentes (Figure 5.1), est composé de trois couches, avec une variation latérale de l'épaisseur de la deuxième couche. Ainsi, ce modèle comprend les paramètres suivants :

- Un substratum de résistivité égale à 500 Ω .m à une profondeur de 8m,
- Une couche conductrice de résistivité égale à 5 Ω .m à 2.8 m de profondeur avec variation latérale de l'épaisseur, surmontant le substratum résistant,
- Une couche superficielle de résistivité égale à 50 Ω .m. et d'une épaisseur de 3.56 m.

- Trois blocs de résistivités différentes (5 et 500Ω.m) ont été placés dans la couche superficielle pour simuler les effets des inhomogénéités de surface (NSIs).

Rappelons qu'une inhomogénéité de surface, en prospection électrique, est un corps géologique, d'habitude de petite taille et proche de la surface du sol, dont la résistivité est anormale comparativement à son environnement.

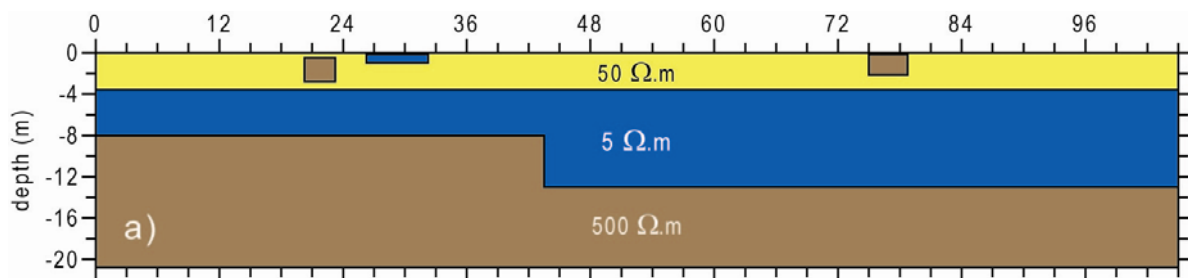


Figure 5.1. Modèle 2D utilisé pour générer les pseudosections de résistivités apparentes, en présence des inhomogénéités de surface (NSIs). L'espaceur 'a' varie de 3m à 9m pour n=1,2...8.

La pseudosection des résistivités apparentes calculées à partir du modèle 2D de la [Figure 5.1](#), est présentée en [Figure 5.2](#). Remarquons la présence d'anomalies caractéristiques en forme de V inversé associées aux NSIs (l'apex de ce V inversé se trouve aux alentours de 20m, 30m et 75m). Ces anomalies sont très gênantes et souvent considérées comme bruit cohérent en prospection électrique. La présence de ces anomalies dans les données affecte grandement le processus d'inversion.

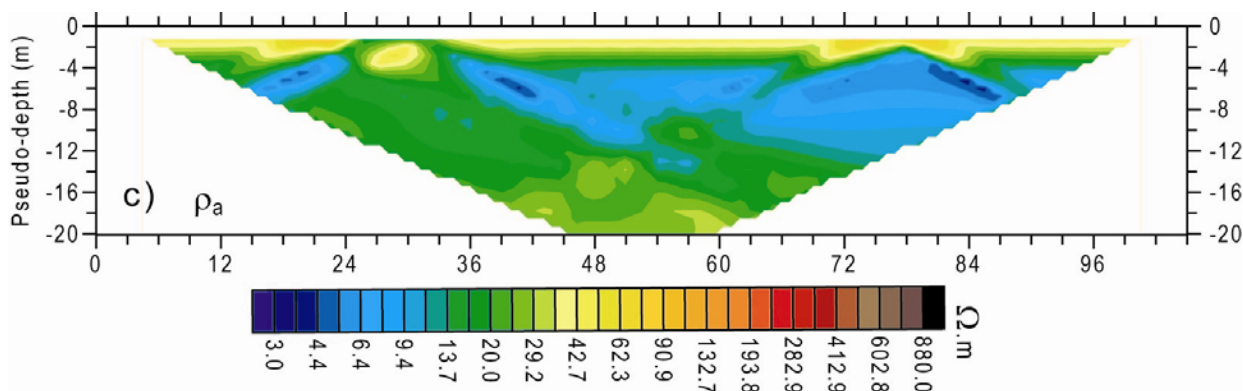
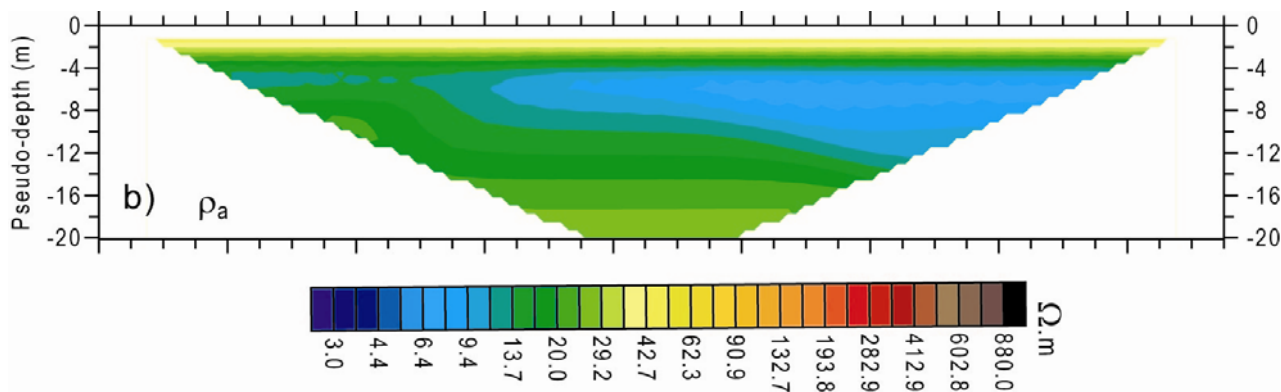


Figure 5.2. Pseudosection de résistivités apparentes calculées, correspondant au modèle de la figure 5.1, en présence des NSIs.

Afin de pouvoir évaluer les performances de chaque filtre, nous avons construit un modèle sans inhomogénéités de surface, tout en gardant les mêmes paramètres que ceux du modèle 2D de la Figure 5.1. La pseudosection des résistivités apparentes calculées résultante est donnée en Figure 5.3.



La section inversée a été utilisée comme référence lors de la comparaison des résultats des différents filtres. Les données synthétiques ont été générées pour un dispositif dipôle-dipôle avec un espacement « a » variant de 3 à 9m et un facteur « n » variant de 1 à 8, ce qui nous donne une profondeur d'investigation de 20.16m avec un total de 552 mesures. La méthode des différences finies a été utilisée pour la génération des pseudosections de résistivités apparentes.

5.1.2. Résultats du filtrage sur les résistivités apparentes calculées

Afin de comparer les performances de chaque filtre, nous avons filtré les résistivités apparentes de la Figure 5.2 par les filtres SD-ROM et médian (MED), avec différentes longueurs (Figure 5.4). La Figure 5.4a correspond à la pseudosection des résistivités apparentes filtrées par le filtre SD-ROM. Il est clair que l'effet des NSIs, caractérisé par un V inversé, a été complètement éliminé du résultat final, ce qui montre l'efficacité du filtre et sa faculté à détecter ces valeurs anormales, sans aucune intervention de l'interpréteur. L'effet du filtre médian (Figures 5.4b et 4c) n'a pas été aussi concluant que pour le cas du filtre SD-ROM. On peut remarquer sur la Figure 5.4b, qui correspond à un filtre médian de longueur 3, que l'anomalie en V inversé produite par les NSIs, n'a pas complètement disparu montrant ainsi les limites du filtre médian. En plus, la rampe de la deuxième couche n'est plus

identifiable. Les performances du filtre médian sont légèrement améliorées avec l'augmentation de la taille de la fenêtre (Figure 5.4c). Toutefois, la résolution a été réduite.

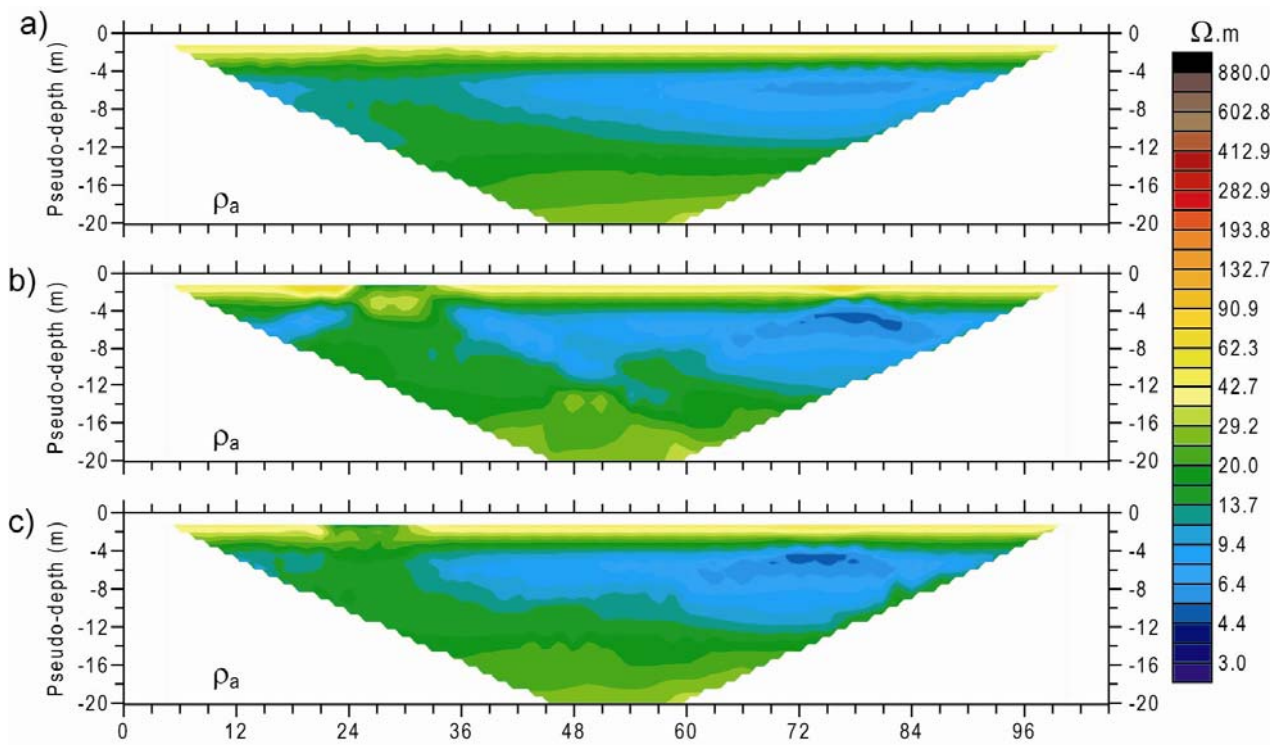


Figure 5.4. Pseudosections de résistivités apparentes calculées, après filtrage par filtre a) SD-ROM, b) médian (fenêtre 3) et c) médian (fenêtre 5).

5.1.3 Effets du filtrage sur le processus d'inversion

Les données de résistivités apparentes ont été filtrées par les filtres MED et SD-ROM, et les sections inversées sont montrées sur la Figure 5.5. La Figure 5.5a a été utilisée comme référence, à cause de l'absence des effets des NSIs et de sa faible erreur quadratique moyenne (RMS), égale à 0.99%. Sur cette figure on peut observer une bonne similitude entre la section inversée et le modèle synthétique (traits noirs). La section inversée du modèle synthétique en présence des NSIs (Figure 5.5b) possède une erreur RMS de 4.19%. Il est à noter, que les zones de résistivités anormales se superposent parfaitement dans l'espace et possèdent des amplitudes analogues au modèle synthétique, comme indiqué par les traits noirs.

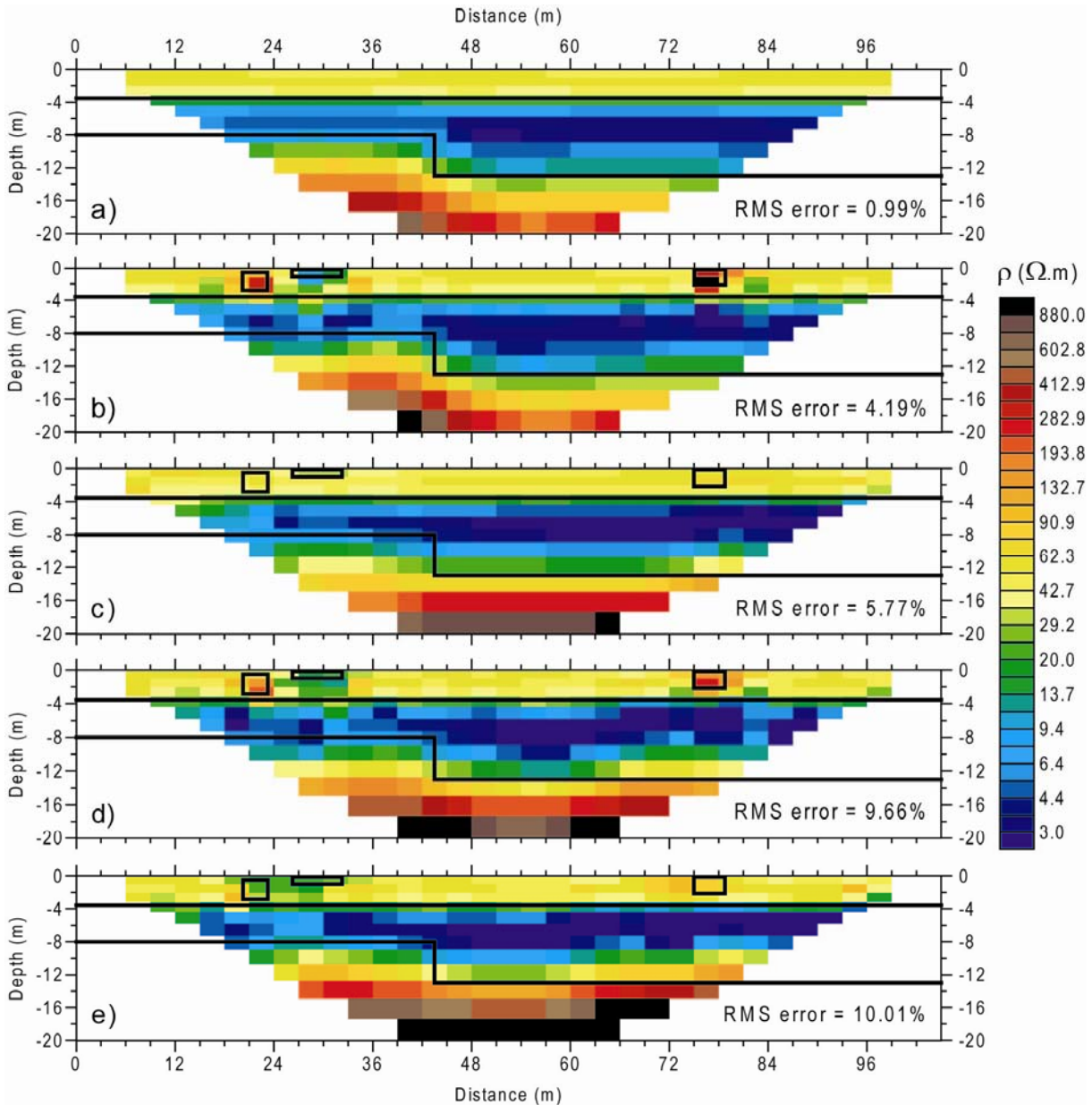


Figure 5.5. Résultats de l'inversion des données synthétiques (Les contours noirs reproduisent le modèle synthétique) : a) en l'absence des NSIs, b) en présence des NSIs, après filtrage par filtres c) SD-ROM, d) médian (fenêtre de longueur 3) et e) médian (fenêtre de longueur 5).

La [Figure 5.5c](#) illustre la section inversée après filtrage par filtre SD-ROM avec une erreur RMS de 5.77%. Tous les paramètres du modèle synthétique ont été récupérés avec une bonne atténuation des effets perturbateurs générés par les NSIs. Notons toutefois que, sur la section correspondant au filtre SD-ROM, le bord de la couche inférieure se trouvant à 43m et à une profondeur de 8m, est décalé vers la gauche du modèle à une distance d'environ 36m.

Le filtre médian est l'outil de comparaison par excellence en traitement de l'image. C'est pour cette raison que nous l'avons utilisé avec le filtre SD-ROM pour l'élimination des bruits associés aux NSIs. Les résultats sont illustrés par les [Figures 5.5d](#) et [5.5e](#) (correspondants au filtres MED de longueur 3 et 5, respectivement); où il est facile de constater que les zones anormales se trouvant à 21,30 et 75m n'ont pas été parfaitement atténuées, et que la forme du substratum est complètement détériorée. Toutefois, il est possible d'améliorer le processus d'inversion par une modification de la taille de la grille (raffinement). Malheureusement, il se traduit par l'augmentation du nombre de paramètres inconnus et aura un impact négatif sur la stabilité de l'algorithme. En plus, il augmentera le coût numérique lors du calcul de la matrice du Jacobien.

5.2. Application du filtre SD-ROM aux données réelles.

5.2.1 Description des données et implantation des filtres

Notre application réelle s'est basée sur des données issues de deux profils multi-électrodes, enregistrées avec un résistivimètre Syscal R1+ Switch 24 (IRIS Instruments) équipé de 24 électrodes. Le premier profil est de direction Nord-Sud tandis que le second est de direction Est-Ouest. Nous avons choisi un dispositif d'acquisition de type dipôle-dipôle avec un espacement inter-électrodes égale à 4m pour les deux profils. Le choix du dispositif dipôle-dipôle est dicté par sa sensibilité aux variations latérales des résistivités ainsi qu'à sa profondeur d'investigation importante, en particulier aux extrémités du profil ([Gharibi et Bentley, 2005](#)). Les inconvénients de ce dispositif résident dans son faible RSB, en particulier lorsqu'on augmente les valeurs du facteur « n », ainsi qu'à l'apparition d'anomalies complexes. La profondeur d'investigation peut être calculée en utilisant la fonction de sensibilité ou les dérivées de Fréchet. En ce qui nous concerne, nous avons choisi d'utiliser la profondeur moyenne formulée par [Edwards \(1977\)](#). C'est ainsi, qu'avec un espacement « a » variant de 3 à 12m et un facteur d'espacement « n » variant de 1 à 5, nous avons pu obtenir une profondeur d'investigation d'environ 17.72m (voir [Tableau 5.1](#)). Un total de 231 mesures a été obtenu durant chaque séquence.

Array type		z_e/a	z_e/L
Wenner alpha		0.519	0.173
Dipole-dipole	n = 1	0.416	0.139
	n = 2	0.697	0.174
	n = 3	0.962	0.192
	n = 4	1.220	0.203
	n = 5	1.476	0.211
	n = 6	1.730	0.216
Equatorial dipole-dipole			
	n = 1	0.451	0.319
	n = 2	0.809	0.362
	n = 3	1.180	0.373
	n = 4	1.556	0.377
Wenner - Schlumberger			
	n = 1	0.52	0.173
	n = 2	0.93	0.186
	n = 3	1.32	0.189
	n = 4	1.71	0.190
	n = 5	2.09	0.190
	n = 6	2.48	0.190
Pole-dipole	n = 1	0.52	
	n = 2	0.93	
	n = 3	1.32	
	n = 4	1.71	
	n = 5	2.09	
	n = 6	2.48	
Pole-Pole		0.867	

Tableau 5.1. Profondeur d'investigation moyenne (Z_e) correspondant à différents dispositifs de longueur totale L (d'après Edwards, 1977).

Pour réaliser les pseudosections de résistivités apparentes, nous avons fait recours au logiciel Res2dinv (Loke et Barker 1996). Ce logiciel détermine automatiquement un modèle de résistivité en deux dimensions à partir des données recueillies d'une campagne d'imagerie électrique (Griffiths et Barker 1993). Le modèle ainsi généré, divise le sous-sol en un nombre de blocs rectangulaires. Puis, un sous programme de modélisation génère les résistivités apparentes calculées, soit par la technique des différences finies, soit par la technique des éléments finis. Le but est de déterminer les résistivités des différents blocs qui permettent de générer une pseudosection de résistivités apparentes similaire aux mesures réelles. L'optimisation a pour but de réduire les différences entre les résistivités apparentes calculées et celles mesurées en ajustant les valeurs de résistivités des différents blocs. On utilise l'EQM ou RMS (en anglais) pour quantifier cette erreur. Le processus d'inversion est réalisé par la

méthode d'optimisation des moindres carrées à contraintes lisses (Smoothness-constrained least-squares method) (deGroot-Hedlin et Constable, 1990 ; Loke et Barker, 1996) selon l'équation suivante :

$$(J^T J + uF) d = J^T g \quad (5.1)$$

Avec

J : matrice des dérivées partielles

d : vecteur de perturbation du modèle

g : vecteur de dispersion

u : facteur de pondération

$$F = f_x f_x^T + f_z f_z^T$$

Où f_x est un filtre « flatness » horizontal

f_z est un filtre « flatness » vertical

On note enfin que cette partie est traduite du manuel du logiciel Res2Dinv (www.geoelectrical.com).

5.2.2. Paramétrisation

Dans le but d'étudier les performances des différents filtres vis-à-vis du bruit, indépendamment de tout autre facteur, nous avons choisi de garder tous les paramètres de l'algorithme d'inversion (Res2dinv) inchangés. En addition, nous avons mis en place des programmes sous Matlab® pour les filtres SD-ROM et MED. Les filtres MED, avec différentes fenêtres, ont été utilisés comme moyen de comparaison avec le filtre SD-ROM.

Notre choix des paramètres du filtre SD-ROM s'est fixé sur l'utilisation d'une fenêtre glissante de longueur 9, avec quatre valeurs seuils (T_1, T_2, T_3, T_4). Concernant le filtre MED, nous avons utilisé des fenêtres de longueurs 3 et 5. Les valeurs seuils ont été choisies après plusieurs tests, selon le niveau du bruit, ainsi que le signal de sortie désiré. Nos tests ont montré qu'une fenêtre de longueur 9 produit les meilleurs résultats (ceci n'exclut pas que, pour d'autres données, d'autres fenêtres produirons de meilleurs résultats).

Le RSB ainsi que l'EQM ont été choisis pour quantifier les performances de chaque filtre. Notons toutefois, qu'une faible EQM n'est pas toujours synonyme du meilleur modèle géologique (ceci est dû en partie à ce que les solutions au problème inverse ne sont généralement pas uniques). Le nombre maximum d'itérations a été fixé à 3 et 4, selon les variations observées sur l'EQM. La fenêtre glissante utilisée est de taille 9, centrée à

l'échantillon $x(i,j)$. Les échantillons du voisinage de $x(i,j)$ ont été réarrangés dans un vecteur 1D appelé $w(k)$ ($k=(i-1)*l+j$ où $l=3$ et $i=j=l$), qui sera ensuite filtré par le filtre SD-ROM. Le vecteur ainsi résultant pourra être utilisé itérativement afin d'améliorer le processus de filtrage. Le vecteur filtré est réarrangé pour revenir au cas 2D, et la fenêtre est déplacée à la position suivante jusqu'à la fin de la séquence.

5.2.3. Géologie de la région d'étude

Le site de mesures, où l'on a recueilli les données de résistivité électrique, est une vallée alluviale de direction Est-Ouest, localisée dans la région de Khalfoun (W. de Sétif) à 280 km à l'Est d'Alger (Algérie). Le site est une plaine plate, et sa lithologie est constituée principalement de deux couches :

- 1- une couche superficielle d'âge Quaternaire d'une épaisseur d'environ 9m, incluant une couche argilo-carbonatée avec des éléments rocailloux caractérisées par des résistivités inférieures à $10 \Omega.m$, surmontant des dépôts argileux à éléments rocailloux avec des épaisseurs variant de 1 à 3m et des résistivités supérieures à $20 \Omega.m$.
- 2- Un substratum d'âge Mio-Pliocène constitué principalement d'argiles rouges et caractérisé par des résistivités inférieures à $10 \Omega.m$.

5.3 Résultats et discussions

5.3.1 Cas du premier profil

Dans le but de démontrer la robustesse du filtre SD-ROM, nous l'avons comparé avec le filtre MED standard. Les résistivités mesurées du premier profil (Figure 5.6) ont été utilisées comme entrée des différents filtres. Trois valeurs anormales sont visibles à la limite droite inférieure à des distances de 58, 62 et 66m respectivement, et sont montrées par des flèches. Ces valeurs ont été identifiées comme du bruit à cause des mesures anormalement faibles du potentiel.

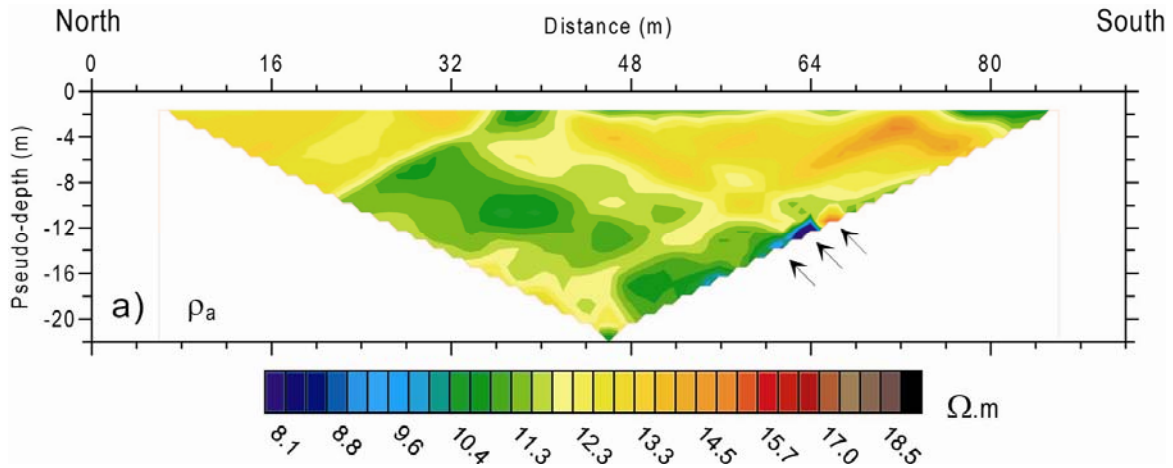


Figure 5.6. Données réelles tirées du premier profil de direction Nord-Sud (Les flèches indiquent les valeurs anormales).

5.3.2. Analyse des pseudosections différences

Nous avons utilisé une technique assez répandue en traitement de l'image pour montrer l'efficacité de chaque filtre dans l'élimination du bruit. La technique en question est la différence, où « l'image différence », est obtenue par calcul des différences entre les points correspondants des deux images. Le but de cette méthode est d'améliorer la visibilité et de mettre en relief les différences entre les deux images. Les résultats des sections différences sont illustrés sur la [Figure 5.7](#).

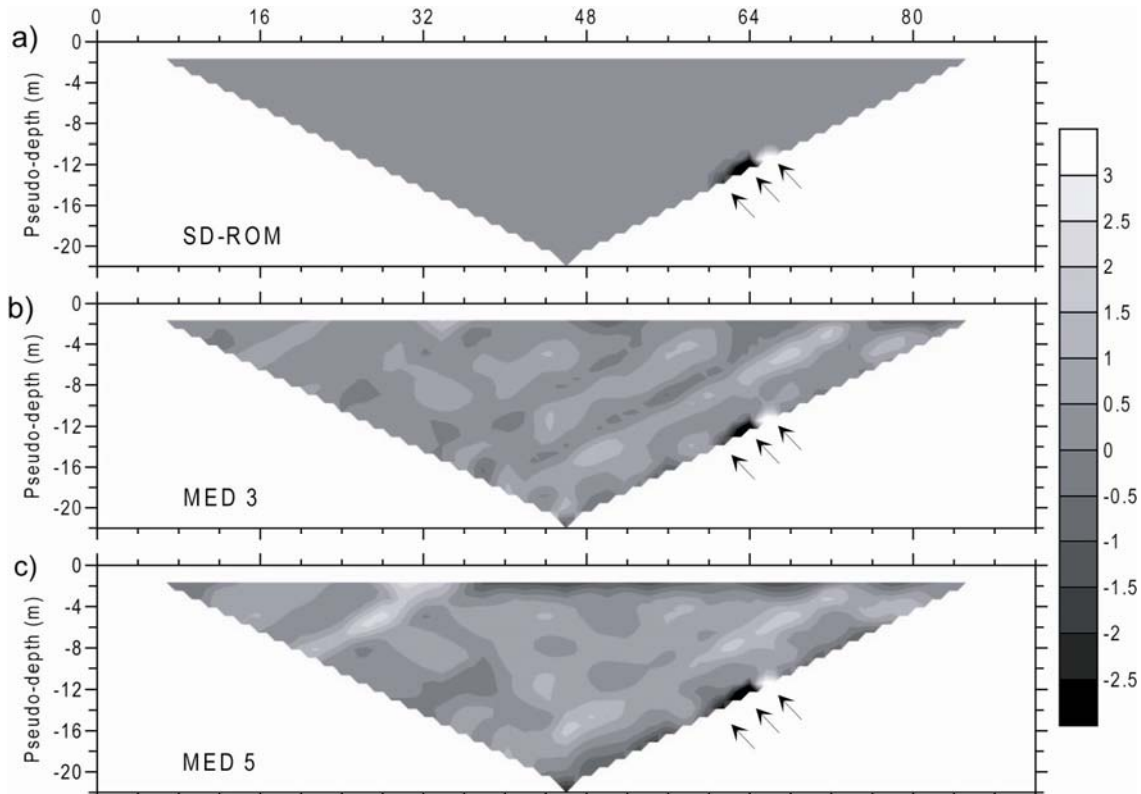


Figure 5.7. Pseudosections différences après filtrage a) SD-ROM, b) médian (fenêtre 3) et c) médian (fenêtre 5).

La [Figure 5.7a](#) montre la différence calculée entre les résistivités apparentes brutes et les valeurs filtrées par le filtre SD-ROM. Il est clair que les impulsions identifiées sur la [Figure 5.6](#) ont été parfaitement localisées et isolées, sans effets indésirables sur le signal. Notons que le nombre et l'amplitude des impulsions localisées et éliminées par le filtre SD-ROM dépendent du choix des valeurs seuils et de la taille de la fenêtre utilisée. Dans notre cas, la taille de la fenêtre a été fixée à 9, les valeurs seuils ont été déterminées par des tests. Les [Figures 5.7b et 5.7c](#) montrent les pseudosections différences calculées entre les données filtrées, par le filtre MED avec fenêtres de tailles 3 et 5, et les données non filtrées. Là aussi, nous remarquons que les impulsions (flèches) ont été localisées et éliminées. Néanmoins, les pseudosections différences montrent une distribution de valeurs associées au signal mais considérées par le filtre MED comme bruit, ce qui montre clairement l'inefficacité du filtre MED à différencier entre le bruit et le signal, à l'inverse du filtre SD-ROM qui a montré une grande robustesse. Ces valeurs extraites sont dues au comportement même du filtre MED. En effet, la particularité du filtre MED standard est qu'il remplace l'échantillon central par la médiane, même si ce dernier n'est pas affecté par le bruit. Ces valeurs extraites du signal, si elles ne sont pas corrigées, peuvent affecter grandement l'inversion et conduire ainsi, à une interprétation erronée.

5.3.3. Evaluation des performances des filtres

Lors de l'évaluation des performances des différents filtres, nous avons aussi utilisé le rapport RSB en dB. Dans notre cas, le RSB est calculé en utilisant la formule suivante :

$$RSB(dB) = 10 \log_{10} \left(\frac{\sum_i (x(i)^2)}{\sum_i (F(i) - x(i))^2} \right) \quad (5.1)$$

Avec $x(i)$ les résistivités considérées non affectées par le bruit et $F(i)$ les sorties filtrées ($i=1,2,\dots,n=231$). $x(i)$ ont subies un filtrage manuel. Nous avons aussi utilisé la déviation standard (σ) afin de calculer l'étendue des différentes valeurs. Une déviation standard importante indique un éloignement des valeurs de la moyenne, tandis qu'une petite déviation indique un regroupement autour de la moyenne.

Le **Tableau 5.1** résume les résultats, en termes de RSB et de déviation standard, des données filtrées par les filtres MED et SD-ROM.

Données du premier profil	MED (fenêtre 5)	MED (fenêtre 3)	SD-ROM
RSB en dB	18.27	29.35	48.07
Déviati on standard (σ)	1.53	0.81	0.95

Tableau 5.2. Etude comparative, en termes de RSB et déviation standard, des différents filtres du premier profil.

Il est clair que le filtre SD-ROM surpasse la performance des filtres MED, où les grandes valeurs du RSB signifient un bon filtrage. Les faibles valeurs de la déviation standard peuvent être interprétées par l'élimination des pics des données par les filtres MED.

Comme nous l'avons mentionné plus haut, le bruit a une grande influence sur le processus d'inversion, et son atténuation est primordiale avant inversion.

5.3.4. Impact du filtrage sur l'inversion

Les résultats d'inversion des données issues du premier profil sont illustrés par la [Figure 5.8](#).

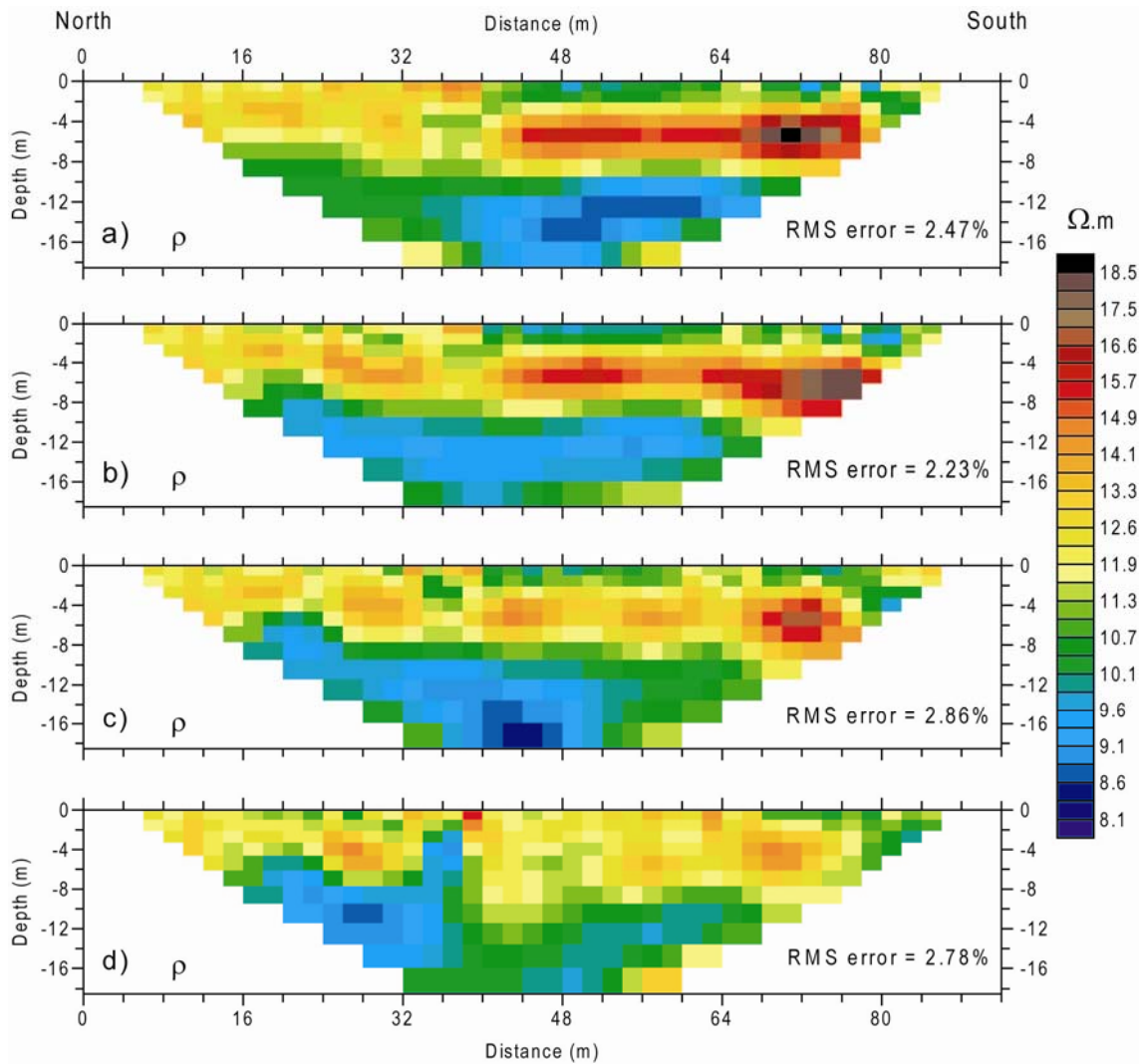


Figure 5.8. Pseudosections inversées issues du premier profil après filtrage a) manuel, b) SD-ROM, c) médian (fenêtre 3) et d) médian (fenêtre 5).

L'inversion de la section supposée non-contaminée par le bruit est montrée sur la [Figure 5.8a](#), où les impulsions ont été localisées et éliminées. Cette pseudosection a été utilisée comme référence lors des comparaisons avec les sections inversées issues des différents filtres. La [Figure 5.8a](#) possède une erreur RMS relativement faible de 2.47%. Comme attendu, le filtre SD-ROM donne la meilleure section inversée ([Figure 5.8b](#)) avec une erreur RMS de 2.23%, laquelle est très proche de la section de référence ([Figure 5.8a](#)).

5.3.5. Interprétation géologique

Le modèle géologique ainsi obtenu est en accord avec les études géotechniques qui ont montré une zone de faible résistivité ($< 9 \Omega.m$) près de la surface, correspondant aux argiles ainsi qu'une zone plus résistante à 3.8m de profondeur, juste en-dessous de la couche superficielle, correspondant à des dépôts alluviaux ou bien à une couche argilo-sableuse proche de la surface dans la partie nord de la zone d'étude. Le substratum argileux est défini par de faibles résistivités ($< 9 \Omega.m$) à 9.8m de profondeur. Les résultats obtenus par les filtres MED produisent une erreur de 2.86% dans le cas d'une fenêtre de longueur 3 (Figure 5.8c), et une erreur RMS de 2.78% dans le cas d'une fenêtre de longueur 5 (Figure 5d). La qualité d'un filtre donné dépend de son habilité à éliminer un type de bruit donné, en préservant, au mieux, le signal d'intérêt et sa résolution. Pour des fenêtres larges (p.ex. 5), le filtre MED tend à détruire la résolution de l'image comme montré en Figure 5.8d.

5.3.6. Cas du second profil

Notre second exemple correspond à des données de résistivités entachées par du bruit relativement fort (Figure 5.9).

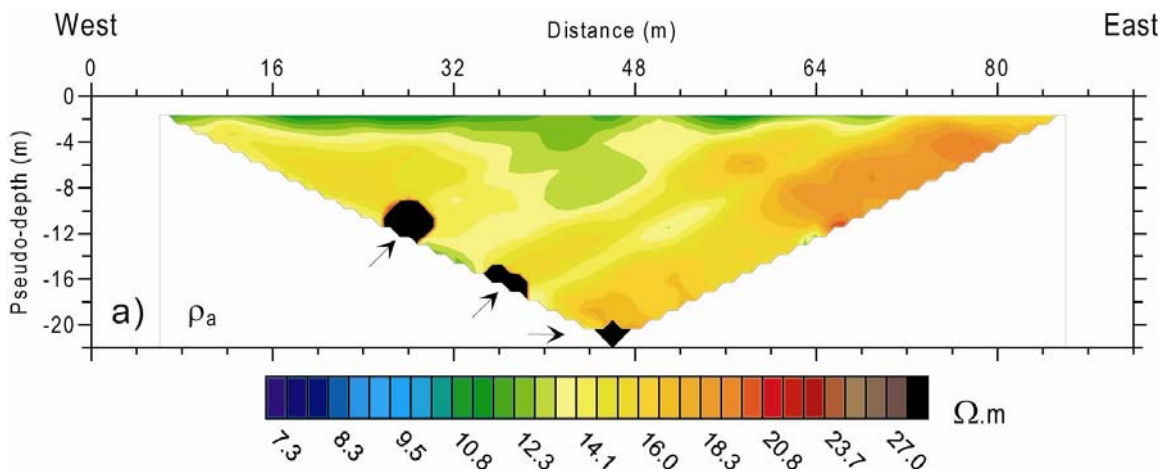


Figure 5.9. Données réelles issues du second profil de direction Est-Ouest (Les flèches indiquent des valeurs de résistivités anormalement élevées).

Ces anomalies (indiquées par des flèches) sont localisées à des distances de 28, 36 et 44m de l'origine, et possèdent des résistivités de 232, 733 et 1677 $\Omega.m$ respectivement. Elles sont très élevées comparées aux résistivités de l'arrière plan.

5.3.7 Examen des pseudosections différences

Nous avons appliqué la même procédure de traitement que pour les données du premier profil. On peut constater sur les pseudosections différences, que le filtre SD-ROM a réussi à détecter et à éliminer les impulsions (flèches), sans aucun changement des autres parties de la pseudosection, contribuant ainsi à l'amélioration du signal d'intérêt (Figure 5.10a). Ces impulsions possèdent des valeurs positives et sont caractérisées par une couleur blanche sur la pseudosection différence.

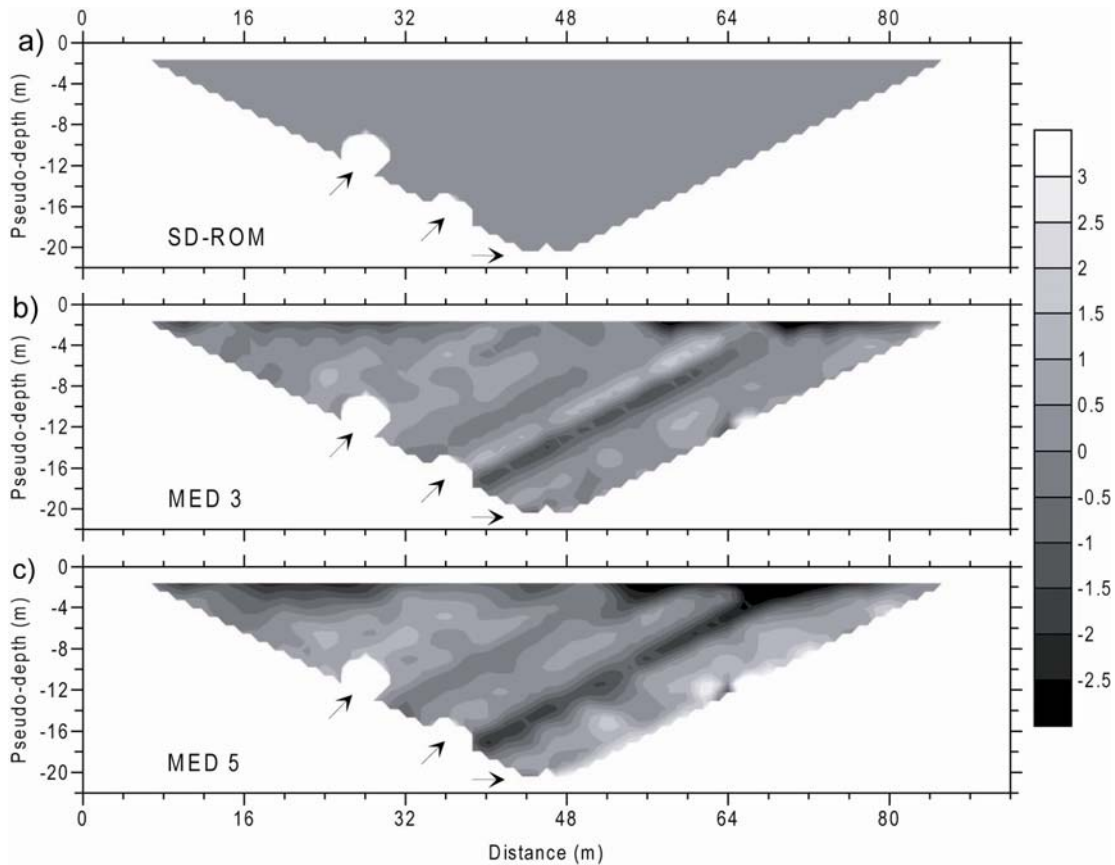


Figure 5.10. Pseudosections différences après filtrage SD-ROM (a), médian (fenêtre 3) (b) et médian (fenêtre 5) (c).

Les Figures 5.10b et 10c montrent que les filtres MED ont pu aussi isoler et éliminer les impulsions liées aux bruits (indiquées par des flèches). Néanmoins, comme le montre les pseudosections différences, les autres parties de la pseudosection ont été affectées par les filtres MED.

5.3.8. Evaluation quantitative des performances des filtres

Le [tableau 5.3](#) montre que le filtre SD-ROM produit le meilleur RSB et une déviation élevée par rapport au filtre MED (fenêtre 3). Le moins performant des filtres, est le filtre MED de longueur 5, dont la longueur a pour effet d'exagérer le lissage.

Données réelles du second profil	MED (fenêtre 5)	MED (fenêtre 3)	SD-ROM
RSB en dB	15.82	24.44	49.54
Déviatiion standard (σ)	2.32	1.77	1.96

Tableau 5.3. Etude comparative, en termes de RSB et de déviation standard, des différents filtres du second profil.

5.3.9. Impact du filtrage sur l'inversion

La [Figure 5.11](#) présente les résultats d'inversion des données du second profil. La [Figure 5.11a](#) illustre la section inversée de référence (après traitement manuel des bruits impulsif) avec une erreur RMS égale à 4.06%.

On comparant la section inversée après passage du filtre SD-ROM ([Figure 5.11b](#)) avec la section de référence ([Figure 5.11a](#)) nous pouvons constater la grande similitude entre les deux, avec une erreur RMS de 4.1% pour le filtre SD-ROM. L'interprétation géologique du modèle présenté en [Figure 5.11b](#) met en évidence une couche superficielle conductrice, correspondant à des argiles de résistivités inférieures à 10 Ω .m, surmontant une couche alluvionnaire d'âge Quaternaire à 3.8m de profondeur et ayant des résistivités supérieures à 19 Ω .m, suivie d'un substratum argileux de faibles résistivités à 9.8m de profondeur. Il est toutefois important de noter la présence d'une structure conductrice verticale révélée par l'inversion à 38m de l'origine du profil et que nous avons interprétée comme un prolongement du substratum argileux vers la surface.

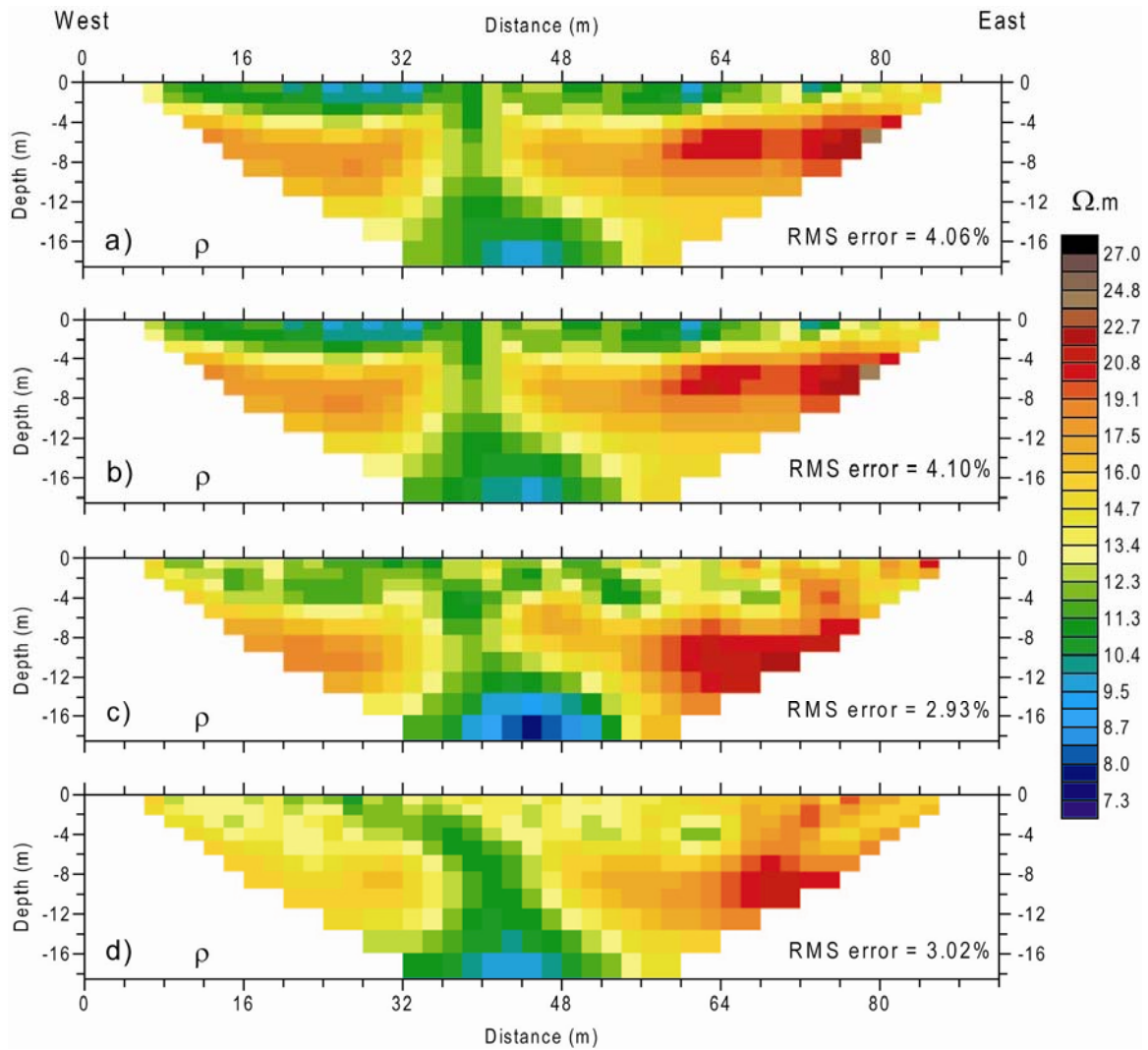


Figure 5.11. Pseudosections inversées issues du second profil après filtrage a) manuel, b) SD-ROM, c) médian (fenêtre 3) et d) médian (fenêtre 5).

Les [Figures 5.11c et 11d](#) montrent les résultats d'inversion, après application des filtres MED de longueurs 3 et 5, avec des erreurs RMS égales à 2.93% et 3.02 %, respectivement. Comme mentionné plus haut, une faible erreur RMS n'est pas toujours synonyme d'un modèle correspondant à la réalité géologique. Ainsi, le filtre MED a donné une erreur RMS plus faible que celle du filtre SD-ROM. Néanmoins, une comparaison entre les pseudosections inversées ([Figures 5.11c et 11d](#)) après filtrage MED et la section de référence ([Figure 5.11a](#)) montre des différences notables, en particulier près de la surface. Toutefois, l'interprétation géologique reste pratiquement la même que celle donnée par le filtre SD-ROM. Notons, que la structure verticale mise à jour par les [Figure 5.11a et 11b](#), n'a pas été proprement identifiée

sur les pseudosections issues du filtrage MED. Là aussi, nous avons remarqué que c'est le filtre SD-ROM qui donne le modèle géologique de subsurface le plus plausible.

5.3.10. Conclusion

Les inhomogénéités de surface, tout comme les bruits aléatoires impulsifs, génèrent des anomalies dans les données de résistivités électriques, pouvant affecter grandement les résultats de l'inversion.

Dans ce chapitre, nous avons appliqué avec succès le filtre SD-ROM pour l'élimination des bruits impulsifs et des effets des NSIs des données électriques, synthétiques et réelles. Le filtre SD-ROM appartient à la famille des filtres dits de statistiques d'ordre ; le principal avantage de ces filtres est leurs relatives simplicités et leurs habilités à distinguer le signal du bruit. Nous avons démontré que l'application du filtre SD-ROM améliore grandement le processus d'inversion. Les résultats obtenus sont très encourageants et l'efficacité du filtre SD-ROM a été vérifiée, comparé aux filtres MED. Nos travaux ont montré que l'utilisation d'une fenêtre de longueur 9, dans le cas du filtre SD-ROM, donne les résultats les plus optimaux, et qu'une utilisation itérative du filtre donne de meilleurs résultats lors de l'inversion.

Nous recommandons d'autres recherches afin d'investir les effets des inhomogénéités en 3D sur les mesures des résistivités. En plus, nous recommandons le recours à de nouvelles méthodes, telles les techniques évolutionnaires à réseaux de neurones et à base de logique floue, pour le choix des seuils, dans le cas du filtre SD-ROM.

Conclusion générale, nouveautés scientifiques et perspectives

L'objectif principal de ce travail de thèse était d'introduire une classe particulière de filtres adaptatifs et non-linéaires à base de statistiques d'ordre au domaine de la géophysique. Les filtres ainsi proposés, sont couramment utilisés en traitement de l'image, mais trouvent peu ou pas d'application en géophysique. Nous avons ainsi souligné les points essentiels liés à leurs utilisations ; du point de vue de la paramétrisation, du choix du critère de détection du bruit et de la résolution recherchée. Nous nous sommes basés lors de l'application de ces filtres sur des données issues de deux domaines de la géophysique : la sismique réflexion et la méthode électrique à courant continu.

Notre principale contribution s'est portée essentiellement sur l'étude et l'implémentation de trois filtres tirés du domaine du traitement de l'image. Cette implémentation s'est rendue possible grâce à l'analogie faite entre les représentations des données géophysiques (section sismique et pseudosection de résistivités apparentes) et les images.

Le début de cette thèse traite de différentes méthodes de filtrages en géophysique qui existent dans la littérature. Les méthodes de filtrage linéaire en temps ont pour inconvénient de garder leurs paramètres inchangés pendant l'opération de filtrage. Tandis que, les méthodes opérant dans le domaine fréquentielle sont incapables de séparer la partie bruit de la partie signal, lorsque leurs plages de fréquences se superposent, ou lorsqu'elles possèdent des caractéristiques semblables, telle que la vitesse. Les autres méthodes de filtrage sont le plus souvent complexes, et se comportent comme une boîte noire, ce qui rend le suivi des calculs et la validation des résultats très difficiles.

Nous avons ensuite décrit les trois filtres : diffusion anisotropique, trilatéral et SD-ROM sous leur aspect théorique et algorithmique. Ces filtres ont pour objectif d'atténuer le bruit afin d'améliorer l'inversion et de préserver la résolution en vue de faciliter l'interprétation. Ces filtres opèrent dans le domaine spatio-temporel et sont associés aux filtres non-linéaires et adaptatifs. Le principal avantage de ces filtres réside dans leur habilité à distinguer le bruit du signal avec un minimum d'intervention humaine, tout en ayant la capacité d'adaptation en cas de changement des données d'entrée. Les expérimentations menées dans cette thèse ont montré une grande robustesse de ces filtres vis-à-vis du bruit aléatoire et ont souligné leur importance lors de la manipulation de données en amplitude préservée.

Le premier filtre, dit de diffusion anisotropique, tire son fonctionnement de la solution des équations aux dérivées partielles (PDE). Il reprend le principe du phénomène physique de la diffusion. Nous avons étudié et appliqué ce filtre sur des données de sismique réflexion synthétiques et réelles. Pour rendre l'algorithme de filtrage fonctionnel, nous avons considéré la section sismique comme une fonction à deux variables à la manière d'une image 2D, où l'intensité d'un pixel est remplacée par l'amplitude d'un événement sismique. La paramétrisation du filtre se base essentiellement sur la détermination de la constante de diffusion et le pas d'échantillonnage espace-temps. Nous avons simulé des données de sismique réflexion. Le modèle ainsi proposé, tentait d'imiter la complexité du sous-sol, en incorporant des variations latérales d'épaisseurs et des discontinuités au niveau des réflecteurs. En plus, pour rapprocher nos données de la réalité, nous avons ajouté du bruit aléatoire à différents niveaux. D'après nos expérimentations, le filtre montrait une grande robustesse vis-à-vis du bruit. Nous avons ainsi constaté que le nombre maximum d'itérations doit satisfaire le compromis entre l'atténuation de bruit et la préservation d'amplitude. Nous avons donc pu montrer, qu'une augmentation du nombre d'itérations rendait les données plus lisses, en revanche le signal d'intérêt est détérioré. Nous l'avons démontré par une comparaison qualitative des sections différences. De plus, en mesurant le rapport RSB en dB et l'erreur MSE, nous avons mis en évidence que les données filtrées sont de meilleures qualités. L'analyse conjointe dans le domaine temps fréquence, à l'aide des spectrogrammes, a confirmé nos résultats. Nous avons aussi expérimenté le pouvoir de récupération des données perdues (exemple d'une trace morte) par le filtre anisotropique. Là aussi, le filtre a pu reconstituer la trace à partir de son voisinage. Ce résultat est très intéressant, car la trace reconstituée tire ces valeurs de son voisinage immédiat, ce qui diminue le risque de biaiser le résultat final. Les simulations ont montré qu'il était possible de reconstituer jusqu'à 6 traces mortes successives. Les performances du filtre diminuent en augmentant le bruit, mais le résultat final reste visuellement bon.

Nous avons ensuite appliqué le filtre de diffusion anisotropique sur des données réelles n'ayant subi aucun traitement à l'exception d'une sommation brute. Nous avons pu constater, qu'après seulement trois itérations, la qualité de la section initiale a été grandement améliorée, et que les bruits présents ont été fortement atténués. Le temps de calcul d'une telle opération est très faible, comparé à d'autres méthodes de traitement. Ce qui rend la technique très séduisante, en particulier lors de la manipulation de données de très forte densité, comme pour le cas des données 3D. Notons enfin, qu'une telle opération peut être pratiquée avant ou après sommation.

La deuxième méthode s'intéresse à l'application du filtre trilatéral (première fois en géophysique). L'objectif de l'utilisation de ce filtre est de supprimer le bruit, tout en gardant les bords inchangés. En sismique, il se traduit par une atténuation des bruits aléatoires et/ou cohérents, une meilleure résolution et une amélioration de la continuité des réflecteurs. Ce filtre se base sur une distribution préférentielle de poids. L'idée est que tous les échantillons proches dans l'espace, similaires en magnitude et possédant une certaine statistique (ROAD) se voient attribués de forts poids, et sont ainsi privilégiés. Le filtre trilatéral est un filtre de voisinage et se base sur le filtre bilatéral de Tomasi et Manduchi (1998). Nous avons exploité ce filtre pour atténuer le bruit aléatoire présent sur des points de tirs synthétiques et réels. Les résultats de simulation ont montré une bonne atténuation du bruit avec un recouvrement optimal du signal, même en présence de bruit aussi important que le signal sismique. Nous avons aussi appliqué le filtre trilatéral sur des données réelles d'une campagne 3D du Sud algérien. Nous avons modifié la stratégie de filtrage, au lieu de calculer la statistique ROAD avec un ordre croissant, nous avons choisi l'ordre décroissant. Ce qui nous a permis de séparer la partie bruit cohérent, caractérisée par le Ground roll de très fortes amplitudes, par simple soustraction qui a fourni le résultat filtré.

Nous avons également exploité le filtre SD-ROM pour supprimer le bruit impulsionnel présent sur des données électriques. Notre application du filtre SD-ROM est une première en géophysique. En effet, ce filtre est devenu une référence en traitement de l'image et de la parole. Malheureusement, nous n'avons enregistré aucune application de ce filtre en géophysique. Son mode de fonctionnement et son implémentation faciles l'ont rendu très populaire. Le filtre SD-ROM procède en deux temps : 1) identification 2) estimation. Nous l'avons utilisé sur des données électriques à courant continu, synthétiques et réelles. Les bruits dans la partie synthétique, ont été simulés par les inhomogénéités de surface (NSIs). Ces NSIs produisent des anomalies de résistivités, sous forme de V inversé. Ces bruits sont assimilés à des bruits cohérents. Les simulations ont montré que le filtre SD-ROM détecte et remplace efficacement les valeurs anormales. Ces mêmes résultats ont été comparés à ceux produits par le filtre médian classique. Nous avons aussi procédé à l'inversion des données filtrées, et là aussi, le filtre SD-ROM permettait d'avoir des modèles avec l'erreur RMS la plus faible. Les données réelles utilisées sont issues de deux profils électriques multi-électrodes enregistrés dans la région de Sétif (Algérie). Ces données sont entachées par des valeurs anormales « impulsions ». L'application du filtre SD-ROM a permis une localisation parfaite (comme indiqué par les pseudosections différences) et une bonne restitution des valeurs, comparé au filtre médian. L'opération de filtrage est suivie par une inversion afin de

construire un modèle du sous-sol, sous forme de blocs de résistivités. Les données inversées, qui ont subi auparavant un filtrage par SD-ROM, ont été les plus proches des données de référence filtrées à la main. Les résultats d'inversion des filtres médians n'ont pas été concluants.

Nouveautés scientifiques et contributions

Durant cette thèse, nous avons exploré, pour la première fois en géophysique, plusieurs méthodes de filtrage non-linéaires et adaptatives qui permettent d'améliorer la qualité des données géophysiques. Les premières contributions de notre travail de recherche apparaissent au deuxième chapitre de ce mémoire. Ces contributions peuvent se résumer aux points suivants :

- Nous avons pu réaliser une recherche inter- disciplinaire croisant le domaine du traitement de l'image au domaine de la géophysique
- Nous avons innové, à notre connaissance, c'est une première en géophysique, en introduisant le filtre SD-ROM. Nous avons appliqué le filtre SD-ROM pour éliminer les bruits impulsifs des données de prospection électrique. Ce procédé de filtrage a grandement amélioré les résultats d'inversion.
- Le filtre trilatéral est l'autre innovation. Nous avons ainsi introduit ce filtre pour la première fois afin d'atténuer le bruit aléatoire des données sismiques. Les résultats ont montré sa grande robustesse vis-à-vis du bruit.
- Le filtre anisotropique trouve très peu d'applications en sismique. Or, nous l'avons appliqué avec succès pour le filtrage des données bruitées, et nous l'avons utilisé comme moyen de restauration des traces mortes (contribution).

Perspectives

Les perspectives de recherches portent à la fois sur des améliorations et sur des applications de nos filtres.

La perspective qui vient le plus aisément à l'esprit est le développement de moyens automatiques, ne nécessitant aucune intervention humaine, dans le choix des seuils et des constantes des différents filtres. Les méthodes évolutionnistes et biologiques sont les plus séduisantes, en particulier les algorithmes génétiques et les fourmis artificielles. Ce choix des paramètres reste la principale difficulté dans la paramétrisation des filtres, car il se base sur l'expérience de l'utilisateur et sur la nature des données à l'entrée de chaque filtre.

Une autre perspective consiste à appliquer les filtres à d'autres domaines de la géophysique, en particulier en magnétisme et en gravimétrie. Dans ce dernier cas, la séparation des anomalies régionales et locales constitue la principale tâche de filtrage. Or, un filtre, comme le filtre trilatéral, pourra être facilement implémenté à condition de bien le paramétrer et l'adapter pour la séparation.

Enfin, d'autres méthodes de haut niveau (logique floue) peuvent être utilisées pour piloter les filtres proposés.

Liste des travaux scientifiques

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

1. Publications Internationales :

1. Application of signal dependent rank-order mean filter to the removal of noise spikes from 2D electrical resistivity imaging data. *Near Surface Geophysics*. **7** (3), 159-169, **2009**. Published on behalf of the European Association of Geoscientists and Engineers (EAGE) by Wiley-Blackwell.
2. Acoustic Impedance Inversion by Feedback Artificial Neural Network. *Journal of Petroleum Science and Engineering*, in press, doi: 10.1016/j.petrol.2009.09.012, published by Elsevier.
3. Application of Radial Basis Function Artificial Neural Network to Seismic Data Inversion. *Computers & Geosciences*, 35, 2338-2344, **2009**, published by Elsevier.
4. Application of feedback connection artificial neural network to seismic data filtering. *Comptes rendus Geoscience*, 340, 335-344, **2008** Published by Elsevier Masson SAS on behalf of Académie des Sciences, Paris-France.
5. Tomographic Velocity Images by Artificial Neural Networks. *Journal of Engineering and Applied Sciences* **2**(4):775-782, **2007**.
6. Contribution of seismic attributes in the direct detection of hydrocarbon, case history of a permit in the south Algerian Sahara. *Bulletin du Service Géologique de l'Algérie* 14 (1):37-51, **2003**.

2. Communications Internationales

1. Trilateral filter: A new tool for coherent and random noise removal from seismic noisy data. *4ème congrès maghrébin de géophysique appliqué (CMGA 4), Hammamet (Tunisie)*. 26-28 Mars **2009**.
2. Seismic data inversion by Elman artificial neural network. *4ème congrès maghrébin de géophysique appliqué (CMGA 4), Hammamet (Tunisie)*. 26-28 Mars **2009**.
3. Application des diagraphies de test aux réservoirs Dévonien du champ de Tin Fouye Tabankort. bassin d'Illizi, Algérie. *4ème congrès maghrébin de géophysique appliqué (CMGA 4), Hammamet (Tunisie)*. 26-28 Mars **2009**.

4. Application of Elman artificial neural networks to filter seismic data. 4th international symposium on hydrocarbons & chemistry. Ghardaïa, Algeria, March 21-23, **2008**.
5. Tomographic Velocity Images by Artificial Neural Networks. 2nd international symposium on hydrocarbons & chemistry. Ghardaïa, Algeria, March 21-23, **2004**.

3. Communications Nationales

1. Enhancement of geophysical field data inversion via median filtering, application to 2D and 3D resistivity imaging. 7^{ème} journées Scientifiques et Techniques (JST 7) de SONATRACH. Oran 28-29 Novembre **2006**.
2. Introduction à l'utilisation des attributs sismiques pour la détection directe des hydrocarbures (Etude de cas au Sahara Algérien). 1^{ères} journées internationales sur l'exploration pétrolière en Afrique (1^{ères} JIEPA). 6-8 Juin **2005** Sidi-Fredj (Algérie).
3. Application des attributs sismiques à la détection directe des hydrocarbures. 4^{ème} congrès national de la physique et ses applications (4^{ème} CNPA). Sidi Fredj, 21-23 Novembre **2000** (Algérie).

Bibliographie

1. Abreu E., Lightstone M., Mitra S.K. et Arakawa K. 1996. A new efficient approach for the removal of impulse noise from highly corrupted images. *IEEE Transactions on Image Processing* special issue on Non-linear Image Processing **5**, 1012-1025.
2. Ames W.F. 1992. Numerical methods for partial differential equations. *Academic Press Inc.*, Boston, third edition.
3. Al-Yahia K.M. 1991. Application of the partial Karhunen-Loève transform to suppress random noise in seismic sections. *Geophysical prospecting* **39**, 77-93.
4. Barash D. 2002. A fundamental relationship between bilateral filtering, adaptative smoothing, and the nonlinear diffusion equation. *IEEE transactions on Pattern Analysis and Machine Intelligence PAMI* **24**, 844-7.
5. Bednar J.B. 1983. Applications of median filtering to deconvolution, pulse estimation and statistical editing of seismic data. *Geophysics* **48**, 1598-1610.
6. Bernstone C., Dahlin T., Ohlsson T. et Hogland W. 2000. DC-resistivity mapping of internal landfill structures: two pre-excavation surveys. *Environmental Geology* **39**, 360-371.
7. Bobachev A.A., Marchenko M.N., Modin I.N., Pervago E.V., Urusova A.V. et Shevnin V.A. 1996. New approaches to electrical soundings of horizontally inhomogeneous media. *Physics of the solid earth* **31**, 1075-1086.
8. Canales L.L. 1984. Random noise attenuation. 54th Ann. Mtg. Soc. Expl. Geoph., Expanded Abstracts, Session S10.1, 525-527.
9. Cassano E. et Rocca F. 1973. Multichannel linear filters for optimal rejection of multiple reflections. *Geophysics* **38**, 1053-1061.
10. Cassano E. et Rocca F. 1974. After-stack multichannel filters without mixing effects. *Geophysical prospecting* **22**, 330-334.
11. Chambers J.E., Meldrum P.I., Ogilvy R.D. et Wilkinson P.B. 2005. Characterisation of a NAPL-contaminated former quarry site using electrical impedance tomography. *Near Surface Geophysics* **3**, 79-90.
12. Chambers J.E., Kuras O., Meldrum P.I., Ogilvy R.D. et Hollands J. 2006. Electrical resistivity tomography applied to geologic, hydrogeologic, and engineering investigations at a former waste-disposal site. *Geophysics* **71**, B231-B239.

13. Chandra C., Moore M.S. et Mitra S.K. 1998. An efficient method for the removal of impulse noise from speech and audio signals. *Proceeding of the IEEE symposium on circuits and systems* **4**, 206-208.
14. Choudhury P. et Tumblin J. 2003. The trilateral filter for high contrast images and meshes. In Proceedings of the Eurographics Symposium on Rendering EGSR 2003, Eurographics, 186-196.
15. Claerbout J.F. 1975. Slant –stacks and radial traces: *Stanford Expl. Project Report SEP-5*, 1-12.
16. Claerbout J.F 1983. Ground roll and radial traces: *Stanford Expl. Project Report SEP-35*, 43-53.
17. Claerbout J.F 1985. Imaging the earth's interior: *Stanford Expl. Project*, 216-219.
18. Cooper J. 1998. Introduction to partial differential equations with MATLAB. *Birkhäuser*, Boston, pp 540.
19. Cornacchiulo D. et Bagtzoglou A.C. 2004. Geostatistical reconstruction of gaps in near-surface electrical resistivity data. *Vadoze Zone Journal* **3**, 1215-1229.
20. Dahlin T. 1996. 2D resistivity surveying for environmental and engineering applications. *First Break* **14**, 275-283.
21. Dahlin T. et Zhou B. 2004. A numerical comparison of 2D resistivity imagings with 10 electrode arrays. *Geophysical Prospecting* **52**, 379-398.
22. Daily W., Ramirez A. et Johnson R. 1998. Electrical impedance tomography of a perchloroethylene release. *Journal of Environmental and Engineering Geophysics* **2**, 189–201.
23. Deighan A.J. et Watts 1997. Ground-Roll suppression using the wavelet transform. *Geophysics* **62**, 1896-1903.
24. Deriche R. et Faugeras O. 1996. Les EDP en traitement des images et vision par ordinateur. *Traitement du signal* **13**(6).
25. Dey A. et Morrison H.F. 1979. Resistivity modelling for arbitrarily shaped three-dimensional structures. *Geophysics* **44**, 753-780.
26. Djarfour N., Aïfa T., Baddari k., Mihoubi A. et Ferahtia J. 2008. Application of feedback connexion artificial neural network to seismic data filtering. *Comptes rendus Geoscience*, **340**, 335-344.
27. Edwards L.S. 1977. A modified pseudosection for resistivity and induced-polarization. *Geophysics* **42**, 1020-1036.

28. Essenreiter R. 1999. Identification and attenuation of multiple reflections with neural networks. *Phd Thesis*, der Fakultät für Physik der Universität Karlsruhe (TH), 103pp.
29. Fatti J.L., Smith G.C., Vail P.J., Strauss P.J. et Levitt P.R. 1994. Detection of gas in sandstone reservoir using AVO analysis: A 3-D seismic case history using the Geostack technique. *Geophysics* **59**, 1362-1376.
30. Fehmers G.C. et Hocker C. 2003. Fast structure interpretation with structure oriented filtering. *Geophysics* **68**, 1286-1293.
31. Ferahtia J., Djarfour N., Baddari K. et Guèrin R. 2009. Application of signal dependent rank-order mean filter to the removal of noise spikes from 2D electrical resistivity imaging data. *Near surface geophysics* **7**, 159-169.
32. Galbraith J.N. et Wiggins R.A. 1968. Characteristics of optimum multichannel stacking filters. *Geophysics* **33**, 36-48.
33. Garnett R., Huegerich T., Chui C. et He.W. 2005. A universal noise removal algorithm with an impulse detector. *Image processing, IEEE Transactions on* **14**, 1747-1754.
34. Gharibi M. et Bentley L.R. 2005. Resolution of 3-D electrical resistivity images from inversions of 2-D orthogonal lines. *Journal of Environmental and Engineering Geophysics* **10**, 339–349.
35. Gersztenkorn A. et Scales J.A. 1988. Smoothing seismic tomograms with alpha-trimmed means. *Geophysical Journal* **92**, 67–72.
36. Gonzalez R.C. et Woods R.E. 2002. *Digital image processing*. 2nd ed. Addison Wesley.
37. Guitton A. 2002. Coherent noise attenuation using inverse problems and prediction error filters. *First break* **20**, 161-167.
38. Gulunay N. 1986. FXDECON and complex wiener prediction filter. *56th Annual Internat. Mtg., Soc. Expl. Geophys.* Expanded Abstracts. Session POS2.10, 279-281.
39. Gonzalez R.C. et Woods R.E. 1992. *Digital image processing*. 2nd ed. Addison Wesley.p??
40. Grellier S., Guèrin R., Robain H., Bobachev A., Vermeersch F., Tabbagh A., 2008. Monitoring of leachate recirculation in a bioreactor landfill by 2D electrical resistivity imaging. *The Journal of Environmental and Engineering Geophysics*. doi:10.2113/JEEG13.4.351.

41. Guérin R., Begassat P., Benderitter Y., David J., Tabbagh A. et Thiry M. 2004. Geophysical study of the industrial waste land in Mortagne-du-Nord (France) using electrical resistivity. *Near Surface geophysics* **2**, 137-143.
42. Griffiths D.H. et Barker R.D. 1993. Two-dimensional resistivity imaging and modelling in areas of complex geology. *Journal of Applied Geophysics* **29**, 211-226.
43. deGroot-Hedlin C. et Constable S. 1990. Occam's inversion to generate smooth, two-dimensional models from magnetotelluric data. *Geophysics* **55**, 1613-1624.
44. Hagen D.C. 1982. The application of principal components analysis to seismic data sets. *Geoexploration* **20**, 93-111.
45. Haines S.S., Guitton A., et Biondi B. 2007. Seismoelectric data processing for surface surveys of shallow targets. *Geophysics* **72**, G1-G8.
46. Hale D. 2001. Atomic images- a method for meshing digital imaging. *Proceedings of the 10th international meshing roundtable*, 185-196.
47. Harris P.E. et White R.E. 1997. Improving the performance of f-x prediction filtering at low signal-to-noise ratios. *Geophysical prospecting* **45**, 269-302
48. Hashemi H., Javaherian A. et Babuska R. 2008. A semi-supervised method to detect seismic random noise with fuzzy GK clustering. *Journal of Geophysics and Engineering* **5**, 457-468.
49. Hayes M.H. 1996. Statistical digital signal processing and modelling. John Wiley & Sons, Inc.
50. Haykin S. 1996. Adaptive filter theory. Prentice-Hall, pp 989.
51. Henderson R.P., et Webster J.G. 1978. An impedance camera for spatially specific measurements of thorax. *IEEE Transaction on Biomedical Engineering* **25**, 250-254.
52. Henley D.C. 2003. Coherent noise attenuation in the radial trace domain. *Geophysics* **68**, 1408-1416.
53. Jones I.F. et Levy S. 1987. Signal-to-noise ratio enhancement in multi-channel seismic data via the Karhunen-Loève transform. *Geophysical prospecting* **35**, 12-32.
54. Kalman R.E. 1960. A new approach to linear filtering and prediction problems. *Transactions of the ASME, Journal of Basic Engineering* **82**, 35-45.
55. Karsli H. et Bayrak Y. 2003. Using the Wiener-Levinson algorithm to suppress ground-roll. *Journal of Applied Geophysics* **55**, 187-197.

56. Kolmogorov A.N. 1939. Sur l'interpolation et extrapolation des suites stationnaires. C. R. Acad. Sci. Paris **208**, 2043-2045. (English translation in Kailath, 1977).
57. LaBrecque D.J., Miletto M., Daily W., Ramirez A. et Owen E. 1996. The effect of noise on Occam's inversion of resistivity tomography data. *Geophysics* **61**, 538-548.
58. Larner K., Chambers R., Yang M., Lynn W. et Wai W. 1983. Coherent noise in marine seismic data. *Geophysics* **48**, 854-886.
59. Li X.G., Ulrych J. 1996. Coherent noise filtering using a 2-D Gabor transform. *SEG Technical program Expanded Abstracts*, 1180-1183.
60. Linville, A.F. et Meek R.A. 1995. A procedure for optimally removing localized coherent noise. *Geophysics* **60**, 191-203
61. Loke M.H., Acworth I. et Dahlin T. 2003. A comparison of smooth and blocky inversion methods in 2D electrical imaging surveys. *Exploration geophysics* **34**, 182-187.
62. Loke H. et Barker R.D. 1996. Rapid least-squares inversion of apparent resistivity pseudosections by a quasi Newton method. *Geophysical Prospecting* **44**, 131-152.
63. Lytle R.J. et Dines K.A. 1978. *An impedance camera: a system for determining the spatial variation of electrical conductivity*. California University, Livermore (USA). Lawrence Livermore Lab. Technical report. Lawrence Livermore National Laboratory UCRL-52413.
64. March D.W. et Bailey A.D. 1983. A review of the two-dimensional transform and its use in seismic processing. *First Break* **1**, 9-21.
65. Margrave G. 1991. "rnoise" toolbox for MATLAB, CREWES project. University of Calgary, Canada.
66. Mayne W.H. 1962. Common reflection point horizontal data stacking techniques. *Geophysics* **27**, 927-938.
67. McAndrew A. 2004. Introduction to digital image processing. 1st Ed. *Thomson*.
68. Meyerhoff J.N. 1966. Horizontal stacking and multichannel filtering applied to common-depth-point seismic data. *Geophysical prospecting* **14**, 441-454.
69. Morton K.W. et Mayers D.F. 1994. Numerical solution of partial differential equations: An introduction. *Cambridge University Press*. pp 227.
70. Naess O.E. et Bruland L. 1985. Stacking methods other than simple summation. In Fitch, A.A. (Ed.). *Developement in geophysical exploration methods-6*. London. Elsevier Applic. Sci. Publishers.

71. Manglem K.S., Singh P.K.B. et Singh S.B. 2002. Impulse noise removal using fuzzy rank selection technique. *Second International Conference on Electrical and Computer Engineering*, Dhaka (Bangladesh), 140-143.
72. Manolakis D.G., Ingle V.K. et Kogon S.M. 2005. Statistical and adaptive signal processing. *Artech house*. pp 796.
73. Moore M.S., Gabbouj M. et Mitra S.K. 1999. Vector SD-ROM filter for removal of impulse noise from colour images. *EURASIP conference DSP for multimedia communications and services*, Krakow (Poland).p??
74. Moore M.S. et Mitra S.K. 2000. Statistical threshold design for the two-state signal dependent rank order mean filter. *Proceedings of the International Conference on Image Processing* **1**, 904-907.
75. Pelton W.H., Rijo L. et Swift C.M.Jr. 1978. Inversion of two-dimensional resistivity and induced polarization data. *Geophysics* **43**, 788-803.
76. Perona P. et Malik J. 1990. Scale-space and edge detection using anisotropic detection. *IEEE transactions on pattern analysis and machine intelligence* **12**, 629-639.
77. Ristau J.P. et Moon W.M. 2001. Adaptive filtering of random noise in 2-D geophysical data. *Geophysics* **66**, 342-349.
78. Ritz M., Robain H., Pervago E., Albouy Y., Camerlynck C., Descloitres M. et Mariko A. 1999. Improvement to resistivity pseudosection modelling by removal of near-surface inhomogeneity effects: application to a soil system in south Cameroon. *Geophysical Prospecting* **47**, 85-101.
79. Robinson E. 1984. Statistical pulse compression. *Proc. IEEE* **72**, 1276-1289.
80. Sasaki Y. 1989. Two-dimensional joint inversion of magnetotelluric and dipole-dipole resistivity data. *Geophysics* **54**, 254-262.
81. Sasaki Y. 1992. Resolution of resistivity tomography inferred from numerical simulation. *Geophysical Prospecting* **40**, 453-464.
82. Spitzer R., Nitsche F. O. et Green A. 2001. Reducing source-generated noise in shallow seismic data using linear and hyperbolic τ -p transformations. *Geophysics* **66**, 1612-1621.
83. Tarantola A. 1987. Inverse problem theory. *Elsevier*. pp 630.
84. Tatham R.H., Keenay J.K. et Nojonen I. 1982. Application of the tau-p transform (slant stack) in processing reflection data. *52nd Annual international SEG meeting*, 3-4.

85. Tatham R.H. 1984. Multi-channel filtering of seismic data. *Proceeding of the IEEE*, **72**, 1357-1369.
86. Tomasi C. et Manduchi R. 1998. Bilateral filtering for gray and color images. *Proc. IEEE 6th. Int. Conf. Computer vision, New Delhi, India*, 839-846.
87. Trickett S.R. 2003. F-xy Eigenimage noise suppression. *Geophysics* **68**, 751-759.
88. Tukey J.W. 1974. Nonlinear (nonsuperposable) methods for smoothing data," in Congress Records. EASCON, p. 673.
89. Ulrych T.J., Sacchi M.D. et Freire S.L.M. 1999. Eigenimage processing of seismic sections, in Kirlin, R.L., and Done W.J. (Ed.). Covariance analysis for seismic processing. *Soc. Expl. Geophys.* 241-274.
90. Ursin B. et Zheng Y. 1985. Identification of seismic reflections using singular value decomposition. *Geophysical prospecting* **33**, 773-799.
91. Voss D. et Hearn S. 2003. Seismic multiple attenuation based on pre-stack reflectivity modelling. *ASEG 16th Geophysical Conference and Exhibition*, extended abstracts.
92. Weiglen A.B. 1999. Multiple attenuation: an overview of recent advances and the road ahead. *The leading edge*, 40-44.
93. Wenkai L. 2002. Adaptive noise attenuation of seismic image using singular value decomposition and texture direction detection. *Image Processing 2002*. Proceeding. 2002 International conference on. 465-468.
94. Wiener N. 1949. Extrapolation, Interpolation and Smoothing of Time Series, with Engineering Applications, New York: Wiley. Originally appears in 1942 as a classified National Defense Research Council Report. Also published under the title Time Series Analysis by MIT Press.
95. Wilkinson P.B., Meldrum P.I., Chambers J.E., Kuras O. et Ogilvy R.D. 2006a. Improved strategies for the automatic selection of optimised sets of electrical resistivity tomography measurement configurations. *International Geophysical Journal* **167**, 1119-1126.
96. Wilkinson P.B., Kuras O., Meldrum P.I., Chambers J.E. et Ogilvy R.D. 2006b. Comparison of the spatial resolution of standard and optimised electrical resistivity tomography arrays. *Proceedings of the 12th meeting of the EAGE Near Surface Geophysics Conference*, Helsinki (Finland).
97. Yilmaz O. 1987. Seismic data processing. *Society of Exploration Geophysicists*, 526pp.

98. Zhou B. et Dahlin T. 2003. Properties and effects of measurement errors on 2D resistivity imaging surveying. *Near Surface Geophysics* **1**, 105-117.

Near Surface Geophysics

Volume 7 · Number 3 · June 2009

- 147 Integration of multidimensional archaeogeophysical data using supervised and unsupervised classification
E.G. Ernenwein
- 159 Application of signal dependent rank-order mean filter to the removal of noise spikes from 2D electrical resistivity imaging data
J. Ferahtia, N. Djarfour, K. Baddari and R. Guérin
- 171 Imaging near-surface inhomogeneities using weighted multipath summation
V. Shtivelman, S. Keydar and M. Mikenberg
- 179 Comparison of instruments for geoelectrical soil mapping at the field scale
R. Gebbers, E. Lück, M. Dabas and H. Domsch
- 191 Imaging and monitoring tree-induced subsidence using electrical resistivity imaging
G.M. Jones, N.J. Cassidy, P.A. Thomas, S. Plante and J.K. Pringle
- 207 High-resolution magnetic gradient and electrical resistivity tomography survey at the Plaka Petrified Forest Park in Lesvos Island, Greece
G. Vargemezis, N. Zouros, P. Tsourlos and I. Fikos
- 217 Gravity anomalies of pull-apart basins having finite strike length with depth dependent density: a ridge regression inversion
V. Chakravarthi
- 227 Hydrocarbon contaminated soil: geophysical-chemical methods for designing remediation strategies
D. Coria, V. Bongiovanni, N. Bonomo, M. de la Vega and M.T. Garea



EAGE

EUROPEAN
ASSOCIATION OF
GEOSCIENTISTS &
ENGINEERS

www.nearsurfacegeophysics.org

Application of signal dependent rank-order mean filter to the removal of noise spikes from 2D electrical resistivity imaging data

Jalal Ferahtia^{1*}, Nouredine Djarfour¹, Kamel Baddari¹ and Roger Guérin²

¹ Laboratoire de Physique de la Terre (LABOPHYT), Université M'hamed Bougara, FHC Boumerdès 35000, Algeria

² Université Pierre et Marie Curie-Paris 6, UMR 7619 Sisyphe, Case 105, 4 place Jussieu, 75252 Paris Cedex 05, France

Received January 2008, revision accepted March 2009

ABSTRACT

It is well-known that when inverting two-dimensional (2D) electrical resistivity data, a major source of errors is the presence of noise and in particular noise spikes. The popular median filter is often applied to the removal of single spikes. However, when the signal is highly corrupted with successive spikes, the median filter performance is poor. This paper deals with the use of the signal dependent rank-order mean filter for the detection and removal of noise spikes from highly corrupted 2D electrical resistivity imaging data. In addition to its computational simplicity, this filter is shown to be extremely robust, even in the presence of very strong noise, especially when it is applied recursively. The signal dependent rank-order mean filter was tested on 2D synthetic resistivity data contaminated by near-surface inhomogeneities and the results confirmed efficient removal of the disturbances normally associated with near-surface inhomogeneities. The signal dependent rank-order mean filter was also applied to field data and demonstrated its ability to significantly improve the accuracy of the inversion process and to produce good visual results in the inverted electrical sections.

INTRODUCTION

Electrical resistivity imaging was originally proposed by Henderson and Webster (1978) as a medical imaging tool and was introduced by Lytle and Dines (1978) for the purpose of geophysical studies. In the recent past, many authors have established the basis of the electrical resistivity imaging method (Edwards 1977; Pelton *et al.* 1978; Dey and Morrison 1979). This method is however still a relatively new geophysical imaging technique (Griffiths and Barker 1993). Electrical resistivity imaging is used in environmental and pollution studies (Daily *et al.* 1998), engineering problems (Chambers *et al.* 2006; Grellier *et al.* 2008), landfill site investigations (Bernstone *et al.* 2000; Guérin *et al.* 2004; Chambers *et al.* 2005), as well as for the characterization of near surface geological characteristics (Dahlin 1996). The electrical resistivity imaging technique measures the Earth's resistivity by injecting a direct current (DC) into the ground and measuring the resulting potential. There are many different types of arrays used by the electrical resistivity imaging technique, the most common being the Wenner array, the Schlumberger array, the dipole-dipole array, the pole-pole array, the pole-dipole array and recently, more sophisticated optimized arrays such as those proposed by

Wilkinson *et al.* (2006a, 2006b). Each array has its advantages and disadvantages (Dahlin and Zhou 2004), depending on various parameters such as the depth of investigation, sensitivity to changes in horizontal and vertical resistivity, noise level and signal strength. The resistivity measurements (i.e., apparent resistivity pseudosections) are inverted to produce a subsurface model in the form of 2D or 3D sections. Robust algorithms (Sasaki 1989; deGroot and Constable 1990; Sasaki 1992; Loke and Barker 1996; Loke *et al.* 2003) and the development of computers have improved the speed and accuracy of the inversion process. However, as for any measurement of physical properties, coherent noise and random noise are an integral part of the collected data. If not corrected for, noise can compromise the inversion process and the resulting resolution, making it very difficult to estimate the correct geological model (LaBrecque *et al.* 1996; Ritz *et al.* 1999; Zhou and Dahlin 2003). It is therefore important to understand its origin and to make appropriate efforts to minimize its influence.

In the case of electrical resistivity data, there are different kinds of noise: i) coherent noise due to geological structures such as near-surface inhomogeneities (Ritz *et al.* 1999), 3D effects or electrode effects, ii) random and spiky noise such as telluric currents, electrode polarization (Cornacchiulo and Bagtzoglou 2004), or noise resulting from measurements with a poor signal-to-noise ratio. Near-surface inhomogeneities are defined as small

* jalelferahtia@yahoo.fr

shallow geological bodies characterized by anomalous resistivity values compared to background resistivities.

Bobachev *et al.* (1996) and Ritz *et al.* (1999) demonstrated that near-surface inhomogeneities can distort electric sounding curves as well as pseudosections, leading to erroneous geological interpretations. The electrical inversion process can thus be adversely affected by near-surface inhomogeneities, particularly when inspecting deep structures, such that distortions that occur before inversion must be filtered out.

Strategies for electrical resistivity noise attenuation are needed in order to preserve the signal of interest. Common methods are manual ‘cleaning’, or by the application of smoothing linear filters such as the low-pass filter. However, these methods have the disadvantages of being time-consuming, in addition to gaps introduced in the field data and of being sensitive to anomalous observations. In image and speech processing, the median filter is widely used for noise reduction (Gonzalez and Woods 1992), as it is particularly effective in the presence of impulse noise, also referred to as salt-and-pepper noise. This filter belongs to the class of statistical filters that are non-linear spatial filters. The principle of a median filter is quite simple: inside a sliding window of odd size l , the data subset is sorted and the central point sample of rank p is then replaced by the median value of the set. However, when the signal is highly corrupted by successive impulses, a large window is needed and this leads to the replacement of a larger number of uncorrupted samples by the median value, thus lowering the high-frequency content of the signal. In image processing, this leads to the removal of details and to the blurring of images (Chandra *et al.* 1998; Manglem *et al.* 2002). Gersztenkorn and Scales 1988 used a combination of median and mean filters to attenuate the noise in order to improve the efficiency of inversion algorithm.

In the present paper, we propose the use of the signal dependent rank-order mean filter (Abreu *et al.* 1996; Chandra *et al.* 1998; Moore *et al.* 1999; Moore and Mitra 2000) for the removal of noise spikes from highly corrupted 2D electrical resistivity data, thereby avoiding the disadvantages of the median filter.

This paper is structured as follows: firstly, we briefly describe the signal dependent rank-order mean filter theory. Secondly, we use forward and inverse modelling to generate synthetic data from a 2D resistivity model and test different filters with this data. We then report on the efficient use of the signal dependent rank-order mean filter on field data. Finally, we provide various concluding remarks and suggestions for future work.

SIGNAL DEPENDENT RANK-ORDER MEAN FILTER

Theory

The signal dependent rank-order mean filter is a non-linear filter, which is extensively used in image and speech processing and belongs to the class of decision-based filters, or state-conditioned filters. The signal dependent rank-order mean filter is shown to

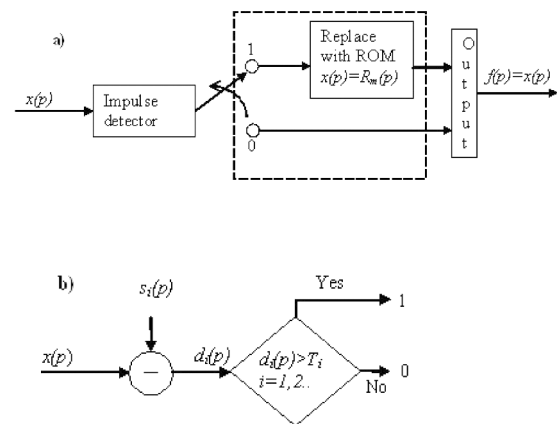


FIGURE 1 Schematic diagrams of: a) the signal dependent rank-order mean filter, b) the impulse detector. $x(p)$ represents the central sample of a 3×3 window, $S_i(p)$ an eight element rank-order observation vector containing the neighbouring of $x(p)$ (excluding $x(p)$, itself) inside the window, $d_i(p)$ is the rank-order differences vector, T_i threshold values ($i = 1, \dots, 4$), $f(p)$ is the output and $R_m(p)$ defines the rank-order mean (ROM).

be sufficiently robust in the presence of both Gaussian noise and noise spikes (Abreu *et al.* 1996; Moore *et al.* 1999). The main advantage of this filter is its ability to distinguish between unwanted noise and useful signals.

The signal dependent rank-order mean filter proceeds in two steps: firstly it detects corrupted samples and secondly it replaces the corrupted samples with an appropriate estimate. A schematic diagram of the signal dependent rank-order mean filter is provided in Fig. 1.

Let $\rho(i)$ denote a set of observed apparent resistivities ($i = 1, 2 \dots n$), where $x(p)$ is a sample of rank p . Firstly, $x(p)$ is tested by the impulse detector (Fig. 1b), which determines whether or not $x(p)$ is a corrupted sample. If $x(p)$ is corrupted, the filter switches to position 1 and the corrupted sample is replaced by the rank-order mean R_m . Otherwise, the filter remains at position 0 and the uncorrupted sample passes through it, with no alteration. In image processing, a window of dimensions $n \times n$ is used and the same filtering principle is applied.

Algorithm

In the following, we describe the algorithm of a 1D sliding window signal dependent rank-order mean filter, according to the guidelines suggested by Chandra *et al.* (1998):

1. Consider $\rho_{i,j}$ to be the measured apparent resistivities,
2. Select a window T of size $l \times l$ (in this case 3×3 , i.e., $l = 3$) centred on sample p ,
3. Rearrange the elements of window T into a 1D vector $w_i(p)$ that corresponds to left-to-right, top-to-bottom mapping from the 3×3 window onto the 1D vector:

$$w_i(p) = [x_{-4}(p), x_{-3}(p), x_{-2}(p), x_{-1}(p), x(p), x_1(p), x_2(p), x_3(p), x_4(p)]^T \tag{1}$$

4. Exclude the central sample $x(p)$ from vector $w_i(p)$, such that the resulting vector $\hat{w}_i(p)$ contains eight elements:
 $\hat{w}_i(p) = [x_{.4}(p), x_{.3}(p), x_{.2}(p), x_{.1}(p), x_1(p), x_2(p), x_3(p), x_4(p)]^T$, (2)
5. Rearrange the elements of vector $\hat{w}_i(p)$ in ascending order:
 $\hat{w}_i(p) \text{ rearranged} = S_i(p) = [s_1(p), s_2(p), s_3(p), s_4(p), s_5(p), s_6(p), s_7(p), s_8(p)]^T$, (3)
 such that $s_1(p) \leq s_2(p) \dots \leq s_8(p)$,
6. Define rank-order differences $d_i(p)$ between the elements of $S_i(p)$ and $x(p)$:

$$d_i(p) = \begin{cases} s_i(p) - x(p) & \text{if } x(p) \leq R_m \\ x(p) - s_{9-i}(p) & \text{otherwise} \end{cases} \quad \text{for } i = 1, \dots, 4, \quad (4)$$

where R_m is the rank-order mean (ROM):

$$R_m = [s_4(p) + s_5(p)]/2. \quad (5)$$

- It can be seen in expression (4) that $d_i(p)$ is used to detect the presence of corrupted samples,
7. If any of the corresponding $d_i(p)$ values is greater than the corresponding threshold T_i , then identify sample $x(p)$ as being corrupted.
- Steps 6 and 7 above define the impulse detector.
8. Replace the corrupted sample $x(p)$ by the rank-order mean R_m (expression (5)) and
 9. Move window T to the next position and repeat operations 1–9 until T reaches position $n-l$.

The threshold values T_i can be obtained statistically or by means of empirical tests (Moore and Mitra 2000). In the case of a sliding window of size 9, four threshold values are needed. For a larger window size, a larger number of thresholds is required. The performance of this algorithm can be enhanced by applying it recursively to the same data set (Chandra *et al.* 1998).

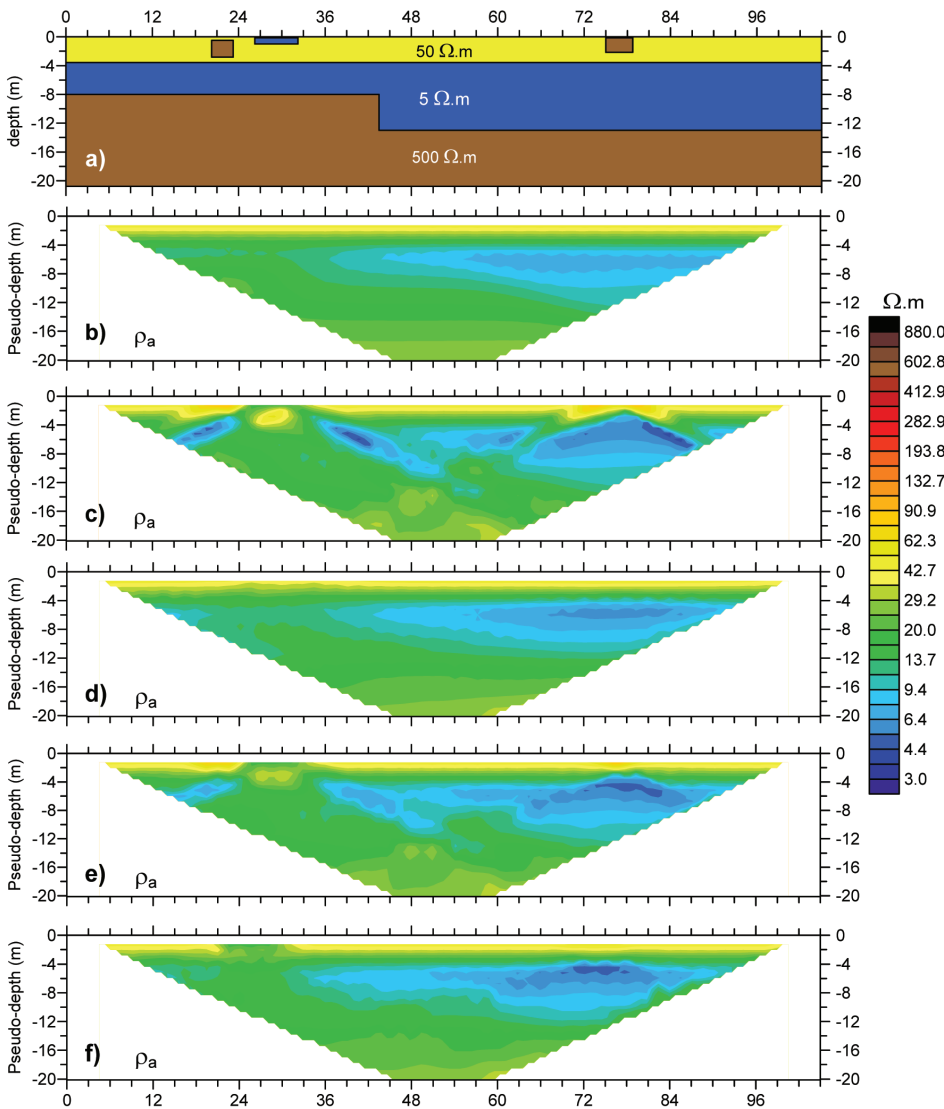


FIGURE 2

a) Model used to generate synthetic 2D resistivity data, in the presence of near-surface inhomogeneities. The ‘a’ spacing varies from 3–9 m for $n = 1, 2, \dots, 8$. Calculated apparent resistivity pseudosection for the synthetic model: b) without near-surface inhomogeneities, c) in the presence of near-surface inhomogeneities, d) after signal dependent rank-order mean filtering, e) after median filtering (window size 3), f) after median filtering (window size 5).

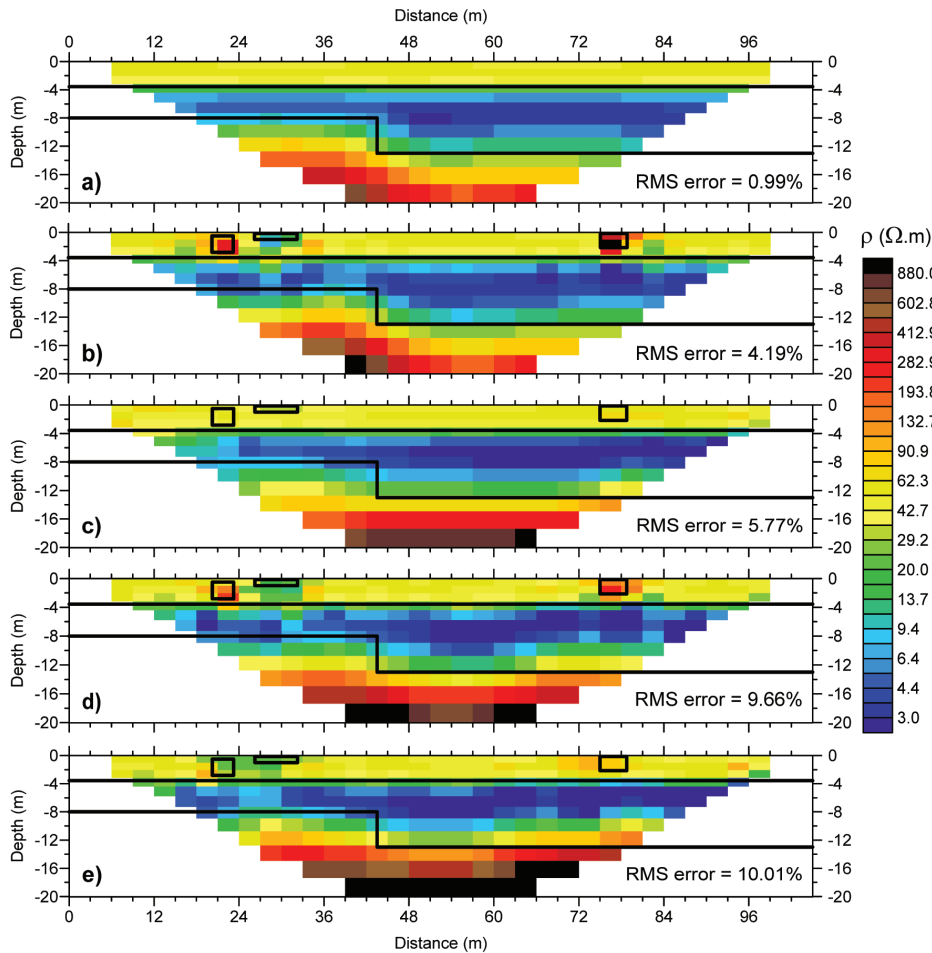


FIGURE 3

Inverted sections of synthetic data (black outlines indicate the synthetic model): a) without near-surface inhomogeneities, b) with near-surface inhomogeneities, c) after signal dependent rank-order mean filtering, d) after median filtering (window size 3), e) after median filtering (window size 5).

APPLICATION TO SYNTHETIC DATA

Model description

The signal dependent rank-order mean filter was tested on synthetic 2D electrical resistivity data generated by the forward modelling program Res2dmod. The 2D resistivity model (Fig. 2a) comprised three layers, with the following parameters: a basement with a resistivity of $500 \Omega.m$ at a depth of 8 m and lateral variations in thickness, covered by overburden layers comprising a conductive strata with a resistivity of $5 \Omega.m$ at 2.8 m depth and characterized by lateral variations in thickness and a superficial layer with a resistivity of $50 \Omega.m$ and a thickness of 3.56 m. Three blocks with different resistivities ($5 \Omega.m$ and $500 \Omega.m$) were embedded in the superficial layer, in order to simulate near-surface inhomogeneities. In addition, we constructed a model with no near-surface inhomogeneities, with the same parameters as the synthetic model. The synthetic data were generated for the dipole-dipole array, using spacings varying from 3–9 m and a separation factor 'n' ranging from 1–8, such that the median depth of investigation was about 20.16 m. A total of 552 measurements were simulated. We used the finite-difference method to generate apparent resistivity values.

Synthetic data analysis

Figure 2(b) shows the computed apparent resistivity values of the model without near-surface inhomogeneities, where its inverted section was used as a reference when comparing the performance of different filters. The calculated apparent resistivity pseudosection shows strongly inverted V-shaped anomalies associated with the near-surface inhomogeneities (Fig. 2c). As expected, the apparent resistivity pseudosection after filtering by the signal dependent rank-order mean filter (Fig. 2d) gives better results than the median filter (Figs 2e and 2f). As seen in Fig. 2(d), the signal dependent rank-order mean filter successfully removed inverted V-shaped anomalies associated with near-surface inhomogeneities. In addition, all three layers of the initial model are well imaged compared to the apparent resistivity pseudosection without near-surface inhomogeneities (Fig. 2b). However, the step at around 43 m is quite distorted and shifted to the left of the pseudosection. Numerous anomalous values related to near-surface inhomogeneities are still visible with the use of a median filter with a window size of 3 (Fig. 2e). This indicates that the median filter fails to remove near-surface inhomogeneity effects. In comparison with a median filter with a window size of 3, a median filter with a window size of 5 (Fig. 2f) provides better

results. Neither the step nor the layers form can be clearly imaged by median filters compared to the signal dependent rank-order mean filter.

The synthetic apparent resistivity values were filtered with median and signal dependent rank-order mean filters and the resulting inverse sections are shown in Fig. 3. Figure 3(a) is used as a reference, because it is free of near-surface inhomogeneity effects and has a low root mean square (rms) error of 0.99%. One can note the good agreement between the inverted section and the synthetic model (black outlines). The inverse section of the synthetic model with near-surface inhomogeneity shown in Fig. 3(b) has an rms error of 4.19% and can be seen to have anomalously high and low resistive zones associated with near-surface inhomogeneities, which coincide in magnitude and location with the synthetic model, as indicated by the black outlines. Figure 3(c) shows an inverted section after signal dependent rank-order mean filtering, with an rms error of 5.77%. All of the features in the synthetic model are well retrieved, with satisfactory removal of the anomalous resistivities produced by near-surface inhomogeneities. Note however, that the basement edge located at a horizontal distance of 43 m and at a depth of 8 m in the synthetic model is shifted to the left by about 36 m in the signal dependent rank-order mean filtered section.

Median filters with different window sizes were also used to remove resistivity anomalies associated with near-surface inhomogeneities. The resulting sections (Figs 3d and 3e) have an rms error of respectively 9.66% and 10.01%, for a median filter with a window size of 3 and 5. One can see that the anomalous resistivity zones located at 21 m, 30 m and 75 m are not completely removed and that the shape of the basement is totally altered. It is important to notice that changing the mesh grid may improve inversion results. However, this will increase the number of unknown parameters of the inversion process, which has an impact on the stability of the algorithm.

APPLICATION TO REAL DATA

The surveyed site used to validate our approach is an east-west striking alluvial valley, located in the region of Sétif, 280 km east of Algiers (Algeria).

Geology of the site

The area is a flat plain and its lithology consists of two main layers:

1. A top layer dating from Quaternary age whose thickness is about 9 m, it includes an encrusted, 6 m thick carbonate-clay layer with stony elements characterized by resistivities of less than 10 Ω m, overlaying silty, stony alluvial deposits with thicknesses varying from 1–3 m and resistivities greater than 20 Ω m.
2. A substratum consisting mainly of reddish clay dating from the Mio-Pliocene age, characterized by resistivities below 10 Ω m.

Parameters and implementation

The data presented in this study were collected from two multi-electrode profiles, using a Syscal R1+ Switch 24 resistivity meter (IRIS Instruments) equipped with 24 electrodes. The first profile (PE1) was taken in the north-south direction, whereas the second one (PE2) was made in the east-west direction. A dipole-dipole array with a 4 m electrode spacing was chosen for both profiles. The choice of the dipole-dipole array was guided by its sensitivity to horizontal changes in resistivity and its superior depth coverage at the profile extremities (Gharibi and Bentley 2005). The shortcomings of this type of array include its low signal-to-noise ratio (SNR), especially in the case of a large spacing factor 'n' and the appearance of complex anomalies. The depth of investigation achieved by different arrays can be estimated using the sensitivity function or the Fréchet derivative. In our case, we used the median depth of investigation given by Edwards (1977), assuming 'a' spacings varying from 3–12 m and a spacing factor 'n' ranging from 1–5, thus leading to a value of approximately 17.72 m. A total of 231 measurements were made during each sequence. The inversion process was computed with the Res2dinv software package (Loke and Barker 1996), using the L_2 -norm (smoothness constrained) optimization method. Default settings of the algorithm were used during the inversion process, for all data sets. In addition, we developed a Matlab® code to determine the signal dependent rank-order mean and median filter settings. This code allows the user to select the thresholds T_i and to visualize the resulting filtering effects.

For comparison, the median filter was used with a different window size, in order to assess the performance of the signal dependent rank-order mean technique.

Our implementation of the signal dependent rank-order mean filter used a sliding window of size 9, with four threshold values (T_1, T_2, T_3, T_4). We also used sliding windows of size 3 and 5 for the median filter. The threshold values (T_i) chosen for the signal dependent rank-order mean filter were determined by extensive testing and were changed for every data set. The choice of threshold values depends on the noise level and the desired signal. Our trials showed that a sliding window of size 9 produced the best results.

We used the SNR and rms error as quantitative criteria, to evaluate the performance of each filter. It is important to note that the model with the lowest rms error is not always the best and most geologically reasonable model; this is because the solutions to inverse problems are not generally unique. The maximum number of iterations was set to 3 or 4, depending on the observed variations in rms error. We used a sliding window of size 9 (3×3), centred at $x(i,j)$. The neighbouring samples of $x(i,j)$ are rearranged into a 1D vector $w(k)$ ($k = (i-1)*l + j$ where $l=3$ and $i = j = l$). This vector is then filtered by the signal dependent rank-order mean filter. The resulting filtered vector can be used as an input to the filter, in order to improve the filtering process. The one-dimensional filtered vector is then transformed back to the 2D case and the window is translated to the next position, until it reaches the final sample.

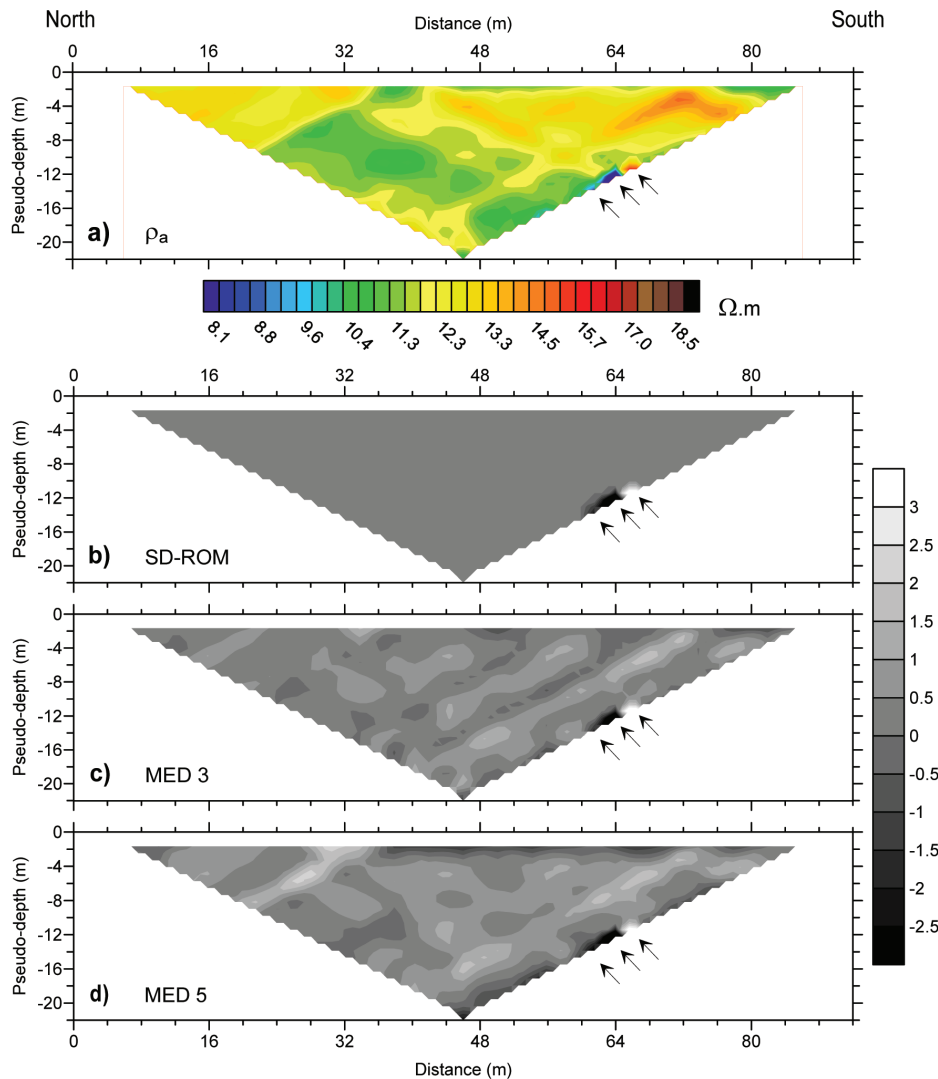


FIGURE 4 First field data set, with arrows indicating the location of data spikes to be removed: a) measured apparent resistivities. Pseudosections showing the differences between the measured apparent resistivities and those computed after: b) signal dependent rank-order mean filtering, c) median filtering (window size 3), d) median filtering (window size 5)

Field data analysis

In order to demonstrate the robustness of the signal dependent rank-order mean filter, we compared its performance with that of the standard median filter. Measured apparent (unfiltered) resistivities, taken from the first profile (PE1) shown in Fig. 4(a), were used as input to both filters. Three anomalous values were identified near to the bottom-right corner, at distances of 58 m, 62 m and 66 m respectively and are indicated by arrows in this figure. These values were identified because of their anomalous potential measurements.

To demonstrate the effectiveness of each filter, we have used the image-difference method. The ‘image difference’ is obtained by computing the difference between corresponding points on each of the two images, so that those parts that have changed can be clearly identified in the resulting difference image. The aim of this method is to enhance the difference between two images.

Figure 4(b) thus illustrates the computed difference between the unfiltered apparent resistivities and the signal dependent

rank-order mean filtered values. It can be seen that the impulses identified in Fig. 4(a) (arrows) have been perfectly located and isolated, without any undesirable effect on the signal. The number and magnitude of the impulses to be located and isolated by the signal dependent rank-order mean filter depend on the choice of threshold values and window size. In the present case, the window size was fixed to 9 and threshold values were determined following extensive tests. Figures 4(c) and 4(d) illustrate the difference pseudosections computed between median filtered data with a sliding window of size 3 and 5 respectively and unfiltered data. The impulses (arrows) have also been isolated and removed. However, difference pseudosections show distribution of values extracted from the signal considered as noise by the median filter, this indicates clearly the less efficiency of the median filter to discriminate the noise from signal as the signal dependent rank-order mean filter does with high performance. These extracted values are induced by the rank ordering process. Indeed, the median filter replaces the central sample by the

median, even if this sample is uncorrupted by noise. These extracted values, if unaccounted for, can noticeably affect the inversion process and will lead to erroneous interpretations.

A more quantitative criterion for the evaluation of the performance of different filters is the data SNR in dB. For the resistivity data studied here, the SNR (dB) was calculated using:

$$\text{SNR} = 10 \log_{10} \left(\frac{\sum_i (x(i))^2}{\sum_i (F(i) - x(i))^2} \right) \quad (6)$$

where $x(i)$ are the resistivities considered to be unaffected by noise (these were filtered manually) and $F(i)$ are the filtered outputs ($i = 1, 2, \dots, n = 231$). We also used the standard deviation (σ) of our data set to measure the spread of its values. A large standard deviation indicates that the data are often far from their mean value and a small deviation indicates that they are in

general closely clustered to the mean.

Table 1 summarizes the results, in terms of SNR and standard deviation, of the corrupted data using the signal dependent rank-order mean filter. The results achieved with the median filter are also included. This table clearly shows that the signal dependent rank-order mean filter outperforms the median filter: larger SNR values signify better filtering. The small deviation values provided by the median filter can be explained by the clipping of spikes in noisy data.

As described above, noise has a considerable influence on the inversion process and its elimination prior to the inversion process is crucial. The results produced by inverting the first data set are summarized in Fig. 5. The inversion of the section considered to be free of noise produced the result shown in Fig. 5(a), for which noise spikes were manually located and then removed. This result

TABLE 1

SNR in dB and standard deviation (σ) of noisy data filtered by the signal dependent rank-order mean filter and by the median filter, for the first field data set

First data set	Median (window size 5)	Median (window size 3)	Signal dependent rank-order mean
SNR in dB	18.27	29.35	48.07
Standard deviation (σ)	1.53	0.81	0.95

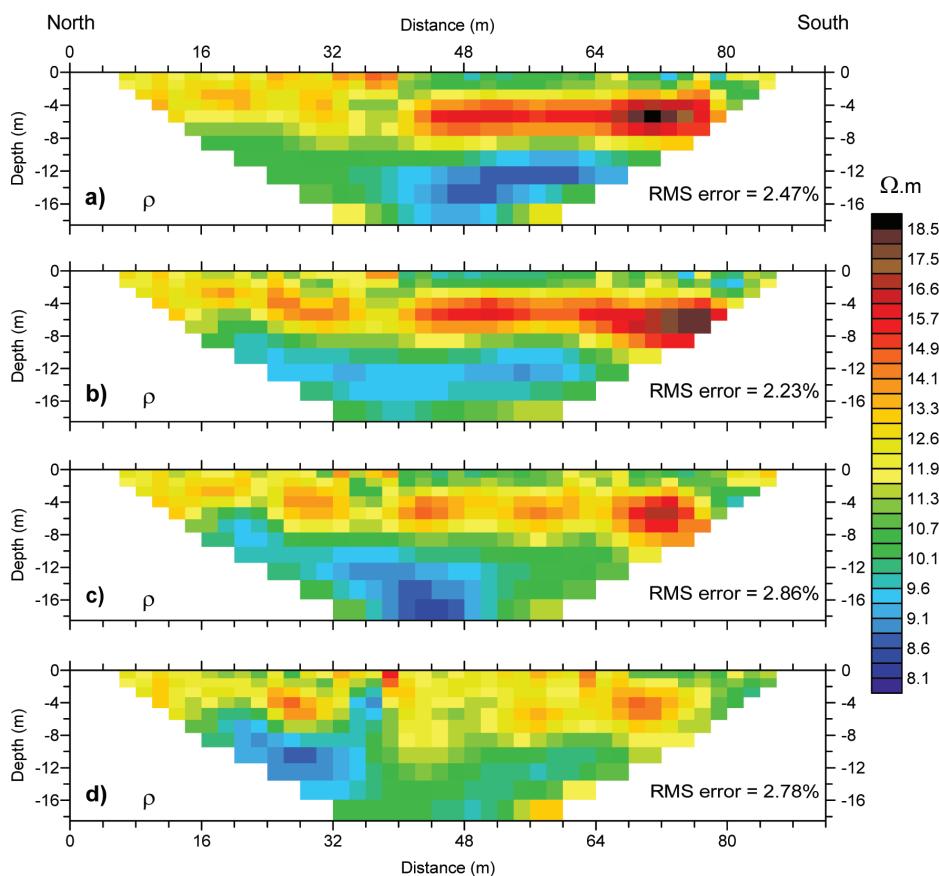


FIGURE 5

Inverted resistivity sections computed from the first field data set: a) manual filtering, b) after signal dependent rank-order mean filtering, c) after median filtering (window size 3), d) after median filtering (window size 5).

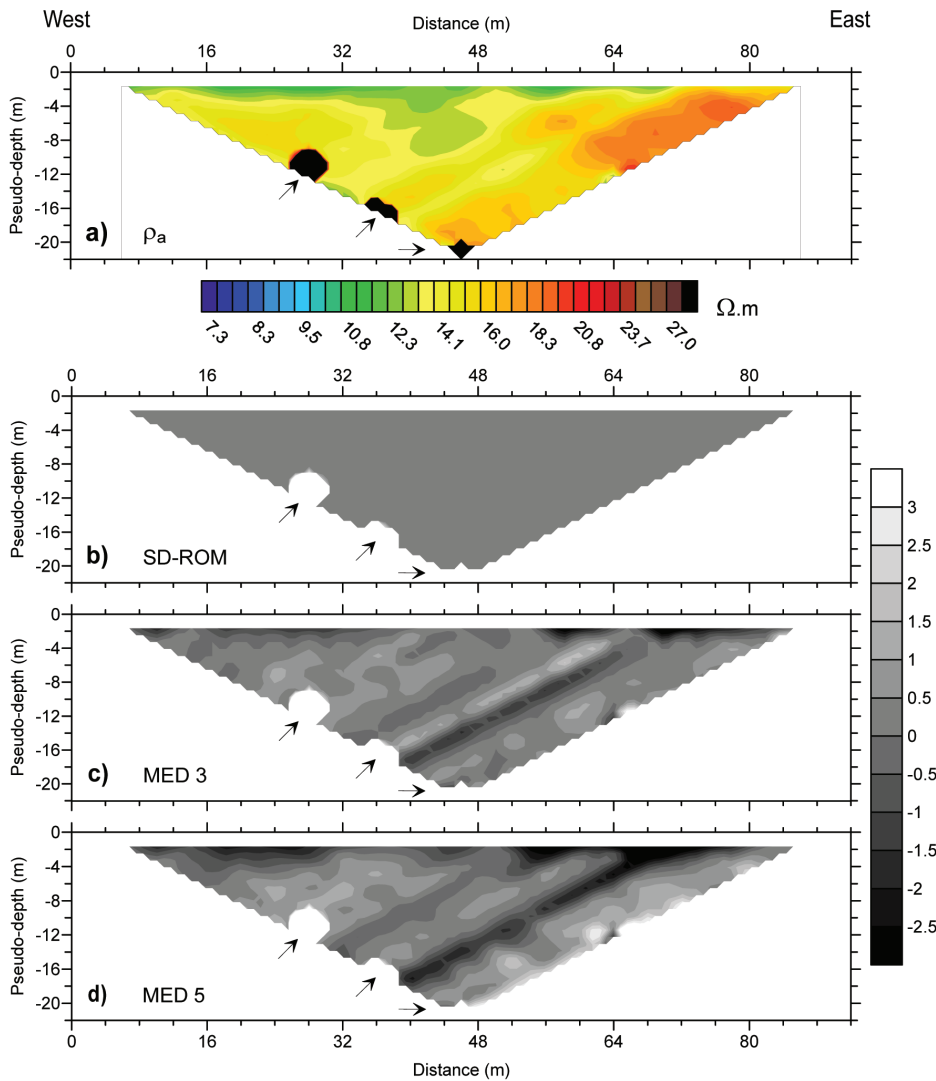


FIGURE 6 Second field data set, with arrows indicating the location of data spikes to be removed: a) measured apparent resistivities. Pseudosections showing the differences between the measured apparent resistivities and those computed after b) signal dependent rank-order mean filtering, c) median filtering (window size 3), d) median filtering (window size 5).

provides a tool that can be used as a reference, for comparison with the inverted images produced following noise removal achieved with different filters. The inverted section of Fig. 5(a) has a relatively low rms error, of about 2.47%. As expected, the signal dependent rank-order mean filter provides the best inverted section (Fig. 5b), with an rms error of 2.23%, which is very close to that of the reference section (Fig. 5a). Figure 5(b) corresponds to the inverted section, with smooth variations in the resistivity values. The geological model of the subsurface is in good agreement with the geotechnical survey data, which revealed a low resistive zone (<9 Ωm) near the surface, corresponding to clay (cold colours) and a more resistive zone at a depth of 3.8 m, just below the surface layer, corresponding to alluvial deposits or a sandy clay layer (hot colours), which approaches the surface on the northern side of the zone. The clayey substratum is defined by the low resistive zone (<9 Ωm) at a depth of 9.8 m. The results obtained using the median filter lead to an rms error of 2.86% in the case of a window of size 3 (Fig. 5c) and an rms error of 2.78% in the case of a win-

dow of size 5 (Fig. 5d). The quality of a filter depends on its ability to remove specific types of noise, whilst preserving the signal of interest and resolution. For a large window size (e.g., 5), the median filter tends to destroy the resolution of the image, as can be seen in Fig. 5(d).

Our second example, based on the second profile (PE2), corresponds to resistivity data corrupted by significant levels of noise (Fig. 6a). These anomalies (indicated by arrows) are located at distances of 28 m, 36 m and 44 m from the origin and have resistivities of: 232 Ωm, 733 Ωm and 1677 Ωm, which are very high when compared to the background resistivities. This data set was subjected to the same study procedure as the first.

It can be seen from the difference pseudosections (Fig. 6), the signal dependent rank-order mean filter helps with the detection and removal of the previously identified impulses (arrows), without changing the other parts of the pseudosection, thereby enhancing the signal of interest (Fig. 6b). These impulses are positive and are shown in white on the difference pseudosection.

Inspection of Figs 6(c) and 6(d) reveals that the median filter also isolated and removed the noise spikes (indicated by arrows), although difference pseudosections show that other parts are also significantly affected by median filters. As stated in the foregoing section, these extracted values can affect the inversion process and hence be incorrectly interpreted as geological features.

The performance figures given in Table 2 show that the signal dependent rank-order mean filter yields a better SNR and a higher standard deviation than the median filter (window size 3). The poorest results are given by the median filter with a window size of 5, due to the excessive smoothing that occurs when the window size is too large.

Figure 7 presents the inversions obtained from the second field data set. Figure 7(a) represents the 'reference' inverted section, after manual removal of the noise spikes and has an rms error of 4.06%.

When the inverted section produced by the signal dependent rank-order mean filter (Fig. 7b) is compared with the 'reference' inverted section described above (Fig. 7a), they can be seen to be highly similar, with the former being characterized by an rms error of 4.1%. Geological interpretation of the model represented by Fig. 7(b) reveals a conductive superficial layer, corresponding to clays with resistivities of less than 10 Ωm , capping a Quaternary layer of alluvial deposits at a depth of 3.8 m, with resistivities greater than 19 Ωm , followed by a weakly resistive clayey substratum, lying at a depth of 9.8 m. It is important to notice the vertical conductive structure at 38 m revealed by the inversion, which is supposed to be an extension of the clayey substratum to the surface.

Figures 7(c) and 7(d) show the results obtained, with rms errors of respectively 2.93% and 3.02%, by applying median filters of size 3 and 5. As described above, the lowest rms error

TABLE 2

SNR in dB and standard deviation (σ) of noisy data filtered by the signal dependent rank-order mean filter and by the median filter, for the second field data set

Second data set	Median (window size 5)	Median (window size 3)	Signal dependent rank-order mean
SNR in dB	15.82	24.44	49.54
Standard deviation (σ)	2.32	1.77	1.96

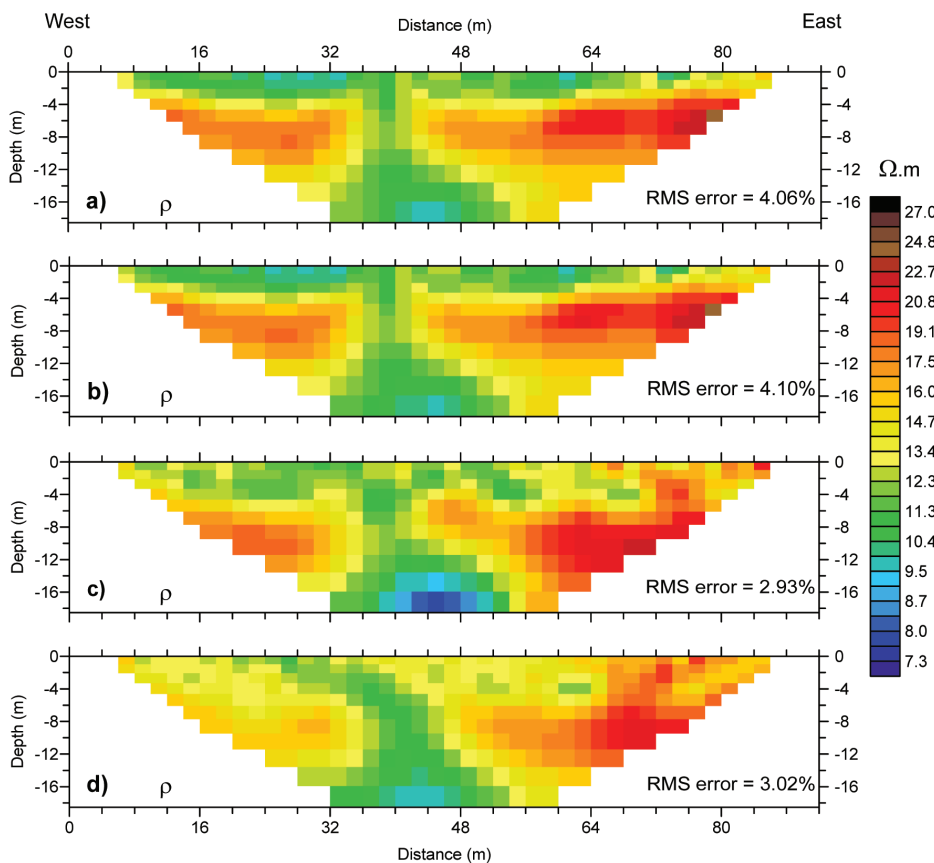


FIGURE 7

Inverted resistivity sections computed from the second field data set: a) manual filtering, b) after signal dependent rank-order mean filtering, c) after median filtering (window size 3), d) after median filtering (window size 5).

is not always synonymous with the best geological model. In the case of the present examples, although the median filter produces an rms error lower than that of the signal dependent rank-order mean filter, a comparison between the inverted, median-filtered sections (Figs 7c and 7d) and the inverted reference section (Fig. 7a) reveals notable differences, particularly at the surface. Nevertheless, the same geological interpretation is found as for the data filtered by the signal dependent rank-order mean filter. However, the vertical conductive structure revealed by the inversion on Figs 7(a) and 7(b) has not been properly located on inverted pseudosections corresponding to median filtered data. In this example again, the signal dependent rank-order mean filter is found to provide the most accurate sub surface model.

CONCLUSION

Near-surface inhomogeneities, as well as spiky random noise, generate anomalies in recorded resistivity data, which can adversely affect the result of the inversion process.

In this paper, we have successfully applied the signal dependent rank-order mean technique to the removal, from synthetic and 2D field electrical resistivity data, of noise induced by close to surface small geological structures (near-surface inhomogeneities) as well as spiky noise. The signal dependent rank-order mean filter belongs to the family of order-statistical or decision based filters; the major advantage of this type of filter is its computational simplicity and its ability to distinguish noise from a set of real data. It has been shown that application of the signal dependent rank-order mean filter to the measured data leads to significant improvements in the inversion process. The results obtained are very encouraging and the efficiency of the signal dependent rank-order mean filter was verified by comparing it to the standard median filter. Our trials show that a window of size 9 for the signal dependent rank-order mean filter produces optimal results and that further improvements are achieved in the inversion when it is used recursively. We have also demonstrated improved filtering in terms of SNR, in addition to the improvements in the inversion process, expressed in terms of an rms error. We found that data filtered by signal dependent rank-order mean filter produces smoothed apparent resistivity values.

Further research is needed in order to investigate the role of 3D inhomogeneities on resistivity measurements and the inversion process. In addition, we propose to explore new methods, in particular evolutionary neural fuzzy techniques, for the choice of noise-identification thresholds.

REFERENCES

- Abreu E., Lightstone M., Mitra S.K. and Arakawa K. 1996. A new efficient approach for the removal of impulse noise from highly corrupted images. *IEEE Transactions on Image Processing* **5**, 1012–1025.
- Bernstone C., Dahlin T., Ohlsson T. and Hogland W. 2000. DC-resistivity mapping of internal landfill structures: Two pre-excavation surveys. *Environmental Geology* **39**, 360–371.
- Bobachev A.A., Marchenko M.N., Modin I.N., Pervago E.V., Urusova A.V. and Shevnin V.A. 1996. New approaches to electrical soundings of horizontally inhomogeneous media. *Physics of the Solid Earth* **31**, 1075–1086.
- Chambers J.E., Kuras O., Meldrum P.I., Ogilvy R.D. and Hollands J. 2006. Electrical resistivity tomography applied to geologic, hydrogeologic and engineering investigations at a former waste-disposal site. *Geophysics* **71**, B231–B239.
- Chambers J.E., Meldrum P.I., Ogilvy R.D. and Wilkinson P.B. 2005. Characterisation of a NAPL-contaminated former quarry site using electrical impedance tomography. *Near Surface Geophysics* **3**, 79–90.
- Chandra C., Moore M.S. and Mitra S.K. 1998. An efficient method for the removal of impulse noise from speech and audio signals. *Proceedings of the IEEE Symposium on Circuits and Systems* **4**, 206–208.
- Cornacchiulo D. and Bagtzoglou A.C. 2004. Geostatistical reconstruction of gaps in near-surface electrical resistivity data. *Vadoze Zone Journal* **3**, 1215–1229.
- Dahlin T. 1996. 2D resistivity surveying for environmental and engineering applications. *First Break* **14**, 275–283.
- Dahlin T. and Zhou B. 2004. A numerical comparison of 2D resistivity imagings with 10 electrode arrays. *Geophysical Prospecting* **52**, 379–398.
- Daily W., Ramirez A. and Johnson R. 1998. Electrical impedance tomography of a perchloroethylene release. *Journal of Environmental and Engineering Geophysics* **2**, 189–201.
- Dey A. and Morrison H.F. 1979. Resistivity modelling for arbitrarily shaped three-dimensional structures. *Geophysics* **44**, 753–780.
- Edwards L.S. 1977. A modified pseudosection for resistivity and induced-polarization. *Geophysics* **42**, 1020–1036.
- Gersztenkorn A. and Scales J.A. 1988. Smoothing seismic tomograms with alpha-trimmed means. *Geophysical Journal* **92**, 67–72.
- Gharibi M. and Bentley L.R. 2005. Resolution of 3-D electrical resistivity images from inversions of 2-D orthogonal lines. *Journal of Environmental and Engineering Geophysics* **10**, 339–349.
- Gonzalez R.C. and Woods R.E. 1992. *Digital Image Processing*, 2nd edn. Addison Wesley.
- Grellier S., Guérin R., Robain H., Bobachev A., Vermeersch F. and Tabbagh A. 2008. Monitoring of leachate recirculation in a bioreactor landfill by 2D electrical resistivity imaging. *Journal of Environmental and Engineering Geophysics* **13**, 351–359.
- Griffiths D.H. and Barker R.D. 1993. Two-dimensional resistivity imaging and modelling in areas of complex geology. *Journal of Applied Geophysics* **29**, 211–226.
- deGroot-Hedlin C. and Constable S. 1990. Occam's inversion to generate smooth, two-dimensional models from magnetotelluric data. *Geophysics* **55**, 1613–1624.
- Guérin R., Begassat P., Benderitter Y., David J., Tabbagh A. and Thiry M. 2004. Geophysical study of the industrial waste land in Mortagne-du-Nord (France) using electrical resistivity. *Near Surface Geophysics* **2**, 137–143.
- Henderson R.P. and Webster J.G. 1978. An impedance camera for spatially specific measurements of thorax. *IEEE Transaction on Biomedical Engineering* **25**, 250–254.
- LaBrecque D.J., Miletto M., Daily W., Ramirez A. and Owen E. 1996. The effect of noise on Occam's inversion of resistivity tomography data. *Geophysics* **61**, 538–548.
- Loke M.H., Acworth I. and Dahlin T. 2003. A comparison of smooth and blocky inversion methods in 2D electrical imaging surveys. *Exploration geophysics* **34**, 182–187.
- Loke H. and Barker R.D. 1996. Rapid least-squares inversion of apparent resistivity pseudosections by a quasi Newton method. *Geophysical Prospecting* **44**, 131–152.

- Lytle R.J. and Dines K.A. 1978. *An impedance camera: A system for determining the spatial variation of electrical conductivity*. California University, Livermore (USA). Lawrence Livermore Lab. Technical report. Lawrence Livermore National Laboratory UCRL-52413.
- Manglem K.S., Singh P.K.B. and Singh S.B. 2002. Impulse noise removal using fuzzy rank selection technique. *Second International Conference on Electrical and Computer Engineering*, Dhaka, Bangladesh, 140–143.
- Moore M.S., Gabbouj M. and Mitra S.K. 1999. Vector SD-ROM filter for removal of impulse noise from colour images. *EURASIP Conference DSP for Multimedia Communications and Services*, Krakow, Poland.
- Moore M.S. and Mitra S.K. 2000. Statistical threshold design for the two-state signal dependent rank order mean filter. *Proceedings of the International Conference on Image Processing* **1**, 904–907.
- Pelton W.H., Rijo L. and Swift C.M.Jr. 1978. Inversion of two-dimensional resistivity and induced polarization data. *Geophysics* **43**, 788–803.
- Ritz M., Robain H., Pervago E., Albouy Y., Camerlynck C., Descloitres M. *et al.* 1999. Improvement to resistivity pseudosection modelling by removal of near-surface inhomogeneity effects: application to a soil system in south Cameroon. *Geophysical Prospecting* **47**, 85–101.
- Sasaki Y. 1989. Two-dimensional joint inversion of magnetotelluric and dipole-dipole resistivity data. *Geophysics* **54**, 254–262.
- Sasaki Y. 1992. Resolution of resistivity tomography inferred from numerical simulation. *Geophysical Prospecting* **40**, 453–464.
- Wilkinson P.B., Meldrum P.I., Chambers J.E., Kuras O. and Ogilvy R.D. 2006a. Improved strategies for the automatic selection of optimised sets of electrical resistivity tomography measurement configurations. *International Geophysical Journal* **167**, 1119–1126.
- Wilkinson P.B., Kuras O., Meldrum P.I., Chambers J.E. and Ogilvy R.D. 2006b. Comparison of the spatial resolution of standard and optimised electrical resistivity tomography arrays. *Near Surface 2006*, Helsinki, Finland.
- Zhou B. and Dahlin T. 2003. Properties and effects of measurement errors on 2D resistivity imaging surveying. *Near Surface Geophysics* **1**, 105–117.

ABEM France

www.abemfrance.eu



geophysical equipment & software

RENTAL - SALES - SERVICE - TRAINING



ABEM France
offers a large range
of geophysical equipment
and interpretation
software.

Contact us :
Email info@abemfrance.eu
Phone +33.2.99.85.16.54




Photo: Peter Jansson, Stockholm University

Enjoy the use of GPR, experience the MALÅ GPR Systems

With the best technology available built on Swedish quality products, the userfriendly MALÅ GPR Systems are offering durability, flexibility and various alternatives to be applied to at a reasonable cost.

Experience the MALÅ GPR Systems for your GPR surveys!


Contact us now for further information or visit our website www.malags.com

Head Office
MALÅ Geoscience AB
Skolgatan 11, SE-930 70
Malå Sweden
Phone: +46 953 345 50,
Fax: +46 953 345 67
E-mail: sales@malags.com

Sales Offices

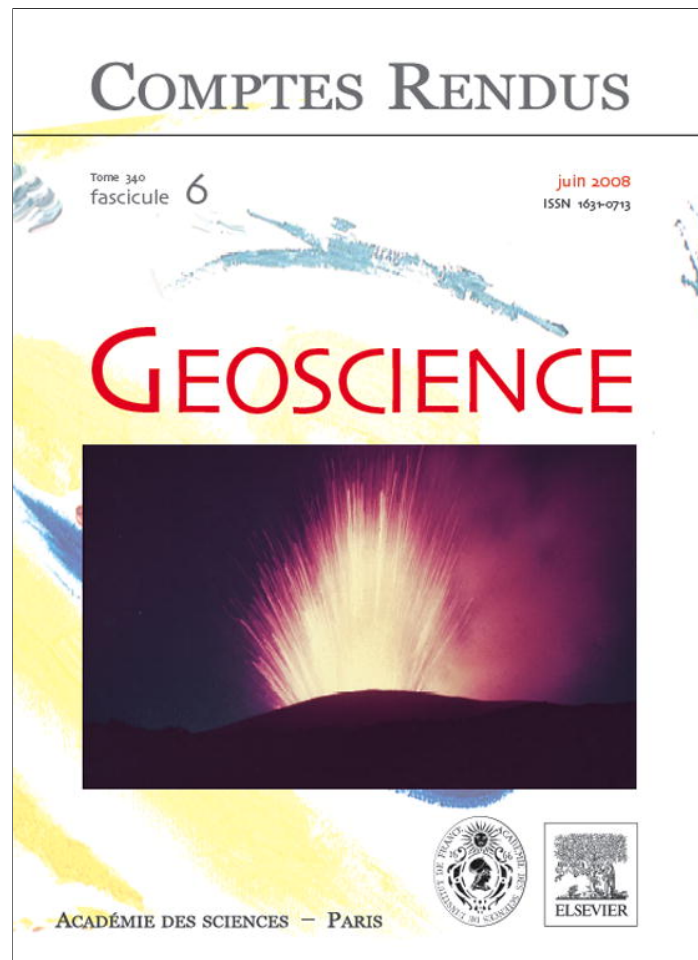
USA: MALÅ Geoscience USA, Inc., 2040 Savage Rd. Box 80430, Charleston, SC 29416
Phone: +1 843 852 5021, Fax: +1 843 769 7397, E-mail: sales.usa@malags.com

China: MALÅ Geoscience (China), Room 2604, Yuan Chen Xin BLDG, No.12 Yu Min Road Chao Yang District, Beijing 100029
Phone: +86 10 8225 0728, Fax: +86 10 8225 0815, E-mail: sales@malags.com



www.malags.com

Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



ELSEVIER

Available online at www.sciencedirect.com

C. R. Geoscience 340 (2008) 335–344

<http://france.elsevier.com/direct/CRAS2A/>

Internal Geophysics
**Application of feedback connection artificial neural
 network to seismic data filtering**

Noureddine Djarfour^a, Tahar Aïfa^{b,*}, Kamel Baddari^a,
 Abdelhafid Mihoubi^a, Jalal Ferahtia^a

^aLaboratoire de physique de la Terre (LABOPHYT), université M'hamed-Bougara, 35000 Boumerdès, Algeria

^bGéosciences-Rennes, CNRS UMR6118, université Rennes-1, Campus de Beaulieu, 35042 Rennes cedex, France

Received 22 November 2007; accepted after revision 10 March 2008

Available online 3 June 2008

Presented by Jean-Louis Le Mouël

Abstract

The Elman artificial neural network (ANN) (feedback connection) was used for seismic data filtering. The recurrent connection that characterizes this network offers the advantage of storing values from the previous time step to be used in the current time step. The proposed structure has the advantage of training simplicity by a back-propagation algorithm (steepest descent). Several trials were addressed on synthetic (with 10% and 50% of random and Gaussian noise) and real seismic data using respectively 10 to 30 neurons and a minimum of 60 neurons in the hidden layer. Both an iteration number up to 4000 and arrest criteria were used to obtain satisfactory performances. Application of such networks on real data shows that the filtered seismic section was efficient. Adequate cross-validation test is done to ensure the performance of network on new data sets. **To cite this article: N. Djarfour et al., C. R. Geoscience 340 (2008).**

© 2008 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

Résumé

Application du réseau de neurones artificiel récurrent au filtrage des données sismiques. Le réseau de neurones artificiel (RNA) de type Elman (rétroconnexion) a été utilisé pour filtrer des données sismiques. La boucle de rétroaction qui caractérise ce réseau présente l'avantage de stocker la trace de l'événement précédent pour être utilisée dans le traitement courant. Cette récurrence facilite l'apprentissage de la structure neuronale proposée par rétropropagation (méthode des gradients). Des essais ont été réalisés sur des données sismiques synthétiques (avec 10 % et 50 % de bruit aléatoire et gaussien) et réelles, en utilisant respectivement 10 à 30 neurones et un minimum de 60 neurones dans la couche cachée. Un nombre d'itération fixé à 4000 et un critère d'arrêt ont été utilisés pour satisfaire de meilleures performances. L'application de tels réseaux sur des données réelles montre que le filtrage de la section sismique est efficace. Le test de validation croisée est réalisé pour assurer la performance du réseau sur de nouvelles données. **Pour citer cet article : N. Djarfour et al., C. R. Geoscience 340 (2008).**

© 2008 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

Keywords: Elman's ANN; Gaussian and Random noise; Filtering; Training; Back-propagation; Seismic

Mots clés : RNA d'Elman ; Bruit gaussien et aléatoire ; Filtrage ; Apprentissage ; Rétropropagation ; Sismique

* Corresponding author.

E-mail address: tahar.aifa@univ-rennes1.fr (T. Aïfa).

Version française abrégée

Introduction

Dans la séquence de traitement des données sismiques, la reconnaissance et l'atténuation du bruit est une étape importante. Les méthodes conventionnelles utilisent les propriétés physiques qui différencient le signal du bruit, telles que vitesse apparente, nombre d'onde, temps de trajet... par l'application d'algorithmes de filtrage ou de prédiction du bruit par modélisation [1,8,9,13]. Le bruit prédit sera ainsi soustrait des données sismiques. Cependant, de telles méthodes correspondent à une solution dans laquelle le signal désiré est de forme connue. Par conséquent, leurs applications supposent qu'on dispose simultanément de données *a priori* sur le signal et le bruit, ce qui n'est pas toujours le cas. Dans la séquence de traitement des données sismiques, la reconnaissance et l'élimination du bruit par les différentes techniques de filtrage restent un défi majeur. Malgré les progrès significatifs effectués au cours de ces dernières années, les performances des filtres proposés sont encore loin d'éliminer totalement le bruit, notamment dans les conditions d'acquisition des données sur des structures géologiques complexes.

Dans ce travail, nous proposons d'utiliser le réseau de neurones artificiel (RNA) pour filtrer les données sismiques. L'utilisation du RNA est motivée par sa capacité à extraire des informations utiles à partir de données hétérogènes ou imprécises. De plus, le modèle RNA permet d'approximer des relations arbitraires complexes non linéaires et d'obtenir une fonction de transfert. L'accent sera mis sur l'aptitude du RNA à l'apprentissage et à la généralisation sur des données synthétiques et réelles de sismiques réflexions.

RNA et ses applications

Un réseau de neurones artificiel (RNA) est constitué de neurones élémentaires (« agents »), connectés entre eux par l'intermédiaire de poids jouant le rôle de synapses [11]. 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 (Fig. 1a). Le concept du neurone est fondé sur la sommation pondérée du potentiel d'action qui lui parvient des neurones voisins ou du milieu extérieur (Fig. 1b). Le neurone s'active suivant la valeur de cette sommation pondérée. Si cette somme dépasse un seuil, le neurone est activé, puis transmet une réponse (sous forme de potentiel d'action) dont la valeur est celle de son activation. Si le neurone n'est pas activé, il ne

transmet rien [3,4]. Un RNA est généralement constitué de trois couches. La première couche représente l'entrée ; la seconde, dite couche cachée, constitue le cœur du réseau de neurones ; la troisième représente la couche de sortie (Fig. 1a). Un réseau de neurones fonctionne en deux temps : (1) il détermine d'abord ses paramètres suivant un algorithme d'apprentissage (conception) qui fait appel généralement à la rétropropagation du gradient conjugué [10,12], (2) puis il est appliqué comme fonction de transfert.

Applications

Nous avons d'abord essayé d'entraîner un réseau de neurones de type « connexion directe » (*feed-forward*), sans succès par manque de convergence de la rétropropagation. Calderi3n-Macías et al. [2] ont mentionné que le RNA de type *feed-forward*, utilisant la rétropropagation seule, ne peut converger ; ils ont adopté un apprentissage hybride par la technique du recuit simulé pour entraîner ce genre de réseau. Dans ce travail, nous avons choisi un RNA de type Elman (Fig. 1b) (connexion récurrente) pour le filtrage des données sismiques. Ce type de RNA possède une couche spéciale appelée « couche de contexte » [5], qui copie l'activité des neurones de la couche cachée ; les sorties des neurones de la « couche de contexte » sont utilisées en entrée de la couche cachée. Il y a alors conservation de la trace des activités internes du réseau sur un pas de temps. L'intérêt de ces modèles réside dans leur capacité à apprendre des relations complexes à partir de données numériques [11]. Les fonctions d'activation de la couche cachée de la structure proposée sont de type sigmoïde, permettant au RNA de se comporter comme un réseau non linéaire dans la transformation qu'il réalise. L'algorithme de rétropropagation du gradient conjugué est utilisé pour entraîner la structure neuronale proposée [2,7].

Données synthétiques

Étant donné que l'objectif de cette phase de simulation est le filtrage des données sismiques des bruits aléatoires et gaussiens par RNA, un modèle de subsurface de six couches à stratification horizontale a été utilisé (Tableau 1). Le sismogramme utilisé pour l'apprentissage du RNA ne simule que l'onde P, calculée à incidence normale. Sur la base de ce modèle, une section sismique a été calculée par convolution avec une impulsion de Ricker de fréquence centrale 20 Hz ; les trajets sont calculés suivant le principe de Snell–Descartes [6].

La section sismique utilisée comme sortie désirée du RNA dans la phase d'apprentissage est donnée sur la

(Fig. 2a). Remarquons que les simulations ont été réalisées sous Matlab, qui n'accepte pas d'indice négatif ou nul pour les distributions de vitesses, engendrant ainsi un décalage temporel pour la première couche qui démarre à 0 ms. Les graphes de la (Fig. 2b et d), qui représentent l'entrée du RNA, sont obtenus respectivement par l'addition de 50 % de bruit aléatoire et gaussien à la section sismique précédente (Fig. 2a). Nous avons réalisé des tests en présence de 10 % de bruit aléatoire et gaussien (Tableau 2a et b), mais nous ne présentons ici que les résultats venant de l'addition de 50 % de bruit. Le nombre d'itération a été fixé à 4000, avec un critère d'arrêt (Tableau 2a). La section sismique filtrée par RNA après apprentissage respectivement pour le bruit aléatoire (Fig. 2c) et gaussien (Fig. 2e) est compatible avec le modèle de sortie désiré (Fig. 2a). Mis à part quelques traces dans lesquelles le bruit gaussien (Fig. 2e) a été réduit, le filtrage par RNA est plus efficace dans le cas d'un bruit aléatoire. Dans un exemple de performance obtenue (Fig. 2f) pendant la phase d'apprentissage avec 10 neurones (Tableau 2a et b) dans la couche cachée, l'erreur correspond à l'erreur cumulée, calculée par comparaison de la sortie RNA avec la sortie désirée. La performance avec 4000 itérations prédit une erreur croissante quand le nombre de neurones augmente. On remarque que la convergence est plus rapide lorsqu'on a affaire à du bruit aléatoire (Fig. 2f). Si l'on compare les Fig. 2c et e, la qualité du filtrage est meilleure dans le cas du bruit aléatoire.

Données réelles

Deux types de données ont été préparés par filtrage dans le domaine spectral, sous forme de couples constitués d'entrée (traces d'un point de tir) et de sortie désirée (traces filtrées), l'un pour réaliser l'apprentissage et l'autre pour tester le réseau obtenu et déterminer ses performances. Les données d'apprentissage (Fig. 3a), composées d'une collection de 48 traces à 2000 échantillons par trace, traitées en amplitudes égalisées, contiennent plusieurs types de bruits (bruits aléatoires, arrivées directes, réfractées, des ondes de surface...), qui ont été filtrés dans le domaine ($f-x$) (Fig. 3b), pour fournir au RNA un exemple représentatif de bruits qu'il doit apprendre à identifier et à atténuer.

Plusieurs essais d'apprentissage ont été réalisés (Tableau 2c). Les performances obtenues montrent la simplicité d'apprentissage de ce type de RNA par l'algorithme de rétropropagation du gradient conjugué et sont donc satisfaisants (cf. Fig. 3c).

Deux tests ont ensuite été réalisés. Le premier, qui est un test d'apprentissage, a été accompli en utilisant les

données d'apprentissage (Fig. 3d) pour constituer la sortie du réseau après apprentissage, en lui présentant les traces bruitées utilisées dans cette phase. En comparant la sortie du réseau après apprentissage (Fig. 3d) avec les données bruitées (Fig. 3a), on remarque la reproductibilité de la sortie désirée (Fig. 3b) par l'atténuation du bruit. Le second est un test de généralisation. Dans ce cas, le deuxième modèle de données (Fig. 4a) constitue l'entrée du RNA entraîné, tandis que les données filtrées dans le domaine ($f-x$) (Fig. 4b) sont la sortie désirée, que l'on comparera à la sortie du RNA. Nous avons obtenu une section sismique (Fig. 4c), où l'on observe l'atténuation du bruit si on la compare à la section filtrée donnée dans la Fig. 4b, qui est le résultat attendu. Des horizons peuvent être isolés dans la partie supérieure de la section (< 1200 ms), où le filtrage était effectif : réduction des zones énergétiques (< 400 ms), atténuation des ondes de surface, des arrivées directes et réfractées à différents niveaux.

Conclusion

L'apprentissage du RNA constitue un moyen de synthétiser automatiquement une fonction de transfert non linéaire entre les données bruitées et les données filtrées, qui permet au RNA de se comporter comme un filtre. Le RNA de type Elman offre l'avantage de simplicité d'apprentissage par la rétro-propagation du gradient conjugué.

La parfaite concordance entre la sortie du RNA en phase de généralisation et la section filtrée, résultat attendu, témoigne de la réussite de la structure neuronale proposée en phase de généralisation. Ces résultats obtenus démontrent clairement la faisabilité de la méthode, ainsi que l'intérêt d'exploiter la capacité de filtrage des données sismiques offerte par l'utilisation de l'approche neuronale.

1. Introduction

The filtering of noise is one of the crucial steps in a data processing sequence, especially in the light of the complexity of observations and the interference of noise types. Noise occurs not only during the initial phase of data acquisition from different sources (e.g., instrumental, natural) and processing related to techniques of sampling and methodology of filtering, but can also be visible in the final document. Methods used to attenuate the noise operate either by exploiting physical properties that differentiate noise from signal, such as apparent velocity, wave number, travel time... by the use of filtering algorithms, or by predicting the noise through

modelling techniques [8,9,13]. Such predicted noise will be subtracted from the recorded seismic data. However, such methods correspond to a solution when the desired signal is of a known shape. Consequently, their application simultaneously assumes *a priori* information on signal and noise. Such information can be a minimum phase analytical signal (e.g., Ricker) that we usually use for seismic for instance.

In spite of significant progress regarding the performance of the proposed filters, we are still far from reducing the noise completely either from synthetic or real data. So far there is no noise attenuation technique that works universally in all acquisition data. Therefore, recognizing and removing the noise remains one of the great challenges in signal processing. The method proposed in this paper is an attempt to filter the noise by means of the artificial neural network (ANN). The use of ANN is motivated by its remarkable ability to extract useful information from heterogeneous or inaccurate data. Moreover, the ANN can approximate arbitrary complicated non-linear relationships and get a good non-linear transfer function. The emphasis is placed on the capacity of the ANN for training and generalization.

We will first address the ANN by means of several tests performed on synthetic and real seismic shot data, then make an adequate cross-validation and compare the quality of filtering at different stages of training and generalization. Filtering by means of Elman's ANN has not yet been applied to the seismic processing sequence.

2. Artificial neural network principles and its application

An ANN is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. The most basic element of the human brain is a specific type of cell, which provides us with the ability to remember, think, and apply previous experiences to our every action [11]. These cells are known as neurons. The power of the brain comes from the number of these basic components and the multiple connections between them. The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems [4]. Each neuron is linked to certain of its neighbours with varying coefficients of connectivity that represent the strengths of these connections [3]. Each neuron transforms all signals that it receives into an output signal, which is communicated to other neurons. For example, the artificial neuron i multiplies each input signal by the

connection coefficient w_i , and adds up all these weighted entries in order to obtain a total simulation. Using an activation function f (or transfer function), it calculates its activity at the output, which is communicated to the following neurons (Fig. 1a).

Neurons are grouped into layers. In a multilayer network, there are usually an input layer, one or more hidden layers and an output layer (Fig. 1b). The layer that receives the inputs is called the input layer. It typically performs no function on the input signal. The network outputs are generated from the output layer. Any other layers are called hidden layers because they are internal to the network and have no direct contact with the external environment. The 'topology' or structure of a network defines how the neurons in different layers are connected. The choice of the topology to be used depends closely on the properties and the requirements of the application.

2.1. Artificial neural network's training

The most widely used training method is known as back-propagation method. This latter produces a least-square fit between the actual network output and desired results by computing a local gradient in terms of the network weights [12]. The design of neural network models capable of training originates in the work of the

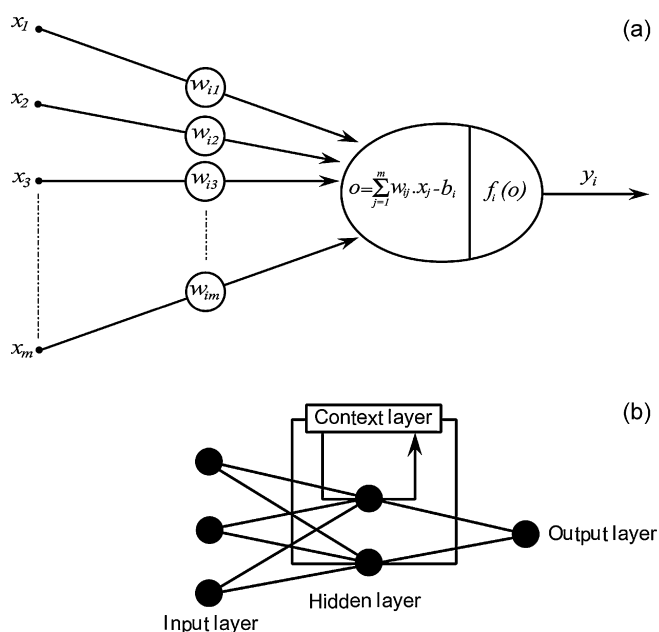


Fig. 1. Architecture of an artificial neuron and a multilayered feed-back neural network. (a) Artificial neuron. (b) Multilayered artificial neural network.

Fig. 1. Configuration d'un neurone artificiel et un réseau de neurones artificiel à connexion récurrente. (a) Neurone artificiel. (b) Réseau de neurones artificiels multicouches.

neurophysiologist Hebb [11]. Let us start from the simplest and best-known of facts: the neurons are connected to each other by synapses. Hebb made the assumption that the force of a synapse increases when the neurons it connects act in the same way at the same time. Conversely, it decreases when the neurons have different activities at the same time. Consequently, the intensity of this force varies, more or less, according to the simultaneous activity of each interconnected neuron. A learned form will correspond to a state of the network in which certain neurons will be active while others will not. It is the same principle which is applied to artificial neuron training. All the neurons that are connected to each other constitute a network. When one presents to the network a form to be learned, the neurons simultaneously enter into a state of activity that causes a slight modification of the synaptic forces. What follows is a quantitative reconfiguration of the whole of the synapses, some of them becoming very strong (high value of synaptic force), and the others becoming weak. The learned pattern is not directly memorized at an accurate location. It corresponds to a particular energy state of the network, a particular configuration of the activity of each neuron, in a very large range of possible configurations. This configuration is supported by the values of the synaptic forces [10,12].

Let y_j^s represent the output of the j^{th} neuron at layer s . w_{ij}^s is the weight connecting the i^{th} neuron in layer $s-1$ to the j^{th} neuron at layer s and o_j^s is the weighted sum at the input of the j^{th} neuron of layer s . This output can be expressed by the following equation [7]:

$$y_j^s = f(o_j^s) = f\left(\sum_{i=1}^N w_{ji}^s \cdot y_i^{s-1} - b_j\right) \quad (1)$$

where N is the number of neurons in layer $s-1$. This relation allows, by knowing the input of the first layer of the network, to calculate gradually the value of the global output of the network, thus ensuring forward propagation. When one compares this output with the desired output [7], one can calculate the error function, given by:

$$e = \frac{1}{2} \sum_{k=1}^M (y_k - \bar{y}_k)^2 \quad (2)$$

where y_k is the desired output, \bar{y}_k the obtained output of the ANN and M is the number of neurons in the output layer. In the back-propagation method, the direction in which weights w_{ij}^s are updated is given by the opposite gradient of e with respect to every element of the weight. The weights update starts in the output layer and processing will be spread towards the first layer, from where term ‘back-propagation’ comes. The back-propagation

algorithm consists in minimizing e . In our case, we used the gradient steepest descent [3] to minimize e .

With each example, one has to modify the connection weights so that we may drastically reduce the value of e . The modification of the connection weights is carried out using the following relation:

$$\Delta w_{ij}^s = -\mu \cdot (e_j^s \cdot y_i^{s-1})_n + (\Delta w_{ji}^s)_{n-1} \quad (3)$$

where μ is known as the learning rate parameter to speed the ANN training and is usually a small number [2], say between 0 and 0.5. The amount e_j^s is the local error of j^{th} neuron in layer s , defined as:

$$e_j^s = \bar{f}'(o_j^s) \sum_{k=1}^N e_k^{s+1} w_{kj}^{s+1} \quad (4)$$

In these equations, \bar{f}' is the derivative of f .

Weights and threshold terms are first initialized to random values [2]. In general, there are no strict rules for determining the network configuration for optimum training and prediction.

2.2. Seismic data filtering

One can model the seismic signal by means of the following equation [6,8,9]:

$$T(t) = W(t) * R(t) + B(t) \quad (5)$$

where (*) represents the convolution product, $T(t)$ the actually measured signal, $W(t)$ the emitted signal, $R(t)$ the reflectivity set (the effective signal to be evaluated) and $B(t)$ the seismic noise. Generally, one distinguishes the natural noise, which exists apart from all seismic activities and the artificial noise, produced by the recording or the processing of seismic data. These two types of noise can be coherent or random. All the matter of filtering consists of starting from a measured signal $T(t)$, to obtain a filtered function that estimates as adequately as possible the convolution product $W(t) * R(t)$. Many efforts in signal processing were directed towards the definition of powerful filtering tools, which must be flexible and able to adapt with the variation of the noise type. In fact, very encouraging results are available today for the filtering and the improvement of the signal-to-noise ratio. These results are obtained either by temporal or frequency filtering. Consequently, several filters are proposed, which can be mainly the adapted filter of Wiener, Kalman types, etc. [1,8,9,13]. All the conventional techniques for noise attenuation fundamentally require *a priori* information on characteristics of signal and noise. This information is usually not well known and the parameters that

characterize the signal and noise can often only be estimated inaccurately. Here we make an attempt to overcome some of these limitations by designing a network that can learn characteristics from signal and noise during processing in the training phase. In fact, by the use of ANN technique, there is no *a priori* information on signal and noise characteristics.

3. Applications

Because a three-layered neural network could approximate any arbitrary complex non-linear relationship, we unsuccessfully tried to train feed-forward ANN; this failure was due to the non-convergence of the back-propagation. Calderi3n-Mac3as et al. [2] mentioned that feed-forward ANN using back-propagation alone could not converge; they used hybrid training by very fast simulated annealing technique to train this kind of network. In the present study, we adopted a three-layered feedback architecture of neural network (Fig. 1b), which is of Elman network type. It is one hidden layer with feedback connection from the hidden layer output to its input that constitutes the context representation layer. This feedback loop allows Elman networks learning to recognize and generate temporal or spatial patterns. This internal looping ensures a recirculation of information inside the hidden layer of this Elman ANN. In fact, the delay in this connection stores values from the previous time step, which can be used in the current time step [5]. In networks of this type, the hidden layer is activated simultaneously by the input layer and the context layer. The result of the processing of this hidden layer at its output is communicated to output of Elman ANN and will be stored in the context layer. By analogy with human brain thinking, this delayed information leads to hesitations before the making of a final decision by a human brain.

The activation functions of the hidden layer are of sigmoid type, given by the following equation [2,3]:

$$f(x) = \frac{1 - e^{-2\alpha x}}{1 + e^{+2\alpha x}} \quad (6)$$

where α is a parameter that controls the steepness of the function near $x = 0$.

This function can produce outputs with a reasonable discriminating power and is a differentiable function that is essential for the back-propagation of errors. The number of neurons in the hidden layer depends on the desired optimum performance, which could be selected on a trial and error basis.

All further simulations and applications on real data were performed under the Matlab environment.

3.1. Synthetic data set

As an application of the method described above, we designed a synthetic subsurface model that is 500 m wide and 500 m deep. The main goal of this phase of simulation is to train the Elman ANN to filter seismic data from random and Gaussian noise in which subsurface consists of six homogeneous horizontal layers. A linearly increasing velocity structure is given with interval velocities varying with depth from 1200 m s^{-1} to 4000 m s^{-1} (Table 1). The synthetic seismograms used for network training include only P-wave primary reflections computed at normal incidence. Using this model, a set of trace gathers was computed using convolution equation (5), for which ray paths are computed by the Snell–Descartes' law. The convolved seismic wavelet corresponds to a Ricker wavelet with a central frequency of 20 Hz. The synthetic seismic section used as the desired output of the ANN in the training phase is shown in (Fig. 2a). We may notice that Matlab does not agree with index lower or equal to zero for the velocity distributions, generating thus a temporal shift for the first layer beginning with 0 ms. In the input of the ANN (Fig. 2b), the previous seismic section (Fig. 2a) was mixed with 50% of random noise. We also did tests on synthetic data with additional 10% of random and Gaussian noise (Table 2a and b), but here we represent and discuss only results coming from 50% of noise. The ANN with different numbers of neurons in the hidden layer was trained using the back-propagation algorithm [2], with an iteration number of 4000 (Table 2a).

In an example of obtained performance, during the training phase with 10 neurons (Table 2a and b) in the hidden layer ((Fig. 2f), the plotted error corresponds to the accumulated errors obtained by comparing the filtered shot gather (namely the desired output) with the network output. Two criteria to arrest the calculation are adopted: if the performance is around 10^{-6} or if the number of iterations is 4000. Different tests were then

Table 1
Simulation parameters used to generate the training data for ANN

Tableau 1
Paramètres de simulation utilisés pour générer les données d'apprentissage du RNA

No. of layer	Velocity (m s^{-1})	Thickness (m)
1	1200	30.6
2	1700	42.5
3	2100	105
4	2700	135
5	3200	160
6	4000	200

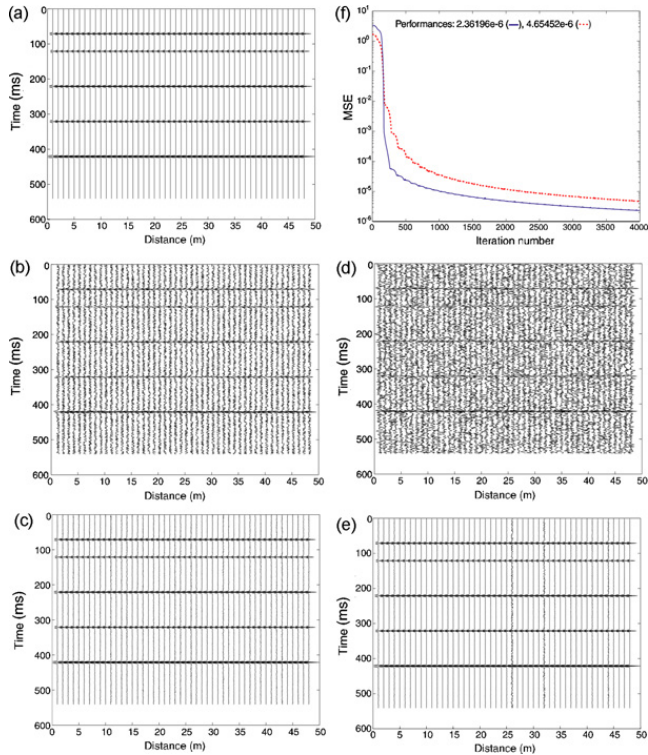


Fig. 2. Training on synthetic data. (a) Used model (see Table 1), considered as desired output. ANN input (same section as in (a) with 50% of random (b) and Gaussian (d) noise respectively with their ANN output after training (c,e). (f) Mean square error (MSE) vs. number of iterations for a training stage. The full (dashed) line curve corresponds to 50% of added random (Gaussian) noise.

Fig. 2. Apprentissage sur données synthétiques. (a) Modèle utilisé (voir Tableau 1), considéré comme sortie désirée. Entrée du RNA (même section qu'en (Fig. 2a) avec 50 % de bruit aléatoire (b) et gaussien (d) et leur sortie du RNA respectivement après apprentissage (c,e). (f) Erreur quadratique moyenne en fonction de nombre d'itérations, dans la phase d'apprentissage. Courbe pleine (en tiretés) correspondant à 50 % de bruit aléatoire (gaussien) ajouté.

performed with different numbers of neurons in the hidden layer; the prediction error increases as the number of neurons grows if we keep 4000 iterations. One can notice that the performance curve for random noise converges more quickly than that for the Gaussian noise ((Fig. 2f). If we compare both diagrams in Fig. 2c and e, the quality of filtering is better on random noise.

The result of filtering the seismic section by the ANN after training is shown in (Fig. 2c). The obtained seismic section by ANN after training (Fig. 2c) is compatible with the seismic model, which is also the desired output (Fig. 2a). When we add 50% of Gaussian noise to our model (Fig. 2d); one can see that the obtained section using ANN after training (Fig. 2e) is similar to the seismic model (Fig. 2a). We therefore demonstrated that the ANN technique clearly improved the seismic section, even if some noise with reduced amplitude still remains.

Table 2

Results of training on synthetic data with random (a) and Gaussian (b) noise and on real data (c)

Tableau 2

Résultats de l'apprentissage sur des données synthétiques avec bruit aléatoire (a), gaussien (b) et sur des données réelles (c)

Percentage of noise	Number of neurons in the hidden layer	Number of iterations	Performance
(a)	10%	10	1.66204×10^{-6}
		15	3.84408×10^{-6}
		20	6.30692×10^{-6}
	50%	10	9.92197×10^{-6}
		15	1.39624×10^{-5}
		20	2.36196×10^{-6}
(b)	10%	10	2.87078×10^{-5}
		15	4.60794×10^{-5}
		20	4.98179×10^{-5}
	50%	10	8.00859×10^{-5}
		15	1.66596×10^{-6}
		20	3.63946×10^{-6}
(c)	50%	25	6.33773×10^{-6}
		30	9.83710×10^{-6}
		30	1.09455×10^{-5}
	90%	50	4.65452×10^{-6}
		60	9.39854×10^{-6}
		70	1.77530×10^{-8}
Real data	50	4000	368.4620
	60	1201	2.33044×10^{-08}
	70	794	1.77530×10^{-08}
	80	648	6.17601×10^{-08}
	90	573	1.09518×10^{-08}

The ANN technique was also applied using the same procedure, with the same model, to filter Gaussian noise at different rates of noise, respectively 10% and 50%. The example in which 50% of Gaussian noise was added to the seismic section reveals that, in the second (~120 ms) and the fourth (~320 ms) horizons, the signal-to-noise ratio is weak. According to our model, i.e. for a constant rate of noise (here 50%), when the velocity contrasts are lower, the signal-to-noise ratio is weaker, and conversely. The result of ANN application (Fig. 2e) is in agreement with the model itself (Fig. 2a). It shows, apart from a few traces where the noise has been reduced significantly, a similarity with the model that can be distinguished (Fig. 2a). Comparing this result with the one obtained with 50% of random noise, we can conclude that better filtering is efficient in presence of random noise.

3.2. Real seismic data

The three following steps must be given chronologically if we wish to implement the use of ANN technique to filter real seismic data.

3.2.1. Data preparation

In order to implement the ANN to filter seismic data, two sets of models are used. The first model is used to train the proposed network structure, while the second one is used to test whether the network can filter the new data using the derived weights. The seismic data used in this study were extracted from a survey of an oil field in the South of Algeria. Each model, which constitutes a shot gather recorded using a Vibroseis source and a 48-trace recorder with a shift of 20 m for the source and receivers, was processed with equalized amplitudes. The sampling rate is 2 ms and each trace contains arrivals between 100 ms and 4000 ms. In order to prepare the data for training, each shot was subjected to an FFT, its frequency components were analyzed on the x direction. If the events recorded within the analyzed traces show coherent aligned signals, the horizontal frequency will be regular along the x direction. Deviations from such regular coherency are caused by the presence of noise. A predictive operator is built in the x direction to reduce the effect of noise by horizontalizing the coherent events. Such an operation

will isolate the noise from the signal; it will be removed easily by subtraction from the original data. Finally, the result will be given in the temporal domain through an inverse FFT algorithm. As an example, one can see that the shot gather (Fig. 3a) shows different kinds of noise (random noise, direct and refracted waves, ground roll), which are reduced significantly by application of such a filtering procedure (Fig. 3b). The projective filter in the ($f-x$) domain separates the signal, assuming that it is predictable in x , from non-predictable noise within all the frequency spectra. We note that the ground roll has been severely reduced, direct and refracted waves were also filtered so that the first horizons can be easily identified between 400 and 900 ms (Fig. 3b). We may also notice that energetic waves were also reduced before 200 ms. We distinguish different levels of reflections between 200 and 500 ms. Some horizons are also detected, for instance at around 1300 and 1800 ms (Fig. 3b), even if they were lacking or presenting a very low signal-to-noise ratio in the original data (Fig. 3a).

3.2.2. Training phase

The shot gather selected from the training set (Fig. 3a) together with the filtered one (Fig. 3b) is used as the desired output of the ANN in the training phase. Since the efficiency of any network application depends mainly on the database used for training and testing the

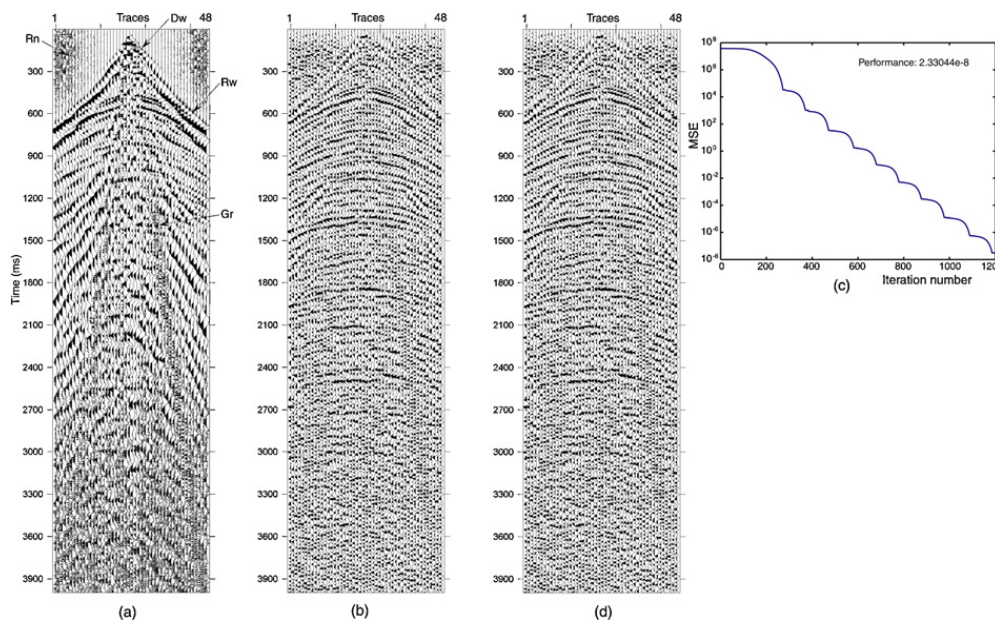


Fig. 3. (a) Full shot gathers. Input of ANN. (b) Corresponding filtered shot gather, which is the desired output for the ANN in the training phase. (c) The performance obtained during the training phase. (d) Output of ANN in the training test. Rn: random noise, Dw: direct wave, Rw: refracted wave, Gr: ground roll.

Fig. 3. (a) Point de tir. Entrée du RNA. (b) point de tir filtré correspondant qui représente la sortie désirée du RNA. (c) Performance obtenue pendant la phase d'apprentissage. (d) Sortie du RNA dans le test d'apprentissage. Rn : bruit aléatoire, Dw : onde directe, Rw : onde réfractée, Gr : onde de surface.

network, the training data set should contain all the possible noise types that can appear in the field.

The ANN with different numbers of neurons in the hidden layer was trained using the back-propagation method. Weights were randomly initialized and the learning rate is at 0.05 [10,11], the number of neurons in the hidden layer is 50, 60, 70 and 90, and the generalization performance is reported in Table 2c. All networks were trained for an identical maximum number of iterations (4000). The performance obtained during the training phase with 60 neurons in the hidden layer is shown in Fig. 3c. Tests performed with different numbers of neurons in the hidden layer show that the prediction error decreases as the number of neurons grows, if we keep 4000 iterations (Table 2c).

However, final training errors are similar if we deal with a number ranging between 60 and 90 neurons in the hidden layer. For 50 neurons, the performance is not really accepted.

To test the learning ability of the proposed ANN structure, we conducted tests on its capability to produce outputs for the set of inputs that were used in the training. The result obtained is shown on section of Fig. 3d. If we compare the output of the network (Fig. 3d) after the training, using the noisy shot gather (Fig. 3a), with the filtered one (Fig. 3b), one can easily notice that the network behaves like a filter. It significantly reduced several types of noise (random noise, direct and refracted waves, ground roll). The network fits the training data accurately and the 60 neurons in the hidden layer are therefore able to perform the task of training satisfactorily.

3.2.3. Generalization phase

Once the training is carried out, i.e. the architecture and the activation functions, the synaptic weights of the network are fixed, the obtained network is used like a classical non-linear function. Now that we have tested the network obtained on the training data, it is important to see what it can do with new data, not mentioned

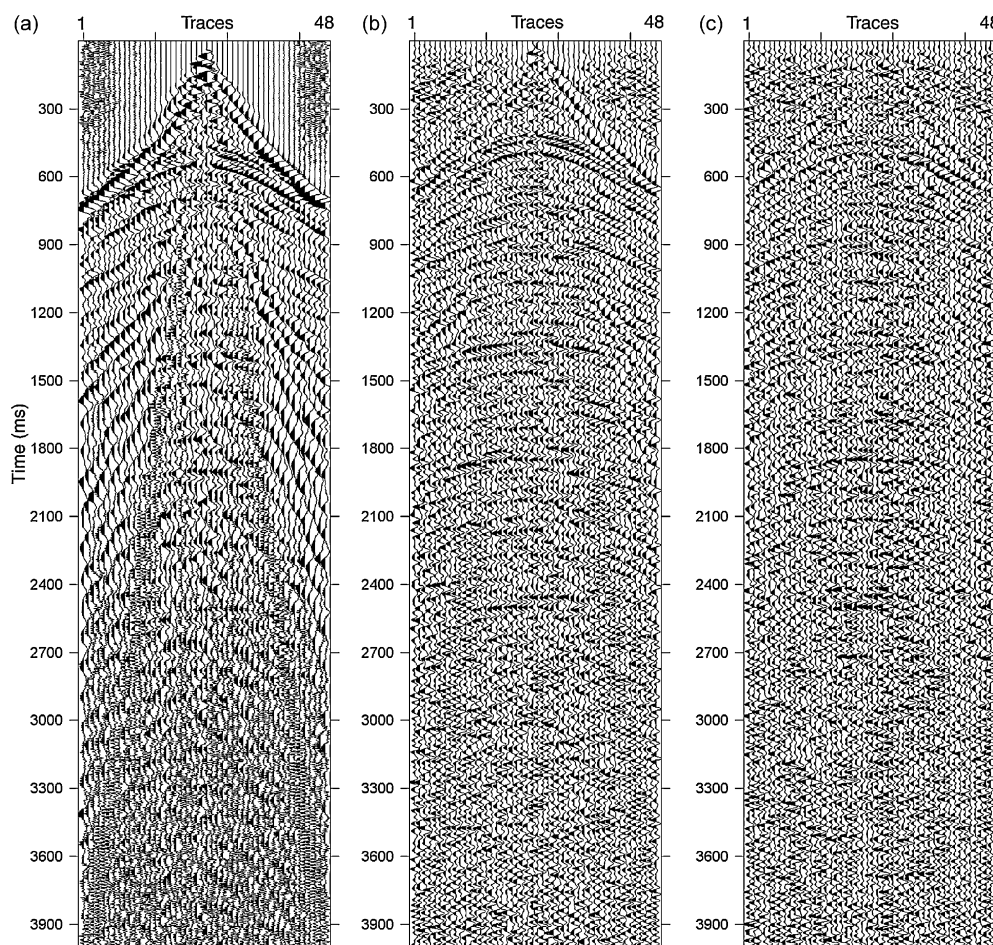


Fig. 4. (a) Full shot gather. Input of the ANN. (b): corresponding filtered shot gather which is the desired output of ANN during the generalisation test. (c): output of ANN during the generalization test. Same notations as in Fig. 3.

Fig. 4. (a) Point de tir. Entrée du RNA. (b) Point de tir filtré, qui correspond à la sortie désirée du RNA pour le test de généralisation. (c) Sortie du RNA pendant le test de généralisation. Mêmes notations que sur la Fig. 3.

before. If this ANN does not give reasonable filtered outputs for this test set, the training period is not achieved. Indeed, this testing is critical to ensure that the network has not simply memorized a given set of data, but has learned the task of data filtering.

The test for its generalization ability is carried out by investigating its capability to filter the new data that were not included in the training process. In this study, for a test network generalization, we plotted output sets and desired outputs of data that were chosen for this stage of network development. To test the trained ANN, new data were used with some noisy events (Fig. 4a). One can see that the noise is attenuated on the obtained output ANN (Fig. 4c). In this case, the trained network is said to have a good generalization performance.

Let's compare both ANN output in the generalization phase (Fig. 4c) and the obtained filtered section (Fig. 4b), which is the desired ANN output, using the previous procedure in data preparation. Isolated horizons can be seen mostly in the upper part of the section (< 1200 ms), where filtering was effective: reduction of energetic areas (< 400 ms), attenuation of direct and refracted waves and ground roll at different levels.

4. Conclusion

In this work, the neural networks of Elman type were addressed to filter seismic data. The results show that networks could establish relationship between noisy and filtered traces, 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 that is based on straight downhill in the weight space till the first valley is reached. This makes the choice of an 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 this network was done successfully, and the Elman network has the ability to obtain strict convergence for a complicated set of data in case of filtering 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 signal-to-noise ratio on data sets that were not learned.

Several tests were performed with synthetic data in the presence of two kinds of noise (random and

Gaussian) with rates of 10% and 50%. The results show that the used ANN structure (i.e. Elman type) was able to recognize and reduce the noise. The same results were also obtained on real seismic data.

We think that feedback 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 types of noise in data contaminated by multiple events classified as noise (e.g., direct and refracted waves, ground roll, random noise) and reduce them significantly.

The main advantage of using ANN for filtering seismic data is that once the ANN has been trained, it has the ability to quickly filter new data.

Acknowledgements

We thank both the two anonymous referees who suggested some remarks that have allowed us to improve the final version of the manuscript.

References

- [1] R. Boite, H. Leich, Les filtres numériques. Analyse et synthèse des filtres unidimensionnels, Masson, Paris, 1991.
- [2] C. Calderón-Macías, M.K. Sen, P.L. Stoffa, Neural networks for parameter estimation in geophysics, *Geophys. Prospect.* 48 (2000) 21–47.
- [3] S. Chen, S. Billings, Neural networks for non-linear dynamic systems modelling and identification, *Int. J. Control* 56 (1992) 319–346.
- [4] E. Davalo, P. Naim, Des réseaux de neurones, Eyrolles, Paris, 1993.
- [5] J.L. Elman, Finding structure in time, *Cogn. Sci.* 14 (1990) 179–211.
- [6] G. Henry, La sismique réflexion : principes et développements, Éditions Technip, Paris, 1997.
- [7] C. Manoj, N. Nagarajan, The application of artificial neural networks to magnetotelluric time-series analysis, *Geophys. J. Int.* 153 (2003) 409–423.
- [8] J.-L. Mari, F. Glangeaud, F. Coppens, Traitement du signal pour géologues et géophysiciens. 1. Prospection sismique, Éditions Technip, Paris, 2001.
- [9] J.-L. Mari, F. Glangeaud, F. Coppens, Traitement du signal pour géologues et géophysiciens. 2. Techniques de base, Éditions Technip, Paris, 2001.
- [10] R. Parisi, E. DiClaudio, G. Orlandi, B.D. Rao, A generalized learning paradigm exploiting the structure of feedforward neural networks, *IEEE Trans. Neural Netw.* 7 (6) (1996) 1450–1460.
- [11] J.-M. Renders, Algorithmes génétiques et réseaux de neurones, Éditions Hermès-Lavoisier, Paris, 1995.
- [12] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning representations by back-propagation errors, *Nature* 393 (1986) 533–536.
- [13] A.W.M. Van den Eenden, N.A.M. Verhoeckx, Traitement numérique du signal, Masson, Paris, 1992.