



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>

Intégration multi-échelles des données de réservoir et quantification des incertitudes

THÈSE

présentée et soutenue le 28 mai 2014
pour l'obtention du grade de

Docteur de l'Université de Lorraine

Spécialité Géosciences

par

Théophile Gentilhomme

Composition du jury :

| | |
|--------------------------------|--|
| <i>Rapporteurs :</i> | Mikaele LE RAVALEC Jaime GOMEZ-HERNANDEZ |
| <i>Examineurs :</i> | Erwann GLOAGUEN Philippe ACKERER Dean OLIVER Rémi MOYEN |
| <i>Directeur de thèse :</i> | Guillaume CAUMON |
| <i>Co-directeur de thèse :</i> | Jean-Jacques ROYER |
| <i>Invité :</i> | Philippe DOYEN Thierry COLÉOU |

UMR 7359 - GeoRessources

ENSG - Université de Lorraine
TSA 70605, 54518 Vandœuvre-lès-Nancy cedex - FRANCE

À Charlotte,
Ma joie,
Mon inspiration,
Ma plus grande découverte

Remerciements

« La reconnaissance est la mémoire du cœur. »

Hans Christian Andersen, 1805-1875

Je tiens tout d'abord à remercier mes directeurs de thèse, Pr. Guillaume Caumon et Pr. Jean-Jacques Royer, qui m'ont fait confiance et soutenu tout au long de ma thèse tout en me laissant une grande liberté dans l'organisation et la direction de mes travaux de recherche.

Mes remerciements les plus sincères vont également aux professeurs Dean Oliver et Trond Mannseth, pour m'avoir fait découvrir les subtilités du monde de l'optimisation, pour leur disponibilité et leurs conseils, pour leur accueil et pour l'intérêt qu'ils ont manifesté pour mon travail. Mes séjours à Bergen, où je m'y suis senti comme un membre à part entière du CIPR, m'ont été très bénéfiques, autant d'un point de vue professionnel que personnel.

Je tiens également à remercier l'entreprise qui a financé ma thèse, CGG, et tout particulièrement Philippe Doyen pour m'avoir fait confiance et pour sa supervision, et Rémi Moyen pour nos discussions, pour avoir toujours répondu à mes appels et pour ses réponses précises et détaillées à mes questions pas toujours claires.

Je remercie Paradigm et Schlumberger pour les licences de Gocad et d'Eclipse qui ont été des outils importants de mon travail de thèse.

Un grand merci à Thierry Coléou pour sa formidable aptitude à partager son savoir et sa passion, pour sa gentillesse et sa patience, pour m'avoir emmené aux JIPs de l'IFP, pour son aide et sa disponibilité.

J'adresse ma gratitude à l'ensemble de mes collègues et amis de Nancy, Paris, Bergen ou Londres, pour leur aide et surtout pour le climat sympathique dans lequel ils m'ont permis de travailler. Merci donc à Jeanne pour ses féroces reviews, à Guillaume pour son aide précieuse pour les MPS, à Flo pour m'avoir permis de collaborer avec lui, à Gautier toujours prêt à rendre service, à Pablo pour nos parties de tir au Crabe (marque déposée), à Kristian et Lina pour leur aide avec Eclipse, à Yan pour son aide avec Brugge, à Zoya toujours stressée, à Nico, Charline, Jérémy, Pauline, Fatima, Paul, Fifi, John, Benjamin, Romain, Christophe, Pierre, Pierre M, Christine, Kristin, Thierry, Fabien, Sveen, Nahim, Randi, Martha, Annie, Yves, Jean-Philippe.

Enfin, je tiens à remercier tendrement ma famille qui a toujours été présente pour me soutenir et m'aider à m'épanouir. Merci à Nanou, Guigui et Charly pour avoir fait découvrir à votre petit frère les sciences et plein d'autres choses.

Merci à ma Charlotte de m'aider au quotidien à donner le meilleur de moi-même. *Last but not least*, un grand merci à mes chers parents pour m'avoir inculqué des valeurs qui me permettent d'avancer tous les jours dans la bonne direction, pour votre soutien sans faille et pour toutes les autres choses qui ne peuvent s'exprimer par des mots.

Résumé

Dans ce travail de thèse, nous proposons de suivre une approche multi-échelles de création de propriétés spatiales de réservoir, permettant d'intégrer des données directes (observation de puits) ou indirectes (sismique et données de production) de résolutions différentes. Deux paramétrisations sont utilisées pour paramétrer le problème : les ondelettes et les pyramides gaussiennes. Dans les deux cas, les propriétés spatiales sont exprimées à l'aide de bases mathématiques possédant à la fois des composantes spatiales et fréquentielles. À l'aide de ces paramétrisations, nous démontrons les avantages de l'approche multi-échelles sur deux types de problèmes d'estimations des incertitudes basés sur la minimisation d'une distance.

Le premier problème traite de la simulation de propriétés à partir d'une méthode de géostatistique multipoints, l'algorithme d'échantillonnage direct (*Direct Sampling* ou DS), permettant de générer un ensemble de réalisations conditionnées par des observations directes et/ou interprétées (carte de porosité sismiques), ayant pour objectif d'échantillonner un espace d'incertitude restreint par l'utilisation d'un modèle géologique conceptuel (image d'entraînement). Il est montré que l'approche multi-échelles basée sur les pyramides gaussiennes améliore la qualité des réalisations générées, respecte davantage les données et réduit les temps de calculs par rapport à l'approche DS standard, et de ce fait, effectue un meilleur échantillonnage de l'espace nul.

Le second problème traite de la préservation des modèles *a priori* lors de l'assimilation des données d'historique de production. L'objectif de cette méthode est de minimiser l'amplitude des perturbations des modèles initiaux et l'insertion de bruit. Des ondelettes de seconde génération sont utilisées pour re-paramétriser le problème. Nous développons une transformée en ondelette 3D applicable à des grilles stratigraphiques complexes de réservoir, possédant des cellules mortes ou de volume négligeable. Afin d'estimer les incertitudes liées à l'aspect mal posé du problème inverse, une méthode d'optimisation par ensemble est intégrée dans l'approche multi-échelles de calage historique. À l'aide de plusieurs exemples d'applications, nous montrons que l'inversion multi-échelles permet de mieux préserver les modèles *a priori* et est moins assujettie au bruit que les approches standards, tout en respectant aussi bien les données de conditionnement.

Enfin, nous proposons un *workflow* complet d'intégration séquentielle des données sismiques et de production à l'aide de la méthode multi-échelles de calage historique développée dans cette thèse. L'application de ce *workflow* est illustré par un cas synthétique 3D dont nous détaillons la construction.

Table des matières

| | |
|--|------------|
| Remerciements | iii |
| Résumé | v |
| Introduction | 1 |
| 1 Simulations numériques et problèmes inverses | 9 |
| 1.1 Résumé et contribution | 9 |
| 1.2 Simulations numériques | 9 |
| 1.2.1 Simulation de la réponse sismique | 11 |
| Propagation, réflexion, réfraction et conversion des ondes sismiques | 11 |
| Traitement du signal sismique brut | 14 |
| Méthodes numériques de modélisation de la réponse sis- mique par convolution 1D | 15 |
| Physique des roches et propriétés élastiques | 17 |
| 1.2.2 Simulation de la production de fluide | 17 |
| Modèles mathématiques | 18 |
| Résolution numérique | 20 |
| Paramètres d'entrée et de sortie | 21 |
| 1.2.3 Utilisation de modèles de substitution | 21 |
| 1.3 Inversions des données de réservoir | 22 |
| 1.3.1 Problème inverse et méthodes d'optimisation | 22 |
| Modèle direct | 22 |
| Données observées | 23 |
| Notion de problème inverse | 24 |
| Modèle <i>a priori</i> | 24 |
| Distribution de probabilité et probabilité maximale . . . | 25 |
| Méthodes d'optimisation | 26 |
| 1.3.2 Optimisation de Gauss-Newton | 27 |
| Méthode de Newton | 27 |
| Méthode de Gauss-Newton | 27 |
| Recherche linéaire et région de confiance | 29 |
| Critères de convergence | 31 |
| 1.3.3 Algorithmes d'échantillonnage : quantification des incer- titudes | 31 |
| Cas linéaires | 32 |

| | |
|--|-----------|
| Échantillonnage par minimisation : méthode RML pour problèmes inverses non linéaires | 33 |
| Méthodes d'ensemble | 34 |
| 2 Re-paramétrisations multi-échelles | 39 |
| 2.1 Résumé et contributions | 39 |
| 2.2 Choix de la paramétrisation | 39 |
| 2.2.1 Analyse linéaire de résolution des paramètres | 40 |
| 2.2.2 Paramétrisations adaptatives et approches multi-échelles | 41 |
| 2.3 Une paramétrisation multi-échelles adaptée aux problématiques des géosciences : les ondelettes | 42 |
| 2.3.1 Les ondelettes de première génération | 44 |
| 2.3.2 Lifting en ondelettes et ondelettes de seconde génération | 51 |
| Ondelettes et filtres bi-orthogonaux | 52 |
| Lifting en ondelettes | 53 |
| <i>Split, Predict, Update</i> | 55 |
| <i>Lifting</i> et ondelettes de Haar déséquilibrées | 57 |
| Ondelettes basées sur l'interpolation : exemple de l'ondelette "Chapeau mexicain" | 59 |
| Ondelettes construites à partir de la méthode de subdivision par interpolation de la moyenne | 61 |
| 2.3.3 Décompositions multidimensionnelles | 63 |
| 2.4 Pyramides gaussiennes | 64 |
| 2.5 Conclusions partielles et perspectives | 66 |
| 3 Intégration des données statiques par méthodes géostatistiques multipoints multi-échelles | 69 |
| 3.1 Résumé et contributions | 69 |
| 3.1.1 Résumé | 69 |
| 3.1.2 Contributions | 70 |
| 3.2 Méthodes géostatistiques | 70 |
| 3.2.1 Géostatistiques deux-points | 71 |
| Fonctions de corrélations et matrice de covariance | 71 |
| Estimation : exemple du krigeage simple | 72 |
| Simulation stochastique : exemple de la Simulation Séquentielle Gaussienne (SGS) | 73 |
| 3.2.2 Géostatistiques multipoints | 74 |
| 3.2.3 <i>Direct Sampling</i> | 75 |
| 3.3 <i>Direct Sampling</i> Multi-échelles (MS-DS) | 77 |
| 3.3.1 MS-DS basé sur les ondelettes (MS-DSw) et intégration de données secondaires | 79 |
| Décomposition de l'image d'entraînement et adaptation des données primaires à l'échelle grossière | 79 |
| Simulation des coefficients d'échelle | 81 |
| Simulation conditionnée par les coefficients d'échelle | 84 |
| Résultats et discussions | 85 |

| | | |
|----------|---|------------|
| 3.3.2 | MS-DS basé sur la décomposition en pyramides gaussiennes (MS-DSgp) | 88 |
| | Simulation des données de conditionnement | 89 |
| | Simulation d'une échelle r | 91 |
| | Traitements aux échelles intermédiaires | 94 |
| | Résultats et discussions | 95 |
| 3.4 | Conclusions partielles et perspectives | 103 |
| 4 | Approche multi-échelles de calage des données de production basée sur des méthodes d'optimisation par ensemble | 105 |
| 4.1 | Résumé et contributions | 105 |
| 4.1.1 | Résumé | 105 |
| 4.1.2 | Contributions | 107 |
| 4.2 | Paramétrisation | 108 |
| 4.3 | Optimisation d'un sous-ensemble de coefficients d'ondelettes . . | 108 |
| 4.3.1 | Notations | 108 |
| 4.3.2 | Méthode d'optimisation | 109 |
| | RML Levenberg-Marquardt basé ensemble | 109 |
| | Lissage initial des champs de propriétés | 111 |
| 4.4 | L'approche multi-échelles | 112 |
| 4.4.1 | Localisation multi-échelles | 114 |
| | Localisation multi-échelles dans l'espace de Dirac | 115 |
| | Régularisation du gain de Kalman dans l'espace des ondelettes | 116 |
| 4.4.2 | Optimisation et contrôle du paramètre de Levenberg-Marquardt | 118 |
| 4.4.3 | Validation de la méthode | 120 |
| 4.5 | Tests numériques | 121 |
| 4.5.1 | Réservoir 2D | 121 |
| 4.5.2 | Le cas Brugge | 125 |
| 4.6 | Discussions et conclusions | 132 |
| 4.7 | Algorithmes | 134 |
| 5 | Intégration séquentielle des données sismiques et de production : exemple d'application | 139 |
| 5.1 | Résumé et contributions | 139 |
| 5.1.1 | Résumé | 139 |
| 5.1.2 | Contributions | 140 |
| 5.2 | Présentation du jeu de données synthétiques | 140 |
| 5.2.1 | Modèle géologique fin | 140 |
| 5.2.2 | Création des données sismiques | 142 |
| | Modèle pétro-élastique (PEM) | 142 |
| | Simulation des amplitudes sismiques | 143 |
| 5.2.3 | Création des données de production | 145 |
| 5.3 | Inversion des données | 150 |
| 5.3.1 | Inversion sismique stochastique des paramètres élastiques | 150 |
| 5.3.2 | Interprétation des modèles élastiques | 150 |

| | | |
|----------|---|------------|
| 5.3.3 | <i>Upscaling</i> et conversion temps-profondeur | 152 |
| | <i>Upscaling</i> | 152 |
| | Conversion temps-profondeur | 157 |
| 5.3.4 | Inversions des données de production | 157 |
| 5.4 | Perspectives et conclusions partielles | 160 |
| 6 | Conclusions générales et perspectives | 163 |
| 6.1 | Contributions de la thèse | 163 |
| 6.2 | Études complémentaires | 165 |
| | Résumé | |

Table des figures

| | | |
|------|--|----|
| 1 | Évolution d'un système pétrolier | 2 |
| 2 | Systèmes de chenaux | 3 |
| 3 | Grille discrète de réservoir | 4 |
| 1.1 | Acquisition sismique et schéma de réflexion/réfraction des ondes | 12 |
| 1.2 | Convolution sismique 1D | 16 |
| 1.3 | Contrôle de la mise à jour de Gauss-Newton dans l'espace des paramètres | 29 |
| 2.1 | Analyse temps-fréquence de la représentation de Dirac et de Fourier | 42 |
| 2.2 | Décomposition en coefficients d'ondelettes | 43 |
| 2.3 | Analyse temps-fréquence et temps-échelle | 45 |
| 2.4 | Bases d'ondelettes (droite) et de fonctions d'échelle (gauche) . . | 47 |
| 2.5 | Algorithme de décomposition en cascade (<i>filter bank</i>) | 50 |
| 2.6 | Schéma de <i>lifting</i> en ondelettes direct et inverse | 56 |
| 2.7 | Algorithme de la transformée de Haar déséquilibrée | 58 |
| 2.8 | Reconstruction après perturbation des coefficients d'ondelette . | 59 |
| 2.9 | Ondelette basée sur l'interpolation linéaire | 60 |
| 2.10 | Subdivision par interpolation de la moyenne | 61 |
| 2.11 | Ondelettes de Haar améliorées | 63 |
| 2.12 | Transformée 3D en ondelettes | 64 |
| 2.13 | Transformation anisotrope | 65 |
| 2.14 | Pyramide gaussienne | 66 |
| 2.15 | Pyramide gaussienne et représentation temps-fréquence | 67 |
| 3.1 | Résultat d'un krigeage | 73 |
| 3.2 | Réalisations générées par une SGS | 74 |
| 3.3 | Illustration de l'algorithme d'échantillonnage direct. Tirée de [100] | 76 |
| 3.4 | Chemin semi-aléatoire | 77 |
| 3.5 | Résultats de l'algorithme d'échantillonnage direct conditionnés par des données de puits | 78 |
| 3.6 | Relations spatiales et d'échelles des coefficients d'ondelettes . . . | 80 |
| 3.7 | Image d'entraînement et données de conditionnement à l'échelle grossière | 82 |
| 3.8 | Simulation des coefficients d'échelle | 83 |
| 3.9 | MS-DSw voisinage de simulation | 84 |
| 3.10 | MS-DSw sans conditionnement | 86 |
| 3.11 | MS-DSw conditionné : données initiales | 87 |

| | | |
|------|---|-----|
| 3.12 | Simulations conditionnées par les données secondaires 1 | 87 |
| 3.13 | Simulations conditionnées par les données secondaires 2 | 88 |
| 3.14 | MS-DSgp simulation des données de conditionnement à l'échelle grossière r | 90 |
| 3.15 | MS-DSgp simulation des échelles | 92 |
| 3.16 | Contrôle du balayage de l'image d'entraînement par la position des coefficients | 93 |
| 3.17 | Contrôle du balayage de l'image d'entraînement par la valeur du coefficient grossier | 93 |
| 3.18 | Traitements intermédiaires | 94 |
| 3.19 | Images d'entraînement | 97 |
| 3.20 | Test de minimisation | 98 |
| 3.21 | Réalisations MD-DSgp et DS standard | 100 |
| 3.22 | Matrice des dissimilarités | 101 |
| 3.23 | Représentation MDS | 102 |
| 3.24 | Validation de la reproduction MDS | 102 |
| | | |
| 4.1 | Lissage initial des réalisations | 112 |
| 4.2 | Présentation schématique de la méthode multi-échelles | 114 |
| 4.3 | Localisation multi-échelles | 118 |
| 4.4 | Porosité et perméabilité de référence pour le test 1 | 122 |
| 4.5 | Test 1 : évolution des fonctions objectifs | 124 |
| 4.6 | Test 1 : calages et prédictions | 125 |
| 4.7 | Test 1 : comparaisons des propriétés moyennes | 126 |
| 4.8 | Test 1 : réalisations avant et après optimisation | 127 |
| 4.9 | Test 1 : différences moyennes avec le modèle <i>a priori</i> | 128 |
| 4.10 | Test 1 : analyses variographiques des réalisations | 129 |
| 4.11 | Test 2 : évolution de la fonction objectif | 130 |
| 4.12 | Test 2 : calage des données de production | 131 |
| 4.13 | Test 2 : réalisations avant et après optimisation | 132 |
| 4.14 | Test 2 : différences avec le modèle <i>a priori</i> | 133 |
| | | |
| 5.1 | Intégration séquentielle des données | 140 |
| 5.2 | Modèle géologique fin | 141 |
| 5.3 | Contacts fluides | 143 |
| 5.4 | Représentation schématique d'un modèle de grès cimentés | 143 |
| 5.5 | Évolution de V_p en fonction de la porosité et de la quantité de ciment | 144 |
| 5.6 | Ondelettes synthétiques | 146 |
| 5.7 | Sismique synthétique | 147 |
| 5.8 | Grille de simulation des données de production | 148 |
| 5.9 | Perméabilités relatives | 148 |
| 5.10 | Estimation de l'ondelette 30 – 40° | 151 |
| 5.11 | Résolution verticale d'une réalisation d'impédance | 152 |
| 5.12 | Réalisations d'impédances P , couche 24 (modèle fin) | 153 |
| 5.13 | Relations Poro- I_p - I_s | 153 |
| 5.14 | Relations Perméabilité - Porosité | 154 |

| | | |
|------|--|-----|
| 5.15 | Réalisations pétrophysiques | 155 |
| 5.16 | Grille de réservoir utilisée lors de l'inversion des données de production | 156 |
| 5.17 | Perméabilités équivalentes | 156 |
| 5.18 | Localisation des données de pression, puits P7 | 158 |
| 5.19 | Évolution de la fonction objectif et du nombre de paramètres . . | 158 |
| 5.20 | Calage des données de production | 159 |
| 5.21 | Réalisations initiales et finales | 161 |
| 6.1 | Re-paramétrisation en fonction des amplitudes de variation d'impédances | 166 |

Introduction

« Le commencement est la moitié du tout. »

Pythagore, 580-495 av. J.C.

Problématique générale

L'exploitation de l'eau, des matières minérales et des hydrocarbures, ainsi que son impact sur l'environnement, est devenue une problématique majeure de développement des sociétés modernes. Dans un contexte de raréfaction des ressources et d'exploitation responsable de gisements de plus en plus complexes (exploitation en eaux profondes) ou non conventionnels (gaz de schistes), la caractérisation des structures géologiques du sous-sol, conditionnant l'écoulement des fluides, est une étape primordiale des processus d'optimisation de la production et de minimisation des risques de pollution. Dans ce travail, nous nous intéressons à la caractérisation des propriétés pétrophysiques des roches constitutives de réservoirs d'hydrocarbures, telles que la porosité et la perméabilité, afin de pouvoir prédire l'évolution de la composition fluide et de maximiser la récupération lors de l'exploitation d'un champ pétrolier, tout en minimisant les coûts de production.

Qu'est-ce qu'un réservoir pétrolier ?

Un réservoir pétrolier est une formation géologique enfouie, composée de roches poreuses et perméables saturées par des fluides interstitiels récupérables. C'est un élément d'un système pétrolier qui permet l'accumulation d'hydrocarbures. Afin d'obtenir un réservoir exploitable, le système pétrolier doit comprendre plusieurs étapes : la création d'hydrocarbures, leur migration et leur piégeage dans le réservoir (voir figure 1).

La formation des hydrocarbures nécessite la présence de matière organique (préservée) soumise à des températures extrêmes dues à son enfouissement. En fonction du type de matière organique (organismes micro cellulaires, plancton ou végétaux), différents types d'hydrocarbures pourront être produits à partir de cette roche organique appelée « roche mère ». Généralement, le stockage de ces hydrocarbures ne s'effectue pas au même endroit, ni au même moment que leur création. Les molécules hydrocarbures étant plus légères que l'eau, elles migrent vers la surface jusqu'à ce qu'elles atteignent des zones réservoirs poreuses où elles peuvent se concentrer. Cependant, les hydrocarbures ne pourront

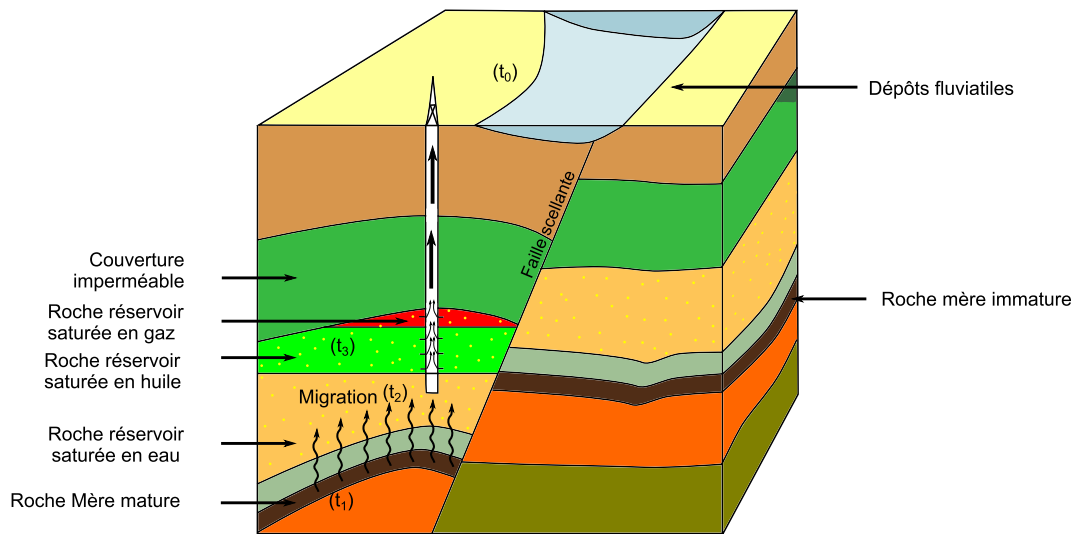


FIGURE 1 – Évolution d'un système pétrolier. (t_0) Dépôts fluviaux précurseurs possibles de roches réservoirs. (t_1) Formation de molécules d'hydrocarbures. (t_2) Migration des hydrocarbures. (t_3) Accumulation des hydrocarbures dans le réservoir.

s'accumuler que s'ils sont arrêtés par des barrières imperméables (couche argileuse ou faille scellante) et ne peuvent pas migrer davantage. Un piège pétrolier est alors composé d'une couverture imperméable et d'un réservoir.

Un réservoir est créé soit par l'accumulation de roches poreuses et perméables lors des processus de sédimentation (réservoirs clastiques), soit par la fracturation (réservoirs fracturés) ou la dégradation chimique d'une roche (p. ex. karst dans une roche carbonatée). Ainsi, la variation spatiale des propriétés d'un réservoir est directement liée aux processus de création des interstices poreux. Par exemple, dans un réservoir fracturé, l'orientation préférentielle de la perméabilité est directement liée à la direction des contraintes locales ou régionales qui ont provoqué la fracturation de la roche [18]. De même, un paléokarst présentera une perméabilité bien plus importante dans la direction originelle de formation des cavités et des galeries [37].

Dans ce travail, nous nous intéressons particulièrement aux réservoirs clastiques issus du dépôt de matériaux associés à des systèmes fluviaux (continentaux figure 2 ou turbiditiques). Ces systèmes permettent le dépôt d'éléments sableux qui développent une porosité interstitielle importante créée par l'espace entre les grains. Ces dépôts sableux sont généralement associés à des dépôts plus argileux ayant une porosité et une perméabilité moindre (p. ex. plaines d'inondations, méandres abandonnés) qui limitent les écoulements dans le réservoir. Ainsi, les propriétés intrinsèques des roches et la structure du réservoir dépendent du type de système sédimentaire et de l'agencement des dépôts associés.

Géomodélisation des réservoirs

Un géomodèle est un ensemble d'informations et/ou d'interprétations regroupées de manière cohérente vis-à-vis de processus géologiques théoriques. Il

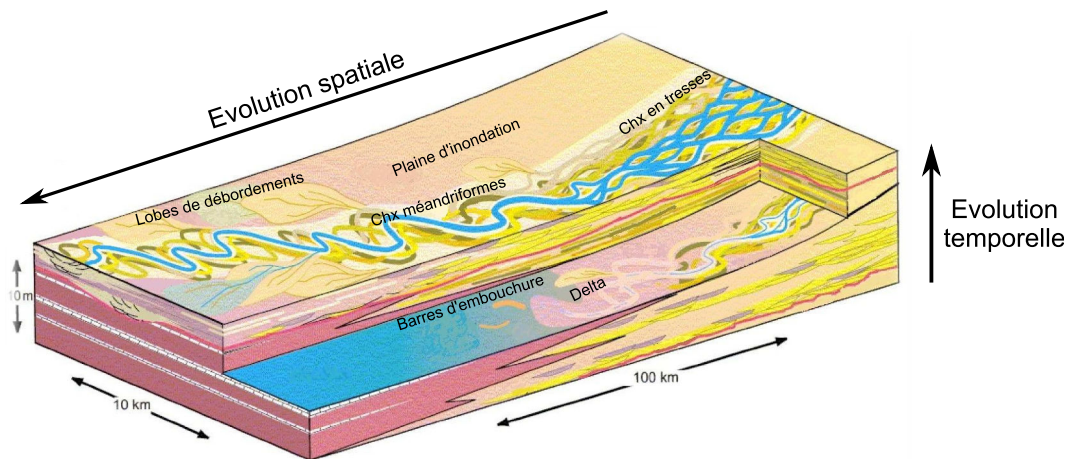


FIGURE 2 – Systèmes de chenaux. Évolution temporelle et spatiale des systèmes fluviaux et du type de dépôts sédimentaires associés. Modifié de [57].

permet une visualisation efficace des données et des éléments constitutifs du modèle, ce qui facilite les contrôles de qualité des données et du modèle, ainsi que l'analyse des résultats simulés.

En étude réservoir, l'un des objectifs du géomodèle est de définir la géométrie du domaine étudié. Les limites des réservoirs sont principalement conditionnées par les horizons géologiques et les failles. Ces derniers sont interprétés à partir des données sismiques et de puits. Afin de pouvoir modéliser les structures et les propriétés des réservoirs, ainsi que l'évolution de leurs états dynamiques, des modèles numériques 3D sont créés à partir de logiciels de géomodélisation. Des grilles sont générées afin de discrétiser le domaine étudié en volumes élémentaires (p. ex. figure 3). Chacun de ces éléments est associé à une ou plusieurs valeurs locales des propriétés du réservoir. À partir de cette représentation, il est possible d'effectuer des calculs numériques permettant de simuler le comportement spatial et temporel du réservoir.

Un des objectifs de la géomodélisation est d'obtenir la meilleure approximation possible des propriétés afin de pouvoir prédire les écoulements dans le réservoir et d'optimiser l'exploitation des fluides. Pour obtenir un modèle réaliste, l'estimation des faciès et des propriétés doit prendre en compte l'ensemble des données disponibles et doit être compatible avec les processus responsables de la création du réservoir.

Utilisation de données d'observations indirectes

Les réservoirs étant généralement enfouis sous des centaines voire des milliers de mètres d'eau et de roche, seuls quelques puits peuvent être forés. De ce fait, peu d'information est disponible pour créer les modèles numériques. Pour cette raison, des mesures indirectes, telles que les données sismiques ou de production de puits, sont utilisées pour caractériser les propriétés du sous-sol et assigner des valeurs à chacun des éléments du modèle. Ces mesures enregistrent la réponse d'un phénomène physique provoqué par la perturbation du réservoir (p. ex. production de fluide ou envoi d'ondes). En connaissant les lois physiques

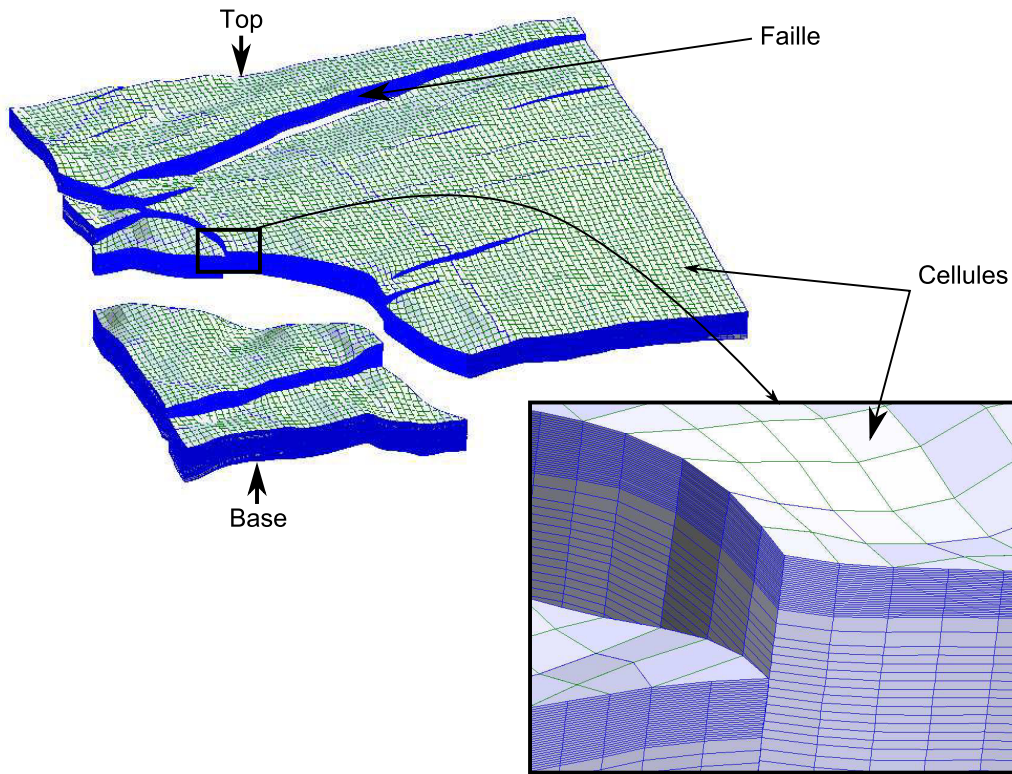


FIGURE 3 – Grille discrète de réservoir. Le domaine d’étude, défini par les failles et les horizons (toit et base) du réservoir, est discrétisé en cellules hexaédriques.

contrôlant cette réponse, il est possible de déduire certaines propriétés du système. On parle alors de résolution de problèmes inverses ou d’inversions des données [131]. Dans ce travail, en plus des observations directes, nous utilisons deux types de données indirectes : les données sismiques et les données de production.

Les données sismiques enregistrent les temps de propagation et les amplitudes des ondes envoyées dans le sous-sol. Les vitesses de propagation étant liées à la nature des milieux, il est possible de déduire certaines caractéristiques des roches traversées par les ondes (p. ex. les propriétés élastiques). En effet, les roches denses et consolidées sont associées à des vitesses de propagation élevées et à une meilleure conservation de l’énergie des ondes. Au contraire, des roches poreuses et peu consolidées sont associées à des vitesses de propagation plus faibles et à une plus grande dissipation de l’énergie. Ce phénomène peut être expérimenté lors de l’arrivée d’un train : en collant l’oreille sur le rail (test non recommandé !), il est possible de prévoir son arrivée bien avant qu’il soit entendu à travers l’air qui est bien moins dense que l’acier¹. Cependant, les données sismiques ne portent de l’information que dans une certaine gamme de fréquences. De plus, le signal peut n’être que partiellement récupéré dans certaines configurations (p. ex. zones faillées, dômes de sel). Ainsi, pour avoir une caractérisation complète des propriétés, il est nécessaire d’utiliser d’autres sources d’information.

¹vitesses à 20°C du son dans l’air : 333 m/s ; eau : 1480 ; acier : 5600 à 5900

Les données de production enregistrent aux puits les variations de pressions, les débits, les proportions et les compositions des fluides produits lors de l'exploitation du réservoir. Les écoulements étant liés à la perméabilité et la porosité du milieu, il est possible, connaissant les lois de la mécanique des fluides, d'en déduire les propriétés des roches traversées par les fluides. Cependant, le contenu fréquentiel de ces données est hétérogène dans l'espace. En effet, les paramètres proches des puits ayant une influence plus importante sur la production, ils peuvent être caractérisés de manière plus précise ; loin des puits, seules des caractéristiques grossières ou ayant un impact très important sur les écoulements (p. ex. drains et barrières de transmissibilité extrêmes) peuvent être extraites des données.

Dans le Chapitre 1, nous présentons des méthodes dites d'optimisation (ou de minimisation) permettant d'estimer les paramètres du modèle grâce à ces données.

Plusieurs modèles possibles

Il est important de comprendre que les observations indirectes ne permettent de caractériser que partiellement les différents aspects des roches du réservoir et qu'elles sont généralement bruitées et limitées. De plus, les données directes ne sont disponibles que de façon ponctuelle dans le réservoir. En conséquence, les informations disponibles ne permettent généralement pas de résoudre l'ensemble des paramètres du modèle. On parle alors de problèmes mal posés. Un problème mal posé est un problème qui accepte plusieurs solutions équivalentes². Ainsi, différents modèles de réservoirs peuvent reproduire les données observées.

Il est impératif de prendre en compte cet aspect lors des études de réservoirs : les prédictions fournies par un modèle ne sont valables que dans un intervalle de confiance qui dépend du nombre et de la qualité des données assimilées. Généralement, le calcul analytique de cet intervalle est impossible et des méthodes approximatives doivent être utilisées. Une approche classique d'estimation des incertitudes consiste à construire plusieurs modèles différents qui honorent les données observées. Cependant, nous verrons qu'il peut être difficile et coûteux de générer plusieurs de ces modèles, et que le processus de calage peut être instable et peut introduire du bruit. De plus, les données étant de nature différentes, elles sont rarement assimilées simultanément. Généralement, un premier modèle est créé à partir des puits et de la sismique, puis est perturbé afin de reproduire les données de production. Cette approche ne garantit pas que le modèle final respecte l'ensemble des données disponibles. Or pour être prédictifs et quantifier correctement les incertitudes, les modèles doivent être cohérents avec l'ensemble des données et doivent être compatibles avec les processus géologiques à l'origine de la création du réservoir.

Dans cette optique, ce travail de thèse cherche à répondre aux questions suivantes :

1. Comment créer des modèles de réservoir compatibles avec un concept géologique ?

²ou qui n'a pas de solution

2. Comment intégrer des données de nature et de résolution différentes ?
3. Comment préserver la structure initiale des modèles de réservoir lors de l'intégration des données indirectes ?

Contributions

Dans ce travail, nous proposons de suivre une approche multi-échelles de création de champs de propriétés contraints par des données d'observation directes et/ou indirectes. Au lieu d'estimer des valeurs correspondantes aux cellules d'un modèle, nous travaillons avec des paramètres qui décomposent localement les propriétés en différentes gammes de fréquences que l'on appelle « échelles ». Plus précisément, nous utilisons des paramétrisations en ondelettes et en pyramides gaussiennes. Ces paramétrisations sont définies à la fois dans l'espace et en fréquence. Elles sont particulièrement bien adaptées aux problèmes rencontrés en géosciences. En effet, elles permettent d'ajuster la paramétrisation des problèmes à la résolution locale des données. De plus, elles permettent de fournir des approximations grossières des champs de propriétés à partir d'un nombre très limité de paramètres.

La réduction du nombre de paramètres permet de stabiliser le processus d'inversion. En effet, le problème inverse devient mieux posé, c'est-à-dire que les données disponibles permettent de résoudre l'ensemble des paramètres décrivant le modèle. De plus, en réduisant la dimension du problème, les processus d'estimation des variables deviennent beaucoup plus efficaces, ce qui permet l'application de méthodes plus complexes produisant des modèles de bonne qualité.

Les données comportant des informations à différentes résolutions, il est tout de même nécessaire d'intégrer l'ensemble des paramètres associés aux différentes fréquences des propriétés. Cependant, l'estimation des échelles grossières permet de stabiliser le processus d'inversion, de régulariser l'estimation des échelles fines et de préserver les modèles initiaux.

Dans le Chapitre 2, après un développement théorique, nous introduisons une paramétrisation en ondelettes applicable aux grilles stratigraphiques utilisées dans les études de réservoirs. Les avantages de l'approche multi-échelles sont ensuite utilisés dans le Chapitre 3 afin de générer des modèles fluviaux à partir des données d'observation directes et de données sismiques interprétées. Dans le Chapitre 4, nous proposons une méthode multi-échelles d'intégration des données de production qui permet de limiter les perturbations des modèles initiaux et l'addition de bruit. Enfin, un exemple d'intégration séquentielle des données sismiques et de production, à partir des méthodes développées dans cette thèse, est exposé dans le Chapitre 5.

Contexte de la thèse

Cette thèse CIFRE a été sponsorisée par la Compagnie Générale de Géophysique (CGG) et effectuée au sein de l'université de Lorraine dans le laboratoire GéoRessources en collaboration avec le Centre pour la Recherche Pétrolière Intégrée de l'université de Bergen (Uni-CIPR, Norvège). Ce travail a abouti à la

publication/soumission de deux articles dans des revues scientifiques [87, 64] et de plusieurs articles de conférences [65, 63, 103, 115].

Le développement des applications a été effectué sous Linux en C++ dans le géomodeleur Gocad 2009.3 de Paradigm et représente environ la moitié du temps de travail de cette thèse. Les écoulements ont été simulés avec le logiciel Eclipse de Schlumberger. La modélisation et les inversions sismiques ont été effectuées grâce à la plateforme Stratavista³ de CGG.

³logiciel interne CGG

Chapitre 1

Simulations numériques et problèmes inverses

« Nous nommons vrai un concept qui concorde avec le système général de tous nos concepts, vraie une perception qui ne contredit pas le système de nos perceptions ; la vérité est cohérence. »

Miguel de Unamuno, 1864-1936

1.1 Résumé et contribution

Dans la première partie de ce chapitre nous présentons les données et les phénomènes physiques associés, ainsi que les méthodes de simulations utilisées dans ce travail de thèse. Dans la deuxième partie, nous présentons des méthodes qui permettent de caractériser les paramètres d'un modèle de réservoir à partir de ces données d'observation. Le développement de la Section 1.3 est centré sur les méthodes d'optimisation de Gauss-Newton et les méthodes d'ensemble correspondantes, ces algorithmes étant utilisés dans le Chapitre 4.

Ce chapitre introduit les problématiques et les notions abordées dans ce travail de thèse, fait un état de l'art des techniques disponibles pour répondre aux problèmes posés et argumente les choix des méthodes utilisées par la suite. Néanmoins, ce chapitre ne contient pas de contributions majeures.

1.2 Simulations numériques

Une simulation numérique est une séquence de calculs informatiques qui permet de reproduire un processus physique décrit par un modèle mathématique et un ensemble de paramètres donnés. Le terme simulation est souvent associé à un processus physique caractérisé par des équations aux dérivées partielles fonctions du temps et de l'espace. On distingue alors deux types de paramètres : les paramètres statiques et dynamiques (variables d'état) du modèle. Les paramètres statiques sont fixes et donnés au début de la simulation. Les variables dynamiques décrivent l'évolution de l'état d'un système au cours du

temps ; elles doivent être initialisées et sont modifiées par le modèle numérique au cours de la simulation.

Lorsque le modèle mathématique ne dépend pas du temps ou ne contient pas de dérivées partielles, la réponse physique est uniquement contrôlée par les paramètres statiques du modèles (voir Section 1.2.1). Dans les deux cas, on appelle « modèle direct » (*forward model*), la simulation d’une réponse physique définie par un modèle mathématique et « résolution du problème direct » le processus de calcul de cette réponse. Dans cette thèse, nous nous intéressons principalement aux simulations d’écoulements de fluides dans des milieux poreux perméables et à la modélisation de la réponse sismique pour un milieu élastique continu donné.

Les équations du modèle mathématique sont généralement décrites de façon continue. Lorsque les paramètres d’entrée peuvent également être décrits par des fonctions mathématiques continues (par exemple des fonctions constantes), il est parfois possible (souvent en faisant quelques approximations) de calculer la réponse directe analytiquement. Malheureusement dans de nombreux cas, les distributions des paramètres, la forme du domaine et les conditions aux limites sont complexes et multidimensionnelles (champs des propriétés spatiales, telles que la porosité, la perméabilité), ce qui ne permet pas un calcul analytique de la solution. Dans ce cas, il est nécessaire de discrétiser le modèle mathématique afin de résoudre le problème numériquement.

La discrétisation est une méthode d’analyse numérique qui permet d’approximer la solution d’un problème continu par une solution d’un problème discret obtenue par ordinateur. Une fonction continue sera par exemple approchée par une quantité finie de valeurs associées à des points disposés dans le domaine étudié. Il existe différentes méthodes de discrétisation, plus ou moins adaptées aux différents modèles mathématiques, incluant la méthode des éléments finis, la méthode des différences finies, la méthode des volumes finis.

Le domaine discrétisé ou maillage est composé d’un ensemble fini de formes géométriques 2D (polygones) ou 3D (polyèdres, tels que les hexaèdres, prismes, tétraèdres, ...) contraints par la géométrie du domaine. Les maillages doivent remplir certains critères de qualité (p. ex. angles des polygones, orthogonalité des hexaèdres, aspect ratio, asymétrie) en fonction du type de méthode numérique, afin de limiter la diffusion d’erreurs numériques. La création d’un maillage peut s’avérer très difficile lorsque le domaine a une géométrie complexe [107]. Ainsi, la faculté d’approximation de la géométrie du domaine par des éléments de bonne qualité conditionnera fortement la qualité d’estimation de la réponse physique du modèle direct.

Dans un certain nombre de maillages constitués uniquement d’hexaèdres, il est possible de traiter chaque élément par un indice unique, (i, j) ou (i, j, k) , ce qui facilite l’identification, l’accès et le positionnement relatif de chaque cellule. On parle alors de grilles structurées. À l’inverse, lorsque les éléments ne sont pas identifiables par un indice unique, on parle de grilles non-structurées. Les grilles structurées sont généralement utilisées en différences finies car elles ne nécessitent pas d’information géométrique supplémentaire sur la structure de la grille. Elles sont également largement utilisées en simulations d’écoulements et de la réponse sismique. Bien qu’elles soient plus difficiles à construire que les

grilles non-structurées, elles permettent un contrôle plus simple de la qualité des cellules. De plus, elles permettent de remplir l'espace efficacement et supportent un niveau d'asymétrie et d'étirement élevé avant d'impacter le calcul de la solution numérique. Néanmoins, lorsque la géométrie du domaine est complexe, les grilles structurées introduisent des approximations fortes dans la discrétisation du milieu. Il peut être alors utile d'utiliser des maillages irréguliers de type Voronoi [102] pour les simulations d'écoulements. Dans cette thèse, nous utilisons uniquement des grilles régulières hexaédriques car elles demeurent le standard pour les simulations d'écoulements en milieu industriel.

Une fois le domaine discrétisé, il est nécessaire de définir les valeurs des propriétés nécessaires à la résolution du problème direct pour chacune des cellules. Afin d'approcher et de prédire au mieux la réponse réelle observée, les valeurs assignées aux cellules de la grille doivent refléter au mieux les propriétés réelles associées aux volumes des cellules. Les propriétés réelles n'étant généralement pas connues sur l'ensemble du domaine, il est nécessaire d'utiliser des mesures indirectes du réservoir afin d'obtenir un modèle proche de la réalité. Cet aspect est traité dans la Section 1.3 et les Chapitres 4 et 5. Dans ce qui suit, nous considérons que les propriétés sont connues et nous nous intéressons aux modèles directs de simulations des fluides et de la réponse sismique.

1.2.1 Simulation de la réponse sismique

Les données sismiques permettent de caractériser les propriétés élastiques de la roche par l'analyse des temps de propagation, des chemins, des réflexions et des transformations des ondes sismiques. Ces ondes sont créées artificiellement soit par des canons à air en milieu aquatique, soit par des explosions ou des camions vibreurs en terrestre. Elles se propagent dans le sol, se réfléchissent aux interfaces des différentes formations et sont ensuite récupérées grâce à des capteurs appelés « géophones » ou « hydrophones ». Le signal brut (sans traitement) est alors constitué des temps d'arrivée des ondes et de leurs amplitudes (énergie) en fonction du temps.

Propagation, réflexion, réfraction et conversion des ondes sismiques

On distingue généralement deux types d'ondes utilisées en interprétation sismique : les ondes P (primaires) et les ondes S (secondaires). Les ondes P, appelées également ondes de compression, accompagnent leur passage par des dilatations et des compressions des roches du sous-sol parallèles à la direction de propagation. Comme l'indique leur nom, les ondes P sont plus rapides que les ondes S. Ces dernières, aussi appelées ondes de cisaillement, accompagnent leur passage par des mouvements perpendiculaires à la direction de propagation. Les ondes S ne se propagent pas dans les milieux liquides. Pour cette raison, elles ne sont généralement pas récupérées lors des acquisitions sismiques marine. Néanmoins, leurs vitesses de propagation peuvent être déduites à partir des ondes P par analyse du signal en fonction de l'angle d'incidence des ondes (cf. équation 1.3 ci-après).

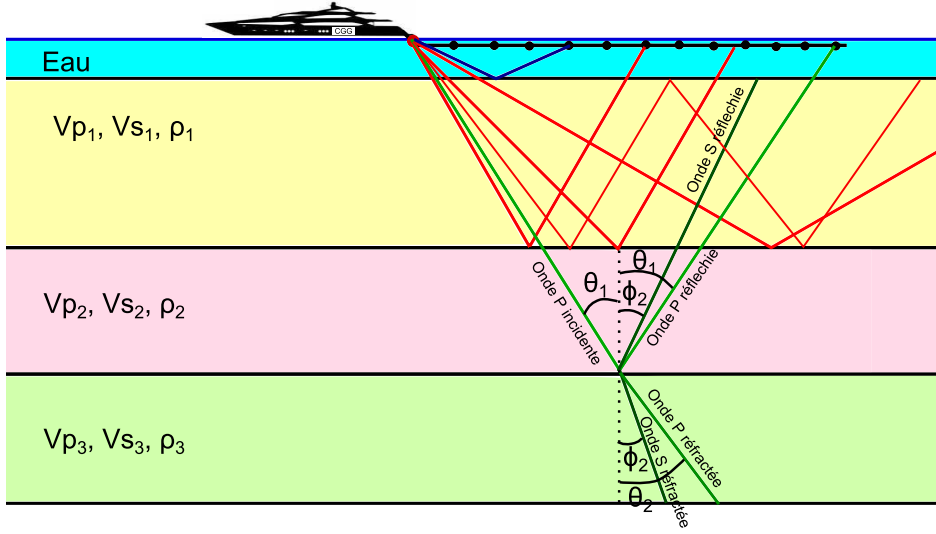


FIGURE 1.1 – Acquisition sismique et schéma de réflexion/réfraction des ondes. Acquisition sismique marine : les ondes S ne sont pas récupérées (pas de capteur au fond de l'eau). Les rayons rouges ne représentent que des réflexions directes (et une multiple). Le rayon vert illustre le processus de conversion des ondes P en ondes S (voir équation 1.3 de Zoeppritz).

Leurs vitesses sont données par les formules suivantes [62, 16] :

$$V_p = \sqrt{\frac{K + 4/3\mu}{\rho}} \quad (1.1)$$

$$V_s = \sqrt{\frac{\mu}{\rho}} \quad (1.2)$$

où K représente le module de compressibilité du milieu (roche + fluide), μ le module de cisaillement et ρ la densité du milieu. Ces paramètres élastiques sont liées aux types de lithologies présentes dans le milieu (cf. modèle pétro-élastique ci-après). Lorsque les ondes rencontrent des interfaces de propriétés, elles se réfléchissent et se réfractent de la même manière que la lumière (figure 1.1). C'est grâce à ce phénomène qu'il est possible de récupérer les ondes envoyées et de caractériser les événements géologiques constitutifs du sous-sol.

Les équations de Zoeppritz forment un modèle couramment utilisé en sismique qui décrit les phénomènes de réflexion, réfraction et conversion des ondes à une interface (figure 1.1). Loin de la source, le front d'onde peut être considéré comme plan et l'on appelle « chemin du rayon » (*raypath*) la propagation qui est perpendiculaire à ce front. Lorsqu'une onde P vient heurter l'interface entre deux couches, elle génère des ondes P et S réfléchies et réfractées. Les équations de Zoeppritz¹ permettent alors de quantifier l'énergie répartie entre ces différentes ondes, en fonction de l'angle d'incidence et du contraste des propriétés

¹ici le cas isotrope

élastiques. La forme matricielle des équations est donnée par [143] :

$$\begin{bmatrix} R_p \\ R_s \\ T_p \\ T_s \end{bmatrix} = \begin{bmatrix} -\sin \theta_1 & -\cos \phi_1 & \sin \theta_2 & \cos \phi_2 \\ \cos \theta_1 & -\sin \phi_1 & \cos \theta_2 & -\sin \phi_2 \\ \sin 2\theta_1 & \frac{V_{p1}}{V_{s1}} \cos 2\theta_1 & \frac{\rho_2 V_{s2}^2}{\rho_1 V_{s1}^2} \cos 2\phi_1 & \frac{\rho_2 V_{s2} V_{p1}}{\rho_1 V_{s1}^2} \cos 2\phi_2 \\ -\cos 2\phi_1 & \frac{V_{s1}}{V_{p1}} \sin 2\phi_1 & \frac{\rho_2 V_{p2}}{\rho_1 V_{p1}} \cos 2\phi_2 & \frac{\rho_2 V_{s2}}{\rho_1 V_{p1}} \sin 2\phi_2 \end{bmatrix}^{-1} \begin{bmatrix} \sin \theta_1 \\ \cos \theta_1 \\ \sin 2\theta_1 \\ \cos 2\phi_1 \end{bmatrix} \quad (1.3)$$

où R_p , R_s correspondent aux coefficients de réflexion des ondes P et S ; T_p , T_s correspondent aux coefficients de réfraction (transmission) des ondes P et S, et où θ_1 , θ_2 , ϕ_1 , ϕ_2 sont respectivement, l'angle d'incidence des ondes P, l'angle de réflexion des ondes S converties, l'angle de réfraction des ondes P et l'angle de réfraction des ondes S converties.

Les ondes réfractées sont principalement utilisées en génie civil où elles permettent de caractériser des structures peu profondes (<300m). On parle alors de sismique réfraction. La sismique réfraction est uniquement applicable lorsque la vitesse des couches augmente avec la profondeur. Pour ces raisons, elle est très peu utilisée en exploration d'hydrocarbures.

La sismique utilisée dans l'industrie pétrolière s'intéresse principalement aux réflexions des ondes P aux interfaces. On parle alors de sismique réflexion. Les ondes S sont rarement enregistrées en acquisition marine car elles nécessitent la mise en place de dispositifs sous-marins, les ondes S ne se propageant pas dans les milieux fluides. On ne s'intéresse alors qu'aux termes des équations de Zoeppritz relatifs à la réflexion des ondes P. Il existe plusieurs approximations linéaires de l'équation de Zoeppritz pour la réflexion des ondes P [19, 113, 112]. Entre autre, elles permettent d'avoir une compréhension plus intuitive de l'effet d'ouverture de l'angle d'incidence sur la réponse sismique (AVA, *Amplitude Versus Angle*). Par exemple, l'approximation de Fatti [53] est donnée par l'équation 1.4 :

$$\begin{aligned} R_p(\theta) &= (1 + \tan^2 \theta) \frac{\Delta I_p}{2I_p} - 8 \left(\frac{V_s}{V_p} \right)^2 \sin^2 \theta \frac{\Delta I_s}{2I_s} \\ &\quad - \left[\frac{1}{2} \tan^2 \theta - 2 \left(\frac{V_s}{V_p} \right)^2 \sin^2 \theta \right] \frac{\Delta \rho}{\rho} \\ &= A \frac{\Delta I_p}{2I_p} + B \frac{\Delta I_s}{2I_s} + C \frac{\Delta \rho}{\rho} \end{aligned} \quad (1.4)$$

où I_p , I_s sont les impédances des ondes P et S, ρ la densité du milieu, $\Delta I = I_2 - I_1$ la différence d'impédance aux interfaces de milieux différents. Ainsi, le coefficient de réflexion étant directement associé à l'amplitude de la sismique, il est possible de trouver les propriétés élastiques de l'interface dès lors qu'il est possible de discriminer les réflexions sismiques provenant d'au moins trois angles différents (cf. Chapitre 5).

Traitement du signal sismique brut

Le traitement est une étape importante de l'imagerie sismique. Il a pour but d'enlever le bruit contenu dans les données et de relier les données aux événements géologiques présents dans le sous-sol afin de pouvoir être interprétés par un géologue. Il existe de nombreuses corrections, mais il est classique de distinguer les étapes majeures suivantes :

1. **Correction de l'amplitude** : l'amplitude du signal sismique varie considérablement au cours de l'enregistrement, reflétant la dissipation de l'énergie du front d'onde avec la distance à la source due à l'expansion du front d'onde, l'absorption préférentielle des hautes fréquences et l'atténuation par réfractions. Cette correction a pour but de ramener le signal à une amplitude quasi-constante.
2. **Sommation du signal** (*CMP stacking*) : un même événement (appelé *Common MidPoint* ou CMP) est souvent illuminé par plusieurs rayons sismiques ayant des angles d'incidences différents. Il est alors possible, pour une gamme d'angles donnée, de moyenner l'ensemble des signaux sismiques relatifs au même CMP, ce qui permet de diminuer le bruit par rapport au signal. Néanmoins, une sommation trop large conduira à la perte d'informations importantes sur la relation entre l'amplitude sismique et les angles d'incidence, ce qui empêche toute analyse AVO². Plusieurs corrections doivent être appliquées avant cette phase de sommation, telles que la NMO (*Normal Move Out*), les corrections statiques ou résiduelles, ...
3. **Déconvolution** : cette étape permet de déduire des séries de réflectivités en déconvoluant le signal sismique par un filtre que l'on appelle ondelette. L'ondelette est un filtre représentatif de l'énergie et des fréquences portées par le signal sismique enregistré. Ainsi, il est possible d'obtenir un sismogramme $s(t)$ par convolution (opérateur $*$) de coefficients de réflexions $r(t)$ et d'une ondelette donnée, $w(t)$ (figure 1.2) :

$$s(t) = r(t) * w(t) \quad (1.5)$$

Inversement, lorsque la trace sismique est connue, il est possible d'obtenir les coefficients de réflexion $r(t)$ par inversion ou déconvolution. La déconvolution est utilisée pour éliminer les démultiplications du signal, tels que les réflexions multiples, les fantômes et réverbérations, ainsi que pour augmenter la résolution du signal en altérant la forme et le contenu fréquentiel de l'ondelette. Il existe plusieurs méthodes de déconvolution (déconvolution prédictive, homomorphique, d'entropie maximum ou minimum, ...) qui peuvent être appliquées avant ou après l'étape de sommation des traces (*post-* ou *pre-stack*).

4. **Migration** : ce processus a pour but de relier les événements observés par la sismique, aux événements géologiques présents dans le sous-sol.

²Amplitude Versus Offset

La migration peut être faite en temps ou en profondeur, avant ou après sommation.

Plus d'informations sur le traitement sismique peut être trouvées dans [98, 139].

L'interprétation quantitative des structures ou des propriétés est généralement basée sur l'analyse d'un signal traité. Il est important de garder à l'esprit que les approximations et les choix qui ont été faits durant le traitement ont un impact sur l'interprétation des données sismiques.

Méthodes numériques de modélisation de la réponse sismique par convolution 1D

Pour un jeu de paramètres donné (p. ex. vitesses, densités), le modèle direct doit générer un signal sismique (sismogramme) synthétique. Le signal produit peut soit reproduire une réponse sismique brute, soit une sismique après traitement.

Le premier cas, appelé modélisation complète de la forme de l'onde (*full waveform*), permet de prendre en compte plus de physique (c.-à-d. les réflexions, conversions, réfractions, atténuations) et des géométries plus complexes. Néanmoins, elle est extrêmement coûteuse en temps de calcul et n'est actuellement pas utilisée dans les études opérationnelles.

La seconde approche prend uniquement en compte les réflexions directes et repose (généralement) sur une hypothèse d'horizontalité des interfaces. Cette approche est souvent utilisée afin d'interpréter quantitativement les propriétés d'un réservoir. Dans ce travail de thèse, nous utilisons un modèle de convolution 1D de simulation de sismique traitée.

Paramètres d'entrée et de sortie Les paramètres d'entrée nécessaires au modèle direct par convolution 1D utilisé dans cette thèse sont :

- une grille stratigraphique en temps, échantillonnée régulièrement latéralement au pas sismique (25m ou 12.5m), mais verticalement irrégulière (cellules d'épaisseur variables). (p. ex. Chapitre 5 Section 5.2.2) ;
- des champs de propriétés élastiques (c.-à-d. Vp , Vs , ρ ou impédances $Ip = \rho Vp$, $Is = \rho Vs$ correspondantes) ;
- une ondelette 1D estimée à partir de la sismique traitée et des puits, pour chaque réponse simulée par gamme d'angles d'incidences (p. ex. Chapitre 5 figure 5.6) ;
- un modèle de physique des roches (voir Section 1.2.1).

Le modèle direct produit en sortie, pour chaque gamme d'angles, un cube d'amplitudes échantillonné régulièrement dans les trois directions (latéralement au pas sismique, verticale en fonction du contenu fréquentiel des ondelettes utilisées), représentatif des premières réflexions des ondes sismiques (p. ex. Chapitre 5 figure 5.7).

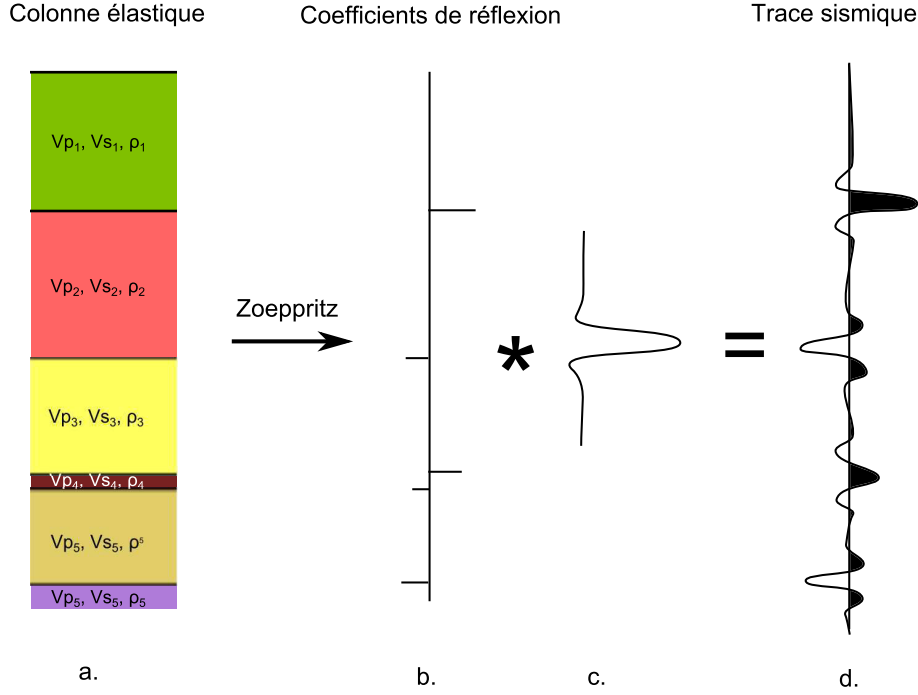


FIGURE 1.2 – Convolution sismique 1D. (a) Séquence de milieux élastiques différents. (b) Coefficients de réflexion résultants. (c) Ondelette. (d) Signal sismique généré. Le contenu fréquentiel de l'ondelette étant limité, la couche fine numéro 4 ne peut pas être caractérisée par la sismique.

Méthode par convolution 1D Cette méthode de simulation considère que chaque trace sismique 1D (séquences verticales des amplitudes pour une position latérale donnée) peut être simulée indépendamment, et que toutes les interfaces traversées sont horizontales. À partir des propriétés élastiques, il est possible de calculer, pour chaque gamme d'angles, les coefficients de réflexion associés aux interfaces en utilisant les modèles mathématiques présentés dans la Section 1.2.1 (figure 1.2-b). Les amplitudes sont ensuite générées en convoluant chaque séquence avec les ondelettes (figure 1.2-d).

Cette méthode est très efficace et peut être facilement parallélisable, ce qui en fait un modèle direct idéal pour les processus itératifs d'inversion sismique (cf. Section 1.3.1).

Linéarisation du modèle direct L'équation 1.4 peut être étendue en une fonction continue [125] et exprimée par rapport aux impédances I_p et I_s :

$$\begin{aligned} Rp(\theta) &= \frac{1}{2\cos^2\theta} \Delta \ln(I_p) - 4 \frac{I_s^2}{I_p^2} \sin^2\theta \Delta \ln(I_s) \\ &= A(\theta) \Delta \ln(I_p) + B(\theta) \Delta \ln(I_s) \end{aligned} \quad (1.6)$$

En regroupant les coefficients $A(\theta)$ et $B(\theta)$ en une matrice $((n-1) \times 2(n-1))$, \mathbf{A}_θ , il est possible pour un vecteur $(2n \times 1)$ de logarithmes d'impédances \mathbf{m} de

taille $2n$, de calculer un vecteur $((n-1) \times 1)$ de réflectivités \mathbf{r}_θ tel que :

$$\mathbf{r}_\theta = \mathbf{A}_\theta \mathbf{D} \mathbf{m} \quad (1.7)$$

où \mathbf{D} est la matrice $(2(n-1) \times 2n)$ de différence finie qui transforme le vecteur \mathbf{m} en vecteur de variation de logarithmes. Finalement, il est possible d'exprimer l'équation 1.5 de manière discrète comme le produit matriciel du vecteur de réflectivités avec une matrice $((n \times n_\theta) \times (n-1))$, \mathbf{W} , définie par les ondelettes associées aux n_θ angles θ et les positions des interfaces. Ainsi, les amplitudes sismiques, \mathbf{s}_θ peuvent être calculées à l'aide d'une expression linéaire fonction du logarithme des impédances :

$$\begin{aligned} \mathbf{s}_\theta &= \mathbf{W} \mathbf{A}_\theta \mathbf{D} \mathbf{m} \\ &= \mathbf{G}_s \mathbf{m}, \end{aligned} \quad (1.8)$$

où \mathbf{G}_s représente le modèle direct linéarisé de simulation des amplitudes sismiques.

Physique des roches et propriétés élastiques

On appelle « modèle pétro-élastique » ou PEM, un modèle mathématique permettant d'estimer les propriétés élastiques des roches à partir de ses propriétés pétro-physiques et de son contenu fluide. Ce modèle repose sur un ensemble de relations mathématiques (empiriques ou dérivées de la physique) permettant de décrire le comportement élastique d'une roche en fonction de sa composition et de sa structure interne.

Bien que l'on puisse trouver un certain nombre de PEM dans la littérature [7], il est généralement conseillé d'utiliser des relations spécifiques pour chaque réservoir étudié. Des expériences en laboratoire peuvent être effectuées afin d'établir un PEM précis et discriminer les éléments ayant un impact important sur le comportement élastique de la roche, tels que la composition minéralogique, la forme des grains et leur agencement, la cimentation ou la proportion d'argile [7]. Un exemple de PEM est donné dans le Chapitre 5 Section 5.2.2.

Un PEM est généralement une application subjective, c.-à-d que pour un vecteur de paramètres élastiques, il existe plusieurs combinaisons possibles de paramètres caractérisant la structure de la roche (différentes combinaisons de porosité, cimentation, fluides, etc...). Ainsi, une même réponse élastique peut s'expliquer par plusieurs états différents. L'interprétation des propriétés élastiques en termes de propriétés pétrophysiques est donc un problème inverse mal posé (voir Section 1.3.1).

1.2.2 Simulation de la production de fluide

Un réservoir pétrolier est formé d'un ensemble de roches sédimentaires dont les interstices (porosité) sont remplis par différents fluides (eau, huile, gaz) à une pression proportionnelle à son enfouissement, initialement ségrégués en fonction de leurs densités respectives. La production des hydrocarbures dépend de la mobilité des fluides piégés dans le réservoir. Lorsqu'un puits est foré et que des hydrocarbures sont extraits, la pression du réservoir est localement diminuée ce

qui provoque un mouvement des fluides vers le puits, proportionnellement au gradient de pression (cf. loi de Darcy).

Le mouvement des fluides est à la fois contrôlé par des effets microscopiques (tels que la viscosité des fluides, les tensions aux interfaces) et macroscopiques (tels que les lithologies, le réseau fracturé, les failles). Dans ce travail de thèse, nous nous intéressons plus particulièrement aux effets macroscopiques des variations spatiales des propriétés pétrophysiques des réservoirs. Deux propriétés sont fondamentales dans tous les modèles réservoirs utilisés pour la simulation des écoulements : la porosité de la roche, ϕ , qui est une quantité sans dimension représentant la fraction d'espace vide disponible pour le remplissage par les fluides et la perméabilité, K , qui mesure la capacité de la roche à transmettre les fluides. La perméabilité (communément exprimée en Darcy³), est généralement corrélée à la porosité. Elle peut être anisotrope lorsque la structure interne de la roche possède une orientation privilégiée de connections des pores ou que des structures de plus grande échelle privilégient les mouvements des fluides dans certaines directions. Dans ce cas, la perméabilité est exprimée par un tenseur. L'estimation des propriétés d'un réservoir est décrite dans les Chapitres 3 et 4.

L'écoulement des fluides dans un réservoir est décrit par un modèle mathématique comprenant un ensemble de lois physiques fondamentales qui dépendent (en partie) de ces variables. Dans ce qui suit, nous présentons les modèles les plus communs d'écoulements isothermes, sans effets géomécaniques.

Modèles mathématiques

Écoulements à deux phases. Dans les réservoirs d'hydrocarbures, l'espace poreux est généralement rempli par deux ou plusieurs fluides non miscibles, c.-à-d qu'ils forment deux phases distinctes séparées par une interface. L'eau est presque toujours présente dans les réservoirs (aquifères) et l'injection d'eau est couramment utilisée pour améliorer la récupération des hydrocarbures. Bien que physiquement séparés à l'échelle microscopique, les fluides sont supposés présents à la même localisation à l'échelle macroscopique. Pour un système eau-huile de volume poreux, on appelle saturation en eau, S_w , et huile, S_o , les fractions volumiques occupées par chaque phase, avec $S_w + S_o = 1$. Chacune des phases p respecte la loi de conservation de la masse :

$$\frac{\partial \rho_p S_p \phi}{\partial t} + \nabla \cdot (\rho_p \vec{v}_p) = q_p, \quad (1.9)$$

où ρ_p correspond à la densité (kg/m^3) de la phase p , \vec{v} à la vitesse de filtration (m/s) et q_p (m^3) correspond au terme d'injection/production de la phase p dans le système.

Lorsque plusieurs phases sont présentes dans le système, la vitesse de filtration d'une phase est affectée par les interactions avec les autres phases. Ce phénomène physique est modélisé à l'échelle macroscopique par les perméabilités relatives, $kr_p(S_p)$, qui dépendent des saturations. Ces dernières modifient la

³En unité SI, 1 Darcy $\approx 0.987 \times 10^{-12} m^2$. La définition d'une unité Darcy implique la transmission d'un fluide de viscosité égale à 1 cp à travers une roche homogène à la vitesse de 1 cm/s sous un gradient de pression de 1 atm/cm à travers une surface de 1 cm^2

valeur de la perméabilité absolue K , afin de tenir compte des interactions entre les fluides. La vitesse d'écoulement de la phase p est alors reliée à la variation de pression par la loi de Darcy :

$$\vec{v}_p = -\frac{Kkr_p}{\mu_p}(\nabla P_p - \rho_p \vec{g}), \quad (1.10)$$

où μ_p représente la viscosité dynamique de la phase p , P_p sa pression et \vec{g} le vecteur de gravité. Due aux tensions aux interfaces grains-fluide, la pression diffère pour chacune des phases. La différence de pression, appelée pression capillaire $P_{c_p}(S_p)$, $p = \{w, o\}$, est également fonction des saturations.

Les équations 1.9 et 1.10 peuvent être par la suite reformulées afin d'explicitier les équations de pression et de saturation. Dans le cas de fluides non miscibles et incompressibles, l'équation elliptique de pression est donnée par :

$$\nabla \cdot \vec{v} = q, \quad \vec{v} = -K(\lambda_o + \lambda_w)\nabla P + K(\lambda_w \rho_w + \lambda_o \rho_o)\vec{g}, \quad (1.11)$$

où les indices (w, o) représentent respectivement la phase la plus mouillante (généralement l'eau) et l'huile, $\vec{v} = \vec{v}_w + \vec{v}_o$, la vitesse totale, q , le flux volumique global et $\lambda_p = \frac{kr_p}{\mu_p}$ la mobilité de la phase p . L'équation parabolique de saturation est donnée par :

$$\phi \frac{\partial S_w}{\partial t} + \nabla \cdot [f_w \vec{v} + f_w K \lambda_o (\rho_w - \rho_o) \vec{g} + f_w K \lambda_o \nabla P_{c_w}] = \frac{q_w}{\rho_w}, \quad (1.12)$$

où $f_w = \frac{\lambda_w}{\lambda_w + \lambda_o}$ représente le débit fractionnaire (*fractional-flow*) fonction de la saturation en eau (S_w). Les équations 1.11 et 1.12 sont non linéaires et peuvent être résolues numériquement de façon séquentielle : la pression et la vitesse sont calculées dans un premier temps en utilisant l'équation 1.11, et sont ensuite considérées connues dans l'équation 1.12 tout en incrémentant d'un pas de temps.

Écoulements multiphasiques et multi-composants. Les équations décrivant les écoulements à deux phases sont facilement généralisables à plusieurs phases, mais l'obtention des perméabilités relatives devient plus difficile. Chaque phase peut être constituée d'un à plusieurs composants, et les lois de conservations peuvent être exprimées pour chaque composant. En fonction de certaines conditions (PVT, équilibre des phases), les composants peuvent migrer d'une phase à l'autre ou former une nouvelle phase (cf. modèle *black-oil*), ce qui modifie le comportement d'écoulement des fluides. En fonction du type de réservoir étudié et des mécanismes de récupération, il peut être nécessaire de modéliser ces comportements, ce qui complique le modèle.

Le modèle *black-oil* Ce modèle est largement utilisé dans l'industrie pétrolière pour les réservoirs conventionnels. Il décrit le comportement de 2 à 3 phases en fonction des conditions de pression et de température. En condition de surface, les hydrocarbures sont constitués de 2 phases distinctes : l'huile et le gaz dont les compositions chimiques restent constantes. Aux conditions du

réservoir, le gaz peut être partiellement ou complètement dissout dans l'huile⁴. Le modèle *black-oil* est défini par les lois de conservation suivantes :

$$\frac{\partial}{\partial t} \left(\frac{\phi \rho_p^s}{B_p} S_p \right) + \nabla \cdot \left(\frac{\rho_p^s}{B_p} \vec{v}_l \right) = q_p \quad (1.13)$$

$$\frac{\partial}{\partial t} \left(\frac{\phi \rho_g^s}{B_g} S_g + \frac{\phi R s_o \rho_g^s}{B_o} S_o \right) + \nabla \cdot \left(\frac{\rho_g^s}{B_g} \vec{v}_g + \frac{R s_o \rho_g^s}{B_o} \vec{v}_l \right) = q_g \quad (1.14)$$

où $B_p = V_p/V_p^s$ correspond au facteur de volume qui est le ratio des volumes occupés aux conditions de surface et de réservoir par une phase p , \vec{v}_l et \vec{v}_g correspondent aux vitesses d'écoulement pour les phases fluides et gazeuse, $R s_o$ est le coefficient de solubilité du gaz dans l'huile et où l'exposant s signifie que la variable associée est exprimée aux conditions de surface.

Modèle de puits Le but du modèle de puits est de représenter l'écoulement des fluides dans la cavité du puits. Dans sa forme la plus simple, un puits peut être considéré comme une cavité verticale, cimentée et perforée à des intervalles spécifiques le reliant au réservoir, par lesquels sont produits ou injectés des fluides. La taille des perforations étant très petite par rapport aux cellules du modèle discret intersectées par le puits, il est difficile de résoudre numériquement les équations de façon précise, et des solutions analytiques ou semi-analytiques sont utilisées pour modéliser la répartition du flux [106].

Résolution numérique

Discretisation des équations Les équations du modèle mathématique (Section 1.2.2) comportant à la fois des dérivées partielles en temps et en espace, elles sont discrétisées dans ces deux domaines afin de résoudre le problème numériquement. Il existe de nombreuses méthodes de discrétisation [49], mais les méthodes de volumes finis sont généralement utilisées dans l'ingénierie des réservoirs. Le principe général de ces méthodes est de définir un volume de contrôle et d'imposer la conservation des équations sur ce volume (voir [56] pour plus d'informations). On distingue deux grands types de schéma numériques temporels :

- les méthodes explicites : ces méthodes calculent l'état d'un système pour un temps $t + 1$ à partir de son état courant au temps t . Ce type de schéma est très rapide, mais peut être instable lorsqu'un pas de temps trop grand est utilisé. De plus, il accumule les erreurs d'un pas de temps à l'autre.
- les méthodes implicites : ces méthodes calculent l'état d'un système pour un temps $t + 1$ à partir de son état courant t ainsi que son état au temps $t + 1$. Ce schéma nécessite la résolution d'un problème inverse (inversion de matrices) ce qui le rend moins efficace que les méthodes explicites, mais il est beaucoup plus stable et numériquement précis.

⁴Il existe des modèles plus généraux dans lesquels l'huile peut être partiellement dissoute dans le gaz, les hydrocarbures peuvent être dissouts dans l'eau et l'eau peut être dissoute dans les hydrocarbures

La plupart des simulateurs commerciaux utilisent des schémas entièrement implicites pour résoudre les équations de la Section 1.2.2. Cependant, il est parfois possible d'utiliser des méthodes hybrides. Par exemple, la méthode IMPES (implicite en pression, explicite en saturation) résout l'équation de pression implicitement et utilise ensuite les résultats pour résoudre explicitement l'équation de saturation. Plus de détails sur la discrétisation et les méthodes de résolution peuvent être trouvés dans [26].

Conditions aux limites Afin d'obtenir une solution unique, il est nécessaire de définir les conditions aux limites du modèle discret. Des contraintes de flux (condition de Dirichlet) ou de pression (condition de Neumann) sont imposées aux frontières du réservoir et aux puits.

Paramètres d'entrée et de sortie

Les paramètres d'entrée nécessaires à la simulation de la production des fluides d'un réservoir incluent (liste non exhaustive) :

- une grille discrète adaptée à la géométrie du réservoir et remplie de propriétés pétrophysiques (p. ex. porosité, *net-to-gross*) ;
- pour chaque face des cellules du modèle discret, les transmissibilités déduites du tenseur de perméabilité, de la transmissibilité des failles ou du réseau de fractures ;
- la localisation, le schéma de production et les paramètres de contrôle des puits ;
- les conditions initiales du réservoir (telles que la saturation ou les contacts fluides) et les conditions aux limites ;
- un modèle d'aquifère ;
- les paramètres des fluides (p. ex. la viscosité) et les paramètres d'interaction fluide-fluide et fluide-roche (p. ex. les perméabilités relatives).

Les données simulées incluent (liste non exhaustive) :

- champs des variations de la saturation, de la pression, de la température, ... dans le réservoir ;
- les variations de composition des phases ;
- la production des fluides et les pressions aux puits.

Dans ce travail de thèse nous nous intéressons particulièrement aux simulations de la production aux puits.

1.2.3 Utilisation de modèles de substitution

Lorsqu'un grand nombre de simulations est nécessaire, par exemple lors de la résolution de problèmes inverses et/ou de la quantification des incertitudes (Section 1.3), il est possible d'utiliser un modèle direct approximatif mais

plus rapide. Par exemple, le temps de simulation des écoulements peut aller de quelques minutes à quelques jours en fonction de la taille du modèle et du schéma de production. Dans ce cas, seul un nombre très limité d'évaluations est possible. De plus, dans certaines études de quantification des incertitudes, il n'est pas nécessaire d'obtenir une réponse exacte.

L'annexe 6.2 présente la création d'un modèle approximatif de production par déplétion naturelle, basé sur la connectivité dynamique des propriétés pétrophysiques du réservoir. Ce travail a été développé en collaboration avec Pierre Monamicq dans le cadre de son projet de Master 2. Ce modèle a pour but de classer des réalisations issues de processus de simulation géostatistique (Chapitre 3) ou d'inversion stochastique (Chapitre 5) et ne nécessite pas une réponse précise des écoulements. De plus, le modèle approximatif développé est linéaire, ce qui facilite certains développements théoriques et peut permettre dans certains cas d'exprimer analytiquement la solution d'un problème donné (p. ex. densité de probabilité).

D'une manière générale, les modèles approximatifs les plus simples ne sont valides que pour un cas précis, et plus le modèle est robuste, plus le temps d'exécution est important. Les avantages et inconvénients relatifs à l'utilisation de tels modèles par rapport aux simulations à physique complète est traité dans [144].

1.3 Inversions des données de réservoir

S'il est possible de connaître ponctuellement les propriétés des réservoirs grâce aux observations directes de puits, la grande majorité du domaine étudié reste inconnu. Néanmoins, les données présentées dans la section précédente (c.-à-d. les amplitudes sismiques et les courbes de productions de puits) étant conditionnées par la structure et les propriétés du réservoir, elles permettent de les caractériser indirectement. En effet, puisque la réponse physique peut être simulée (Section 1.2) pour un ensemble de paramètres donnés, il est possible d'obtenir des informations sur les propriétés du réservoir en comparant les données simulées aux données réelles. On parle alors de problème inverse.

Dans cette section, nous introduisons dans un premier temps la notion de problème inverse et les méthodes de résolution. Le cas particulier de l'optimisation convexe dans le cadre gaussien est par la suite détaillé dans la Section 1.3.2. Enfin, les méthodes de quantification des incertitudes, et en particulier les méthodes d'ensemble utilisées dans ce travail de thèse (type filtres de Kalman), sont présentées dans la Section 1.3.3.

1.3.1 Problème inverse et méthodes d'optimisation

Modèle direct

La notion de modèle direct est présentée dans la Section 1.2, dans laquelle sont introduits les modèles de simulation des amplitudes sismiques et des données de production. Nous rappelons qu'un modèle direct, $g(\cdot)$, est un ensemble d'équations permettant de simuler un phénomène physique à partir de paramètres d'entrée, \mathbf{m} , décrivant les propriétés et l'état d'un système à un temps

donné. Ainsi, il est possible de simuler un phénomène, \mathbf{d}^{sim} , tel qu'il aurait pu être observé dans les conditions réelles décrites par les paramètres \mathbf{m} [131] :

$$\mathbf{d}^{sim} = g(\mathbf{m}) \quad (1.15)$$

Le modèle direct doit en principe être un problème bien posé continu, c.-à-d. que pour un vecteur de paramètre $\mathbf{m} \in \Omega$ défini dans l'espace des paramètres possibles, le modèle direct doit avoir une solution unique et que cette solution est une fonction continue des données du problème. Cependant, il arrive en pratique que pour un vecteur de paramètres $\mathbf{m} \in \Omega$, le simulateur numérique n'arrive pas à converger (problème de discrétisation, ...).

Lorsque le modèle direct est une application linéaire (voir Section 1.2.1), \mathbf{G} , l'équation 1.15 s'écrit :

$$\mathbf{d}^{sim} = \mathbf{G}\mathbf{m} \quad (1.16)$$

On pourra noter qu'il est alors possible, lorsque la matrice \mathbf{G} est inversible, d'obtenir analytiquement les paramètres \mathbf{m} à partir des observations \mathbf{d}^{obs} .

Données observées

On appelle « données observées », l'observation d'un phénomène physique, à une position donnée, provoqué par la stimulation d'un système étudié. Dans les applications réelles, les observations sont souvent contaminées par du bruit. Les données observées, \mathbf{d}^{obs} , peuvent alors être exprimées sous la forme :

$$\mathbf{d}^{obs} = g(\mathbf{m}^{real}) + \epsilon \quad (1.17)$$

où ϵ représente le bruit d'observation et \mathbf{m}^{real} correspond aux paramètres réels. Il est important de remarquer que nous considérons que le modèle direct est capable de simuler parfaitement la réponse physique et que le milieu réel peut être parfaitement décrit par les paramètres \mathbf{m}^{real} . Bien évidemment, cette hypothèse est irréaliste. Pour cette raison, nous considérerons que les erreurs de modélisation et de discrétisation sont incluses dans le vecteur ϵ .

Il est difficile d'obtenir un modèle cohérent de bruit⁵. De ce fait, une distribution gaussienne est généralement utilisée pour modéliser le bruit :

$$\epsilon \sim \mathcal{N}(0, \mathbf{C}_D) \quad (1.18)$$

où \mathbf{C}_D est une matrice de covariance souvent considérée diagonale (modèle de bruit indépendant).

Ainsi, la probabilité de mesurer des données \mathbf{d}^{obs} connaissant l'ensemble des paramètres, \mathbf{m} du système est donnée par la relation :

$$p(\mathbf{d}^{obs}|\mathbf{m}) = p(\epsilon = \mathbf{d}^{obs} - g(\mathbf{m})) \propto \exp \left[-\frac{1}{2}(\mathbf{d}^{obs} - g(\mathbf{m}))^T \mathbf{C}_D^{-1}(\mathbf{d}^{obs} - g(\mathbf{m})) \right] \quad (1.19)$$

où T est la transposée du vecteur des écarts entre les données observées et simulées $\mathbf{d}^{obs} - g(\mathbf{m})$, et $p(\mathbf{d}^{obs}|\mathbf{m})$ la probabilité conditionnelle des observations

⁵lorsque le bruit peut être modélisé, il est généralement enlevé des données

simulées conditionnée par le modèle **m**.

Notion de problème inverse

Les problèmes inverses [131] ont pour but de déduire les causes (paramètres du système) à partir des effets d'un phénomène observé (données observées) et sont opposés aux problèmes directs qui simulent les effets en connaissant les causes (lois physiques). La résolution d'un problème inverse dépend donc en partie de la capacité du modèle direct à simuler correctement le phénomène étudié.

Les problèmes inverses sont pratiquement toujours mal posés, c.-à-d. qu'il existe une infinité de solutions pouvant reproduire les données observées. Ceci peut être dû à un nombre insuffisant de données indépendantes par rapport au nombre de paramètres à résoudre ou parce que les données observées sont bruitées et incohérentes avec le modèle direct⁶.

La question se pose alors : comment choisir un modèle parmi l'ensemble des modèles possibles ? Pour répondre à cette question, nous pouvons décrire le problème d'un point de vu probabiliste en utilisant l'approche Bayésienne. Ainsi, on peut considérer que chaque modèle est associé à une probabilité d'existence. Il est alors possible de choisir le modèle le plus probable. Par exemple, en régression linéaire, on choisira la droite qui respecte le mieux les données (le critère de respect des données restant à définir, voir Section 1.3.1). Malheureusement, dans la plupart des cas, une infinité de modèles sont possibles du fait du nombre insuffisant des données disponibles. Afin d'obtenir une solution unique, il est possible d'utiliser un modèle *a priori* associé à une fonction de densité de probabilité (PDF). Cette PDF peut ensuite être modifiée en intégrant les données observées bruitées. La PDF résultante, appelée fonction de densité de probabilité *a posteriori*, représente alors la solution du problème inverse.

En pratique, il est plus facile de travailler sur des modèles. Il est possible de choisir le modèle le plus probable (maximum *a posteriori* ou MAP), ou bien de travailler sur un ensemble de modèles échantillonnés sur la PDF *a posteriori* afin d'avoir une estimation des incertitudes. Généralement, ces PDF sont quelconques, mais afin de pouvoir exprimer mathématiquement la solution, il est nécessaire d'utiliser des modèles spécifiques. Dans ce qui suit, on considèrera notamment que la PDF *a priori* est une gaussienne multivariée.

Dans ce travail de thèse sont présentés deux exemples de problèmes inverses : l'inversion des données sismiques (Section 1.3.3 et 5.3.1) et l'inversion des données de production (Chapitre 4).

Modèle *a priori*

Le modèle *a priori* décrit les connaissances que l'on a du réservoir avant l'intégration des données à inverser. Cependant, le modèle *a priori* peut lui-même provenir d'un processus d'inversion : des modèles dérivés de l'inversion sismique peuvent être utilisés comme modèles *a priori* pour l'assimilation des données de production. Lorsqu'il s'agit de la première intégration de données,

⁶Le modèle direct n'est le plus souvent lui-même qu'une approximation des phénomènes physiques étudiés

le modèle *a priori* est généralement construit par méthodes géostatistiques en utilisant les observations directes de puits (c.-à-d. diagraphies, carottes ; voir Chapitre 3).

Bien que de manière générale la distribution de probabilité *a priori* soit non gaussienne, il est souvent nécessaire de faire cette hypothèse. Le modèle *a priori* est alors défini par la PDF suivante :

$$f_M(\mathbf{m}) = \frac{1}{\sqrt{(2\pi)^n |\mathbf{C}_M|}} \exp \left[-\frac{1}{2} (\mathbf{m} - \mathbf{m}_{pr})^T \mathbf{C}_M^{-1} (\mathbf{m} - \mathbf{m}_{pr}) \right] \quad (1.20)$$

où \mathbf{m}_{pr} et \mathbf{C}_M représentent le vecteur des moyennes et la matrice covariance de la distribution *a priori*, et $|\mathbf{C}_M|$ le déterminant de \mathbf{C}_M .

Distribution de probabilité et probabilité maximale

En combinant les équations 1.19 et 1.20 et en utilisant le théorème de Bayes, il est possible d'obtenir l'expression de la distribution *a posteriori* (à une constante de normalisation près) :

$$\begin{aligned} p(\mathbf{m} | \mathbf{d}^{obs}) &\propto p(\mathbf{d}^{obs} | \mathbf{m}) p(\mathbf{m}) \\ &\propto \exp \left[-\frac{1}{2} (\mathbf{d}^{obs} - g(\mathbf{m}))^T \mathbf{C}_D^{-1} (\mathbf{d}^{obs} - g(\mathbf{m})) \right. \\ &\quad \left. - \frac{1}{2} (\mathbf{m} - \mathbf{m}_{pr})^T \mathbf{C}_M^{-1} (\mathbf{m} - \mathbf{m}_{pr}) \right] \\ &\propto \exp \left[-\frac{1}{2} O(\mathbf{m}) \right] \end{aligned} \quad (1.21)$$

avec

$$O(\mathbf{m}) = \frac{1}{2} (\mathbf{d}^{obs} - g(\mathbf{m}))^T \mathbf{C}_D^{-1} (\mathbf{d}^{obs} - g(\mathbf{m})) + \frac{1}{2} (\mathbf{m} - \mathbf{m}_{pr})^T \mathbf{C}_M^{-1} (\mathbf{m} - \mathbf{m}_{pr}) \quad (1.22)$$

Lorsque le modèle direct $g(\cdot)$ est linéaire, il est possible de calculer une solution analytique (Section 1.3.3). Malheureusement, la plupart du temps $g(\cdot)$ est non linéaire et il est nécessaire d'utiliser des méthodes de minimisation afin de générer un ou plusieurs modèles représentatifs de la distribution 1.21.

Le modèle le plus probable (Maximum *a posteriori* ou MAP) est communément utilisé dans les processus d'inversion en géosciences (p. ex. dans les inversions déterministes des données sismiques ou de production). Le MAP est obtenu par minimisation d'une fonction objectif, représentée dans le cas étudié par la fonction 1.22 :

$$\mathbf{m}_{MAP} = \min_{\mathbf{m}} O(\mathbf{m}) \quad (1.23)$$

Le modèle direct étant non linéaire, la fonction objectif $O(\mathbf{m})$ peut avoir plusieurs minima locaux. Dans ce cas, les méthodes de minimisation utilisant la pente locale de la fonction objectif (cf. Gradient conjugué, méthodes de Newton) ne peuvent pas assurer une convergence vers le MAP. De plus, le résultat de l'optimisation est fortement dépendant du modèle *a priori* utilisé.

Dans le cas linéaire, lorsque le nombre de données assimilées est grand, on peut montrer [131] que $O(\mathbf{m})$ suit la loi du χ^2 de moyenne $O(\mathbf{m}_{MAP}) = \frac{n^d}{2}$ et

de variance n^d , où n^d représente le nombre de données assimilées. On peut alors faire l'hypothèse que ce résultat s'applique aux cas non linéaires, tel que :

$$E[O(\mathbf{m}_{MAP})] = \frac{n^d}{2} \quad (1.24)$$

Dans ce travail on considère que tout modèle acceptable se trouve dans un intervalle de confiance de cinq fois l'écart-type autour de la moyenne [105], tel que :

$$n^d - 5\sqrt{2n^d} \leq 2O(\mathbf{m}) \leq n^d + 5\sqrt{2n^d} \quad (1.25)$$

Lorsque le résultat de la minimisation ne remplit pas les conditions de l'équation 1.25, il est probable que l'algorithme ait convergé vers un minimum local.

Méthodes d'optimisation

Le but de l'optimisation est de trouver les valeurs des paramètres d'un modèle qui minimisent (ou maximisent) une fonction objectif. En fonction de la nature du problème, de ses contraintes et de la forme de la fonction objectif, différentes méthodes d'optimisation seront utilisées. Dans ce travail de thèse nous nous intéressons plus spécifiquement aux méthodes d'optimisation non linéaires, continues et non contraintes.

On pourra distinguer ici deux classes de méthodes d'optimisation : les méthodes d'optimisation globale et locale.

Optimisation globale Les méthodes d'optimisation globale ont pour but de trouver le minimum global d'une fonction objectif quelconque. Elles incluent notamment les méthodes stochastiques⁷(p. ex. recuit simulé, méthodes de Monte-Carlo) et les méthodes (méta)heuristiques (p. ex. essaims de particules (*particle swarms*), les algorithmes génétiques ou évolutionnistes). Ces méthodes sont très robustes, mais ne peuvent prendre en compte qu'un nombre réduit de paramètres et exigent un grand nombre d'itérations pour converger, ce qui les rend difficilement applicables lorsque le modèle direct est coûteux en temps de calcul et/ou le problème est de grande dimension. Voir [90] pour plus d'informations.

Optimisation locale Les méthodes locales recherchent un minimum autour d'un modèle initial. Ces algorithmes modélisent localement, à chaque itération, la forme de la fonction objectif afin de mettre à jour le modèle pour qu'il se rapproche du minimum le plus proche. On peut distinguer deux groupes de méthodes : les méthodes basées sur le gradient (ou une approximation), telles que le gradient conjugué, les approximations quadratiques (*Sequential Quadratic Programming* ou SQP), les méthodes de Newton, ou les méthodes BFGS [21, 105], et les méthodes sans gradient, telles que les méthodes de recherche directe (Direct Search methods [109]).

Ces méthodes convergent vers un minimum qui peut être local. Si la fonction objectif est convexe, ce minimum est global. Cependant, les méthodes basées gradient peuvent inclure un grand nombre de paramètres et convergent généralement beaucoup plus rapidement que les méthodes globales. Pour ces raisons,

⁷il est important de différencier les méthodes stochastiques d'optimisation et les approches stochastiques qui ont pour but de caractériser la PDF *a posteriori*

elles sont souvent utilisées en calage historique sous hypothèse de convexité. Lorsque cette hypothèse est irréaliste, le résultat de l'optimisation est extrêmement conditionné par le modèle initial. C'est pourquoi le modèle *a priori* doit intégrer l'ensemble de l'information disponible avant optimisation afin de converger vers une solution réaliste. Il est également possible d'optimiser plusieurs modèles initiaux afin de limiter l'influence du modèle initial (voir Section 1.3.3).

Dans ce travail de thèse, nous utilisons principalement des méthodes d'optimisation locale basées sur les méthodes de Newton.

1.3.2 Optimisation de Gauss-Newton

Méthode de Newton

Les méthodes de Newton font parties des méthodes d'optimisation locale. Elles utilisent le gradient (première dérivée) et le Hessien (seconde dérivée) de la fonction objectif afin de la modéliser localement. Ces méthodes ont l'avantage de converger très rapidement, ce qui limite les appels du modèle direct. Lorsque le problème est non linéaire, il est nécessaire d'itérer. Pour chaque itération, une perturbation est calculée afin de se rapprocher d'un minimum modélisé par la méthode de Newton, jusqu'à ce que le modèle atteigne un minimum de la fonction objectif.

Soit \mathbf{m}^l un vecteur de paramètres représentant la dernière approximation d'un minimum. En définissant $\delta\mathbf{m} = \mathbf{m} - \mathbf{m}^l$, il est possible d'approximer la fonction objectif par son développement de Taylor au second ordre, $Q(\mathbf{m})$, tel que :

$$O(\mathbf{m}) \approx Q(\mathbf{m}) = O(\mathbf{m}^l) + \nabla O(\mathbf{m}^l)^T \delta\mathbf{m} + \frac{1}{2} \delta\mathbf{m}^T \mathbf{H}_l \delta\mathbf{m} \quad (1.26)$$

où $\nabla O(\mathbf{m}^l)^T$ et $\mathbf{H}_l = \nabla_{\mathbf{m}} [(\nabla O(\mathbf{m}^l))^T]$ représentent le gradient et le Hessien de la fonction objectif. Lorsque la matrice Hessienne est définie positive, la fonction $Q(\mathbf{m})$ atteint son minimum (unique) quand sa dérivée en fonction de $\delta\mathbf{m}$ est nulle. Ainsi, la mise à jour du vecteur \mathbf{m}^l est donné par :

$$\mathbf{H}_l \delta\mathbf{m}^l = -\nabla O(\mathbf{m}^l) \quad (1.27)$$

$$\mathbf{m}^{l+1} = \mathbf{m}^l + \delta\mathbf{m}^l \quad (1.28)$$

où $\delta\mathbf{m}^l$ est la solution de l'équation 1.27, sous réserve que le Hessien soit inversible.

Méthode de Gauss-Newton

Le gradient et le Hessien de la fonction objectif 1.22 sont donnés par :

$$\nabla O(\mathbf{m}) = \mathbf{C}_m^{-1}(\mathbf{m} - \mathbf{m}_{pr}) + \mathbf{G}^T \mathbf{C}_D^{-1}(g(\mathbf{m}) - \mathbf{d}^{obs}) \quad (1.29)$$

et

$$\mathbf{H} = \mathbf{C}_m^{-1} + \mathbf{G}^T \mathbf{C}_D^{-1} \mathbf{G} + (\nabla \mathbf{G})^T \mathbf{C}_D^{-1} (g(\mathbf{m}) - \mathbf{d}^{obs}) \quad (1.30)$$

où \mathbf{G} est une $(n \times n^d)$ matrice de dérivées partielles des simulations du modèle direct $g(\cdot)$ par rapport aux paramètres \mathbf{m} , telle que tout élément $g_{ij} = \mathbf{G}(i, j)$ est donné par :

$$g_{ij} = \frac{\partial g_i(\mathbf{m})}{\partial \mathbf{m}_j} \quad (1.31)$$

Cette matrice est appelée matrice des sensibilités. Elle permet d'estimer l'influence des variations des paramètres sur les réponses simulées (linéarisation du modèle direct).

Le gradient de la matrice de sensibilité étant généralement difficile à obtenir, le troisième terme de l'équation 1.30 est négligé dans la méthode de Gauss-Newton et la matrice Hessienne est donnée par :

$$\mathbf{H} = \mathbf{C}_m^{-1} + \mathbf{G}^T \mathbf{C}_D^{-1} \mathbf{G} \quad (1.32)$$

Ainsi, l'obtention du Hessien est directe lorsque \mathbf{G} est connue.

En substituant les équations 1.32 et 1.29 dans l'équation 1.28, on obtient l'expression de la perturbation de Gauss-Newton, $\delta \mathbf{m}^l$:

$$\delta \mathbf{m}^l = -(\mathbf{C}_m^{-1} + \mathbf{G}^T \mathbf{C}_D^{-1} \mathbf{G})^{-1} [\mathbf{C}_m^{-1}(\mathbf{m} - \mathbf{m}_{pr}) + \mathbf{G}^T \mathbf{C}_D^{-1} (g(\mathbf{m}) - \mathbf{d}^{obs})] \quad (1.33)$$

Lorsque le nombre de données est inférieur au nombre de paramètres ($n^d \ll n$), il est préférable de reformuler l'équation 1.33 à l'aide des formules d'inversion de matrices de Sherman-Woodbury-Morrison [71] afin d'inverser la matrice $(n^d \times n^d)$, $(\mathbf{C}_D + \mathbf{G}_l \mathbf{C}_M \mathbf{G}_l^T)^{-1}$, au lieu de la matrice Hessienne $(n \times n)$:

$$\begin{aligned} \delta \mathbf{m}^l = & -(\mathbf{m}^l - \mathbf{m}_{pr}) \\ & - \mathbf{C}_M \mathbf{G}_l^T (\mathbf{C}_D + \mathbf{G}_l \mathbf{C}_M \mathbf{G}_l^T)^{-1} (g(\mathbf{m}^l) - \mathbf{d}^{obs} - \mathbf{G}_l(\mathbf{m}^l - \mathbf{m}_{pr})) \end{aligned} \quad (1.34)$$

L'équation 1.34 permet de mettre à jour un modèle \mathbf{m}^l basé sur une approximation quadratique locale, $Q(\mathbf{m})$, de la fonction objectif 1.22, $O(\mathbf{m})$ et de la linéarisation G_l du modèle direct pour un état donné à l'itération l . Lorsque cette dernière est effectivement quadratique (p. ex. la distribution est gaussienne et le modèle direct est linéaire), le minimum (global) de $Q(\mathbf{m})$ coïncide avec le minimum de $O(\mathbf{m})$ et l'algorithme de Gauss-Newton ne nécessite qu'une seule itération⁸. Néanmoins, la fonction objectif étant généralement de forme quelconque, l'approximation quadratique n'est valable que localement. Pour cette raison, il est nécessaire de contrôler l'amplitude de la mise à jour afin d'assurer un rapprochement vers un minimum à chaque itération.

⁸Gauss-Newton étant basé sur une approximation du Hessien, le minimum de la fonction quadratique et celui donné par l'algorithme ne sont pas exactement les mêmes, surtout lorsque la différence entre les données observées et simulées est importante

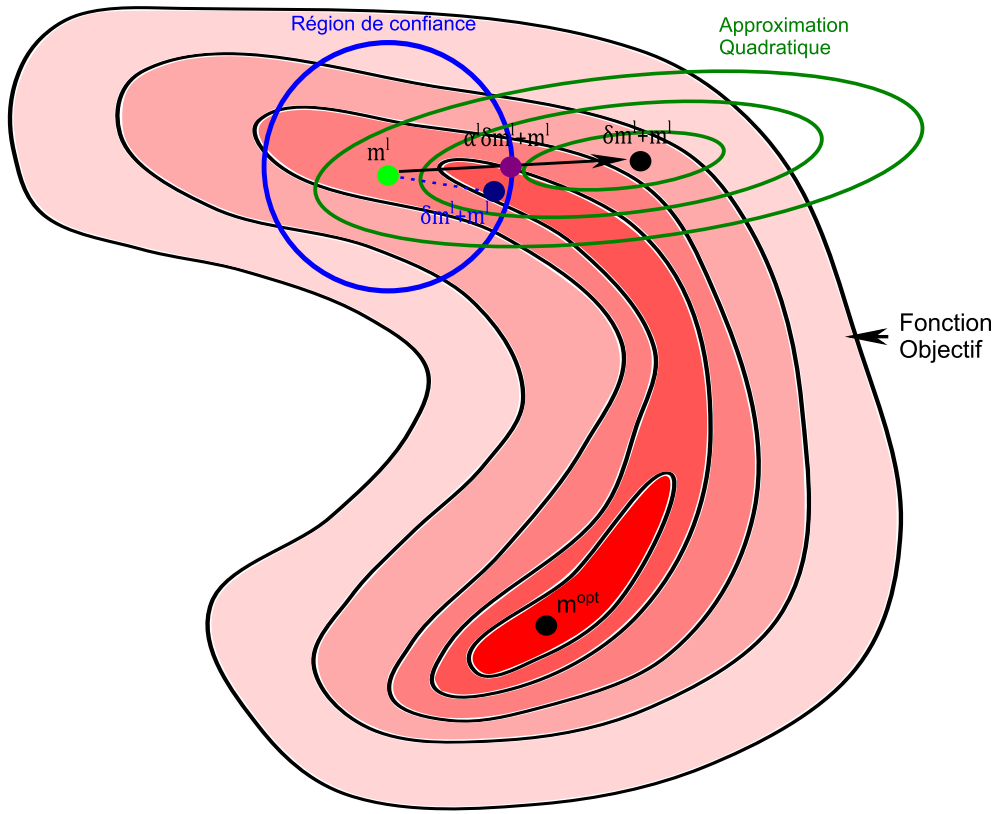


FIGURE 1.3 – Contrôle de la mise à jour de Gauss-Newton dans l'espace des paramètres. Point Vert : état initial. Point Noir : mise à jour complète de Gauss-Newton. Point Violet : contrôle de la mise à jour par recherche linéaire. Point Bleu : contrôle de la mise à jour par région de confiance.

Recherche linéaire et région de confiance

Les méthodes de contrôle de la mise à jour du modèle peuvent se diviser en deux techniques : la recherche linéaire (*line search*) et la recherche dans une région de confiance (*trust region*). La figure 1.3 illustre le principe des deux méthodes.

Recherche linéaire Dans ce groupe de méthodes, la direction de la mise à jour est préservée et seule son amplitude est modifiée. L'équation 1.28 s'écrit alors :

$$\mathbf{m}^{l+1} = \mathbf{m}^l + \alpha^l \delta \mathbf{m}^l \quad (1.35)$$

Le but de l'algorithme est de trouver un facteur α^l tel que :

$$\alpha^l = \min_{\alpha} O(\mathbf{m}^l + \alpha \delta \mathbf{m}^l) \quad (1.36)$$

Le calcul exact de α^l n'est généralement pas nécessaire. Afin que la mise à jour proposée respecte $O(\mathbf{m}^{l+1}) < Q(\mathbf{m}^l)$ et ainsi assure la convergence de l'optimisation, le facteur α doit être choisi selon certaines conditions (cf. inégalité d'Armijo [5], de Goldstein-Price, de Wolfe [137]). La plupart des méthodes de

recherches linéaires utilisent des interpolations quadratiques et cubiques le long de la ligne de recherche (détails et implémentations sont disponibles dans [40]). Néanmoins, ces méthodes nécessitent une évaluation supplémentaire du modèle direct, ce qui peut s'avérer coûteux dans les problèmes de calage historique. D'autres méthodes proposent de définir le facteur α^l en fonction des évaluations précédentes [60]. Par exemple, la méthode du pas réduit (*restricted-step*) évalue la qualité de prédiction de la fonction quadratique par rapport à la fonction objectif réelle (les détails de l'algorithme peuvent être trouvés dans [105]). Lorsque le modèle de la fonction objectif à une itération l est valide, le facteur α est augmenté afin de se rapprocher de 1 (pas complet). Lorsque le modèle quadratique diffère beaucoup de la fonction réelle (p. ex. quand la valeur de la fonction objectif augmente ou que la valeur prédite par $Q(\mathbf{m})$ est très différente de la valeur $O(\mathbf{m})$) la valeur de α est réduite.

Trust region Ces méthodes cherchent à minimiser le modèle de la fonction objectif dans une région de confiance, considérant qu'au-delà, le modèle n'est plus valable. Contrairement aux algorithmes de recherche linéaire, la direction donnée par le modèle n'est pas préservée. Ces méthodes incluent (liste non exhaustive), l'algorithme de Levenberg-Marquardt, la méthode de jambe de chien (*Dog-Leg method*) ou la méthode de gradient conjugué tronqué (*Truncated Conjugate gradient*). L'algorithme de Levenberg-Marquardt étant utilisé dans ce travail de thèse, il est présenté plus en détails dans ce qui suit. Le lecteur pourra se référer au livre de Conn et al. [38] pour plus de détails sur les algorithmes d'optimisation susceptibles d'être utilisés.

Levenberg-Marquardt L'algorithme de Levenberg-Marquardt remplace le Hessien de l'équation 1.27 par une version amortie à l'aide du paramètres d'amortissement λ . Dans sa forme basique, la perturbation \mathbf{m}^{l+1} est obtenue par résolution du système :

$$(\lambda \mathbf{I} + \mathbf{H}_l) \delta \mathbf{m}^l = -\nabla O(\mathbf{m}^l) \quad (1.37)$$

La valeur du paramètre d'amortissement λ a un effet direct sur le comportement de l'algorithme. Il contrôle à la fois la direction de recherche et la taille du pas. Lorsque λ est grand, l'effet du Hessien devient négligeable ($\lambda \mathbf{I} + \mathbf{H}_l \approx \lambda \mathbf{I}$)⁹ et seul le gradient conditionne la mise à jour du modèle. L'effet du Hessien devient de plus en plus important à mesure que λ diminue. Ainsi, l'algorithme de Levenberg-Marquardt est un intermédiaire entre une optimisation par le gradient et une optimisation de Gauss-Newton classique. Le contrôle du paramètre λ dépend généralement de la valeur courante de la fonction objectif et son initialisation dépend du problème donné (voir Chapitre 4 Section 4.4.2). Loin du minimum (en début d'optimisation), la fonction objectif prend des valeurs élevées et l'approximation quadratique n'est pas valide. En conséquence, la valeur de λ est élevée. À proximité du minimum, comme la méthode de Gauss-Newton converge rapidement [105], la valeur de λ doit être faible. Des processus de contrôle optimal peuvent être utilisés pour calculer la valeur de λ à chaque

⁹certaines formes de l'algorithme de Levenberg-Marquardt utilisent la diagonale du Hessien ou d'une matrice de covariance *a priori* au lieu de la matrice identité

itération (voir [40]), mais peuvent être difficiles à appliquer aux problèmes de calage historique de grandes dimensions. Une méthode spécifique de contrôle du paramètre d'amortissement dans le cadre d'une approche multi-échelles est proposée dans le Chapitre 4 Section 4.4.2.

Critères de convergence

Afin d'arrêter le processus d'optimisation, il est important de définir des critères de convergence. Les critères suivants sont utilisés dans ce travail de thèse :

- La variation de la valeur de la fonction objectif est inférieure à une valeur ϵ_0 donnée :

$$\frac{|O(\mathbf{m}^l) - O(\mathbf{m}^{l+1})|}{|O(\mathbf{m}^{l+1})|} < \epsilon_0 \quad (1.38)$$

- La fonction objectif respecte les conditions d'admissibilité données par l'équation 1.25
- L'algorithme dépasse un nombre maximum d'itérations définies par l'utilisateur

En plus de ces conditions, il est possible d'utiliser des critères basés sur la variation des paramètres et la pente du gradient [105].

1.3.3 Algorithmes d'échantillonnage : quantification des incertitudes

Les processus d'optimisation présentés dans la section précédente ne permettent pas de quantifier les incertitudes associées aux problèmes inverses. Dans le meilleur des cas, le processus de minimisation génère le modèle le plus probable, mais aucune information sur la distribution *a posteriori* n'est disponible. Dans cette section, nous nous intéressons aux incertitudes liées à l'estimation des paramètres et à la prédiction du modèle après assimilation des observations.

La forme de la PDF *a posteriori* étant généralement quelconque, il n'est pas possible de l'exprimer analytiquement. Même lorsque cela est possible (p. ex. cas gaussien linéaire), il est souvent plus simple en pratique de travailler avec des modèles numériques. Pour cette raison, l'estimation des incertitudes est souvent basée sur l'utilisation d'un ensemble fini de réalisations échantillonnées de la PDF *a posteriori*.

Différents algorithmes peuvent être utilisés pour échantillonner une PDF de façon exacte ou approchée. Les méthodes d'échantillonnage par parcours aléatoires sur des chaînes de Markov, dites de Monte-Carlo-Markov-Chain (MCMC) [30, 67, 105], sont souvent considérées comme les seules approches pouvant caractériser correctement une PDF. On parle de chaîne de Markov lorsque la simulation d'un état \mathbf{m}^{l+1} dépend uniquement de l'état précédent \mathbf{m}^l . Dans les algorithmes MCMC, les chaînes de Markov sont utilisées pour évaluer la distribution de probabilité. Un premier modèle est choisi aléatoirement afin d'initialiser la chaîne. Puis, pour tout modèle \mathbf{m}^l , on génère un modèle

\mathbf{m}^{l+1} à partir de probabilités de transitions choisies par l'utilisateur. Un critère d'acceptation basé sur les évaluations des états \mathbf{m}^l et \mathbf{m}^{l+1} est ensuite appliqué. Les états les plus probables sont souvent acceptés, tandis que les états peu probables ont une faible probabilité d'être acceptés. Les échantillons (acceptés) de la chaîne permettent alors de caractériser correctement la distribution lorsque sa longueur tend vers l'infini.

Néanmoins, afin d'échantillonner correctement une PDF, ces algorithmes exigent un nombre important d'évaluations, notamment lorsque le problème a beaucoup de paramètres ; ils sont donc difficilement utilisables lorsque le coût d'évaluation du modèle direct est important. Dans ce cas, les méthodes d'échantillonnages basées sur la minimisation [105] peuvent s'avérer plus efficaces pour estimer la PDF. En effet, dans ces méthodes, la simulation de chaque échantillon se fait de manière indépendante. Ainsi, pour un ensemble d'échantillons limité, les réalisations seront moins corrélées que celles produites par une méthode de MCMC, ce qui permet d'obtenir une estimation moins biaisée des incertitudes. Cependant, lorsque le problème est non linéaire, les méthodes d'échantillonnage par minimisation exigent l'application d'un processus d'optimisation pour chaque réalisation générée, ce qui limite le nombre d'échantillons. Dans ce qui suit, la méthode RML (*Randomized Maximum Likelihood*) d'échantillonnage par minimisation est décrite plus en détails.

Lorsque la PDF *a posteriori* réelle est difficile à évaluer, il est possible d'utiliser des méthodes approximatives d'échantillonnage. Un certain nombre d'algorithmes d'optimisation globale (voir Section 1.3.1) tels que le recuit simulé ou les algorithmes génétiques ont été utilisés comme méthodes d'échantillonnage [2, 120, 116]. Néanmoins, de même que lorsqu'elles sont utilisées en minimisation, ces méthodes sont beaucoup trop coûteuses en temps de calcul pour être appliquées à des problèmes de grandes dimensions.

Sous hypothèse de convexité, il est possible d'utiliser des méthodes d'optimisation locale basées sur un ensemble de réalisations [50, 1, 51]. Ces méthodes échantillonnent la PDF par minimisation simultanée d'un ensemble de réalisations. Cet ensemble fournit alors une estimation empirique de la PDF. Ces méthodes peuvent être considérées comme des méthodes de Gauss-Newton pour lesquelles une estimation globale du gradient et du Hessien est calculée statistiquement à partir de l'ensemble de réalisations. L'expression mathématique du modèle direct n'intervenant pas de ces calculs, ces méthodes sont très flexibles en termes de paramétrisation du problème inverse (voir Chapitre 4). Grâce à leur convergence rapide, les méthodes d'ensemble peuvent être appliquées à des problèmes de grande taille associés à un modèle direct coûteux en temps de simulation.

Cas linéaires

Lorsque le modèle direct peut être exprimé sous forme d'application linéaire (p. ex. approximation linéaire de la réponse sismique, Section 1.2.1) et que la distribution *a priori* est supposée gaussienne multivariée (Section 1.3.1), la PDF *a posteriori* peut être exprimée analytiquement (équation 1.21).

Lorsque la taille du problème est modeste, il est possible d'échantillonner la PDF et de calculer le maximum *a posteriori* en utilisant la décomposition

de Cholesky (plus de détails disponibles dans [71, 105]). Lorsque la taille du problème est trop grande, la décomposition de Cholesky devient impossible. C'est par exemple le problème qui est rencontré lors de l'inversion sismique Bayésienne [23].

Plusieurs méthodes sont proposées afin de résoudre ce problème. La première approche consiste à re-paramétriser le problème afin d'obtenir des paramètres indépendants. Buland et al. [22] expriment la PDF *a posteriori* dans le domaine de Fourier ; il est alors possible d'inverser chaque fréquence indépendamment.

La seconde approche, suivie dans ce travail de thèse, utilise un procédé similaire aux méthodes géostatistiques¹⁰ de simulation séquentielle de champs gaussiens (*Sequential Gaussian Simulation ou SGS*) [122, 20, 45]. Les traces sismiques sont simulées de façon séquentielle : pour chaque trace, une distribution de probabilité est localement construite en prenant en compte les données sismiques et les traces précédemment simulées. Ainsi, l'inversion est décomposée en plusieurs problèmes semi-indépendants ayant une taille permettant l'échantillonnage.

Échantillonnage par minimisation : méthode RML pour problèmes inverses non linéaires

La méthode RML (*Randomized Maximum Likelihood*) d'échantillonnage de la PDF *a posteriori* pour des problèmes non linéaires est présentée dans cette section. Il est démontré [105] que cette méthode échantillonne parfaitement dans le cas linéaire gaussien. Le principe de cette méthode est de générer une réalisation non conditionnée¹¹ et de la corriger en utilisant les données observées. Dans ce qui suit, on considère un modèle *a priori* (et de données) gaussien¹² (équations 1.20 et 1.19). L'algorithme RML est alors défini par les étapes suivantes :

1. Générer une réalisation de paramètres du modèle $\mathbf{m}_i \sim \mathbf{N}(\overline{\mathbf{m}}_{pr}, \mathbf{C}_M)$
2. Générer une réalisation de données bruitées $\mathbf{d}_i \sim \mathbf{N}(\mathbf{d}^{obs}, \mathbf{C}_D)$
3. Minimiser la fonction objectif 1.39 :

$$O_i(\mathbf{m}) = \frac{1}{2}(\mathbf{d}_i - g(\mathbf{m}))^T \mathbf{C}_D^{-1}(\mathbf{d}_i - g(\mathbf{m})) + \frac{1}{2}(\mathbf{m} - \mathbf{m}_i)^T \mathbf{C}_M^{-1}(\mathbf{m} - \mathbf{m}_i) \quad (1.39)$$

Le groupe de réalisations générées par la méthode RML donne alors une approximation de la PDF *a posteriori*¹³. Néanmoins, la dernière étape de l'algorithme nécessitant un processus d'optimisation¹⁴, seul un nombre très réduit

¹⁰les méthodes géostatistiques sont des processus d'optimisation. Les méthodes déterministes (p. ex. le krigeage) sont des recherches de MAP, tandis que les méthodes stochastiques (p. ex. la SGS) sont des processus d'échantillonnage.

¹¹non conditionnée par les données utilisées dans le processus courant d'inversion, mais la réalisation peut être conditionnée par d'autres données

¹²l'algorithme est généralisable aux cas non gaussiens.

¹³il est possible d'intégrer dans la méthode RML, un processus d'acceptation/rejet similaire aux méthodes MCMC afin d'assurer un échantillonnage correct

¹⁴n'importe quelle méthode d'optimisation peut être utilisée.

de modèles peuvent être générés en pratique. Dans ce cas, une méthode RML basée ensemble (enRML, voir Section 1.3.3) peut être plus appropriée.

Méthodes d'ensemble

Dans cette section, nous décrivons quelques méthodes d'optimisation locale (Section 1.3.1) basées sur un ensemble de réalisations (enOpt). Nous ne présentons pas les méthodes stochastiques d'optimisation globale (telles que les essais de particule [47]) qui utilisent également un ensemble de réalisations.

Les techniques enOpt forment un groupe de méthodes d'optimisation dérivées de la méthode de Gauss-Newton. Comme présenté dans la Section 1.3.2, le calcul du gradient et du Hessian dépend essentiellement de la matrice de sensibilité \mathbf{G} . Dans les méthodes déterministes standard, cette matrice peut être calculée par des méthodes de différences finies, par des méthodes directes ou par des méthodes adjointes [105, 29]. Seules les méthodes adjointes sont utilisables pour des modèles de grandes dimensions¹⁵, mais elles nécessitent une connaissance approfondie des équations du modèle direct et ne permettent pas une grande flexibilité de paramétrisation. Les méthodes enOpt évitent ces problèmes en estimant de manière statistique une matrice de sensibilité globale utilisée dans la mise à jour de chacune des réalisations de l'ensemble.

Notations Dans ce qui suit, on considère un ensemble de n^e réalisations constituées de n paramètres et on dénote par \mathbf{m}_{pr}^j la j ème réalisation *a priori* de l'ensemble, \mathbf{m}_l^j la j ème réalisation à l'itération l , \mathbf{d}_l^j le vecteur $((n^d \times 1))$ de données simulées par la réalisation \mathbf{m}_l^j , $\overline{\mathbf{m}}_{pr}$ le vecteur moyen de l'ensemble *a priori*, $\overline{\mathbf{m}}_l$ le vecteur moyen de l'ensemble à l'itération l , et par $\overline{\mathbf{d}}_l$ le vecteur moyen de l'ensemble des données simulées. On dénote par $\Delta\mathbf{M}_{pr}$ la matrice $(n \times n^e)$ de déviation *a priori* dont les colonnes correspondent aux différences entre chaque réalisation *a priori* \mathbf{m}_{pr}^j et la moyenne $\overline{\mathbf{m}}_{pr}$, $\Delta\mathbf{M}_l$ la matrice $(n \times n^e)$ de déviation dont les colonnes correspondent aux différences entre chaque réalisation \mathbf{m}_l^j et la moyenne $\overline{\mathbf{m}}$, et \mathbf{D}_l la matrice $(n^d \times n^e)$ de déviation des réponses simulées dont les colonnes correspondent aux différences entre les vecteurs \mathbf{d}_l^j et $\overline{\mathbf{d}}_l$.

Matrice des sensibilités globale La matrice des sensibilités relie linéairement les variations des paramètres du modèle aux variations des données simulées par le modèle direct. Ainsi on peut poser :

$$\Delta\mathbf{D}_l = \mathbf{G}_l \Delta\mathbf{M}_l \quad (1.40)$$

La matrice $\Delta\mathbf{M}_l$ n'étant généralement pas inversible (ou même carrée), sa pseudo-inverse¹⁶ est utilisée pour obtenir une estimation statistique de \mathbf{G}_l :

$$\mathbf{G}_l = \Delta\mathbf{D}_l (\Delta\mathbf{M}_l)^+ \quad (1.41)$$

où l'exposant « + » dénote la pseudo-inverse.

¹⁵la complexité de la méthode adjointe dépend du nombre de données à assimiler et non pas du nombre de paramètres

¹⁶il est possible d'utiliser la décomposition en valeur singulière tronquée ou non-tronquée ((T)SVD). Lorsque la TSVD est utilisée, il est important de normaliser les paramètres optimisés afin de ne pas perdre des informations importantes

Matrice de covariance *a priori* L'ensemble des réalisations est utilisé pour estimer la covariance *a priori* du modèle :

$$\mathbf{C}_M = \frac{\Delta \mathbf{M}_{pr} \Delta \mathbf{M}_{pr}^T}{n^e - 1} \quad (1.42)$$

Intégration séquentielle des données Un groupe de méthodes enOpt, dont fait partie l'*ensemble Kalman Filter* (enKF) [50], propose une intégration séquentielle des données dans le temps. Ainsi, pour chaque temps t d'une séquence donnée, l'ensemble des paramètres du modèle, ainsi que les variables d'état, sont mis à jour (étape d'*Update*) en utilisant les données observées. L'ensemble mis à jour au temps t est ensuite considéré comme modèle *a priori* lors de l'assimilation au temps $t + 1$. Ainsi, l'équation 1.33 se simplifie pour chaque *Update* :

$$\delta \mathbf{m} = -\mathbf{C}_{md}(\mathbf{C}_D + \mathbf{C}_{dd})^{-1}(g(\mathbf{m}^l) - \mathbf{d}^{obs}) \quad (1.43)$$

$$= \mathbf{K}^e(g(\mathbf{m}^l) - \mathbf{d}^{obs}) \quad (1.44)$$

où

$$\mathbf{m}_{pr} - \mathbf{m} = 0$$

et

$$\mathbf{C}_{md} = \mathbf{C}_M \mathbf{G}^T = \frac{\Delta \mathbf{M}_{pr} \Delta \mathbf{M}_{pr}^T (\Delta \mathbf{M}_{pr})^T + \Delta \mathbf{D}^T}{n^e - 1} = \frac{\Delta \mathbf{M}_{pr} \Delta \mathbf{D}^T}{n^e - 1}$$

correspond à la matrice de covariance croisée entre les paramètres et les données, et

$$\begin{aligned} \mathbf{C}_{dd} &= \mathbf{G} \mathbf{C}_M \mathbf{G}^T = \frac{\Delta \mathbf{D} (\Delta \mathbf{M}_{pr})^+ \Delta \mathbf{M}_{pr} \Delta \mathbf{M}_{pr}^T (\Delta \mathbf{M}_{pr})^T + \Delta \mathbf{D}^T}{n^e - 1} \\ &= \frac{\Delta \mathbf{D} \Delta \mathbf{D}^T}{n^e - 1} \end{aligned}$$

à la matrice de covariance des données simulées. On retrouve alors l'étape *Update* enKF pour les paramètres du modèle (les variables d'état ne sont pas considérées ici) dans lequel \mathbf{K}^e représente l'estimation du gain de Kalman. Ainsi, lorsque la matrice de covariance *a priori* est estimée à partir de l'ensemble courant, la matrice \mathbf{G} n'a pas besoin d'être calculée explicitement.

Intégration simultanée des données Contrairement à l'approche séquentielle, l'ensemble des données est intégré en une seule fois [52, 32]. Pour les algorithmes itératifs [33, 35], la matrice de sensibilité est calculée à chaque itération en utilisant l'équation 1.42. La matrice de covariance \mathbf{C}_M est calculée à partir de l'ensemble *a priori*, mais n'évolue pas durant les itérations. Ainsi, les termes $\mathbf{C}_M \mathbf{G}^T$ et $\mathbf{G} \mathbf{C}_M \mathbf{G}^T$ ne se simplifient pas et doivent être calculés explicitement, ce qui rend le calcul de la perturbation $\delta \mathbf{m}^l$ moins efficace que dans l'approche séquentielle présentée précédemment. Néanmoins, il est possible dans certains cas, d'approximer la matrice de covariance *a priori* par la matrice de covariance courante (voir Chapitre 4 Section 4.3.2).

Un deuxième désavantage des approches itératives, par rapport aux approches séquentielles, vient du fait que le modèle direct est évalué à chaque itération pour chacun des membres de l'ensemble. En pratique, le redémarrage du modèle direct (nécessaire dans l'approche séquentielle) peut avoir un coût important ce qui limite l'avantage de l'approche séquentielle par rapport à l'approche itérative¹⁷. D'une manière générale, l'utilisation de méthodes itératives est pertinente lorsque le problème est hautement non linéaire et que le modèle initial est loin du minimum [74]. Une description détaillée de la méthode de Levenberg-Marquart basée ensemble, utilisée dans ce travail de thèse, est fournie dans le Chapitre 4 Section 4.3.2.

Méthode RML basée ensemble (enRML) L'algorithme RML peut s'adapter efficacement aux méthodes enOpt. Pour chaque membre \mathbf{m}^j de l'ensemble, on cherche à optimiser l'équation 1.39 en utilisant le gradient global calculé à partir de l'ensemble. L'enRML propose originellement [74] une intégration séquentielle des données, mais pour chaque étape d'*Update*, plusieurs itérations sont effectuées, ce qui permet entre autre d'éviter les problèmes liés à la mise à jour des variables d'état¹⁸. Néanmoins, cette approche est doublement pénalisée car elle exige des évaluations multiples du modèle direct et elle nécessite plusieurs redémarrages du simulateur. Dans ce cas, il est préférable d'utiliser les méthodes itératives avec intégration simultanée des données (voir Chapitre 4 Section 4.3.2).

Limitations La capacité d'approximer les matrices de covariances et de covariances croisées par un ensemble de réalisations de taille très inférieure au nombre de paramètres permet d'appliquer les méthodes d'ensembles à des problèmes de grandes dimensions. Néanmoins, cette caractéristique pose deux problèmes majeurs. Premièrement, les estimations des matrices de covariances (équation 1.44) et/ou de la matrice des sensibilités (équation 1.42) ont tendance à être corrompues par de fausses corrélations (*spurious correlations*) qui sont dues à la taille limitée de l'ensemble. Ces fausses corrélations provoquent des changements non nécessaires des paramètres du modèle et des variables d'état, et peuvent aboutir à une réduction importante de la variabilité de l'ensemble (*collapse*), ce qui affecte grandement l'estimation des incertitudes. Deuxièmement, les méthodes enKF ne peuvent pas assimiler un nombre de données indépendantes plus important que le rang de la matrice $\Delta \mathbf{M} < n^e$. Enfin, étant donné que l'étape d'*Update* est une combinaison linéaire des réponses simulées de l'ensemble, elle est limitée à l'espace couvert par l'ensemble. De plus, la mise à jour des variables d'état dans la méthode enKF standard peut conduire à des états physiques impossibles. Cependant, les méthodes itératives pallient ces derniers problèmes et souffrent uniquement des problèmes de fausses corrélations. Des méthodes de régularisation peuvent alors être utilisées afin de limiter la dégradation des réalisations causée par ces corrélations erronées.

¹⁷pour un simulateur d'écoulements, les coûts de redémarrage peuvent prendre un temps équivalent à une simulation ou plus.

¹⁸la mise à jour des variables d'état peut créer des configurations irréalistes ce qui pose des problèmes de convergence pour le simulateur.

Localisation des perturbations Dans les applications réelles, lorsque la taille de l'ensemble est relativement petite par rapport à la dimension du problème, un processus de régularisation, appelé *localisation*, aide à éviter les perturbations parasites des variables qui ne sont pas réellement sensibles aux observations. Des matrices de projection (*screening matrices*) sont généralement appliquées par multiplications éléments-par-éléments (multiplication d'Hadamard) au gain de Kalman [32, 31] ou indirectement aux matrices de covariances [76, 82, 31]. La localisation basée sur la distance (*distance-based localisation*) est utilisée dans les applications pétrolières et atmosphériques afin de limiter l'influence des données, sur la mise à jour des paramètres, au-delà d'une certaine distance. Des fonctions de localisation, habituellement définies empiriquement, sont utilisées pour créer une matrice de projection. Leurs portées dépendent de la taille de l'ensemble, ce dernier conditionnant la qualité des approximations [76, 82, 58]. Lorsque les fonctions de localisation ne permettent pas de régulariser correctement le processus, il est possible d'effectuer des analyses de sensibilité plus poussées (p. ex. en utilisant des lignes de courant en monophasique) afin de définir des zones de localisation [31], mais ceci peut s'avérer être une tâche très complexe. De plus, pour certains paramètres, tels que les courbes de perméabilités relatives ou les contacts fluides, une interprétation spatiale n'a pas de sens.

Dans ce cas, des méthodes plus générales ont été proposées [3, 142, 34]. Ces méthodes permettent d'attribuer un facteur de confiance pour chaque corrélation, en analysant plusieurs estimations du gain de Kalman calculées à partir d'ensembles différents (technique de *bootstrap*). Ces méthodes peuvent être améliorées lorsque la paramétrisation permet de faire ressortir des paramètres importants. Chen et Oliver [34] appliquent la méthode de *bootstrap* aux coefficients d'ondelettes et montrent que l'estimation du gain de Kalman pour les basses fréquences est moins sujet au bruit, même lorsque un petit ensemble est utilisé. Dans ce cas cependant, la localisation automatique ne parvient pas à identifier les vraies corrélations relatives aux échelles fines. Des informations supplémentaires sur la localisation ainsi que le développement d'une régularisation multi-échelles sont disponibles en Section 4.4.1.

Chapitre 2

Re-paramétrisations multi-échelles

« Le véritable voyage de découverte ne consiste pas à chercher de nouveaux paysages, mais à avoir de nouveaux yeux. »

Marcel Proust, 1871-1922

2.1 Résumé et contributions

Dans ce chapitre nous discutons des avantages de la re-paramétrisation lors de l'intégration des données de réservoirs et dans les processus d'optimisation. Nous nous intéressons particulièrement aux approches multi-échelles et à la paramétrisation en ondelettes dont nous introduisons les bases théoriques. Nous décrivons par la suite un nouveau schéma de *lifting* en ondelettes (*lifting scheme*) qui permet la création d'ondelettes non stationnaires, appelées ondelettes de seconde génération, applicables à des grilles stratigraphiques complexes ; ces ondelettes sont utilisées dans le Chapitre 4 dans le processus de calage des données de production. Enfin, nous présentons une paramétrisation multi-échelles en pyramide gaussienne utilisée dans le Chapitre 3 lors de la génération de réalisations géostatistiques multipoints.

Ce chapitre inclut la contribution suivante :

1. développement et implémentation (C++) d'une transformation 3D en ondelettes de seconde génération adaptée aux grilles stratigraphiques (Section 2.3.2)

2.2 Choix de la paramétrisation

Cette section traite de la régularisation par re-paramétrisation lors de la résolution d'un problème inverse non linéaire. Bien que nous nous appuyions principalement sur l'exemple de l'inversion des données de production, les concepts débattus ici sont applicables à d'autres problèmes d'optimisation non linéaire, tels que les méthodes géostatistiques multipoints présentées dans le Chapitre 3.

Nous distinguons ici deux types de paramètres qui peuvent être parfois confondus :

Paramètres de simulation Ce sont les paramètres nécessaires au modèle direct pour produire une réponse simulée (équation 1.15). On les dénote par le vecteur¹ $\mathbf{m} \in \mathbb{R}^{n^{sim}}$. La taille du vecteur est généralement fixe pendant toute l'étude du réservoir. Dans les exemples étudiés dans ce travail de thèse, ce vecteur est composé (entre autres) de valeurs de porosité et de perméabilités définies dans chacune des cellules de la grille de simulation, ainsi que de paramètres caractérisant le comportement des fluides (p. ex. perméabilités relatives).

Paramètres d'optimisation Ce sont les paramètres modifiés durant les processus d'optimisation (Section 1.3). Ils sont directement reliés aux paramètres de simulation. Lorsque le lien est linéaire, le vecteur de paramètres d'optimisation, $\gamma \in \mathbb{R}^n$ est donné par :

$$\gamma = \mathbf{A}\mathbf{B}\mathbf{m} \quad (2.1)$$

où \mathbf{B} est une matrice ($n \times n^{sim}$) de sous échantillonnage et \mathbf{A} est une matrice ($n \times n$) de transformation. La matrice \mathbf{B} montre que certains paramètres du modèles direct ne sont pas touchés par l'optimisation. Afin de simplifier les notations, on considère par la suite $\mathbf{B} = \mathbf{I}$. On parle de re-paramétrisation lorsque la matrice \mathbf{A} n'est pas la matrice identité. À ce stade, l'optimisation de γ ou de \mathbf{m} est équivalente². Cependant, certaines paramétrisations permettent de réduire plus facilement le nombre de paramètres optimisés afin de :

- réduire les coûts de calculs et stabiliser les processus d'inversion
- éviter la sur-paramétrisation et l'inversion de bruit
- éviter des minima locaux lors de l'optimisation
- mettre en avant des paramètres importants
- préserver certains paramètres pendant l'optimisation

D'une manière générale, le choix de la paramétrisation dépend du problème, de la méthode d'optimisation et du but recherché.

2.2.1 Analyse linéaire de résolution des paramètres

Le contenu informationnel porté par les données de production et/ou les données sismiques est souvent limité et ne permet pas de caractériser l'ensemble des paramètres du modèle. Lorsque la relation entre les paramètres optimisés et la réponse simulée est linéaire, tel que $\mathbf{d}^{sim} = \mathbf{G}\gamma$ (p. ex. Section 1.3.1) et que les données sont mesurées avec un bruit modélisé par la matrice de covariance \mathbf{C}_D (Section 1.3.1), Tarantola [131] définit une matrice \mathbf{R} qui mesure la résolution des paramètres γ par les données comme :

$$\mathbf{R} = \mathbf{C}_M \mathbf{G}^T (\mathbf{G} \mathbf{C}_M \mathbf{G}^T + \mathbf{C}_D)^{-1} \mathbf{G} \quad (2.2)$$

¹le modèle direct est supposé discret.

²lorsque que des processus d'optimisation linéaires ou basés sur des linéarisations successives (p. ex. Gauss-Newton) sont utilisées

La trace de \mathbf{R} équivaut au nombre de paramètres qui sont résolus par les données, le restant étant uniquement caractérisé par l'information *a priori*. Le but de la re-paramétrisation est alors de trouver au mieux un sous-ensemble de paramètres qui est entièrement (à l'incertitude des données près) caractérisé par les données. Un développement sur l'effet de la re-paramétrisation et la réduction de la taille du problème est disponible dans [95].

La résolution des paramètres étant directement liée à la sensibilité des données et à la covariance des paramètres initiaux, il est possible de trouver une paramétrisation optimale en utilisant les matrices \mathbf{G} et \mathbf{C}_M . Plusieurs approches basées sur la décomposition en valeurs singulières (SVD) de sous-produits des matrices \mathbf{G} et \mathbf{C}_M sont proposées dans la littérature [121, 17, 114, 140, 135, 111, 46].

Bien que ce type de paramétrisation soit optimale lorsque le problème est linéaire et qu'une bonne estimation de la covariance *a priori* est disponible, ces méthodes ont plusieurs limites. Tout d'abord, lorsque le nombre de paramètres est très important, il est coûteux d'effectuer une décomposition en valeurs singulières. Ensuite, la covariance *a priori* étant rarement connue avec précision, elle peut conduire à une paramétrisation mal adaptée au problème. Enfin, la matrice des sensibilités \mathbf{G} n'est valable (dans le cas des problèmes non linéaires) que pour une configuration de paramètres donnée : la sensibilité des paramètres évoluant avec les itérations, la paramétrisation devrait également évoluer.

2.2.2 Paramétrisations adaptatives et approches multi-échelles

Les approches multi-échelles sont communément utilisées en calage historique [9, 72, 10, 95, 14, 12] dans le but d'adapter la paramétrisation aux données et à l'état du système, stabiliser l'inversion et éviter la sur-paramétrisation. Bien que cela n'ait pas été formellement prouvé, il est considéré que les méthodes multi-échelles aident à éviter les minima locaux lors de l'optimisation [92, 97, 29] ou produisent une solution optimale plus simple [91]. La plupart de ces méthodes commencent par optimiser un nombre limité de paramètres de grandes échelles. Des paramètres supplémentaires sont ensuite choisis grâce à l'analyse des résultats de l'étape d'optimisation : des indicateurs de raffinement sont calculés à partir de la matrice des sensibilités ou bien du gradient afin de maximiser la réduction de la fonction objectif lors des prochaines itérations.

Plusieurs paramétrisations multi-échelles sont utilisées dans le domaine des géosciences [104, 14]. Les méthodes basées sur des zonations adaptatives démarrent avec un modèle grossier et affinent localement le découpage des propriétés en fonction de leur influence sur les écoulements [9, 72]. Un défi de taille pour ces méthodes est de pouvoir intégrer plusieurs sources d'informations *a priori* exprimées à des échelles différentes. Les techniques multi-échelles de paramétrisations dérivées de la géostatistique [61] ou la stratégie de prédiction-correction [54] permettent de préserver des variabilités spatiales, mais ne peuvent pas gérer des modèles *a priori* qui contiennent des informations précises, telles que des structures géologiques ou des transitions de propriétés dérivées de la sismique. Ainsi, il est parfois nécessaire d'utiliser des paramétrisations qui permettent d'exprimer les propriétés à toutes les résolutions.

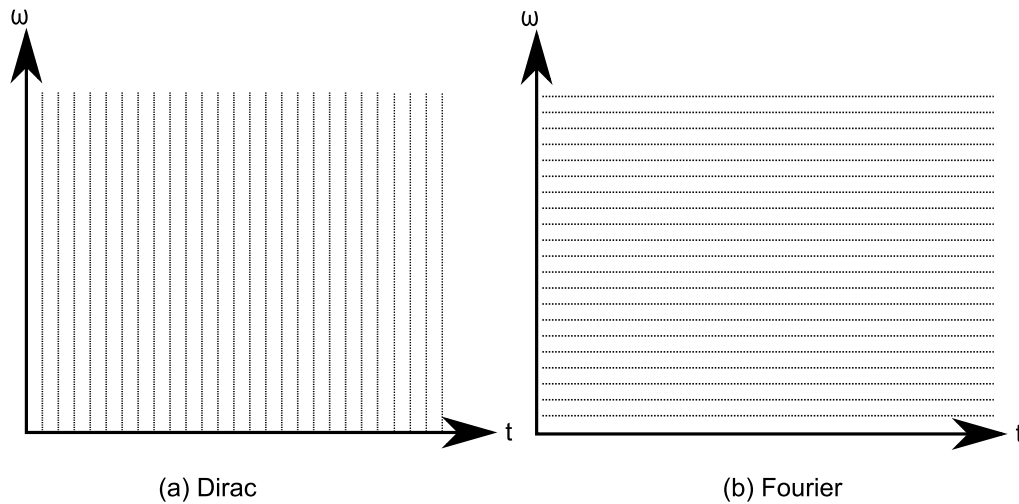


FIGURE 2.1 – Analyse temps-fréquence de la représentation de Dirac et de Fourier. Schéma temps-fréquence de la représentation de Dirac (a) et de Fourier (b). Modifié de [11]

Les approches les plus communes effectuent un changement de base : des caractéristiques intéressantes peuvent découler de ces transformations, telles qu’une décomposition en échelles indépendantes, une représentation compressée ou une décorrélation des paramètres. Des références de quelques méthodes basées sur la décomposition en composantes principales (ou en valeurs singulières) ou par transformation de Karhunen-Loeve sont données dans la section précédente. Des approches plus récentes, issues du domaine de compression numérique, ont l’avantage d’être indépendantes de la connaissance *a priori* et sont très efficaces en termes de temps de calcul. Deux transformations sont principalement utilisées dans l’inversion des données historiques : la transformée discrète en cosinus (DCT) [85, 88, 13, 15] et la transformée discrète en ondelettes (DWT) [95, 119, 34]. Pour la DCT, le domaine est caractérisé par des fonctions globales basées sur des cosinus, alors que la DWT utilise des fonctions complexes, appelées ondelettes, localisées dans l’espace. Ces paramétrisations fournissent des approximations grossières des champs de propriétés avec un nombre limité de paramètres et peuvent, au besoin, inclure plus de détails dans le modèle.

2.3 Une paramétrisation multi-échelles adaptée aux problématiques des géosciences : les ondelettes

Les ondelettes ont été introduites au milieu du XX^e siècle afin de décrire un signal à la fois par ses composantes en temps et en fréquence. Jusqu’alors, un signal $f(t)$ était soit représenté en temps sur la base continue des fonctions de Dirac, soit en fréquence dans la base de Fourier. La première représentation permet une analyse précise en temps, mais ne contient aucune information fréquentielle (figure 2.1-a). Inversement, la décomposition de Fourier est très précise en fréquence, mais ne contient pas d’information sur le contenu en temps (figure 2.1-b). L’analyse fréquentielle de Fourier devient limitée lorsque

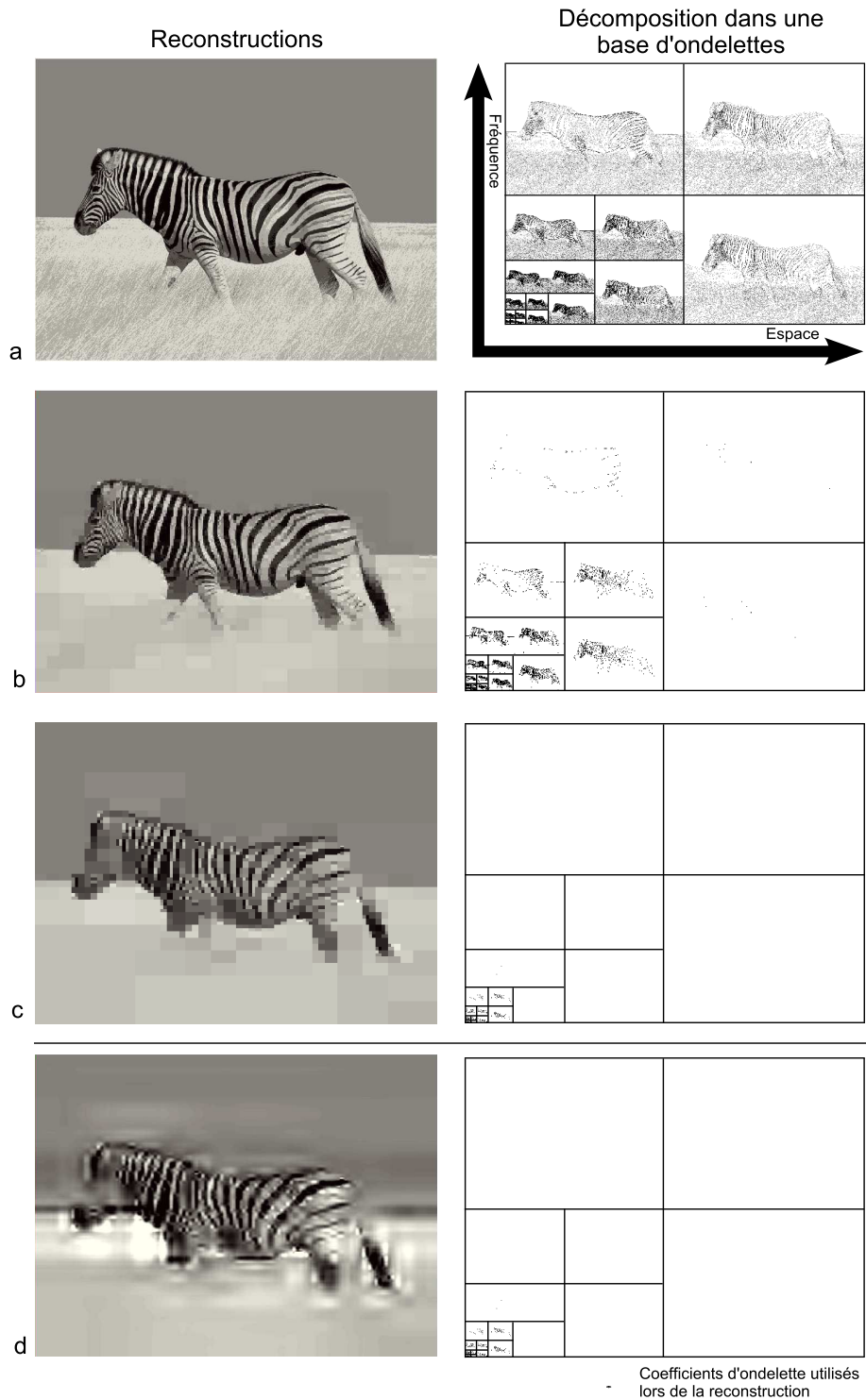


FIGURE 2.2 – Décomposition en coefficients d'ondelettes. La colonne de droite montre les reconstructions partielles de l'image en utilisant les coefficients d'ondelettes représentés dans la colonne de gauche. Utilisation des ondelettes de Haar (a-b-c) et de Haar améliorée (d) (ondelette primaire de degré 3, cf. Section 2.3.2). Reconstruction utilisant : (a) 100% des coefficients (complète), (b) 2% des coefficients, (c) et (d) 0.01%. Images générées à partir du code C++ développé dans cette thèse.

l'on doit traiter des signaux non stationnaires. En effet, l'analyse de Fourier d'une musique nous donnera l'ensemble des notes jouées, mais sera incapable de distinguer les différents mouvements qui constituent le morceau. De même, l'analyse de l'image présentée dans la figure 2.2-a nous indiquera qu'il existe des transitions rapides de couleurs, correspondant aux rayures du zèbre et à l'herbe, mais sera incapable de préciser à quels endroits ses fréquences sont importantes ou inversement négligeables.

Les transformations en ondelettes répondent à ce problème. Un signal est décomposé sur des bases de fonctions finies, appelées ondelettes, qui sont localisées dans l'espace. Contrairement aux fonctions globales de la base de Fourier (sinus et cosinus), les ondelettes sont définies uniquement sur une portion finie du domaine (support de l'ondelette). Ainsi, les coefficients possèdent à la fois une composante spatiale et fréquentielle.

Cette propriété est exploitée en compression numérique. Le but de cette technique est de sous échantillonner les coefficients d'un signal afin de réduire sa taille, tout en préservant certaines de ses propriétés. En compression d'images, l'aspect visuel est le critère conditionnant l'échantillonnage des paramètres. Ainsi, les fréquences nécessaires pour caractériser une variation rapide de couleur (par exemple les rayures et les contours du zèbre, figure 2.2) seront localement préservées, tandis que certaines fréquences pourront être négligées lorsque la coloration est homogène (p. ex. ciel en arrière plan du zèbre, figure 2.2-b-c-d). Cette capacité d'adaptation du contenu fréquentiel en fonction de la localisation fait la force de la décomposition en ondelettes.

Les données utilisées en caractérisation des réservoirs (Chapitre 4.1) sont souvent hétérogènes à la fois dans l'espace, dû à la mesure ponctuelle des données, et en fréquence. En effet, les données sismiques sont généralement homogènes dans l'espace, à l'exception des zones faillées où le signal n'est que partiellement récupéré, mais contiennent de l'information uniquement dans une certaine bande de fréquences. Le contenu informatif des données de production est généralement le plus important proche des puits et s'estompe avec l'éloignement à la mesure. Ainsi, le contenu fréquentiel des propriétés caractérisé par les données dynamiques est variable selon la position dans l'espace. De la même façon, les données diagraphiques de puits sont généralement très détaillées, mais de manière ponctuelle. Grâce à la paramétrisation temps-fréquence en ondelettes, l'intégration de différents types de données peut être efficacement gérée. Ces avantages sont utilisés dans les Chapitre 3, lors de l'intégration de données d'échelles différentes, et dans le Chapitre 4 lors de mise à jour de modèles issus de l'inversion sismique par des données de production.

2.3.1 Les ondelettes de première génération

Analyse temps-fréquence et ondelette génératrice

La première transformée en ondelettes fut proposée par Alfréd Haar en 1909 comme exemple de base orthogonale de l'espace des fonctions de carré intégrables (espace $L^2(\mathbb{R})$). Elle ne fut reconnue comme telle que bien plus tard, lorsque le terme « ondelette » fut introduit à la suite des travaux de Morlet et Grossmann [73]. Plus tard, afin de contourner les limites de la transformée

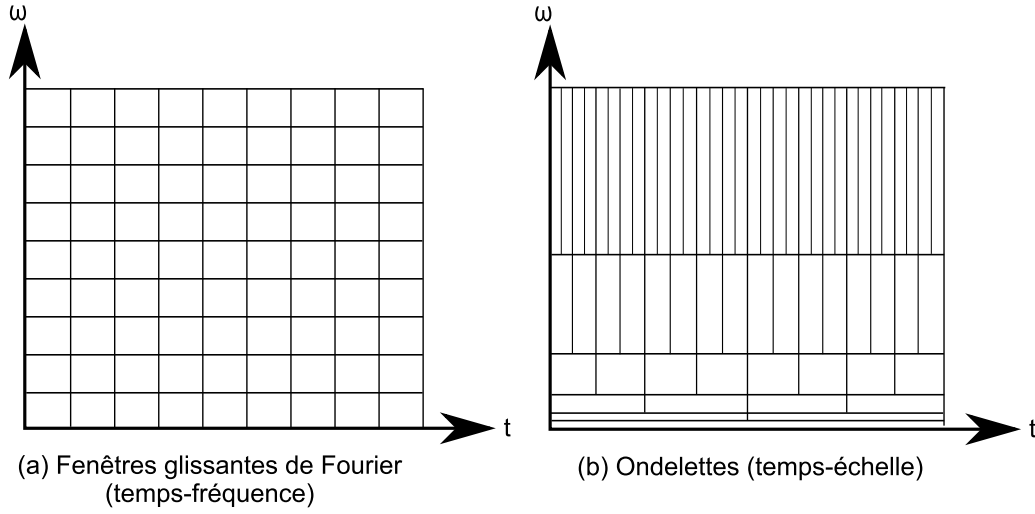


FIGURE 2.3 – Analyse temps-fréquence et temps-échelle. Analyse temps-fréquence (a) et temps-échelle (b). Adapté de [11]

de Fourier sur l'analyse en temps, Gabor [59] proposa d'appliquer une analyse temps-fréquence, semblable à Fourier, basée sur des fonctions gaussiennes dans des fenêtres spatiales de largeur constante (indépendante de la fréquence, voir figure 2.3-a). Ces fonctions satisfont approximativement les conditions d'admissibilité des ondelettes et sont pour cette raison parfois appelées ondelettes de Gabor. Finalement, c'est avec les travaux de Morlet et Grossmann [73] que les ondelettes apparaissent pour la première fois. Les auteurs présentent une approche temps-échelle dans laquelle la largeur spatiale des fonctions de base est inversement proportionnelle à la fréquence (figure 2.3-b) et ils proposent de construire des bases de fonctions, $\psi_{r,k}(t)$, par translation et dilatation d'une fonction génératrice, $\psi(t) \in L^2(\mathbb{R})$, que l'on appelle ondelette mère :

$$\psi_{r,k}(t) = \frac{1}{\sqrt{s_0^r}} \psi \left(\frac{t - k\tau_0 s_0^r}{s_0^r} \right) \quad (2.3)$$

où $(\tau_0, s_0^r) \in \mathbb{R}^{+*} \times \mathbb{R}$ correspondent respectivement aux facteurs de translation et de dilatation, lesquels prenant souvent les valeurs $\tau_0 = 1$ et $s_0^r = 2$. Les indices r et k font référence à l'échelle et à la position dans le temps (ou l'espace) de l'ondelette. On désigne par la suite cette catégorie d'ondelettes comme « ondelettes de première génération » (*First Generation Wavelets* ou FGW).

Lorsque r est décrémenté de 1, le spectre de fréquences couvert par les ondelettes est divisé par un facteur 2 et est translaté des hautes aux basses fréquences (figure 2.5). En pratique, l'analyse temps-échelle à fenêtres variables est beaucoup plus pratique que celle par fenêtres glissantes de Fourier : il est plus simple de construire une grande variété de bases de fonctions, la transformation repose sur la forme d'une fonction unique, les différents échelles (largeur des fenêtres en fréquence) sont adaptées à l'échantillonnage des signaux et enfin, elle permet une décomposition efficace des signaux discrets.

Analyse multi-résolution

Quelques années plus tard, Mallat [96] intégra la décomposition en ondelettes dans la théorie de l'*analyse multi-résolution*. Ce cadre d'analyse permet d'exprimer les ondelettes et leurs propriétés à l'aide de filtres discrets. Une *analyse multi-résolution* d'un espace dans $L^2(\mathbb{R})$ est une séquence de sous-espaces de Hilbert³ $(V_r)_{r \in \mathbb{Z}}$ vérifiant la condition suivante :

$$\{0\} \subset V_0 \subset V_1 \subset \dots \subset V_r \subset V_{r+1} \subset \dots \subset L^2(\mathbb{R}) \quad (2.4)$$

Pour chacun de ces sous-espaces, une base de fonctions, $\Phi_{r,k}$, appelées fonctions d'échelle (*scaling functions*), est définie.

Dans qui suit, on développe l'exemple présenté dans [123, 124] du cas particulier de Haar dans l'intervalle $[0, 1]$. La manière la plus simple de créer une base pour chaque sous-espace V_r est de définir une famille de fonctions par translation et dilatation d'une fonction créneau, tel que :

$$\Phi_{r,k}(t) = \Phi(2^r t - i), \quad i = 0, \dots, 2^r - 1, \quad (2.5)$$

où la fonction créneau est définie par :

$$\Phi(t) = \begin{cases} 1, & \forall t \in [0, 1[\\ 0, & \text{sinon} \end{cases} \quad (2.6)$$

La figure 2.4-gauche représente les bases de fonctions pour l'espace V_0, V_1, V_2 . Pour satisfaire la relation 2.4 d'inclusions successives des sous-espaces, toute fonction $\Phi_{r-1,k}$ doit pouvoir être exprimée par rapport aux fonctions $\Phi_{r,k}$, ce qui équivaut à dire qu'il existe une suite de coefficients $(h[i])_{i \in \mathbb{Z}}$ telle que :

$$\Phi_{r-1,k} = \sum_i h[i] \Phi_{r,i} \quad (2.7)$$

Il est facile de constater que les fonctions d'échelle définies dans l'équation 2.5 (cas de Haar), satisfont cette relation. Par exemple, $\Phi_{0,0} = 1 \times \Phi_{1,0} + 1 \times \Phi_{1,1}$ et $\Phi_{1,1} = 0 \times \Phi_{2,0} + 0 \times \Phi_{2,1} + 1 \times \Phi_{2,2} + 1 \times \Phi_{2,3}$. Ainsi, il est possible de relier les fonctions d'échelle de résolutions différentes grâce à un filtre discret h .

Bien évidemment, un signal quelconque défini sur une base V_r ne pourra pas, en général, être exprimé sur la base grossière associée V_{r-1} sans perte d'informations. Dans ce cas, les ondelettes sont introduites naturellement comme le moyen de stocker les différences entre deux espaces consécutifs V_r et V_{r-1} . Ainsi, il est possible de définir un espace de fonctions, W_{r-1} , orthogonalement complémentaire de V_{r-1} dans V_r tel que :

$$V_r = V_{r-1} \oplus W_{r-1} \quad (2.8)$$

Étant donné que l'espace W_{r-1} est par définition un sous-espace de V_r , il existe

³espace vectoriel de fonctions continues muni d'un produit scalaire

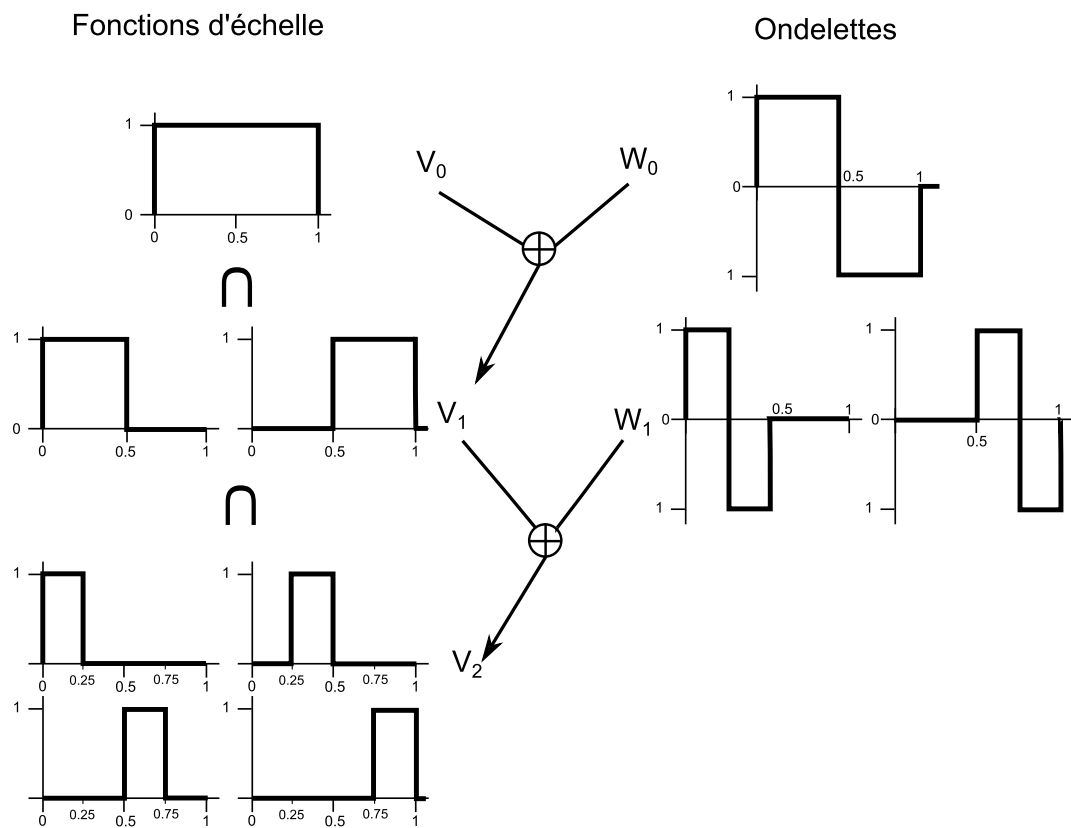


FIGURE 2.4 – Bases d'ondelettes (droite) et de fonctions d'échelle (gauche). Relation entre les sous-espaces, V_0 , W_0 , V_1 , W_1 et V_2

une séquence $(g)_{k \in \mathbb{Z}}$ de telle sorte que :

$$\psi_{r-1,k} = \sum_i g[i] \psi_{r,i} \quad (2.9)$$

Avant de pouvoir trouver cette séquence il est nécessaire de définir les ondelettes $\psi_{r-1,k}$. D'après l'équation 2.8, il apparaît que les ondelettes sont contraintes par la forme des fonctions d'échelle, étant donné qu'elles doivent être orthogonales à ces dernières. Malheureusement, il existe une multitude de fonctions satisfaisant cette condition et d'autres contraintes doivent être prises en compte. Par exemple, il est généralement souhaitable d'avoir des ondelettes avec un support minimum, ce qui permet de choisir une solution unique [124]. Les ondelettes de Haar associées aux fonctions d'échelle définies par l'équation 2.5 sont données par :

$$\psi_{r,k}(t) = \psi(2^r t - i), \quad i = 0, \dots, 2^r - 1, \quad (2.10)$$

où la fonction $\psi(t)$ est définie par :

$$\psi(t) = \begin{cases} 1, \forall t \in [0, \frac{1}{2}[\\ 1, \forall t \in [\frac{1}{2}, 1[\\ 0, \text{sinon} \end{cases} \quad (2.11)$$

La figure 2.4-droite représente les ondelettes de Haar des espaces W_1, W_2 . Il est simple de vérifier que toute fonction de W_{r-1} est orthogonale à toute fonction de V_{r-1} et que toute fonction de V_r peut être exprimée par une combinaison linéaire de fonctions de $V_{r-1} \oplus W_{r-1}$.

L'équation 2.9 fait apparaître un deuxième filtre discret, g , qui relie les ondelettes aux fonctions d'échelle. Les relations établies dans les équations 2.7 et 2.9, parfois appelées équations des deux échelles, sont simples mais ont une importance capitale, notamment dans le calcul des filtres associés à la transformée rapide en ondelettes.

Décompositions récursives

Ainsi, la transformée en ondelettes d'un signal $x(t)$ discrétisé en $\lambda_{r+1,k}$ initialement décrit dans la base V_{r+1} , est algorithmé itératif qui consiste à remplacer les composants de V_{r+1} par une représentation dans $V_r \oplus W_r$ jusqu'à une échelle

grossière r_{min} , telles que :

$$\begin{aligned}
 & V_{r+1} \\
 & = \\
 & V_r \oplus W_r \\
 & = \\
 & V_{r-1} \oplus W_{r-1} \oplus W_r \\
 & = \\
 & \dots \\
 & = \\
 & V_{r_{min}} \oplus W_{r_{min}} \oplus \dots \oplus W_r,
 \end{aligned} \tag{2.12}$$

ce qui revient à faire des approximations successives du signal $\lambda_{r+1,k}$ par projection sur les bases $V_{r,k}$ et de sauvegarder les composantes perdues grâce aux coefficients d'ondelettes $\gamma_{r,k}$ associés à la base $W_{r,k}$ tel que :

$$\begin{aligned}
 \lambda_{r+1} &= \sum_k \lambda_{r+1,k} \Phi_{r+1,k} \\
 &= \sum_k \lambda_{r,k} \Phi_{r,k} + \sum_k \gamma_{r,k} \psi_{r,k} \\
 &= \sum_k \lambda_{r-1,k} \Phi_{r,k} + \sum_k \gamma_{r-1,k} \psi_{r-1,k} + \sum_k \gamma_{r,k} \psi_{r,k} \\
 &\quad \dots \\
 &= \sum_k \lambda_{r_{min},k} \Phi_{r_{min},k} + \sum_k \gamma_{r_{min},k} \psi_{r_{min},k} + \dots + \sum_k \gamma_{r,k} \psi_{r,k}
 \end{aligned} \tag{2.13}$$

On nomme « phase d'analyse » la décomposition d'un signal (approximé ou non) $\lambda_{r,k}$ en coefficients d'ondelettes $\gamma_{r-1,k}$ et d'échelle $\lambda_{r-1,k}$. Inversement, on nomme « phase de synthèse » la reconstruction d'un signal à partir des coefficients d'ondelettes et d'échelle. Afin de procéder efficacement à la décomposition et reconstruction d'un signal, on peut démontrer que l'analyse et la synthèse peuvent être associées à des filtres discrets, et que la transformée en ondelettes peut être effectuée par applications successives de ces filtres aux coefficients d'échelle $\lambda_{r,k}$, sans avoir à construire explicitement les ondelettes.

Comme les fonctions d'échelle et les ondelettes sont ici supposées orthogonales au sens strict (ondelettes de Haar), on peut calculer les coefficients $\lambda_{r,k}$ et

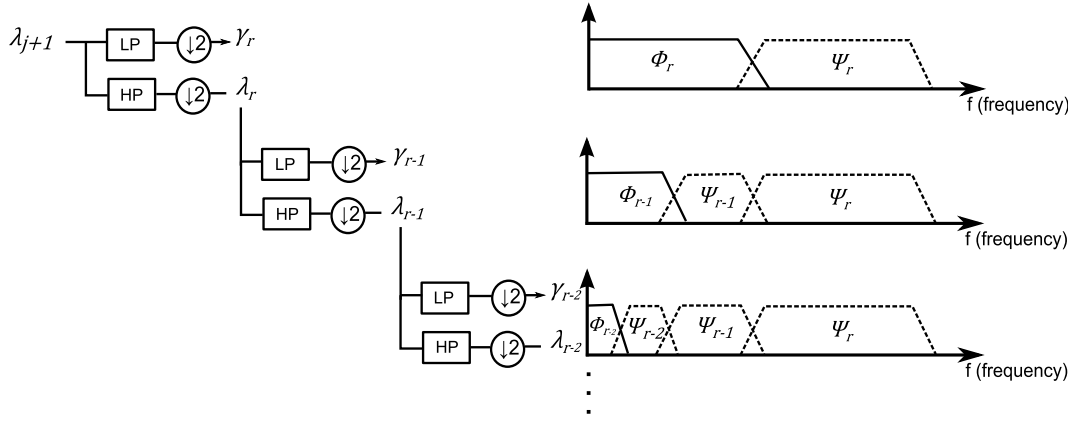


FIGURE 2.5 – Algorithme de décomposition en cascade (*filter bank*). Gauche : décomposition en coefficients d'échelle et d'ondelettes par application successive de filtres passe-bas (LP, correspond à h pour la transformée de Haar) et passe-bande (HP, correspond à g pour la transformée de Haar). Droite : couverture spectrale des fonctions d'échelle et d'ondelettes

$\gamma_{r,k}$ en utilisant le produit scalaire :

$$\begin{aligned}
 \lambda_{r-1,k} &= \langle x(t), \Phi_{r-1,k}(t) \rangle \\
 &= \left\langle \sum_n \lambda_{r,n} \Phi_{r,n}(t), \Phi_{r-1,k}(t) \right\rangle \\
 &= \sum_n \lambda_{r,n} \langle \Phi_{r,n}(t), \Phi_{r-1,k}(t) \rangle
 \end{aligned} \tag{2.14}$$

$$\begin{aligned}
 \gamma_{r-1,k} &= \langle x(t), \psi_{r-1,k}(t) \rangle \\
 &= \left\langle \sum_n \lambda_{r,n} \Phi_{r,n}(t), \psi_{r-1,k}(t) \right\rangle \\
 &= \sum_n \lambda_{r,n} \langle \Phi_{r,n}(t), \psi_{r-1,k}(t) \rangle
 \end{aligned} \tag{2.15}$$

En utilisant les équations 2.7 et 2.9 ainsi que la propriété d'orthogonalité, on peut montrer [24] que :

$$\langle \Phi_{r,n}(t), \Phi_{r-1,k}(t) \rangle = h(n - 2k) \tag{2.16}$$

$$\langle \Phi_{r,n}(t), \psi_{r-1,k}(t) \rangle = g(n - 2k) \tag{2.17}$$

et ainsi,

$$\lambda_{r-1,k} = \sum_n h(n - 2k) \lambda_{r,k} \tag{2.18}$$

$$\gamma_{r-1,k} = \sum_n g(n - 2k) \lambda_{r,k} \tag{2.19}$$

Ces relations directes entre les coefficients d'ondelettes et d'échelle au niveau $r - 1$ et les coefficients d'échelle de niveau r permet une décomposition récur-

sive efficace ($O(n)$) en cascade, comme présentée dans la figure 2.5. De façon similaire, on peut démontrer que la phase de synthèse (reconstruction) peut également s'exprimer en fonction des filtres discrets h et g ⁴, tel que :

$$\lambda_{r,k} = \sum_n \left(h(n-2k)\lambda_{r-1,k} + \sum_n g(n-2k)\lambda_{r-1,k} \right) \quad (2.20)$$

Fonctions duales

Dans le cas des bases orthogonales, les coefficients des filtres d'analyse et de synthèse sont les mêmes. Néanmoins, il existe un certain nombre de familles d'ondelettes qui ne sont pas orthogonales au sens strict (par exemple les ondelettes de spline, les « chapeaux mexicains », ...). Dans ce cas, deux couples d'ondelettes et de fonctions d'échelle duales, $(\tilde{\psi}, \tilde{\Phi})$ et (ψ, Φ) , sont utilisées respectivement pour la phase d'analyse (décomposition) et de synthèse (reconstruction). Deux filtres duaux peuvent être déduits de ces fonctions duales afin de pouvoir suivre une transformation/reconstruction en cascade.

Intuitivement, les ondelettes et les filtres associés doivent être les mêmes dans le cas orthogonal étant donné que la transformation linéaire inverse est donnée par la transposée de la base. Lorsque la base n'est pas orthogonale, il est clair que la transformée inverse n'aura pas la même forme que la transformée directe. Une famille particulière d'ondelettes non orthogonales, mais dites « bi-orthogonales », fait l'objet d'études plus approfondies dans la section suivante.

2.3.2 Lifting en ondelettes et ondelettes de seconde génération

Le cadre théorique et la terminologie des ondelettes ont été introduits dans la section précédente. Nous avons vu que la puissance des transformées en ondelettes vient de l'analyse temps-échelle d'un signal. Néanmoins, l'application des ondelettes de première génération peut s'avérer difficile dans certains cas pratiques : la conception des ondelettes de première génération n'est pas originellement adaptée aux signaux finis, ne peut pas gérer un échantillonnage irrégulier et son cadre théorique peut être difficile à appréhender pour une personne non initiée à la théorie de Fourier et à l'analyse multi-résolutions.

Le *lifting* en ondelettes (*Lifting scheme*) introduit par Sweldens [128] est une méthode permettant de construire de nouvelles ondelettes, dites de seconde génération. Ce processus est utilisé pour décomposer efficacement un signal sur une base formée d'ondelettes non stationnaires, sans avoir à définir ces dernières explicitement. La différence majeure entre les deux formulations vient du fait que la construction des ondelettes de deuxième génération ne repose plus sur la transformation de Fourier et que ces ondelettes ne respectent plus les propriétés d'invariance par translation et dilatation, ce qui permet une plus grande flexibilité lors de leur création. Ainsi, les ondelettes de seconde génération peuvent être adaptées en fonction de l'échantillonnage du signal, de sa localisation (proche des bords, ...) et peuvent prendre en compte des contraintes extérieures (via l'utilisation d'un produit scalaire pondéré). De plus, les transformations par *lifting* peuvent être jusqu'à deux fois plus efficaces que celles

⁴dans le cas de bases orthogonales

basées sur la convolution. Les calculs étant entièrement effectués « sur place », le *lifting* en ondelettes ne nécessite pas de mémoire auxiliaire. Finalement, si le cadre théorique reste très spécialisé, l'introduction au *lifting* en ondelettes est plus naturelle, ce qui permet une meilleure compréhension et une implémentation plus facile des transformations.

Tout le cadre théorique présenté dans la section précédente reste valide pour les ondelettes de seconde génération. Néanmoins, contrairement à l'exemple développé précédemment, le *lifting* en ondelettes s'appuie sur des bases d'ondelettes non orthogonales (au sens strict), mais bi-orthogonales qui possèdent des propriétés permettant la construction d'ondelettes à support fini. Ainsi, la formulation du *lifting* en ondelettes et des ondelettes de deuxième génération s'appuie sur l'étude de couples de filtres duaux et de fonctions duales, $(h, g)/(\tilde{h}, \tilde{g})$ et $(\psi, \Phi)/(\tilde{\psi}, \tilde{\Phi})$.

Ondelettes et filtres bi-orthogonaux

On dénote par \tilde{V}, \tilde{W} les sous-espaces définis par les fonctions duales $\tilde{\Phi}, \tilde{\psi}$. Les espaces $(\tilde{V}_r)_{r \in \mathbb{Z}}$ respectent les mêmes relations d'inclusions successives que $(V_r)_{r \in \mathbb{Z}}$ avec $\tilde{V}_{r+1} = \tilde{V}_r \oplus \tilde{W}_r$ et $V_{r+1} = V_r \oplus W_r$. Néanmoins, les bases duales ne sont pas orthogonales ($V_r \not\perp W_r$ et $\tilde{V}_r \not\perp \tilde{W}_r$), mais bi-orthogonales, c.-à-d :

$$V_r \perp \tilde{W}_r, \tilde{V}_r \perp W_r \quad (2.21)$$

$$\langle \Phi_{r,k}, \tilde{\Phi}_{r,k'} \rangle = \langle \psi_{r,k}, \tilde{\psi}_{r,k'} \rangle = \delta_{k,k'}, \quad (2.22)$$

où $\delta_{k,k'}$ une fonction de Dirac qui prend la valeur 1 lorsque $k = k'$ et 0 partout ailleurs. De manière équivalente, l'équation 2.22 ainsi que les relations d'inclusion des sous-espaces impliquent :

$$\begin{aligned} \sum_l g_{r,m,l} \tilde{g}_{r,m',l} &= \delta_{m,m'} \\ \sum_l h_{r,k,l} \tilde{h}_{r,k',l} &= \delta_{k,k'} \\ \sum_l h_{r,k,l} \tilde{g}_{r,m,l} &= 0 \\ \sum_l g_{r,m,l} \tilde{h}_{r,k,l} &= 0 \end{aligned} \quad (2.23)$$

Il est à noter que $g_{r,m,l} = g_{l-2k}$, $h_{r,m,l} = h_{l-2k}$, $\tilde{h}_{r,m,l} = \tilde{h}_{l-2k}$, $\tilde{g}_{r,m,l} = \tilde{g}_{l-2k}$ dans le cas des ondelettes stationnaires de première génération. Ceci n'est plus valable dans le cas des ondelettes non stationnaires de seconde génération. Ainsi, lors du *lifting* en ondelettes, les filtres appliqués aux coefficients d'échelle $\lambda_{r+1,k}$ et d'ondelettes $\gamma_{r+1,m}$ dépendent de la position et de l'échelle, et la transformée

rapide est donnée par l'application récursive des filtres :

$$\begin{aligned}\lambda_{r,k} &= \sum_l \tilde{h}_{r,k,l} \lambda_{r+1,l} \\ \gamma_{r,m} &= \sum_l \tilde{g}_{r,m,l} \lambda_{r+1,l},\end{aligned}\tag{2.24}$$

et de manière similaire, la transformée inverse est donnée par :

$$\lambda_{r+1,l} = \sum_k h_{r,k,l} \lambda_{r,k} + \sum_m g_{r,m,l} \lambda_{r,m}\tag{2.25}$$

Afin d'améliorer la lisibilité, nous utilisons les opérateurs de filtrage $H_r, G_r, \tilde{H}_r, \tilde{G}_r$ définis dans [128] qui intègrent l'ensemble des filtres pour la résolution r , de telle sorte que les équations 2.24 et 2.25 s'écrivent sous la forme :

$$\lambda_r = \tilde{H}_r \lambda_{r+1}, \quad \gamma_r = \tilde{G}_r \lambda_{r+1}\tag{2.26}$$

$$\lambda_{r+1} = H_r^T \lambda_r + G_r^T \gamma_r\tag{2.27}$$

Avec ces notations, la condition de bi-orthogonalité (équation 2.23) peut s'écrire sous la forme matricielle :

$$\begin{aligned}\tilde{H}_r H_r^T &= \tilde{G}_r G_r^T = 1, \quad \tilde{G}_r H_r^T = \tilde{H}_r G_r^T = 0 \\ H_r^T \tilde{H}_r + G_r^T \tilde{G}_r &= 1\end{aligned}\tag{2.28}$$

où H_r^T et G_r^T sont les applications transposées de H_r et G_r .

Lifting en ondelettes

Le *lifting* en ondelettes et la construction d'ondelettes de seconde génération reposent sur le théorème suivant ([128]) :

Théorème 1. (Lifting) *Soit $\{H_r^{old}, \tilde{H}_r^{old}, G_r^{old}, \tilde{G}_r^{old}\}$ un ensemble initial d'opérateurs bi-orthogonaux de filtrage. Alors, il est possible de trouver un nouvel ensemble d'opérateurs bi-orthogonaux de filtrage $\{H_r, \tilde{H}_r, G_r, \tilde{G}_r\}$ tel que :*

$$\begin{aligned}H_r &= H_r^{old} \\ \tilde{H}_r &= \tilde{H}_r^{old} + S_r \tilde{G}_r^{old} \\ G_r &= G_r^{old} - S_r^T H_r^{old} \\ \tilde{G}_r &= \tilde{G}_r^{old}\end{aligned}\tag{2.29}$$

où S_r est un opérateur de $L^2(\mathbb{M})(r)$ dans $L^2(\mathbb{K})(j)$, avec $\mathbb{M}(r)$ et $\mathbb{K}(j)$ représentant des ensembles d'indices pour la résolution r .

Preuve :

Sous forme matricielle, le système 2.29 s'écrit :

$$\begin{bmatrix} \tilde{H}_r \\ \tilde{G}_r \end{bmatrix} = \begin{bmatrix} 1 & S_r \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \tilde{H}_r^{old} \\ \tilde{G}_r^{old} \end{bmatrix}; \quad \begin{bmatrix} H_r \\ G_r \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -S_r^T & 1 \end{bmatrix} \begin{bmatrix} H_r^{old} \\ G_r^{old} \end{bmatrix}$$

alors les conditions de bi-orthogonalité 2.28 sont vérifiées par :

$$\begin{bmatrix} \tilde{H}_r \\ \tilde{G}_r \end{bmatrix} [H_r^T, G_r^T] = \begin{bmatrix} 1 & S_r \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -S_r \\ 0 & 1 \end{bmatrix}$$

et

$$\begin{aligned} [H_r^T, G_r^T] \begin{bmatrix} \tilde{H}_r \\ \tilde{G}_r \end{bmatrix} &= [H_r^{oldT}, G_r^{oldT}] \begin{bmatrix} 1 & -S_r \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & S_r \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \tilde{H}_r^{old} \\ \tilde{G}_r^{old} \end{bmatrix} \\ &= [H_r^{oldT}, G_r^{oldT}] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \tilde{H}_r^{old} \\ \tilde{G}_r^{old} \end{bmatrix} = 1 \end{aligned}$$

l'ondelette d'origine étant orthogonale.

L'effet du *lifting* sur les opérateurs a des répercussions sur la forme des ondelettes, et seule la fonction primaire d'échelle, Φ_r , reste inchangée (l'ondelette duale est modifiée par la relation d'affinement 2.9; voir [128] pour plus de précisions. La puissance du *lifting* vient du fait qu'il est possible de créer des ondelettes à filtres bi-orthogonaux à partir d'un ensemble de fonctions d'échelle, uniquement en adaptant l'opérateur S_r , de façon à ce que les ondelettes (primaires) résultantes répondent à certaines spécifications. Par exemple, il est possible de choisir S_r afin de contraindre l'ordre de l'ondelette générée (nombre de moments nuls (*vanishing moments*)).

De façon similaire, il est possible d'améliorer les ondelettes duales en appliquant un *lifting* dual sur les filtres H_r, \tilde{G}_r via l'opérateur de *lifting* \tilde{S}_r tel que :

$$\begin{aligned} H_r &= H_r^{old} + \tilde{S}_r \tilde{G}_r^{old} \\ \tilde{H}_r &= \tilde{H}_r^{old} \\ G_r &= G_r^{old} \\ \tilde{G}_r &= \tilde{G}_r^{old} - \tilde{S}_r^T H_r^{old} \end{aligned} \tag{2.30}$$

De même que le *lifting*, le *lifting* dual des opérateurs de filtrage a des répercussions sur la forme des ondelettes (duales et primaires), et seule la fonction d'échelle duale, $\tilde{\Phi}_r$, reste inchangée. Ainsi, comme pour le *lifting*, l'opérateur \tilde{S}_r peut être adapté afin de répondre à certaines spécifications. Lifting dual et *lifting* peuvent s'appliquer successivement, sans annulation des effets recherchés, malgré les impacts respectifs qu'ils peuvent avoir sur les ondelettes duales et primaires (voir [128]).

Split, Predict, Update

En pratique, il n'est pas nécessaire de calculer explicitement les ondelettes produites par le *lifting*. Le cadre théorique assure que les opérations successives (*lifting* et *lifting* dual) sur une ondelette de base permet d'obtenir un ensemble d'ondelettes bi-orthogonales dont on peut contrôler les propriétés.

Dans ce qui suit, nous nous intéressons au *lifting* de deux ondelettes initiales : l'ondelette paresseuse (*Lazy wavelet*) et l'ondelette de Haar. L'ondelette paresseuse, comme son nom l'indique, ne fait pas grand-chose. Elle se contente de séparer les coefficients $\lambda_{r,k}$ d'indices pairs et impairs à l'aide de deux opérateurs de sous échantillonnage E et D , tel que :

$$\begin{aligned}\lambda_{r,2k} &= E\lambda_{r,k} \\ \lambda_{r,2k+1} &= D\lambda_{r,k}\end{aligned}\tag{2.31}$$

On peut facilement démontrer que ces opérateurs répondent aux conditions d'orthogonalité (et par conséquence de bi-orthogonalité) dont les filtres sont donnés par :

$$\begin{aligned}H_r^{old} &= \tilde{H}_r^{old} = E \\ G_r^{old} &= \tilde{G}_r^{old} = D\end{aligned}\tag{2.32}$$

Étant donné que l'objet principal de cette ondelette est de découper un signal en deux, l'ondelette paresseuse est appelée opérateur de découpage ou *Split operator*. Il est à noter que cette ondelette d'ordre 0 (aucun moment nul) produit une bien mauvaise décomposition. Il est possible d'imager d'autres opérateurs de découpages (autres que pair/impair) mieux adaptés à la nature du signal à décomposer. En général, il est souhaitable que le découpage sépare les coefficients initiaux en deux séquences corrélées afin d'obtenir la meilleure approximation du signal (c.-à-d la meilleure compression) après application du *lifting* dual.

Le *lifting* dual (opérateur \tilde{S}) a pour but d'augmenter le nombre de moments nuls N de l'ondelette duale. L'ondelette duale mesure la différence entre l'approximation polynômiale de degré $N - 1$ donnée par la projection d'un signal sur la base de fonction d'échelle V_r et sa représentation dans l'espace V_{r+1} . En pratique, plus cette projection donne une bonne prédiction (approximation) du signal, plus la compression du signal sera bonne puisque les coefficients produits seront négligeables (proche de zéro). Ainsi, l'opérateur de *lifting* dual \tilde{S} est souvent appelé opérateur de Prédiction (*Predict operator*) car il permet de prédire (approximer) les valeurs des coefficients λ_{r+1} à partir des coefficients λ_r générés par l'application des ondelettes initiales (*Lazy* ou Haar), ce qui revient à lifter le couple de filtres (H, \tilde{G}) . En conséquence, l'étape de prédiction peut s'écrire :

$$\gamma_{r,m} = \gamma_{r,m}^{old} - \mathcal{P}(\lambda_{r,k}^{old}),\tag{2.33}$$

où l'opérateur $\mathcal{P}(\cdot)$ exprime l'effet du *lifting* dual sur les coefficients $\gamma_{r,m}^{old}$. L'ordre de l'ondelette duale est lié au nombre de coefficients $\lambda_{r,k}^{old}$ utilisés dans la pré-

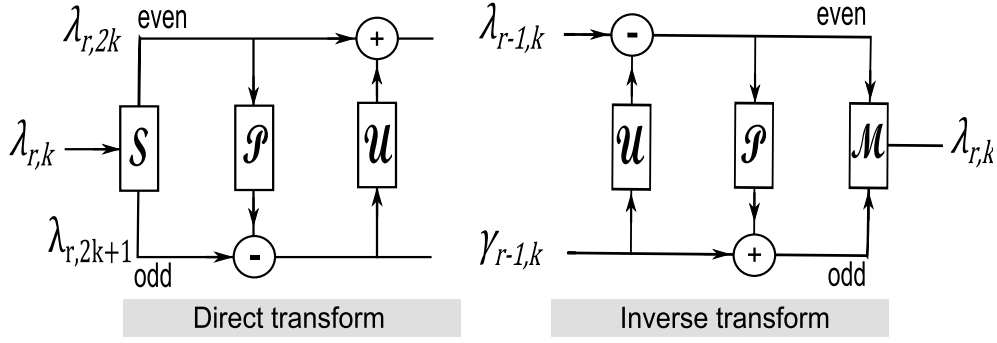


FIGURE 2.6 – Schéma de *lifting* en ondelettes direct et inverse. Découpage (Split), Prédiction (P) et mise à jour (U). Une étape de décomposition et reconstruction de coefficients d'échelle r .

diction. Ainsi, lorsque N coefficients sont utilisés, il est possible de prédire les coefficients $\gamma_{r,m}^{old}$ en construisant un polynôme de degré $N-1$, ce qui veut dire que l'ondelette duale est d'ordre N . L'opérateur de prédiction étant généralement linéaire (il doit être au moins inversible), il permet de générer une séquence de coefficients de filtre (*filter coefficients*) associés à la méthode d'interpolation utilisée.

De même, le *lifting* est souvent utilisé pour augmenter le nombre de moments nuls \tilde{N} de l'ondelette (primaire), et ainsi préserver les \tilde{N} moments des séquences successives $\lambda_{r,k}$. Cette étape permet d'éviter les effets d'*aliasing*. Par exemple, l'ordre primaire de l'ondelette paresseuse est $\tilde{N} = 0$ et la séquence $\lambda_{r,k}$ n'est qu'un sous-échantillonnage de la séquence $\lambda_{r,k}$, ce qui donne une piètre approximation du signal. Ainsi, l'opérateur de *lifting* S est appelé opérateur de mise à jour (*Update operator*) car il permet d'améliorer l'approximation λ_r de la séquence λ_{r+1} . En conséquence, l'étape de mise à jour peut s'écrire :

$$\lambda_{r,k} = \lambda_{r,k}^{old} - \mathcal{U}(\gamma_{r,m}), \quad (2.34)$$

où l'opérateur $\mathcal{U}(\cdot)$ exprime l'effet du *lifting* primaire sur les coefficients d'échelle $\lambda_{r,k}^{old}$. L'ordre de l'ondelette primaire est lié au nombre de moments préservés. Une séquence de coefficients, appelés coefficients de *lifting* (*Lifting coefficients*), peut être déduite par la résolution d'un système de dimension $\tilde{N} \times \tilde{N}$ correspondant à la contrainte de préservation des moments. Une implémentation efficace de calcul de ces coefficients est décrite dans [55], pour des ondelettes basées sur l'interpolation polynomiale. Lorsque l'ondelette primaire initiale possède déjà un moment nul, il n'est pas obligatoire d'appliquer l'opérateur de mise à jour (voir Interpolation de la moyenne 2.3.2).

Les opérateurs *Split*, *Predict* et *Update* étant inversibles, les phases d'analyse et de synthèse du *lifting* en ondelettes sont souvent représentées par les diagrammes de la figure 2.6. Il est important de préciser qu'il s'agit ici du cas spécial du *lifting* primaire et dual de l'ondelette paresseuse. Dans ce qui suit, nous analysons l'évolution des coefficients lors du schéma de *lifting* afin de comprendre concrètement l'effet des différents opérateurs sur les coefficients.

***Lifting* et ondelettes de Haar déséquilibrées**

L'ondelette de Haar est une ondelette d'ordre 1 connue pour sa simplicité et ses propriétés d'orthogonalité. Nous proposons dans cette section de développer l'exemple des ondelettes bi-orthogonale de Haar déséquilibrées, construites à partir du *lifting* de l'ondelette paresseuse. Cette transformée prend en compte l'échantillonnage du signal ; le but de ces ondelettes est d'obtenir des approximations pondérées qui prennent en compte de l'importance des différents échantillons.

Considérons une séquence initiale de coefficients $\lambda_{r+1,k}$ de la base V_{r+1} . Imaginons que ces valeurs représentent des moyennes locales sur des intervalles irréguliers $I_{r+1,k}$. Par exemple, $\lambda_{r+1,k}$ pourrait représenter une propriété dans une grille stratigraphique à cellules de volumes variables.

La première phase du schéma de *lifting* est d'appliquer les filtres de l'ondelette paresseuse :

$$\begin{aligned}\lambda_{r,k}^{old} &= E\lambda_{r+1,k} = \lambda_{r+1,2k} \\ \gamma_{r,m}^{old} &= D\lambda_{r+1,m} = \lambda_{r+1,2k+1}\end{aligned}\tag{2.35}$$

L'ordre de l'ondelette duale est ensuite augmenté par *lifting* afin de pouvoir approximer le signal grâce à des polynômes constants. Pour ce faire, il est possible d'utiliser les coefficients d'échelle $\lambda_{r,k}^{old}$ afin de modifier les coefficients $\gamma_{r,m}^{old}$ (cf. *lifting* dual 2.30). Le but de l'opérateur de prédiction est d'obtenir la meilleure approximation possible des coefficients de la séquence $\gamma_{r,m}^{old}$ en utilisant N coefficients de la séquence $\lambda_{r,k}^{old}$. Dans le cas de l'ondelette de Haar, $N = 1$, ce qui signifie que la prédiction des coefficients impairs repose sur la construction d'un polynôme constant. La valeur du polynôme pour l'échantillon impair est ensuite soustraite aux coefficients $\gamma_{r,m}^{old}$:

$$\gamma_{r,m} = \gamma_{r,m}^{old} - \lambda_{r,k}^{old}\tag{2.36}$$

Ainsi, si le signal est localement constant, le coefficient $\gamma_{r,m}$ sera nul. Inversement, si le coefficient $\gamma_{r,m} = 0$, la fonction primaire d'échelle permettra la reconstruction d'un polynôme constant.

Le but de l'ondelette déséquilibrée de Haar est de préserver la moyenne pondérée à chaque approximation, tel que :

$$\lambda_{r,k} = \frac{I_{r+1,2k} \cdot \lambda_{r+1,2k} + I_{r+1,2k+1} \cdot \lambda_{r+1,2k+1}}{I_{r,k}}\tag{2.37}$$

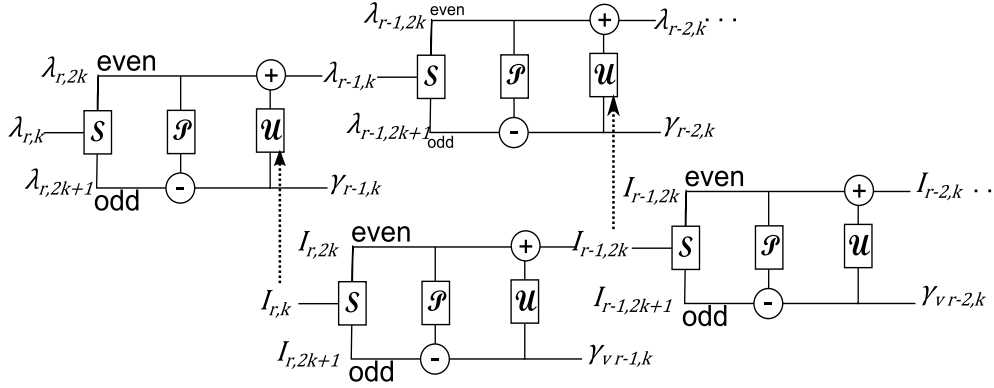


FIGURE 2.7 – Algorithme de la transformée de Haar déséquilibrée. L’opérateur de mise à jour à la résolution r nécessite les intervalles $I_{r,k}$ donnés par la décomposition de la propriété d’importance des intervalles (c.-à-d. les volumes des cellules). Dans cette thèse, nous proposons d’effectuer une transformée standard de Haar de l’extension des intervalles (p. ex. volumes des cellules) en parallèle de la transformée déséquilibrée afin d’obtenir les pondérations utilisées dans l’étape d’*Update* à toutes les échelles.

avec $I_{r,k} = I_{r+1,2k} + I_{r+1,2k+1}$, ce qui est accompli grâce à la mise à jour suivante :

$$\begin{aligned} \lambda_{r,k} &= \lambda_{r,k}^{old} + \frac{I_{r+1,2k+1} \times \gamma_{r,m}}{I_{r,k}} = \\ &= \frac{(I_{r+1,2k} + I_{r+1,2k+1})\lambda_{r+1,2k} + I_{r+1,2k+1}(\lambda_{r+1,2k+1} - \lambda_{r+1,2k})}{I_{r,k}} = \\ &= \frac{I_{r+1,2k} \cdot \lambda_{r+1,2k} + I_{r+1,2k+1} \cdot \lambda_{r+1,2k+1}}{I_{r,k}} \end{aligned} \quad (2.38)$$

Au lieu de calculer directement les intervalles $I_{r,k}$ à partir de la grille d’échantillonnage, il est possible d’effectuer une transformation (reconstruction) standard de Haar de la propriété d’importance de l’intervalle (longueur (1D), surface (2D) ou volume (3D)) en parallèle de la transformée déséquilibrée de Haar comme représenté par la figure 2.7.

La transformée déséquilibrée de Haar s’avère très utile lorsque des grilles stratigraphiques sont utilisées. Elle permet d’obtenir de meilleures approximations aux différentes échelles (moyennes pondérées par les volumes des cellules) et permet de gérer les cellules inactives (mortes ou à volumes négligeables) sans augmenter le nombre de paramètres actifs. En effet, les valeurs de ces cellules n’entrant pas dans le calcul de la moyenne, les coefficients d’ondelettes associés peuvent être négligés.

Cependant, les ondelettes de Haar étant discontinues, des artefacts en forme de blocs peuvent apparaître lors de reconstructions partielles ou lorsque les coefficients sont perturbés (figure 2.8-b). Il est néanmoins possible de résoudre ce problème en augmentant l’ordre de l’ondelette duale par *lifting* et ainsi obtenir des reconstructions plus lisses des propriétés (figure 2.8-c, Section 2.3.2).

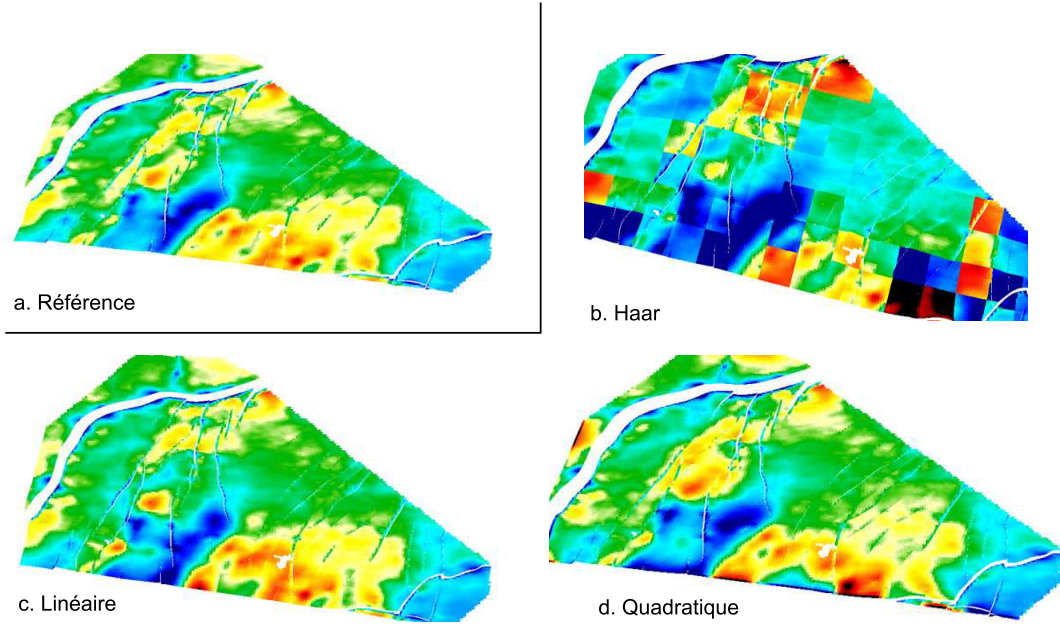


FIGURE 2.8 – Reconstruction après perturbation des coefficients d'ondelette. (a) Propriété initiale. Décomposition et reconstruction en utilisant (b) l'ondelette de Haar, (c) l'ondelette basée sur l'interpolation linéaire, (d) l'ondelette basée sur l'interpolation quadratique (Haar améliorée)

Ondelettes basées sur l'interpolation : exemple de l'ondelette "Chapeau mexicain"

Les ondelettes basées sur l'interpolation appliquent le(s) *lifting* (s) sur l'ondelette paresseuse et utilisent les méthodes d'interpolation - subdivision [41] lors de l'étape de prédiction, durant laquelle $\tilde{N} = 2D$, $D \in \mathbb{N}$ coefficients $\lambda_{r,k}^{old}$ sont utilisés pour construire un polynôme de degré impair (linéaire, cubique, ...). La figure 2.9 illustre la décomposition d'un signal 1D échantillonné de façon régulière, par une ondelette construite à partir de l'interpolation linéaire ($N = 2, \tilde{N} = 2$). La première étape d'analyse est présentée en détails. De (a) à (b), le signal est subdivisé en échantillons pairs et impairs par l'application de l'ondelette paresseuse. Ensuite, les échantillons impairs sont prédits à l'aide des échantillons pairs par interpolation linéaire. On peut remarquer que le coefficient $\gamma_{r-1,2}$ ne possède pas deux voisins ; la prédiction se fait alors par extrapolation des deux coefficients pairs les plus proches, ce qui affecte la forme de l'ondelette duale. Les valeurs prédites sont finalement soustraites aux coefficients impairs (b à c). Les résidus $\gamma_{r-1,m}$ reflètent ainsi les différences entre le signal et son approximation linéaire donnée par les fonctions d'échelle. Enfin, l'ondelette primaire est liftée de manière à ce que la séquence $\lambda_{r-1,k}$ préserve les premier et second moments. Les coefficients d'échelle ont été calculés selon la méthode décrite par [55].

Lorsque le signal est échantillonné de façon régulière, les coefficients de filtrage et de *lifting* n'ont besoin d'être calculés qu'une seule fois pour chacune des configurations (affectée ou non par les bords), alors qu'il est nécessaire de les calculer pour chaque position dans le cas irrégulier, ce qui peut se révé-

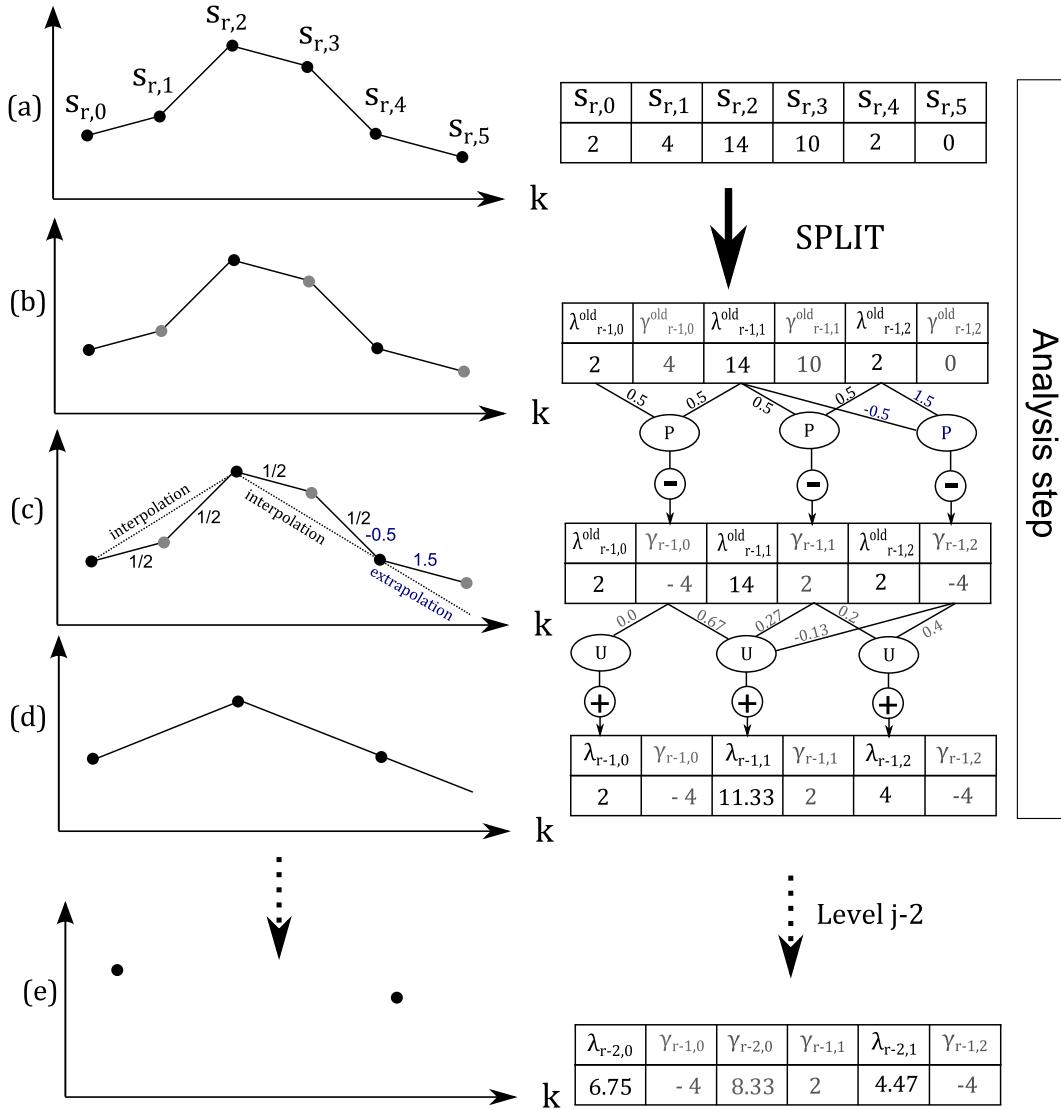


FIGURE 2.9 – Ondelette basée sur l'interpolation linéaire. Étape d'analyse de la transformée en ondelettes de seconde génération (*Lifting scheme*). Gauche : Représentation graphique du signal durant les étapes de *lifting*. Droite : évolutions numériques des coefficients. (a) Séquence initiale. (b) Décomposition de l'ondelette paresseuse. Les points noirs/gris représentent les coefficients pairs/impairs. (c) Étape de prédiction (*lifting dual*) : l'interpolation/extrapolation linéaire est utilisée pour prédire les coefficients impairs avec les coefficients pairs. Les coefficients de filtrage sont indiqués sur les branches. (d) Étape de mise à jour (*Update*) : la séquence grossière préserve les premier et second moments. Les coefficients de *lifting* sont indiqués sur les branches. (e) Décomposition complète du signal.

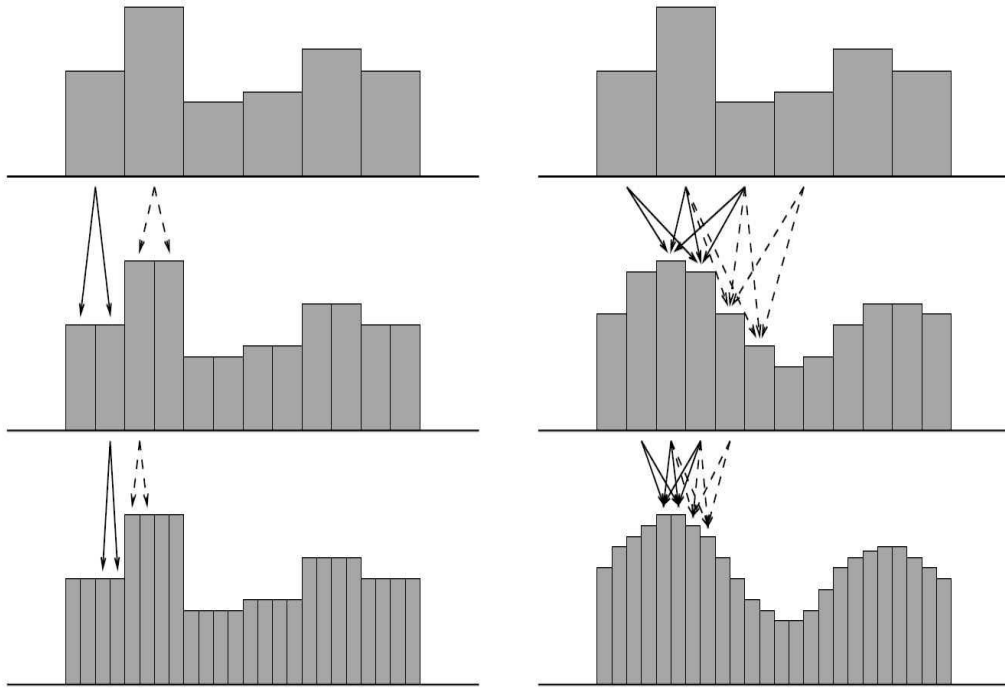


FIGURE 2.10 – Subdivision par interpolation de la moyenne. Une séquence est subdivisée en faisant l’hypothèse que chaque valeur correspond à la moyenne d’une fonction sous-jacente sur un intervalle $I_{r,k}$. Gauche : on considère que la fonction sous-jacente est une fonction constante. Droite : on considère que la fonction sous-jacente est un polynôme de degré 3. Tirée de [68]

ler coûteux. Le processus d’analyse est ensuite appliqué récursivement sur les coefficients d’échelle afin d’obtenir la décomposition complète du signal (e).

Ondelettes construites à partir de la méthode de subdivision par interpolation de la moyenne

Ces ondelettes sont construites à partir du *lifting* dual d’une ondelette de Haar. Nous considérons ici le cas où aucun *lifting* primaire (étape d’*Update*) n’est effectué : seul l’ordre N de l’ondelette duale est augmenté, tandis que l’ordre de l’ondelette primaire reste inchangé ($\tilde{N} = 1$). Ainsi, les approximations successives $\lambda_{r,k}$ préservent uniquement la moyenne (premier moment), mais la fonction d’échelle primaire est capable de reconstruire des polynômes de degré $N - 1$. L’étape de synthèse est schématisée par la figure 2.11. Bien évidemment, l’étape d’analyse nécessite d’inverser l’ordre des opérateurs et de changer les signes.

L’étape de prédiction est basée sur la méthode de subdivision par interpolation de la moyenne [44]. Cette méthode part du principe qu’un coefficient $\lambda_{r,k}$ n’est pas une observation ponctuelle d’une fonction (signal) mais sa moyenne sur un intervalle $I_{j,k} = [x_{r,k}, x_{r,k+1}]$. Le but est de trouver une approximation de cette fonction sous-jacente afin de contrôler la subdivision d’une séquence $\lambda_{r,k}$ en séquence $\lambda_{r+1,k}$. Dans cette section, nous considérons que cette fonction est

un polynôme, $p(x)$, de degré $N - 1$ (avec $N = 2D + 1$). La figure 2.10 montre la subdivision d'un signal en utilisant un polynôme constant ($N = 1$, gauche) et un polynôme de degré 2 ($N = 3$, droite). Dans le premier cas, les valeurs après subdivision restent inchangées car on considère que la fonction sous-jacente est constante, tandis que dans le deuxième cas, la subdivision reforme un polynôme de degré 2 ayant des transitions lisses. Ainsi, il est possible d'exprimer localement une séquence $\lambda_{r,k}$ telle que :

$$\begin{aligned}\lambda_{r,k-1} &= \int_{x_{r,k-1}}^{x_{r,k}} p(x) dx \\ \lambda_{r,k} &= \int_{x_{r,k}}^{x_{r,k+1}} p(x) dx \\ \lambda_{r,k+1} &= \int_{x_{r,k+1}}^{x_{r,k+2}} p(x) dx\end{aligned}\tag{2.39}$$

Étant donné que $p(x)$ est un polynôme de degré 2, sa primitive $P(x)$ sera également un polynôme. Il est possible de le trouver en résolvant le système suivant dérivé de l'équation 2.39 :

$$\begin{aligned}0 &= P(x_{j,k-1}) \\ \lambda_{j,k} &= P(x_{j,k}) \\ \lambda_{j,k} + \lambda_{j,k+1} &= P(x_{j,k+1}) \\ \lambda_{j,k} + \lambda_{j,k+1} + \lambda_{j,k+2} &= P(x_{j,k+2})\end{aligned}\tag{2.40}$$

Les ondelettes d'ordre supérieurs peuvent être construites de manière similaire, mais nécessitent la résolution d'un système de plus grande taille. Grâce à la primitive $P(x)$, il est maintenant possible de subdiviser la moyenne $\lambda_{r,k}$ telle que :

$$\begin{aligned}\lambda_{r+1,2k} &= 2(P(x_{r,k+1}/2) - P(x_{r,k})) \\ \lambda_{r+1,2k+1} &= 2(P(x_{r,k+1}) - P(x_{r,k+1}/2))\end{aligned}\tag{2.41}$$

Revenons maintenant au schéma présenté par la figure 2.11. Le but de l'opérateur de prédiction est d'évaluer les coefficients $\gamma_{r,m}^{old}$ grâce aux coefficients $\lambda_{r,k}^{old}$ produits par l'ondelette initiale. Comme l'ondelette de Haar est utilisée, les coefficients $\lambda_{r,k}$ correspondent à des moyennes locales et le processus de subdivision peut être utilisé pour estimer la différence $\gamma_{r,m}^P = \lambda_{r+1,2k+1} - \lambda_{r+1,2k}$ qui est une approximation de $\gamma_{r,m}$. Ainsi, si le signal est un polynôme de degré deux, on obtiendra $\gamma_{r,m} = \gamma_{r,m}^{old} - \gamma_{r,m}^P = 0$. Inversement, si $\gamma_{r,m} = 0$, l'étape de prédiction assurera que la reconstruction de Haar donnera bien une séquence polynomiale.

En utilisant les équations 2.40 et 2.41, il est possible d'exprimer la prédiction $\gamma_{r,m}^P$ en fonction des coefficients $\lambda_{j,k}$ et ainsi de déduire les coefficients du filtre (c.-à-d. définir \tilde{S}). Cependant, comme les polynômes sont construits à partir de positions $x_{j,k}$ quelconques, les filtres déduits seront également exprimés

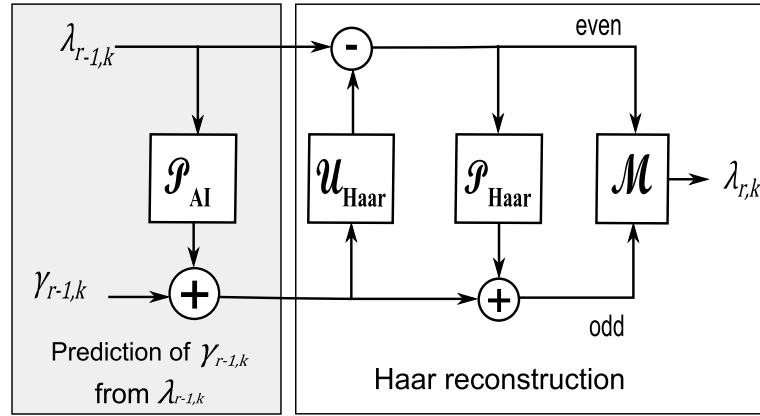


FIGURE 2.11 – Ondelettes de Haar améliorées. Ondelettes construites à partir de l’interpolation de la moyenne, étape de synthèse. Modifiée d’après [129]

en fonction de $x_{j,k}$ et devront être calculés pour chaque position, ce qui peut être coûteux. Pour cette raison, nous proposons dans cette thèse de suivre une approche différente.

Dans la Section 2.3.2, nous avons introduit l’algorithme de la transformée de Haar déséquilibrée qui permet de calculer efficacement des coefficients d’échelle correspondant aux moyennes locales, une propriété que l’on désire ici conserver. Ainsi, au lieu de calculer les coefficients du filtre à chaque position, il est possible de procéder à une transformée de Haar déséquilibrée et considérer ensuite que les coefficients d’échelle ont été calculés sur une grille régulière. Les filtres ne sont alors définis que pour un nombre limité de configurations (bords, centre) et la transformée préserve bien les moyennes pondérées des coefficients d’échelle. De plus, les cellules à volumes nuls n’entrant pas dans les calculs des coefficients, ils sont toujours égaux à zéro, ce qui évite une augmentation du nombre de paramètres actifs après la transformation en ondelettes.

2.3.3 Décompositions multidimensionnelles

Les exemples de décomposition présentés jusqu’ici sont appliqués à un signal 1D. Pour des signaux 2D ou 3D, il est possible d’utiliser des ondelettes multidimensionnelles [4] ou effectuer des décompositions indépendantes pour chaque direction. La première approche reste encore peu utilisée dans les applications car de telles ondelettes sont plus difficiles à créer et peuvent être moins efficaces en termes de temps de calcul. L’avantage des ondelettes 2D et 3D vient du fait qu’elles peuvent être capables de mieux capturer certains motifs ou anisotropies exprimés par le signal, ce que conduit à une meilleure compression. Elles peuvent également être utilisées sur des grilles irrégulières.

La deuxième approche, qui est utilisée dans ce travail, est beaucoup exploitée en compression d’images pour sa simplicité d’implémentation et de son efficacité. Un signal 3D (ou 2D) est divisé en plusieurs signaux 1D alignés sur le premier axe de la grille. Une phase d’analyse à l’échelle r est ensuite appliquée indépendamment à chaque signal 1D, ce qui génère une séquence de coefficients d’échelle et d’ondelettes (figure 2.12-b). La segmentation 1D suivie d’une phase d’analyse est répétée pour chacune des directions, à la fois sur les

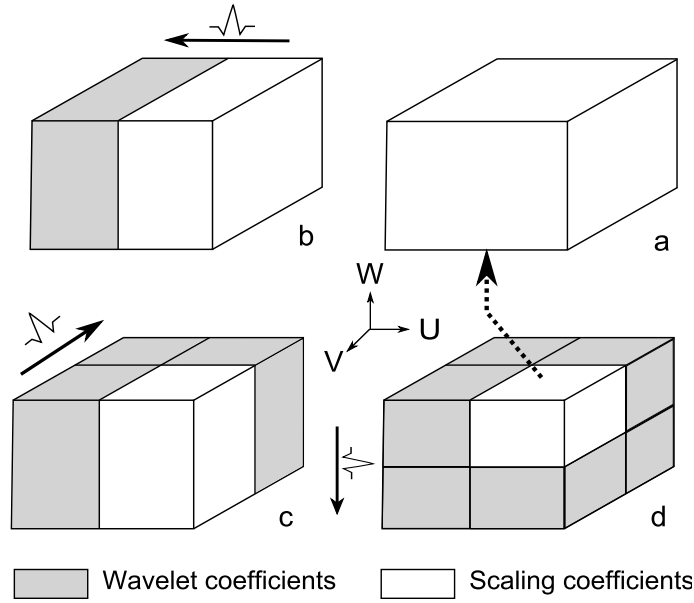


FIGURE 2.12 – Transformée 3D en ondelettes. (a) Coefficients d'échelle à la résolution r . (b) Séquence de coefficients après analyse des signaux 1D alignés sur la direction U . (c) Après analyse des signaux 1D alignés sur la direction V . (d) Après analyse des signaux 1D alignés sur la direction W .

coefficients d'échelle et d'ondelettes (figure 2.12-c-d). L'ensemble du processus (figure 2.12 a-d) est réitéré sur le bloc de coefficients d'échelle restant, jusqu'à ce que l'échelle la plus grossière soit atteinte.

Différentes ondelettes peuvent être utilisées pour les différentes directions en fonction de la variation spatiale des propriétés. Dans les modèles de réservoir 3D, les transitions latérales des propriétés au sein de chaque formation peuvent être assez lisses, alors qu'elles peuvent être abruptes entre deux couches différentes. Dans ce cas, des ondelettes lisses (d'ordre élevé) peuvent être utilisées latéralement, tandis qu'une ondelette discontinue (p. ex. l'ondelette de Haar) peut être utilisée verticalement (voir figure 2.13). Une anisotropie latérale peut également être prise en compte en utilisant des ondelettes différentes en fonction de la corrélation associée à chaque direction. En général, plus la distance de corrélation est grande, plus les ondelettes d'ordre élevé seront efficaces. Il est important de préciser que cette méthode ne permet que de gérer des anisotropies alignées sur les directions principales de la grille.

2.4 Pyramides gaussiennes

En traitement d'images, les pyramides correspondent à des types de représentation multi-échelles, où l'image est décomposée par l'application successive de filtres passe-bas et par sous-échantillonnage. Elles sont principalement utilisées dans la synthèse de textures [78] et le mélange d'images [138].

Ces décompositions ont beaucoup de point communs avec les transformées en ondelettes : les pyramides utilisent des filtres locaux et sont considérées comme les précurseurs de l'analyse multi-résolution. Néanmoins, elles ne sont

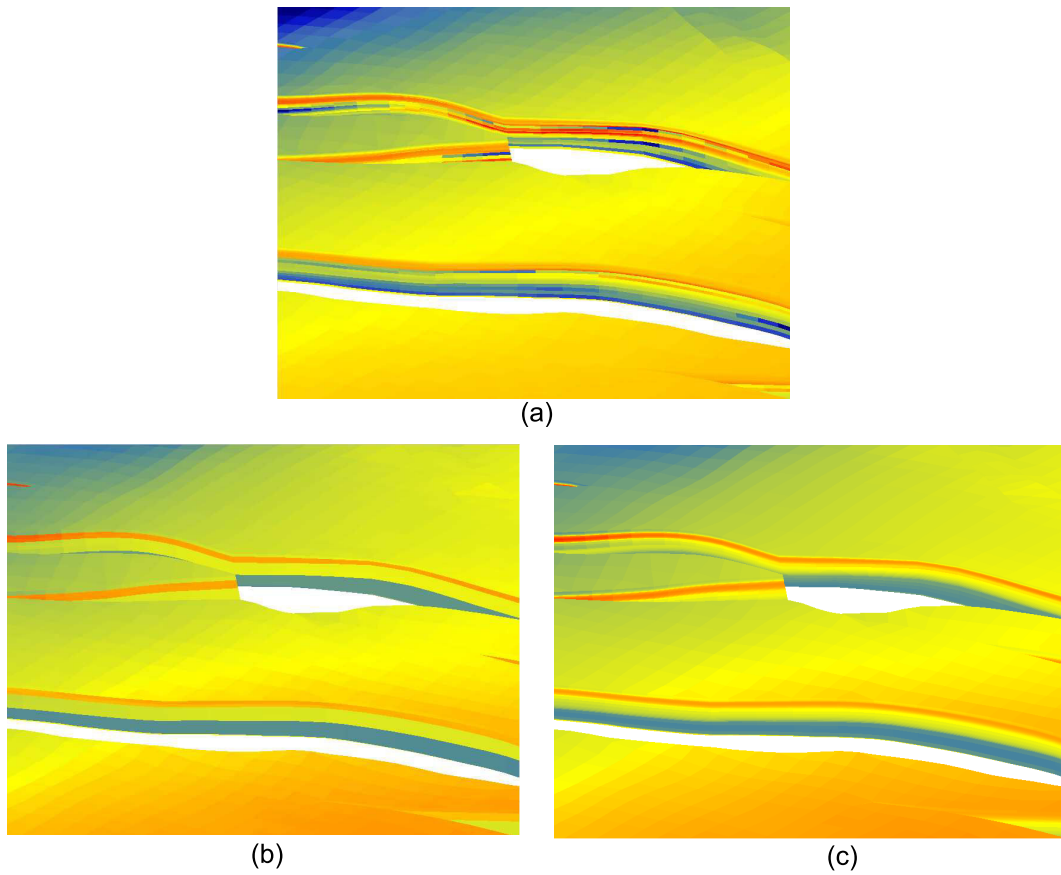


FIGURE 2.13 – Transformation anisotrope. (a) Propriété initiale. Reconstructions à partir de 0.005% des coefficients pour une transformation : (b) anisotrope (latéralement ondelettes quadratiques lisses, verticalement ondelette de Haar) (c) isotrope (ondelettes quadratiques). On observe que les transitions verticales sont discontinues dans le cas (b) alors qu'elles sont continues dans le cas (c).

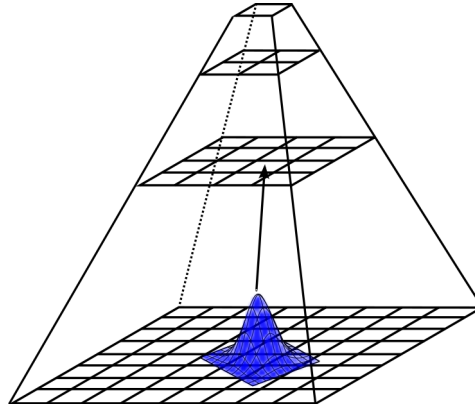


FIGURE 2.14 – Pyramide gaussienne. Chaque niveau est créé par l'application d'un filtre gaussien d'un niveau plus fin

pas utilisées en compression d'images car elles augmentent le nombre de coefficients.

Les pyramides gaussiennes utilisent des filtres gaussiens afin de créer une succession d'approximations d'un signal donné :

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}, \quad (2.42)$$

où $g(x)$ est un filtre gaussien 1D défini par l'écart-type σ , lequel contrôle le support de la fonction. On peut remarquer que ces fonctions servent également de base pour la représentation de Gabor. Ces filtres sont similaires aux filtres inter-échelles \tilde{h} utilisés dans les transformés en ondelettes rapides. De la même manière que pour la transformée en ondelettes 2D (ou 3D), ces filtres peuvent être appliqués successivement sur des segments 1D alignés sur les directions principales de la grille (figure 2.12). Après application du filtre, le signal est sous-échantillonné (figure 2.14). Dans le cas des pyramides de Laplace, tout comme pour les ondelettes, l'approximation grossière est retranchée au signal afin que les coefficients résultant ne soient représentatifs que d'une bande limitée de fréquences (filtre passe-bande, figure 2.3-b), alors que le contenu fréquentiel des coefficients de la pyramide de Gauss inclut toujours les basses fréquences (filtre passe-bas, figure 2.15). Nous verrons que cette caractéristique des pyramides gaussiennes rend plus flexible le conditionnement d'un modèle à des données de différentes résolutions lors des simulations géostatistiques multi-échelles (Chapitre 3).

2.5 Conclusions partielles et perspectives

Dans ce chapitre nous avons présenté des re-paramétrisations en ondelettes qui forment des bases éparées dont les fonctions sont localisées à la fois dans l'espace et en fréquence. Ces ondelettes permettent de paramétrer efficacement des signaux hétérogènes en s'adaptant localement à leur contenu fréquentiel. Elles peuvent également simplifier les problèmes posés en travaillant à partir de modèles approximatifs de dimension réduite. De plus, dans ce chapitre nous

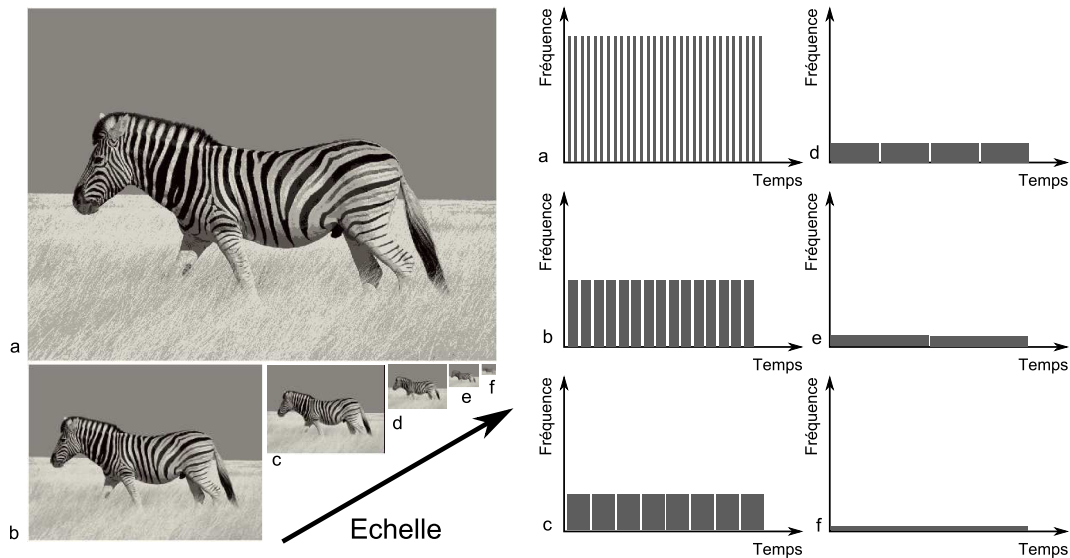


FIGURE 2.15 – Pyramide gaussienne et représentation temps-fréquence. Décomposition de l'image 2.2. Gauche : approximations successives. Droite : couvertures temps-fréquence associées

proposons une décomposition 3D adaptée aux grilles complexes de réservoir qui permet de prendre en compte efficacement les volumes des cellules. Ces transformations en ondelettes, ainsi qu'une décomposition 2D en pyramides gaussiennes ont été intégrées à un *plugin* Gocad. Le code créé est générique et optimisé afin de pouvoir être utilisé efficacement dans différentes applications. Dans les chapitres suivants, ces transformations sont utilisées pour une intégration optimale des données de réservoir.

D'autres perspectives intéressantes pourraient exploiter les propriétés de ces ondelettes. En géosciences, et notamment en surveillance sismique des réservoirs, le nombre de données spatiales peut être extrêmement important et des problèmes de stockage peuvent survenir. Les propriétés compressives des ondelettes pourraient permettre de réduire efficacement la taille des données tout en conservant les composantes importantes du signal.

De même, l'affichage de données en modélisation 3D peut poser problèmes lorsque la dimension augmente. Il pourrait être intéressant d'utiliser la représentation en ondelettes pour optimiser la visualisation des données en fonction de la distance aux objets et/ou de la partie affichée. Des approches similaires pourraient également être utiles en visualisation des incertitudes. En effet, le contenu fréquentiel local pourrait dépendre des incertitudes associées aux données. Les endroits incertains auraient un graphisme grossier tandis que les parties certaines comporteraient plus de détails.

Enfin, dans ce chapitre il a été montré qu'il était possible de construire des ondelettes adaptées à l'échantillonnage du signal. En allant plus loin, il est possible d'imaginer d'utiliser des ondelettes adaptées au signal lui-même [36]. Il serait alors possible d'adapter ou de modifier directement les fonctions de base directement en plus de leurs coefficients. De plus, ces transformations ont un potentiel énorme en termes de taux de compressions des données. Ce-

pendant, ces applications étant non linéaires (dépendent du signal), elles sont plus difficilement utilisables en algèbre linéaire (p. ex. pour une optimisation de Gauss-Newton).

Chapitre 3

Intégration des données statiques par méthodes géostatistiques multipoints multi-échelles

« Our most precise description of nature must be in terms of probabilities. »

Richard Phillips Feynman, 1918-1988

3.1 Résumé et contributions

3.1.1 Résumé

Ce chapitre traite de la simulation géoStatistique MultiPoints (MPS) utilisée pour créer des champs de propriétés continues ou discrètes ayant des structures géologiques complexes, à partir des données d'observations directes de puits, de données secondaires interprétées (p. ex. carte de porosité sismique) et de modèles conceptuels appelés images d'entraînement (*Training Images* ou TI). Ces modèles peuvent être utilisés comme modèle *a priori* dans les processus d'inversion et ont pour but d'intégrer les connaissances géologiques du champ étudié. Nous nous intéressons plus particulièrement à la simulation de systèmes chenalisés 2D, dont la difficulté principale est de générer des objets continus plus ou moins connectés entre eux. Ce chapitre introduit dans un premier temps les différentes méthodes MPS existantes, puis propose une présentation plus détaillée de l'algorithme d'échantillonnage direct (*Direct Sampling* ou DS) utilisé dans ce travail de thèse. Cependant, il est possible d'étendre l'approche multi-échelles à d'autres algorithmes multipoints.

Dans un deuxième temps, nous introduisons une nouvelle approche multi-échelles basée sur une paramétrisation en ondelettes (Section 2.3). Cette méthode permet de générer efficacement des réalisations à la fois conditionnées par des données de puits et des données secondaires (par exemple de la sismique interprétée). Cependant, la décomposition en gammes de fréquences (échelles)

indépendantes rend difficile l'intégration des données ponctuelles de puits. De plus, la simulation par échelle indépendante peut produire des artefacts.

Afin d'exploiter les avantages de l'approche multi-échelles tout en évitant les inconvénients de la décomposition en échelles indépendantes, nous proposons de re-paramétriser le problème à l'aide de pyramides gaussiennes (Section 2.4). Cette approche permet un meilleur conditionnement des réalisations par les données dures, une utilisation de TI de grandes tailles et l'application de traitements complémentaires aux cours de la simulation des échelles intermédiaires, tout en conservant des temps de calculs raisonnables.

Les approches multi-échelles (MS-DS) sont comparées à l'algorithme DS standard en termes de qualité des réalisations produites et de temps de calcul. Nous montrons que pour des temps de calculs semblables, l'approche multi-échelles basée sur les pyramides gaussiennes génère des réalisations plus réalistes et moins bruitées que l'approche standard.

3.1.2 Contributions

Ce chapitre contient les contributions suivantes :

- une adaptation multi-échelles de l'algorithme MPS d'échantillonnage direct (DS), basée sur une paramétrisation en ondelettes (Section 3.3.1)
- une adaptation multi-échelles de l'algorithme DS, basée sur une paramétrisation en pyramides gaussiennes (Section 3.3.2)
- des schémas multi-échelles de conditionnement des réalisations par des données dures ponctuelles et des données secondaires disponibles pour l'ensemble du domaine étudié (Sections 3.3.1 et 3.3.2)
- une méthode de recherche préférentielle dans l'image d'entraînement permettant l'utilisation de TI de grandes tailles (Section 3.3.2)
- une méthode de traitement basée sur la connectivité au cours de la simulation des échelles intermédiaires adaptée à la simulation 2D de chenaux (Section 3.3.2)
- une analyse comparative entre la méthode MS-DS proposée et la méthode DS standard (Section 3.3.2)

3.2 Méthodes géostatistiques

La simulation de variables spatialement corrélées est fondamentale dans le domaine des géosciences et notamment dans la modélisation des comportements physiques des réservoirs (voir modélisation directe dans la Section 1.2). Les méthodes géostatistiques permettent de générer des champs de propriétés conditionnés par des données observées (données de puits) et/ou interprétées (carte de porosité) grâce à des modèles de corrélations spatiales. Contrairement aux méthodes inverses, les géostatistiques n'utilisent pas de modèles directs. Cependant, de manière similaire à la résolution des problèmes inverses, les méthodes géostatistiques peuvent se diviser en deux grands groupes : les méthodes déterministes et les méthodes stochastiques.

Les méthodes déterministes cherchent à générer, sous certaines hypothèses (modèle gaussien, hypothèse de stationnarité, . . .), le champ de propriété le plus probable en fonction des données observées et du modèle de corrélation utilisé, ce qui correspond à la recherche du maximum *a posteriori* lors l'optimisation d'un problème inverse (Section 1.3.1).

Cependant, le problème étant généralement sous-conditionné et les données étant bruitées, il existe un grand nombre de modèles qui peuvent satisfaire les données. Les méthodes stochastiques cherchent alors à estimer les incertitudes liées à la nature du problème. De façon semblable à la résolution d'un problème inverse, les algorithmes stochastiques cherchent à caractériser la distribution des modèles possibles en échantillonnant la PDF conditionnée par les données (voir Section 1.3.3).

Dans les deux cas, la forme de la PDF (et donc le MAP) dépend du modèle de corrélation décrivant les relations entre les variables spatiales modélisées. Lorsque le modèle est linéaire, il est possible de calculer analytiquement la PDF et le modèle le plus probable. C'est généralement l'hypothèse faite par les géostatistiques deux-points, pour lesquelles les relations spatiales sont modélisées par une matrice de covariance.

Cependant, ces modèles linéaires sont incapables de reproduire des configurations géologiques complexes [86] comme des systèmes fluviaux ou des réseaux de fractures. Dans ce cas, il est possible d'utiliser des méthodes basées sur des concepts géologiques. Les méthodes génétiques génèrent des modèles très réalistes en utilisant des règles de dépôts liées à des processus sédimentaires [94]. Les méthodes basées-objets [42] permettent de simuler certains types d'objets géologiques ou de formes connues. Néanmoins, ces méthodes ne sont pas génériques et il peut s'avérer difficile de contraindre les modèles avec les données observées. Pour cette raison, des algorithmes MPS sont parfois utilisés en complément de ces méthodes.

Les méthodes MPS utilisent des images d'entraînement pour modéliser les relations spatiales entre les variables. Ces méthodes, utilisées dans ce travail de thèse, sont généralisables à un grand nombre de concepts géologiques différents et permettent de générer des modèles respectant les données observées.

Dans les deux exemples mentionnés ci-dessus, les modèles sont non-linéaires et il n'est généralement pas possible de calculer analytiquement le modèle le plus probable. Des processus de simulation sont alors utilisés pour générer des réalisations qui échantillonnent l'espace des solutions.

3.2.1 Géostatistiques deux-points

Un développement complet sur les géostatistiques multivariée peut être trouvé dans [136].

Fonctions de corrélations et matrice de covariance

Dans le cas des géostatistiques à deux-points, la matrice de covariance \mathbf{C}_M est entièrement définie par des fonctions de covariances $C(h)$ qui dépendent uniquement de la distance de séparation des variables $Z_i = Z(x_i)$ de positions

x_i :

$$C_M(x_i, x_j) = C(h = |x_i - x_j|) \quad (3.1)$$

Dans ce cas, l'hypothèse de stationnarité doit être valable sur l'ensemble du domaine. Généralement, les fonctions de covariance sont définies à partir de modèles de variogrammes (p. ex. cubiques, exponentiels, gaussiens), $\gamma(h)$, qui peuvent provenir de l'analyse des données disponibles :

$$C(h) = C(0) - \gamma(h) = \gamma(\infty) - \gamma(h), \quad (3.2)$$

où $\gamma(\infty)$ représente le plateau du variogramme¹.

Estimation : exemple du krigeage simple

La forme de base du krigeage [101] est donnée par :

$$Z^*(x) - m(x) = \sum_{d=1}^{n^d} \lambda_d [Z(x_d) - m(x_d)] \quad (3.3)$$

où Z^* représente la valeur estimée à la position x , $m(x)$ la moyenne à la position x , n^d le nombre de données de conditionnement et $Z(x_d)$ les valeurs des données aux positions x_d . Le but du krigeage est de déterminer les poids λ_d qui minimisent la variance d'estimation,

$$\sigma_E^2(x) = \text{Var}[Z^*(x) - Z(x)] \quad (3.4)$$

sous la contrainte de non-biais, $E[Z^*(x) - Z(x)] = 0$.

Dans le cas du krigeage simple, $Z(x)$ est supposée stationnaire à moyenne connue, $m(x) = m$, ce qui respecte cette contrainte. Il est alors possible d'approximer la variance d'estimation en fonction des points et de la covariance (voir [136] pour le développement complet) :

$$\sigma_E^2(x) = \sum_{d_1=1}^{n^d} \sum_{d_2=1}^{n^d} \lambda_{d_1} \lambda_{d_2} C(x_{d_1} - x_{d_2}) + C(0) - 2 \sum_{d_1=1}^{n^d} \lambda_{d_1} C(x_{d_1} - x) \quad (3.5)$$

En dérivant l'équation 3.5 en fonction des λ_i et annulant les dérivées (recherche du minimum), on obtient le système linéaire suivant à résoudre :

$$\sum_{d_2=1}^{n^d} \lambda_{d_2} C(x_{d_1} - x_{d_2}) = C(x_{d_1} - x) \quad (3.6)$$

ou sous sa forme matricielle :

$$\mathbf{C}_M \boldsymbol{\lambda}_{d_2} = \mathbf{c} \quad (3.7)$$

où \mathbf{c} représente le vecteur de covariances entre les points de données et le point

¹Sous l'hypothèse d'une variable régionalisée stationnaire d'ordre 2

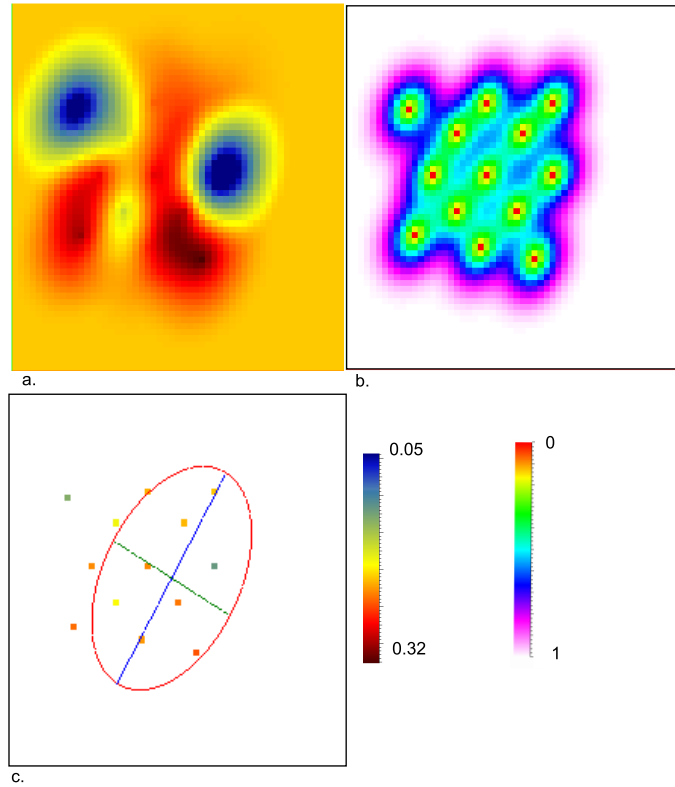


FIGURE 3.1 – Résultat d'un krigeage. (a) Propriété Krigée (porosité). (b) Variance de krigeage. (c) Données et extensions du variogramme utilisé.

d'estimation.

À partir des poids λ , il est possible de calculer la variance d'estimation pour une cellule i :

$$\sigma_{Kg}^2(i) = C(0) - \lambda^T \mathbf{c} \quad (3.8)$$

La figure 3.1 montre le résultat d'un krigeage conditionné par des données de puits. Les propriétés générées par krigeage sont toujours lisses et continues. La variance d'estimation augmente à mesure que l'on s'éloigne des données et devient très important pour des distances aux points d'échantillonnage plus grandes que la portée de la covariance.

Simulation stochastique : exemple de la Simulation Séquentielle Gaussienne (SGS)

Lorsque l'on souhaite effectuer une analyse des incertitudes, il est possible d'utiliser des schémas stochastiques afin de générer des réalisations de la PDF. Lorsque la distribution des propriétés est supposée gaussienne², il est possible d'utiliser l'algorithme de simulation séquentielle gaussienne défini par les étapes suivantes :

²lorsqu'elles ne le sont pas, il est possible de transformer les variables par anamorphose gaussienne et appliquer la transformée inverse après simulation

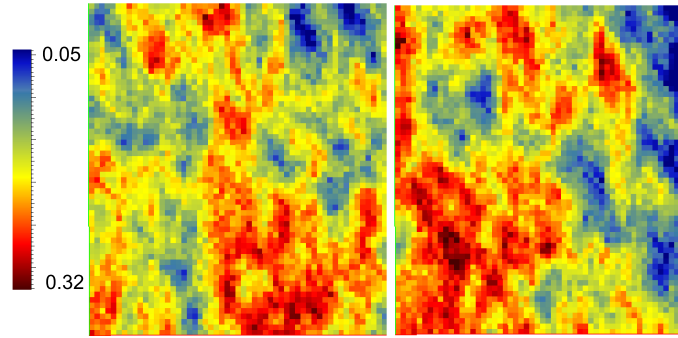


FIGURE 3.2 – Réalisations générées par une SGS. Une distribution gaussienne, ainsi que les données et le variogramme de la figure 3.1 ont été utilisés par le processus de simulation.

1. générer un chemin aléatoire pour visiter toutes les cellules de la grille de simulation
2. pour chaque cellule du chemin aléatoire :
 - (a) utiliser le krigeage pour estimer une distribution conditionnelle de probabilité locale³ à partir des données de conditionnement
 - (b) sélectionner aléatoirement une valeur de propriété dans cette distribution par tirage de type Monte-Carlo et l'assigner à la cellule
 - (c) inclure la valeur simulée dans les données de conditionnement

La figure 3.2 présente des exemples de réalisations générées à partir d'une SGS.

Comme nous le verrons dans la Section 3.2.2, ce schéma de simulation est également extensible au cas non gaussien.

3.2.2 Géostatistiques multipoints

Les méthodes géostatistiques basées sur un modèle linéaire de covariance, telles que les méthodes deux-points présentées dans la section précédente, ne permettent pas de reproduire des variabilités spatiales complexes. Afin de remédier à cette limite, des méthodes non linéaires multipoints ont été proposées (une revue des différentes méthodes est disponible dans [83, 99]). Les statistiques multipoints et le concept d'image d'entraînement (TI) ont été introduits par Guardiano et Srivastava [75]. Une TI est un modèle de référence (2D ou 3D) qui décrit les relations spatiales et la distribution des variables simulées. Les différents motifs contenus dans la TI sont supposés être semblables à ceux présents dans le champ étudié [75].

La première implémentation suffisamment rapide pour être utilisée en 3D de simulation MPS fut proposée par Strebel [126]. Dans cette méthode, appelée *snesim* (*single normal equation simulation*), les configurations de la TI sont probabilisées grâce à un arbre de recherche qui enregistre leurs fréquences d'apparitions dans la TI, ce qui permet d'estimer efficacement la probabilité d'un

³moyenne et covariance

faciès pour un voisinage donné. Le déroulement de l'algorithme s'appuie sur les processus séquentiels de simulation stochastique (voir Section 3.2.1) : toutes les cellules de la grille de simulation sont parcourues dans un ordre aléatoire et la valeur de chacune des cellules est tirée dans une densité de probabilité conditionnelle générée à partir du voisinage de la cellule et de l'arbre de recherche. Le voisinage est composé de données et de valeurs précédemment simulées contenues à l'intérieur d'une forme géométrique appelée *template*. L'ensemble des positions et des valeurs du voisinage est appelé événement de données (*data event* ou DE). Généralement, la qualité des simulations va dépendre de la taille du DE, du *template* et de la TI : plus le conditionnement est important et plus la TI sera grande, meilleure sera la simulation. Cependant, la taille de l'arbre de recherche est directement lié à la taille du *template* et l'utilisation d'un voisinage important peut conduire à des problèmes de mémoire et d'efficacité de calcul des PDF conditionnelles.

Arpat et Caers [6] proposent un algorithme (*simpat*) qui utilise directement les motifs (*pattern*) de l'image d'entraînement. Un motif est une configuration particulière de pixels (valeurs de cellules) définie dans une image d'entraînement pour un *template* donné. Contrairement aux méthodes qui simulent les valeurs cellule-par-cellule, ces algorithmes extraient un motif entier de l'image d'entraînement et l'insèrent directement dans la grille de simulation. Différentes évaluations de distance pour une configuration donnée peuvent être utilisées pour choisir le motif le plus semblable dans l'image d'entraînement. Zhang et al. [141] proposent l'algorithme *filtersim* afin de réduire le temps de recherche grâce à une classification des différents motifs présents dans l'image d'entraînement. Le calcul de la distance ne se fait plus directement à partir de l'image d'entraînement mais est défini pour chacun des groupes de motifs. Un motif est ensuite choisi dans le groupe le plus semblable. Néanmoins, les méthodes basées sur l'utilisation de motifs produisent généralement des images peu réalistes comportant des artéfacts aux frontières des motifs insérés, même si un certain nombre de méthodes sont proposées pour améliorer ce problème [70, 28, 81, 130].

La plupart de méthodes décrites ci-dessus sont uniquement applicables à la simulation de propriétés discrètes et nécessitent l'utilisation d'une base de données. Mariethoz et al. [100] proposent un algorithme d'échantillonnage direct (*Direct Sampling* ou DS) qui peut être utilisé à la fois pour simuler des propriétés discrètes et continues, ce qui permet d'intégrer facilement des données secondaires. Au lieu d'utiliser une base de données intermédiaire, l'algorithme DS va échantillonner directement l'image d'entraînement pour un DE donné. L'algorithme DS permet une bonne reproduction des structures de l'image d'entraînement, mais la recherche directe peut s'avérer très coûteuse.

L'algorithme DS a été retenu dans ce travail de part sa facilité d'implémentation et sa bonne reproduction des structures discrètes et continues.

3.2.3 *Direct Sampling*

La figure 3.3 illustre les différentes étapes de l'algorithme. Pour chaque cellule non simulée, un DE est choisi à partir du voisinage de la cellule. Cet événement contient un nombre maximum d'éléments qui doivent être le plus proche possible de la cellule centrale. L'image d'entraînement est ensuite scan-

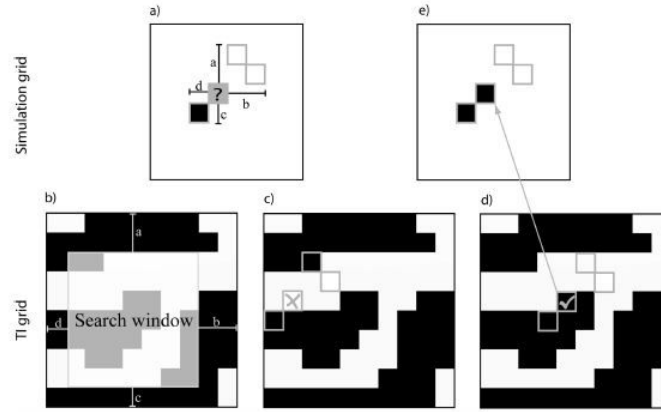


FIGURE 3.3 – Illustration de l’algorithme d’échantillonnage direct. Tirée de [100]. (a) Définition de l’événement de données (voisinage) dans la grille de simulation. La cellule marquée par un point d’interrogation représente l’élément simulé. Les cellules noires et blanches représentent des données de conditionnement (points de données ou éléments déjà simulés). (b) Définition d’une fenêtre de recherche dans l’image d’entraînement en fonction des dimensions a, b, c, d, du DE. (c) Recherche aléatoire dans l’image d’entraînement jusqu’à ce qu’on obtienne (d) une configuration semblable au DE. (e) La valeur centrale de l’image d’entraînement est assignée à la cellule simulée.

née afin de trouver une configuration compatible avec cet événement. Différentes mesures de distances peuvent être utilisées, mais sont pour la plupart basées sur les différences entre les valeurs des éléments de l’événement et les valeurs correspondantes de l’image d’entraînement. Ainsi, il est possible de changer la définition de la distance en fonction de la nature des propriétés simulées (discrètes ou continues) et le type des données de conditionnement (données dures ou secondaires). La recherche s’arrête dès que l’on trouve une configuration ayant une distance inférieure à un seuil donné par l’utilisateur ou que l’on a parcouru toute la grille. La valeur de la cellule centrale est alors attachée à la grille de simulation.

Un aspect important de l’algorithme DS est qu’il n’utilise pas de *template* de simulation. L’évènement de données peut alors inclure des éléments très éloignés de la cellule centrale, ce qui évite l’assignation de valeurs aléatoires lors de la simulation des premières cellules⁴. De plus, l’extension spatiale du DE varie naturellement au cours du processus. En effet, il est plus facile de trouver des voisins proches lorsque la densité de points simulés augmente. Ainsi, il n’est pas nécessaire d’utiliser des méthodes multi-grilles [134].

Afin d’améliorer la qualité de conditionnement des réalisations par les données ponctuelles, il est possible d’utiliser des chemins (semi-)structurés de visites des cellules de la grille. Liu et Journel [93] proposent d’adapter le chemin en fonction du contenu informatif des données. L’idée est de simuler en premier les éléments bien conditionnés par les données. Dans ce travail, nous proposons de générer un chemin semi-aléatoire qui utilise une carte de distances aux données : les cellules les plus proches des observations ont une probabilité plus

⁴cte aspect dépend également de la taille et la richesse de la TI

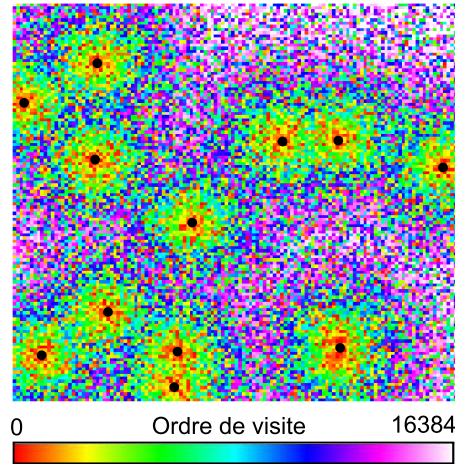


FIGURE 3.4 – Chemin semi-aléatoire. Les ronds noirs correspondent aux données. Les variables proches des données sont simulées en priorité.

forte d’être simulées au début du processus (figure 3.4). Les données ont ainsi plus de poids sur le développement des structures simulées et sont donc mieux respectées comme le montre la figure 3.5-d.

Des traitements peuvent également être appliqués afin d’améliorer la qualité des réalisations [83]. Un traitement en temps réel est proposé par [100] ; il a pour objectif de re-simuler les valeurs des cellules dès qu’une configuration non conforme à la TI est détectée. Plus simplement, il est possible d’effectuer un second passage sur les valeurs simulées afin de remplacer les configurations les plus mauvaises (figure 3.5). Néanmoins, cette technique a un impact limité sur la structure générale du champ de propriété et peut s’accompagner de phénomènes indésirables, tels que l’érosion ou la déconnexion d’objets. Ces méthodes étant généralement très coûteuses, il est souvent nécessaire de limiter les traitements à un nombre réduit d’éléments [127] (voir Section 3.3.2).

3.3 Direct Sampling Multi-échelles (MS-DS)

La qualité d’échantillonnage de la densité de probabilité associée à un modèle multipoints dépend de la qualité des réalisations générées. L’un des objectifs de cette section est d’utiliser des approches multi-échelles afin d’améliorer l’échantillonnage de la PDF en générant des modèles ayant une forte probabilité d’apparition, c.à.d, respectant mieux l’image d’entraînement. L’idée est de décomposer la simulation MPS sur plusieurs échelles. Dans un premier temps, les échelles grossières sont simulées, puis servent à conditionner les échelles plus fines. Un effet de régularisation est ainsi obtenu et permet l’utilisation d’un DE de plus petite taille. De plus, très peu de paramètres sont simulés aux échelles grossières et intermédiaires. Il est alors possible de procéder à des traitements supplémentaires efficaces.

Un avantage inhérent aux approches multi-échelles vient du fait qu’il est possible d’intégrer différentes sources d’informations à des échelles différentes. Chatterjee et al. [27] re-paramétrisent le problème dans l’espace des ondelettes (Section 2.3) et proposent de cosimuler les coefficients d’échelles à partir de

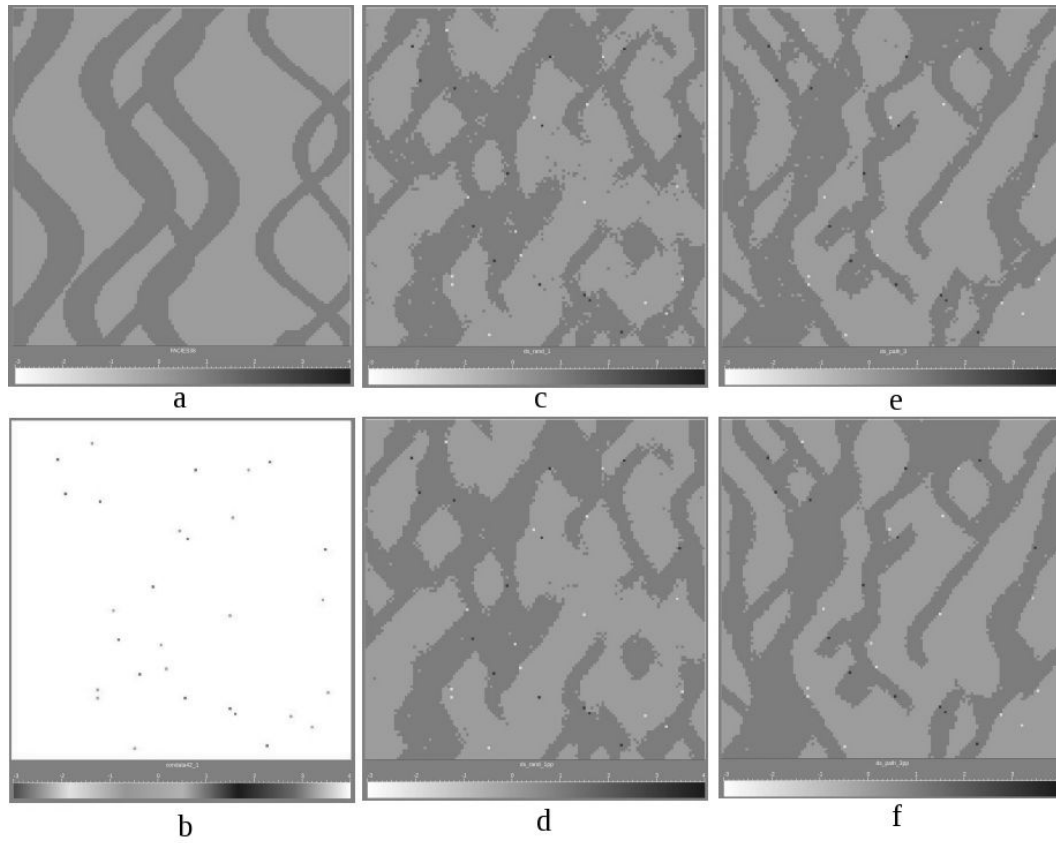


FIGURE 3.5 – Résultats de l'algorithme d'échantillonnage direct conditionnés par des données de puits. (a) Image d'entraînement ; (b) Données de conditionnement ; (c) Sans traitement + Chemin aléatoire (d) Sans traitement + Chemin organisé ; (e) Traitement + Chemin aléatoire ; (f) Traitement + Chemin organisé

données secondaires et des données primaires (de puits) définies à la même résolution. Ils utilisent ensuite un algorithme MPS (basé-motifs) pour simuler les coefficients d'ondelettes restants (échelles fines). Finalement, ils appliquent la transformée d'ondelettes inverse pour obtenir les champs de propriétés. De façon similaire, Gloaguen et al. [70] effectuent une cosimulation des coefficients d'ondelettes à partir des données, en utilisant une image d'entraînement décomposée en coefficients d'ondelettes pour modéliser les dépendances spatiales. La limitation principale de ces méthodes réside dans l'utilisation des géostatistiques deux-points lors de l'intégration des données. Lorsque les données secondaires sont de mauvaise qualité et/ou que peu de données primaires sont disponibles, les géostatistiques deux-points ne permettent pas de reproduire des structures complexes. Or, les coefficients (d'ondelettes ou d'échelles) co-simulés servant de base à la simulation, la qualité des réalisations reste fortement liée à la qualité des données, cela malgré l'utilisation d'une image d'entraînement.

3.3.1 MS-DS basé sur les ondelettes (MS-DSw) et intégration de données secondaires

Cette section s'appuie sur l'article de Romanenko et al. [115] issu d'un travail de master supervisé durant cette thèse. Elle propose d'intégrer des données primaires et secondaires à l'aide d'un schéma MPS multi-échelles basé sur l'utilisation des ondelettes de seconde génération. L'approche est semblable à celle proposée par Chatterjee et al. [27], mais utilise l'algorithme d'échantillonnage direct conditionné [100] pour simuler les coefficients d'échelles et sa forme non conditionnée pour générer les coefficients d'ondelettes. Les données secondaires pouvant être définies à une échelle bien plus grossière que les données primaires, une approche de simulation échelle-par-échelle, ainsi qu'un conditionnement spécifique des données primaires est proposé. L'algorithme se décompose alors en trois grandes étapes :

1. décomposition de l'image d'entraînement et conditionnement aux données primaires à l'échelle r des données secondaires
2. simulation des coefficients d'échelle r par l'algorithme d'échantillonnage direct (DS)
3. simulation DS récursive des coefficients d'ondelettes pour les échelles r à r_{max} (échelle la plus fine), conditionnée par les coefficients d'échelle grossière

Décomposition de l'image d'entraînement et adaptation des données primaires à l'échelle grossière

Le *lifting* en ondelettes décrit dans la Section 2.3.2 est appliqué à l'image d'entraînement afin de la décomposer partiellement en coefficients d'échelle correspondant aux données secondaires et en coefficients d'ondelettes (voir figure 3.7 (b) et (d)). Des ondelettes basées sur une interpolation quadratique sont employées afin d'éviter l'apparition d'artéfacts discontinus lors de la reconstruction des propriétés. De plus, ces ondelettes ayant un support plus important que les ondelettes de Haar, elles permettent de mieux identifier la présence de bords,

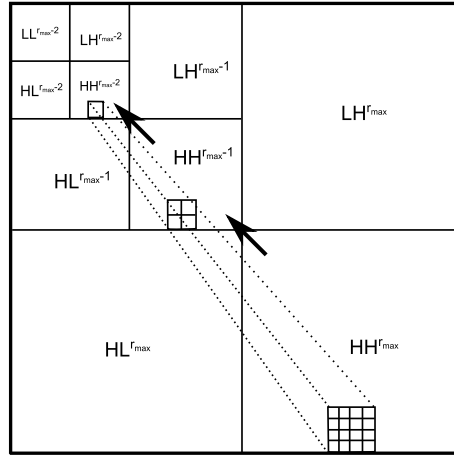


FIGURE 3.6 – Relations spatiales et d'échelles des coefficients d'ondelettes. Les motifs représentent la décomposition fréquentielle d'un même lieu. Les flèches indiquent la direction des échelles fines vers les échelles grossières. Les codes LL, LH, HL et HH représentent respectivement, les coefficients d'échelles, les composants horizontaux, les composants verticaux et les composants transversaux. Les exposants associés indiquent la résolution des coefficients. Modifiée de [27].

ce qui est un avantage lors de la simulation d'objets discontinus. Il est important de préciser que la transformation n'est pas normalisée ; de ce fait, les valeurs des coefficients d'échelle correspondent aux moyennes locales des propriétés. Les relations spatiales et d'échelles des coefficients d'ondelettes sont rappelées dans la figure 3.6. Dans les approches multi-grilles, les données de conditionnement sont assignées aux nœuds les plus proches de la grille sous-échantillonnée [126]. Cette approche est ici impossible car les transformées directes et inverses sont utilisées pour passer d'une échelle à une autre.

Les coefficients d'échelle r représentant des moyennes locales sur une surface⁵ de 2^r par 2^r cellules (voir figure 3.6), les données de conditionnement utilisées à l'échelle r sont estimées à partir des moyennes locales, $av_{hd}(x)$, des points contenus dans cette surface. Cependant, lorsque peu de données primaires sont incluses, les estimations sont de mauvaise qualité et ne doivent pas être utilisées comme des données dures.

Pour évaluer le poids des données à l'échelle r , il est nécessaire de prendre en compte la corrélation spatiale des propriétés. Un moyen d'évaluer le contenu en information d'une donnée primaire sur ces voisins (à l'échelle fine) est d'utiliser la variance normalisée du krigeage $\sigma_{Kg}^2(i)$ (équation 3.8) issue d'un modèle de covariance. Dans ce travail, cette variance est approximée par un modèle de distance isotrope. On introduit deux paramètres de distance : la distance de corrélation maximale max_{cl} qui représente la distance de séparation à partir de laquelle les corrélations entre les variables du modèle sont nulles (portée du variogramme) et la distance de corrélation optimale min_{cl} qui représente la distance de séparation en dessous de laquelle la corrélation entre les variables est proche de 1. Une carte d'influence, $infl_{hd}$, est alors calculée à l'échelle fine

⁵on ne considère ici que le cas 2D

à partir des données primaires de la façon suivante :

$$infl_{hd} = \begin{cases} 1 & \text{if } dist \leq min_{cl} \\ 0 & \text{if } dist > max_{cl} \\ \frac{max_{cl}-dist}{max_{cl}-min_{cl}} & \text{else} \end{cases}$$

Les poids, $weight_{hd} \in [0, 1]$, des données de conditionnement sont des moyennes locales (sur la surface associée aux coefficients d'échelle, figure 3.6) de la carte de d'influence.

Lorsque le poids $weight_{hd}$ est supérieur à un certain seuil critique $weight_{crit}$, la valeur $av_{hd}(x)$ est considérée comme une donnée dure. En dessous de ce seuil, elle est considérée comme une donnée semi-dure et est traitée comme une donnée secondaire (voir équation 3.9 ci-dessous). Bien évidemment, lorsque le coefficient d'échelle n'est associé à aucune donnée, aucune valeur initiale n'est attribuée. Cette approche de poids évite de fixer une valeur estimée à partir de statistiques linéaires qui peut ne pas être compatible avec le modèle multipoints. Cependant, elle oblige les réalisations à respecter les données certaines.

Simulation des coefficients d'échelle

Une fois l'image d'entraînement décomposée et les données de conditionnement calculées (et pondérées) pour l'échelle grossière, les coefficients d'échelle peuvent être simulés en utilisant l'algorithme DS.

Dans un premier temps, un chemin semi-aléatoire de visite des cellules est calculé en fonction des poids $weight_{hd}$ des données⁶. Ensuite, pour chaque cellule libre, un événement de donné est choisi selon deux critères : une taille maximale et une distance de séparation maximale, s_{max}^r , entre la cellule simulée et les données de conditionnement. Cette distance de séparation est exprimée en nombre de cellules en fonction de l'échelle r et de la distance maximale de max_{cl} , telle que, $s_{max}^r = max_{cl}/2^r$. Cette restriction additionnelle a le même rôle qu'un *template* de recherche et permet de réduire les temps calculs par rapport à l'approche DS standard.

L'image d'entraînement est alors scannée en utilisant une fonction de distance qui évalue la différence entre l'événement de données, $DE(x)$, associé à la cellule x simulée et les différentes configurations $DE(y)$, centrées en y , disponibles dans l'image d'entraînement. Afin de pouvoir simuler sans distinction des variables discrètes et continues, une norme l_2 est utilisée dans le calcul de

⁶si le poids des données secondaires varie en fonction de la localisation, ils peuvent également être utilisés dans le calcul du chemin

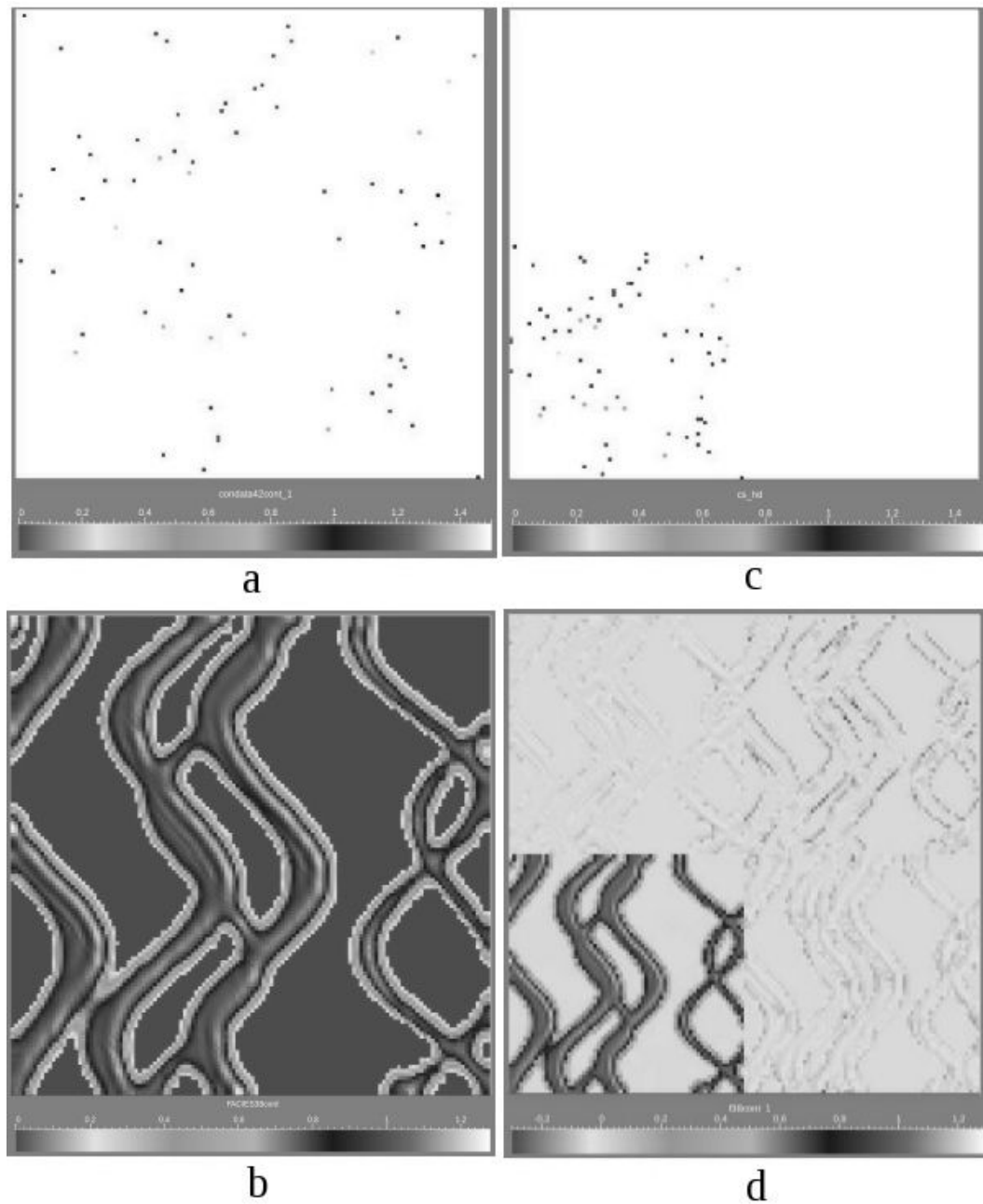


FIGURE 3.7 – Image d’entraînement et données de conditionnement à l’échelle grossière. (a) Données de conditionnement à l’échelle fine. (b) Image d’entraînement initiale. (c) Données de conditionnements à l’échelle grossière. (d) Image d’entraînement décomposée jusqu’à l’échelle grossière.



FIGURE 3.8 – Simulation des coefficients d'échelle. Résultats de la simulation.

la distance $d(DE(x), DE(y))$:

$$d(DE(x), DE(y)) = \sqrt{\frac{\sum_{i \in \Gamma} (Z(x_i) - Z(y_i))^2}{n_1} + \frac{\sum_{i \in \Omega} weight_{hd}^i (Z(x_i) - Z(y_i))^2}{n_2} + \frac{\sum_{i \in \Theta} weight_{sd}^i (Z(x_i) - Z(y_i))^2}{n_3}} \quad (3.9)$$

où $Z(x)$ et $Z(y)$ représentent les variables de la grille de simulation et de l'image d'entraînement, Γ l'ensemble des indices des données dures ou des valeurs déjà simulées, Ω l'ensemble des indices des données semi-dures, Θ l'ensemble des indices des données secondaires et $weight_{sd}^i$ les poids des données secondaires.

Le balayage de l'image d'entraînement s'arrête dès que l'on trouve une configuration dont la distance est inférieure à un seuil donné par l'utilisateur. La valeur de la cellule centrale y de l'image d'entraînement est alors assignée à la cellule x de la grille de simulation.

A la fin de cette étape, l'échelle grossière de la propriété simulée est cohérente avec les données dures, secondaire et de l'image d'entraînement (figure 3.8). Les coefficients d'ondelettes doivent à leur tour être simulés afin de pouvoir reconstruire la réalisation.

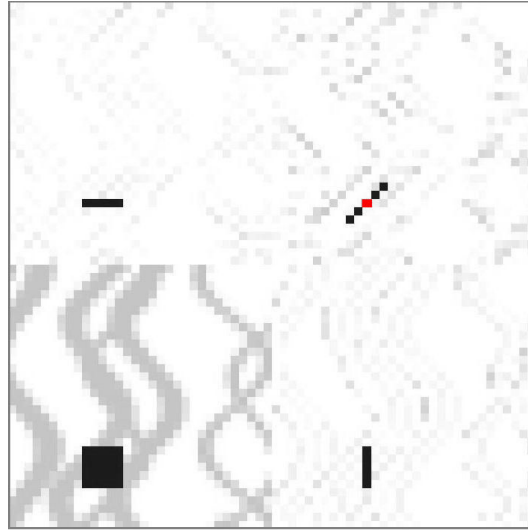


FIGURE 3.9 – MS-DSw voisinage de simulation. (Rouge) Coefficient d’ondelette simulé. (Noir) Voisinage utilisé. (Gris) Valeurs des coefficients de l’image d’entraînement.

Simulation conditionnée par les coefficients d’échelle

En utilisant les coefficients d’échelle générés précédemment, les coefficients d’ondelettes de même niveau⁷ sont simulés à partir de l’image d’entraînement décomposée. L’algorithme MS-DSw utilise un DE multi-échelles (figure 3.9) : la simulation d’un coefficient d’ondelette prend en compte le coefficient d’échelle associé, ses voisins et leurs coefficients d’ondelettes associés. Les coefficients d’ondelettes voisins capturent les variations des propriétés dans les différentes directions (verticale, horizontale et transversale) et permettent d’identifier les bords des objets simulés. L’extension du voisinage est fonction de la corrélation maximale définie par l’utilisateur (voir la section précédente). L’utilisation des coefficients d’échelle permet de conditionner localement la simulation des coefficients d’ondelettes même lorsqu’aucun coefficient n’est encore disponible, ce qui évite de simuler des valeurs aléatoires.

Les coefficients d’ondelettes sont simulés sans conditionnement par les données dures⁸. En effet, les données étant définies à l’échelle la plus fine, il existe une infinité de combinaisons de coefficients d’ondelettes possibles pour reconstituer les valeurs ponctuelles des données dures, même lorsque les coefficients d’échelle sont définis. Il est alors difficile de privilégier une combinaison plus qu’une autre. Cependant, il est possible de sélectionner aléatoirement ou à partir de l’image d’entraînement, une combinaison possible pour chaque réalisation afin de créer des conditionnements différents. Cette méthode n’est pas utilisée ici, mais une approche similaire est proposée dans la Section 3.3.2.

Néanmoins, il est impératif que les réalisations respectent les données observées. Le schéma proposé est composé des étapes suivantes :

⁷de même niveau en termes de transformation, mais de fréquences différentes, les coefficients d’ondelettes étant d’une résolution plus grande.

⁸les données secondaires étant déjà intégrées.

1. simulation des coefficients d'ondelettes conditionnés par les coefficients d'échelle $r - 1$, mais sans utiliser les données dures
2. reconstruction des coefficients d'échelle de résolution r plus fine en utilisant les coefficients d'ondelettes simulés
3. calcul des valeurs $av_{hd}(x)$ (données semi-dures, voir Section 3.3.1) et de leurs poids $weight_{hd}$ associés
4. mise à jour des coefficients d'échelle r par les données semi-dures en fonction de leurs poids :

$$Z(x_i) = Z(x_i) * (1 - weight_{hd}) + value_{hd}(x_i) * weight_{hd} \quad (3.10)$$

Cette procédure est répétée jusqu'à la reconstruction de l'échelle la plus fine. La modification des données est proportionnelle à l'incertitude d'estimation. Ainsi, lors de la dernière mise à jour, les valeurs des données dures sont automatiquement attachées à la grille de simulation.

Résultats et discussions

Nous proposons d'étudier les résultats de l'application de l'algorithme MS-DSw sur deux situations différentes. La première teste la régularisation multi-échelles lors d'une simulation non conditionnée. Le second test utilise des données de conditionnement primaires et secondaires. Les résultats sont comparés avec l'approche DS standard.

Simulation sans conditionnement La figure 3.10 présente une simulation sans données à partir d'une image d'entraînement, de même taille (128 par 128 cellules) que la grille de simulation, présentant un grand nombre de configurations possibles. Une distance de corrélation $max_{cl} = 12cellules$ et un seuil d'acceptation proche de zéro sont utilisés lors de la simulation. Grâce au conditionnement des coefficients d'échelle, l'algorithme MS-DSw parvient à obtenir un minimum proche de l'optimum global puisque la réalisation générée est très similaire à l'image d'entraînement. En effet, lorsqu'aucune donnée n'est disponible, la configuration la plus probable est celle proposée par l'utilisateur. On peut tout de même observer que la reconstruction n'est pas parfaite, ce qui montre que le conditionnement par les coefficients d'échelle ne permet pas de discriminer à chaque fois une configuration unique.

Simulations conditionnées Les données de conditionnement et l'image d'entraînement utilisée sont présentées dans la figure 3.11. Les images d'entraînement et de référence ont les mêmes dimensions (128 par 128) et ont été obtenues par simulation basée-objets de chenaux. Les données comportent 70 observations directes (données dures) échantillonnées aléatoirement dans le modèle de référence, deux jeux de données secondaires d'échelle r_{max-1} et r_{max-2} créés à partir du modèle de référence (ajout de bruit corrélé suivi d'un lissage).

Les simulations de l'algorithme MS-DSw avec les paramètres $min_{cl} = 2$, $max_{cl} = 12$, $weight_{crit} = 0.95$ et $weight_{sd} = 0.3$ sont comparées à celles de l'algorithme DS conditionné standard [100] avec les paramètres

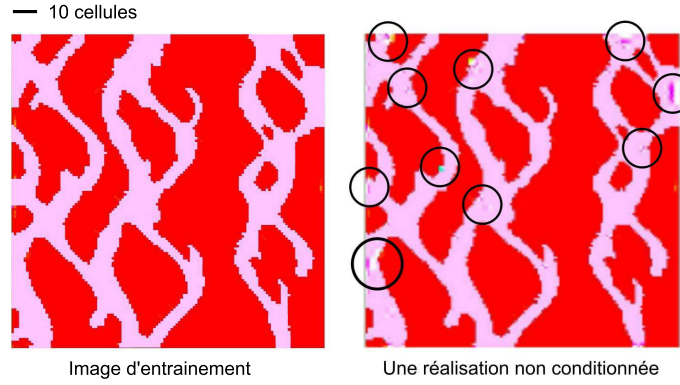


FIGURE 3.10 – MS-DSw sans conditionnement. Image d’entraînement (gauche) et une réalisation (droite). Les cercles noirs pointent les différences entre la réalisation et l’image d’entraînement.

$correlationlength = 12$, $dataeventsiz = 0.05 \times nu \times nv \times nw$, où (nu, nv, nw) sont les dimensions de la grille de simulation.

Les résultats montrent que dans les deux cas, les chenaux principaux de l’image de référence sont bien reproduits (figure 3.12). Les chenaux de la moitié inférieure et de la partie centrale sont mal représentés car ils sont mal définis pas les données secondaires et non échantillonnés par les données dures. Les simulations honorent correctement les données. Dans le cas des simulations utilisant les données secondaires 1, les résultats sont similaires pour les deux méthodes (figure 3.12). Lorsque les données secondaires 2 sont utilisées, les résultats obtenus par l’algorithme MS-DSw sont de moins bonne qualité que ceux obtenus par l’approche standard (figure 3.13). Des artefacts peuvent apparaître dans les modèles, en particulier près des données de conditionnement. De plus, des valeurs non physiques peuvent être simulées (valeurs négatives, ...). Ceci peut être facilement résolu par un traitement *a posteriori*, mais indique une faiblesse de l’algorithme.

Limites importantes de la méthode La qualité insuffisante des résultats obtenus par la méthode MS-DSw est due à plusieurs raisons :

- La méthode de conditionnement par les données n’est pas appropriée. Lors de la simulation des coefficients d’ondelettes, les données ne sont pas prises en compte, ce qui aboutit à la création de champs de propriétés qui ne sont pas compatibles avec les données. La mise à jour des coefficients d’échelle arrive trop tardivement : les propriétés ne peuvent être modifiées que localement, ce qui provoque des discontinuités irréalistes dans les structures ou la création de formes déconnectées. De plus, le conditionnement utilise des statistiques linéaires (moyennes locales) qui s’hybrident difficilement avec les méthodes MPS.
- La transformée en ondelettes décompose les champs de propriétés en coefficients indépendants par résolutions, ce qui pose un problème de propagation de l’erreur. En effet, une fois simulée, une résolution particulière ne peut plus être modifiée ; seuls les coefficients d’ondelettes, qui ont une

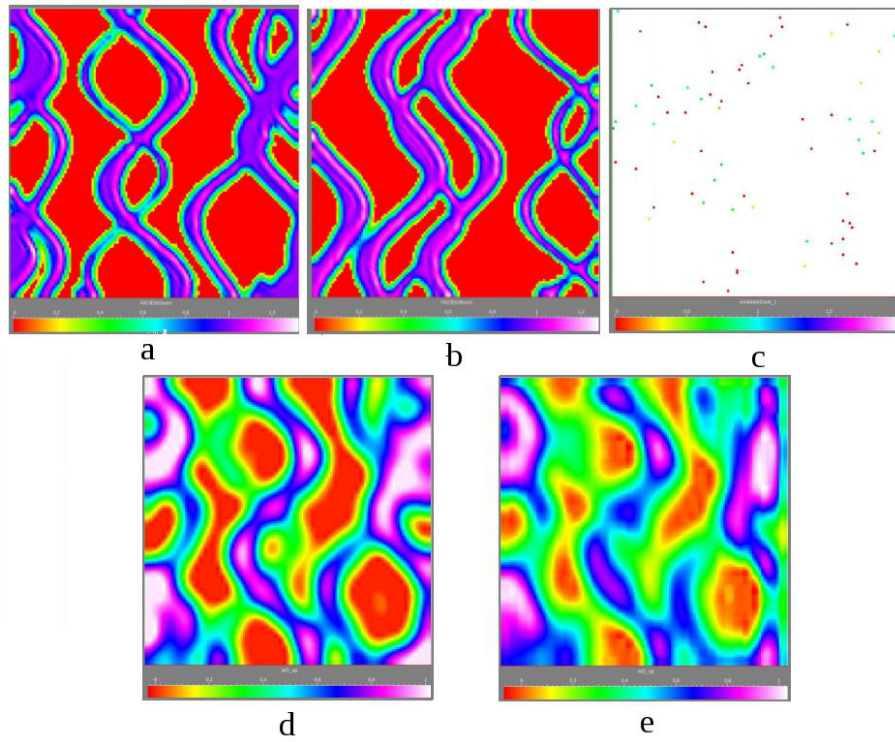


FIGURE 3.11 – MS-DSw conditionné : données initiales. (a) Image de référence. (b) Image d’entraînement. (c) Données dures. (d) Données secondaires 1. (e) Données secondaires 2.

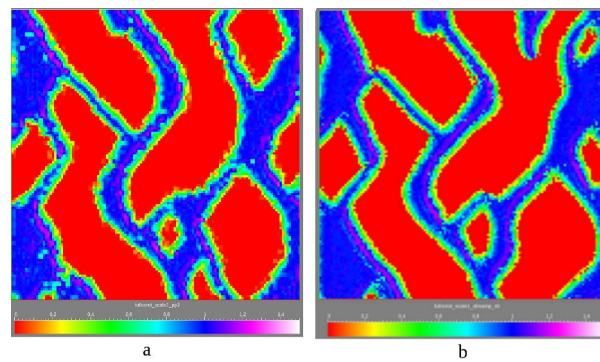


FIGURE 3.12 – Simulations conditionnées par les données secondaires 1. (a) MS-DSw. (b) DS standard.

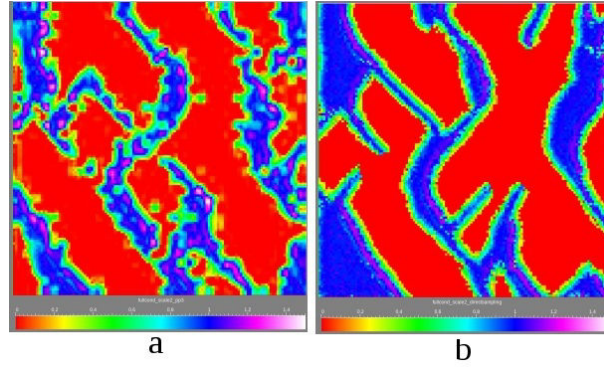


FIGURE 3.13 – Simulations conditionnées par les données secondaires 2. (a) MS-DSw. (b) DS standard.

gamme de fréquences bien définies, sont manipulés. Ceci suppose implicitement que les différentes résolutions peuvent être simulées de manière indépendante. Or cette hypothèse est fautive : un objet, tel qu'un chenal, est défini par plusieurs résolutions qui peuvent être corrélées [39, 69]. Ainsi, lorsqu'une échelle est mal simulée ou s'avère incompatible avec les données fines, elle ne peut pas être corrigée par la suite. Le conditionnement inter-résolutions a un effet bénéfique lorsque les échelles sont correctement simulées, mais peut avoir un effet négatif qui tend à propager les erreurs lorsqu'elles sont incorrectes.

Aspects positifs de la méthode Malgré les limitations ci-dessus, principalement dues à la paramétrisation utilisée, l'approche multi-échelles possède des points positifs :

- La régularisation entre les résolutions permet de créer des modèles de moindre énergie (configurations simulées plus proches de celles disponibles dans l'image d'entraînement).
- L'approche multi-échelles permet de réduire les temps de calculs, même lors de l'application de traitements intermédiaires grâce à la réduction de la dimension du problème. Les temps de simulation moyens par réalisation, pour le cas décrit ici, sont de 35s pour l'algorithme MS-DSw, 67s pour l'algorithme DS et 153s pour l'algorithme DS avec traitement *a posteriori*.

3.3.2 MS-DS basé sur la décomposition en pyramides gaussiennes (MS-DSgp)

Nous avons vu dans la Section précédente que l'approche multi-échelles permet l'application efficace de traitements additionnels et régularise la simulation d'une échelle grâce à son approximation grossière. Néanmoins, la paramétrisation en ondelettes s'est avérée inadéquate pour les raisons que nous rappelons ici : il est difficile d'intégrer les données de conditionnement d'échelle fine et il est impossible de modifier les résolutions déjà simulées, ce qui provoque une propagation des erreurs.

Afin de tirer avantage de l'approche multi-échelles tout en évitant les inconvénients de la simulation par échelles indépendantes, nous proposons dans cette section une approche basée sur la paramétrisation en pyramides gaussiennes. Cette paramétrisation est présentée dans la Section 2.4. Sa différence majeure par rapport aux ondelettes vient du fait que chaque échelle intègre l'ensemble des résolutions plus grossières. Grâce à cette paramétrisation, il est possible, à tout moment, de corriger les résolutions grossières précédemment simulées, ce qui évite l'effet de propagation de l'erreur, tout en bénéficiant de l'effet de régularisation des échelles supérieures. De plus, l'échelle la plus fine est représentée par les coefficients de la propriété d'origine. Ceci permet de gérer plus facilement l'intégration des données d'échelles différentes mais à le désavantage d'augmenter le nombre global de paramètres.

Cependant, le nombre de paramètres manipulés aux échelles grossières et intermédiaires reste très réduit et des traitements complexes peuvent être appliqués. Dans cette section, nous nous intéressons particulièrement à la continuité des chenaux simulés. Un traitement basé sur l'analyse de la connectivité est appliqué au cours de la simulation des résolutions intermédiaires.

Dans cette section, seules des données dures sont utilisées pour conditionner les réalisations ; l'intégration de données secondaires peut néanmoins s'effectuer de la même façon que dans la Section 3.3.1. L'algorithme MS-DSgp se compose alors de trois étapes principales et une optionnelle :

1. décomposition de l'image d'entraînement
2. simulation des données de conditionnement, hd_r , à l'échelle r à partir des données fines $r = r_{max}$, de l'échelle plus grossière $r - 1$ et de l'image d'entraînement.
3. simulation des variables d'échelle r par l'algorithme DS conditionné par les coefficients de l'échelle précédente $r - 1$ et des données hd_r ⁹.
4. (optionnelle) traitement et simulations itératives des variables simulées en (3).
5. retour en (2) avec $r = r + 1$ jusqu'à ce que $r = r_{max}$.

Simulation des données de conditionnement

Contrairement à la méthode MS-DSw, les données aux échelles grossières ne sont pas estimées par des moyennes locales mais sont simulées à partir de l'image d'entraînement par une méthode MPS. Dans la figure 3.14, on cherche à simuler la donnée grossière (identifiée par un point d'interrogation) associée aux données fines et aux coefficients d'échelle grossière $r - 1$ entourés en bleu. Dans ce cas particulier, on prend également en compte les voisins adjacents à l'échelle r et leurs données fines associées. Cependant, il est important de garder à l'esprit que plus le DE est grand, plus le temps de calcul est élevé.

Ce DE est ensuite utilisé pour scanner l'image d'entraînement. Une fonction de distance est définie de la même manière que dans l'algorithme standard, mais

⁹sauf pour la moyenne qui est toujours donnée par (2) ou que toutes les variables d'échelles grossières sont associées à des données fines

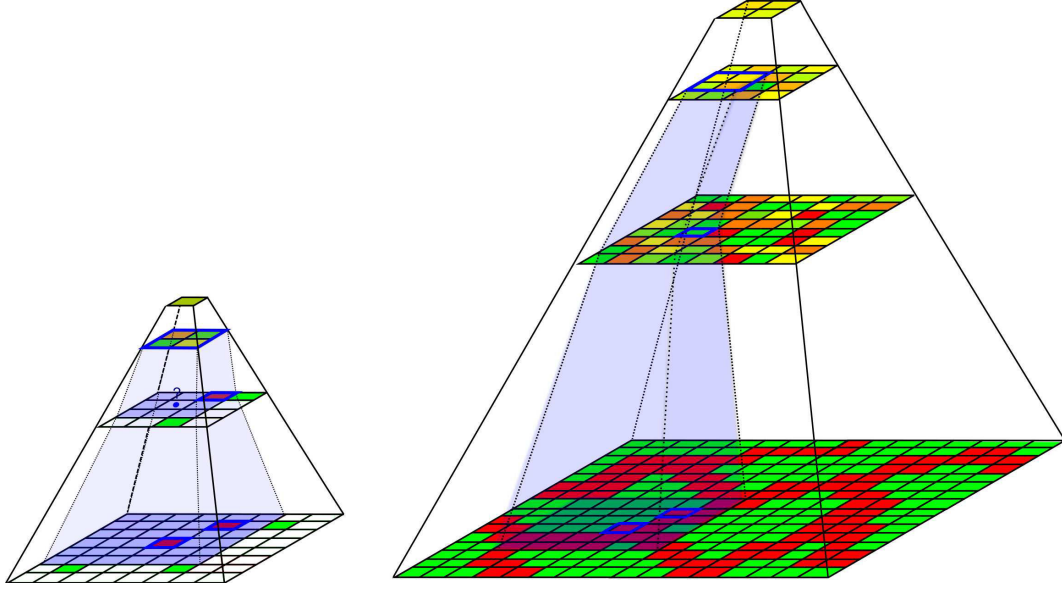


FIGURE 3.14 – MS-DSgp simulation des données de conditionnement à l'échelle grossière r . Recherche de la meilleure configuration dans la pyramide d'entraînement (droite) en fonction de l'événement contenant les données fines ($r = r_{max}$) et l'échelle grossière $r - 1$.

elle peut être ajustée en fonction de l'échelle des coefficients. Dans le cas présent, la fonction de distance se compose de trois éléments principaux de même poids reliés aux échelles $r - 1$, r et r_{max} :

$$d(DE(x), DE(y)) = \sqrt{\frac{\sum_{i \in \Lambda_{r_{max}}} (Z(x_i) - Z(y_i))^2}{|\Lambda_{r_{max}}|}} + \sqrt{\frac{\sum_{i \in \Lambda_r} (Z(x_i) - Z(y_i))^2}{|\Lambda_r|}} + \sqrt{\frac{\sum_{i \in \Lambda_{r-1}} (Z(x_i) - Z(y_i))^2}{|\Lambda_{r-1}|}} \quad (3.11)$$

où $\Lambda_{r_{max}}$, Λ_r , Λ_{r-1} représentent les ensembles de coefficients d'échelles r_{max} , r , $r - 1$ inclus dans le DE et $|\cdot|$ représente le cardinal de l'ensemble. Le balayage de l'image d'entraînement s'arrête dès que la distance est inférieure à une valeur seuil donnée.

Ce procédé permet de simuler des données de conditionnement compatibles avec l'image d'entraînement, mais peut s'avérer coûteux lorsque la densité de données est importante. Il est également possible de simuler simultanément des données semi-dures adjacentes aux données dures afin de démarrer avec une configuration compatible avec les données¹⁰. Ces données semi-dures peuvent par la suite être re-simulées.

¹⁰cela est similaire à l'utilisation d'un chemin conditionné, mais dans ce cas particulier les données fines sont prises en compte dans la simulation

Simulation d'une échelle r

De même que pour la méthode MS-DSw, les coefficients de résolutions $r + 1$ sont conditionnés par l'échelle grossière r . Cependant, dans l'algorithme MS-DSgp présenté ici, les données simulées sont déjà incorporées à l'échelle r et conditionnent directement la simulation des coefficients restants (figure 3.15). La fonction de distance utilisée est semblable à l'équation 3.11 sans le terme $\Lambda_{r_{max}}$ des données fines.

Dans bien des cas, il est avantageux d'utiliser des images d'entraînement plus grandes que la grille de simulation afin d'obtenir un plus grand nombre de configurations possibles et de meilleures statistiques. Cependant, l'utilisation d'images d'entraînement de grandes dimensions peut s'avérer très coûteuse. Nous proposons ici un contrôle du balayage de l'image d'entraînement pour la simulation d'une échelle r , basé sur l'analyse des coefficients simulés à l'échelle $r - 1$.

La figure 3.16 illustre le fonctionnement de ce contrôle. L'idée est qu'il est plus probable et potentiellement plus cohérent de choisir un coefficient d'échelle r dans le même secteur d'où provient son coefficient d'échelle $r - 1$ associé. Ainsi, la surface balayée de l'image d'entraînement est limitée, ce qui permet de réduire le nombre d'évaluations de l'événement de donnée et d'augmenter les performances de l'algorithme. Cette méthode nécessite l'enregistrement des positions dans l'image d'entraînement des coefficients assignés dans la grille de simulation, ce qui représente peu de mémoire.

Il est également possible d'utiliser la valeur des coefficients d'échelle $r - 1$ afin d'améliorer le balayage. Par exemple, imaginons qu'un coefficient d'échelle grossière ait une valeur proche 1 et que l'on simule des variables discrètes $Z \in \{0, 1\}$. Clairement, d'après la valeur de l'échelle grossière il est plus probable de trouver des configurations compatibles dans l'image d'entraînement ayant elles-même un coefficient d'échelle $r - 1$ proche de 1. Ainsi, lors du balayage de l'image d'entraînement, il est possible d'évaluer un DE uniquement dans les cas où les valeurs des coefficients $r - 1$ sont assez proches. La figure 3.17 illustre ce principe.

La notion de « valeurs proches » doit évoluer au cours du processus multi-échelles. Lors des simulations des premières échelles, le modèle est encore peu conditionné et il existe un grand nombre de configurations possibles pour un état grossier donné. Il est alors nécessaire de ne pas trop contraindre le balayage de l'image d'entraînement afin de ne pas limiter les options possibles. Une sélection trop importante des positions visitées peut conduire à une propagation des erreurs, la valeur des coefficients grossiers devenant le critère de choix principal. De plus, le nombre de paramètres étant largement réduit pour les échelles grossières, il n'est pas nécessaire d'être très sélectif lors du balayage. La contrainte doit devenir plus active pour les résolutions fines. Ainsi un DE à l'échelle r , comportant un coefficient d'échelle $r - 1$ de valeur V_{r-1}^{sim} , est évalué si est seulement si la valeur V_{r-1}^{TI} de l'image d'entraînement est dans l'intervalle :

$$V_{r-1}^{sim} - \alpha_r^V \leq V_{r-1}^{TI} \leq V_{r-1}^{sim} + \alpha_r^V \quad (3.12)$$

où α_r^V est une valeur qui évolue avec la résolution. Dans ce travail, α_r^V est définie

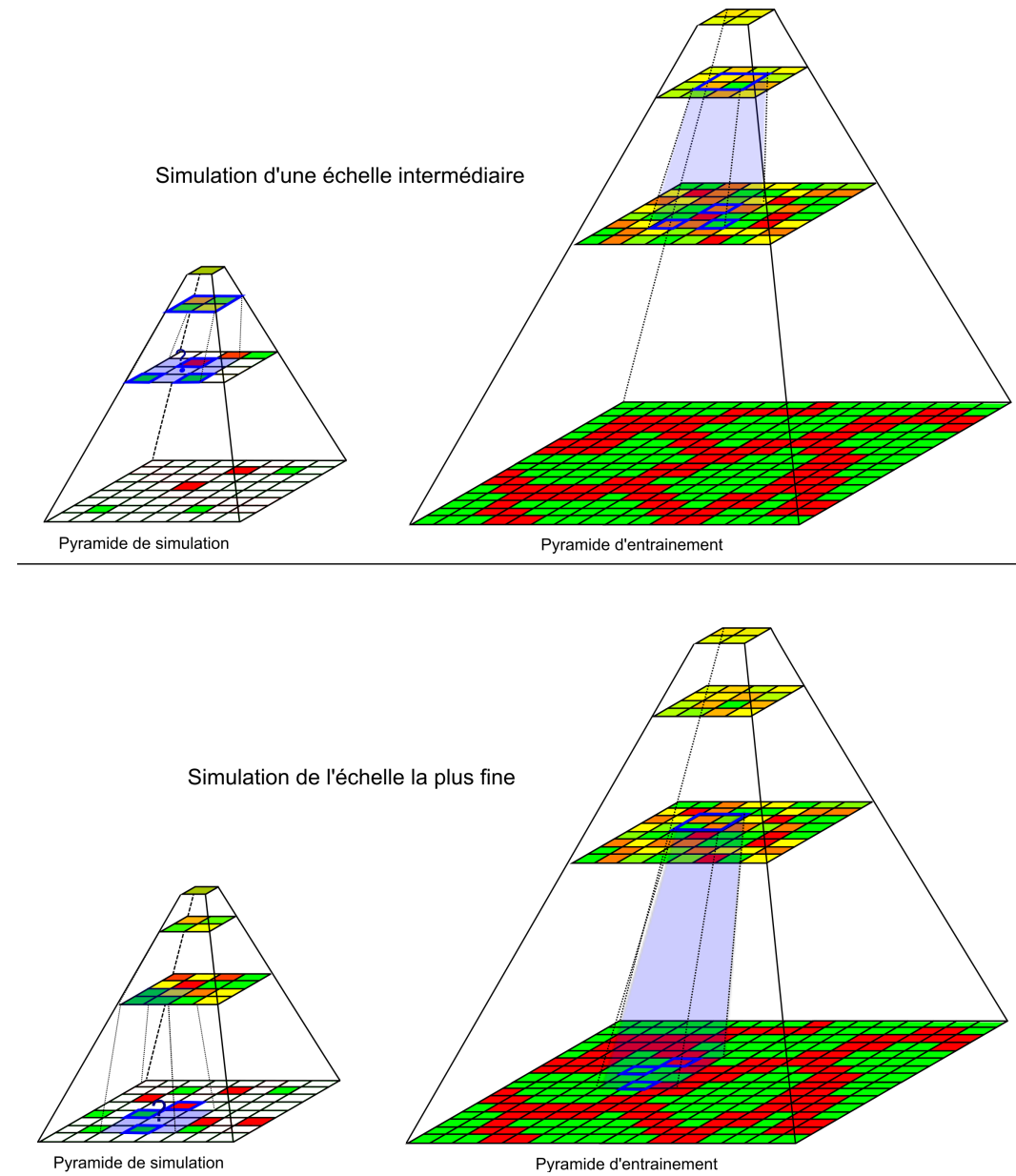


FIGURE 3.15 – MS-DSgp simulation des échelles. Simulation d'une échelle intermédiaire r conditionnée par l'échelle $r - 1$ (haut). Simulation de l'échelle r_{max} (bas) : les données de conditionnement sont les données dures.

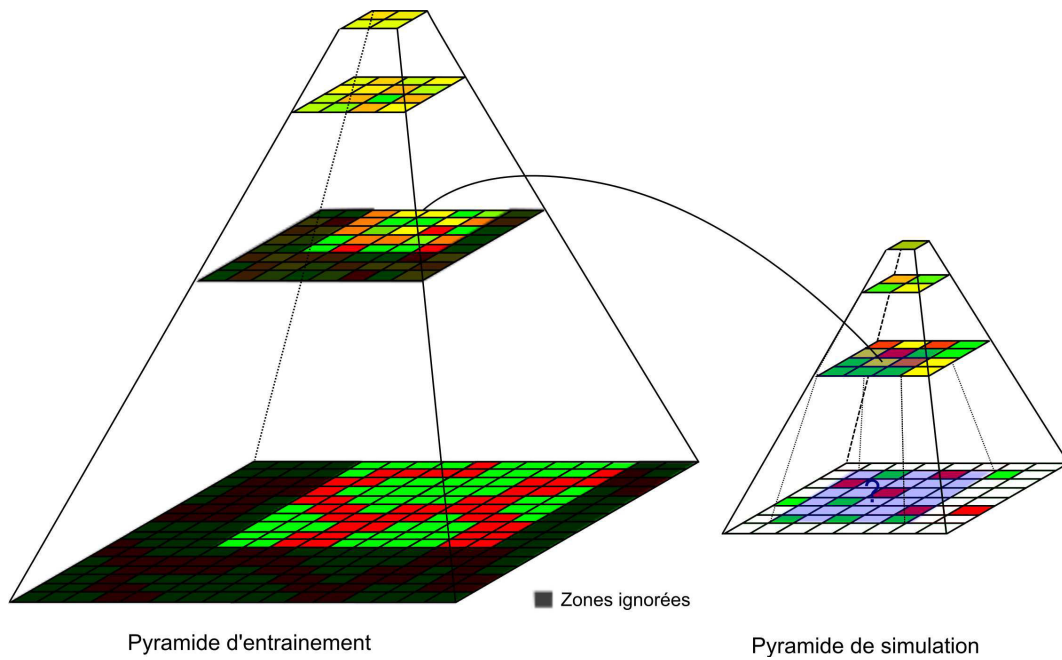


FIGURE 3.16 – Contrôle du balayage de l'image d'entraînement par la position des coefficients. Le balayage de l'image d'entraînement est contrôlé par la position dans l'image d'entraînement des coefficients de l'échelle précédente. Seules les positions voisines sont parcourues (cellules éclairées).

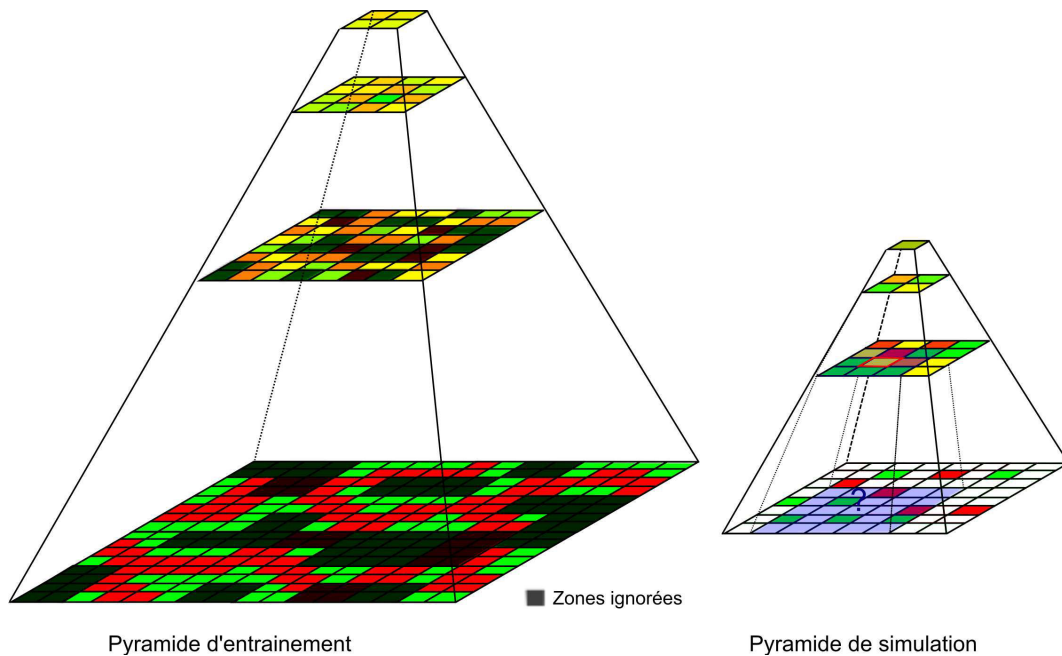


FIGURE 3.17 – Contrôle du balayage de l'image d'entraînement par la valeur du coefficient grossier. Le balayage de l'image d'entraînement est contrôlé par la valeur du coefficient de l'échelle précédente. Seule les positions dans l'image d'entraînement ayant des valeurs proches de ce coefficient sont parcourues (cellules éclairées).

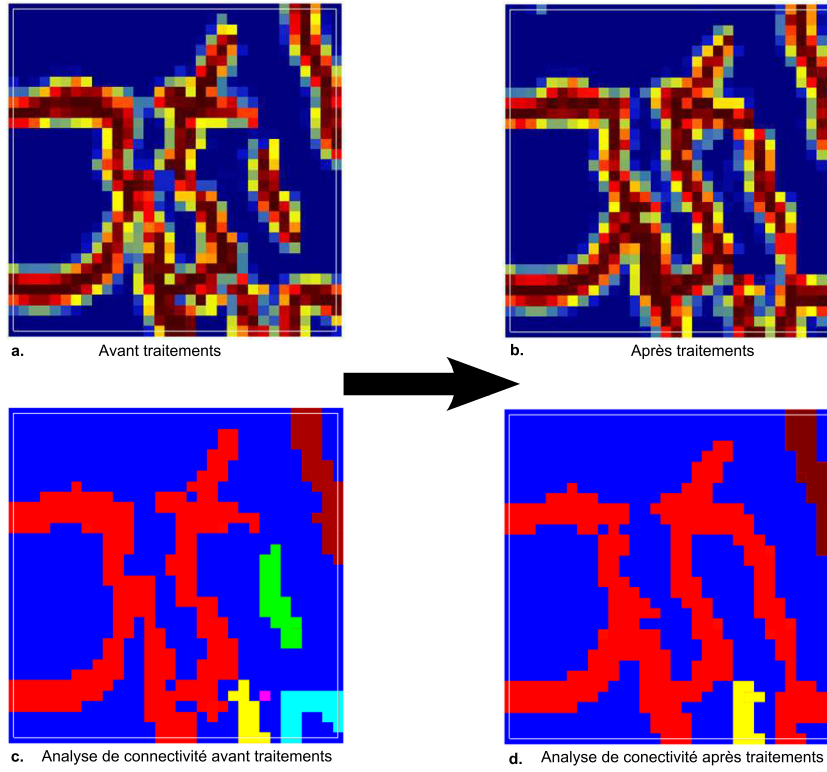


FIGURE 3.18 – Traitements intermédiaires. Traitements basés sur l’analyse de la connectivité et la re-simulation des variables.

par :

$$\alpha_r^V = 0.5 + \frac{r_{max} - r}{2r_{max}} \quad (3.13)$$

mais doit être adaptée en fonction des propriétés simulées.

Traitements aux échelles intermédiaires

Ces traitements permettent d’améliorer la qualité des réalisations. Ils s’effectuent au cours des échelles intermédiaires pour lesquelles le nombre de paramètres est fortement réduit par rapport à la dimension initiale du problème. Ainsi, il est possible d’appliquer des traitements complexes sans augmenter significativement les temps de calculs. Dans ce travail, nous nous intéressons particulièrement à la connectivité des chenaux simulés.

Traitement par re-simulation des variables Un premier traitement simple consiste à re-simuler les variables de la pyramide de simulation à partir de l’image d’entraînement. Il est possible d’utiliser un *template* plus petit afin de réduire le temps de calcul et de choisir aléatoirement les éléments du DE. Ce traitement permet d’enlever des configurations « bruitées » (p. ex., une valeur de faciès A se retrouve au milieu de faciès B), mais ne permet pas de changer fondamentalement la structure du champ initial, sauf lorsque la dimension du problème est très réduite.

La re-simulation des variables est souvent associée à des phénomènes d’éro-

sions ou de dilatation des structures, notamment lorsque le DE utilisé est de grande taille. Cependant, lors de la re-simulation des variables, les coefficients de l'échelle grossière supérieure sont inclus dans le calcul de la distance et ne sont pas modifiés, ce qui empêche une perturbation trop importante des modèles.

Analyse de la connectivité L'implémentation actuelle de ce traitement n'est valable que pour des modèles bimodaux (continu ou discret) d'objets connectés. Cependant, il est possible de l'étendre aux cas multi-modaux en un minimum de changements. Ce traitement n'est pas appliqué à l'échelle la plus fine. Il comporte les étapes suivantes :

1. recherche d'une valeur de séparation des modes. Dans ce travail, nous utilisons la valeur moyenne théorique des deux modes. Les exemples d'applications étant basés sur la simulation de variables discrètes $Z \in \{0, 1\}$, la valeur de séparation est fixée à 0.5 pour toutes les échelles.
2. calcul des corps connectés dans le modèle simulé en utilisant la valeur de séparation (figure 3.18).
3. balayage du modèle simulé à la recherche de déconnexions verticales ou horizontales (ou transversales). Une déconnexion est une configuration où deux corps connectés distincts sont séparés (verticalement, horizontalement ou transversalement) par une seule cellule.
4. analyse des valeurs encadrant la discontinuité et (possible) modification de la valeur centrale. Il est possible de favoriser certaines directions par rapport à d'autres en fonction des connaissances *a priori* sur l'élongation des structures. La valeur moyenne ou maximale des cellules déconnectées est ensuite assignée à la variable centrale afin d'établir la connexion entre les deux corps.

L'analyse des déconnexions étant basée sur les variables adjacentes, ce traitement n'est pas efficace à l'échelle fine.

Chemin de visite des variables Lors de la re-simulation des variables, il est possible d'utiliser un chemin basé sur les valeurs des distances associées aux variables assignées à la grille de simulation ou à l'analyse de connectivité. Le chemin pourra par exemple visiter en priorité les configurations les plus problématiques ou bien les variables modifiées lors de l'analyse de la connectivité afin de valider les valeurs attribuées.

Ces différents traitements peuvent être appliqués successivement et plusieurs fois par résolution en fonction de la dimension du problème. Le traitement aux échelles grossières permet d'effectuer des modifications majeures sur un modèle approximatif qui conditionnent ensuite les échelles fines grâce à la régularisation inter-échelles.

Résultats et discussions

Dans ce qui suit, nous utilisons une image d'entraînement de taille (512×512) bien supérieure à la grille de simulation (128×128). L'approche standard

exigeant un temps de calcul trop important pour une telle image d'entraînement, une sous image de dimension (128×128) est utilisée pour les simulations DS standards (voir figure 3.19).

Test de minimisation Dans ce premier test, les données de conditionnement sont extraites de l'image d'entraînement utilisée dans l'approche standard (figure 3.20). Dans ce cas, les réalisations de probabilité maximale devrait être semblables à la partie de l'image d'entraînement échantillonnée. La figure 3.20 montre une réalisation de l'algorithme MS-DSgp utilisant l'image d'entraînement complète et une réalisation de l'algorithme DS standard utilisant la partie de l'image d'entraînement échantillonnée (128 par 128). L'algorithme MS-DSgp utilise un DE contenant un nombre maximal de 15 éléments et génère la simulation en 80s, tandis que le DS standard utilise un DE contenant un nombre maximal de 25 éléments et génère la simulation en 186s. Bien que l'image d'entraînement utilisée dans l'approche standard respecte parfaitement les données, l'algorithme ne parvient pas à obtenir une réalisation ayant une énergie minimale. Dans l'approche multi-échelles, l'algorithme trouve le minimum global (ou proche) et génère un modèle très proche de celui présent dans l'image d'entraînement. Il est à noter que cela peut poser un problème de variabilité des modèles simulés. Cependant, nous rappelons cette réalisation a été générée pour un seuil d'acceptation des DE très faible (proche de zéro). Lorsque ce seuil est plus élevé, l'algorithme produit des réalisations plus variées car des DE ayant une distance non nulle sont acceptés lors du parcours de la TI.

Validation de la méthode Les méthodes MS-DSgp et DS standard sont utilisées pour générer 50 réalisations à partir des données dures présentées dans la figure 3.21 et les images d'entraînement de la figure 3.19. Là encore, pour des raisons de temps de calculs, l'approche standard utilise une image d'entraînement de 128 par 128 cellules. Plusieurs tailles maximales d'évènements de données (15, 25, 40)¹¹ sont utilisées dans l'approche standard afin d'analyser l'évolution de la qualité des réalisations produites et de comparer les temps de calculs avec l'approche multi-échelles. Pour toutes les approches, une valeur seuil proche de zéro est utilisée lors du balayage de l'image d'entraînement afin que ce paramètre n'ait pas d'influence majeure sur les résultats des simulations.

La qualité des réalisations est analysée de manière qualitative et quantitative. Les critères quantitatifs utilisés sont :

- **Analyse de dissimilarité des réalisations avec l'image d'entraînement** : les outils d'analyse développés dans [118] sont utilisés pour comparer les résultats. Cette analyse intègre plusieurs indicateurs de comparaisons essentiellement dérivés de l'analyse des corps connectés. La liste de ces indicateurs est donnée dans le tableau 3.1. Contrairement à ce qui est présenté dans [118], les valeurs de dissimilarité sont calculées à partir de la distance de Hellinger [79]. Cette méthode de calcul est proposée par [117]. Ces valeurs sont ensuite analysées à partir d'une matrice de dissimilarité et d'une représentation MDS (*MultiDimensional Scaling*, voir [118]).

¹¹en dessous de 15, les réalisations sont beaucoup trop discontinues ; au-dessus de 40, l'image d'entraînement devient trop petite et les temps de calculs beaucoup trop importants

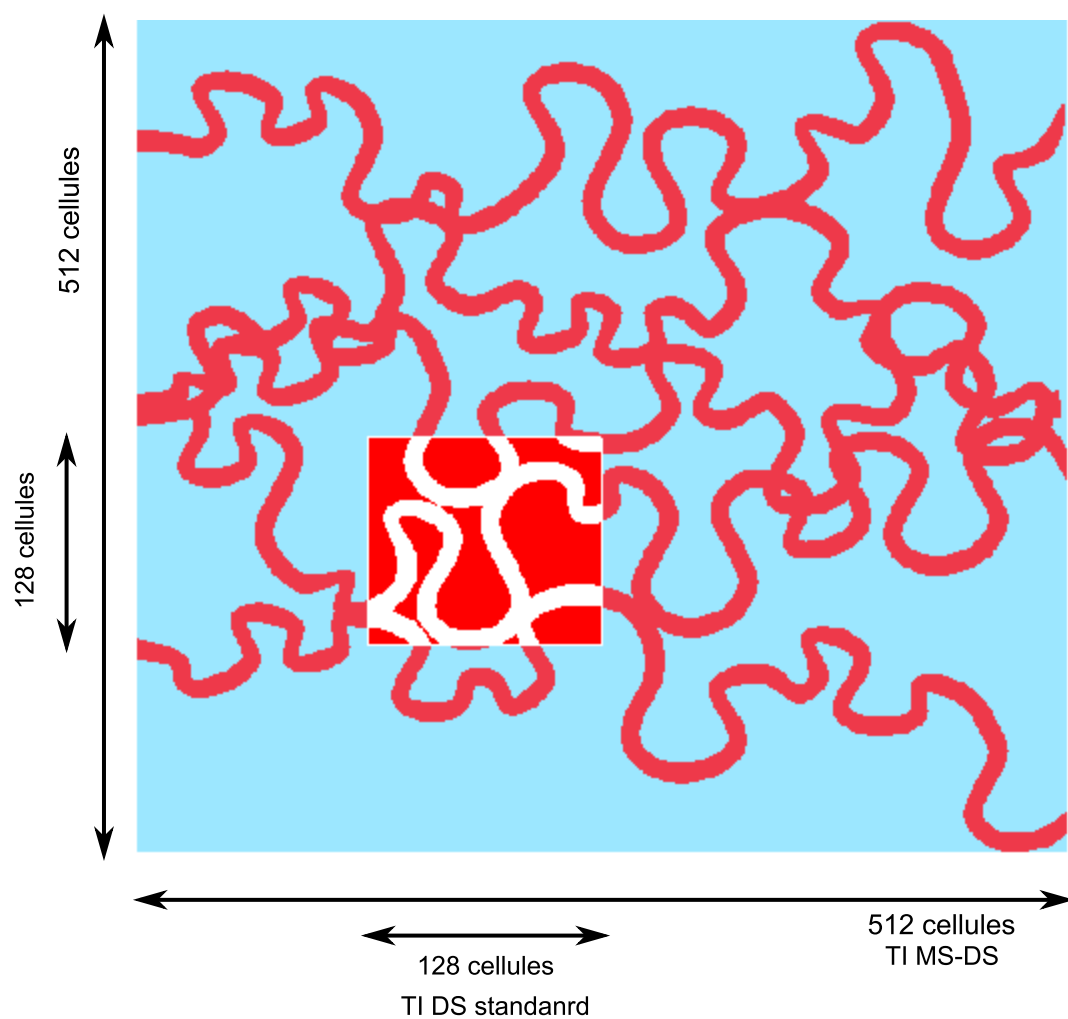


FIGURE 3.19 – Images d’entraînement. L’image la plus grande est utilisée dans l’algorithme MS-DSgp, tandis que seule une partie (rouge et blanc) est utilisée dans l’approche standard.

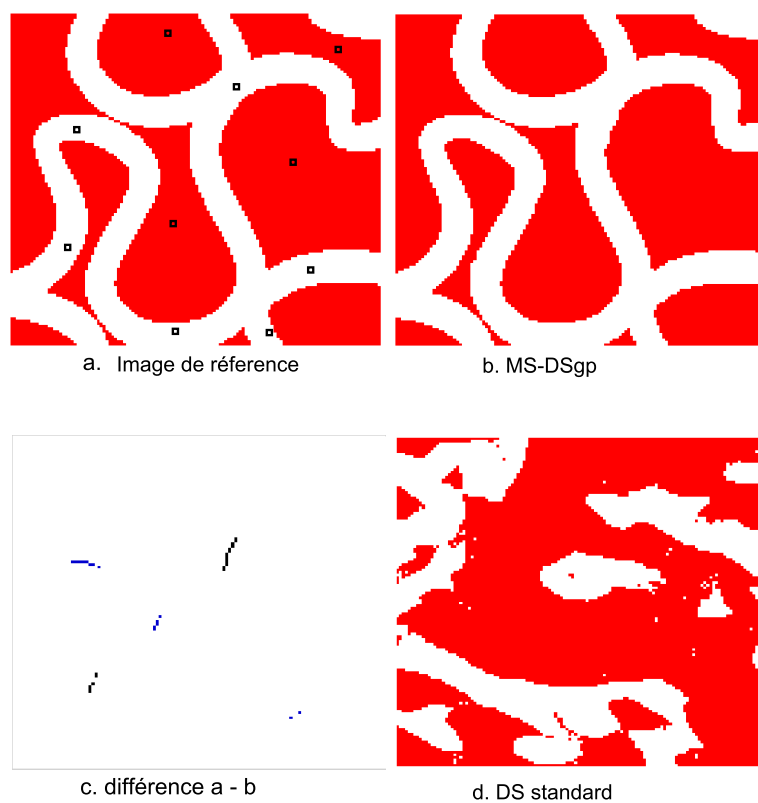


FIGURE 3.20 – Test de minimisation. L’approche multi-échelles parvient à s’approcher du modèle optimal alors que l’approche standard échoue. Les carrés noirs représentent les données de conditionnement utilisées.

| Indicateur | Description |
|---|--|
| Proportion de faciès | $p_f = \frac{n_f}{n}$, où n_f et n représente le nombre de cellules de faciès f et la taille de la grille |
| Probabilité de connexion | Indicateur global de connectivité |
| Densité de corps connectés | Nombre de corps connectés par unité de volume |
| Proportion de corps unitaires | Proportion de corps constitués d'une seule cellule |
| Analyses de formes des corps connectés | Analyses basées sur les volumes et les formes des objets (sphère, boîte) |

Tableau 3.1 – Indicateurs de comparaison des réalisations et de l'image d'entraînement. Voir [118] pour plus d'informations. Ces indicateurs sont utilisés dans le calcul de la dissimilarité, elle-même représentée sous la forme de matrice (figure 3.22).

| MS-DSgp | DS-t15 | DS-t25 | DS-t40 |
|---------|--------|--------|--------|
| 81 s | 76 s | 226 s | 576 s |

Tableau 3.2 – Temps de simulation moyens par réalisation. Seul le DS standard utilisant un nombre maximum de 15 éléments a une efficacité comparable à l'approche multi-échelles en termes de temps de calcul. Il est à noter que la méthode multi-échelles utilise une image d'entraînement de dimension (512×512) alors que les approches DS standards utilisent une image (128×128) .

- **Analyse du respect au conditionnement** : il s'agit d'une analyse simple du conditionnement. Pour chaque donnée, la moyenne des valeurs des cellules adjacentes est calculées et est retranchée à la valeur de la donnée (la valeur de conditionnement étant toujours honorée par les réalisations). On considèrera qu'une donnée est statiquement bien respectée si l'écart moyen est faible.

Résultats Le tableau 3.2 donne les temps de calculs moyens pour chacun des tests. L'approche MS-DSgp a un temps de calcul comparable à la simulation DS-t15 alors que cette dernière utilise une image d'entraînement 16 fois plus petite, ce qui démontre l'efficacité de l'approche multi-échelles.

La figure 3.21 présente un échantillon de réalisations produites par chacune des méthodes. Les réalisations MS-DSgp sont d'une qualité clairement supérieure aux autres : elles reproduisent correctement les corps connectés de l'image d'entraînement, ne sont pas bruitées et respectent les données de conditionnement. Pour l'approche DS standard, la meilleure qualité est obtenue pour la simulation DS-t15. Ceci peut s'expliquer par l'utilisation d'une image d'entraînement de petite taille. En effet, lorsque la taille du DE augmente, le nombre de configurations différentes disponibles dans l'image d'entraînement diminue, ce qui nuit la répétabilité des structures et à la qualité des réalisations. Les réalisations DS-t40 forment des morceaux d'objets bien dessinés mais déconnectés. Grâce à l'utilisation d'une image d'entraînement de plus grande taille, l'algorithme MS-DSgp parvient à trouver un plus grand nombre de configurations compatibles avec les données. Malheureusement, il est impossible d'utiliser

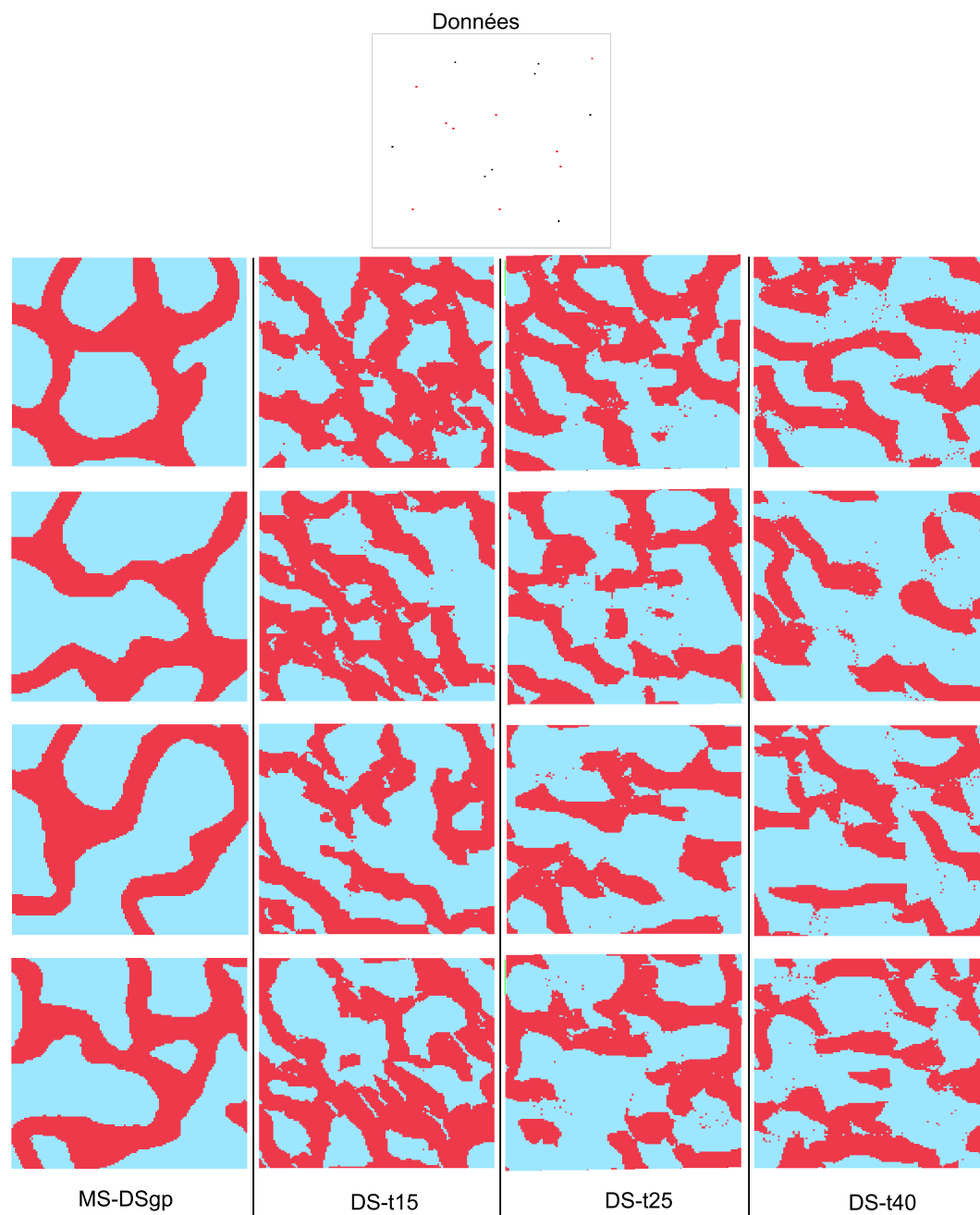


FIGURE 3.21 – Réalisations MD-DSgp et DS standard. Données de conditionnement (haut) et quelques réalisations associées.

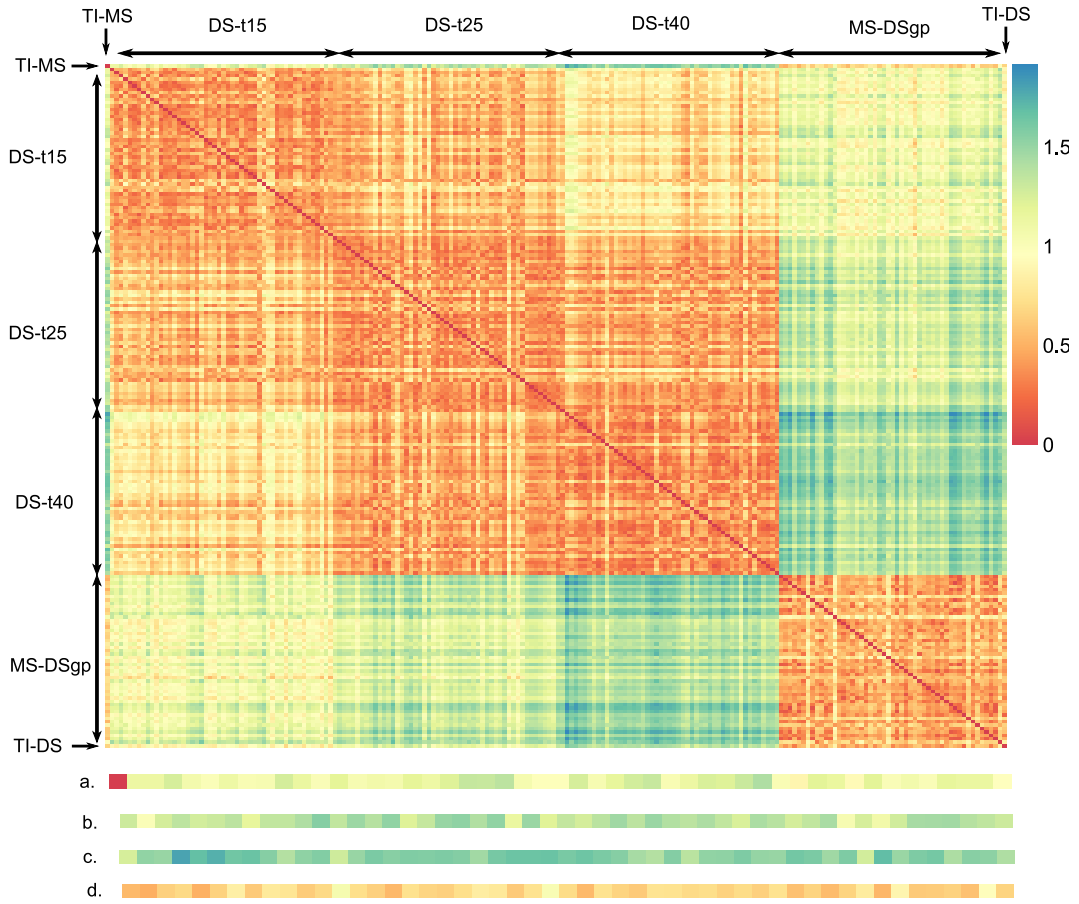


FIGURE 3.22 – Matrice des dissimilarités. En dessous, agrandissement des dissimilarités entre l'image d'entraînement (512×512) et les réalisations (première ligne de la matrice) (a) DS-t15, (b) DS-t25, (c) DS-40 et (d) MS-DSgp. La teinte rouge dénote une forte similarité tandis que la verte en dénote une faible.

cette image dans les approches DS standards.

La figure 3.22 représente les valeurs de dissimilarité entre les différents modèles. Nous pouvons observer une nette ségrégation entre les modèles MS-DSgp et les modèles DS standards. Les valeurs de similarité sont beaucoup plus grandes entre les images d'entraînement et les réalisations MS-DSgp, ce qui démontre un meilleur respect des structures géologiques représentées par l'image d'entraînement. La figure 3.23 est une représentation de la matrice des dissimilarités par un nombre réduit de dimensions (ici 2). Elle a pour objectif de représenter les modèles comme des points d'un espace euclidien de telle façon que les distances entre les points équivalent aux dissimilarités entre les modèles. Deux dimensions sont ici utilisées pour classer les modèles. Afin de vérifier qu'elles sont suffisantes pour « résumer » les informations fournies par la matrice des dissimilarités, une analyse de la courbe d'éboulis (*scree plot*) et du diagramme de Shepard est effectuée dans la figure 3.24. L'analyse MDS montre clairement que les modèles issus de la simulation MS-DSgp ont des caractéristiques plus proches de l'image d'entraînement que ceux générés par la méthode standard.

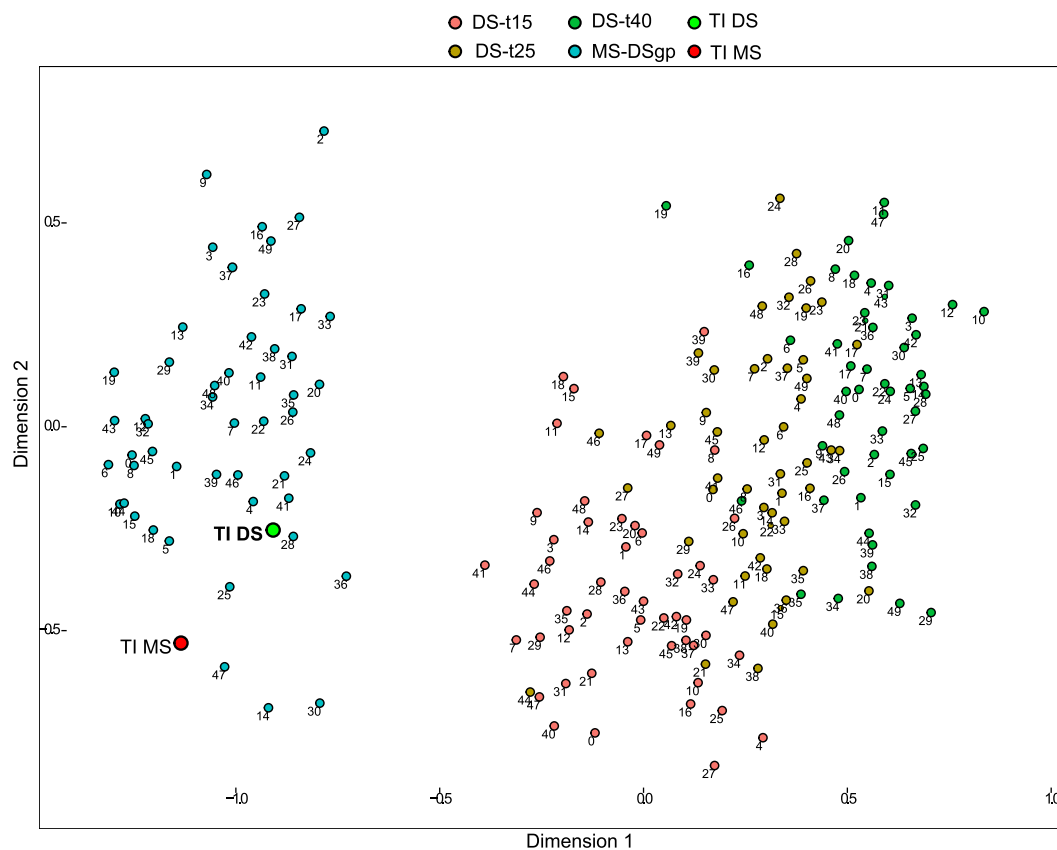


FIGURE 3.23 – Représentation MDS. Séparation claire des réalisations MS-DSgp et des réalisations DS standards. Les réalisations MS-DSgp sont proches des images d'entraînement.

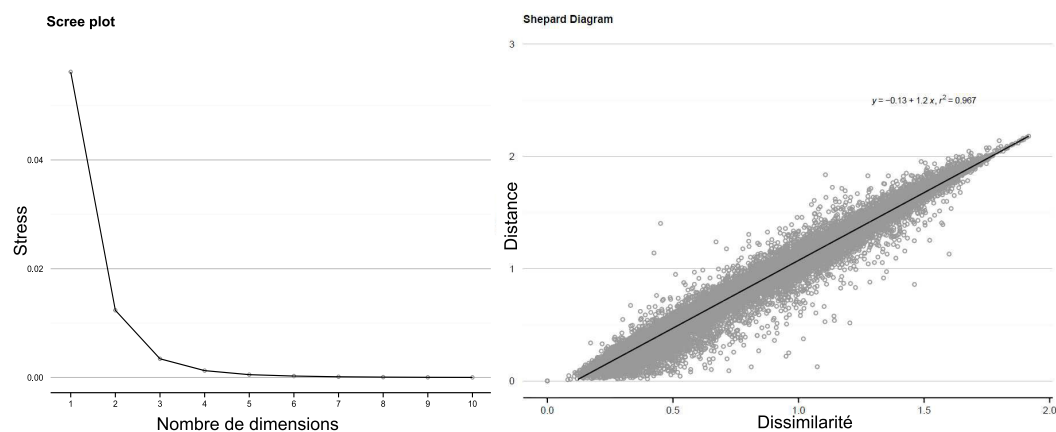


FIGURE 3.24 – Validation de la reproduction MDS. (a) Le *scree plot* représente la contribution des dimensions dans la reproduction des dissimilarités. (b) Le diagramme de Shepard représente la distance entre les modèles (dans la représentation MDS) par rapport à leur dissimilarité réelle. Une corrélation de 0.967 prouve que le digramme MDS reproduit bien les dissimilarités.

| MS-DSgp | DS-t15 | DS-t25 | DS-t40 |
|---------|--------|--------|--------|
| 0.024 | 0.122 | 0.145 | 0.151 |

Tableau 3.3 – Qualité du respect des données de conditionnement. Écarts moyens pour l'ensemble des données de la différence entre la valeur des données de puits et la moyenne des valeurs simulées dans son voisinage direct. Une valeur faible indique un bon respect des données.

Le tableau 3.3 donne les valeurs des indicateurs de la qualité du respect du conditionnement. Là encore, la méthode MS-DSgp obtient un meilleur résultat, ce qui montre que les données contraignent efficacement le développement des chenaux et sont bien intégrées dans les structures.

3.4 Conclusions partielles et perspectives

L'effet de régularisation inter-échelles permet de générer des réalisations de meilleure qualité que les approches standards. D'un point de vue de minimisation, l'effet multi-échelles aide à éviter les minima locaux, ce qui permet de créer des modèles fortement probables pour une image d'entraînement et un conditionnement donnés. Ainsi, l'espace des incertitudes est mieux échantillonné par l'ensemble des réalisations générées.

Cependant, nous avons montré que cet effet de régularisation pouvait biaiser les simulations lorsque les différentes échelles étaient simulées indépendamment. En effet, les échelles grossières ne pouvant être modifiées après leur simulation, elles peuvent propager une erreur de modélisation. Ceci implique que les structures simulées sont définies à des échelles différentes non indépendantes. Ainsi, nous pensons que la paramétrisation en ondelettes n'est pas adaptée aux simulations échelle-par-échelle pour les approches MPS.

Afin d'éviter ce problème, nous proposons d'utiliser une paramétrisation en pyramides gaussiennes. Cette paramétrisation présente les mêmes avantages que les ondelettes mais permet d'intégrer l'ensemble des échelles grossières. L'effet de régularisation inter-échelles aide alors à conditionner la simulation des variables à une échelle sans bloquer la correction des échelles supérieures, ce qui évite la propagation des erreurs et des incohérences avec les échelles fines. De plus, les pyramides gaussiennes permettent de mettre en place une méthode efficace de conditionnement des données de puits.

L'utilisation d'un DE de taille réduite et la diminution de la dimension du système, améliore significativement l'efficacité de l'algorithme, ce qui permet la simulation d'un grand nombre de réalisations. Ainsi une meilleure estimation des incertitudes peut être obtenue. De plus, des traitements intermédiaires peuvent être mis en place tout en n'affectant que raisonnablement les temps de calculs. Dans les exemples présentés ici, un traitement basé sur la connectivité des structures est appliqué afin d'améliorer la reproduction des caractéristiques de l'image d'entraînement.

Le conditionnement de l'exploration de l'image d'entraînement par les échelles grossières permet l'utilisation d'images de grandes tailles tout en maintenant un temps de calcul bas : un plus grand nombre de configurations est

disponible ce qui permet la génération de réalisations plus variées.

Bien que la méthode MS-DSgp donne des résultats très satisfaisants pour la simulation 2D de faciès binaires, il est nécessaire de généraliser l'approche au cas 3D et/ou multi-faciès. Ces extensions devraient bénéficier des mêmes avantages que l'approche 2D. En particulier, la re-paramétrisation multi-échelles dans le cas 3D réduira encore plus efficacement les dimensions intermédiaires. Cependant, les problématiques 3D multi-faciès étant différentes de celles rencontrées dans ce travail, il sera nécessaire d'adapter les traitements appliqués aux échelles grossières.

Chapitre 4

Approche multi-échelles de calage des données de production basée sur des méthodes d'optimisation par ensemble

« There is nothing wrong with change, if it is in the right direction. »

Winston Churchill, 1874-1965

4.1 Résumé et contributions

Ce chapitre est basé sur l'article [64] soumis à la revue *Computational Géosciences* en décembre 2013.

4.1.1 Résumé

Dans les processus standard de géomodélisation, l'intégration des données statiques et dynamiques est généralement effectuée de manière séquentielle : des modèles pétrophysiques sont initialement construits à partir des données géophysiques et de diagraphies et sont ensuite perturbés afin de caler les données historiques de production. Pour cette dernière étape, les méthodes d'optimisation par ensemble (voir Section 1.3.3) sont devenues populaires grâce à leur flexibilité, leur efficacité calculatoire et leur capacité à intégrer les données dynamiques en utilisant un large éventail de paramètres, ainsi qu'à quantifier les incertitudes *a posteriori*. Un gradient global est calculé statistiquement à partir d'un ensemble de réalisations et est utilisé pour mettre à jour chaque membre de l'ensemble. Cependant, pour des problèmes de grande dimension, la taille de l'ensemble est limitée puisqu'une simulation des écoulements doit être exécutée pour chaque réalisation. Pour cette raison, le gradient calculé peut être bruité et conduire à de mauvaises mises à jour. Ces dernières détériorent

les caractéristiques géologiques des modèles *a priori* et réduisent considérablement la variabilité de l'ensemble. L'objectif de ce chapitre est de conditionner un ensemble de réalisations, généré à partir d'une assimilation préalable de données et de connaissances, avec l'histoire de production en utilisant une méthode efficace d'optimisation basée ensemble, tout en minimisant la perturbation des paramètres ainsi que l'ajout de bruit. De cette manière, les modèles finaux restent compatibles avec toutes les données disponibles et par conséquent fournissent une bonne estimation des incertitudes. Ainsi, les modèles résultants peuvent être considérés comme fiables lors des prises de décisions relatives au développement du réservoir.

Les basses résolutions des champs de propriétés ont généralement un impact important sur l'écoulement des fluides dans le réservoir et de ce fait, sont bien caractérisées par les données de production [34, 13, 119]. De plus, les contrastes de propriétés définies par une modélisation de faciès ou bien une inversion sismique sont peu affectés par la perturbation de ces basses fréquences : les structures géologiques, telles que les chenaux, sont caractérisées par des transitions rapides des propriétés et les données sismiques ne portent pas d'information sur les basses fréquences (Section 1.2.1). En conséquent, notre méthode postule que le calage des données de production peut être grandement amélioré par la seule modification des basses résolutions, avec un impact limité sur le modèle *a priori* et sans introduire de bruit dans les hautes fréquences. Pour cette raison, une re-paramétrisation multi-échelles du problème est appropriée.

L'information portée par les données de production est spatialement hétérogène [119], les écoulements étant plus affectés par la variation des paramètres proches des puits où de vraies corrélations existent entre les données et les hautes fréquences. Loin des mesures de puits, les hautes fréquences ne sont généralement pas résolues par les données (sauf lorsqu'elles ont un impact important sur les écoulements, par exemple dans le cas où de fines barrières de perméabilité sont présentes). Ainsi, la paramétrisation doit être définie à la fois en fréquence et dans l'espace. Dans ce travail, nous décomposons les champs de propriétés en utilisant des ondelettes de seconde génération adaptée aux grilles de réservoirs (voir Section 2.3). De plus, les modèles de réservoir sont constitués de milliers voire de millions de paramètres alors que les données sont généralement éparées et incertaines, ce qui rend le problème mal posé (Section 2.2). La re-paramétrisation aide alors à stabiliser le processus d'inversion en réduisant le nombre de paramètres tout en préservant les structures principales des champs de propriétés grâce aux qualités de compression des ondelettes.

À partir de cette paramétrisation, nous proposons une approche d'inversion multi-échelles : différents sous-ensembles de coefficients d'ondelettes sont successivement mis à jour en fonction des données de production et en utilisant la méthode itérative RML (*Randomized Maximum Likelihood* [105]) basée ensemble de Levenberg-Marquardt (LM-enRML, Section 1.3.3) comme méthode d'optimisation. L'algorithme procède à l'optimisation d'un ensemble initial de paramètres basses résolutions. La paramétrisation est ensuite progressivement raffinée et une nouvelle optimisation est effectuée à chaque ajout de paramètres. Néanmoins, comme les réalisations *a priori* peuvent contenir des hétérogénéités fines incertaines qui peuvent avoir un impact important sur la réponse aux

écoulements, nous proposons d'atténuer ces hautes fréquences pendant l'optimisation des coefficients de grandes échelles afin d'éviter de biaiser leurs estimations. Les valeurs originales des coefficients atténués sont par la suite facilement rétablies lorsqu'ils sont inclus dans l'optimisation. La fonction objectif étant fortement améliorée pendant les premières optimisations des grandes échelles, les perturbations aux échelles les plus fines sont ensuite limitées.

Bien que la re-paramétrisation aide à identifier des paramètres sensibles, elle ne permet pas d'enlever les fausses corrélations produites par l'ensemble (voir Section 1.3.3). Dans les méthodes standard, ces effets sont atténués par la localisation des perturbations afin d'éviter la détérioration de l'ensemble (voir Section 1.3.3). Dans ce travail de thèse, nous proposons une nouvelle approche de localisation appliquée à la fois en fréquence et dans l'espace sur le gain de Kalman. Nous utilisons l'hypothèse selon laquelle les écoulements sont plus sensibles aux basses résolutions tandis que seules les hétérogénéités fines proches des puits ont un impact significatif sur la réponse simulée. Ainsi, il est possible d'adapter la localisation par rapport à l'ensemble des paramètres optimisés. Au début du processus, l'optimisation ne contenant que des paramètres de grandes échelles, les corrélations sont considérées valables et aucune régularisation n'est appliquée, ce qui permet une mise à jour globale des champs de propriétés. La localisation devient par la suite active lorsque des paramètres moins sensibles sont inclus dans l'optimisation.

L'algorithme est testé sur deux problèmes de calage historique. Le premier est un réservoir synthétique chenalisé associé à environ 6 années d'historique de production mesurée par 9 puits producteurs et 4 puits injecteurs. Le second problème est le cas comparatif *Brugge* comprenant 10 années de production. Les résultats de l'approche multi-échelles sont comparés à l'optimisation LM-enRML avec une paramétrisation de Dirac (c.-à-d sans re-paramétrisation, voir Section 2.3) et en utilisant une localisation basée sur la distance aux mesures (Section 1.3.3).

4.1.2 Contributions

La méthode multi-échelles proposée dans ce chapitre inclut les contributions suivantes :

1. nous utilisons des ondelettes lisses de seconde génération adaptées aux grilles réservoirs complexes à géométrie irrégulière afin de re-paramétriser les champs de propriétés (Section 2.3.2) ;
2. nous proposons d'optimiser de façon séquentielle un sous-ensemble dynamique de coefficients d'ondelettes, comportant initialement un nombre limité de paramètres de grandes échelles et progressivement raffiné afin de permettre des changements plus importants d'un nombre réduit de paramètres sensibles de résolutions plus fines (Section 4.3.2) ;
3. nous proposons un schéma d'atténuation inversible des effets des coefficients non inclus dans les premières optimisations afin d'éviter de biaiser l'estimation des paramètres optimisés. Ce schéma rétablit par la suite leur

influence dès qu'ils sont ajoutés dans le processus d'optimisation (Section 4.3.2 et Section 4.4). Ainsi, il est possible d'incorporer des connaissances *a priori* à toutes les échelles tout en suivant un processus d'inversion multi-échelles.

4. nous proposons une méthode de localisation espace-fréquence et un contrôle spécifique de l'optimisation LM-enRML qui dépendent de l'ensemble de paramètres optimisés (Section 4.4.1 et Section 4.4.2) ;
5. l'application à deux cas tests comportant des hétérogénéités complexes montre que notre approche modifie les paramètres du modèle initial de manière plus réaliste que les approches standard, tout en calant aussi bien les données de production (Section 4.5).

4.2 Paramétrisation

Nous utilisons ici les ondelettes de seconde génération adaptées aux grilles complexes de réservoirs décrites dans la Section 2.3.2. Dans les exemples présentés dans ce chapitre, nous utilisons des transformées en ondelettes 2D, soit pour des grilles 2D, soit pour des grilles 3D constituées de couches géologiquement indépendantes.

Dans ce qui suit, on dénote par $\mathbf{m} \in \mathbb{R}^n$ le vecteur des propriétés (paramètres de simulation, voir Section 2.2) de taille n défini dans l'espace de Dirac (base de fonctions constantes par cellules du modèle, voir Section 2.3), $\boldsymbol{\gamma} \in \mathbb{R}^n$ le vecteur de coefficients d'ondelettes (vecteur d'optimisation), \mathbf{W} la transformée directe en ondelettes, telle que $\boldsymbol{\gamma} = \mathbf{W}\mathbf{m}$ et \mathbf{W}^{-1} sa transformée inverse.

4.3 Optimisation d'un sous-ensemble de coefficients d'ondelettes

Dans cette section nous décrivons une méthode d'optimisation d'un sous-ensemble de coefficients d'ondelettes à partir de la méthode RML de Levenberg-Marquardt basée ensemble (LM-enRML). Pour plus d'informations sur l'optimisation et les problèmes inverses se référer au Chapitre 1.

4.3.1 Notations

Soit n et k des entiers tels que $1 \leq k \leq n$. On dénote par $Q_{k,n}$ l'ensemble de toutes les séquences de k entiers constituées d'éléments strictement croissants et choisis dans $\{1, \dots, n\}$. Soit $\mathbf{A} = (a_{ij})$ une matrice $(m \times n)$ constituée d'éléments $a_{ij} \in \mathbb{R}$. Soit r et c des entiers tels que $1 \leq r \leq m, 1 \leq c \leq n$, et $\alpha \in Q_{r,m}, \beta \in Q_{c,n}$ deux séquences d'indices i et j de la matrice \mathbf{A} , telles que $\alpha = (i_1, \dots, i_r), \beta = (j_1, \dots, j_c)$. Alors, la sous-matrice $\mathbf{A}[\alpha, \beta]$ contient des entrées (l, t) égales à a_{i_l, j_t} et possède r lignes et c colonnes. On dénote par « $*$ », la séquence croissante $Q_{n,n}$ constituée de tous les éléments de $\{1, \dots, n\}$, telle que $\mathbf{A}[*] = \mathbf{A}$. On dénote par $\mathbf{A}(\alpha, \beta)$ la sous-matrice de \mathbf{A} dont les lignes et colonnes sont respectivement les compléments de α et β .

Dans ce qui suit, n et n_r représentent le nombre total de coefficients d'ondelettes et le nombre de coefficients associés à la résolution r , tel que $n = \sum n_r$. On dénote par $\{i_\gamma\}_r$ l'ensemble des indices des coefficients d'ondelettes de résolution r et par $q(\{i_\gamma\}_r) \in Q_{n_r, n}$ sa séquence croissante correspondante. Pour finir, on dénote par $\{I_\gamma\}_{r, r'} = \{i_\gamma\}_r \cup \dots \{i_\gamma\}_{r'}$ l'ensemble des indices des coefficients de résolution r à r' et $q_{r, r'} \in Q_{n_{r'} + \dots + n_r, n}$ sa séquence croissante correspondante. Avec ces notations, l'univers (ensemble de tous les indices) est donné par $U = \{I_\gamma\}_{0, r_{max}}$, où $r = 0$ et $r = r_{max}$ correspondent respectivement à la résolution la plus grossière et la plus fine.

4.3.2 Méthode d'optimisation

RML Levenberg-Marquardt basé ensemble

La méthode RML Levenberg-Marquardt basée ensemble (LM-enRML) [33, 35] est une adaptation de l'algorithme de Levenberg-Marquardt (Section 1.3.2) pour laquelle un gradient global, calculé à partir d'un ensemble, est utilisé pour mettre à jour simultanément plusieurs réalisations [74]. L'ensemble initial est composé d'échantillons d'une distribution *a priori* qui peut être conditionnée par d'autres données (p. ex. sismique, données de puits) et est considéré être représentatif des incertitudes avant l'assimilation des données de production.

Chaque membre de l'ensemble $\mathbf{m} \in \mathbb{R}^n$ est transformé en coefficients d'ondelettes $\gamma \in \mathbb{R}^n$ en appliquant le schéma de *lifting* adapté aux grilles stratigraphiques, présenté dans la Section 2.3.2. L'optimisation est alors appliquée à un sous-vecteur $\gamma^* \in \mathbb{R}^{n^*}$ de coefficients sélectionnés par l'algorithme multi-échelles (cf. Section 4.4) tel que $n^* \leq n$, $\gamma^* = \gamma[\alpha]$ et $\gamma = \begin{bmatrix} \gamma^* \\ \gamma^c \end{bmatrix} = \mathbf{W}\mathbf{m}$, où $\alpha \in Q_{n^*, n}$ et $\gamma^c = \gamma(\alpha)$ correspondent respectivement aux indices des coefficients sélectionnés dans γ et au vecteur complémentaire de γ^* . Les paramètres de simulation peuvent ensuite être reconstruits en utilisant la transformée inverse : $\mathbf{m} = \mathbf{W}^{-1}\gamma$.

Afin d'échantillonner la PDF *a posteriori*, l'équation 1.39 (Section 1.3.3) est minimisée pour chaque membre de l'ensemble en perturbant les paramètres d'optimisation. On peut alors réécrire cette équation sous la forme :

$$S_j^{rml} \left(\begin{bmatrix} \gamma_j^* \\ \gamma_j^c \end{bmatrix} \right) = \frac{1}{2} (\gamma_j^* - \gamma_{pr, j}^*)^T \mathbf{C}_{\gamma^*}^{-1} (\gamma_j^* - \gamma_{pr, j}^*) + \frac{1}{2} (g(\mathbf{m}_j) - \mathbf{d}_j^0)^T \mathbf{C}_D^{-1} (g(\mathbf{m}_j) - \mathbf{d}_j^0) \quad (4.1)$$

où $\gamma_{pr, j}^* = \gamma_{pr, j}[\alpha]$, $\mathbf{C}_{\gamma^*}^{-1} = \mathbf{C}_\gamma[\alpha]^{-1}$ et $g(\cdot)$ sont respectivement les sous-vecteurs de la j^{eme} réalisation *a priori* $\gamma_{pr, j}$, l'inverse de la sous-matrice de covariance \mathbf{C}_γ et le modèle direct (c.-à-d le simulateur d'écoulements). Le modèle direct étant ici non linéaire, le problème ne peut pas se résoudre analytiquement et il est nécessaire d'utiliser un processus itératif.

Les méthodes de Gauss-Newton, et en particulier l'algorithme de Levenberg-Marquardt, sont généralement très efficaces lorsque la matrice des sensibilités (voir Section 1.3.2) est rapide à calculer. Le problème re-paramétrisé consiste

alors à trouver pour la l^{eme} itération, une perturbation $\delta\gamma_l^*$ obtenue par résolution du système :

$$[(\lambda_l + 1)\mathbf{C}_{\gamma^*}^{-1} + \mathbf{G}_l^T \mathbf{C}_D^{-1} \mathbf{G}_l] \delta\gamma_l^* = -[\mathbf{C}_{\gamma^*}^{-1}(\gamma_l^* - \gamma_{pr}^*) + \mathbf{G}_l^T \mathbf{C}_D^{-1}(g(\mathbf{m}_l) - \mathbf{d}^0)], \quad (4.2)$$

où le terme de droite et le premier terme de gauche de l'égalité correspondent respectivement au gradient et au Hessian de $S^{rml}(\gamma)$. Nous rappelons que le Hessian est modifié par un facteur λ afin de réduire l'influence des données dans les premières itérations et ainsi limiter l'amplitude des changements lorsque le modèle se trouve loin de l'optimum (voir « région de confiance » à la Section 1.3.2). La matrice $(n^d \times n^*)$ des sensibilités \mathbf{G}_l est alors calculée à partir de n^e réalisations en résolvant le système :

$$\Delta \mathbf{D}_l = \mathbf{G}_l \Delta \mathbf{\Gamma}_l \quad (4.3)$$

Nous rappelons que chaque colonne de la matrice $(n^d \times n^e)$, $\Delta \mathbf{D}_l$, et de la matrice $(n^* \times n^e)$, $\Delta \mathbf{\Gamma}_l$, se compose respectivement, pour chaque réalisation, de la déviation au vecteur moyen (de l'ensemble) des données simulées et des coefficients d'ondelettes (Section 1.3.3). Une décomposition en valeurs singulières est alors utilisée pour résoudre le système. Grâce à la réduction de paramètres $(n^* \leq n)$ et à la taille réduite de l'ensemble $(n^e \ll n)$, le temps de calcul de \mathbf{G} est relativement rapide.

Néanmoins, dans ce chapitre nous utilisons la forme approximative de la méthode de Levenberg-Marquardt [35], dans laquelle la matrice de covariance \mathbf{C}_{γ^*} utilisée dans le calcul du Hessian est approximée par une matrice définie positive, $\mathbf{P}_l = \Delta \mathbf{\Gamma}_l \Delta \mathbf{\Gamma}_l^T / (n^e - 1)$. En utilisant la formule d'inversion matricielle de Sherman-Woodbury-Morrison [71] et en substituant $\mathbf{C}_{\gamma^*}^{-1}$ par \mathbf{P}_l^{-1} dans le Hessian, l'équation 4.2 s'écrit :

$$\delta\gamma_l^* = -[(\lambda_l + 1)\mathbf{P}_l^{-1} + \mathbf{G}_l^T \mathbf{C}_D^{-1} \mathbf{G}_l]^{-1} \mathbf{C}_{\gamma^*}^{-1}(\gamma_l^* - \gamma_{pr}^*) - \mathbf{P}_l \mathbf{G}_l^T [(\lambda_l + 1)\mathbf{C}_D + \mathbf{G}_l \mathbf{P}_l \mathbf{G}_l^T]^{-1} (g(\mathbf{m}_l) - \mathbf{d}^0) \quad (4.4)$$

Enfin, le terme de pénalisation de l'éloignement au modèle *a priori* (premier terme de l'équation 4.4) est négligé afin que la mise à jour ne soit conditionnée que par les données :

$$\delta\gamma_l^* = -\mathbf{P}_l \mathbf{G}_l^T [(\lambda_l + 1)\mathbf{C}_D + \mathbf{G}_l \mathbf{P}_l \mathbf{G}_l^T]^{-1} (g(\mathbf{m}_l) - \mathbf{d}^0) \quad (4.5)$$

Il est montré dans [35] que cette forme donne de meilleurs calages, est numériquement plus stable et est plus efficace que la forme d'origine (équation 4.4). D'après notre expérience, cette approximation donne de bien meilleurs calages lorsque les réponses simulées de l'ensemble initial ne couvrent pas les données observées, et la qualité des réalisations produites ne souffre pas de la perte de la contrainte *a priori*.

Lorsque le même ensemble est utilisé pour calculer la matrice des sensibilités et la matrice \mathbf{P}_l , l'équation 4.5 devient similaire à la mise à jour de l'ensemble

smoother [52] dans laquelle il n'est pas nécessaire de calculer explicitement la matrice des sensibilités (voir détails des simplifications dans la Section 1.3.3) :

$$\delta\gamma_l^* = -\mathbf{C}_l^{\gamma^d}[(\lambda_l + 1)\mathbf{C}_D + \mathbf{C}^{dd}]^{-1}(g(\mathbf{m}_l) - \mathbf{d}^0) = -\mathbf{K}_l^e(g(\mathbf{m}_l) - \mathbf{d}^0) \quad (4.6)$$

où $\mathbf{C}_l^{\gamma^d} = \Delta\mathbf{\Gamma}_l\Delta\mathbf{D}_l^T/(n^e - 1)$ est la matrice de covariance croisée entre les paramètres optimisés et les réponses simulées, $\mathbf{C}^{dd}_l = \Delta\mathbf{D}_l\Delta\mathbf{D}_l^T/(n^e - 1)$ est la matrice de covariance des données simulées et \mathbf{K}_l^e est similaire à l'estimation du gain de Kalman.

Lissage initial des champs de propriétés

Bien que nous optimisions uniquement un sous-ensemble de n^* coefficients, tous les coefficients sont utilisés lors de la reconstruction des paramètres de simulation. Cependant, il est généralement utile d'atténuer l'impact des coefficients qui n'ont pas (encore) été sélectionnés (principalement des hautes fréquences) sur la réponse dynamique afin d'éviter tout biais d'estimation et d'être bloqué dans un minimum local. À l'emplacement des puits, les hautes résolutions peuvent avoir un impact important sur les écoulements et peuvent biaiser la caractérisation des basses fréquences si elles ne sont pas enlevées. Pour cette raison, les effets des coefficients d'ondelettes qui ne sont pas inclus dans l'optimisation sont atténués avant de débiter le processus d'inversion : chacune des réalisations γ est multipliée par un vecteur d'atténuation $\mathbf{s}_0 \in \mathbb{R}^n$ tel que :

$$\gamma^s = \mathbf{s}_0 \circ \gamma = \begin{bmatrix} \gamma^* \\ \gamma_s^c \end{bmatrix}, \quad (4.7)$$

où \circ correspond au produit d'Hadamard¹ (ou Schur) et γ_s^c représente une version atténuée de γ^c . L'application de la transformée en ondelettes inverse sur le vecteur γ^s reconstruit des champs de propriétés lissées comme le montre la figure 4.1. Généralement, \mathbf{s}_0 n'est pas constant mais ses éléments varient en fonction de la localisation et résolution des coefficients d'ondelettes associés. Ici, la valeur des éléments de \mathbf{s}_0 diminue linéairement avec la résolution des paramètres, de 1 (pas d'atténuation) à une valeur minimum (atténuation maximale), s_{min} ², telle que :

$$\forall r \in \{0, \dots, r_{max}\}, \forall j_r \in \{i_\gamma\}_r$$

$$\mathbf{s}_0(j_r) = \min \left(1, \frac{(1 - s_{min}) \times r + s_{min} \times r_0 - r_{max}}{(r_0 - r_{max})} \right) \quad (4.8)$$

où r est la résolution associée aux coefficients d'ondelettes, des plus grossiers ($r = 0$) aux plus fins ($r = r_{max}$), r_0 est la résolution de départ de l'optimisation (voir Section 4.4) et $\{i_\gamma\}_r$ est un ensemble d'indices, j_r , de coefficients d'ondelettes pour la résolution r .

Les valeurs de l'atténuation doivent être plus grandes que zéro afin de pouvoir par la suite calculer les matrices des sensibilités ou de covariances croi-

¹multiplication élément-par-élément de deux matrices de même dimensions

²égale à 0.1 dans les applications présentées dans ce chapitre

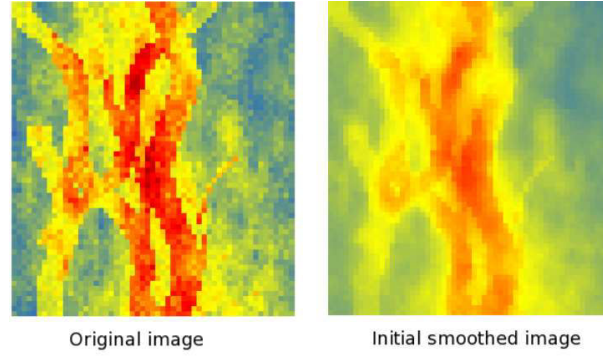


FIGURE 4.1 – Lissage initial des réalisations. Porosité initiale (gauche) et sa version lissée (droite) utilisée au début de l’optimisation après atténuation des hautes fréquences.

sées (équations 4.3 or 4.6) des coefficients d’ondelettes après raffinement de la paramétrisation et afin de pouvoir restaurer leurs valeurs initiales. Grâce aux propriétés compressives des bases d’ondelettes, toutes les structures importantes restent visibles après atténuation (figure 4.1). Les fréquences atténuées sont facilement rétablies pendant le processus multi-échelles, en multipliant les coefficients d’ondelettes par \mathbf{s}_0^{-1} , tel que $\gamma = \mathbf{s}_0^{-1} \circ \gamma^s$ (voir Section 4.4 et l’algorithme 2 de la Section 4.7), où \mathbf{s}_0^{-1} est la multiplication élément-par-élément de \mathbf{s}_0 .

En général, le facteur d’atténuation minimale, s_{min} , associé aux basses fréquences doit être petit afin d’empêcher de biaiser l’estimation des coefficients de plus grandes échelles. L’augmentation de s_{min} peut ralentir la diminution de la fonction objectif lors de l’optimisation des basses fréquences, tandis qu’une atténuation trop importante peut provoquer une augmentation de la fonction objectif lors du rétablissement des valeurs initiales des coefficients d’ondelettes pendant les dernières itérations (voir Section 4.4). On peut remarquer que tous les éléments de \mathbf{s}_0 correspondants aux coefficients γ^* sont égaux à 1 puisqu’ils sont inclus dans l’optimisation (équation 4.8).

4.4 L’approche multi-échelles

Nous proposons une approche multi-échelles basée sur l’optimisation LM-enRML et l’utilisation des ondelettes qui permet d’inclure des connaissances *a priori* détaillées. Un schéma général de la méthode est donné dans la figure 4.2. La boucle centrale correspond aux itérations à travers les différentes échelles (algorithme 1). Avant d’entrer dans la boucle (étape 0), quelques coefficients basses résolutions (de 0 à r_0), $\gamma_{r_0}^*$, sont sélectionnés, tels que, $\gamma_{r_0}^* = \gamma[q_{0,r_0}]$ et une atténuation réversible (Section 4.3.2) est appliquée aux coefficients restants, $\gamma_{r_0}^c = \gamma(q_{0,r_0})$. La sélection initiale devrait uniquement inclure des coefficients associés aux ondelettes, Ψ^* , qui couvrent au moins un élément de l’ensemble des positions des mesures (c.-à-d les positions des puits), P_d , telle que,

$$\forall \psi_{r,k} \in \Psi^*, \exists x_d \in P_d \mid (\psi_{r,k} * \delta_{x_d})(x) > \epsilon, \quad (4.9)$$

où $(*)$ est un opérateur de convolution, $\psi_{r,k}$ représente les ondelettes de résolution r et de position k , δ_{x_d} est une fonction de Dirac centrée sur la position x_d et ϵ est une valeur minimale d'impact de l'ondelette. De plus, la paramétrisation initiale devrait inclure un nombre suffisant de paramètres pour que les structures présentes dans les réalisations *a priori* restent discernables après l'atténuation des hautes fréquences (figure 4.1), et afin d'éviter de converger, pendant l'optimisation des larges échelles, vers un minimum qui n'est pas compatible avec les informations *a priori* contenues dans les échelles plus fines. Des tests qualitatifs de reconstruction après lissage ont été effectués dans ce travail afin de sélectionner l'ensemble initial de coefficients (résolution de démarrage, voir Section 4.5). L'objectif principal de cette initialisation est de (1) trouver un nombre réduit de paramètres qui aideront à améliorer le calage aux données et (2) préserver les paramètres les moins sensibles.

Dans la première étape (étape 1), les paramètres γ_r^* sont optimisés en utilisant l'algorithme LM-enRML (Section 4.3). Bien que l'ensemble ne soit généralement pas assez grand pour obtenir une estimation parfaite du gain de Kalman lorsque tous les paramètres sont inclus, il permet de d'obtenir une bonne approximation pour les coefficients de grandes échelles. Nous insistons sur le fait que la transformée en ondelettes ne permet pas d'enlever les fausses corrélations induites par l'utilisation d'un ensemble réduit, mais permet de faire ressortir des paramètres sensibles. Dans les applications géologiques ou météorologiques, seul un nombre limité de configurations basses résolutions peuvent être cohérentes avec la physique, les processus de dépôt et les connaissances *a priori*, ce qui rend possible une meilleure approximation de la distribution des paramètres avec un petit ensemble. La réponse simulée est également très sensible à ces coefficients étant donné qu'ils sont associés à des fonctions de base qui recouvrent de grandes parties du modèle. Néanmoins, pendant les premières itérations de résolutions, comme tous les paramètres nécessaires au calage des données ne sont pas inclus, il est recommandé de ne pas chercher à converger afin d'éviter de générer des modèles irréalistes. De plus, l'optimisation des grandes échelles est uniquement effectuée après l'atténuation des hautes fréquences afin d'éviter tout biais dans leurs estimations (figure 4.7).

Ensuite (étape 2), de nouveaux paramètres de résolution plus fine (résolution $r + 1$) sont ajoutés dans le processus avec leurs valeurs initiales, et le vecteur d'atténuation \mathbf{s}_{r+1} est mis à jour :

$$\gamma_{r+1}^* = \mathbf{s}_r [q_{0,r+1}]^{-1} \circ \gamma^s [q_{0,r+1}] \quad (4.10)$$

$$\forall j \in \{I_\gamma\}_{0,r+1}, \mathbf{s}_{r+1}(j) = 1 \quad (4.11)$$

L'influence des échelles plus fines est également partiellement réintroduite en multipliant les coefficients restant par un facteur ϵ_s^3 afin de restaurer progres-

³égale à 1.15 dans les applications présentées dans ce chapitre

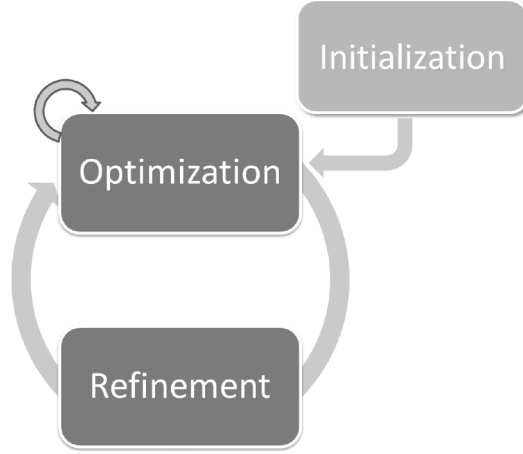


FIGURE 4.2 – Présentation schématique de la méthode multi-échelles. Étape 0 : initialisation de la paramétrisation et atténuation des hautes fréquences (lissage). Étape 1 : Optimisation des paramètres sélectionnés. Étape 2 : Raffinement de la paramétrisation et restauration des valeurs initiales des coefficients ajoutés.

sivement leurs valeurs d'origine (voir Algorithme 2 de la Section 4.7) :

$$\forall j \in \{I_\gamma\}_{0,r+1}^c, \quad \gamma_{s,r+1}^c(j) = \gamma_{s,r}^c(j) \times \epsilon_s \quad (4.12)$$

$$\mathbf{s}_{r+1}(j) = \mathbf{s}_r(j) \times \epsilon_s, \quad (4.13)$$

où $\{I_\gamma\}_{0,r+1}^c$ est l'ensemble complémentaire des indices des coefficients d'ondelettes (c.-à-d les indices de $\gamma_{s,r+1}^c$). Pour finir, une nouvelle optimisation est effectuée avec les nouveaux paramètres d'optimisation. Afin d'éviter le lancement de nouvelles simulations avec les réalisations restaurées, les résultats de la dernière évaluation de la fonction objectif avant raffinement sont utilisés lors de la première mise à jour de la nouvelle paramétrisation. Pour cette raison, les valeurs du vecteur d'atténuation doivent être plus grandes que zéro.

Ce processus est répété jusqu'à ce que l'échelle la plus fine soit atteinte et qu'un critère d'arrêt de l'optimisation soit rempli (Section 4.4.2). À cause de l'atténuation initiale, le processus ne devrait pas se terminer avant d'atteindre l'échelle la plus fine, même si le calage est suffisamment bon, puisque cela pourrait conduire à la perte d'information *a priori*.

4.4.1 Localisation multi-échelles

L'objectif de la localisation (Section 1.3.3), appliquée aux méthodes d'ensemble, est de minimiser les différences entre le gain de Kalman réel, \mathbf{K} , et son estimation par l'ensemble, \mathbf{K}^e , affectée par de fausses corrélations. Ceci peut être accompli en multipliant l'estimation du gain de Kalman par une matrice de projection \mathbf{A} , qui minimise la norme suivante :

$$f(\mathbf{A}) = \|\mathbf{A} \circ \mathbf{K}^e - \mathbf{K}\|_f^2 \quad (4.14)$$

où $\|\cdot\|_f^2$ est la norme de Frobenius⁴.

Nous définissons dans ce travail, deux matrices de projection, $\mathbf{A}_{\mathbf{gb}}$ et $\mathbf{A}_{\mathbf{wt}}$, appliquées de manière séquentielle dans les espaces de Dirac et des ondelettes afin de régulariser l'estimation du gain de Kalman. La première régularisation est basée sur des fonctions de distance, définies dans l'espace de Dirac, qui varient avec la résolution. La seconde régularisation est dérivée de cartes de pseudo-sensibilités définies par l'utilisateur dans la grille de simulation et qui sont ensuite interprétées dans le domaine espace-temps. Bien qu'elles soient toutes deux reliées à la position et à la résolution des coefficients, elles ont des objectifs différents. La localisation dans l'espace de Dirac limite les changements des paramètres à des zones prédéfinies, sans distinction d'échelles, alors que la localisation dans l'espace des ondelettes limite uniquement les changements de certaines fréquences dans certaines zones, mais permet des modifications globales des champs de propriétés (basses fréquences uniquement).

Localisation multi-échelles dans l'espace de Dirac

La localisation basée sur la distance a pour but de régulariser l'estimation du gain de Kalman en limitant l'effet des données aux zones dont elles sont sensibles. De manière équivalente, la localisation basée sur la distance peut être vue comme la décomposition d'un problème de grande taille, en plusieurs problèmes semi- indépendants moins affectés par les fausses corrélations. En effet, seul un nombre réduit de paramètres ayant des corrélations importantes avec les données est inclus dans chacun des sous-problèmes.

Nous proposons d'adapter la localisation en fonction des paramètres d'optimisation sélectionnés. Une fonction de localisation, $\mathbf{A}_{\mathbf{gb}}$, semblable à celle proposée par Furrer and Bengtsson [58] est utilisée pour régulariser le gain de Kalman dans l'espace (discret) de Dirac :

$$\mathbf{A}_{\mathbf{gb}}(h, r) = \frac{1}{1 + [f(0, r)^2 / f(h, r)^2] / n^e} \quad (4.15)$$

où $f(h, r)$ est une fonction de distance qui dépend à la fois de la distance de séparation (avec les données) h et du niveau de résolution r de la boucle multi-échelles. Dans les premières itérations du processus multi-échelles, la dimension du problème est grandement réduite puisque seuls quelques paramètres basses résolutions sont optimisés. La distribution de ces paramètres est généralement mieux échantillonnée par l'ensemble (discussion Section 4.4) et ces paramètres sont souvent très sensibles, ce qui permet une meilleure estimation du gain de Kalman. Nous rappelons également que le lissage initial empêche d'introduire un biais induit par les hautes fréquences. Dans cette configuration, il n'y a pas besoin d'appliquer de régularisation, ce qui permet d'effectuer des mises à jour globales des champs de propriétés sans introduire de bruit aux résolutions plus fines. Avec les raffinements successifs, le nombre de paramètres moins sensibles augmente et l'estimation du gain de Kalman devient plus affectée par les fausses corrélations. En conséquent, les portées des fonctions de distance diminuent et les changements sont appliqués dans des zones plus réduites (voir figure 4.3,

⁴ $\|\mathbf{A}\|_f^2 = \sqrt{\text{tr}(\mathbf{A}\mathbf{A}^T)}$

haut). À la résolution la plus fine, la localisation devrait être équivalente à celle proposée par Furrer et Bengtsson [58].

La localisation étant définie dans l'espace de Dirac et à cause de l'utilisation du produit d'Hadamard dans l'équation 4.14, il n'est pas possible de l'appliquer directement dans l'espace des ondelettes ($\mathbf{W}(\mathbf{B} \circ \mathbf{C}) \neq (\mathbf{W} \cdot \mathbf{B}) \circ (\mathbf{W} \cdot \mathbf{C})$, où \mathbf{B}, \mathbf{C} sont des matrices de dimensions appropriées). Au lieu de cela, le gain de Kalman est re-transformé dans l'espace de Dirac où la localisation est appliquée. Il est ensuite transformé une nouvelle fois vers l'espace des ondelettes afin d'effectuer la mise à jour de l'ensemble réduit des coefficients optimisés (voir Algorithme 3 de la Section 4.7) :

$$\begin{aligned} \mathbf{K}^{\text{tmp}} &= \mathbf{W} \left(\mathbf{A}_{\text{gb}} \circ \left(\mathbf{W}^{-1} \cdot \mathbf{K}_{n,n^d}^{\gamma^*} \right) \right) \\ \mathbf{K}^{\gamma^*} &= \mathbf{K}^{\text{tmp}} [q_{0,r}, *] \end{aligned} \quad (4.16)$$

où \mathbf{K}^{tmp} est une matrice $(n \times n^d)$ temporaire, $\mathbf{K}_{n,n^d}^{\gamma^*}$ est une matrice $(n \times n^d)$ définie par, $\mathbf{K}_{n,n^d}^{\gamma^*} [q_{0,r}, *] = \mathbf{K}^e$ et $\mathbf{K}_{n,n^d}^{\gamma^*} (q_{0,r}, *) = \mathbf{0}_{n-n^*,n^d}$, où \mathbf{K}^e est la matrice de Kalman $(n^* \times n^d)$ pour les coefficients d'ondelettes sélectionnés et $\mathbf{0}_{n-n^*,n^d}$ est une zéro-matrice $((n - n^*) \times n^d)$ associée aux coefficients non sélectionnés, et \mathbf{K}^{γ^*} est le $(n^* \times n^d)$ gain de Kalman gain localisé, exprimé dans l'espace des ondelettes. Ainsi, même la perturbation d'un paramètre global, tel que la moyenne, n'aura un impact qu'à l'intérieur des zones de localisation définies pour la résolution courante. Excepté dans les cas stationnaires stricts, il est important de limiter les perturbations lorsque la paramétrisation est raffinée. En effet, avec l'inclusion de paramètres de hautes fréquences, l'estimation loin des données des paramètres basses fréquences est biaisée par la configuration locale. En d'autres termes, nous évitons d'extrapoler incorrectement des caractéristiques locales. La transformée en ondelettes étant rapide et facilement parallélisable, la transformation et la reconstruction du gain de Kalman prend un temps raisonnable comparé aux temps de simulations des écoulements pour l'ensemble des réalisations.

Régularisation du gain de Kalman dans l'espace des ondelettes

Pour des problèmes de grandes dimensions, il n'est généralement pas possible, même en utilisant les dépendances spatiales, d'assurer une suppression totale des fausses corrélations, à moins que les fonctions de localisation n'incluent qu'un nombre extrêmement limité de paramètres, ce qui risque au contraire de supprimer des corrélations réelles et conduire à un mauvais calage ou à la création d'artéfacts. Généralement, les fonctions de localisation doivent être plus grandes qu'un motif de puits⁵[31]. Dans cette thèse, un autre type de régularisation, inspirée de l'analyse décrite dans [34], est utilisée en plus de la localisation décrite précédemment.

Nous proposons d'appliquer une régularisation du gain de Kalman dans l'espace des ondelettes. Généralement, les paramètres proches des puits ont un impact significatif sur la réponse aux écoulements, ce qui permet de caractériser

⁵un puits et ses voisins directs

avec plus de précisions les hétérogénéités fines, alors que loin des puits, seul le contenu basses-fréquences peut être déduit des données (sauf dans le cas de réservoirs très hétérogènes). En conséquence, des cartes de pseudo-sensibilités sont définies dans un premier temps dans la grille de simulation et ont pour objectif d'identifier les zones pour lesquelles les données de production sont très informatives. Des fonctions de distance à faible portée⁶ sont utilisées pour générer ses cartes (figure 4.3, en bas à gauche). Ensuite, cette information interprétative est décomposée dans l'espace des ondelettes : on considère que la validité des corrélations croisées entre les données et les paramètres est fonction de la sensibilité des zones couvertes par les ondelettes. Ainsi, des opérateurs semblables aux fonctions d'échelles (voir Section 2.3.1) sont utilisés afin d'intégrer les valeurs des cartes de sensibilités sur le support des ondelettes afin d'estimer l'importance de chacun des coefficients.

Par exemple, si la transformée de Haar est utilisée pour décomposer un champ de propriété, et si on dénote par $m_{j+1,k}^s$ la carte initiale de pseudo-sensibilités, alors, les valeurs $m_{j,k}^s$ associées aux coefficients d'ondelettes $\gamma_{j,k}$ sont calculées par :

$$m_{j,k}^s = m_{j+1,k}^s + m_{j+1,k+1}^s \quad (4.17)$$

Cette opération est répétée sur les valeurs $m_{j,k}^s$ restantes jusqu'à ce que la résolution la plus grossière soit atteinte (voir l'algorithme en cascade figure 2.5). Pour des propriétés 2D ou 3D, ce processus est appliqué successivement le long des directions de la grille, de telle sorte que les valeurs finales correspondent à des sommes sur des volumes ou des surfaces. Ainsi, l'importance des coefficients d'ondelettes augmente rapidement lorsque l'échelle diminue (voir figure 4.3, bas) et seules les hautes fréquences sont affectées par cette régularisation (les corrélations avec les grandes échelles sont supposées valables [34]). Les valeurs sont finalement nivelées à 1 afin de pouvoir être utilisées dans la matrice $(n^* \times n^d)$ de projection, $\mathbf{A}_{\mathbf{wt}}$, du gain de Kalman dans l'espace des ondelettes :

$$\mathbf{K}_{loc}^{\gamma^*} = \mathbf{A}_{\mathbf{wt}} \circ \mathbf{K}^{\gamma^*} \quad (4.18)$$

où $\mathbf{K}_{loc}^{\gamma^*}$ est le gain de Kalman après régularisation dans l'espace de Dirac et dans l'espace des ondelettes. Nous précisons que même si le terme pseudo-sensibilités est utilisé pour définir les cartes dans l'espace de Dirac, la matrice $\mathbf{A}_{\mathbf{wt}}$ n'est pas une approximation de la sensibilité dans l'espace des ondelettes, mais est uniquement utilisée pour identifier les corrélations valables entre les paramètres et les données.

Lorsque des ondelettes d'un ordre plus élevé sont utilisées, la somme de l'équation 4.17 doit inclure plus d'éléments (en fonction du support de des ondelettes) avec différents poids qui peuvent être déduits des fonctions d'échelle associées. Étant donné que les fonctions d'échelle, dans le cadre des ondelettes de seconde génération, changent avec l'échantillonnage du signal et ne sont pas représentées explicitement dans le schéma de *lifting*, il peut être difficile de définir précisément les fonctions d'intégration correspondantes. Dans ce cas, des approximations peuvent être faites. Dans ce travail, une ondelette basée sur une

⁶les valeurs de pseudo sensibilités diminuent rapidement de 1 à 0

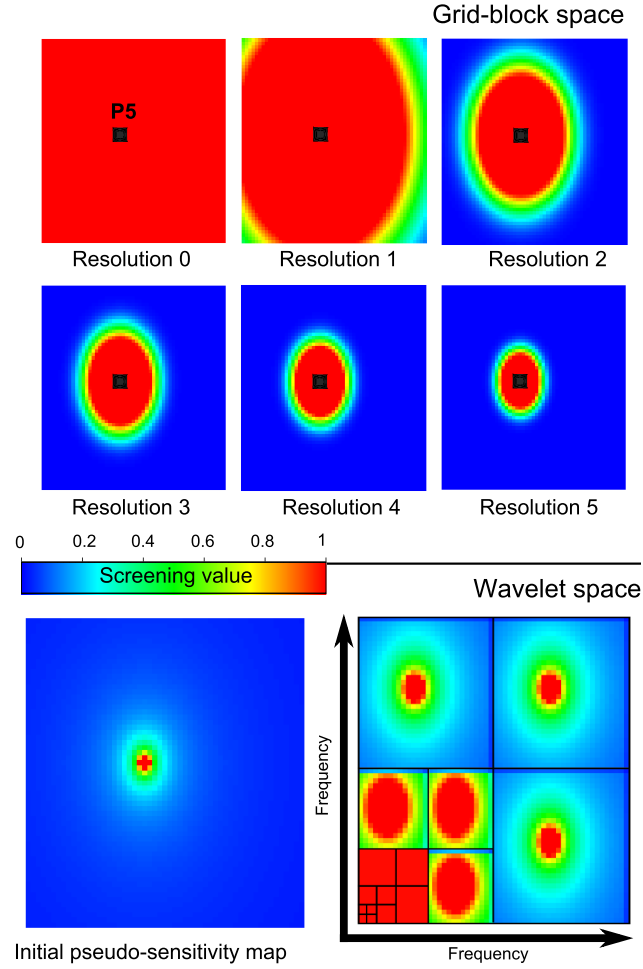


FIGURE 4.3 – Localisation multi-échelles. Localisation multi-échelles pour les données de pression au puits P5. Appliquée au gain de Kalman dans l'espace de Dirac à différentes résolutions du processus multi-échelles (haut). Appliquée au gain de Kalman dans l'espace des ondelettes (bas). Voir algorithme 3 de la Section 4.7.

interpolation quadratique de la moyenne est utilisée pour décomposer les propriétés, mais le modèle de Haar est utilisé pour calculer la matrice de projection (équation 4.17), les deux ondelettes étant assez semblables (Section 2.3.2).

4.4.2 Optimisation et contrôle du paramètre de Levenberg-Marquardt

Dans notre méthode multi-échelles, une optimisation est effectuée après chaque raffinement de la paramétrisation. Pendant les premières résolutions (itérations), la paramétrisation ne comprend que des paramètres de grandes échelles et le problème est simplifié grâce au lissage initial des réalisations, ce qui permet une convergence rapide. Cependant, le problème étant généralement sous-paramétré à ce stade, il n'est pas souhaitable de converger car cela pourrait conduire à des solutions irréalistes. De plus, les composants initialement lissés, réintroduits lors de l'étape de raffinement, peuvent être incompatibles avec l'état optimal actuel. Pour cette raison, une seule itération LM-enRML

est effectuée à chaque niveau de résolution, jusqu'à ce que les résolutions les plus fines soient atteintes. Cependant, si la valeur de la fonction objectif est fortement affectée par la réintroduction des composants lissés et que la majorité des réalisations ne sont pas améliorées, une mise à jour LM-enRPML supplémentaire est effectuée. Habituellement, cela se produit au cours des dernières étapes de raffinement durant lesquelles un grand nombre de paramètres sont ajoutés dans l'optimisation. Pour cette raison, nous itérons deux fois aux résolutions intermédiaires lorsque le nombre de paramètres est de deux (ou plus) ordres de grandeur supérieur à la taille de l'ensemble.

L'initialisation et le contrôle du paramètre de Levenberg-Marquardt sont basés sur la valeur initiale et l'évolution de l'erreur moyenne des prédictions de l'ensemble par rapport aux données, S_{av} (voir l'équation 4.19 ci-dessous) et à l'écart type (dans l'ensemble) de cette erreur, σ_S , comme décrit dans [35]. Lorsque S_{av} et σ_S diminuent, λ est divisé par un coefficient ω , mais est limité à une valeur minimum. Lorsque S_{av} diminue mais que σ_S augmente, ce qui reflète une amélioration hétérogène des réalisations de l'ensemble, la mise à jour est acceptée mais la valeur de λ reste inchangée. Lorsque S_{av} augmente, la mise à jour est rejetée et λ est multiplié par ω .

Cependant, après un raffinement et une restauration des valeurs initiales des coefficients d'ondelettes, une augmentation raisonnable de la fonction objectif est tolérée : si la valeur de S_{av} après l'étape de raffinement est supérieure à sa valeur avant l'optimisation des paramètres à la résolution précédente, la mise à jour est rejetée et seulement la moitié des composants lissés sont réintroduits lors de la mise à jour suivante. Lorsque la mise à jour est acceptée mais que la fonction objectif n'est pas améliorée, la valeur de λ est inchangée (voir algorithme 1 de la Section 4.7). Un rejet à un stade précoce peut refléter une atténuation initiale des coefficients trop importante ou l'utilisation d'une information *a priori* incohérente avec les données. Dans ce cas, une incorporation de paramètres hautes fréquences dans des zones sensibles (autours des puits par exemple) dès le début de l'optimisation peut aider à stabiliser le comportement de la fonction objectif ; cela peut néanmoins affecter la préservation du modèle *a priori*.

Dans les approches standard, ω est habituellement constant. Dans notre approche multi-échelles, la valeur de ω est réglée en fonction des résolutions optimisées. Durant les premières itérations d'échelles, la valeur de ω est maintenue élevée afin de garder un λ relativement grand et ainsi limiter l'amplitude des mises à jour des grandes échelles. La valeur de ω est ensuite augmentée lors des étapes de raffinement et atteint son maximum à la résolution la plus fine (voir algorithme 1 de la Section 4.7). Dans les exemples décrits par la suite, la valeur initiale de ω est égale à 2 et augmente linéairement jusqu'à 10 et la valeur de λ est limitée à 0.01. Lorsque le même contrôle du paramètre de Levenberg-Marquardt est utilisé dans l'approche standard, l'optimisation peine à réduire la fonction objectif pendant les premières itérations, sans que cela ait un impact sur la qualité des modèles finaux, alors que la mise à jour des paramètres sensibles de grandes échelles permet de réduire significativement la fonction objectif même lorsque la valeur de λ est élevée.

Plusieurs critères d'arrêt sont utilisés pour la dernière optimisation (à

l'échelle la plus fine) : nombre maximum d'itérations, augmentation maximale de λ à une itération ou bien lorsque la valeur de la fonction objectif tombe dans l'incertitude des données ($S_{av} < n^d$). Pour les optimisations intermédiaires, étant donné que l'on ne cherche pas à converger, les critères d'arrêts sont uniquement basés sur le nombre maximum d'itérations et l'amélioration du calage des données.

4.4.3 Validation de la méthode

Les méthodes de re-paramétrisation et les approches multi-échelles sont généralement efficaces lorsque le problème inverse est sous conditionné. Afin d'illustrer les avantages de notre approche, deux cas tests de taille médium sont utilisés dans ce chapitre. Étant donné la taille des problèmes, la validation des résultats n'est pas basée sur l'analyse de la distribution conditionnelle échantillonnée par une méthode de Monte Carlo (voir Section 1.3.3). Les résultats obtenus par l'approche multi-échelles (MS-LM-enRML) sont comparés avec la forme approximée de l'algorithme de Levenberg-Marquardt (GB-LM-enRML) utilisant une paramétrisation standard (cellules de la grille) [35] et une localisation basée sur la distance. Étant donné que le même algorithme est utilisé pour optimiser les coefficients d'ondelettes, les bénéfices de l'approche multi-échelles peuvent être évalués en comparant les résultats des deux algorithmes.

La qualité du calage des données, la prédiction de l'ensemble et la qualité des réalisations générées sont utilisées comme critères de comparaison. L'erreur de calage, donnée par la fonction objectif RML (second terme de l'équation 4.1), est évaluée afin d'assurer que les réalisations sont des échantillons possibles de la distribution *a posteriori* :

$$S_{av}(\mathbf{m}_j) = \frac{1}{n^e} \sum_{j=1}^{n^e} (g(\mathbf{m}_j) - \mathbf{d}^0_j)^T \mathbf{C}_D^{-1} (g(\mathbf{m}_j) - \mathbf{d}^0_j) \quad (4.19)$$

où \mathbf{d}^0_j est une réalisation échantillonnée sur la distribution des données. Il est généralement impossible d'atteindre $S_{av} \ll n^d$, même lorsque les incertitudes sont surestimées car les échantillons \mathbf{d}^0_j ne sont pas physiquement plausibles à cause de l'hypothèse d'indépendance des données (\mathbf{C}_D est diagonale). Des mesures quantitatives de qualité des réalisations générées sont difficiles à mettre en place étant donné que la notion de modèle géologique cohérent est généralement abstraite. Cependant, l'analyse spatiale des réalisations donne un aperçu intéressant sur le comportement des processus d'inversion. Pour cette raison, un examen qualitatif des réalisations ainsi qu'une analyse variographique sont proposés dans ce chapitre.

L'ensemble *a priori* pouvant incorporer d'autres sources d'informations, telles que de la sismique et/ou des interprétations géologiques, il est généralement important de minimiser les perturbations des modèles et d'éviter les effets des fausses corrélations. Des cartes de perturbations absolues moyennes des réalisations de l'ensemble sont générées afin d'analyser la qualité des changements. Pour un calage similaire aux données, une perturbation la plus faible possible est préférable car cela dénote une meilleure préservation de l'informa-

tion *a priori*.

Lorsque le modèle réel est connu, il est possible de le comparer avec les éléments de l'ensemble. L'analyse directe cellule-par-cellule n'est généralement pas adaptée étant donné la nature mal posée du problème inverse et l'utilisation de mesures locales. Alternativement, il peut être possible d'utiliser des algorithmes de reconnaissance de formes [132]. Au lieu de cela, nous effectuons une reconnaissance qualitative des structures conditionnant les écoulements.

Les temps d'exécution pour les deux méthodes sont similaires et sont principalement contrôlés par les simulations d'écoulements de l'ensemble. Les transformées en ondelettes impliquées dans la méthode multi-échelles nécessitent des calculs supplémentaires, mais ceci est contrebalancé par la réduction du nombre de paramètres qui permet d'améliorer l'efficacité des mises à jour de Levenberg-Marquardt. Les implémentations de ces deux méthodes sont basées sur le même code source de manière à pouvoir comparer objectivement les deux algorithmes. Dans le cas standard, la transformée en ondelettes est juste remplacée par l'identité et le processus d'optimisation démarre à la résolution la plus fine.

4.5 Tests numériques

Les deux méthodes sont appliquées à deux problèmes de calage historique. Le premier est un réservoir synthétique 2D composé de 13 puits et contenant environ 8 années d'historique de production. Les 6 premières années sont utilisées pour contraindre l'inversion et les 2 années restantes sont utilisées pour tester la prédiction des ensembles. Le second problème est le cas bien connu de Brugge, composé de 9 couches géologiquement indépendantes, incluant 10 années d'historique provenant de 30 puits [108]. Le processus de calage historique prend environ 3 heures dans le premier cas et environ une journée dans le second pour les deux méthodes testées, les simulations étant tournées simultanément sur 25 processeurs.

Des ondelettes basées sur l'interpolation de la moyenne (Section 2.3.2) sont utilisées comme base de fonctions pour toutes les applications MS-LM-enRML. Une localisation basée sur la distance et un contrôle de λ sont appliqués pour la méthode GB-LM-enRML comme décrit dans [35], alors qu'une localisation et un contrôle multi-échelles de λ (Sections 4.4.1 et 4.4.2) sont utilisés dans l'algorithme MS-LM-enRML. Dans les deux cas, une fonction de covariance sphérique, $f(\cdot)$, est utilisée dans la fonction de localisation (équation 4.15).

4.5.1 Réservoir 2D

Le premier test est un réservoir synthétique 2D composé de 3355 cellules actives contenant les champs de porosité et de perméabilité générés à partir d'une méthode géostatistique basée objet et d'une méthode de simulation gaussienne séquentielle (SGS). Les propriétés utilisées pour générer le jeu de données synthétique sont présentées dans la figure 4.4. Une relation non-linéaire stochastique est utilisée pour générer le champ de perméabilité à partir de la porosité. Les distributions finales des propriétés sont bi-modales comme pourrait l'être un système fluvial comportant des dépôts de plaines d'inondations.

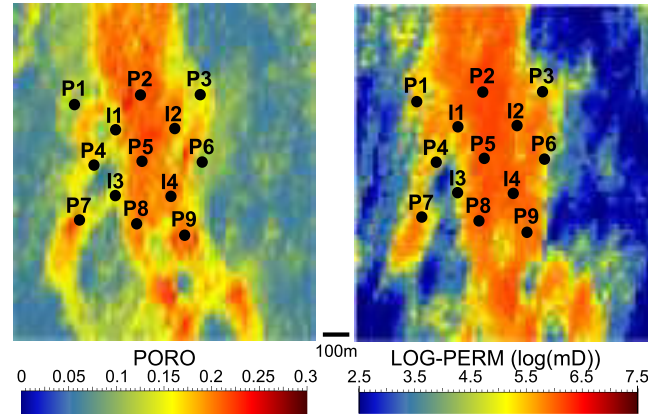


FIGURE 4.4 – Porosité et perméabilité de référence pour le test 1.

Un total de 9 puits producteurs et 4 puits injecteurs, arrangés selon un schéma à cinq points (*five-spot pattern*), sont utilisés pour générer 8 années d'historique de production. Les contraintes primaires des puits sont les productions d'huile et les injections d'eau. Une limite de pression de fond de puits (BHP) de 120 bars est également utilisée pour chaque puits. La proportion de production d'eau (*water cut* ou WCT), la pression de fond de puits (BHP) et le ratio huile/gaz (GOR) sont utilisés pour contraindre l'inversion des champs de propriétés, avec un nombre total d'environ 700 points de données. L'écart type du bruit associé aux mesures est de 2 bars pour le BHP, 2.5 % pour le WCT et $3 \text{ sm}^3/\text{sm}^3$ pour le GOR.

Pour les deux méthodes testées, un maximum de 20 mises à jour des paramètres d'optimisation sont permises. Un ensemble de 50 réalisations, initialement conditionnées par des données de diagraphie bruitées et générées par une combinaison de méthodes objet et de SGS, est utilisé pour calculer le gradient. L'orientation des chenaux est connu *a priori* avec une incertitude de 30° . Le paramètre de Levenberg-Marquardt λ est initialisé avec une valeur égale à 1000 dans les deux cas.

En raison de la taille relativement modeste du problème (en comparaison avec un cas réel), des ensembles de 50 réalisations sont utilisés lors de l'optimisation. Les résultats dépendent fortement de l'ensemble initial, surtout lorsque ce dernier est petit ; afin d'éviter un biais dans l'analyse des résultats, nous utilisons 5 ensembles indépendants pour assimiler les données de production avec les deux méthodes. Néanmoins, les réalisations contenues dans les 5 ensembles différents ont été générées à partir du même processus de simulation.

La figure 4.5 montre l'évolution de la fonction objectif S_{av} pour chacune des méthodes et pour les 5 ensembles. Dans les premières itérations, la diminution du décalage entre les données simulées et les mesures est équivalente pour les deux méthodes, même si l'algorithme MS-LM-enRML utilise une paramétrisation réduite. L'algorithme MS-LM-enRML n'itère qu'une ou deux fois par résolution avant d'atteindre l'échelle la plus fine (1 itération des résolutions 1 à 4 et 2 itérations à la résolution 5). La réintroduction des composants lissés au cours de l'étape de raffinement finale peut provoquer une légère augmentation de S_{av} (ensembles 1, 3 et 4) qui est rapidement corrigée à l'itération suivante.

Après avoir atteint la résolution la plus fine, tous les groupes utilisant l'approche multi-échelles s'arrêtent parce que l'erreur de calage tombe sous l'incertitude des données ($S_{av} < n^d$), tandis que la plupart des cas standards convergent légèrement au-dessus de cette limite.

La figure 4.6 montre les calages initiaux et finaux des données ainsi que les prédictions pour les puits P1, P2, P6 and P9. On peut noter qu'il est impératif d'utiliser un schéma itératif d'optimisation dans ce cas précis : les calages initiaux ne couvrant pas les données, on peut supposer une forte non-linéarité du problème. Les résultats de l'ensemble numéro 4 sont utilisés pour comparer les deux méthodes car les calages finaux sont semblables. Un bon calage des données est obtenu dans les deux cas. Les prédictions sont également très similaires mais échouent parfois à prédire correctement l'évolution de la production. Les performances des prédictions peuvent légèrement varier en fonction de la méthode, mais l'analyse approfondie des résultats des 5 ensembles ne permet pas de faire ressortir des différences majeures entre les deux méthodes.

La comparaison des moyennes des logarithmes de perméabilité avec le cas réel est présentée dans la figure 4.7. L'échelle des couleurs est affichée afin de faire ressortir les structures principales conditionnant les écoulements et les contrastes de propriété. Dans les deux cas, MS- et GB-LM-enRML, les formes des structures sont caractérisées par les données de production (p. ex. le bras perméable sur la droite). Cependant, nous pouvons observer que les contrastes de propriété sont définis de façon plus précise avec l'approche multi-échelles.

La figure 4.8 montre un groupe de réalisations choisies parmi les 5 ensembles avant et après optimisation et la figure 4.9 donne les perturbations moyennes des modèles *a priori* de porosité et de perméabilité, calculées à partir des réalisations des 5 ensembles. Les réalisations générées par l'approche multi-échelles préservent clairement mieux les structures spatiales des modèles *a priori* et sont beaucoup moins bruitées, tout en calant aussi bien les données que la méthode standard. De plus, on peut remarquer que les perturbations importantes sont principalement localisées aux endroits clés pour le calage des observations (p. ex., pour caractériser la barrière aux écoulements localisée entre les puits P7 et P8).

La figure 4.10 montre les variogrammes, correspondant aux réalisations de la figure 4.8, d'orientation Nord 0° (orientation des chenaux) avec une tolérance de 20° et une bande passante égale à 20% de l'extension du modèle. Les variogrammes *a posteriori* associés à la méthode MS-LM-enRML restent semblables aux variogrammes *a priori*, alors que les variogrammes de l'approche GB-LM-enRML sont très affectés par l'assimilation des données, ce qui reflète une meilleure stabilisation du processus d'inversion par l'approche multi-échelles. La préservation des propriétés est uniquement possible lorsque le modèle *a priori* contient de l'information (l'existence d'un système fluvial et son orientation, des données de puits, ...) qui est cohérente avec la réalité. Lorsque l'ensemble initial est mal conditionné, des changements plus importants sont nécessaires, mais l'approche multi-échelles aide tout de même à stabiliser l'inversion (voir le cas Brugge, Section 4.5.2).

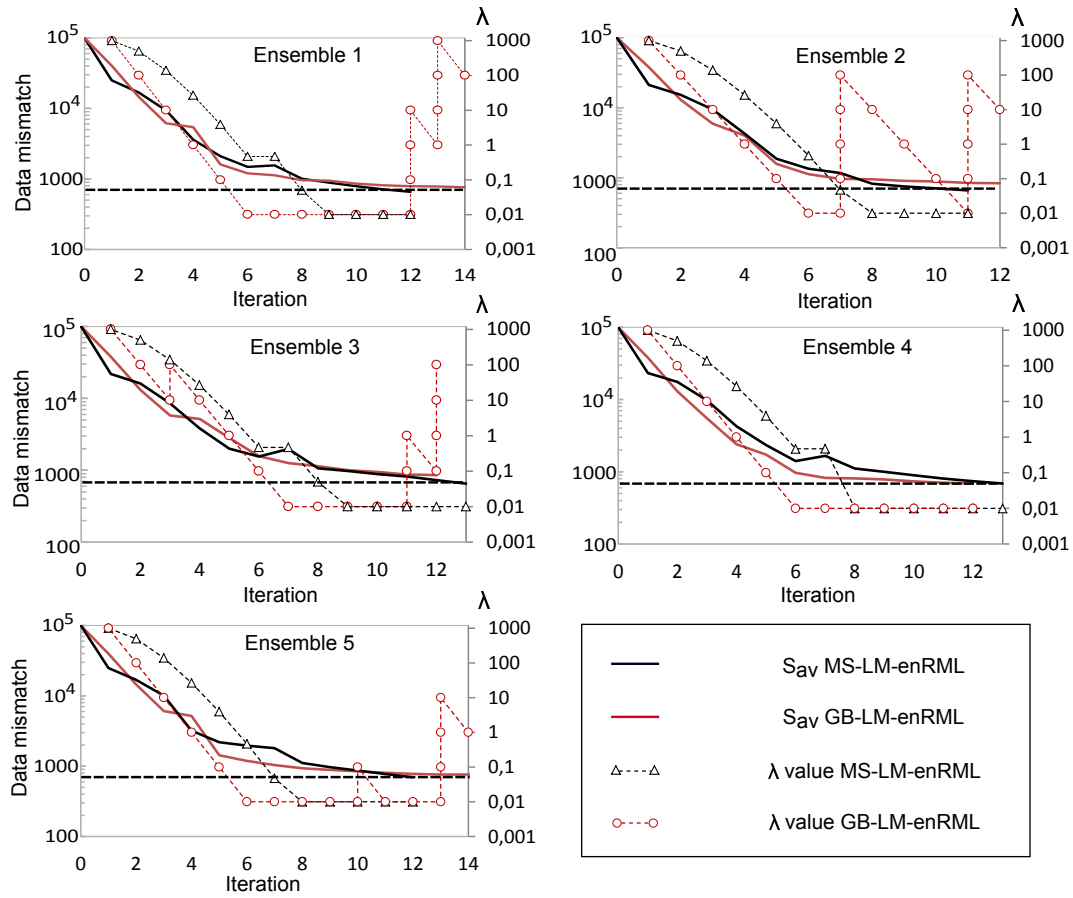


FIGURE 4.5 – Test 1 : évolution des fonctions objectifs. L'évolution des moyennes des erreurs de calage (équation 4.19) et des valeurs λ au cours des itérations pour les 5 ensembles. Le nombre de données est d'environ 700 (ligne horizontale en pointillé)

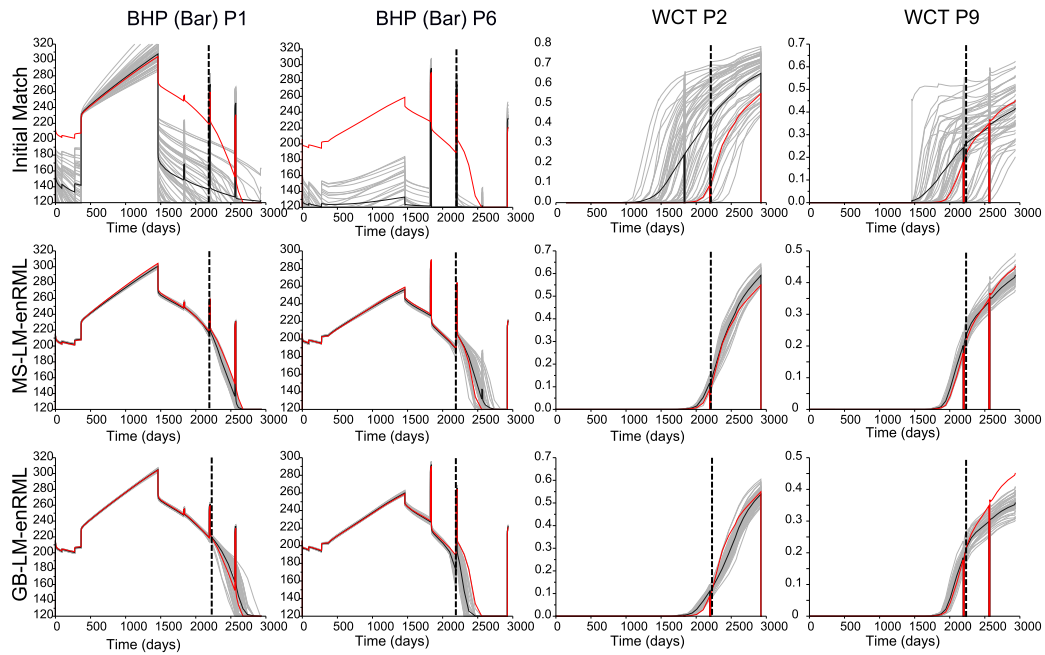


FIGURE 4.6 – Test 1 : calages et prédictions. Les calages (jours 0 à 2200) et les prédictions (jours 2200 à 3000) des productions d’eau (WCT) et de la pression pour 4 puits. La courbe rouge représente les données mesurées. Les courbes grises montrent les prédictions de l’ensemble. La courbe noire représente la moyenne de l’ensemble.

4.5.2 Le cas Brugge

Le cas Brugge est un modèle 3D ($139 \times 48 \times 9$) conçu par l’Organisation néerlandaise pour la recherche scientifique appliquée (TNO). Le modèle géologique initial a été construit dans une grille fine composée de 20 millions de cellules. Une première mise à l’échelle (*Upscaling*) des propriétés pétrophysiques a été réalisée pour générer les données de production à partir de 30 puits verticaux (20 producteurs et 10 injecteurs) à partir d’un modèle de simulation contenant un demi-million de cellules. Le réservoir ne contient que de l’huile et de l’eau. Ces données sont ensuite utilisées pour tester différents algorithmes d’optimisation sur un modèle encore plus grossier contenant 44 550 cellules actives [108]. Les contraintes primaires des puits sont les taux de production de fluides (FPR) fixés à 2000 barils/jour pour les producteurs et les taux d’injection d’eau (WIR) fixés à 4000 barils/jour pour les injecteurs. Une limite de pression de fond de puits est fixée à 725 psi pour tous les puits. Cette contrainte est active pour le puits P9 pendant les 10 années.

Les données utilisées dans ce travail sont : la pression du fond de puits pour les 30 puits, les taux de production d’eau (WCT) et les taux d’injection/production de fluides correspondant à un historique de 10 ans pour un nombre total d’environ 1 330 points de données. Les écarts types utilisés pour modéliser le bruit contenu dans les différents types de données sont : 30 psi pour la BHP, 3,5 % pour le WCT, 2 b/j pour le FPR et WIR si elles sont des contraintes actives, et 20 barils/jour pour le FPR du puits P9.

Les propriétés mises à jour sont : la porosité, les perméabilités latérales et

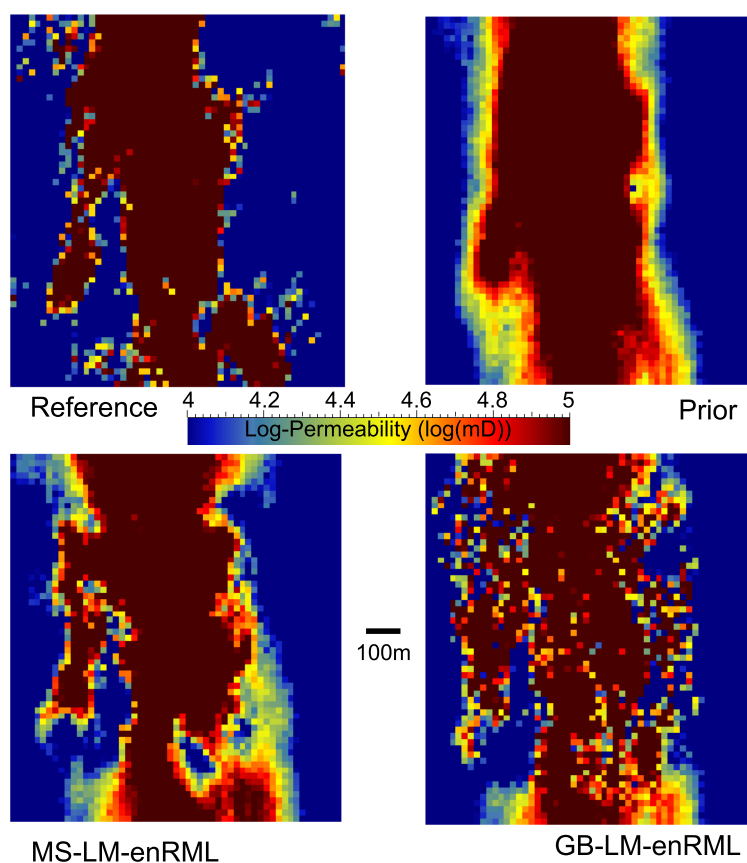


FIGURE 4.7 – Test 1 : comparaisons des propriétés moyennes. Comparaison des perméabilités moyennes des ensembles après optimisation avec la méthode MS-LM-enRML (en bas à gauche), la méthode GB-LM-enRML (en bas à droite), le modèle réel (en haut à gauche) et la moyenne initiale (en haut à droite).

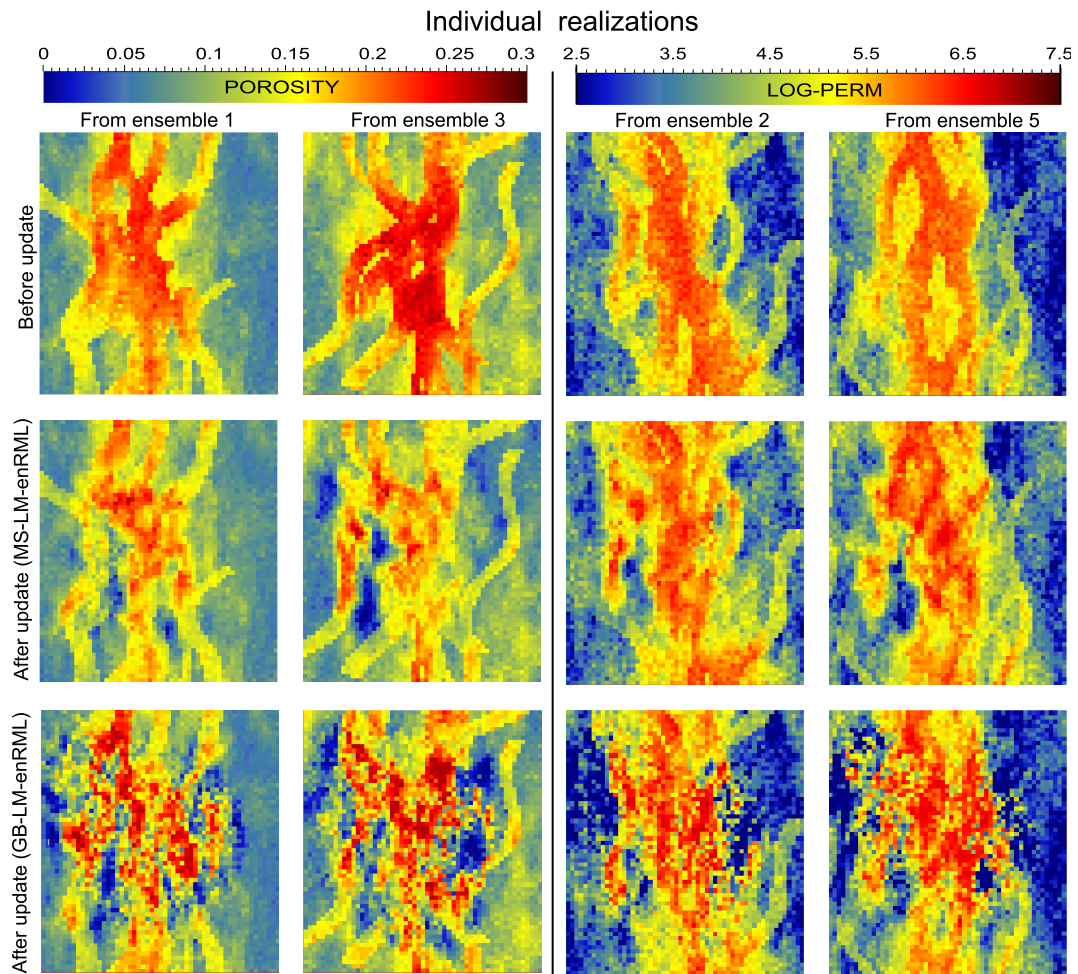


FIGURE 4.8 – Test 1 : réalisations avant et après optimisation. Réalisations individuelles de porosité/log-perméabilité avant et après l'intégration des données de production

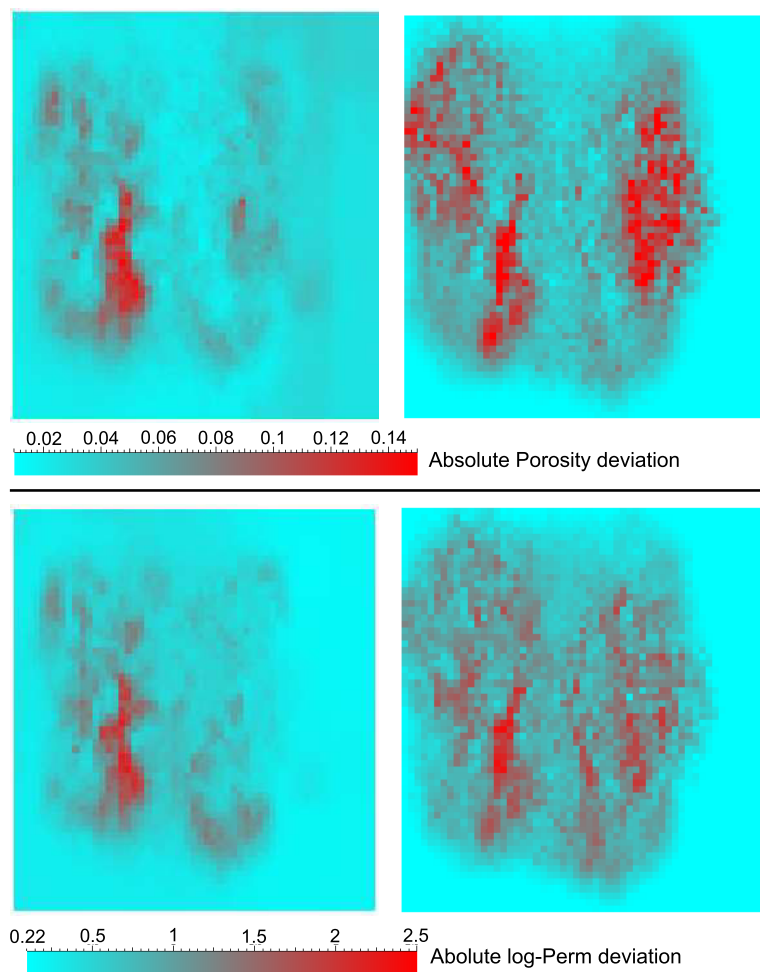


FIGURE 4.9 – Test 1 : différences moyennes avec le modèle *a priori* . Perturbations moyennes du modèle *a priori* calculées à partir de toutes les réalisations des 5 ensembles. MS-LM-enRML (gauche) et GB-LM-enRML (droite).

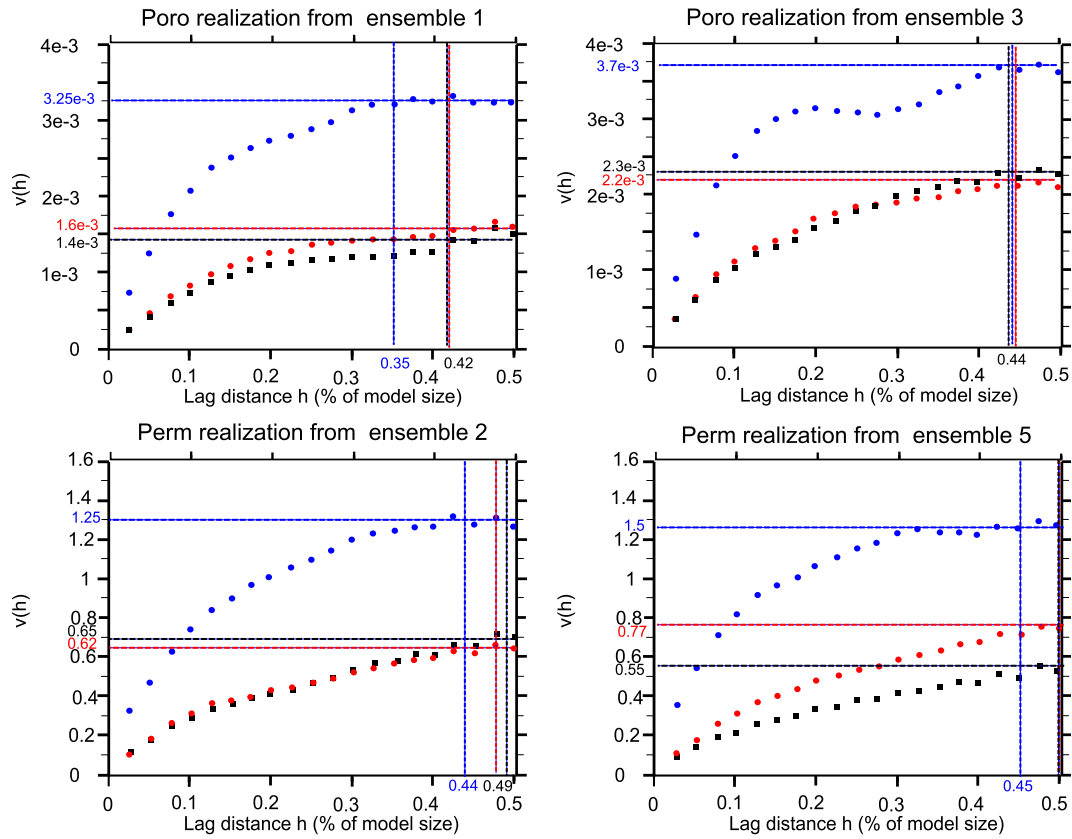


FIGURE 4.10 – Test 1 : analyses variographiques des réalisations. Analyses variographiques (North 0) des réalisations présentées dans la figure 4.8. Les carrés noirs représentent les variogrammes des réalisations initiales. Les points rouges et bleus représentent les variogrammes des réalisations *a posteriori* générées à partir des méthodes MS-LM-enRML et GB-LM-enRML. Les lignes verticales correspondent aux portées et seuils interprétés

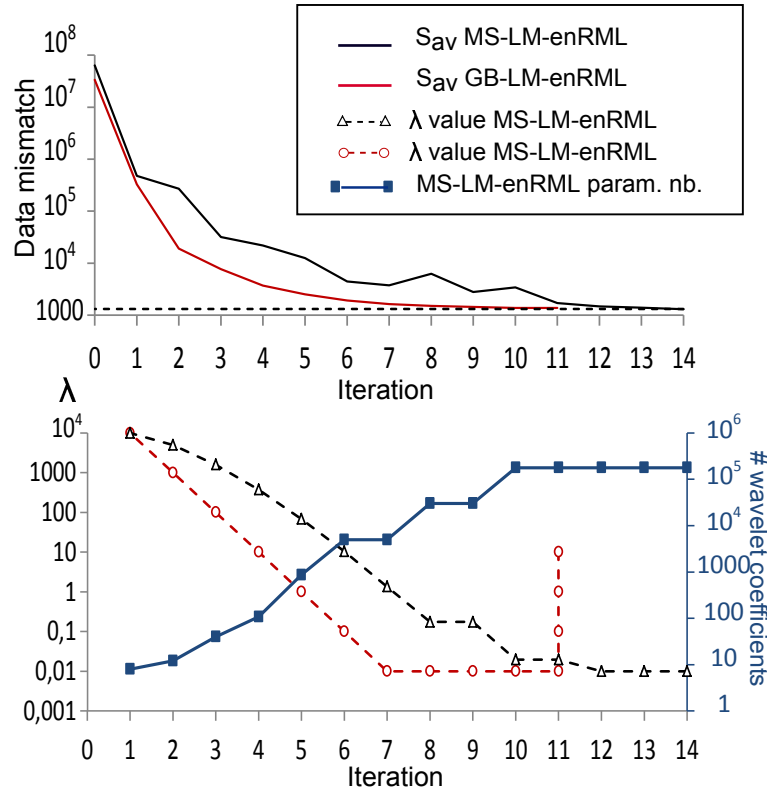


FIGURE 4.11 – Test 2 : évolution de la fonction objectif. Haut : évolution de la fonction objectif pour les méthodes MS-LM-enRML (courbe noire) et GB-LM-enRML (courbe rouge). Bas : évolution de la valeur de λ pour chaque méthode et du nombre de paramètres inclus dans l’optimisation MS-LM-enRML (courbe bleue) au cours des itérations.

verticales, le net-to-gross (NTG), les trois points des courbes de perméabilité relatives du modèle de Corey et le contact initial huile-eau. Pour l’approche MS-LM-enRML, les coefficients d’ondelettes des propriétés transformées sont inclus dans l’optimisation, tandis que les valeurs des cellules sont utilisées dans l’approche standard GB-LM-enRML. Il est à noter que les propriétés moyennes de différentes couches ne sont pas incluses dans l’optimisation GB-LM-enRML comme c’est le cas dans [34, 35] car cela peut être considéré comme une paramétrisation multi-échelles partielle. Cependant, cela devrait permettre d’améliorer les résultats de l’optimisation GB-LM-enRML. Le nombre total de paramètres du modèle est d’environ 178 200. Un ensemble (fourni par TNO [108]) de 104 réalisations, générées à partir de simulations objets et de SGS, est utilisé pour les deux méthodes. Les réalisations initiales ne sont pas conditionnées par des données et sont donc très disparates, ce qui rend la préservation des modèles plus difficile. Les points initiaux et finaux des courbes de Corey ainsi que le contact initial huile-eau sont générés à partir des distributions uniformes définies dans [32].

La figure 4.11 montre l’évolution de la fonction objectif (haut) et de la valeur de λ (bottom) au cours des itérations. La méthode GB-LM-enRML donne de meilleurs résultats au cours des premières itérations et s’arrête pour une er-

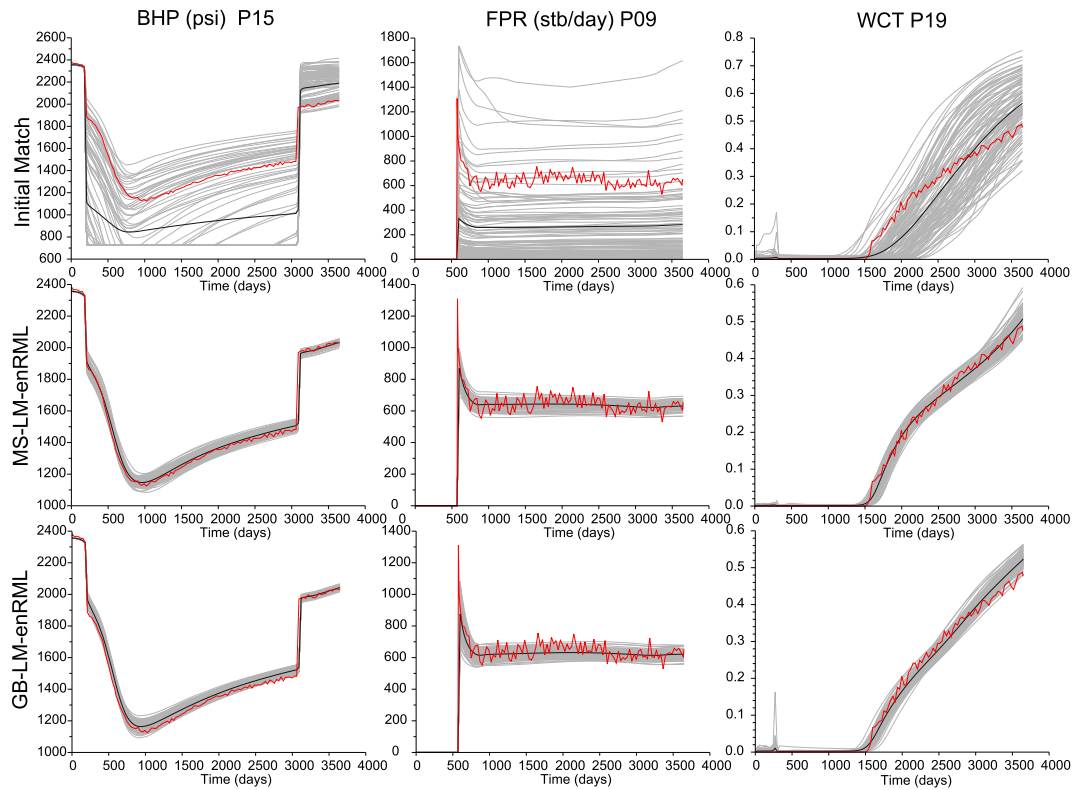


FIGURE 4.12 – Test 2 : calage des données de production. Calages des données de pression, de production de fluides et de production d’eau (WCT) pour trois puits. La courbe rouge correspond aux données observées. Les courbes noires et grises correspondent à la moyenne des prédictions et aux prédictions individuelles de l’ensemble

reur moyenne de calage de 1 389 à cause de plusieurs mises à jour infructueuses, alors que l’approche MS-LM-enRML s’arrête au bout de 14 itérations car l’erreur moyenne de calage tombe sous l’incertitude des données. Dans les deux cas, la valeur initiale de λ est de 10 000 et sa valeur minimum est de 0.01. Durant les premières itérations, seul un nombre réduit de paramètres est inclus dans l’optimisation MS-LM-enRML, ce qui explique la différence d’efficacité entre les deux méthodes, même si l’on obtient tout de même une amélioration significative de la fonction objectif. Nous pouvons observer qu’après les deux premiers raffinements, la réintroduction des composants initialement lissés provoque une légère augmentation de la fonction objectif. La dissemblance des réalisations *a priori* (générées à partir de processus différents sans conditionnement par des données secondaires) peut expliquer le besoin de gradients individuels (pour des sous-groupes de réalisations) après la réintroduction des composants atténués. Lorsque la paramétrisation entière est utilisée depuis le début de l’optimisation, cet effet est atténué par la projection linéaire de l’ensemble de paramètres lors des premières mises à jour, ce qui permet d’obtenir des modèles plus semblables, mais provoque également une diminution de la variabilité de l’ensemble et un endommagement des réalisations *a priori*.

La figure 4.12 montre les calages initiaux et finaux des données pour trois puits producteurs. Les deux méthodes testées donnent de très bons calages.

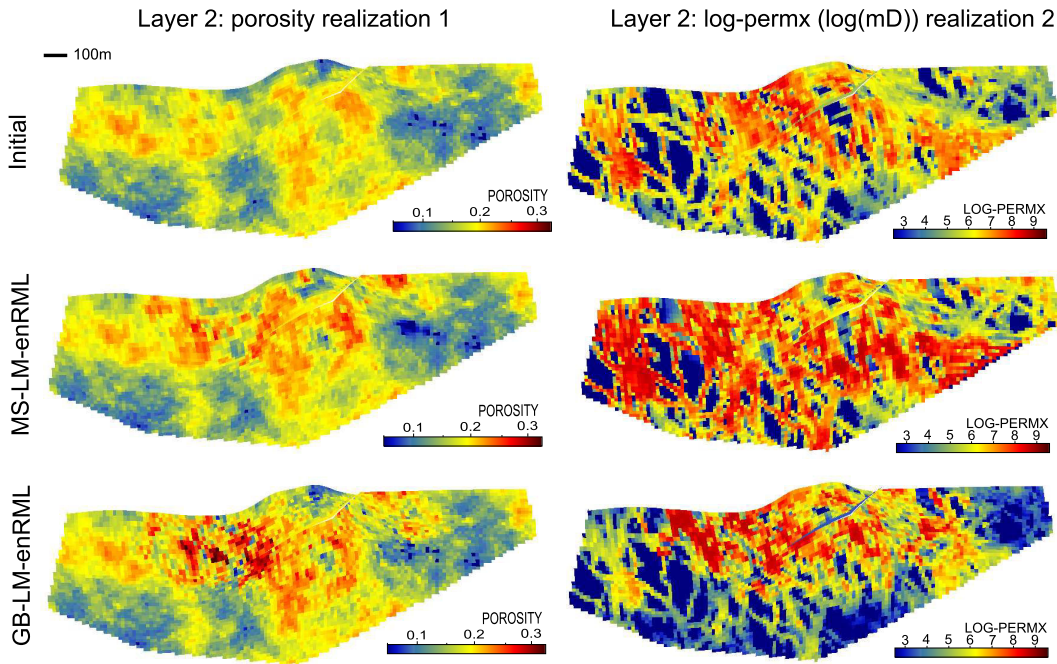


FIGURE 4.13 – Test 2 : réalisations avant et après optimisation. Une réalisation de porosité (gauche) et une réalisation de log-perméabilité (droite) de l'ensemble initial (colonne du haut) et les réalisations finales générées par l'approche MS-LM-enRML (colonne du milieu) et GB-LM-enRML (colonne du bas). Seules les propriétés de la seconde couche sont représentées

La figure 4.13 montre des réalisations de perméabilité et de porosité initiales (générées par des processus différents) et finales, mises à jour par les méthodes MS-LM-enRML et GB-LM-enRML, tandis que les cartes de différences moyennes entre les modèles *a priori* et *a posteriori* sont présentées dans la figure 4.14. Seule la couche la plus sensible (seconde couche) est représentée. Pour un calage équivalent, l'approche multi-échelles effectue des mises à jour plus réalistes et évite l'apparition de valeurs extrêmes ou bien l'ajout de bruit et minimise l'amplitude des changements. Nous pouvons également observer que la méthode MS-LM-enRML permet une modification globale des champs de propriétés lors des premières itérations. Ainsi, les changements effectués semblent moins artificiels et la continuité initiale des propriétés est préservée (perméabilité des chenaux de la réalisation 2 dans la figure 4.13).

4.6 Discussions et conclusions

La méthode adaptative multi-échelles basée sur des ondelettes de seconde génération se révèle être utile pour stabiliser l'inversion et éviter les effets néfastes liés à l'utilisation de méthodes d'optimisation par ensemble. Dans les exemples présentés ici, une meilleure préservation du modèle *a priori* est obtenue avec l'approche multi-échelles par rapport à la méthode GB-LM-enRML pour un calage équivalent des données de production. Cependant, cette approche multi-échelles repose sur l'hypothèse que l'écoulement est plus affecté par les basses résolutions que par les hautes. Dans des réservoirs où les écou-

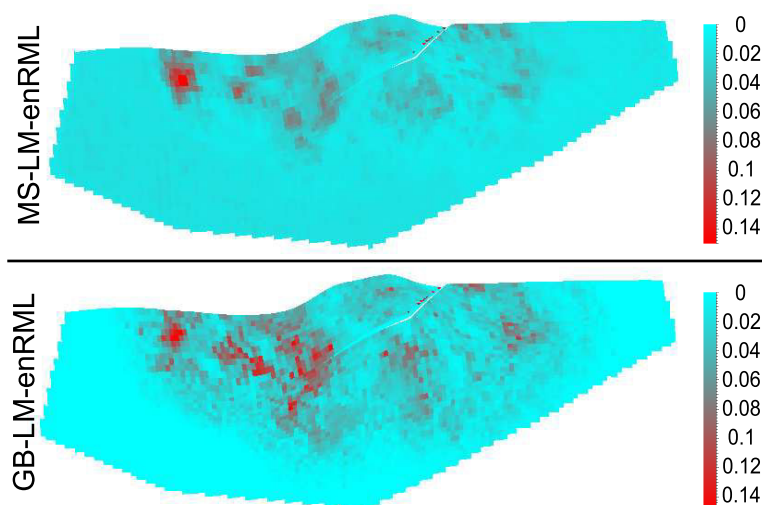


FIGURE 4.14 – Test 2 : différences avec le modèle *a priori* . différences moyennes avec le modèles *a priori* (différences absolues cellule-par-cellule). MS-LM-enRML (haut) et GB-LM-enRML (bas).

lements sont principalement contrôlés par des structures fines (p. ex. des réservoirs fracturés ou stratifiés), l’approche multi-échelles pourrait conduire à une estimation erronée des propriétés du réservoir. Dans ce cas, il peut être utile d’inclure des fréquences élevées dès le début du processus d’inversion (p. ex., pour les réservoirs stratifiés, des transformations 2D indépendantes peuvent être utilisés pour chaque couche).

L’approche multi-échelles et le lissage des réalisations ont tendance à rendre le problème plus linéaire grâce à l’atténuation des effets des hautes fréquences et la réduction du nombre de paramètres. Une réduction importante de la fonction objectif est obtenue par l’optimisation de paramètres de grandes échelles au cours des premières itérations, ce qui limite par la suite les perturbations des échelles plus fines. Cependant, la restauration des valeurs initiales des coefficients atténués peut légèrement affecter la convergence de l’algorithme comparé à l’approche standard. Dans l’implémentation proposée ici, le lissage initial dépend uniquement de la résolution des paramètres. Cependant, il est possible d’imaginer une procédure d’atténuation plus avancée, basée sur la localisation des paramètres et/ou leurs incertitudes *a priori*. En effet, l’atténuation des paramètres devrait être appliquée en premier lieu dans les zones sensibles (p. ex. à proximité des puits) et être moins importante dans les zones ayant peu d’impact sur la réponse simulée. De plus, lorsque les incertitudes du modèle ont déjà été réduites par l’intégration d’autres sources de données, l’atténuation devrait être moins importante puisque les réalisations sont plus semblables. L’impact sur les écoulements des structures identifiées avec certitude devrait également pris en compte lors des premières mises à jour du modèle⁷.

La régularisation du gain de Kalman par localisation des perturbations peut s’avérer être une tâche très difficile et une mauvaise localisation peut conduire

⁷en considérant que les différentes informations soient cohérentes

à des mises à jour irréalistes ou à un calage insuffisant aux données. Dans l'approche multi-échelles, les effets de la localisation pendant les premières itérations sont limités car seuls des paramètres sensibles sont inclus dans l'optimisation. Cependant, lorsque le modèle initial contient de l'information à des échelles fines, il est nécessaire de raffiner la paramétrisation jusqu'à la résolution la plus fine. La régularisation devient alors essentielle. Pour cette raison, une localisation multi-échelles des coefficients d'ondelettes est proposée dans ce travail de thèse afin de stabiliser l'inversion et de mieux préserver l'information *a priori* lorsqu'un grand nombre de paramètres est inclus dans l'optimisation. Néanmoins, de même que pour les localisations standard, la définition de cette régularisation n'est pas toujours triviale. Des analyses plus poussées des matrices de corrélation croisée [31] ou utilisant des techniques analogues au *bootstrap sampling* pourraient aider à trouver des fonctions de localisation plus appropriées.

Lorsque peu d'information est utilisée pour générer l'ensemble *a priori*, les réintroductions successives des composants lissés peuvent affecter l'efficacité de l'optimisation, en particulier lorsque les réalisations sont très différentes les unes des autres. Dans ce cas, l'utilisation d'un gradient global pourrait être inappropriée. Une optimisation utilisant l'ensemble des paramètres pourrait alors être plus adaptée. Il est également possible de re-paramétriser le problème et négliger certaines interprétations *a priori* afin que l'ensemble des réalisations convergent vers le même optimum. Néanmoins, dans ces deux cas, l'ensemble final fournira une mauvaise estimation des incertitudes.

4.7 Algorithmes

| Symbole | Description |
|------------------------|--|
| n^e | Nombre de réalisations |
| n^d | Nombre de données |
| l_r^{max} | Nombre maximum d'itérations pour sortir de l'optimisation à la résolution r |
| n^* | Nombre de paramètres en cours d'optimisation |
| r | Résolution maximale des paramètres en cours d'optimisation |
| r_{max} | Résolution maximale des paramètres |
| l_r | Indice d'itération pour la résolution r |
| I_r | Groupe d'indices de coefficients d'ondelettes inclus dans l'optimisation au cours de la résolution r |
| ω_r | Facteur de contrôle de λ pour la résolution r |
| $SFact$ | Facteur de contrôle de des-atténuation |
| \mathbf{M}^{pr} | Ensemble des (valeurs des) variables <i>a priori</i> du modèle |
| \mathbf{M}^{up} | Ensemble des (valeurs des) variables <i>a posteriori</i> du modèle |
| $\mathbf{\Gamma}^{pr}$ | Ensemble <i>a priori</i> des coefficients d'ondelettes |
| $\mathbf{\Gamma}^s$ | Ensemble des coefficients d'ondelettes atténués |
| $\mathbf{\Gamma}^*$ | Ensemble des coefficients d'ondelettes utilisés dans l'optimisation |
| \mathbf{d}_{obs} | Données observées |
| \mathbf{D}^0 | Ensemble des données perturbées |
| $g(\cdot)$ | Modèle direct (simulateur d'écoulements) |
| \mathbf{D}_l | Ensemble des données simulées ($\mathbf{D}_l = g(W^{-1} \cdot \mathbf{\Gamma}_l^s)$) |
| $Floor(\cdot)$ | Retourne le plus grand entier le plus proche |
| $zeros(n_r, n_c)$ | Génère une matrice nulle de dimensions $(n_r \times n_c)$ |
| \mathbf{A}_r^{gb} | Matrice de localisation dans l'espace de Dirac pour la résolution r |
| \mathbf{A}_r^{wt} | Matrice de localisation dans l'espace des ondelettes |
| \mathbf{C}_D | Matrice de covariance du bruit des données |
| Sav_r | Meilleure valeur de la fonction objectif pour la résolution r |
| Sav_l | Moyenne de la fonction objectif à l'itération l |
| λ | Paramètre de régularisation de Levenberg-Marquardt |
| \mathbf{W} | Transformée directe en ondelettes |
| \mathbf{W}^{-1} | Transformée inverse en ondelettes |
| \mathbf{S} | Vecteur d'atténuation des coefficients d'ondelettes |

Tableau 4.2 – Paramètres de l'algorithme. Les scalaires et les fonctions, les vecteurs, et les matrices sont respectivement écrites en minuscules, gras, et en majuscules.

ALGORITHME 4.1 - L'algorithme multi-échelles. Pour plus de simplicité, seules les variables spatiales sont considérées.

Input: $\mathbf{M}^{\text{Pr}}, d_{\text{obs}}, r_0, \mathbf{A}^{\text{gb}}, \mathbf{A}^{\text{wt}}, \omega_0, \epsilon_s$
Output: \mathbf{M}^{up}
begin initialization
 $r \leftarrow r_0$;
 $l = 1$; $SFact = 1$;
 $\mathbf{D}^0 \leftarrow d_{\text{obs}} + \epsilon$, with $\epsilon \sim \mathcal{N}(0, \mathbf{C}_D)$;
 $\mathbf{\Gamma}_0 \leftarrow W \cdot \mathbf{M}^{\text{Pr}}$;
 $I_r \leftarrow \{I_\gamma\}_{0, r_0}$; // Initial set of coef. indices
 $n^* \leftarrow |I_r|$; // Cardinality of the set
 $\mathbf{S}_r \leftarrow \mathbf{S}_0$; // Initial smoothing
 $\mathbf{\Gamma}^s \leftarrow \mathbf{S}_r \circ \mathbf{\Gamma}_0$;
 $\mathbf{\Gamma}^* \leftarrow \mathbf{\Gamma}^s [q(I_r), *]$; // Sub-matrix of $\mathbf{\Gamma}^s$
 $\mathbf{D}_0 \leftarrow g(\mathbf{W}^{-1} \cdot \mathbf{\Gamma}^s)$; $Sav_0 \leftarrow (\mathbf{D}_0 - \mathbf{D}^0)^T \mathbf{C}_D (\mathbf{D}_0 - \mathbf{D}^0)$;
 $Sav_{r_0} \leftarrow Sav_0$;
 $\lambda \leftarrow 10^{\text{Floor}(\log_{10}(\frac{Sav_0}{2N^d}))}$; $\omega_{r_0} \leftarrow \omega_0$;
end
while $r \leq N^r$ **do**
 while $l \leq l_r^{\text{max}}$ **do**
 if $l = 1$ **and** $r \neq r_0$ **then** Algorithm 2
 Refinement($\mathbf{\Gamma}_{r-1}^s, I_{r-1}, \mathbf{S}_{r-1}, SFact, \epsilon_s$);
 end
 $\mathbf{K} \leftarrow \text{KalmanGain}(\mathbf{\Gamma}_l^*, \mathbf{D}_l, \lambda_l)$; // see [35]
 $\mathbf{K} \leftarrow \text{MSLocalization}(\mathbf{K}, \mathbf{A}_r^{\text{gb}}, \mathbf{A}^{\text{wt}}, I_r)$; // Algorithm 3
 $\delta \mathbf{\Gamma}^* \leftarrow -\mathbf{K}(\mathbf{D}_l - \mathbf{D}^0)$;
 $\mathbf{\Gamma}_l^* \leftarrow \mathbf{\Gamma}_{l-1}^* + \delta \mathbf{\Gamma}^*$; $\mathbf{D}_l \leftarrow g(\mathbf{\Gamma}^s)$;
 $Sav_l = (\mathbf{D}_l - \mathbf{D}^0)^T \mathbf{C}_D (\mathbf{D}_l - \mathbf{D}^0)$;
 if $Sav_l \leq Sav_{l-1}$ **then**
 $SFact = 1$; $l = l + 1$;
 if StopCriteria(r, Sav_l) **then**
 Break;
 else
 $\lambda_{l+1} \leftarrow \lambda_l / \omega_r$;
 end
 else
 if $l = 1$ **and** $r - 2 \leq r_0$ **then** post refinement special case
 if $Sav_l \leq Sav_{r-2}$ **then**
 $l = l + 1$; $SFact = 1$;
 else
 $\lambda_l \leftarrow \lambda_l \times \omega_r$;
 $\mathbf{D}_l \leftarrow \mathbf{D}_{l-1}$;
 $SFact = SFact / 2$;
 end
 else
 $\lambda_l \leftarrow \lambda_l \times \omega_r$;
 $\mathbf{\Gamma}_l^* \leftarrow \mathbf{\Gamma}_{l-1}^*$;
 $\mathbf{D}_l \leftarrow \mathbf{D}_{l-1}$;
 end
 end
 end
 $Sav_r \leftarrow Sav_{l-1}$; $Sav_0 \leftarrow Sav_{l-1}$;
 $\mathbf{D}_0 \leftarrow \mathbf{D}_{l-1}$;
 $\omega_{r+1} \leftarrow \omega_r + \frac{10 - \omega_{r_0}}{r_{\text{max}} - r_0}$;
 $r = r + 1$; $l = 1$;
end

 ALGORITHME 4.2 - Fonction de raffinement.

Input: $\Gamma_{r-1}^s, I_{r-1}, \mathbf{S}_{r-1}, SFact, \epsilon_s$
Output: $\Gamma_r^s, I_r, \mathbf{S}_r$
 $I_r \leftarrow I_{r-1} \cup \{i_\gamma\}_r; n^* \leftarrow |I_r|;$
 $\mathbf{S}_{tmp}[n^*, n^e];$ // Temporary matrix
for $e = 1$ **to** $e = n^e$ **do**
 for $p = 1$ **to** $p = n$ **do** partial re-introduction for all frequencies
 $\mathbf{S}_{tmp}(p, e) = \min(\mathbf{S}_{r-1}(p, e)^{-1}, \epsilon_s) / SFact;$
 end
 forall the $p \in I_r$ **do**
 $\mathbf{S}_{tmp}(p, e) = \mathbf{S}_{r-1}(p, e)^{-1} / SFact;$
 end
end
 $\Gamma_r^s \leftarrow \mathbf{S}_{tmp} \circ \Gamma_{r-1}^s;$
 $\mathbf{S}_r \leftarrow \mathbf{S}_{tmp} \circ \mathbf{S}_{r-1};$

 ALGORITHME 4.3 - MSLocalization : localisation multi-échelles du gain de Kalman.

Input: $\mathbf{K}, \mathbf{A}_r^{gb}, \mathbf{A}^{wt}, I_r$
Output: \mathbf{K}
 $\mathbf{K} = \mathbf{A}_r^{wt} \circ \mathbf{K};$ // Localization in wavelet space
 $\mathbf{K}^{gb} \leftarrow \text{zeros}(n, n^e);$ // Temporary Kalman gain
for $d = 1$ **to** $d = n^d$ **do**
 forall the $p \in I_r$ **do**
 $\mathbf{K}^{gb}(p, d) = \mathbf{K}(p, d);$
 end
end
 $\mathbf{K}^{gb} = \mathbf{W}^{-1} \cdot \mathbf{K}^{gb};$
 $\mathbf{K}^{gb} = \mathbf{W} \cdot (\mathbf{A}_r^{gb} \circ \mathbf{K}^{gb});$
 $\mathbf{K} \leftarrow \mathbf{K}^{gb}[q(I_r), *];$ // extract sub-matix

Chapitre 5

Intégration séquentielle des données sismiques et de production : exemple d'application

« Longue est la route par le précepte, courte et facile par l'exemple. »

Sénèque, 4 av. J.C.- 65 ap. J.C.

5.1 Résumé et contributions

5.1.1 Résumé

Ce chapitre développe un exemple complet du *workflow* proposé dans ce travail de thèse. Les descriptions théoriques et les analyses détaillées des résultats sont disponibles dans les chapitres précédents. La figure 5.1 présente une vue d'ensemble des différentes étapes suivies.

Avant l'intégration des données indirectes, les observations directes de puits, ainsi que l'ensemble des connaissances *a priori* sont utilisées pour construire un premier modèle¹. En général, des méthodes géostatistiques sont utilisées pour produire ce modèle *a priori* (voir Chapitre 3). Afin d'être utilisable dans un processus d'inversion stochastique, le modèle *a priori* doit fournir une estimation des incertitudes initiales et un modèle de covariance afin de décrire la fonction de densité de probabilité (PDF) associée. Dans ce chapitre, un modèle gaussien, décrit par des variogrammes et une réalisation moyenne, est utilisé comme modèle *a priori*.

Les données sismiques sont ensuite utilisées pour modifier la PDF et réduire les incertitudes. Le processus d'inversion stochastique (décrit dans les Sections 1.2.1 et 1.3.3) est appliqué et génère un ensemble de réalisations représentatives des incertitudes liées à la sismique et au modèle direct. Ces réalisations sont interprétées en propriétés pétrophysiques, puis converties en profondeur à l'aide des observations de puits et d'un modèle de vitesses.

¹regroupant potentiellement plusieurs réalisations

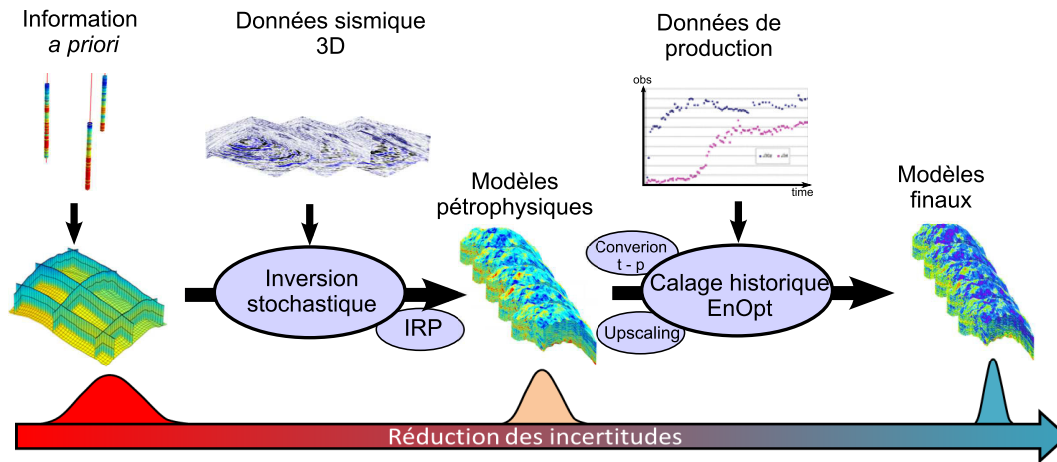


FIGURE 5.1 – Intégration séquentielle des données. IRP (*Inverse Rock Physics*) : transformation des modèles élastiques en modèles pétrophysiques. Conversion t - p : conversion temps - profondeur. *Upscaling* : génération de modèles grossiers à partir des modèles fins.

Enfin, la méthode multi-échelles proposée dans le Chapitre 4 est utilisée pour assimiler les données de productions. Cette méthode de calage historique minimise les perturbations des réalisations de l'ensemble, ce qui permet une meilleure préservation du conditionnement sismique initial et de la variabilité de l'ensemble. Les réalisations finales restent ainsi (autant que possible) en cohérence avec l'ensemble des données disponibles et donnent une bonne estimation des incertitudes.

5.1.2 Contributions

Ce chapitre inclut les contributions suivantes :

- développement d'une méthode d'intégration séquentielle des données sismiques et de production
- illustration de la méthode sur un cas synthétique 3D

5.2 Présentation du jeu de données synthétiques

Dans cette section, nous détaillons la création des données synthétiques.

5.2.1 Modèle géologique fin

Le modèle géologique initial est construit dans une grille de dimensions $(160 \times 200 \times 30)$ (environ 1 million de cellules). Le réservoir a une extension latérale de 5 km par 4 km et une épaisseur de 40 m.

Le réservoir est constitué de chenaux sableux d'orientation Nord-Sud, de lobes de débordements et de plaines d'inondations créés à partir d'une simulation objet (figure 5.2). La position des chenaux évolue verticalement, ce qui a un impact sur la connectivité des corps sableux. Le modèle de faciès sert de base à la génération de la porosité, perméabilité et du ratio net sur total (NTG).

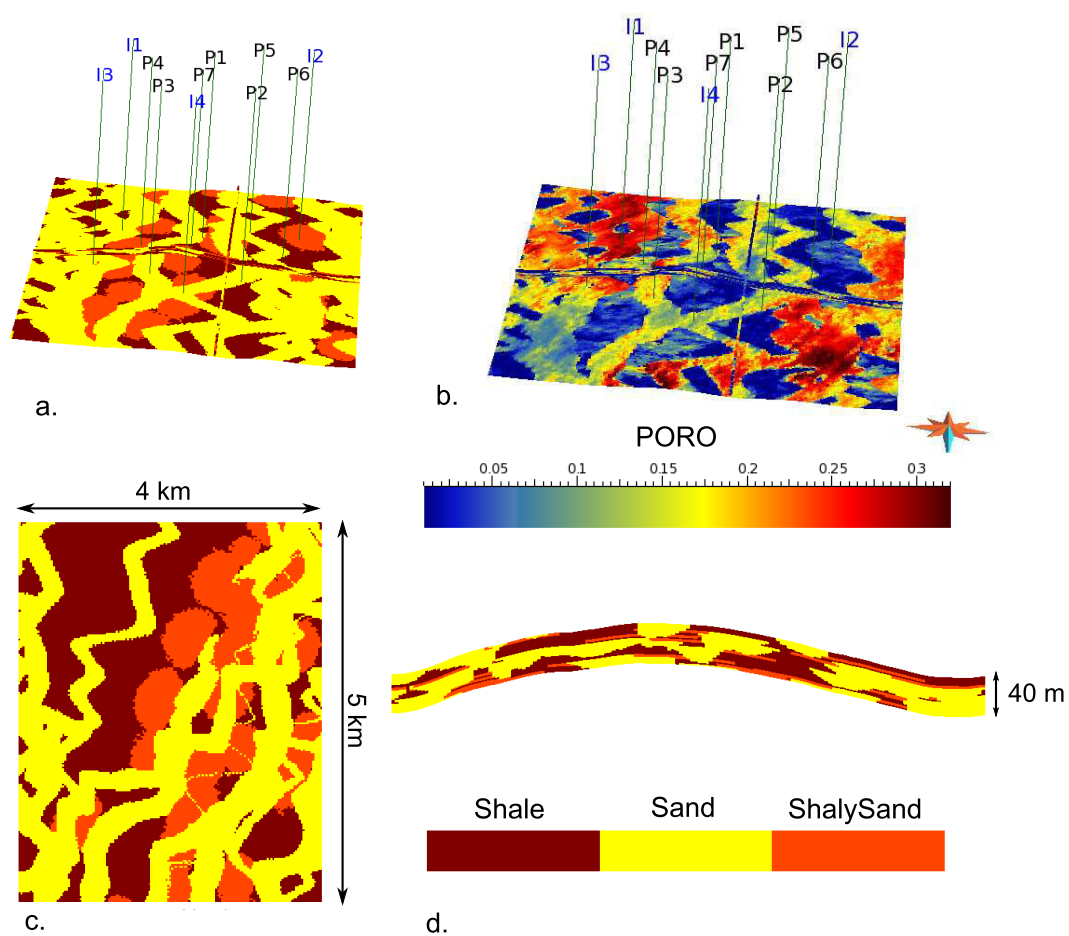


FIGURE 5.2 – Modèle géologique fin. (a) Vue d'ensemble du modèle de faciès (couche 16) et (b) son modèle de porosité associé. (c) Vue du haut du modèle de faciès. (d) Coupe Est-Ouest.

| | Chenaux | Lobes de débordements | Plaines d'inondations |
|--|--|--|---|
| Type de distribution | (Log)-gaussienne | Bimodale gaussienne | Log-Normale |
| Paramètres de la distribution (moyenne,écart-type) | Porosité (0.25, 0.045) NTG (0.86, 0.11) Log-Perm (5.538, 0.51) | Porosité $0.3 \times (0.05, 0.02)$; $(0.15, 0.03)$ NTG $0.3 \times (0.3, 0.05)$; $(0.5, 0.1)$ Log-Perm $0.5 \times (2.3, 1.3)$; $0.8 \times (4.5, 0.6)$ | Porosité (0.01, 0.018) NTG(0.05, 0.06) Perm (1, 3.57) |
| Valeur (min,max) | Porosité (0.07, 0.31) NTG (0.31, 1) Log-Perm (2, 6.9) | Porosité (0.01, 0.28) NTG (0.17, 0.85) Log-Perm (-3.2, 6.5) | Porosité (0.01, 0.18) NTG (0.01, 0.47) Log-Perm (-4, 4.5) |

Tableau 5.1 – Paramètres de simulations des propriétés. Distributions par faciès. La perméabilité est donnée en log(mD).

Différentes distributions (tableau 5.1) sont utilisées pour chacun des faciès et un variogramme anisotrope (orientation principale Nord 20°, distance de corrélation maximale/minimale 0.3/0.15 fois l'extension du modèle) est utilisé pour simuler les propriétés à partir de l'algorithme SGS (Section 3.2.1).

Trois fluides sont présents dans le réservoir : eau, huile et gaz (figure 5.3). Les contacts initiaux eau-huile et huile-gaz sont respectivement à la profondeur de 1260 m et 1176 m. Le chapeau de gaz (*gas cap*) permet en partie de maintenir la pression dans le réservoir. Le couvercle du réservoir est constitué d'argiles imperméables qui arrêtent la progression des fluides. Les propriétés pétrophysiques de la couverture et de la base du réservoir sont générées par une simulation SGS. Bien qu'elles n'interviennent pas dans le modèle d'écoulements, ces formations ont un impact sur le signal sismique du réservoir.

5.2.2 Création des données sismiques

Modèle pétro-élastique (PEM)

Les propriétés pétrophysiques sont ensuite transformées en propriétés élastiques, (V_p, V_s, ρ) , à partir d'un modèle pétro-élastique (cf. Section 1.2.1) de sables à grains cimentés (cf. figure 5.4). Dans ce modèle, la structure de la roche est décrite par des grains sphériques identiques (modèle de Hertz-Mindlin) dont leur surface et leurs contacts sont en partie cimentés. La vitesse de propagation des ondes va dépendre de la densité des grains (c.-à-d de la porosité), de la composition minéralogique (c.-à-d du NTG), de la quantité de ciment et des fluides interstitiels (c.-à-d des saturations). Avec l'augmentation de la porosité,

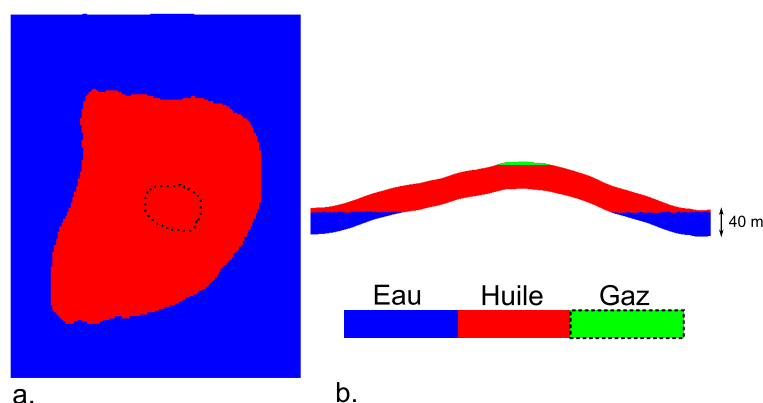


FIGURE 5.3 – Contacts fluides. (a) Vue de dessus, couche 16. La forme en pointillé représente la projection du contact huile-gaz. (b) Coupe Est-Ouest.

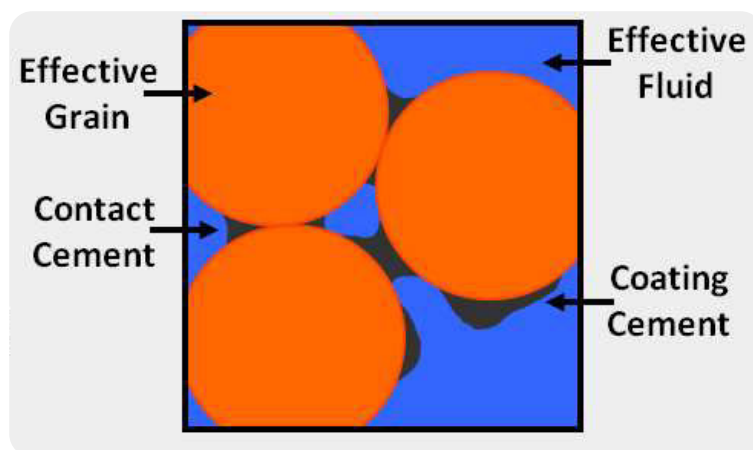


FIGURE 5.4 – Représentation schématique d'un modèle de grès cimentés. Tirée de la documentation interne de CGG.

le nombre de contacts entre grains diminue et la roche devient moins rigide. Inversement, la rigidité de la roche augmente avec la quantité de ciment² (voir figure 5.5). Plus la roche est rigide, moins le contenu fluide a un impact sur les vitesses de propagation. Les paramètres d'entrée du modèle pétro-élastique sont donnés dans le tableau 5.2. Les modules de compression et de cisaillement pour la matrice sont calculés à partir de la théorie de Hertz-Mindlin, du modèle cimenté de Dvorkin et des limites de Hashin-Shtrikman [77]. Le module de compression pour les fluides est donné par le modèle de Wood. Voir [7] pour une description détaillée de ces modèles.

Simulation des amplitudes sismiques

Les champs de vitesses et de densité créés à partir du PEM sont dans un premier temps transformés en cube de coefficients de réflexions (voir Sec-

²sauf si la porosité est supérieure à la porosité critique, auquel cas il n'existe plus de contacts entre les grains

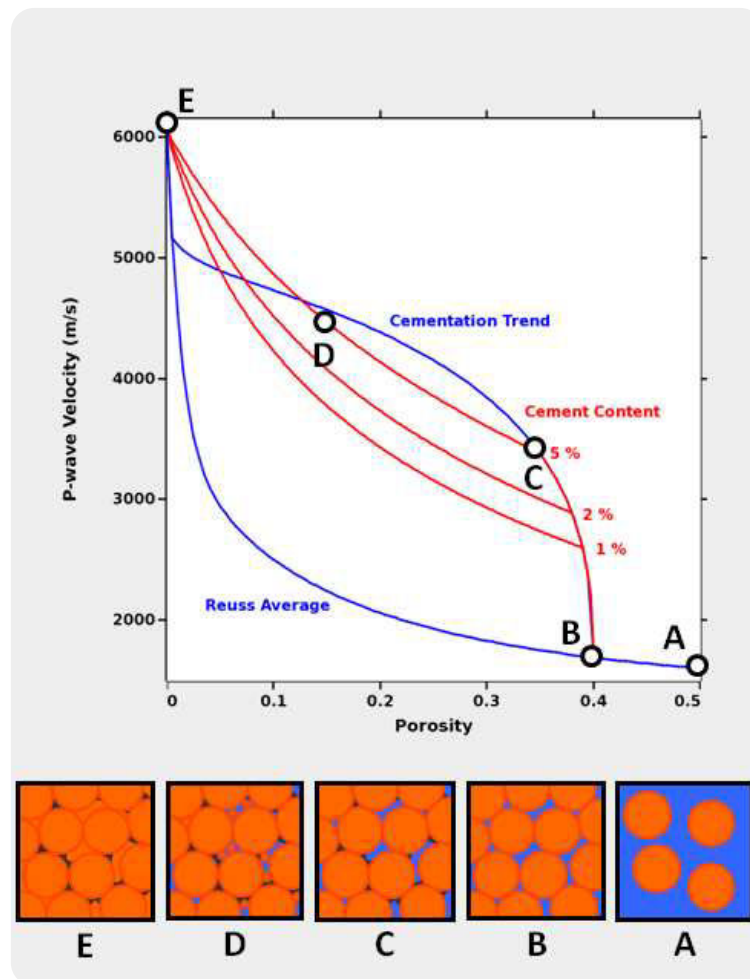


FIGURE 5.5 – Évolution de V_p en fonction de la porosité et de la quantité de ciment. Tirée de la documentation interne de CGG.

| Paramètres | Valeurs |
|-----------------------------------|---|
| Porosité | MG |
| Volume d'argiles | MG (1-NTG) |
| Fluide interstitiel | MG |
| Proportion de ciment | 0.5 |
| Densité des fluides (g/cm^3) | eau : 1 huile : 0.7 gaz : 0.15 |
| Densité des minéraux (g/cm^3) | Quartz, ciment : 2.65 Argiles : 2.55 |
| Modules de cisaillement (GPa) | Quartz, ciment : 45 Argile : 9 |
| Modules de compression (GPa) | Quartz, ciment : 37 Argile : 25 |
| Porosité critique | 0.4 |

Tableau 5.2 – Paramètres du PEM. MG : Propriété donnée par le modèle géologique. Voir [7] pour plus d'informations sur les autres paramètres.

tion 1.2.1). Ces coefficients sont ensuite bruités³ et lissés afin d'obtenir un signal sismique plus réaliste. Finalement, les ondelettes de la figure 5.6 sont utilisées pour générer le signal sismique synthétique *pre-stack* (voir Section 1.2.1) pour 4 gammes d'angles d'incidence. Les amplitudes sismiques résultantes sont présentées dans la figure 5.7. Malgré l'ajout de bruit, le signal sismique généré est de bonne qualité et il est possible de discerner certains chenaux.

5.2.3 Création des données de production

Les données de production sont générées à partir d'un modèle mis à l'échelle aux dimensions ($80 \times 100 \times 15$), comportant 94 035 cellules actives (figure 5.8). Certaines zones de l'aquifère sont désactivées afin de diminuer le temps de calcul. Une simulation *black-oil* (cf. Section 1.2.2) est effectuée avec les perméabilités relatives présentées dans la figure 5.9. Les perforations des puits et le plan de production du champ sont donnés dans le tableau 5.3. Les données se composent de 12 années d'historique de pression de fond de puits (BHP), de production d'eau (WCT) et de ratio de production huile/gaz (GOR).

La plupart des puits ont une contrainte de production/injection constante, sauf le puits P7 qui est proche de la nappe de gaz et le puits P6 qui a une arrivée d'eau précoce.

Du bruit aléatoire (distribution uniforme) est ajouté aux données. Les écarts types associés (et utilisés lors du calage) sont donnés dans le tableau 5.4. Les données de productions restent néanmoins très cohérentes avec le modèle réel et sont considérées de bonne qualité.

Le jeu de données se compose de 790 observations de pression, d'arrivée d'eau et de gaz.

³bruit corrélé, simulé à partir d'une SGS. Bien évidemment, en conditions réelles le bruit est souvent beaucoup plus complexe et peut provenir d'une multitude de phénomènes différents.

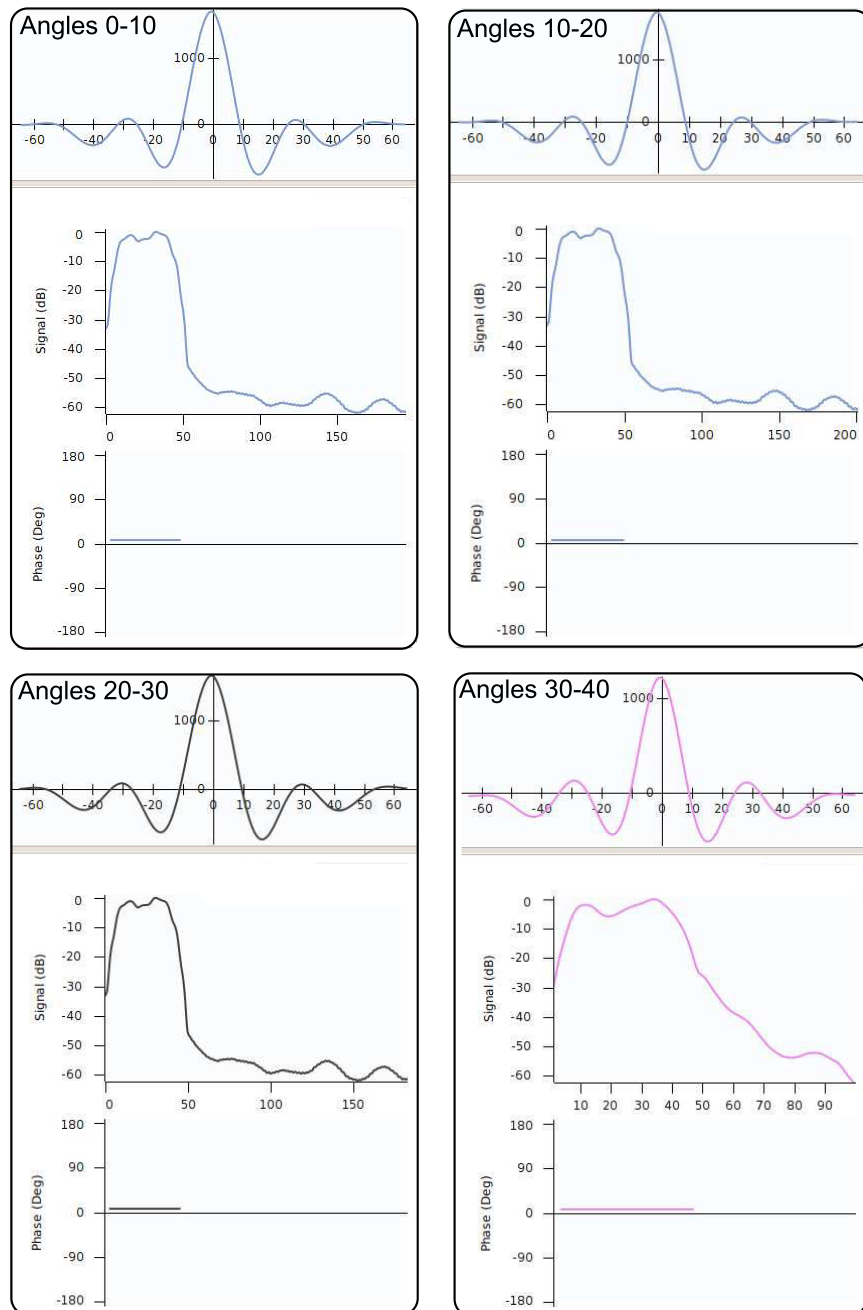


FIGURE 5.6 – Ondelettes synthétiques. Amplitudes, contenus fréquentiels et rotations de la phase des ondelettes utilisées pour générer le signal sismique *pre-stack* pour 4 gammes d'angles.

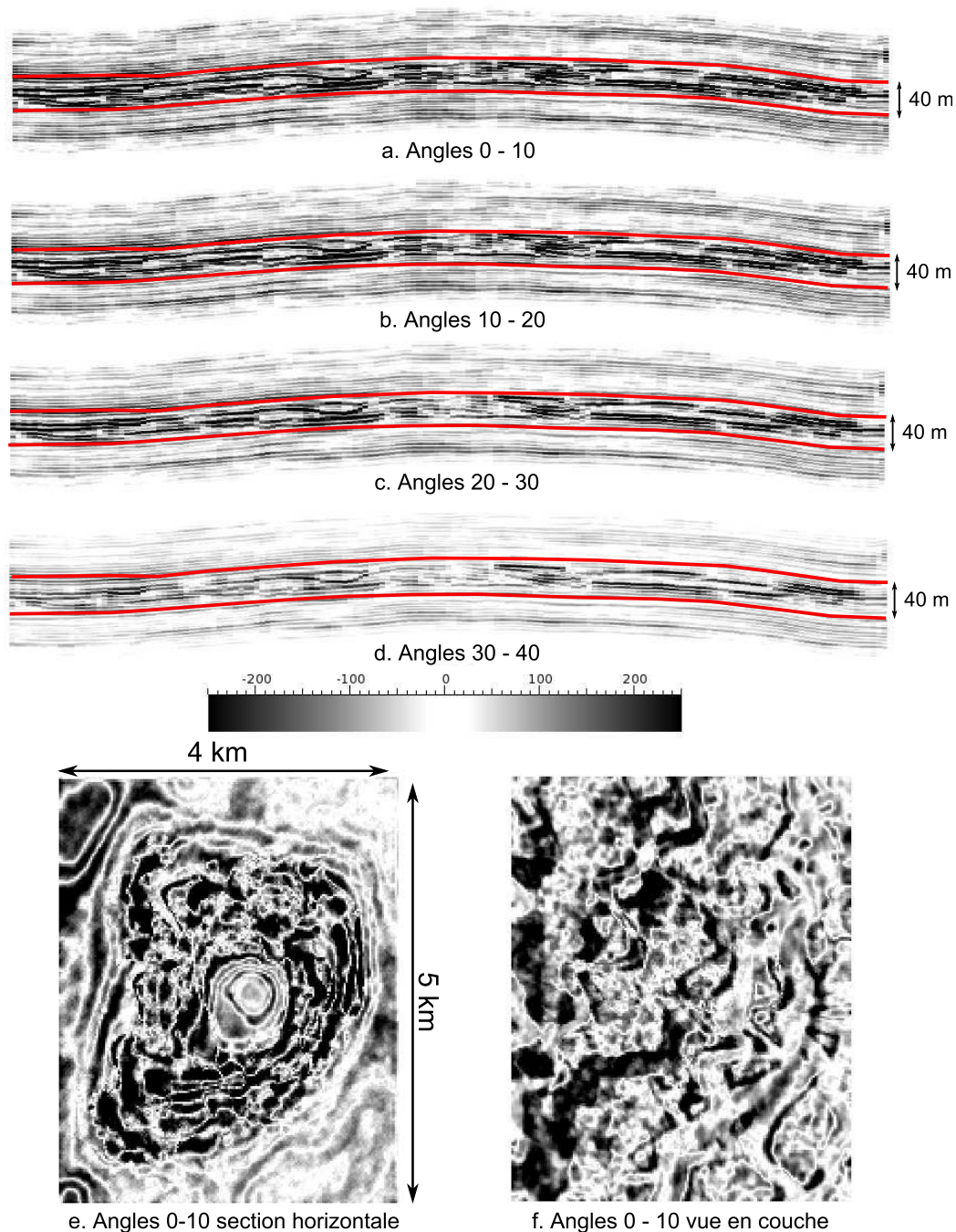


FIGURE 5.7 – Sismique synthétique. (a-d) Section sismique (Est-Ouest) pour les 4 angles. (e) Coupe horizontale. (f) Représentation en couche. Les lignes rouges indiquent les limites (top et base) du réservoir.

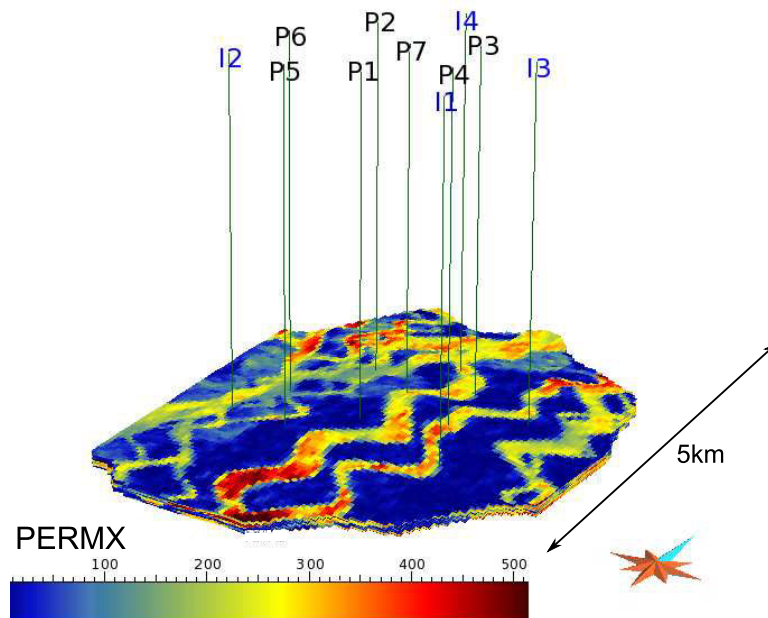


FIGURE 5.8 – Grille de simulation des données de production. *Upscaling* $2 \times 2 \times 2$ de la grille géologique. Certaines cellules de l'aquifère, éloignées des puits, ont été désactivées.

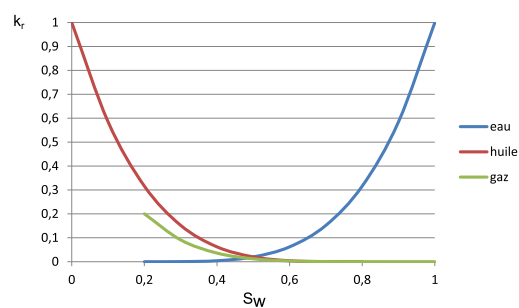


FIGURE 5.9 – Perméabilités relatives. Pour obtenir la saturation en huile ou gaz, prendre $1 - S_w$.

| Nom du puits | Type de puits | Contrainte(s) primaire(s) | Perforations (couches) |
|--------------|---------------|---|------------------------|
| I1 | Inj. | $150 \text{ sm}^3/\text{jour}$ | 3 à 11 ; 14 à 15 |
| I2 | Inj. | $150 \text{ sm}^3/\text{jour}$ | 1 à 15 |
| I3 | Inj. | $200 \text{ sm}^3/\text{jour}$ | 4 à 15 |
| I4 | Inj. | $200 \text{ sm}^3/\text{jour}$ | 1 à 15 |
| P1 | Pro. | $150 \text{ sm}^3/\text{jour}$ | 5 à 12 ; 15 |
| P2 | Pro. | $150 \text{ sm}^3/\text{jour}$ | 1 à 7 ; 13 à 15 |
| P3 | Pro. | $150 \text{ sm}^3/\text{jour}$ | 1 et 2 ; 9 à 15 |
| P4 | Pro. | $150 \text{ sm}^3/\text{jour}$ | 1 à 11 ; 14 et 15 |
| P5 | Pro. | $150 \text{ sm}^3/\text{jour}$ | 5 à 15 |
| P6 | Pro. | 0-1826 jours : $150 \text{ sm}^3/\text{jour}$ 1826-2192 : $130 \text{ sm}^3/\text{jour}$ 2192-fin : 0 | 1 à 8 |
| P7 | Pro. | 0-1826 jours : $100 \text{ sm}^3/\text{jour}$ 1826-fin : $150 \text{ sm}^3/\text{jour}$ | 9 à 15 |

Tableau 5.3 – Puits et plan de production. Historique des changements du contrôle des puits.

| Type de données | Écarts-types du bruit | Unités |
|---------------------------------|-----------------------|---------------------------|
| Pression de fond de puits (BHP) | 3 | Bar |
| Production d'eau (WCT) | 2.5 | % |
| Ratio huile-gaz (GOR) | 3 | sm^3/sm^3 |
| GOR P7 | 6 | sm^3/sm^3 |

Tableau 5.4 – Données et bruit. Écarts-types du bruit aléatoire ajouté aux données simulées.

5.3 Inversion des données

Les données synthétiques, présentées dans la section précédente, sont ensuite assimilées en suivant le *workflow* de la figure 5.1. L'inversion sismique s'effectue dans la grille dans laquelle ont été générées les données, tandis que le processus de calage historique s'effectue dans une grille plus grossière, la géométrie du domaine restant cependant la même.

5.3.1 Inversion sismique stochastique des paramètres élastiques

L'inversion sismique stochastique (Section 1.2.1 et 1.3.3) est utilisée pour obtenir les impédances élastiques I_p et I_s . La sismique synthétique est supposée être contaminée par un bruit homogène tel que l'incertitude sismique soit de 8%. Les variations spatiales du modèle *a priori* sont modélisées latéralement de façon isotrope par un variogramme cubique de portée égale à 150 m et verticalement par un variogramme exponentiel ayant une portée de 12 ms (environ 4 couches) interprété à partir des observations de puits.

Le réservoir étudié étant composé d'objets discrets (chenaux), le modèle de corrélation deux-points (variogrammes) utilisé par l'inversion Bayésienne [23] est mal adaptée (cf. Chapitre 3). Afin de ne pas générer des artéfacts près des puits, dus à l'interpolation d'un nombre limité de données ponctuelles, les observations élastiques aux puits ne sont pas utilisées lors de l'inversion afin que les propriétés inversées ne dépendent que des données sismiques. De même, un modèle homogène (égal à l'impédance moyenne aux puits) est utilisé comme moyenne *a priori*.

Les ondelettes utilisées dans l'inversion sont estimées à partir de la sismique et des données de puits. L'ondelette associée aux angles 30 – 40° est présentée dans la figure 5.10. Malgré un léger décalage de la phase, l'estimation de l'ondelette est de très bonne qualité.

Les figures 5.12 et 5.11 montrent les résultats de l'inversion stochastique. L'un des avantages de l'inversion stochastique est de pouvoir inclure un modèle *a priori* vertical fin et ainsi produire directement des réalisations à l'échelle du modèle de réservoir (constitué de couches fines) cohérentes avec les données sismiques.

5.3.2 Interprétation des modèles élastiques

Les réalisations élastiques sont ensuite interprétées en termes de propriétés pétrophysiques. Une analyse statistique permet de trouver une relation entre les données de porosité de puits et les impédances générées par l'inversion des données sismiques. Le lien étant incertain, plusieurs relations aléatoires sont générées pour chaque réalisation (voir figure 5.13) afin d'échantillonner les incertitudes liées à l'interprétation pétro-élastique et obtenir des réalisations cohérentes avec les paramètres élastiques.

Une relation entre la porosité et la perméabilité (équation 5.1) est déduite à partir des données d'observation de puits.

$$PERM = 10^{A \times \log_{10}(PORO) + B} \quad (5.1)$$

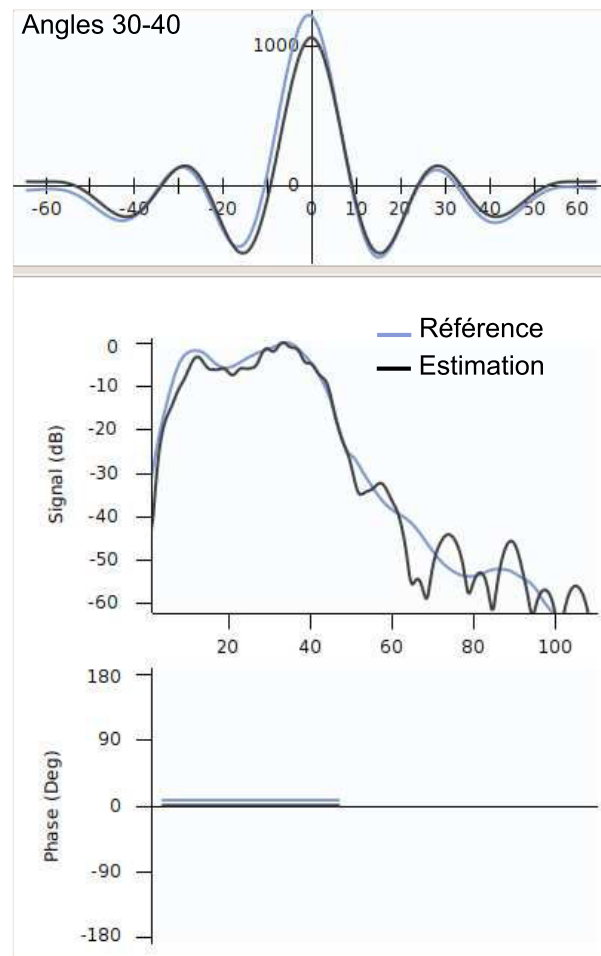


FIGURE 5.10 – Estimation de l’ondelette 30–40°. Comparaison de l’ondelette estimée pour la gamme d’angles 30 – 40° avec l’ondelette utilisée pour générer les données synthétiques

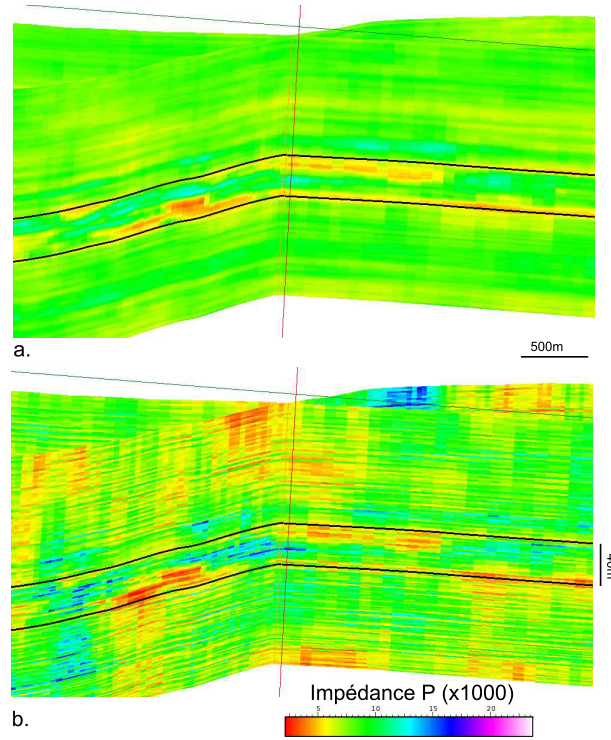


FIGURE 5.11 – Résolution verticale d'une réalisation d'impédance. (a) Inversion déterministe (moyenne des réalisations) (b) Réalisation d'inversion stochastique.

Cette relation étant incertaine, les paramètres A et B de l'équation sont tirés aléatoirement pour chaque réalisation dans les distributions $A \mathcal{N}(2.26, 0.1)$ et $A \mathcal{N}(3.79, 0.2)$ (figure 5.14).

Le NTG est cosimulé à partir de la porosité et des paramètres élastiques, tandis que la perméabilité verticale est simulée à partir de la perméabilité horizontale, ces deux dernières étant reliées par une relation de proportionnalité. Quelques réalisations pétrophysiques sont présentées dans la figure 5.15.

Des procédés plus complexes, impliquant une modélisation des lithologies ou l'utilisation de densités de probabilités empiriques [7, 45], auraient pu être utilisés pour transformer les champs de propriétés élastiques. Cette étape n'étant pas l'objet principal de ce travail de thèse, et bien qu'elle soit une étape primordiale du *workflow*, nous nous contentons ici d'un processus d'interprétation simple à mettre en œuvre.

5.3.3 Upscaling et conversion temps-profondeur

Upscaling

La figure 5.16 présente la grille grossière ($40 \times 50 \times 15$) utilisée pour le calage de l'historique de production. Un processus d'*upscaling* (aussi appelé homogénéisation) est utilisé pour obtenir des modèles de propriétés grossiers équivalents aux modèles issus de l'interprétation sismique. Cette étape peut s'avérer primordiale dans les études réservoirs et fait l'objet de recherches [110, 80, 66]. Cependant, l'efficacité et la précision de l'algorithme d'*upscaling* sont moins

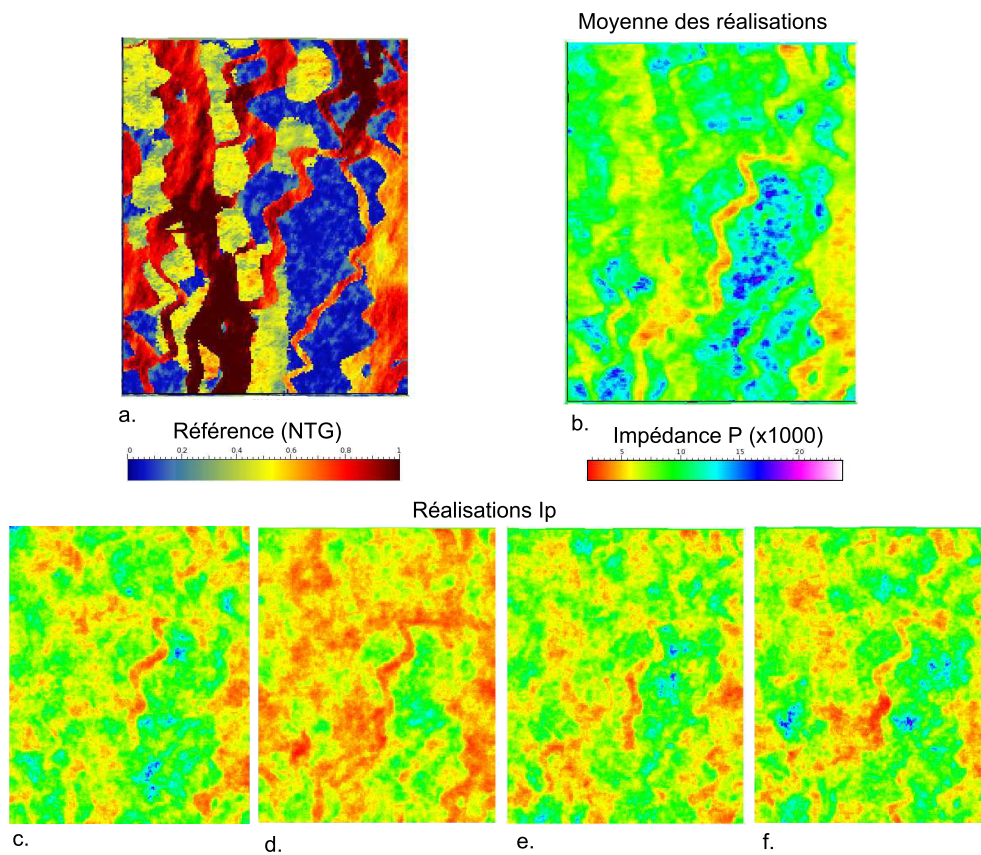


FIGURE 5.12 – Réalisations d'impédances P, couche 24 (modèle fin). (a) Image de référence des structures (NTG). (b) Moyenne des réalisations d'impédances P. (c-f) Réalisations d'impédances P.

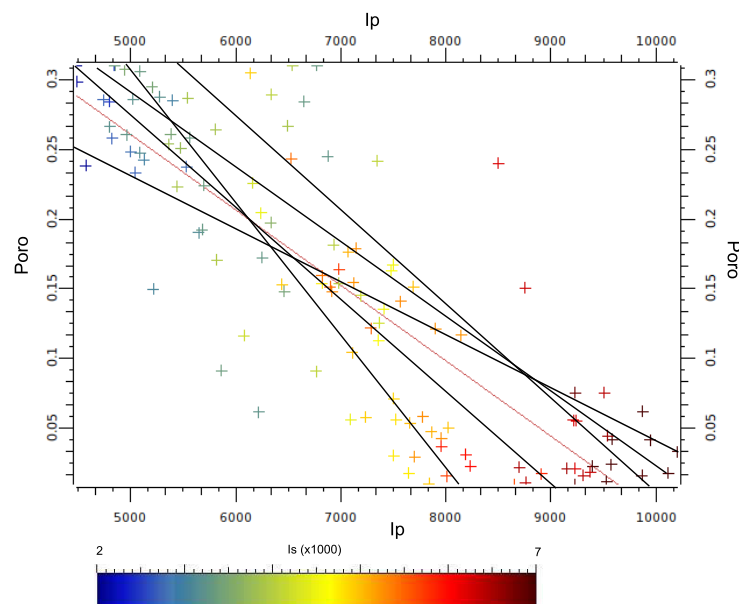


FIGURE 5.13 – Relations Poro- I_p - I_s . Les croix représentent les observations aux puits. Elles sont colorées avec les valeurs de I_s . Les différentes lignes représentent des relations linéaires individuelles simulées pour chaque réalisation.

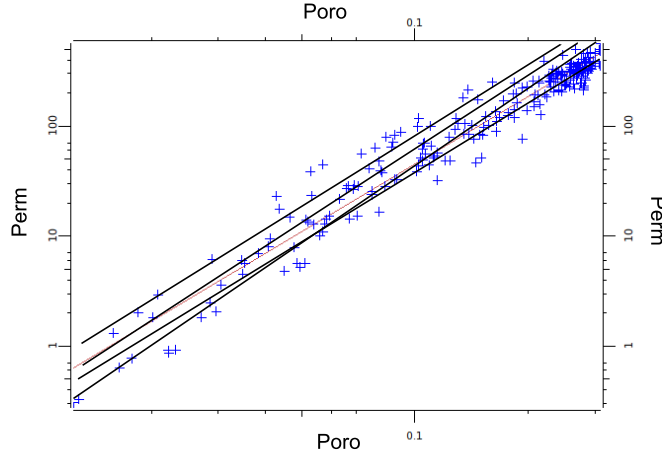


FIGURE 5.14 – Relations Perméabilité - Porosité. Les croix représentent les observations aux puits. Les différentes lignes représentent des relations individuelles simulées pour chaque réalisation.

importantes lorsque les modèles grossiers sont ensuite utilisés dans un processus d'optimisation. En effet, l'erreur d'*upscaling* est implicitement corrigée lors du calage historique. Pour cette raison, nous utilisons ici des méthodes simples.

Une simple moyenne arithmétique est utilisée pour créer les champs grossiers à partir des propriétés fines de porosité et de NTG. La perméabilité étant un tenseur, la propriété équivalente dépend de la direction. La figure 5.17 présente deux configurations d'écoulements pour un même champ de perméabilités absolues. La perméabilité équivalente, k_e , est différente dans les deux cas. Pour un écoulement en parallèle (cas a), la perméabilité équivalente est donnée par la moyenne arithmétique :

$$k_e = \frac{h_1.k_1 + h_2.k_2 + h_3.k_3}{h_1 + h_2 + h_3} \quad (5.2)$$

alors que pour un écoulement en série (cas b), la perméabilité équivalente est donnée par la moyenne harmonique :

$$k_e = \frac{h_1 + h_2 + h_3}{\frac{h_1}{k_1} + \frac{h_2}{k_2} + \frac{h_3}{k_3}} \quad (5.3)$$

Ainsi, pour la perméabilité verticale, la perméabilité équivalente est obtenue en effectuant une moyenne arithmétique latérale, suivi d'une moyenne harmonique verticale. La perméabilité latérale incluant deux directions, un procédé différent est utilisé. Latéralement, la moyenne géométrique est utilisée pour approximer la perméabilité équivalente, tandis qu'une moyenne arithmétique est appliquée verticalement (cas d'écoulements parallèles).

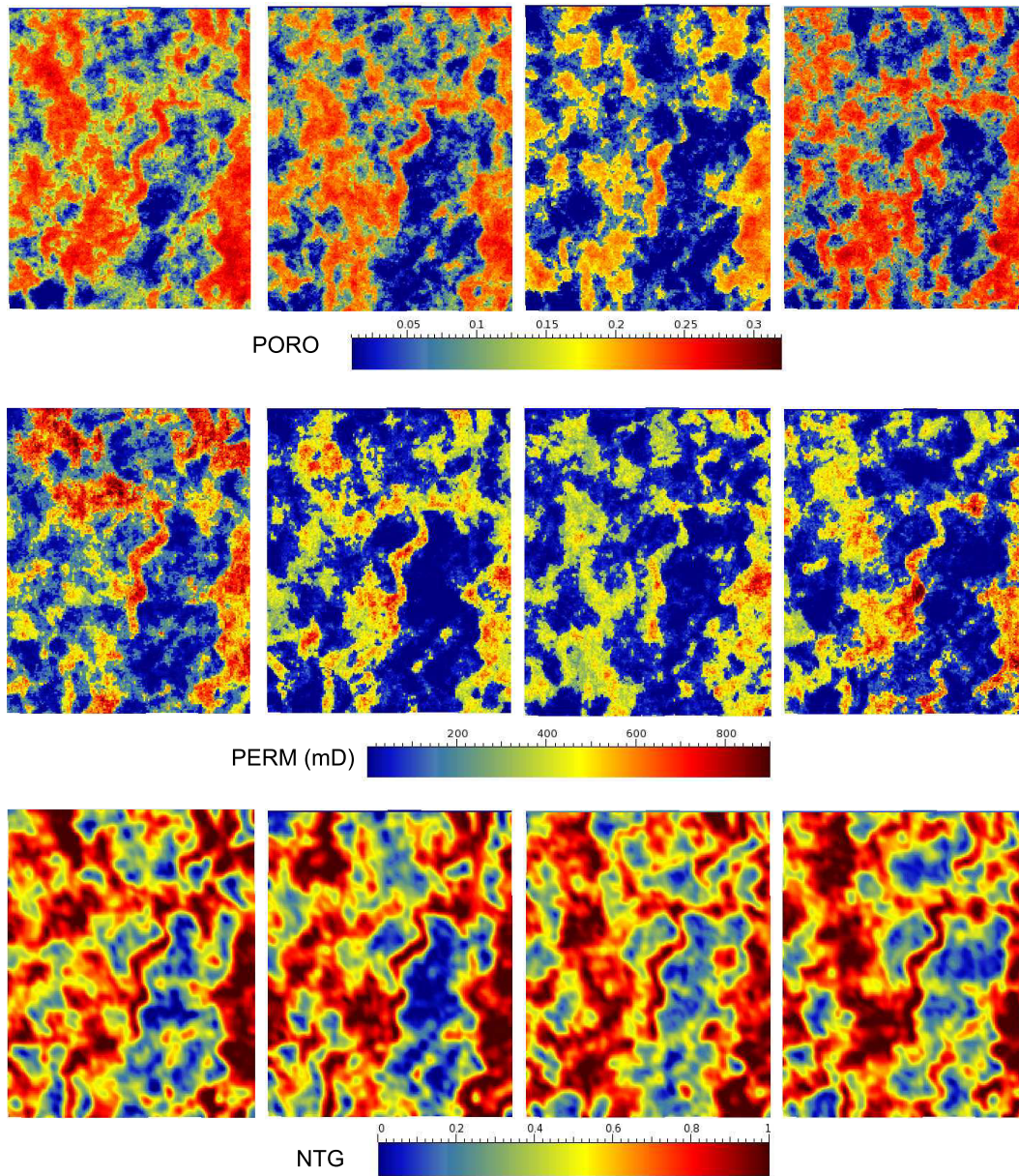


FIGURE 5.15 – Réalisations pétrophysiques. Interprétations des réalisations (c-f) de la figure 5.12.

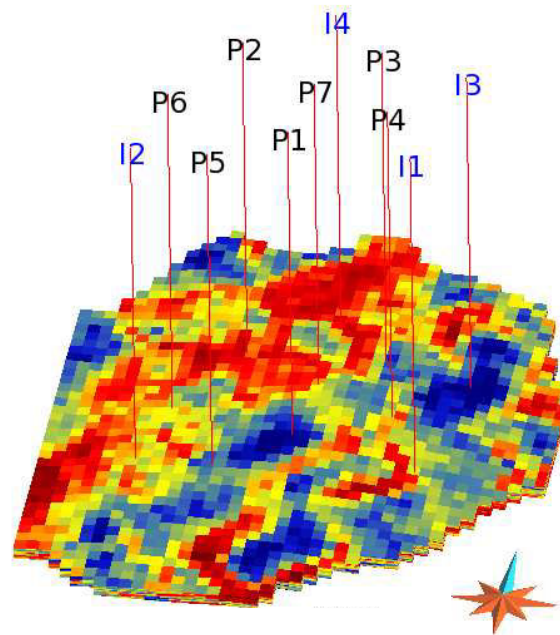


FIGURE 5.16 – Grille de réservoir utilisée lors de l'inversion des données de production. La résolution verticale reste fine tandis que les dimensions latérales sont divisées par deux par rapport à la grille utilisée pour générer les données.

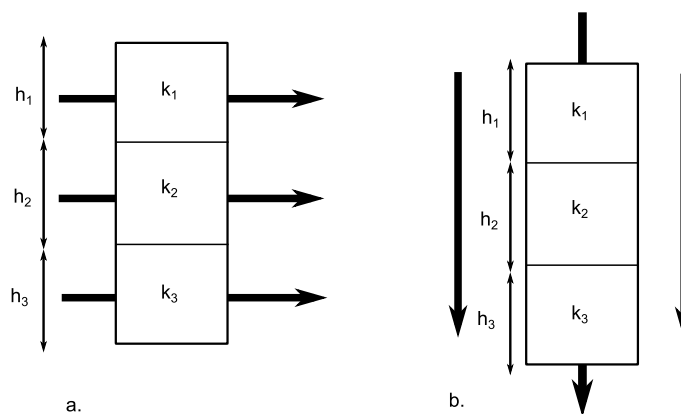


FIGURE 5.17 – Perméabilités équivalentes. (a) Flux en parallèle. (b) Flux en série.

Conversion temps-profondeur

La problématique de conversion temps-profondeur n'étant pas l'objet de ce travail de thèse, un modèle de vitesses constantes est utilisé pour convertir le modèle. Cependant, en pratique, cette étape a une influence importante sur les éléments structuraux du modèle et l'épaisseur des couches, ce qui peut avoir un impact de premier ordre sur la réponse aux écoulements. Ainsi, il peut être utile d'intégrer cette étape dans le processus de calage historique [133].

5.3.4 Inversions des données de production

Les modèles obtenus sont mis à jour par les données de production en utilisant la méthode multi-échelles présentée dans le Chapitre 4. Afin de donner assez de liberté au début de l'optimisation, des transformations indépendantes sont effectuées pour chacune des couches. Le réservoir étant composé d'objets sédimentaires plus ou moins connectés (chenaux, lobes de débordement), il est nécessaire de démarrer avec une paramétrisation permettant de les discriminer. Cependant, le nombre de paramètres au début de l'optimisation reste très réduit par rapport à une paramétrisation complète (figure 5.19).

Les paramètres inversés incluent les coefficients des ondelettes quadratiques présentées dans la Section 2.3.2 et les contacts initiaux eau/huile - huile/gaz (nécessaires après l'*upscaling*). L'ensemble initial est composé de 40 réalisations dérivées de la sismique.

Les mises à jour de l'algorithme LM-enRML utilisent les méthodes multi-échelles de contrôle développées dans la Section 4.4. La localisation multi-échelles utilise des fonctions de distances isotropes centrées sur les puits pour la pression et l'arrivée de gaz, alors que des fonctions anisotropes alignées entre les différents couples de puits producteurs/injecteurs sont utilisées pour les arrivées d'eau (figure 5.18). Les réalisations étant conditionnées par les données sismiques, leur lissage initial est moins accentué que dans les exemples de la Section 4.5. Des tests empiriques de lissage sont effectués afin de régler l'atténuation initiale (voir section 4.3.2).

L'évolution de la fonction objectif, ainsi que le nombre de paramètres inclus dans l'optimisation sont présentés dans la figure 5.19. L'erreur de calage initial est relativement basse, indiquant que les données sismiques sont cohérentes avec les données de production et permettent de réduire les incertitudes *a priori*. Nous pouvons observer que l'erreur de calage diminue très rapidement grâce à la modification d'un nombre réduit de paramètres de basses résolutions. Les étapes de raffinement ne provoquent pas d'augmentation de la fonction objectif. En effet, les réalisations initiales étant conditionnées par des données sismiques, elles sont plus cohérentes et une atténuation plus faible des hautes fréquences a été appliquée avant l'optimisation.

La figure 5.20 montre les calages initiaux et finaux de l'ensemble des réalisations. Le calage des données est quasi parfait et nous pouvons remarquer que la variabilité de l'ensemble reste assez importante.

La figure 5.21 présente quelques champs de perméabilité et de porosité avant et après assimilation des données de production. L'approche multi-échelles permet une bonne préservation des propriétés. Bien que la paramétrisation initiale

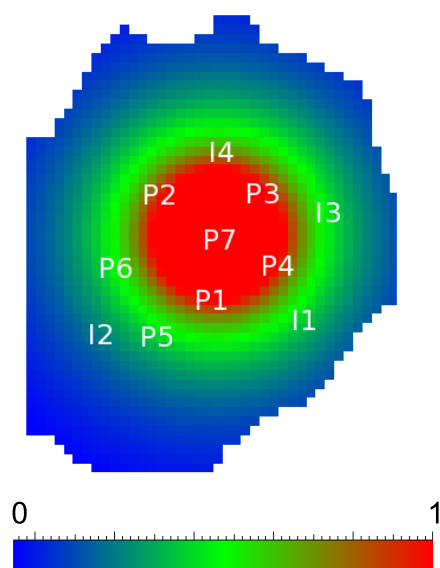


FIGURE 5.18 – Localisation des données de pression, puits P7. Exemple de localisation pour la résolution la plus fine. Voir Section 4.4.1 pour plus d'informations sur la localisation multi-échelles.

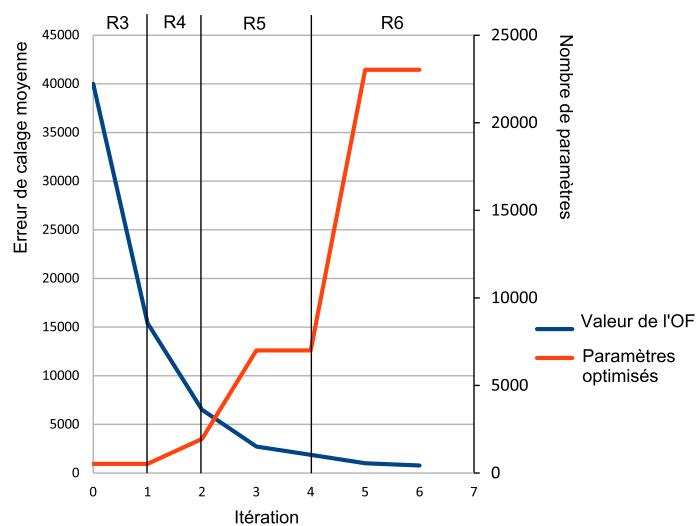


FIGURE 5.19 – Évolution de la fonction objectif et du nombre de paramètres. L'algorithme multi-échelles itère à travers 4 échelles différentes (R1, R2, R3, R4).

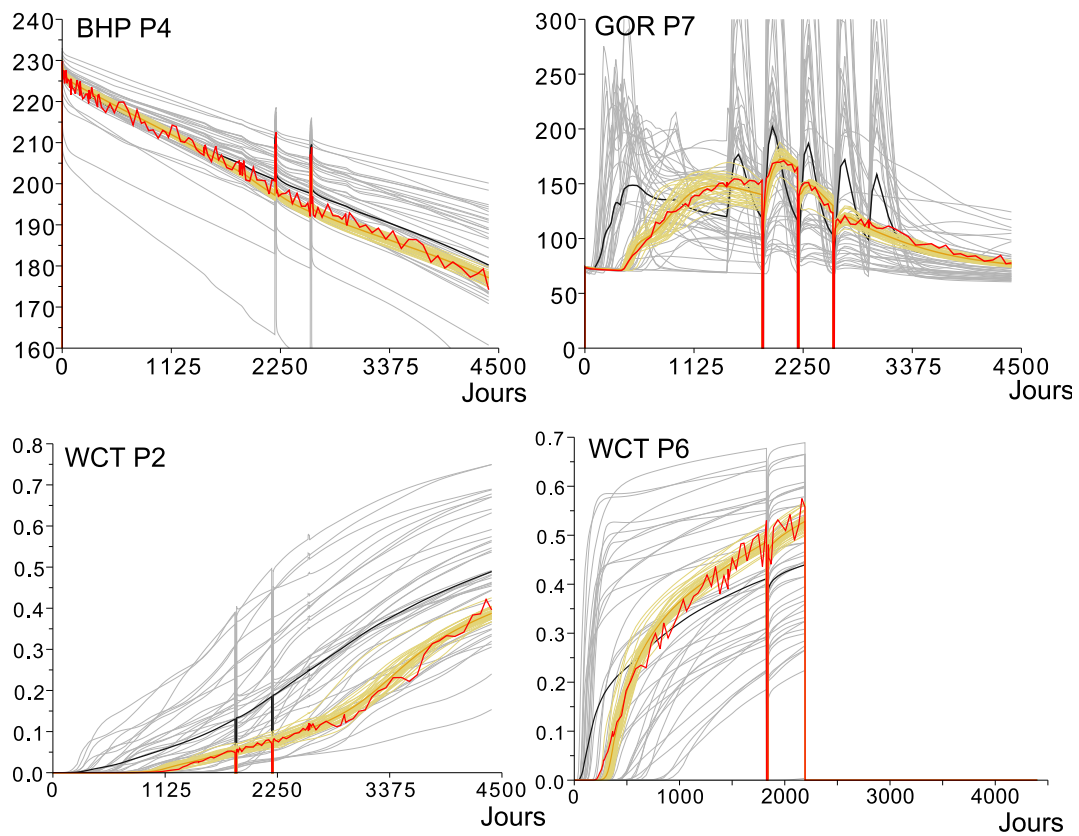


FIGURE 5.20 – Calage des données de production. La courbe rouge représente les données bruitées, les courbes grises les prédictions initiales, les courbes oranges les prédictions finales.

soit définie de façon indépendante pour chaque couche, l'effet de régularisation latérale permet de limiter les perturbations des moyennes/hautes fréquences verticales qui peuvent impacter le calage à la sismique.

5.4 Perspectives et conclusions partielles

Le *workflow* présenté dans ce chapitre permet de créer des modèles cohérents avec les données sismiques et de production. L'approche séquentielle est possible grâce à la méthode de calage historique multi-échelles qui minimise les perturbations des modèles et privilégie la modification des fréquences qui n'impactent pas le calage aux données sismiques.

Cependant, afin de vérifier le bon fonctionnement de cette méthode, il peut être nécessaire d'effectuer un contrôle de qualité en utilisant les données sismiques. Par exemple, pour chacune des réalisations, il est possible de générer des amplitudes sismiques (à partir d'un PEM et des ondelettes utilisées lors de l'inversion) et d'analyser leur calage initial à l'aide de méthodes statistiques de contrôle de qualité [65]. Le même procédé peut être appliqué à la fin du calage des données de production. Les analyses de qualité peuvent être effectuées entre les modèles mis à jour et les données sismiques, mais également entre les réponses initiales et postérieures au calage de l'historique de production. Lorsque l'erreur initiale est comparable à l'erreur finale, on peut considérer que le calage historique affecte raisonnablement le calage aux données sismiques. Une déviation importante des modèles peut refléter une incohérence des données (données de mauvaise qualité, modélisation des erreurs mal adaptée), une contamination importante par les fausses corrélations (mauvaise localisation, ensemble trop petit) ou une erreur lors de l'inversion des données sismiques (mauvais modèle *a priori* ou mauvaise transformation des modèles élastiques). Ces critères peuvent également servir à comparer des méthodes de calage historique différentes. En effet, cette approche permet une analyse fréquentielle et spatiale beaucoup plus informative que l'analyse variographique ou que les cartes de déviation des propriétés utilisées dans le Chapitre 4.

Bien que l'exemple synthétique présenté dans ce chapitre donne une bonne illustration de l'application du *workflow* proposé dans cette thèse, les processus de génération des données et d'inversions partageant la même grille et les mêmes modèles directs, le problème posé ne reflète pas les conditions réelles d'étude d'un réservoir. De plus, certaines étapes primordiales, telles que la conversion temps-profondeur ou l'*upscaling*, ont été simplifiées, mais peuvent avoir un impact important sur le schéma d'inversion proposé. Ainsi, une validation avec un cas d'étude réel reste nécessaire.

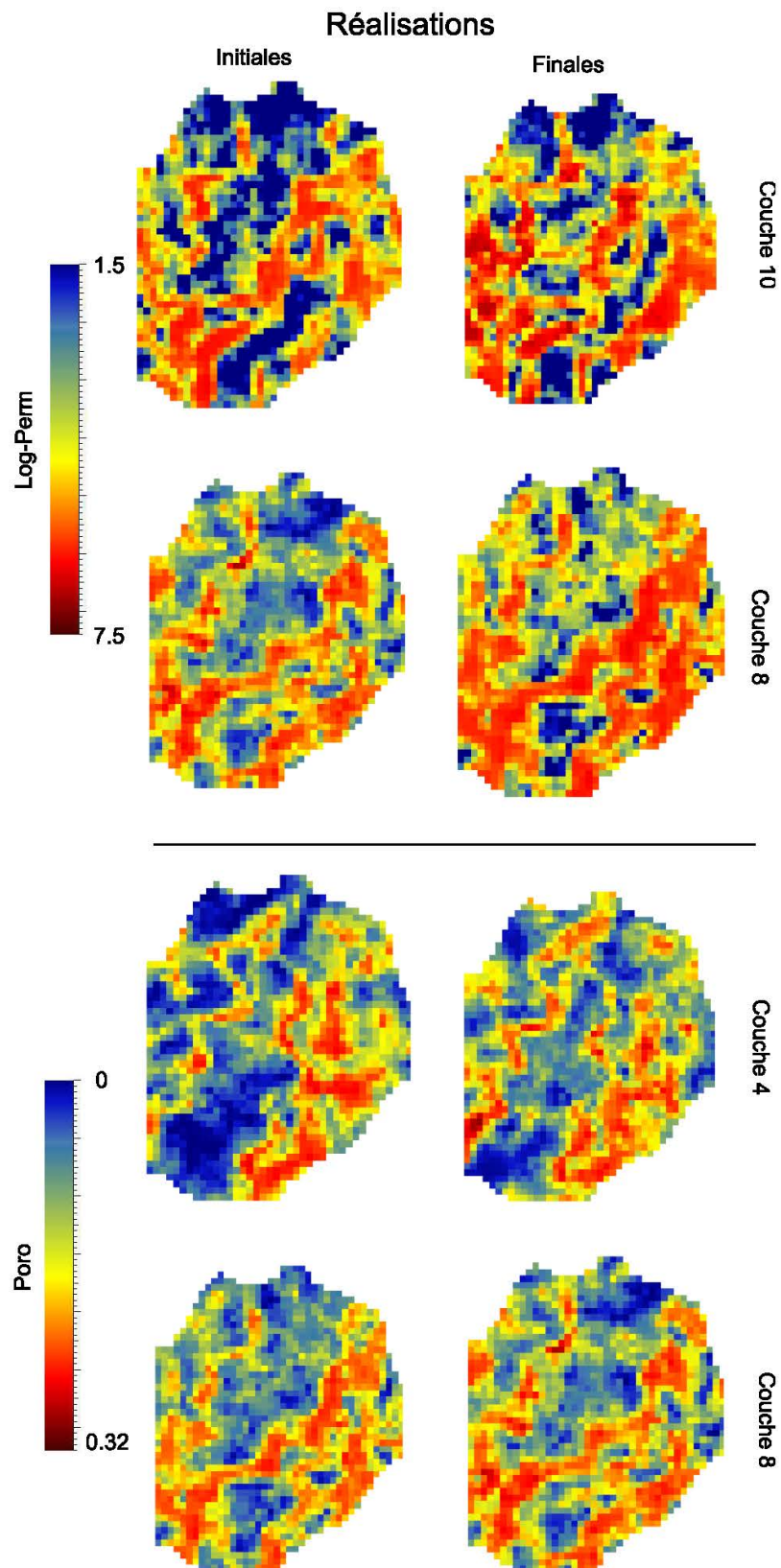


FIGURE 5.21 – Réalisations initiales et finales. Log-Perméabilité (haut) et Porosité (bas). On note une très bonne préservation des modèles initiaux.

Chapitre 6

Conclusions générales et perspectives

« Life is like riding a bicycle. To keep your balance, you must keep moving. »

Albert Einstein, 1879-1955

6.1 Contributions de la thèse

Des paramétrisations adaptées aux données de réservoir Que ce soit dans la résolution de problèmes inverses ou la simulation de réalisations géostatistiques, l'utilisation de paramétrisations définies à la fois en fréquence et dans l'espace facilite l'intégration des données de résolutions et de natures différentes.

Ainsi, il est possible d'adapter les contraintes appliquées sur les paramètres estimés lors des processus d'inversion en fonction de l'information localement disponible. Lorsque le problème est mal posé, l'utilisation des ondelettes permet de réduire efficacement le nombre de paramètres. Grâce à leurs propriétés de compression, elles fournissent une bonne approximation des champs de propriétés avec peu de paramètres. À l'aide de la transformation en ondelettes développée dans ce travail, il est possible d'appliquer cette paramétrisation à des grilles complexes de réservoir utilisées dans les applications réelles. De plus, cette transformation prenant en compte les volumes des cellules, elle évite d'augmenter le nombre de paramètres actifs et permet de produire des approximations grossières non biaisées par des valeurs aberrantes que peuvent contenir des cellules fines.

Cependant, la paramétrisation en ondelettes s'est avérée mal adaptée aux simulations géostatistiques par échelles indépendantes¹. Afin de conserver les propriétés espace-fréquence, il a été proposé d'utiliser des pyramides gaussiennes. Les paramètres de chacune des échelles intégrant le contenu fréquentiel des échelles plus grossières, il est possible à chaque étape de modifier l'ensemble des composants déjà simulés.

¹lors de l'inversion des données de production, le nombre de paramètres s'accroît et les basses fréquences sont toujours incluses.

Effets de régularisation de l'approche multi-échelles Les exemples développés dans ce travail permettent d'illustrer les effets bénéfiques de l'approche multi-échelles. Dans le cas de la méthode multipoint, les avantages de la régularisation inter-échelles s'expriment à travers le conditionnement de la simulation des paramètres fins par les échelles grossières. Cet effet permet d'orienter de façon cohérente la simulation des valeurs, même lorsque le voisinage est réduit.

Dans le cas de l'inversion des données de production, et en particulier lors de l'utilisation de méthodes d'optimisation par ensemble, la paramétrisation multi-échelles permet de faire ressortir les paramètres sensibles du modèle. Ainsi, il est possible de perturber en priorité certains paramètres et préserver les composants moins sensibles, ce qui limite l'ajout de bruit et l'effet des corrélations fausses lors de la mise à jour des paramètres. De plus, l'utilisation de modèles grossiers rend le problème plus linéaire. Ainsi, l'optimisation des échelles grossières rapproche les modèles de l'état optimal tout en aidant à éviter les minima locaux. Après raffinement, lorsque l'optimisation intègre un plus grand nombre de paramètres, les modèles se trouvant plus proche du minimum, le comportement devient naturellement plus linéaire et les modifications nécessaires sont moindres. Les exemples développés dans ce travail montrent que l'inversion multi-échelles minimise la perturbation des modèles initiaux.

Respect des données et estimation des incertitudes La minimisation des perturbations des modèles initiaux aide à préserver une cohérence avec les données *a priori* et évite la dégradation de l'ensemble. De plus, l'approche multi-échelles est particulièrement bien adaptée lorsque les propriétés initiales sont conditionnées par des données sismiques. En effet, la réponse sismique n'étant pas affectée par les très basses fréquences, il est possible de les modifier lors du processus d'inversion des données de production sans impacter le calage initial des modèles. Toutefois, les relations entre les propriétés pétrophysiques et élastiques étant généralement non linéaires, il est difficile de prédire l'effet de la modification de certaines fréquences du modèle d'écoulement sur la réponse sismique.

La méthode MPS multi-échelles génère également des réalisations moins bruitées qui respectent mieux les données et les connaissances *a priori* que les approches standards. En effet, les simulations sont plus fidèles aux concepts géologiques portés par l'image d'entraînement et les données s'intègrent mieux dans les structures modélisées (chenaux). De plus, le contrôle multi-échelles du balayage de l'image d'entraînement permet d'utiliser des images beaucoup plus grandes, intégrant un plus grand nombre de configurations géologiques possibles.

Ainsi, dans les deux cas (inversion et simulation géostatistiques) les modèles créés par l'approche multi-échelles sont plus probables, car ils respectent mieux l'ensemble des données et des concepts utilisés, et échantillonnent mieux l'espace des possibles, grâce à l'obtention d'une plus grande variabilité pour l'ensemble des réalisations. L'estimation finale des incertitudes est donc de meilleure qualité que dans les approches standard.

6.2 Études complémentaires

En plus des perspectives données à la fin de chacun des chapitres, nous proposons les études complémentaires suivantes basées sur l'extension des outils et algorithmes développés dans ce travail.

Tests de sensibilités Les méthodes proposées impliquent un certain nombre de paramètres de contrôle. Afin d'analyser plus en détails le comportement des algorithmes, il serait nécessaire d'effectuer des tests de sensibilités. Basé sur les résultats de ces tests, il serait alors possible de proposer des réglages adaptés aux différents problèmes étudiés et d'améliorer l'efficacité des méthodes.

Application à des données réelles Bien que les tests synthétiques permettent une vérification théorique des algorithmes développés, il est nécessaire de les valider sur des données réelles. Cette étape nécessite un travail additionnel de nettoyage des données et d'analyse des paramètres sensibles qui peut s'avérer long et difficile. De plus, les modèles utilisés dans les cas réels ont généralement une taille importante et/ou un schéma de production complexe qui exige une puissance de calcul importante.

Généralisation de l'approche au calage de données 4D La sismique 4D est de plus en plus utilisée en surveillance des réservoirs. Ces données fournissent des informations sur les variations spatiales et temporelles de l'état d'un réservoir et notamment sur l'évolution des saturations. Lorsque l'huile produite est substituée par un autre fluide (eau ou gaz), le comportement élastique de la roche est modifié. Ainsi, en répétant les acquisitions sismiques et en les comparant les unes aux autres, il est possible de faire ressortir les zones affectées par la production. Les variations 4D sont fortement hétérogènes : elles dépendent de la proximité avec les puits et de la perméabilité des faciès. À l'aide la paramétrisation en ondelettes développées dans ce travail de thèse, il serait possible d'adapter la paramétrisation en fonction de l'intensité du signal observé. Dans les zones présentant des variations importantes, la paramétrisation pourrait être raffinée (cf. figure 6.1). Ceci permettrait d'éviter une sur-paramétrisation du problème et l'inversion de bruit.

Transformation fréquentielle des données Dans cette thèse, seules les propriétés spatiales sont décomposées en différentes gammes de fréquences. Il pourrait être également possible de décomposer les données afin de faire ressortir les variations importantes. De plus, la capacité de compression des ondelettes permet de réduire le nombre de points de données, ce qui peut améliorer l'efficacité des processus d'inversion [8]. Cette re-paramétrisation pourrait être particulièrement bien adaptée aux données 4D étant donné qu'elles expriment des variations relatives et sont spatialement corrélées. La transformation en ondelettes devra être effectuée en 4D. Ainsi, lorsque peu ou pas de changements 4D seront observés pour une position donnée, la majorité des coefficients auront une valeur faible et pourront être négligés. Cependant, dans certains cas, l'absence de variation peut être une information à part entière ; il faut donc être prudent dans le choix des coefficients utilisés.

Modèle de variabilité *a priori* non gaussien Lors des processus d'inversions présentés dans ce travail, les modèles *a priori* reposent sur une approxima-

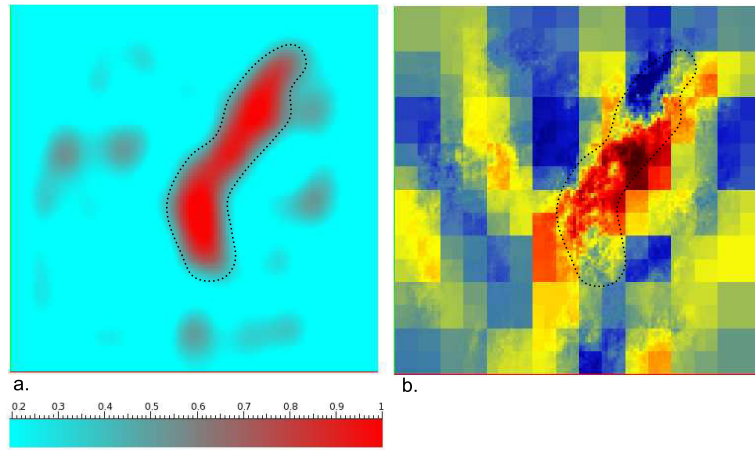


FIGURE 6.1 – Re-paramétrisation en fonction des amplitudes de variation d'impédances. (a) Carte d'amplitudes de variation d'impédances. (b) Propriété re-paramétrisée. Les ondelettes de Haar sont utilisées et les paramètres non inclus ont une valeur nulle afin de voir l'adaptation de la paramétrisation. La paramétrisation inclut uniquement des paramètres d'échelles fines dans les zones à fortes variations.

tion gaussienne de la variabilité spatiale. Cette hypothèse n'est cependant pas adaptée lorsque les réservoirs sont constitués de structures discontinues et peut conduire à la création de modèles peu réalistes. Dans ce cas, il peut être intéressant d'utiliser d'autres formes de régularisation *a priori*. Par exemple, il peut être utile de contraindre l'inversion en minimisant l'énergie des coefficients d'ondelettes afin de promouvoir l'émergence de formes discrètes décrites par un nombre réduit de paramètres [84]. Cette méthode basée sur les techniques d'acquisition comprimée (*compressed sensing* [43]) permet de trouver la solution la plus parcimonieuse d'un système sous déterminé et évite ainsi de générer des modèles bruités. Cependant, il peut être difficile d'adapter cette technique à une optimisation par ensemble, la contrainte de parcimonie devant être spécifique pour chacune des réalisations. En effet, cette contrainte tend à favoriser l'émergence de coefficients non négligeables alors qu'elle maintient une valeur proche de zéro pour les coefficients de faible amplitude. Or les différentes réalisations de l'ensemble présentant des champs de propriété différents, les valeurs des coefficients d'ondelettes varient d'une réalisation à l'autre. Ainsi, une contrainte globale ne peut pas être appliquée pour l'ensemble des réalisations. Il serait possible de calculer cette contrainte à partir de la moyenne des coefficients des réalisations. Cependant, ceci aurait tendance à diminuer fortement la variabilité de l'ensemble étant donné que l'émergence des mêmes coefficients serait favorisée pour toutes les réalisations.

D'autres techniques permettent de contraindre l'inversion par une image d'entraînement en suivant une approche MPS [25, 48, 89]. Il peut être alors possible d'adapter les approches multi-échelles présentées dans ce travail afin d'assimiler les données de production de manière cohérente avec un modèle conceptuel décrit par une image d'entraînement. L'approche multi-échelles pourrait permettre de rendre la fonction objectif plus continue que dans le cas de l'inversion en facies des valeurs des cellules de la grille (cas standard). En effet, l'approxi-

mation grossière d'un champ de propriété discrète est lisse, ce qui permettrait d'éviter des sauts de la fonction objectif (p. ex. deux segments connectés ou déconnectés par une seule cellule). Ainsi, les premières optimisations aux échelles grossières seraient proches du cas continu. Par la suite, lorsqu'un plus grand nombre de paramètres de hautes fréquences est inclus dans l'optimisation et que l'on se rapproche du cas discret (redéfinition des contrastes rapides de propriété), les optimisations aux échelles précédentes auront permis de s'approcher de l'optimum et l'amplitude des perturbations nécessaires seront moindre, ce qui pourrait permettre de limiter (localement) les sauts de la fonction objectif.

Transformées anisotropes Pour certaines propriétés, telles que les perméabilités, l'approximation grossière par une moyenne arithmétique n'est pas adaptée et/ou doit dépendre de l'orientation. Il serait possible d'adapter les ondelettes selon les directions de la grille. Par exemple, la décomposition 3D du champ de perméabilité alignée avec le premier axe de la grille (PERMX) pourrait utiliser une ondelette basée sur une moyenne harmonique dans cette direction (X) et des ondelettes basées sur une moyenne arithmétique pour les autres directions (Y et Z). Ainsi, l'approximation grossière de cette propriété serait plus représentative de l'échelle fine. Cependant, lorsque le tenseur de perméabilité n'est pas aligné avec les axes principaux de la grille, la transformation 3D par décompositions indépendantes pour chaque direction de la grille n'est plus adaptée. Il serait alors possible d'utiliser des ondelettes spécifiques 3D prenant en compte l'anisotropie du tenseur.

Appendices

Modèle approximation de production par déplétion naturelle

Cet article a été proposé lors du 32ème Gocad meeting en Septembre 2012.

Résumé

Les processus modélisation stochastique des réservoirs peuvent générer un grand nombre de réalisations pétrophysiques afin de caractériser les incertitudes liées aux données géologiques, sismiques et/ou de production de puits. Dans ce travail, nous développons un modèle de réponse approximatif de production par déplétion naturelle des réservoirs. Ce modèle est ensuite utilisé pour éliminer les réalisations incohérentes avec les données afin d'améliorer la quantification des incertitudes. Nous utilisons le concept de connectivité dynamique, que nous implémentons en 3D à l'aide de l'algorithme de *fast marching*, dont nous contrainsons le champ de vitesse en prenant en compte la conservation de la masse et la baisse de pression dans le réservoir. Cet algorithme permet de calculer de manière efficace, à partir de réalisations statiques, un volume approximatif de drainage des puits, ce dernier étant lié à la structure (locale) du réservoir. Nous montrons qu'il est possible d'obtenir une relation linéaire entre les volumes de drainage estimés et les données de production, à partir d'un nombre réduit de simulations (*full physics*) numériques des écoulements. Une estimation de la production primaire peut ensuite être obtenue pour l'ensemble des réalisations en calculant leurs volumes de drainage estimés et en appliquant la relation. À partir de ces estimations et des données, il est possible d'estimer la probabilité conditionnelle d'une réalisation par rapport aux données observées. Cette probabilité peut être utilisée dans des processus de sélections de modèles, de contrôle de qualité ou de quantifications des incertitudes.

Efficient Ranking of Stochastic Reservoir Models: A Dynamic Approach Suitable for Natural Depletion

Pierre Monamicq (University of Lorraine), Théophile Gentilhomme (CGG Veritas, CRPG-CNRS),
Guillaume Caumon (CRPG-CNRS, University of Lorraine)
September 2012

Abstract

Stochastic reservoir modeling can generate a huge amount of petrophysical realizations sampling the uncertainties associated with geological, seismic and well data. In this work, we develop a fast method to reduce the uncertainties by eliminating realizations inconsistent with early production data. Based on the work of Hird and Dubrule [1998b] and Hovadik and Larue [2007], we introduce a 3D dynamic connectivity using the fast marching algorithm. In particular, we propose specific strategies to account for material balance concepts and pressure decline during reservoir depletion in setting up the fast marching velocity. This algorithm allows a fast computation, on each static realization, of an estimated drained volume which characterizes the reservoir dynamic behavior in terms of mass balance. We show that a relationship can be obtained between these estimated drained volumes and observed produced volumes by simulating flow only on a few realizations. An estimation of primary recoveries can be obtained for all the other realizations by computing their estimated drained volume and applying the relationship. Using this estimation and observed data, the likelihood of each realization with respect to observed produced volumes can be assessed. This quantitative ranking of realizations allows a reduction of uncertainties and can also be used as quality control, for instance in the frame of petro-elastic seismic inversion.

Introduction

In reservoir modeling, the correct evaluation of geological uncertainties is one of the crucial points to perform history matching on meaningful parameters and ensure model predictivity. In this context, stochastic modeling methods such as stochastic seismic inversion are often used to generate a large number of possible models honoring the data at hand. However it is practically impossible to use each of these realizations in geomodeling and then reservoir studies, due to high computation time. Therefore only a few realizations are usually selected for making reservoir forecasts and development decisions.

Nevertheless, uncertainties in petrophysical modeling and then reservoir production forecasts cannot be correctly assessed if the entire space of uncertainties associated with the available knowledge is not explored. Moreover it will be very difficult to match production data if the chosen realization is too far from the real reservoir. In that case, history matching is not only long and arduous, but also may either make the model inconsistent with previous geological and geophysical interpretations or use inappropriate reservoir parameters in artificially forcing a match.

Early production data can be integrated into the reservoir description process in order to reduce uncertainties. This permits to increase the likelihood of using more appropriate models for further studies by ensuring the consistency with early production data. Several parameters can be used to account for production data into geomodeling processes. One of the most important static discriminators is original oil in place (*OOIP*) because it is very easy to obtain from a realization by integrating on all reservoir grid cells the product of the cell volume V , the net-to-gross NtG , the porosity ϕ , the water saturation Sw and the formation volume factor Bo :

$$STOOIP = \sum_{cells} (V * NtG * \phi * (1 - Sw) / Bo) \quad (1)$$

Produced volumes being highly dependant on the development plan (number of wells and well placement), there is no general relation between oil in place and produced volumes. Corre and Thore [2000] and Charles [2001] show that there is usually a strong linear relationship between original oil in place and cumulative recovery for a given field and a given development plan, but this observation only holds when considering volumes produced after a long time. Indeed, oil production is mainly governed by the pressure drawdown and the permeability of the reservoir whereas oil in place essentially depends on reservoir pressure and porosity. Consequently, for early production time, the relationship between *OOIP* and cumulative recoveries strongly depends on the relationship between porosity and permeability.

Another approach can be taken: estimate the *OOIP* in the reservoir from early production data. With sufficient production (in general a minimum of 10 to 20% of *OOIP*), material balance analysis permits to identify and reliably extrapolate a trend to the *OOIP* volume. Material balance analysis does not take into account fluid flow inside the reservoir but considers fluid and rock compressibility under the effect of pressure change with fluid injection and production. In practice this method could be very efficient for gas reservoirs: Kleppe [2011] shows that, in the case of a closed gas reservoir, initial volume of gas-in-place can be estimated by extrapolating the pressure curve even with a limited production history. In a more general case, the complete black oil material balance equation can be used along with ultimate recovery estimated by decline curve analysis. However, this method can only be used at reservoir scale if global connectivity is high. Otherwise, this method must be used on a connected zone of the reservoir around wells.

Hovadik and Larue [2007] defined a static or geometric connectivity called geobody connectivity as the probability of two reservoir cells belonging to the same permeable geobody. A permeable geobody is defined as a group of connected cells inside a region of a stratigraphic grid that has one or more rock property values that fall within given property ranges. It is based on the concept of flow unit from Hearn et al [1986], which is defined by the potential for lateral and vertical continuity of flow through a geological unit. These concepts cannot be used directly as selection criteria because they are defined at global scale. However they constitute the basis for advanced connectivity indicators such as reservoir-to-well connectivity or dynamic connectivity.

Reservoir-to-well connectivity has a very important influence on hydrocarbon recoveries because it is closely related to the drainage volume of wells. It is defined by Hovadik and Larue [2007] as the ratio of the permeable geobody volume intersected by wells to the total volume of geobodies. Because the computation of this connectivity can only approximate the drainage volume, the results must be calibrated against a reference value obtained by flow simulating one example realization. However, given that all hydrocarbons in a flow unit are not equally producible due to their distances from wells or high permeability contrasts, this kind of static method tends to overestimate the recovery. Therefore, this kind of criterion may be used for a relative ranking of realizations but does not seem appropriate for a direct correlation with production data. Thus a more advanced connectivity must be defined.

The method developed in this article is based on the dynamic connectivity introduced by Hird and Dubrule [1998a] and further used e.g. by Ballin et al. [2002], Hovadik and Larue [2007]. We will present the general method they developed, based on the computation of a resistivity index to quantify this dynamic connectivity. As proposed by Hovadik and Larue [2007], we extend the initial 2D resistivity index of Hird and Dubrule [1998b] using a fast marching method in three dimensions and introduce corrections to the fast marching velocity field to better reflect material balance during production (Section 2). This method is applied to several synthetic cases in Section 3. Finally, we also discuss a possible strategy to compute the likelihood of a stochastic model given the connectivity method.

1 Dynamic connectivity characterization

1.1 Definition of the resistivity index

Comparing criteria computed on geomodels with production data requires quantitatively characterizing the dynamic behavior of the geomodels. This could be achieved by computing estimated drained volumes on each geomodel. First the concept that allows a computation of a virtual drainage zone on geomodels is presented and then we will detail how to use this tool to discriminate realizations.

The dynamic connectivity as defined by Hird and Dubrule [1998a] adds the notion of time progression to classical connectivity parameters. It is based on the concept of connectivity in terms of least resistive paths, which is similar to the common minimum path length problem Cherkassky and Godberg [1993]. The idea is to calculate the shortest paths from a cell to all the other cells in a geobody in terms of resistivity or time. Thus we can estimate the hydrocarbon pore volume connected to a well and then link this volume to primary recovery by using a relevant cutoff.

Hird and Dubrule [1998b] defined a parameter called resistivity index on 2D gridded cross-sections in order to characterize the dynamic connectivity of the reservoir model. This index is defined for each cell (u, v, w) and computed by minimizing the weighted harmonic sum of permeabilities k from a source (well) cell to the cell under consideration:

$$RI_{(u,v,w)} = \min \left(\sum_{producer}^{(u,v,w)} \frac{\Delta L}{kA} \right), \quad (2)$$

where ΔL is the distance between two cells and A is the interface area between two cells. Following classical finite volume hypotheses, flow is supposed to occur only through face connections between reservoir grid blocks.

In this work, we apply Equation (2) without explicitly including relative permeabilities. This means that our approach is only physically valid when no significant change in fluid saturation occurs during reservoir depletion.

1.2 General principle of the method

The estimated drained volume of each realization is computed by using a resistivity index cutoff that stops the propagation of the front on the models [Hird and Dubrule, 1998b, Ballin et al., 2002]. This cutoff is an arbitrary limit that must be the same for each realization in order to compare the dynamic responses of these realizations. Even if the choice of the cutoff is arbitrary, it must be carefully defined in a certain range of values to avoid extreme cutoff values that will not permit to characterize the dynamic behavior of all realizations. For example if the cutoff is too high, the front will rapidly reach the limits of the geomodel for realizations with high connectivity. In practice, the cutoff can be determined using one realization so that the estimated drained volume is equal to the produced volume obtained by flow simulating the realization. Then this cutoff is used to compute the estimated drained volumes for all the other realizations:

$$Drained\ volume = \sum_{cells}^{t_f(cell) < t_{cutoff}} V_{oil}(cell), \quad (3)$$

where $t_f(cell)$ is the arrival time of the front at the given cell).

The main idea of the method is to use the link, provided by the concept of dynamic connectivity, between estimated drained volumes computed on geomodels and produced volumes obtained from flow simulations. Hird and Dubrule [1998a] show that a linear statistical relationship can be obtained between estimated volumes and actually produced volumes. This linear relationship is obtained by plotting the estimated drained volume computed on a few realizations versus the recoveries obtained by flow simulating these realizations. Our goal is therefore to make sure this linear correlation is

physically valid so as to reduce the number of flow simulations required to estimate the correlation coefficient to an ideal number of 1.

2 Improved characterization of the dynamic connectivity using the fast marching algorithm

The resistivity index (2) is computed as a cumulative weighted harmonic sum of permeabilities. As proposed by Hovadik and Larue [2007], we implemented a more realistic and faster approach for characterizing the dynamic connectivity of 3D models by using the fast marching algorithm introduced by Sethian [1996]. In this article the term front propagation time always refers to the time computed by the fast marching method on the static geomodels and production time refers to the depletion time of the real reservoir. Moreover we call estimated drained volume, the volume behind the propagation front, and produced volume or recovery the volume obtained from a complete flow simulation.

Reservoir connectivity can be characterized by solving the Eikonal equation, which describes the propagation of a front in a heterogeneous medium progressing outward from a source cell until an ending condition is met. The objective is to assign an arrival time to each cell of the grid. This arrival time can be related to a producing time for the cell under consideration. The Eikonal equation, which is a non-linear partial differential equation, is solved using the High Accuracy Fast Marching Method with a propagation velocity taken as a function of spatial position. This propagation velocity field is approximated by a speed function that needs to be carefully defined on each cell with the available data in order to represent how fluids flow through the model.

With this method, the front is more realistic than with the initial RI (figure 1) and the computation time is significantly reduced as compared to Eq. (2). This is very interesting for processing a large number of realizations.

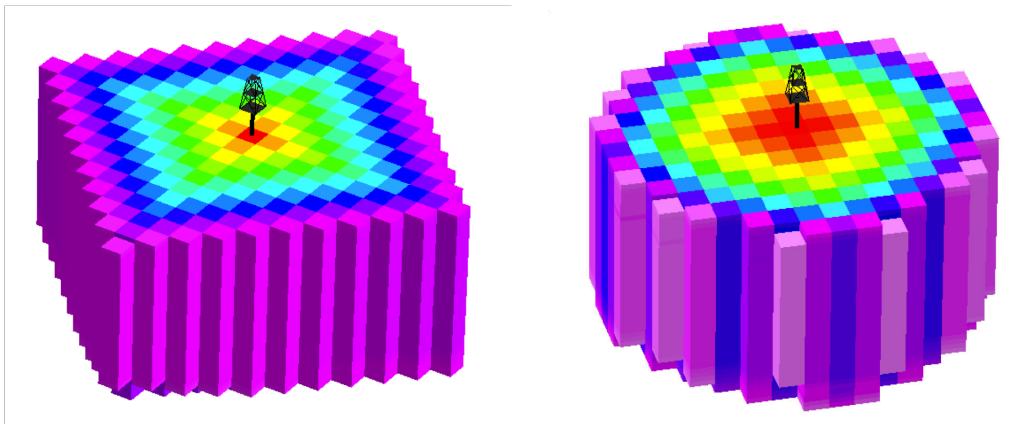


Figure 1: *Front propagation comparison between resistivity index method and fast marching method.*

However, in the case of natural depletion, two major drawbacks that affect the estimation of a drained volume appear with this method. The first problem is to ensure material balance when computing the arrival time of the propagating front. The second one is the non-consideration of the pressure decline induced by natural depletion, which causes a decrease of the oil flow rate that must be taken into account. The objective is to correct these effects in order to approach production mechanisms at best. Indeed, a necessary condition to get the good relationship between estimated drained volumes and recoveries is to have a computed volume profile similar to the actual production profile for each realization. Production mechanisms being well modeled, drained volumes of realizations will thereby differ only by the distribution of petrophysical properties.

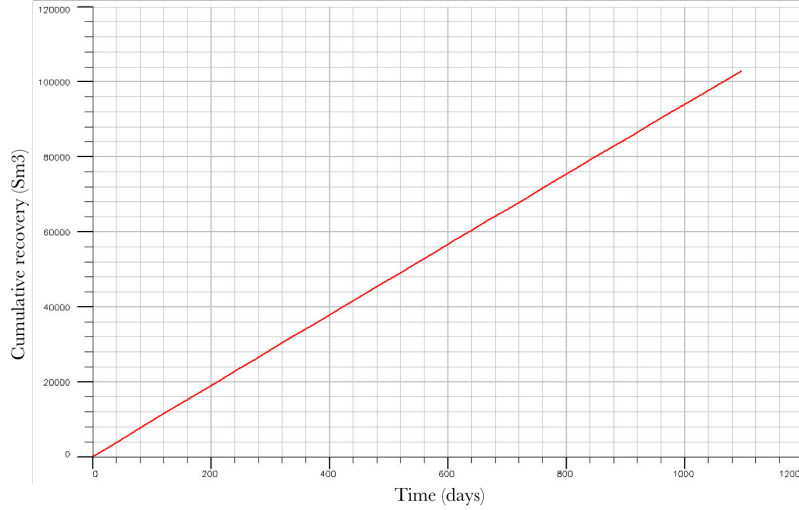


Figure 2: *Cumulative oil recovery versus production time for constant reservoir pressure on a homogeneous reservoir.*

2.1 Equivalence between arrival time and production time: Ensuring material balance

To use an equivalence between the arrival time of the front at a given cell and the producing date of this cell, a correction must be applied to the computation of the propagation front. Consider a homogeneous reservoir with constant reservoir pressure during production. In this case, the true cumulative recovery linearly increases with time (figure 2). With the fast marching algorithm only constrained by the permeability field as an approximation of the velocity field, the computed propagation time varies linearly with the distance (figure 3).

That means that the estimated drained volume (the volume behind the front) varies in a polynomial way with the propagation time (figure 4). Considering that the front can be assimilated to a closed surface, this increase will be cubic if the shape of the front is a sphere and quadratic if it is a cylinder. Therefore, a function of the distance to the source (well) is added to the expression of the velocity as a new constraint for the fast marching algorithm:

$$V = k * f(\text{distance}) \quad f(d) = d^i, \quad 0 < i < 3 \quad (4)$$

The shape of the front being dependent on the geometrical environment, this correction function is automatically modified, if needed, at each step of the front propagation by an optimization loop in order to adapt the propagation to the reservoir geometry and to respect the constraint of constant flow rate. This loop optimizes the parameter i in Eq. 4). For instance, when the front reaches a permeability barrier, the propagation stops in that direction, so it is necessary to modify the parameter i so as to ensure a constant flow rate. Using this correction, we are able to obtain a drained volume profile similar to the actual production profile (straight line) in the case of a reservoir with constant pressure (figure 5).

2.2 Accounting for depletion

Pressure usually decreases during reservoir production if no pressure support (aquifer, gas cap) is strong enough to maintain the reservoir pressure. This causes a decrease of the oil flow rate which depends on the distribution of petrophysical properties. Pressure decline is strongly linked to oil production: When only petrophysical distributions vary, the pressure profile as a function of produced volume is the same for all realizations (figure 6). We propose to account for this effect in the fast marching method at each step of the front propagation by ponderating the velocity of the front with a coefficient

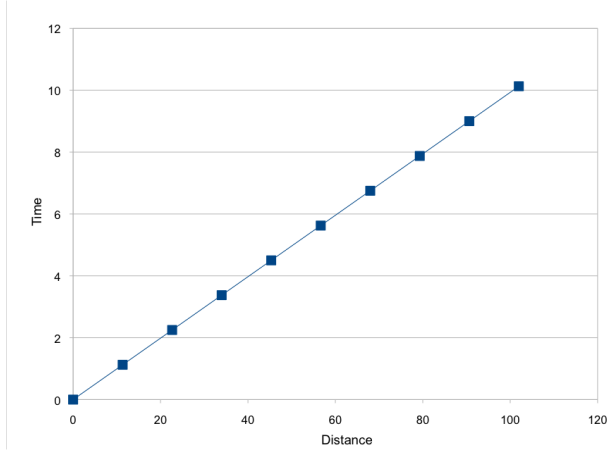


Figure 3: *Front propagation time computed by the fast marching method as a function of distance on a homogeneous reservoir.*

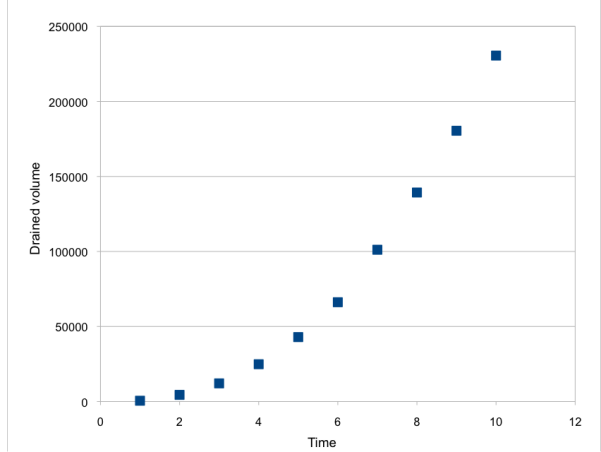


Figure 4: *Drained volume computed on the geomodel as a function of front propagation time. on a homogeneous reservoir*

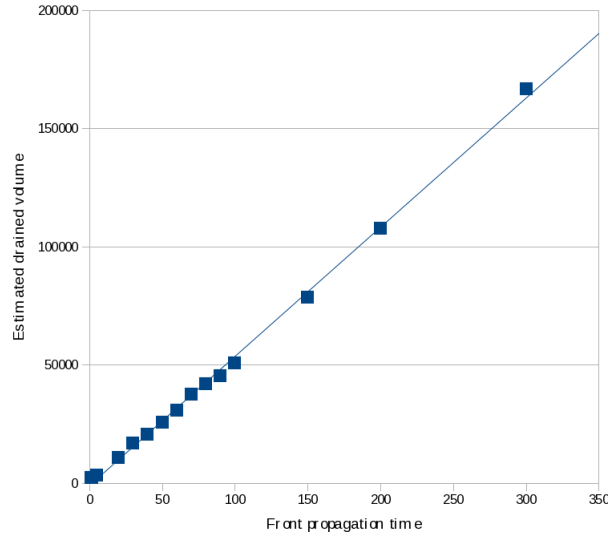


Figure 5: *Drained volume, computed with the corrected fast marching method, as a function of propagation time for one example realization.*

which depends on the drained volume. As a first correction for the synthetic cases presented below, we multiplied the velocity by the normalized drained volume at each step of the front computation.

3 Tests and discussions on synthetic cases

We considered three synthetic cases. The first simple case consists in 25 realizations of a homogeneous reservoir (constant permeability and porosity for each realization) with constant reservoir pressure. The reservoir is produced by natural depletion by one well. The propagation time cutoff is calibrated on an intermediate realization. Estimated drained volumes are then computed on each realization and plotted versus the actual produced volumes obtained by flow simulations. The production profile being well modeled thanks to the correction (figure 5), we obtain the linear relationship (figure 7).

Next we built a synthetic reservoir model of 10^5 cells, with one well in the middle of the reservoir. We created 45 different petrophysical realizations. We simulated permeabilities with sequential Gaussian simulations constrained to well. A wide range of different realizations is obtained by using

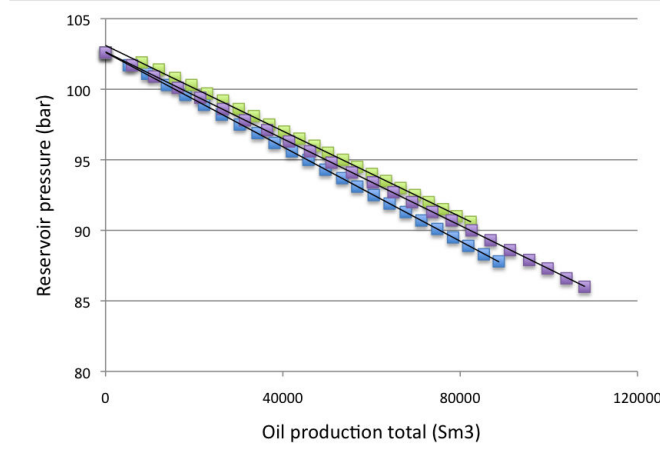


Figure 6: *Reservoir pressure as a function of recovery for three very different realizations.*

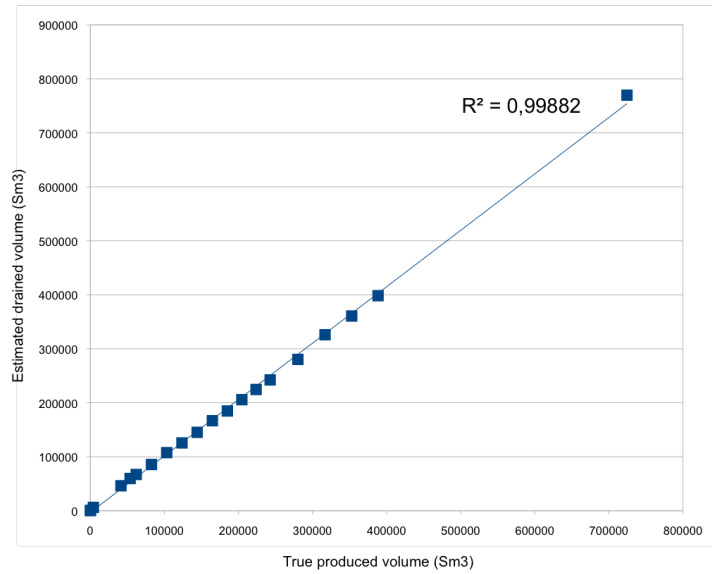


Figure 7: *Estimated drained volume, computed on several realizations of a homogeneous reservoir with constant pressure using the same time cutoff, plotted versus the actual produced volume obtained from flow simulation.*

very different variograms and distributions (figure 8). A logarithmic relationship (figure 5) is applied to model the porosity from the permeability.

$$\phi = \left(\frac{k}{2500 * 0.85^{-3.64}} \right)^{\frac{1}{3}} \quad (5)$$

The reservoir is produced during two years by natural depletion. A first case with constant reservoir pressure is studied in order to show the validity of the geometrical adaptative correction for a heterogeneous correction (figure 9 left). The reservoir pressure being not maintained in the second case, it permits to check the correction of the effect of pressure decline (figure 9 right). Using the same workflow as in the homogeneous case, we obtain the relation between drained volumes and produced volumes for both cases. However, while realizations with low spatial correlation and very high spatial correlation are well aligned, we observe that the effect of pressure decline is inversed for some intermediate realizations (circled blue dots significantly below the line). This underlines the difficulty to appropriately model the effect of pressure decline during the front propagation. In practice, if we need to deal with a very wide range of realizations, a better method for taking the pressure decline

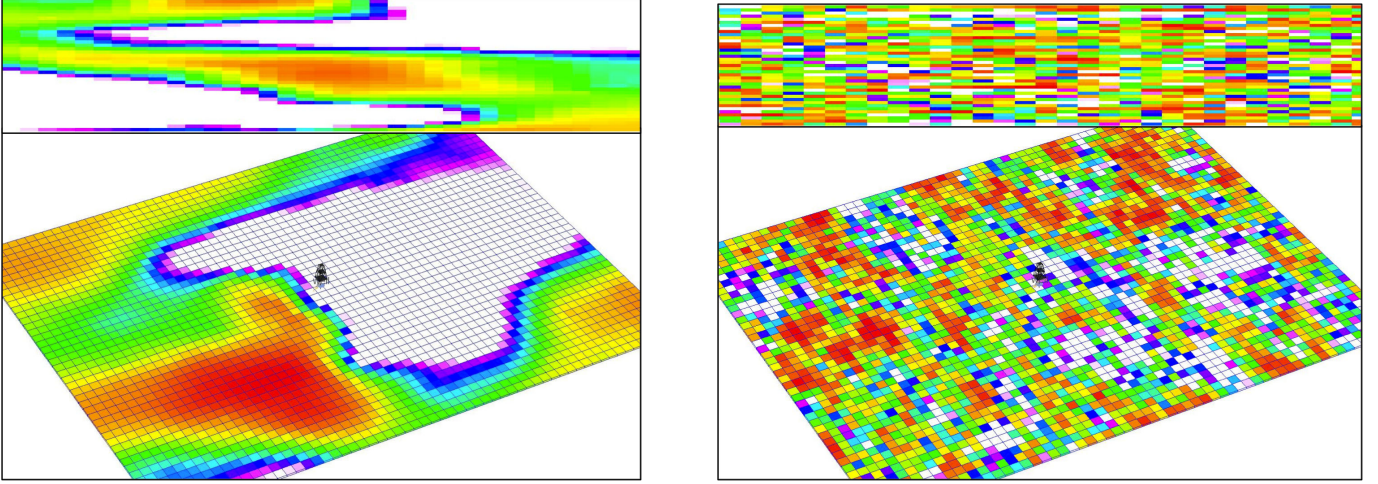


Figure 8: *Two example extreme realizations of a heterogeneous isotropic model obtained by SGS with two different variograms. Cross-section above the map view.*

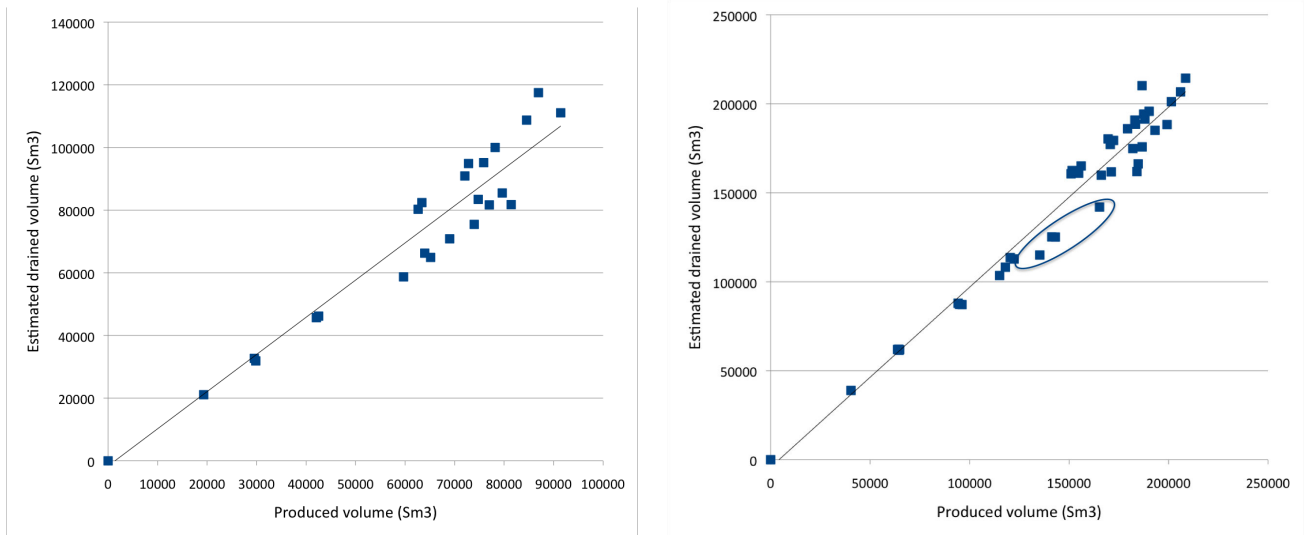


Figure 9: *Experimental linear relationship between produced and estimated drained volumes. Heterogeneous model with constant pressure on the left and with pressure decline on the right.*

into account must be developed. However, in most cases, realizations obtained from stochastic seismic inversion don't cover such a wide range and pressure decline is generally weak in early production. Thus an acceptable correction of the pressure decline can be achieved and used for all the realizations.

Finally, a more complex and bigger reservoir populated with meandering channels is used to test the validity of the method (figure 10). Channels are simulated using an object-based method. Sequential Gaussian simulations constrained to a well are used to model permeabilities to create 15 realizations. Porosity is modeled using the same logarithmic relationship as previously. Using the proposed workflow, we obtain a linear relationship between estimated and simulated drained volumes (figure 11).

4 Method for discriminating realizations using the dynamic connectivity and early production data

With the proposed corrections, the propagation of the front in the geomodel well represents the production mechanism and is able to characterize the dynamic behavior of each 3D realization in

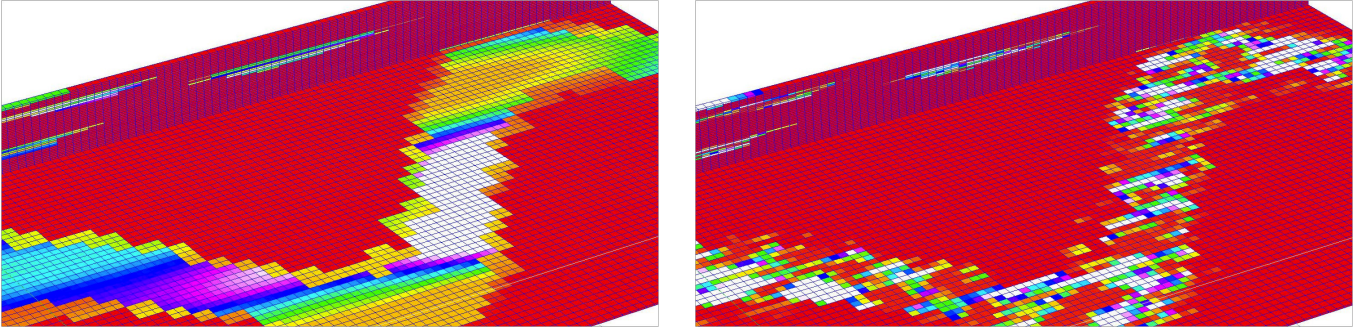


Figure 10: Two examples of petrophysical realizations of a reservoir model with meandering channels. High spatial correlation on the left. Lower spatial correlation on the right.

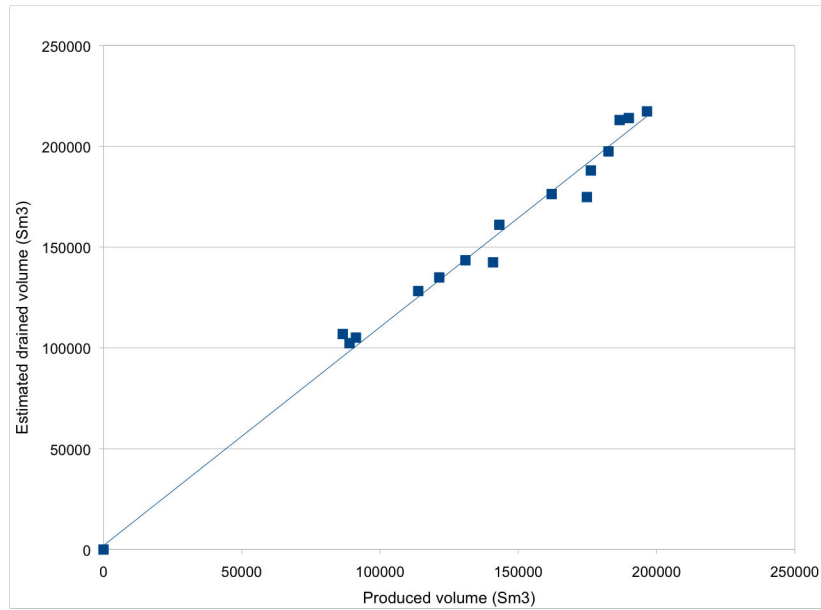


Figure 11: Linear relationship between estimated drained volumes computed on realizations of a channelized reservoir and produced volumes obtained from flow simulations.

natural depletion. The obtained linear relationship can be used to predict the cumulative recovery of realizations. However, we observe that the dispersion around the straight line is more important in the heterogeneous case than in the homogeneous case due to petrophysical heterogeneities, particularly for realizations with high spatial correlations. As a consequence, one of the most important questions is how many realizations need to be flow simulated to obtain a reliable relationship. We propose to use static reservoir-to-well connectivity, as defined by Hovadik and Larue [2007], in order to rank realizations and perform flow simulations on realizations sampled from this ranking. However, even if all realizations were simulated to get the best relationship, we would still have errors on the recovery prediction because of the dispersion. As a consequence, it is necessary to express the consistency of a realization with production data in terms of probabilities.

In order to discriminate realizations, a model likelihood based on production data can be attributed to each realization. This probability distribution function may be assumed to follow a Gaussian law centered on the measured produced volume. The standard deviation is given by the user taking into account the dispersion around the linear relationship as obtained from several full flow simulations and uncertainties on the measured produced volume. The proposed method is summarized in figure 12: Some realizations are simulated (blue dots) in order to evaluate the linear relationship. Then the recovery of a realization (145000 Sm³ in this example) can be estimated by computing the estimated drained volume (140,000) on this realization and applying the relationship. Finally the likelihood

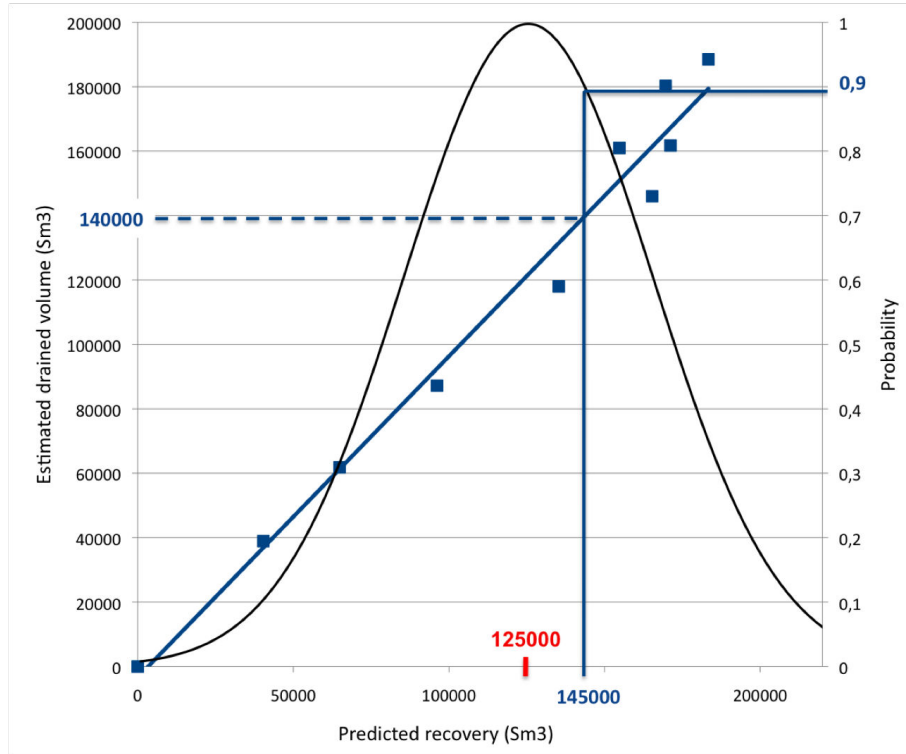


Figure 12: *Description of the method for associating a probability of consistency with observed recoveries for one example realization.*

probability may be attributed to the realization (0,9 in this example) using a Gaussian law centered on the true measured recovery (125,000 Sm³ in red).

In more complete uncertainty studies involving sensitivity analyses, this strategy could be used to reduce the uncertainty due to stochastic effects [Fetel and Caumon, 2008]. More generally, this likelihood probability could be used in inverse problem frameworks such as distance-based approaches [Scheidt and Caers, 2009] to efficiently solve history matching problems.

Conclusion

Based on the work of Hird and Dubrule [1998b] and Hovadik and Larue [2007], we developed a fast method for predicting early recoveries of a huge number of realizations from simple drained volumes computed on the geomodels. This method is based on the concept of dynamic connectivity which provides the link between estimated drained volumes and recoveries. Production is modeled by the propagation of a front from the producing well using the fast marching method. In cases where variations of saturation –hence relative permeability effects– can be neglected, a linear relationship is obtained between estimated drained volumes and produced volumes by flow simulating only a few realizations. This method allows discriminating realizations by probabilizing their consistency with early production data, which is very interesting in order to reduce uncertainties associated with the seismic data.

References

- P. Ballin, R. Solano, K. B. Hird, and R. F. Voltz. New reservoir dynamic connectivity measurement for efficient well placement strategy analysis under depletion. In *SPE 77375, SPE Annual Technical Conference and Exhibition.*, 2002.

- T. Charles. Experience with the quantification of subsurface uncertainties. *SPE 68703, SPE Asia Pacific Oil and Gas Conference.*, 2001.
- B. Cherkassky and A. Godberg. Shortest paths algorithms: Theory and experimental evaluation. *Mathematical Programming*, 73:129–174, 1993.
- B. Corre and P. Thore. Integrated uncertainty assessment for project evaluation and risk analysis. *SPE 65205, SPE European Petroleum Conference.*, 2000.
- E. Fetel and G. Caumon. Reservoir flow uncertainty assessment using response surface constrained by secondary information. *Journal of Petroleum Science and Engineering*, 60(3-4):170–182, 2008. doi: doi:10.1007/s10596-007-9063-9.
- K. Hird and O. Dubrule. Quantification of reservoir connectivity for reservoir description applications. *SPE Reservoir Evaluation and Engineering*, 1:12–17, 1998a.
- K. B. Hird and O. Dubrule. Quantification of reservoir connectivity for reservoir description applications. *SPE Reservoir Evaluation & Engineering*, 1(1):12–17, 1998b. doi: 10.2118/30571-PA.
- J. Hovadik and D. Larue. Static characterizations of reservoirs: refining the concepts of connectivity and continuity. *Petroleum Geoscience*, 13:195–211, 2007.
- J. Kleppe. Material balance equations. Technical report, Norwegian University of Science and Technology., 2011. URL <http://www.ipt.ntnu.no/~kleppe/TPG4150/matbal.pdf>.
- C. Scheidt and J. K. Caers. Representing spatial uncertainty using distances and kernels. *Mathematical Geosciences*, 41(4):397–419, 2009.
- J. Sethian. A fast marching level set method for monotonically advancing fronts. *Applied Mathematics*, 93:1591–1595, 1996.

Bibliographie

- [1] AANONSEN, S. I., NÆVDAL, G., OLIVER, D. S., REYNOLDS, A. C., AND VALLÈS, B. The ensemble Kalman filter in reservoir engineering—a review. *SPE Journal* 14, 3 (2009), 393–412. (Cité page 32)
- [2] ALABERT, F. G. Constraining description of randomly heterogeneous reservoirs to pressure test data : a monte carlo study. In *SPE Annual Technical Conference and Exhibition* (1989). (Cité page 32)
- [3] ANDERSON, J. L. Exploring the need for localization in ensemble data assimilation using a hierarchical ensemble filter. *Physica D : Nonlinear Phenomena* 230, 1–2 (2007), 99–111. (Cité page 37)
- [4] ANTOINE, J. P. *Two-dimensional wavelets and their relatives*. Cambridge University Press, 2004. (Cité page 63)
- [5] ARMIJO, L. Minimization of functions having lipschitz continuous first partial derivatives. *Pacific Journal of mathematics* 16, 1 (1966), 1–3. (Cité page 29)
- [6] ARPAT, G., AND CAERS, J. A multiple-scale, pattern-based approach to sequential simulation. In *Geostatistics Banff 2004*, O. Leuangthong and C. Deutsch, Eds., vol. 14 of *Quantitative Geology and Geostatistics*. Springer Netherlands, 2005, pp. 255–264. (Cité page 75)
- [7] AVSETH, P., MUKERJI, T., AND MAVKO, G. *Quantitative seismic interpretation : Applying rock physics tools to reduce interpretation risk*. Cambridge, 2010. (Cité pages 17, 143, 145, and 152)
- [8] AWOTUNDE, A. A., AND HORNE, R. N. A wavelet approach to adjoint state sensitivity computation for steady state differential equations. *Water Resources Research* 47, 3 (2011), W03502. (Cité page 165)
- [9] BEN AMEUR, H., CHAVENT, G., AND JAFFRÉ, J. Refinement and coarsening indicators for adaptive parametrization : application to the estimation of hydraulic transmissivities. *Inverse Problems* 18, 3 (2002), 775. (Cité page 41)
- [10] BEN AMEUR, H., CLÉMENT, F., WEIS, P., AND CHAVENT, G. The multidimensional refinement indicators algorithm for optimal parametrization. *J. Inverse Ill-Posed Probl.* 16, 2 (2008), 107–126. (Cité page 41)

- [11] BERNARD, C. Ondelettes et problèmes mal posés : la mesure du flot optique et l'interpolation irrégulière. *Mémoire de thèse, École Polytechnique, Palaiseau* (1999). (Cité pages 42 et 45)
- [12] BHARK, E., DATTA-GUPTA, A., AND JAFARPOUR, B. Subsurface flow model calibration with a spectral-domain parameterization adaptive to grid connectivity and prior model information. *Mathematical Geosciences* 44, 6 (2012), 673–710. (Cité page 41)
- [13] BHARK, E., JAFARPOUR, B., AND DATTA-GUPTA, A. An adaptively scaled frequency-domain parameterization for history matching. *Journal of Petroleum Science and Engineering* 75, 3–4 (2011), 289–303. (Cité pages 42 et 106)
- [14] BHARK, E., REY, A., DATTA-GUPTA, A., AND JAFARPOUR, B. A multiscale workflow for history matching in structured and unstructured grid geometries. *SPE Journal* 17, 3 (2012), 828–848. (Cité page 41)
- [15] BHARK, E. W., JAFARPOUR, B., AND DATTA-GUPTA, A. A generalized grid connectivity-based parameterization for subsurface flow model calibration. *Water Resources Research* 47, 6 (2011). (Cité page 42)
- [16] BIOT, M. A. Theory of elastic waves in a fluid-saturated porous solid. 2. higher frequency range. *J. Acoust. Soc. Am.* 28 (1956), 179–191. (Cité page 12)
- [17] BISSELL, R. Calculating optimal parameters for history matching. In *European Conference on the Mathematics of Oil Recovery* (1994). (Cité page 41)
- [18] BONNEAU, F., HENRION, V., CAUMON, G., RENARD, P., AND SAUSSE, J. A methodology for pseudo-genetic stochastic modeling of discrete fracture networks. *Computers & Geosciences* 56, 0 (2013), 12 – 22. (Cité page 2)
- [19] BORTFELD, R. Seismic waves in a mathematical model of the surface layer. *Geophysical Prospecting* 9, 3 (1961), 350–369. (Cité page 13)
- [20] BOURNARD, R. A mathematical description of total's sauternes/geosi algorithm – cgg internal document. CGG RDI London, June 2007. (Cité page 33)
- [21] BOYD, S. P., AND VANDENBERGHE, L. *Convex optimization*. Cambridge university press, 2004. (Cité page 26)
- [22] BULAND, A., KOLBJØRNSSEN, O., AND OMRE, H. Rapid spatially coupled AVO inversion in the fourier domain. *Geophysics* 68, 3 (2003), 824–836. (Cité page 33)
- [23] BULAND, A., AND OMRE, H. Bayesian linearized AVO inversion. *Geophysics* 68, 1 (2003), 185–198. (Cité pages 33 et 150)

- [24] BURRUS, C. S., GOPINATH, R. A., GUO, H., ODEGARD, J. E., AND SELESNICK, I. W. *Introduction to wavelets and wavelet transforms : a primer*, vol. 23. Prentice hall Upper Saddle River, 1998. (Cité page 50)
- [25] CAERS, J. History matching under training-image-based geological model constraints. *SPE journal* 8, 3 (2003), 218–226. (Cité page 166)
- [26] CAO, H. *Development of techniques for general purpose simulators*. PhD thesis, Stanford University, 2002. (Cité page 21)
- [27] CHATTERJEE, S., AND DIMITRAKOPOULOS, R. Multi-scale stochastic simulation with a wavelet-based approach. *Computers & Geosciences* 45, 0 (2012), 177–189. (Cité pages 77, 79, and 80)
- [28] CHATTERJEE, S., DIMITRAKOPOULOS, R., AND MUSTAPHA, H. Dimensional reduction of pattern-based simulation using wavelet analysis. *Mathematical Geosciences* 44, 3 (2012), 343–374. (Cité page 75)
- [29] CHAVENT, G. *Nonlinear least squares for inverse problems : theoretical foundations and step-by-step guide for applications*. Springer, 2009. (Cité pages 34 et 41)
- [30] CHEN, M.-H., SHAO, Q.-M., AND IBRAHIM, J. G. *Monte Carlo methods in Bayesian computation*. Springer New York, 2000. (Cité page 31)
- [31] CHEN, Y., AND OLIVER, D. S. Cross-covariances and localization for EnKF in multiphase flow data assimilation. *Computational Geosciences* 14 (2010), 579–601. (Cité pages 37, 116, and 134)
- [32] CHEN, Y., AND OLIVER, D. S. Ensemble-based closed-loop optimization applied to Brugge field. *SPE Reservoir Evaluation & Engineering* 13, 1 (2010), 56–71. (Cité pages 35, 37, and 130)
- [33] CHEN, Y., AND OLIVER, D. S. Ensemble randomized maximum likelihood method as an iterative ensemble smoother. *Mathematical Geosciences* 44, 1 (2012), 1–26. (Cité pages 35 et 109)
- [34] CHEN, Y., AND OLIVER, D. S. Multiscale parameterization with adaptive regularization for improved assimilation of nonlocal observation. *Water Resources Research* 48, 4 (2012). (Cité pages 37, 42, 106, 116, 117, and 130)
- [35] CHEN, Y., AND OLIVER, D. S. Levenberg-marquardt forms of the iterative ensemble smoother for efficient history matching and uncertainty quantification. *Computational Geosciences* 17, 4 (2013), 689–703. (Cité pages 35, 109, 110, 119, 120, 121, 130, and 136)
- [36] CLAYPOOLE, R. L., DAVIS, G. M., SWELDENS, W., AND BARANIUK, R. G. Nonlinear wavelet transforms for image coding via lifting. *Image Processing, IEEE Transactions on* 12, 12 (2003), 1449–1459. (Cité page 67)

- [37] COLLON-DROUAILLET, P., PELLERIN, J., AND BARDY, G. Modeling karstic networks : introduction of speleogenetic hypothesis in stochastic simulations. In *H2Karst - 9th Conference on Limestone Hydrogeology* (2011). (Cité page 2)
- [38] CONN, A. R., GOULD, N. I., AND TOINT, P. L. *Trust region methods*. No. 1. Siam, 2000. (Cité page 30)
- [39] CROUSE, M., NOWAK, R., AND BARANIUK, R. Wavelet-based statistical signal processing using hidden markov models. *Signal Processing, IEEE Transactions on* 46, 4 (Apr 1998), 886–902. (Cité page 88)
- [40] DENNIS, J. J. E., AND SCHNABEL, R. B. *Numerical methods for unconstrained optimization and nonlinear equations*, vol. 16. Siam, 1983. (Cité pages 30 et 31)
- [41] DESLAURIERS, G., AND DUBUC, S. Symmetric iterative interpolation processes. *Constructive Approximation* 5, 1 (1989), 49–68. (Cité page 59)
- [42] DEUTSCH, C., AND WANG, L. Hierarchical object-based stochastic modeling of fluvial reservoirs. *Mathematical Geology* 28, 7 (1996), 857–880. (Cité page 71)
- [43] DONOHO, D. Compressed sensing. *Information Theory, IEEE Transactions on* 52, 4 (April 2006), 1289–1306. (Cité page 166)
- [44] DONOHO, D. L. Smooth wavelet decompositions with blocky coefficient kernels. *Recent advances in wavelet analysis* 308 (1993). (Cité page 61)
- [45] DOYEN, P. *Seismic reservoir characterization : An earth modelling perspective*, vol. 2. EAGE publications, 2007. (Cité pages 33 et 152)
- [46] EFENDIEV, Y., HOU, T., AND LUO, W. Preconditioning Markov chain Monte Carlo simulations using coarse-scale models. *SIAM Journal on Scientific Computing* 28, 2 (2006), 776–803. (Cité page 41)
- [47] ENGELBRECHT, A. P. *Fundamentals of computational swarm intelligence*, vol. 1. Wiley Chichester, 2005. (Cité page 34)
- [48] ESKANDARIDALVAND, K., AND SRINIVASAN, S. Reservoir modelling of complex geological systems : A multiple-point perspective. *Journal of Canadian Petroleum Technology* 49, 8 (2010), 59–68. (Cité page 166)
- [49] EUVRARD, D. *Résolution numérique des équations aux dérivées partielles*, vol. 988. Masson, 1994. (Cité page 20)
- [50] EVENSEN, G. Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. *Journal of Geophysical Research : Oceans* 99, C5 (1994), 10143–10162. (Cité pages 32 et 35)

- [51] EVENSEN, G. The ensemble Kalman filter : theoretical formulation and practical implementation. *Ocean Dynamics* 53, 4 (2003), 343–367. (Cité page 32)
- [52] EVENSEN, G., AND VAN LEEUWEN, P. J. An ensemble Kalman smoother for nonlinear dynamics. *Monthly Weather Review* 128, 6 (2000), 1852–1867. (Cité pages 35 et 111)
- [53] FATTI, J. L., SMITH, G. C., VAIL, P. J., STRAUSS, P. J., AND LEVITT, P. R. Detection of gas in sandstone reservoirs using AVO analysis ; a 3-d seismic case history using the geostack technique. *Geophysics* 59, 9 (1994), 1362–1376. (Cité page 13)
- [54] FENG, T., AND MANNSETH, T. Improvements on a predictor-corrector strategy for parameter estimation with several data types. *Inverse Problems* 25, 10 (2009), 105012–105032. (Cité page 41)
- [55] FERNANDEZ, G., PERIASWAMY, S., AND SWELDENS, W. LIFTPACK : A software package for wavelet transforms using lifting. In *Proceedings of SPIE* (1996), vol. 2825, p. 396. (Cité pages 56 et 59)
- [56] FERZIGER, J. H., AND PERIĆ, M. *Computational methods for fluid dynamics*, vol. 3. Springer Berlin, 1996. (Cité page 20)
- [57] FRYKMAN, P. Geological model & fluvial environment with sand deposits, [http ://www.co2geonet.com/](http://www.co2geonet.com/), 2014. (Cité page 3)
- [58] FURRER, R., AND BENGTSSON, T. Estimation of high-dimensional prior and posterior covariance matrices in Kalman filter variants. *Journal of Multivariate Analysis* 98, 2 (2007), 227–255. (Cité pages 37, 115, and 116)
- [59] GABOR, D. Theory of communication. part 1 : The analysis of information. *Electrical Engineers - Part III : Radio and Communication Engineering, Journal of the Institution of* 93, 26 (1946), 429–441. (Cité page 45)
- [60] GAO, G., ZAFARI, M., AND REYNOLDS, A. Quantifying uncertainty for the punq-s3 problem in a bayesian setting with rml and enkf. *SPE J* 11, 4 (2006), 506–515. (Cité page 30)
- [61] GARDET, C., RAVALEC, M., AND GLOAGUEN, E. Multiscale parameterization of petrophysical properties for efficient history-matching. *Mathematical Geosciences* (2013), 1–22. (Cité page 41)
- [62] GASSMANN, F. Elastic waves through a packing of spheres. *Geophysics* 16, 4 (1951), 673–685. (Cité page 12)
- [63] GENTILHOMME, T., MANNSETH, T., OLIVER, D., CAUMON, G., AND MOYEN, R. Smooth multi-scale parameterization for integration of seismic and production data using second-generation wavelets. In *EC-MOR XIII-13th European Conference on the Mathematics of Oil Recovery* (2012). (Cité page 7)

- [64] GENTILHOMME, T., OLIVER, D. S., MANNSETH, T., MOYEN, R., CAUMON, G., AND DOYEN, P. Multi-scale ensemble-base history-matching using second generation wavelets. *Computational Geosciences* (in review). (Cité pages 7 et 105)
- [65] GENTILHOMME, T., AND THIERRY, C. Use of quality and anomaly indicators in the qc of seismic inversion. In *Proc. 31st Gocad Meeting, Nancy* (2011). (Cité pages 7 et 160)
- [66] GHOLAMI, V., AND MOHAGHEGH, S. Intelligent upscaling of static and dynamic reservoir properties. In *SPE annual technical conference and exhibition 2009* (2009). (Cité page 152)
- [67] GILKS, W. R., RICHARDSON, S., AND SPIEGELHALTER, D. J. *Markov chain Monte Carlo in practice*, vol. 2. CRC press, 1996. (Cité page 31)
- [68] GIRARDI, M., AND SWELDENS, W. A new class of unbalanced Haar wavelets that form an unconditional basis for L_p on general measure spaces. *Journal of Fourier Analysis and Applications* 3, 4 (1997), 457–474. (Cité page 61)
- [69] GLOAGUEN, E., AND DIMITRAKOPOULOS, R. Two-dimensional conditional simulation based on wavelet decomposition of training images. *Mathematical Geosciences* 41, 6 (2009), 679–701. (Cité page 88)
- [70] GLOAGUEN, E., AND DIMITRAKOPOULOS, R. Conditional wavelet based simulation of non-stationary geologic fields using geophysical and model analogue information. *Advances in Orebody Modelling and Strategic Mine Planning* (2010), 91–100. (Cité pages 75 et 79)
- [71] GOLUB, G. H., AND VAN LOAN, C. F. *Matrix Computations*, second ed. Johns Hopkins University Press, Baltimore, 1989. (Cité pages 28, 33, and 110)
- [72] GRIMSTAD, A. A., MANNSETH, T., NÆVDAL, G., AND URKEDAL, H. A. Adaptive multiscale permeability estimation. *Computational Geosciences* 7, 1 (2003), 1–25. (Cité page 41)
- [73] GROSSMANN, A., AND MORLET, J. Decomposition of hardy functions into square integrable wavelets of constant shape. *SIAM Journal on Mathematical Analysis* 15, 4 (1984), 723–736. (Cité pages 44 et 45)
- [74] GU, Y., AND OLIVER, D. S. An iterative ensemble Kalman filter for multiphase fluid flow data assimilation. *SPE Journal* 12, 4 (2007), 438–446. (Cité pages 36 et 109)
- [75] GUARDIANO, F., AND SRIVASTAVA, R. Multivariate geostatistics : Beyond bivariate moments. In *Geostatistics Tróia '92*, A. Soares, Ed., vol. 5 of *Quantitative Geology and Geostatistics*. Springer Netherlands, 1993, pp. 133–144. (Cité page 74)

- [76] HAMILL, T. M., WHITAKER, J. S., AND SNYDER, C. Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Monthly Weather Review* 129, 11 (2001), 2776–2790. (Cité page 37)
- [77] HASHIN, Z., AND SHTRIKMAN, S. A variational approach to the theory of the elastic behavior of multiphase materials. (Cité page 143)
- [78] HEEGER, D. J., AND BERGEN, J. R. Pyramid-based texture analysis/synthesis. In *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques* (1995), ACM, pp. 229–238. (Cité page 64)
- [79] HELLINGER, E. Neue begründung der theorie quadratischer formen von unendlichvielen veränderlichen. *Journal für die reine und angewandte Mathematik* 136 (1909), 210–271. (Cité page 96)
- [80] HOLDEN, L., AND NIELSEN, B. Global upscaling of permeability in heterogeneous reservoirs; the output least squares (ols) method. *Transport in Porous Media* 40, 2 (2000), 115–143. (Cité page 152)
- [81] HONARKHAH, M., AND CAERS, J. Stochastic simulation of patterns using distance-based pattern modeling. *Mathematical Geosciences* 42, 5 (2010), 487–517. (Cité page 75)
- [82] HOUTEKAMER, P. L., AND MITCHELL, H. L. A sequential ensemble Kalman filter for atmospheric data assimilation. *Monthly Weather Review* 129, 1 (2001), 123–137. (Cité page 37)
- [83] HU, L. Y., AND CHUGUNOVA, T. Multiple-point geostatistics for modeling subsurface heterogeneity : A comprehensive review. *Water Resources Research* 44, 11 (2008). (Cité pages 74 et 77)
- [84] JAFARPOUR, B., GOYAL, V., MCLAUGHLIN, D., AND FREEMAN, W. Compressed history-matching : exploiting transform-domain sparsity for regularization of nonlinear dynamic data integration problems. *Mathematical Geosciences* 42, 1 (2010), 1–27. (Cité page 166)
- [85] JAFARPOUR, B., AND MCLAUGHLIN, D. Reservoir characterization with the discrete cosine transform. *SPE Journal* 14, 1 (2009), 182–201. (Cité page 42)
- [86] JOURNEL, A., AND ALABERT, F. Non-gaussian data expansion in the earth sciences. *Terra Nova* 1, 2 (1989), 123–134. (Cité page 71)
- [87] LALLIER, F., CAUMON, G., BORGOMANO, J., VISEUR, S., FOURNIER, F., ANTOINE, C., AND GENTILHOMME, T. Relevance of the stochastic stratigraphic well correlation approach for the study of complex carbonate settings : application to the malampaya buildup (offshore palawan, philippines). *Geological Society, London, Special Publications* 370, 1 (2012), 265–275. (Cité page 7)

- [88] LI, L., AND JAFARPOUR, B. A sparse Bayesian framework for conditioning uncertain geologic models to nonlinear flow measurements. *Advances in Water Resources* 33, 9 (2010), 1024–1042. (Cité page 42)
- [89] LI, L., SRINIVASAN, S., ZHOU, H., AND GÓMEZ-HERNÁNDEZ, J. A pilot point guided pattern matching approach to integrate dynamic data into geological modeling. *Advances in Water Resources* 62, Part A, 0 (2013), 125–138. (Cité page 166)
- [90] LIBERTI, L., AND MACULAN, N. *Global Optimization : Volume 84, From Theory to Implementation*, vol. 84. Springer, 2006. (Cité page 26)
- [91] LIEN, M., BROUWER, D., MANNSETH, T., AND JANSEN, J.-D. Multiscale regularization of flooding optimization for smart field management. *SPE Journal* 13, 2 (2008), 195–204. (Cité page 41)
- [92] LIU, J. A multiresolution method for distributed parameter estimation. *Scientific Computing* 14, 2 (1993), 389–405. (Cité page 41)
- [93] LIU, Y., AND JOURNEL, A. Improving sequential simulation with a structured path guided by information content. *Mathematical Geology* 36, 8 (2004), 945–964. (Cité page 76)
- [94] LOPEZ, S., COJAN, I., RIVOIRARD, J., AND GALLI, A. *Process-Based Stochastic Modelling : Meandering Channelized Reservoirs*. Wiley-Blackwell, 2009, pp. 139–144. (Cité page 71)
- [95] LU, P., AND HORNE, R. A multiresolution approach to reservoir parameter estimation using wavelet analysis. In *SPE annual technical conference and exhibition* (2000). (Cité pages 41 et 42)
- [96] MALLAT, S. G. A theory for multiresolution signal decomposition : the wavelet representation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 11, 7 (1989), 674–693. (Cité page 46)
- [97] MANNSETH, T. Permeability identification from pressure observations : Some foundations for multiscale regularization. *Multiscale Modeling & Simulation* 5, 1 (2006), 21–44. (Cité page 41)
- [98] MARI, J.-L., COPPENS, F., AND GLANGEAUD, F. *Traitement du signal pour géologues et géophysiciens*. Editions Technip, 1997. (Cité page 15)
- [99] MARIETHOZ, G., AND LEFEBVRE, S. Bridges between multiple-point geostatistics and texture synthesis : Review and guidelines for future research. *Computers & Geosciences*, 0 (2014). (Cité page 74)
- [100] MARIETHOZ, G., RENARD, P., AND STRAUBHAAR, J. The direct sampling method to perform multiple-point geostatistical simulations. *Water Resources Research* 46, 11 (2010). (Cité pages xi, 75, 76, 77, 79, and 85)
- [101] MATHERON, G. *Traité de géostatistique appliquée. 1 (1962)*, vol. 1. Editions Technip, 1962. (Cité page 72)

- [102] MERLAND, R., LEVY, B., AND CAUMON, G. Voronoi grids conformal to 3D structural features. In *13th European Conference on the Mathematics of Oil Recovery (ECMOR)* (Biarritz, France, September 2012). (Cité page 11)
- [103] MONAMICQ, P., GENTILHOMME, T., AND CAUMON, G. Efficient ranking of stochastic reservoir models : A dynamic approach suitable for natural depletion. In *Proc. 32nd Gocad Meeting, Nancy* (2012). (Cité page 7)
- [104] OLIVER, D. S., AND CHEN, Y. Recent progress on reservoir history matching : a review. *Computational Geosciences* 15, 1 (2011), 185–221. (Cité page 41)
- [105] OLIVER, D. S., REYNOLDS, A. C., AND LIU, N. *Inverse theory for petroleum reservoir characterization and history matching*. Cambridge University Press, 2008. (Cité pages 26, 30, 31, 32, 33, 34, and 106)
- [106] PEACEMAN, D. W. *Fundamentals of Numerical Reservoir Simulation*. Elsevier Science Inc., 1991. (Cité page 20)
- [107] PELLERIN, J., LEVY, B., CAUMON, G., AND BOTELLA, A. Automatic surface remeshing of 3D structural models at specified resolution : a method based on voronoi diagrams. *Computers & Geosciences* 62 (2013), 103–116. (Cité page 10)
- [108] PETERS, L., ARTS, R., BROUWER, G., GEEL, C., CULICK, S., LORENTZEN, R., CHEN, Y., DUNLOP, N., VOSSEPOEL, F., XU, R., ET AL. Results of the Brugge benchmark study for flooding optimization and history matching. *SPE Reservoir Evaluation & Engineering* 13, 3 (2010), 391–405. (Cité pages 121, 125, and 130)
- [109] POWELL, M. J. *A direct search optimization method that models the objective and constraint functions by linear interpolation*. Springer, 1994. (Cité page 26)
- [110] RENARD, P., AND DE MARSILY, G. Calculating equivalent permeability : a review. *Advances in Water Resources* 20, 5-6 (1997), 253–278. (Cité page 152)
- [111] REYNOLDS, A. C., HE, N., CHU, L., AND OLIVER, D. S. Reparameterization techniques for generating reservoir descriptions conditioned to variograms and well-test pressure data. *SPE Journal* 1, 4 (1996), 413–426. (Cité page 41)
- [112] RICHARDS, P. G., AND AKI, K. *Quantitative seismology : Theory and methods*. Freeman, 1980. (Cité page 13)
- [113] RICHARDS, P. G., AND FRASIER, C. W. Scattering of elastic waves from depth-dependent inhomogeneities. *Geophysics* 41, 3 (1976), 441–458. (Cité page 13)

- [114] RODRIGEZ, J. Reparameterization techniques for generating reservoir descriptions conditioned to variograms and well-test pressure data. *Computational Geosciences* 10, 1 (2006), 119–136. (Cité page 41)
- [115] ROMANENKO, Z., GENTILHOMME, T., AND CAUMON, G. Multi-scale multi-point simulation using wavelet transform. In *33st Gocad Meeting, Nancy, France* (2013). (Cité pages 7 et 79)
- [116] ROMERO, C., CARTER, J., GRINGARTEN, A., AND ZIMMERMAN, R. A modified genetic algorithm for reservoir characterisation. In *International Oil and Gas Conference and Exhibition in China* (2000). (Cité page 32)
- [117] RONGIER, G. personal communication. (Cité page 96)
- [118] RONGIER, G., COLLON-DROUAILLET, P., RENARD, P., STRAUBHAAR, J., AND SAUSSE, J. Reproduction assessment of connected geobodies in multiple-point simulation. In *Proc. 33rd Gocad Meeting* ("sep" 2013). (Cité pages 96 et 99)
- [119] SAHNI, I., AND HORNE, R. Multiresolution wavelet analysis for improved reservoir description. *SPE Reservoir Evaluation & Engineering* 8, 1 (2005), 53–69. (Cité pages 42 et 106)
- [120] SEN, M. K., DATTA-GUPTA, A., STOFFA, P., LAKE, L., AND POPE, G. Stochastic reservoir modeling using simulated annealing and genetic algorithm. *SPE Formation Evaluation* 10, 1 (1995), 49–56. (Cité page 32)
- [121] SHAH, P., GAVALAS, G., AND SEINFELD, J. Error analysis in history matching : the optimum level of parameterization. *SPE Journal* 18, 6 (1978), 219–228. (Cité page 41)
- [122] SHRESTHA, R. K., AND BOECKMANN, M. K. Stochastic seismic inversion for reservoir modeling. In *2002 SEG Annual Meeting* (2002). (Cité page 33)
- [123] STOLLNITZ, E. J., DEROSE, A. D., AND SALESIN, D. H. Wavelets for computer graphics : a primer, part 1. *Computer Graphics and Applications, IEEE* 15, 3 (1995), 76–84. (Cité page 46)
- [124] STOLLNITZ, E. J., DEROSE, A. D., AND SALESIN, D. H. Wavelets for computer graphics : a primer, part 2. *Computer Graphics and Applications, IEEE* 15, 4 (1995), 75–85. (Cité pages 46 et 48)
- [125] STOLT, R., AND WEGLEIN, A. Migration and inversion of seismic data. *Geophysics* 50, 12 (1985), 2458–2472. (Cité page 16)
- [126] STREBELLE, S. Conditional simulation of complex geological structures using multiple-point statistics. *Mathematical Geology* 34, 1 (2002), 1–21. (Cité pages 74 et 80)
- [127] SUZUKI, S., AND STREBELLE, S. Real-time post-processing method to enhance multiple-point statistics simulation. *Petroleum geostatistics* (2007), 10–14. (Cité page 77)

- [128] SWELDENS, W. The lifting scheme : A construction of second generation wavelets. *SIAM J. Math. Anal.* 29, 2 (Mar. 1998), 511–546. (Cité pages 51, 53, and 54)
- [129] SWELDENS, W., AND SCHRÖDER, P. Building your own wavelets at home. In *Wavelets in the Geosciences*, R. Klees and R. Haagmans, Eds., vol. 90 of *Lecture Notes in Earth Sciences*. Springer Berlin Heidelberg, 2000, pp. 72–107. (Cité page 63)
- [130] TAHMASEBI, P., HEZARKHANI, A., AND SAHIMI, M. Multiple-point geostatistical modeling based on the cross-correlation functions. *Computational Geosciences* 16, 3 (2012), 779–797. (Cité page 75)
- [131] TARANTOLA, A. *Inverse problem theory : Methods for data fitting and model parameter estimation*. Elsevier, Amsterdam, The Netherlands, 1987. (Cité pages 4, 23, 24, 25, and 40)
- [132] TILLIER, E., DA VEIGA, S., AND DERFOUL, R. Appropriate formulation of the objective function for the history matching of seismic attributes. *Comput. Geosci.* 51 (Feb. 2013), 64–73. (Cité page 121)
- [133] TILLIER, E., LE RAVALEC, M., AND ROGGERO, F. Using time domain seismic data for history matching processes. *First Break* 29, 3 (2011), 47–53. (Cité page 157)
- [134] TRAN, T. T. Improving variogram reproduction on dense simulation grids. *Computers & Geosciences* 20, 7–8 (1994), 1161 –1168. (Cité page 76)
- [135] VOGEL, C. R., AND WADE, J. G. Iterative svd-based methods for ill-posed problems. *SIAM J. Sci. Comput.* 15, 3 (May 1994), 736–754. (Cité page 41)
- [136] WACKERNAGEL, H. *Multivariate geostatistics*. Springer, 2003. (Cité pages 71 et 72)
- [137] WOLFE, P. Convergence conditions for ascent methods. *Siam Review* 11, 2 (1969), 226–235. (Cité page 29)
- [138] XIONG, Y., AND TURKOWSKI, K. Registration, calibration and blending in creating high quality panoramas. In *Applications of Computer Vision, 1998. WACV '98. Proceedings., Fourth IEEE Workshop on* (1998), pp. 69–74. (Cité page 64)
- [139] YILMAZ, Ö. *Seismic data analysis*, vol. 1. Society of Exploration Geophysicists Tulsa, 2001. (Cité page 15)
- [140] ZHANG, F., REYNOLDS, A., AND OLIVER, D. Evaluation of the reduction in uncertainty obtained by conditioning a 3d stochastic channel to multiwell pressure data. *Mathematical Geology* 34, 6 (2002), 715–742. (Cité page 41)

- [141] ZHANG, T., SWITZER, P., AND JOURNEL, A. Filter-based classification of training image patterns for spatial simulation. *Mathematical Geology* 38, 1 (2006), 63–80. (Cité page 75)
- [142] ZHANG, Y., AND OLIVER, D. S. Improving the ensemble estimate of the Kalman gain by bootstrap sampling. *Mathematical Geosciences* 42, 3 (2010), 327–345. (Cité page 37)
- [143] ZOEPPRITZ, K. Erdbebenwellen vii. viib. Über reflexion und durchgang seismischer wellen durch unstetigkeitsflächen. nachrichten der königlichen gesellschaft der wissenschaften zu göttingen. *Mathematisch-physikalische Klasse* (1919), 66–84. (Cité page 13)
- [144] ZUBAREV, D. Pros and cons of applying proxy-models as a substitute for full reservoir simulations. In *SPE Annual Technical Conference and Exhibition* (2009). (Cité page 22)

Intégration multi-échelles des données de réservoir et quantification des incertitudes

Dans ce travail, nous proposons de suivre une approche multi-échelles pour simuler des propriétés spatiales des réservoirs, permettant d'intégrer des données directes (observation de puits) ou indirectes (sismique et données de production) de résolutions différentes. Deux paramétrisations sont utilisées pour résoudre ce problème : les ondelettes et les pyramides gaussiennes. À l'aide de ces paramétrisations, nous démontrons les avantages de l'approche multi-échelles sur deux types de problèmes d'estimations des incertitudes basés sur la minimisation d'une distance.

Le premier problème traite de la simulation de propriétés à partir d'un algorithme de géostatistiques multipoints. Il est montré que l'approche multi-échelles basée sur les pyramides gaussiennes améliore la qualité des réalisations générées, respecte davantage les données et réduit les temps de calculs par rapport à l'approche standard.

Le second problème traite de la préservation des modèles *a priori* lors de l'assimilation des données d'historique de production. Pour re-paramétriser le problème, nous développons une transformée en ondelette 3D applicable à des grilles stratigraphiques complexes de réservoir, possédant des cellules mortes ou de volume négligeable. Afin d'estimer les incertitudes liées à l'aspect mal posé du problème inverse, une méthode d'optimisation basée ensemble est intégrée dans l'approche multi-échelles de calage historique. À l'aide de plusieurs exemples d'applications, nous montrons que l'inversion multi-échelles permet de mieux préserver les modèles *a priori* et est moins assujettie au bruit que les approches standards, tout en respectant aussi bien les données de conditionnement.

Mots clés : Caractérisation des réservoirs, Multi-échelles, Re-paramétrisation, Schéma de *lifting*, Ondelettes de seconde génération, Géostatistiques multipoints, Problèmes inverses, Calage historique, Optimisation, Sismique, Données de production

Multiscale reservoir data integration and uncertainty quantification

In this work, we propose to follow a multiscale approach for spatial reservoir properties characterization using direct (well observations) and indirect (seismic and production history) data at different resolutions. Two decompositions are used to parameterize the problem : the wavelets and the Gaussian pyramids. Using these parameterizations, we show the advantages of the multiscale approach with two uncertainty quantification problems based on minimization.

The first one concerns the simulation of property fields from a multiple points geostatistics algorithm. It is shown that the multiscale approach based on Gaussian pyramids improves the quality of the output realizations, the match of the conditioning data and the computational time compared to the standard approach.

The second problem concerns the preservation of the prior models during the assimilation of the production history. In order to re-parameterize the problem, we develop a new 3D grid adaptive wavelet transform, which can be used on complex reservoir grids containing dead or zero volume cells. An ensemble-based optimization method is integrated in the multiscale history matching approach, so that an estimation of the uncertainty is obtained at the end of the optimization. This method is applied on several application examples where we observe that the final realizations better preserve the spatial distribution of the prior models and are less noisy than the realizations updated using a standard approach, while matching the production data equally well.

Keywords : Reservoir characterization, Multiscale, Reparameterization, Lifting scheme, Second generation wavelets, Multiple points geostatistics, Inverse problems, History-Matching, Optimization, Seismic, Production data