Incertitudes structurales en géomodélisation : échantillonnage et approche inverse

Nicolas Cherpeau

▶ **To cite this version:**

Nicolas Cherpeau. Incertitudes structurales en géomodélisation : échantillonnage et approche inverse. Sciences de la Terre. Université de Lorraine, 2012. Français. NNT : 2012LORR0141 . tel-01749341

HAL Id: tel-01749341
https://hal.univ-lorraine.fr/tel-01749341

Submitted on 29 Mar 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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

Université de Lorraine **École doctorale RP2E**

THÈSE

présentée et soutenue publiquement le 4 avril 2012
pour l'obtention du grade de

Docteur de l'Université de Lorraine
Spécialité Géosciences

par

Nicolas CHERPEAU

Incertitudes structurales en géomodélisation : échantillonnage et approche inverse

Directeur de thèse :	Bruno Lévy
Co-directeur de thèse :	Guillaume Caumon

Composition du jury :

Rapporteurs :	Jean-Pierre Gratier
	Klaus Mosegaard
Examinateurs :	Mary Ford
	Ghislain de Marsily
	Benoît Noetinger
Invité :	Jef Caers

Centre de Recherches Pétrographiques et Géochimiques
ENSG, rue du Doyen Marcel Roubault - 54500 Vandoeuvre les Nancy

Remerciements

Mes premiers remerciements vont bien sûr à Guillaume, pour m'avoir accompagné tout au long de ces 3 ans. Je souhaite à tout thésard d'avoir un encadrant de cette qualité, tant pour le côté scientifique qu'humain. Merci Guillaume !

Je remercie également Bruno, bien que l'encadrement fût plus distant, il n'en demeura pas moins riche et agréable. Je remercie aussi Jef, pour sa collaboration et son accueil à Stanford. Je souhaite également remercier Jean-Jacques, notamment pour ses relectures rapides et efficaces du manuscrit.

Je tiens bien entendu à remercier tous les membres du jury, Klaus Mosegaard et Jean-Pierre Gratier pour leur rôle de rapporteurs ; Mary Ford, Ghislain de Marsily et Benoît Noetinger pour avoir été examinateurs.

Dans le quotidien du thésard, les heures passées au labo, en dehors de l'écran, sont certes nombreuses (lesquelles ?) et ne se retrouvent pas directement dans le manuscrit mais y participent pourtant, indiscutablement. Celles-ci furent rendues bien agréables grâce à l'ensemble de l'équipe, formant une troupe fort sympathique et joyeuse. Je remercie donc (honneur aux anciens) : Marco, Lise, Thomas, Vincent, Pauline(s), Gautier, Romain, Jeanne, Jérémy, Fifi, Charline pour les diverses pauses récréatives (je ne m'attarderai pas sur leur contenu, j'en oublierais).

Je tiens également à remercier le reste de l'équipe : Fatima, Pierre et Christophe pour leur gentillesse, patience et aide précieuse.

Dans le quotidien du thésard, les heures passées en dehors du labo, sont également nombreuses et ne se retrouvent pas non plus dans le manuscrit, mais c'est peut-être mieux ainsi. De ce côté, je tiens à remercier mes collocs (Bid et Flo) et autres crpgiens (Julien bien évidemment) pour la chouette vie nancéenne.

Enfin, je souhaite remercier ma famille de m'avoir permis d'arriver jusqu'ici, et Lise d'être là, tout simplement.

Table des matières

Table des matières v

Table des figures vii

Introduction 1

1 Modèle stochastique de failles 13
1.1 Qu'est-ce qu'une faille ? . . . 14
1.1.1 Une faille n'est pas toujours une simple surface de glissement . . . 14
1.1.2 Les failles en géomodélisation . . . 15
1.2 La modélisation implicite dans le cadre des incertitudes structurales . . . 17
1.2.1 Définition d'une surface implicite . . . 17
1.2.2 Avantages de la modélisation implicite . . . 17
1.3 Les arbres binaires comme représentation de la topologie d'un réseau de failles 19
1.3.1 Une surface implicite comme partition de l'espace . . . 19
1.3.2 Représentation graphique des relations de filiation entre failles . . . 20
1.3.3 Propriétés des arbres binaires . . . 21
1.4 Description d'un réseau de failles . . . 22
1.4.1 Familles de failles . . . 22
1.4.2 Mise en place de familles de failles . . . 23
1.4.3 Relations entre familles de failles . . . 23
1.5 Algorithme général et données d'entrée . . . 25
1.5.1 Algorithme général . . . 25
1.5.2 Paramètres d'une famille de failles . . . 26
1.5.3 Gestion des données . . . 28
1.5.4 Critères d'arrêt de la simulation . . . 29
1.5.5 Gestion des familles cogénétiques . . . 30
1.6 Modélisation d'un objet faille . . . 32
1.6.1 Entrées . . . 32
1.6.2 Aperçu des étapes de simulation d'un objet faille . . . 32
1.6.3 Attribution des points de données . . . 32
1.6.4 Orientation de la future faille . . . 34
1.6.5 Centre de faille . . . 34
1.6.6 Géométrie initiale . . . 38
1.6.7 Regroupement d'ensembles de points . . . 39
1.6.8 Géométrie finale . . . 39
1.6.9 Extension latérale . . . 41
1.6.10 Choix de l'ordre des étapes de modélisation d'une faille . . . 43
1.6.11 Mise à jour du modèle . . . 45
1.7 Perspectives . . . 50
1.7.1 Concepts géologiques supplémentaires . . . 50
1.7.2 Intégration de données supplémentaires . . . 53

1.7.3 Simulation d'un objet faille modulable 56
1.8 Conclusions 58

2 Modélisation structurale stochastique 59
2.1 Introduction 61
2.2 Stochastic structural modelling method 62
2.2.1 Fault characterization from interpreted data 62
2.2.2 Simulation overview 64
2.2.3 Fault simulation 65
2.2.4 Computing the stratigraphy 72
2.2.5 Simulation choices 75
2.3 Application to a Middle East field case 75
2.3.1 Input data 75
2.3.2 Fault network characterization 75
2.3.3 Stratigraphic uncertainty 77
2.3.4 Results 77
2.4 Conclusions 78

3 Prise en compte des structures géologiques lors de la résolution de problèmes inverses 83
3.1 Discrétisation d'un réseau de failles 85
3.1.1 Calcul des primitives géométriques 86
3.1.2 Calcul des nœuds du graphe de connectivité 94
3.1.3 Calcul des connexions du graphe de connectivité 95
3.1.4 Insertion de puits dans le graphe de connectivité 96
3.1.5 Comparaison avec d'autres supports de discrétisation 99
3.1.6 Perspectives 100
3.2 Méthode de modélisation stochastique inverse de la géométrie et de la connectivité de failles grâce à des données d'écoulement de fluides 103
3.2.1 Introduction 105
3.2.2 Related Work and Challenges of Structural Parameters Inversion . . 106
3.2.3 Parameterization of Fault Networks 109
3.2.4 Inversion Methodology 111
3.2.5 Applications 114
3.2.6 Discussion and Conclusions 122
3.3 Gestion des nombres aléatoires 125
3.3.1 Vecteur de nombres aléatoires 125
3.3.2 Structure de données gérant les nombres aléatoires 125
3.4 Perspectives 128
3.4.1 Familles de failles cogénétiques et modélisation inverse 128
3.4.2 Échantillonnage de l'espace d'incertitudes 128
3.5 Conclusions 129

Conclusions générales 131

Bibliographie 143

A Publications **145**
Stochastic simulations of fault networks in 3D structural modeling 146
Stochastic simulation of fault networks from 2D seismic lines 154
Assessing the impact of fault connectivity uncertainty in reservoir studies using explicit discretization 159

Résumé **169**

Table des figures

1 Étude montrant l'incertitude d'interprétation et le possible biais humain. . . 5
2 Types d'incertitudes en modélisation structurale 7
3 Incertitude de type 1 induisant une incertitude de type 3 8
4 Faille paramétrique de la méthode HAVANA 10

1.1 Complexité structurale d'une faille 15
1.2 Représentations d'une faille en géomodélisation 16
1.3 Dôme de sel construit par méthode implicite 17
1.4 Données utilisées et contraintes DSI associées pour le calcul d'un champ scalaire 18
1.5 Opérations booléennes en modélisation implicite 19
1.6 Représentation sous forme d'arbre binaire d'une faille découpant l'espace en deux blocs 20
1.7 Réseau de failles et représentation topologique en arbre binaire associée . . 20
1.8 Conséquences des relations de filiation dans un arbre binaire d'un point de vue géologique 21
1.9 Non unicité de la représentation d'un ensemble de failles dans un arbre binaire 22
1.10 Carte structurale avec différentes familles de failles 24
1.11 Évolution spatiale et temporelle d'une zone de faille en relai 25
1.12 Relations entre familles de failles dans un arbre topologique 26
1.13 Paramètres décrivant la sinuosité d'une faille 28
1.14 Incertitude lors du pointage d'une faille sur un profil sismique 29
1.15 Aperçu des étapes de simulation d'un objet faille 33
1.16 Éloignement et répulsion entre failles 36
1.17 interactions entre failles et probabilité de nucléation 38
1.18 Méthode de perturbation 3D d'une surface implicite 41
1.19 Méthode de perturbation 2D d'une surface implicite 42
1.20 Exemples de perturbation géométrique 43
1.21 Arbre topologique et failles d'extension finie 44
1.22 Calcul du seuil de déviation d'une faille 46
1.23 Résultats montrant l'effet de la probabilité de déviation 47
1.24 Cas possibles pour l'intersection entre deux failles 48
1.25 Paramètres et probabilité de nucléation de failles en relai 51
1.26 Nucléation de failles secondaires dans une zone en relai 52
1.27 Exemple de modélisation d'une faille listrique 52
1.28 Structures en fleur 53
1.29 Modélisation d'une structure en fleur 54
1.30 Évolution d'une structure en fleur 54
1.31 Rejet observé dans une zone de failles 55
1.32 Zones non faillées et orientation d'une faille 57
1.33 Zones non faillées et géométrie d'une faille 57

2.1 Slip information and fault location uncertainties 63

2.2 Uncertainties about fault trace information 64
2.3 Partial branching between two faults 65
2.4 Flow chart showing the fault simulation steps 66
2.5 Fault represented as an implicit surface 67
2.6 Fault trace filtering based on orientation 68
2.7 Isolated fault displacement distribution 69
2.8 Fault trace filtering based on displacement information 70
2.9 Data clustering based on angle criterion 71
2.10 Method for modeling stratigraphic uncertainty 74
2.11 Reference interpretation of the Middle East field case 76
2.12 Gross rock volumes estimation distributions for unit U_1, U_2, U_3 and total volume 79
2.13 Example of stochastic models (top reservoir surfaces) 80
2.14 Example of stochastic models (cross-sections) 81

3.1 Approche générale de la résolution d'un problème inverse 85
3.2 Configurations possibles d'un tétraèdre par rapport à une faille 89
3.3 Calcul de l'intersection de plusieurs failles dans un tétraèdre 91
3.4 Intersection partielle dans un tétraèdre 92
3.5 Intersection partielle et mise à jour des liens 93
3.6 Calcul de la transmissibilité entre deux primitives géométriques 95
3.7 Calcul de la transmissibilité au niveau de l'intersection entre deux failles . . 96
3.8 Vue 2D d'un graphe de connectivité 97
3.9 Supports de discrétisation utilisés pour la comparaison des simulations d'écoulements 100
3.10 Comparaison des simulations d'écoulement entre différents supports de discrétisation 101
3.11 Effet de la résolution lors d'une simulation d'écoulement 102
3.12 Typical modeling and inversion workflow 107
3.13 Fault length distribution coupled with existence probability parameter . . . 113
3.14 Reference model of first example 115
3.15 Misfit repartition computed from equation 3.9 116
3.16 Misfit values in function of perturbation amplitude p 117
3.17 Probability maps of faults belonging to family A 118
3.18 Probability maps of faults belonging to family B 119
3.19 Misfit evolution for one seed model 120
3.20 Initial interpretations and reference for second example 121
3.21 Misfit evolution of seven structural scenarios 122
3.22 Comparison of initial interpretations and end-member models 123
3.23 Structure de données gérant les nombres aléatoires 126

Introduction

La géomodélisation a pour but de représenter les structures géologiques du sous-sol, ainsi que les propriétés des roches. Elle est utilisée dans de nombreux domaines des géosciences, notamment pour :

- quantifier les ressources naturelles tant énergétiques, pétrolières, minières, qu'hydriques du sous-sol.
- comprendre les processus géologiques de mise en place des objets géologiques, leurs interactions, dans une optique aussi bien scientifique qu'industrielle. En effet, parvenir à estimer un volume de minerai ou d'hydrocarbures requiert une compréhension globale de la zone étudiée, et les enjeux sont trop importants pour se permettre de négliger de tels aspects.

Les modèles se construisent de nos jours en trois dimensions, pour permettre une approche globale, afin de confronter les différentes données d'entrée et mettre à jour leurs relations, pour relever notamment les incohérences géologiques et les corriger le plus tôt possible dans le processus de modélisation.

Les données utilisées lors de la construction des modèles sont hétérogènes car elles sont acquises à différentes échelles et résolutions, parfois à différentes époques :

- des données précises mais locales :
  - les mesures effectuées sur le terrain, pendage et azimuth pour l'orientation des surfaces par exemple ;
  - les données de puits ;
  - les profils miniers (galeries, puits).
- des données précises mais partielles (en surface) :
  - les photos satellites ;
  - les données radar ;
  - les données acquises grâce à la télédétection par laser (LIDAR).
- des données globales mais avec une résolution limitée :
  - les profils sismiques, aussi bien 2D que 3D ;
  - les données acquises par des méthodes géophysiques de surface (*e.g.* gravimétrie, électromagnétisme).

Des données au géo-modèle

La construction d'un modèle 3D commence par l'interprétation des différentes données d'entrée afin de mettre en évidence les structures géologiques majeures (failles, horizons stratigraphiques clairement identifiables) de la zone étudiée. En effet, celles-ci délimitent des compartiments distincts qu'il convient d'identifier au plus tôt afin de définir des volumes de roche aux propriétés similaires. Les failles jouent particulièrement un rôle crucial en introduisant des discontinuités verticales et latérales dans les couches géologiques. Par conséquent, le réseau de failles est modélisé en premier, puis les horizons stratigraphiques sont construits dans les différents blocs de failles en maintenant la cohérence du rejet observé de part et d'autre des failles.

Une fois le cadre structural défini, les propriétés des roches sont alors modélisées, le plus souvent par des méthodes géostatistiques pour palier le manque d'observations exhaustives. L'espace est alors discrétisé afin de résoudre les équations numériques représentant le phénomène physique étudié (*e.g.* propagation d'ondes, écoulement de fluides).

Les sources d'incertitude en géomodélisation

Avant de discuter les sources d'incertitude, il convient de définir le mot *incertitude* qui sera employé tout au long de ce mémoire. La définition suivante est celle proposée par Zimmermann [2000] :

Définition 1. ***Incertitude***
"L'incertitude implique que dans une situation donnée, une personne ne dispose pas d'information qualitative et quantitative appropriée pour décrire, formuler ou prédire numériquement et de manière déterministe un système, son comportement ou tout autre de ses caractéristiques."

Comme nous l'avons vu précédemment, les données d'entrée ne sont jamais à la fois exhaustives et précises, c'est-à-dire que l'on ne peut pas décrire le sous-sol de manière déterministe sans ambiguïté. L'*incomplétude* et la *résolution limitée* des données sont donc les deux principales sources d'incertitude en géomodélisation, inhérentes à leur acquisition.

D'autres sources d'incertitude viennent s'ajouter et rendent l'interprétation des données ambiguë. Dans le cadre de la modélisation structurale, le traitement des données sismiques brutes afin de produire des profils interprétables en profondeur est une source importante d'incertitude. D'après Thore et al. [2002], la migration des réflecteurs à leur vraie position ainsi que la conversion d'un espace en temps à un espace en profondeur sont les deux principales sources d'incertitude dans la chaîne de traitement sismique. En effet, le modèle de vitesse est en général approximatif, son utilisation introduit donc une erreur non négligeable dans le positionnement des structures géologiques, même si une réévaluation du modèle de vitesse peut être faite lors de l'acquisition de données de forage.

Une interprétation des données nécessaire, mais introduisant des biais

Les données seules ne suffisent pas à la construction d'un modèle. Comme l'a très justement souligné Frodeman [1995], la géologie est une *science interprétative* par nature à cause des incertitudes inhérentes à l'acquisition des données. Dans d'autres sciences, des hypothèses peuvent être testées par des expériences sous des conditions contrôlées et des variables limitées en nombre. Au contraire, le géologue doit interpréter des données partielles produites par des expériences naturelles sous des conditions inconnues et avec des variables peu ou mal connues. L'interprétation des données dans le cadre d'un raisonnement géologique et selon un modèle géologique conceptuel est donc primordiale [Frodeman, 1995].

La géomodélisation fait partie intégrante de ce processus et l'interprétation y

joue un rôle très important lors de la construction d'un modèle. Cette interprétation repose sur différents points :

- des concepts géologiques en rapport avec le contexte régional afin de mieux contraindre la modélisation et écarter les scénarios incohérents d'un point de vue géologique ;
- des lois physiques qui permettent de déterminer les paramètres manquants ou non observés des structures géologiques ;
- des analogues géologiques de la zone étudiée pour guider l'interprétation.

Un biais humain lors de l'interprétation

L'interprétation et l'utilisation de concepts géologiques sont très subjectives et dépendent de la personne qui interprète. Bond et al. [2007] ont réalisé une enquête sur l'interprétation d'une même section sismique 2D synthétique (figure 1*b*) par 412 géo-scientifiques. Seulement 21% d'entre eux ont trouvé le bon contexte tectonique (figure 1*c*), ce qui montre d'une part que les incertitudes lors de l'interprétation sont grandes et d'autre part qu'un biais humain existe. En effet, Bond et al. [2007] ont entre autre montré qu'une personne a tendance à interpréter les données dans son domaine d'expertise (figure 1*d* et 1*e*) : c'est le *biais de disponibilité*. D'autres biais perturbent l'interprétation des données, par exemple :

- le *biais de confirmation* où la personne cherche à faire correspondre les données à une hypothèse préalablement établie ;
- le *biais d'optimisme* où la personne surestime les chances et minimise les risques ;
- le *biais d'ancrage* où la personne se focalise sur un point ou une donnée précise et ignore ou minimise l'importance du reste des données.
- le *biais de représentativité* où la personne se base sur des faits et ignore les connaissances a priori.

Le choix d'un contexte tectonique et plus généralement l'interprétation des données sont donc issus à la fois d'une décision consciente et inconsciente, créant un *biais cognitif*. Ces biais affectent de manière plus générale les décisions de la vie quotidienne et ont été décrits dans la littérature des sciences cognitives [Krueger and Funder, 2004]. Tversky and Kahneman [1974] mettent en évidence le biais de représentativité en demandant à des personnes de dire si un individu est avocat ou ingénieur d'après la description suivante, en sachant que cet individu a été choisi au hasard dans un groupe de 100 personnes dont 70 sont ingénieurs et 30 sont avocats :

"Dick est âgé de 30 ans. Il est marié et n'a pas d'enfant. Ayant des capacités intellectuelles importantes et une grande motivation, il a promis de réussir dans sa carrière. Il est très apprécié par ses collègues".

Cette description ne donne aucun indice sur le fait que Dick soit avocat ou ingénieur, la proportion de réponse devrait donc être de 70% ingénieurs et 30% avocats eu égard des statistiques a priori. Pourtant, Dick est considéré comme avocat dans 50% des cas, c'est-à-dire que les personnes ont fait leur jugement d'après des données personnalisées mais non informatives pour la question demandée, au lieu d'utiliser les probabilités a priori. La proportion 70-30 est bien respectée quand aucune information n'est donnée. Le biais de représentativité incite également l'individu à ne pas prendre en compte la taille de l'échantillon, problème bien connu en géostatistique. Pour illustrer ce problème, Tversky and Kahneman [1974] posent par exemple la question suivante :

"Deux hôpitaux sont dans une même ville. Le plus grand hôpital enregistre en moyenne 45 naissances par jour alors que 15 bébés naissent en moyenne dans le plus petit. Comme vous le savez, environ 50% des bébés sont des garçons, mais le pourcentage exact est différent chaque jour. Pendant une période de un an, les deux hôpitaux enregistrent le nombre de jours où plus de 60% des naissances sont des garçons. Quel hôpital a enregistré le plus de ces jours ? 3 choix sont possibles : le plus petit, le plus grand ou environ le même nombre de jours pour les deux ?"

95 personnes sont interrogées, 21 d'entre elles choisissent comme réponse le plus petit hôpital (la bonne réponse), 21 le plus grand, et 53 disent que les deux hôpitaux enregistrent le même nombre de jours avec plus de 60% de garçons. Elles ont basé leur jugement sur le fait que les deux hôpitaux sont représentatifs de la population générale, mais ont ignoré le fait que la variabilité d'un échantillon augmente quand la taille de l'échantillon diminue.

Réduire le risque lors de la construction des modèles 3D

Nous avons vu qu'il y a des incertitudes inhérentes à la géomodélisation d'une part ; et qu'un biais humain cognitif est possiblement introduit lors de l'interprétation des données d'autre part. Ces deux aspects montrent que les modèles déterministes sont hautement risqués, dans le sens où la probabilité que le modèle soit prédictif est très faible. L'utilisation des géo-modèles est une aide à la décision dans de nombreux domaines, dans le secteur des énergies fossiles mais aussi pour l'évaluation des risques naturels ou des ressources en eau par exemple. Dans tous les cas, les enjeux - financiers ou humains - sont considérables et les risques doivent donc être maîtrisés.

L'approche multi-scénarios

Cette idée a très vite émergé dans la communauté géologique, puisque Chamberlin [1890] et Gilbert [1886] avaient déjà proposé il y a plus d'un siècle de considérer plusieurs hypothèses de travail. Chamberlin [1890] illustre ce concept avec la question de la formation des Grands Lacs d'Amérique du Nord. Ses étudiants étaient farouchement partagés pour expliquer leur origine :

- certains pensaient que leur origine était due à la présence de vallées préglaciaires ;
- d'autres soutenaient que ces grandes excavations étaient dues à une érosion glaciaire ;
- enfin, il se pouvait aussi que leur origine soit due à une flexion de la croûte.

Ces trois origines possibles constituent pour Chamberlin trois hypothèses de travail à considérer (l'origine des Grands lacs est en fait due à une combinaison de ces trois hypothèses...).

Cette approche a plus récemment été soutenue par Bond et al. [2008] dans le domaine de l'exploration d'hydrocarbures. Ils proposent de garder et de continuer la suite du processus de modélisation avec plusieurs modèles conceptuels différents (*i.e.* différents contextes tectoniques dans leur étude) suite à l'interprétation et non un seul modèle comme actuellement dans l'industrie.

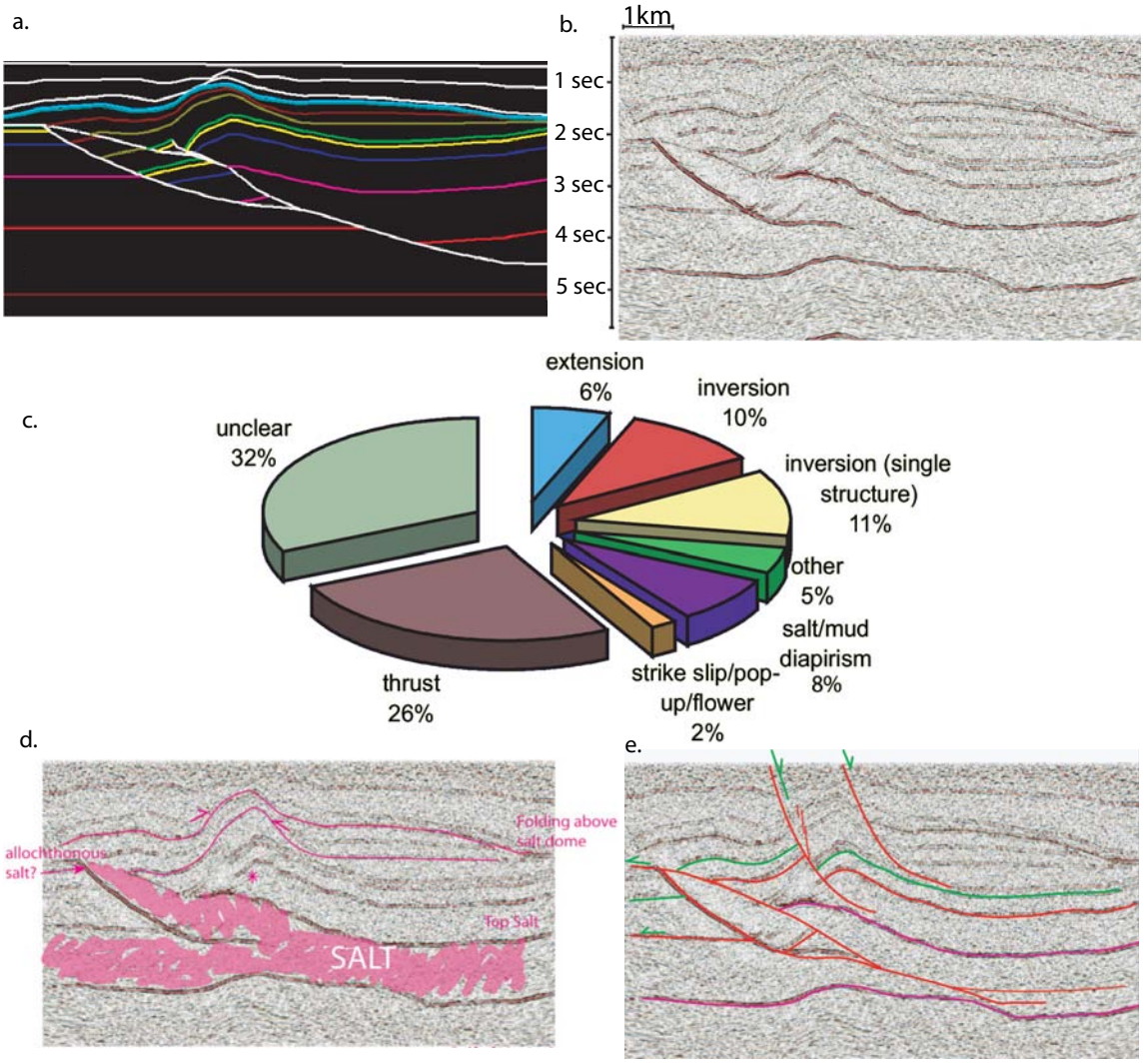

FIGURE 1 – Étude montrant l'incertitude d'interprétation et le possible biais humain. Un modèle synthétique (*a.*) a été utilisé pour générer une section sismique synthétique (*b.*). Le diagramme (*c.*) montre l'étendue des contextes tectoniques interprétés par les géoscientifiques. L'étude montre aussi qu'il existe un biais selon l'expérience de la personne. *d.* L'interprétation d'un étudiant en tectonique salifère. *e.* L'interprétation d'un expert avec plus de 15 ans d'expérience en tectonique d'extension. D'après Bond et al. [2007].

Réduire le biais humain pour obtenir des modèles conceptuels différents

Il est difficile pour une personne de proposer plusieurs scénarios à partir de l'interprétation d'un jeu de données, par exemple un contexte tectonique différent dans l'étude de Bond et al. [2007]. Une des solutions consiste à mettre en présence plusieurs personnes avec des domaines d'expertise différents afin de concilier les vues et limiter le biais de disponibilité [Bond et al., 2008, Wood and Curtis, 2004].

L'effet du biais de disponibilité peut aussi être réduit en présentant aux per-

sonnes concernées différents contextes possibles en amont de l'interprétation, ce qui a pour effet de diminuer le poids relatif du domaine d'expertise de la personne par rapport aux autres domaines.

Une autre étude de Bond et al. [2011] a montré que l'interprétation de données sismiques est améliorée lorsque le niveau d'étude académique de la personne qui interprète est élevé (master ou doctorat). La méthode d'interprétation est également discriminante, l'interprétation des horizons et l'utilisation de schémas d'évolution chronologique étant les deux techniques augmentant significativement la pertinence des résultats d'interprétation. La méthode d'interprétation et le niveau d'étude sont liés puisqu'une personne a plus de chance d'avoir appris ces techniques au cours de ses études académiques. L'étude a cependant montré que ces deux variables ne sont pas totalement corrélées et que le niveau d'étude, indépendamment de la technique d'interprétation utilisée, est discriminant. Les auteurs concluent qu'un master ou un doctorat sont en eux-mêmes précieux pour le développement d'un raisonnement critique requis lors de l'interprétation.

Étudier la variabilité ou l'incertitude intrinsèque de chaque modèle conceptuel

Nous avons vu jusqu'ici que plusieurs modèles conceptuels peuvent être proposés pour les mêmes données (voir le diagramme des contextes tectoniques dans l'étude de Bond et al. [2007] en figure 1c). En outre, pour chaque contexte (extensif, compressif, tectonique salifère par exemple), l'ambiguïté des données ne permet pas de résoudre avec certitude la géométrie et la topologie des structures géologiques présentes.

Définition 2. ***Topologie***
Le mot *topologie* employé dans ce mémoire renvoie aux relations entre les objets dans une représentation de l'espace, le sous-sol à travers un géo-modèle dans notre cas. Pour un objet donné, la *topologie* concerne aussi son existence, c'est-à-dire sa présence ou son absence dans une représentation de l'espace. Les règles de recoupement entre failles, leur relation de parenté font notamment partie de la topologie d'un réseau de failles.

Il convient donc d'étudier la variabilité, due aux incertitudes, de chaque modèle conceptuel. C'est dans ce cadre que s'inscrit le projet de recherche mené au cours de cette thèse. Une méthode permettant d'échantillonner les incertitudes des structures géologiques a été développée. L'échantillonnage des incertitudes topologiques constitue l'originalité du travail de recherche par rapport aux méthodes existantes, ce qui permet un balayage plus large de la variabilité pour un même modèle conceptuel (contexte tectonique). La méthode a été développée dans le cas de contextes faillés en général, extensif ou compressif.

Les incertitudes en modélisation structurale

Types d'incertitudes structurales

Les incertitudes structurales peuvent être classées selon leur origine (échantillonnage, observation, mesure) et le degré d'incertitude (la variabilité due à l'incertitude). La classification suivante synthétise la description plus générale de Bardossy and Fodor [2001], Mann

[1993] et est adaptée à la modélisation structurale [Wellmann et al., 2010]. On distingue trois principaux types d'incertitudes :

- type 1 : incertitudes dues aux erreurs et à l'imprécision des mesures, par exemple concernant la position et l'orientation des surfaces. Cette incertitude concerne principalement la géométrie des structures ($2a$).
- type 2 : incertitudes dues à la variabilité naturelle inhérente des objets géologiques [Bardossy and Fodor, 2001], plus ou moins importante selon les objets géologiques ($2b$). Ainsi, la variabilité naturelle d'une surface de faille est beaucoup plus faible que celle d'une surface de minéralisation d'un gisement d'uranium de type roll-front.
- type 3 : incertitudes dues au manque de connaissances, sur l'existence d'une structure par exemple ($2c$).

Seule l'incertitude de type 2 est une propriété de la nature indépendante de l'acquisition, les types 1 et 3 sont dus à l'investigation géologique et à l'activité humaine [Bardossy and Fodor, 2001]. Les incertitudes de type 1 et 2 concernent principalement la géométrie des structures (ou leur topologie propre, l'auto-intersection par exemple) alors que le type 3 peut affecter la topologie d'un modèle comme illustré figure $2c$.

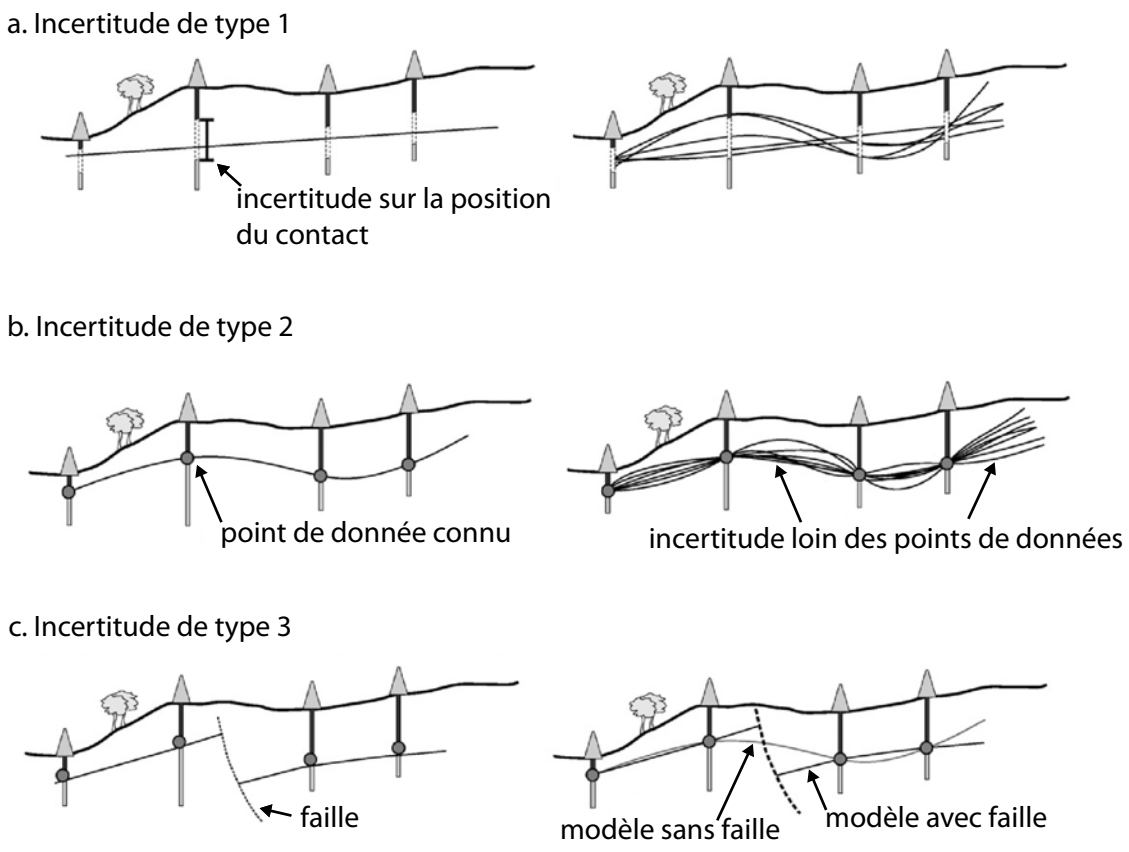

FIGURE 2 – Types d'incertitudes en modélisation structurale. *a.* L'incertitude de type 1 est due aux imprécisions et erreurs de mesure. La position de l'horizon est ici incertaine, sa position n'est connue que dans un intervalle. (*b.*). L'incertitude de type 2 est due à la variabilité naturelle intrinsèque des objets géologiques. *c.* L'incertitude de type 3 concerne l'existence des objets et leurs relations. Modifié d'après Wellmann et al. [2010].

Les trois types d'incertitude ne sont pas indépendants et sont tous présents lors de la construction d'un modèle. Ainsi, des incertitudes de type 1 ou 2 peuvent induire des incertitudes de type 3 selon le degré d'incertitude (figure 3).

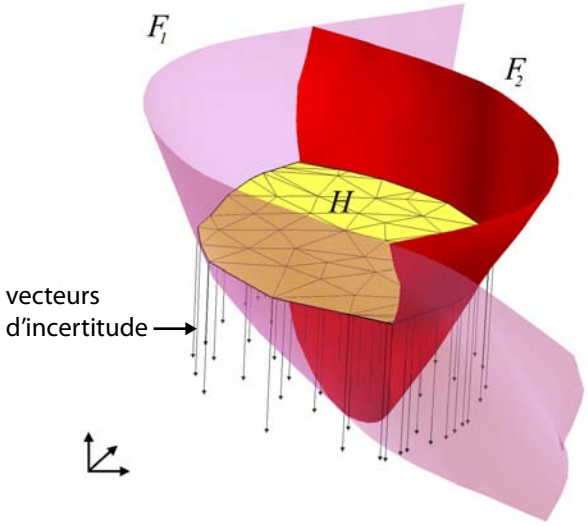

FIGURE 3 – Incertitude de type 1 induisant une incertitude de type 3. Dans le cas initial, l'horizon H se situe entre les failles $F1$ et $F2$. Une simple translation de H le long de vecteurs d'incertitude peut complètement changer la configuration : H pourrait être en contact avec seulement $F1$ ou $F2$ et ne plus se situer entre les deux failles. D'après Cherpeau et al. [2010a].

Méthodes d'échantillonnage

L'investigation géologique ne permet d'approcher que partiellement la réalité, celle-ci peut-être donc être considérée comme un évènement aléatoire avec une certaine probabilité d'occurrence. L'approche probabiliste permet ainsi de rendre compte de la connaissance partielle du sous-sol, la probabilité étant une mesure de l'incertitude. Les méthodes probabilistes, ou stochastiques, génèrent, à partir de probabilités associées aux paramètres d'entrée, des réalisations équiprobables du phénomène étudié.

Les méthodes d'échantillonnage des incertitudes structurales ont principalement été développées dans le contexte d'exploration d'hydrocarbures, pour prendre en compte les différentes incertitudes dues à l'acquisition, au traitement et à l'interprétation des données sismiques [Thore et al., 2002].

Méthodes d'échantillonnage des incertitudes géométriques (type 1 et 2)

La majorité des méthodes proposées dans la littérature se sont concentrées sur les incertitudes de type 1 et 2 [Abrahamsen, 1992, Charles et al., 2001, Caumon et al., 2007, Lecour et al., 2001, Thore et al., 2002, Wellmann et al., 2010]. Abrahamsen [1992] propose un modèle stochastique pour les réflecteurs sismiques et le champ de vitesse à l'échelle du réservoir. Des champs aléatoires gaussiens de moyenne nulle représentent l'incertitude

sur la position exacte des réflecteurs et sur les valeurs de vitesse. Un krigeage bayésien [Omre, 1987] permet alors de prendre en compte l'incertitude sur l'information a priori (*e.g.* courbure des surfaces, vitesse par intervalle).

Les méthodes proposées par Lecour et al. [2001], Thore et al. [2002] utilisent la méthode de perturbation par champs de probabilité [Srivastava, 1992] pour perturber des surfaces triangulées (failles ou horizons) le long de vecteurs d'incertitude définis en chaque point de la surface. L'amplitude du déplacement est déterminée par un tirage de Monte-Carlo dans une loi de distribution prédéterminée (loi triangulaire par exemple, nulle aux extrémités du vecteur, maximum à la position de référence). Le champ de probabilité fournit les valeurs pour le tirage de Monte-Carlo et assure ainsi une déformation cohérente pour des points adjacents.

L'amplitude des perturbations respecte une enveloppe d'incertitude déterminée par les incertitudes sismiques, en contrôlant la norme des vecteurs d'incertitude. Dans cette approche, les incertitudes de type 1 sont modélisées par les vecteurs d'incertitude aux points de données, la norme des vecteurs contrôlant le degré de confiance dans la position initiale. Le variogramme utilisé pour générer le champ de probabilité permet quant à lui de contrôler la longueur d'onde des perturbations et donc permet de refléter la variabilité naturelle des surfaces considérées (incertitude de type 2).

Cette méthode de perturbation est surfacique et la mise à jour du modèle volumique n'est pas directe. Caumon et al. [2007] ont proposé une méthode de perturbation volumique qui permet de déformer une grille stratigraphique, en propageant le déplacement à l'ensemble des nœuds de la grille. Une attention particulière est toutefois nécessaire pour conserver une grille cohérente par rapport aux objets géologiques.

La modélisation implicite [Calcagno et al., 2008, Frank et al., 2007, Guillen et al., 2008] permet dans une certaine mesure de s'affranchir des contraintes de maillage (pour plus de détails, se référer à la section 1.2). Dans ce cadre, Wellmann et al. [2010] ont proposé une méthode d'échantillonnage des incertitudes de type 1. À partir des données initiales et de leurs incertitudes, plusieurs jeux de données sont générés, chacun représentant un ensemble de données aussi exactes que possible. Ces données sont ensuite automatiquement utilisées pour générer des modèles stochastiques, le processus de modélisation étant guidé par un modèle de référence construit au préalable.

Méthodes d'échantillonnage des incertitudes topologiques (type 3)

Du fait de leur difficulté de mise en oeuvre, peu de méthodes permettant de modifier la topologie des modèles ont été développées. Holden et al. [2003], Hollund et al. [2002] ont proposé une méthode, HAVANA, permettant de simuler des réseaux de failles à topologie variable dans des grilles stratigraphiques. L'avantage pratique d'une telle approche est de permettre la mise à jour rapide de la grille stratigraphique afin d'évaluer l'impact des incertitudes sur l'écoulement de fluides. En revanche, les failles majeures sont représentées par des plans bilinéaires parallèles aux piliers de la grille, ce qui limite la flexibilité et le réalisme d'une telle approche (figure 4). Les failles plus petites sont représentées par un modèle plan elliptique et sont prises en compte par la perturbation du champ de perméabilité ou par des multiplicateurs de transmissibilité lors de la simulation des écoulements de fluides. Chaque faille est considérée comme un opérateur de rejet et le déplacement se fait le long des piliers

de la grille, ce qui limite la modélisation de la composante décrochante du rejet étant donné que les piliers sont subverticaux.

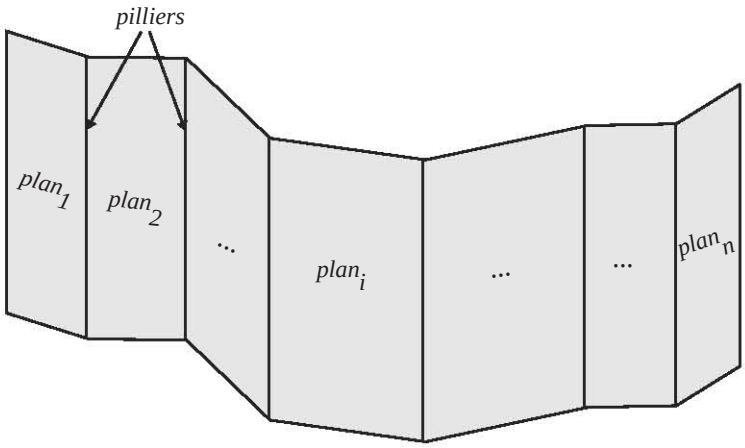

FIGURE 4 – Faille paramétrique de la méthode HAVANA. La faille est représentée par un ensemble de plans bilinéaires alignés sur les piliers d'une grille stratigraphique.

Suzuki et al. [2008] ont proposé une approche différente pour échantillonner les incertitudes topologiques. Leur méthode consiste à perturber la géométrie de plusieurs interprétations structurales. La distance de Hausdorff (mesure de la similarité entre deux formes géométriques) est calculée pour l'ensemble des paires de modèles obtenus, ce qui permet de n'évaluer la réponse dynamique que de modèles éloignés selon la distance de Hausdorff. Un algorithme stochastique [Sambridge, 1999] explore ensuite le voisinage des modèles évalués. Ce voisinage est défini par des cellules de Voronoï calculées à partir de la distance de similarité. Chaque cellule est ensuite explorée selon une probabilité d'autant plus élevée que l'écart du modèle définissant la cellule avec les observations est faible.

Limites des méthodes actuelles

Nous avons vu précédemment que des incertitudes dues à l'imprécision des données, à la variabilité naturelle et au manque de connaissance sur les structures géologiques sont intimement liées et qu'il est difficile d'isoler un type d'incertitude pour le caractériser individuellement. Or les méthodes actuelles ne permettent pas une telle approche. Les méthodes d'échantillonnage géométriques [Caumon et al., 2007, Lecour et al., 2001, Thore et al., 2002] conservent la topologie des modèles et ne peuvent prendre en compte les incertitudes importantes relatives aux relations entre objets géologiques. De plus, ces méthodes reposent sur un modèle de référence que l'on déforme plus ou moins, en préservant sa topologie globale, de manière analogue à un objet qui vibre. D'après l'étude de Bond et al. [2007], plusieurs biais peuvent affecter la construction d'un modèle de référence. Par conséquent, partir d'un modèle biaisé et approximatif ne peut qu'aboutir à des prédictions hasardeuses. De même, l'approche de Suzuki et al. [2008] est intéressante mais ne permet d'échantillonner qu'un nombre limité d'interprétations structurales dépendant des moyens humains et du temps disponibles pour réaliser de telles interprétations.

La méthode HAVANA ne repose pas sur une interprétation de référence mais souffre

d'un manque de réalisme géologique, la flexibilité de la méthode étant limitée à la géométrie de la grille.

Le problème des incertitudes structurales en géomodélisation reste donc ouvert, les méthodes actuelles ne permettant pas de les appréhender de manière satisfaisante. La suite du mémoire reprend le travail effectué au cours de ma thèse pour répondre à cette problématique. La contribution principale est le développement d'un modèle stochastique de faille présenté en chapitre 1.

Ce modèle a ensuite été étendu pour générer des modèles structuraux incluant le déplacement des failles et des horizons stratigraphiques ; puis utilisé dans un contexte d'exploration avec peu de données pour estimer les incertitudes sur des volumes de roches (chapitre 2).

Le problème de cohérence des modèles intervient également lorsque des données dynamiques sont disponibles. Or les structures jouent également un rôle sur l'écoulement des fluides, en plus des propriétés hydrauliques des roches. La paramétrisation des structures géologiques permet de les inclure dans le problème inverse de caractérisation du sous-sol par données dynamiques. Dans ce contexte, le chapitre 3 présente une paramétrisation de réseaux de failles et ouvre de nouvelles portes vers la prise en compte de structures géologiques lors de la résolution de problèmes inverses.

À la fin, un chapitre de conclusions résume les principaux résultats obtenus et tente de proposer de nouvelles voies de recherche en guise de perspective. Une annexe rassemble les différents travaux (publications, résumés étendus) écrits pendant ce travail de thèse.

Chapitre 1

Modèle stochastique de failles

Sommaire

1.1 Qu'est-ce qu'une faille ? 14
1.1.1 Une faille n'est pas toujours une simple surface de glissement 14
1.1.2 Les failles en géomodélisation 15
1.2 La modélisation implicite dans le cadre des incertitudes structurales 17
1.2.1 Définition d'une surface implicite 17
1.2.2 Avantages de la modélisation implicite 17
1.3 Les arbres binaires comme représentation de la topologie d'un réseau de failles 19
1.3.1 Une surface implicite comme partition de l'espace 19
1.3.2 Représentation graphique des relations de filiation entre failles 20
1.3.3 Propriétés des arbres binaires 21
1.4 Description d'un réseau de failles 22
1.4.1 Familles de failles 22
1.4.2 Mise en place de familles de failles 23
1.4.3 Relations entre familles de failles 23
1.5 Algorithme général et données d'entrée 25
1.5.1 Algorithme général 25
1.5.2 Paramètres d'une famille de failles 26
1.5.3 Gestion des données 28
1.5.4 Critères d'arrêt de la simulation 29
1.5.5 Gestion des familles cogénétiques 30
1.6 Modélisation d'un objet faille 32
1.6.1 Entrées 32
1.6.2 Aperçu des étapes de simulation d'un objet faille 32
1.6.3 Attribution des points de données 32
1.6.4 Orientation de la future faille 34
1.6.5 Centre de faille 34
1.6.6 Géométrie initiale 38
1.6.7 Regroupement d'ensembles de points 39
1.6.8 Géométrie finale 39
1.6.9 Extension latérale 41
1.6.10 Choix de l'ordre des étapes de modélisation d'une faille 43
1.6.11 Mise à jour du modèle 45
1.7 Perspectives 50
1.7.1 Concepts géologiques supplémentaires 50
1.7.2 Intégration de données supplémentaires 53

1.7.3 Simulation d'un objet faille modulable 56
1.8 Conclusions 58

L'objectif de ce chapitre est de présenter le modèle stochastique de faille développé au cours du projet de thèse. Nous avons vu en introduction que la plupart des méthodes actuelles génèrent des modèles stochastiques en perturbant la géométrie d'un modèle de référence. Notre objectif est de générer des modèles stochastiques pour prendre en compte les incertitudes structurales, mais en générant également des modèles avec des topologies différentes en essayant de partir des données brutes.

Avant de modéliser les volumes de roches, la construction des modèles commence par la modélisation des structures, les failles principalement. En effet, celles-ci partitionnent la zone étudiée en différents blocs. Par ailleurs, les propriétés des roches sont principalement acquises lors de leur dépôt, il est donc plus judicieux de les quantifier par couche géologique [Mallet, 2004]. Or les failles affectent la géométrie de ces couches par les discontinuités spatiales qu'elles engendrent. La modélisation des surfaces de faille intervient donc en premier lieu lors de la modélisation structurale. Les horizons stratigraphiques sont ensuite modélisés dans chaque bloc de faille en honorant les contraintes de rejet imposées par les failles. Le développement de la méthode a donc commencé par la génération stochastique d'objets faille.

1.1 Qu'est-ce qu'une faille ?

Définition 3. ***Faille***
Le mot *faille* vient de l'ancien français *faillir*, manquer, car le mineur, après une faille, ne retrouve plus le filon ou la couche qu'il exploitait [Foucault and Raoult, 2005]. Cette origine, observée par le mineur, traduit bien ce qu'est réellement une faille : une cassure de terrain avec déplacement relatif des parties séparées [Foucault and Raoult, 2005], ce qui induit une discontinuité latérale et verticale des couches géologiques.

1.1.1 Une faille n'est pas toujours une simple surface de glissement

Une faille peut correspondre à un volume de roche broyée entre deux compartiments rocheux, incorporée au fur et à mesure par le déplacement relatif des deux blocs de faille (figure 1.1*a*).

Une faille se propage à travers la roche de manière irrégulière [Childs et al., 2009, Walsh et al., 2003]. Une faille peut être segmentée, en particulier dans des milieux stratifiés de rhéologies différentes. Par exemple, des bancs ductiles peuvent accommoder une partie du déplacement des bancs fragiles (figure 1.1*b*). Ces "sauts" sont des zones de concentration des contraintes et peuvent éventuellement disparaître lorsque celles-ci augmentent et provoquent la fracturation puis l'incorporation de la roche dans le volume faillé (figure 1.1*c*).

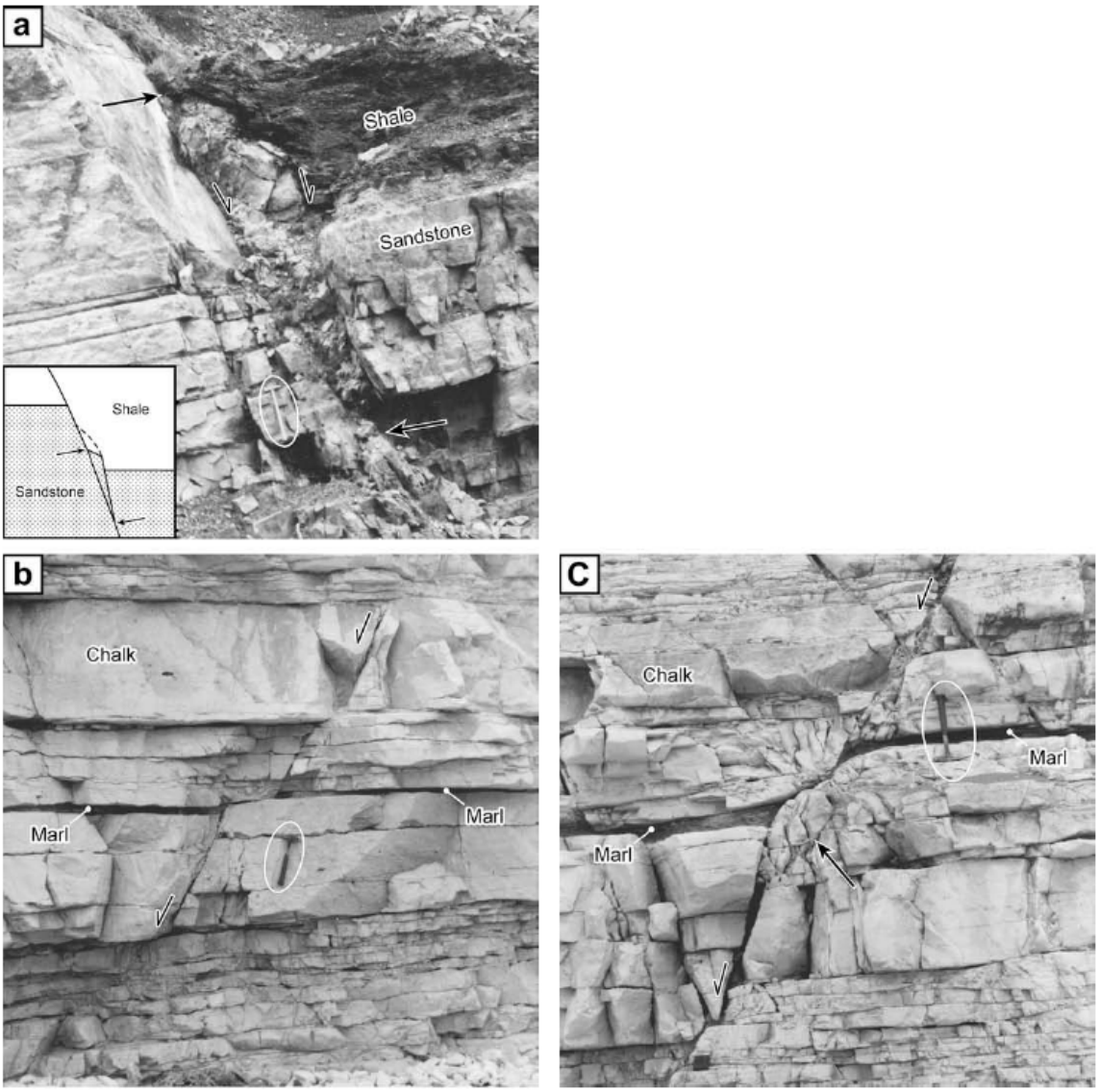

FIGURE 1.1 – Complexité structurale d'une faille. *a.* Une faille peut correspondre à un volume de roche incorporé au fur et à mesure du déplacement relatif des deux blocs de part et d'autre de la faille. *b.* Une faille est une surface irrégulière et peut être segmentée au niveau de bancs ductiles, de marnes ici. Ces bancs accommodent une partie du déplacement et sont des zones de concentration des contraintes. *c.* L'augmentation des contraintes provoque la fracturation et l'incorporation des roches au reste du volume de faille. Ces aspérités peuvent éventuellement disparaître mais affectent la géométrie finale de la faille. D'après Childs et al. [2009].

1.1.2 Les failles en géomodélisation

Une faille devrait être représentée de manière volumique pour une description réaliste. Cependant, d'un point de vue structural, l'épaisseur de la faille apparaît négligeable par rapport à son étendue dans le plan de glissement et les failles sont communément représentées

de manière surfacique (figure 1.2a, b).

En revanche, le volume de roche joue un rôle prépondérant pour l'écoulement fluide, donnant à la faille ses propriétés hydrauliques de drain ou barrière selon le matériel incorporé. On s'intéressera donc au volume de la faille et à son matériel pour la représentation dynamique de celle-ci, afin de modéliser l'écoulement fluide de la zone d'intérêt (figure 1.2c).

La section suivante présente succinctement la modélisation implicite utilisée au cours du projet de recherche pour représenter les surfaces de faille et ses avantages pour la modélisation des incertitudes structurales.

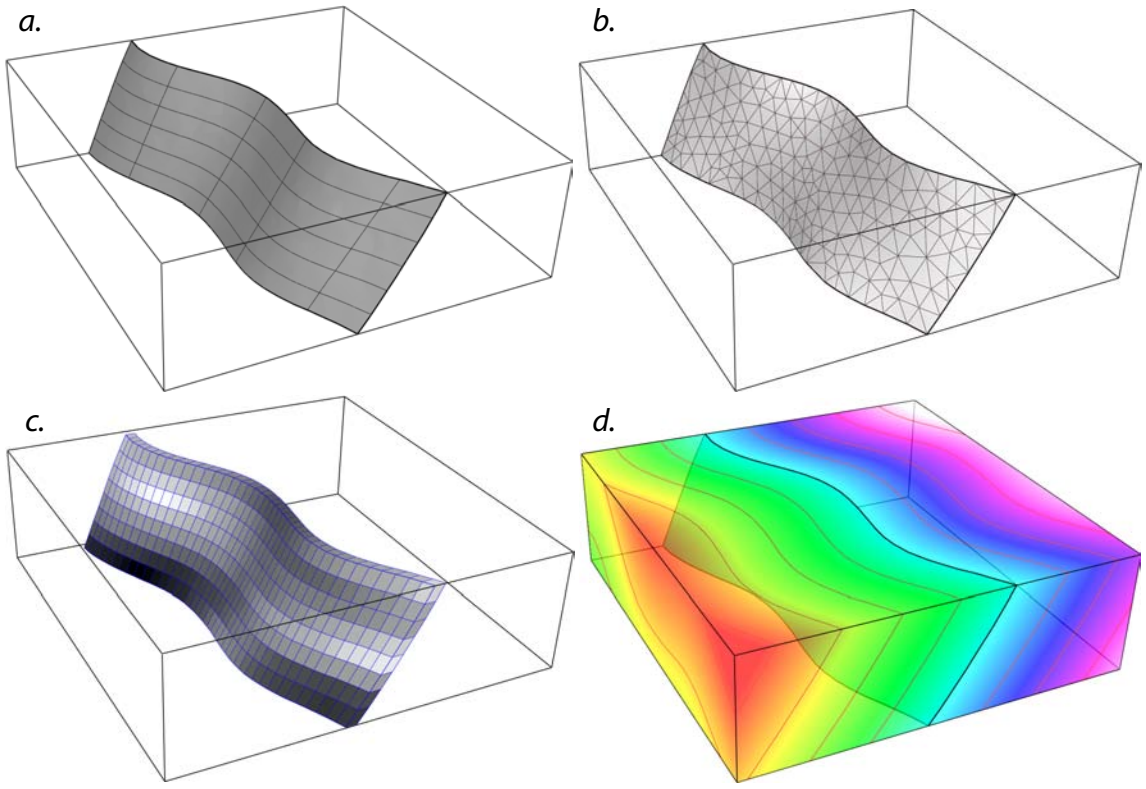

FIGURE 1.2 – Représentations d'une faille en géomodélisation. *a.* Objet faille proposé par Lecour et al. [2001]. Les lignes horizontales sont des génératrices, les lignes verticales constituent la "colonne vertébrale". *b.* Représentation par surface triangulée utilisée en modélisation explicite. *c.* Représentation volumique dans une grille hexaédrique, ce qui permet d'estimer les propriétés hydrauliques de la faille. *d.* Représentation par surface implicite (voir section 1.2 pour plus de détail).

1.2 La modélisation implicite dans le cadre des incertitudes structurales

1.2.1 Définition d'une surface implicite

Par opposition à la modélisation explicite qui décrit les objets par un ensemble de points avec des liens d'adjacence, la modélisation implicite [Calcagno et al., 2008, Frank et al., 2007, Guillen et al., 2008] propose de décrire une surface par une équipotentielle f d'un champ scalaire volumique $\mathcal{F}(x, y, z)$ représentant par exemple le temps géologique [Mallet, 2004] ou une distance à un objet [Ledez, 2003] :

$$\mathcal{F}(x, y, z) = f \tag{1.1}$$

Une faille peut ainsi être représentée par une surface implicite (figure 1.2*d*). Une pile stratigraphique conforme est représentée par un même champ scalaire, la différence d'amplitude du gradient correspond aux variations latérales d'épaisseur. Au contraire, toute discontinuité, surface érosive ou faille, est définie par un champ scalaire propre.

Calcagno et al. [2008], Chilès et al. [2004] utilisent le cokrigeage pour calculer un champ de potentiel 3D représentant les interfaces géologiques. Les failles sont modélisées en introduisant une fonction de dérive discontinue dans le système de krigeage.

FIGURE 1.3 – *a*. Dôme de sel construit par méthode implicite. *b*. Détail du dôme avec les points de données visibles. D'après Frank et al. [2007].

Mallet [2004] définit une correspondance entre l'espace actuel G et l'espace de dépôt $\overline{G}$ (espace géochronologique). Toute particule $\mathbf{x} \in G$ est localisée dans $\overline{G}$ par des paléo-coordonnées (u, v, t) grâce à une transformation $\mathbf{u}$:

$$\mathbf{x} = (x, y, z) \in G \overset{\mathbf{u}}{\rightarrow} \mathbf{u}(\mathbf{x}) = (u, v, t) \in \overline{G} \tag{1.2}$$

Frank et al. [2007], Moyen et al. [2004] ont implémenté cette méthode sur des maillages tétraédriques en utilisant l'interpolateur DSI [Mallet, 1992], ce qui permet de modéliser des surfaces à géométrie complexe (figure 1.3).

1.2.2 Avantages de la modélisation implicite

Les avantages d'une telle approche dans le cadre de la modélisation des incertitudes sont multiples :

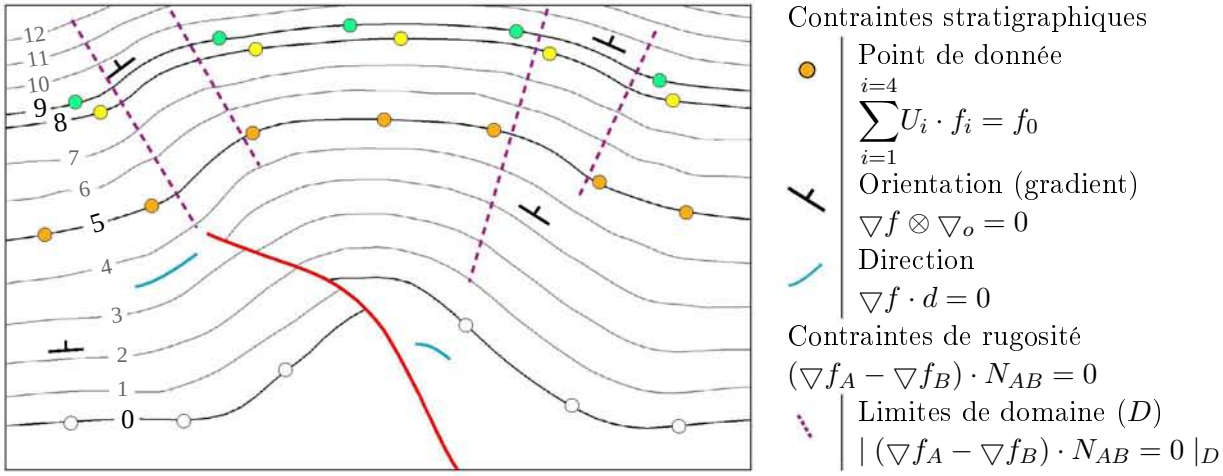

FIGURE 1.4 – Données utilisées et contraintes DSI associées pour le calcul d'un champ scalaire f. Des points de données, associées à une valeur f_0 (correspondant à un âge relatif d'une pile stratigraphique par exemple), contraignent les valeurs du champ scalaire. Des données d'orientation et de direction peuvent également être prises en compte. Enfin, des contraintes de rugosité, sur l'ensemble du modèle ou par domaine, sont ajoutées (A et B sont deux tétraèdres voisins). D'après Durand-Riard [2010].

- l'intégration directe de différents types de données (orientation, direction axiale, données dures) en contraintes d'interpolation (figure 1.4) facilite l'automatisation de la construction des surfaces par rapport à la construction surfacique indirecte [Caumon et al., 2009] ;
- les contacts entre surfaces, horizon-horizon, faille-horizon ou faille-faille, ne sont explicités que lors de la visualisation surfacique du modèle, ce qui rend l'édition et la modification des surfaces plus faciles par rapport à la modélisation explicite, tout en garantissant le maintien des conditions de cohérence des contacts et de validité d'un modèle géologique [Caumon et al., 2004b] ;
- la description des intersections entre surfaces est simplifiée en utilisant des concepts de Géométrie de Construction de Solides. Ainsi, l'intersection d'une surface implicite A ($\mathcal{F}_A(x,y,z) = f_A$) par une surface implicite B ($\mathcal{F}_B(x,y,z) = f_B$) est définie par une intersection booléenne entre A et un demi-espace de B (figure 1.5). Par exemple, la surface tronquée $A_{|B^+}$ est définie par :

$$\mathcal{F}_A(x,y,z) = f_{A_{|\mathcal{F}_B(x,y,z) \geq f_B}} \tag{1.3}$$

C'est dans ce cadre que le modèle stochastique de faille a été développé et que, plus généralement, le projet de thèse s'inscrit. Après avoir souligné l'intérêt de l'approche implicite pour décrire la relation entre objets, la section suivante détaille la structure utilisée pour décrire l'ensemble des relations de filiations d'un réseau de failles.

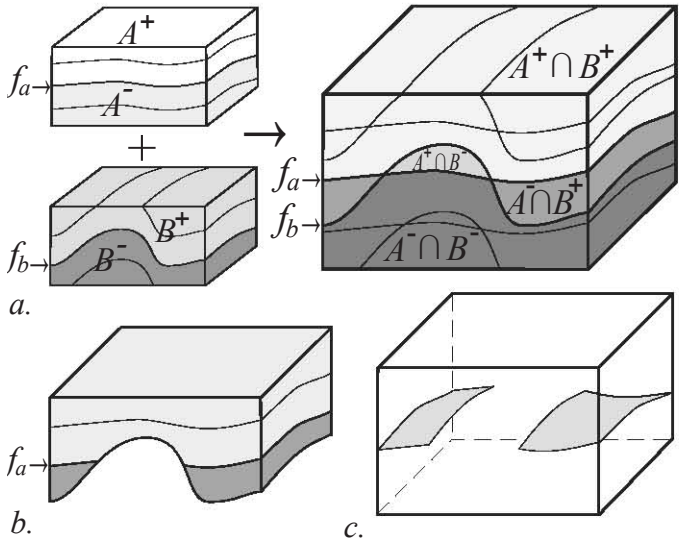

FIGURE 1.5 – Opérations booléennes en modélisation implicite. *a.* Gauche : deux surfaces A ($\mathcal{F}_A(x,y,z) = f_A$) et B ($\mathcal{F}_B(x,y,z) = f_B$) et leurs demi-espaces. Droite : 4 régions spatiales définies par l'intersection des deux champs scalaires. *b.* Exemple d'opération booléenne : $\mathcal{F}_A(x,y,z) \cap \{\mathcal{F}_B(x,y,z) \geq f_B\}$. *c.* Exemple d'intersection : $\mathcal{F}_A(x,y,z) = f_A \cap \{\mathcal{F}_B(x,y,z) \geq f_B\}$ (équation 1.3). D'après Cherpeau et al. [2010a].

1.3 Les arbres binaires comme représentation de la topologie d'un réseau de failles

Cette partie s'appuie sur les travaux publiés dans Cherpeau et al. [2010a], disponibles en annexe A.

1.3.1 Une surface implicite comme partition de l'espace

Une surface implicite A est définie par une fonction volumique $\mathcal{F}_A(x,y,z)$ et une isovaleur f_A de $\mathcal{F}_A$. Tout point $\mathbf{p}(x_p, y_p, z_p)$ de l'espace peut donc être facilement situé par rapport à A, en évaluant $\mathcal{F}_A(x_p, y_p, z_p)$, la valeur de $\mathcal{F}_A$ en $\mathbf{p}$. $\mathcal{F}_A(x,y,z)$ est parfois appelé puissance du point x, y, z par rapport à la surface A, qui doit posséder des faces internes et externes et être orientable.

Toute surface implicite A divise donc l'espace $\mathcal{M}$ en deux parties distinctes B^- et B^+ [Tertois and Mallet, 2006], facilement identifiables :

- $B^- = \{(x,y,z) \in \mathcal{M} | \mathcal{F}_A(x,y,z) < f_A\}$
- $B^+ = \{(x,y,z) \in \mathcal{M} | \mathcal{F}_A(x,y,z) > f_A\}$

Cette relation peut être visualisée sous forme d'arbre binaire, une faille implicite F étant la racine de l'arbre, les deux blocs situés de part et d'autre de la faille représentant les feuilles de l'arbre (figure 1.6). Une feuille correspond donc à un volume dans le modèle et n'a pas de nœuds fils dans l'arbre.

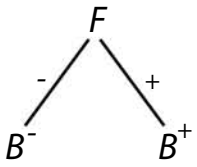

FIGURE 1.6 – Représentation sous forme d'arbre binaire d'une faille F découpant l'espace en deux blocs B^- et B^+.

1.3.2 Représentation graphique des relations de filiation entre failles

Il est également possible de décrire les relations de filiation entre failles de la même manière, la faille principale étant la racine de l'arbre, la faille secondaire (celle qui se branche) une des deux feuilles, selon le côté du branchement. En généralisant, un réseau de failles peut être représenté sous forme d'arbre binaire, comme illustré figure 1.7. Chaque faille correspond à un nœud de l'arbre alors que les feuilles sont des partitions de l'espace.

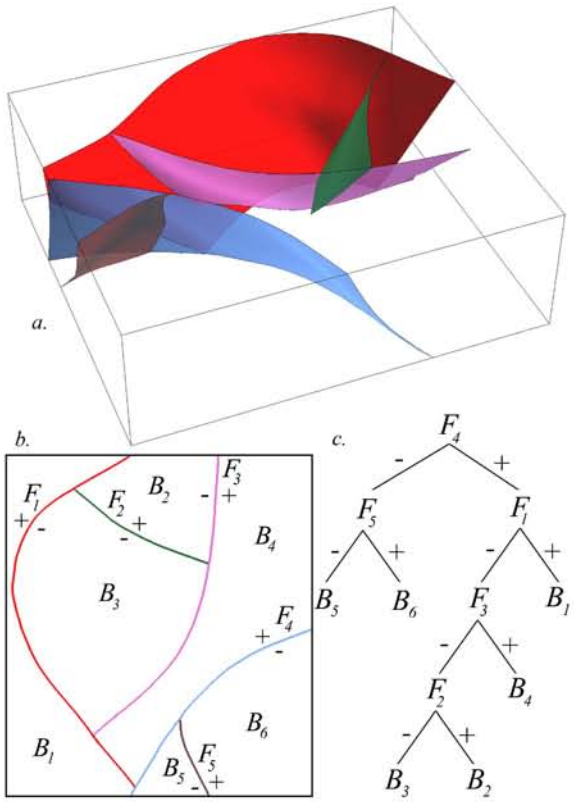

FIGURE 1.7 – Réseau de failles et représentation topologique en arbre binaire associée. *a.* Réseau de failles en 3D. *b.* Le même modèle vu de dessus afin d'expliciter la topologie. Chaque faille F_i coupe le modèle en deux blocs B_i. *c.* Arbre binaire représentant les relations entre failles. La faille F_4 est la faille principale (racine de l'arbre), les autres failles étant des failles secondaires. D'après Cherpeau et al. [2010a].

Cette représentation est analogue aux arbres de partition binaire de l'espace (*BSP tree*

an anglais) utilisés dans l'imagerie numérique pour décrire les scènes 3D. L'arbre est ici utilisé pour décrire une partition de l'espace, non pas avec des plans mais avec des surfaces de faille, non planes par nature.

1.3.3 Propriétés des arbres binaires

1.3.3.1 Informations sur les blocs de faille

Un arbre topologique représente une partition de l'espace Ω, chaque feuille B de l'arbre étant une partie de Ω, l'ensemble des feuilles correspond donc à Ω :

$$\bigcup_{i=1}^{n} B_i = \Omega \tag{1.4}$$

Chaque bloc de faille B est délimité par une ou plusieurs failles appartenant à la branche parente de B dans l'arbre binaire associé. Ainsi, en figure 1.7, le bloc B_3 est délimité par les failles F_1, F_2, F_3 situées en tant que nœuds parents de B_3 dans l'arbre binaire.

1.3.3.2 Informations sur les failles

Chaque faille dans l'arbre binaire se branche potentiellement sur ses failles parentes dans l'arbre. Ceci signifie que si une relation père-fils est inversée dans l'arbre, la relation entre les deux failles correspondantes est également changée : la faille parente devient la faille secondaire et inversement (figure 1.8).

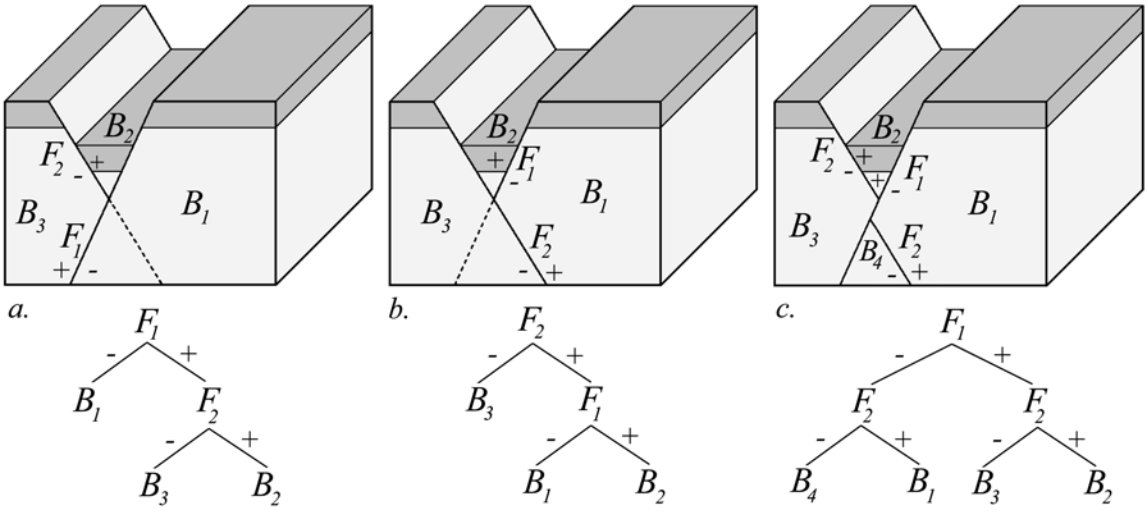

FIGURE 1.8 – Conséquences des relations de filiation dans un arbre binaire d'un point de vue géologique. *a.* La faille F_1 est la racine de l'arbre, F_2 est la faille secondaire. *b.* Inversion de la relation père-fils. F_1 est maintenant la faille secondaire. *c.* Une faille F_2 présente des deux côtés de F_1 signifie que F_2 est recoupée par F_1. F_2 est représentée par des nœuds fils de F_1 dans l'arbre bien qu'elle soit plus ancienne. D'après Cherpeau et al. [2010a].

Cette relation de filiation faille parente-faille secondaire dans l'arbre change dans le cas d'une faille F recoupant des failles plus anciennes $S_{old} = \{F_i | i \in [1, n]\}$. En effet, les failles

S_{old} sont de part et d'autre de la faille F dans le modèle surfacique, ce qui se traduit par des nœuds de faille S_{old} présents dans les deux sous-branches de F dans l'arbre topologique (figure 1.8c).

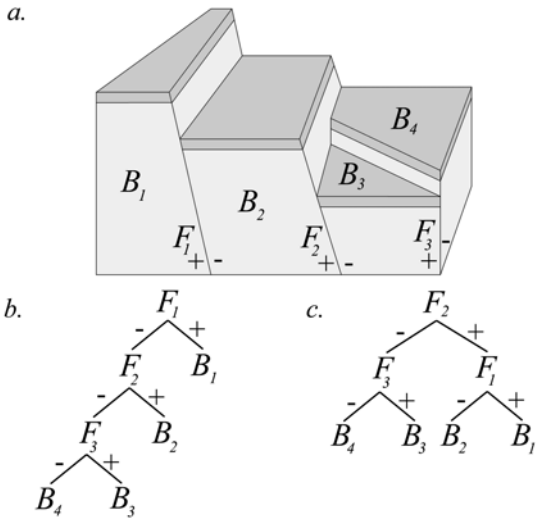

FIGURE 1.9 – Non unicité de la représentation d'un ensemble de failles dans un arbre binaire. *a.* Diagramme 3D représentant 3 failles. Les failles F_1 et F_2 ne sont pas en contact. Par conséquent, elles peuvent être père ou fils dans l'arbre binaire. *b.* Arbre binaire avec F_1 en tant que faille parente. *c.* F_2 est maintenant la racine de l'arbre mais aucun recoupement n'a changé. Il se peut que le recoupement change plus en profondeur, en dehors du diagramme en fonction de la géométrie. D'après Cherpeau et al. [2010a].

Les failles peuvent également être isolées les unes par rapport aux autres. Dans ce cas, la relation de filiation indiquée par leur position dans l'arbre est relative puisque l'inversion de leur position ne change rien d'un point de vue recoupement, les failles n'étant pas en contact. Par conséquent, plusieurs arbres peuvent décrire le même réseau de failles (figure 1.9), chacun déterminant néanmoins les relations entre failles.

1.4 Description d'un réseau de failles

Cette partie a pour but de présenter l'organisation et la mise en place d'un réseau de failles en général afin de les modéliser par la suite.

1.4.1 Familles de failles

Les failles sont induites par des contraintes mécaniques imposées aux roches. Ainsi, quand le seuil de rupture est atteint, une faille se propage de manière à relâcher les contraintes. Plusieurs failles peuvent donc se propager dans les mêmes directions, en réponse à un état de contraintes similaire. Ces failles présentent des caractéristiques communes et nous les regrouperons sous le terme de *famille de failles* dans le reste du mémoire.

1.4.2 Mise en place de familles de failles

Un réseau de failles peut être constitué de plusieurs familles de failles selon l'histoire géologique de la zone étudiée. En effet, plusieurs évènements tectoniques, associés à des états de contraintes différents, peuvent se succéder et chacun engendre une ou plusieurs familles de failles différentes. Deux familles de failles peuvent ainsi être définies sur la carte structurale présentée en figure 1.10, liées à deux évènements tectoniques distincts dans le temps. La seconde famille de failles orientée NO-SE est postérieure et les failles se terminent sur les deux failles antérieures de la première famille (orientée NNE-SSO).

La mise en place d'une zone de faille en relai crée éventuellement deux familles de failles selon la maturité de la zone. Dans un premier temps, des segments de faille se mettent en place et constituent une première famille (figure 1.11*a*). Ensuite, lorsque le déplacement augmente, des fractures en échelon se développent dans la zone de recouvrement mais peuvent évoluer en failles connectant les segments déjà en place (figure 1.11*b*,*c* et *d*). D'un point de vue descriptif, ces failles de connexion représentent une seconde famille.

Les familles de failles peuvent donc se mettre en place suite à des évènements tectoniques indépendants et séparés dans l'échelle de temps géologique, comme illustré en figure 1.10. Elles peuvent également se mettre en place dans un laps de temps plus restreint correspondant à un seul évènement majeur comme dans le cas d'une zone de faille en relai. Une rampe mature contient une seule faille irrégulière d'un point de vue structural. Cependant, dans un but de représentation et de modélisation, nous adoptons une démarche plus descriptive et considérons deux familles de failles aux caractéristiques distinctes.

1.4.3 Relations entre familles de failles

Nous avons vu qu'il existe une relation de filiation entre différentes familles de failles, les failles postérieures venant se brancher sur les failles déjà en place. Cette topologie particulière se retrouve dans l'arbre topologique associé à un réseau de failles. Ainsi, les failles de la famille antérieure se situent en haut de l'arbre et ne sont jamais des nœuds fils de failles postérieures, sauf quand une faille postérieure recoupe une faille antérieure (figure 1.8*c*).

Lorsque deux familles de failles se mettent en place lors d'un même évènement tectonique, par exemple des failles conjuguées en contexte extensif, il n'y a pas de règle de recoupement précise, les failles d'une famille jouant successivement le rôle de faille majeure ou secondaire par rapport à l'autre famille de failles. Nous regroupons ces familles de failles sous le terme de famille *cogénétique* pour signifier qu'elles se sont mises en place suite à un même évènement et qu'il n'y a pas de règle de recoupement entre elles. Ces failles alternent donc entre nœuds fils et père dans l'arbre topologique (cas de la famille B dans la figure 1.12). Dans la suite du mémoire, le terme de famille cogénétique correspond à un ensemble de sous-familles cogénétiques, en particulier pour les parties relatives à la méthode de simulation.

Les relations entre familles de failles et leur position relative dans l'arbre topologique sont résumées dans la figure 1.12.

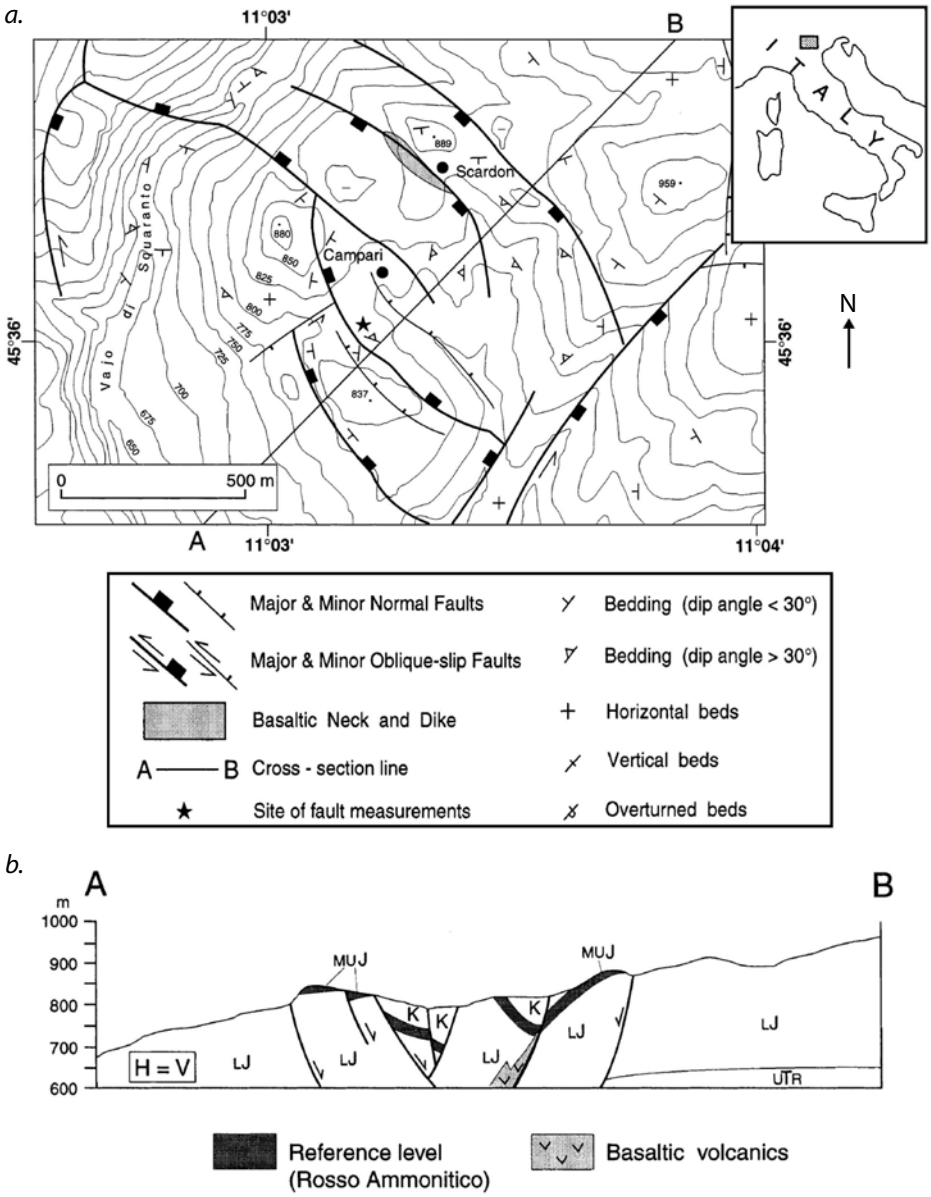

FIGURE 1.10 – Carte structurale avec différentes familles de failles dans les Alpes du Sud au Nord de l'Italie. *a.* En première approximation, deux familles de failles peuvent être définies, la première représentant les deux failles orientées NNE-SSO dans la partie Sud-Est de la carte. La deuxième famille correspond aux failles conjuguées orientées NO-SE, formant un graben. *b.* Coupe AB située sur la carte, suggérant que les failles de la deuxième famille pendant vers le Nord-Ouest ont un pendage plus faible que celles pendant vers le Sud-Est. Dans ce cas, il peut être intéressant de séparer ces failles en deux familles distinctes pour affiner leur description, ce qui donne trois familles au total. Modifié d'après Zampieri [2000].

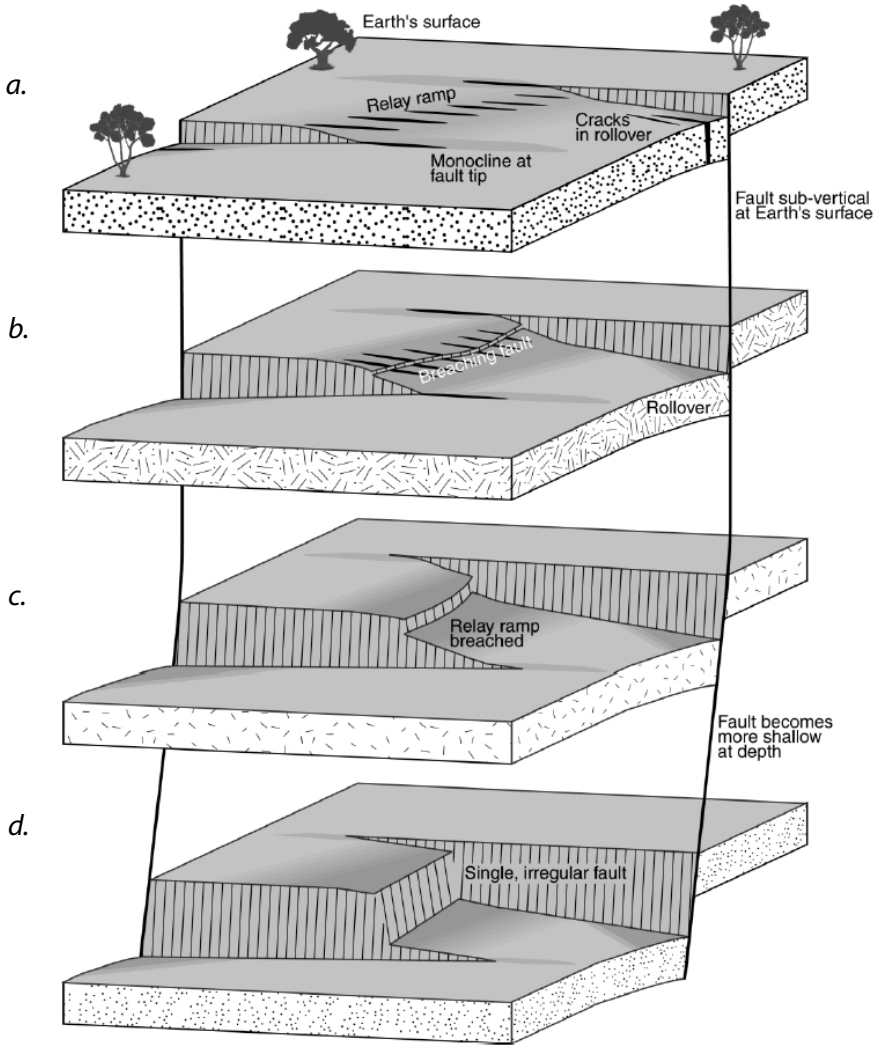

FIGURE 1.11 – Évolution spatiale et temporelle d'une zone de faille en relai. *a.* Des segments de faille se recouvrent et sont connectés par des fractures en échelon. *b.* La rampe se développe, le déplacement augmente ce qui provoque le développement d'une faille de connexion. *c.* La rampe continue de se fracturer et le déplacement de la faille connectant les deux segments augmente. *d.* Finalement, les deux segments sont connectés et on peut considérer une seule faille irrégulière. Cependant, d'un point de vue descriptif, deux familles de failles peuvent être identifiées, les segments et les failles de connexion entre les segments. D'après Peacock and Parfitt [2002].

1.5 Algorithme général et données d'entrée

1.5.1 Algorithme général

Les relations entre familles de failles décrites précédemment permettent de définir un algorithme de simulation d'un réseau de failles qui honore ces relations.

L'algorithme 1 est très général, il reflète néanmoins la mise en place d'un réseau de failles dans le temps. Plusieurs éléments restent à définir :

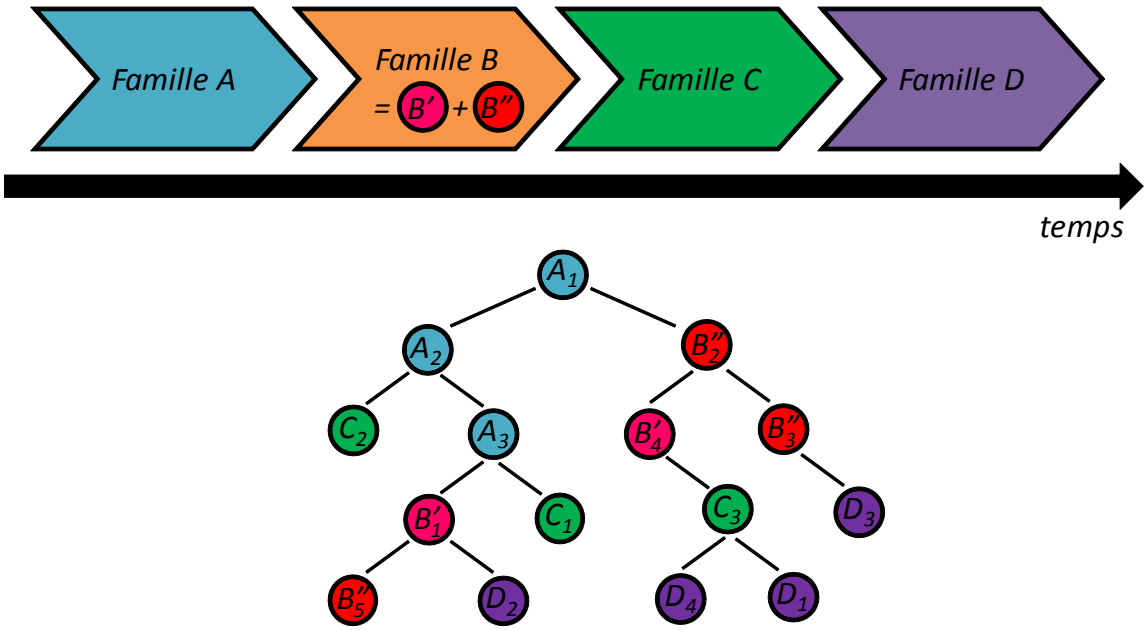

FIGURE 1.12 – Relations entre familles de failles dans un arbre topologique. Quatre familles de failles se succèdent dans le temps, A étant la plus ancienne, D la plus récente. La famille B est une famille cogénétique regroupant deux sous-familles B' et B''. Les failles d'une famille postérieure sont toujours des nœuds fils des familles antérieures dans l'arbre. En revanche, pour B' et B'', il n'y a pas de règle de filiation dans l'arbre.

- Comment est caractérisée une famille de failles ? (section 1.5.2)
- Comment modéliser un objet faille ? (section 1.6)

Dans une moindre mesure :

- Pour une famille de failles cogénétique, quelle stratégie utiliser pour sélectionner une sous-famille à simuler (fonction SélectionSousFamille) ? (section 1.5.5)
- Quels sont les critères d'arrêt de la simulation d'une famille de failles ? (section 1.5.4)

Ces quatre éléments modulent le comportement de l'algorithme. Nous allons voir dans la suite de ce chapitre la réponse choisie à ces quatre questions, en commençant par la caractérisation d'une famille de failles.

1.5.2 Paramètres d'une famille de failles

1.5.2.1 Des distributions statistiques pour caractériser les failles

Nous avons vu qu'une famille de failles englobe des failles similaires. Chaque famille de failles est définie par des distributions statistiques définissant les caractéristiques communes des failles :

- pendage ;
- azimuth ;
- extension latérale, dans la direction azimutale ;
- extension en profondeur dans la direction du pendage.

$\mathcal{E}$ = ensemble de familles de failles
pour *chaque famille de failles* $\mathcal{S}$ *dans* $\mathcal{E}$*, de la plus ancienne à la plus récente* **faire**
  si $\mathcal{S}$ *est cogénétique* **alors**
    $S^{liste} \leftarrow$ sous-familles de failles de $\mathcal{S}$
    tant que $S^{liste} \neq \emptyset$ **faire**
      $\mathcal{S}' \leftarrow$ `SélectionSousFamille(`S^{liste}`)` `(section 1.5.5)`
      simuler une nouvelle faille F de S' `(section 1.6)`
      si *critère d'arrêt valide* `(section 1.5.4)` **alors**
        retirer $\mathcal{S}'$ de S^{liste}
      finsi
    fin
  sinon
    tant que *critère d'arrêt non valide* **faire**
      simuler une nouvelle faille F de $\mathcal{S}$ `(section 1.6)`
    fin
  finsi
fin

Algorithme 1: Algorithme général de simulation d'un réseau de failles.

L'orientation des surfaces est donnée par un couple *(azimuth, pendage)* défini par "la règle de la main droite", c'est-à-dire que l'azimuth d'une faille est définie quand celle-ci est inclinée vers la droite (*e.g.* N30E, 45SE équivaut à (30,45), N30E, 45NW à (210,45)). Cette notation a l'avantage d'être facilement manipulable de manière informatique.

Ces paramètres définissent l'orientation et la dimension des failles. D'autres paramètres caractérisent la sinuosité des failles (figure 1.13) :

- longueur d'onde de la sinuosité dans la direction azimutale ;
- longueur d'onde de la sinuosité dans la direction du pendage ;
- amplitude de la sinuosité ;
- angle de glissement par rapport à la direction du pendage, lié à la composante décrochante du rejet.

Un nombre théorique de failles est également associé à chaque famille de failles. Celui-ci sert de condition d'arrêt de la simulation lorsqu'il n'y a pas d'autres critères définis.

1.5.2.2 Intérêt des familles cogénétiques

Les failles conjuguées orientées NO-SE sur la carte structurale en figure 1.10*a* ne peuvent pas être représentées par des familles séparées puisque les règles de recoupement ne sont pas constantes. En revanche, elles peuvent être simulées par une seule et même famille de failles, en utilisant une distribution bimodale (un pôle centré en N135, l'autre en N315) pour caractériser l'azimuth. Cependant, les informations fournies par la coupe (figure 1.10*b*) suggèrent que le pendage des failles qui pendent vers le Nord-Ouest est plus faible que le pendage des failles pendant vers le Sud-Est. Dans ce cas, il n'est pas possible de représenter cette différence en distribution bimodale puisque des combinaisons non observées seraient alors possibles avec la distribution bimodale de l'azimuth. Ces failles doivent donc être représentées par deux familles distinctes, regroupées dans une famille cogénétique pour

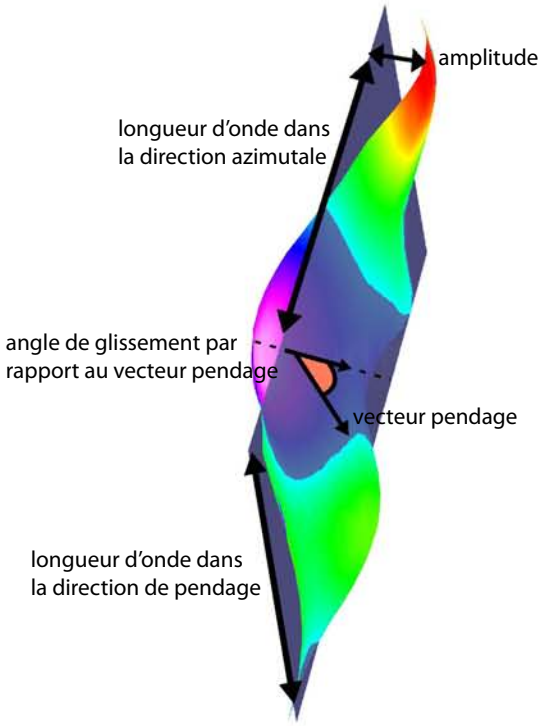

FIGURE 1.13 – Paramètres décrivant la sinuosité d'une faille. Le plan violet sert uniquement à définir les différents paramètres.

garantir l'alternance et le caractère aléatoire de leur recoupement.

1.5.3 Gestion des données

1.5.3.1 Des ensembles de points comme données

La méthode prend en entrée des ensembles de points représentant les failles observées. Ces ensembles peuvent provenir de l'interprétation de profils sismiques, de photos aériennes, ou provenir de relevés de terrain par exemple. L'utilisateur interprétateur pointe les failles visibles et crée ainsi autant d'ensembles de points que de failles observées.

1.5.3.2 Incertitude des points

Nous l'avons vu lors de l'introduction, les données comportent des incertitudes qui peuvent être dues à une imprécision du pointage sur un profil sismique, à la résolution limitée des données interprétées, au géo-référencement approximatif d'un relevé de terrain par exemple. Une incertitude est donc associée à chaque point de donnée pour prendre en compte de telles imprécisions. L'incertitude peut correspondre à une sphère de position autour du point de référence associée à une densité de probabilité radiale. Une valeur d'incertitude différente pour chaque point permet d'intégrer au plus près les observations, comme par exemple une incertitude croissante avec la profondeur lors de l'interprétation d'une faille sur un profil sismique (figure 1.14).

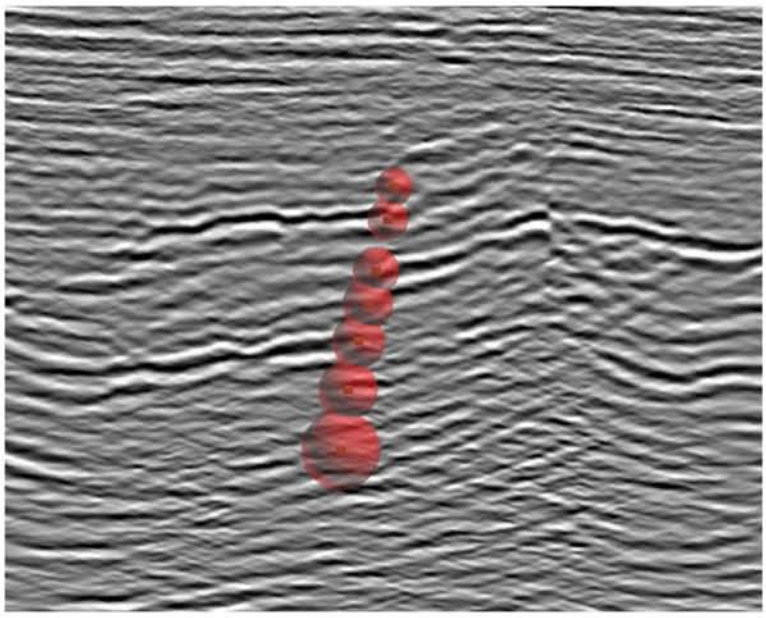

FIGURE 1.14 – Incertitude lors du pointage d'une faille sur un profil sismique. Les points orangés représentent la faille pointée sur le profil. Les sphères rouges associées à chaque point représentent l'incertitude du pointage, croissante avec la profondeur car la faille est de plus en plus difficilement localisable.

1.5.3.3 Ensembles de points et familles de failles

L'interprétation d'une faille en un ensemble de points peut suffire à caractériser l'orientation globale de la faille et donc définir la famille de failles à laquelle les points appartiennent. D'autres indices, comme le rejet de part et d'autre d'une faille peuvent indiquer le type de faille, normale ou inverse, et donc restreindre les familles de failles possibles pour un ensemble de points donné. Dans ce cas, cette information est attachée à l'ensemble de points, c'est-à-dire qu'une liste de familles de failles est fournie en entrée avec cet ensemble de points et seules ces familles peuvent honorer ces points. Par défaut, cette liste contient toutes les familles de failles définies en entrée.

1.5.3.4 Information attachée aux ensembles de points

L'observation du rejet, au moins vertical, permet de mieux caractériser les failles, notamment leur taille puisqu'une relation est généralement admise entre l'extension et le rejet d'une faille [Kim and Sanderson, 2005]. Cette information peut donc être précisée pour chaque point et est utilisée lors de la simulation pour éviter des cas irréalistes (pour plus de détails, voir la section 1.6.7).

1.5.4 Critères d'arrêt de la simulation

Deux conditions sont nécessaires (mais non suffisantes) pour l'arrêt de la simulation d'une famille de failles $\mathcal{F}$:

- le nombre de failles simulées doit être au moins égal au nombre de failles défini en entrée pour $\mathcal{F}$;
- tous les ensembles de points ne pouvant être honorés que par une faille de $\mathcal{F}$ doivent avoir été attribués.

Lorsque ces deux conditions sont remplies, l'algorithme 2 détermine alors une probabilité d'arrêt p_{arret} qui dépend du nombre d'ensembles de points restants à honorer $nbpts_{global}$, du nombre de failles restantes $nf_{restant} = nf_{global} - nf_{global_{sim}} - n$ pouvant honorer ces points et du rapport entre le nombre de failles théorique $nf_{\mathcal{F}}$ et déjà simulées $nf_{\mathcal{F}_{sim}}$ de $\mathcal{F}$:

$$\begin{aligned} p_{arret} &= 1 \text{ si } \frac{nf_{\mathcal{F}_{sim}}}{nf_{\mathcal{F}}} \times \frac{nf_{restant}}{nbpts_{global}} > 1 \\ &= \frac{nf_{\mathcal{F}_{sim}}}{nf_{\mathcal{F}}} \times \frac{nf_{restant}}{nbpts_{global}} \text{ sinon} \end{aligned} \tag{1.5}$$

1.5.5 Gestion des familles cogénétiques

Une famille de failles cogénétique $\mathcal{F}_{cogenetique} = \{\mathcal{F}_1, ...\mathcal{F}_n\}$ est un ensemble de n familles supposées s'être mis en place lors du même évènement tectonique. Lors de la simulation d'une famille cogénétique, l'algorithme alterne aléatoirement la simulation de failles des différentes sous-familles. Plusieurs stratégies peuvent être envisagées pour calculer la probabilité p_i de chaque sous-famille $\mathcal{F}_i$ d'être sélectionnée pour simuler la prochaine faille. Le calcul de p_i peut par exemple dépendre du nombre d'ensembles de points à honorer pour la famille $\mathcal{F}_i$. La probabilité p_i peut également être inversement proportionnelle au ratio du nombre de failles déjà simulées sur le nombre théorique de failles de $\mathcal{F}_i$. Dans ce cas, les familles ayant proportionnellement simulé peu de failles sont favorisées. D'autres combinaisons sont possibles. Dans notre cas, l'algorithme attribue simplement une probabilité uniforme à l'ensemble des sous-familles : $p_i = \frac{1}{n}$.

$\mathcal{F}$ = famille de failles courante

$\mathcal{E}_{\mathcal{F}}$ = liste d'ensembles de points ne pouvant être attribués qu'à $\mathcal{F}$
$nbpts_{\mathcal{F}}$ = nombre d'ensembles de points ne pouvant être attribués qu'à $\mathcal{F}$
$nf_{\mathcal{F}}$ = nombre théorique de failles de $\mathcal{F}$
$nf_{\mathcal{F}_{sim}}$ = nombre de failles de $\mathcal{F}$ déjà simulées

$\mathcal{E}_{global}$ = liste d'ensembles de points pouvant être attribués à une ou plusieurs familles de failles , dont $\mathcal{F}$ ($\mathcal{E}_{\mathcal{F}} \subseteq \mathcal{E}_{global}$)
$nbpts_{global}$ = nombre d'ensembles de points pouvant être attribués à une ou plusieurs familles de failles , dont $\mathcal{F}$
$\mathcal{L}_{\mathcal{F}}$ = liste de familles de failles pouvant honorer $\mathcal{E}_{global}$
nf_{global} = nombre global de failles correspondant à la somme de tous les nombres de failles de $\mathcal{L}_{\mathcal{F}}$
$nf_{global_{sim}}$ = nombre global de failles déjà simulées de $\mathcal{L}_{\mathcal{F}}$

si $nf_{\mathcal{F}_{sim}} < nf_{\mathcal{F}}$ *ou* $nbpts_{\mathcal{F}} > 0$ **alors**
| continuer la simulation
sinon si $nbpts_{global} = 0$ **alors**
| arrêt de la simulation
finsi
$n \leftarrow 0$
pour chaque *famille de failles* $\mathcal{F}_i$ *de* $\mathcal{L}_{\mathcal{F}}$ **faire**
| $n \leftarrow$ ajouter le nombre d'ensembles de points ne pouvant être attribués qu'à $\mathcal{F}_i$
fin
si $nf_{global} - nf_{global_{sim}} < n$ **alors**
| continuer la simulation
sinon
| $p_{arret} \leftarrow \frac{nf_{\mathcal{F}_{sim}}}{nf_{\mathcal{F}}}$ `(supérieur à 1)`
| $p_{arret} \leftarrow p_{arret} \times \frac{nf_{global} - nf_{global_{sim}} - n}{nbpts_{global}}$
| **si** $p_{arret} > 1$ **alors**
| | arrêter la simulation
| **sinon**
| | arrêt de la simulation avec une probabilité p_{arret}
| **finsi**
finsi

Algorithme 2: Algorithme gérant l'arrêt de la simulation d'une famille de failles.

1.6 Modélisation d'un objet faille

Cette partie décrit la simulation d'un objet faille, en détaillant les étapes qui mènent à l'ajout d'une nouvelle faille dans le modèle courant.

1.6.1 Entrées

- **Un ensemble de familles de failles** avec des paramètres concernant la position, l'orientation, la sinuosité et la dimension des failles.
- **Le réseau de failles courant** : la simulation des failles est séquentielle, c'est-à-dire qu'elles sont simulées dans l'ordre chronologique relatif donné par l'âge des familles de failles. Le réseau de failles s'enrichit d'une faille à chaque nouvelle simulation et le modèle de sortie obtenu après la simulation d'une faille devient le modèle d'entrée pour la prochaine faille.
- **La liste des ensembles de points non attribués** : au cours de la simulation, les ensembles de points honorés par une faille sont retirés de la liste.

1.6.2 Aperçu des étapes de simulation d'un objet faille

Un objet faille est ajouté au modèle en 8 étapes :

1. Tirage aléatoire de points de données à honorer ;
2. Tirage aléatoire de l'orientation (azimuth et pendage) ;
3. Tirage aléatoire du centre de faille ;
4. Calcul d'une première géométrie grâce aux paramètres définis précédemment ;
5. Attribution de nouveaux points de données selon leur cohérence, proximité avec la géométrie initiale ;
6. Mise à jour de la géométrie de la faille afin d'honorer les nouveaux points attribués et de modéliser la sinuosité de la faille ;
7. Modélisation de l'extension de la faille ;
8. Détermination des intersections avec les failles déjà simulées.

Les sections suivantes détaillent chacune des 8 étapes permettant de modéliser un objet faille. Le lecteur pourra se référer à la figure 1.15 pour un aperçu de chaque étape et de l'ensemble de la simulation d'un objet faille.

1.6.3 Attribution des points de données

Plusieurs paramètres sont à prendre en compte lors de l'attribution (ou la non-attribution) d'un ensemble de points à la prochaine faille à simuler ; la méthode est décrite par l'algorithme 3. L'algorithme prend en compte les autres failles pouvant honorer les mêmes jeux de points, ce qui permet de ne pas les attribuer en priorité aux premières failles simulées.

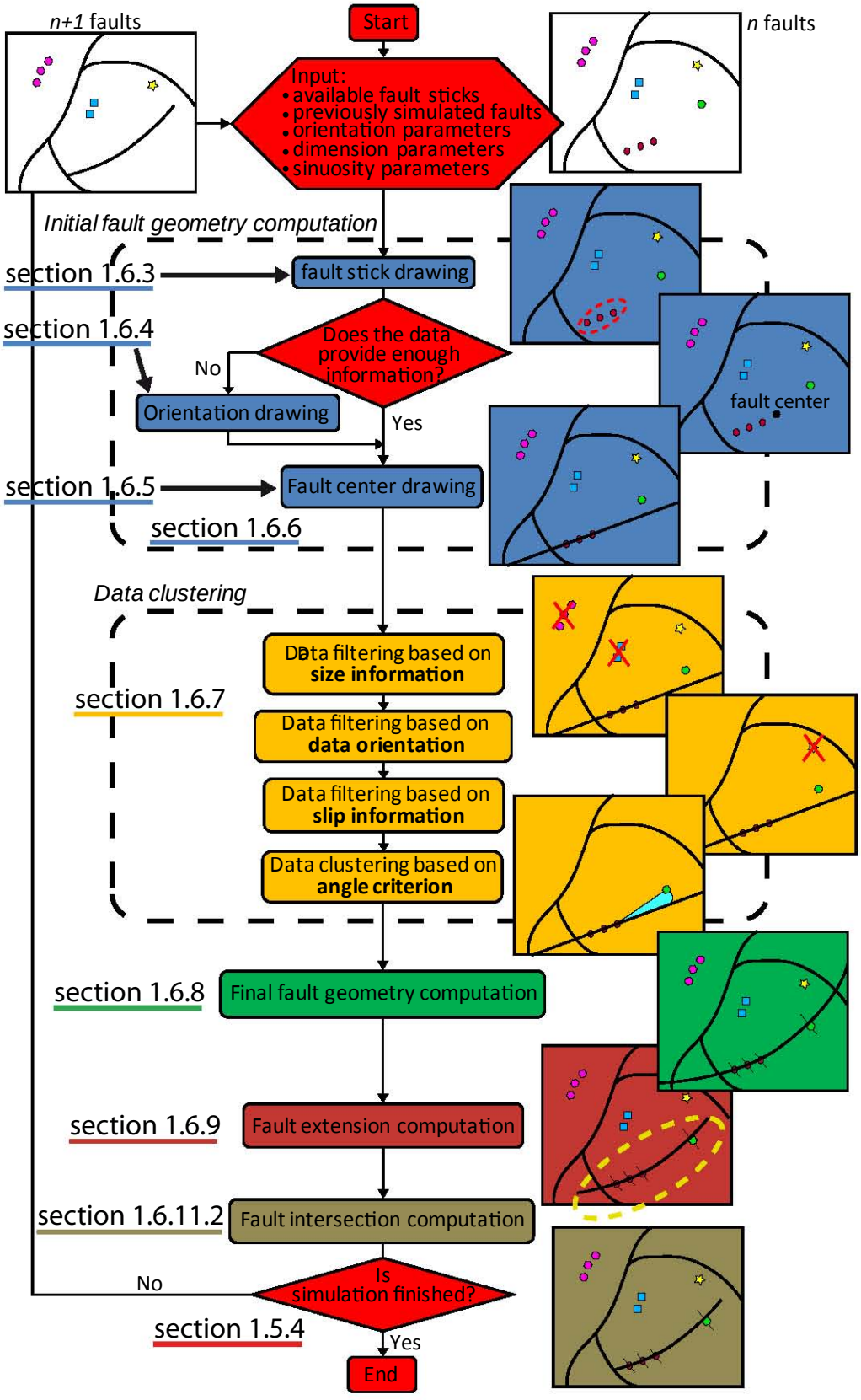

FIGURE 1.15 – Aperçu des étapes de simulation d'un objet faille.

$\mathcal{F}$ = famille de failles à simuler

$\mathcal{E}_{\mathcal{F}}$ = liste d'ensembles de points ne pouvant être attribués qu'à $\mathcal{F}$
$nbpts_{\mathcal{F}}$ = nombre d'ensembles de points ne pouvant être attribués qu'à $\mathcal{F}$
$nf_{\mathcal{F}}$ = nombre théorique de failles de $\mathcal{F}$
$nf_{\mathcal{F}_{sim}}$ = nombre de failles de $\mathcal{F}$ déjà simulées

$\mathcal{E}_{global}$ = liste d'ensembles de points pouvant être attribués à une ou plusieurs familles de failles , dont $\mathcal{F}$ ($\mathcal{E}_{\mathcal{F}} \subseteq \mathcal{E}_{global}$)
$nbpts_{global}$ = nombre d'ensembles de points pouvant être attribués à une ou plusieurs famille de failles , dont $\mathcal{F}$
$\mathcal{L}_{\mathcal{F}}$ = liste de familles de failles pouvant honorer $\mathcal{E}_{global}$
nf_{global} = nombre global de failles correspondant à la somme de tous les nombres de failles de $\mathcal{L}_{\mathcal{F}}$
$nf_{global_{sim}}$ = nombre global de failles déjà simulées de $\mathcal{L}_{\mathcal{F}}$

si $nf_{\mathcal{F}} - nf_{\mathcal{F}_{sim}} < nbpts_{\mathcal{F}}$ **alors**
  retourner aléatoirement un ensemble de points de $\mathcal{E}_{\mathcal{F}}$
sinon
  retourner aléatoirement un ensemble de points de $\mathcal{E}_{global}$, avec une probabilité $p = \max(0, 1 - \frac{nbpts_{global}}{nf_{global} - nf_{global_{sim}}})$ de ne rien retourner
finsi

Algorithme 3: Algorithme gérant l'attribution des points de données à une faille

1.6.4 Orientation de la future faille

Deux cas de figure existent selon le résultat de la première étape de simulation. Si aucun point de donnée n'est attribué à la faille, le pendage et l'azimuth sont tirés aléatoirement dans les lois de distribution de la famille de failles correspondante. En revanche, si un ensemble de points S a été attribué à la faille, celui-ci contient déjà une information d'orientation. Comme cette information peut ne pas être suffisante (par exemple dans le cas de points alignés), l'algorithme calcule le plan médian de S. Si l'orientation du plan n'est pas contenue dans l'intervalle défini par les distributions statistiques, une nouvelle orientation est alors tirée aléatoirement. En revanche, si l'orientation moyenne est correcte eu égard aux données d'entrée, l'orientation de la faille correspond alors à celle de S.

1.6.5 Centre de faille

Le centre de faille représente le point de nucléation à partir duquel la faille s'est propagée. C'est un paramètre important puisqu'il détermine l'emplacement et le voisinage de la future faille F.

1.6.5.1 Restriction sur la position des centres de faille

Si aucun ensemble de points S n'a été attribué à la future faille F, son centre c_F peut, a priori, se situer n'importe où dans l'espace 3D. Au contraire, si un ensemble S a été attribué à F, la probabilité de c_F est nulle au delà d'un ellipsoïde $e(g_S, r_1, r_2, r_3)$ et en dehors du

plan de faille $p(g_S, pendage_F, azimuth_F)$, avec g_S le barycentre de S. Les axes r_1, r_2, r_3 sont calculés à partir de la moyenne des distributions d'entrée caractérisant la longueur et la hauteur de la famille de failles de F. La région $\mathcal{R}$ correspondant aux positions possibles de c_F est définie ainsi :

$$\begin{aligned} \mathcal{R} &= e(g_S, r_1, r_2, r_3) \cap p(g_S, pendage_F, azimuth_F) \text{ si } S \neq \emptyset \\ &= \Omega \text{ sinon} \end{aligned} \tag{1.6}$$

La connaissance exhaustive et multi-échelle des distributions spatiales des contraintes et du comportement mécanique des roches permettrait de prédire la position de nucléation des failles. Cependant, cette information ne peut être approchée que ponctuellement à l'état actuel et reste sujette à des incertitudes importantes. Sans connaissance *a priori* de la zone étudiée, le centre c_F est tiré aléatoirement avec une probabilité uniforme dans $\mathcal{R}$. Il existe cependant des moyens de définir cette probabilité, selon les données disponibles et le contexte géologique, par exemple si des attributs structuraux ont été préalablement calculés.

1.6.5.2 Probabilité de nucléation

Probabilité de nucléation à partir d'attributs sismiques
La présence de failles peut être probabilisée par des attributs sismiques, en mesurant la perte locale de cohérence du signal sismique [Bahorich and Farmer, 1995, Marfurt et al., 1998], le changement local de pendage-azimuth [Rijks and Jauffred, 1991] ou la courbure des réflecteurs sismiques. Chaque position de l'espace a alors une valeur de probabilité représentant la présence possible d'une faille.

Probabilité de nucléation à partir d'observations
Les études de réseaux de failles à l'affleurement permettent d'approcher la mise en place d'un réseau de failles et donc la nucléation de nouvelles failles. Ackermann and Schlische [1997] observent deux familles de failles sur des affleurements dans la carrière Solite située dans le bassin de type rift de Danville (Californie) (figure 1.16). La première famille correspond aux failles les plus grandes distribuées uniformément. La deuxième famille correspond à des failles plus petites qui sont éloignées des failles de la première famille. La zone d'exclusion correspond à un ellipsoïde autour des failles de la première famille. La position des failles antérieures a donc influencé la nucléation des failles de la deuxième famille.

Ackermann et al. [2001] montrent grâce à des modèles expérimentaux que la répartition des centres des failles évolue au cours de la mise en place d'un système de failles normales entre aléatoire, groupé et uniforme. Clifton and Schlische [2001] simulent la mise en place d'une zone de rift grâce à un modèle expérimental. Ils observent qu'au cours de l'extension du système, les nouvelles failles se développent près des failles existantes. Après l'analyse d'une carte structurale du Yorkshire du Sud en Angleterre, Munthe et al. [1994] en déduisent que les failles ont tendance à se regrouper mais que les failles majeures semblent être légèrement répulsives.

Ces observations, basées sur l'analyse de données de terrain ou de modèles analogiques, montrent que la probabilité d'occurrence de failles, si elle dépend initialement de l'état de contraintes des roches (sans indication, la probabilité peut être considérée

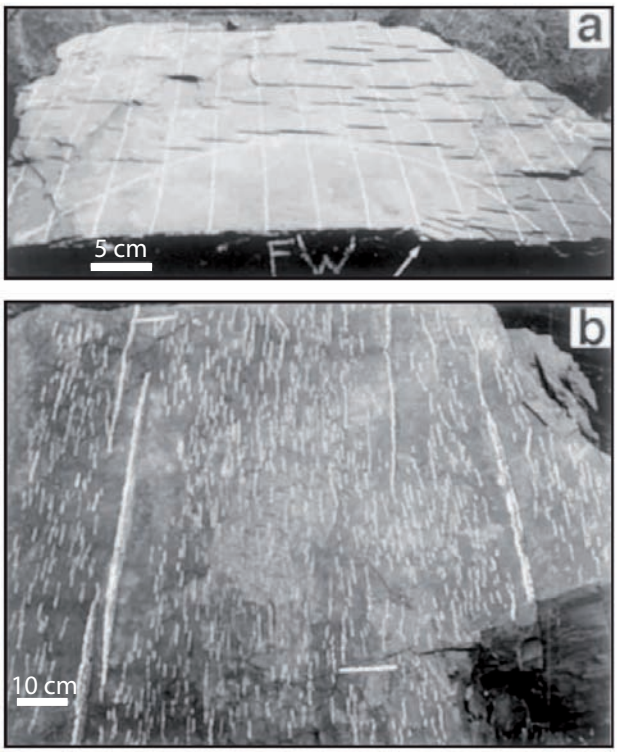

FIGURE 1.16 – Éloignement et répulsion entre failles. Haut : le mur de la faille principale est indiqué en bas de la photo. Les failles secondaires se sont mises en place à l'écart de la faille principale, la zone d'exclusion formant un ellipsoïde. Bas : les zones d'exclusion autour des failles majeures sont nettement visibles (vue de dessus). D'après Ackermann and Schlische [1997].

uniforme), évolue ensuite avec la mise en place des premières failles. En effet, les failles existantes perturbent le champ de contraintes et modifient l'emplacement des futures failles.

Prise en compte de l'interaction entre failles lors de la nucléation

Afin de prendre en compte les interactions possibles entre failles (répulsion ou attraction), une grille régulière 3D de dimension n_u, n_v, n_w est définie et couvre l'ensemble du domaine. Une probabilité d'interaction $p_{int}(u, v, w)$ est ensuite définie dans cette grille. Sans information, la probabilité est uniforme ($p_{int}(u, v, w) = \frac{1}{n_u \times n_v \times n_w}$ pour chaque cellule de la grille). Les interactions entre failles sont prises en compte avec le champ de distances euclidiennes $\mathcal{D}$ des familles de failles concernées :

$$\forall u \in [1, n_u], v \in [1, n_v], w \in [1, n_w] :$$

$$p_{int}(u,v,w) = \frac{1}{k} \sum_{i=1}^{m} \omega_i \times \frac{d_i}{\max(\mathcal{D}_i^{\|f_i\|}) - \min(\mathcal{D}_i^{\|f_i\|})}$$

avec :

- k une constante de normalisation ($\sum_{uvw} p_{int}(u,v,w) = 1$)
- m le nombre de familles de failles influençant la nucléation de la faille courante
- ω_i le poids relatif de la famille i
- f_i la puissance du champ de distance (contrôle le degré d'attraction/répulsion)
- $d_i = \mathcal{D}_i(u,v,w)^{\|f_i\|} - \min(\mathcal{D}_i^{\|f_i\|})$ si $f_i > 0$ (répulsion)
  $= \max(\mathcal{D}_i^{\|f_i\|}) - \mathcal{D}_i(u,v,w)^{\|f_i\|}$ si $f_i < 0$ (attraction)

(1.7)

La distance minimale $\min(\mathcal{D}_i)$ est nulle lorsque les cellules de la grille sont intersectées par une faille (étape de rasterisation). La relation précédente s'écrit alors sous forme simplifiée :

$$\begin{aligned} p_{int}(u,v,w) &= \frac{1}{k} \sum_{i=1}^{m} \omega_i \times \frac{\mathcal{D}_i(u,v,w)^{\|f_i\|}}{\max(\mathcal{D}_i^{\|f_i\|})} \text{ si } f_i > 0 \text{ (répulsion)} \\ &= \frac{1}{k} \sum_{i=1}^{m} \omega_i \times \left(1 - \frac{\mathcal{D}_i(u,v,w)^{\|f_i\|}}{\max(\mathcal{D}_i^{\|f_i\|})}\right) \text{ si } f_i < 0 \text{ (attraction)} \end{aligned} \tag{1.8}$$

L'équation 1.7 permet de prendre en compte des interactions avec plusieurs familles de failles et de leur attribuer un poids relatif ω_i. La valeur de $\omega_i(u,v,w)$ peut être choisie constante dans tout le domaine (figure 1.17*a.*) mais il peut aussi être intéressant d'utiliser une pondération par la distance afin de mieux prendre en compte la proximité des failles environnantes (figure 1.17*b.*) :

$$\forall u,v,w : \omega_i(u,v,w) = 1 - \frac{\mathcal{D}_i(u,v,w)}{\sum_{j=1}^{m} \mathcal{D}_j(u,v,w)} \tag{1.9}$$

Le degré d'interaction est modélisé par le coefficient f dont le signe définit l'attraction (négatif) ou la répulsion (positif) (figure 1.17*b.*). Les failles d'une même famille peuvent également interagir. Le champ de distances euclidiennes utilisé est alors calculé à partir des failles ajoutées précédemment au modèle. La distance euclidienne est calculée grâce à l'algorithme de Saito and Toriwaki [1994] implémenté par Ledez [2003]. Après la simulation d'une faille F, le champ de distances de la famille correspondante $\mathcal{D}_\mathcal{F}$ est mis à jour en prenant la valeur minimale entre $\mathcal{D}_\mathcal{F}$ et la distance euclidienne à F ($\mathcal{D}_F$) :

$$\forall u,v,w : \mathcal{D}_\mathcal{F}(u,v,w) = \min(\mathcal{D}_F(u,v,w), \mathcal{D}_\mathcal{F}(u,v,w)) \tag{1.10}$$

Cette opération évite de recalculer les distances aux failles antérieures, seules les cellules plus proches de F sont mises à jour.

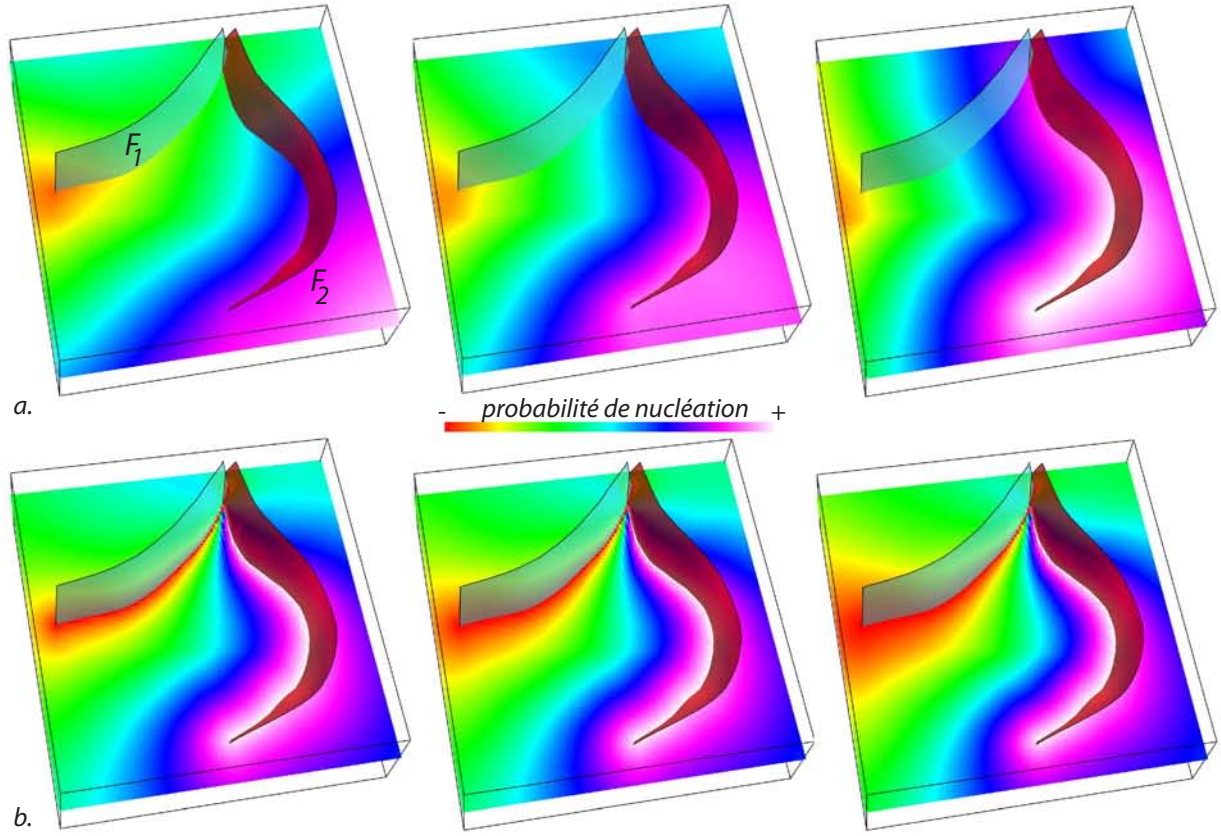

FIGURE 1.17 – interactions entre failles et probabilité de nucléation. Les cartes représentent la probabilité de nucléation d'une troisième faille dans le modèle. La probabilité a été calculée en tenant compte d'une répulsion avec la faille F_1 et d'une attraction avec F_2.
a. La probabilité est calculée avec un poids ω constant dans tout le domaine. De gauche à droite : $\omega_{F_1} = 0.7$ et $\omega_{F_2} = 0.3$; $\omega_{F_1} = \omega_{F_2} = 0.5$; $\omega_{F_1} = 0.3$ et $\omega_{F_2} = 0.7$.
b. La probabilité est calculée avec un poids ω pondéré par la distance à F_1 et F_2 (équation 1.9). De gauche à droite : $f_{F_1} = 0.8$, $f_{F_1} = 1$, $f_{F_1} = 1.3$ ($f_{F_2} = -1$ dans les 3 cas). La répulsion est de plus en plus forte quand la puissance du champ de distance f augmente.

La méthode n'utilise pas de carte initiale de probabilité (*e.g.* dérivée d'attributs sismiques). Un moyen de prendre en compte cette information serait de pondérer la probabilité de nucléation p_{nuc} par la carte de probabilité m et la probabilité d'interaction p_{int} :

$$\forall u, v, w : p_{nuc}(u, v, w) = \lambda m(u, v, w) + (1 - \lambda) p_{int}(u, v, w) \tag{1.11}$$

1.6.6 Géométrie initiale

À ce stade de la simulation, la future faille est caractérisée par son orientation (pendage et azimuth), son centre et éventuellement un ensemble de points. Il est donc possible de modéliser une première géométrie à partir de ces informations. Si aucune donnée n'a été attribuée à la faille, la géométrie initiale correspond à un plan passant par le centre de faille et orienté selon le pendage et l'azimuth. La fonction volumique scalaire représentant la faille

est simplement la distance signée à ce plan.

La construction du champ scalaire est différente dans le cas où des données ont été attribuées à la faille. La construction utilise des contraintes DSI décrites en figure 1.4 (page 18) afin d'honorer à la fois les points de données et l'orientation de la faille lors de l'interpolation du champ scalaire.

1.6.7 Regroupement d'ensembles de points

Lors de l'interprétation des données brutes, une même faille peut être interprétée sur différents supports (topographie, coupe interprétative, section sismique, forage, etc.), ou être interprétée sur le même support en différents endroits, sans que l'appartenance au même objet des différents points interprétés soit évidente pour l'interprétateur. Un jeu de données peut donc contenir plusieurs ensembles de points appartenant en réalité à la même faille.

L'étape de regroupement a pour but d'attribuer, en plus de l'ensemble de points initialement honoré par la faille courante F, des ensembles de points supplémentaires à F. Tous les ensembles de points non attribués sont alors considérés comme candidats au regroupement. Cette étape est divisée en deux sous-étapes :

- une première étape de filtrage écarte les points incohérents d'après les règles décrites ci-après ;
- de nouveaux points sont ensuite attribués aléatoirement à F parmi les candidats restants.

Cette étape est détaillée au chapitre 2, section 2.2.3.2 (page 67), seuls les principes sont rappelés ici. L'étape de filtrage des candidats est basée sur trois critères : l'éloignement des points candidats au centre de F ; leur orientation par rapport à F et la compatibilité de la valeur maximale de déplacement attaché aux points candidats. La probabilité d'attribution est inversement proportionnelle à l'angle formé par les points candidats, la surface F et les points déjà honorés par F.

1.6.8 Géométrie finale

Le but de cette étape est triple :

- donner à la faille courante F une sinuosité en utilisant les paramètres d'entrée de la famille de F ;
- honorer les ensembles de points nouvellement attribués à F lors de l'étape de regroupement ;
- prendre en compte l'incertitude attachée aux points de données (initialement et nouvellement attribués à F).

1.6.8.1 Méthode de perturbation

L'étape consiste à perturber la géométrie de la faille courante F représentée par le champ scalaire $\mathcal{F}(x, y, z)$ et i_F l'isovaleur correspondant à la surface F. Le champ $\mathcal{F}(x, y, z)$ est perturbé en ajoutant un bruit aléatoire corrélé $\mathcal{R}(x, y, z)$ calculé en fonction des trois objectifs définis ci-dessus. Cette méthode de perturbation de surface implicite a été proposée par Caumon et al. [2007] pour perturber la géométrie d'horizons stratigraphiques d'un modèle de référence. Le champ de perturbation $\mathcal{R}(x, y, z)$ est simulé en pratique par une

Simulation Séquentielle Gaussienne. Le variogramme utilisé est gaussien afin d'obtenir une perturbation lisse. Les autres paramètres du variogramme (portée, orientation) sont établis à partir des données de sinuosité. La distribution cible est gaussienne, centrée en zéro (perturbation moyenne nulle) et l'écart-type est égal au tiers de l'amplitude de la sinuosité de sorte que 99.7% des valeurs de perturbation sont inférieures à l'amplitude maximale définie en entrée. La prise en compte d'une anisotropie locale permettrait de modéliser une sinuosité variable.

1.6.8.2 Incertitude des points de données

La simulation du champ de perturbation est conditionnée aux points de données en prenant en compte leur incertitude. La valeur de $\mathcal{R}(x,y,z)$ en chaque point de donnée $pt(x_{pt},y_{pt},z_{pt})$, avec une incertitude u_{pt} (en unité de distance) est définie ainsi :

$$\begin{aligned} &\mathcal{R}(x_{pt},y_{pt},z_{pt}) = i_F - \mathcal{F}(x_{pt},y_{pt},z_{pt}) + u \\ &\quad \text{avec :} \\ &\qquad \cdot\ u \leftarrow \mathcal{T}(0,\mathcal{F}(x_b,y_b,z_b)) \\ &\qquad \cdot\ b = pt + u_{pt} \times \bigtriangledown\mathcal{F}(x_{pt},y_{pt},z_{pt}) \end{aligned} \tag{1.12}$$

La valeur u représente l'incertitude (en valeur du champ de $\mathcal{F}$) et est tirée aléatoirement dans une loi triangulaire symétrique $\mathcal{T}$ centrée en zéro dont les extrema sont définis par la valeur de $\mathcal{F}$ au point b. Le calcul du point b permet de passer d'une incertitude en unité de distance (u_{pt}) à une incertitude en valeur de $\mathcal{F}$ ($\mathcal{F}(x_b,y_b,z_b)$).

L'équation 1.12 se simplifie pour les points de données pt déjà honorés par F. En effet, $\mathcal{F}(x_{pt},y_{pt},z_{pt}) = i_F$ donc la valeur du champ de perturbation en pt est simplement égale à l'incertitude u.

Le tirage aléatoire de u dans $\mathcal{T}$ est corrélé à travers les différents points de données afin d'obtenir une perturbation cohérente. La méthode utilise un champ de probabilité [Srivastava, 1992, Lecour et al., 2001] pour fournir les nombres aléatoires utilisés lors du tirage de u.

1.6.8.3 Perturbation 2D *vs.* perturbation 3D

Le champ de perturbation $\mathcal{R}(x,y,z)$ doit être généré en tout point (x,y,z) du domaine. La simulation séquentielle gaussienne peut être effectuée directement dans cet espace en trois dimensions (figure 1.18). Cependant, le but est de perturber la géométrie d'une surface (définie par $\mathcal{F}(x_b,y_b,z_b) = i_F$) et l'ajout de $\mathcal{R}(x,y,z)$ peut éventuellement modifier sa topologie (*e.g.* création de "bulles"). Une deuxième méthode de génération de $\mathcal{R}(x,y,z)$ a été proposée (figure 1.19) : (1) une grille régulière 2D $\mathcal{G}$ est définie parallèlement au plan de faille ; (2) le champ $\mathcal{R}(x,y,z)$ est généré dans $\mathcal{G}$; puis (3) transféré dans le reste du domaine par projection. Cette méthode a également l'avantage d'être moins coûteuse en temps de calcul puisque le champ de perturbation est seulement généré en 2 dimensions dans une grille régulière (la définition du voisinage est ainsi simplifiée).

Cette méthode a également l'avantage de lier directement les paramètres de sinuosité d'une faille présentée en figure 1.13) avec les paramètres variographiques utilisés pour générer le champ de perturbation $\mathcal{R}(x,y,z)$.

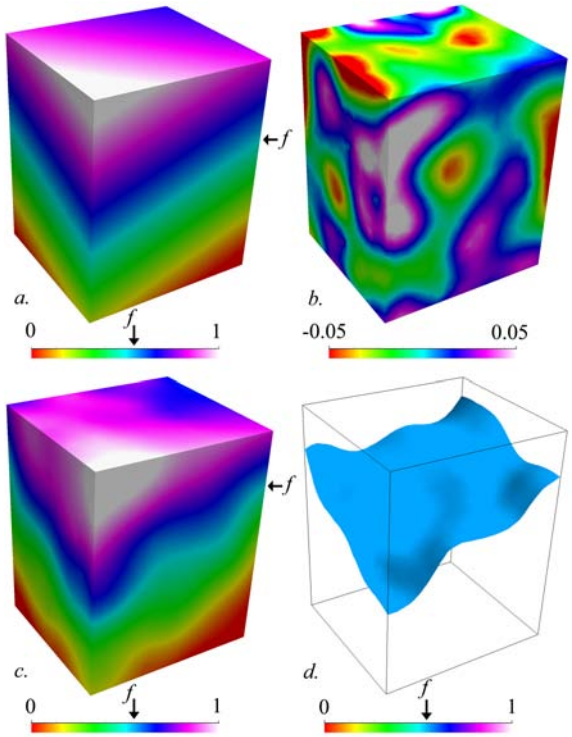

FIGURE 1.18 – Méthode de perturbation 3D d'une surface implicite. *a.* Le champ scalaire de référence $\mathcal{F}(x, y, z)$ définit la surface initiale (isovaleur f, surface plane). *b.* Un champ aléatoire corrélé généré par une Simulation Séquentielle Gaussienne correspond au champ de perturbation $\mathcal{R}(x, y, z)$. *c.* La surface est maintenant définie par l'isovaleur f dans le champ $\mathcal{F}(x, y, z) + \mathcal{R}(x, y, z)$. *d.* Surface perturbée.

1.6.8.4 Exemples de perturbation

La figure 1.20 montre des exemples de perturbation d'une faille initiale (1.20*a.*) dans le cas d'incertitudes faibles (1.20*b.*) et fortes (1.20*c.*).

1.6.9 Extension latérale

1.6.9.1 Faille elliptique

Les failles sont des objets géologiques finis par nature : leurs terminaisons latérale et verticale s'opèrent quand la contrainte mécanique devient inférieure au seuil de rupture des roches. Il n'y a qu'en modélisation que la notion de faille "infinie" prend un sens, pour signifier qu'une faille traverse complètement l'espace 3D $\mathcal{M}$ de la zone étudiée, comme par exemple une faille représentée par une surface implicite ($\mathcal{F}(x, y, z) = f$).

Tertois [2007] propose de limiter l'extension d'une faille par un polygone construit à partir de l'enveloppe convexe des données. Le polygone est projeté sur la surface implicite et des tests d'intersections avec le maillage tétraédrique permettent alors de limiter l'extension de la faille. Dans ce travail, une autre approche a été retenue : un sous-domaine $\mathcal{D} \subset \mathcal{M}$ est défini afin de restreindre l'existence d'une faille :

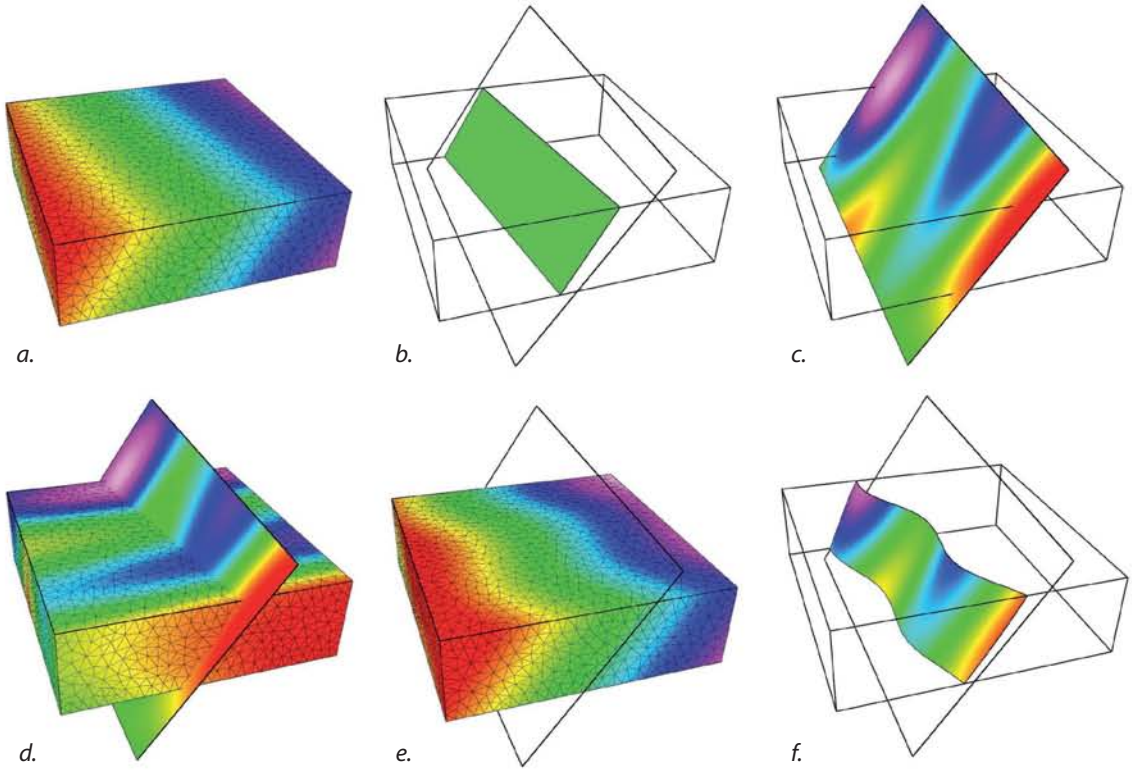

FIGURE 1.19 – Méthode de perturbation 2D d'une surface implicite. Champ scalaire initial $\mathcal{F}(x, y, z)$ $(a.)$ et surface correspondante $(b.)$. $c.$ Le champ de perturbation $\mathcal{R}(x, y, z)$ est généré sur une grille 2D parallèle au plan de faille. $d.$ La perturbation est transférée à l'ensemble du domaine par projection. $e.$ Le champ $\mathcal{R}(x, y, z)$ est ajouté au champ de la faille $\mathcal{F}(x, y, z)$. $f.$ Nouvelle géométrie de la faille.

$$\mathcal{F}(x, y, z) = f \ \forall (x, y, z) \in \mathcal{D} \subseteq \mathcal{M} \tag{1.13}$$

La frontière Ω_D du domaine $\mathcal{D}$ est définie par une isovaleur du champ scalaire de $\mathcal{D}$ $(\mathcal{D}(x, y, z) = d)$, la faille existe seulement d'un côté de Ω_D, par exemple :

$$\mathcal{F}(x, y, z) = f_{|\mathcal{D}(x,y,z) \leq d} \tag{1.14}$$

Les failles isolées ont une forme circulaire à elliptique [Barnett et al., 1987, Walsh et al., 2003], la frontière Ω_D est donc définie par un ellipsoïde centré sur le centre de faille et dont les axes v_1, v_2, v_3 sont définis ainsi :

- v_1 est horizontal et parallèle à la direction d'azimuth de la faille ;
- v_2 est parallèle au vecteur pendage ;
- $v_3 = v_1 \times v_2$ est le vecteur orthogonal à v_1 et v_2.

Les normes des vecteurs v_1 et v_2 sont tirées aléatoirement dans les distributions statistiques définies en entrée. La géométrie très elliptique d'une faille dont l'extension verticale a été restreinte dans une unité stratigraphique [Benedicto et al., 2003] peut être approchée en choisissant $\|v_1\| >> \|v_2\|$.

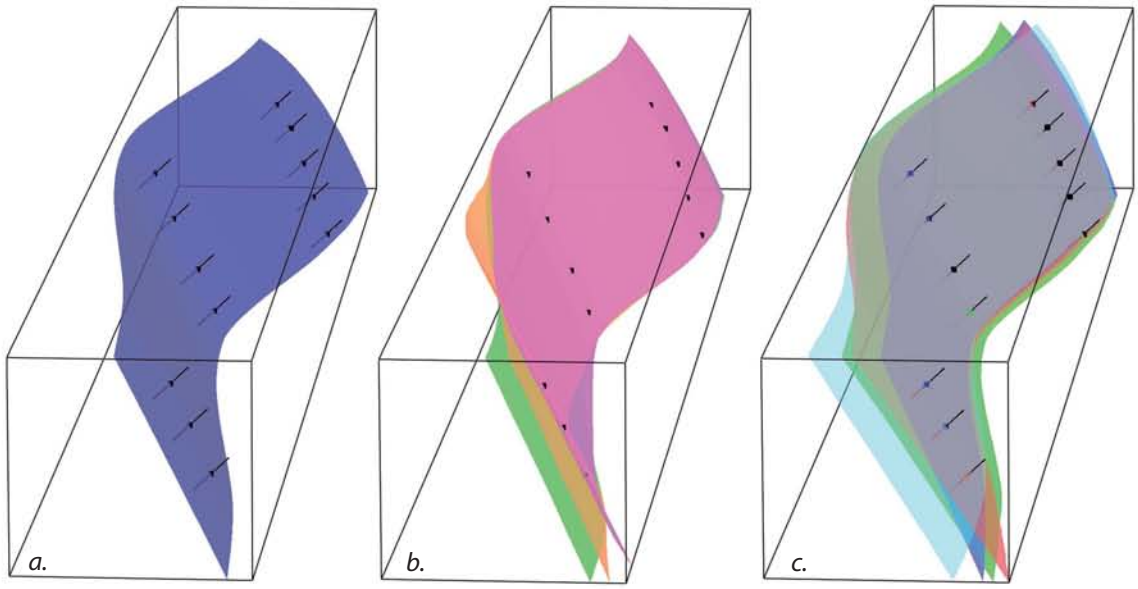

FIGURE 1.20 – Exemples de perturbation géométrique. *a.* Faille initiale, l'incertitude est représentée par des vecteurs d'incertitude normaux à la faille, dont la norme est calculée à partir de l'incertitude attachée aux points de données. *b.* Différentes réalisations en considérant une incertitude nulle aux points de données. Seule l'incertitude sur la variabilité de la faille est échantillonnée. *c.* Les différentes réalisations prennent en compte les vecteurs d'incertitude. La variabilité des géométries est plus importante qu'en *b* car la méthode échantillonne également l'incertitude des données.

Cette approche paraît plus simple à implémenter que celle proposée par Tertois [2007]. Elle engendre cependant un coût de stockage supplémentaire puisqu'un second champ scalaire est nécessaire pour définir une faille. Elle permet toutefois de définir un domaine en 3D qui pourrait par exemple être utilisé pour définir la zone de déplacement non nul lié à la faille, en vue d'une modélisation du rejet (pour plus de détails, voir la partie Perspectives 1.7.2.1 de ce chapitre). L'ellipsoïde pourrait également servir de zone de nucléation interdite pour des failles secondaires, de manière à approcher les observations faites par Ackermann and Schlische [1997] (figure 1.16 page 36).

1.6.9.2 Adaptation de l'arbre topologique

La notion d'arbre topologique présentée en section 1.3 est étendue afin de prendre en compte l'extension limitée des failles. En effet, une faille F_A peut se situer des deux côtés d'une faille existante F_B. Dans ce cas, F_A est représentée par deux nœuds dans l'arbre, situés dans les deux sous-branches du nœud de F_B (figure 1.21).

1.6.10 Choix de l'ordre des étapes de modélisation d'une faille

Les étapes décrites dans les sections précédentes permettent de modéliser un objet faille, leur ordre est rappelé ici :

1. Tirage aléatoire de points de données à honorer ;

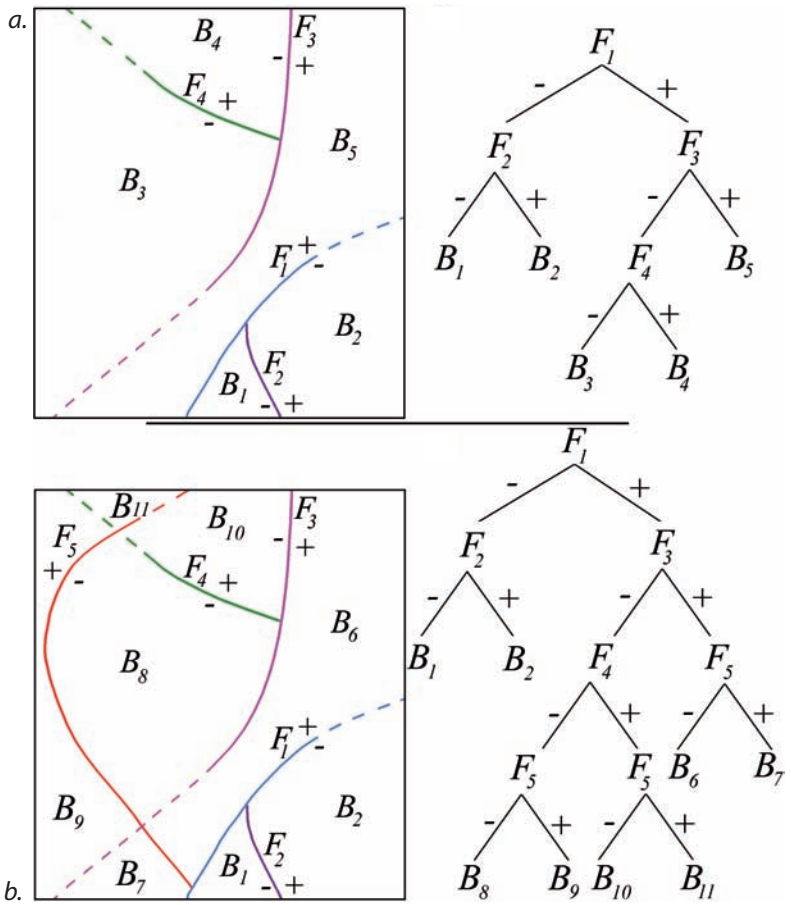

FIGURE 1.21 – Arbre topologique et failles d'extension finie (vue en carte). *a.* Un réseau de failles en 2D et l'arbre topologique correspondant. Les lignes pointillées représentent les failles en dehors de leur domaine d'existence. *b.* Une faille F_5 est ajoutée au modèle. Elle se situe de part et d'autre des failles F_3 et F_4. F_5 remplace les blocs B_3, B_4 et B_5 dans l'arbre.

2. Tirage aléatoire de l'orientation (azimuth et pendage) ;
3. Tirage aléatoire du centre de faille ;
4. Calcul d'une première géométrie grâce aux paramètres définis précédemment ;
5. Attribution de nouveaux points de données selon leur cohérence, proximité avec la géométrie initiale ;
6. Mise à jour de la géométrie de la faille afin d'honorer les nouveaux points attribués et modéliser la sinuosité de la faille ;
7. Modélisation de l'extension de la faille ;
8. Détermination des intersections avec les failles déjà simulées.

L'ordre de certaines de ces étapes pourrait être changé. L'extension de la faille pourrait par exemple être simulée avant l'attribution de nouveaux points de données. Dans ce cas, les conditions de regroupement seraient plus strictes car plus de points seraient en dehors de l'ellipsoïde limitant l'extension de la faille. En effet, l'étape actuelle utilise la dimen-

sion maximale comme condition de filtrage, ce qui permet de sélectionner plus de points candidats.

De même, l'attribution de nouveaux points pourrait être favorisée en précédant l'étape de nucléation (étape 3). Dans ce cas, l'ellipsoïde de recherche de points candidats au regroupement serait mobile autour des points déjà attribués, ce qui permettrait un balayage plus large. En revanche, le tirage aléatoire du centre de faille serait plus contraint car moins de positions seraient alors possibles. Le choix a été fait de favoriser le centre de faille car il apparaît être plus important en déterminant l'emplacement de la future faille. Ce choix a également été motivé dans l'optique d'ajout de contraintes supplémentaires lors du tirage du centre de faille, comme par exemple favoriser la nucléation aux endroits où un déplacement important a été observé (pour plus de détails, voir la partie Perspectives 1.7).

Ce choix rejoint celui de tirer aléatoirement un ensemble de points à honorer lors de la première étape. Ces choix ont pour but de favoriser les observations dans un premier temps, pour contraindre par la suite la simulation des paramètres plus incertains.

1.6.11 Mise à jour du modèle

Lorsqu'une faille vient d'être simulée, elle est ajoutée au réseau de failles déjà simulées. Il reste néanmoins à assurer la cohérence générale en vérifiant que la faille ne recoupe aucun ensemble de points non attribué et en évaluant les intersections avec les failles existantes.

1.6.11.1 Conflits possibles avec des points de données

La construction de la géométrie de la faille ne tient pas compte des ensembles de points non attribués. La faille nouvellement construite peut donc passer à travers un ensemble de points qui a pourtant été interprété comme appartenant à une seule faille. L'option la plus simple consiste à séparer l'ensemble de points intersecté en deux parties distinctes en considérant qu'ils appartiennent à deux failles différentes. Pourtant, dans certains cas, une légère perturbation de la géométrie de la faille pourrait éviter que la faille ne recoupe ces points. L'algorithme calcule l'angle α correspondant à la déviation nécessaire de la surface de faille et attribue ainsi un seuil de déviation $s(\alpha)$ d'autant plus élevé que l'angle est faible (figure 1.22). Un tirage aléatoire détermine ensuite si la faille est déviée ou non : si la valeur tirée est inférieure au seuil, alors la faille est déviée sinon les points sont séparés en deux sous-ensembles.

La distribution du seuil de déviation $s(\alpha)$ est identique pour toutes les familles de failles. Elle pourrait cependant être adaptée à chaque famille, en étant d'autant plus stricte (déviation moins probable) que l'incertitude sur l'orientation des failles de la famille est faible.

Le champ scalaire de la faille à dévier est réinterpolé en ajoutant une contrainte d'interpolation sur les points à éviter de sorte que l'isovaleur correspondant à la faille soit inférieure ou supérieure aux valeurs des points à éviter. Il est alors nécessaire de répéter l'étape de génération de la géométrie finale. Si la nouvelle géométrie engendre de nouveaux conflits, les points situés de part et d'autre de la faille sont alors directement séparés en deux sous-ensembles.

La figure 1.23 montre des exemples obtenus en utilisant le seuil de déviation $s(\alpha)$.

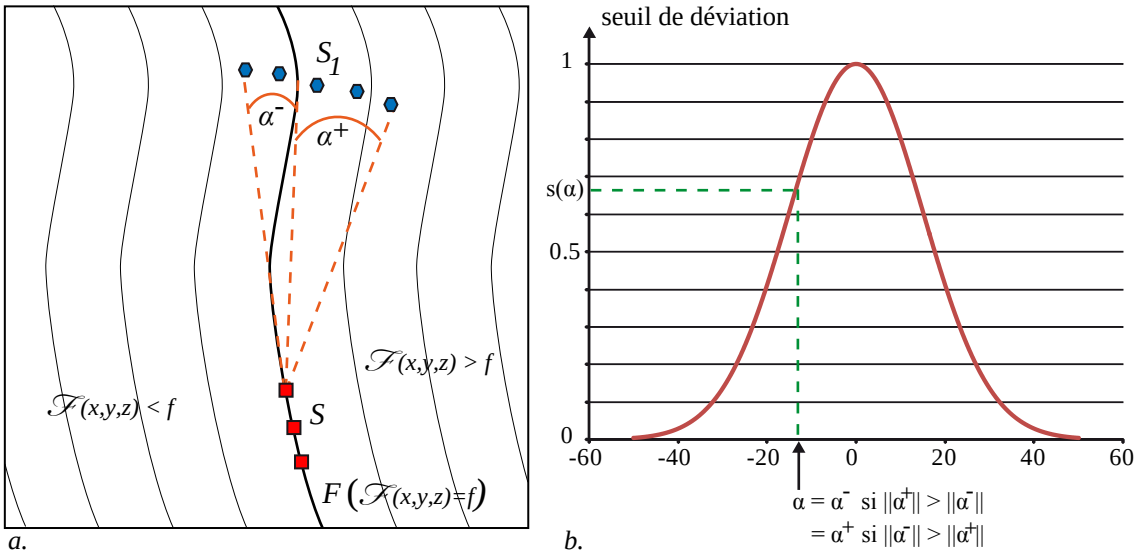

FIGURE 1.22 – Calcul du seuil de déviation d'une faille. *a.* L'algorithme calcule les angles maximum α^- et α^+ de part et d'autre de la faille. *b.* L'angle α correspond à l'angle dont la valeur absolue est la plus petite. L'algorithme attribue alors un seuil de déviation $s(\alpha)$ d'après une loi normale $\mathcal{N}(0, 15)$ mise à l'échelle.

Cette approche nécessite de réinterpoler le champ scalaire de la faille et de simuler une nouvelle perturbation dans le cas où la faille est déviée. Une autre approche consisterait à simuler une nouvelle perturbation avec des contraintes aux points à éviter, de sorte que la somme du champ scalaire de la faille et du champ de perturbation soit supérieure ou inférieure à l'isovaleur de la faille.

1.6.11.2 Intersections

Cas possibles

L'algorithme simule les failles dans un ordre chronologique relatif, la dernière faille simulée A peut donc se brancher sur les failles existantes (*i.e.* simulées précédemment) ou recouper des failles plus anciennes. L'intersection de A ($\mathcal{F}_A(x, y, z) = f_A$ dans le domaine $\mathcal{D}_A$) avec une faille B existante ($\mathcal{F}_B(x, y, z) = f_B$ dans le domaine $\mathcal{D}_B$) est définie par :

$$I_{A-B} : \mathcal{F}_A(x, y, z) = f_A \cap \mathcal{F}_B(x, y, z) = f_B \quad \forall (x, y, z) \in \mathcal{D}_A$$

Il existe 4 configurations d'intersection (figure 1.24) :
- cas 1 : $I_{A-B} = \emptyset$ signifie que A ne recoupe pas B et existe seulement d'un côté de B.
- cas 2 : $I_{A-B} \neq \emptyset$ et $I_{A-B} \cap \mathcal{D}_B = \emptyset$ signifie que A et B ne se coupent pas mais la faille A est définie des deux côtés de B.
- cas 3 : $I_{A-B} \cap \mathcal{D}_B \neq \emptyset$ et $I_{A-B} \notin \mathcal{D}_B$: A recoupe partiellement B.
- cas 4 : $I_{A-B} \subset \mathcal{D}_B$: A se branche sur B.

Le cas 3 correspondant à une intersection partielle a reçu peu d'intérêt dans la littérature quant à son réalisme géologique. L'algorithme rapporte cette situation au cas 2 ou 4 avec une probabilité pour l'un ou l'autre définie arbitrairement par l'utilisateur.

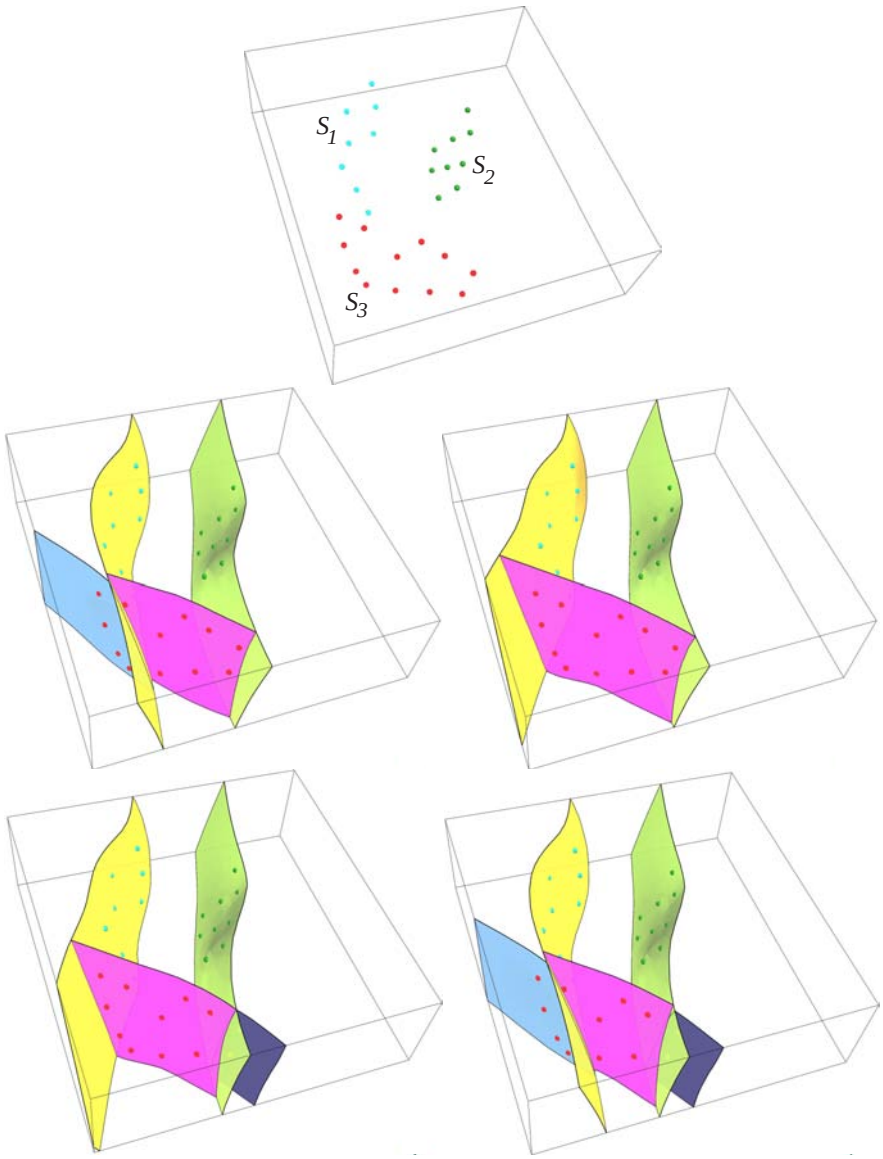

FIGURE 1.23 – Résultats montrant l'effet de la probabilité de déviation. 3 ensembles de points ont été définis et correspondent a priori à 3 failles distinctes. Les failles honorant S_1 et S_2 sont les failles majeures. Elles recoupent S_3 et sont déviées selon la probabilité définie en figure 1.22. L'algorithme génère aléatoirement l'une des 4 configurations.

Algorithme de recherche d'intersections

La méthode la plus simple et naïve pour trouver les intersections possibles entre la faille dernièrement simulée F et les autres failles est de tester l'intersection de F avec chacune d'entre elles. L'opération de test d'intersection étant relativement coûteuse en temps de calcul, l'algorithme retenu utilise l'arbre topologique associé au réseau de failles.

Soit R la région dans laquelle se trouve le centre de F. L'algorithme 4 commence par tester les intersections de F avec les failles définissant R. Pour la faille F_5 dans la figure

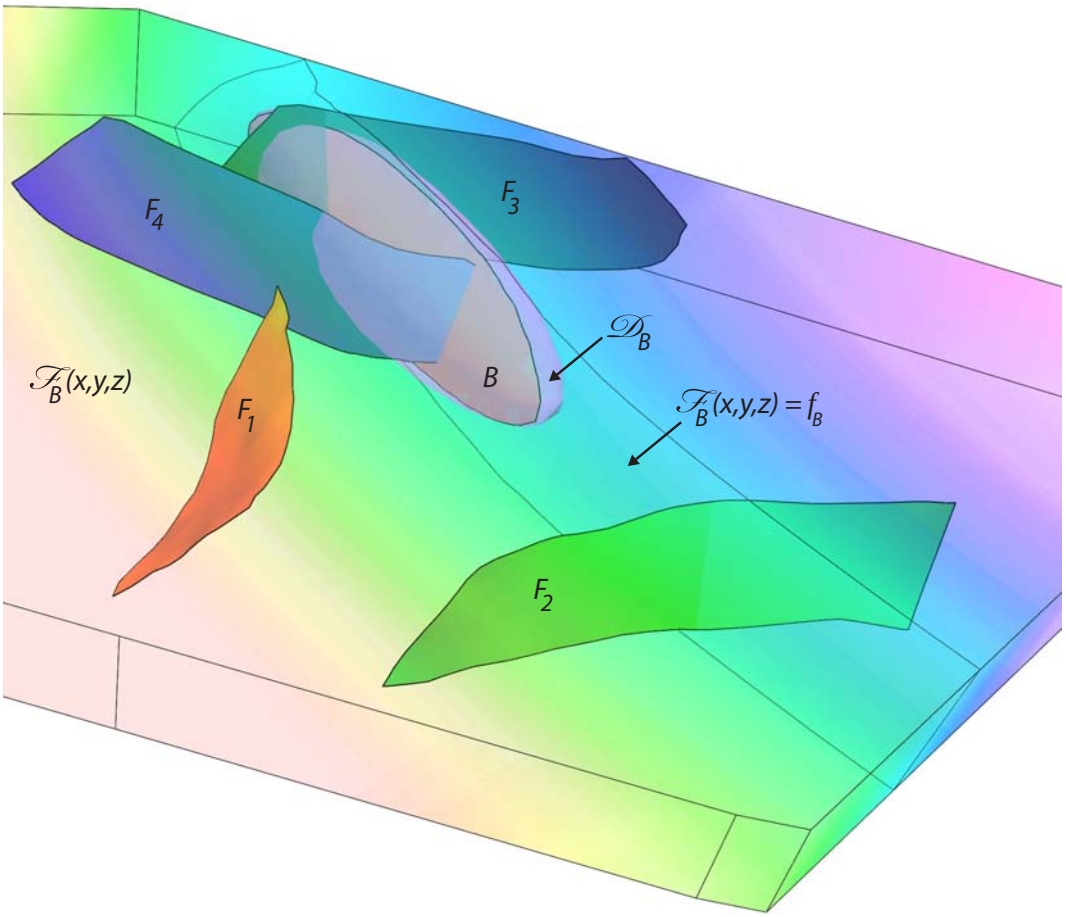

FIGURE 1.24 – Cas possibles pour l'intersection entre deux failles. Les failles F_1, F_2, F_3, F_4 correspondent respectivement aux cas d'intersection 1, 2, 3 et 4 par rapport à la faille antérieure B.

1.21*a.*, supposons que $R = B_3$. L'algorithme teste alors les intersections I_{5-3} et I_{5-4}. Ces deux intersections correspondent au cas 2 (aucune intersection), l'algorithme teste donc les régions B_4 et B_5. Il n'y a pas d'intersection dans la région B_4 et I_{5-1} correspond au cas 1. Par conséquent, l'algorithme s'arrête et les blocs B_3, B_4 et B_5 sont remplacés dans l'arbre par des nœuds correspondant à F_5 (figure 1.21*b.*).

L'algorithme 4 a l'avantage de ne tester dans un premier temps que les intersections avec les failles définissant la région contenant le centre de la faille concernée F. Il se propage seulement aux régions voisines si nécessaire, ce qui limite le nombre de tests. Une structure de donnée de type octree (subdivision récursive de l'espace en huit sous-espaces permettant d'accélérer l'évaluation de la position des objets) pourrait également être utilisée pour trouver rapidement les failles susceptibles d'être intersectées. Cependant, l'algorithme d'intersection, en partant au plus près du centre F permet de tester les failles proches et ne se propage pas en cas d'intersection. En utilisant un octree, plusieurs failles pourraient se trouver dans la même région et il faudrait alors vérifier en cas d'intersections multiples quelle intersection prédomine. En effet, une intersection limite l'extension de F, il se pourrait donc que les autres intersections n'existent en fait pas, étant du côté tronqué de F.

Entrées :
F= faille
R_0 = région initiale contenant le centre de F

Définitions :
S = liste de régions
S' = liste de failles
S'' = liste de régions où F existe

$S \leftarrow R_0$
tant que $S \neq \emptyset$ **faire**
  $R \leftarrow$ dernière région de S
  $S'' \leftarrow R$
  pour chaque *faille* F_i *définissant* R **faire**
    calculer l'intersection I_{F-F_i}
    si $I_{F-F_i} = \emptyset$ **alors**
      ne rien faire `(cas 1)`
    sinon si $I_{F-F_i} \neq \emptyset$ *et* $I_{F-F_i} \cap \mathcal{D}_{F_i} = \emptyset$ **alors**
      $S' \leftarrow F_i$ `(cas 2)`
    sinon si $I_{F-F_i} \cap \mathcal{D}_{F_i} \neq \emptyset$ *et* $I_{F-F_i} \notin \mathcal{D}_{F_i}$ **alors**
      choisir entre le cas 2 et 4 `(cas 3)`
    sinon
      vérifier que $I_{F-F_i} \subset \mathcal{D}_{I_{F-F_i}}$ `(cas 4)`
      ne rien faire
    finsi
  fin
  enlever R de S
  pour chaque *faille* F_i *dans* S' **faire**
    pour chaque *région* R *ayant* F_i *comme frontière* **faire**
      si $I_{F-F_i} \cap R \neq \emptyset$ **alors**
        $S \leftarrow R$
      finsi
    fin
  fin
fin
Sorties : S''

Algorithme 4: Détermination de l'intersection d'une faille.

1.7 Perspectives

Le travail effectué a permis d'élaborer une méthode de simulation de réseaux de failles qui combine l'interprétation de données avec une connaissance *a priori* des failles. La partie suivante détaille des pistes possibles pour l'intégration de données et de concepts géologiques supplémentaires afin de mieux contraindre le processus de simulation. Ces apports permettraient d'améliorer la cohérence géologique des réalisations en fonction du contexte tectonique et ainsi l'échantillonnage de l'espace d'incertitudes.

1.7.1 Concepts géologiques supplémentaires

1.7.1.1 Modélisation de zones de faille en relai

Dans certains contextes tectoniques, notamment de rifting, des zones de faille en relai se mettent en place [Childs et al., 2009, Kim et al., 2004, Peacock and Sanderson, 1995](figure 1.11 page 25). La disposition des failles principales peut être décrite par (figure 1.25*a*.) [Huggins et al., 1995, Imber et al., 2004] : (1) un espacement correspondant à la distance entre deux failles dans la direction perpendiculaire à l'azimuth moyen ; (2) un recouvrement plus ou moins important dans la direction d'azimuth. Ces deux caractéristiques pourraient servir lors de la nucléation d'une faille, afin de définir des zones de nucléation à forte probabilité (figure 1.25).

Dans une zone en relai, les failles secondaires se développent principalement dans la zone de recouvrement des failles principales afin d'accommoder une partie du déplacement et de relaxer les contraintes. La simulation d'une telle disposition pourrait être atteinte en favorisant la nucléation des failles secondaires dans les zones d'intersection des domaines d'extension des failles principales (figure 1.26). Il est cependant important de disposer des informations sur la géométrie de ces structures en profondeur pour mieux contraindre la géométrie des horizons.

1.7.1.2 Modélisation de failles listriques

Dans le cas où aucune information n'est disponible, la construction d'une faille suppose une faille plane, en calculant la distance à un plan calculé à partir de l'orientation. Cependant, la modélisation d'une faille à géométrie particulière n'est possible que si des données interprétées, en profondeur en particulier, contraignent suffisamment l'interpolation du champ scalaire de la faille. Il pourrait cependant être intéressant d'obtenir de telles géométries sans ou avec peu de données (par exemple avec des données de surface ou des plans nodaux issus de l'analyse et du traitement sismologiques).

Plusieurs options peuvent être envisagées :

- l'obtention d'une géométrie listrique peut être obtenue grâce au champ de perturbation $\mathcal{R}(x, y, z)$ (ajouté au champ $\mathcal{F}(x, y, z)$ de la faille initiale plane). En effet, l'utilisation d'une tendance croissante avec la profondeur permettrait d'obtenir une géométrie listrique. Si la dérive est linéaire, seul le pendage de la faille est modifié. L'obtention d'une courbure plus ou moins prononcée peut être obtenue avec une fonction de type z^a avec a négatif (axe z vers le haut) ;
- une fonction de type z^a pourrait être ajoutée directement au champ de distances avant d'ajouter un champ de perturbations (figure 1.27).

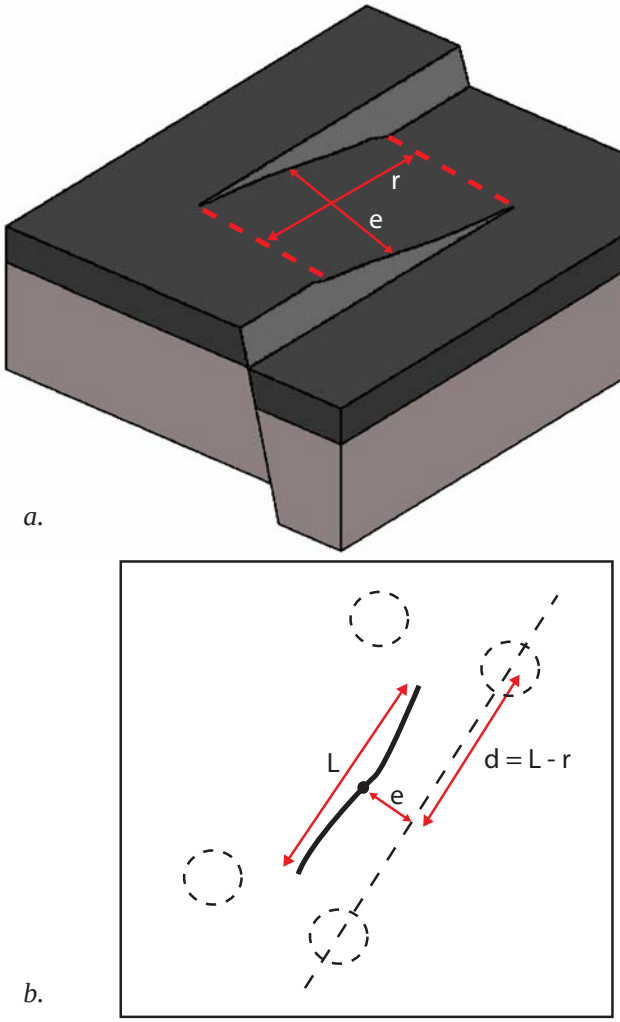

FIGURE 1.25 – Paramètres et probabilité de nucléation de failles en relai. *a.* Deux failles en relai peuvent être décrites par rapport à leur éloignement et leur recouvrement (schéma modifié d'après Peacock and Parfitt [2002]). *b.* Ces deux paramètres pourraient servir à définir des zones à forte probabilité lors de la nucléation d'une nouvelle faille (vue de dessus).

1.7.1.3 Modélisation de structures en fleur

Des structures en fleur peuvent se mettent en place en contexte principalement décrochant. Dans le cas de transpression, la structure est dite positive car formée de failles inverses qui tendent à surélever les roches (figure 1.28*a.*). En contexte de transtension, les failles sont normales et la structure est qualifiée de négative du fait de l'abaissement des roches (figure 1.28*b.*). Le terme structure "en fleur" vient de sa ressemblance en coupe avec les pétales d'une fleur [Fossen, 2010].

Ces structures formées par deux (ou plusieurs) failles sont difficiles à interpréter et à modéliser du fait de la complexité des troncations (figures 1.29*a.*, 1.30). En effet, les deux failles s'arrêtent mutuellement l'une sur l'autre, si bien que la représentation de filiation dans un arbre topologique n'est pas possible, les deux failles étant à la fois père et fils dans

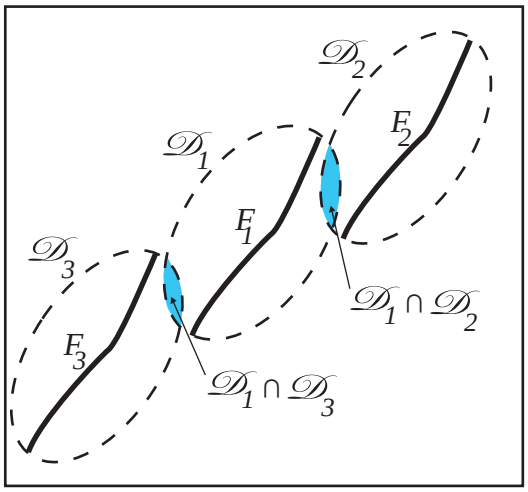

FIGURE 1.26 – Nucléation de failles secondaires dans une zone en relai. Vue de dessus de trois failles principales. L'intersection des domaines d'extension de deux failles représente une zone à forte probabilité de failles secondaires.

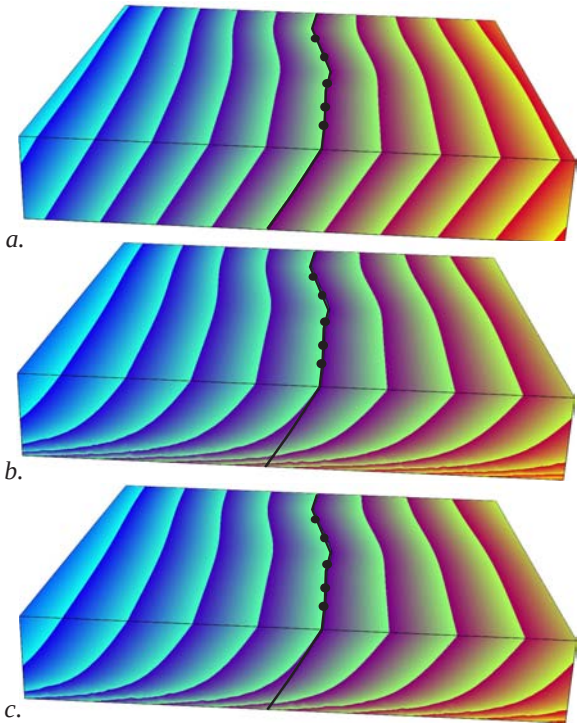

FIGURE 1.27 – Exemple de modélisation d'une faille listrique. *a.* Champ scalaire initial calculé à partir de l'orientation de la faille et des données de surface (points noirs). *b, c.* Deux exemples obtenu après ajout d'une fonction $z^{-2.1}$ (*b.*) et $z^{-2.2}$ (*c.*) au champ scalaire initial. La trace de la faille initiale est visible en noir.

l'arbre. De plus, quelle faille modéliser en premier ? Étant donné la spécificité d'une telle structure, il semble préférable de modéliser directement les deux champs scalaires *a priori*

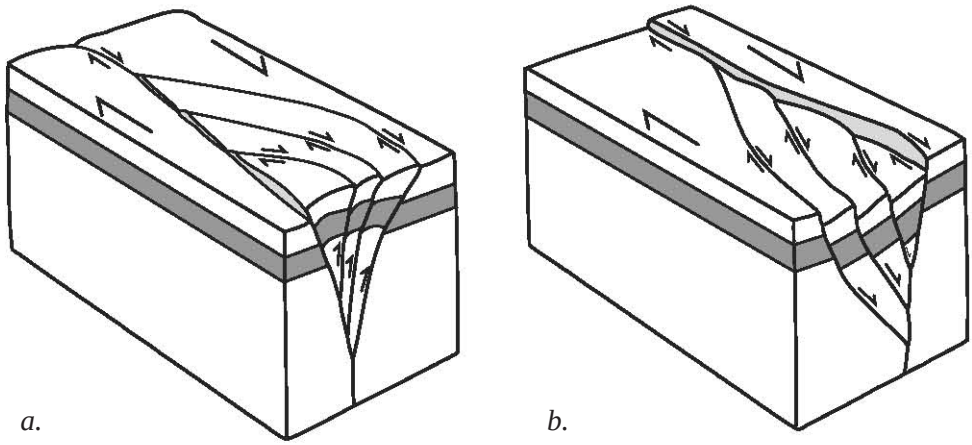

FIGURE 1.28 – Structures en fleur. *a.* Structure en fleur positive en régime transpressif. *b.* Structure en fleur négative en régime transtensif. Modifié d'après Drew [2003].

plutôt que de les modéliser de manière indépendante. Une fois les deux champs obtenus (figure 1.29*b.*), la discrétisation des deux failles par propagation à partir de leur centre respectif permettrait d'obtenir la géométrie souhaitée (figure 1.29*c, d.*).

Cette géométrie particulière s'observe également pour des fractures déviées lors de leur propagation par le champ de contraintes d'une fracture voisine. Dans ce cadre, Henrion [2011] a développé une approche pseudo-génétique pour la simulation stochastique de milieux fracturés qui simule un tel comportement.

D'après Walsh et al. [1999], une structure en fleur peut dégénérer en une seule surface de faille principale qui accommode les contraintes et le déplacement. La deuxième faille forme alors une lentille inactive (figure 1.30). Dans ce cas, la représentation de l'intersection est simplifiée puisqu'il n'y a alors qu'une seule faille principale, et la méthode actuelle convient pour ce type de structure. En revanche, le problème de génération des deux champs scalaires correspondant à cette géométrie, reste ouvert.

1.7.2 Intégration de données supplémentaires

1.7.2.1 Prise en compte du déplacement lors de la nucléation

Lors de la modélisation d'une zone de failles, des données de déplacement peuvent être disponibles, par exemple à partir de l'interprétation des horizons stratigraphiques sur une coupe sismique, de part et d'autre de la zone faillée. Dans ce cas, seule la composante verticale du déplacement est accessible. Ces données fournissent néanmoins une information sur la présence de failles :

- une zone à rejet faible ne contient pas de failles majeures, du moins probablement pas le centre d'une faille majeure où le déplacement est souvent submaximal à maximal ;
- une zone à rejet élevé contient probablement une ou plusieurs failles majeures ;

Cette information pourrait être utilisée lors de la nucléation d'une faille, en privilégiant les zones à déplacement élevé. La simulation du rejet sur la faille courante permettrait de mettre à jour le rejet observé. Le rejet pourrait alors constituer un critère d'arrêt supplémentaire, la méthode ajoutant des failles tant que le rejet résiduel est supérieur à un seuil

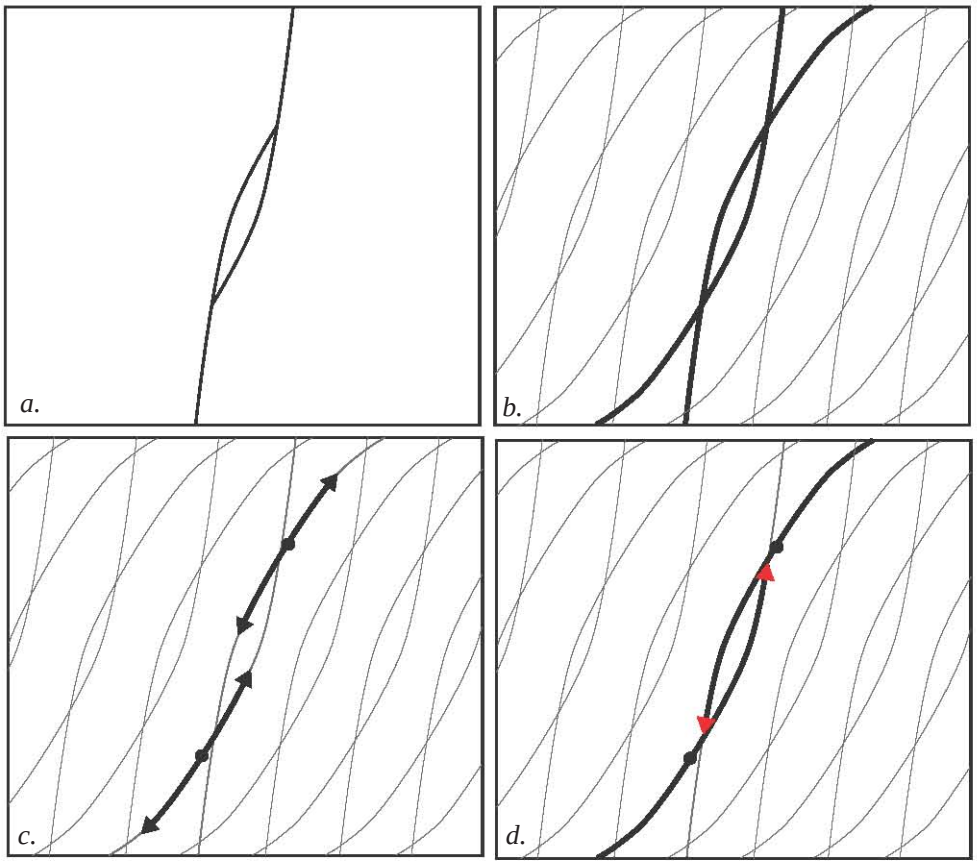

FIGURE 1.29 – Modélisation d'une structure en fleur. *a.* Exemple de structure en fleur vue en carte. *b.* Champs scalaires des deux failles. *c.* Discrétisation par propagation à partir des centres de failles. *d.* La propagation s'arrête lorsque la faille voisine est déjà discrétisée.

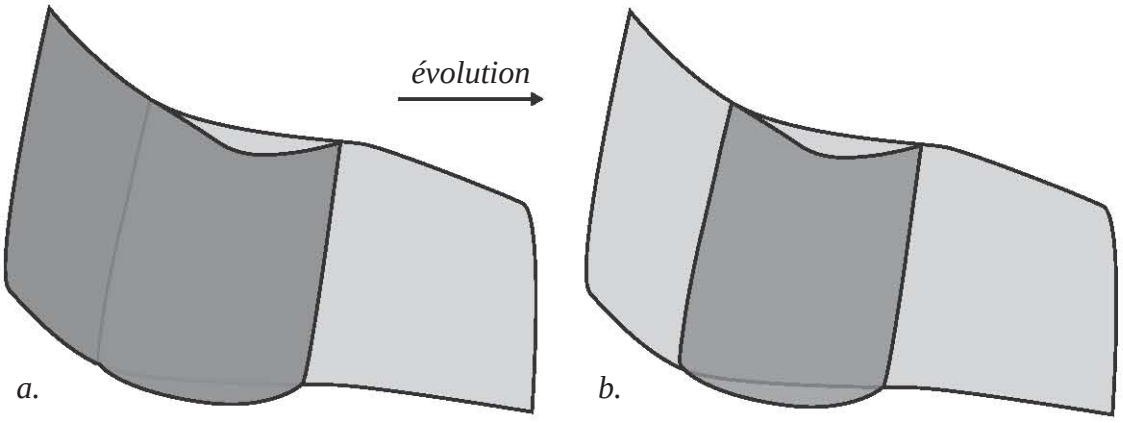

FIGURE 1.30 – Évolution d'une structure en fleur. Une structure en fleur (*a.*) peut dégénérer en une seule faille active et une seconde faille inactive en forme de lentille (*b.*).

donné. La simulation du rejet pourrait donc constituer une étape supplémentaire dans la simulation d'un objet faille. Pour cela, Bouziat et al. [2011] proposent de déformer les ho-

rizons stratigraphiques dans un domaine 3D autour d'une faille afin de modéliser son rejet. Un champ de déplacement vectoriel est calculé, en s'appuyant sur les travaux de von Funck et al. [2006] dans la communauté de l'infographie. Røe et al. [2010] déforment également des horizons grâce à un champ vectoriel 3D calculé de manière analytique. Le domaine 3D limitant l'extension d'une faille pourrait définir la zone de rejet non nul.

Le rejet d'une faille peut être interpolé à partir du déplacement attaché aux points de données associés à la faille. Si l'information n'est pas disponible, ou si aucun point de donnée n'est associé à la faille, un rejet théorique peut être calculé à partir d'une loi de type $d = cL^n$ avec d le déplacement maximal au centre, c une constante et L la longueur de la faille (une estimation de c à partir d'analogues peut être faite à partir des travaux de Kim and Sanderson [2005]).

L'information du rejet pourrait être stockée sur l'axe médian MA de la zone de failles à modéliser. Chaque cellule de la grille servant à définir la probabilité de nucléation p serait ensuite projetée sur MA. La probabilité p pourrait alors être calculée à partir de la valeur du rejet et de son orientation par rapport au type de la faille courante. Une fois la faille entièrement simulée (rejet y compris), les valeurs du rejet seraient projetées sur MA et soustraites aux valeurs stockées sur MA.

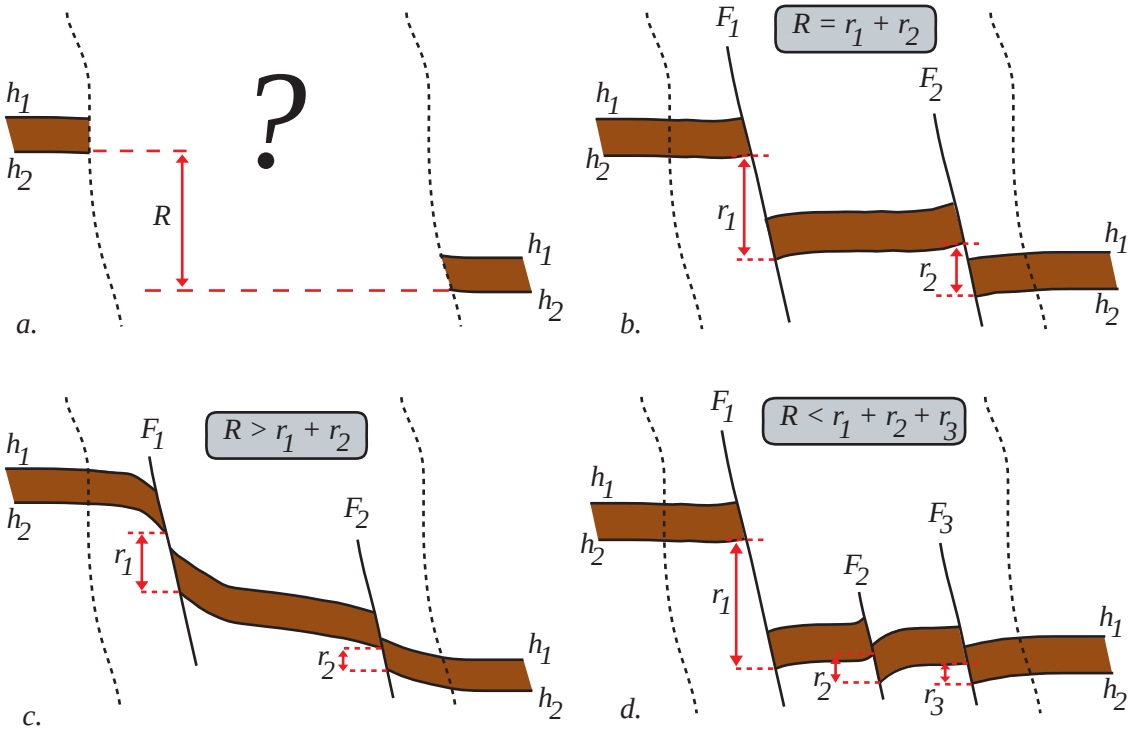

FIGURE 1.31 – Rejet observé dans une zone de failles. *a.* Un rejet vertical R est observé à travers une zone faillée. *b.* Les rejets r_1 et r_2 sont égaux à R. *c.* Des blocs basculés accommodent une partie du rejet, la somme des rejets de F_1 et F_2 est donc inférieure au rejet R. *d.* Le rejet R est inférieur au rejet réel des trois failles, car la participation de F_2 au rejet total est nulle.

Le rejet observé de part et d'autre d'une zone de failles peut ne pas correspondre à la somme des rejets des failles présentes sur une même ligne transversale à la zone :

- des blocs basculés peuvent accommoder une partie du rejet observé, ce qui tend à surestimer le rejet des failles (figure 1.31*c.*) ;
- la participation d'une faille au rejet observé peut être nulle (faille F_2 de la figure 1.31*d.*).

La méthode de calcul de l'axe médian proposée par Van Dortmont et al. [2006] a été implémentée en début de thèse mais n'a pas été utilisée par la suite dans ce cadre spécifique. En revanche, l'axe médian de milieux complexes (*e.g.* milieu poreux) peut servir de base pour leur modélisation stochastique à partir de son champ de distances [Henrion et al., 2010]. L'axe médian représente néanmoins un support en deux dimensions permettant de stocker de l'information relative à une zone en trois dimensions. Cette information (*e.g.* rejet, orientation locale) pourrait être utilisée lors de la simulation d'un réseau de failles afin de contraindre leur géométrie. Dans le cas de données sismiques, il reste cependant à définir une méthode efficace pour extraire une zone à faible cohérence représentant le volume faillé. Dans une approche indépendante, Kadlec et al. [2008] proposent de calculer la transformée d'axe médian directement dans un volume sismique afin d'identifier les failles.

1.7.2.2 Honorer des zones non faillées

La méthode utilise en entrée des données interprétées afin de fournir des informations relatives aux failles, notamment des points de données à honorer. Cependant, elle n'exploite pas l'information contraire "d'absence" de failles. L'interprétation d'horizons stratigraphiques continus sur une ligne sismique suggère par exemple l'absence de failles sismiques. Cette information pourrait être utilisée à des étapes différentes de la simulation d'une faille :

- l'algorithme pourrait empêcher l'association d'ensembles de points qui sont de part et d'autre d'une zone non faillée ;
- après la nucléation d'une faille en dehors des zones non faillées, l'algorithme pourrait déterminer l'intervalle des orientations possibles pour une faille de taille moyenne. Dans ce cas, l'orientation de la faille interviendrait après l'étape de nucléation (figure 1.32) ;
- après la simulation d'une faille F, si l'intersection des volumes non faillés V_{nf} avec F n'est pas nulle, deux options sont possibles : (1) supprimer la faille et en simuler une nouvelle ; (2) modifier la géométrie de F pour que $V_{nf} \cap F = \emptyset$. Le calcul des distances de Hausdorff relatives entre F et V_{nf} pourrait servir de critère entre les deux options, la deuxième option étant privilégiée si la distance d est inférieure à la sinuosité de F par exemple. Dans ce cas, les distances doivent être envisagées séparément des deux côtés de F afin de déterminer la modification de moindre coût (figure 1.33). D'autres tests sont également nécessaires pour évaluer la possibilité d'une modification du champ d'extension $\mathcal{D}_F$ de F de sorte que $\mathcal{D}_F \cap V_{nf} = \emptyset$. Un angle de déviation pourrait également être utilisé, de la même manière que lorsque une faille recoupe des ensembles de points (section 1.6.11.1, figure 1.22 page 46).

1.7.3 Simulation d'un objet faille modulable

La méthode de simulation d'un objet faille présentée en section 1.6 détaille la succession des étapes aboutissant à un nouvel objet faille dans le modèle initial. Le choix est fait de favoriser les points de données en plaçant les étapes d'attribution d'un ensemble de

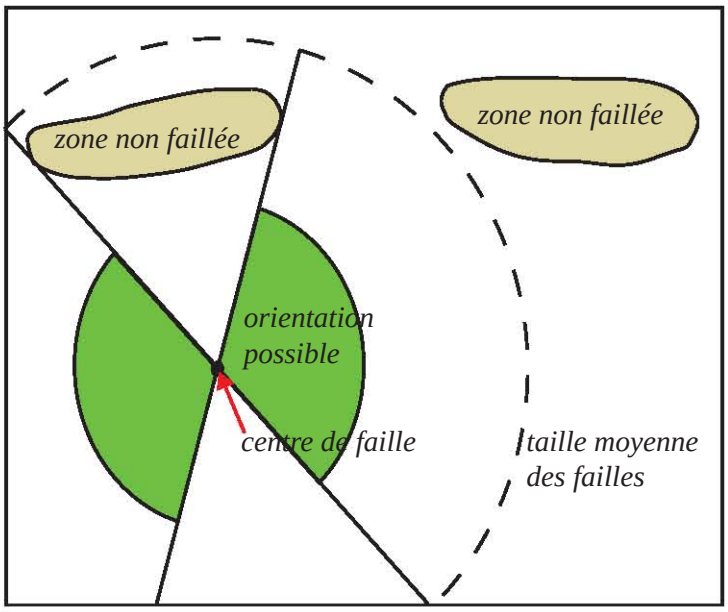

FIGURE 1.32 – Zones non faillées et orientation d'une faille. Après la nucléation d'une faille, l'algorithme pourrait déterminer quelles sont les orientations possibles afin d'honorer les zones non faillées.

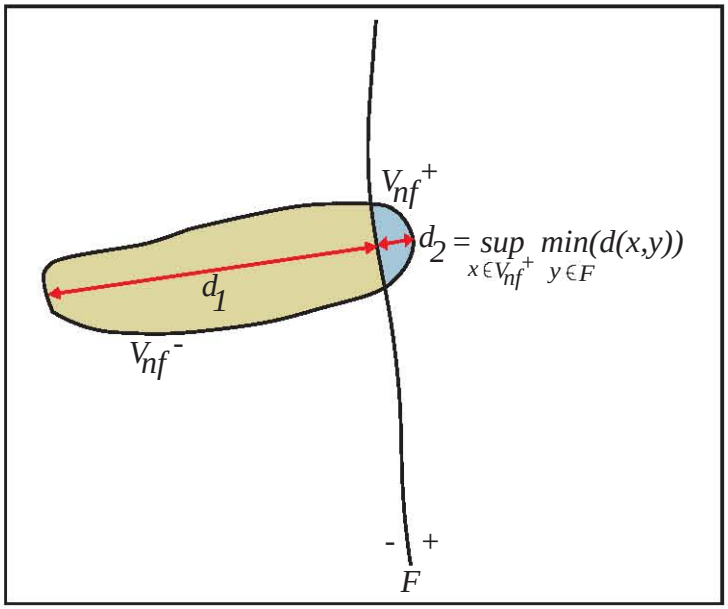

FIGURE 1.33 – Zones non faillées et géométrie d'une faille. Le calcul des distances de Hausdorff relatives, dans les blocs V_{nf}^+ et V_{nf}^- séparément permettrait de déterminer l'amplitude de la déformation nécessaire pour honorer une zone non faillée. L'algorithme pourrait alors choisir entre perturber la géométrie de la faille ou la supprimer et en simuler une nouvelle.

points et de regroupement en début de simulation d'un objet faille. L'étape de simulation de l'extension d'une faille est en fin de simulation et donc est contrainte par les étapes

précédentes. Cet ordre convient si peu d'information concernant la taille des failles est disponible. En revanche, dans le cas où la taille des failles serait bien caractérisée, il pourrait être intéressant de simuler celle-ci dans les premières étapes afin qu'elle contraigne le reste des étapes moins renseignées. La simulation pourrait donc être modulable, avec un ordre des étapes de simulation variant afin d'exploiter au maximum l'information disponible. Cette approche nécessite une implémentation souple de l'algorithme, chaque étape de simulation s'adaptant selon les informations qui lui sont fournies en entrée.

1.8 Conclusions

Le modèle stochastique de failles proposé permet de générer des réseaux de failles à partir de données interprétées et d'une connaissance *a priori* des failles. Cette connaissance statistique des failles peut provenir de l'analyse des données, du contexte tectonique régional, d'analogues ou encore de concepts géologiques.

Les méthodes existantes pour prendre en compte les incertitudes structurales [Charles et al., 2001, Caumon et al., 2007, Lecour et al., 2001, Thore et al., 2002, Wellmann et al., 2010] perturbent un modèle de référence. Elles permettent donc l'échantillonnage d'incertitudes faibles grâce à des perturbations géométriques. La méthode proposée, tout comme la méthode HAVANA [Holden et al., 2003], ne s'appuie sur aucun modèle préexistant, ce qui permet de prendre en compte des incertitudes plus grandes en échantillonnant des topologies différentes. Elle diffère de la méthode HAVANA par le processus et les règles de simulation ainsi que par la représentation d'un objet faille, qui ne dépend pas dans notre cas du support de modélisation. En effet, l'approche implicite offre un cadre de modélisation plus souple qui permet de s'affranchir des contraintes de maillage. La contrepartie est une demande en capacité informatique plus importante puisqu'il est nécessaire de stocker en tout point du domaine la valeur du champ scalaire de chaque faille.

Ce chapitre a permis d'aborder l'échantillonnage des incertitudes relatives aux failles. La génération stochastique de modèles structuraux, incluant des horizons stratigraphiques, est abordée au chapitre suivant.

Chapitre 2

Modélisation structurale stochastique

Sommaire

2.1 Introduction 61
2.2 Stochastic structural modelling method 62
2.2.1 Fault characterization from interpreted data 62
2.2.2 Simulation overview 64
2.2.3 Fault simulation 65
2.2.4 Computing the stratigraphy 72
2.2.5 Simulation choices 75
2.3 Application to a Middle East field case 75
2.3.1 Input data 75
2.3.2 Fault network characterization 75
2.3.3 Stratigraphic uncertainty 77
2.3.4 Results 77
2.4 Conclusions 78

La modélisation des failles constitue la première étape du processus de géomodélisation et les couches géologiques, déplacées par les failles, sont ensuite modélisées afin d'obtenir une représentation volumique du sous-sol. Notre méthode suit également ce raisonnement et une méthode de simulation stochastique de failles, développée au cours de la thèse, a été décrite au chapitre 1. Elle permet, à partir de données interprétées et d'une connaissance *a priori*, de générer des réseaux de failles par la simulation successive de familles de failles. Ce chapitre aborde la modélisation stochastique des horizons stratigraphiques nécessaires pour délimiter les différentes unités géologiques. En effet, la géométrie des horizons stratigraphiques est, de la même manière que la géométrie des failles, incertaine en raison de leur variabilité naturelle et de l'incertitude des données interprétées. La connaissance limitée du rejet des failles constitue un facteur d'incertitude supplémentaire.

La première partie de ce chapitre détaille la méthode de simulation stochastique de modèles structuraux en intégrant les éléments présentés au chapitre précédent afin de prendre en compte les incertitudes liées aux failles. Les incertitudes structurales sont d'autant plus importantes que la quantité de données est faible. La deuxième partie de chapitre présente une application de la méthode dans un contexte incertain avec peu de données, afin d'évaluer les incertitudes sur les volumes de roches d'un réservoir d'hydrocarbures. Ce travail est en cours de préparation pour une publication dans *Petroleum Geoscience*.

Stochastic structural modelling in sparse data situations

Nicolas Cherpeau and Guillaume Caumon

Article en préparation pour *Petroleum Geoscience*

Resource estimation is a key step for decision-making in natural resource management because it determines the economic viability of a deposit with regard to exploration and production costs. Uncertainties on fault network layout may lead to risky reserves estimation in compartmentalized reservoirs, especially when few data is available. In such context, this paper suggests to generate stochastic structural models using prior information such as fault orientation, size-displacement relationships and interpretations of both faults and stratigraphic horizons on particular cross-sections. Interpreted fault data on 2D seismic lines provide significant information to constrain fault simulation and reduce spatial uncertainty. Indeed, the fault slip is generally assumed to be related to the fault length. Consequently a global fault size range can be inferred from observed displacements and is an input parameter for the fault simulation algorithm. For each simulated fault network, stratigraphic modelling is performed to honour interpreted horizons using an implicit approach. Geometrical uncertainty on stratigraphic horizons can then be simulated by adding a correlated random noise to the stratigraphic scalar field. Implicit modelling provides a convenient framework for such perturbations as horizon-to-fault contacts are not explicitly represented, hence are consistently maintained. The method is applied to a Middle East field where structural models are generated from interpreted 2D seismic lines. Generated models show the impact of the fault network and the stratigraphy on the gross rock volumes estimation uncertainties.

KEYWORDS : structural modeling ; uncertainty ; stochastic method

2.1 Introduction

Structural uncertainty is often a major but neglected factor in E&P studies. Structural uncertainty typically stems from the lack of observations or ambiguities of measurement and associated processing (*e.g.* limited resolution, lack of impedance contrasts, velocity uncertainty in seismic reflection data). In such context, geological interpretation is very important and relies on regional knowledge about the tectonic and sedimentary evolution of the area through time, in addition to physical laws describing rock behaviour. However, the interpretation of seismic data may lead to various interpretations by experts [Bond et al., 2007]. Consequently, a given subsurface model should not be considered as the truth but as a possible representation of the subsurface at a given stage in the modelling chain from available observations. Additional data may then be used to update or even to falsify the model in case of inconsistency with new data [Tarantola, 2006].

One solution to evaluate risks consists in using a set of possible models consistent with geological concepts and observations to make predictions [Scheidt and Caers, 2009, Seiler et al., 2010, Suzuki et al., 2008]. In practice, geostatistical realizations are commonly used to sample petrophysical uncertainties. Some authors [Nivlet et al., 2002] suggested rather using arithmetic interval techniques. However, structural uncertainties impact spatial layout of reservoir layers, hence resource estimation. Most techniques [Caumon et al., 2007, Lecour et al., 2001, Mallet and Tertois, 2010, Thore et al., 2002] start from an initial interpretation and perturb the geometry of the structures, keeping the topology frozen. Our method does not rely on an initial interpretation, which avoids introducing an interpretation bias when building the initial model. In addition, it allows, in principle, to observe a larger variability

in the simulated models because the input data is only the raw interpretation and simulated models are not constrained by any reference model.

The Havana modelling tool proposed by [Holden et al., 2003] enables to generate fault networks with different topologies on pillar-based reservoir grids, mainly to evaluate the impact of faults on fluid flow. In this paper, we focus on fault-related and stratigraphic uncertainties at the structural modeling step and use the technique proposed by Cherpeau et al. [2010b] that generates stochastic fault networks with various fault connections. Each fault is modelled by (1) geometric parameters such as a fault centre, an orientation, a size, a sinuosity and (2) topological parameters which determine the fault truncations. The proposed method uses a general fault representation : each fault is represented as a set of polygons computed from the equipotential of a volumetric scalar field. The method is extended to model also the stratigraphy. The workflow in composed of two main steps :

1. The stochastic fault modelling tool [Cherpeau et al., 2010b,a] generates realistic fault networks ;
2. One or several structural models are generated from a fault array and horizon data by computing the stratigraphic field(s).

This new methodology is applied on a Middle East field in order to illustrate its potential on an exploration play. Stochastic structural models are generated from interpreted seismic lines and are used to estimate gross rock volume uncertainty. The results are compared to those obtained by only perturbing the stratigraphy of the reference model built from 3D seismic data in order to highlight the influence of fault-related uncertainties.

2.2 Stochastic structural modelling method

2.2.1 Fault characterization from interpreted data

The first step in the stochastic modelling workflow consists in characterizing faults from the various input data (*e.g.* seismic profiles, cross-sections, wells). Similar faults are grouped into fault families, each one representing a given faulting event, *e.g.* a NW-SE extension or a W-E transtension. Fault families are ranked chronologically to represent the time sequence of the fault events and to model observed fault truncations : a fault terminates on or crosses existing faults.

2.2.1.1 Location information

The interpretation of faults on input data gives clues about fault location : at least one fault should pass through an interpreted fault trace (a set of interpreted fault data points). However, due to limited data resolution, georeferencing errors or picking approximations, the exact fault location may be uncertain. Therefore, an uncertainty is attached to fault traces to account for such approximations.

White squares in figure 2.1 represent a fault trace. An uncertainty range is attached to each point, which defines a 3D sphere of possible fault location. A radial probability distribution function is associated to this sphere as in Lecour et al. [2001], Wellmann et al. [2010]. If no information is provided, any fault from any fault family may honor a given fault trace.

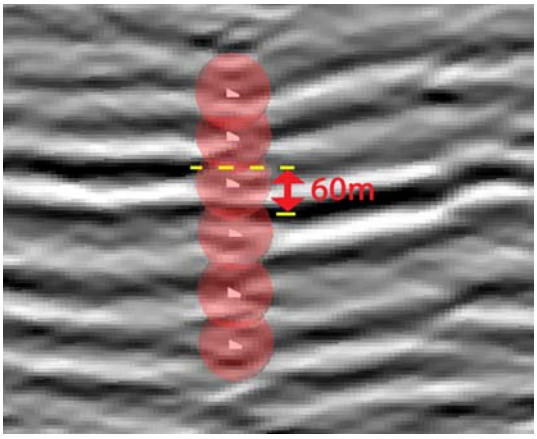

FIGURE 2.1 – Slip information and fault location uncertainties. White points have been picked at possible fault occurrence and a sphere (in red) is defined around each point to define possible fault locations. The fault slip may also be estimated from interpreted stratigraphic horizons. Only the vertical slip component can be estimated.

2.2.1.2 Orientation information

When interpreting fault traces on seismic data, especially on 2D vertical sections, only the apparent orientation is available. Indeed, the apparent dip is reduced as compared to the true fault dip, unless the seismic section is strictly perpendicular to the fault strike (equation 2.1). The apparent strike provides little information about the true fault strike as faults with various strikes may correspond to the apparent strike on a 2D section (figure 2.2).

$$\begin{aligned} &\tan(dip_{apparent}) = \tan(dip_{true}) \times \sin(\alpha) \\ &\text{with } \alpha \text{ the angle between the cross-section and the fault strike} \end{aligned} \tag{2.1}$$

Consequently, interpreted fault traces provide a range of possible fault orientations, hence it may not be possible to assign fault traces to a given fault or fault family (figure 2.2). Due to the no linearity of equation 2.1, the smaller α, the greater the uncertainty about the true dip. As angle α is unknown, little information can therefore be deduced from apparent orientation. However, in case the global fault orientation is believed to be almost perpendicular to seismic sections (*e.g.* from regional context), fault traces can be restricted to some fault families based on apparent orientation, *e.g.* to distinguish opposed-dipping faults. Consequently, a list of fault families may be assigned to the fault traces, meaning that only faults of these families can honour these fault traces.

2.2.1.3 Slip information

Interpreted fault traces do not only provide location and orientation information but also vertical slip information when stratigraphic horizons have been interpreted (figure 2.1). Semi-automatic methods may also provide such information. For instance, Liang et al. [2010] propose to compute slip vectors based on cross-correlation of adjacent seismic traces.

In the current implementation presented in this work, the slip information attached

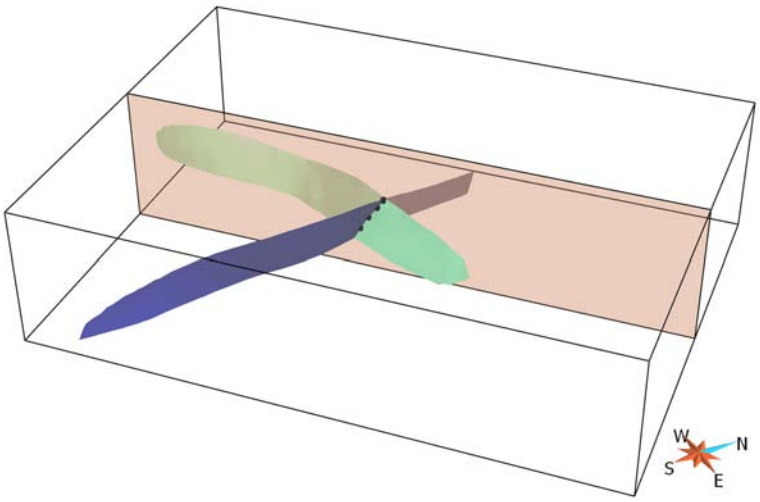

FIGURE 2.2 – Uncertainties about fault trace information. Many faults may honour a fault trace interpreted on a 2D section. Example of two fault surfaces representing two different faulting events honouring the same fault trace.

to a fault trace is a positive value representing the vertical slip component. Indeed, the discrimination between fault types is made when defining fault families. Then, if information is sufficient to discriminate the fault type for a given fault trace, the fault trace can be restricted to belong to specific fault families corresponding to this particular fault type. A signed slip value could be relevant in tectonically complex domains where the slip may have changed due to fault reactivations.

2.2.2 Simulation overview

Cherpeau et al. [2010a] present a theoretical framework for stochastic fault modelling which is based on a realistic 3D fault representation and a binary space partition tree for accounting for the topology and spatial layout of the model, *i.e.* fault connections. Indeed, large fault-related uncertainties may entail changes of fault connections and thus require a flexible fault representation. In [Cherpeau et al., 2010b], this method is extended to the conditional simulation case, *e.g.* to honour fault traces interpreted from seismic data. These methods are briefly reviewed and extended in the following.

Characteristics fault shape parameters can be summarized through statistical distributions of size, orientation and sinuosity. Faults are simulated sequentially in the input chronological order (*i.e.* the oldest fault family is simulated first). In case of full intersection, the younger fault is considered branching on the older ones, to model the growth of the younger fault that stops when reaching another already existing fault. To our knowledge, the partial branching case between faults has received little attention in the literature. In such case, we propose the fault contact to be considered either as crossing or branching (figure 2.3).

From a set of fault traces and fault parameters distributions, each fault F is modelled using the following steps :

1. Draw a fault trace S to honour among the available fault traces (interpreted fault

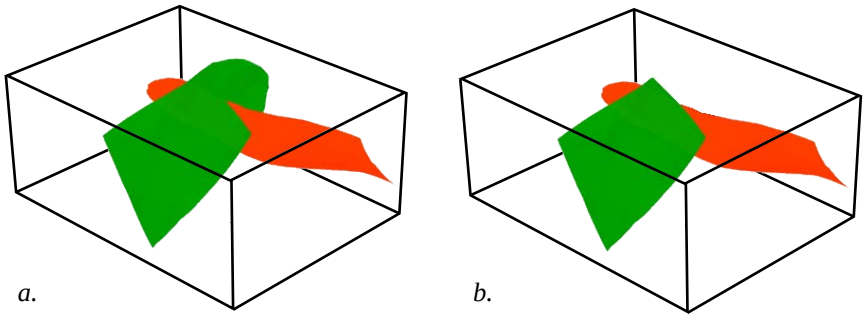

FIGURE 2.3 – Partial branching between two faults. In case of partial branching, the algorithm may consider it as a crossing contact (a) or branching contact (b).

data points) ;

2. Draw orientation parameters if S does not provide enough information ;
3. Draw a fault centre in the neighbourhood of S ;
4. Generate a first geometry using information of previous steps ;
5. Cluster some fault traces with S, so that F passes through several fault traces ;
6. Perturb the initial fault geometry while accounting for data uncertainty ;
7. Limit the existence of F into a 3D closed domain ;
8. Work out intersections with other faults.

Steps 1 to 7 define the geometry of the fault being built, while the last step determines the fault truncations and affects the topology of the whole fault network. Once faults have been simulated, a mesh conformable to faults is built and the stratigraphy is interpolated from available horizon data. Uncertainty about stratigraphy is handled by adding a correlated random perturbation to the stratigraphic field, *i.e.* several stratigraphic fields may be obtained for a given fault network [Caumon, 2010].

2.2.3 Fault simulation

Once fault characterization has been performed and fault families defined, faults are simulated sequentially. This section describes the fault simulation steps and is summarized by the flow chart in figure 2.4.

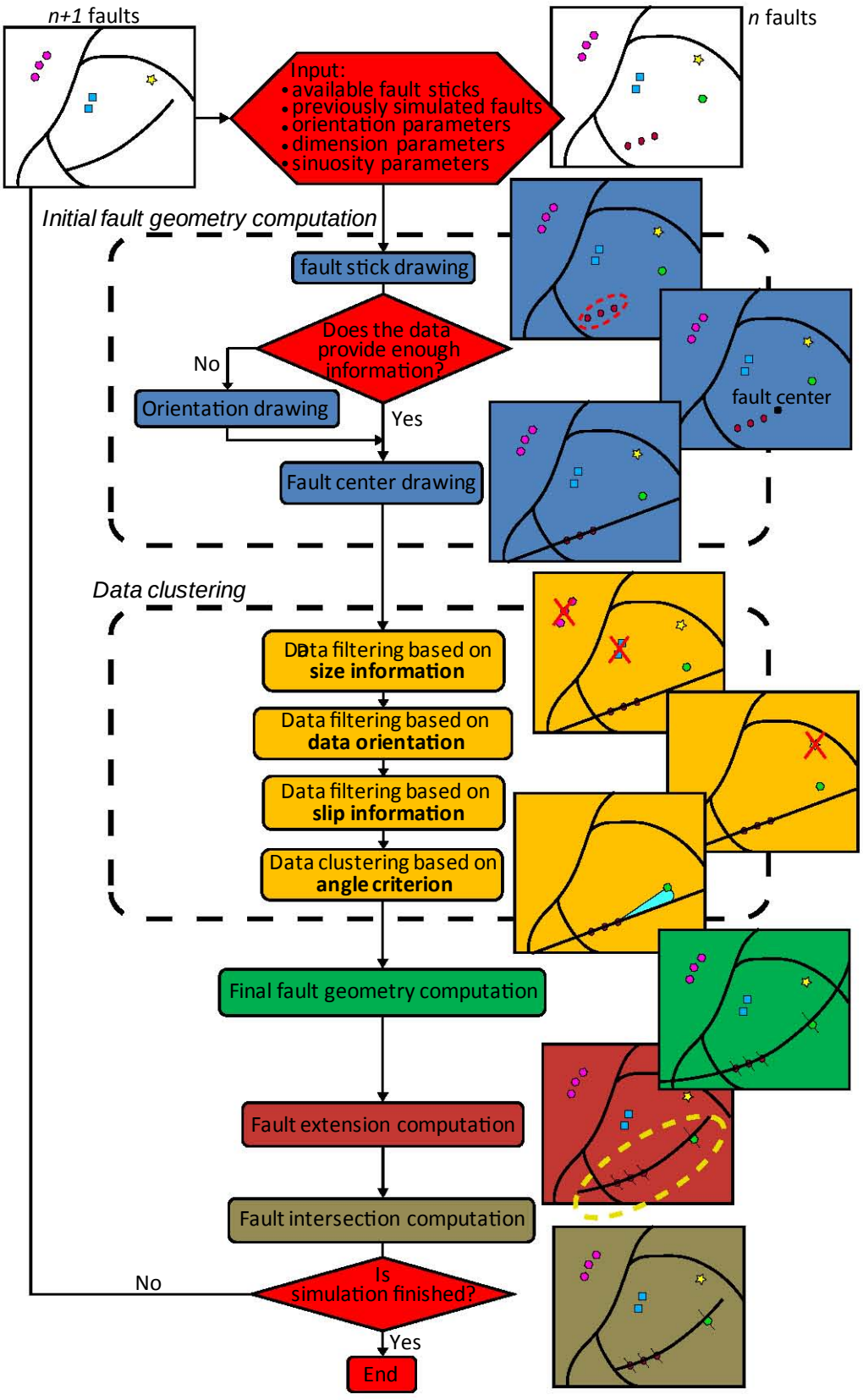

FIGURE 2.4 – Flow chart showing the fault simulation steps.

2.2.3.1 Initial fault geometry computation

A fault trace S is first randomly drawn. If the fault trace does not provide enough information about the fault orientation (as shown in figure 2.2), the dip and the strike are simulated from input statistical distributions. The next step consists in drawing a fault centre in the neighbourhood of S. The fault centre is drawn in a volumetric probability distribution function, which is uniform if no input information is provided. This distribution may also account for the distance to other faults, *e.g.* to simulate secondary faults occurring in the vicinity of existing major faults (conditional distribution law).

At this point, fault parameters (fault trace S, orientation, fault centre) are used to constrain the interpolation of an initial scalar field representing the fault surface [Frank et al., 2007]. Indeed, fault surfaces are modelled as implicit surfaces [Calcagno et al., 2008, Frank et al., 2007, Mallet, 2004]. An implicit surface F is defined by an equipotential f of a volumetric scalar field $D_F(x, y, z)$:

$$D_F(x, y, z) = f \tag{2.2}$$

Conform stratigraphic series (figure 2.10a) are represented by the same volumetric function $D(x, y, z)$ whereas each fault is defined by its own scalar field (figure 2.5). The interpolation [Mallet, 1992] is performed in a tetrahedral mesh and scalar field values are stored on tetrahedra vertices.

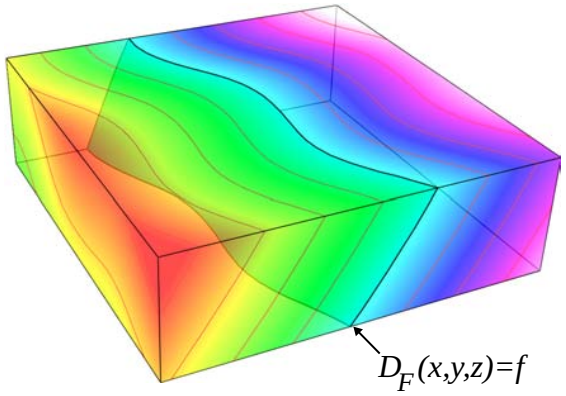

FIGURE 2.5 – Fault represented as an implicit surface. The fault is defined by the isopotential f of a volumetric scalar field $D_F(x, y, z)$.

2.2.3.2 Data clustering strategy

Once an initial fault surface has been computed, available fault traces are considered as candidates for possible inclusion in the fault surface being built. Therefore, the algorithm first performs a three-step filtering in order to ensure plausible fault geometry and slip profile. All available fault traces may be considered, but the algorithm could also filter those that are not in the same fault block than the fault centre to avoid crossing contacts between faults.

Data filtering based on size information

Fault traces that are further away from the current fault centre than the target fault size are

eliminated. Indeed, their inclusion in the fault would lead to a fault too long as compared to the input size distribution. However, a tolerance could be included as the fault centre position may have no direct geological meaning (*e.g.* for faults that grow by segment linkage).

Data filtering based on data orientation

In general fault traces with opposed dip should not be clustered, as it would imply significant changes in fault dip and unrealistic stratigraphy geometry. For a given fault trace S, we propose to compute the distance between the furthest point p_f and the closest point p_c projected onto the fault surface, respectively p_f^{impact} and p_c^{impact}. The angle α' between the vectors $v_S(p_f, p_c)$ and $v_S^{impact}(p_f^{impact}, p_c^{impact})$ is then computed. Clustering is not allowed if α' is deemed too large, which ensures an acceptable dip change in case of clustering (figure 2.6). In case S is sub horizontal, this strategy avoids clustering data points that would entail a too large strike change for the fault being built. In practice, α' is not be chosen greater than 30 ° to ensure realistic orientation changes. However, α' could also depend on fault orientation uncertainty or be larger to allow for listric fault geometries.

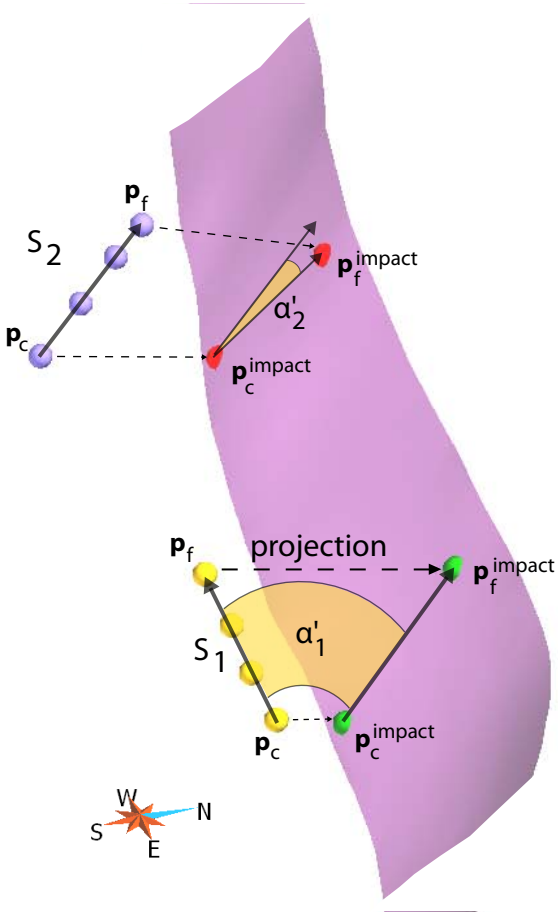

FIGURE 2.6 – Fault trace filtering based on orientation. The closest and furthest points of each trace S_1 and S_2 from fault F are projected onto F. Angle α'_1 is too large so F cannot honor S_1 whereas α'_2 is small which enables F to possibly honor S_2. The filtering ensures reasonable dip or strike changes in case the fault trace is subhorizontal.

Data filtering based on slip information

The slip attached to fault traces provides useful information that can help avoiding geologically inconsistent configurations. Indeed, the slip amplitude is generally assumed to be sub-maximum to maximum at the fault centre and null at fault tip-line [Barnett et al., 1987, Walsh et al., 2003] (figure 2.7).

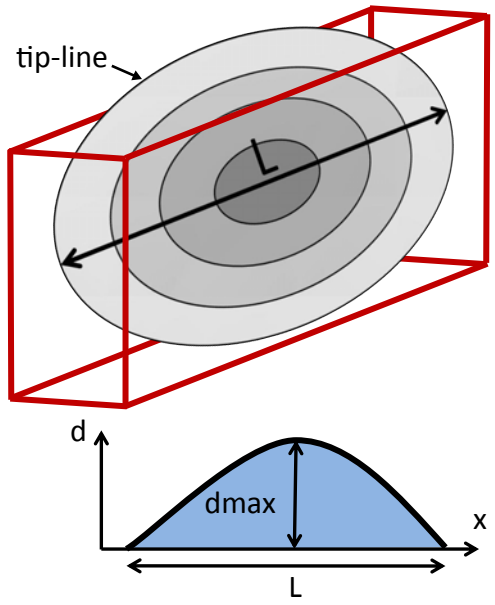

FIGURE 2.7 – Isolated fault displacement distribution. The fault displacement is maximum at the centre of the fault and null at tip-line.

Therefore, a fault honouring a trace S with maximum displacement d_S may or may not be clustered with a trace S' with maximum displacement $d_{S'}$ depending on the following configurations :

1. If S and S' are on both sides of the fault centre, no rule applies, thus clustering is possible (trace S'_0 in figure 2.8) ;
2. If S' is between S and the fault tip-line, $d_{S'}$ should be smaller than d_S (trace S'_2 in figure 2.8) ;
3. If S' is between S and the fault centre, $d_{S'}$ should be larger than d_S (trace S'_1 in figure 2.8).

These rules must be strictly applied in case of isolated faults. However, these rules must not be applied when faults are close and interact together because the displacement profile may be more complex. They do not apply either if faults have grown by segment linkage [Walsh et al., 2003]. Consequently, a tolerance ϵ is introduced to allow for more general fault displacement profiles, which gives the following updated rules in cases 2 and 3 :

- If S' is between S and the fault tip-line, $d_{S'}$ must satisfy :
$$d_{S'} \leq d_S \times (1 + \epsilon) \quad (2)$$
- If S' is between S and the fault centre, $d_{S'}$ must satisfy :
$$d_{S'} \geq d_S \times (1 - \epsilon) \quad (3)$$

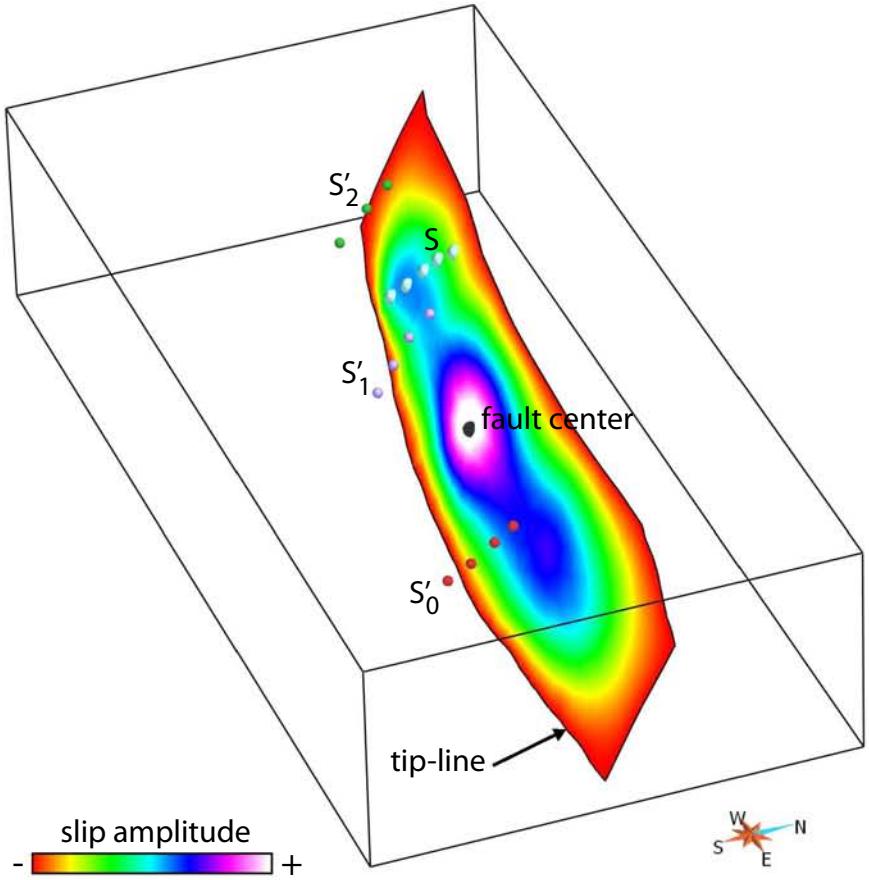

FIGURE 2.8 – Fault trace filtering based on displacement information. The fault honors the fault trace S with attached displacement d_S and three other fault traces S'_0, S'_1, S'_2 (with respectively displacement $d_{S'_1}$, $d_{S'_2}$ and $d_{S'_3}$) are candidates for clustering. Fault trace S'_0 is not on the same side of the fault center than S, therefore S'_0 is not filtered out because no rule applies. Fault trace S'_1 is between S and the fault center, so $d_{S'_1}$ should be larger than d_S (equation 3) for considering S'_1 as candidate for clustering. Fault trace S'_2 is between S and the fault tip-line, so $d_{S'_2}$ should be smaller than d_S (equation 2) to consider S'_2 as candidate for clustering.

In case these relations are not honoured, S' is not considered for possible clustering with S. Other rules could also be considered to constrain data clustering. Indeed, if the difference of displacement observed at two fault traces is large, these two fault traces unlikely belong to the same fault if they are close together because it would imply the stratigraphy to be highly curved along the fault strike.

Data clustering probability

The three previous processing steps aim at filtering geologically inconsistent fault traces. Then, remaining fault traces are considered as candidates for possible clustering with the fault being built. The data clustering probability of each candidate relies on the angle it forms with the initial fault surface (figure 2.9) [Cherpeau et al., 2010b]. Small angles entail small fault surface changes as compared to large angles, hence are considered more likely,

i.e. the smaller the angle, the higher the probability of clustering.

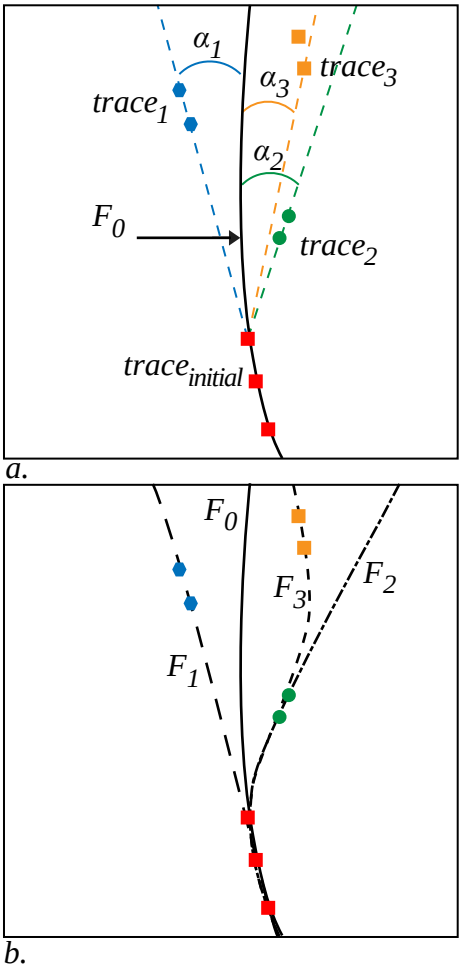

FIGURE 2.9 – Data clustering based on angle criterion (map view). (*a*) The angles α between data points and the initial fault surface are computed (map view). The probability of clustering increases when α decreases. (*b*) Different possible data clusterings corresponding to different fault surfaces may be generated. From Cherpeau et al. [2010b].

2.2.3.3 Estimating the final fault geometry

Once possible data clustering has been performed, the fault scalar field is updated so that the fault equipotential passes through all selected fault traces. Then, the fault scalar field is stochastically perturbed to give the fault its final geometry. The perturbation consists in adding a correlated random field generated by a conditional Sequential Gaussian Simulation, whose variographic parameters are computed from input sinuosity parameters. The value of the random field is set to zero at conditional fault trace locations if no uncertainty exists. If a perturbation range is attached to fault traces, the value of the perturbation is drawn from a triangular law, whose mode is the current location. As in Lecour et al. [2001], a probability field simulation [Srivastava, 1992] is used to provide the random numbers in order to ensure a correlated drawing through data points.

2.2.3.4 Modelling the fault extension

The next simulation step consists in modelling the extension of the fault. The fault existence domain corresponds to a 3D ellipsoid centred on the fault centre. Without any further information and based on previous studies [Barnett et al., 1987, Walsh et al., 2003], the 3D ellipsoid is built from size parameters simulated from the input statistical distributions given by the users.

The simulated fault may intersect previously simulated faults, hence the final step consists in honouring fault intersections.

2.2.3.5 Stopping criterion

A theoretical number of faults is given in input for each fault family. It can be estimated by the supposed number of fault traces per fault, regional analogues or other prior information. The simulation process continues as long as the number of simulated faults is smaller than the input number of faults or some fault traces are still available, *i.e.* fault traces not included in any simulated fault.

2.2.4 Computing the stratigraphy

The method uses implicit surfaces to represent fault surfaces. Consequently, the tetrahedral mesh supporting the fault simulation is not conformable to the simulated faults. However, these discontinuities need to be introduced into the mesh to be able to compute the stratigraphic field with displacement along fault surfaces [Clément and Caumon, 2010]. This work uses a resampling and point insertion Delaunay method implemented in the SKUA framework to compute fault blocks from fault surfaces [Jayr et al., 2008, Lepage, 2003]. Then, one or several stratigraphic fields, depending on the number of stratigraphic unconformities, are interpolated through the faulted model with horizon data guiding and constraining the process [Frank et al., 2007].

2.2.4.1 Stratigraphy-related uncertainty

Stratigraphic horizon uncertainties can be divided into two terms :

1. A first order uncertainty relative to the position of a horizon far from the available information, mainly due to velocity uncertainty Abrahamsen [1992], Thore et al. [2002] ;
2. A second order uncertainty due to fault throw imprecision [Suzuki et al., 2008].

The first order uncertainty is not affected by faults, as it is continuous in the depositional space [Mallet, 2004]. The second order uncertainty is relative to the data precision and may be non negligible in low or poor resolution areas, in particular in the vicinity of faults.

2.2.4.2 Stratigraphy perturbation method

These uncertainties are handled by adding a correlated random noise representing uncertainties to the stratigraphic scalar field (figure 2.10(a) and (b)) [Caumon et al., 2007, Caumon, 2010]. This correlated noise is computed in the depositional space for the first order uncertainties ($R_1(x, y, z)$) (figure 2.10(c)), which ensures a coherent perturbation through faults. A second random field $R_2(x, y, z)$ is computed in the xyz Cartesian space to model

second order uncertainties (figure 2.10(d)) with locally varying anisotropy aligned on the strata orientation. This second random field is discontinuous across faults to account for fault throw uncertainty. In both cases, the random field is set to zero at data location to guarantee data conditioning.

Specific care is needed to ensure a homogeneous perturbation of the stratigraphic scalar field $D_S(x, y, z)$. Indeed, the random perturbation has to be scaled by the norm of the gradient of $D_S(x, y, z)$ so that the amplitude of perturbation is constant over the domain of interest [Caumon, 2010]. The perturbed stratigraphic field $D_{S'}(x, y, z)$ is obtained as follows (figure 2.10(e)) :

$$D_{S'}(x,y,z) = D_S(x,y,z) + R_1(x,y,z) \times \| \bigtriangledown D_S(x,y,z) \| + R_2(x,y,z) \times \| \bigtriangledown D_S(x,y,z) \| \quad (2.3)$$

Consequently, several stochastic stratigraphic fields can be generated for a given fault network realization. Mallet and Tertois [2010] also account for geometric uncertainty about faults and horizons but their methodology relies on a preferred deterministic geometric model, whereas here we use stochastic fault network realizations. Fault-related uncertainties have already been accounted for during the first simulation step, only the stratigraphy is perturbed here.

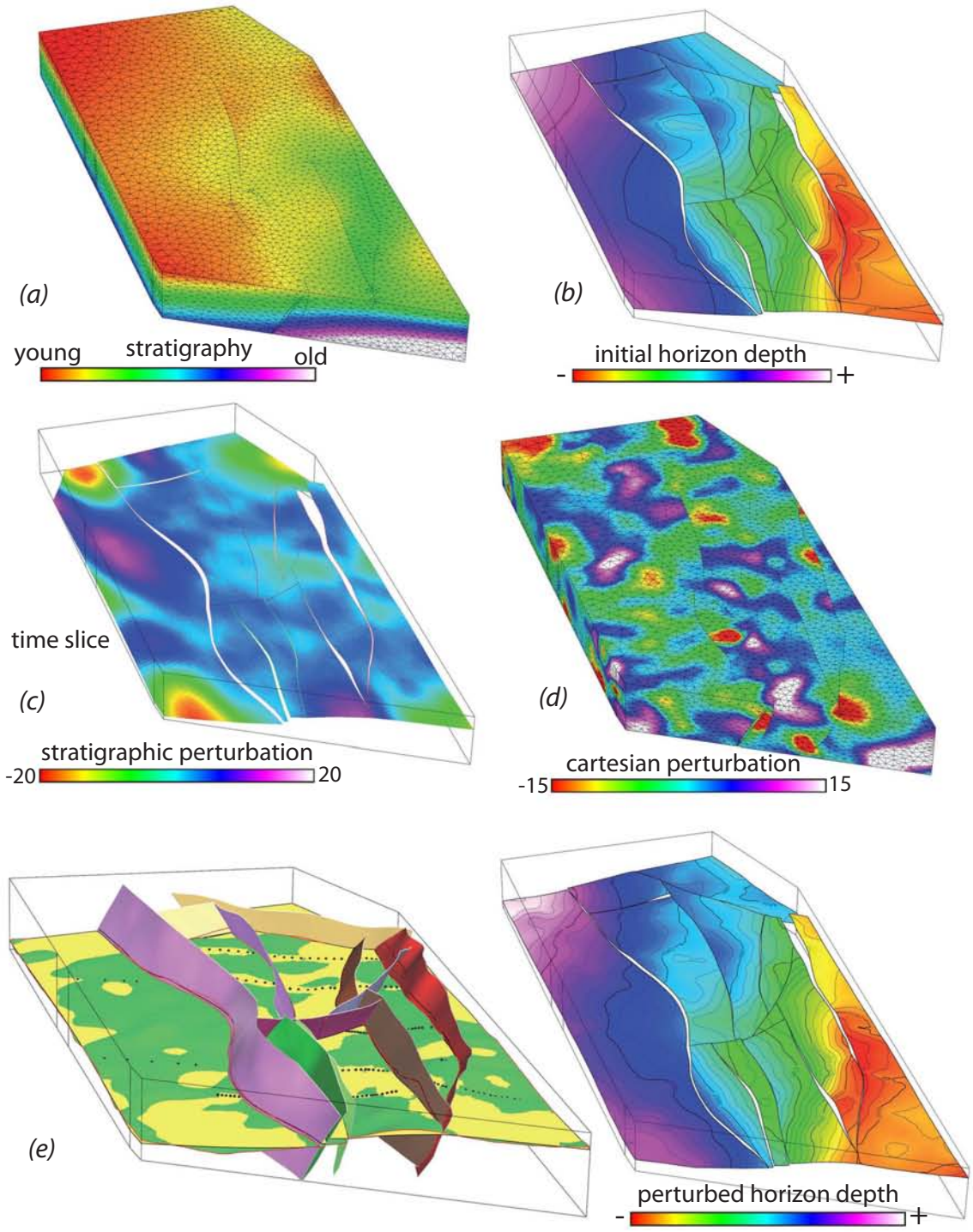

FIGURE 2.10 – Method for modeling stratigraphic uncertainty. (*a*) Reference stratigraphic field. (*b*) Initial horizon painted by the depth and contour lines. (*c*) A first-order correlated random perturbation is generated in the depositional space. Consequently, the perturbation field is continuous in a time slice. (*d*) Second-order correlated random perturbation in the Cartesian space to model fault throw uncertainty. (*e*) Left : initial (yellow) and final (green) horizons. Right : final horizon painted by the depth and contour lines.

2.2.5 Simulation choices

The simulation process presented in this work generates stochastic structural models to account for possibly large structural uncertainties. Each fault is simulated in a number of steps that sequentially simulate a characteristic (fault trace inclusion, orientation, centre, sinuosity, dimension) of the fault being built. The order of the simulation steps is arbitrary, *e.g.* the fault extension could be simulated before the data clustering step. In this case, the clustering rules would be stricter as the method actually takes the maximum fault size from input distribution. The fault centre could also be drawn after the data clustering to allow for more fault trace inclusion possibilities. However, we consider the fault centre as being a crucial parameter as it determines the neighborhood of the future fault. This choice is also motivated by the perspective of including more input information during the fault nucleation step, *e.g.* to favor nucleation in high displacement areas and avoid nucleation in low displacement areas.

In general, the simulation steps should first randomly draw the fault characteristics the most informed from observations. Then, these well constrained fault parameters should help constraining the simulation of the less informed fault characteristics, along with geological rules, physical laws or prior geological knowledge.

2.3 Application to a Middle East field case

In this section, we propose to illustrate the methodology on a Middle East field to show the impact of both fault and horizon uncertainties on gross rock volume (GRV) estimations in a sparse data situation. The studied area is a horst with normal faults and three main horizons overlying three stratigraphic reservoir units.

2.3.1 Input data

Both horizons and faults are interpreted on five fictive seismic lines, using the reference interpretation built from a 3D seismic dataset (figure 2.11). For each fault trace S, the maximum vertical offset of horizons on both sides of the fault is measured and attached to S. Interpreted data is composed of twenty-five fault traces and data points corresponding to stratigraphic horizons. Uncertainty about fault location is set at 10m for each fault trace (sphere of 10m radius around interpreted data points), with a triangular radial probability function centred on the reference location. No uncertainty is attached to horizon data. For confidentiality reasons, orientation has been changed from original data, GRV unit is arbitrary and scale is not given.

2.3.2 Fault network characterization

2.3.2.1 Fault families

From the interpreted fault traces and regional context, faults are grouped into two fault families :

1. $family_1$ is oriented [N80-N100] and dipping [60-70]. Most traces seem oriented along that direction (green data points in figure 2.11) and the minimum number of faults is set at 5 ;

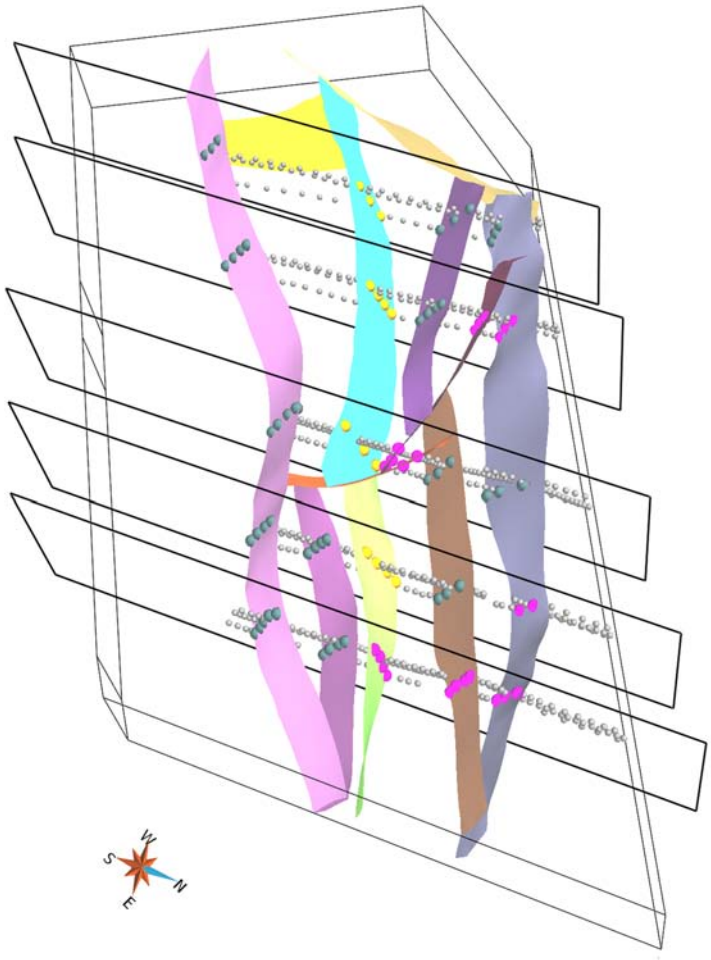

FIGURE 2.11 – The reference interpretation has been built from a full 3D seismic survey. Faults and horizons have been interpreted on five fictive seismic lines. Fault traces have been grouped according to their orientation. SW dipping traces (green colour) are set to belong to $family_1$. NE dipping traces (yellow colour) are set to belong to $family_2$. Other fault traces (pink colour) are not classified, hence can correspond to a fault from any fault family. Scale is not given and orientation has been modified for confidentiality reasons.

2. $family_2$ corresponds to opposed-dipping faults, hence is oriented [N260-N280] and dips [60-70] (yellow data points in figure 2.11) and the minimum number of faults is set at 3.

From regional reservoir analogues, a third fault family, $family_3$, oriented [N160-N180] with dip uniformly distributed in [60-70] is also expected. The number of faults for $family_3$ is set at 2. The relative age of the two first fault families cannot be deduced from the available data and tectonic history, so families are considered cogenetic, *i.e.* there is no systematic truncation rule between faults from different fault families. The third fault family, $family_3$, is considered branching on the two first families, so the algorithm simulates first $family_1$ and $family_2$.

2.3.2.2 Fault size

The maximum displacement d_{max} and fault size L are generally assumed to be correlated through a power law (equation 2.4) [Kim and Sanderson, 2005, Yielding et al., 1996]. Therefore slip information provides useful information about fault size :

$$d_{max} = cL^n \text{ with } n \in [0.5, 2] \tag{2.4}$$

The factor c corresponds to the displacement at unit length. In the absence of regional analogue data, we used the coefficient $c = 10^{-1.7}$ provided by Kim and Sanderson [2005] for normal faults. The exponent n is generally assumed to be equal to 1.

The observed slips range from 20 to 260m which corresponds to maximum fault elongation of 13km. However, the maximum observed displacement is only a rough approximation due to several factors :

1. The maximum slip may not be observed on 2D sections, so the fault size may be under-estimated ;
2. Some faults may be missed during interpretation or misinterpreted, *e.g.* two close faults may be interpreted as one single fault which tends to over-estimate the fault size.

Therefore, random variables could be used for parameters c and n in order to account for uncertainty about the maximum observed slip.

2.3.3 Stratigraphic uncertainty

Gaussian variograms are used for both continuous and discontinuous stratigraphic uncertainties in order to obtain smooth perturbations. Variogram ranges are equal to 2,500m horizontally and 600m vertically for the continuous term. The discontinuous term is modelled by a variogram with ranges equal to 1,000m horizontally and 300m vertically. The perturbation fields are simulated using a Sequential Gaussian Simulation conditioned to horizon data. The perturbation amplitude is considered centred Gaussian in both cases, with $\sigma = 17m$ and $\sigma = 8m$ for respectively the continuous and discontinuous uncertainty.

2.3.4 Results

2.3.4.1 Fault networks

190 fault networks have been simulated using interpreted fault traces and input fault families. All realizations have a larger number of faults than the reference one (figure 2.13) because most simulated faults contain less fault traces than the reference model. This may be explained by :

1. too restrictive clustering rules (displacement profile, orientation deviation) ;
2. fault size underestimation, due to an approximated maximum slip or inappropriate parameters c and n in equation 2.4. Indeed, using equation 2.4 and assuming n equal to one for a normal fault, parameter c is equal to $10^{-1.9}$ for the largest fault in the reference model. In this case, a displacement equal to 100m corresponds to a fault of 8000m long instead of 5000m using $c = 10^{-1.7}$.

2.3.4.2 Gross Rock Volumes (GRV)

For each stochastic fault network, five stratigraphic fields have been computed by adding the stratigraphic field obtained by interpolating horizon data through the fault blocks with a random noise composed of a first and second order perturbation. Consequently, 950 structural models have been generated using the method presented in this work. These models, referred to as case A, are then used to study the uncertainty on the gross rock volumes above a constant water-oil contact at 2800m depth in the three distinct stratigraphic units.

In a second case B, 950 models have been run perturbing only the stratigraphy of the reference model and using the same method and variograms than case A. The goal is to compare stratigraphic uncertainty with the proposed topological uncertainty method (case A).

The GRV values for case B are centred on the GRV values of the reference model, the difference is only of 0.01% for the total GRV estimation. In case A, the total GRV mean is about 1% greater than the reference estimation but 4% greater for unit U_1 maybe due to lower volumes than other units U_2 and U_3. The GRV interquartile range is about five times larger for all three units in case A, which suggests a higher uncertainty about the reference estimation when topological uncertainty is accounted for.

Top reservoir surfaces of a pessimistic scenario (small gross rock volume), a medium scenario (volume similar to the reference) and an optimistic scenario (volume larger than the reference) are presented in figure 2.13 and cross-sections of these scenarios are shown in figure 2.14.

Making decisions using only one deterministic model is highly risky whereas uncertainty estimation enables to be much more confident and makes the decision easier. In the present case, the reference model corresponds to a medium scenario. However, P10 and P90 values are much larger when uncertainties about faults are accounted for, which suggests higher risks for the potential recovery. These results show that considering fault geometry and connectivity uncertainty may be crucial when few data is available and the presented method samples various possible geological scenarios.

2.4 Conclusions

The presented work sets the basis for a full stochastic structural modelling workflow. The method starts by the characterization of faults using hard data, geological concepts and physical laws to help determining fault parameters that cannot be directly measured. Then, the stochastic fault model samples both geometrical and topological fault uncertainties. During the next step, one or several stratigraphic fields, depending on the number of unconformities, are computed and stratigraphic uncertainties are considered by adding a correlated random perturbation in both depositional and Cartesian spaces. Consequently, the method generates structural models with various fault connections, number of faults, fault slips and stratigraphic units without any simplification of fault and horizon geometry. This set of models can then be used for making predictions such as gross rock volumes as presented on a Middle East field case study, which enables to better evaluate risks and be more confident for decision-making. Such stochastic approach is even more relevant in sparse data situations, because of the large uncertainties and broad range of possible models

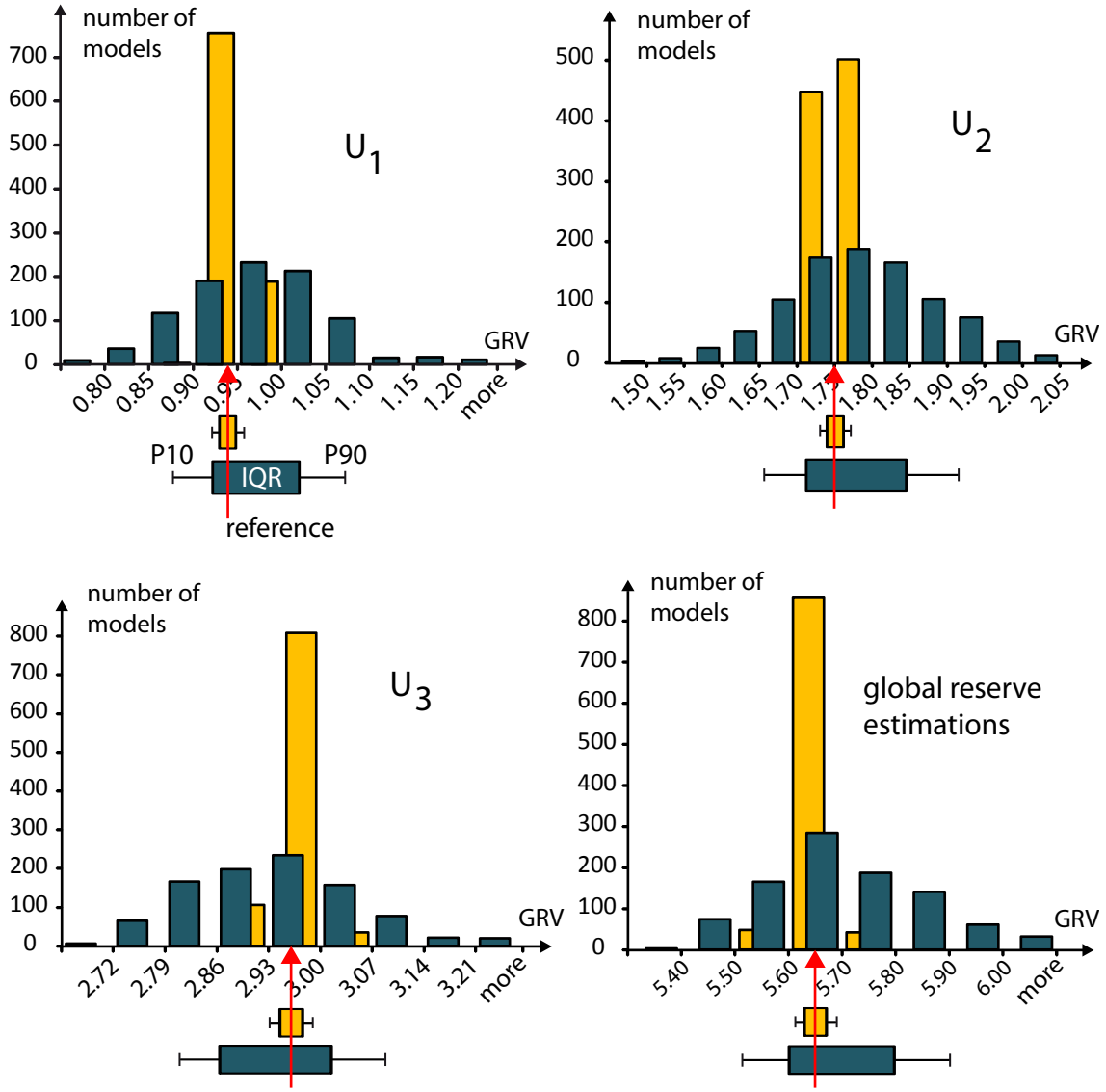

FIGURE 2.12 – Gross rock volumes estimation distributions for unit U_1, U_2, U_3 and total volume (unit is arbitrary for confidentiality reasons) for case A (blue) and B (yellow). P10, P90 and interquartile range (IQR) are represented for both cases. The reference model (red arrows) corresponds to a medium case for all units. However, considering topological uncertainties about both faults and stratigraphy (case A) enables to consider a larger range of possible models as compared to considering only the stratigraphic uncertainty (case B).

that cannot be sampled by geometrical perturbations of a deterministic model. The set of all possible structural models should then be used for the rest of the modelling workflow and ultimately be used to evaluate human or economical risks. In practice, the computational time required for such approach does not allow to consider all possible models but a subset of models. These models should be carefully selected to cover the uncertainty space, which requires relevant parameters to discriminate models.

In the presented workflow, the algorithm first generates fault surfaces, using fault displa-

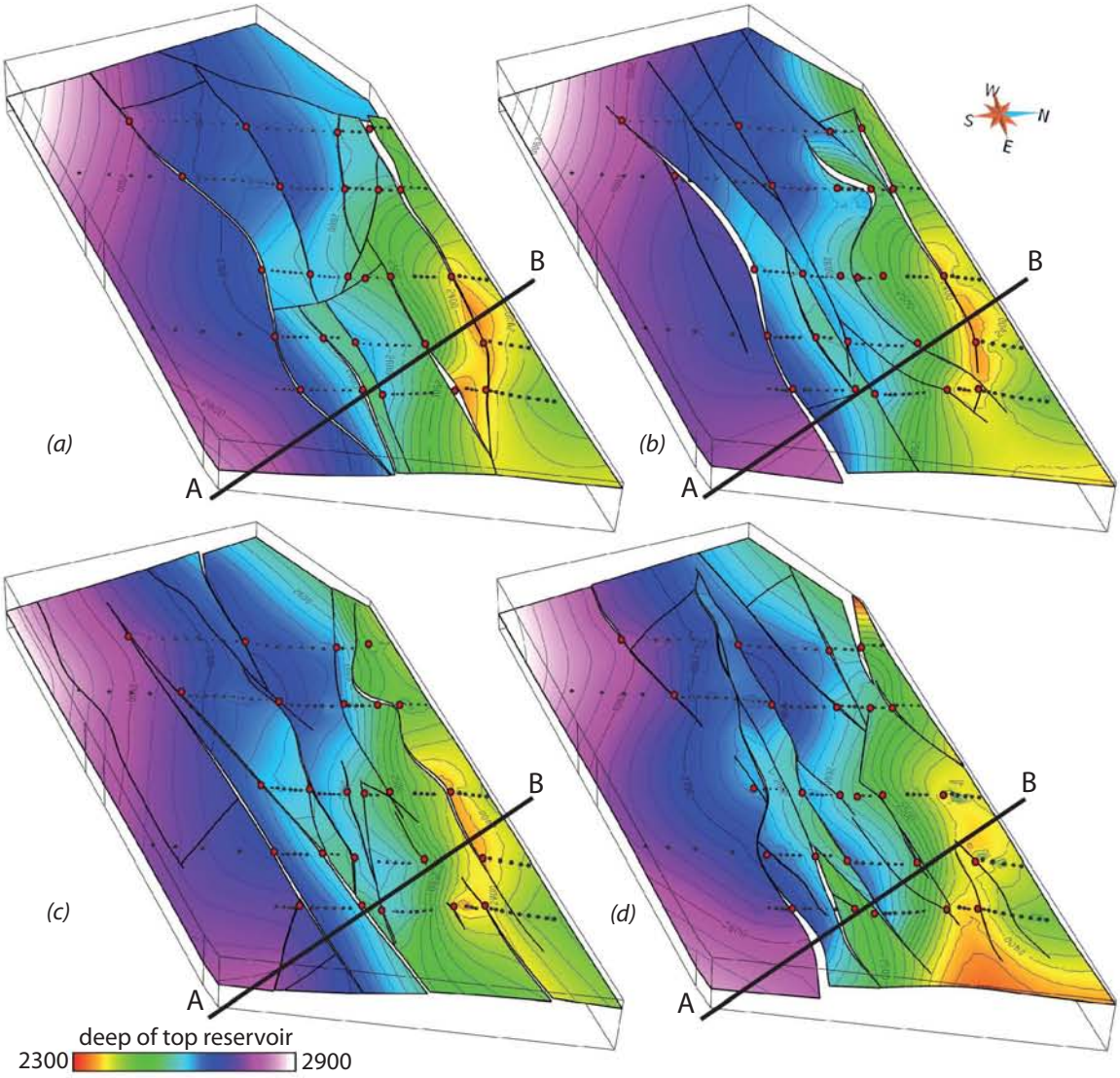

FIGURE 2.13 – Example of stochastic models (top reservoir surfaces). Top horizon is painted by the depth and contour lines for : (*a.*) the reference model, (*b.*) a pessimistic scenario, (*c.*) a medium scenario and (*d.*) an optimistic scenario. The black lines represent cross-sections of figure 2.14. Black points correspond to horizon data of the top reservoir. Red points correspond to the locations of the reference faults at the top reservoir (fault data are not visible because they are below the top reservoir). The difference in GRV estimation may be explained by a deeper top reservoir in the southern part of the studied area for the pessimistic scenario and a higher top reservoir in the eastern part for the optimistic scenario. These zones contains few data so uncertainty is large, which allows more variability in the simulated models.

cement information at the data clustering step. Then, the fault displacement is computed using horizon data. However, no check is made during fault simulation to ensure that simulated faults are consistent with horizon data, which may lead to unrealistic results. One solution could consist in adding an extra step at the end of a fault object simulation to

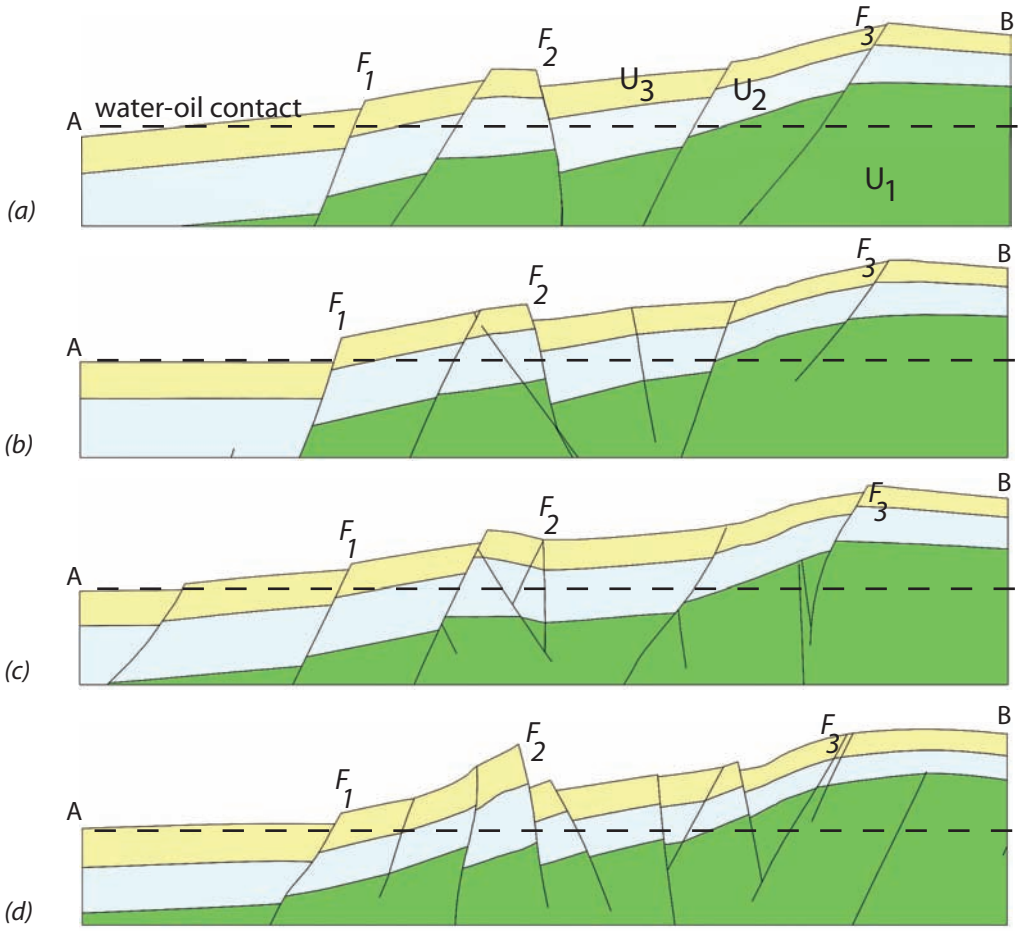

FIGURE 2.14 – Example of stochastic models (cross-sections). Cross-sections AB (location is shown in figure 2.13) corresponding to : (*a.*) the reference model, (*b.*) a pessimistic scenario, (*c.*) a medium scenario and (*d.*) an optimistic scenario. Some faults may be interpreted on all cross-sections due to the proximity of fault traces, *e.g.* faults F_1 and F_3 from $family_1$ and fault F_2 from $family_2$. The slip of fault F_1 is larger for the pessimistic scenario so that the hanging wall is below the water-oil contact. The optimistic scenario seems to contain more faults of $family_2$ than other scenarios, which entails the uplift of tilted foot walls.

simulate a fault displacement according to the fault size. Then, the algorithm could compare the global measured displacement to the current simulated displacement field and use this information to : (1) nucleate faults where observed displacement is still greater than the simulated one ; (2) avoid nucleation where no displacement has been observed and (3) stop the fault simulation when all displacement is honoured given a tolerance threshold. Instead of interpolating a stratigraphic field constrained to horizon data, the displacement of geological layers could also be modelled accounting for the rock behaviours, *e.g.* to model drag folds, directly in the vicinity of faults [Røe et al., 2010].

These improvements would constrain the simulation process to obtain more realistic results, in particular for the location and slip profile of faults. The geomechanical restoration of the simulated models could also be used to falsify inconsistent models.

Other uncertainty modelling techniques could be coupled to the presented method, *e.g.* methods accounting for well correlation uncertainties [Lallier et al., accepted], to better sample the geological model uncertainty space. Additional data could then be used to falsify models, *e.g.* mass balance calculations or production curves. Cherpeau et al. [2012] propose a parameterization of faults integrated in a stochastic inverse method and use dynamic information to infer fault characteristics and reduce fault-related uncertainties. A further step is to include this full stochastic structural method in such inverse problem to consider geological features all together and not only faults separately.

Chapitre 3

Prise en compte des structures géologiques lors de la résolution de problèmes inverses

Sommaire

3.1 Discrétisation d'un réseau de failles 85
3.1.1 Calcul des primitives géométriques 86
3.1.2 Calcul des nœuds du graphe de connectivité 94
3.1.3 Calcul des connexions du graphe de connectivité 95
3.1.4 Insertion de puits dans le graphe de connectivité 96
3.1.5 Comparaison avec d'autres supports de discrétisation 99
3.1.6 Perspectives 100
3.2 Méthode de modélisation stochastique inverse de la géométrie et de la connectivité de failles grâce à des données d'écoulement de fluides 103
3.2.1 Introduction 105
3.2.2 Related Work and Challenges of Structural Parameters Inversion . . . 106
3.2.3 Parameterization of Fault Networks 109
3.2.4 Inversion Methodology 111
3.2.5 Applications 114
3.2.6 Discussion and Conclusions 122
3.3 Gestion des nombres aléatoires 125
3.3.1 Vecteur de nombres aléatoires 125
3.3.2 Structure de données gérant les nombres aléatoires 125
3.4 Perspectives 128
3.4.1 Familles de failles cogénétiques et modélisation inverse 128
3.4.2 Échantillonnage de l'espace d'incertitudes 128
3.5 Conclusions 129

Les chapitres précédents ont abordé l'échantillonnage des incertitudes structurales grâce au développement d'une méthode stochastique de simulation de réseaux de failles (chapitre 1) et de modèles structuraux (chapitre 2). Les modèles générés peuvent être utilisés pour évaluer les incertitudes des caractéristiques du sous-sol, comme le volume de roches dans l'étude du cas du champ pétrolier *Nancy* au chapitre 2. Un modèle structural n'a pas seulement pour but de quantifier les volumes du sous-sol, il sert également à prédire le comportement des hétérogénéités structurales et pétrophysiques en réponse à un processus physique, par exemple la propagation d'ondes ou l'écoulement de fluides. Lorsque ces phénomènes sont quantifiés par des mesures ou des observations, il est alors intéressant d'utiliser ces données externes pour caractériser les hétérogénéités. Celles-ci ne peuvent néanmoins être utilisées directement puisqu'elles ne fournissent pas d'information directe sur le sous-sol mais la réponse au processus physique considéré. La formalisation du problème est appelée *problème inverse*.

La résolution du problème direct (*i.e.* la simulation du phénomène physique) fournit généralement une solution unique pour un état du sous-sol donné. En revanche, une infinité de modèles peuvent respecter les observations, c'est-à-dire que la solution d'un problème inverse n'est pas un modèle unique mais une collection de modèles cohérents avec les observations et l'information disponible a priori [Tarantola, 2004]. Par exemple, connaissant la répartition de la densité à l'intérieur de la Terre, il est possible de prédire un champ de gravité unique. Par contre, il existe une infinité de distributions de masse à l'intérieur de la Terre qui donnent exactement le champ de gravité observé (théorème de Newton).

La plupart des problèmes rencontrés en géosciences sont fortement non-linéaires avec une infinité de solutions, c'est pourquoi l'approche probabiliste est privilégiée lors de la résolution de problèmes inverses, en particulier les techniques de Monte Carlo [Mosegaard and Sambridge, 2002, Sambridge and Mosegaard, 2002]. Celles-ci sont utilisées aussi bien pour des problèmes d'optimisation (*e.g.* méthode du recuit-simulé, algorithmes génétiques) que pour l'échantillonnage de l'espace d'incertitude. L'algorithme de Metropolis [Metropolis et al., 1953] est la méthode d'échantillonnage la plus connue et permet d'échantillonner la fonction de probabilité *a posteriori* (probabilité des paramètres du modèle étant donné les observations). Cet algorithme est utilisé dans de nombreux domaines scientifiques et des variantes ont été proposées afin d'améliorer son efficacité pour l'exploration de l'espace d'incertitudes (*e.g.* Smith and Marshall [2008] comparent plusieurs méthodes dans le domaine de l'hydrologie). Ces approches deviennent accessibles grâce à la rapidité, en augmentation permanente, des ordinateurs actuels disponibles pour de tels calculs.

La résolution stochastique du problème inverse consiste donc à échantillonner des modèles suivant la distribution *a posteriori* des paramètres du modèle. C'est un processus itératif : à chaque étape les paramètres du modèle courant sont perturbés de manière à produire un nouveau modèle dont la réponse est calculée (résolution du problème direct). En essence, de nouveaux modèles sont générés jusqu'à la convergence de l'algorithme (figure 3.1).

L'approche peut être découpée en 3 étapes [Tarantola, 2004] :

1. *paramétrisation du système* : le système est décrit par un ensemble de paramètres qui le détermine complètement ;
2. *modélisation directe* : un processus physique est simulé, pour un système donné, afin de fournir des prédictions sur les variables observées ;

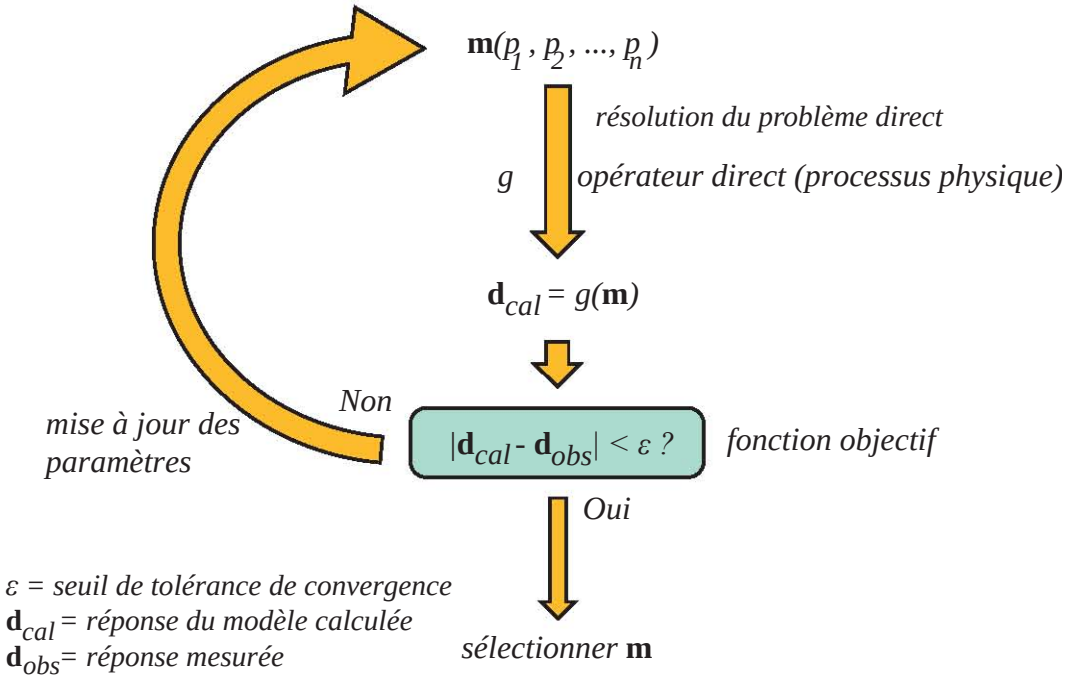

FIGURE 3.1 – Approche générale de la résolution d'un problème inverse. Un modèle $\mathbf{m}$ est complètement décrit par un ensemble de paramètres. Un opérateur g permet de simuler un processus physique étant donné $\mathbf{m}$. Les prédictions $\mathbf{d}_{cal}$ du modèle $\mathbf{m}$ sont ensuite comparées aux données observées $\mathbf{d}_{obs}$. Une fonction objectif est alors évaluée à partir de $\mathbf{d}_{cal}$ et $\mathbf{d}_{obs}$. Si le modèle reproduit les observations, il est alors sélectionné en tant qu'échantillon de la probabilité *a posteriori*. Les paramètres du modèle courant sont ensuite perturbés de manière à fournir un nouveau modèle.

3. *modélisation inverse* : des caractéristiques des paramètres du modèle sont inférées à partir des observations.

Dans le cadre du problème inverse, je me suis intéressé dans cette thèse à l'utilisation de données d'écoulement afin de réduire les incertitudes sur les réseaux de failles. Le problème direct est alors une simulation d'écoulement qui permet de fournir des données de saturation et de pression en tout point du modèle. Cette étape est abordée dans la première partie de ce chapitre. La paramétrisation d'un réseau de failles ainsi qu'une stratégie de modélisation inverse sont proposées dans la deuxième partie de ce chapitre, constituée d'un article paru dans *Mathematical Geosciences* [Cherpeau et al., 2012].

3.1 Discrétisation d'un réseau de failles

La résolution d'un problème inverse nécessite dans un premier temps de résoudre le problème direct $g(\mathbf{m})$. Dans notre cas, il s'agit d'évaluer l'écoulement de fluides dans un milieu faillé. Or, le modèle stochastique de failles génère des surfaces de failles implicites, une faille étant une équipotentielle f d'un champ scalaire volumique ($\mathcal{F}(x, y, z) = f$). Ces surfaces implicites sont générées dans un maillage tétraédrique couvrant l'ensemble du domaine d'intérêt. Une étape de discrétisation de ces résultats est donc nécessaire afin de

pouvoir simuler les écoulements de fluides.

Cette partie détaille la discrétisation d'un réseau de failles publiée dans Cherpeau et al. [2011] (disponible en annexe A). Un réseau de failles est discrétisé en un ensemble de nœuds connectés entre eux, représentant les failles et la matrice. La structure de données utilisée pour représenter ce graphe de connectivité a été développée par Vitel [2007] et est constituée de deux sous-structures :

- un *ensemble de nœuds*, chaque nœud représente un volume de contrôle avec les propriétés pétrophysiques (porosité, perméabilité) ;
- un *ensemble de connexions*, chaque connexion relie deux nœuds et stocke la transmissibilité entre les deux nœuds.

L'algorithme de discrétisation est réalisé en trois étapes :

1. un ensemble de *primitives géométriques* est calculé, chaque primitive représentant un volume de matrice ou de faille ;
2. les *nœuds* sont créés à partir du barycentre de chaque primitive géométrique ;
3. les *connexions* sont créés entre deux primitives ayant une face commune.

3.1.1 Calcul des primitives géométriques

3.1.1.1 Positions possibles d'un tétraèdre par rapport à une faille

Les primitives géométriques sont calculées à partir du maillage tétraédrique, il est donc nécessaire de pouvoir évaluer la position d'un tétraèdre par rapport à une faille. Chaque faille F est définie par une équipotentielle f d'un champ scalaire volumique ($\mathcal{F}(x, y, z) = f$) dont les valeurs sont stockées sur les nœuds du maillage tétraédrique. La position d'un tétraèdre T par rapport à F est donc évaluée à partir de la valeur de $\mathcal{F}(x, y, z)$ aux sommets de T. L'extension de la faille est limitée par un second champ scalaire ($\mathcal{D}_F(x, y, z) = d_F$) :

$$\mathcal{F}(x, y, z) = f | \mathcal{D}_F(x, y, z) \leq d_F \tag{3.1}$$

Un tétraèdre T peut être dans trois configurations différentes par rapport à F :

- *non-intersecté* : T n'est pas intersecté par F ou leur intersection est en dehors de $\mathcal{D}_F$: $T \cap F = \emptyset$ ou $T \cap F \cap \mathcal{D}_F = \emptyset$
- *intersecté* : l'intersection de T avec F est incluse dans $\mathcal{D}_F$: $T \cap F \neq \emptyset$ et $T \cap F \subset \mathcal{D}_F$
- *partiellement intersecté* : l'intersection de T avec F est partiellement incluse dans $\mathcal{D}_F$: $T \cap F \neq \emptyset$ et $T \cap F \cap \mathcal{D}_F \neq T \cap F$

3.1.1.2 Structures de données

Plusieurs structures de données sont utilisées lors du calcul des primitives géométriques. Le terme *primitive* désigne les polyèdres de la matrice ainsi que les polygones des failles.

Structure représentant un polyèdre de matrice

La structure *PFace* correspond à une face de polyèdre et stocke les informations suivantes :

- un tableau des sommets de la face ;
- un lien vers la face voisine (éventuellement) ;
- un lien vers la primitive (structure *Polyedre* ou *Polygone*) à laquelle elle appartient ;

- un lien vers la primitive (de type *Polyedre* ou *Polygon* dans le cas d'une faille) vers laquelle elle pointe (éventuellement) (peut être différente de la primitive de la face voisine dans le cas d'intersection entre failles, voir section 3.1.1.6) ;
- l'identifiant signé (positif si du côté $\mathcal{F}(x, y, z) > f$) de la faille dont elle est la voisine (éventuellement).

La structure *Polyedre* correspond à un polyèdre définissant un volume de matrice et stocke les informations suivantes :

- la liste des structures *PFace* constituant le polyèdre ;
- un lien vers le tétraèdre contenant le polyèdre ;
- son volume.

Structure représentant un polygone de faille

La structure *Polygon* correspondant à un polygone de faille contient une liste de structures *PSegment* qui sont un type spécialisé de la structure *PFace*. En effet, les sommets de *PSegment* définissent un segment et non une face. Le reste de la structure *Polygon* contient un lien vers son tétraèdre ainsi que sa surface.

3.1.1.3 Étapes principales de l'algorithme

Une faille peut se brancher ou recouper les failles qui lui sont antérieures. Par conséquent, les failles sont discrétisées dans l'ordre de leur simulation, de la plus ancienne à la plus récente de manière à pouvoir calculer les primitives géométriques (polyèdres et polygones) résultant du branchement ou du recoupement entre deux failles. Chaque faille F est discrétisée en cinq étapes :

1. l'algorithme recherche les tétraèdres intersectés ou partiellement intersectés par F et les divise en deux ensembles (section 3.1.1.4) :
   - S_{cut} : ensemble des tétraèdres déjà intersectés par une autre faille, c'est-à-dire composé de plusieurs primitives géométriques ;
   - S_{uncut} : ensemble des tétraèdres qui ne sont pas intersectés par d'autres failles.
2. L'ensemble de tétraèdres S_{uncut} est intersecté par F (section 3.1.1.5) ;
3. Pour chaque tétraèdre de S_{cut}, l'algorithme sélectionne les primitives à couper à partir des relations entre failles, puis les coupe par F (section 3.1.1.6) ;
4. Les tétraèdres partiellement intersectés par F sont intersectés par le domaine d'existence de F (section 3.1.1.7) ;
5. Les faces communes des nouvelles primitives sont connectées (section 3.1.1.8).

3.1.1.4 Recherche des tétraèdres intersectés par une faille

L'algorithme 5 se propage de tétraèdre intersecté en tétraèdre intersecté à partir du centre de faille. Il se propage à travers les faces intersectées tant que la frontière du domaine d'existence $\mathcal{D}_F$ ou qu'une faille parente ne sont pas rencontrées.

3.1.1.5 Calcul de primitives géométriques à partir d'un tétraèdre intersecté par une faille

Le champ scalaire $\mathcal{F}(x, y, z)$ d'une faille F est linéaire à l'intérieur d'un tétraèdre. L'intersection d'un tétraèdre T par une faille F consiste donc à couper T par un plan. L'en-

Entrées :
F= faille
c_F = centre de F
T_F = tétraèdre contenant c_F
$\mathcal{R}_F$ = région définie par l'ensemble des failles parentes de F

Définitions :
S = liste de tétraèdres
S_{cut} = liste de tétraèdres complètement intersectés
S_{uncut} = liste de tétraèdres partiellement intersectés

$S \leftarrow T_F$
tant que $S \neq \emptyset$ **faire**
  $T \leftarrow$ dernier tétraèdre de S
  si T *est composé de plusieurs primitives* **alors**
    $S_{cut} \leftarrow T$
  sinon
    $S_{uncut} \leftarrow T$
  finsi
  S_{faces} = liste de faces
  pour chaque *face* f_i *de* T **faire**
    si f_i *n'est pas intersectée* **alors**
      continuer
    sinon
      calculer $(p_1, p_2) = f_i \cap F$
      si $(p_1 \in \mathcal{D}_F$ *et* $p_1 \in \mathcal{R}_F)$ *ou* $(p_2 \in \mathcal{D}_F$ *et* $p_1 \in \mathcal{R}_F)$ **alors**
        $S_{faces} \leftarrow f_i$
      finsi
    finsi
  fin
  pour chaque *face* f_i *de* S_{faces} **faire**
    T_i = tétraèdre connecté à f_i $(T_i \neq T)$
    si $T_i \notin S$ **alors**
      $S \leftarrow T_i$
    finsi
  fin
fin
Sorties :
S_{cut}
S_{uncut}

Algorithme 5: Recherche des tétraèdres intersectés par une faille.

semble des configurations possibles peut se ramener à quatre cas distincts selon la valeur de $\mathcal{F}(x, y, z)$ aux sommets de T :

1. n sommets où $\mathcal{F}(x, y, z) < f$ et $4 - n$ sommets où $\mathcal{F}(x, y, z) > f$, avec $n = 1$ ou $n = 3$ (figure 3.2*a*) ;
2. 2 sommets où $\mathcal{F}(x, y, z) < f$ et 2 sommets où $\mathcal{F}(x, y, z) > f$ (figure 3.2*b*) ;
3. n sommets où $\mathcal{F}(x, y, z) < f$, 1 sommet où $\mathcal{F}(x, y, z) = f$ et $3 - n$ sommets où $\mathcal{F}(x, y, z) > f$, avec $n = 1$ ou $n = 2$ (figure 3.2*c*) ;
4. 1 sommet : $\mathcal{F}(x, y, z) < f$, 2 sommets : $\mathcal{F}(x, y, z) = f$, 1 sommet : $\mathcal{F}(x, y, z) > f$ (figure 3.2*d*).

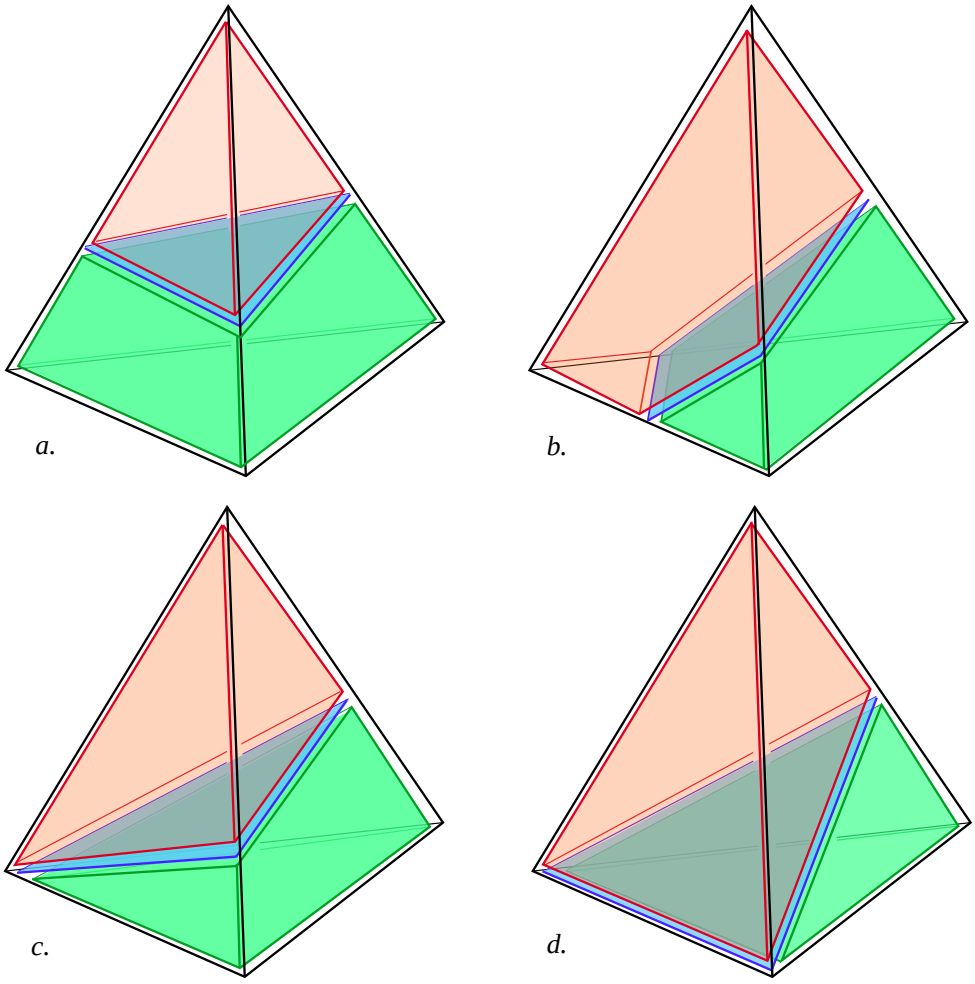

FIGURE 3.2 – Configurations possibles d'un tétraèdre par rapport à une faille. L'algorithme génère trois primitives géométriques différentes selon les configurations a, b, c ou d : deux polyèdres de matrice (rouge et vert) et un polygone de faille (bleu).

L'algorithme génère les primitives géométriques selon l'une des quatre configurations possibles (figure 3.2). Dans les quatre cas, deux polyèdres de matrice et un polygone de faille sont construits. Le point d'intersection $pt(x_{pt}, y_{pt}, z_{pt})$ de F ($\mathcal{F}(x, y, z) = f$) entre deux

sommets $A(x_a, y_a, z_a)$ et $B(x_b, y_b, z_b)$ d'un tétraèdre est calculé à partir de l'interpolation linéaire de leur valeur :

$$\begin{aligned} \lambda &= \frac{f - \mathcal{F}(x_b, y_b, z_b)}{\mathcal{F}(x_a, y_a, z_a) - \mathcal{F}(x_b, y_b, z_b)} \\ (x_{pt}, y_{pt}, z_{pt}) &= \lambda(x_a, y_a, z_a) + (1 - \lambda)(x_b, y_b, z_b) \end{aligned} \tag{3.2}$$

3.1.1.6 Calcul de primitives géométriques à partir d'un tétraèdre intersecté par plusieurs failles

Lorsque le tétraèdre coupé par F contient déjà des primitives géométriques (polyèdres ou polygones), l'algorithme sélectionne d'abord les primitives à couper grâce aux relations entre failles. En effet, il est facile d'évaluer la position d'une primitive de type *Polyedre* si elle est adjacente à la faille parente F_{parent} car l'une de ses faces (structure $PFace$) contient un identifiant signé qui donne le côté par rapport à F_{parent}. Si la primitive n'est pas adjacente à F_{parent}, sa position est évaluée par rapport à la position de son barycentre.

Lors de l'intersection d'un polygone PG_1 correspondant à une faille F_1, avec un polygone PG_2 correspondant à une faille F_2 plus récente, c'est le polygone PG_1 qui est coupé en deux polygones. Il y a donc une intersection triple au niveau du segment (correspondant à une structure $PSegment$) commun entre les trois polygones (PG_1, PG_2 et PG_3 figure 3.3). Dans ce cas, chaque segment contient un lien vers un polygone voisin, de manière à former un cycle. Par exemple, soient trois segments s_1, s_2, s_3 contenus respectivement dans les structures PG_1, PG_2 et PG_3 de type *Polygone*. Chaque segment contient un lien vers son polygone parent, et un lien vers l'un des polygones voisins. s_1 pointe vers PG_2, s_2 vers PG_3 et s_3 vers PG_1 : un cycle est formé. Ces cycles sont facilement repérables par l'algorithme et seront utilisés lors du calcul des connexions entre polygones de failles (section 3.1.3.2).

Si une faille postérieure recoupe une faille antérieure, alors le cycle correspondant à l'intersection sera constitué de quatre segments appartenant chacun à l'un des quatre polygones résultant de l'intersection.

3.1.1.7 Intersection d'un tétraèdre par le domaine d'existence d'une faille

Lorsqu'un tétraèdre est coupé par le domaine d'existence $\mathcal{D}_F$ d'une faille F, une étape supplémentaire est nécessaire afin de recouper les polygones correspondant à la faille par $\mathcal{D}_F$. La position des polygones correspondant à F est donc évaluée : si le polygone n'est pas coupé par $\mathcal{D}_F$, aucune opération n'est nécessaire. En revanche, dans le cas où un polygone PG est intersecté par $\mathcal{D}_F$, celui-ci est coupé en deux et la partie se situant à l'extérieur de $\mathcal{D}_F$ est supprimée. Les deux faces appartenant à des polyèdres de matrice de part et d'autre de F, adjacentes à PG sont également coupées en deux, les deux faces extérieures à $\mathcal{D}_F$ pointant l'une vers l'autre. Ceci permettra de créer une connexion entre les nœuds des deux polyèdres de matrice (figure 3.4).

Le tétraèdre voisin T_{voisin} du tétraèdre T intersecté par $\mathcal{D}_F$ mais qui n'est pas dans $\mathcal{D}_F$ ne fait pas partie des ensembles S_{cut} et S_{uncut}. Pourtant, les faces des primitives de T_{voisin} adjacentes à celles de T qui ont été coupées, doivent également être coupées afin de maintenir la cohérence entre faces (la face f_2 doit être coupée en figure 3.5*b*). Ces faces sont identifiées par leur lien avec les faces coupées appartenant aux primitives de T.

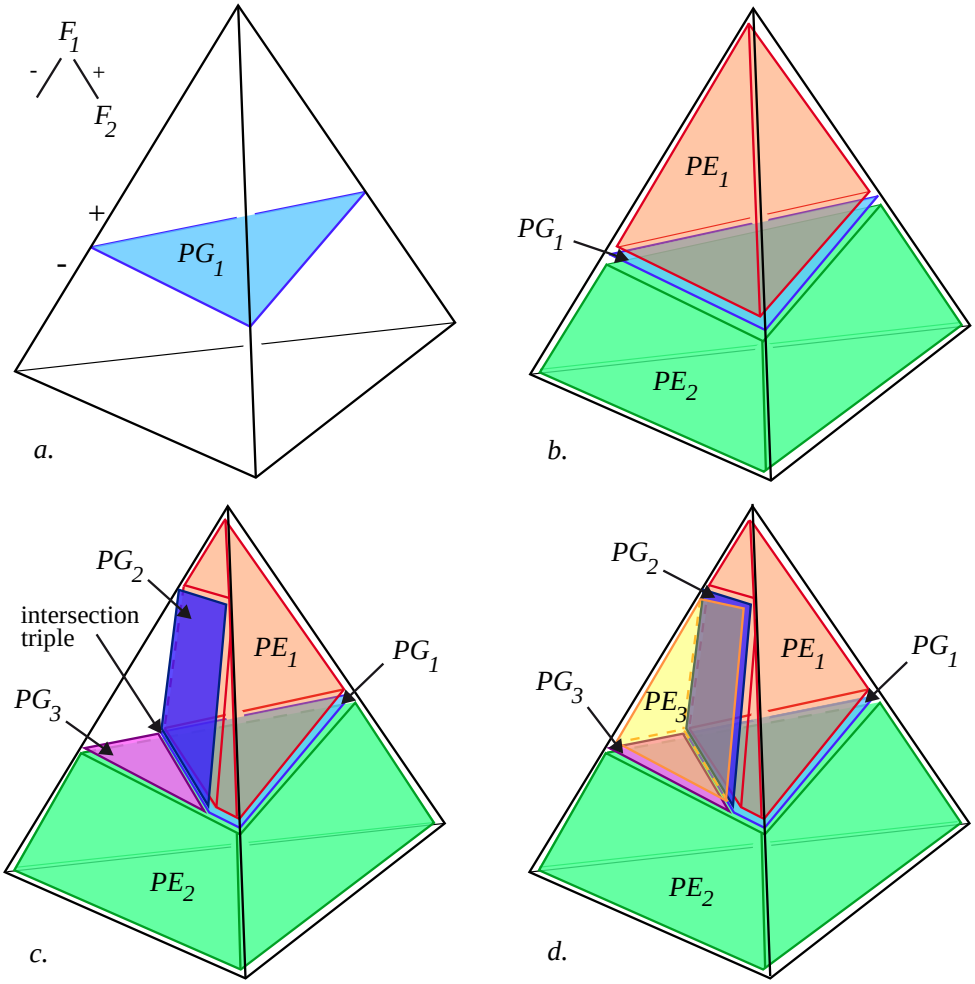

FIGURE 3.3 – Calcul de l'intersection de plusieurs failles dans un tétraèdre. *a* La primitive géométrique PG_1 de type *Polygone* correspond à la faille F_1. *b* Avant l'intersection avec une seconde faille F_2, le tétraèdre contient trois primitives : PG_1, PE_1 et PE_2. *c* Le polyèdre PE_1 situé du côté de F_2 dans l'arbre est coupé, ainsi que le polygone PG_1 correspondant à la première faille. il y a donc une intersection triple entre les polygones PG_1, PG_2 et PG_3. *d.* Le tétraèdre contient maintenant six primitives : deux polygones PG_1 et PG_3 appartenant à F_1, un polygone PG_2 appartenant à F_2 et trois polyèdres PE_1, PE_2 et PE_3 correspondant à la matrice.

3.1.1.8 Création des liens entre faces nouvellement créées

Après l'intersection des tétraèdres des ensembles S_{cut} et S_{uncut}, les nouvelles primitives sont connectées entre elles afin de garantir la cohérence de l'ensemble. En effet, à chaque nouvelle faille discrétisée, il est important de connaître les liens entre primitives afin de trouver celles dont les faces doivent être intersectées.

Dans la figure 3.5, les faces f_3, f_4, f_5, f_6 ont été générées suite à l'intersection de la faille F_2 et les liens sont créés lors de la dernière étape (figure 3.5*d*).

Pour un tétraèdre donné, l'algorithme compare les sommets des faces de ses primitives avec ceux des primitives appartenant aux tétraèdres voisins. Afin d'éviter les erreurs nu-

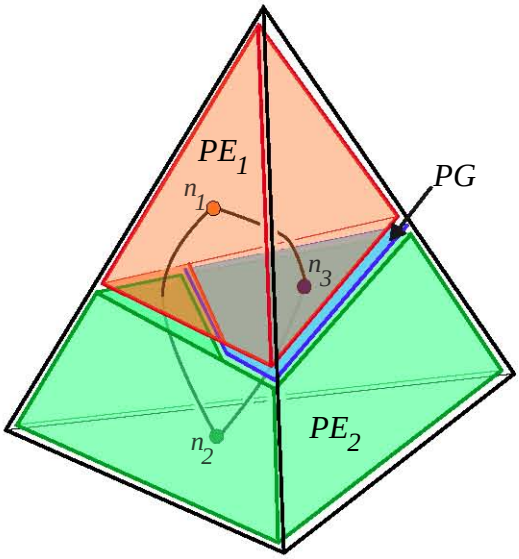

FIGURE 3.4 – Intersection partielle dans un tétraèdre. Le polygone de la faille PG est recoupé par le domaine d'existence de sorte qu'il ne traverse pas complètement le tétraèdre. Les faces des deux polyèdres adjacents PE_1 et PE_2 sont également coupées en deux. Les faces extérieures au domaine d'existence pointent l'une vers l'autre. Lors de la création des connexions entre les nœuds n_1, n_2 et n_3 appartenant respectivement à PE_1, PE_2 et PG, ces liens permettent de relier les deux polyèdres PE_1 et PE_2 de part et d'autre de la faille.

mériques, les sommets ne sont pas comparés par rapport à leur coordonnées x, y, z mais par rapport à leur adresse mémoire. Par conséquent, chaque sommet n'existe qu'en un seul exemplaire, et les différentes faces stockent un lien vers le sommet partagé.

3.1.1.9 Stratégies pour éviter les erreurs numériques

Le processus de discrétisation requiert de nombreux calculs numériques pour l'intersection de primitives géométriques, de faces, mais également pour évaluer la position des sommets des faces par rapport aux surfaces de failles. Nous avons vu que les sommets des faces sont partagés, c'est-à-dire qu'il n'existe qu'une seule instance par sommet avec des coordonnées identiques. Les faces ne stockent que des liens vers ces sommets partagés. Cependant, la position de ces sommets par rapport aux failles est évaluée à différents moments de l'algorithme, et surtout dans différentes configurations. Par exemple, la position d'un sommet st commun à deux primitives appartenant à des tétraèdres différents T_1 et T_2 sera évaluée différemment dans T_1 et T_2. En effet, une faille F recoupant T_1 et T_2 sera représentée par un plan P_1 dans T_1 et P_2 dans T_2. Les deux plans ne sont pas identiques et la position de st peut donc être évaluée différemment par rapport à P_1 et P_2.

Pour chaque sommet, l'algorithme stocke sa position par rapport à une faille donnée afin d'éviter ces erreurs numériques. De cette manière, la position de chaque sommet n'est évaluée qu'une seule fois. Même si une erreur d'arrondi est commise lors de la première évaluation de la position, l'algorithme utilisera ce résultat tout au long du processus de discrétisation et reste ainsi cohérent.

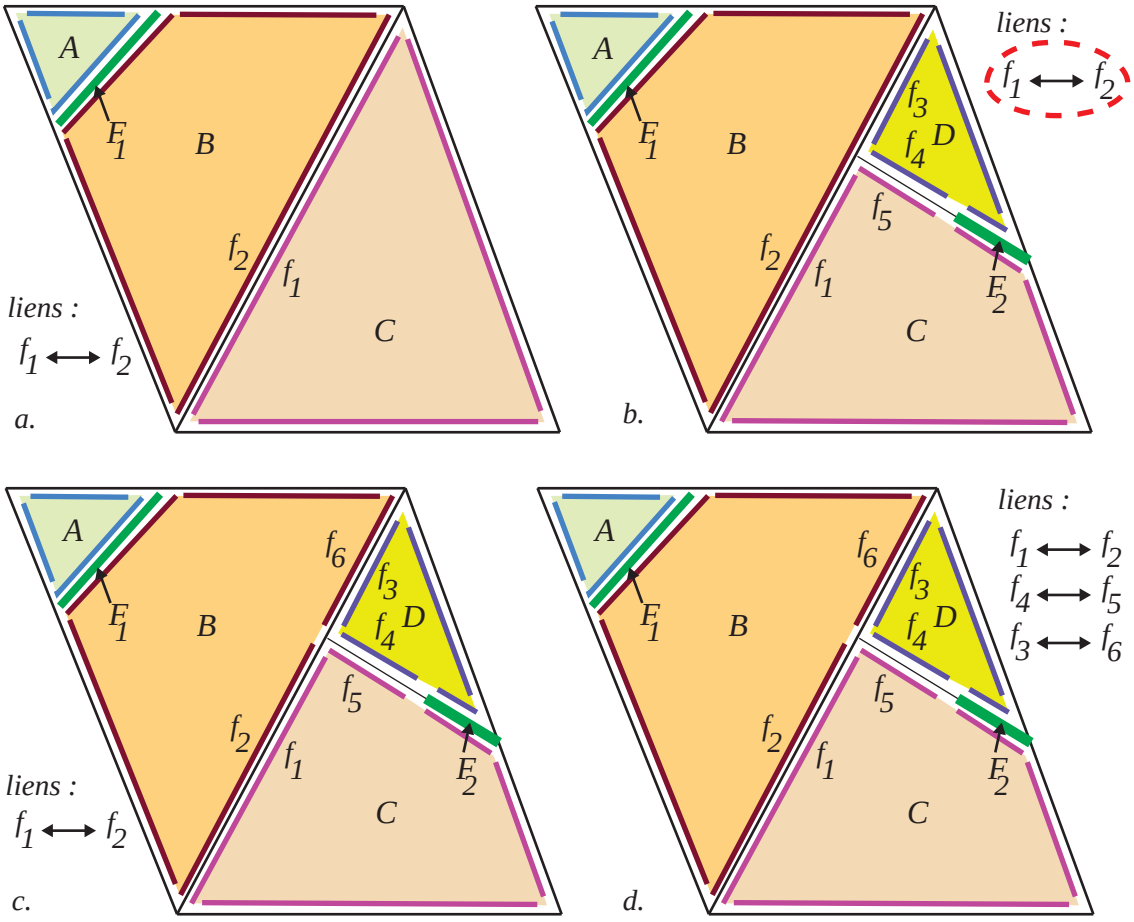

FIGURE 3.5 – Intersection partielle et mise à jour des liens (vue de dessus). *a.* Le tétraèdre de gauche contient trois primitives : A, B et la primitive correspondant à la faille F_1. L'autre tétraèdre ne contient que C. *b.* Une faille F_2 recoupe la primitive C. La primitive D est créée. La faille F_2 ne recoupe pas complètement le tétraèdre, les faces f_4 et f_5 sont voisines. Le lien entre f_1 et f_2 existe toujours mais f_2 doit être intersectée par F_2 pour garder une cohérence entre faces. *c.* La face f_2 est coupée en deux, f_6 est créée. *d.* Les faces voisines nouvellement créées sont connectées, les faces f_3, f_6 et f_4, f_5 sont liées.

Lors du calcul du point d'intersection pt entre le segment formé par deux sommets consécutifs pt_a, pt_b d'une face et un plan d'une faille F, l'algorithme associe le couple (pt_a, pt_b) avec le point pt et la faille F. Lorsque la même intersection doit être calculée dans une face ou une primitive (potentiellement d'un autre tétraèdre) différentes, l'algorithme utilise alors l'information stockée et renvoie le même point pt. Cette stratégie limite les erreurs numériques et permet ainsi de partager les points créés au fur et à mesure du processus de discrétisation.

La robustesse des algorithmes géométriques est cruciale dans de nombreux domaines. En effet, les algorithmes géométriques reposent sur des théorèmes et une géométrie réelle, mais sont appliqués sur des nombres flottants utilisés par les ordinateurs. Kettner et al. [2008] montrent à travers l'exemple de deux algorithmes simples (calcul de l'enveloppe convexe

d'un ensemble de points et triangulation de Delaunay) les raisons et la manière aboutissant à un résultat erroné dans un algorithme géométrique. Ces problèmes sont largement étudiés en géométrie algorithmique et deux types de solutions ont été proposées : les approches arithmétiques (augmentation de la précision) ou les approches géométriques (préservation de propriétés géométriques et topologiques des objets concernés) [Schirra, 2000, Yap, 2004]. Parmi les approches arithmétiques, le calcul géométrique exact [Li et al., 2005] garantit que chaque évaluation d'un prédicat géométrique sur les données d'entrée est correcte. Le calcul par intervalle permet de ne réaliser un calcul exact (coûteux en temps de calcul) que dans le cas d'un résultat indéterminé. Cette approche a par exemple été implémentée dans la librairie *CGAL* [Fabri et al., 2000] pour garantir la robustesse des algorithmes géométriques, avec un surcoût en temps de calcul raisonnable. Son utilisation, en complément des stratégies déjà mises en place, permettrait d'améliorer la robustesse de la discrétisation proposée. La robustesse ne peut être négligée en vue de l'utilisation de la discrétisation lors des très nombreuses évaluations du problème direct requises par les méthodes de type Monte Carlo pour la résolution des problèmes inverses.

3.1.2 Calcul des nœuds du graphe de connectivité

Après la première étape, l'algorithme a créé un ensemble de primitives, polyèdres et polygones, connectées entre elles par leurs faces communes. La deuxième étape consiste à générer l'ensemble des nœuds du graphe de connectivité.

3.1.2.1 Calcul du volume des primitives géométriques

L'algorithme commence par calculer le volume de chaque primitive. Comme chaque polyèdre est convexe, chaque face à n sommets est parcourue et le volume signé du tétraèdre formé par un point pris au hasard pt_1, deux points consécutifs pt_i, pt_{i+1} et le barycentre g de la primitive, est ajouté au volume V :

$$V = \sum_{i=2}^{n-1} \frac{1}{6}(pt_1 - pt_i) \otimes (pt_1 - pt_{i+1}) \cdot (pt_1 - g) \tag{3.3}$$

Un algorithme plus général de calcul du volume d'un polyèdre quelconque, n'utilisant pas le barycentre, a été proposé par Gelder [1995]. Le calcul du volume des primitives correspondant aux failles est différent puisqu'elles ne contiennent pas de faces (structure $PFace$) mais des segments (structure $PSegment$). Le volume est égal à l'aire de la surface délimitée par les segments multipliée par un facteur donné en entrée correspondant à l'épaisseur de la faille. Ce volume est ensuite retranché aux volumes des primitives adjacentes proportionnellement à leur aire de contact avec l'élément faille considéré.

3.1.2.2 Fusion des primitives ayant un volume trop petit

L'étape de création des primitives peut générer des éléments ayant un très faible volume. Ces éléments participent peu à l'écoulement des fluides mais entraînent des erreurs numériques lors du calcul de l'écoulement, dues à leur faible volume en comparaison de la majorité des éléments. Après l'étape de calcul des volumes, l'algorithme fusionne les primitives ayant un volume trop faible par rapport à un seuil donné. Chaque primitive est

prioritairement fusionnée avec sa voisine ayant le volume le plus faible. Les polyèdres représentant la matrice ne peuvent être fusionnés qu'entre eux, tout comme les polygones représentant une faille ne peuvent être fusionnés qu'avec un polygone de la même faille.

Lors de la fusion entre deux primitives, les faces communes sont supprimées et les liens avec les primitives voisines sont mis à jour.

3.1.2.3 Création des nœuds

L'algorithme itère sur chaque primitive et crée un nœud à partir de son barycentre. À cette étape, les propriétés pétrophysiques du tétraèdre contenant la primitive sont assignées au nœud. Chaque réseau de failles comporte au final trois types de nœuds :

- des *nœuds de matrice* correspondant à des tétraèdres non intersectés par des failles ;
- des *nœuds de matrice* représentant un polyèdre d'un tétraèdre coupé par une ou plusieurs failles ;
- des *nœuds* correspondant aux polygones de failles.

La perméabilité particulière des failles peut être prise en compte lors de cette étape, en assignant des valeurs de perméabilité aux nœuds de failles différentes de celles de la matrice.

3.1.3 Calcul des connexions du graphe de connectivité

La dernière étape consiste à créer les connexions entre les nœuds voisins. Une connexion est créée entre chaque paire de faces liées.

3.1.3.1 Calcul de la transmissibilité entre deux faces

La transmissibilité T_{ij} entre deux nœuds représentant deux primitives i et j est calculée, pour un schéma d'approximation du flux à deux points, par la formule suivante [Karimi-Fard et al., 2004] :

$T_{ij} = \frac{T_i \times T_j}{T_i + T_j}$
avec :
$T_i = \frac{A_{ij} \times k_i}{\|c\vec{c}_i\|} \vec{n_i} \cdot \vec{g_i}$
A_{ij} l'aire de la face commune entre i et j
k_i la perméabilité de l'élément i

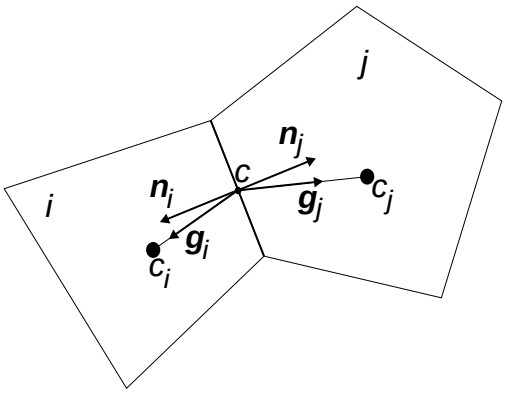

FIGURE 3.6 – Calcul de la transmissibilité entre deux primitives géométriques.

Le flux q_{ij} de l'élément i vers l'élément j s'exprime dans un schéma d'approximation du flux à deux points par :

$$q_{ij} = T_{ij}(p_j - p_i)$$
avec : (3.4)
p_i la pression de l'élément i

3.1.3.2 Calcul de la transmissibilité au niveau de l'intersection entre failles

La discrétisation exacte du branchement d'une faille F_A sur une faille F_B devrait engendrer un petit élément représentant l'intersection. L'analogie aux résistances électriques, valable en régime permanent pour un fluide incompressible, proposée par Karimi-Fard et al. [2004] est utilisée afin de calculer une transmissibilité équivalente et évite ainsi un petit volume qui pourrait engendrer des problèmes numériques lors de la simulation d'écoulement (figure 3.7).

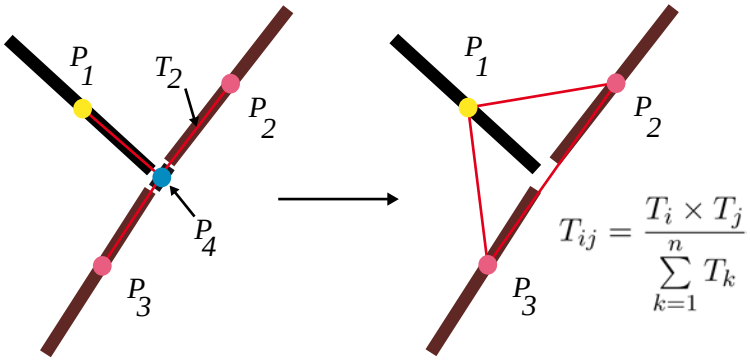

FIGURE 3.7 – Calcul de la transmissibilité au niveau de l'intersection entre deux failles. La méthode utilise la transformation étoile-polygone [Karimi-Fard et al., 2004] pour calculer les transmissibilités entre les éléments P_1, P_2 et P_3, ce qui permet de ne pas représenter le petit volume P_4.

Les intersections entre failles sont repérées par le cycle de liens formé autour des trois segments (quatre si F_A recoupe F_B) formant l'intersection (section 3.1.1.6).

Un exemple en 2D d'un graphe de connectivité est présenté en figure 3.8.

La liste de connectivité ainsi que les propriétés des nœuds (volume, perméabilité, porosité) peuvent ensuite être fournis en entrée à un simulateur d'écoulement supportant ce type de structure. Dans notre travail, l'ensemble des simulations d'écoulement est effectué sur le simulateur *GPRS* (pour *GeneralPurposeReservoirSimulator*), développé à l'université de Stanford.

3.1.4 Insertion de puits dans le graphe de connectivité

Un puits est représenté par les nœuds du graphe de connectivité dont les primitives associées sont intersectées par la trajectoire du puits. L'ajout d'un puits A se fait en trois étapes :

- les tétraèdres intersectés par A sont ajoutés dans une liste S_A (section 3.1.4.1) ;
- pour chaque tétraèdre de S_A, l'algorithme sélectionne la liste L_P des primitives intersectées par A ;

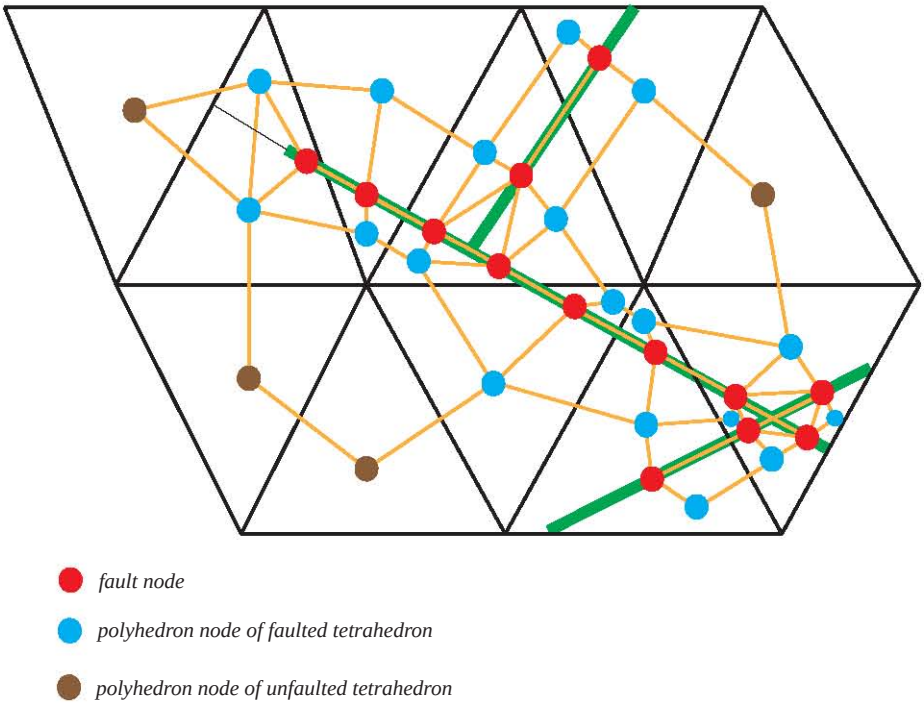

FIGURE 3.8 – Vue 2D d'un graphe de connectivité.

– un coefficient d'échange est calculé pour chaque primitive de L_P (section 3.1.4.2).

3.1.4.1 Recherche des primitives géométriques intersectées par un puits

L'algorithme 6 de recherche des tétraèdres coupés par un puits A commence par chercher un premier tétraèdre intersecté *seed*. L'algorithme se propage ensuite à partir de *seed* en testant les intersections de chacune de ses quatre faces et en ajoutant leur voisin en cas d'intersection. Tous les tétraèdres intersectés sont ajoutés à S_A. Seules les primitives des tétraèdres de S_A qui sont intersectées par A sont utilisées pour le calcul du terme d'échange.

3.1.4.2 Calcul de la transmissibilité d'un puits

L'échange de matière entre un puits A et un bloc de matrice c est caractérisé par la transmissibilité WI_A calculée grâce à la formule de Peaceman [Peaceman, 1983] :

Entrées :
A= puits
$\mathcal{T}$ = maillage tétraédrique

Définitions :
S_A = liste de tétraèdres intersectés par A
S = liste de tétraèdres
$S_{primitive}$ = liste de primitives géométriques

$S \leftarrow$ un tétraèdre de $\mathcal{T}$ intersecté par A
tant que $S \neq \emptyset$ **faire**
  $T \leftarrow$ dernier tétraèdre de S
  $S_A \leftarrow T$
  pour chaque *face* f_i *de* T **faire**
    si f_i *est intersectée par* A **alors**
      T_i = tétraèdre connecté à f_i ($T_i \neq T$)
      si $T_i \notin S$ **alors**
        $S \leftarrow T_i$
      finsi
    finsi
  fin
fin
pour chaque *tétraèdre* T *de* S_A **faire**
  pour chaque *primitive* P *de* T **faire**
    si P *est intersectée par* A **alors**
      $S_{primitive} \leftarrow P$
    finsi
  fin
fin
Sorties :
$S_{primitive}$

Algorithme 6: Algorithme de recherche des primitives géométriques intersectées par un puits.

$$
\begin{aligned}
&WI_A = \frac{2\pi\sqrt{k_x k_y}h}{\ln(\frac{r_0}{r_w}) + s} \\
&avec : \\
&\quad k_x, k_y \text{ la perméabilité en } x, y \\
&\quad h \text{ la hauteur de } c \\
&\quad r_0 = 0.28\frac{\sqrt{\sqrt{k_y/k_x}\Delta x^2 + \sqrt{k_x/k_y}\Delta y^2}}{\sqrt[4]{k_y/k_x} + \sqrt[4]{k_x/k_y}} \\
&\quad r_w \text{ le diamètre de } A \\
&\quad s \text{ représentant l'endommagement, la stimulation} \\
&\quad (e.g. \text{ fracturation, acidification}) \text{ et la complétion de } A
\end{aligned}
\tag{3.5}
$$

Dans notre cas, la forme des primitives géométriques étant quelconque, un cube d'aire équivalente est utilisé pour calculer la transmissibilité WI d'un puits.

3.1.5 Comparaison avec d'autres supports de discrétisation

Une discrétisation exacte du flux entre deux éléments requiert que la droite joignant les nœuds des deux éléments voisins soit perpendiculaire à leur face commune. Les supports de discrétisation remplissant ce critère sont qualifiés de "PErpendicular BIssector" (PEBI). Or un maillage tétraédrique ne remplit pas cette condition. La discrétisation proposée utilise le calcul de transmissibilité proposé par Karimi-Fard et al. [2004] pour prendre en compte l'angle formé entre la normale de la face et la droite reliant deux nœuds voisins.

Nous comparons ici la discrétisation d'un maillage tétraédrique non faillé (figure 3.9*d.*) avec deux types de maillage de type PEBI : une grille régulière (figure 3.9*c.*) et un maillage de Voronoï barycentrique [Du et al., 1999] (figure 3.9*b.*). La comparaison est effectuée avec des maillages de même résolution et avec un champ de perméabilité uniforme et isotrope. La discrétisation d'un réseau de failles avec une perméabilité équivalente à celle de la matrice est également comparée aux 3 autres cas. La perméabilité est isotrope et égale à 100mD pour l'ensemble du domaine. La porosité est également uniforme et égale à 0.2. La saturation initiale en eau est de 0.25. Le domaine comporte un puits injecteur et un puits producteur (figure 3.9*a.*).

La figure 3.10 présente l'évolution de la proportion d'eau produite au puits producteur pour les quatre discrétisations. Les courbes ne sont pas strictement identiques, pourtant l'ensemble est homogène. Les courbes diffèrent au moment de l'arrivée de l'eau autour de 3000 jours, puis seule la grille régulière se différencie des autres courbes pour le reste de la simulation. Après 5000 jours, la saturation en eau du puits producteur ne diffère que de 0,1% entre le maillage tétraédrique seul et le réseau de faille discrétisé. Les différences sont de 2,2% et 0,3% entre le maillage tétraédrique et respectivement la grille régulière et le maillage de Voronoï barycentrique. Ces différences paraissent négligeables au vu des incertitudes cumulées au cours du processus de modélisation. Les résultats se rapprochent de ceux obtenus par Karimi-Fard et al. [2004].

L'effet de la résolution du maillage est présenté en figure 3.11. Les courbes restent très proches malgré une résolution divisée par deux pour le maillage tétraédrique.

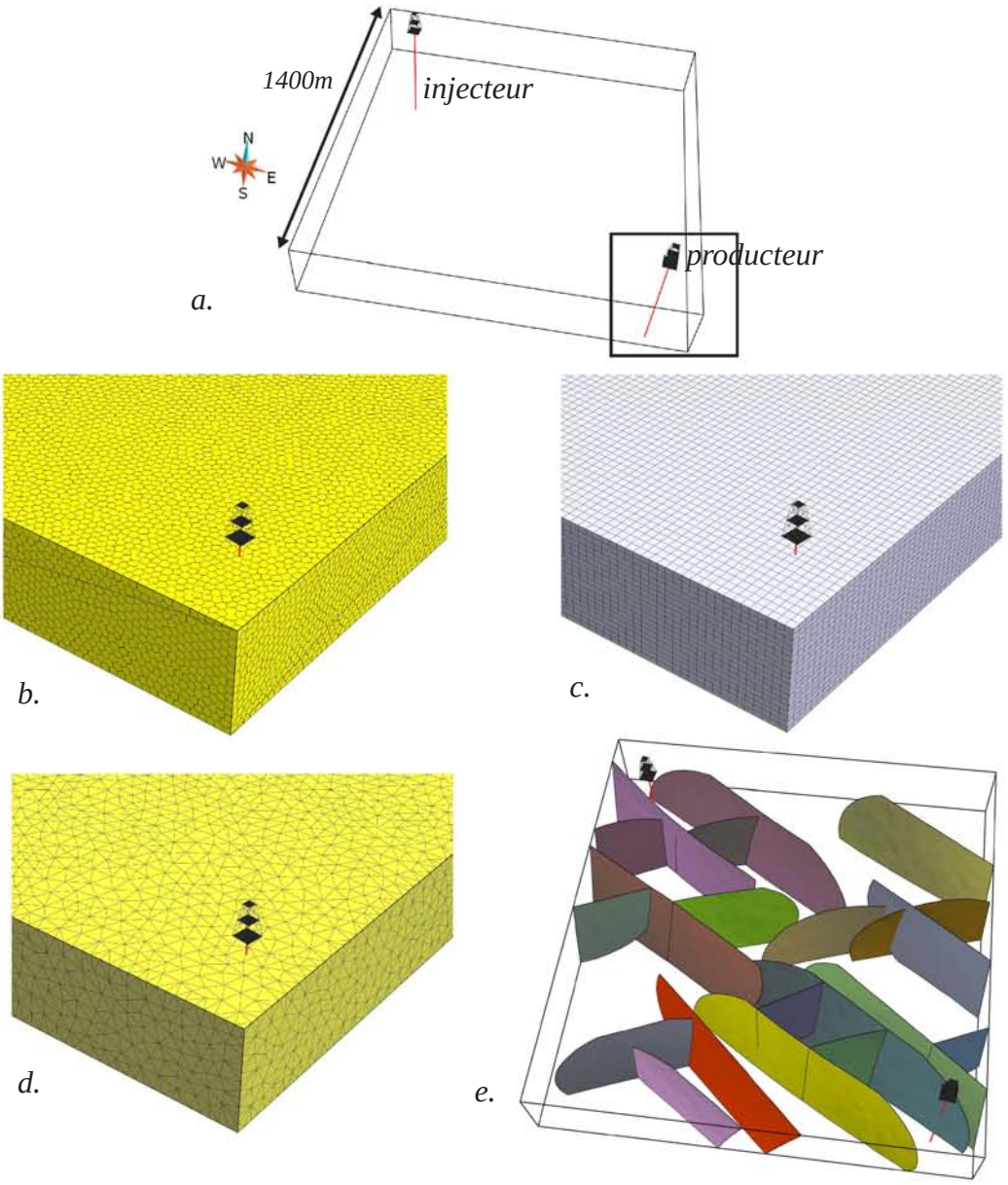

FIGURE 3.9 – Supports de discrétisation utilisés pour la comparaison des simulations d'écoulements. *a.* Le domaine fait environ 1400x1400x250m. Un puits producteur et injecteur sont diamétralement opposés. *b.* Maillage de Voronoï barycentrique. *c.* Grille régulière. *d.* Maillage tétraédrique. *e.* Réseau de failles.

3.1.6 Perspectives

La discrétisation proposée représente explicitement les failles par des nœuds du graphe de connectivité, ce qui permet de prendre en compte leur géométrie exacte. Leur participation à l'écoulement n'est pas simplement prise en compte à travers des coefficients d'échange, leur géométrie n'est pas simplifiée par une rasterisation, ce qui permet, a priori, de pouvoir étudier l'impact de la connectivité d'un réseau de failles sur la réponse du sous-sol aux écoulements de fluides.

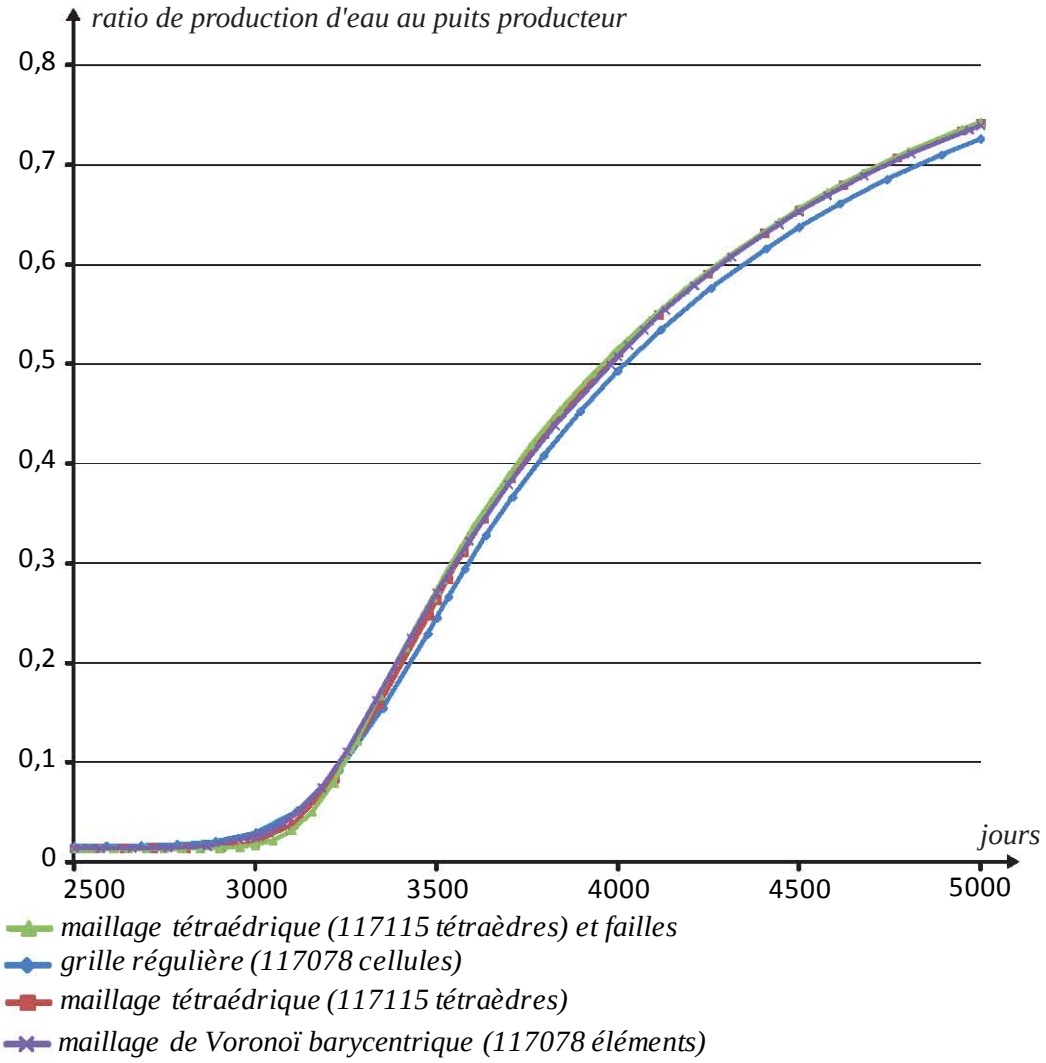

FIGURE 3.10 – Comparaison des simulations d'écoulement entre différents supports de discrétisation. Les courbes montrent l'évolution au cours du temps de la proportion d'eau produite au puits producteur. Les courbes ne sont pas strictement identiques mais l'ensemble est proche.

Les failles impactent également l'écoulement des fluides par le déplacement des couches géologiques qu'elles induisent. Les horizons stratigraphiques doivent être représentés afin de prendre en compte ces aspects. Or le maillage tétraédrique n'étant pas conforme aux failles, il n'est donc pas possible de prendre en compte le déplacement des failles puisque cela requiert d'interpoler un champ stratigraphique discontinu de part et d'autre des failles. En revanche, avec un maillage conforme aux failles, il serait possible de discrétiser simplement les failles par les faces des tétraèdres représentant la discontinuité. Les principaux horizons stratigraphiques pourraient alors être discrétisés dans une approche similaire à celle proposée ici, ce qui permettrait d'assigner des propriétés pétrophysiques différentes par unité stratigraphique (avec éventuellement un modèle pétrophysique associé).

Cette méthodologie reste néanmoins limitée par le maillage tétraédrique, notamment

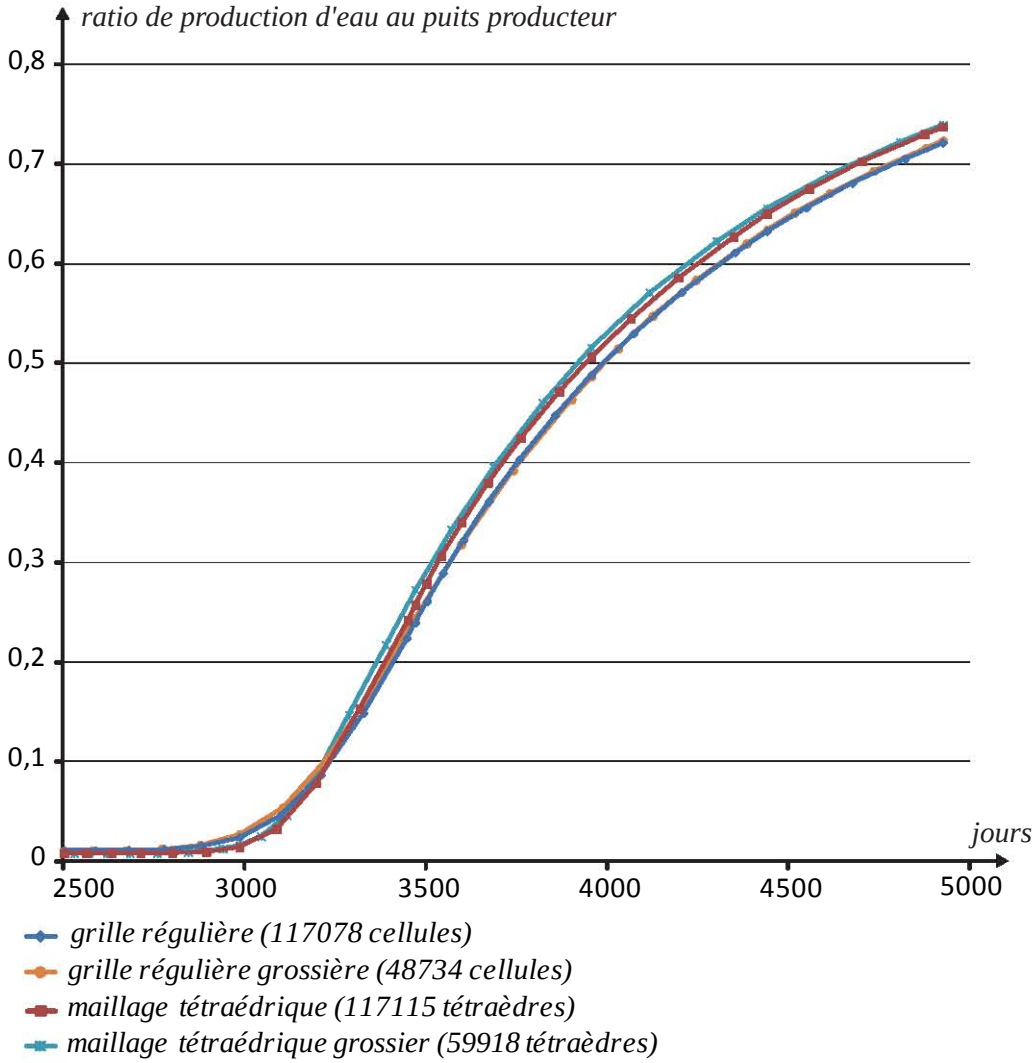

FIGURE 3.11 – Effet de la résolution lors d'une simulation d'écoulement. Malgré une résolution divisée par deux, les résultats pour les deux maillages tétraédriques sont très proches. Les courbes des grilles régulières sont presque identiques.

pour prendre en compte l'anisotropie des hétérogénéités pétrophysiques. Un support de discrétisation s'adaptant aussi bien aux hétérogénéités structurales que pétrophysiques, tout en garantissant une discrétisation aussi exacte que possible (*e.g.* grille PEBI orthogonale aux axes du tenseur de perméabilité) est nécessaire. Dans ce cadre, Merland et al. [2011] proposent une méthode prometteuse basée sur le calcul de maillages de Voronoï barycentriques (PEBI par nature) grâce à une méthode d'optimisation globale. Cette méthode permet de respecter à la fois les discontinuités structurales et l'anisotropie des propriétés pétrophysiques, qui sont prises en compte par l'alignement des cellules sur les lignes de courant calculées à fine échelle. Une mise à l'échelle des propriétés pétrophysiques à partir du modèle pétrophysique fin est toutefois nécessaire et reste délicate à mettre en oeuvre afin de représenter le comportement dynamique des hétérogénéités à fine échelle.

3.2 Méthode de modélisation stochastique inverse de la géométrie et de la connectivité de failles grâce à des données d'écoulement de fluides

La section précédente a montré comment est modélisé le problème direct $g(\mathbf{m})$, à savoir l'écoulement de fluides. Cette partie détaille la paramétrisation du système (un réseau de failles) (section 3.2.3) et une stratégie d'inversion (section 3.2.4). Ces trois éléments permettent une approche inverse pour la caractérisation d'un réseau de failles à partir de données d'écoulement de fluides. Des exemples d'application de la méthodologie inverse sont présentés en section 3.2.5.

La suite de cette partie est constituée d'un article paru dans *Mathematical Geosciences*.

Method for stochastic inverse modeling of fault geometry and connectivity using flow data

Nicolas Cherpeau, Guillaume Caumon, Jef Caers and Bruno Lévy

Article publié dans *Mathematical Geosciences*

This paper focuses on fault-related uncertainties in the subsurface, which can significantly affect the numerical simulation of physical processes. Our goal is to use dynamic data and process-based simulation to update structural uncertainty in a Bayesian inverse approach. We propose a stochastic fault model where the number and features of faults are made variable. In particular, this model samples uncertainties about connectivity between the faults. The stochastic 3D fault model is integrated within a stochastic inversion scheme in order to reduce uncertainties about fault characteristics and fault zone layout, by minimizing the mismatch between observed and simulated data.

The stochastic fault model uses a priori information such as fault orientation, location, size and sinuosity, to sample both geometrical and topological uncertainties with realistic fault descriptions. Each fault object is parameterized by the random vector used to simulate fault features. Then, during inversion, the random vector of the current model is stochastically perturbed, producing a new parameter vector used as input by the stochastic fault model to produce a new model. Even if the topology varies from one model to another, the algorithm produces correlated models so that their flow responses evolve quite smoothly.

The methodology is applicable in general and illustrated on a synthetic two-phase flow example. A first set of models is generated to sample the prior uncertainty space. Then, models minimizing reference water-saturation data misfit are used as seeds to generate continuous Markov chains of models with discrete states. Posterior models reduce uncertainties about fault position, while the topology varies from one model to another. A second example highlights the interest of the parameterization when interpreted data are available, by perturbing geological scenarios and falsifying those that do not match two-phase flow observations.

KEYWORDS : structural modeling ; uncertainty ; topology ; inverse modeling

3.2.1 Introduction

In subsurface studies, uncertainties about faults and fault network primarily stem from the lack or poor quality of observation data. Due to the inherent uncertainties and possible human bias during interpretation of geological data [Bond et al., 2007], model predictions seldom match physical observations such as reservoir pressure, seismograph measurements or gravity anomalies. One may want to use such data to infer some characteristics of the model, which is known as the inverse problem.

In this paper, we propose a general method to reduce structural uncertainty about fault networks. The method generates models with varying number of faults, fault geometry and fault connectivity, *i.e.* how faults truncate each others. The method accounts for large to small scale faults (km to decameter scale), for which uncertainties may be large due to the lack of data (incomplete sampling or sensing), imaging ambiguities (*e.g.* in potential field and seismic methods), limited resolution and positioning errors (borehole trajectory, time-to-depth conversion, etc.). Therefore, in the context of hydrocarbon reservoir management, our method is applicable at several stages :

- In exploration of "green fields", in which a limited amount of data are generally available (a set of seismic lines, a few borehole data). In this case, we focus on large-scale, kilometric faults or larger.
- In "brown field" management, in which 3D seismic data are generally available but only provide an approximate view on structures because of resolution, acquisition

or processing issues. In this case, the method is applicable to aid the interpretation of large to mid-scale faults in poorly imaged areas or to simulate sub-seismic faults following geological priors.

In such uncertain situations, the method mainly relies on prior information about faults, which may come from the interpretation of available data, structural geological concepts, regional context or reservoir analogs.

We apply this method to the inversion of reservoir or hydrogeologic flow data. Indeed, faults impact fluid flow in different ways :

1. Individually, faults have a direct impact on the permeability field due to fracturing, alteration, cementation or shale smearing nearby the fault. Uncertainties about fault rock hydraulic properties in flow models are generally high and may have a significant impact on fluid flow [Manzocchi et al., 1999].
2. At the reservoir scale, fault zone complexity and spatial layout directly impact fluid flow by creating connected flow conduits or flow barriers isolating fault blocks, depending on the nature of the fault rock. In this paper, we focus on uncertainties about fault layout and geometry including fault connectivity uncertainties.
3. The displacement of geological layers along fault surfaces perturbs both the horizontal and vertical connectivity of transmissive layers. 3D structural models that represent geological layers and structures such as faults are of prime importance for accurately representing fault displacement, trends, anisotropy of permeability and porosity heterogeneities. The rock properties greatly depend on the nature of the rocks, thus are often estimated in a coordinate system representing the depositional space [Mallet, 2004, Gringarten et al., 2008].

The proposed methodology considers geological structures as a reduced set of parameters, thus opens the ways to the management of structural uncertainty in classical stochastic inversion methods like MCMC (figure 3.12). A simple non-conditional case study (section 3.2.5.1) shows the applicability of the proposed parameterization to reduce fault location uncertainty using a rejection sampler. Indeed, the goal is to sample the posterior distribution and obtain very different realistic models, hence rejection sampling is more appropriate than optimization techniques.

In sparse data situations, several fault arrays may be interpreted, all consistent with available data. In this context, we propose to perturb geological scenarios while maintaining consistency with interpreted data, in order to falsify scenarios that do not match two-phase flow observations (section 3.2.5.2).

3.2.2 Related Work and Challenges of Structural Parameters Inversion

The inverse problem is known as history matching or model calibration in the case observed data concern the reservoir/aquifer dynamic flow response. The idea behind inverse modeling is obvious : if a model does not reproduce the observations, it is unlikely that its predictions are correct. Moreover, if a model matches the observations, it has a higher chance to make relevant predictions. However, only stochastic approaches provide a relevant framework for risk assessment, by considering all possible scenarios and not just one deterministic calibrated model.

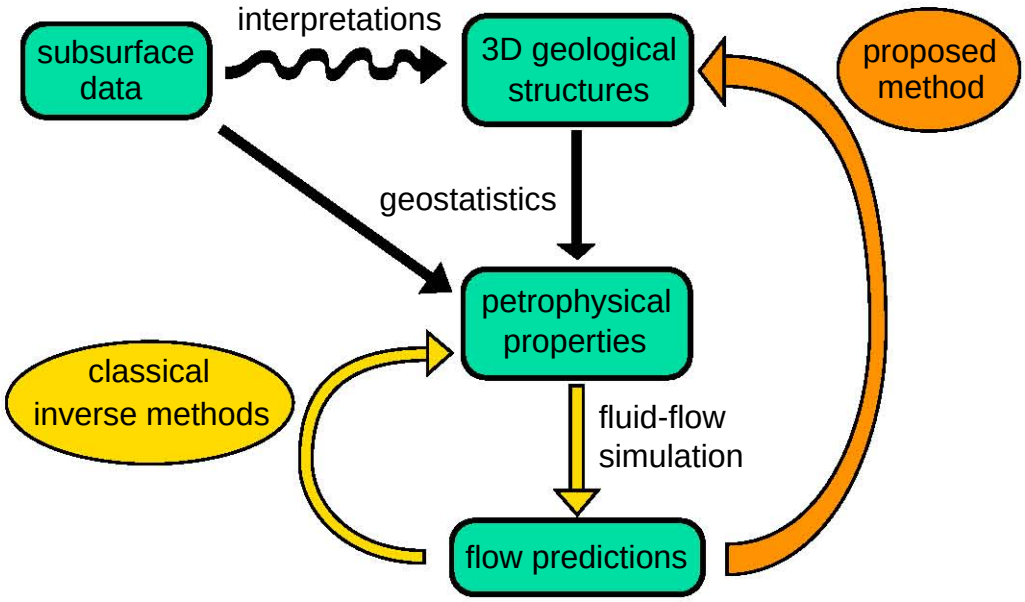

FIGURE 3.12 – Typical modeling and inversion workflow. 3D structural modeling is based on the interpretation of subsurface data, which may entail uncertainties due to model approximations, sparse/low resolution data and possible interpretive bias. Once geological structures have been modeled, petrophysical heterogeneities are computed using geostatistical methods conditionning to subsurface data. Fluid-flow simulation then enables to compute flow predictions that can be compared to observations. The proposed parameterization opens the way to inversion strategies by updating not only petrophysical properties, as proposed by classical inverse methods, but also 3D geological structures.

Most history matching techniques focus on permeability and porosity heterogeneities, either in the field of hydrogeology [Carrera et al., 2005, de Marsily et al., 2005, Zimmerman et al., 1998], or for the hydrocarbon industry [Oliver and Chen, 2011]. Other techniques focus on connectivity features [Western et al., 2001] or sedimentary objects such as channels [Ronayne et al., 2008] that may have a great impact on fluid flows by creating high-conductivity preferred flow paths. However, uncertainties about geological structures are neglected in most cases. Jenni et al. [2004] propose to handle uncertainties about fractures using gradual deformations to update field scale planar fractures while honoring a fracture density map. In this work, we propose an approach similar to Jenni et al. [2004], but use a 3D flexible and more realistic representation of fault objects in terms of geometry and topology. Suzuki et al. [2008] focus on larger structural uncertainties. They select models that match reflection seismic data in a large set of prior structural models using a stochastic search method and the Hausdorff distance as metric to avoid running the forward problem for each ensemble member. The prior set of models is generated by perturbing the geometry of different seismic interpretations representing different geological scenarios. As Suzuki et al. [2008], we also focus on large structural uncertainties but our parameterization is flexible enough to automatically generate models with varying number of faults and fault

connections.

Other uncertainties about reservoir boundaries do not only affect reservoir volume but also control the spatial layout of geological layers and thus possibly affect fluid flows. In this context, Seiler et al. [2010] use the Ensemble Kalman Filter (EnKF) to update top and bottom stratigraphic horizons of a reservoir. They use a corner-point grid to represent the most likely initial state and deform it to match simulated top and bottom horizons. As the grid is deformed, their algorithm keeps the same number of cells over realizations and thus a state vector with same dimensions as required by the EnKF.

However, none of these methods is flexible enough to sample posterior, *i.e.* after data inversion, realistic 3D models of geological structures. Geological surfaces (faults, fractures or stratigraphic horizons) are either planar or their geometry is limited by the gridding method, *e.g.* pillar-based grids. Our fault representation does not suffer such limitations, for implicit modeling [Calcagno et al., 2008, Frank et al., 2007, Guillen et al., 2008] is used to represent faults as isopotentials of scalar fields, namely implicit surfaces.

Inverting structural parameters using dynamic data raises both conceptual and technical issues :

1. One needs a robust method to sample structural uncertainties. While it is a common practice to generate several geostatistical realizations of facies or rock properties, few methods propose to sample uncertainties about geological structures, either geometrical [Abrahamsen, 1992, Caumon et al., 2007, Lecour et al., 2001, Mallet and Tertois, 2010, Thore et al., 2002, Wellmann et al., 2010] or topological [Cherpeau et al., 2010a, Holden et al., 2003]. The algorithm used in this work is adapted from Cherpeau et al. [2010a] in order to meet the needs of inverse modeling. It generates stochastic fault networks from prior information about fault orientation, size, sinuosity and location.

2. It is not an easy task to describe a structural model as a set of relevant parameters. In this paper, we focus on fault network parameters, considering both individual fault parameters such as orientation, size and location and also the number of faults in the model. Connections between faults are not explicitly parameterized but result from the fault network generation process (section 3.2.3).

3. From a technical standpoint, modeling realistic geological structures requires a flexible 3D representation while modeling physical phenomena requires meshes for discretization of partial differential equations describing the physical phenomenon of interest [Caumon et al., 2004a]. Consequently, using dynamic data to constrain geological structures calls for either back and forth and tedious conversions between the two representations, or a simplified geological representation that only enables to update the reservoir model, as proposed in the seminal work of Holden et al. [2003]. Our implicit fault surfaces can be discretized into a control volumes connectivity graph that can be fed into a flow simulator [Cherpeau et al., 2011]. Hence, the geological representation does not suffer from any simplifications while the flow response can be computed from the connectivity graph.

3.2.3 Parameterization of Fault Networks

3.2.3.1 Overview of the Stochastic Fault Model

The fault generation process is a modified version of the method proposed by Cherpeau et al. [2010a,b] to sample fault-related uncertainties in sparse data situations. The algorithm sequentially simulates faults in the domain of interest. Faults are grouped into fault families representing different faulting events in terms of timing and regional stress, resulting in different input parameters about fault orientation, size, localization and sinuosity. Moreover, fault families are simulated in a given chronological order reflecting the timing of faulting events. Consequently, new simulated faults are branching faults in case of intersections with previously simulated faults. When the timing is uncertain, fault families are considered cogenetic so that truncation is randomly simulated.

Each fault object is simulated through a number of substeps and the global simulation process is as follows :

For each fault family S, from the oldest to the youngest :
For each fault F in S :
1. Draw dip and strike from prior distributions.
2. Draw fault center coordinates from the fault center probability distribution function f built from input parameters (section 3.2.3.1).
3. Generate a fault geometry from orientation parameters, fault center and input sinuosity parameters.
4. Draw an extension along the strike direction and along the dip direction (section 3.2.3.1).
5. Look for intersections with previously simulated faults.

The simulation can also be conditional to interpreted fault traces. In this case, extra simulation steps assign data to faults and randomly cluster data points using geometrical concepts [Cherpeau et al., 2010b]. The algorithm assigns data points to faults in different orders and generates different data combinations, hence produces different models that all honor hard data. The following sections describe some substeps of the fault object simulation.

Fault center coordinates

The fault nucleation step is a key step during the simulation process for it determines the neighborhood of a fault thus possible intersections with neighbor faults. In some geological contexts, major faults tend to repulse each other, meaning that new faults occur far from existing ones, whereas secondary faults tend to occur around major faults. If one wants to mimic such behaviors, the fault center probability distribution function f may be non-uniform in the domain of interest and may depend on neighbor faults. Consequently, the fault center is drawn from $f(u, v, w)$ defined in a Cartesian grid, by successive drawing of $u \in [1; n_u]$, $v \in [1; n_v]$ and $w \in [1; n_w]$ coordinates, with n_u, n_v, n_w the number of voxels along the three orthogonal axis a_u, a_v and a_w respectively. The probability $p(U \in cell_a)$

that the u coordinate is located in an elementary cell, $cell_a$ centred on a is :

$$\begin{aligned} p(U \in cell_a) &= \int_{u \in cell_a} \int_v \int_w f(u,v,w) du dv dw \\ &= \frac{1}{V} \sum_{j=1}^{n_v} \sum_{k=1}^{n_w} f(a,j,k) v_{ajk} \end{aligned} \tag{3.6}$$

with :

V the total volume of interest

v_{ajk} the volume of the elementary cell at (a,j,k)

Once u coordinate has been drawn, $p(V \in cell_b | U \in cell_a)$ with $b \in [1; n_v]$ is computed according to :

$$p(V \in cell_b | U \in cell_a) = \int_{v \in cell_b} \int_w k f(a,v,w) dv dw = \frac{1}{V} \sum_{i=1}^{n_w} k f(a,b,i) v_{abi}$$

with k a normalization constant, (3.7)

$$k = \frac{1}{\int_v \int_w f(a,v,w) dv dw} = \frac{1}{\sum_{j=1}^{n_v} \sum_{k=1}^{n_w} f(a,j,k) v_{ajk}}$$

Then, w can be drawn in the same manner, *i.e.* by computing a discrete pdf along the axis a_w given $u = a$ and $v = b$ coordinates. Continuous values of these discrete probability density functions are obtained by linear interpolation between discrete values, which can then be back-transformed into x, y, z coordinates.

Fault extension

In order to model laterally terminating and synsedimentary faults, the algorithm restricts the existence of a fault by a 3D ellipsoid $\mathcal{E}(x, v_1, v_2, v_3)$ with x the fault center coordinates, v_1, v_2, v_3 three orthogonal vectors. The vectors v_1 and v_2 are respectively computed from extension values randomly drawn from length distribution along the strike direction and height distribution along the dip direction. The third vector v_3 is fully defined by $v_3 = v_1 \times v_2$. This elliptical fault shape is not a restriction of our algorithm and is consistent with the literature. Indeed, a fault's tipline is considered to be circular to elliptical for an isolated fault [Barnett et al., 1987, Walsh et al., 2003] but may evolve to elongated ellipses if the vertical fault height is restricted by the thickness of a geological layer [Benedicto et al., 2003], which can be approximated by choosing $||v_1|| >> ||v_2||$.

Fault number parameter

In case of large fault-related uncertainties, not only is the fault connectivity uncertain but also the number of faults occurring in the studied area. The algorithm associates, for each fault family, a theoretical number of faults n with a probability of fault existence $p_e \in [0, 1]$, so that the output number of faults m varies in $[0, n]$ according to the following binomial probability density function $Q(m)$:

$$Q(m) = \binom{n}{m} p_e^m \times (1 - p_e)^{n-m} \tag{3.8}$$

The following section presents the parameterization of our stochastic fault model to fit the needs of an inversion scheme.

3.2.3.2 Model Parameterization

The stochastic fault model sequentially simulates faults in the domain of interest. Each step during the simulation of a fault object stochastically set a fault parameter. Finally, each fault can be described by six parameters : dip, strike, length, height, sinuosity and center. During the fault simulation, one or more random numbers are used by the stochastic steps to generate each fault parameter, *i.e.* only one random number is used per dip, strike, height, length parameters whereas three random numbers are required to draw the fault center position. Consequently, each fault can be parameterized by a vector with eight components $F(r_1, ..., r_8)$. A model m with n faults is fully parameterized by a random vector with $8n$ components $m = \{F^1, ..., F^n\}$ or $m = \{(r_1^1, ..., r_8^1); ...; (r_1^n, ..., r_8^n)\}$.

3.2.4 Inversion Methodology

The parameterization described in section 3.2.3.2 opens ways to inversion schemes in which geological structures are part of the inversion loop, as shown in figure 3.12. In this section, we present how the parameterization can be used to sample fault networks using two-phase flow response as a misfit to minimize. The inversion scheme relies on a version of the Metropolis algorithm [Metropolis et al., 1953] proposed by Mosegaard and Tarantola [1995]. The Metropolis algorithm was originally designed to sample Gibbs-Boltzmann distributions in chemical science and has been since successfully applied in several fields such as mathematic, physic and biology. It is a *Markov Chain Monte Carlo* method, *Markov Chain* implicitly meaning first-order Markov chain, *i.e.* current state only depends on previous state, and *Monte Carlo* meaning that it is a random sampling method. The Metropolis algorithm is an *importance sampling* algorithm designed to solve highly nonlinear inverse problems by sampling the posterior probability density function defined by the product of the likelihood function and the prior probability density function. A general description of Monte Carlo methods as solutions to inverse problems can be found in Mosegaard and Sambridge [2002].

3.2.4.1 Random Walk

Inverse methods search for solutions that minimize the misfit between predictions and observations and thus need random walks or perturbation processes to explore the parameter space. In this space, the neighborhood of point i is defined as all the points j that can be reached from i in a single move. We define the neighborhood $\mathcal{N}_i$ of a model $m_i(r_1^i, ..., r_{8n}^i)$ in 8n-D search space as the set of all models $m_j(r_1^j, ..., r_{8n}^j)$ that can be obtained through the following steps :
For each random number r_k^i :

1. Compute $b_{low} = \max(0, (r_k^i(1-p)))$ with p a perturbation amplitude given in input ;
2. compute $b_{up} = \min(1, (r_k^i(1+p)))$;
3. define a uniform distribution D_k $\mathcal{U}[b_{low}, b_{up}]$;
4. randomly draw r_k^j from D_k.

Consequently, the neighborhood $\mathcal{N}_i$ corresponds to an infinite set of models. However, any point B in parameter space can be reached from any point A provided a sufficient number of steps, *i.e.* the random walk covers the whole parameter space. Moreover, the proposed transition process from model i to model j is symmetric, *i.e.* the probability that a jump from model m_j to model m_i takes place is equal to the probability that a jump from model m_i to model m_j takes place : $p(m_i|m_j)p(m_j) = p(m_j|m_i)p(m_i)$.

For a point i in parameter space, the perturbation amplitude p defines the size of the neighborhood around i, thus controls the speed of exploration of the parameter space. Moreover, p reflects the trade-off between *exploitation* and *exploration* of the parameter space. Indeed, if p is small, it means that the algorithm searches for models in a close neighborhood and thus tries to *exploit* current position in parameter space. If p is large, the neighborhood is large and thus the algorithms tends to *explore* the parameter space. However, high-dimensional spaces tend to be empty spaces and the probability of moving from a high-misfit point i to a low-misfit point j may be high if the move, *i.e.* the perturbation amplitude p, is too large.

Parameter variation speed

Another way of generating a new model could be to directly perturb the fault parameters of the current model, *i.e.* to bypass the statistical distributions. When perturbing random numbers, all numbers evolve at the same "speed" in the range $[0, 1]$ depending on the input perturbation amplitude p. Then, each fault parameter is drawn from the corresponding distribution, meaning that fault parameters evolve at a speed proportional to the spread of their distribution. Indeed, the narrower the distribution, the slower the parameter variation speed. Consequently, the method accounts for the prior information and uncertainty. This could hardly be achieved by perturbing directly the fault parameters.

Structural parameters evolution

The parameterization and perturbation process described above ensures that fault parameters evolve smoothly. However, the fault sinuosity is described by a single random number used as a seed for generating a correlated random noise using a geostatistical algorithm, *e.g.* Sequential Gaussian Simulation. The method is similar to Caumon et al. [2007], the generated random field is added to the scalar field representing the fault surface. The input variogram ranges needed for the geostatistical method is computed from sinuosity parameters. In contrast with gradual deformation methods, the fault sinuosity does not evolve smoothly since perturbing the seed entails the generation of a new perturbation field.

Moreover, the fault connectivity is not parameterized as it results from the stochastic fault simulation algorithm. However, if two faults are no longer connected, the fault positions evolve smoothly so that the flow has a small volume of matrix to pass trough to connect the two disconnected faults. Consequently, even if some fault connections are lost from one model to another, the corresponding flow responses are still correlated.

Fault number evolution

Adding or removing a fault from one model to another may entail a large perturbation of the model predictions, *e.g.* if the new or removed fault is localized in a flow-influencing zone, the flow-response change may be too large to consider correlated models. In order to get smooth transitions, the existence of a fault for a given fault family is coupled with its length distribution so that the probability of a zero-length fault is $1 - p_e$ with p_e the probability of existence. Consequently, the length distribution is modified as illustrated in figure 3.13.

The ϵ parameter should be chosen carefully in function of the perturbation amplitude p, for it controls the fault growth/contraction rate. Indeed, if p is too large as compared to ϵ, the fault may get its normal size or disappear during one iteration.

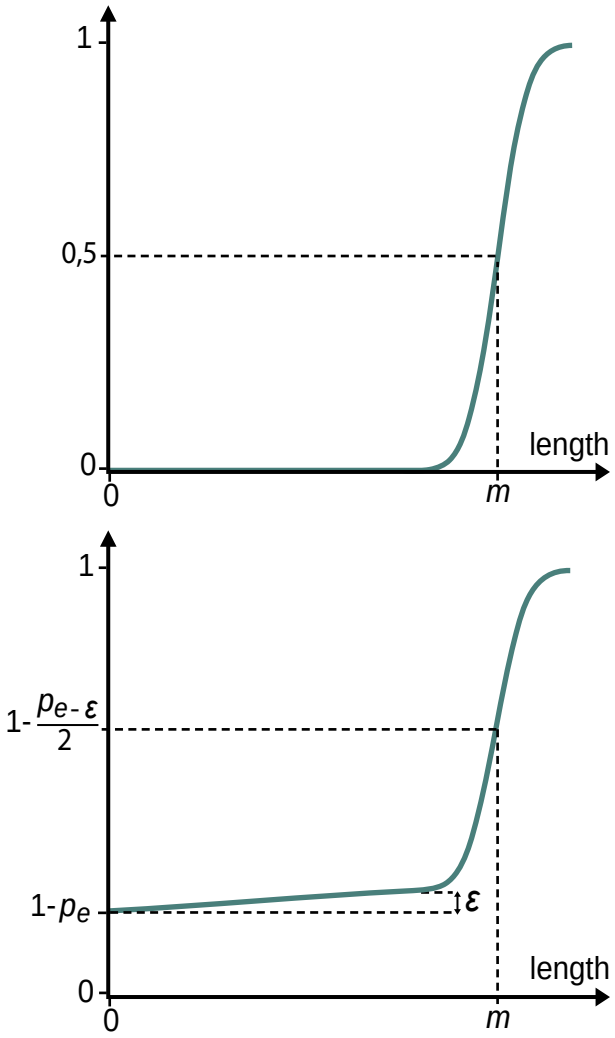

FIGURE 3.13 – Fault length distribution coupled with existence probability parameter. At the top, an initial gaussian CDF representing the fault length with mean m. At the bottom, the distribution has been modified to account for the existence probability p_e. Parameter ϵ should be carefully set in function of the perturbation amplitude p.

3.2.4.2 Acceptance Rule

The acceptance rule modifies the random walk (presented in section 3.2.4.1) by introducing a probability p_{ij} to move from point i to point j, so that the posterior probability density function is sampled. We use the Metropolis rule as the acceptance rule. For each model m, we run the forward problem, *i.e.* a flow simulation, and compute some water-cuts $w_t(m)$ at production wells. The misfit function $S(m)$ is the root mean squared error with

the observations w_t^{obs} :

$$S(m) = \sqrt{\frac{\sum_{t=1}^{l}(w_t(m) - w_t^{obs})^2}{l}} \tag{3.9}$$

The likelihood function is of the form $L(m) = \exp(-\frac{S(m)}{s^2})$, with s^2 the total noise variance. Then, the probability p_{ij} of moving from point i to point j is :

$$\begin{aligned} p_{ij} &= 1 && \text{if } L(m_j) \geq L(m_i) \\ &= \frac{L(m_j)}{L(m_i)} && \text{otherwise} \\ &= \exp(-\frac{S(m_j) - S(m_i)}{s^2}) \end{aligned} \tag{3.10}$$

Note that the misfit function could be any function of interest for the study at hand. We here use it for the two examples presented in section 3.2.5. Moreover, the methodology could be used in other scientific areas. It could for instance be used in geophysical inversions to better constrain subsurface structure. In this case, a wave propagation method would replace the flow simulation during the forward modeling step in the inverse methodology.

3.2.5 Applications

3.2.5.1 Updating Prior Structural Uncertainty

This case study aims at understanding the behavior of the method for a simple case and assessing the amount of structural information introduced by a simple history matching problem. Fault occurrence and location are not constrained by any interpreted data.

Input Data

The studied area of size 1400x1400x200m has one injector and one producer wells on both sides of a fault zone composed of two fault families :

1. The first fault family corresponds to major faults aligned along the main flow path :
   - strike : uniform, $[110 - 130]^\circ$
   - dip : Gaussian, $mean = 70^\circ$, $\sigma = 3^\circ$
   - length : Gaussian, $mean = 800m$, $\sigma = 50m$
   - height : Gaussian, $mean = 350m$, $\sigma = 5m$
   - number of faults : 3
   - probability of existence $p_e = 0.8$
   - sinuosity wavelength along strike direction : 500m
   - sinuosity wavelength along dip direction : 500m
2. The second fault family corresponds to branching faults :
   - strike : uniform, $[50 - 70]^\circ$
   - dip : Gaussian, $mean = 70^\circ$, $\sigma = 3^\circ$
   - length : Gaussian, $mean = 500m$, $\sigma = 30m$
   - height : Gaussian, $mean = 300m$, $\sigma = 5m$
   - number of faults : 2
   - probability of existence $p_e = 0.7$
   - sinuosity wavelength along strike direction : 350m

– sinuosity wavelength along dip direction : 300m

The total number of faults varies in [0,5] according to equation 3.8. The matrix has a constant porosity of 0.2 and a constant permeability of 100mD. Faults are considered flow barriers with a permeability set at 0.01mD, hence uncertainty about fault transmissibility is not investigated.

Inversion Scheme

The Metropolis algorithm is designed to sample asymptotically the posterior distribution function so that the starting point or seed to generate the Markov chain of models is meaningless [Mosegaard and Tarantola, 1995]. Then, after the burn-in time is over, samples of the posterior can be picked according to a "waiting time" to get independent samples.

To speed up the process, we propose to generate first 2000 samples from the prior and then to use the 100 models that best match flow data as seeds for generating 100 independent Markov chains of models using the acceptance rule described in section 3.2.4.2. Each chain stops when the misfit is below a given threshold or 500 models are generated. This process does not guarantee unbiased sampling of the posterior distribution, but corresponds to a practical use of the method in which only models deemed acceptable are retained for further use.

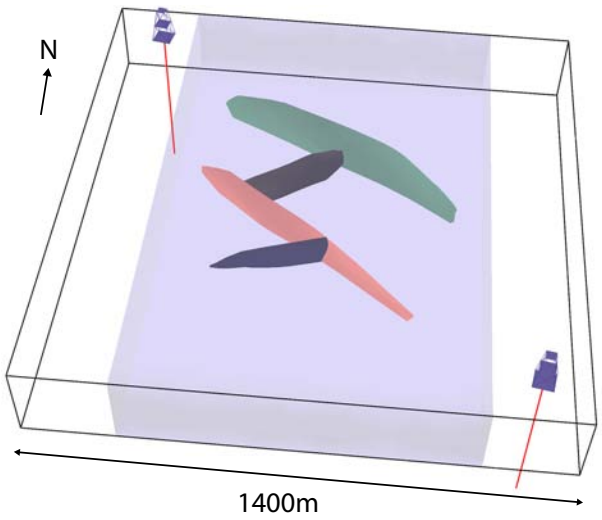

FIGURE 3.14 – The reference model has 4 faults. The shaded zone represents the fault zone thus ensures faults to be inside this zone.

Reference Model and selected models

The reference model is randomly drawn among the 2000 prior models and is presented in figure 3.14. Then, the misfit is computed for other prior models using equation 3.9 and water-cut data. Results are presented in figure 3.15. The misfit threshold to accept posterior models is set to 10 (from 51 for the worst of the 100 selected models) ; only four models are below that threshold among the 2000 models.

Sensitivity to Perturbation Amplitude Parameter

The perturbation amplitude parameter p controls the exploitation versus exploration of the

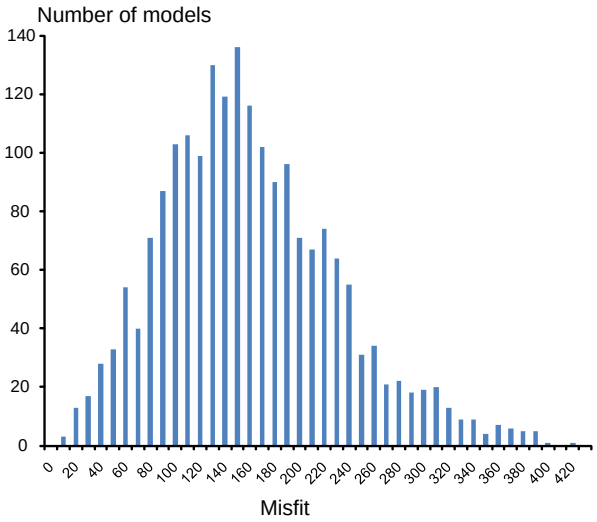

FIGURE 3.15 – Misfit repartition computed from equation 3.9. Only the 100 selected models are used as seeds to generate Markov chain of models and four of them have a misfit below 10 thus are considered as posterior models.

parameter space. To test the sensitivity of our methodology to parameter p, the best model is perturbed using different values of p. 100 models are generated for each value of p, and their flow-response misfit is computed and summarized in figure 3.16. Statistics are almost constant for small perturbations ranging from 0.0001 to 0.03 maybe due to the impact of the fault sinuosity seed that generates a new sinuosity whatever the perturbation value. Moreover, maximum values are unexpectedly constant, maybe due to a fault configuration in unstable state, *e.g.* a fault center coordinate close to a neighbor fault may switch of side even for small changes. Minimum misfit values are also almost constant due to the perturbation process that does not exclude small perturbations even for large p values. Consequently, we focus on centile values that do not depend on extreme values. For p ranging from 0.04 to 1, the median misfit increases linearly with a correlation coefficient of 0.997.

These results greatly depend on prior information, thus cannot be generalized. Indeed, as described in section 3.2.4.1, the parameters evolution speed depends on the parameters distribution, *e.g.* a Dirac distribution always generates the same value thus the speed is null whatever the perturbation amplitude is. Given these results, the perturbation amplitude p is set to 0.03 for the case study which ensures small misfit changes while enabling exploration of the parameter space.

Results

Any perturbation among the fault parameters may entail a fault network topology change and impact the flow response. Moreover, parameters are connected, *e.g.* a change of fault position increasing the distance between two faults may be balanced by a fault growth decreasing the distance between the two faults, thus cannot be considered individually. In case the fault size is small as compared to other faults, changes of fault location or orientation may entail insignificant changes in flow response. Consequently, we propose to compare fault occurrence probability maps (figures 3.17 and 3.18) instead of posterior distributions.

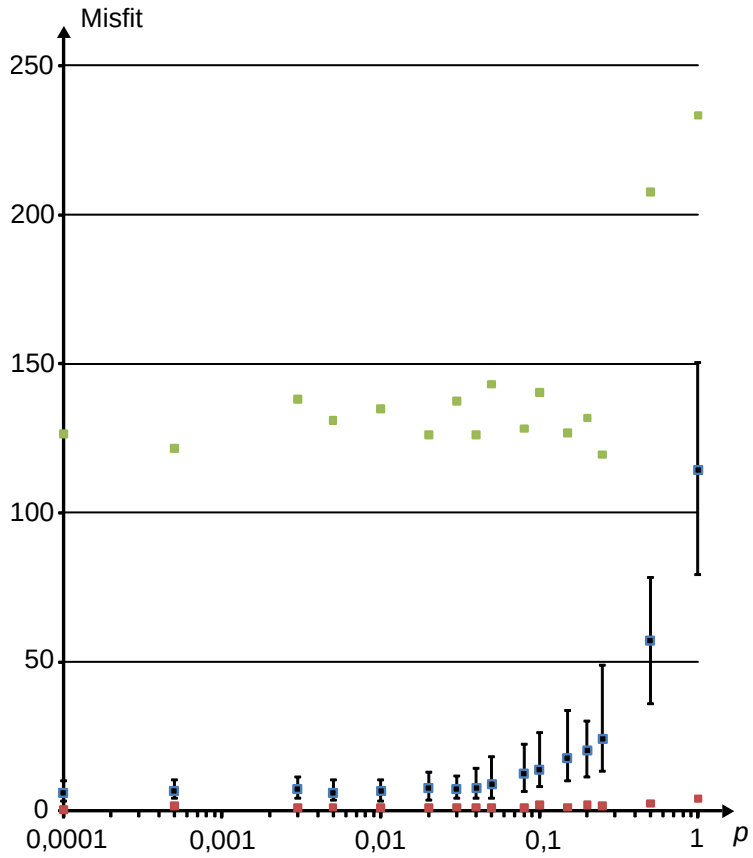

FIGURE 3.16 – Misfit values in function of perturbation amplitude p in logarithmic scale. Vertical bars represent interquartile ranges with median value. For p values ranging from 0.0001 to 0.03, median values are almost constant, then increase with larger interquartile ranges. Minimum (red squares) and maximum (green squares) values are almost constant for all p values due to the perturbation process and maybe unstable fault position.

Only three Markov chains among the 100 chains did not manage to reduce the misfit under the input threshold. An example of misfit evolution for a seed model is shown in figure 3.19. The average acceptance rate is about 43% for all Markov chains, which is relevant with prior studies [Gelman et al., 1996, Mosegaard and Sambridge, 2002]. Both prior probability maps do not show any particular feature but a large zone with medium probability. Posterior probability maps show specific features that are present in the two subsets which suggests relative convergence and stability of results.

Posterior probability maps relatively well resolve fault location uncertainty. A typical realization would have two faults of family A on both sides of the main flow path and one fault of family B playing the role of flow barrier along the flow path, which is relatively similar to the reference model. Other faults located elsewhere in the model do not impact the flow response, which may explain why these areas are not very well resolved on the posterior probability maps. Indeed, any hard data constrain fault location and the impact of conditioning is still to be evaluated.

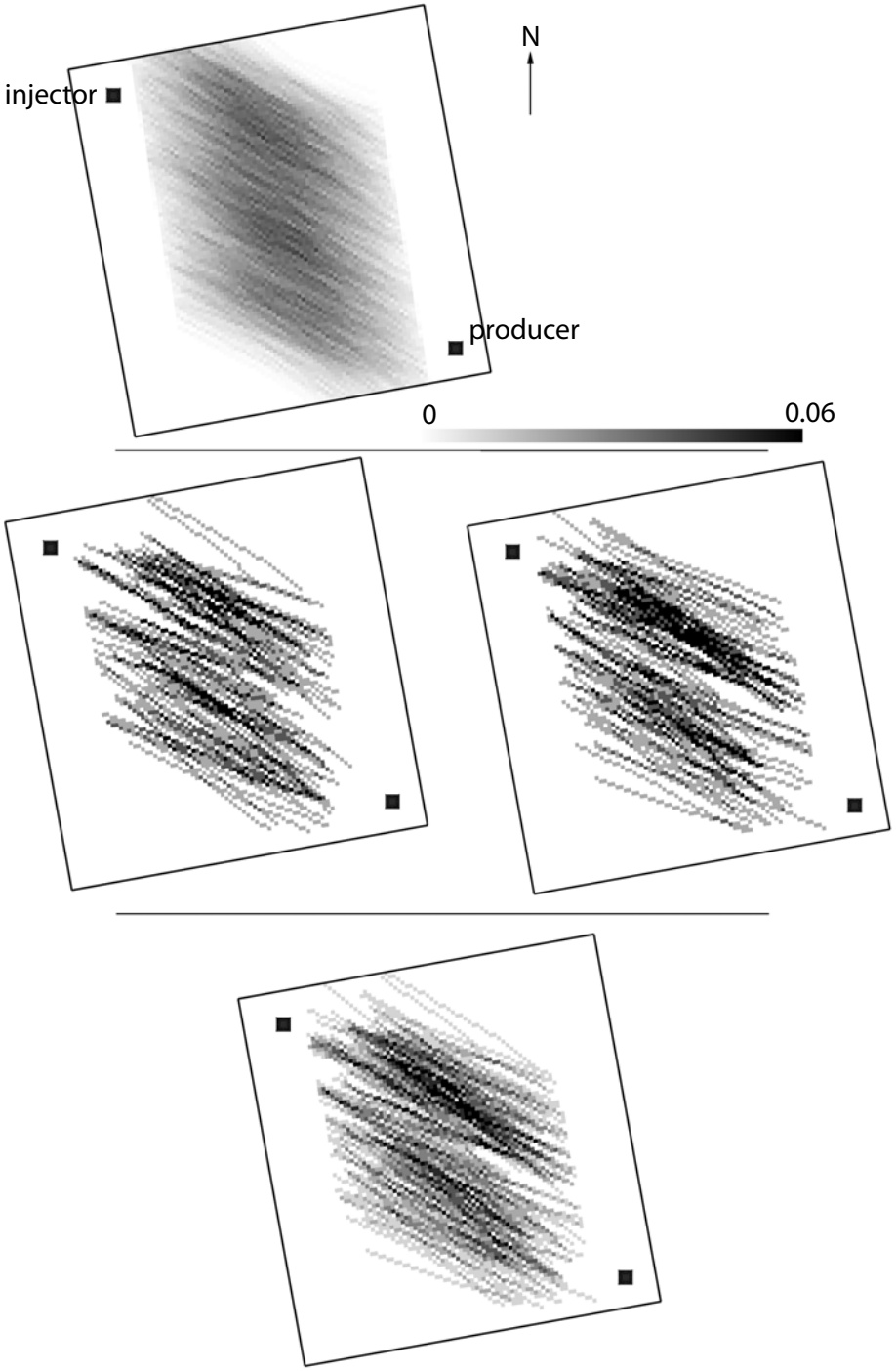

FIGURE 3.17 – Probability maps of faults belonging to family A. Top : the map corresponds to the 2000 prior models. Middle : posterior models computed using the inversion scheme from the best 50 models (left) and from the 50 following models (right). The two sets of models show similar characteristics, *i.e.* two high-probability areas on both sides of the main flow path. Bottom, the probability map computed from the 100 posterior models.

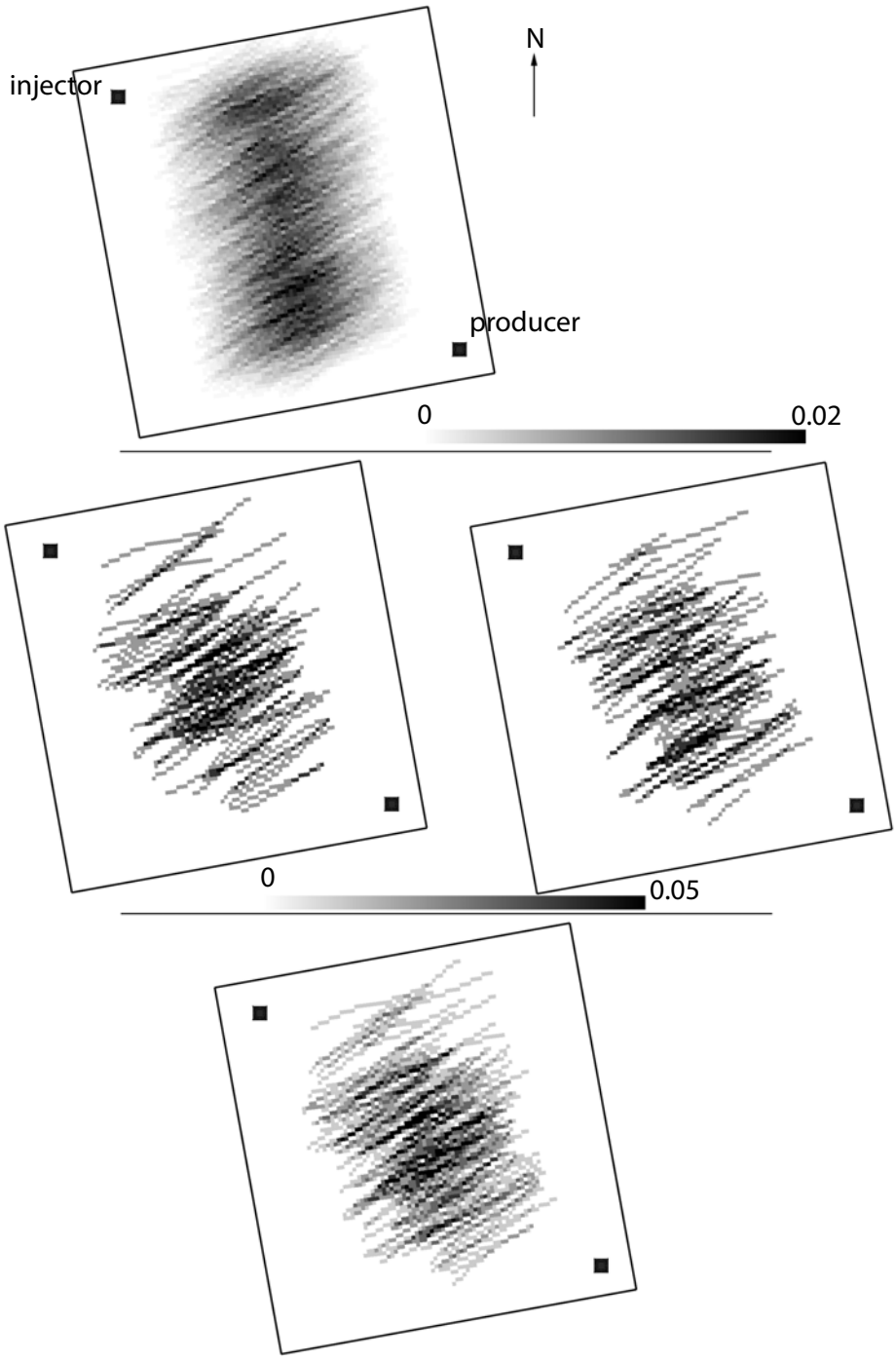

FIGURE 3.18 – Probability maps of faults belonging to family B. Top : the map obtained from the 2000 prior models. Middle : posterior models computed using the inversion scheme from the best 50 models (left) and from the 50 following models (right). Bottom : posterior probability map built from the 100 posterior models. The posterior map highlights an area on the main flow path which is related to the reference with a fault playing the role of flow barrier.

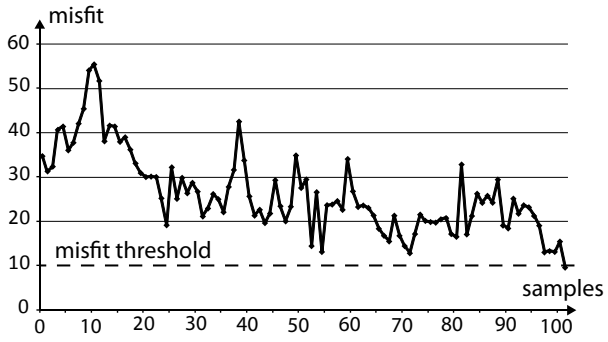

FIGURE 3.19 – Misfit evolution for one seed model. The Markov chain stops when the misfit is below the input threshold and the last sample is used to compute posterior statistics.

3.2.5.2 Analysis of Structural Scenarios

The interpretation of subsurface data may lead to different interpretations [Bond et al., 2007] and thus different geological scenarios. Fluid-flow data represent additional data that may help to analyze different interpretations and to falsify those that do not match observations [Tarantola, 2006]. In this section, we propose to analyze different structural scenarios built from the interpretation of seismic sections by generating a family of models around each interpretation. The sinuosity and size of each deterministic fault of the initial interpretation are perturbed. Other fault parameters such as fault center location and orientation are frozen. The initial interpretations are used as seeds to generate Markov chains of models as in section 3.2.5.1 and conditioning to seismic interpretations is made during neighborhood search. Only models that match flow data are retained and other interpretations are falsified. The study is not meant to be statistical but practical, hence scenarios that do not have any neighbor model with an acceptable misfit to water-cut data after a large number of iterations are falsified.

Data Overview

Eight structural scenarios have been built from four interpreted cross-sections, one of them is considered as the reference scenario (figure 3.20). There are one injector in the center and four producers in the corners of the studied area of size 1400x1400x200m. As in section 3.2.5.1, the misfit is the root mean squared error (equation 3.9) of water-cut data. Matrix porosity is set at 0.2, matrix permeability is 5mD and fault permeability is 0.01mD.

Results

Initial interpretations are perturbed until an acceptable misfit is reached. This arbitrary choice is motivated by a practical need to obtain a match in limited time, a more rigorous version of MCMC could be run in principle to better sample the search space, at the expense of computational efficiency.

The evolution of the misfit from one model to another is presented in figure 3.21. The average acceptance rate for all scenarios is about 33%. Scenarios I_1, I_4, I_5 and their corresponding chain of models have a flow response that remains far from the reference thus are falsified. Scenarios I_0, I_2, I_3, I_6 have a neighbor model within an acceptable misfit range, respectively I'_0, I'_2, I'_3, I'_6 (figure 3.22), which are considered consistent with available water-cut

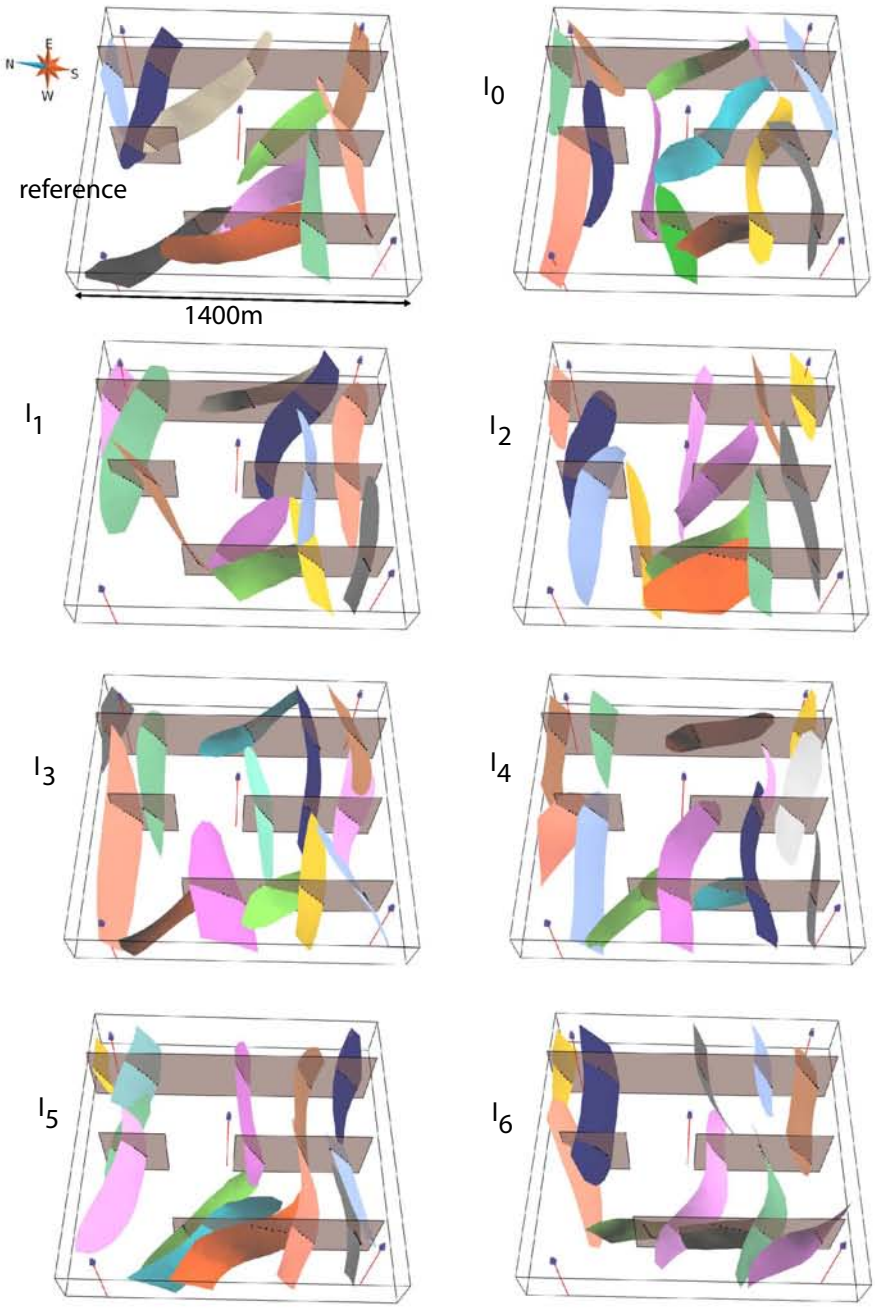

FIGURE 3.20 – Initial interpretations and reference. Only fault sinuosity and fault size are perturbed during the inversion scheme, other parameters are frozen.

data.

Conclusions would have been much different by only considering the initial interpretations misfit, *e.g.* interpretations I_2, I_3, I_6 have a larger initial misfit than falsified interpretations I_1, I_4, I_5.

The low-misfit models I'_0, I'_2, I'_3, I'_6 can then be used for further modeling steps with a higher confidence in their corresponding structural models.

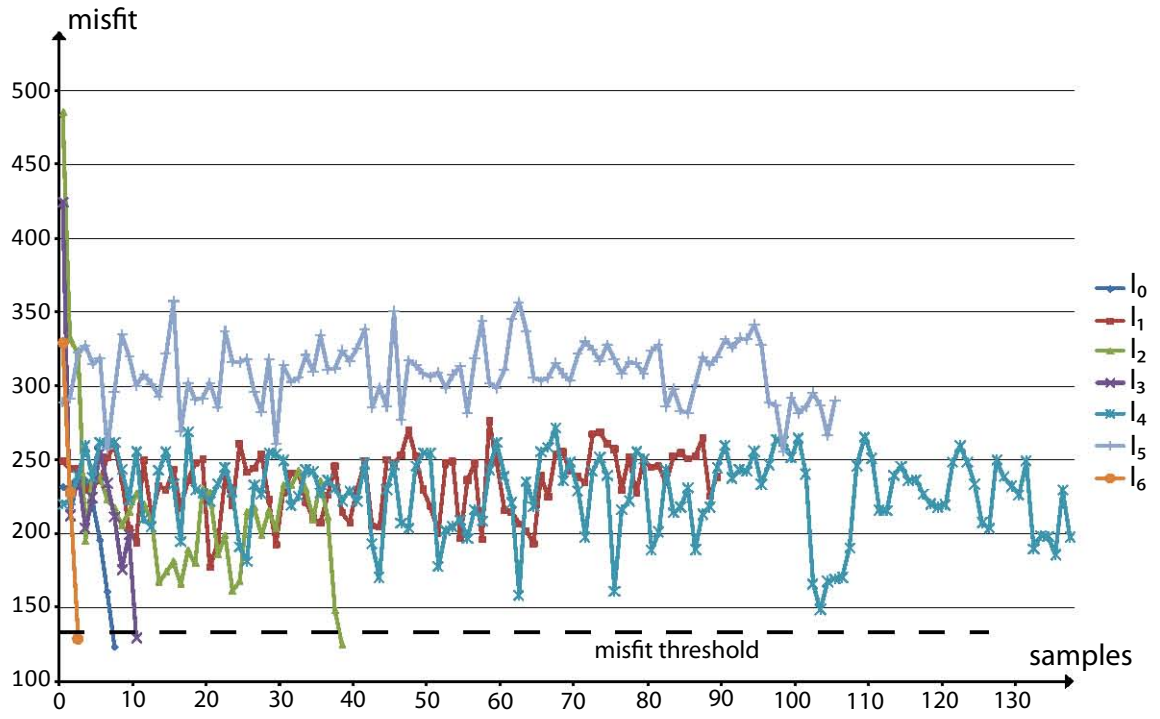

FIGURE 3.21 – Misfit evolution of seven structural scenarios using a version of Metropolis algorithm proposed by Mosegaard and Tarantola [1995] and water-cut data. Scenarios I_0, I_2, I_3, I_6 have a neighbor model consistent with available data. Other scenarios I_1, I_4, I_5 are falsified.

3.2.6 Discussion and Conclusions

The proposed parameterization of faults enables to sample posterior fault networks with various number of faults and fault connections. The modeling method uses a realistic 3D representation of fault surfaces that goes beyond planar elliptical surfaces and enables to model realistic fault arrays. It considers faults as a random vector corresponding to fault parameters and opens the way to inversion of geological structures using realistic descriptions. It goes beyond traditional inversion schemes by perturbing both fault geometry and fault network topology.

The first presented example shows how fault position uncertainty can be reduced using our parameterization and a version of the Metropolis algorithm to sample posterior models. The sampled 100 posterior models suggest specific features about fault location. The second example shows the interest of the methodology when conditioning data are available, by analyzing structural scenarios and falsifying those that do not have any neighbor model consistent with observations. The simulation method enables to permanently honor data points during the perturbation process, which reduces the neighborhood around selected scenarios and ensures the consistency of the proposed models.

The smooth evolution of fault parameters, even for the number of faults coupled with size evolution, enables to generate correlated flow responses from one sample to another. Only the fault sinuosity does not evolve smoothly from one sample to another but gradual deformation could help to obtain more continuous transitions [Hu, 2000]. Other fault-related

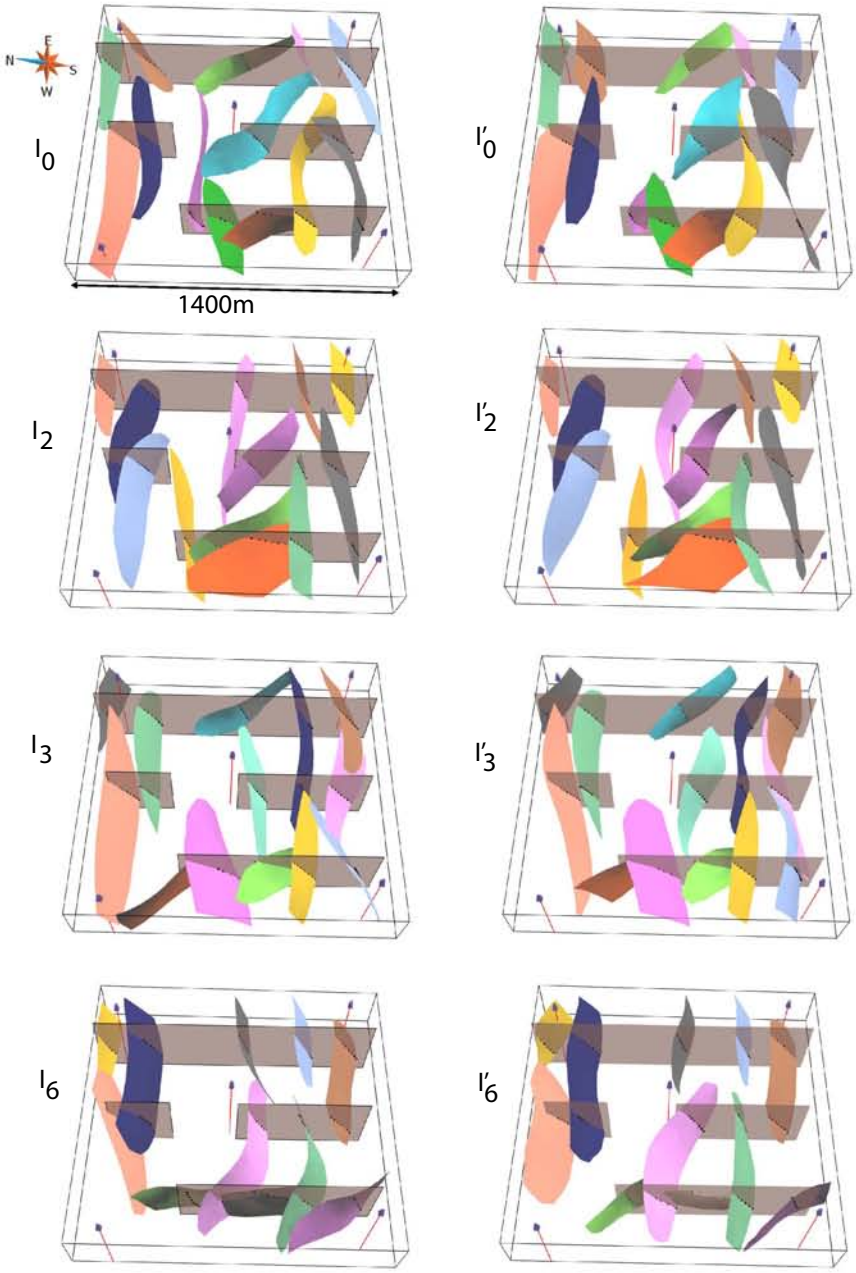

FIGURE 3.22 – Left : initial interpretations used as seeds for generating Markov chains of models. Right : the end-member of each chain, consistent with water-cut data. Other scenarios did not produce any acceptable model thus have been falsified.

uncertainties could be considered to complete the method, such as fault transmissibility which may have a significant impact on the flow response. In the presented examples, faults occur in a homogeneous geological layer but rock heterogeneities could also be taken into account. In such cases, fault displacement should also be modeled for it may connect or disconnect transmissive layers. This point will be addressed in a future contribution.

In our first simple MCMC example, the forward problem, *i.e.* the flow simulation, re-

presents about 91% of the computation time (about 9% for fault simulation, the time for discretization and generation of flow input files is negligible) for an average of 1min20s per MCMC iteration on a 3.4Ghz PC with about 8000 flow elements and 16000 connections per discretized model. When complexity increases as in the second example with a larger number of faults and wells, this computation time is even more important (about four times longer than first example) even if time proportions remain relatively constant : the inversion spends 88% of the computation time to solve flow equations, 12% to simulate faults (maybe due to extra simulation steps needed for data conditioning, other steps are insignificant). These performance issues and the large number of parameters do not enable to rigorously sample the posterior. The presented inversion strategy is parallel by nature since Markov chains of models can be computed independently, which reduces the total computation time of the presented examples. Other methodologies could also be used in order to reduce the computation time, *e.g.* Mariethoz et al. [2010] propose a stochastic stopping criterion to reduce the number of forward runs. Distance-based optimization [Sambridge, 1999, Scheidt and Caers, 2009, Suzuki et al., 2008] could also be used either separately or in conjunction with the proposed method for more efficient sampling.

Interpretations of inversion results are difficult to analyze due to the amount of data and correlation between parameters. As argued by Mosegaard and Tarantola [1995], there is still a need for methods to summarize information and reveal specific features of posterior models, that go beyond the computation of posterior distributions of individual inverse parameters. In that sense, Wellmann and Regenauer-Lieb [in press] propose to use information entropy to compare results and evaluate uncertainty reduction of geological units repartition after data inversion.

Beyond practical solutions and issues about the computation cost of inversions, one may wonder whether the search for fully sampled posterior distributions is meaningful. In the case of structural modeling, one only needs some models roughly covering the uncertainty space to continue the modeling workflow and make predictions. The Metropolis sampler enables such spread over the posterior probability function that cannot be obtained with optimization techniques. If other uncertainties are to be taken into account, *e.g.* about fault rock properties or fault slip, the empty space problem becomes even more difficult to handle and calls for much more models, which makes the computation time quickly becoming unreasonable in practice.

3.3 Gestion des nombres aléatoires

La méthodologie inverse présentée dans la section précédente s'appuie largement sur les nombres aléatoires utilisés lors de la simulation stochastique puisqu'ils constituent le coeur de la paramétrisation. Leur gestion est donc primordiale pour assurer le bon fonctionnement de la méthode inverse.

3.3.1 Vecteur de nombres aléatoires

L'ensemble des nombres aléatoires utilisés lors de la simulation d'un réseau de failles A peut être stocké dans un vecteur v_A, rempli au fur et à mesure de la simulation successive des failles. L'utilisation de ce même vecteur, pour des paramètres d'entrée identiques, permet de simuler à nouveau A. Cependant, la perturbation du vecteur v_A peut amener l'algorithme à utiliser un nombre différent de nombres aléatoires, par exemple :

- le changement de la géométrie d'une faille peut modifier les intersections entre failles. Or l'algorithme utilise des nombres aléatoires pour considérer un contact partiel soit en branchement, soit en croisement ;
- la modification d'une faille peut l'amener à recouper des ensembles de points. Dans ce cas, des nombres aléatoires sont utilisés pour décider de dévier et simuler à nouveau une géométrie de faille ou diviser en deux sous-ensembles les points recoupés (section 1.6.11.1 page 45) ;
- la nucléation ou la suppression (taille nulle) d'une faille modifie le nombre d'intersections à évaluer.

L'utilisation d'un vecteur aléatoire v_A de taille constante, parcouru linéairement du début à la fin, ne garantit donc pas la proximité (visuelle) entre deux simulations, la seconde étant obtenue après perturbation de v_A. En effet, des décalages peuvent se créer lors de la lecture informatique du vecteur aléatoire et les nombres aléatoires peuvent être utilisés lors d'étapes différentes d'une simulation à une autre après perturbation.

3.3.2 Structure de données gérant les nombres aléatoires

Une structure de données, appelée *rng* (pour random number generator), plus souple qu'un vecteur aléatoire a été mise en place afin de gérer correctement les nombres aléatoires requis lors de la simulation. La structure *rng* est divisée en sous-structures s_{F_i} pour chaque faille F_i du réseau de failles. Lors de la simulation de la $i^{\text{ème}}$ faille, l'algorithme accède directement à la $i^{\text{ème}}$ sous-structure s_{F_i}, indépendamment de la taille des $(i-1)^{\text{ème}}$ sous-structures précédentes. De cette manière, les nombres aléatoires ne peuvent pas être utilisés pour des failles différentes d'une simulation à l'autre, après perturbation.

La structure de données doit également assurer que pour une même faille F_i, *i.e.* une même sous-structure s_{F_i}, les nombres utilisés lors d'une étape de simulation ne sont pas utilisés pour une étape différente. Pour cela, la sous-structure s_{F_i} associe un mot clé représentant l'étape de simulation avec un vecteur v de nombres aléatoires. Ainsi, lors de la simulation du pendage de la $i^{\text{ème}}$ faille, l'algorithme requiert un nombre aléatoire à la structure *rng* avec les paramètres i et *pendage*. Le gestionnaire de nombres aléatoires *rng* accède alors à la $i^{\text{ème}}$ sous-structure s_{F_i}. La structure s_{F_i} accède ensuite au vecteur $v_{pendage}$ et retourne la valeur courante. En effet, chaque vecteur v connait le nombre d'accès, ce

qui permet de fournir les nombres successivement au cours des accès. Dans notre mise en oeuvre en C++, les associations entre indices et paramètres sont codées à l'aide de tables de hachage (*std : :map*).

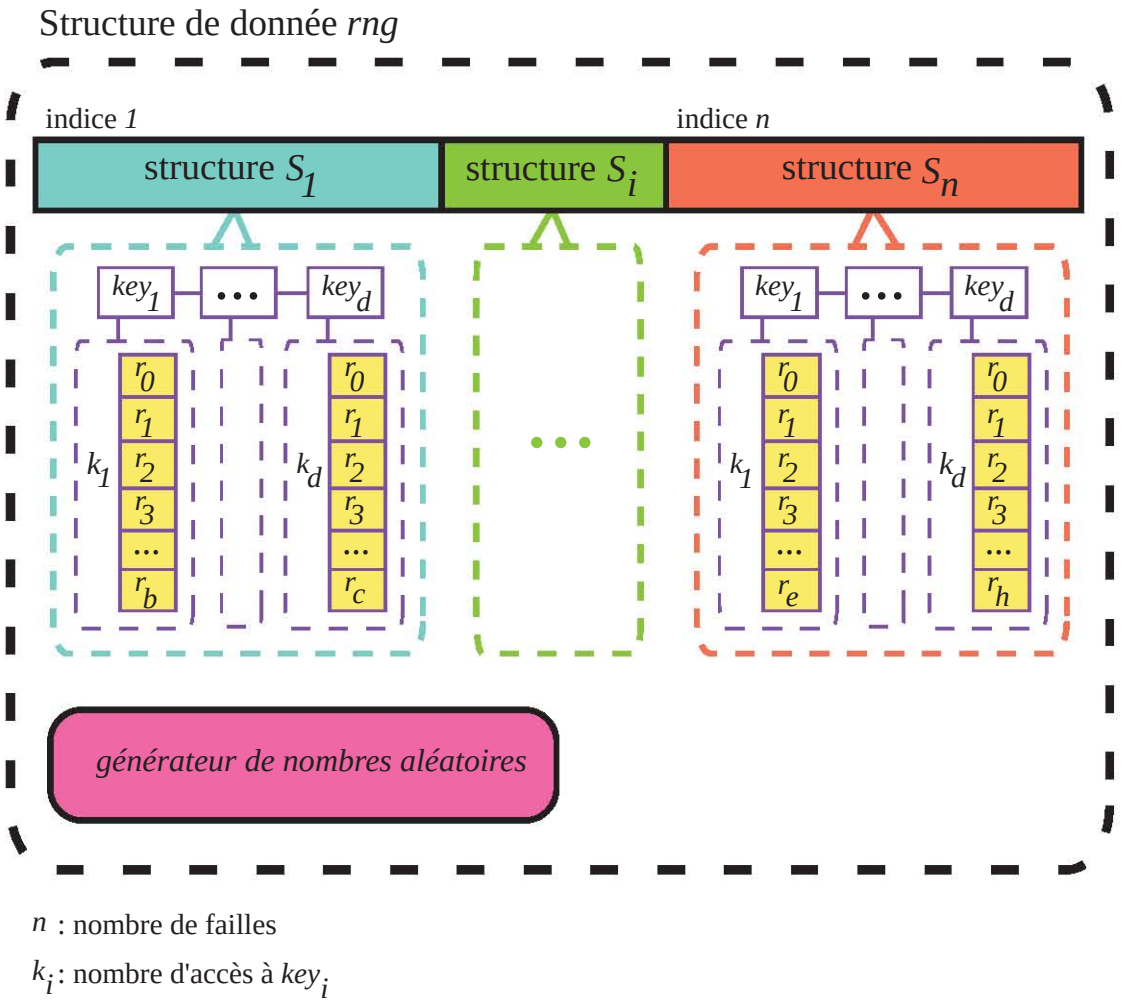

FIGURE 3.23 – Structure de données *rng* gérant les nombres aléatoires.

Si, au cours de la simulation, l'algorithme a besoin de nombres aléatoires supplémentaires (par exemple pour évaluer une nouvelle intersection partielle) stockés dans le vecteur v_{key} associé à la clé *key* (par exemple "intersection") de la structure s_{F_i} pour la $i^{\text{ème}}$ faille, la structure *rng* utilise un générateur de nombres aléatoires pour fournir un nouveau nombre qui est alors ajouté au vecteur v_{key}. La structure *rng* est présentée en figure 3.23 et l'accès aux nombres aléatoires est détaillé dans l'algorithme 7.

Entrées :
i = accès à la $i^{\text{ème}}$ faille
key = clé représentant une étape de simulation (exemples : "pendage", "intersection")

Définitions :
S = sous-structure de rng
k_{key} = nombre d'accès associé à key
g = générateur de nombres aléatoires
nb = nombre aléatoire

$S \leftarrow rng(i)$: $i^{\text{ème}}$ sous-structure de rng
$v \leftarrow S(key)$: vecteur de S associé à la clé key
$n \leftarrow$ taille de v
si $k_{key} < n$ **alors**
  $nb \leftarrow v_{key}[k_{key}]$
  $k_{key} \leftarrow k_{key} + 1$
sinon
  $nb \leftarrow$ nombre aléatoire généré par g
  $k_{key} \leftarrow k_{key} + 1$
  Ajouter nb à la fin de v_{key}
finsi

Sorties :
nb

Algorithme 7: Fonction de gestion des nombres aléatoires.

3.4 Perspectives

3.4.1 Familles de failles cogénétiques et modélisation inverse

La structure proposée pour gérer les nombres aléatoires n'est pas adaptée pour inclure des familles de failles cogénétiques dans la modélisation inverse. La perte de cohérence intervient lorsque l'ordre de simulation entre les différentes familles est modifié. Dans ce cas, une structure de nombres aléatoires d'une faille peut être utilisée par des familles différentes, ce qui revient à simuler une faille avec les nombres aléatoires d'une faille d'une autre famille. Le problème pourrait être résolu en adaptant la structure existante, en répartissant par exemple les structures des failles dans des structures correspondant aux familles de failles.

3.4.2 Échantillonnage de l'espace d'incertitudes

La méthode présentée dans ce mémoire génère des modèles structuraux stochastiques. Le but d'une telle approche est de couvrir l'ensemble des scénarios géologiquement possibles pour un ensemble d'informations (données interprétées, lois physiques, concepts géologiques) donné. Cette stratégie est séduisante en théorie puisqu'elle assure couvrir l'ensemble des possibles, pourvu que l'échantillonnage soit correct. Deux conditions sont nécessaires à un tel échantillonnage : (1) l'échantillonnage doit explorer l'ensemble de l'espace d'incertitudes qui peut être très grand notamment lorsque le nombre de variables est important (proportionnel m^n où n est le nombre de variables et m le nombre de catégories des variables) ; (2) le nombre de modèles générés doit être suffisant pour représenter cet espace. Ces deux conditions sont d'autant plus difficiles à remplir en pratique, l'échantillonnage étant souvent limité par la capacité informatique et le temps requis pour de tels calculs. En effet, un échantillonnage exhaustif est en général impossible à réaliser et seule une fraction de modèles est générée en pratique.

Dans les exemples d'application de la méthodologie inverse présentés en section 3.2.5, les échantillons sont séparés en deux sous-ensembles qui sont ensuite utilisés pour calculer des statistiques sur la position des failles. Si les deux sous-ensembles fournissent des résultats similaires, alors l'échantillonnage est considéré comme suffisant. Cette stratégie nécessite néanmoins de sur-échantillonner l'espace d'incertitudes afin de s'assurer de la stationnarité des résultats.

La définition d'une *métrique d'incertitude* permettrait de quantifier la dispersion des échantillons et donnerait des indications sur le nombre de modèles nécessaire pour garantir un échantillonnage correct de l'espace d'incertitudes.

Dans ce cadre, Aaditya Satija, actuellement en thèse avec Jef Caers à l'université de Stanford, s'intéresse au calcul de distances entre les arbres topologiques de réseaux de failles différents. Sa méthode compare chaque niveau successif entre deux arbres afin de quantifier leur similitude [Satija and Caers, 2010]. Aaditya teste également l'efficacité d'une telle métrique pour le calage historique, ce qui permettrait de n'évaluer le problème direct que sur quelques échantillons représentatifs, de manière similaire à Suzuki et al. [2008]. D'autres paramètres structuraux sont explorés, comme le nombre de failles, leur orientation ou leur position par rapport aux puits [Satija and Caers, 2011]. La définition de paramètres représentatifs permettrait d'appliquer des techniques de calcul de distances et de noyaux, comme proposé par Scheidt and Caers [2009] pour les faciès pétrophysiques.

3.5 Conclusions

Ce chapitre a abordé la modélisation inverse, présentée en trois étapes : la *paramétrisation du système* (réseau de failles), la *simulation du problème direct* (discrétisation et simulation d'écoulement de fluides) ainsi qu'une *stratégie d'inversion* basée sur l'algorithme de Metropolis. Cette méthodologie a permis d'utiliser des données d'écoulements de fluides afin de réduire les incertitudes sur les paramètres des failles. Les nombres aléatoires utilisés lors de la simulation stochastique sont gérés par une structure particulière qui maintient la cohérence du vecteur aléatoire entre deux modèles successifs.

La discrétisation proposée évite la rasterisation des surfaces de failles dans une grille cartésienne, traditionnellement utilisée pour la simulation des écoulements. En revanche, seul l'impact de la perméabilité des failles a été étudié et les incertitudes sur le rejet des failles, la stratigraphie ainsi que les champs de perméabilité et porosité ne sont pas prises en compte. Une conversion automatique de support de modélisation serait alors requise, et des algorithmes de conversion robustes et efficaces sont indispensables dans la perspective d'une inversion bayésienne nécessitant de nombreuses évaluations du problème direct.

La prise en compte de paramètres supplémentaires, comme le rejet des failles, augmente le nombre de paramètres incertains et la dimensionnalité du problème. L'intégration des incertitudes structurales dans les méthodes inverses n'utilisant traditionnellement que les propriétés des roches apparaît néanmoins nécessaire pour aborder le problème inverse de la manière la plus générale possible, étant donné l'interdépendance des paramètres.

Les failles ont également un impact sur la réponse du sous-sol à d'autres processus physiques, comme la propagation d'ondes ou la résistivité principalement due aux fluides qu'elles contiennent, et d'autres données pourraient être utilisées dans le cadre d'une méthode inverse.

conclusions générales

La connaissance limitée du sous-sol ne permet pas de représenter sans ambiguïtés ses hétérogénéités aussi bien pétrophysiques que structurales. Par conséquent, de nombreuses incertitudes affectent la construction des géomodèles et leur prédictibilité. Les travaux présentés dans ce mémoire ont permis d'aborder les incertitudes structurales par le développement d'une méthode de génération stochastique de modèles structuraux 3D.

La méthode proposée génère dans un premier temps les failles, organisées en familles, dans un ordre chronologique relatif (chapitre 1). La simulation successive de familles de failles permet de représenter la mise en place d'un réseau de failles suite à des évènements tectoniques distincts dans le temps. Chaque objet faille est construit au cours d'étapes successives, chaque étape permettant de modéliser une caractéristique spécifique de la faille courante (orientation, centre, sinuosité, dimension). Les horizons stratigraphiques sont ensuite construits par interpolation d'un champ scalaire stratigraphique. La méthode s'appuie sur la modélisation implicite, définissant une surface par une équipotentielle d'un champ scalaire volumique, pour représenter les interfaces géologiques. Cette approche facilite la description topologique d'un modèle et réduit fortement les contraintes de maillage de l'approche explicite.

L'échantillonnage des incertitudes relatives aux failles et aux horizons stratigraphiques en contexte incertain a permis de mettre en évidence la variabilité des estimations de volumes de roches, par rapport à l'échantillonnage restreint aux incertitudes stratigraphiques (chapitre 2).

Les incertitudes structurales n'affectent pas seulement les observations statiques mais également la réponse dynamique du sous-sol. Le chapitre 3 présente une méthode de modélisation stochastique inverse de la géométrie et de la connectivité des failles en utilisant des données d'écoulement de fluides. Dans ce cadre, la paramétrisation des failles en un nombre restreint de paramètres, leur discrétisation explicite en un graphe de connectivité et une stratégie d'inversion ont été proposées. Cette méthodologie permet de considérer les failles dans le problème inverse et les résultats obtenus dans des cas simples sont prometteurs dans l'optique d'une modélisation inverse plus globale.

La méthode proposée diffère des méthodes existantes par le processus de génération d'une réalisation. En effet, la plupart des méthodes existantes perturbent la géométrie d'un modèle de référence explicite tout en conservant sa topologie. La méthode proposée ne repose sur aucun modèle de référence, chaque réalisation étant générée à partir de l'interprétation des données et d'une connaissance géologique *a priori*. Cette approche permet donc de s'affranchir d'un modèle de référence, toujours subjectif et susceptible de comporter des biais d'interprétation. La démarche proposée reste également subjective lors de l'interprétation des données mais essaie de s'affranchir des biais introduits lors de la construction d'un modèle de référence.

L'interprétation des données est une étape primordiale aussi bien dans le cas d'une modélisation structurale déterministe que stochastique. En effet, elle garantit la cohérence des modèles par rapport aux observations en informant la géométrie des structures. Celle-ci est extrapolée, soit de manière subjective dans le cas déterministe, soit en échantillonnant la

variabilité naturelle des structures géologiques dans une approche stochastique. L'approche stochastique permet également de prendre en compte l'incertitude des données interprétées, à condition qu'un modèle d'incertitude leur soit associé. Cela nécessite donc une démarche plus complète lors de l'interprétation parcimonieuse des données. L'interprétateur ne doit pas seulement pointer les structures géologiques sur les différents supports (profil sismique, coupe interprétative, topographie,...) mais aussi préciser une mesure de l'incertitude, par exemple sous forme d'un intervalle de confiance ou d'une distribution statistique.

Dans le cas d'études réservoirs, le stockage des données géoréférencées permet d'estimer l'incertitude des données interprétées *a posteriori*, même si les outils automatiques ou semi-automatiques d'interprétation n'incluent pas de mesure directe de l'incertitude. Pourtant, l'extraction automatique d'un horizon sismique et de l'incertitude associée permettrait de faciliter la mise en oeuvre d'une approche stochastique. En revanche, dans le cas d'un projet de cartographie, l'approche stochastique ne peut s'envisager *a posteriori* une fois les mesures de terrain effectuées, puisqu'une mesure fine de l'incertitude nécessite l'observation des données brutes, c'est-à-dire le terrain lui-même. Le raisonnement scientifique serait donc différent de la démarche traditionnelle et la cartographie de terrain serait alors envisagée dans la perspective d'une modélisation stochastique. Chaque relevé d'orientation, de contact lithologique, de schistosité, de stries sur un miroir de faille pourrait être probabilisé, tant d'un point de vue de la mesure que de la localisation de l'observation. De nouvelles observations pourraient également aider la modélisation stochastique future : la position d'une faille disparaissant à l'affleurement pourrait par exemple être encadrée en cartant la limite de la couverture empêchant l'observation de la faille.

La méthode développée au cours de la thèse ajoute une connaissance *a priori* des structures géologiques aux données interprétées. Ces règles, qui limitent l'association entre points de données par exemple, permettent de limiter l'échantillonnage aux modèles aussi cohérents que possible. Elles permettent d'obtenir des géométries réalistes d'un point de vue géologique théorique, par exemple en limitant la sinuosité et le profil de déplacement le long d'une faille. Pourtant, la compatibilité des structures entre elles, étant donné un contexte tectonique, n'est pas garantie. Elle est pourtant nécessaire pour la validité géologique d'un modèle et son utilisation future, scientifique ou industrielle. La simulation du rejet des failles au cours de la simulation, en contrôlant le type de déformation, ainsi que l'intégration de données sur l'absence de failles permettraient dans un premier temps d'obtenir des modèles plus réalistes, au moins d'un point de vue géométrique.

La validité mécanique d'un modèle pourrait être évaluée de manière globale grâce à la restauration géomécanique. Chaque réalisation (modèle simulé) serait alors restaurée : les unités géologiques seraient successivement dépliées, de la plus récente à la plus ancienne, pour revenir dans le temps à l'état de dépôt. L'enregistrement de la déformation permettrait d'évaluer la vraisemblance du modèle dans le contexte tectonique supposé. Cependant, cette évaluation interviendrait *a posteriori*, une fois l'ensemble des structures simulées, ce qui limite l'efficacité pratique et théorique d'une telle approche. D'un point de vue théorique, elle nécessiterait de sur-échantillonner l'espace d'incertitudes puisqu'un grand nombre de modèles seraient falsifiés lors de la restauration. En pratique, le temps de calcul des codes de restauration géomécanique n'est pas négligeable. Comment garantir alors un maximum de compatibilité structurale au cours de la simulation stochastique

d'un modèle ? Des tests locaux, par exemple sur la courbure d'une surface par exemple, sur les variations d'épaisseur d'une unité géologique, ou sur l'analyse du déplacement le long d'une faille permettraientt de valider ou d'invalider les structures en cours de simulation. Lors de l'ajout d'une faille dans le modèle en cours de simulation, une première géométrie pourrait être simulée puis elle pourrait être optimisée selon des tests locaux de compatibilité [Thibault et al., 1996]. Une autre approche consisterait à inclure des principes mécaniques lors de la modélisation de la géométrie initiale des structures, par exemple lors du déplacement des horizons stratigraphiques dans le voisinage d'une faille [Bouziat et al., 2011], ou pour garantir la développabilité des horizons modélisés [Thibert et al., 2005].

La simulation stochastique de modèles structuraux et de géomodèles en général, en couvrant l'ensemble des scénarios géologiques possibles étant donné une connaissance limitée du sous-sol, n'a d'autre but que d'approcher la réalité géologique, insaisissable par essence. Une approche stochastique, intégrée dans un raisonnement scientifique spécifique, garantissant la validité géologique des scénarios proposés, permettra néanmoins de mieux l'appréhender.

Bibliographie

P. Abrahamsen. Bayesian kriging for seismic depth conversion of a multi-layer reservoir. In A. Soares, editor, *Geostatistics Tróia '92. Volume 1. Forth International Geostatistical Congress. Tróia, Portugal, 13-18 September, 1992*, volume 5 of *Qualitative Geology and Geostatistics*, pages 385–398. Kluwer Academic Publ., 1992. 8, 72, 108

R. V. Ackermann and R. W. Schlische. Anticlustering of small normal faults around larger faults. *Geology*, 25(12) :1127–1130, 1997. doi : 10.1130/0091-7613(1997)025<1127:AOSNFA>2.3.CO;2. 35, 36, 43

R. V. Ackermann, R. W. Schlische, and M. O. Withjack. The geometric and statistical evolution of normal fault systems : an experimental study of the effects of mechanical layer thickness on scaling laws. *Journal of Structural Geology*, 23(11) :1803 – 1819, 2001. ISSN 0191-8141. doi : DOI:10.1016/S0191-8141(01)00028-1. 35

M. Bahorich and S. Farmer. 3-D seismic discontinuity for faults and stratigraphic features : The coherence cube. *The Leading Edge*, 14(10) :1053–1058, 1995. doi : 10.1190/1.1437077. 35

G. Bardossy and J. Fodor. Traditional and new ways to handle uncertainty in geology. *Natural Resources Research*, 10(3) :179–187, 2001. doi : 10.1023/a:1012513107364. 6, 7

J. A. M. Barnett, J. Mortimer, J. H. Rippon, J. J. Walsh, and J. Watterson. Displacement geometry in the volume containing a single normal fault. *AAPG Bulletin*, 71(8) :925–937, 1987. 42, 69, 72, 110

A. Benedicto, R. A. Schultz, and R. Soliva. Layer thickness and the shape of faults. *Geophys. Res. Lett.*, 30(20), 2003. doi : 10.1029/2003GL018237. 42, 110

C. Bond, A. Gibbs, Z. Shipton, and S. Jones. What do you think this is ? "conceptual uncertainty" in geoscience interpretation. *GSA Today*, 17(11) :4–10, 2007. 3, 5, 6, 10, 61, 105, 120

C. Bond, Z. Shipton, A. Gibbs, and S. Jones. Structural models - optimizing risk analysis by understanding concept uncertainty. *First Break*, 26(6) :65–71, 2008. 4, 5

C. Bond, R. Lunn, Z. Shipton, and A. Lunn. What makes an expert effective at interpreting seismic images ? *Geology*, 2011. doi : 10.1130/G32375.1. 6

A. Bouziat, G. Laurent, and G. Caumon. Vector field based fault modeling and stratigraphic horizons deformation. In *Gocad Meeting Proceedings*, 2011. 54, 133

P. Calcagno, J. Chilès, G. Courrioux, and A. Guillen. Geological modelling from field data and geological knowledge : Part i. modelling method coupling 3d potential-field interpolation and geological rules. *Physics of the Earth and Planetary Interiors*, 171(1-4) :147 – 157, 2008. ISSN 0031-9201. Recent Advances in Computational Geodynamics : Theory, Numerics and Applications. 9, 17, 67, 108

J. Carrera, A. Alcolea, A. Medina, J. Hidalgo, and L. J. Slooten. Inverse problem in hydrogeology. *Hydrogeology Journal*, 13 :206–222, 2005. ISSN 1431-2174. 10.1007/s10040-004-0404-7. 107

G. Caumon. Towards stochastic time-varying geological modeling. *Mathematical Geosciences*, 42(5) :555–569, 2010. doi : 10.1007/s11004-010-9280-y. 65, 72, 73

G. Caumon, O. Grosse, and J.-L. Mallet. High resolution geostatistics on coarse unstructured flow grids. In O. Leuangthong and C. V. Deutsch, editors, *Geostatistics Banff, Proc. of the seventh International Geostatistics Congress*. Kluwer, Dordrecht, 2004a. 108

G. Caumon, C. H. Sword, and J.-L. Mallet. Building and editing a Sealed Geological Model. *Mathematical Geology*, 36(4) :405–424, 2004b. 18

G. Caumon, A.-L. Tertois, and L. Zhang. Elements for stochastic structural perturbation of stratigraphic models. In *Proc. Petroleum Geostatistics*. EAGE, 2007. 8, 9, 10, 39, 58, 61, 72, 108, 112

G. Caumon, P. Collon-Drouaillet, C. Le Carlier de Veslud, S. Viseur, and J. Sausse. Surface-based 3d modeling of geological structures. *Mathematical Geosciences*, 41 :927–945, 2009. ISSN 1874-8961. 10.1007/s11004-009-9244-2. 18

T. C. Chamberlin. The method of multiple working hypotheses. *Science*, 15 :92–96, 1890. 4

T. Charles, J. M. Guéméné, B. Corre, G. Vincent, and O. Dubrule. Experience with the quantification of subsurface uncertainties. *Paper presented at SPE Asia Pacific Oil and Gas Conference and Exhibition, Jakarta, Indonesia, SPE 68703, 17–19 April*, 2001. 8, 58

N. Cherpeau, G. Caumon, and B. Lévy. Stochastic simulations of fault networks in 3D structural modeling. *Comptes Rendus Geoscience*, 342(9) :687 – 694, 2010a. ISSN 1631-0713. doi : 10.1016/j.crte.2010.04.008. 8, 19, 20, 21, 22, 62, 64, 108, 109

N. Cherpeau, G. Caumon, and B. Lévy. Stochastic simulation of fault networks from 2D seismic lines. *SEG Technical Program Expanded Abstracts*, 29(1) :2366–2370, 2010b. doi : 10.1190/1.3513325. 62, 64, 70, 71, 109

N. Cherpeau, G. Caumon, J. Caers, and B. Lévy. Assessing the impact of fault connectivity uncertainty in reservoir studies using explicit discretization. In *SPE Reservoir Characterization and Simulation Conference and Exhibition (RCSC)*, 2011. SPE 148085. 86, 108

N. Cherpeau, G. Caumon, J. Caers, and B. Lévy. Method for stochastic inverse modeling of fault geometry and connectivity using flow data. *Mathematical Geosciences*, 44 :147–168, 2012. ISSN 1874-8961. doi : 10.1007/s11004-012-9389-2. 82, 85

C. Childs, T. Manzocchi, J. J. Walsh, C. G. Bonson, A. Nicol, and M. P. Schöpfer. A geometric model of fault zone and fault rock thickness variations. *Journal of Structural Geology*, 31(2) :117 – 127, 2009. ISSN 0191-8141. doi : DOI:10.1016/j.jsg.2008.08.009. 14, 15, 50

J. P. Chilès, C. Aug, A. Guillen, and T. Lees. Modelling the geometry of geological units and its uncertainty in 3d from structural data : The potential-field method. In *Proc. Orebody Modelling and Strategic Mine Planning*, pages 313–320, 2004. 17

J. Clément and G. Caumon. Improvement and edition of tetrahedral meshes in numerical geology. In *Gocad Meeting Proceedings*, June 2010. 72

A. E. Clifton and R. W. Schlische. Nucleation, growth, and linkage of faults in oblique rift zones : Results from experimental clay models and implications for maximum fault size. *Geology*, 29(5) :455–458, 2001. doi : 10.1130/0091-7613(2001)029<0455:NGALOF>2.0.CO;2. 35

G. de Marsily, F. Delay, J. Gonçalvès, P. Renard, V. Teles, and S. Violette. Dealing with spatial heterogeneity. *Hydrogeology Journal*, 13 :161–183, 2005. ISSN 1431-2174. 10.1007/s10040-004-0432-3. 107

L. J. Drew. Low-sulfide quartz gold deposit model. Technical report, U.S. Geological Survey, 2003. 53

Q. Du, V. Faber, and M. Gunzburger. Centroidal voronoi tesselations : Applications and algorithms. *SIAM Review*, 41 :637–676, 1999. 99

P. Durand-Riard. *Gestion de la complexité géologique en restauration géomécanique 3D.* PhD thesis, Institut National Polytechnique de Lorraine, Nancy, 2010. 18

A. Fabri, G.-J. Giezeman, L. Kettner, S. Schirra, and S. Schönherr. On the design of cgal a computational geometry algorithms library. *Softw. Pract. Exper.*, 30 :1167–1202, September 2000. ISSN 0038-0644. doi : 10.1002/1097-024X(200009)30:11<1167::AID-SPE337>3.0.CO;2-B. 94

H. Fossen. *Structural geology.* Cambridge University Press, 2010. 51

A. Foucault and J.-F. Raoult. *Dictionnaire de Géologie.* Dunod, 6ème edition, 2005. 14

T. Frank, A.-L. Tertois, and J.-L. Mallet. 3D-reconstruction of complex geological interfaces from irregularly distributed and noisy point data. *Computers & Geosciences*, 33(7) :932 – 943, 2007. ISSN 0098-3004. 9, 17, 67, 72, 108

R. Frodeman. Geological reasoning : geology as an interpretive and historical science. *Geological Society of America Bulletin*, 107(8) :960–968, 1995. cited By (since 1996) 59. 2

A. V. Gelder. Efficient computation of Polygon Area and Polyhedron Volume. In *Graphics Gems V*, pages 35–41. Academic Press, 1995. 94

A. Gelman, G. O. Roberts, and W. R. Gilks. Efficient Metropolis jumping rules. In *Bayesian statistics, 5 (Alicante, 1994)*, Oxford Sci. Publ., pages 599–607. Oxford Univ. Press, New York, 1996. 117

G. K. Gilbert. The inculcation of scientific method by example. *American Journal of Science*, 31 :284–299, 1886. 4

E. Gringarten, B. Arpat, S. Jayr, and J.-L. Mallet. New Geologic Grids for Robust Geostatistical Modeling of Hydrocarbon Reservoir. In J. Ortiz and X. Emery, editors, *Proc. eighth Geostatistical Geostatistics Congress*, volume 2, pages 647–656. Gecamin ltd, 2008. 106

A. Guillen, P. Calcagno, G. Courrioux, A. Joly, and P. Ledru. Geological modelling from field data and geological knowledge : Part ii. modelling validation using gravity and magnetic data inversion. *Physics of the Earth and Planetary Interiors*, 171(1-4) :158 – 169, 2008. ISSN 0031-9201. Recent Advances in Computational Geodynamics : Theory, Numerics and Applications. 9, 17, 108

V. Henrion. *Approche pseudo-génétique pour la simulation stochastique de la géométrie 3D de réseaux fracturés et karstiques.* PhD thesis, Institut National Polytechnique de Lorraine, 2011. 53

V. Henrion, G. Caumon, and N. Cherpeau. Odsim : An object-distance simulation method for conditioning complex natural structures. *Mathematical Geosciences*, 42 :911–924, 2010. 56

L. Holden, P. Mostad, B. F. Nielsen, J. Gjerde, C. Townsend, and S. Ottesen. Stochastic structural modeling. *Math. Geol.*, 35(8) :899–914, 2003. 9, 58, 62, 108

K. Hollund, P. Mostad, B. F. Nielsen, L. Holden, J. Gjerde, M. G. Contursi, A. J. McCann, C. Townsend, and E. Sverdrup. Havana - a fault modeling tool. In A. G. Koestler and R. Hunsdale, editors, *Hydrocarbon Seal Quantification. Norwegian Petroleum Society Conference, Stavanger, Norway*, volume 11 of *NPF Special Publication*, pages 157–171. Elsevier Science, 2002. 9

L. Hu. Gradual deformation and iterative calibration of gaussian-related stochastic models. *Mathematical Geology*, 32 :87–108, 2000. ISSN 0882-8121. doi : 10.1023/A:1007506918588. 122

P. Huggins, J. Watterson, J. J. Walsh, and C. Childs. Relay zone geometry and displacement transfer between normal faults recorded in coal-mine plans. *Journal of Structural Geology*, 17(12) :1741 – 1755, 1995. ISSN 0191-8141. doi : DOI:10.1016/0191-8141(95)00071-K. 50

J. Imber, G. W. Tuckwell, C. Childs, J. J. Walsh, T. Manzocchi, A. E. Heath, C. G. Bonson, and J. Strand. Three-dimensional distinct element modelling of relay growth and breaching along normal faults. *Journal of Structural Geology*, 26(10) :1897 – 1911, 2004. ISSN 0191-8141. doi : DOI:10.1016/j.jsg.2004.02.010. 50

S. Jayr, E. Gringarten, A.-L. Tertois, J.-L. Mallet, and J.-C. Dulac. The need for a correct geological modelling support : the advent of the uvt-transform. *First Break*, 26(10) : 73–79, 2008. 72

S. Jenni, L. Y. Hu, R. Basquet, G. de Marsily, and B. Bourbiaux. History matching of stochastic models of field-scale fractures : Methodology and case study. In *SPE Annual Technical Conference and Exhibition*, Houston, Texas, 2004. 107

B. J. Kadlec, G. A. Dorn, H. M. Tufo, and D. A. Yuen. Interactive 3-D computation of fault surfaces using level sets. *Visual Geosciences*, 13(1) :133–138, 2008. 56

M. Karimi-Fard, L. J. Durlofsky, and K. Aziz. An efficient discrete-fracture model applicable for general-purpose reservoir simulators. *SPE Journal*, 9 :227–236, June 2004. Paper 79699. 95, 96, 99

L. Kettner, K. Mehlhorn, S. Pion, S. Schirra, and C. Yap. Classroom examples of robustness problems in geometric computations. *Comput. Geom. Theory Appl.*, 40 :61–78, May 2008. ISSN 0925-7721. doi : 10.1016/j.comgeo.2007.06.003. 93

Y.-S. Kim and D. J. Sanderson. The relationship between displacement and length of faults : a review. *Earth-Science Reviews*, 68(3-4) :317 – 334, 2005. ISSN 0012-8252. doi : 10.1016/j.earscirev.2004.06.003. 29, 55, 77

Y.-S. Kim, D. C. P. Peacock, and D. J. Sanderson. Fault damage zones. *Journal of Structural Geology*, 26(3) :503 – 517, 2004. ISSN 0191-8141. doi : DOI:10.1016/j.jsg.2003.08.002. 50

J. I. Krueger and D. C. Funder. Towards a balanced social psychology : causes, consequences, and cures for the problem-seeking approach to social behavior and cognition. *Behavioral and Brain Sciences*, 27(3) :313–327, 2004. 3

F. Lallier, G. Caumon, J. Borgomano, S. Viseur, F. Fournier, and C. Antoine. Relevance of the stochastic stratigraphic well correlation approach for the study of complex carbonate settings : Application to the Malampaya buildup (Offshore Palawan, Philippines). In *Geological Society London Special Publication*, accepted. 82

M. Lecour, R. Cognot, I. Duvinage, P. Thore, and J.-C. Dulac. Modeling of stochastic faults and fault networks in a structural uncertainty study. *Petroleum Geoscience*, 7 :S31–S42, 2001. 8, 9, 10, 16, 40, 58, 61, 62, 71, 108

D. Ledez. *Modélisation d'objets naturels par formulation implicite.* PhD thesis, INPL, Nancy, France, 2003. 17, 37

F. Lepage. *Three-Dimensional Mesh Generation for the Simulation of Physical Phenomena in Geosciences.* PhD thesis, INPL, Nancy, France, 2003. 72

C. Li, S. Pion, and C. Yap. Recent progress in exact geometric computation. *Journal of Logic and Algebraic Programming*, 64(1) :85–111, 2005. doi : 10.1016/j.jlap.2004.07.006. 94

L. Liang, D. Hale, and M. Maučec. Estimating fault displacements in seismic images. *SEG Technical Program Expanded Abstracts*, 29(1) :1357–1361, 2010. doi : 10.1190/1.3513094. 63

J.-L. Mallet. Discrete smooth interpolation. *Computer-Aided Design*, 24(4) :263–270, 1992. 17, 67

J.-L. Mallet. Space-time mathematical framework for sedimentary geology. *Mathematical geology*, 36(1) :1–32, 2004. 14, 17, 67, 72, 106

J.-L. Mallet and A.-L. Tertois. Solid earth modeling and geometric uncertainties. In *SPE Annual Technical Conference and Exhibition*, Florence, Italy, 2010. 61, 73, 108

C. J. Mann. *Uncertainty in Geology*, chapter 20, pages 241–254. Oxford University Press, New York, NY, 1993. 6

Manzocchi, T., Walsh, J. J., Nell, P. A. R., and Yielding, G. Fault transmissibility multipliers for flow simulation models. *Pet. Geol.*, 5 :53–63, 1999. 106

K. J. Marfurt, R. L. Kirlin, S. L. Farmer, and M. S. Bahorich. 3-D seismic attributes using a semblance-based coherency algorithm. *Geophysics*, 63(4) :1150–1165, 1998. doi : 10.1190/1.1444415. 35

G. Mariethoz, P. Renard, and J. Caers. Bayesian inverse problem and optimization with iterative spatial resampling. *Water Resources Research*, 46, 2010. doi : 10.1029/2010WR009274. 124

R. Merland, B. Levy, G. Caumon, and P. Collon-Drouaillet. Building centroidal voronoi tessellations for flow simulation in reservoirs using flow information. In *SPE Reservoir Simulation Symposium (SPE141018)*, Feb. 2011. doi : 10.2118/141018-MS. 102

N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics*, 21(6) :1087–1092, 1953. 84, 111

K. Mosegaard and M. Sambridge. Monte Carlo analysis of inverse problems. *Inverse Problems*, 18(3) :R29–R54, 2002. 84, 111, 117

K. Mosegaard and A. Tarantola. Monte carlo sampling of solutions to inverse problems. *Journal of Geophysical Research*, 100(B7) :12431–12447, 1995. doi : 10.1029/94JB03097. 111, 115, 122, 124

R. Moyen, J.-L. Mallet, T. Frank, B. Leflon, and J.-J. Royer. 3D-parameterization of the 3D geological space - the GeoChron model. In *Proc. European Conference on the Mathematics of Oil Recovery (ECMOR IX)*, 2004. 17

K. L. Munthe, L. Holden, P. Mostad, and C. Townsend. Modelling sub-seismic fault patterns using a marked point process. In *ECMOR IV. Proceedings of the 4th European Conference on the Mathematics of Oil Recovery in Røros, Norway, 7-10 June 1994. Topic B : Heterogeneity Description and Assessment of Uncertainty.* EAGE, 1994. 35

P. Nivlet, F. Fournier, and J.-J. Royer. A new nonparametric discriminant analysis algorithm accounting for bounded data errors. *Mathematical Geology*, 34(2) :223–246, Feb. 2002. 61

D. Oliver and Y. Chen. Recent progress on reservoir history matching : a review. *Computational Geosciences*, 15 :185–221, 2011. ISSN 1420-0597. 10.1007/s10596-010-9194-2. 107

H. Omre. Bayesian kriging-merging observations and qualified guesses in kriging. *Mathematical Geology*, 19 :25–39, 1987. ISSN 0882-8121. 10.1007/BF01275432. 9

D. W. Peaceman. Interpretation of well-block pressures in numerical reservoir simulation with nonsquare grid blocks and anisotropic permeability. *SPE Journal*, 23(3) :531–543, 1983. 97

D. Peacock and E. Parfitt. Active relay ramps and normal fault propagation on kilauea volcano, hawaii. *Journal of Structural Geology*, 24(4) :729 – 742, 2002. ISSN 0191-8141. doi : DOI:10.1016/S0191-8141(01)00109-2. 25, 51

D. C. P. Peacock and D. J. Sanderson. Strike-slip relay ramps. *Journal of Structural Geology*, 17(10) :1351 – 1360, 1995. ISSN 0191-8141. doi : DOI:10.1016/0191-8141(95)97303-W. 50

E. J. H. Rijks and J. C. E. M. Jauffred. Attribute extraction : An important application in any detailed 3-d interpretation study. *The Leading Edge*, 10(9) :11–19, 1991. doi : 10.1190/1.1436837. 35

P. Røe, F. Georgsen, A. Syversveen, and O. Lia. Fault displacement modelling using 3D vector fields. In *12th European Conference on the Mathematics of Oil Recovery*. EAGE, 2010. 55, 81

M. J. Ronayne, S. M. Gorelick, and J. Caers. Identifying discrete geologic structures that produce anomalous hydraulic response : An inverse modeling approach. *Water Resources Research*, 44, W08426, 2008. doi : 10.1029/2007WR006635. 107

T. Saito and J. I. Toriwaki. New algorithms for euclidean distance transformation of an n-dimensional digitized picture with applications. *Pattern recognition*, 27(11) :1551–1565, 1994. 37

M. Sambridge. Geophysical inversion with a neighbourhood algorithm-I : Searching a parameter space. *Geophysical Journal International*, 138 :479–494, 1999. 10, 124

M. Sambridge and K. Mosegaard. Monte Carlo methods in geophysical inverse problems. *Reviews of Geophysics*, 40(3) :29, Nov. 2002. 84

A. Satija and J. Caers. History matching subject to prior structural uncertainty : Challenges and proposals. In *SRCF meeting*, 2010. 128

A. Satija and J. Caers. Applying the distance-based approach to model flow uncertainty in structurally complex reservoirs : initial investigation. In *SRCF meeting*, 2011. 128

C. Scheidt and J. Caers. Representing spatial uncertainty using distances and kernels. *Mathematical Geosciences*, 41 :397–419, 2009. ISSN 1874-8961. doi : 10.1007/s11004-008-9186-0. 61, 124, 128

S. Schirra. Robustness and precision issues in geometric computation. In J.-R. Sack and J. Urrutia, editors, *Handbook of Computational Geometry*, pages 597–632. Elsevier Science Publishers B.V. North-Holland, Amsterdam, 2000. 94

A. Seiler, S. Aanonsen, G. Evensen, and J. C. Rivenæs. Structural Surface Uncertainty Modeling and Updating Using the Ensemble Kalman Filter. *SPE Journal*, 15(4) :1062–1076, 2010. ISSN 1086-055X. 61, 108

T. J. Smith and L. A. Marshall. Bayesian methods in hydrologic modeling : A study of recent advancements in Markov chain Monte Carlo techniques. *Water Resources Research*, 44, 2008. doi : 10.1029/2007WR006705. 84

R. M. Srivastava. Reservoir characterization with probability field simulation. In *SPE Annual Technical Conference and Exhibition*, pages 927–938, Washington D. C., 1992. SPE 24753. 9, 40, 71

S. Suzuki, G. Caumon, and J. Caers. Dynamic data integration for structural modeling : model screening approach using a distance-based model parameterization. *Computational Geosciences*, 12 :105–119, 2008. 10, 61, 72, 107, 124, 128

A. Tarantola. *Inverse Problem Theory*. SIAM : Society for Industrial and Applied Mathematics, 2004. ISBN 978-0898715729. 352 p. 84

A. Tarantola. Popper, bayes and the inverse problem. *Nature Physics*, 2 :492–494, 2006. 61, 120

A.-L. Tertois. *Création et modification de modèles géologiques par champs de potentiel. Application au modèle GeoChron*. PhD thesis, Institut national polytechnique de Lorraine, 2007. 41, 43

A.-L. Tertois and J.-L. Mallet. Distance maps and virtual fault blocks in tetrahedral models, 2006. 26^{th} Gocad Meeting, Nancy, June. 19

M. Thibault, J.-P. Gratier, M. Léger, and J.-M. Morvan. An inverse method for determining three dimensional fault with thread criterion : strike slip and thrust faults. *Journal of Structural Geology*, 18 :1127–1138, 1996. 133

B. Thibert, J.-P. Gratier, and J.-M. Morvan. A direct method for modeling and unfolding developable surfaces and its application to the ventura basin (california). *Journal of Structural Geology*, 27(2) :303–316, 2005. 133

P. Thore, A. Shtuka, M. Lecour, T. Ait-Ettajer, and R. Cognot. Structural uncertainties : determination, management and applications. *Geophysics*, 67(3) :840–852, 2002. 2, 8, 9, 10, 58, 61, 72, 108

A. Tversky and D. Kahneman. Judgment under uncertainty : Heuristics and biases. *Science*, 185(4157) :1124–1131, 1974. ISSN 00368075. 3

M. A. M. M. Van Dortmont, H. M. M. Van De Wetering, and A. C. Telea. Skeletonization and distance transforms of 3d volumes using graphics hardware. In *DGCI'06*, volume 4245, pages 617–629. LNCS, 2006. 56

S. Vitel. *Méthodes de discrétisation et de changement d'échelle pour les réservoirs fracturés 3D*. PhD thesis, INPL, Nancy, France, 2007. 86

W. von Funck, H. Theisel, and H.-P. Seidel. Vector field based shape deformations. *ACM Trans. Graph.*, 25(3) :1118–1125, 2006. ISSN 0730-0301. doi : http://doi.acm.org/10.1145/1141911.1142002. 55

J. J. Walsh, J. Watterson, W. R. Bailey, and C. Childs. Fault relays, bends and branch-lines. *Journal of Structural Geology*, 21(8-9) :1019 – 1026, 1999. ISSN 0191-8141. 53

J. J. Walsh, W. R. Bailey, C. Childs, A. Nicol, and C. G. Bonson. Formation of segmented normal faults : a 3-D perspective. *Journal of Structural Geology*, 25(8) :1251 – 1262, 2003. ISSN 0191-8141. doi : 10.1016/S0191-8141(02)00161-X. 14, 42, 69, 72, 110

J. F. Wellmann and K. Regenauer-Lieb. Uncertainties have a meaning : Information entropy as a quality measure for 3-D geological models. *Tectonophysics*, in press. ISSN 0040-1951. doi : 10.1016/j.tecto.2011.05.001. 124

J. F. Wellmann, F. G. Horowitz, E. Schill, and K. Regenauer-Lieb. Towards incorporating uncertainty of structural data in 3d geological inversion. *Tectonophysics*, 490(3-4) :141 – 151, 2010. ISSN 0040-1951. doi : 10.1016/j.tecto.2010.04.022. 7, 8, 9, 58, 62, 108

A. W. Western, G. Blöschl, and R. B. Grayson. Toward capturing hydrologically significant connectivity in spatial patterns. *Water Resources Research*, 37(1) :83+, 2001. ISSN 0043-1397. doi : 10.1029/2000WR900241. 107

R. Wood and A. Curtis. Geological prior information and its applications to geoscientific problems. *Geological Society Special Publication*, (239) :1–14, 2004. cited By (since 1996) 9. 5

C. K. Yap. Robust geometric computation. In J. E. Goodman and J. O'Rourke, editors, *Handbook of Discrete and Computational Geometry*, chapter 41, pages 927–952. Chapmen & Hall/CRC, Boca Raton, FL, 2nd edition, 2004. 94

G. Yielding, T. Needham, and H. Jones. Sampling of fault populations using sub-surface data : a review. *Journal of Structural Geology*, 18(2-3) :135 – 146, 1996. ISSN 0191-8141. doi : 10.1016/S0191-8141(96)80039-3. 77

D. Zampieri. Segmentation and linkage of the lessini mountains normal faults, southern alps, italy. *Tectonophysics*, 319(1) :19 – 31, 2000. ISSN 0040-1951. doi : 10.1016/S0040-1951(00)00025-1. 24

D. A. Zimmerman, G. de Marsily, C. A. Gotway, M. G. Marietta, C. L. Axness, R. L. Bras, J. Carrera, P. B. Davies, D. P. Gallegos, J. G. Hernández, A. L. Gutjahr, P. K. Kitanidis, A. M. Lavenue, D. McLaughlin, B. S. RamaRao, C. Ravenne, R. L. Beauheim, Y. Rubin, S. P. Neuman, P. Grindrod, and G. Dagan. A comparison of seven geostatistically based inverse approaches to estimate transmissivities for modeling advective transport by groundwater flow. *Water Resources Research*, 34(6) :1373–1413, 1998. 107

H. J. Zimmermann. An application-oriented view of modeling uncertainty. *European Journal of Operational Research*, 122(2) :190 – 198, 2000. ISSN 0377-2217. doi : DOI:10.1016/S0377-2217(99)00228-3. 2

Annexe A

Publications

Cette annexe regroupe les différentes publications en premier auteur écrites au cours de la thèse, hormis celles des Gocad Meetings :

- N. Cherpeau, G. Caumon, and B. Lévy. Stochastic simulations of fault networks in 3D structural modeling. *Comptes Rendus Geoscience*, 342(9) :687 – 694, 2010

- N. Cherpeau, G. Caumon, and B. Lévy. Stochastic simulation of fault networks from 2D seismic lines. *SEG Technical Program Expanded Abstracts*, 29(1) :2366–2370, 2010

- N. Cherpeau, G. Caumon, J. Caers, and B. Lévy. Assessing the impact of fault connectivity uncertainty in reservoir studies using explicit discretization. In *SPE Reservoir Characterization and Simulation Conference and Exhibition (RCSC)*, 2011. SPE 148085.

Stochastic simulations of fault networks in 3D structural modeling

Nicolas Cherpeau[*,a], Guillaume Caumon[a], Bruno Lévy[b]

[a]*CRPG - CNRS, Nancy université, rue du Doyen Marcel Roubault, 54501 Vandoeuvre-lès-Nancy, France*
[b]*Centre INRIA Nancy Grand-Est, Campus scientifique, 615 rue du Jardin Botanique, 54600 Villers les Nancy, France*

Abstract

3D Structural modeling is a major instrument in geosciences, e.g. for the assessment of groundwater and energy resources or nuclear waste underground storage. Fault network modeling is a particularly crucial step during this task, for faults compartmentalize rock units and play a key role in subsurface flow, whether faults are sealing barriers or drains.

Whereas most structural uncertainty modeling techniques only allow for geometrical changes and keep the topology fixed, we propose a new method for creating realistic stochastic fault networks with different topologies. The idea is to combine an implicit representation of geological surfaces which provides new perspectives for handling topological changes with a stochastic binary tree to represent the spatial regions. Each node of the tree is a fault, separating the space in two fault blocks. Changes in this binary tree modify the fault relations and therefore the topology of the model.

Résumé

Simulations stochastiques de réseaux de failles en modélisation structurale 3D. La modélisation structurale est largement utilisée en géoscience, notamment pour l'évaluation des ressources énergétiques et hydriques du sous-sol. La caractérisation des failles est l'une des étapes clés du processus de modélisation étant donné leur importance dans les écoulements de subsurface.

Alors que la plupart des techniques de modélisation d'incertitudes structurales existantes perturbent seulement la géométrie des objets, nous proposons une nouvelle méthode de simulation stochastique de réseaux de failles, incluant des changements topologiques. Cette méthode associe une modélisation implicite des surfaces géologiques, avec un arbre binaire permettant d'agencer les régions spatiales du modèle. Chaque noeud de l'arbre représente une faille, séparant l'espace en deux blocs. Des changements dans l'arbre binaire modifient les relations entre failles et par conséquent la topologie du modèle.

Key words: 3D structural modeling, Structural uncertainties, Implicit modeling, Fault network, Constructive Solid Modeling

Mots clés : Modélisation structurale 3D, Incertitudes structurales, Modélisation implicite, Réseau de failles, Modélisation solide

1. Introduction

A 3D structural model of the subsurface helps visualizing, understanding and quantifying geophysical processes and assessing natural resources. Indeed, geological structures control to some extend the spatial layout of subsurface heterogeneities. Therefore, most geostatistical petrophysical modeling methods use distances which follow folded and faulted structures [1–3]. Moreover, physical modeling codes are best run on grids conforming to geological structures [4–6].

However, surface geology and borehole data only provide limited information about subsurface structures and even exhaustively sampled 3D seismic surveys cannot remove interpretational uncertainties. Uncertainties are due to the inherent incompleteness and the limited resolution of geological data sets. Geology is by nature an interpretive science [7] and geological concepts and physical laws help geoscientists reducing uncertainties. However, geoscientists may introduce a prior geological knowledge bias and/or human bias [8] when interpreting subsurface data.

These uncertainties have a wide range of consequences in quantitative geosciences, and Gilbert [9] and Chamberlin [10] in their pioneering work already argued for multiple hypotheses. Consequently, uncertainties can be assessed by considering not just one (probably wrong) deterministic model, but a set of possible structural models. Depending on the amount and quality of observations, three levels of uncertainty can be defined by comparing such possible 3D structural models:

- Low uncertainty when all models have the same layout of structural surfaces, and show relatively small geometric variations.
- Medium uncertainty when the geometry of surfaces changes more significantly and the connection between geological surfaces may vary locally.
- High uncertainty when the number of structural interfaces, their spatial layout and their geometry are globally variable, except at some observation points.

[*]Corresponding author
Email address: nicolas.cherpeau@gocad.org (Nicolas Cherpeau)

From a modeling standpoint, these three degrees of uncertainty correspond to increasing difficulty. Therefore, multi-realization methods have mostly focused on low and medium uncertainty, primarily to assess risk in hydrocarbon reservoir management (rock volume estimates [11], underground flow [12], well plannning [13], history matching [14]). Most of these methods proceed by perturbing a reference interpretation to generate a large number of equiprobable geometric models of geological structures. In faulted formations, this may tend to underestimate variability by perturbing only fault geometry and leaving the first-order fault connectivity constant. Therefore, existing methods are best suited for large scale studies (pluri-decametric) with high-resolution 3D seismic data. For smaller objects or sparser data, stochastic approaches should also sample large uncertainties when generating possible models, some of which should be falsified whenever a new observation is made [15].

In this paper, we propose to sample both the connectivity and the geometry of fault networks in order to account for large uncertainties due to limited data quantity and resolution. Faults indeed are key elements in 3D structural models, and often have a first-order impact on the modeling output.

After a review of uncertainty modeling methods using both traditional explicit modeling or recent implicit approach (section 2), we introduce our stochastic fault simulation method (section 3).

2. Structural uncertainties: state of the art

Most existing structural uncertainty modeling methods proceed by perturbing a reference interpretation. The result is a large number of equiprobable geometries representing the uncertainty relative to geological structures.

2.1. Explicit geometrical perturbation techniques and limitations

In explicit modeling, geological interfaces are represented as polygonal surfaces. Lecour *et al.* [16] propose to modify the geometry of a surface by perturbing its nodes along an uncertainty bar defined at each node. The perturbation is correlated along the surface using the probability field method [17], in order to obtain realistic geometries. In the case of fault perturbation, curvature sign along the sliding direction is preserved to ensure fault compliance.

Other techniques have been developed to perturb the geometry of structural models, keeping the topology fixed [14, 18–20]. For a whole model, each surface is perturbed according to its structural uncertainties and all connections (horizon to horizon, horizon to fault and fault to fault) are stored. Once all geological objects have been perturbed, connections are honored so that the topology is preserved [16]. In practice, this raises a number of challenges to maintain model consistency in the case of large uncertainties, for interferences between surfaces and large mesh distorsions may occur. Moreover, large geometrical uncertainties may require topology changes, for instance horizon to fault connections (Fig. 1).

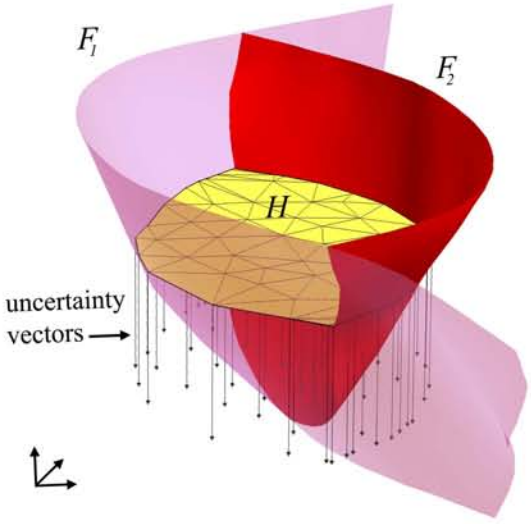

Figure 1: Topology changes induced by geometrical changes. In the reference case, the horizon H is in between faults F_1 and F_2. A simple drift of H along uncertainty vectors may completely change the configuration: H could possibly occur on either side of F_1 or F_2 and not anymore in between the two faults.

Fig. 1. Changements topologiques induits par des perturbations géométriques. Dans le cas initial, l'horizon H se situe entre les failles F_1 et F_2. Une simple translation de H le long de vecteurs d'incertitude peut complètement changer la configuration : H pourrait être en contact avec seulement F_1 ou F_2 et ne plus se situer entre les deux failles.

2.2. Topological perturbation

Topology is all about connections and relations between objects. In a structural model, topological changes may be introduced in different manners:

- Adding or removing geological objects.
- In the case of faults, changing the truncation rule between two faults.
- During geometrical perturbations, a simple drift of an horizon along uncertainty vectors may induce a topological change (especially fault to horizon connections, Fig. 1).

Few methods have been proposed to change the topology of a structural model. A fault modeling tool, referred to as Havana has been proposed in [21, 22]. It is mainly designed for the oil industry and works directly on a corner-point reservoir grid used for flow simulations. This choice allows for directly observing the effects of structural uncertainties on flow simulations. However, resevoir grids have known shortcomings to accurately represent geological structures. Consequently, subseismic faults are added by simply modifying the permeability field. Main faults are bilinear planes parallel to the pillars of the reservoir grid or stair-stepped faults.

In our approach, we borrow the idea of fault operator to [21], but use a flexible representation of faults. This makes it possible to account for large structural uncertainties without making simplifications due to the grid orientation. For accuracy, this method uses an implicit representation of structural interfaces, as in [23–25].

2.3. Implicit modeling: new perspectives for 3D modeling

An implicit surface (e.g. Fig. 3*d*) is described by an isovalue f of a monotonic volumetric function $\mathcal{F}(x, y, z)$ (e.g. Fig. 3*a*,*c*)(e.g. the geological time [26], the signed distance to an object [27]):

$$\mathcal{F}(x, y, z) = f \tag{1}$$

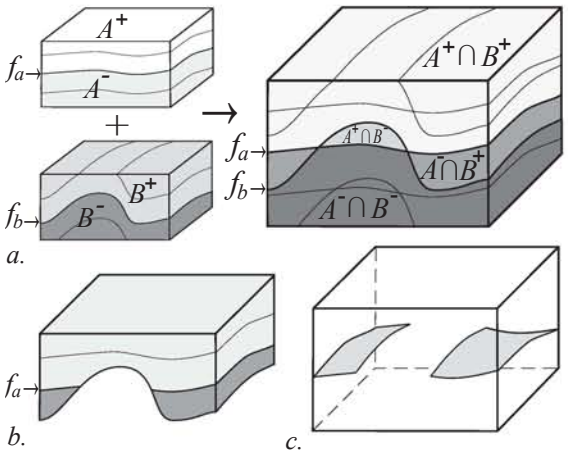

Figure 2: Boolean operations in implicit modeling. *a*. Left: two surfaces A ($\mathcal{F}_A(x, y, z) = f_A$) and B ($\mathcal{F}_B(x, y, z) = f_B$) and their corresponding half-spaces. Right: boolean sum of the two scalar fields defining four spatial regions. *b*. Example of boolean operation: $\mathcal{F}_A(x, y, z) \cap \mathcal{F}_B(x, y, z) \geq f_B$. *c*. Example of truncation corresponding to the operation: $\mathcal{F}_A(x, y, z) = f_A \cap \mathcal{F}_B(x, y, z) \geq f_B$ (equation 2).

Fig 2. Opérations booléennes en modélisation implicite. *a*. Gauche : deux surfaces A ($\mathcal{F}_A(x, y, z) = f_A$) et B ($\mathcal{F}_B(x, y, z) = f_B$) et leurs demi-espaces. Droite: 4 régions spatiales définies par l'intersection des deux champs scalaires. *b*. Exemple d'opération booléenne: $\mathcal{F}_A(x, y, z) \cap \mathcal{F}_B(x, y, z) \geq f_B$. *c*. Exemple d'intersection : $\mathcal{F}_A(x, y, z) = f_A \cap \mathcal{F}_B(x, y, z) \geq f_B$ (équation 2).

A conforming stratigraphic column is then represented as a set of isopotentials of a same scalar field, whereas unconformities and faults are defined by their own scalar field. Consequently, horizons can be perturbed independently of faults. Truncations by faults are only honored when an explicit surface is extracted from an isovalue of a scalar field.

Implicit modeling provides a means to easily truncate some surfaces by others using Constructive Solid Geometry (CSG) concepts. For instance, a surface A ($\mathcal{F}_A(x, y, z) = f_A$) can be truncated by another implicit surface B ($\mathcal{F}_B(x, y, z) = f_B$), by making a boolean intersection between A and either half-spaces of B (Fig. 2). For instance, a truncated surface $A_{|B^+}$ is defined by:

$$\mathcal{F}_A(x, y, z) = f_A | \mathcal{F}_B(x, y, z) \geq f_B \tag{2}$$

The reference scalar field of an implicit surface ($\mathcal{F}(x, y, z) = f$) can be perturbed by introducing a correlated random field $\mathcal{R}(x, y, z)$ [18]. The new surface is defined by the same isovalue f in the field $\mathcal{F}(x, y, z) + \mathcal{R}(x, y, z)$ (Fig. 3).

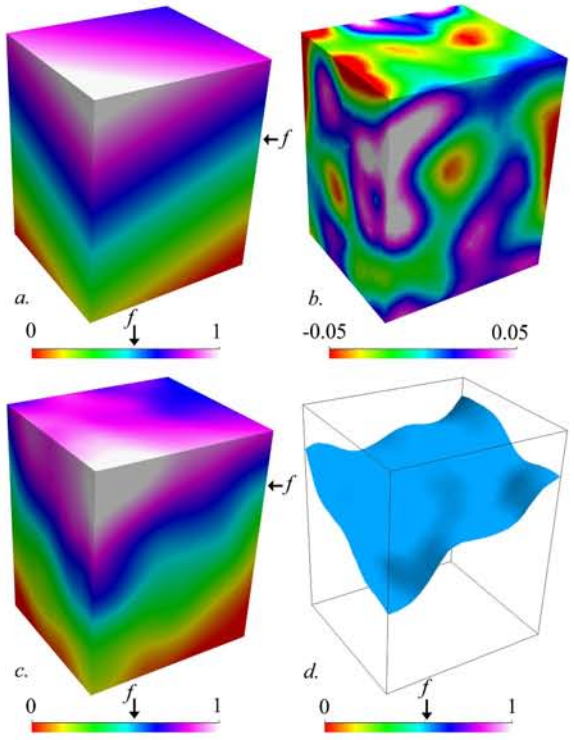

Figure 3: Geometrical perturbation method in implicit modeling. *a*. The reference scalar field $\mathcal{F}(x, y, z)$ defining the initial implicit surface (isovalue f, planar surface). *b*. A correlated random field $\mathcal{R}(x, y, z)$ generated by Sequential Gaussian Simulation represents the perturbation field. *c*. The surface is now defined by the isovalue f in the field $\mathcal{F}(x, y, z) + \mathcal{R}(x, y, z)$. *d*. View of the perturbed surface.

Fig 3. Méthode de perturbation géométrique en modélisation implicite. *a*. Le champ scalaire de référence $\mathcal{F}(x, y, z)$ définit la surface initiale (isovaleur f, surface plane). *b*. Un champ aléatoire corrélé généré par une Simulation Séquentielle Gaussienne correspond au champ de perturbation. *c*. La surface est maintenant définie par l'isovaleur f dans le champ $\mathcal{F}(x, y, z) + \mathcal{R}(x, y, z)$. *d*. Surface perturbée.

3. A new method for stochastic simulations of fault networks

The proposed stochastic fault simulation method takes advantage of the implicit surface perturbation and CSG operations between faults. To represent how faults partition the domain and interact one with another, we propose using a binary tree.

3.1. Binary trees as descriptors of fault relationships

Indeed, each fault divides the model $\mathcal{M}$ in two distinct fault blocks B^- and B^+ [28], defined by:

- $B^- = \{(x, y, z) \in \mathcal{M} | \mathcal{F}(x, y, z) < f\}$
- $B^+ = \{(x, y, z) \in \mathcal{M} | \mathcal{F}(x, y, z) > f\}$

Then, in the binary tree representing a fault network, each node represents a fault and each leave (node without children) represents a fault block (Fig. 4).

Each fault in the tree is potentially a branching fault for its parent faults in the tree. Switching a parent and its child in the tree changes their relationship: the main fault becomes the

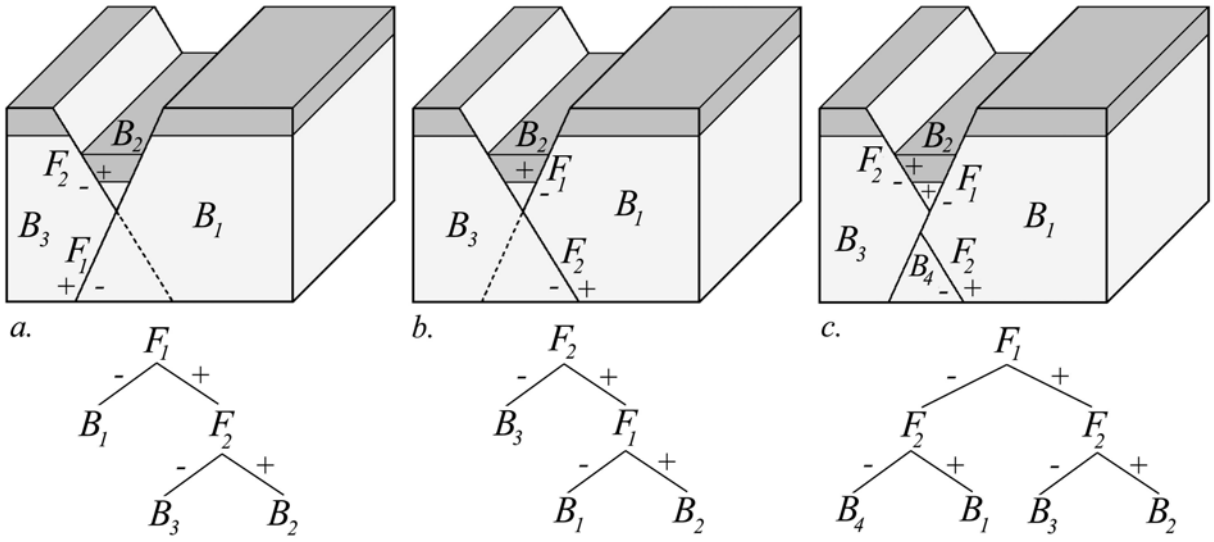

Figure 5: Consequences of filiation relationships in a binary tree from a geological point of view. *a*. The fault F_1 is the root, F_2 is a branching fault. *b*. Inversion of the relationship. F_1 is now a branching fault. *c*. A fault F_2 being on both sides of F_1 means that F_2 is cut by F_1. F_2 is the child of F_1 in the tree although it is the oldest fault.

Fig. 5. Conséquences des relations de filiation dans un arbre binaire d'un point de vue géologique. *a*. La faille F_1 est la racine de l'arbre, F_2 est la faille secondaire. *b*. Inversion de la relation père-fils. F_1 est maintenant la faille secondaire. *c*. Une faille F_2 présente des deux côtés de F_1 signifie que F_2 est recoupée par F_1. F_2 est la fils de F_1 dans l'arbre bien qu'elle soit plus ancienne.

branching fault and inversely (Fig. 5). A special case may occur when a fault F cuts a set of older faults $S_{old} = \{F_i | i \in [1, n]\}$. In this case, the oldest faults S_{old} are on both sides of the cutting fault F in the 3D model, hence they appear in the two branches of F in the tree (Fig. 5*c*).

When modeling a given fault array, some faults may not be truncated by other faults. Therefore, individual faults can be considered either parent or child in the binary tree, there is no consequence in term of truncation since faults are not in contact. Consequently, a given fault network may be described by several binary trees (Fig. 6). As there is no one-to-one correspondence between a fault network and a binary tree, the latter should be considered as a hierarchy of fault events, the oldest fault being the root of the tree, except for faults offset by more recent faults. Nevertheless, a tree fully defines the topology of the fault network.

3.2. Fault array simulations

3.2.1. Fault network composition

We consider a fault network composed of one or several fault families. Fault families regroup faults having similar structural parameters (orientation, type). If only a few data about a fault family is available, it can be described by statistical input parameters:

- a relative age
- a number of faults
- orientation (dip and strike) distributions
- perturbation parameters

Alternatively, fault families may be described by any scalar field (e.g. coming from the interpolation of field, borehole or seismic data) and perturbation parameters defining the associated uncertainties. The key input parameter is the relative age of faults since it determines the order of simulation and thus the place in the binary tree (section 3.2.3).

To simulate a fault array, only one binary tree is needed, containing all the faults of the network. Different fault networks are obtained by generating different binary trees.

3.2.2. Method for simulating a fault

For a given fault family, each fault is simulated as follows:

- Add a leaf randomly in the tree (i.e. choose randomly a fault block).
- Define the initial fault surface by a reference field $\mathcal{F}(x, y, z)$ from statistical parameters or input scalar field.
- Perturb the initial geometry by simulating a noise $\mathcal{R}(x, y, z)$ using a Sequential Gaussian Simulation, set to 0 perturbation at data location (method presented in Fig. 3).
- Define the fault by the field $\mathcal{F}'(x, y, z) = \mathcal{F}(x, y, z) + \mathcal{R}(x, y, z)$.
- Draw an isovalue f from the range of values that occurs in the selected block to define the fault by $\mathcal{F}'(x, y, z) = f$.

For a model $\mathcal{M}$ containing n fault blocks $B_{i|i\in[1,n]}$, let be p_i the probability for the block B_i to be selected for containing the new fault, with $\sum_{i=1}^{n} p_i = 1$. Different strategies are possible to

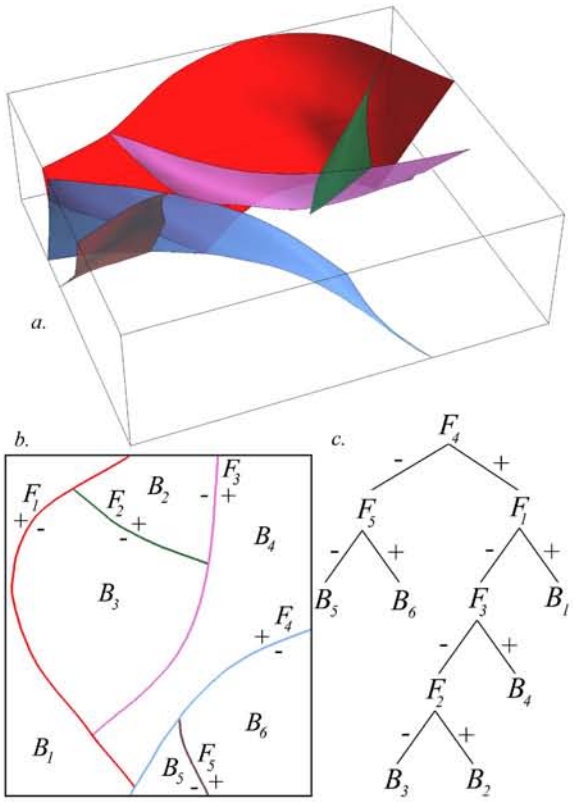

Figure 4: A fault network and its topological representation as a binary tree. *a*. Fault network in 3D. *b*. The same model in top view to make the topology more explicit. Each fault F_i divides the model in two blocks B_i. *c*. Binary tree representing the fault relationships. The fault F_4 is the main and oldest fault (root node of the tree), other faults are branching faults.

Fig. 4. Un réseau de failles et sa représentation topologique en arbre binaire. *a*. Réseau de failles en 3D. *b*. Le même modèle vu de dessus afin d'expliciter la topologie. Chaque faille F_i coupe le modèle en deux blocs B_i. *c*. Arbre binaire représentant les relations entre failles. La faille F_4 est la faille principale (racine de l'arbre), les autres failles étant des failles secondaires.

define $p_{i|i\in[1,n]}$. For instance, a uniform probability law can be used:

$$p_{i|i\in[1,n]} = \frac{1}{n} \tag{3}$$

Using this strategy, no block is priviledged and fault blocks totally different in volume may be obtained. Another strategy consists in balancing the probability by the volume V_i of the block B_i:

$$p_{i|i\in[1,n]} = \frac{V_i}{\sum_{i=1}^{n} V_i} \tag{4}$$

The volume of a block defined by several implicit faults can be calculated using the method presented in [29].

Once a block $B_{selected}$ has been drawn, an isovalue f corresponding to the fault ($\mathcal{F}'(x, y, z) = f$) can be drawn in $B_{selected}$ in different manners. As for selecting a block, each location in the block can be equiprobable. Therefore the isovalue can be drawn from a uniform law whose extremities are the minimum and maximum values of the field $\mathcal{F}'(x, y, z)$ in the block $B_{selected}$. However, as observed in [21, 22], main sub-seismic faults tends to repulse each other. A simple way to simulate such a behavior is to draw the isovalue corresponding to the fault from a symetric triangular or truncated Gaussian law so that medium positions in the block are priviledged.

3.2.3. Order of simulation

In the general case, fault families are simulated in chronological order, the oldest one first. Consequently, all the faults belonging to the same family are branching faults for the faults belonging to older families (Fig. 7).

3.2.4. Case of cogenetic faulting

Faults with different structural characteristics (i.e., belonging to different families) may initiate simultaneously at the geological time scale and thus truncate each other. One example of such a situation are conjugate faults, corresponding to steeply opposed-dipping faults. In this case, fault families cannot be simulated one after the other. Instead, let be S_{coeval} the set of all the faults to be simulated belonging to cogenetic families. The method is iterative: a fault is drawn in the set S_{coeval} and added randomly in the binary tree, until S_{coeval} is empty. Consequently, faults belonging to a given family may be parent or child for the other cogenetic families, hence different truncation rules between families (Fig. 8*d*).

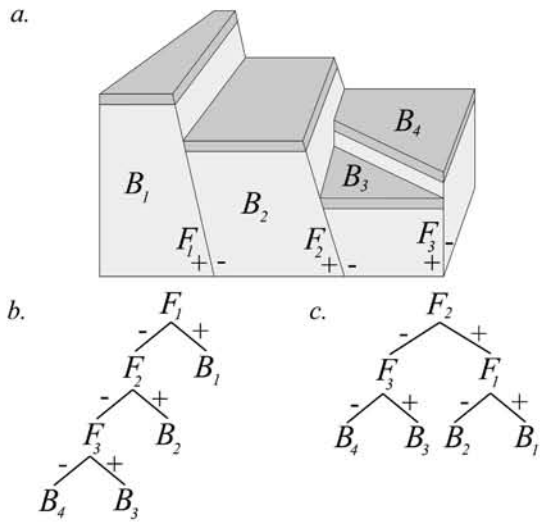

Figure 6: Non uniqueness of the representation of a fault array in a binary tree. *a*. 3D diagram representing three faults. The faults F_1 and F_2 are not in contact so they can either be parent or child in the binary tree. *b*. Representation of the three faults in a binary tree with F_1 as the parent fault. *c*. F_2 is now the root of the tree without any change about their truncation in the diagram. There may be a consequence of the change deeper, outside the diagram.

Fig. 6. Non unicité de la représentation d'un ensemble de failles dans un arbre binaire. *a*. Diagramme 3D représentant 3 failles. Les failles F_1 et F_2 ne sont pas en contact. Par conséquent, elles peuvent être père ou fils dans l'arbre binaire. *b*. Arbre binaire avec F_1 en tant que faille parente. *c*. F_2 est maintenant la racine de l'arbre mais aucun recoupement n'a changé. Il se peut que le recoupement change plus en profondeur, en dehors du diagramme.

3.2.5. Case of cross-cutting faults

A special situation may occur when a fault is offset by other faults (Fig. 5*c*). Indeed, in this case, the parent fault in the tree is the youngest one. This situation can be obtained by simulating the youngest fault F_{young} ($F_{young}(x, y, z) = f_{young}$) first and then simulating the shifted fault $F_{shifted}$ in a block (adding it at one leaf as usually). Then, depending on the displacement of the cutting fault F_{young}, $F_{shifted}$ may be added to another leaf. Assuming no rotation of F_{young}'s displacement, $F_{shifted}$ is defined by the same scalar field but by different isovalues depending on the fault block:

- $F_{shifted}(x, y, z) = f_{-}|F_{young}(x, y, z) \leq f_{young}$
- $F_{shifted}(x, y, z) = f_{+}|F_{young}(x, y, z) \geq f_{young}$

4. Perspectives and conclusion

We have introduced a new framework for modeling structural uncertainties. In particular, the method goes beyond the perturbation of deterministic 3D structural interpretations, but considers the uncertainty about connectivities between structural interfaces (Fig. 8). We expect this method to provide a basis for further advances in subsurface uncertainty management, including:

- Taking into account additional data, e.g. field and seismic data, size distribution, slip information. In the case seismic data is available, the global vertical displacement is known for the fault zone. The sum of the simulated displacement for all the faults present laterally in the zone must be equal to the displacement known from seismic data. Some work has also be done to underline the correlation between the displacement of a fault and its size [30, 31]. Such correlations could be used either to simulate the displacement or laterally dying and synsedimentary faults, depending on the available data.

- How to model particular fault zones? Using our method, realistic fault arrays are obtained but not particular configurations. One may want to simulate, e.g. a fault relay zone, for there are evidence such a structure occurs in the studied area although uncertainties remain. For instance, the probability of a breaching fault to connect two segments of a relay zone must only be defined for the block in between the two segments. It means that some blocks cannot be selected during the simulation, i.e. the binary tree structure is constrained.

- Flower structures between two faults is another challenge since in this case, each fault is both parent and child of the other fault, corresponding to cycles in the tree. To solve this problem, the geometry of a flower structure may be obtained by considering one main fault depending on the sliding direction and an inactive lentil bounded to the sliding surface [32].

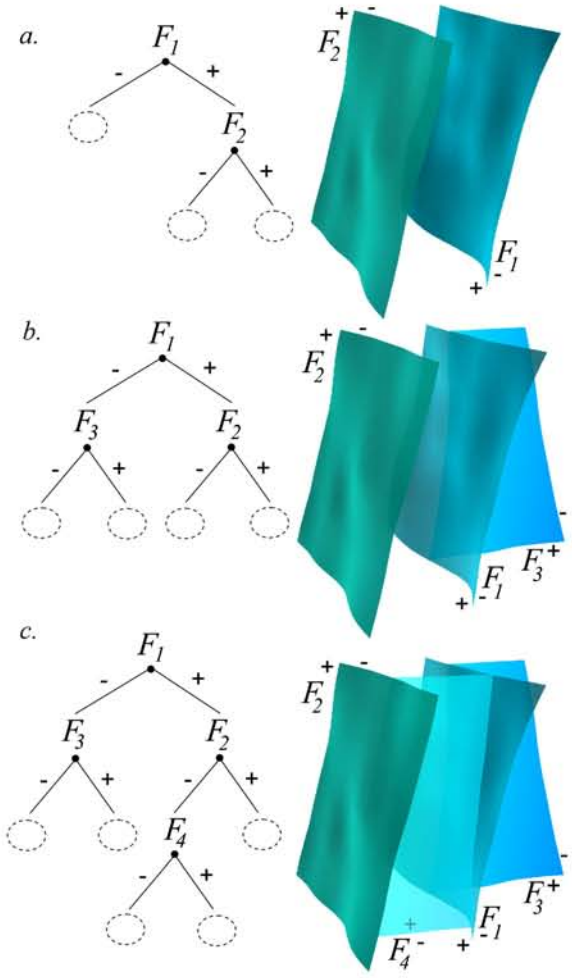

Figure 7: Steps for generating a fault network. *a*. Two faults belonging to the same family and the corresponding tree have been simulated. Two faults belonging to another fault family remain to simulate. *b*. One leaf of the tree is drawn depending on the random function used (e.g. uniform, balanced by block's volume). Once a block is selected, the algorithm generates a field for the next fault and perturbs it using the method presented in Fig.3. Then, an isovalue is drawn among the range of values present in the selected block. *c*. The same process is repeated to add another fault.

Fig. 7. Étapes de la génération d'un réseau de failles. *a*. Deux failles appartenant à la même famille ainsi que l'arbre binaire correspondant ont déjà été simulés. Deux failles appartenant à une autre famille doivent encore être ajoutées. *b*. Un bloc est tiré selon une fonction aléatoire (ex., uniforme, proportionnelle au volume des blocs). Une fois le bloc sélectionné, un champ scalaire est simulé pour la prochaine faille en utilisant la méthode présentée en Fig.3. Une isovaleur est ensuite tirée parmi les valeurs présentes dans le bloc. *c*. Le même procédé est répété afin d'ajouter une faille supplémentaire.

The method takes advantage of implicit modeling for both topology description and changes. Each fault is fully described by a monotonic volumetric function $\mathcal{F}'(x, y, z)$ corresponding to the sum of a reference field $\mathcal{F}(x, y, z)$, computed from statistical or hard data, and a spatially correlated random field $\mathcal{R}(x, y, z)$ corresponding to the associated uncertainties. We believe this implicit representation provides a convenient and robust way of guaranteeing the model consistency.

A binary tree has been introduced to describe the spatial relationships between faults. The tree should be considered as

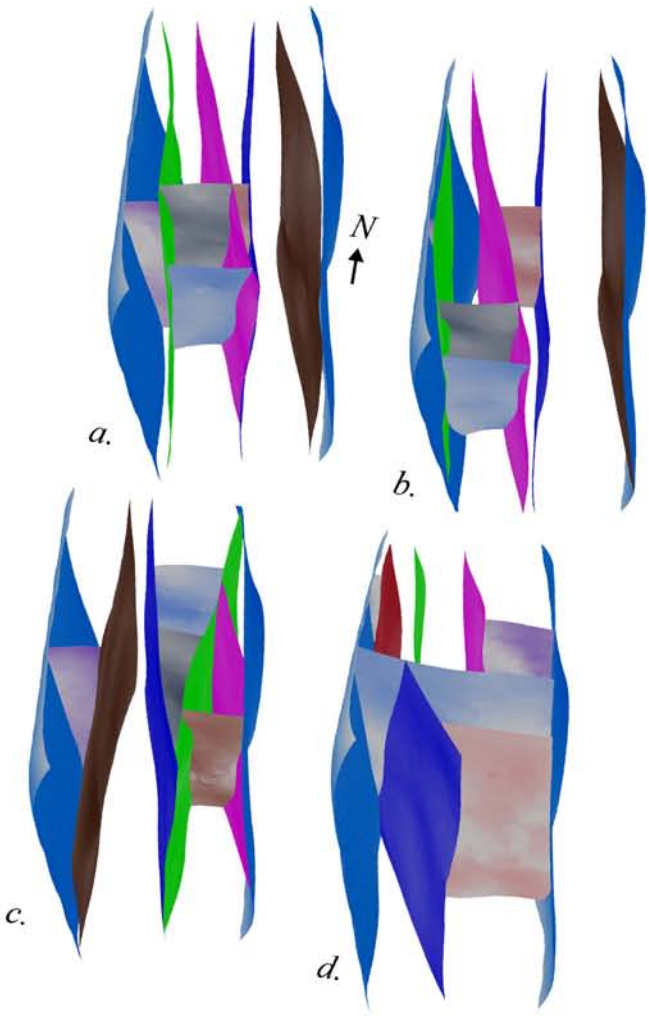

Figure 8: Simulation results illustring different levels of structural uncertainties in a vertical fault zone. $a - b$: low uncertainty, only the geometry of fault surfaces is uncertain, the connectivity is fixed. $c - d$: large uncertainty, both the geometry and the connectivity are uncertain. In c, the fault family oriented NS is the main family, the second fault family oriented WS is branching on the first family. In d, even the faulting events are uncertain, fault families are considered as cogenetic families.

Fig. 8. Résultats de simulation illustrant différents niveaux d'incertitude dans une zone de faille verticale. $a - b$: faible incertitude, seule la géométrie des failles est incertaine, la connectivité est fixe. $c - d$ incertitude importante, la géométrie ainsi que la connectivité des failles sont incertaines. Pour le cas c, la famille de faille orientée Est-Ouest est secondaire par rapport à la famille de faille orientée Nord-Sud. Pour le cas d, l'ordre chronologique des deux évènements de fracturation est incertain, ce qui revient à considérer les deux familles de faille cogénétiques.

a descriptor of fault events, a fault being a branching fault for its ascending branch in the tree. During the fault simulation, faults are added in the tree according to their relative age in order to obtain a fault network honoring input structural constraints. Fault arrays with different topologies are obtained by simulating different binary trees.

Such a multi-realization approach enables to assess the inherent uncertainties to geological data sets. Existing methods are mostly used in the oil industry to reduce economical risks. However, we believe it has a strong potential in other application fields such as potential field interpretation or seismic waveform inversion.

Acknowledgements

This research was performed in the frame of the gOcad research project. The companies and universities members of the gOcad consortium are acknowledged for their support, as well as Paradigm Geophysical for also providing the Gocad software and API. This is CRPG contribution 2020.

References

[1] J.-P. Chilès, P. Delfiner, Geostatistics: Modeling Spatial Uncertainty, Series in Probability and Statistics, John Wiley and Sons, 1999, 696p.

[2] P. Goovaerts, Geostatistics for natural resources evaluation, Applied Geostatistics, Oxford University Press, New York, NY, 1997, 483p.

[3] N. Remy, A. Boucher, J. Wu, Applied Geostatistics with SGeMS: A User's Guide, Cambridge University Press, 2008, 284p.

[4] C. A. Guzofski, J. P. Mueller, J. H. Shaw, P. Muron, D. A. Medwedeff, F. Bilotti, C. Rivero, Insights into the mechanisms of fault-related folding provided by volumetric structural restorations using spatially varying mechanical constraints, AAPG Bulletin 93 (4) (2009) 479–502.

[5] A. Paluszny, S. K. Matthai, M. Hohmeyer, Hybrid finite elementfinite volume discretization of complex geologic structures and a new simulation workflow demonstrated on fractured rocks, Geofluids 7 (2) (2007) 186–208.

[6] J. F. Thompson, B. K. Soni, N. P. Weatherill, Hand Book of Grid Generation, CRC Press, New York, 1999.

[7] R. Frodeman, Geological reasoning: geology as an interpretive and historical science, Geological Society of America Bulletin 107 (8) (1995) 960–968.

[8] C. Bond, A. Gibbs, Z. Shipton, S. Jones, What do you think this is? "conceptual uncertainty" in geoscience interpretation, GSA Today 17 (11) (2007) 4–10.

[9] G. K. Gilbert, The inculcation of scientific method by example, American Journal of Science 31 (1886) 284–299.

[10] T. C. Chamberlin, The method of multiple working hypotheses, Science 15 (1890) 92–96.

[11] P. Samson, O. Dubrule, N. Euler, Quantifying the impact of structural uncertainties on gross-rock volume estimates, in: NPF/SPE European 3D Reservoir Modelling Conference (SPE 35535), 1996, pp. 381–392.

[12] T. Manzocchi, A. E. Heath, B. Palananthakumar, C. Childs, J. J. Walsh, Faults in conventional flow simulation models: a consideration of representational assumptions and geological uncertainties, Petroleum Geoscience 14 (1) (2008) 91–110.

[13] G. Vincent, B. Corre, P. Thore, Managing structural uncertainty in a mature field for optimal well placement, in: SPE Reservoir Evaluation & Engineering, Vol. 2, 1999, pp. 377–384.

[14] S. Suzuki, G. Caumon, J. Caers, Dynamic data integration for structural modeling: model screening approach using a distance-based model parameterization, Computational Geosciences 12 (2008) 105–119.

[15] A. Tarantola, Popper, bayes and the inverse problem, Nature Physics 2 (2006) 492–494.

[16] M. Lecour, R. Cognot, I. Duvinage, P. Thore, J.-C. Dulac, Modeling of stochastic faults and fault networks in a structural uncertainty study, Petroleum Geoscience 7 (2001) S31–S42.

[17] R. M. Srivastava, R. Froidevaux, Probability field simulation: A retrospective, in: Geostatistics Banff 2004, Springer, 2004, pp. 55–64.

[18] G. Caumon, A.-L. Tertois, L. Zhang, Elements for stochastic structural perturbation of stratigraphic models, in: Proc. Petroleum Geostatistics, EAGE, 2007.

[19] T. Charles, J. M. Guéméné, B. Corre, G. Vincent, O. Dubrule, Experience with the quantification of subsurface uncertainties, Paper presented at SPE Asia Pacific Oil and Gas Conference and Exhibition, Jakarta, Indonesia, SPE 68703, 17–19 April.

[20] P. Thore, A. Shtuka, M. Lecour, T. Ait-Ettajer, R. Cognot, Structural uncertainties: determination, management and applications, Geophysics 67 (3) (2002) 840–852.

[21] L. Holden, P. Mostad, B. F. Nielsen, J. Gjerde, C. Townsend, S. Ottesen, Stochastic structural modeling, Math. Geol. 35 (8) (2003) 899–914.

[22] K. Hollund, P. Mostad, B. F. Nielsen, L. Holden, J. Gjerde, M. G. Contursi, A. J. McCann, C. Townsend, E. Sverdrup, Havana - a fault modeling tool, in: A. G. Koestler, R. Hunsdale (Eds.), Hydrocarbon Seal Quantification. Norwegian Petroleum Society Conference, Stavanger, Norway, Vol. 11 of NPF Special Publication, Elsevier Science, 2002, pp. 157–171.

[23] P. Calcagno, J. Chilès, G. Courrioux, A. Guillen, Geological modelling from field data and geological knowledge: Part i. modelling method coupling 3d potential-field interpolation and geological rules, Physics of the Earth and Planetary Interiors 171 (1-4) (2008) 147 – 157, recent Advances in Computational Geodynamics: Theory, Numerics and Applications.

[24] T. Frank, A.-L. Tertois, J.-L. Mallet, 3d-reconstruction of complex geological interfaces from irregularly distributed and noisy point data, Computers & Geosciences 33 (7) (2007) 932 – 943.

[25] A. Guillen, P. Calcagno, G. Courrioux, A. Joly, P. Ledru, Geological modelling from field data and geological knowledge: Part ii. modelling validation using gravity and magnetic data inversion, Physics of the Earth and Planetary Interiors 171 (1-4) (2008) 158 – 169, recent Advances in Computational Geodynamics: Theory, Numerics and Applications.

[26] J.-L. Mallet, Space-time mathematical framework for sedimentary geology, Mathematical geology 36 (1) (2004) 1–32.

[27] D. Ledez, Modélisation d'objets naturels par formulation implicite, Ph.D. thesis, INPL, Nancy, France (2003).

[28] A.-L. Tertois, J.-L. Mallet, Distance maps and virtual fault blocks in tetrahedral models, 26^{th} Gocad Meeting, Nancy, June (2006).

[29] J.-J. Royer, Conditional integration of a linear function on a tetrahedron, 25^{th} Gocad Meeting, Nancy, June (2005).

[30] J. J. Walsh, J. Watterson, Analysis of the relationship between displacements and dimensions of faults, Journal of Structural Geology 10 (3) (1988) 239–247.

[31] P. A. Gillespie, J. J. Walsh, J. Watterson, Limitations of dimension and displacement data from single faults and the consequences for data analysis and interpretation, Journal of Structural Geology 14 (10) (1992) 1157–1172.

[32] J. J. Walsh, J. Watterson, W. R. Bailey, C. Childs, Fault relays, bends and branch-lines, Journal of Structural Geology 21 (8-9) (1999) 1019 – 1026.

Stochastic simulation of fault networks from 2D seismic lines

Nicolas Cherpeau, Guillaume Caumon, CRPG - Nancy University, and Bruno Lévy, Institut National de Recherche en Informatique et Automatique*

SUMMARY

The structural interpretation of 2D seismic lines and their correlation to deduce a 3D structure, e.g. a fault network, is a hard task for geoscientists. Interpretations may introduce a prior geological knowledge bias and/or human bias (Bond et al., 2007). Consequently, decisions based on one single deterministic model of the subsurface may be highly risky. The goal of our study is to show that the 3D organization of a fault network, especially fault connections, is uncertain and to investigate how these uncertainties evolve depending on the quantity of input data. For this, we realized stochastic simulations of fault networks using a new method that allows both geometry and fault connection changes from statistical parameters about fault geometry and fault traces interpreted on 2D seismic lines. Simulations have been performed for two scenarios in which the quantity of input data was different, and generate a set of possible models whose diversity decreases as more data are added.

INTRODUCTION

Inferring 3D geological structures from a few seismic lines and exploration wells is a big challenge. Bond et al. (2007) show that among 412 geoscientist interpretations of a synthetic seismic image, only 21% of the participants found the right tectonic setting.

The goal of this paper is not to explicitly address such conceptual geological uncertainty, but to focus on topological and geometrical uncertainties about fault networks, under prior knowledge about fault statistical properties. Indeed, fault uncertainty is often first order in structural interpretation and related applications, because faults significantly affect not only gross rock volumes but also fluid flow in the subsurface.

Uncertainty about fault networks goes beyond the geometric uncertainty (Caumon et al., 2007; Charles et al., 2001; Lecour et al., 2001; Thore et al., 2002), because it also concerns fault existence and connectivity. This latter topological uncertainty is significant in 3D modeling from sparse 2D information. In their pioneering work, Hollund et al. (2002) have proposed sampling this uncertainty through multiple realizations on corner-point reservoir grids. Suzuki et al. (2008) consider topological uncertainty by selecting an ensemble of models generated by combining manual interpretations and stochastic model perturbations. The stochastic model used in this paper also generates several alternative models, but is tailored to accurately represent more complex fault networks by using implicit surfaces on a tetrahedral mesh (Cherpeau et al., 2010). Such multi-realization approach can help understanding risks related to structural interpretation, both in underground resource exploration or management and assessment of natural hazards.

This abstract briefly explains the features of this stochastic model, then presents its application to a Middle East field: from the 3D seismic of the area, a few 2D seismic lines have been extracted and interpreted; the stochastic fault simulation is then used and the various ouptut fault networks are compared to the reference model obtained by interpreting the 3D seismic.

STOCHASTIC MODEL OVERVIEW

We hereby present the stochastic model used to generate fault surfaces, based on previous work (Cherpeau et al., 2010). The method is applied and adapted to model laterally dying and synsedimentary faults and to account for fault locations interpreted from available data.

Fault geometry characteristics

Faults to simulate with similar structural parameters are grouped into fault families. Each fault family has a number of faults to simulate and statistical parameters used to define the fault geometry:

- orientation parameters (dip and strike)
- size parameters (extension along strike vector and height along dip vector)
- sinuosity parameters
- a relative age to determine major faults versus posterior branching faults

Data point management

Let *trace* denote an interpreted fault trace corresponding to one or several data points. From one simulation to another, faults from different families may honor *trace*. If additional interpretation is made, *trace* can be restricted to a particular fault family A, i.e. only faults belonging to A can honor *trace*, so that uncertainty is reduced (let $trace_A$ denote such fault trace). Consequently, the set of fault locations $\mathscr{S}$ is composed of two subsets $\mathscr{S}'$ and $\mathscr{S}''$:

- $\mathscr{S}' = \left\{ trace_{i\in[1,n]} \right\}$ the set of unassigned fault traces.
- $\mathscr{S}'' = \left\{ \mathscr{S}'_{j\in[1,m]} \right\}$, m being the number of fault families and $\mathscr{S}'_j = \left\{ trace_{j,k\in[1,p]} \right\}$ the set of fault traces assigned to family j.

Fault object simulation

Fault surfaces are generated sequentially in the domain of interest $\mathscr{M}$. The fault simulation steps, for a fault F from fault family A, are as follows:

1. To randomly draw a fault trace $trace_{initial}$ from the set $\mathscr{S}_F = \mathscr{S}'_A + \mathscr{S}''$.

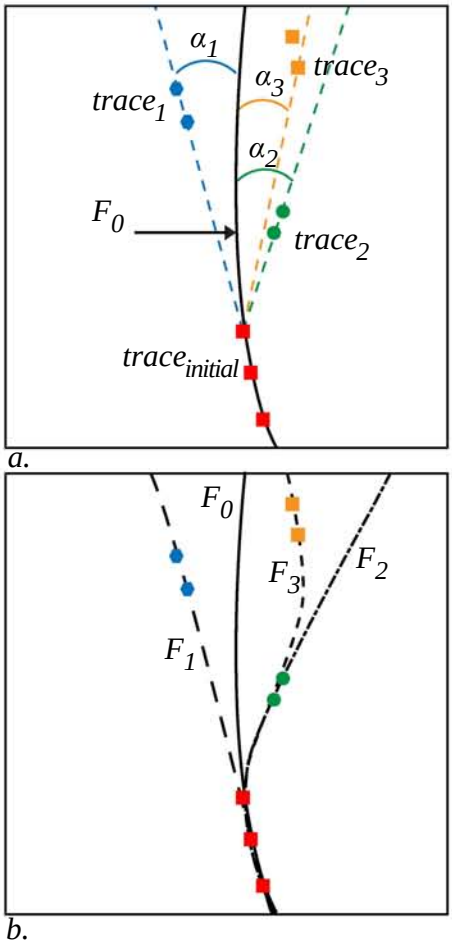

Figure 1: *a*. Initial fault geometry F_0 honoring fault trace $trace_{initial}$. For each fault trace $trace_i$, the maximum angle α_i is computed, determining the probability p_i that $trace_{initial}$ and $trace_i$ belong to the same fault surface. *b*. Example of data clusterings. Fault geometry F_1 corresponds to the cluster [$trace_{initial}$, $trace_1$], F_2 to [$trace_{initial}$, $trace_2$], F_3 to [$trace_{initial}$, $trace_2$, $trace_3$]. The cluster [$trace_{initial}$, $trace_1$, $trace_2$] is not possible because α_1 and α_2 are too different.

2. To simulate a dip and a strike from input parameters.
3. To nucleate the fault by randomly drawing a fault center coordinates.
4. To generate an initial fault geometry from the dip and strike, honoring $trace_{initial}$, if any.
5. To eventually cluster $trace_{initial}$ with other fault traces from $\mathscr{S}_F$ to define $trace_{final}$ (see below for details) .
6. To update the fault geometry to honor $trace_{final}$.
7. To perturb the fault geometry from input sinuosity parameters.
8. To simulate the fault extension domain $\mathscr{D} \subseteq \mathscr{M}$, from size-related parameters.
9. To define the fault existence domain $\mathscr{D}' \subseteq \mathscr{D}$, by looking for truncation with previously simulated faults. A Binary Space Partition Tree (BSP Tree) of $\mathscr{M}$ is used to speed up the process.
10. To update the BSP tree.

Fault trace clustering

When a fault trace $trace_{initial}$ is drawn for a fault F from family A, an initial fault geometry is simulated from the dip and strike parameters so that the fault surface honors $trace_{initial}$ in a given uncertainty range. Then, for each fault trace $trace_i$ belonging to $\mathscr{S}_F$, a probablity p_i that $trace_i$ and $trace_{initial}$ belong to the same fault surface is computed. The probability p_i depends on the geometry of F and the location of $trace_i$, i.e. the smaller is the angle α_i in figure 1*a*, the higher is p_i. When a fault trace $trace_1$ with angle α_1 is drawn, only fault traces with an angle similar to α_1 are considered to avoid inconsistent fault geometries. Consequently, the final fault trace $trace_{final}$ associated to F may be composed of several fault traces: $trace_{final} = \left\{trace_{initial}, trace_{i\in[1,n]}\right\}$ (figure 1*b*). Other elements such as the distance between $trace_{initial}$ and $trace_i$ or the fault displacement, could also come into play when computing this probability.

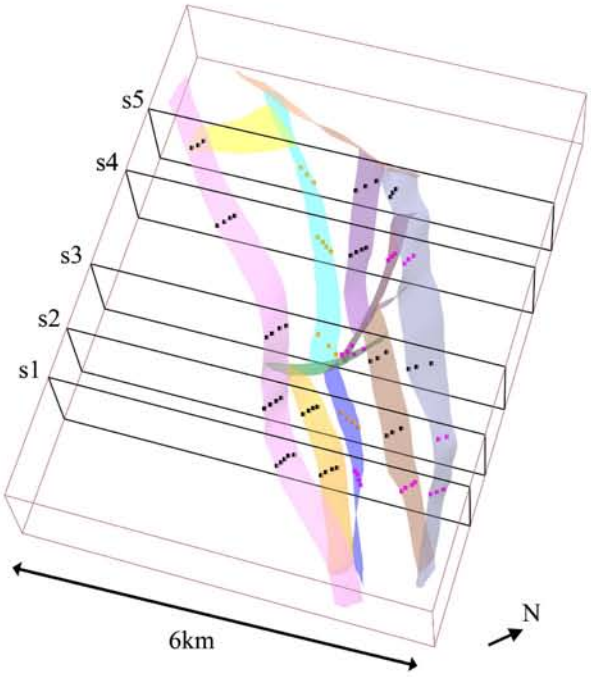

Figure 2: The studied area is about 6x8x1km. The deterministic fault network and 5 fictive seismic lines are shown. Fault traces have been classified according to their orientation. SW dipping traces are set to belong to the same fault family. NE dipping traces are set to belong to another fault family. Other fault traces are not classified, hence can correspond to a fault from any fault family.

The algorithm may stop when all data points have been processed or may continue until some target number of faults have been simulated. New simulated faults are branching faults for previously simulated ones. By repeating the process, fault

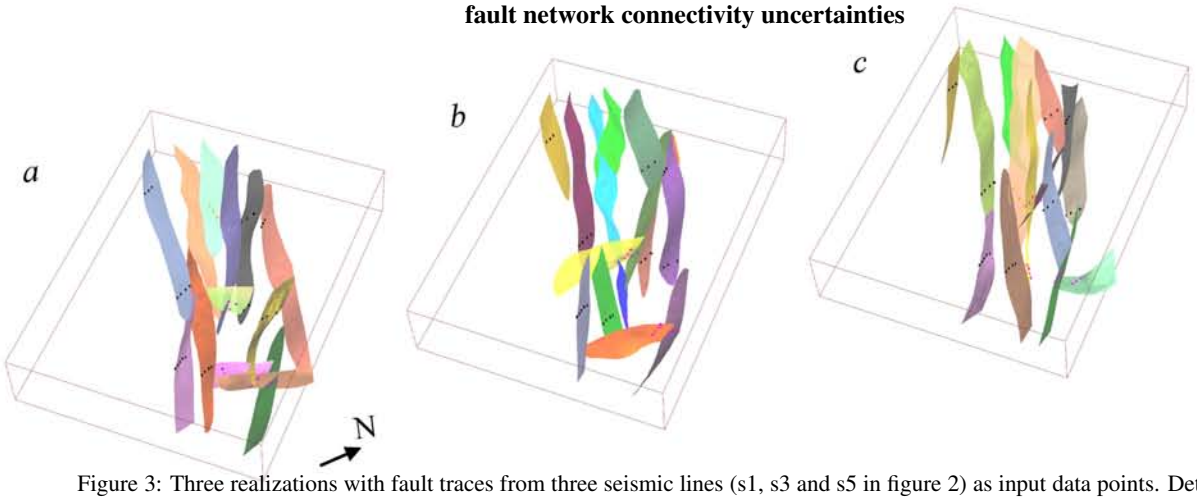

Figure 3: Three realizations with fault traces from three seismic lines (s1, s3 and s5 in figure 2) as input data points. Details in the text.

traces are drawn in a different order and clustered in different ways, thus the method generates fault networks with different fault geometries and fault connections.

INPUT DATA

The studied area is about 6x8x1km. A deterministic fault network has been interpreted from a full 3D seismic survey (figure 2). To emulate an exploration context with sparse data, we interpreted only five SW-NE seismic lines extracted from the 3D data set, and used the 3D interpretation as the deterministic reference to discuss stochastic simulation results.

Fault network characterization

From the seismic lines, faults are grouped into 3 fault families:

- the first one, $family_1$, is oriented [N110-N130] and dipping [60-70]. Most traces seem oriented along that directions (black data points in figure 2).
- the second one, $family_2$, correspond to opposed-dipping faults to $family_1$ and thus is oriented [N290-N310] and dips [60-70] (orange data points in figure 2).
- the third fault family, $family_3$, corresponds to transversal faults to the fault zone and is oriented [N190-N210] and dips [60-70].

As we cannot deduce the relative age of the three fault families from the seismic lines, families are considered as cogenetic, i.e. there is no systematic truncation rule between faults from different fault families.

Displacement information

The displacements mesured on seismic lines fall into two groups: one group has a displacement of about 160m and the other 60m. The relationship between the maximum displacement of a fault d_{max} and the maximum extension of the fault surface L is generally assumed to be correlated through a power law (Kim and Sanderson, 2005):

$$d_{max} = cL^n | n \in [0.5, 2] \quad (1)$$

The factor c corresponds to the displacement at unit length (more details can be found in Kim and Sanderson (2005)). The exponent n is generaly assumed to be equal to 1. From figure 5a in Kim and Sanderson (2005), $c \cong 10^{-1.7}$ for normal faults.

From equation 1, with $n = 1$ and $c = 10^{-1.7}$, a displacement of 60m corresponds to a maximum fault size of about 3km, a displacement of 160m to a maximum fault size of 8km. The maximum displacement is maybe not reached on the seismic lines and thus the maximum fault size may be underestimated.

Data point interpretation

Black data points in figure 2 are clearly dipping SW and thus are restricted to $family_1$, i.e only faults belonging to $family_1$ can go through these points. Similarly, orange data points are dipping SE and thus correspond to faults of $family_2$. Other data points may belong to any fault family.

RESULTS

To show the evolution of the fault network organization depending on the data avalaible, simulations have been performed according to two scenarios with different input data.

Scenario 1

In this scenario, input data is limited to seismic lines s1, s3 and s5 (figure 2). According to the displacement information, the maximum fault size is taken to be in the range $[2000, 9000]m$.

The three realizations shown in figure 3 respect the global layout of the deterministic fault network $network_{ref}$ (figure 4f). Major bording faults present in $network_{ref}$ are missing in the realizations because of a too small fault input size. However, all three realizations present at least one major structure going through the studied area, composed of several connected and aligned faults. This major structure corresponds to the major southern bording fault only for realization a in figure 3. A few data points are part of a unique fault, maybe due to the large interval between seismic lines as compared to the fault input size.

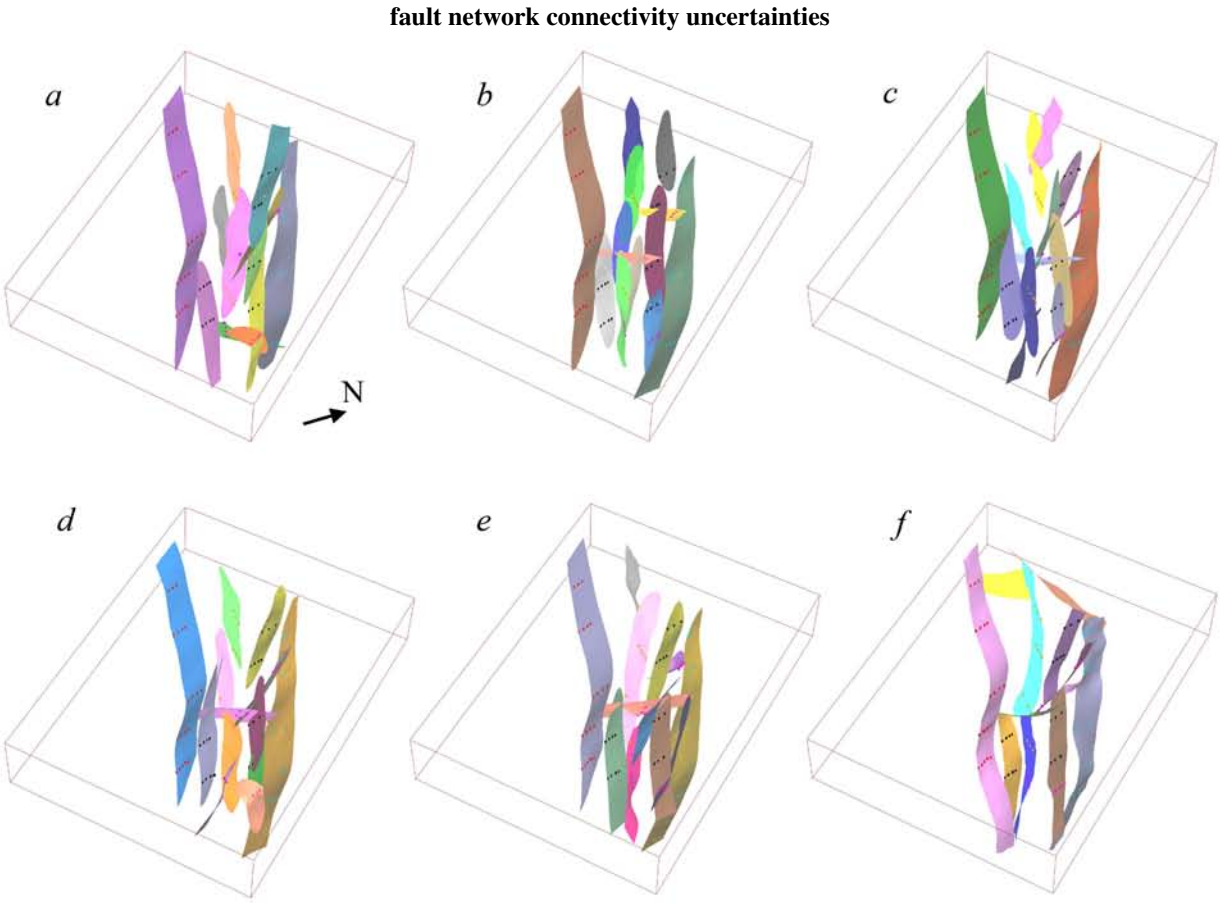

Figure 4: $a-e$: five realizations with fault traces from five seismic lines (figure 2) as input data points. Fault traces are also regrouped for the two bording faults using the displacement information. f: the reference fault network.

Scenario 2
As compared to scenario 1, we added seimic lines s2 and s4 (figure 2). We also used the diplacement information to cluster fault traces where a high displacement is observed to belong to two different faults, one in the southern part and the other in the northern part of the studied area. The maximum fault size for the rest of the faults is decreased to be in the range $[2000, 4000]m$.

The two major bording faults are present in the realizations due to the fault traces clustering (figure 4). Realizations are globally closer to the deterministic fault network built from the 3D data set than realizations from scenario 1.

CONCLUSIONS

We have proposed to sample both geometrical and topological uncertainties of fault networks. The method accounts for prior geological knowledge such as fault relative ages, fault displacement, and is conditional to partial interpretations. The implementation of the method takes advantage of implicit fault representation and uses a BSP Tree to efficiently generate models.

The stochastic fault method was applied using two scenarios with different quantity of 2D input data. As more input data is available, the uncertainty space reduces, i.e. results are closer to the deterministic model built from a 3D data set.

In the presence of 3D seismic data, we suggest our simulation method would also be useful to account for uncertainty due to migration errors and limited seismic resolution. Indeed, ambiguous data interpretation by different geoscientists may lead to different understandings and decisions for the same subsurface area.

Extensions of this method are currently under way to also use an uncertainty model for horizons to provide fully stochastic models of structural frameworks.

ACKNOWLEDGMENTS

The authors would like to thank the companies and universities members of the gOcad research consortium for funding this work, especially Paradigm Geophysical for providing the Gocad software and API, and Total for providing the data.

REFERENCES

Bond, C., A. Gibbs, Z. Shipton, and S. Jones, 2007, What do you think this is? "conceptual uncertainty" in geoscience interpretation: GSA Today, **17**, 4–10.

Caumon, G., A.-L. Tertois, and L. Zhang, 2007, Elements for stochastic structural perturbation of stratigraphic models: Presented at the Proc. Petroleum Geostatistics, EAGE.

Charles, T., J. M. Guéméné, B. Corre, G. Vincent, and O. Dubrule, 2001, Experience with the quantification of subsurface uncertainties: Paper presented at SPE Asia Pacific Oil and Gas Conference and Exhibition, Jakarta, Indonesia, SPE 68703, 17–19 April.

Cherpeau, N., G. Caumon, and B. Lévy, 2010, Stochastic simulations of fault networks in 3D structural modeling: Comptes Rendus Geoscience, **342**, 687 – 694.

Hollund, K., P. Mostad, B. F. Nielsen, L. Holden, J. Gjerde, M. G. Contursi, A. J. McCann, C. Townsend, and E. Sverdrup, 2002, Havana - a fault modeling tool: Hydrocarbon Seal Quantification. Norwegian Petroleum Society Conference, Stavanger, Norway, Elsevier Science, 157–171.

Kim, Y.-S., and D. J. Sanderson, 2005, The relationship between displacement and length of faults: a review: Earth-Science Reviews, **68**, 317 – 334.

Lecour, M., R. Cognot, I. Duvinage, P. Thore, and J.-C. Dulac, 2001, Modeling of stochastic faults and fault networks in a structural uncertainty study: Petroleum Geoscience, **7**, S31–S42.

Suzuki, S., G. Caumon, and J. Caers, 2008, Dynamic data integration for structural modeling: model screening approach using a distance-based model parameterization: Computational Geosciences, **12**, no. 1, 105–119.

Thore, P., A. Shtuka, M. Lecour, T. Ait-Ettajer, and R. Cognot, 2002, Structural uncertainties: determination, management and applications: Geophysics, **67**, 840–852.

Assessing the impact of fault connectivity uncertainty in reservoir studies using explicit discretization

Nicolas Cherpeau, Guillaume Caumon, SPE, CRPG - CNRS, Université de Lorraine-ENSG, rue du Doyen Marcel Roubault, 54501 Vandoeuvre-lès-Nancy, France; Jef Caers, Department of Energy Resources Engineering, Stanford university, Stanford, CA 94305, USA; and Bruno Lévy, Centre INRIA Nancy Grand-Est, Campus scientifique, 615 rue du Jardin Botanique, 54600 Villers les Nancy, France

Abstract

Reservoir production curves provide significant information about reservoir compartmentalization and connectivity, but are most often used at late stages of reservoir modeling workflows, leading to possible inconsistencies between static and dynamic reservoir models. In this paper, we propose to use this information during structural modeling to reduce fault-related uncertainties induced by subsurface imaging and interpretation ambiguities.

We first generate stochastic fault networks from prior information about faults hierarchy, size, orientation, roughness and localization. Then, the results are discretized into a set of nodes and pipes, called a pipe network, which represents the control volumes connectivity graph. This discretization data structure accurately represents fault surfaces and fault connections while limiting the number of nodes to represent the domain of interest as a connectivity graph that can be fed to a flow simulator. In a third step, we select models that best match the production curves.

The proposed method is flexible enough to assess high uncertainties that may require to add or remove faults, or to perturb the fault connectivity, without need for simplification of the fault network into a corner-point reservoir grid. The method is applied to a synthetic example of high uncertainties relative to a poorly imaged 3D fault zone.

Introduction

Faults play a key role in subsurface fluid flow, *e.g.* by modifying the spatial repartition of geological units, by perturbing the permeability field at the fault rock, forming flow conduits or barriers. Consequently, assessing fault-related uncertainties in reservoir studies is very important to history match to the available data. Most history matching strategies focus on the reservoir grid properties. When dealing with faults, methods are defined to address fault-related uncertainties concerning fault rock properties [Manzocchi *et al.* (2008)]. However, in complex and poorly imaged areas, the history matching problem may require more flexible methods to address the uncertainty of sub-seismic faulting and fault connectivity. Indeed these cases require modifying the topology of the reservoir grid which can be a very time consuming process. Holden *et al.* (2003) have proposed a fault model to account for fault connectivity changes directly in the corner-point reservoir grid. Their method is interesting but can only simulate pillar-aligned faults which significantly limits the range of uncertainty which can be addressed. Other works have been done to update structural parameters, such as top and bottom reservoir boundaries [Seiler *et al.* (2010)], fracture networks [Jenni *et al.* (2004)] or to handle fault displacement uncertainty [Røe *et al.* (2010)].

We propose here another method to incorporate a stochastic fault model in the general reservoir workflow in order to assess fault connectivity uncertainties. The fault stochastic model is based on recent work [Cherpeau *et al.* (2010a), Cherpeau *et al.* (2010b)] and enables to generate realistic fault networks in a tetrahedral mesh, using implicit modeling [Calcagno *et al.* (2008), Frank *et al.* (2007)]. We propose to combine the discretization of these implicit fault surfaces with matrix rock into a set of nodes and pipes, called a pipe network [Vitel (2007)]. The results, i.e. the pipe network, can then be used as an input for flow simulators accepting unstructured grids. This method enables to preserve structural accuracy while keeping an acceptable number of nodes to represent the domain of interest.

The method is illustrated by a synthetic example, showing how the fault connectivity in a fault zone impacts the reservoir flow response.

1 – Stochastic fault model overview

The stochastic fault model is described by [Cherpeau *et al.* (2010a), Cherpeau *et al.* (2010b)]. Faults to simulate with similar structural parameters are grouped into fault families. Each fault family is defined by statistical parameters describing the fault geometry:

- Dip and strike distributions
- Extension along the strike and dip directions
- Roughness parameters
- A relative age to determine major faults versus posterior branching faults

Faults are generated sequentially in the domain of interest according to the input chronological order representing the timing of faulting events, *e.g.* an N-S extension then a W-E compression. Each fault object is defined by six parameters (dip, strike, fault center, roughness, height, length) that are successively randomly drawn from input parameters. The simulation process can also be conditional to interpreted fault data [Cherpeau *et al.* (2010b)]. In this case, an extra simulation step randomly assigns data points to faults to simulate.

Each realization is a set of implicit fault surfaces defined in a tetrahedral mesh representing the zone of interest. Fault connections result from the simulation process, i.e. faults generated lately are considered to be younger than faults generated first and thus are branching faults in case of intersection.

Only one tetrahedral mesh is required to run simulations and is used as support by the discretization process because faults are not explicitely represented into the mesh.

2 – Discretization procedure

Discretization data structure

The data structure used for the flow simulations (pipe network) [Vitel (2007)] is composed of:

- A set of nodes, each one representing a control volume and storing petrophysical properties.
- A set of pipes connecting the nodes, used to store the transmissibility between the nodes.

Each realization is discretized into a pipe network composed of three types of nodes:

- Nodes corresponding to fault surfaces storing fault properties (*e.g.* transmissibility multiplier, SGR).
- Nodes corresponding to unfaulted elements:
  - Nodes corresponding to polyhedra of faulted tetrahedra.
  - Nodes corresponding to unfaulted tetrahedra.

The discretization process is run in three main steps:

1. A set of connected geometrical primitives corresponding to both faults and matrix elements is computed.
2. Nodes are created from the barycenter of the geometrical primitives.
3. Pipes are created between connected faces of different elements.

Tetrahedron position

Each fault F is defined by an isovalue f of a scalar field $F(x,y,z)$ stored on the tetrahedron vertices. The fault extension is limited into a 3D domain D_F defined by a scalar field $E_F(x,y,z)$ and isovalue e_f. For instance, the domain of fault existence D_F may correspond to a 3D ellipsoid computed from fault height and length. Consequently, a fault F is defined by:

$$F(x,y,z) = f \mid \forall (x,y,z) \in D_F : E_F(x,y,z) \leq e_F$$

For a given tetrahedron T, three configurations exist:

1. Non-interstected: T is not intersected by F or T is outside D_F: $T \cap F = \phi$ or $T \cap D_F = \phi$
2. Fully intersected: T is intersected by F and T is included in D_F: $T \cap F \neq \phi$ and $T \subset D_F$
3. Partly intersected: T is intersected by F and T is partly in D_F: $T \cap F \neq \phi$ and $T \cap D_F \neq \phi$ and $T \cap D_F \neq T$

Fault discretization steps

The algorithm discretizes faults according to the simulation order so that oldest faults are discretized first. For each fault, the algorithm proceeds as follows:

1. Find both fully and partly intersected tetrahedra and split them into two sets:
   - S_{uncut}: tetrahedra that have not been cut by other faults yet.
   - S_{cut}: tetrahedra already cut by other faults, *i.e.* composed of several polyhedra (Figure 1b).
2. Cut the set of tetrahedra S_{uncut}.
3. For each tetrahedron in S_{cut}, select polyhedra to cut according to fault connections, then perform a geometrical cut (Figure 1c and Figure 1d).
4. Connect newly created polyhedra.

The geometrical cut of a tetrahedron from S_{uncut} creates two matrix polyhedra connected to the fault polygon (Figure 1a and Figure 1b). Each polyhedron is a set of faces and faces know their face mate belonging to a neighbor polyhedron.

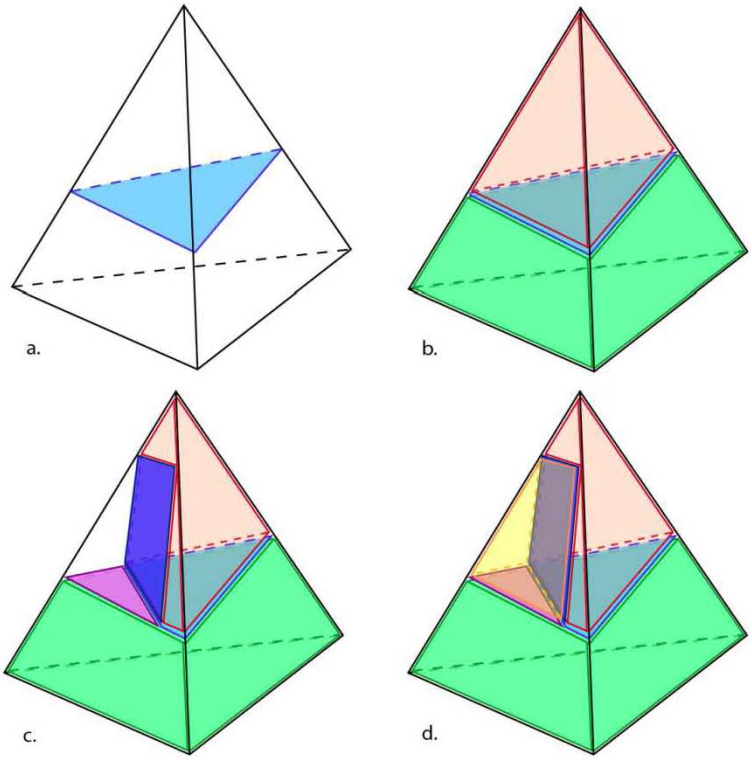

Figure 1: Geometrical cut of a tetrahedron. a. A fault polygon is computed from the fault scalar field values. b. Two polyhedra are created on both sides of the fault polygon. c. A branching fault only cuts the red polyhedron and splits the existing fault polygon into two parts. d. After the cut, the tetrahedron contains three polyhedra and three fault polygons.

If a polyhedron is partly intersected by a fault, an extra cut is performed to respect fault extension. In this case, the two resulting polyhedra on both sides of the fault polygon are connected through a common face (Figure 2).

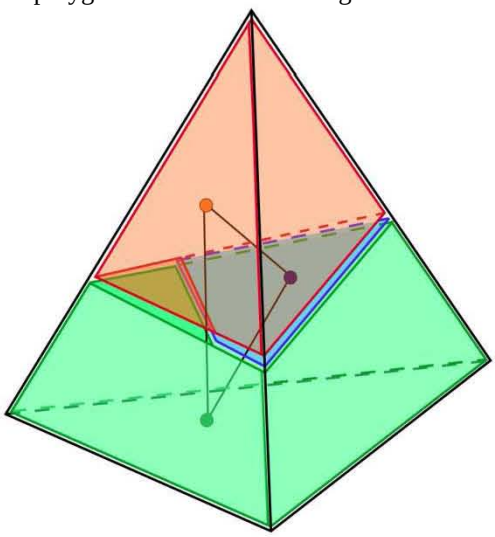

Figure 2: Partial cut of a tetrahedron. The fault polygon (blue) does not fully intersect the tetrahedron so that the two resulting matrix polyhedra are connected.

The first discretization step is finished when both unfaulted and faulted tetrahedra are discretized into a set of connected geometrical primitives.

Nodes computation

Given a set of connected geometrical primitives, the algorithm first tries to merge too small elements in order to avoid numerical issues during flow simulation. Nodes are then computed from the barycenter of each geometrical primitive. Each node stores the control volume of its polyhedron as well as the porosity and permeability. At this step, fault rock property can be modeled by assigning permeability values different from the matrix to fault nodes.

Pipes computation

Each polyhedron is a set of faces knowing their face mate thus the neighbor polyhedron, if any. Consequently, a pipe is created between two nodes of neighbor polyhedra. The transmissibility T_{ij} between polyhedra i and j is computed as follows:

$T_{ij} = \frac{T_i \times T_j}{T_i + T_j}$, $T_i = \frac{A_i \times k_i}{\|\overrightarrow{cc_i}\|} \boldsymbol{n}_i . \boldsymbol{g}_i$ with A_i the area of the common face and k_i the permeability of element *i* (Figure 3).

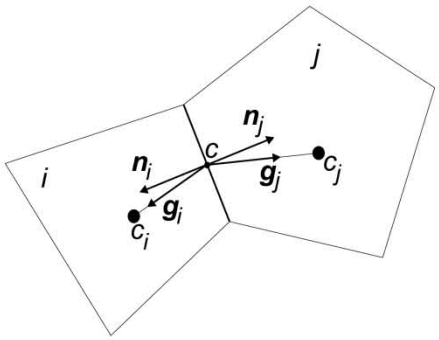

Figure 3: transmissibility computation between two adjacent polyhedra.

A star-polygon transformation [Karimi-Fard *et al.* (2004)], based on resistor network analogy, is used to avoid small elements at intersections between two faults.

The final number of nodes is related to the mesh resolution but is larger than the number of tetrahedra and depends on the number of faults, fault size and fault network complexity (Figure 4).

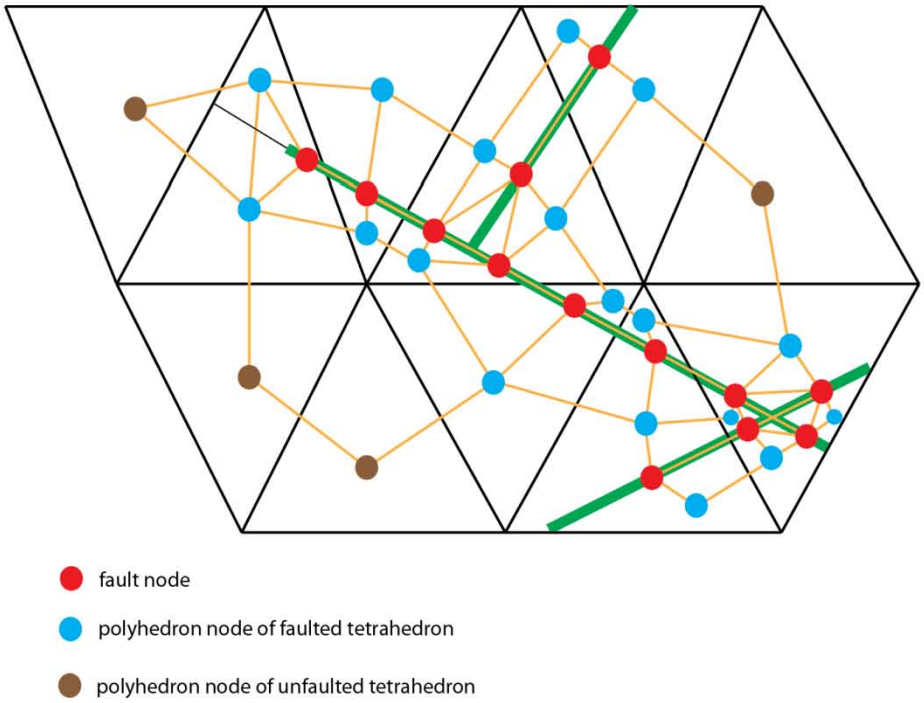

Figure 4: 2D view of a pipe network. A star-polygon transformation avoids small elements at fault intersections.

The discretization scheme relies on tetrahedra wich do not respect orthogonality, thus may entail approximations during the fluid flow simulation. However, the primary goal of the method is to assess first order uncertainties about faults by rejecting unlikely structural models, which does not require an exact discretization procedure. Other discretization strategy could be used for better accuracy, *e.g.* MPFA methods could be used to overcome the non-orthogonality issues.

3 – Proposed workflow

The proposed worklow (Figure 5) starts by simulating stochastic fault networks. Then, each realization is discretized into a pipe network. Flow simulations are performed at the next step, using a flow simulator accepting unstructured grids. Flow responses can then be used either for model selection and perturbation, or could possibly be used to perturb the initial model parameters [Cherpeau *et al.* (submitted)].

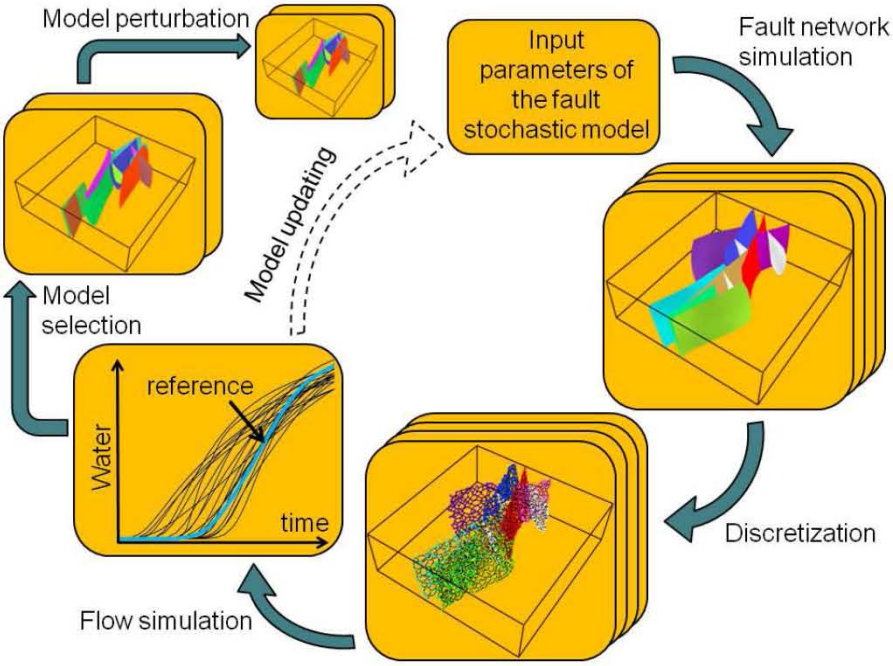

Figure 5: Proposed workflow to assess reservoir connectivity uncertainties. The first step consists in describing the fault network, i.e. to define the model parameters. The next step is the simulation of fault networks. Then, the results are discretized into pipe networks. These pipe networks can then be input into a flow simulator supporting unstructured grids. Finally, comparing the realization flow responses enables to select models close to the observations and gently perturb them to decrease their misfit. A future step of the workflow could be to update the model parameters to include geological structures in the history matching problem [Cherpeau *et al.* (submitted)].

4 – Assessing fault connectivity uncertainty

The goal of the study is to generate fault networks matching production data in a poorly imaged 3D area using the worklow presented in Figure 5. One injector and one producer wells are on both sides of a fault zone which is approximately 300-500m wide, 1400m long in the N-S direction and 300m high in vertical dimension (Figure 6).

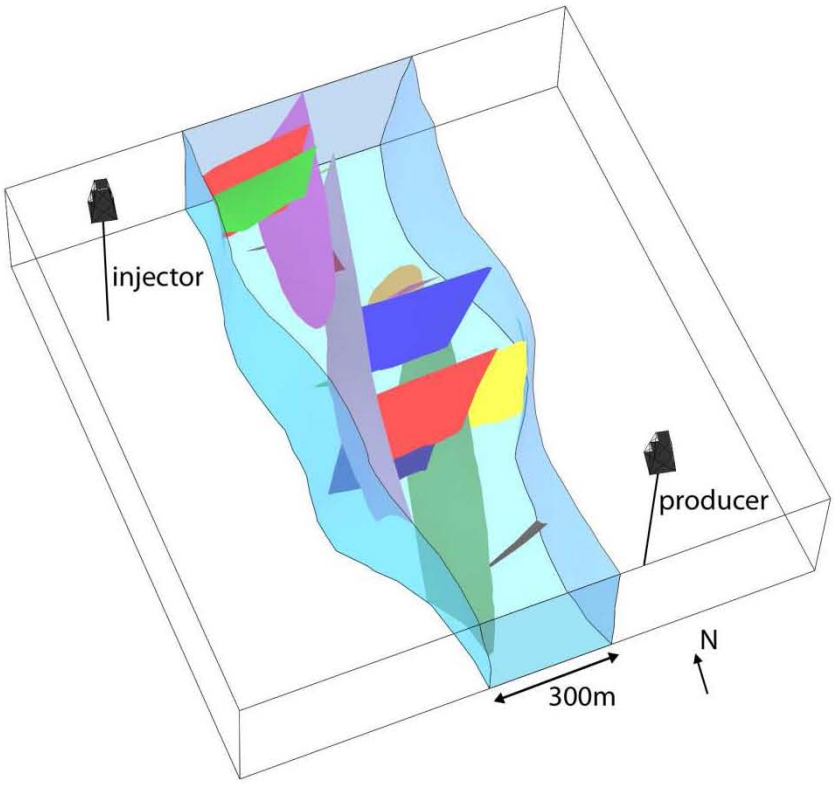

Figure 6: Fault network used as reference. The fault zone is oriented N-S and is about 300-500m wide, 1400m long and 300m high.

Model parameters

Three fault families are defined:

- The first family corresponds to major faults aligned along the fault zone:
  - Number of faults: 5, each fault has a probability of existence equal to 0.9
  - Strike: Uniform, N[170-190]
  - Dip: Gaussian, mean=60^o, σ=3
  - Length: Gaussian, mean=1000m, σ=100m
  - Height: Gaussian, mean=300m, σ=10m
  - Roughness amplitude: 50m
- The two other fault families represent a second tectonic event creating opposed-dipping normal faults:
  - Family 2:
    - Number of faults: 7, probability of existence equals to 0.8
    - Strike: Uniform, N[60-80]
    - Dip: Gaussian, mean=58^o, σ=2
    - Length: Gaussian, mean=600m, σ=50m
    - Height: Gaussian, mean=200m, σ=5m
    - Roughness amplitude: 40m
  - Family 3:
    - Number of faults: 7, probability of existence equals to 0.8
    - Strike: Uniform, N[240-260]
    - Dip: Gaussian, mean=58^o, σ=2
    - Length: Gaussian, mean=600m, σ=50m
    - Height: Gaussian, mean=200m, σ=5m
    - Roughness amplitude: 40m

Results

One hundred realizations are run using the fault parameters defined above. The number of faults ranges from 12 to 19. Realizations are then discretized into pipe networks. The mesh has 6045 tetrahedra and the number of nodes in the pipe network ranges from 7455 to 8712 with a mean of 8169 nodes, which corresponds to approximately one third more nodes than the number of tetrahedra.

Faults are considered as flow barriers with permeability of 0.01mD and matrix permeability is set to 100mD and porosity to 0.2. One realization is randomly drawn to be the reference (Figure 6). Flow simulations have been performed on the one hundred models and the root mean square misfit to water saturation data is computed and used as criterium to rank models. Misfit values range from 10 to 230; the ten best models are selected with misfit values in [10-30]. Selected models have different number of faults, fault connections and fault layout, which emphasizes the importance of fault-related uncertainty assessment methods (Figure 7).

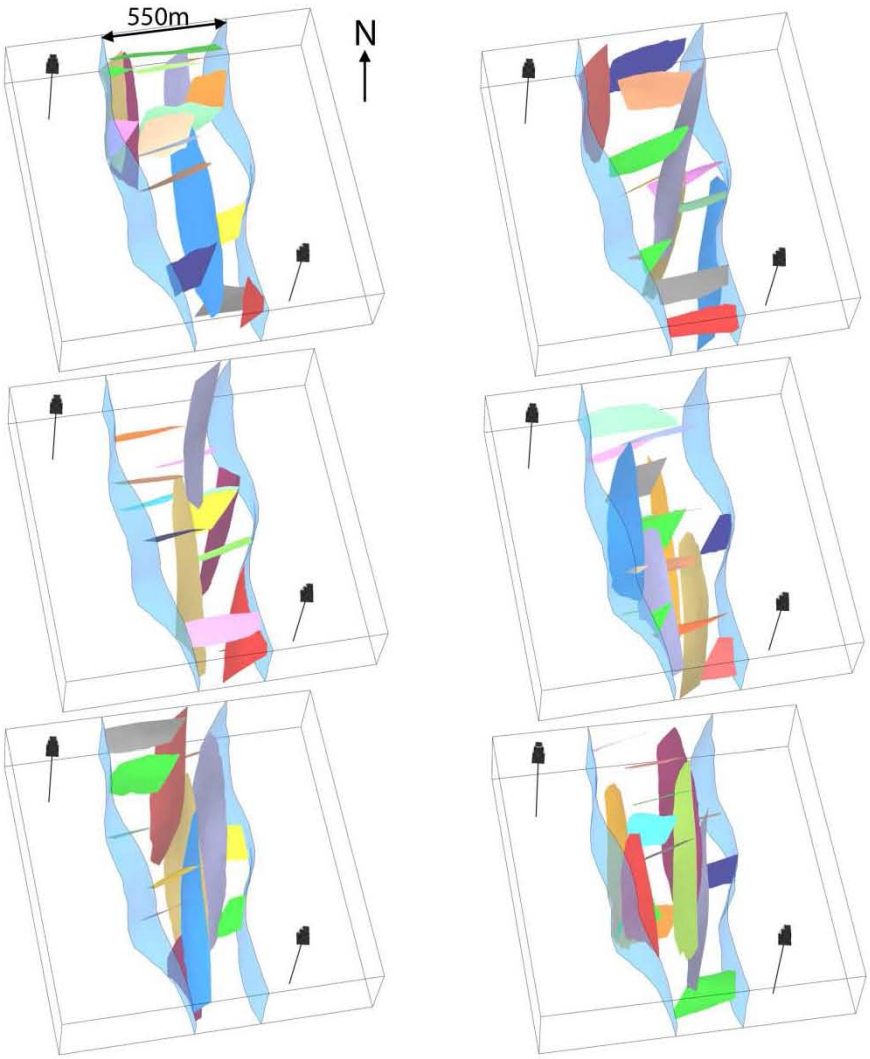

Figure 7: six realizations close to the reference in term of flow response. Models have different number of faults, fault layout and fault connections.

Perspectives

The proposed method enables to sample fault-related uncertainties by generating very different models. Each of these models can then be used for further optimization and calibration, *e.g.* by freezing the topology and perturbing the fault geometry [Mallet and Tertois (2010)] or by perturbing fault parameters using the fault parameterization proposed by Cherpeau *et al.* (submitted) until the misfit to production data is below a given threshold (Figure 8).

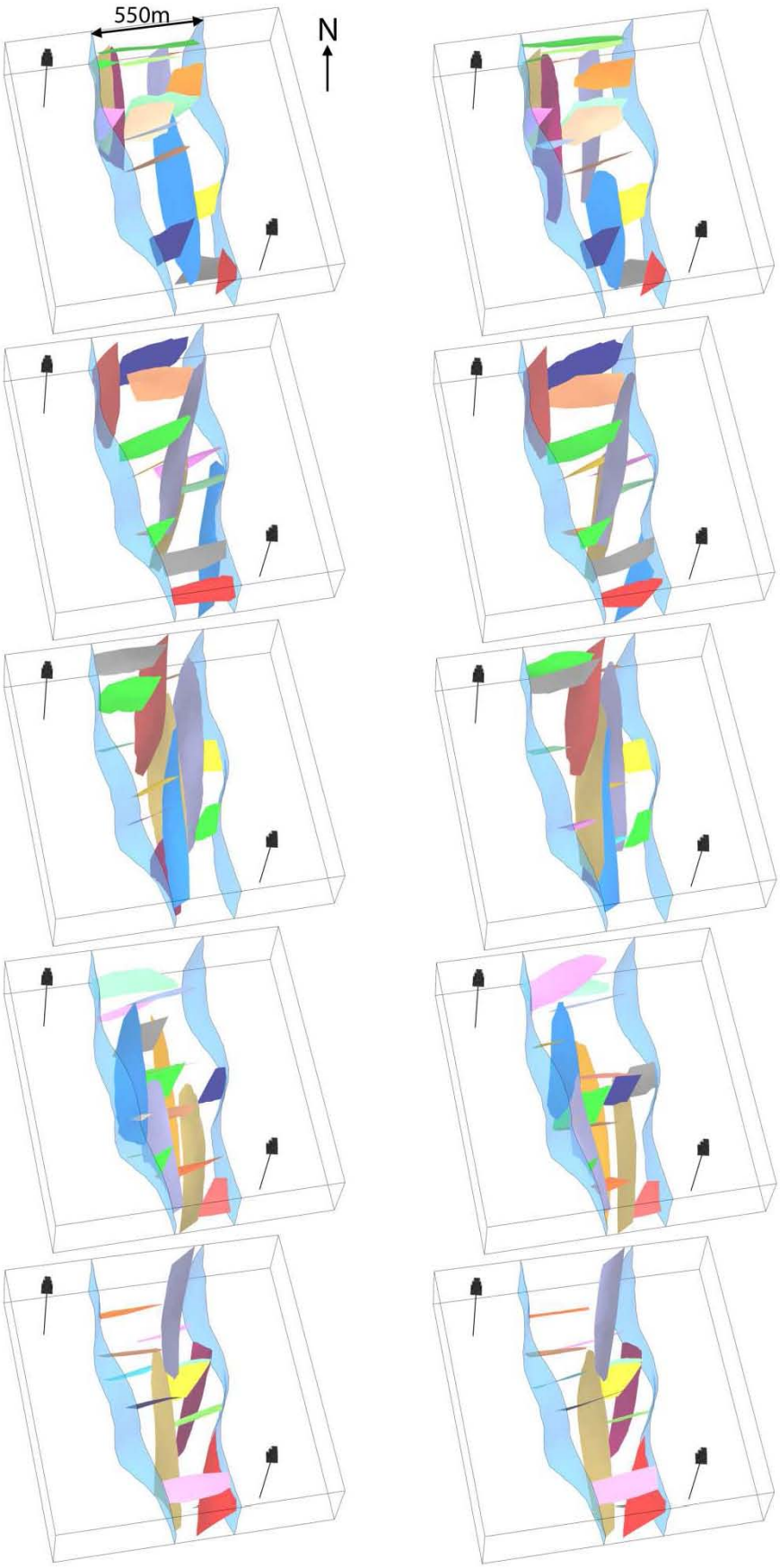

Figure 8: From top to bottom: five realizations close to the reference in term of flow response. Left: initial models. Right: models obtained after perturbation of fault parameters, closer to the reference water-saturation curve. Models may look more or less similar depending on the number of iteration.

Conclusions

The presented method aims at reducing risks in reservoir studies by considering uncertainties about fault geometry and fault connections. The method uses the advantage of implicit modeling to model realistic fault surfaces during the structural modeling step. At late stages during reservoir studies, support for fluid flow simulation is required. Most history matching techniques use pillar-based grids to perform such simulations, which limits the range of uncertainties and representation for faults. Instead, the proposed method uses an explicit discretization to represent the exact fault geometry and fault connections. Both the matrix and faults are represented as a set of control volumes connected by pipes storing the transmissibility. This discretization step is fast and avoids generating an extra grid with simplified fault surfaces.

The proposed methodolgy extends the classical history matching problem by using fluid flow data to handle uncertainties about fault surfaces and could be integrated into a more general history matching workflow.

Acknowledgments

The authors would like to acknowledge all the members of the Gocad consortium for funding this research, especially Paradigm Geophysical for providing the Gocad Software and API. Stanford University is also acknowledged for the GPRS flow simulator.

P. Calcagno, J. Chilès, G. Courrioux, and A. Guillen. Geological modelling from field data and geological knowledge: Part i. modelling method coupling 3d potential-field interpolation and geological rules. *Physics of the Earth and Planetary Interiors*, 171 (1-4): 147 – 157, 2008. ISSN 0031-9201. Recent Advances in Computational Geodynamics: Theory, Numerics and Applications.

N. Cherpeau, G. Caumon, and B. Lévy. Stochastic simulations of fault networks in 3D structural modeling. *Comptes Rendus Geoscience*, 342 (9): 687 – 694, 2010. ISSN 1631-0713. doi: 10.1016/j.crte.2010.04.008.

N. Cherpeau, G. Caumon, and B. Lévy. Stochastic simulation of fault networks from 2D seismic lines. *SEG Technical Program Expanded Abstracts*, 29 (1): 2366–2370, 2010. doi: 10.1190/1.3513325.

N. Cherpeau, G. Caumon, J. Caers, and B. Lévy. Topological uncertainties in structural geology and assimilation of dynamic data: parametrization and sampling. *Water Resources Research*, submitted.

T. Frank, A.-L. Tertois, and J.-L. Mallet. 3D-reconstruction of complex geological interfaces from irregularly distributed and noisy point data. *Computers & Geosciences*, 33 (7): 932 – 943, 2007. ISSN 0098-3004.

L. Holden, P. Mostad, B. F. Nielsen, J. Gjerde, C. Townsend, and S. Ottesen. Stochastic structural modeling. *Math. Geol.*, 35 (8): 899–914, 2003.

S. Jenni, L. Y. Hu, R. Basquet, G. de Marsily, and B. Bourbiaux. History matching of stochastic models of field-scale fractures: Methodology and case study. In *SPE Annual Technical Conference and Exhibition*, Houston, Texas, 2004.

M. Karimi-Fard, L. J. Durlofsky, and K. Aziz. An efficient discrete-fracture model applicable for general-purpose reservoir simulators. *SPE Journal*, 9: 227–236, June 2004. Paper 79699.

J.-L. Mallet and A.-L. Tertois. Solid earth modeling and geometric uncertainties. In *SPE Annual Technical Conference and Exhibition*, Florence, Italy, 2010.

T. Manzocchi, A. E. Heath, B. Palananthakumar, C. Childs, and J. J. Walsh. Faults in conventional flow simulation models: a consideration of representational assumptions and geological uncertainties. *Petroleum Geoscience*, 14 (1): 91–110, 2008.

P. Røe, F. Georgsen, A. Syversveen, and O. Lia. Fault displacement modelling using 3D vector fields. In *12th European Conference on the Mathematics of Oil Recovery*. EAGE, 2010.

A. Seiler, S. Aanonsen, G. Evensen, and J. C. Rivenæs. Structural Surface Uncertainty Modeling and Updating Using the Ensemble Kalman Filter. *SPE Journal*, 15 (4): 1062–1076, 2010. ISSN 1086-055X.

S. Vitel. *Méthodes de discrétisation et de changement d'échelle pour les réservoirs fracturés 3D*. PhD thesis, INPL, Nancy, France, 2007.

Incertitudes structurales en géomodélisation : échantillonnage et approche inverse

Résumé

La modélisation du sous-sol est un outil indispensable pour décrire, comprendre et quantifier les processus géologiques. L'accès au sous-sol et son observation étant limités aux moyens d'acquisition, la construction de modèles tridimensionnels du sous-sol repose sur l'interprétation de données éparses à résolution limitée. Dans ce contexte, de nombreuses incertitudes affectent la construction de tels modèles, dues aux possibles biais humains cognitifs lors de l'interprétation, à la variabilité naturelle des objets géologiques et aux incertitudes intrinsèques des données utilisées. Ces incertitudes altèrent la prédictibilité des modèles et leur évaluation est donc nécessaire afin de réduire les risques économiques et humains liés à l'utilisation des modèles.

Le travail de thèse s'est déroulé dans le cadre plus spécifique des incertitudes sur les structures géologiques. Les réponses apportées sont multiples : (1) une méthode stochastique de génération de modèles structuraux à géométrie et topologie changeantes, combinant une connaissance *a priori* des structures géologiques aux données interprétées, a été développée ; (2) le réalisme géologique des structures modélisées est garanti grâce à la modélisation implicite, représentant une surface par une équipotentielle d'un champ scalaire volumique ; (3) la description des failles en un nombre restreint de paramètres incertains a permis d'aborder la modélisation inverse, ce qui ouvre la voie vers l'assimilation de données géophysiques ou d'écoulement fluides grâce à des méthodes bayesiennes.

Mots clés : géomodélisation, incertitude, modèle stochastique, topologie

---

Structural uncertainties in geomodeling : sampling and inverse approach

Abstract

Subsurface modeling is a key tool to describe, understand and quantify geological processes. As the subsurface is inaccessible and its observation is limited by acquisition methods, 3D models of the subsurface are usually built from the interpretation of sparse data with limited resolution. Therefore, uncertainties occur during the model building process, due to possible cognitive human bias, natural variability of geological objects and intrinsic uncertainties of data. In such context, the predictability of models is limited by uncertainties, which must be assessed in order to reduce economical and human risks linked to the use of models.

This thesis focuses more specifically on uncertainties about geological structures. Our contributions are : (1) a stochastic method for generating structural models with various fault and horizon geometries as well as fault connections, combining prior information and interpreted data, has been developped ; (2) realistic geological objects are obtained using implicit modeling that represents a surface by an equipotential of a volumetric scalar field ; (3) faults have been described by a reduced set of uncertain parameters, which opens the way to the inversion of structural objects using geophysical or fluid flow data by baysian methods.

Keywords : geomodeling, uncertainty, stochastic model, topology

---

Centre de Recherches Pétrographiques et Géochimiques
ENSG, rue du Doyen Marcel Roubault - 54500 Vandoeuvre-lès-Nancy