



HAL
open science

Algorithmes de visualisation des incertitudes en géomodélisation sur GPU

Thomas Viard

► **To cite this version:**

Thomas Viard. Algorithmes de visualisation des incertitudes en géomodélisation sur GPU. Autre. Institut National Polytechnique de Lorraine, 2010. Français. NNT : 2010INPL042N . tel-01748833

HAL Id: tel-01748833

<https://hal.univ-lorraine.fr/tel-01748833v1>

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>

ALGORITHMES DE VISUALISATION DES INCERTITUDES EN GÉOMODÉLISATION SUR GPU

THÈSE

présentée et soutenue publiquement le 5 octobre 2010

pour l'obtention du

Doctorat de l'Institut National Polytechnique de Lorraine

Spécialité Géosciences

par

Thomas VIARD

Composition du jury

Rapporteurs : Georges-Pierre BONNEAU
Klaus MOSEGAARD

Examineurs : Guillaume CAUMON
Bruno JOBARD

Invité : Marek CZERNUSZENKO

Directeurs : Jean-Jacques ROYER
Bruno LÉVY

Remerciements

Pour arriver à bon port, une thèse requiert de laisser libre cours à la créativité du doctorant, tout en l’orientant dans la bonne direction à grand renfort de discussions animées et d’idées éclairées. Je tiens donc à remercier Jean-Jacques Royer et Bruno Lévy, non seulement pour m’avoir encadré tout au long de ces trois années de thèse, mais aussi pour avoir su me communiquer leur passion pour ces deux mondes différents que sont la géostatistique et la visualisation scientifique. Je souhaite également remercier Guillaume Caumon pour avoir initié cette thèse, et surtout pour sa participation active à la plupart des aspects présentés dans le manuscrit. Ca aura été un vrai plaisir de travailler au sein du groupe Gocad sous ta direction. Je tiens bien entendu à remercier chaleureusement l’ensemble des membres du jury pour avoir accepté d’évaluer mes travaux – outre Guillaume Caumon qui a aussi été examinateur de cette thèse, merci à Georges-Pierre Bonneau et Klaus Mosegaard pour avoir tenu le rôle de rapporteurs, ainsi qu’à Bruno Jobard pour avoir été le second examinateur.

Je tiens également à remercier l’ensemble des membres du consortium Gocad, et plus particulièrement la compagnie ExxonMobil, pour avoir financé ces travaux de recherche. Cette thèse aura été l’occasion d’échanges fructueux avec plusieurs des sponsors du consortium. I especially want to thank Marek Czernuszenko, who offered me an internship at ExxonMobil URC in Houston and gave me the opportunity to work within a professional environment. Besides the official acknowledgements, I also want to thank him for our friendly informal discussions ; I really appreciated our cultural exchanges, and I hope you liked my French bread recipe ! J’ai également eu l’occasion de rendre visite à CGGVeritas à Crawley sur l’invitation de Rémi Moyen, je le remercie chaleureusement pour son accueil là bas, pour ses nombreuses suggestions et critiques constructives envers mon travail, et pour ces sorties “hors travail” qui m’ont fait voir de plus près la vie à Crawley. Je tiens aussi à remercier Alexandre Hugot et Emmanuel Gringarten pour leur participation à cette thèse en ce qui concerne les incertitudes dynamiques, et notamment pour les nombreuses simulations d’écoulement qu’ils ont réalisé sur les modèles Nancy1 et Brugge.

La vie au laboratoire aura été riche en rencontres qui ont égayées le paysage Nancéen. Parmi les “anciens”, merci à Anlor qui a su me convaincre de me lancer dans la thèse, à Luc pour nos parties endiablées de Dota et de karting (et accessoirement, pour m’avoir aidé à surmonter mes premières difficultés techniques), à Lise pour ses petits gâteaux (sans toi, le concours de desserts perd son souffle !), et à Marco pour avoir partagé ses expériences américaines. Parmi ceux qui peuplent encore le labo, je voudrais remercier Vincent pour m’avoir supporté tout ce temps (deux, quatre, six... ouite ?), Pauline pour sa bonne humeur permanente et ses talents de guide Houstonienne (!), Florent pour ses chipoglycémies et ses nombreuses défaites à Volemon, Nico Cherpeau pour son humour tranquille (et au milieu coule une thèse !), Gautier pour ses sabres lasers (wiooooooooooom !), Romain pour ses discussions aquariophilosophiques (avec mes félici-

tations pour le petit Joachim!), Julien pour m'avoir insufflé la tactaqua du baby, l'autre Pauline pour son énergie, son enthousiasme et sa naïveté sans limites, Nacim pour nos échanges franco-canadiens (tabernacle! Y causent pas comme icitte ces québécois...), et Jeanne pour ses gentilles râleries (mais ca va pas du tout ca!). La liste serait incomplète sans mentionner Fatima, qui m'a inlassablement aidé sur les problèmes administratifs divers et variés, Christophe, qui est toujours disponible pour une résolution de bug ou une discussion geeko-technique, et Pierre, qui a dû gérer mes nombreux hackages du serveur de licences en plus de m'aider à gérer mes problèmes Linuxiens. Enfin, merci également aux étudiants que j'ai pu encadrer – Florian, Caroline, Thibault, Philippe, Gautier², les deux Charlotte, Gaétan, Solène et Grégoire – d'abord pour avoir supporté ma volonté de perfectionnisme tout au long de leur projet, et pour leurs efforts et leur gentillesse constante.

Hors du laboratoire, un grand merci à toute l'équipe de l'URC pour leur accueil parmi eux, et merci aux stagiaires Houstoniens (Antoine, Derek, Eitan et tous les autres) – je suis heureux d'avoir pu ne pas comprendre les règles du baseball malgré toutes vos explications!

Parmi mes amis de longue date, un grand merci à Marion pour son soutien indéfectible (bêêêêêêêêh!), Morgane pour son éternelle fraîcheur ([Ton plus aigu] bêêêêêêêêh!), Delph pour son caractère bien trempé (mrou?), Wiwi pour notre entente et nos fréquents débats (alors, c'est quoi un monde idéal en fin de compte?), Raph pour nos échanges percussionnistes, ainsi que Geoffrey et Pierrot les orléanais et Virginie la metzine. Merci aussi aux disparus, particulièrement Serge et Kévin, auxquels je pense régulièrement; j'espère que tout va bien pour vous, les amis. De moins longue date, merci aux membres du Vladimir Express et de l'orchestre de l'INPL pour ces délicieuses pauses musicales – Vincent, Elodie, Armand, Philippe, Marion et les autres, j'ai adoré grattouiller la peau de mon djembé pendant que vous faisiez de la musique.

Je tiens également à remercier toute ma famille pour m'avoir fait avancer jusqu'à aujourd'hui - papa, maman, papi et mamie, Annie et Georget, Nanine, Sylvie et Marcot, Gwen, David (félicitations à toi et à Delphine pour la petite Margaux!), Monique et Georges. Quelle que soit la distance, vous restez sans cesse dans mes pensées. Merci aussi à ma belle famille pour m'avoir soutenu ces dernières années - Michel et Louise, Mélanie, Maguelone. Votre accueil me fait toujours chaud au cœur. Enfin, un grand merci à ma future femme, Rébecca, qui a toujours su être derrière moi quand j'en avais besoin. Plus que quelques mois avant de te dire oui!

Pour conclure ces remerciements, je souhaite faire un grand dougoudouk à Topaze et Elodée. Vos incessantes prises de bec me manquent beaucoup.

*A ma future femme, mes parents,
et mes cochons d'inde.*

“ Du travail, encore du travail !”

Un péon anonyme

“ Οἶδα οὐδὲν εἰδώς.”

Σωκράτους

Table des matières

Table des figures	xi
-------------------	----

Liste des tableaux	xv
--------------------	----

Introduction

1

Caractérisation des incertitudes

1.1	Taxonomie des incertitudes	6
1.2	Processus de géomodélisation et théorie des problèmes inverses	9
1.2.1	Le problème direct	9
1.2.2	Le problème inverse	10
1.2.3	Simulations stochastiques	11
1.3	Exemple de création de géomodèle : la caractérisation d'un réservoir	13
1.3.1	Etapas de caractérisation d'un réservoir	13
1.3.2	Propagation des incertitudes	16
1.4	Métriques d'incertitude	17
1.4.1	Génération d'une densité de probabilité locale	18
1.4.2	Statistiques d'incertitude locale	18
1.4.3	Validité de l'échantillonnage des incertitudes	21
1.5	Conclusion	22

2

Etat de l'art des techniques de visualisation des incertitudes

2.1	Méthodes de visualisation scientifique	26
-----	--	----

2.1.1	Extraction d'isosurfaces	26
2.1.2	Rendu volumique	28
2.2	Cadre théorique de visualisation des incertitudes	32
2.2.1	Variables visuelles	32
2.2.2	Qualité des métaphores visuelles	33
2.2.3	Interférences visuelles	35
2.2.4	Préattentivité	35
2.2.5	Concentration de l'information	36
2.3	Méthodes de visualisation des incertitudes	36
2.3.1	Incertitudes sur des graphiques	37
2.3.2	Incertitudes en archéologie	38
2.3.3	Incertitudes de position particulière dans des écoulements	40
2.3.4	Incertitudes sur la reconstruction de surfaces	41
2.3.5	Incertitudes en médecine	45
2.3.6	Incertitudes en climatologie	49
2.3.7	Incertitudes dans les systèmes d'information géographiques	51
2.3.8	Incertitudes en géomodélisation	55
2.4	Evaluation de la qualité des visualisations	56
2.5	Conclusion	58

3

Visualisation des incertitudes pétrophysiques en géomodélisation

3.1	Méthodes statiques dédiées	62
3.1.1	Modification des attributs de la couleur	62
3.1.2	Modification du codage par couleur	64
3.1.3	Surcharge d'informations	69
3.1.4	Ambiguation	78
3.1.5	Modification de la géométrie	84
3.2	Validation des méthodes statiques	87
3.2.1	Participants	87

3.2.2	Description des données	89
3.2.3	Protocole	89
3.2.4	Analyse des résultats	91
3.2.5	Résultats	92
3.2.6	Discussion	94
3.3	Méthodes dynamiques dédiées	95
3.3.1	Animation par images clefs	95
3.3.2	Types d'interpolations temporelles continues	96
3.3.3	Animation de réalisations non-équiprobables	99
3.4	Validation des méthodes dynamiques	101
3.4.1	Description des données	101
3.4.2	Description des questions	102
3.4.3	Résultats	104
3.4.4	Limites	109
3.4.5	Conclusions	110
3.5	Application aux incertitudes pétrophysiques	111
3.5.1	Propriétés statiques	111
3.5.2	Propriétés dynamiques	114
3.6	Conclusion	117

4

Visualisation des incertitudes structurales en géomodélisation

4.1	Méthodes statiques dédiées - les surfaces fantômes	122
4.1.1	Génération d'un volume enveloppe	123
4.1.2	Population par une métrique d'incertitude	125
4.1.3	Visualisation de la surface fantôme	125
4.1.4	Résultats et discussion	126
4.2	Méthodes dynamiques dédiées - isosurfaces animées	127
4.2.1	Méthodes d'extraction	127
4.2.2	Modélisation de surfaces mourantes	128

4.2.3 Résultats et discussion 131
4.3 Conclusion 132

Conclusions

A
Visualisation des incertitudes et programmation sur GPU

A.1 Outils utilisés 140
A.1.1 Shader par sommets 141
A.1.2 Shader de géométrie 141
A.1.3 Shader par fragments 141
A.2 Structure logicielle 142
A.2.1 Organisation d'un shader 142
A.2.2 Exemple 143
A.3 Prototypage de shaders 144

B
Espaces de couleur

C
Publications

C.1 Adjacent versus coincident representations of geospatial uncertainty : Which
promote better decisions ? 151

Bibliographie **167**

Table des figures

1.1	Chaîne de traitement pour la caractérisation du potentiel associé à un réservoir.	14
1.2	Illustration des correspondances spatiales entre deux simulations de faille.	19
2.1	Visualisation d'un milieu poreux par rendu volumique et par isosurface.	27
2.2	Extraction d'isosurface à partir des liens topologiques entre sommets, arêtes et faces.	28
2.3	Table des cas dans l'algorithme des Marching Cubes.	29
2.4	Fonction de transfert visuelle définissant l'émission et l'absorption particulière en fonction d'un champ scalaire.	30
2.5	Rayon visuel paramétrisé par rapport à la distance t à l'observateur.	31
2.6	Discretisation par tranchage incrémental à différents degrés de finesse.	32
2.7	Illustration des propriétés d'ordre naturel, de perception quantitative, de sélectivité et d'associativité de la variable visuelle "taille".	34
2.8	Illustration du phénomène de préattentivité.	36
2.9	Graphique présentant la quantité cumulée d'hydrocarbures produits avec incertitudes sur la production prédite.	38
2.10	Visualisation des incertitudes sur l'architecture d'un bâtiment en archéologie.	39
2.11	Visualisation de chenaux avec une transparence accrue des zones incertaines.	42
2.12	Visualisation d'une couche géologique avec enveloppe d'incertitude sur la position de la couche.	43
2.13	Visualisation d'un modèle de tumeur synthétique avec perturbation de la position des points dans les zones incertaines.	48
2.14	Encodage de l'incertitude de vecteurs sur la forme d'un glyphe.	50
2.15	Visualisation de chenaux avec brouillard au dessus des zones incertaines.	54
2.16	Visualisation de chenaux avec une saturation décréue dans les zones incertaines.	56

3.1	Visualisation des incertitudes par modification des attributs de la couleur en rendu volumique.	64
3.2	Utilisation du filtre de Sobel pour améliorer la perception dans les zones fortement incertaines.	65
3.3	Relations entre l'espace physique \mathcal{P} et l'espace de couleurs \mathcal{C} définies par la fonction de transfert bidimensionnelle.	65
3.4	Schémas d'interpolation définissant une fonction de transfert bidimensionnelle.	66
3.5	Fonction de transfert bidimensionnelle appliquée à des données de porosité.	68
3.6	Fonction de transfert bidimensionnelle appliquée à des données de porosité en rendu volumique.	68
3.7	Stratégies d'échantillonnage des glyphes au sein du domaine d'intérêt.	70
3.8	Comparaison entre la visualisation des incertitudes par densité d'échantillonnage et par taille des glyphes.	70
3.9	Méthodes d'association entre un motif répétitif et le degré d'incertitude local.	71
3.10	Contrôle du degré d'interférence entre un motif répétitif et la couleur de fond.	72
3.11	Modulation de la largeur du trait d'un motif répétitif en fonction du degré d'incertitude local.	73
3.12	Comparaison de la modulation de fréquence de répétition d'un motif de manière discrète et de manière continue.	74
3.13	Visualisation des incertitudes par intensité d'un motif répétitif en rendu volumique.	75
3.14	Visualisation du degré d'incertitude local par isolignes multiples.	76
3.15	Visualisation du degré d'incertitude local par isosurfaces multiples.	77
3.16	Approximations de la courbe fractale de Koch à différentes itérations.	77
3.17	Nucléation et développement par itérations successives d'une fractale dendritique.	78
3.18	Méthodes de visualisation des incertitudes par résolution variable.	79
3.19	Visualisation des incertitudes par intensité de flou.	80
3.20	Visualisation d'une carte de porosité avec et sans flou.	81
3.21	Visualisation des incertitudes par flou avec et sans motif répétitif.	81
3.22	Visualisation des incertitudes par ajout de bruit.	82
3.23	Visualisation des incertitudes par distorsion.	84
3.24	Pixellisation d'une image afin de garantir l'anonymat de la personne photographiée.	85

3.25	Pixellisation d'une carte de porosité en fonction du degré d'incertitude local.	85
3.26	Visualisation des incertitudes par anamorphose discontinue.	87
3.27	Visualisation des incertitudes par barres d'erreur.	88
3.28	Ressources disponibles pour remplir les tâches de l'étude sur des utilisateurs.	90
3.29	Proportions des réponses à la question 2.	93
3.30	Vitesse d'affichage \mathcal{V} en fonction des réalisations \mathcal{E}_i .	100
3.31	Jeu de données synthétique utilisé dans l'étude sur des utilisateurs des méthodes dynamiques.	102
3.32	Distribution des degrés d'expertise des participants en visualisation des incertitudes, présentés séparément pour chaque groupe.	104
3.33	Visualisation d'un horizon faillé du réservoir Nancy1 présentant à la fois la porosité et les incertitudes associées.	113
3.34	Visualisation de la structure du modèle Brugge.	114
3.35	Evolution conjointe de la saturation en eau et des incertitudes associées dans le modèle Brugge.	116
4.1	Extrusion d'une des facettes de la surface initiale \mathcal{S}_i .	124
4.2	Discrétisation du volume d'intérêt à l'aide d'une grille Cartésienne.	124
4.3	Visualisation des incertitudes structurales associées au réservoir Mandaros par la technique des surfaces fantômes.	127
4.4	Exemples d'intersections de surfaces géologiquement incohérentes d'après Caumon <i>et al.</i> (2004).	129
4.5	Illustration du critère de validité spatial par cascade de conditions implicites.	129
4.6	Exemple de faille inverse simple ne pouvant être représentée par une cascade de conditions implicites.	130
4.7	Représentation d'une faille inverse à l'aide d'un arbre binaire.	131
1	Chaîne simplifiée de modélisation d'un réservoir.	136
A.1	Organisation simplifiée du pipeline graphique.	140
A.2	Comparaison entre les modèles Simple Instruction Simple Donnée et Simple Instruction Multiples Données.	141

Liste des tableaux

1.1	Taxonomie des incertitudes.	7
1.2	Effet de l'intervalle des valeurs sur l'écart type et le coefficient de variation.	20
2.1	Variables visuelles proposées par Bertin (1983), avec extensions proposées par MacEachren (1992) et McGranaghan (1993).	33
2.2	Récapitulatif des propriétés perceptuelles associées aux variables visuelles.	34
3.1	Taxonomie des méthodes de visualisation des incertitudes.	63
3.2	Pourcentage de réponses correctes à la question 1.	93
3.3	Comparaison entre les distributions des réponses à la question 2.	94
3.4	Ecart de perception moyen pour chaque groupe à la question 3.	94
3.5	Comparaison entre les distributions des réponses à la question 3.	95
3.6	Estimation moyenne de la probabilité de connexion entre le puits injecteur et le corps sableux.	105
3.7	Comparaison de la confiance moyenne des animations standard et pondérée.	107
3.8	Comparaison entre les distributions de confiance des animations standard et pondérée.	107
3.9	Comparaison entre les distributions de confiance par question, toutes méthodes d'animation confondues.	107
3.10	Corrélation des réponses avec le degré d'expertise des participants.	108
3.11	Résumé des méthodes de visualisation implémentées en fonction du support.	117
3.12	Résumé des principales contributions pour chacune des méthodes de visualisation implémentées.	118
A.1	Éléments de code source associés à la visualisation des incertitudes par transparence et par saturation.	144

Introduction

La modélisation des objets géologiques, ou géomodélisation, a pour but d'améliorer la compréhension de l'état actuel du sous-sol et de mieux anticiper son évolution sous l'effet d'actions anthropiques. Les applications nécessitant l'élaboration de géomodèles sont diverses : production de réservoirs pétroliers, prévision de séismes, exploitations minières, développement de barrages ou étude du transport de polluants... Dans chacun de ces cas, la création du géomodèle repose sur l'intégration de concepts géologiques (environnement de dépôt des sédiments, histoire tectonique régionale, *etc.*) avec les données recueillies sur le terrain (observations d'affleurements, mesures de forages ou campagnes sismiques par exemple). La mise en évidence d'incohérences permet de jeter un regard critique sur la compréhension du sous-sol ; elle révèle des erreurs dont la source peut être imputable à des problèmes de mesure ou à l'application de concepts erronés.

Cependant, même lorsque données et concepts s'accordent, construire un géomodèle reste un exercice délicat connu sous le nom de *problème inverse*. Il consiste à trouver un ensemble de caractéristiques intrinsèques du domaine étudié, dont la signature reproduit au mieux les observations menées sur le terrain, sans pour autant entrer en contradiction avec les règles géologiques imposées par le milieu. Du fait de données parcellaires (mesures aux puits, carte géologique) ou exhaustives mais imprécises (mesures sismiques), plusieurs géomodèles respectant ces contraintes peuvent généralement être construits ; il y a donc *non-unicité* de la solution. L'existence de plusieurs géomodèles possibles traduit une incertitude sur les caractéristiques spatiales du sous-sol, dont la compréhension est critique dans la plupart des processus de décision en géosciences.

Problématiques générales et organisation du mémoire

De par sa nature même, la notion d'incertitude est délicate à formaliser et est généralement mal définie dans la littérature. Dans le premier chapitre de cette thèse, nous examinons l'ensemble des concepts proposés précédemment ainsi que leur validité dans un contexte géologique. Nous discutons ensuite des méthodes permettant d'échantillonner l'espace des incertitudes en géosciences, en nous appuyant notamment sur la théorie des problèmes inverses et le domaine de la simulation stochastique d'entités géologiques. Enfin, nous présentons un certain nombre de stratégies permettant de convertir les incerti-

tudes locales, représentées sous la forme de densités de probabilité, en grandeurs scalaires à l'aide de métriques d'incertitude.

La caractérisation des incertitudes ne présente qu'un intérêt limité si les géologues sont incapables de percevoir leurs évolutions spatiales. Pour répondre à ce problème, un certain nombre de techniques de visualisation scientifique ont été proposées par le passé ; nous présentons un état de l'art de ces approches dans le second chapitre de cette thèse. Nous commençons par une présentation générale des méthodes de visualisation scientifique avant de dresser un bilan aussi exhaustif que possible sur les approches de visualisation des incertitudes existantes. La qualité de ces approches est en outre discutée, tant d'un point de vue théorique qu'empirique.

Nous présentons ensuite nos propres approches permettant la visualisation des incertitudes pétrophysiques et des incertitudes structurales, respectivement dans les troisième et quatrième chapitres de cette thèse. Outre la description des concepts associés à chacune de ces techniques, nous présentons deux études sur utilisateurs respectivement conduites sur les méthodes de visualisation statique et par animation. Ces études apportent des informations inédites sur la manière selon laquelle la visualisation des incertitudes est perçue par un observateur.

Les algorithmes de visualisation présentés dans cette thèse ont fait l'objet d'un effort de développement soutenu, reposant majoritairement sur le caractère programmable des cartes graphiques modernes. Les aspects techniques de la programmation sur GPU sont présentés dans l'annexe A. Cette annexe introduit tout d'abord les entités utilisées pour la programmation sur GPU, en particulier les programmes de shader capables de s'exécuter directement sur la carte graphique. Notre approche de méta-programmation de shaders et l'organisation qui lui est associée sont ensuite présentées. Enfin, nous discutons des possibilités de prototypage de shaders offertes par les facilités de compilation et d'édition de liens à la volée.

Contexte de travail

Ce travail de recherche se focalise sur les problèmes de compréhension intuitive des incertitudes spatiales, et explore les méthodes de visualisation présentant conjointement géomodèle et incertitudes spatiales. Les algorithmes produits lors de ce travail ont été développés en langage C++ et OpenGL Shading Language (GLSL) au sein du groupe de recherche Gocad¹ ; ils s'appuient sur les dernières évolutions matérielles des cartes graphiques (également nommées GPU pour *Graphic Processing Units*), notamment en termes de programmation. Ils permettent la visualisation d'incertitudes de natures différentes (propriétés pétrophysiques statiques ou dynamiques, géométrie ou topologie des

¹<http://www.gocad.org>

structures géologiques) sur les principaux types de support numérique utilisés en géosciences.

Les travaux présentés dans le cadre de cette thèse ont par ailleurs fait l'objet de publications régulières dans le cadre d'une conférence interne, le Gocad Meeting (Viard *et al.*, 2008a, 2009d,c,e,a, 2010a,c), ainsi qu'un certain nombre de conférences externes telles que la Réunion des Sciences de la Terre (Viard *et al.*, 2008b), le International Geological Congress (IGC, Viard *et al.*, 2008c) ou la International Petroleum and Technology Conference (IPTC, Viard *et al.*, 2009b). Cette thèse a également permis de produire une publication dans Computers and Geosciences, journal international avec comité de relecture traitant des dernières avancées en géosciences (Viard *et al.*, 2010b, Annexe C). Une seconde publication a été soumise dans IEEE Transactions on Visualization and Computer Graphics et est actuellement en cours de relecture (Viard *et al.*, 2010d).

Enfin, cette thèse m'a offert l'opportunité d'encadrer ou de co-encadrer dix étudiants, dont quatre en vue de l'obtention d'un Master II et six préparant un Master I, sur des thématiques plus larges que la seule visualisation des incertitudes telles que le rendu volumique de grilles curvilinéaires (Laurent *et al.*, 2009), la classification de réalisations multiples (Panhaleux *et al.*, 2010) ou les techniques de visualisation de tenseurs (Piquet *et al.*, 2010). Bien que tous ne soient pas valorisables dans le cadre de cette thèse, ces projets m'ont néanmoins permis d'élargir mes connaissances et de porter un regard critique sur des problèmes apparentés à mes travaux.

Chapitre 1

Caractérisation des incertitudes

Sommaire

1.1	Taxonomie des incertitudes	6
1.2	Processus de géomodélisation et théorie des problèmes inverses	9
1.2.1	Le problème direct	9
1.2.2	Le problème inverse	10
1.2.3	Simulations stochastiques	11
1.3	Exemple de création de géomodèle : la caractérisation d'un réservoir	13
1.3.1	Etapes de caractérisation d'un réservoir	13
1.3.2	Propagation des incertitudes	16
1.4	Métriques d'incertitude	17
1.4.1	Génération d'une densité de probabilité locale	18
1.4.2	Statistiques d'incertitude locale	18
1.4.3	Validité de l'échantillonnage des incertitudes	21
1.5	Conclusion	22

Les incertitudes peuvent avoir un impact crucial dans le domaine des géosciences : estimation de menaces telluriques (séismes, éruptions, *etc.*), de pollutions liées aux activités humaines, investissements financiers à haut risque dans le domaine énergétique, *etc.* Dans l'ensemble de ces applications, il est nécessaire de connaître l'état du sous-sol dans le processus de prise de décision. Cependant, la majeure partie du sous-sol est inaccessible à toute observation directe ; de ce fait, les décisions sont souvent prises en se basant sur des connaissances incomplètes ou erronées. La compréhension du degré de confiance associé à chaque information est donc critique pour éviter de construire un raisonnement reposant sur des données biaisées.

Dans une étude de différents domaines affectés par les incertitudes, Skeels *et al.* (2008) opposent les *incertitudes connues* aux *incertitudes inconnues*. Les incertitudes *connues* sont les incertitudes dont le décideur sait qu'elles existent, qu'il soit en mesure de les

quantifier ou non. Les incertitudes *inconnues*, en revanche, sont les incertitudes dont le décideur n'est pas conscient ; elles induisent typiquement un excès de confiance vis-à-vis des données qui peut s'avérer extrêmement préjudiciable. Bien que l'ensemble de ces incertitudes existent en géomodélisation, nous nous concentrons dans ce travail sur les incertitudes connues et quantifiables. Nous commençons par un rappel des différents types d'incertitude rencontrés en géosciences (Section 1.1) avant de présenter la manière dont ces incertitudes sont caractérisées dans le cadre d'une démarche inverse (Section 1.2). Nous présentons ensuite certains opérateurs mathématiques permettant de passer d'une distribution de probabilité à une métrique d'incertitude (Section 1.4).

1.1 Taxonomie des incertitudes

L'*incertitude* est un concept complexe reflétant différents aspects de la méconnaissance d'un problème. La majorité des auteurs s'accordent à considérer l'incertitude comme une métadonnée représentant le manque d'informations sur l'entité étudiée (modèle numérique, paramètres d'un modèle physique, observations *in-situ*, etc.), mais ils encapsulent cette grandeur dans différentes notions telles que le *degré d'erreur*, l'*exactitude*, la *fiabilité*, la *précision*, etc. La signification de ces termes est parfois elle-même sujette à discussion – par exemple, Buttenfield (1993) présente l'exactitude comme une mesure de la conformité à une règle ou à un modèle et la sépare explicitement de la notion d'erreur, tandis que Thomson *et al.* (2005) considèrent l'exactitude et l'erreur comme deux notions opposées, c'est-à-dire la proximité ou l'écart entre une mesure ponctuelle et la valeur réelle. Pour d'autres auteurs comme Gershon (1998), l'incertitude n'est qu'une partie d'un concept plus large d'*imperfection des données*, qui intègre également des notions de présentation imparfaite de l'information, de corruption des données ou de complexité de l'information.

Dans le cadre de cette thèse, nous nous intéressons plus particulièrement aux incertitudes locales associées à des variables régionalisées, c'est-à-dire auxquelles est attachée une coordonnée spatiale x . Plusieurs tentatives de formalisation de la notion d'incertitude locale ont été proposées dans la littérature ; la plupart sont bâties sur le standard de transfert des données spatiales (Spatial Data Transfer Standard en anglais, ou SDTS) proposé par le US Geological Survey (USGS), dans lequel sont présentées les notions d'*exactitude de position*, d'*exactitude des attributs*, de *cohérence logique*, de *complétude* et de *traçabilité* (USGS, 1977). Buttenfield et Weibel (1988) proposent une classification des incertitudes locales à double entrée, où sont pris en compte le type des données (discrètes, catégoriques ou continues) et le type des incertitudes qui leur sont associées (telles que proposées dans la classification du SDTS). Dans une approche comparable, Gahegan et Ehlers (2000) présentent une classification à deux dimensions intégrant le type des incertitudes et l'objectif de l'exercice. Ils considèrent la notion d'*erreur temporelle* en plus de la typologie du SDTS. Mann (1993) et Bardossy et Fodor (2001) proposent une taxonomie dédiée

Tab. 1.1 – Taxonomie des incertitudes. Adapté d’après Thomson et al. (2005).

Type d’incertitude	Description
Erreur	Ecart des données par rapport à la valeur réelle
Précision	Résolution des données
Complétude	Quantité de données manquantes
Cohérence	Compatibilité entre les différentes données
Traçabilité	Historique des transformations subies par les données
Actualité	Degré de mise à jour des données
Crédibilité	Fiabilité de la source des données
Subjectivité	Importance des jugements humains inclus dans les données
Indépendance	Degré de parenté entre les données présentées par différentes sources

aux incertitudes géologiques incluant les notions de *variabilité naturelle* inhérente aux objets géologiques, de *limitation de l’échantillonnage spatio-temporel*, d’*erreur d’observation*, d’*erreur de mesure*, d’*erreur d’évaluation* généralement due à l’incomplétude des données disponibles, de *propagation des erreurs* affectant le résultat final et d’*incertitude conceptuelle* affectant la bonne compréhension des phénomènes mis en jeu dans l’étude. Corre et al. (2000) présentent une classification des incertitudes basée sur leur position dans la chaîne de traitement d’un réservoir, et séparent ainsi les incertitudes *géophysiques* liées à l’interprétation de cubes d’impédance sismiques, les incertitudes *géologiques* associées aux concepts sédimentaires et aux propriétés pétrophysiques, et les incertitudes *dynamiques* reliées à l’écoulement des hydrocarbures. Thomson et al. (2005) étendent davantage les spécifications du SDTS en présentant une typologie qui intègre également les notions d’*actualité*, de *crédibilité*, de *subjectivité* et d’*indépendance des sources d’information* (Table 1.1). Nous nous focalisons dans ce travail sur leur typologie, et proposons une illustration dans le contexte d’une campagne sismique 4D en domaine marin. La sismique 4D est une technique d’étude d’un réservoir dans le temps ; plusieurs images sismiques sont prises à des dates différentes, et les différences entre les images sont interprétées comme la conséquence de la production au sein du réservoir.

Erreur L’erreur correspond à l’écart entre la mesure d’une grandeur physique sur un échantillon et sa valeur réelle. Dans le cadre d’une sismique 4D, l’erreur sur l’amplitude sismique n’est pas mesurable *sensus-stricto* puisque la valeur réelle n’est pas connue ; cependant, les constructeurs des instruments de mesure fournissent généralement un intervalle d’erreur calibré en laboratoire sur des échantillons de référence.

Précision La précision indique le degré de résolution des mesures. En sismique, les résolutions sont typiquement basses – de l’ordre de dix mètres selon les axes verticaux

et horizontaux, tandis que les dimensions d'une couche réservoir peuvent ne pas excéder quelques mètres selon la verticale.

Complétude La complétude représente la quantité d'informations manquantes au sein d'un jeu de données. Les mesures sismiques sont à même d'échantillonner la totalité du volume d'intérêt et présentent donc un ensemble d'informations exhaustif dans l'espace. Cependant, la sismique 4D n'échantillonne qu'un nombre limité de pas de temps et présente donc une complétude temporelle faible.

Cohérence La cohérence indique la compatibilité entre les données compilées. En sismique 4D, il est fréquent que les instruments de mesure changent d'une campagne à l'autre du fait des avancées technologiques, limitant ainsi la répétabilité des mesures. Il arrive également qu'un instrument de mesure soit détérioré en cours de campagne et doive être remplacé. Par ailleurs, des facteurs externes non-répétables peuvent également limiter la cohérence entre des images sismiques successives, tels que les conditions météorologiques, l'orientation des courants ou les cavitations des turbines de bateau.

Traçabilité La traçabilité correspond à la capacité à suivre les transformations qu'ont subies les données. Le traitement des images sismiques est un processus complexe que les compagnies pétrolières sous-traitent fréquemment à des compagnies de services ; ces dernières possèdent en interne leurs propres algorithmes de traitement qu'elles peuvent ne pas divulguer à personnes extérieures, limitant ainsi la traçabilité des données. Les opérations qu'ont subies les données sont cependant généralement connues.

Actualité L'actualité représente le degré de mise à jour des données. Il n'est pas rare que les données collectées soient prétraitées immédiatement après leur acquisition, sur le bateau lui-même, puis fournies à la compagnie pétrolière ; les données sont donc généralement parfaitement à jour, mais sont acquises à des intervalles de temps variables.

Crédibilité La crédibilité indique la confiance qui peut être accordée à la source des données. Il est rare que la compagnie chargée de la sismique ait des intérêts particuliers dans l'exploitation du réservoir ; les informations fournies peuvent donc être considérées comme impartiales. Par ailleurs, les compagnies pétrolières font appel aux compagnies de services pour bénéficier de leur expertise. La crédibilité de la source est donc généralement élevée en sismique 4D.

Subjectivité La subjectivité mesure l'importance des jugements humains qui sont incorporés dans les données. Les données sismiques initiales mesurent le temps d'aller et de retour d'une onde dans le sol (on parle de *temps-double*). Pour intégrer ces données au

sein d'un géomodèle, il est nécessaire d'appliquer une conversion d'unités dite *conversion temps-profondeur* ; cette conversion repose sur un modèle de vitesses au sein du volume rocheux qui n'est lui-même pas connu. Quand les données le permettent, les modeleurs calibrent les informations sismiques par rapport aux puits afin de contraindre le modèle de vitesses. Aux endroits où les puits ne sont pas disponibles, le modèle de vitesse repose sur la formulation d'hypothèses quant à la nature des roches rencontrées, qui sont ensuite confrontées aux temps de retour des ondes sismiques et améliorées par itérations successives ; de nombreux jugements humains interviennent donc dans la création du modèle de vitesses ainsi que dans les algorithmes de traitement, notamment la migration temps-profondeur.

Indépendance L'indépendance indique l'absence de relations entre les sources de plusieurs ensembles d'informations. Ce type d'incertitude n'est généralement pas applicable en sismique 4D, dans la mesure où la compagnie de services est la seule source d'informations vis-à-vis des données sismiques.

L'ensemble des types d'incertitudes proposés par Thomson *et al.* (2005) peuvent coexister dans le domaine des géosciences ; cependant, toutes n'ont pas le même poids sur le degré d'incertitude associé au géomodèle produit. Le sous-sol étant inaccessible à toute observation directe, l'élaboration d'un géomodèle repose fréquemment sur des données parcellaires. De ce fait, le modeleur doit fréquemment formuler des hypothèses sur le contexte géologique étudié lors de la création du modèle : les deux types d'incertitudes ayant le plus fort impact en géosciences sont donc la *complétude* et la *subjectivité*.

1.2 Processus de géomodélisation et théorie des problèmes inverses

L'élaboration d'un géomodèle est le processus de construction d'une représentation plausible du sous-sol \mathcal{M} d'après un ensemble d'observations parcellaires d_{obs} collectées *in-situ*. Cette construction s'appuie généralement sur la *théorie des problèmes inverses* (Tarantola, 1987; Mosegaard et Tarantola, 1995) pour générer des représentations dont le calcul théorique par méthode directe reproduit au mieux les données.

1.2.1 Le problème direct

Le problème direct vise à trouver une fonction ψ qui, à une représentation du sous-sol \mathcal{M} donnée, associe les signaux $d_{\mathcal{S}}$ qui devraient théoriquement être reçus par un ensemble de senseurs \mathcal{S} :

$$\psi : \mathcal{M} \mapsto d_{\mathcal{S}} = \psi(\mathcal{M}) \quad (1.1)$$

La fonction ψ correspond donc à un modèle physique adopté comme loi de comportement. En pratique, la fonction ψ dérive généralement de lois physiques bien connues, telles que les lois de diffusion de la chaleur, de résistance d'un matériau ou d'attraction entre particules. Il n'est pas rare que ces lois soient ajustées à un problème particulier à l'aide de mesures expérimentales ; par exemple, pour s'assurer qu'un barrage voûte résistera à la pression des eaux, les propriétés géomécaniques des roches sur lesquelles s'appuie le barrage doivent être mesurées en laboratoire et traduites par une loi de comportement mécanique du matériau.

1.2.2 Le problème inverse

Le problème inverse vise à trouver une représentation du sous-sol \mathcal{M} qui, connaissant un ensemble de signaux d_{obs} mesurés sur le terrain et une loi de comportement théorique ψ , satisfait le problème direct $d_{obs} \approx \psi(\mathcal{M})$. Cette approche considère que la loi de comportement ψ représente le comportement réel du système avec précision ; de la sorte, si la représentation \mathcal{M} reproduit fidèlement la réalité du sous-sol, l'écart entre les observations d_{obs} et les prédictions $\psi(\mathcal{M})$ peut être considéré comme négligeable.

Non-unicité de la solution A partir d'une loi de comportement ψ bijective, le problème inverse peut être résolu en trouvant la fonction réciproque de ψ , notée ψ^{-1} . Cependant, dans le domaine des géosciences, le sous-sol est inaccessible à l'observation directe et l'ensemble des observations d_{obs} n'est que parcellaire. De ce fait, le problème inverse est sous-contraint et il n'existe pas de définition univoque d'une fonction réciproque ψ^{-1} ; on parle généralement de *non-unicité* de la solution \mathcal{M} satisfaisant le problème inverse.

Résolution du problème inverse Le problème inverse *sensus-stricto* est généralement traité par des algorithmes de type chaîne de Markov - Monte Carlo tels que la méthode de Metropolis-Hastings (Metropolis et Ulam, 1949; Hastings, 1970), qui permettent simultanément d'échantillonner et d'optimiser l'ensemble des solutions possibles.

L'algorithme de Metropolis-Hastings procède à partir d'une solution initiale du problème, c'est-à-dire un modèle \mathcal{M}_0 du sous-sol, qui peut être fourni par le géostatisticien ou généré aléatoirement. Cette solution initiale est alors itérativement perturbée selon une marche aléatoire pour produire une nouvelle *proposition de solution* \mathcal{M}_k , dans le but de se rapprocher d'un modèle du sous-sol dont le problème direct reproduit fidèlement les observations d_{obs} . A chaque itération, la qualité de la proposition \mathcal{M}_k est calculée à l'aide d'une probabilité conditionnelle aux observations $P(\mathcal{M}_k|d_{obs})$ sous la forme :

$$P(\mathcal{M}_k|d_{obs}) = k \cdot P(d_{obs}|\mathcal{M}_k) \cdot P_{prior}(\mathcal{M}_k) \quad (1.2)$$

où k est un facteur de normalisation, $P(d_{obs}|\mathcal{M}_k)$ est la probabilité d'effectuer les observations d_{obs} en supposant que la représentation \mathcal{M}_k reproduise fidèlement la réalité du

sous-sol, et $P_{prior}(\mathcal{M}_k)$ correspond à la probabilité *a-priori* de la représentation \mathcal{M}_k .

La qualité de la proposition \mathcal{M}_k est alors confrontée à la qualité de la proposition précédente. Deux cas de figures peuvent ainsi se présenter :

1. Si $P(\mathcal{M}_k|d_{obs})$ est supérieur ou égal à $P(\mathcal{M}_{k-1}|d_{obs})$, la proposition \mathcal{M}_k est conservée.
2. Si $P(\mathcal{M}_k|d_{obs})$ est inférieur à $P(\mathcal{M}_{k-1}|d_{obs})$, la proposition \mathcal{M}_k est conservée avec une probabilité $\frac{P(\mathcal{M}_k|d_{obs})}{P(\mathcal{M}_{k-1}|d_{obs})}$. De la sorte, il est possible de s'affranchir des problèmes de convergence vers des optima locaux de la solution.

1.2.3 Simulations stochastiques

La théorie des problèmes inverses fournit un cadre général permettant de répondre aux problèmes rencontrés en géosciences ; cependant, du fait de la très forte dimension des entités géologiques, les techniques reposant sur une stratégie d'échantillonnage de Monte Carlo se révèlent extrêmement coûteuses et sont souvent inutilisables en pratique. Pour répondre à ce problème, un grand nombre de techniques de simulation stochastique ont vu le jour ; ces méthodes visent à représenter directement les objets géologiques en respectant un ensemble de règles génétiques, conceptuelles et statistiques. La plupart des méthodes intègrent également un conditionnement aux données permettant d'honorer directement les observations. Ces méthodes peuvent être classées en deux catégories : les méthodes de simulation par pixel (par exemple, Srivastava, 1995; Caers, 2000, 2001; Caers et Zhang, 2002; Chambers et Yarus, 2006) et les méthodes de simulation par objet (par exemple, Journel *et al.*, 1998; Deutsch et Tran, 2002).

Simulations par pixel Les méthodes de simulation par pixel génèrent les champs de valeurs associés aux réalisations à l'aide d'une évaluation locale de la fonction de densité de probabilité, calculée pour chacun des nœuds d'une grille de simulation. La plupart de ces techniques reposent sur l'algorithme général suivant : (les étapes optionnelles sont indiquées entre parenthèses)

- (1.) Prétraitement des données ;
2. Conditionnement aux données, généralement par remplissage de la grille de simulation. Il peut arriver que plusieurs points de données doivent être affectés de manière concurrente à un même nœud de la grille. Dans ce cas, des méthodes de mise à l'échelle (*upscaling* en anglais) peuvent être utilisées ;
3. Définition d'un ordre de parcours aléatoire des nœuds dans la grille ;
4. Parcours des cellules dans l'ordre défini à l'étape 3 :

Pour le nœud(rand_1) au nœud(rand_N) **Faire**

- 4.1. Construction de la fonction de probabilité de densité locale, typiquement d'après les informations portées par les nœuds dans le voisinage ;
- 4.2. Tirage de Monte Carlo dans la fonction de probabilité de densité locale ;

4.3. Affectation de la valeur tirée au nœud, qui sera désormais considéré comme un point de donnée.

Fin Pour

(5.) Post-traitement des données.

De nombreuses méthodes de simulations stochastiques par pixel sont décrites dans la littérature, telles que la simulation séquentielle gaussienne, la simulation séquentielle gaussienne tronquée (Journal et Isaaks, 1984), la simulation par indicatrice (Caers, 2000), la simulation par champs de probabilités (Srivastava, 1992), les statistiques multipoints (Caers et Zhang, 2002), *etc.* Les principales différences entre ces méthodes reposent essentiellement sur l'étape d'évaluation de la fonction de densité de probabilité locale, et dans une moindre mesure sur les étapes de prétraitement et de post-traitement des données.

Simulations par objet Les méthodes de simulation par objet sont populaires dans les domaines où des entités géologiques peuvent être conceptualisées de manière claire, notamment dans la modélisation de chenaux en domaine fluvial (Deutsch et Wang, 1996; Holden *et al.*, 1998; Deutsch et Tran, 2002).

Les méthodes objet visent à représenter la forme des entités géologiques indépendamment de toute grille de simulation, généralement à l'aide d'une représentation paramétrique de la forme associée à l'entité; ainsi, le paramètre de sinuosité d'un chenal peut être représenté par la période d'une fonction trigonométrique de type sinus ou cosinus. Si les entités présentent une variabilité naturelle de forme, les valeurs associées à ces paramètres peuvent être définies sous la forme d'une densité de probabilité.

Les simulations par objet sont généralement supérieures aux méthodes pixel pour recréer des entités géologiques de géométrie complexe. En outre, chaque objet possède son propre système de coordonnées local qui permet de reconstituer son espace de dépôt; cette propriété est extrêmement utile pour appliquer aux objets des algorithmes basés sur des règles génétiques. Cependant, les méthodes objet souffrent de limitations en termes de conditionnement: les stratégies adoptées optimisent typiquement la position des objets dans l'espace de la grille de simulation, ce qui s'avère souvent coûteux en temps de calcul et limite la qualité du conditionnement.

La plupart des méthodes objet suivent l'algorithme général suivant: (les étapes optionnelles sont indiquées entre parenthèses)

Tant que la densité ou le nombre d'objets voulu n'est pas atteint **Faire**

1. Définition de la forme d'un nouvel objet représentant l'entité géologique, typiquement à l'aide d'un tirage de Monte Carlo dans la distribution de chacun des paramètres descriptifs;
- (2.) Application des algorithmes utilisant le système de coordonnées interne à l'objet (par exemple la modélisation de dépôts de sédiments drapés sur le fond d'un chenal);

3. Positionnement de l'objet dans un système de coordonnées global (orientation et dimension). Ce positionnement est souvent optimisé pour satisfaire au mieux les données de terrain, généralement à l'aide de la technique du recuit simulé ou d'un algorithme inverse de type Metropolis-Hastings ;
4. Rastérisation et stockage de l'objet dans une grille de simulation. Si l'objet entre en collision avec d'autres objets sur certains nœuds de la grille, la simulation utilise des règles de troncation prédéfinies pour résoudre le conflit.

Fin Tant que

1.3 Exemple de création de géomodèle : la caractérisation d'un réservoir

En géosciences, les incertitudes affectent différentes étapes de la chaîne de traitement des informations et sont donc propagées d'étape en étape, jusqu'à affecter la conclusion finale. Nous proposons dans cette section un résumé des principales étapes permettant d'évaluer le potentiel associé à un réservoir pétrolier (Figure 1.1), afin de resituer nos approches dans un contexte plus général.

1.3.1 Etapes de caractérisation d'un réservoir

Définition d'un prospect La caractérisation d'un réservoir commence généralement par la définition d'un *prospect*, c'est-à-dire d'une zone dont le contexte tectonique et sédimentaire semble favorable à la formation d'hydrocarbures. Ces prospects sont définis à l'échelle régionale, le plus souvent d'après des études menées précédemment sur la région ou grâce à la connaissance de milieux analogues exhaustivement étudiés.

Collecte des données Après la définition d'un prospect et l'obtention des droits sur ce bloc, des campagnes sismiques sont menées à une échelle plurikilométrique pour repérer les principales structures géologiques. Si des pièges structuraux sont repérés et jugés dignes d'intérêt, des puits d'exploration sont forés en nombre limité pour confirmer la présence d'hydrocarbures et obtenir les premières informations sur les propriétés pétrophysiques des roches. Les données brutes présentent généralement un degré d'incertitude assez faible, typiquement du à l'imprécision des outils de mesure.

Interprétation des données Les données brutes sont ensuite interprétées pour pouvoir générer un premier modèle conceptuel du réservoir. Généralement, cette étape repose sur une conversion des cubes sismique des coordonnées de temps aux coordonnées de profondeur puis sur un pointé des réflecteurs sismiques, ainsi que sur une interprétation des

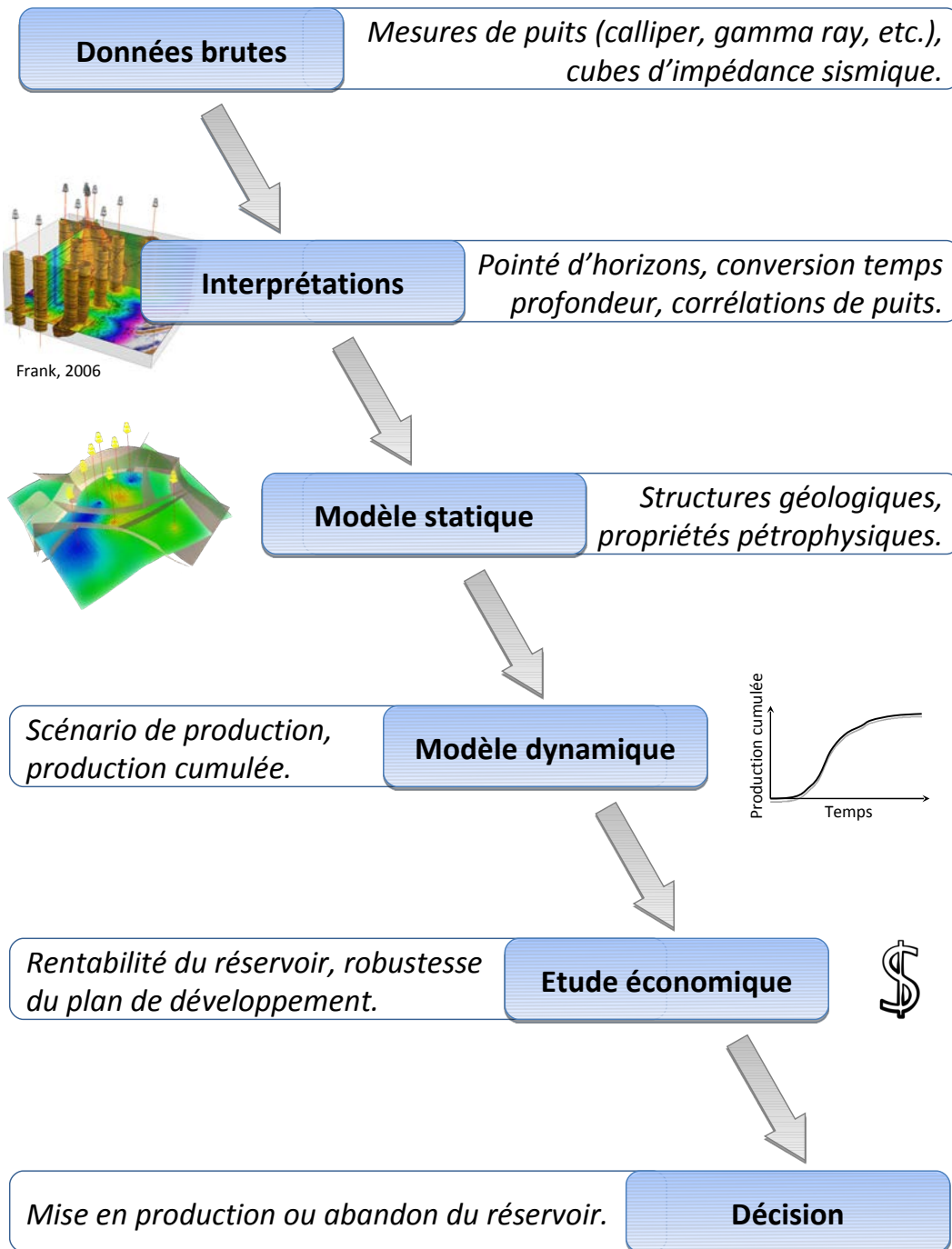


Fig. 1.1 – Chaîne de traitement pour la caractérisation du potentiel associé à un réservoir.

environnements de dépôt d'après les logs de puits. Il arrive également que des corrélations soient effectuées entre plusieurs puits pour relier des marqueurs identifiés comme appartenant à un même évènement géologique.

L'évaluation du contexte géologique est généralement réalisée par un expert géologue. Cette tâche repose donc sur des jugements subjectifs – il n'est pas rare que plusieurs personnes produisent ainsi des interprétations géologiques différentes à partir d'un même faisceau d'indices. [Bond et al. \(2007\)](#) donnent un excellent exemple de ce phénomène dans le cadre d'une interprétation sismique : en demandant à un panel de 412 experts de reconstituer un style tectonique à partir d'une coupe sismique, ils obtiennent six familles d'interprétations dont certaines sont diamétralement opposées. [Bond et al. \(2007\)](#) donnent à ce type d'incertitudes le nom d'*incertitudes conceptuelles*. Les incertitudes conceptuelles peuvent théoriquement être quantifiées en soumettant les données à un grand nombre d'experts et en confrontant leurs interprétations. Cependant, le temps et le nombre d'experts disponibles pour une tâche est généralement restreint ; en pratique, ces incertitudes sont donc largement sous-échantillonnées.

Création du modèle statique Le modèle conceptuel des processus de dépôt et de transformation des roches constituant le réservoir fournit des informations précieuses pour créer un modèle numérique statique du réservoir, représentant l'état actuel du sous-sol. Ce modèle représente essentiellement trois types d'informations : les structures géologiques, les faciès de dépôt et les propriétés pétrophysiques associées. Pour construire une représentation adaptée de ces informations, seules des données incomplètes (mesures de puits), imprécises (cubes sismiques interprétés) ou conceptuelles sont disponibles ; de ce fait, le modèle est généralement soumis à de fortes incertitudes. Pour caractériser ces incertitudes, il est fréquent de construire le modèle à l'aide de méthodes de simulation stochastique (Section 1.2) qui permettent de générer un grand nombre de représentations plausibles du sous-sol.

Création du modèle dynamique Après avoir caractérisé le modèle statique de manière satisfaisante, il est possible d'étudier l'évolution du réservoir (pression, saturations en pétrole et en eau, *etc.*) en réponse à un plan de production prédéfini. Cette étude nécessite de connaître les propriétés dynamiques des fluides présents dans le réservoir afin de réaliser des simulations d'écoulement. En pratique, l'évolution de la viscosité des hydrocarbures en fonction de la pression et de la température est rarement connue avec précision ; il est fréquent de caractériser ces paramètres avec des lois de distribution arbitraires et d'échantillonner ces incertitudes à l'aide de plusieurs simulations d'écoulement. Les simulations d'écoulement sont cependant des processus coûteux en termes de puissance de calcul ; de nombreux auteurs proposent des solutions alternatives permettant de limiter le nombre de simulations, telles que les surfaces de réponse ([Zabalza-Mezghani et al., 2004](#); [Fetel, 2006](#); [Fetel et Caumon, 2008](#)) ou les plans d'expérience adaptatifs ([Scheidt et al., 2007](#)).

Etude économique Le modèle dynamique donne accès à la production cumulée d'hydrocarbures pouvant être extraits du réservoir au cours du temps. Ces informations peuvent être utilisées pour analyser la rentabilité du réservoir, d'après différents paramètres économiques tels que les évolutions prévues du cours du baril ou le coût du transport des hydrocarbures. Par ailleurs, les incertitudes associées au modèle dynamique permettent d'évaluer la robustesse économique du projet (Corre *et al.*, 2000), c'est-à-dire la manière dont la rentabilité du projet est influencée par différents paramètres. Cette dernière information est précieuse pour minimiser le risque financier en ajustant certains choix de développement du réservoir. Enfin, les conclusions de l'étude économique sont confrontées aux fonds d'investissement disponibles et à l'intérêt des autres projets pour décider si le projet doit passer en phase de production, être temporairement mis de côté, ou être définitivement abandonné.

1.3.2 Propagation des incertitudes

Dans la chaîne de caractérisation d'un réservoir, chaque étape utilise implicitement ou explicitement les résultats obtenus lors des étapes précédentes. Cette organisation induit donc une propagation et une accumulation des incertitudes sur l'ensemble de la chaîne. Différentes stratégies ont été proposées pour représenter cette propagation ; nous présentons les approches les plus fréquemment adoptées en nous basant sur les travaux de Massonnat (2000), Corre *et al.* (2000) et Charles *et al.* (2001).

Lorsque les incertitudes sont indépendantes, il est possible d'évaluer les variances associées de manière séparée et de les additionner pour obtenir les incertitudes jointes. Cette propriété est notamment utilisée pour propager les incertitudes dues à la conversion temps-profondeur et au piqué d'horizons sur un cube sismique. Dans le cas général, l'indépendance est cependant une hypothèse forte qui est rarement acceptable en géosciences. Pour propager les incertitudes d'étape en étape, de nombreux auteurs proposent d'échantillonner exhaustivement chaque source d'incertitudes à l'aide de méthodes stochastiques ; l'ensemble des réalisations ainsi généré peut alors être réutilisé dans l'étude de l'étape suivante. Cette stratégie a notamment été adoptée pour la modélisation des incertitudes dans le package professionnel JACTATM/ GOCAD.

Cette approche pose néanmoins un certain nombre de problèmes pratiques, essentiellement dus à l'accroissement exponentiel des ressources nécessaires à l'exploration combinatoire de l'ensemble des réalisations. Prenons un exemple simple où on souhaite étudier le volume total des hydrocarbures présents dans un réservoir. Pour y parvenir, il est nécessaire de modéliser successivement la structure, puis les faciès, et enfin la porosité et la saturation des roches ; pour chaque entité étudiée, admettons que mille réalisations soient nécessaires pour échantillonner correctement les incertitudes associées. En ce cas, il existe $1000^4 = 10^{12}$ combinaisons possibles entre les réalisations, soit un billion de possibilités. Le stockage et la manipulation d'un tel volume de données sont rarement possibles dans

des applications réelles ; par ailleurs, il est fréquent que de plus nombreuses sources d'incertitudes doivent être prises en compte. Nous présentons ci-dessous différentes stratégies permettant de s'affranchir de ces contraintes.

Décimation des réalisations La méthode la plus intuitive pour réduire le nombre de combinaisons possibles consiste à décimer les réalisations à chaque étape. Cette technique repose sur l'idée que certaines réalisations correspondent à l'expression d'une même tendance et sont donc largement redondantes. En identifiant les réalisations similaires, il est donc acceptable de ne conserver qu'un sous-ensemble des réalisations les plus typiques – notamment les réalisations optimistes, pessimistes et médianes.

En pratique, la classification des réalisations de la plus optimiste à la plus pessimiste repose sur l'évaluation d'une fonction objective dont le comportement est intimement lié à la fonction de transfert à utiliser. Par exemple, il est possible de confronter les quantités d'hydrocarbures initialement en place pour approcher la production associée à chacune des réalisations ; cette valeur objective peut alors être utilisée pour limiter le nombre de simulations d'écoulements à effectuer.

Stratégies de stockage Même après décimation, le nombre élevé de réalisations peut s'avérer extrêmement encombrant en termes de stockage. Pour résoudre ce problème, le package JACTATM/ GOCAD utilise une stratégie de génération des réalisations à la volée lorsque les données qui leur sont associées sont nécessaires. Il est ainsi possible de ne stocker en mémoire que les paramètres de la simulation stochastique ; bien que cette technique réduise significativement la vitesse d'accès aux réalisations, elle permet un gain considérable en termes de volume de stockage.

1.4 Métriques d'incertitude

La génération d'un grand nombre de réalisations à l'aide de méthodes géostatistiques permet d'échantillonner l'ensemble des scénarios géologiquement plausibles.² Cependant, l'interprétation de ces réalisations est délicate du fait de la forte cardinalité des ensembles de réalisations géostatistiques. Pour y parvenir, une approche possible est de générer des *métriques d'incertitude* ; ces dernières sont calculées sous la forme de grandeurs scalaires représentatives de la dispersion locale au sein des réalisations.

²En pratique, l'ensemble des scénarios est rarement parfaitement échantillonné, du fait de la haute dimension de l'espace des incertitudes et des capacités de calcul limitées des ordinateurs. Il est néanmoins possible de s'en rapprocher de manière suffisante pour informer les décideurs du degré d'incertitude associé aux données.

1.4.1 Génération d'une densité de probabilité locale

Certaines méthodes d'estimation, telles que le krigeage (Krige, 1966; Matheron, 1969; Gotway *et al.*, 1996; Stein, 1999), présentent conjointement la valeur estimée et l'erreur d'estimation associée. Cette technique repose cependant sur des modèles mathématiques spécifiques à la technique d'estimation, qui ne sont pas nécessairement disponibles lorsque la fonction d'estimation reproduit des phénomènes hautement non linéaires et croit donc en complexité. Dans ce cas, il est nécessaire d'adopter une approche différente pour quantifier les incertitudes locales.

De manière générale, il est possible de générer une distribution locale en tout point du domaine d'intérêt d'après l'ensemble des valeurs associées aux simulations stochastiques satisfaisant le problème inverse; cette distribution est alors entièrement définie par les données et le choix d'un type de simulation lorsque le nombre des réalisations tend vers l'infini. En pratique, seul un nombre discret de réalisations permet d'échantillonner la distribution locale théorique; cette échantillonnage fournit donc une approximation de la densité de probabilité étudiée, dont on suppose que l'erreur décroît avec le nombre d'échantillons (Journel, 1985).

Cette approche suppose qu'en tout point du domaine d'intérêt, il est possible (i) de comparer les réalisations géostatistiques entre elles et (ii) de stocker l'information associée à leur dispersion locale. Certains types de simulations ne remplissent pas ces pré-requis et doivent être traitées de manière différente. Par exemple, une simulation testant la connectivité entre les différentes zones d'un géomodèle présente une très haute dimensionnalité et pose notamment des problèmes de stockage (en chaque nœud, il est nécessaire de stocker l'ensemble des densités de probabilité locales associées à la connectivité avec tous les autres nœuds du modèle). En outre, dans le cadre de simulations générant un ensemble de géométries possibles d'une surface géologique, les informations de correspondance spatiale entre deux surfaces différentes ne sont pas nécessairement maintenues; dans le cas contraire, il peut n'exister aucun moyen de stocker l'information de dispersion en certains points (Figure 1.2).

Par ailleurs, cette approche induit une première simplification de l'ensemble des réalisations : chaque position spatiale étant considérée de manière indépendante, les corrélations spatiales existant au sein des réalisations sont ignorées lors de la création de la distribution locale.

1.4.2 Statistiques d'incertitude locale

Cette section présente quelques métriques fréquemment utilisées en statistiques, et présentant un lien avec la dispersion au sein d'un ensemble de données. Ces dernières permettent de convertir une densité de probabilité en une grandeur scalaire représentative; il s'agit donc d'une simplification de l'ensemble des données.

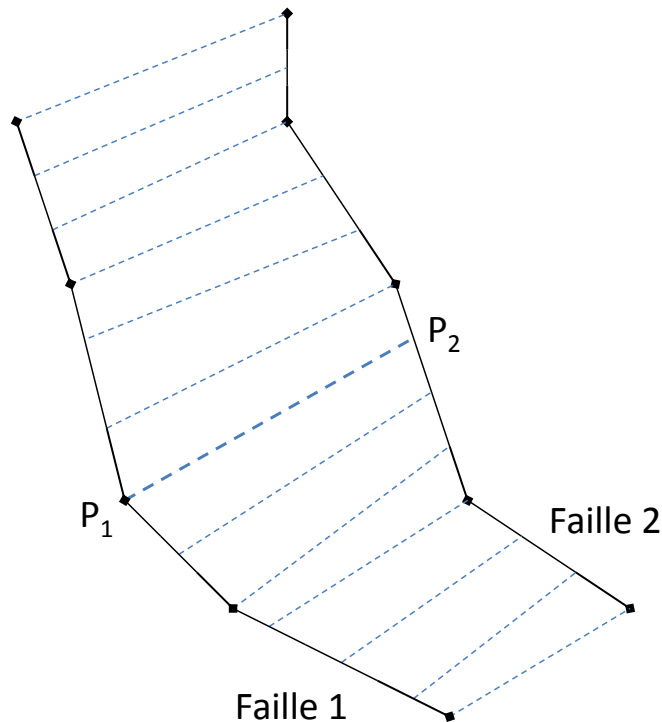


Fig. 1.2 – Illustration des correspondances spatiales entre deux simulations de faille. Le point P_1 sur la première simulation correspond au point P_2 sur la seconde simulation ; une mesure de dispersion peut donc être calculée localement entre les points 1 et 2 (par exemple comme la distance entre les deux points). Cependant, il n'existe aucun nœud sur lequel stocker cette dispersion au niveau de la seconde faille.

Ecart type La variance σ^2 est une des métriques d'incertitude les plus fréquemment utilisées (Chen *et al.*, 2010). Elle correspond au moment d'ordre 2 d'une population de données :

$$\sigma^2 = \frac{1}{N} \cdot \sum_{i=1}^N (x_i - \mu)^2 \quad (1.3)$$

où σ est l'écart type, N est le nombre d'individus au sein de la population, x_i est le $i^{\text{ème}}$ individu et μ est la moyenne des individus.

L'écart type est particulièrement adapté pour quantifier la dispersion de populations dont la distribution est Gaussienne, mais représente moins bien des populations de distributions plus complexes, par exemple multimodales. Par ailleurs, il suffit de quelques individus "aberrants" au sein d'une population homogène pour affecter grandement la valeur de l'écart type, même s'ils sont en nombre très restreint : on qualifie l'écart type d'opérateur *non robuste*.

Coefficient de variation L'écart type repose sur une mesure de l'écart au carré entre l'ensemble des valeurs et leur moyenne. De ce fait, à dispersion égale, un ensemble de

Tab. 1.2 – Effet de l'intervalle des valeurs sur l'écart type et le coefficient de variation.

Ensemble de valeurs	Ecart type	Coefficient de variation
$\Omega_0 = \{1, 3, 4\}$	$\sigma_0^2 = 1,56$	$C_{V0} = 0,47$
$\Omega_1 = 2 \times \Omega = \{2, 6, 8\}$	$\sigma_1^2 = (2 \times \sigma_0)^2 = 6,22$	$C_{V1} = C_{V0} = 0,47$

valeurs élevées présente un écart type plus importante qu'un ensemble de valeurs faibles (Table 1.2). Le coefficient de variation C_V est souvent proposé pour compenser cet effet : il s'agit d'une mesure de l'écart-type normalisée par la valeur moyenne des mesures.

$$C_V = \frac{\sigma}{\mu} \quad (1.4)$$

Le coefficient de variation n'est défini que pour une moyenne μ différent de zéro. Par ailleurs, le coefficient de variation souffre des mêmes limites que l'écart type en termes de robustesse et de capacité à représenter des distributions complexes.

Quantiles Un quantile est la valeur $Q(i)$, $i \in [0, 1]$, telle que, sur l'ensemble des valeurs d'une population Ω , une proportion i des valeurs sont inférieures à $Q(i)$:

$$Q(i) : i \mapsto Q(i) \quad \text{avec} \quad \frac{\text{Card}(x < Q(i), x \in \Omega)}{\text{Card}(\Omega)} = i \quad (1.5)$$

où $\text{Card}()$ est le cardinal d'un ensemble, c'est-à-dire le nombre d'éléments présents dans cet ensemble. On parle de *quartiles* pour indiquer des quantiles multiples de 0,25, et de *centiles* s'ils sont multiples de 0,01.

Le quantile peut être utilisé pour comparer une valeur à un ensemble Ω , c'est-à-dire pour savoir si cette valeur est optimiste, pessimiste ou médiane. Le quantile est une grandeur robuste qui peut être appliquée à tout type de distribution.

Ecart interquartile L'écart interquartile *IQR* mesure l'écart entre le premier et le troisième quartile; il quantifie donc l'intervalle de valeurs autour de la médiane dans lequel 50% des échantillons d'une population se trouvent (Upton et Cook, 1996) :

$$IQR = Q(0,75) - Q(0,25) \quad (1.6)$$

Plus l'écart interquartile est faible, plus une population est resserrée autour de sa médiane, donc moins il y a d'incertitudes sur les valeurs possibles associées à cette population. L'écart interquartile peut être généralisé en écart interquantile :

$$IQR(P) = Q(1 - \frac{P}{2}) - Q(\frac{P}{2}) \quad \text{avec} \quad P \in [0, 1[\quad (1.7)$$

L'écart interquantile est relativement coûteux à calculer, car il nécessite de trier l'ensemble des valeurs de la population – or, la complexité algorithmique des meilleurs algorithmes de tri est en moyenne $o(n^2)$, où n est le nombre d'éléments à trier. Par ailleurs, l'écart interquantile est une grandeur robuste, qui n'est que peu affectée par des valeurs aberrantes.

Résidu Un résidu Res correspond à l'écart entre une valeur particulière et une valeur de référence, qui indique souvent une tendance dans les données. En pratique, la valeur de référence peut par exemple être considérée comme la valeur moyenne μ ou la médiane $Q(0, 5)$ associée à l'ensemble Ω , et la valeur particulière peut être prise au sein de Ω :

$$Res_\mu = x - \mu \quad \text{avec } x \in \Omega \quad (1.8)$$

Cette métrique permet de quantifier l'écart entre un individu et la valeur la plus probable au sein de l'ensemble, donc de savoir si la valeur de l'individu lui-même est probable ou non.

1.4.3 Validité de l'échantillonnage des incertitudes

La validité de l'échantillonnage au sein de l'espace des incertitudes est un problème récurrent en géosciences. Cependant, ce problème est souvent approché en faisant un nombre de réalisations considéré de manière subjective comme "grand", sans connaître le nombre minimal de réalisations à effectuer en réalité. Il est donc tout à fait possible de perdre du temps de calcul en générant trop de réalisations, ou de sous-échantillonner l'incertitude par un nombre trop faible de réalisations.

Validité et métriques d'incertitude La notion de métrique d'incertitude permet d'apporter un cadre théorique pour valider ou invalider un échantillonnage d'incertitudes : le nombre de réalisations est considéré comme suffisant si la métrique d'incertitude converge (Prokhorov, 1956; Van der Vaart et Wellner, 1996), c'est-à-dire que rajouter une réalisation ne change pas significativement la métrique (Corre *et al.*, 2000; Goovaerts, 2006). Les métriques d'incertitude étant des grandeurs locales, il est possible de n'atteindre la convergence qu'en certains points de l'espace. En ce cas, l'échantillonnage n'est valable que sur l'ensemble de ces points.

Convergence d'une métrique La convergence locale C_ν d'une métrique d'incertitude ν peut être évaluée sous la forme suivante :

$$C_\nu = |\nu(\Omega') - \nu(\Omega)| \quad \text{avec } \Omega \subset \Omega' \quad (1.9)$$

où Ω est l'ensemble constitué des K premières réalisations et Ω' est l'ensemble constitué des L premières réalisations, $L > K$. La convergence est considérée comme atteinte si C_ν est localement inférieur à un seuil ϵ . En pratique, le calcul de C_ν est soumis à plusieurs limitations :

1. L'évolution de la valeur de la convergence est sensible à l'ordre selon lequel les réalisations sont agencées ; par exemple, les ensembles $\Omega = \{R_1, R_2\}$ et $\Omega' = \{R_1, R_2, R_3\}$ ne produiront pas une même valeur de convergence que les ensembles $\Omega = \{R_1, R_3\}$

- et $\Omega' = \{R_1, R_3, R_2\}$. Il faut donc que le critère de convergence soit validé entre plusieurs changements d'ensembles successifs, c'est-à-dire que $|\nu(\Omega') - \nu(\Omega)| < \epsilon$, et $|\nu(\Omega'') - \nu(\Omega')| < \epsilon$, *etc.*
2. Les métriques ν non robustes peuvent être fortement affectées par l'ajout d'une valeur extrême, et convergent donc beaucoup plus lentement que les métriques robustes. Il peut arriver que la convergence soit impossible en un temps de calcul raisonnable.
 3. La stratégie permettant de passer de l'ensemble Ω à l'ensemble Ω' est purement empirique : la méthode la plus intuitive consiste à ajouter les réalisations une à une, mais il est également possible de les ajouter par paquets dont la taille est fonction des choix de l'utilisateur. L'ajout de réalisations par paquets permet de limiter le nombre d'évaluations de la métrique ν , et ainsi d'optimiser le temps de calcul si cette métrique est coûteuse à évaluer. Par ailleurs, une telle stratégie permet également d'améliorer la robustesse de la convergence.

1.5 Conclusion

En géosciences, la visualisation des incertitudes spatiales est intimement liée aux méthodes de génération de réalisations multiples en géostatistique, qui traduisent la non-unicité de la solution au vu des informations disponibles. L'objectif de ce chapitre était de présenter les idées fondatrices associées à la caractérisation des incertitudes, ainsi que les techniques adoptées en pratique pour répondre à ce problème. L'ensemble de ces notions est exploité dans cette thèse, de manière directe ou indirecte, pour intégrer nos algorithmes de visualisation des incertitudes dans un contexte plus général.

Nous avons un premier temps présenté un résumé des différents concepts reliés à l'incertitude. La notion d'incertitudes spatiales est en effet généralement mal définie et peut présenter différents sens tels que l'erreur, la complétude, l'imprécision, *etc.* Tous ne sont cependant pas intéressants dans une problématique de géomodélisation ; dans cette thèse, nous nous focalisons sur les incertitudes *quantifiables* sous la forme d'une fonction de densité de probabilité, notamment les incertitudes portant sur les erreurs de mesure, la précision des données et la complétude des informations.

Ces incertitudes peuvent être caractérisées dans le cadre d'une démarche inverse ou à l'aide de simulations stochastiques, qui permettent de générer un grand nombre de représentations plausibles du sous-sol, nommées réalisations, tout en respectant les observations collectées *in-situ* et la connaissance conceptuelle des processus géologiques mis en jeu. Cette approche permet également de propager les incertitudes dans les différentes étapes de la chaîne de géomodélisation, à condition d'être capable de maîtriser l'explosion combinatoire qui résulte de sources d'incertitude multiples.

Pour chaque position spatiale du domaine d'intérêt, la fonction de densité de probabilité associée à la grandeur étudiée peut être approchée par la distribution locale des réalisations. La visualisation d'une fonction de densité de probabilité peut cependant être extrêmement complexe ; de ce fait, nous simplifions cette densité sous la forme d'une unique valeur scalaire, à l'aide d'une métrique d'incertitude adaptée. Cette approche permet également de contrôler le nombre de réalisations nécessaires pour échantillonner correctement les incertitudes spatiales ; en effet, l'échantillonnage peut être validé lorsque la métrique d'incertitude converge.

Chapitre 2

Etat de l'art des techniques de visualisation des incertitudes

Sommaire

2.1	Méthodes de visualisation scientifique	26
2.1.1	Extraction d'isosurfaces	26
2.1.2	Rendu volumique	28
2.2	Cadre théorique de visualisation des incertitudes	32
2.2.1	Variables visuelles	32
2.2.2	Qualité des métaphores visuelles	33
2.2.3	Interférences visuelles	35
2.2.4	Préattentivité	35
2.2.5	Concentration de l'information	36
2.3	Méthodes de visualisation des incertitudes	36
2.3.1	Incertitudes sur des graphiques	37
2.3.2	Incertitudes en archéologie	38
2.3.3	Incertitudes de position particulière dans des écoulements	40
2.3.4	Incertitudes sur la reconstruction de surfaces	41
2.3.5	Incertitudes en médecine	45
2.3.6	Incertitudes en climatologie	49
2.3.7	Incertitudes dans les systèmes d'information géographiques	51
2.3.8	Incertitudes en géomodélisation	55
2.4	Evaluation de la qualité des visualisations	56
2.5	Conclusion	58

L'omniprésence des incertitudes dans le domaine des géosciences requiert des outils adaptés pour les gérer. Nous avons vu précédemment comment modéliser des incertitudes sur une grandeur du sous-sol ; ces informations doivent ensuite être présentées de la manière la plus lisible possible. Dans la majorité des applications, les géomodèles et

leurs incertitudes associées sont présentées de manière indépendante. Pratiquement, cette méthode impose au lecteur une constante gymnastique cérébrale : pour chaque information prise en compte, il doit également aller vérifier sur une carte séparée son degré de fiabilité. Il en résulte naturellement une charge cognitive et perceptuelle accrue lors de l'interprétation des données (Harrower, 2003). Cependant, des techniques spécifiques à la visualisation des incertitudes existent pour présenter conjointement données et incertitudes, bien qu'elles restent encore largement sous-utilisées en géosciences, à l'exception notable des systèmes d'information géographiques (SIG). Après un rappel sur les méthodes générales de visualisation dans le domaine scientifique (Section 2.1), nous introduisons les critères théoriques assurant une bonne perception des modèles (Section 2.2). Nous établissons ensuite un inventaire des méthodes de visualisation des incertitudes existantes (Section 2.3) et concluons par une discussion sur la qualité de ces méthodes d'après les résultats empiriques d'études sur des utilisateurs (Section 2.4).

2.1 Méthodes de visualisation scientifique

L'évolution des techniques de visualisation 3D a principalement été motivée par les besoins de l'industrie du jeu et du cinéma d'animation ; cependant, les applications scientifiques bénéficient également des percées technologiques en matière de matériel graphique. Nous détaillons dans cette section deux des méthodes de visualisation scientifique utilisées pour répondre à une grande variété de problèmes : l'*extraction d'isosurfaces* et le *rendu volumique* (Figure 2.1). Chacune de ces méthodes présente un certain nombre de variantes ou d'optimisations. Plutôt qu'une revue exhaustive, nous nous contentons de discuter les plus importantes pour la problématique de la visualisation scientifique en général, et de la visualisation des incertitudes en particulier.

2.1.1 Extraction d'isosurfaces

Une isosurface \mathcal{I} est une frontière constituée de l'ensemble des points (x, y, z) de l'espace 3D le long desquels un champ scalaire \mathcal{S} est constant et égal à une valeur ω , nommée *isovaleur*. Son équation mathématique est donc :

$$\mathcal{I}(\omega) = \{(x, y, z) \in \mathbb{R}^3 \mid \mathcal{S}(x, y, z) - \omega = 0\} \quad (2.1)$$

De manière pratique, les isosurfaces sont généralement extraites à partir de maillages tridimensionnels sous la forme de surfaces polygonales. Deux familles d'algorithmes peuvent être utilisées lors de l'extraction : les algorithmes se déplaçant le long des arêtes des cellules grâce aux liens topologiques sommets/arêtes, et les algorithmes utilisant des tables indexées.

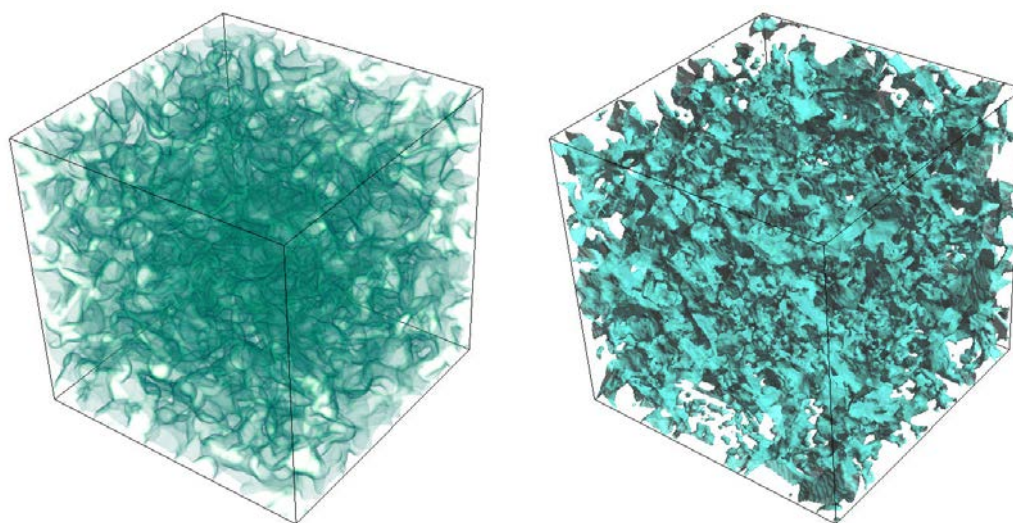


Fig. 2.1 – Visualisation d’un milieu poreux par rendu volumique (gauche) et par isosurface (droite). Le milieu poreux est stocké dans une grille Cartésienne de $200 \times 200 \times 200$ cellules sous forme de propriété booléenne. Données fournies par l’Imperial College, consortium sur la modélisation à l’échelle du pore.

Algorithmes à base de liens topologiques explicites Les algorithmes à base de liens topologiques se déplacent le long des arêtes des cellules grâce aux liens topologiques entre sommets, arêtes et faces (Bloomenthal, 1988; Lévy *et al.*, 2001). Ils testent chaque arête d’une cellule jusqu’à trouver une arête intersectée par l’isosurface. Si une telle arête existe, l’algorithme considère qu’il s’agit du point d’entrée de l’isosurface dans la face. Il tourne alors autour de la face et teste successivement ses arêtes jusqu’à trouver une nouvelle arête intersectée, qui sera l’arête de sortie, puis réitère l’opération sur la face voisine (Figure 2.2).

Les algorithmes utilisant les liens topologiques peuvent traiter tout type de cellule, mais présentent des performances faibles du fait de la lourdeur des structures stockant la topologie et des nombreux accès mémoires induits par la propagation d’arête en arête et de face en face. Dans les grilles à maille prédéfinie – et notamment dans les grilles structurées – les algorithmes utilisant des tables d’index permettent d’accélérer fortement l’extraction de l’isosurface.

Algorithmes utilisant des tables indexées Les algorithmes utilisant des tables indexées précalculent l’ensemble des configurations possibles en fonction de la polarité de la cellule (Lorensen et Cline, 1987; Treece *et al.*, 1999; Lewiner *et al.*, 2003). De manière générale, une cellule composée de N sommets présente 2^N configurations possibles. Par inversion de la polarité, la table des cas peut être ramenée à 2^{N-1} entrées. Certains types de cellules présentent des axes de symétrie, qui permettent de réduire encore davantage

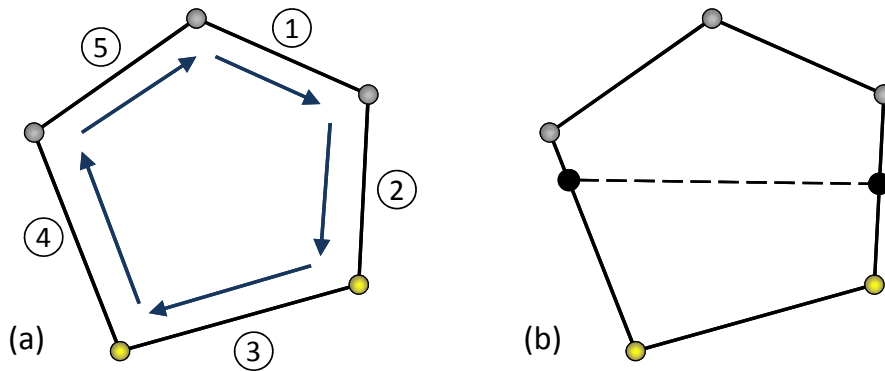


Fig. 2.2 – Extraction d'isosurface à partir des liens topologiques entre sommets, arêtes et faces. Les sommets de couleur grise présentent une valeur inférieure à l'isovaleur ω , et ceux de couleur jaune une valeur supérieure à ω . (a) Face polygonale d'une cellule tridimensionnelle. Les flèches représentent les demi-arêtes. (b) Ligne d'intersection entre l'isosurface et la face.

le nombre de configurations à traiter.

Les algorithmes les plus populaires sont les *Marching Tetrahedra* pour les grilles tétraédriques et les *Marching Cubes* pour les grilles hexaédriques (Figure 2.3). Caumon *et al.* (2005) et Buatois (2008) proposent une extension capable de traiter tout type de grille, nommée *Marching Cells*. Cet algorithme repose sur un inventaire des types de cellules rencontrées dans le maillage; pour chaque type de cellule, une table d'index est précalculée à l'aide des liens topologiques du type courant.

Domaine d'application Les isosurfaces présentent un grand intérêt en visualisation scientifique; elles sont des candidates idéales pour toute application s'intéressant à un seuil de valeurs au sein d'un champ scalaire, mais peuvent également s'avérer utiles pour d'autres applications du fait de leur grande lisibilité. Cependant, elles se limitent à un sous-ensemble du domaine étudié, alors que le domaine doit parfois être considéré dans sa totalité. Dans ces cas, il est préférable d'utiliser des méthodes de rendu volumique pour visualiser le volume dans son ensemble; les méthodes de rendu volumique peuvent être utilisées conjointement aux méthodes d'extraction d'isosurface.

2.1.2 Rendu volumique

Les méthodes de rendu volumique ont pour but de rendre la totalité du volume étudié. Ces méthodes représentent le volume comme un milieu semi-transparent constitué de particules élémentaires émettant et absorbant de la lumière. Le rendu repose alors sur un calcul de l'accumulation de la couleur le long des rayons visés, combiné à une méthode de discrétisation spatiale du volume.

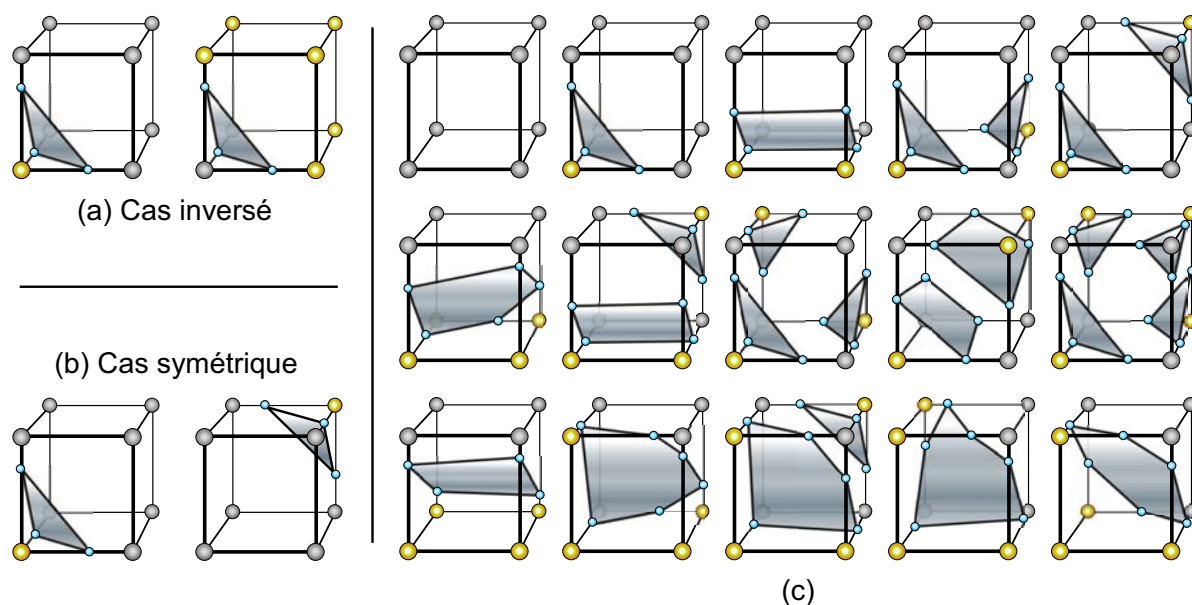


Fig. 2.3 – Table des cas dans l’algorithme des Marching Cubes d’après *Lorensen et Cline (1987)*. Les sommets de couleur grise présentent une valeur inférieure à l’isovaleur ω , et ceux de couleur jaune une valeur supérieure à ω . (a) Une inversion de la polarité ne change pas la topologie de l’isosurface. (b) La symétrie des cellules peut être exploitée pour réduire le nombre de configurations possibles. (c) Ensemble des configurations possibles, après réduction par symétrie et par inversion de polarité. Adapté d’après *Buatois (2008)*.

L’intégrale de couleur L’intégrale de couleur correspond au calcul de l’accumulation de la couleur lors du rendu volumique. L’intégrale s’appuie sur une description des règles selon lesquelles la couleur s’accumule, nommée *modèle optique*, qui est le plus souvent basée sur les lois physiques du transport des ondes lumineuses. Le modèle optique le plus répandu est le modèle d’*émission-absorption* (*Williams et Max, 1992*) ; il existe cependant de nombreux autres modèles optiques, qui traduisent différents mécanismes de transport de la lumière tels que la diffusion ou l’ombrage (*Max, 1995*; *Kniss et al., 2003*).

Le modèle optique d’émission-absorption introduit les notions de *densité optique* ρ_λ et de *chromaticité* κ_λ , où λ est la longueur d’onde considérée. La densité optique (également nommée *absorbance* en physique) et la chromaticité correspondent à la quantité de lumière ayant pour longueur d’onde λ , qui est respectivement absorbée et émise sur une distance infinitésimale au sein du volume. Ces deux grandeurs peuvent varier spatialement au sein du volume, généralement à l’aide d’une fonction de transfert qui dépend elle-même d’un champ scalaire tiers (Figure 2.4). La fonction de transfert permet ainsi de mettre l’emphase sur certaines gammes de valeurs (un signal de Dirac produira par exemple l’équivalent d’une isosurface).

Etant donné un rayon visuel paramétrisé par la distance t à l’observateur, avec t_n le point d’entrée du rayon visuel au sein du volume et t_0 le point de sortie (Figure 2.5),

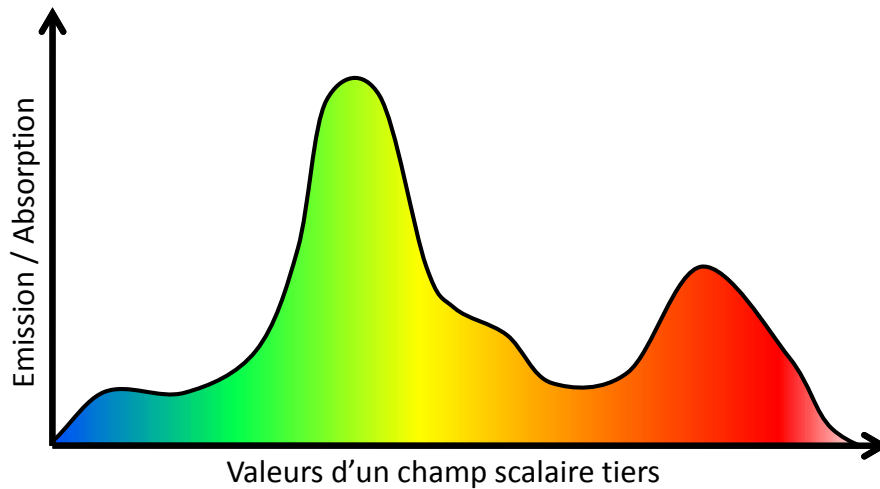


Fig. 2.4 – Fonction de transfert visuelle définissant l'émission et l'absorption particulière en fonction d'un champ scalaire. L'émission est définie par les composantes RGB de la couleur, et l'absorption par l'opacité. Adapté d'après *Buatois (2008)*.

l'intensité $I_\lambda(t_n)$ de la couleur ayant pour longueur d'onde λ qui arrive à l'observateur peut se formaliser comme suit dans le cadre du modèle d'émission-absorption (*Williams et Max, 1992*) :

$$I_\lambda(t_n) = \int_{t_0}^{t_n} \exp\left(-\int_t^{t_n} \rho_\lambda(u) du\right) \cdot \kappa_\lambda(t) \cdot \rho_\lambda(t) dt + I_\lambda(t_0) \cdot \exp\left(-\int_{t_0}^{t_n} \rho_\lambda(t) dt\right) \quad (2.2)$$

où la grandeur $\exp\left(-\int_{t_0}^{t_n} \rho_\lambda(t) dt\right)$ correspond à la transmittance du volume, c'est-à-dire la fraction de la lumière de longueur d'onde λ qui traverse l'ensemble du volume pour parvenir à l'observateur.

Discrétisation de l'espace La majorité des méthodes de rendu volumique considèrent l'espace physique comme un milieu continu dont les caractéristiques sont obtenues par interpolation d'un ensemble discret de points connus. Cependant, le calcul de l'accumulation des couleurs repose sur l'intégration d'une composée de fonctions non linéaires (profils d'émission et d'absorption le long du rayon visuel), pour laquelle il est souvent impossible de trouver une solution analytique. Par ailleurs, même lorsqu'une telle solution existe, l'évaluation de l'intégrale peut nécessiter une puissance de calcul incompatible avec les besoins d'une visualisation interactive. Dans ces situations, l'approche la plus courante repose sur une discrétisation de l'intégrale de couleur le long du rayon visuel, de telle sorte que la fonction à intégrer puisse être considérée comme linéaire pour chaque intervalle. Cette discrétisation met en œuvre différentes stratégies d'échantillonnage de l'espace, dont on distingue deux grandes familles : les méthodes explicites et les méthodes implicites.

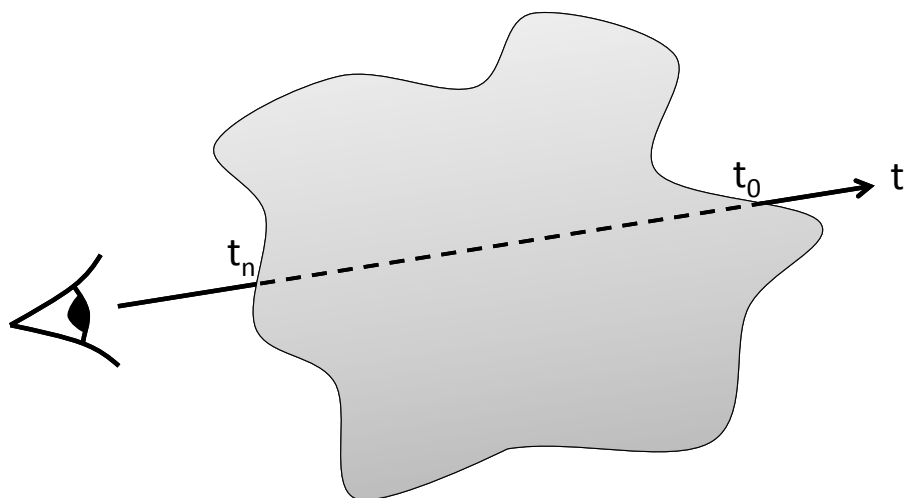


Fig. 2.5 – Rayon visuel paramétrisé par rapport à la distance t à l’observateur. Le point d’entrée du rayon dans le volume est noté t_n , et le point de sortie t_0 . Adapté d’après Williams et Max (1992).

Les méthodes explicites, également nommées *méthodes directes*, reposent sur un *lancer de rayon* dans la direction de vue pour chaque pixel de l’image (Kajiya et Von Herzen, 1984; Kay et Kajiya, 1986; Sabella, 1988; Levoy, 1988, 1990). Les propriétés visuelles du milieu sont alors échantillonnées régulièrement le long de ce rayon, et la contribution de chaque point d’échantillonnage est accumulée selon les règles imposées par le modèle optique. Les méthodes implicites reposent sur une projection des éléments sur l’écran, la contribution de chaque élément étant calculée d’après son volume et ses propriétés optiques. Différentes stratégies de projection existent, par exemple les méthodes par tranchage incrémental (Yagel *et al.*, 1996; Westermann et Ertl, 1998; Engel *et al.*, 2001; Ikits *et al.*, 2004, Figure 2.6), par décomposition des cellules en primitives graphiques (Shirley et Tuchman, 1990; Wittenbrink, 1999; Wylie *et al.*, 2002; Barbier et Bonneau, 2008) ou par approximation de la contribution des éléments à l’écran à l’aide de fonctions ellipsoïdales (Mao *et al.*, 1995; Mueller *et al.*, 1999; Zwicker *et al.*, 2001).

Quelle que soit la méthode de discrétisation, l’ordre d’affichage des primitives graphiques est critique du fait de la non-commutativité des opérateurs de composition (Porter et Duff, 1984). Les éléments doivent donc être triés par ordre de visibilité, c’est-à-dire par ordre d’occlusion par rapport à l’observateur (Williams, 1992; Comba *et al.*, 1999) – la visibilité est une notion complexe, qui est parfois approchée par la distance à l’observateur. L’ordre le plus fréquemment adopté se fait de l’élément le plus lointain à l’élément le plus proche du point de vue; cependant, l’ordre inverse peut permettre d’optimiser le temps de calcul en interrompant l’accumulation quand un seuil d’opacité est atteint.

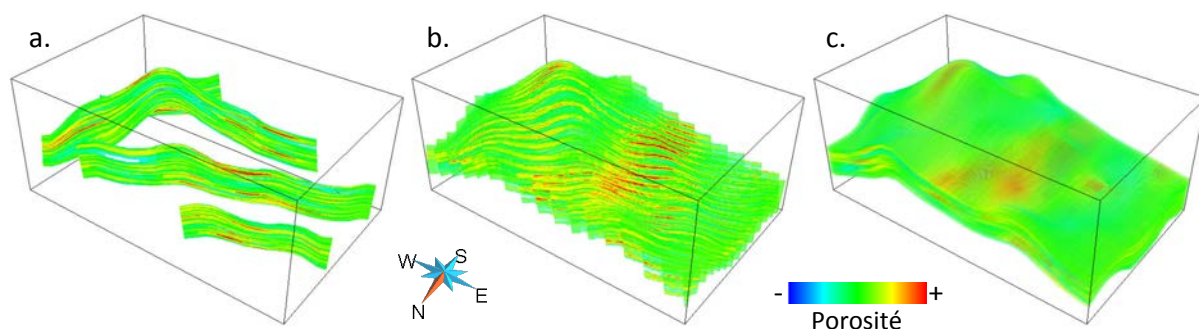


Fig. 2.6 – Discrétisation par tranchage incrémental à différents degrés de finesse. (a) Rendu grossier (4 tranches). (b) Rendu intermédiaire (30 tranches). (c) Rendu fin (1000 tranches). Le jeu de données utilisé représente une propriété de porosité suivant la stratigraphie, stockée sous forme de grille Cartésienne avec 18.000.000 cellules. Données Total.

2.2 Cadre théorique de visualisation des incertitudes

Beaucoup d'auteurs proposent une évaluation de leurs méthodes de visualisation reposant sur des critères théoriques, dont la plupart dérivent des travaux de Bertin (1983), Tufte (2001), Ware (2004) ou Shortridge (1982). Cette section présente les principaux critères utilisés dans ce cadre, en se focalisant tout particulièrement sur la problématique de la visualisation des incertitudes.

2.2.1 Variables visuelles

Le cadre théorique de la visualisation des incertitudes est étudié par Zuk et Carpendale (2006), qui évaluent la qualité de plusieurs méthodes de visualisation des incertitudes en se basant sur les critères de perception définis par Bertin (1983), Tufte (2001) et Ware (2004). En particulier, ils utilisent les variables visuelles définies par Bertin, c'est-à-dire la *position*, la *taille*, la *valeur*, la *texture*, la *couleur*,³ l'*orientation* et la *forme* (Table 2.1). MacEachren (1992) ajoute à cette liste une huitième variable qu'il nomme *focus*; il propose différentes manières de la manipuler, via la rugosité ou le degré de flou de contours, la clarté du remplissage de figurés, l'ajout de brouillard et le changement de la résolution de l'affichage. McGranaghan (1993) étend encore cet ensemble de variables visuelles avec la notion de *réalisme*, qui apparaît de manière récurrente en imagerie de synthèse; il argumente qu'un objet représenté de manière réaliste sera perçu comme plus certain qu'un objet représenté de manière non réaliste.

³La couleur au sens de Bertin n'intègre que deux composantes (longueur d'onde et pureté de la teinte), la valeur étant considérée comme une variable visuelle différente.

Tab. 2.1 – Variables visuelles proposées par Bertin (1983), avec extensions proposées par MacEachren (1992) et McGranaghan (1993).

Variable visuelle	Description
Position	Coordonnées (x,y) d'un élément sur le plan visuel
Taille	Dimensions d'un élément
Valeur	Quantité de noir perçue localement
Texture	Variation locale de l'échelle des éléments
Couleur	Teinte et degré de saturation locaux
Orientation	Angle local des éléments
Forme	Combinaison des attributs de taille et d'orientation
Focus	Pouvoir d'attraction visuelle d'un élément
Réalisme	Similarité perçue d'un élément par rapport à un objet réel

L'ensemble des variables visuelles définies par Bertin (1983) présente un certain nombre de caractéristiques perceptuelles intéressantes telles que l'ordre naturel, la perception quantitative, la sélectivité ou l'associativité (Figure 2.7, Table 2.2) :

1. Une variable visuelle présente un ordre naturel si elle permet de trier des éléments sans qu'il soit nécessaire d'apprendre la convention qui est utilisée pour les représenter. Les variables présentant un ordre naturel sont la position, la taille et la valeur.
2. Une variable visuelle est quantitative si elle permet d'évaluer la grandeur absolue d'un élément. Les variables quantitatives sont la position et la taille.
3. Une variable visuelle est sélective si l'attention de l'observateur peut être dirigée vers elle seule, en excluant toute autre variable visuelle. Les variables sélectives sont la position, la taille, l'orientation, la couleur, la valeur et la texture. La forme n'est par définition pas sélective car elle est une combinaison de la taille et de l'orientation.
4. Une variable visuelle est associative si elle permet de regrouper des éléments comme faisant partie d'une famille, tout en maintenant une séparation possible par le biais des autres variables visuelles. La position, la valeur, la couleur, la taille, la texture, l'orientation et la forme sont des variables visuelles associatives.

Pour favoriser une bonne perception de l'information, les variables visuelles doivent présenter des caractéristiques similaires aux grandeurs qu'elles représentent ; par exemple, il serait incohérent de représenter une grandeur angulaire à l'aide d'une variable visuelle ordonnée.

2.2.2 Qualité des métaphores visuelles

Ware (2004) établit une distinction claire entre les représentations *sensorielles* et les représentations *arbitraires* de l'information. Une représentation sensorielle peut être com-

Tab. 2.2 – Récapitulatif des propriétés associées aux variables visuelles.

	Ordonnée	Associative	Sélective	Quantitative
Position	✓	✓	✓	✓
Taille	✓	✓	✓	✓
Valeur	✓	✓	✓	
Texture		✓	✓	
Couleur		✓	✓	
Orientation		✓	✓	
Forme		✓		

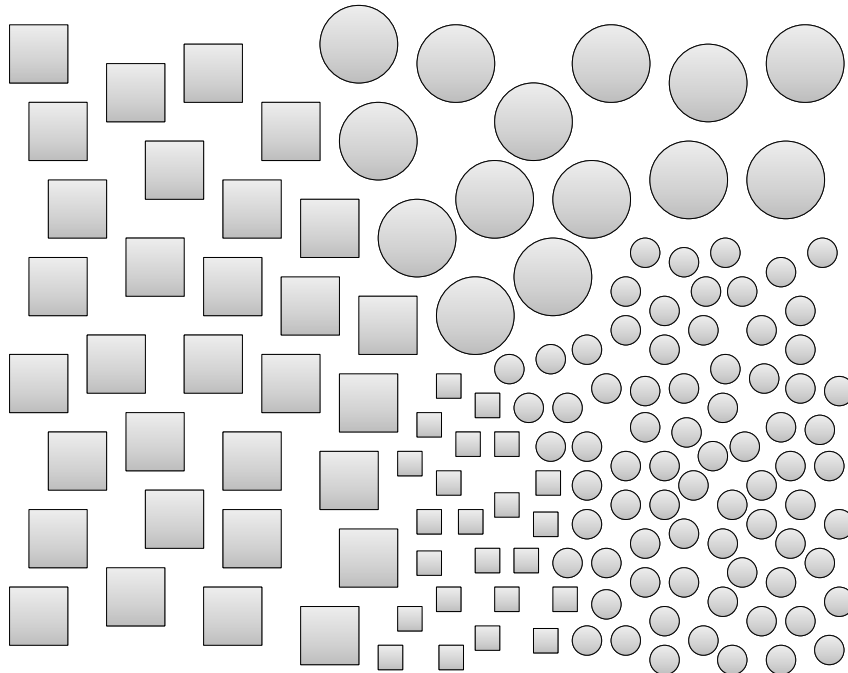


Fig. 2.7 – Illustration des propriétés d'ordre naturel, de perception quantitative, de sélectivité et d'associativité de la variable visuelle "taille". La zone en bas à droite présente des éléments de plus petite dimension que le reste de la scène : la taille est une variable visuelle ordonnée. Ces éléments sont environ deux fois et demie plus petits que les éléments de grande dimension : la taille est une variable quantitative. La dimension des éléments peut être perçue indépendamment de la forme des éléments : la taille est une variable sélective. La zone où sont présents les petits éléments est discernable, sans empêcher de percevoir la forme des éléments qui la composent : la taille est une variable associative.

prise sans qu'il ne soit nécessaire d'apprendre la convention qui est utilisée, typiquement par analogie avec des entités connues du monde réel. Par opposition, les représentations arbitraires font appel à une symbolique ne se référant pas à une base perceptuelle, dont les règles doivent donc être apprises préalablement à l'étude de la scène graphique. De par leur nature même, les représentations sensorielles tendent à être perçues de manière similaire par tout observateur, tandis que la perception des représentations arbitraires peut être influencée par la culture ou les connaissances de l'observateur.

Dans le cadre de nos travaux, nous favorisons l'usage de représentations sensorielles lors de la création de métaphores visuelles rappelant un degré spatial de confiance, afin de produire des visualisations pouvant être comprises de manière plus intuitive comme suggéré par Rheingans et Landreth (1995). Cependant, certains concepts ou modes de visualisation font appel à des notions trop abstraites pour pouvoir reposer sur une base perceptuelle; dans ces cas, nous tentons de nous rapprocher des représentations arbitraires les plus fréquemment utilisées afin de minimiser la charge d'apprentissage pour l'observateur.

2.2.3 Interférences visuelles

La recherche de paramètres visuels ayant entre eux un degré minimal d'interférences est un problème récurrent en visualisation multi-variables, dont la visualisation des incertitudes peut être considérée comme un sous-ensemble (Rheingans et Landreth, 1995). Shortridge (1982) propose une distinction entre les visualisations *coïncidentes séparables* de l'incertitude, pour lesquelles les données peuvent être considérées séparément de l'incertitude, et les visualisations *coïncidentes intégrales* pour lesquelles le degré d'incertitude est pleinement intégré à la représentation même des données. Dans le cas général, les visualisations séparables sont préférables aux visualisations intégrales en géosciences, bien que l'utilisation de variables visuelles non-orthogonales soit acceptable dans certains cas – par exemple, pour sélectionner des zones présentant un fort potentiel en hydrocarbures avec un minimum d'erreur, le caractère “faible potentiel en hydrocarbures” et le caractère “fortement incertain” peuvent être encodés par les mêmes variables visuelles sans nuire à la qualité de la visualisation.

2.2.4 Préattentivité

La préattentivité est un mécanisme du système visuel humain grâce auquel les éléments intéressants d'une scène peuvent être perçus en un laps de temps extrêmement restreint, avant que l'observateur n'ait besoin de se concentrer sur la scène graphique de manière consciente. Healey *et al.* (1996) décrit le laps de temps nécessaire à la préattentivité comme inférieur à 200 millisecondes, ce qui correspond au temps nécessaire pour initier le mouvement des yeux. Ware (2004) affirme que le mécanisme de préattentivité est indépendant du nombre d'éléments distrayants présents dans la scène; il illustre ses

propos à l'aide d'une succession de chiffres, parmi lesquels l'observateur doit compter d'occurrences d'un chiffre en particulier (Figure 2.8). La préattentivité peut être favorisée en utilisant des variables visuelles appropriées, telles que la couleur, la position, l'orientation ou la taille des éléments graphiques (Tory et Möller, 2004).

- a. 67468164883957813438903816538841310395371077727235408733605520
19085778055357365263936061520475717342761978597700545255461395
37474538673892992656736128696164317921446571323388229124503505
6391173953056211928872739980640725246554416524073983569467080

- b. 674681648839578134389038165388413103953710777**2**72354087336055**2**0
19085778055357365**2**639360615**2**047571734**2**761978597700545**2**55461395
3747453867389**2**99**2**6567361**2**86961643179**2**14465713**2**3388**2**291**2**4503505
6391173953056**2**119**2**887**2**7399806407**2**5**2**465544165**2**4073983569467080

Fig. 2.8 – Illustration du phénomène de préattentivité. a : pour compter le nombre d'occurrences du chiffre 2, l'observateur doit parcourir séquentiellement la totalité de la suite de nombres. b : pour compter le nombre d'occurrences du chiffre 2, l'observateur doit parcourir uniquement les nombres coloriés en rouge. Adapté d'après Ware (2004).

2.2.5 Concentration de l'information

La concentration des informations est un concept introduit par Tufte (1983) sous la forme du rapport R , nommé “data-ink ratio” en anglais :

$$R = \frac{\text{Quantité d'encre utilisée pour représenter les données}}{\text{Quantité d'encre utilisée dans la totalité du graphique}} \quad (2.3)$$

La quantité d'encre utilisée pour représenter les données peut être considérée comme la partie essentielle et non-effaçable du graphique. Tufte (1983, 2001) suggère de maximiser rapport R dans le but d'épurer au maximum un graphique des informations inutiles ou redondantes ; cette stratégie permet de n'attirer l'attention de l'observateur que sur les messages importants portés par le graphique. Cette stratégie peut notamment être appliquée en géomodélisation pour condenser les informations présentées.

2.3 Méthodes de visualisation des incertitudes

De nombreux travaux traitent du problème de la visualisation des incertitudes dans des domaines d'applications divers, tels que les systèmes d'information géographiques (SIG), la météorologie, l'océanographie ou la recherche médicale (MacEachren, 1992; Pang *et al.*,

1997; Johnson et Sanderson, 2003; Griethe et Schumann, 2006; Pang, 2006). Certains de ces travaux utilisent des représentations *adjacentes* de l'incertitude, c'est-à-dire des cartes séparées pour les données d'intérêt et les incertitudes ; cependant, la majorité s'oriente vers des représentations *coïncidentes* de l'incertitude, pour lesquelles les incertitudes sont affichées sur la même image que les données (MacEachren *et al.*, 1998). En s'appuyant sur les travaux de Shortridge (1982), MacEachren (1995) établit une distinction entre les représentations *coïncidentes intégrales* de l'incertitude, qui altèrent directement l'affichage des données initiales, et les représentations *coïncidentes séparables*, qui permettent une attention sélective sur l'une ou l'autre des variables présentées. Cette section propose un état de l'art des approches adoptées dans la littérature pour répondre à la visualisation des incertitudes, en se focalisant tout particulièrement sur les représentations coïncidentes de l'incertitude.

2.3.1 Incertitudes sur des graphiques

Les modes de visualisation des incertitudes sur des graphiques sont abondamment décrits dans la littérature. Nous ne proposons qu'une revue limitée des travaux publiés dans ce domaine, car les concepts proposés sont appliqués à des entités unidimensionnelles et présentent rarement des extensions pour les objets 2,5D/3D qui sont quotidiennement utilisés en géosciences.

La technique des barres d'erreur est probablement la méthode de visualisation des incertitudes la plus populaire sur des graphiques ; leur représentation est notamment adaptée aux incertitudes statistiques (Olston et Mackinlay, 2002). Les barres d'erreur peuvent être échantillonnées le long d'un espace de plus haute dimension (2D ou 2,5D) comme présenté dans l'étude de Sanyal *et al.* (2009), et sont alors assimilables à un type de glyphe (les glyphes sont des éléments figurés dont les attributs visuels – le plus souvent, la forme – représentent une ou plusieurs grandeurs scalaires). Sanyal *et al.* (2009) reportent cependant des problèmes de perception significatifs avec cette méthode, par comparaison avec d'autres techniques liant l'incertitude à la taille ou à la couleur de glyphes sphériques.

Olston et Mackinlay (2002) rendent ambiguë la frontière de points, de barres de pourcentage et de lignes en leur donnant une certaine épaisseur ; la zone ambiguë est également rendue avec une couleur différent à la fois de la couleur des zones certaines et de la couleur de fond, afin de clairement séparer le domaine d'incertitude des domaines certains. Leur méthode s'applique à des incertitudes de type borné, c'est-à-dire que la grandeur représentée peut exister en tout point du domaine incertain sans pour autant pouvoir préjuger d'une densité de probabilité sur ce domaine.

Jackson (2008) modifie l'épaisseur et propose un dégradé de couleur pour des courbes ou des points portés par des graphiques afin de représenter leur incertitude. Sa technique utilise des "bandes d'incertitude", qui correspondent à l'ensemble des points du graphique où un élément a une probabilité p non nulle de se trouver ; en tout point de la bande

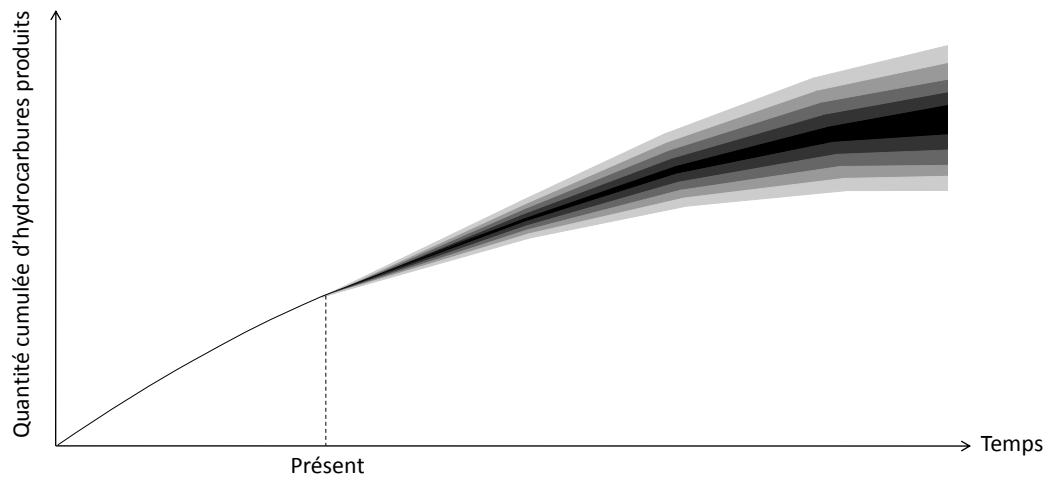


Fig. 2.9 – Graphique présentant la quantité cumulée d'hydrocarbures produits avec incertitudes sur la production prédite. Le domaine incertain est visualisé à l'aide d'une bande d'incertitude (Jackson, 2008).

d'incertitude, la valeur associée à la couleur est directement proportionnelle à la densité de probabilité (Figure 2.9).

Sambridge *et al.* (2006) présentent sur un même graphique un ensemble de profils de vitesses sismiques dans le manteau supérieur de la Terre, indiquant la vitesse des ondes de cisaillement en fonction de la profondeur. Ces profils ont été obtenus par échantillonnage de Monte-Carlo, en éliminant les profils incohérents avec des mesures connues d'ondes de Rayleigh (ondes confinées à un plan vertical). Une telle approche est cependant rarement utilisable en géosciences, car la superposition d'entités de plus haute dimension provoque généralement une importante surcharge perceptuelle et cognitive.

2.3.2 Incertitudes en archéologie

La reconstitution archéologique est généralement soumise à de fortes incertitudes, car les objets et les bâtiments ont souvent été altérés par le temps et doivent donc être restaurés mentalement à partir d'informations indirectes.

Strothotte *et al.* (1999) représentent le degré d'incertitude associé à des reconstitutions de bâtiments en archéologie. Les fondations sont généralement suffisamment bien conservées pour connaître la forme générale du bâtiment, mais les parties hautes ont souvent été détruites et doivent être devinées à partir des informations archéologiques disponibles. Leur méthode fait intervenir un gradient de transparence croissant avec l'altitude sur une reconstitution photo-réaliste du bâtiment ; pour préserver la perception des struc-



Fig. 2.10 – Visualisation des incertitudes sur l'architecture d'un bâtiment en archéologie. Les incertitudes croissent entre les fondations et le sommet de la reconstitution du bâtiment. Adapté d'après Strothotte et al. (1999).

tures même dans les zones fortement transparentes, ils superposent au rendu un dessin du bâtiment non photo-réaliste du bâtiment (Figure 2.10⁴).

Zuk et al. (2005) présentent une méthode de visualisation des incertitudes temporelles utilisant à la fois la transparence selon laquelle un bâtiment est rendu et la position de ce bâtiment par rapport au sol. Cette dernière technique, inspirée des travaux de Pang

⁴Le bâtiment n'est utilisé que pour illustrer la méthodologie de Strothotte et al. (1999) ; il s'agit en réalité d'une église actuellement en fonction située dans la baie de Cardiff au pays de Galles, et non pas d'une vraie reconstruction faite à partir de fondations.

et al. (1997), permet de communiquer efficacement les incertitudes associées aux dates de construction et de destruction des bâtiments ; en effet, un bâtiment émergeant du sol sera considéré comme en cours de construction, et un bâtiment englouti comme en cours de destruction. Cette technique est cependant associée à une sémantique particulière, qui est difficilement applicable à des problématiques en géosciences.

2.3.3 Incertitudes de position particulière dans des écoulements

En physique, la position de particules (atomes ou ensembles assimilés à des points infinitésimaux) est fréquemment sujette à incertitude. Cette caractéristique est fréquemment observée dans le cadre de la modélisation d'écoulements, où les lois de comportement du flux ne sont pas connues avec exactitude.

Lodha *et al.* (1996a) utilisent des glyphes “ligne” pour présenter les écarts entre deux trajectoires possibles d'une particule élémentaire ; ces deux trajectoires sont obtenues à l'aide d'algorithmes de calcul de lignes d'écoulement différents. Les glyphes lignes relient les points les plus proches sur les deux trajectoires possibles et ignorent ainsi toute différence de vitesse particulière d'une méthode à l'autre. Pour pallier à ce problème, Lodha *et al.* (1996a) présentent les glyphes “en barbelés”, qui joignent les positions possibles d'une particule pour un même temps t ; ces glyphes sont similaires aux glyphes “ligne” de par leur forme, mais indiquent également les variations de vitesse de la particule d'une méthode de calcul à l'autre. En effet, si les glyphes sont orthogonaux aux lignes de courant, la vitesse est égale d'une ligne à l'autre – tout autre angle indiquant ainsi une différence de vitesse entre les deux méthodes.

En s'appuyant sur le même jeu de données, Lodha *et al.* (1996a) et Pang *et al.* (1997) tracent un ruban d'incertitude entre les deux trajectoires possibles de la particule ; la largeur du ruban indique alors le degré d'incertitude sur la position de la particule. Cette approche est généralement moins intéressante que les glyphes “en barbelés”, car elle ne permet pas de communiquer les différences de vitesse entre les deux modèles.

Lodha *et al.* (1996a) utilisent également le concept d'enveloppe pour représenter l'incertitude sur la trajectoire d'une particule. Ils calculent la trajectoire moyenne entre deux trajectoires possibles d'une particule, puis génèrent un tube centré autour de cette moyenne dont le diamètre est fonction de l'écart entre les deux modèles. Cette méthode indique ainsi le domaine dans lequel la particule peut raisonnablement se trouver ; son extension en trois dimensions présente un intérêt certain dans la représentation des incertitudes structurales sur la géométrie d'horizons ou de failles.

Lodha *et al.* (2002b) introduisent l'usage de nuages semi-transparents pour représenter la fonction de densité de probabilité associée à la position de particules dans l'espace. Les zones où la probabilité est élevée présentent ainsi une transparence plus élevée que les zones de où la probabilité de présence de la particule est faible. De nouveau, l'extension

de cette technique en trois dimensions est intéressante pour représenter les incertitudes géométriques sur la position d'horizons ou de failles.

Pang *et al.* (1997) animent la trajectoire d'une particule pour représenter l'évolution temporelle de sa position. Leur méthode intègre deux trajectoires possibles ; à chaque instant t , ils présentent l'incertitude sur la position de la particule à l'aide d'un flou cinétique (en photographie, le flou cinétique est l'effet de perte de netteté de l'image dû à un déplacement du sujet photographié ou de l'appareil lui-même). Pang *et al.* (1997) relie ainsi l'incertitude de position au rayon du flou cinétique, les particules les plus incertaines apparaissant ainsi sur un domaine plus large que les particules fiables.

Lodha *et al.* (1996c) étudient l'usage de sons pour représenter un degré d'incertitude local associé à la position d'une particule. Leur système permet de relier différents paramètres du son à l'incertitude, tels que la durée, le volume ou le timbre ; ils présentent un exemple associant la durée du son à l'incertitude de position de la particule. La sonification est généralement peu adaptée en géosciences, car elle ne permet d'échantillonner qu'un point de l'espace à un temps donné et s'avère donc très limitante pour communiquer les évolutions spatiales des incertitudes.

Botchen *et al.* (2005) présentent une technique de visualisation des incertitudes dédiée aux écoulements bidimensionnels, où la trajectoire des particules est représentée sous la forme de stries au sein de textures bidimensionnelles ; cette approche diffère significativement des techniques présentées précédemment, car elle ne requiert pas de tracer la position des particules de manière individuelle. Leur technique est implémentée sous la forme d'un filtre d'image qui produit un épaissement de ces stries au fur et à mesure que la particule s'éloigne de sa position initiale, traduisant ainsi l'accumulation des erreurs sur la position de la particule. Ce concept d'épaissement rejoint la notion d'enveloppe d'incertitude et peut donc s'appliquer à des incertitudes géométriques en géomodélisation.

2.3.4 Incertitudes sur la reconstruction de surfaces

Les problèmes de reconstruction de données sont récurrents dans de nombreux domaines scientifiques, soit du fait de l'échantillonnage discret d'un phénomène continu, soit parce que la collecte de certains échantillons s'est avérée impossible en certains points. Les informations reconstruites présentent cependant une fiabilité inférieure aux échantillons eux-mêmes ; différentes stratégies ont été adoptées pour transmettre visuellement ce degré de fiabilité, notamment dans la reconstruction de surfaces. Ces incertitudes portent tant sur la géométrie de la surface reconstruite que sur ses attributs.

Incertaines de position Wittenbrink *et al.* (1995a) reconstruisent une surface incomplète à l'aide de deux algorithmes différents (interpolations bilinéaire et multi-quadratique). Ils comparent l'écart entre les deux surfaces et expriment cette grandeur grâce à des glyphes "ligne" et "ellipsoïde". Les extrémités des glyphes "ligne" vont d'une surface inter-

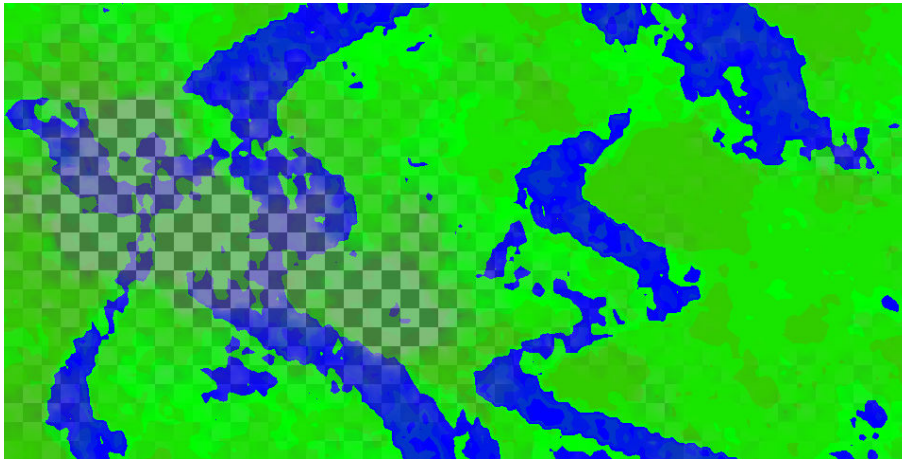


Fig. 2.11 – Visualisation de chenaux avec une transparence accrue des zones incertaines. L'arrière-plan est constitué d'un damier procédural, d'une manière similaire au rendu du logiciel The GIMP (Mattis et Kimball, 2008).

polée à l'autre, leur longueur indiquant ainsi la différence locale entre les deux méthodes. Le grand axe des glyphes "ellipsoïde" représente la distance entre les deux surfaces reconstruites (les axes médian et mineur étant constants) : des ellipsoïdes aplatis indiquent ainsi une faible incertitude spatiale, tandis que des ellipsoïdes allongés marquent un écart important.

Dans une approche similaire, Lodha *et al.* (1996b) proposent des glyphes "de déplacement", dont la hauteur dénote de l'écart entre plusieurs méthodes d'interpolation. En plus des glyphes "ligne" et "ellipsoïde" décrits par Wittenbrink *et al.* (1995a), ils présentent des glyphes de forme hexaédrique et cylindrique ; les glyphes hexaédriques ont par la suite été réutilisés dans les travaux de Jospeh *et al.* (1999). Lodha *et al.* (1996b) introduisent également trois autres types de glyphes : (i) les glyphes "en croix", représentés par deux plans orthogonaux dont la hauteur indique les incertitudes sur la position moyenne et la courbure Gaussienne des surfaces interpolées, (ii) les glyphes triangulaires, pour lesquels les arêtes du triangle représentent les orientations de deux normales possibles partant d'un même point d'origine, et (iii) les glyphes volumiques, représentés par des sphères dont le rayon est proportionnel à l'écart entre deux surfaces interpolées.

Jospeh *et al.* (1999) testent un certain nombre de techniques de visualisation pour représenter la différence entre deux méthodes de reconstruction. Parmi ces techniques, ils accentuent la transparence aux points où l'écart entre les deux surfaces reconstruites est important (Figure 2.11). Jospeh *et al.* (1999) ne garantissent cependant pas un minimum d'opacité dans la visualisation ; cette approche produit ainsi des zones entièrement transparentes aux points où les deux surfaces sont les plus distantes, interférant donc avec la perception des surfaces elles-mêmes.



Fig. 2.12 – Visualisation d’une couche géologique avec enveloppe d’incertitude sur la position de la couche. L’enveloppe d’incertitude englobe l’ensemble des positions que la couche peut occuper.

Wittenbrink (1995) perturbe la rugosité d’une surface reconstruite à l’aide d’une méthode fractale basée sur les systèmes de fonctions itérées ; plus le nombre d’itérations de la fonction croît, plus la surface présente un aspect rugueux, ce qui dénote ainsi d’une incertitude élevée. De manière similaire, Pang *et al.* (1997) perturbent la rugosité de surfaces reconstruites par déplacement de ses sommets, selon une fréquence et une intensité dépendant de l’écart entre les deux méthodes de reconstruction. Ces approches perturbent la géométrie des surfaces, ce qui peut nuire à la perception de leur structure ; bien que leur usage soit intéressant dans une problématique de géomodélisation, il est critique de maintenir une perturbation limitée de ce phénomène pour ne pas altérer la compréhension des structures géologiques.

Pang *et al.* (1997) introduisent le concept de “surfaces épaisses” pour représenter la différence entre deux méthodes de reconstruction ; le volume délimité par deux frontières enveloppe ainsi le domaine dans lequel la surface réelle peut exister. Pang *et al.* (1997) proposent de visualiser ce volume en réalisant une section selon un plan sub-perpendiculaire aux frontières. Cette méthode peut être améliorée en augmentant la transparence des frontières, ou à l’aide de rendu volumique (Figure 2.12) comme présenté dans les travaux de Johnson et Sanderson (2003). Ces derniers présentent deux visualisations d’isosurface, respectivement créées de manière synthétique et à partir de données de résonance magnétique, dans lesquelles une enveloppe d’incertitude est rendue avec une plus faible opacité pour indiquer une incertitude de position.

Lodha *et al.* (1996c) étendent leur méthode de sonification aux technique de reconstruction de surfaces. Ils proposent un exemple où l'incertitude sur la position moyenne de la surface est reliée au pitch⁵ de notes de violon, tandis que l'incertitude sur la courbure Gaussienne est reliée au pitch de notes de tambour. Cette technique ne peut cependant s'appliquer qu'à une unique position à un temps donné, et limite donc fortement l'exploration de l'ensemble d'un modèle.

Incertitudes d'attributs Plutôt que d'étudier les différences de position entre les surfaces reconstruites, Pang *et al.* (1997) s'intéressent à l'écart entre les attributs de la surface d'une méthode de reconstruction à l'autre, notamment en termes de radiosité.

Pang *et al.* (1997) proposent ainsi de modifier le type de lumière qui est renvoyée par la surface pour dénoter de l'incertitude sur la radiosité de la surface. Dans le modèle de Phong (1975), l'éclairage est la somme de trois types de lumières : (i) la lumière spéculaire, qui correspond à la fraction d'une lumière incidente qui est réfléchiée dans une unique direction, (ii) la lumière diffuse, qui correspond à la fraction d'une lumière incidente qui est réfléchiée dans des directions multiples, et (iii) la lumière ambiante, qui peut être considéré comme un niveau d'illumination global de la scène. Pang *et al.* (1997) altèrent les coefficients de réflexion diffuse et spéculaire du matériau en fonction de l'incertitude sur la radiosité d'une surface. Leur approche produit un accroissement de la luminosité renvoyée (altération de la lumière diffuse) ou un effet "miroir" (altération de la lumière spéculaire) dans les zones de forte incertitude.

Pang *et al.* (1997) reprennent également les travaux de Pang et Alper (1995) et proposent la perturbation des normales d'une surface en fonction de l'incertitude. Cette approche s'inspire des techniques de placage de relief (*bump-mapping* en anglais) introduites par Blinn (1978), dans lesquelles une texture est utilisée pour représenter la carte des aspérités affectant une surface. Leur méthode accroît le relief aux points où l'erreur est importante.

Pang *et al.* (1997) modifient par ailleurs la position de patchs de surfaces par rotation ou par translation, en fonction de l'écart de radiosité calculée selon deux méthodes différentes. La rotation des patchs se fait d'un angle variant de 0 (incertitude minimale) à 90 degrés (incertitude maximale); les zones incertaines présentent une forte perturbation de structure tandis que les zones bien connues restent aisément reconnaissables. La translation des patchs permet aisément de voir les contrastes d'incertitude d'un patch à l'autre, mais est moins efficace pour représenter la valeur absolue de l'incertitude locale. Par ailleurs, la translation peut se faire sur un intervalle non-borné et peut donc fortement altérer la perception de la structure initiale des données; son usage est donc déconseillé en géomodélisation, où la compréhension de la structure est primordiale.

Pang *et al.* (1997) combinent les méthodes de modification de géométrie par translation et rotation aux techniques d'animation pour visualiser l'écart de radiosité entre les deux méthodes de reconstruction. Ils proposent deux types d'association entre l'animation et

⁵Le pitch correspond à la modification de vitesse d'un son par rapport à sa vitesse initiale

l'incertitude locale : (i) une association par amplitude du déplacement, pour laquelle le temps de départ et d'arrivée est identique sur tous les patches de surface, l'incertitude étant liée à l'intensité du déplacement, et (ii) une association par fréquence du déplacement, pour laquelle l'intensité du déplacement est identique sur l'ensemble des patches, l'incertitude étant liée à la vitesse de déplacement des patches. Ils proposent également de combiner l'animation aux méthodes par ajout de géométrie, où des surfaces enveloppe oscillent entre la position initiale S_i de la surface (ensemble des positions les plus probables) et la surface frontière S_p (ensemble des positions où la surface a une probabilité p de se trouver).

Enfin, Pang *et al.* (1997) et Brown (2004) réalisent un effet stéréographique à l'aide de lunettes adaptées, permettant à chacun des yeux de recevoir une image différente. En décalant légèrement le point de vue des images envoyées à chacun des yeux, il est ainsi possible de donner à une scène un effet 3D. Pang *et al.* (1997) proposent de paramétrer le décalage de l'image en fonction du degré d'incertitude local – aux points présentant une forte incertitude de radiosité, les deux images seraient rendues à l'identique, annulant ainsi l'effet tridimensionnel. Brown (2004) propose la démarche inverse – décaler les deux images plus fortement que nécessaire – afin de produire une impression de flou dans les zones incertaines. Cependant, Pang *et al.* (1997) affirment qu'une utilisation prolongée de cette méthode produit une importante fatigue visuelle, du fait que le système visuel tente de résoudre des informations contradictoires arrivant à chacun des yeux.

Pang *et al.* (1997) proposent également l'usage de messages subliminaux pour dénoter du degré d'incertitude. Ils expérimentent cette idée à l'aide de messages simples (“irritant” et “éternuer”) placés à l'écran dans les zones les plus incertaines ; bien que convaincus du potentiel associé à ces méthodes, Pang *et al.* (1997) admettent que leurs tests n'ont pas été à même de provoquer les réactions attendues de manière répétée.

2.3.5 Incertitudes en médecine

Ces dernières décennies, la médecine a vu se développer des techniques de plus en plus pointues dans différents domaines, tels que les outils de diagnostic, la compréhension de l'action des maladies, le développement de nouveaux médicaments, la gestion de la santé publique, *etc.* La complexité de ces problèmes s'accompagne souvent d'incertitudes qui peuvent avoir des conséquences critiques en termes de santé. Nous divisons les approches permettant de visualiser ces incertitudes en trois catégories : les incertitudes sur la configuration de molécules, les incertitudes macroscopiques généralement associées aux tissus humains, et les incertitudes statistiques affectant la santé publique.

Incertaines à l'échelle moléculaire Rheingans et Joshi (1999) s'intéressent aux configurations possibles d'une molécule utilisées comme élément actif pour un médicament. Huit configurations possibles peuvent être observées. La plupart des atomes oc-

cupent une position identique d'une configuration à l'autre, mais quelques atomes sont localisés de manière différente et suffisent à modifier radicalement les effets du médicament.

Rheingans et Joshi (1999) introduisent leurs travaux par une visualisation des incertitudes de référence dans le domaine pharmacologique. Dans cette méthode, toutes les familles de configurations sont représentées simultanément. Ils argumentent que ce mode de visualisation est efficace pour appréhender les sections de la molécule qui présentent une forte incertitude, mais qu'il est moins à même de révéler les détails de structure de la molécule du fait de l'encombrement du plan visuel dans les zones incertaines. Pour pallier à ce défaut, Rheingans et Joshi (1999) proposent une méthode de visualisation discrète de l'ensemble des familles, dans laquelle chaque famille est représentée sur un même plan visuel avec une transparence α constamment égale à $\frac{1}{8}$. Cette visualisation produit des contours bien définis pour les atomes dont la position est commune aux huit familles, et plus flous pour les atomes dont la position varie d'une famille à l'autre. Cette approche est cependant peu applicable en géosciences, car les objets géologiques présentent une dimension élevée et leur superposition produirait donc une forte surcharge perceptuelle et cognitive.

Lee et Varshney (2002) et Rheingans et Joshi (1999) utilisent en outre la notion de volume enveloppe pour visualiser l'incertitude sur la configuration de molécules. Dans ces travaux, la position d'un atome est rendue comme une fonction de probabilité à l'aide de méthodes de "splatting", qui représentent la contribution de chaque atome sous la forme d'une fonction Gaussienne dont le rayon dépend de la densité de probabilité locale. Des résultats similaires peuvent également être obtenus à l'aide de méthodes d'extraction d'isosurface, en utilisant la propriété de densité de probabilité si cette dernière est disponible en tout point du domaine étudié.

Incertainitudes macroscopiques Les incertitudes macroscopiques traitent généralement les imprécisions ou erreurs d'interprétation associées à des mesures indirectes de tissus humain. Par exemple, Jones (2003) s'intéressent aux incertitudes sur l'orientation de l'anisotropie associée à des tenseurs d'imagerie par résonance magnétique (IRM). Ils utilisent les glyphes en cône, initialement définis par Basser et Le Bihan (1992), pour représenter ces incertitudes; l'orientation et la taille du cône indiquent respectivement la direction la plus probable et l'intensité de l'anisotropie, tandis que le degré d'ouverture du cône indique les incertitudes sur l'orientation de l'anisotropie.

De manière similaire, Zuk *et al.* (2008) présentent deux glyphes dédiés à la visualisation des incertitudes d'orientation sur des champs de vecteurs bidimensionnels. Leurs glyphes sont un assemblage d'un rectangle et de deux à quatre lignes, où l'orientation du côté le plus long du rectangle indique la direction la plus probable du vecteur, tandis que le degré d'ouverture des lignes indique l'incertitude sur l'orientation qui lui est associée.

Newman et Lee (2004) présentent également trois méthodes s'appuyant sur des glyphes, liant le degré d'incertitude local à la forme de cylindres, de cônes et de groupes de points. Ces approches sont illustrées sur une isosurface représentant l'exosquelette d'une écrevisse, extraite à partir de points d'échantillonnage produits par tomographie, le long de laquelle un champ scalaire d'incertitude synthétique est disponible. Ils valident ces techniques à l'aide d'une étude sur des utilisateurs, qui montre que les glyphes par groupes de points sont perceptuellement supérieurs aux deux autres techniques sur l'ensemble des tâches proposées.

De manière semblable, Rhodes *et al.* (2003) étudient les incertitudes liées à des isosurfaces obtenues par tomographie assistée par ordinateur. Leurs isosurfaces sont extraites à partir de données multi-résolution, dont les versions les plus grossières peuvent induire de larges approximations dans la géométrie de l'isosurface. Rhodes *et al.* (2003) utilisent ainsi des textures à granularité variable pour représenter le degré local d'approximation associé à ces versions grossières de la tomographie. Ils jugent les isosurfaces particulièrement adaptées à la visualisation des incertitudes, car la propriété d'intérêt a par définition une valeur constante le long de l'isosurface (seul le degré d'incertitude est donc variable). Cette caractéristique unique leur permet de contourner les problèmes de rendu multivariable qui apparaissent de manière récurrente en visualisation des incertitudes. Leur approche présente par ailleurs un grand intérêt en géosciences, car les techniques d'extraction d'isosurfaces sont fréquemment utilisées pour définir des domaines d'intérêt (par exemple, les volumes pour lesquels un gisement présente une forte teneur en métaux précieux).

Grigoryan et Rheingans (2004) représentent les frontières d'un modèle de tumeur à l'aide d'un nuage de points placés sur les bordures externes de la tumeur. Chaque point est alors perturbé aléatoirement dans la direction de croissance de la tumeur, avec une intensité de perturbation dépendant de l'incertitude locale. Cette méthode produit des bordures nettes là où les frontières de la tumeur sont bien déterminées, tandis que les zones incertaines sont fortement perturbées et de ce fait, plus difficilement discernables (Figure 2.13). L'insertion d'un grand nombre de points accroît la charge de calcul pour la carte graphique. Cependant, Grigoryan et Rheingans (2004) affirment qu'avec l'augmentation constante de la résolution des modèles, il est fréquent que des polygones occupent un espace inférieur à la dimension d'un pixel à l'écran ; une représentation sous forme de nuage de points n'est donc pas fondamentalement plus coûteuse dans ce cas. Par ailleurs, un modèle surfacique opaque et clos génère de nombreuses occlusions de ses propres primitives, qui sont ainsi inutilement traitées par la carte graphique (Barras, 2010).

Lundström *et al.* (2007) proposent une méthode d'animation sur des données médicales de tomographie, présentant une possible sténose (rétrécissement anormal de vaisseaux sanguins) par des méthodes de rendu volumique direct. L'interprétation de l'image comme étant une sténose dépend de la fonction de transfert utilisée. Plutôt que d'animer des jeux de données, Lundström *et al.* (2007) animent donc la fonction de transfert utilisée pour présenter ces jeux de données. Ils proposent une extension permettant de rendre les fonc-

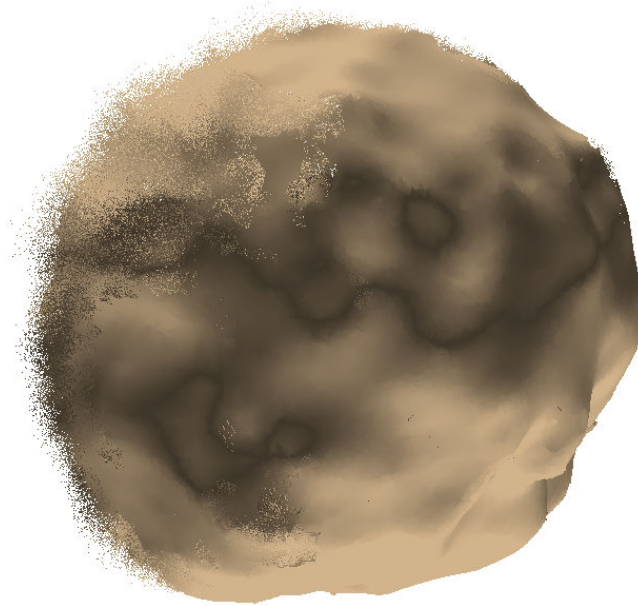


Fig. 2.13 – Visualisation d'un modèle de tumeur synthétique avec perturbation de la position des points dans les zones incertaines. Adapté d'après [Grigoryan et Rheingans \(2004\)](#).

tions de transfert probabilistes, c'est-à-dire que si un matériau a une probabilité de $P_C\%$ d'appartenir à la classe C , il sera rendu avec la fonction de transfert associée à la classe C dans $P_C\%$ des images de l'animation. Leur technique est extrêmement intéressante dans un contexte de géomodélisation, car les incertitudes d'interprétation influence fortement la chaîne de création d'un géomodelle, notamment lors de la détermination de faciès à partir des attributs d'un cube sismique.

Incertainitudes à l'échelle régionale Les incertitudes à échelle régionale sont fréquemment associées à des mesures statistiques effectuées sur des populations entières, et utilisées dans le cadre de programmes de santé publique. Par exemple, [Goovaerts \(2006\)](#) utilise des méthodes d'animation par images clef pour représenter les incertitudes spatiales associées au risque de cancer du sein et du pancréas dans des états faiblement peuplés des USA. De par la faible densité de population et la rareté des cancers, les estimations présentent une fiabilité limitée dans ces zones; pour répondre à cette limitation, [Goovaerts \(2006\)](#) présente une approche par champ de probabilité pour générer différentes réalisations des risques de cancer.

[MacEachren et al. \(1993\)](#) étudient la teneur en azote inorganique dissout dans les eaux de la baie de Chesapeake, où la présence d'azote N_2 peut induire un risque sanitaire si sa concentration est trop élevée. Ils proposent de modifier la taille de patches sur une carte pour indiquer le degré local d'incertitude; les patches dont la taille est fortement réduite sont plus difficiles à voir et indiquent donc une incertitude élevée. Cette technique limite

cependant la perception dans les zones fortement incertaines, où les patches sont assimilables à des points infinitésimaux. MacEachren *et al.* (1993) présentent également des cartes bivariées indiquant à la fois la teneur en azote inorganique et l'incertitude associée à cette teneur. Leur carte bivariée comprend trois couleurs de teintes différentes. Pour chacune de ces trois teintes, le degré d'incertitude local est représenté par trois niveaux différents de saturation et de valeur de la couleur, les zones incertaines présentant une saturation plus élevée et une valeur plus faible que les zones fiables. Cette stratégie est généralement applicable pour visualiser les incertitudes pétrophysiques en géomodélisation.

2.3.6 Incertitudes en climatologie

Les processus intervenant en climatologie mettent en jeu des interactions complexes entre les masses d'air, la topologie, les courants marins, *etc*, qui se traduisent par des phénomènes non-linéaires dont la prédiction est fortement incertaine. Un certain nombre de méthodes ont été développées pour communiquer ces incertitudes dans l'espace.

Wittenbrink *et al.* (1995b, 1996) placent un ensemble de glyphes sur des cartes de courants aériens et marins, échantillonnés régulièrement le long du domaine d'étude. Wittenbrink *et al.* (1995b, 1996) associent l'orientation des flèches figurées à l'orientation locale du courant et leur longueur à l'intensité du courant, tandis que la forme des flèches indique le degré d'incertitude; l'incertitude sur l'orientation est encodée par l'angle des lignes formant la hampe de la flèche, et l'incertitude sur l'intensité du courant est portée par l'aire délimitée par l'enveloppe autour de la pointe de la flèche (Figure 2.14). Leurs glyphes présentent une sémantique très spécifique, qui limite son utilisation à des applications ciblées en géomodélisation (par exemple l'étude d'orientations de stries pour mieux caractériser la tectonique régionale, l'orientation de paléochenaux pour comprendre un paléo-régime d'écoulement, *etc*).

MacEachren (1992) propose de modifier la résolution des données en fonction de l'incertitude; les données incertaines sont ainsi présentées avec une résolution grossière, et les données fiables à la résolution la plus fine possible. Il illustre cette proposition sur le rendu de paléo-lignes de côtes au Crétacé, pour lequel la résolution des lignes de côtes varie en fonction de l'incertitude prédite par le modèle de paléo-courants utilisé. Le changement de résolution est une des quatre méthodes proposées par MacEachren (1992) pour manipuler la variable visuelle *focus*, qu'il ajoute à l'ensemble des variables visuelles de Bertin. Cette approche limite cependant la qualité de l'affichage dans les zones les plus incertaines, ce qui peut s'avérer nuisible à la compréhension d'un modèle.

Boller *et al.* (2010) modifient l'épaisseur de lignes de courant pour représenter l'incertitude sur la trajectoire de particules portées par des courants aériens. Ils proposent d'accroître l'épaisseur des positions fiables afin d'augmenter leur impact visuel; de leur propre aveu, cette convention va à l'encontre des approches classiques pour lesquelles une épaisseur plus importante indique un plus large éventail de positions possibles, donc une

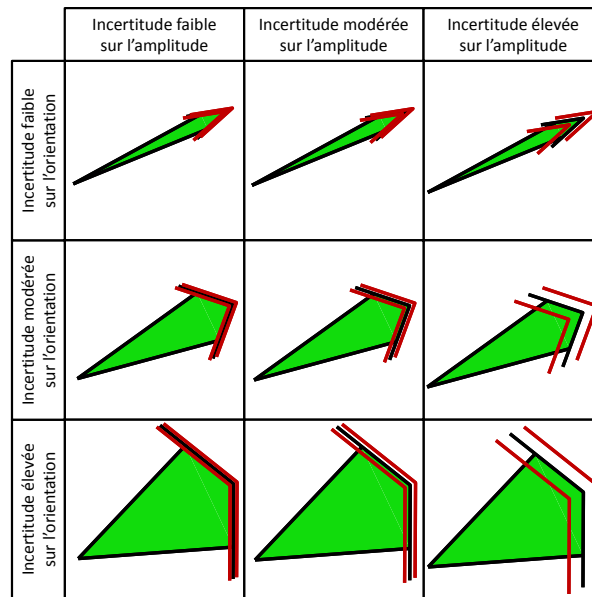


Fig. 2.14 – Encodage de l'incertitude de vecteurs sur la forme d'un glyphe. Les têtes de la flèche (chevrons rouges) indiquent l'incertitude sur l'amplitude du courant, et l'ouverture de la flèche représente l'incertitude sur l'orientation du courant. Adapté d'après [Wittenbrink et al. \(1995b, 1996\)](#).

plus forte incertitude. [Boller et al. \(2010\)](#) présentent ces techniques dans le cadre de l'évaluation des zones de formation des ouragans et apportent des indications suggérant que les ouragans se développent initialement dans une zone calme, dénommée "poche marsupiale".

[Djurcilov et Pang \(2000\)](#) et [Pang \(2006\)](#) proposent la visualisation de lignes incertaines en brisant ponctuellement leur parcours, leur donnant un aspect "en pointillés". Plus la ligne est incertaine, plus la fréquence et la longueur des zones brisées augmente ; cette méthode rend les lignes plus indistinctes lorsque les brisures sont fréquentes. Ils illustrent les lignes brisées sur des données de climatologie, en utilisant une capture du radar NEX-RAD ([Djurcilov et Pang, 1999](#)) sur laquelle la plupart des positions spatiales n'ont pas de données associées.

[Schmidt et al. \(2004\)](#) étudient l'usage combiné de méthodes haptiques et de méthodes visuelles pour faciliter la compréhension de données multivariées présentant une dimensionnalité importante. Ils relient le degré d'incertitude associé à la profondeur des fonds marins au degré de rigidité renvoyé par le matériel haptique ; une incertitude élevée produit ainsi un retour de force important. Cette méthode ne permet qu'une évaluation ponctuelle du degré d'incertitude. [Schmidt et al. \(2004\)](#) étendent la surface perçue en balayant l'aire couverte par la carte, de telle sorte que l'utilisateur puisse se faire une idée raisonnable de l'évolution du degré d'incertitude sur l'ensemble du jeu de données ;

cette amélioration ne permet cependant pas de s'affranchir pleinement des limitations d'échantillonnage perceptuel du modèle et sont donc peu applicables en géosciences.

Djurcilov *et al.* (2001) utilisent un jeu de textures pour représenter l'incertitude associée à des données volumiques de taux de salinités océaniques, dans une approche similaire aux travaux de Interrante (2000). Ils superposent ces textures à une image générée par rendu volumique ; leur jeu de textures présente des niveaux de gris croissant avec l'incertitude. Le choix des textures à utiliser est basé sur une image en niveau de gris du degré d'incertitude, générée de manière similaire au rendu volumique lui-même.⁶ Djurcilov *et al.* (2001) proposent d'autres techniques pour intégrer le degré d'incertitude en traitement *à posteriori* du rendu volumique, par insertion de grains ou par ajout de bruit dans les zones incertaines ; ces techniques permettent à l'utilisateur de contrôler respectivement la taille des grains et le pourcentage de pixels à bruiser. L'approche de Djurcilov *et al.* (2001) est particulièrement intéressante dans une problématique de géomodélisation, car elle permet de visualiser simultanément une propriété d'intérêt et les incertitudes associées dans un contexte entièrement tridimensionnel.

2.3.7 Incertitudes dans les systèmes d'information géographiques

La recherche associée aux systèmes d'information géographiques a une longue tradition de caractérisation des incertitudes simultanément à la collecte des données ; de ce fait, un volume important de publications traite des méthodes permettant de visualiser conjointement ces deux informations. Nous subdivisons les incertitudes associées aux systèmes d'information géographiques en trois familles : les incertitudes sur la position d'entités géographiques, les incertitudes sur la catégorisation des terrains, et les incertitudes sur les attributs d'une carte.

Incertaines de position Les incertitudes de position traitent des erreurs possibles associées aux coordonnées spatiales d'entités géographiques. Par exemple, McGranaghan (1993) utilise des ellipses d'erreur pour indiquer l'incertitude sur la position d'un point dans l'espace. Ces ellipses peuvent être vues comme des enveloppes dans lesquelles le point a une probabilité p de se trouver ; l'aire de l'ellipse indique alors l'intensité de l'incertitude sur la position du point. McGranaghan (1993) propose également de sélectionner les lignes indiquant la présence de cours d'eau dont l'incertitude spatiale est inférieure à un seuil défini par l'utilisateur – son approche épaissit et assombrit les lignes sélectionnées, résulte ainsi en une séparation booléenne des lignes en fonction de leur qualité.

Dans une approche similaire, MacEachren *et al.* (1993) et McGranaghan (1993) modifient l'épaisseur de lignes pour rendre leur position ambiguë : des lignes fines indiquent une position avec précision, tandis que la position de lignes épaisses est plus délicate à in-

⁶Cette approche n'est cependant pas équivalente à une visualisation pleinement intégrée au rendu volumique car la texture n'affecte pas directement le calcul de l'intégrale de couleurs.

interpréter. McGranaghan (1993) présente également les lignes incertaines avec un dégradé latéral au fur et à mesure que la probabilité de voir la ligne occuper la position courante diminue. Il propose deux manières de réaliser le dégradé, par bandes ou de manière continue ; toutes deux donnent aux lignes incertaines un aspect plus flou que les lignes fiables. Ces approches sont intéressantes en géomodélisation, car ces concepts peuvent être étendus à des objets tridimensionnels pour représenter les incertitudes de position de surface géologiques.

McGranaghan (1993) présente l'incertitude sur la position d'un polygone ou d'une ligne de rivage à l'aide d'une superposition de lignes sur un même plan visuel. L'aire couverte par les positions possibles du polygone indique ainsi son degré d'incertitude. La visualisation des lignes de rivage est par ailleurs renforcée par un assombrissement des lignes, indiquant si la campagne géographique dont est extraite la ligne de rivage est à jour ou non.

Lodha *et al.* (2002a) présentent les incertitudes associées à la position de positions fournies par un GPS sous la forme de zones diffuses, représentant la fonction de densité de probabilité associée à la position donnée. La forme de la zone permet également d'intégrer la notion d'incertitude directionnelle si la position GPS indique une entité en déplacement. De nouveau, cette approche est utilisable en géomodélisation pour représenter les incertitudes géométriques d'objets géologiques.

Incertitudes de catégorisation Les incertitudes de catégorisation illustrent la validité selon laquelle une zone du terrain étudié peut être englobée dans une classe conceptuelle. Différentes méthodes de catégorisation ont été proposées et sont utilisées pour quantifier ces incertitudes.

Par exemple, Burt *et al.* (2005) appliquent une méthode de classification floue pour obtenir une probabilité locale d'appartenance d'un terrain à un type de sol. Chaque pixel ou polygone est colorié en fonction de la classe qui lui est la plus probable ; le degré de saturation local est ensuite modifié en fonction du degré de probabilité d'appartenance à cette classe, d'après les modèles de couleurs HSV (Annexe B), $L^*a^*b^*$, $L^*u^*v^*$ ou Ljg. Burt *et al.* (2005) plaident en faveur des modèles de couleur basés sur des critères de perception plutôt que le modèle HSV, car des couleurs différentes ayant même degré de saturation au sens du modèle HSV ne seront pas nécessairement perçues comme ayant le même degré d'incertitude.

MacEachren (1992) propose de manier la variable visuelle *focus* par le biais de la clarté de remplissage de figurés sur une carte géographique. Il présente ainsi une unique catégorisation du type de terrain, représenté par un type de figuré particulier, et altère la netteté de ces figurés en fonction de l'incertitude sur la catégorie du terrain. Il illustre cette méthode sur des surfaces supposées boisées, rendant moins distinct le remplissage des zones où la présence de forêts est incertaine. Cette méthode est voisine des techniques utilisant explicitement un flou Gaussien.

Fisher (1993) introduit une animation de cartes de terrains dans laquelle il sélectionne aléatoirement certains des patches constituant la carte. Le type de terrain associé à ces patches est alors réévalué à l'aide d'un tirage de Monte-Carlo au sein de la fonction de densité cumulée locale des types de terrain, supposée connue à l'avance.

Davis et Keller (1997) animent des cartes de stabilité de sols, où la stabilité du sol est sujette à des incertitudes multiples – type de terrain, déviation des paramètres mécaniques du terrain autour des valeurs moyennes et incertitude sur l'importance du dénivelé local. Ces incertitudes sont échantillonnées à l'aide de deux approches : la logique floue pour le type de terrain, et l'échantillonnage de Monte-Carlo pour les paramètres mécaniques du terrain et le dénivelé local. Leur technique d'animation fait vibrer l'ensemble de ces paramètres ; le degré de stabilité de la pente est affiché à l'aide de la teinte de la couleur (vert pour une bonne stabilité, rouge pour une stabilité critique) et le degré d'incertitude associé à l'aide de la valeur de la couleur (couleurs claires pour les zones fiables, sombres pour les zones incertaines). Davis et Keller (1997) combinent ainsi des incertitudes catégoriques et des incertitudes sur les attributs du terrain.

Incertaines d'attributs Les incertitudes sur les attributs affectent la qualité selon laquelle une variable associée au terrain est caractérisée. Ces incertitudes sont généralement associées à des grandeurs continues.

Cedilnik et Rheingans (2000) étudient une carte satellite où une défaillance du système a empêché la collecte de certaines données, qui ont du être évaluées à l'aide de méthodes géostatistiques. Pour visualiser la perte de qualité associée aux données reconstruites, ils surchargent une grille régulière d'un réseau de *lignes d'annotation*, et déforment ces lignes de manière procédurale en fonction du degré local d'incertitude. Ils proposent trois méthodes de déformation des lignes : la modification locale de l'épaisseur des lignes, de la netteté des lignes, et l'ajout de bruit. Ces méthodes maintiennent une énergie des lignes constante, afin de normaliser la perception de l'incertitude le long du plan visuel. La technique d'annotation présentée dans leurs travaux peut être vue comme une variation de Wright (1942), qui modifie la valeur et la texture de lignes de contour afin d'exprimer les incertitudes sur la position d'une frontière. L'approche proposée par Cedilnik et Rheingans (2000) produit des résultats intéressants, car elle limite les occlusions entre la visualisation de la propriété d'intérêt et la visualisation des incertitudes associées ; de ce fait, les annotations procédurales ont un fort potentiel pour visualiser des incertitudes pétrophysiques en géomodélisation.

MacEachren (1992) augmente l'opacité de "l'atmosphère" sur le plan visuel en fonction du degré local d'incertitude ; sa technique résulte en un brouillard dense dans les zones les plus incertaines (Figure 2.15). Une approche semblable est adoptée par McGranaghan (1993), qui anime des zones de brouillard pour rendre ambiguës les portions incertaines d'une carte. Cette approche présente cependant des risques d'occlusions de la propriété d'intérêt par le brouillard dans les zones les plus incertaines.

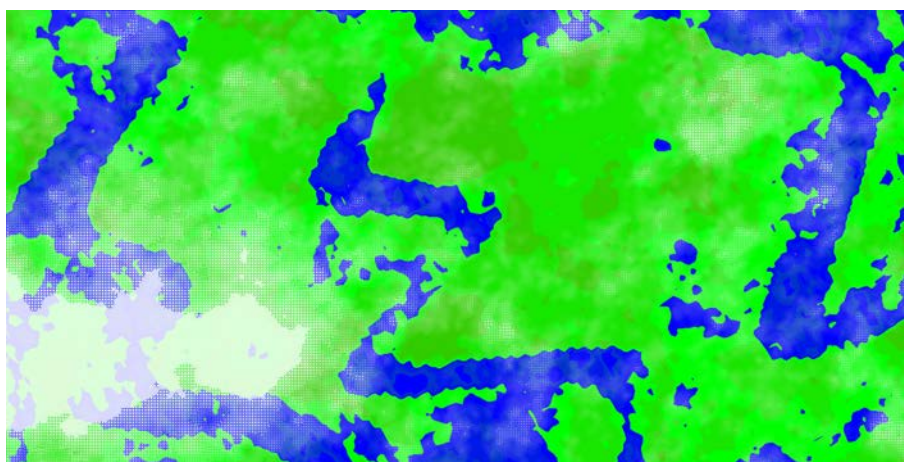


Fig. 2.15 – Visualisation de chenaux avec brouillard au dessus des zones incertaines. Adapté d'après l'idée de MacEachren (1992) et McGranaghan (1993).

Kosara *et al.* (2001) introduisent la notion de *profondeur de champ sémantique* comme nouvelle technique de focus et contexte (Robinson, 2006). Leur méthode utilise le flou inhérent aux objets hors-champ en photographie, dans le but de diminuer l'attention portée aux éléments les moins pertinents d'une scène visuelle. Ils illustrent leur méthode sur un système d'apprentissage de jeu d'échec où seules les pièces menaçant ou défendant une case doivent être perçues. Ils présentent également une carte géographique où seul le réseau routier apparaît dans le champ, les autres informations (courbes de niveau, réseau hydrographique, *etc.*) étant présentées hors-champ. La notion de flou est intuitivement liée au concept d'incertitudes ; cependant, elle dégrade la qualité de la visualisation dans les zones fortement incertaine et peut donc s'avérer limitante dans certains cas.

Dooley et Lavin (2007) affichent un ensemble de 354 cartes de rusticité de terrains sur le sol américain, produites à l'aide de méthodes d'interpolation différentes et à des résolutions différentes. Les cartes sont combinées entre elles par "transparence", à l'aide des opérateurs de composition définis par Porter et Duff (1984). Leur approche présente simultanément les modèles et permet efficacement de dégager des tendances communes à l'ensemble des cartes, mais est plus délicate à interpréter en termes d'incertitudes locales du fait de la surcharge d'informations qu'elle provoque. Pour limiter ce problème, Dooley et Lavin (2007) proposent une animation de cartes isométriques,⁷ où différentes cartes de rusticité de terrains sont présentées. Leur animation permet ainsi d'afficher l'une après l'autre l'ensemble des cartes produites.

Ehlschlaeger *et al.* (1996) présentent une méthode d'animation de surfaces préservant la variabilité spatiale des données dans les images interpolées entre deux réalisations clef. Leur méthode fonctionne sur des données ayant une distribution Gaussienne ou sub-

⁷Les cartes isométriques représentent un champ scalaire continu à l'aide d'un ensemble discret de couleurs, permettant aisément de repérer des seuils dans le champ scalaire.

Gaussienne, à l’aide de techniques de déformation graduelle telles que présentées dans les travaux de Hu (2000). Cette approche est extrêmement intéressante en géomodélisation, car elle permet de visualiser des incertitudes géométriques d’une manière très intuitive sans dégrader la qualité des information lors de l’interpolation temporelle.

2.3.8 Incertitudes en géomodélisation

Malgré le rôle prépondérant des incertitudes en géomodélisation, peu de travaux ont à ce jour été publiés dans ce domaine. Les raisons de ce faible nombre de publications sont probablement culturelles, les ingénieurs de réservoir étant habitués à défendre un unique modèle (Srivastava, 1995); nous présentons dans cette section quelques travaux allant à l’encontre de cette tendance.

Hengl (2003) propose deux méthodes de visualisation, dédiées à des données de type continu ou catégorique, en se basant sur les fonctions de conversion de couleur définies par le modèle HSI. Il illustre l’application à des données continues à l’aide d’un krigeage sur l’épaisseur d’une formation géologique, qui fournit également la variance d’estimation associée à l’épaisseur en tout point. Il représente l’épaisseur grâce à la teinte, et la variance grâce à la “pâleur” de la couleur; une couleur pâle est obtenue par augmentation de l’intensité et diminution de la saturation de la couleur (Figure 2.16). Sa méthode de visualisation dédiée à des données catégoriques est présentée sur une classification floue sur de types de reliefs. Pour chaque pixel ou polygone, la couleur est calculée comme une moyenne pondérée des couleurs $C'_R = f(C_R, p_R)$, où C_R est la couleur associée au relief de type R , p_R est la probabilité d’appartenance du pixel ou polygone à la classe R , et $f()$ est la fonction qui augmente la pâleur d’une couleur en fonction de la probabilité d’appartenance (Hengl *et al.*, 2002; Hengl et Toomanian, 2006).

Zehner *et al.* (2010) illustrent les travaux de Hengl (2003) le long d’isovaleurs de température dans le cadre d’une étude de géothermie. Ils comparent plusieurs approches basées sur le système de couleurs HSV et concluent que les méthodes reliant l’incertitude locale à une teinte gris sombre (faible saturation et faible valeur) présentent plus intuitivement le degré d’incertitude que les méthodes utilisant des teintes pâles, car les zones gris sombre présentent un aspect “sale” qui est intuitivement associé à la notion d’incertitude.

Zehner *et al.* (2010) proposent également des glyphes “ligne” échantillonnés le long d’iso-surfaces de température du sous-sol, et orientés perpendiculairement à ces isosurfaces. La longueur de ces lignes permet ainsi de visualiser les incertitudes sur la position de l’iso-surface en définissant un volume enveloppe, dont les limites supérieure et inférieure sont calculées respectivement par addition et par soustraction de la déviation standard à la valeur moyenne de la propriété d’intérêt.

Srivastava (1994a) propose une méthode d’animation de types de faciès (sable ou argile) basée sur la technique des champs de probabilité. Il génère un champ de probabilité dont la taille est supérieure au modèle, puis présente séquentiellement les réalisations de

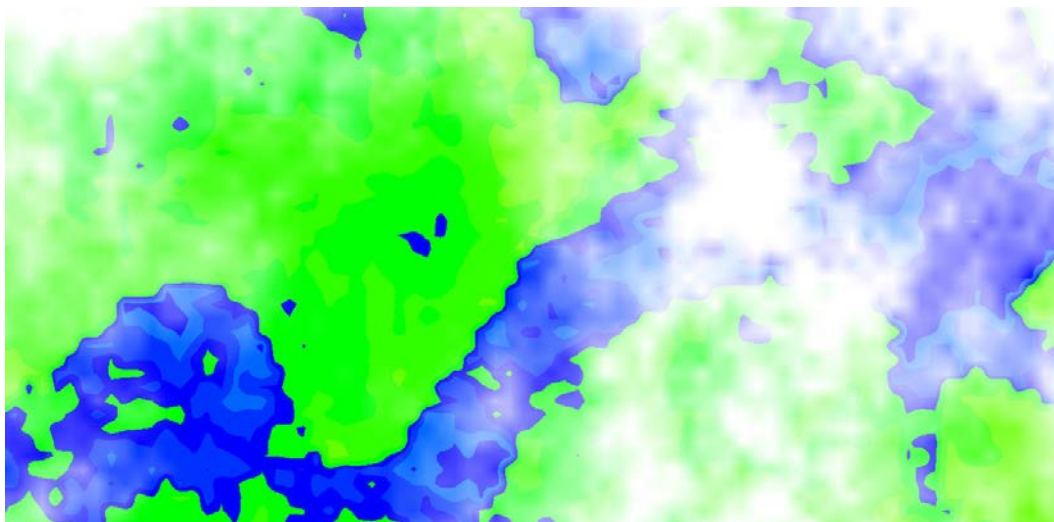


Fig. 2.16 – Visualisation de chenaux avec une saturation décrue dans les zones incertaines. Adapté d'après l'idée de Hengl et al. (2002) et Hengl et Toomanian (2006).

faciès associées en se déplaçant le long du champ de probabilité. Sa méthode permet de classer les réalisations de la plus pessimiste à la plus optimiste à l'aide d'une fonction objective (par exemple, en fonction du volume de sable connecté) ; le problème se résume alors à trouver un chemin dans le champ de probabilité pour lequel la fonction objective varie de manière monotone (Srivastava, 1994b).

2.4 Evaluation de la qualité des visualisations

L'ensemble des méthodes présentées dans la section 2.3 permet de visualiser à la fois des données d'intérêt et leurs incertitudes associées dans l'espace. Cependant, ces caractéristiques ne garantissent pas pour autant qu'une méthode de visualisation des incertitudes soit utilisable dans un cadre pratique ; pour être pleinement validée, une visualisation des incertitudes doit également favoriser une bonne perception des informations présentées. Nombre d'auteurs proposent une évaluation de leurs méthodes de visualisation basée sur leur ressenti personnel. En l'état actuel des connaissances, il est néanmoins possible de valider la qualité des algorithmes de visualisation de manière plus rigoureuse ; deux familles d'approches sont présentées dans la littérature pour y parvenir. La première méthode propose un ensemble de critères de qualité théoriques (Zuk et Carpendale, 2006), destinés à maximiser la perception des informations. Cependant, ces critères généraux peuvent être difficiles à appliquer à des cas concrets, car plusieurs principes peuvent entrer en compétition. Une seconde méthode est présentée dans la littérature ; elle propose une approche purement empirique par le biais d'études statistiques sur un panel d'utilisateurs potentiellement intéressés (Kosara *et al.*, 2003; Tory et Möller, 2004).

Contexte de recherche La représentation des incertitudes spatiales n’a que peu d’intérêt si elle n’apporte aucune amélioration dans le processus de prise de décision (Deitrick et Edsall, 2006) ; ces améliorations peuvent se traduire sous la forme d’une décision prise plus rapidement ou avec plus de précision (Harrower, 2003). Un certain nombre d’études sur des utilisateurs ont été menées pour mieux comprendre la manière dont un utilisateur perçoit des données présentées conjointement avec un taux de confiance spatial, notamment dans le domaine des SIGs (par exemple, Evans, 1997; Aerts *et al.*, 2003; Hope, 2005; Roth, 2009). Cette section résume les plus importantes de leurs découvertes.

Effets de la représentation des incertitudes conjointement aux données initiales McGranaghan (1993) affirmait dans ses travaux que la visualisation d’un degré de confiance surimposé aux données initiales produisait nécessairement une surcharge perceptuelle et cognitive, obscurcissant ainsi le message porté par les données. Cependant, les recherches menées par Leitner et Buttenfield (2000) et Aerts *et al.* (2003) concluent que ce surplus d’information permet au contraire de clarifier les informations présentées, les informations peu fiables étant moins prises en considération que des informations de confiance. Leurs résultats rejoignent les conclusions de Evans (1997) et Edwards et Nelson (2001), dont les études montrent que les utilisateurs perçoivent mieux les informations présentées dans les données avec des représentations coïncidentes qu’avec des représentations adjacentes du degré spatial d’incertitude.

Perception en fonction du degré local d’incertitude L’étude menée par Sanyal *et al.* (2009) demandait aux participants d’identifier des zones de faible et de forte incertitude. En reliant le degré local d’incertitude à la taille de glyphes, les participants ont aisément réussi à localiser des zones de confiance ; en liant l’incertitude locale à la couleur de glyphes ou à la couleur de la surface étudiée, les participants ont plus facilement pu localiser des zones fortement incertaines. Sanyal *et al.* (2009) en concluent que la perception du degré d’incertitude n’est pas uniforme ; la qualité de cette perception peut varier en fonction du degré local d’incertitude lui-même, ainsi que de la méthode utilisée pour le visualiser.

Perception en fonction de la dimension des données L’étude sur des utilisateurs de Sanyal *et al.* (2009) inclut des tâches de nature similaire – identification de zones de faible ou de forte incertitude et dénombrement d’éléments au sein des données – appliquées à des données en une dimension (graphe) ou deux dimensions (surface). De manière générale, Sanyal *et al.* (2009) rapportent que le traitement des questions utilisant des données en deux dimensions a demandé significativement plus de temps aux participants que les tâches analogues sur des données en une dimension. Ils indiquent également que la précision des réponses est supérieure pour les données en une dimension, probablement du fait de la complexité moindre du jeu de données.

Perception en fonction du degré d'expertise Plusieurs études sur des utilisateurs comparent les réponses des participants en fonction de leur expérience vis-à-vis de la tâche demandée. Il n'existe pas de consensus clair quand aux effets de l'expertise sur la prise de décision sous incertitude – Kobus *et al.* (2001) et Hope et Hunter (2007) rapportent que le degré d'expertise des participants affecte la qualité de leurs décisions, tandis que Evans (1997) et Aerts *et al.* (2003) concluent qu'il n'existe pas de différences significatives entre les réponses d'experts et de novices. Roth (2009) suggère que ces résultats contradictoires peuvent être liés à la complexité des tâches demandées aux participants : des tâches simples peuvent aussi bien être effectuées par des novices que par des experts, tandis que des tâches complexes requièrent une expertise suffisante pour être menées à bien.

2.5 Conclusion

La problématique de visualisation des incertitudes n'est pas nouvelle et bien que son application à des problématiques de géomodélisation soit actuellement limitée, de nombreux autres domaines où la caractérisation des incertitudes a des implications importantes, tels que la climatologie, la médecine ou les systèmes d'information géographiques, ont développé des outils de visualisation adaptés à leurs besoins spécifiques. Nous avons proposé dans cette section un état de l'art aussi complet que possible sur ces techniques de visualisation.

Nous avons tout d'abord présenté une introduction aux techniques de visualisation scientifique en général, en ciblant tout particulièrement deux des approches de visualisation volumique les plus populaires : l'extraction d'isosurfaces et le rendu volumique de champs scalaires tridimensionnels. De nombreuses améliorations et optimisations sont présentées dans la littérature ; plutôt qu'une revue exhaustive, nous avons introduit un sous-ensemble de ces techniques pouvant présenter un intérêt manifeste en géomodélisation.

Dans un second temps, nous avons présenté les variables visuelles permettant de représenter visuellement les grandeurs associées à des modèles scientifiques, telle que proposées par Bertin (1983). Les méthodes de visualisation des incertitudes relient souvent explicitement l'incertitude locale à ces variables visuelles ; nous avons ensuite proposé une revue des méthodes de visualisation spécifiquement dédiées aux incertitudes, organisées en fonction du domaine dans lequel elles ont été appliquées. Parmi les méthodes disponibles, toutes ne sont pas adaptées à la géomodélisation ; nous nous efforçons de dégager, pour chaque type de visualisation, les critères pouvant favoriser ou défavoriser leur application à un géomodèle.

Un certain nombre de critères de qualité perceptuels ont par ailleurs été décrits dans la littérature pour évaluer la validité d'une méthode de visualisation ; nous avons en conclu-

sion présenté les approches empiriques de recherche de critères de qualité visuelle, qui sont généralement basées sur les résultats d'études menées directement sur des utilisateurs.

Chapitre 3

Visualisation des incertitudes pétrophysiques en géomodélisation

Sommaire

3.1	Méthodes statiques dédiées	62
3.1.1	Modification des attributs de la couleur	62
3.1.2	Modification du codage par couleur	64
3.1.3	Surcharge d'informations	69
3.1.4	Ambiguation	78
3.1.5	Modification de la géométrie	84
3.2	Validation des méthodes statiques	87
3.2.1	Participants	87
3.2.2	Description des données	89
3.2.3	Protocole	89
3.2.4	Analyse des résultats	91
3.2.5	Résultats	92
3.2.6	Discussion	94
3.3	Méthodes dynamiques dédiées	95
3.3.1	Animation par images clefs	95
3.3.2	Types d'interpolations temporelles continues	96
3.3.3	Animation de réalisations non-équiprobables	99
3.4	Validation des méthodes dynamiques	101
3.4.1	Description des données	101
3.4.2	Description des questions	102
3.4.3	Résultats	104
3.4.4	Limites	109
3.4.5	Conclusions	110
3.5	Application aux incertitudes pétrophysiques	111

3.5.1 Propriétés statiques	111
3.5.2 Propriétés dynamiques	114
3.6 Conclusion	117

La plupart des méthodes de visualisation des incertitudes traitent d’incertitudes sur les attributs spatiaux d’un modèle. Dans ce chapitre, nous discutons de différentes techniques dédiées à la visualisation des incertitudes pétrophysiques ; ces techniques peuvent être catégorisées en méthodes statiques, pour lesquelles l’incertitude est explicitement reliée à une ou plusieurs variables visuelles, et en méthodes dynamiques, pour lesquelles l’incertitude est implicitement portée par l’évolution temporelle de la visualisation du modèle. Chaque approche est validée par le biais d’une étude sur des utilisateurs, mettant en valeur les forces et les faiblesses de ces deux méthodologies.

Dans le cadre de cette thèse, nous adoptons une taxonomie des méthodes de visualisation proche de celle adoptée par Pang *et al.* (1997), dans laquelle les méthodes de visualisation sont classées selon la technique utilisée pour présenter l’information d’incertitude (d’autres critères pourraient également être prises en compte, par exemple le but recherché, ou les variables visuelles exploitées par la visualisation des incertitudes). On distingue ainsi les méthodes par *surcharge d’informations*, *modification des propriétés optiques*,⁸ *modification ou ajout de la géométrie*, *ambiguation*, *affichage de réalisations multiples* par animation (Table 3.1). Pang *et al.* (1997) décrivent d’autres approches telles que les méthodes *psycho-visuelles* et les méthodes *non visuelles*.⁹ Ces techniques présentent un intérêt limité en géosciences et ne sont donc pas utilisées dans cette thèse.

3.1 Méthodes statiques dédiées

Les méthodes statiques dédiées à la visualisation des incertitudes relient explicitement le degré local d’incertitude à une des variables visuelles de la scène graphique. Cette approche requiert généralement que l’incertitude locale ait été échantillonnée en tout point du domaine d’intérêt, et évaluée à l’aide d’une métrique d’incertitude appropriée.

3.1.1 Modification des attributs de la couleur

Les techniques par modification des attributs de la couleur agissent sur les couleurs associées aux données pour présenter le degré d’incertitude local. De ce fait, ces méthodes sont intrinsèquement liées aux cartes de couleurs utilisées ; le choix de la carte de couleurs doit donc être réalisé avec soin pour éviter toute interférence, par exemple une méthode

⁸La modification des propriétés optiques inclut tant la modification des attributs de la couleur que le type de codage couleur adopté.

⁹Il convient cependant de noter que les méthodes non visuelles exploitent d’autres sens que la vue, par exemple le toucher ou l’ouïe, et ne font donc pas partie *sensu-stricto* des méthodes de *visualisation* des incertitudes.

Tab. 3.1 – Taxonomie des méthodes de visualisation des incertitudes.

Méthode de visualisation	Description
Surcharge d'informations	L'incertitude est présentée comme une surcouche imposée sur l'affichage initial
Modification des propriétés optiques	L'incertitude affecte la manière dont les matériaux captent la lumière
Modification ou ajout de géométrie	L'incertitude affecte la position et la forme des objets sur le plan visuel
Ambiguation	Augmente l'ambiguïté dans la représentation des objets incertains
Affichage de réalisations multiples	L'incertitude est portée par une série de représentations possibles d'un modèle
Méthodes psycho-visuelles	Méthodes reposant sur les propriétés du système visuel humain
Méthodes non visuelles	Méthodes explorant d'autres sens que la vue, comme l'ouïe ou le toucher

altérant la saturation en fonction de l'incertitude locale produira une visualisation de mauvaise qualité si elle est associée à une carte de couleurs peu chromatique, par exemple en niveaux de gris.

Modification des composantes de la couleur Nous proposons trois modes de modification de la couleur, altérant la transparence, la saturation et/ou la valeur de la couleur de manière similaires aux techniques présentées par [Strothotte et al. \(1999\)](#) et [Hengl \(2003\)](#). De manière générale, la composante c de la couleur à modifier est altérée avec une intensité i , qui est elle-même fonction du degré d'incertitude local :

$$c \leftarrow \varphi(c, i) \quad \text{avec} \quad \varphi() \text{ la fonction d'altération de la couleur} \quad (3.1)$$

Par défaut, la couleur est représentée sous le format RGBA ([Annexe B](#)). Si l'altération de la transparence peut être réalisée directement en espace RGB, la modification des attributs de saturation et de valeur nécessitent de convertir la couleur de l'espace de couleurs RGB à l'espace de couleurs HSV préalablement à leur altération ([Algorithme 3](#)), à l'aide des opérateurs de conversion définis par [Smith \(1978\)](#). La couleur altérée en espace HSV doit ensuite subir la transformation inverse pour être ramenée dans l'espace de couleurs RGB ([Algorithme 4](#)).

Les techniques par modification des attributs de la couleur sont directement applicables à des sections de grilles, mais peuvent tout aussi bien être appliquées à des approches par rendu volumique ([Figure 3.1](#)). Leur intégration dans un modèle optique explicite ([Section 2.1.2](#)) peut s'avérer complexe ; nous utilisons donc un modèle optique implicite,

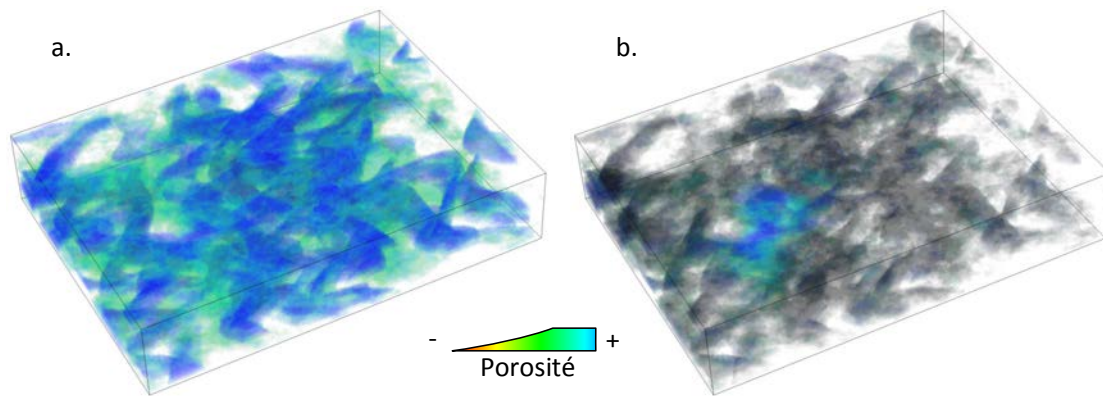


Fig. 3.1 – Visualisation des incertitudes par modification des attributs de la couleur en rendu volumique. (a) Rendu volumique de porosité dans une formation chenalisée, sans visualisation des incertitudes. (b) Rendu volumique de porosité dans une formation chenalisée avec visualisation des incertitudes par modification de la couleur. Les zones fortement incertaines présentent une altération de la valeur et de la transparence.

résultant de l'accumulation des couleurs dans le cadre d'un rendu volumique par tranchage incrémental.

Contours suggestifs Dans ce travail, nous adoptons une approche par rendu suggestif pour limiter les interférences entre la visualisation des incertitudes par modification de la couleur et la propriété d'intérêt, d'une manière similaire à Strothotte *et al.* (1999) : nous superposons les contours de la propriété d'intérêt par-dessus la couleur altérée, afin de conserver la perception de ses caractéristiques essentielles même dans les zones fortement incertaines. Strothotte *et al.* (1999) utilisent comme contours suggestifs les structures supposées de bâtiments, dessinées par un archéologue d'après leurs fondations. Dans notre cas, de telles informations ne sont pas disponibles : nous extrayons donc automatiquement les contours de la propriété d'intérêt à l'aide d'un algorithme de détection de bords, le filtre de Sobel (Danielsson et Seger, 1990), combiné à un algorithme de détection de contours suggestifs (DeCarlo *et al.*, 2003; Burns *et al.*, 2005; Lee *et al.*, 2007, Figure 3.2). Ces contours sont extraits directement depuis la carte graphique pour une efficacité maximale.

3.1.2 Modification du codage par couleur

En visualisation traditionnelle, les propriétés locales sont généralement représentées à l'aide d'une fonction de transfert visuelle 1D (carte de couleurs) associant une grandeur à une couleur. Cependant, certains auteurs proposent des fonctions de transfert à N dimensions, c'est-à-dire associant N grandeurs à une couleur (Kindlmann et Durkin, 1998;

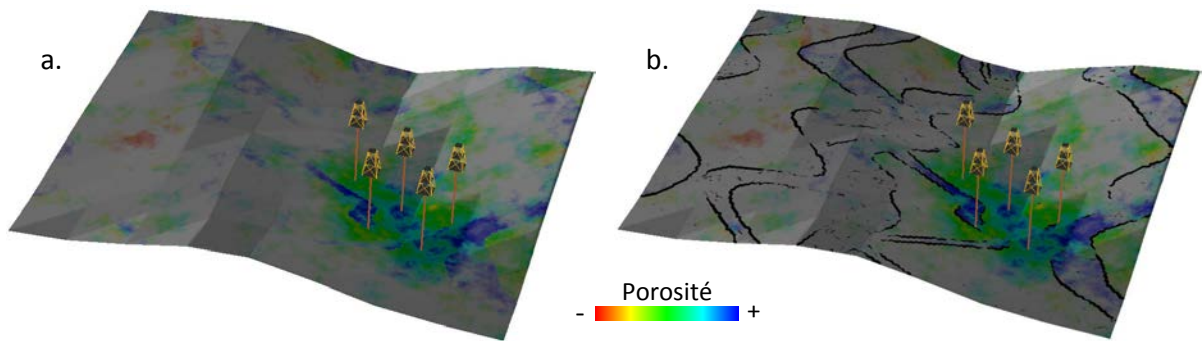


Fig. 3.2 – Utilisation du filtre de Sobel pour améliorer la perception dans les zones fortement incertaines. Les données représentent une porosité estimée conditionnellement aux faciès, avec deux faciès présents (chenal et plaine d’inondation). (a) Altération de la saturation dans les zones incertaines. La couleur représentant la porosité est atténuée par l’intensité de la désaturation. (b) Altération de la saturation dans les zones incertaines avec contours extraits à l’aide du filtre de Sobel. La couleur représentant la porosité est masquée par l’intensité de la désaturation, mais les contours permettent de reconstituer la géométrie des chenaux.

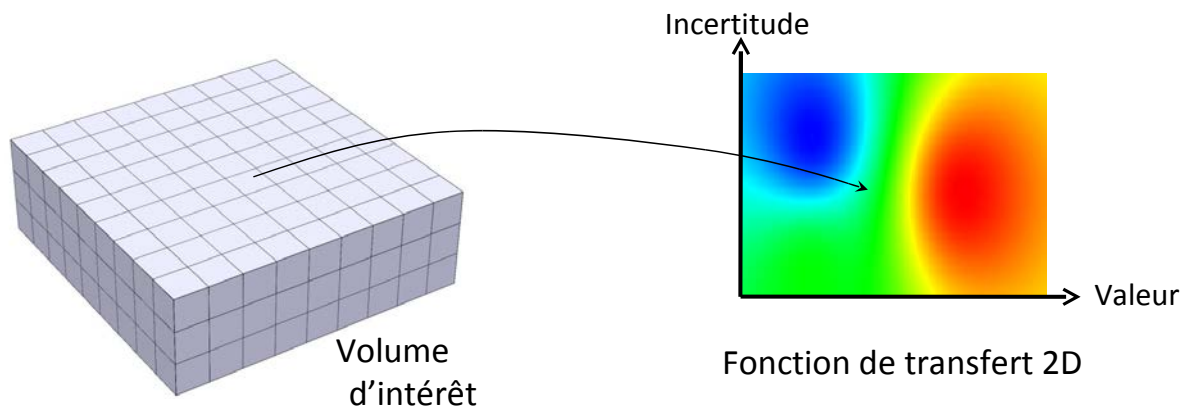


Fig. 3.3 – Relations entre l’espace physique \mathcal{P} (volume d’intérêt) et l’espace de couleurs \mathcal{C} définies par la fonction de transfert bidimensionnelle. La fonction de transfert introduit une surjection de \mathcal{P} dans \mathcal{C} .

Kniss *et al.*, 2001; Djurcilov *et al.*, 2001; Kniss *et al.*, 2005; Roettger *et al.*, 2005). Ces fonctions de transfert sont souvent appliquées à une grandeur d’intérêt et ses dérivées d’ordre N pour délimiter implicitement des matériaux différents au sein des données; cependant, elles sont également adaptées à des problématiques générales de corendu. Nous appliquons cette méthodologie à des fonctions de transfert bidimensionnelles, prenant en compte une grandeur et son incertitude associée (Figure 3.3, Viard *et al.*, 2008a).

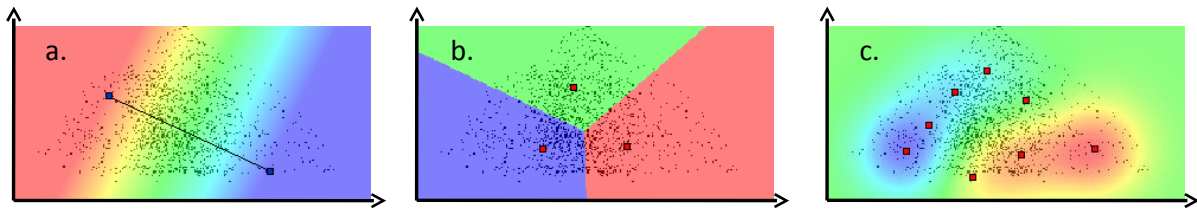


Fig. 3.4 – Schémas d’interpolation définissant une fonction de transfert bidimensionnelle. (a) Interpolation par gradient. Les carrés bleus représentent les points de contrôle (deux points), et la ligne l’orientation du gradient. (b) Interpolation par diagramme de Voronoi. Les carrés rouges représentent les points de contrôle (trois points). (c) Interpolation par krigeage. Les carrés rouges représentent les points de contrôle (huit points).

Définition de la fonction de transfert La définition d’une fonction de transfert adaptée est une étape cruciale pour permettre une visualisation correcte des données, mais qui est souvent délicate à réaliser. La méthode la plus fréquemment adoptée repose sur des itérations successives par essai-erreur, la fonction de transfert étant modifiée de manière incrémentale jusqu’à atteindre un rendu acceptable. Ce processus peut être long et n’aboutit pas toujours à des résultats satisfaisants ; certains auteurs proposent des méthodes de génération de fonctions de transfert automatiques ou semi-automatiques pour minimiser la charge de travail incombant à l’utilisateur (Kindlmann et Durkin, 1998; Roettger *et al.*, 2005), généralement basées sur des algorithmes de détection de bords dans l’espace de la fonction de transfert. Cette méthodologie est rarement applicable en géosciences, car les données à visualiser ne présentent pas nécessairement de frontières clairement définies ; cependant, nous nous inspirons de ces approches pour proposer une définition de la fonction de transfert guidée par les données.

Notre approche repose sur une définition la fonction de transfert par *points de contrôle* au niveau desquels l’utilisateur spécifie explicitement la couleur à utiliser ; en tout point de l’espace de la fonction de transfert où aucun point de contrôle n’est défini, la couleur est interpolée à l’aide d’un schéma d’interpolation prédéfini. Dans sa forme la plus simple, la fonction de transfert est interpolée par calcul du gradient entre deux points de contrôle ; cependant, des schémas plus complexes sont également disponibles (Figure 3.4), tels que l’interpolation de Shepard (distance inverse et distance inverse au carré, Shepard, 1968), le krigeage (Krige, 1966; Matheron, 1969) ou le diagramme de Voronoi des points de contrôle (Hoff *et al.*, 1999; Aurenhammer et Klein, 2000).

Définition de classes optimales Nous superposons le nuage de points représentant l’ensemble des données aux couleurs interpolées dans l’espace de la fonction de transfert. Cette représentation permet à l’utilisateur d’être conscient de la densité locale des données dans cet espace, et donc d’être guidé lors de la sélection des points de contrôle. Cependant, ce mécanisme ne garantit pas que l’utilisateur sélectionne des classes optimales au sein

de ses données ; nous avons donc intégré une méthode de partitionnement automatisée pour aider l'utilisateur dans cette tâche, l'algorithme des nuées dynamiques (*k-means* en anglais, MacQueen, 1966; Kanungo *et al.*, 2002; Gibou et Fedkiw, 2005; Har-Peled et Sadri, 2005; Arthur et Vassilvitskii, 2006). Les nuées dynamiques sont un algorithme de classification strict et non supervisé, amélioré de manière itérative. Les classes sont définies comme un ensemble de centroïdes, partitionnant l'espace à l'aide des domaines de Voronoi qui leurs sont associés.

Nous traitons l'ensemble des points de contrôle définis par l'utilisateur comme l'état initial des centroïdes. Pour chaque centroïde, nous déterminons l'ensemble des données présentes dans le domaine de Voronoi associé et recalculons la position du centroïde courant comme le barycentre de ces données ; cette étape est répétée plusieurs fois jusqu'à la convergence de l'algorithme, c'est-à-dire jusqu'à ce que le déplacement des centroïdes soit inférieur à un seuil ϵ . La position des points de contrôle définit alors une partition optimale des données à nombre de classes fixé.

Exemple Ce paragraphe propose un exemple concret d'application d'une fonction de transfert bidimensionnelle en visualisation des incertitudes géologiques. D'après des données parcellaires de porosité mesurées le long des trajectoires de puits, la porosité dans l'ensemble du domaine d'intérêt est reconstruite par interpolation à l'aide de méthodes de krigeage. Outre une estimation de la porosité, le krigeage permet également de calculer une variance d'estimation σ_{es}^2 , que l'on considère comme une métrique d'incertitude. En appliquant un partitionnement automatique à ces informations, la fonction de transfert définit trois familles (Figure 3.5) :

1. Des données à faible porosité et faible incertitude, colorées en bleu ;
2. Des données à forte porosité et faible incertitude, colorées en rouge ;
3. Des données à forte incertitude, colorées en vert.

Ces familles peuvent alors facilement être utilisées par un géologue pour déterminer les zones d'intérêt dans le modèle – par exemple, si le but est d'estimer les zones potentiellement riches en hydrocarbures, l'attention se concentrera sur les zones de forte porosité et éventuellement sur les zones de forte incertitude, mais pas sur les zones de faible porosité.

Bien que présentées sur des sections de grilles dans cet exemple, les fonctions de transfert bidimensionnelles peuvent également être incorporées dans un moteur de rendu volumique (Figure 3.6), en utilisant un modèle optique implicite résultant de l'accumulation des couleurs par tranchage incrémental. Ces approches peuvent être très utiles pour dégager rapidement des domaines d'intérêt au sein du volume étudié.

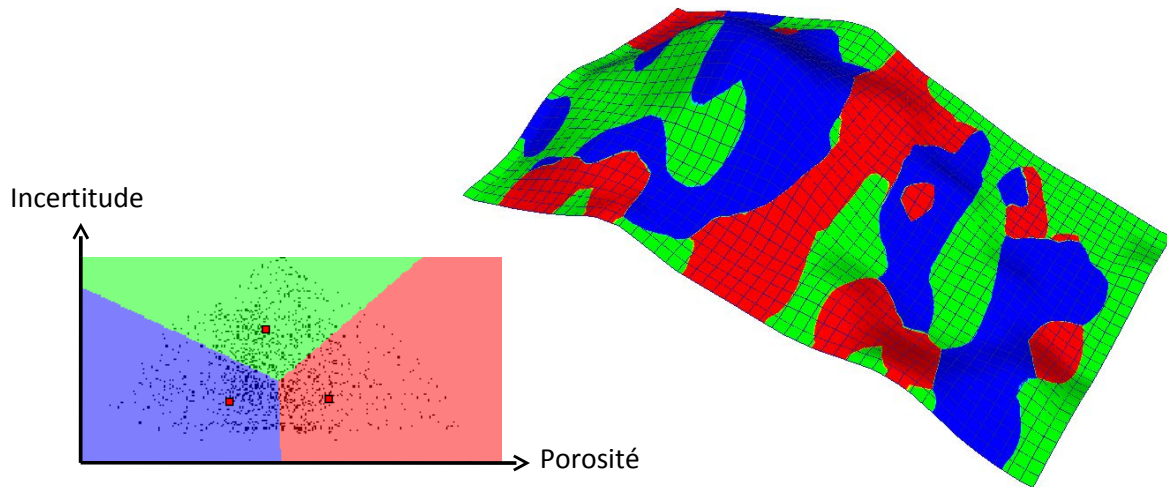


Fig. 3.5 – Fonction de transfert bidimensionnelle appliquée à des données de porosité. La fonction de transfert utilise un schéma d'interpolation par diagramme de Voronoi, qui définit trois classes dans les données : faible porosité, forte porosité et forte incertitude. Ces classes sont reportées dans l'espace physique \mathcal{P} via la surjection définie par la fonction de transfert. Données Total.

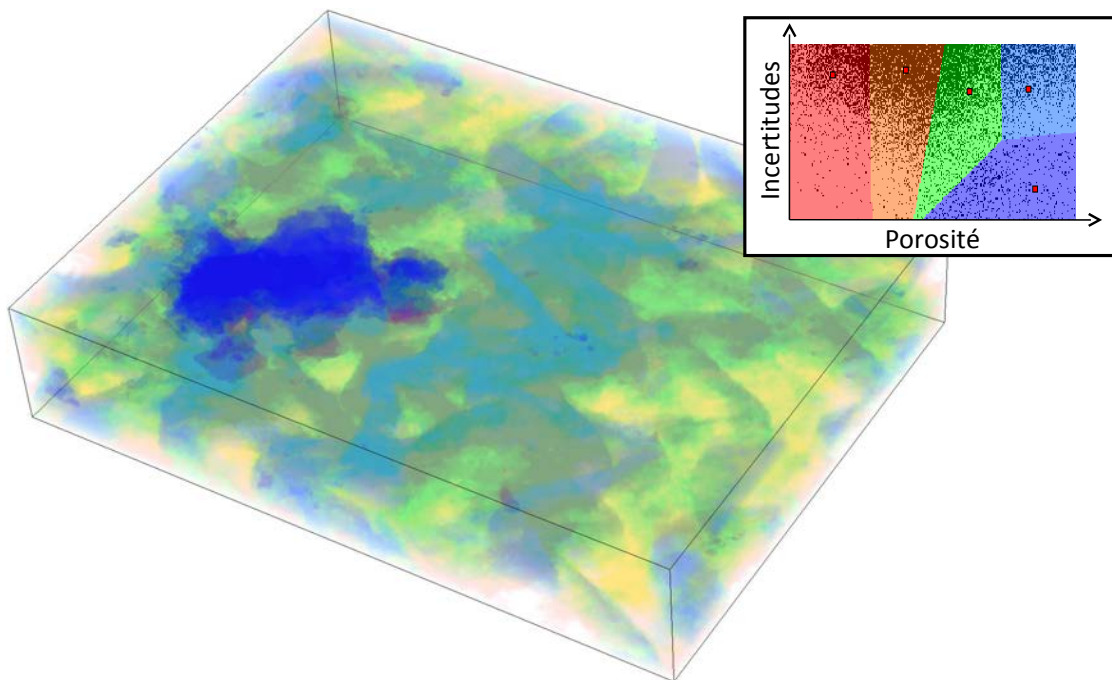


Fig. 3.6 – Fonction de transfert bidimensionnelle appliquée à des données de porosité en rendu volumique. La fonction de transfert utilise un schéma d'interpolation par diagramme de Voronoi qui définit cinq classes dans les données. Les données présentant une forte porosité et une faible incertitude associée (classe bleu foncé) sont mises en valeur par le biais d'une plus forte opacité.

3.1.3 Surcharge d'informations

Les techniques par surcharge d'information incorporent le degré d'incertitude local sous la forme d'une surcouche d'informations, surimposée à la visualisation initiale du modèle. La notion de surcouche d'informations est naturelle sur des données en deux dimensions, mais peut dans certains cas être étendue à des données volumiques. De manière générale, il est recommandé de limiter l'aire totale de l'image qui est masquée par la surcouche d'informations (Cedilnik et Rheingans, 2000), afin de minimiser les interférences entre les données d'intérêt et le degré d'incertitude local.

Glyphes Les glyphes sont des éléments figurés dont les caractéristiques sont utilisées pour représenter des grandeurs physiques ; par exemple, l'orientation d'un glyphe "flèche" peut indiquer la direction d'un courant marin. De manière similaire, les attributs de glyphes peuvent être reliés au degré d'incertitude local. Nous proposons des glyphes en forme de points d'interrogation, dont la sémantique peut être intuitivement reliée à une incertitude élevée. Ce type de figurés impose des contraintes sur la manière dont les glyphes sont reliés au degré d'incertitude local, de telle sorte que les points d'interrogation ne soient visibles que dans les zones incertaines.

Les glyphes peuvent être échantillonnés régulièrement ou aléatoirement au sein du domaine d'intérêt. Un échantillonnage régulier permet de couvrir le domaine d'intérêt aussi exhaustivement que possible ; cependant, l'attention de l'observateur peut être attirée par la structure de l'échantillonnage plutôt que par les glyphes eux-mêmes. Au contraire, un échantillonnage aléatoire ne garantit pas une couverture exhaustive de l'espace, mais limite l'émergence de structures ; l'attention de l'observateur peut tout de même être attirée par des agglomérats de glyphes dans les zones sur-échantillonnées. Nous adoptons une stratégie d'échantillonnage hybride : le domaine d'intérêt est régulièrement subdivisé en un ensemble de sous-domaines, dans lesquels un ou plusieurs glyphes sont placés de manière aléatoire. Cette stratégie garantit que chaque sous-domaine est échantillonné, tout en limitant l'émergence de structures ou d'agglomérats dans la position des glyphes. Cependant, il est possible que plusieurs glyphes entrent en collision, rendant leur lecture individuelle plus délicate (Figure 3.7).

Nous avons comparé deux méthodes différentes : la modification des attributs des glyphes en fonction de l'incertitude, comme proposé dans la littérature, et la modification de la densité de l'échantillonnage des glyphes en fonction de l'incertitude. De manière générale, les méthodes par taille et par densité d'échantillonnage produisent des résultats comparables (Figure 3.8) ; cependant, la taille des glyphes permet de visualiser les évolutions du degré d'incertitude de manière plus continue.

Motifs répétitifs La visualisation des incertitudes à l'aide de motifs répétitifs est une méthode efficace assurant en général une perception correcte des informations présentées.

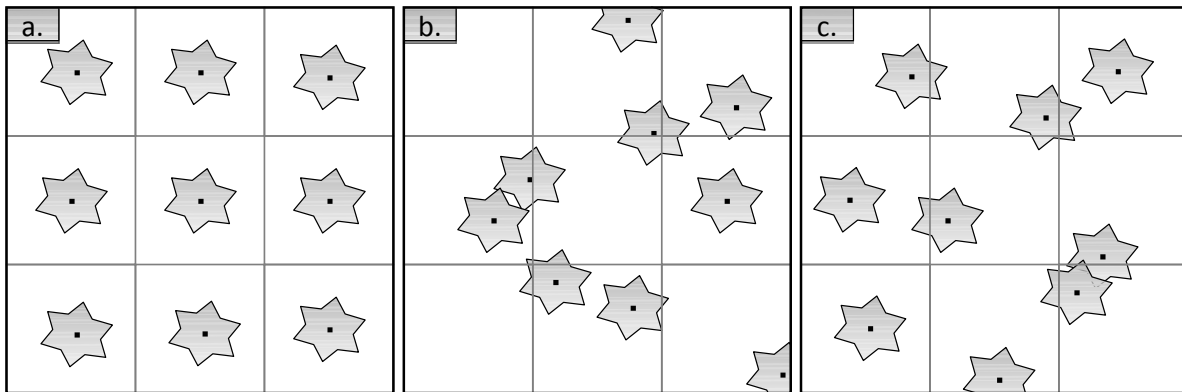


Fig. 3.7 – Stratégies d'échantillonnage des glyphes au sein du domaine d'intérêt. Le domaine d'intérêt est subdivisé en neuf sous-domaines. (a) Echantillonnage régulier. (b) Echantillonnage aléatoire. (c) Echantillonnage hybride. Les échantillonnages aléatoire et hybride peuvent générer des intersections entre glyphes quelle que soit leur taille.

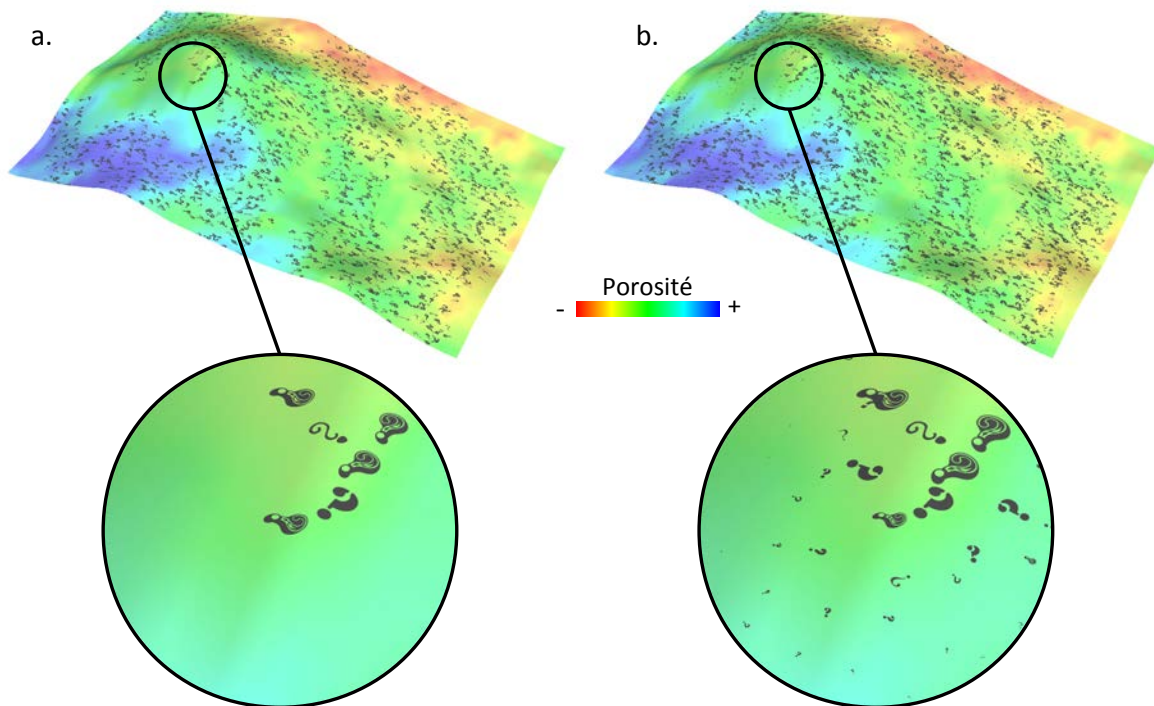


Fig. 3.8 – Comparaison entre la visualisation des incertitudes (a) par densité d'échantillonnage et (b) par taille des glyphes le long d'une surface stratigraphique. Les glyphes utilisent une stratégie d'échantillonnage hybride, et leur orientation est générée aléatoirement. La couleur représente une estimation plausible de la porosité. Données Total.

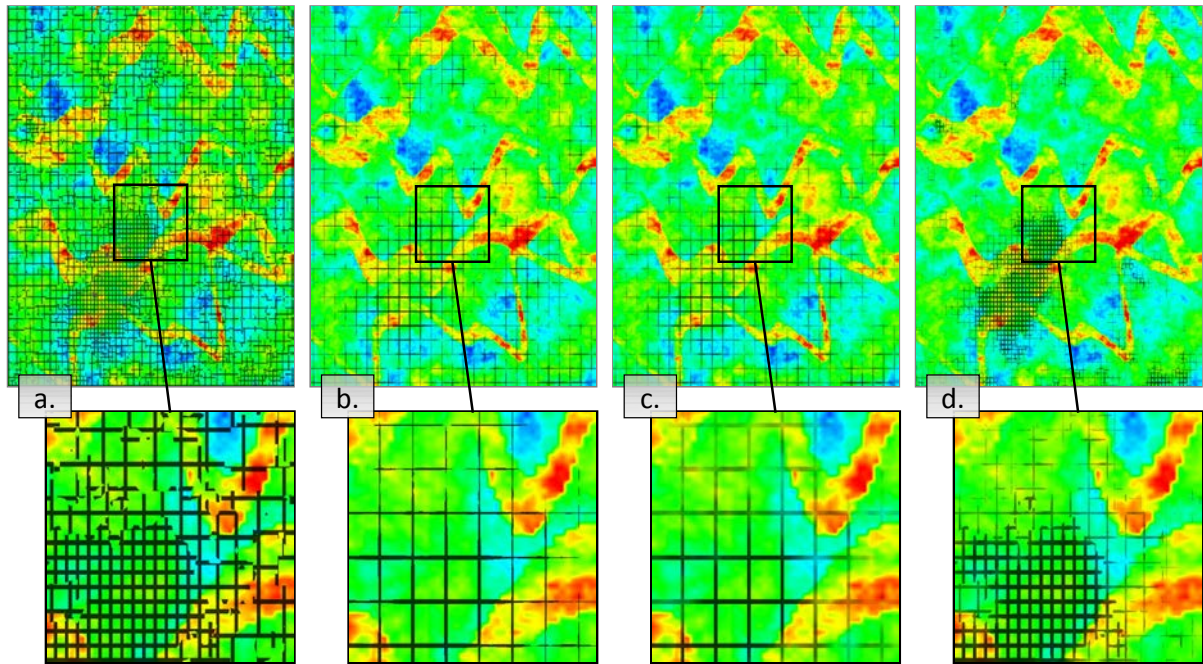


Fig. 3.9 – Méthodes d’association entre un motif répétitif et le degré d’incertitude local. La partie supérieure présente une vue d’ensemble d’une section du modèle, et la partie inférieure une vue de détail. La couleur représente une estimation plausible de porosité. (a) Association par fréquence de répétition du motif. (b) Association par épaisseur du trait du motif. (c) Association par intensité du motif. (d) Combinaison des associations par fréquence de répétition, épaisseur du trait et intensité du motif. Le motif de base représente une croix ; il est d’autant moins visible et/ou fréquent que l’incertitude est élevée.

Cependant, les techniques présentées à ce jour nécessitent l’élaboration d’un jeu de motifs évoluant de manière discrète entre deux extrêmes (Interrante, 2000; Djurcilov *et al.*, 2001; Rhodes *et al.*, 2003) ; cette tâche repose autant sur des compétences techniques qu’artistiques, et peut donc s’avérer difficile à réaliser pour un non-spécialiste dans ces deux domaines. Nous étudions différentes méthodes permettant de représenter l’incertitude à l’aide d’un unique motif répétitif, en reliant le degré d’incertitude locale à l’intensité du motif, la fréquence de répétition du motif ou l’épaisseur du trait (Figure 3.9, Viard *et al.*, 2009d). Ces trois méthodes ne sont pas incompatibles et peuvent être associées selon n’importe quelle combinaison.

La méthode liant l’intensité du motif en fonction de l’incertitude locale repose sur la composition des composantes RGB de la couleur du motif C_m avec les composantes RGB de la couleur de fond C_f , codant pour la valeur de la propriété d’intérêt. Pour chaque pixel de l’écran, cette composition est calculée comme :

$$C \leftarrow (1 - \alpha) \cdot C_f + \alpha \cdot C_m \quad \text{avec} \quad \alpha = \alpha_m \cdot f(v) \quad (3.2)$$

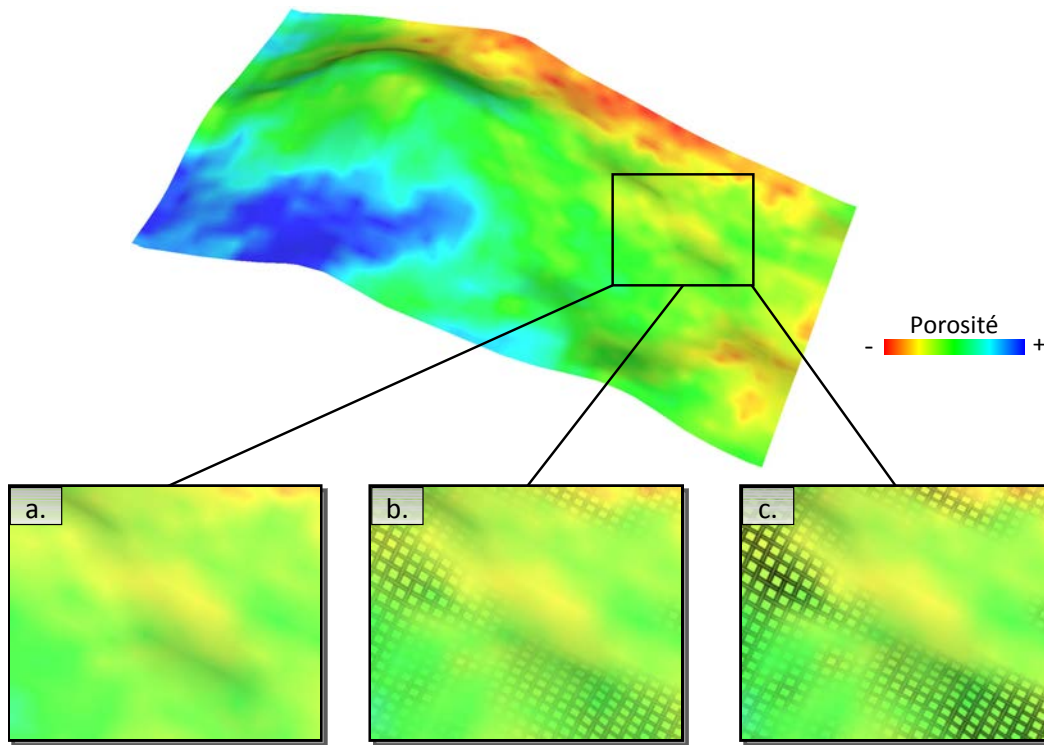


Fig. 3.10 – Contrôle du degré d'interférence entre un motif répétitif et la couleur de fond C_f . La couleur représente une estimation plausible de la porosité le long d'une surface stratigraphique. Partie supérieure : vue d'ensemble de la surface stratigraphique sans motif répétitif. (a) Vue de détail avec $\omega = 0$ (motif non visible). (b) Vue de détail avec $\omega = \frac{1}{2}$. (c) Vue de détail avec $\omega = 1$. Données Total.

où C est la couleur après la composition, α est le coefficient de transparence utilisé pour la composition, α_m est la valeur de transparence locale du motif, et $f()$ est une fonction modulant le degré d'incertitude local v . La fonction $f()$ permet de rendre la texture visible dans les zones d'intérêt, qui peuvent correspondre aux zones de forte ou de faible incertitude.

Cette méthode permet également de contrôler le degré d'interférences entre le motif répétitif et la couleur de fond C_f en modulant le coefficient de transparence α par une valeur constante $\omega \in [0,1]$ définie par l'utilisateur, de telle sorte que $\alpha = \omega \cdot \alpha_m \cdot f(v)$. Plus ω est proche de zéro, moins le motif est visible, même dans les zones d'intérêt (Figure 3.10).

La méthode associant l'épaisseur du trait du motif au degré d'incertitude local nécessite de connaître l'équation analytique décrivant le squelette du motif. Pour tout point de l'écran, il est alors possible de connaître la distance δ entre le centre du pixel et le point le plus proche du squelette. Le pixel est affecté par la texture si δ est inférieur à un seuil τ , défini par $\tau = W \cdot f(v)$, où W est une constante déterminant la largeur maximale du trait (Figure 3.11). La fonction $f()$ permet de d'affiner le trait dans les zones d'intérêt, qui

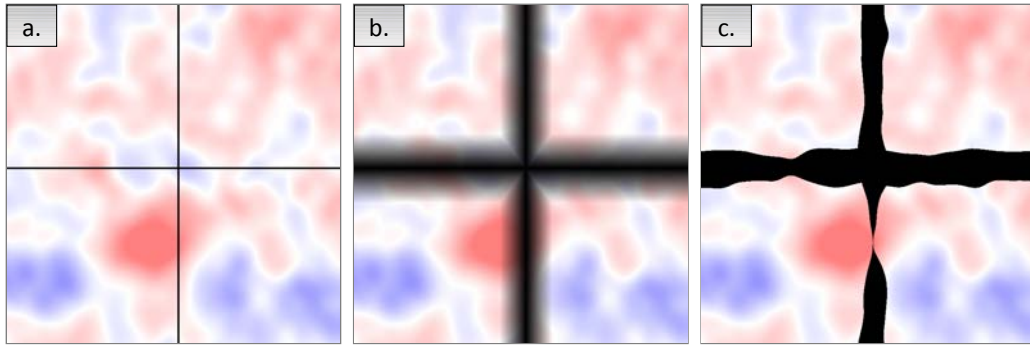


Fig. 3.11 – Modulation de la largeur du trait d'un motif répétitif en fonction du degré d'incertitude local. Le motif unitaire présente une forme de croix, et la couleur de fond représente le degré d'incertitude local. (a) Squelette du motif. (b) Champ des distances δ entre le centre de chaque pixel et le squelette du motif. (c) Motif répétitif érodé par le champ de distance, de telle sorte que la largeur du motif soit proportionnelle à l'incertitude locale.

peuvent présenter une faible ou une forte incertitude locale v selon le choix de la fonction. Cette approche s'inspire des techniques utilisées dans l'algorithme de simulation d'objets ODSim (Object-Distance Simulation), dans lequel les objets sont représentés à l'aide d'un champ de distance tronqué par un champ scalaire de seuils (Henrion *et al.*, 2008, ress).

La méthode modifiant la fréquence de répétition du motif en fonction du degré d'incertitude local procède par distorsion des coordonnées de texture utilisées pour accéder au motif. La distorsion continue de la fréquence du motif produit d'importants artefacts visuels; pour éviter ce problème, nous introduisons des seuils dans la modulation de la fréquence afin de réaliser une distorsion par morceaux (Figure 3.12). Soient f_i la fréquence initiale de répétition de la texture et f la fréquence de répétition modulée par le degré d'incertitude. La relation liant f à f_i est alors :

$$f \leftarrow f_i \cdot \frac{\text{floor}(N_s \cdot f(v))}{N_s} \quad (3.3)$$

où $N_s \in \mathbb{N}$ est le nombre de seuils de modulation de la fréquence initiale f_i . A partir de la fréquence de répétition f , les coordonnées de texture $(u', v')^T$ utilisées pour accéder au motif peuvent être calculées à partir des coordonnées de texture initiales $(u, v)^T \in [0, 1]^2$ comme :

$$(u', v')^T \leftarrow \frac{\text{modulus}((u, v)^T, f)}{f} \quad (3.4)$$

De manière générale, la méthode liant l'incertitude à la fréquence de répétition du motif tend à être moins lisible que les méthodes par épaisseur du trait et par intensité du motif, car les zones de fréquences différentes peuvent avoir une dimension inférieure à celle du motif unitaire, ce qui perturbe fortement la lecture de l'image. En outre, des discontinuités apparaissent entre les zones de fréquences constantes. La méthode par épaisseur du

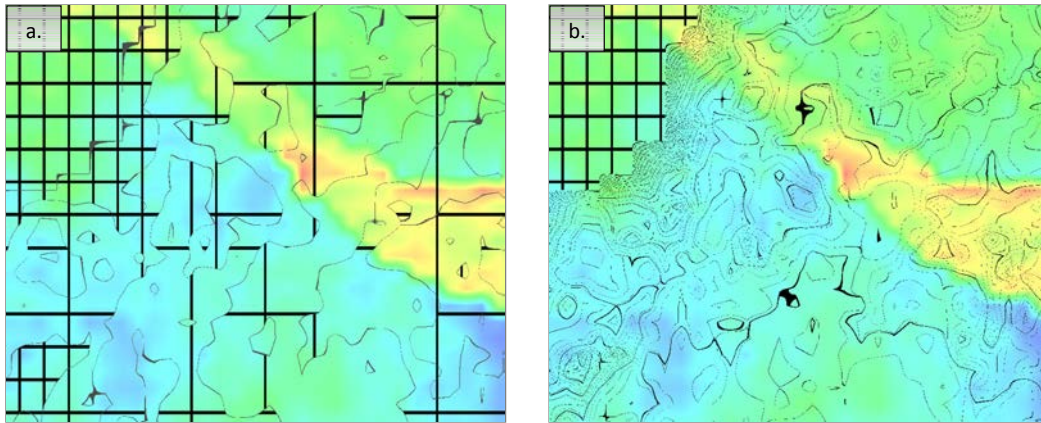


Fig. 3.12 – Comparaison de la modulation de fréquence de répétition d’un motif (a) de manière discrète et (b) de manière continue. Les frontières de seuil sont indiquées par un trait gris pour la modulation discrète (cinq seuils sont utilisés). La couleur de fond représente une estimation plausible de la porosité.

trait est souvent plus aisée à interpréter que celle par intensité du motif, car la perception de la taille évolue linéairement avec la taille de l’objet, tandis que la perception de la transparence dépend du coefficient de transparence utilisé, de la luminance moyenne des objets environnants, et du contraste entre la couleur de la texture et la couleur de fond (Singh et Anderson, 2006; Albert, 2006; Masin, 2006). Cependant, la modulation de l’épaisseur du trait nécessite d’avoir accès au squelette du motif et restreint donc l’ensemble des motifs répétitifs utilisables.

La notion de motif répétitif peut être étendue en trois dimensions au sein d’un moteur de rendu volumique (Figure 3.13), bien que l’ensemble des motifs répétitifs analytiques utilisables pour relier l’incertitude à l’épaisseur du trait soit encore plus restreint qu’en deux dimensions. Dans un moteur de rendu volumique par tranchage incrémental, des artefacts peuvent également apparaître si la distance entre deux tranches successives n’est pas significativement supérieure à l’épaisseur du trait d’un motif unitaire. Cependant, l’extension des motifs répétitifs en trois dimensions donne des résultats généralement satisfaisants en termes de perception des incertitudes locales.

Isovaleurs Les techniques de visualisation d’isovaleurs du degré d’incertitude reposent sur l’extraction de courbes de niveau d’incertitude pour des données surfaciques (2D ou 2,5D) ou d’isosurfaces d’incertitude pour des données volumiques, définissant un contour le long duquel le degré local d’incertitude est constant (section 2.1). Ces isocontours peuvent donc être utilisés pour discriminer les positions de l’espace dont le degré d’incertitude est plus élevé ou plus faible qu’une valeur seuil ; pour percevoir l’évolution de plusieurs valeurs du degré d’incertitude dans l’espace, il est possible de modifier cette valeur seuil dans le

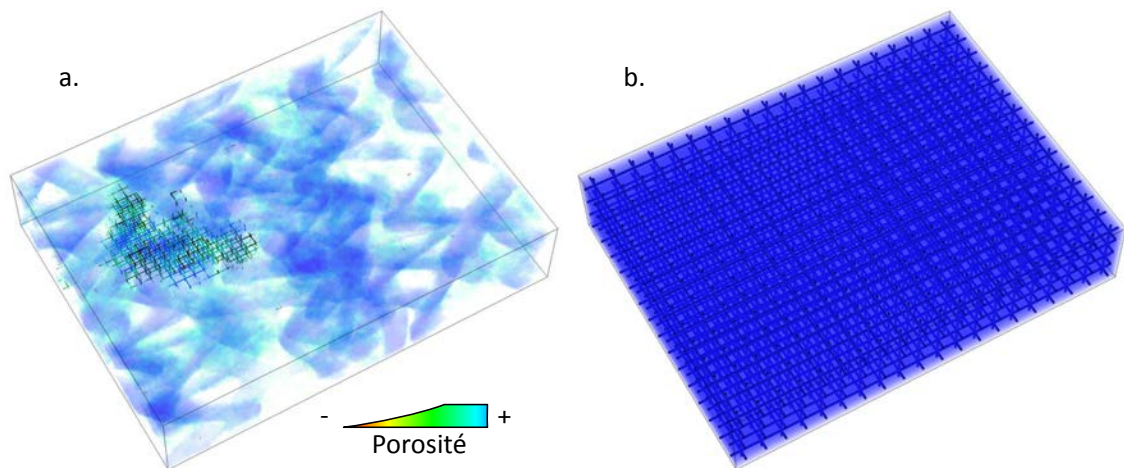


Fig. 3.13 – Visualisation des incertitudes par intensité d’un motif répétitif en rendu volumique. (a) Rendu volumique de chenaux avec visualisation des incertitudes par intensité d’un motif répétitif. Le motif est visible au niveau des zones fiables. (b) Motif répétitif appliqué de manière constante au domaine d’intérêt.

temps (Buatois et Caumon, 2005; Buatois, 2008) ou d’utiliser un ensemble discret de seuils dont les isocontours respectifs peuvent être distingués les uns des autres (par exemple par un code couleur ou un type de figurés, Figure 3.14).

Sur des données volumiques, l’affichage simultané de plusieurs isosurfaces produit fréquemment des occlusions mutuelles qui limitent la perception du modèle, à la fois en ce qui concerne les données d’intérêt et les incertitudes qui leur sont associées. Il est possible d’accroître la transparence des isosurfaces d’incertitude pour limiter ce problème; cependant, pour que les occlusions ne soient pas préjudiciables à la perception, seul un nombre d’isosurfaces très restreint peut en pratique être affiché de la sorte (deux à trois isosurfaces, Figure 3.15).

Résolution Les méthodes par résolution variable s’inspirent de la variable visuelle focus suggérée par MacEachren (1992), en faisant varier la dimension ou la densité d’éléments en fonction du degré local d’incertitude. Nous avons développé deux approches différentes appartenant à cette famille, par résolution du maillage et par dimensions fractales de lignes surimposées.

Les approches par résolution du maillage s’inspirent des techniques de maillages adaptatifs (Berger et Olinger, 1984; Garcia *et al.*, 1999; Plewa *et al.*, 2005) : la densité des cellules est calculée localement comme une fonction du degré d’incertitude. En pratique, les maillages adaptatifs induisent une faible résolution d’échantillonnage des propriétés d’intérêt dans les zones où le maillage est grossier. Cet effet pouvant être indésirable dans certains problèmes en géosciences, nous conservons les propriétés d’intérêt et leurs in-

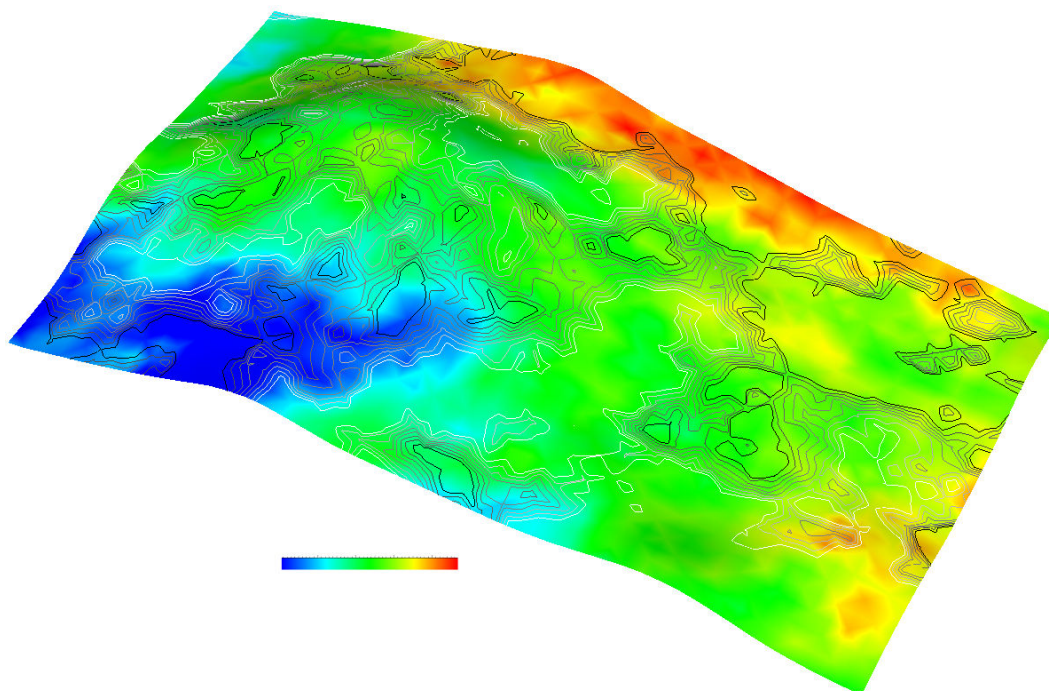


Fig. 3.14 – Visualisation du degré d'incertitude local par isolignes multiples. La couleur représente une réalisation plausible de porosité, affichée le long d'une surface stratigraphique. Le code couleur utilisé pour les isolignes associe une incertitude faible aux couleurs les plus sombres et une incertitude élevée aux couleurs les plus claires.

certitudes associées sur un maillage à haute résolution ; cependant, nous superposons à l'affichage un réseau de mailles adaptatives présentées sous forme de lignes.

Nous proposons deux types de maillages adaptatifs différents, par subdivision récursive d'un maillage régulier et par génération d'un maillage de Voronoi (Aurenhammer et Klein, 2000) dont la densité d'échantillonnage des sites est fonction du degré d'incertitude local (Figure 3.18, images c et d). Ces deux méthodes produisent un maillage fin dans les zones de faible incertitude, et un maillage grossier dans les zones fortement incertaines. La résolution maximale des cellules peut être contrôlée pour chacune des méthodes par la profondeur de récursivité pour le maillage régulier et par le nombre total de sites pour le maillage de Voronoi.

Les approches par dimension fractale relient l'incertitude locale à la qualité de l'approximation d'entités fractales, c'est-à-dire au nombre d'itérations exécutées. Nous avons expérimenté deux types d'entités fractales : une fractale déterministe, la courbe de Koch (Von Koch, 1906; Mandelbrot, 1982; Prusinkiewicz et Sandness, 1988), et une fractale probabiliste produisant des figures dendritiques (Figure 3.18, images e et f). La courbe fractale de Koch est une ligne de longueur infinie, qui est généralement approchée par itérations de raffinement successives d'une géométrie initiale (Figure 3.16). Les figures dendritiques sont générées par nucléation à partir d'un point aléatoire de l'espace, chaque

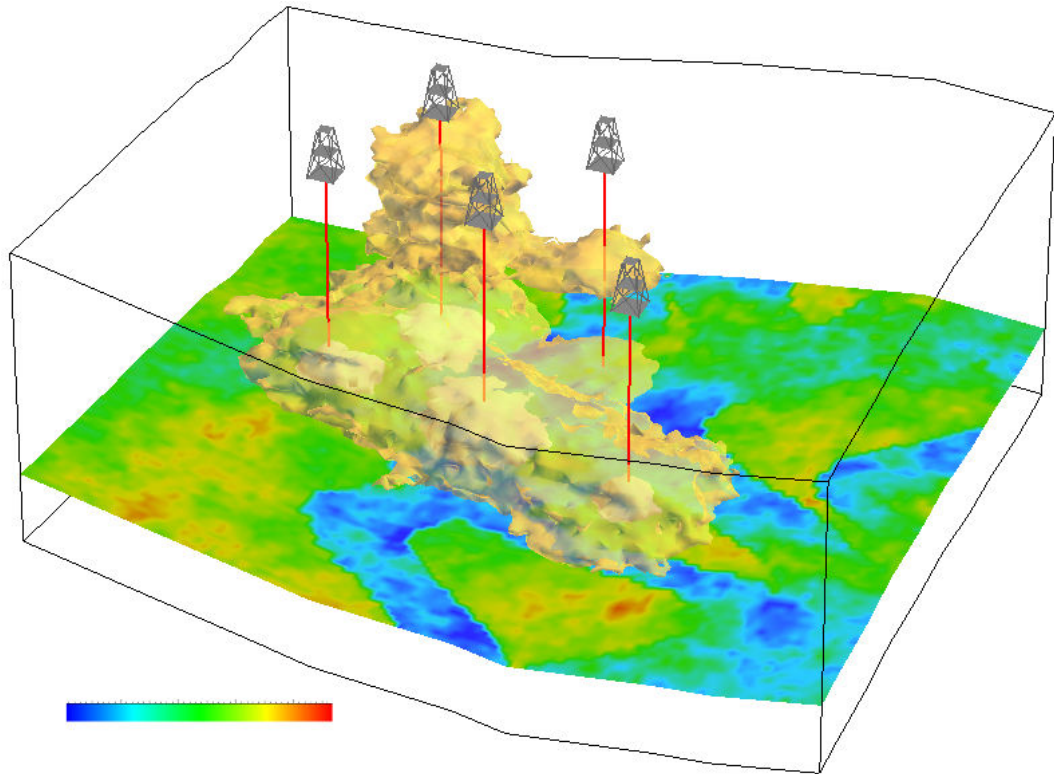


Fig. 3.15 – Visualisation du degré d'incertitude local par isosurfaces multiples (deux isosurfaces). Les données volumiques de porosité et leurs incertitudes associées sont stockées dans une grille curvilinéaire de $80 \times 100 \times 40$ cellules. Seul un sous-ensemble des données est affiché à l'écran le long d'une surface stratigraphique, représentant une réalisation plausible de porosité. L'isosurface grise présente une valeur d'incertitude faible, et l'isosurface jaune une valeur d'incertitude moyenne.

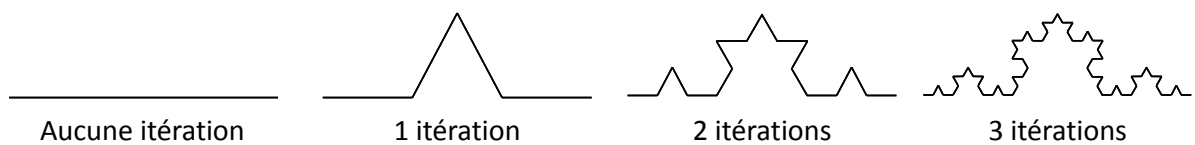


Fig. 3.16 – Approximations de la courbe fractale de Koch à différentes itérations.

itération faisant apparaître une fourche au sein de la dendrite (Figure 3.17). Dans les deux cas, les approximations des fractales sont générées sous forme de lignes superposées à des sections de grilles régulières ; la complexité maximale de ces lignes peut être contrôlée par la profondeur de récursivité.

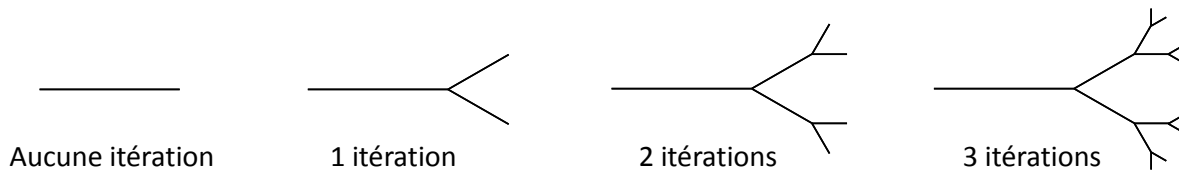


Fig. 3.17 – Nucléation et développement par itérations successives d'une fractale dendritique.

3.1.4 Ambiguation

Les techniques par ambiguation visent à diminuer la facilité selon laquelle un observateur peut percevoir les données initiales. De par leur nature même, ces approches provoquent généralement des interférences dans les zones les plus incertaines du modèle. Elles ne sont donc pas adaptées à des problèmes nécessitant une étude attentive des informations dans les zones les moins connues du modèle, par exemple lors de la mise en place de stratégies visant à acquérir plus de données. Leur action peut cependant aisément être associée à la notion d'incertitude, résultant en une visualisation intuitive du degré d'incertitude local.

Flou Dans notre travail, la visualisation des incertitudes par flou est basée sur la composition d'une image nette et d'une image floue, où le coefficient utilisé pour la composition varie spatialement en fonction du degré local d'incertitude (Figure 3.19).¹⁰ Cette technique permet d'accroître l'intensité du flou perçu dans les zones incertaines, sans pour autant affecter les zones fiables.

La génération d'une image floue repose sur la convolution des couleurs des pixels avoisinants selon un noyau de convolution prédéfini. Le noyau de convolution permet de pondérer la contribution des points avoisinants, généralement afin de privilégier les points les plus proches. A partir d'une image initiale, ce noyau permet alors de générer une image floue selon l'équation analytique suivante :

$$c(x, y) \leftarrow \iint_{-\infty}^{+\infty} f(\delta^x, \delta^y) \cdot c(x + \delta^x, y + \delta^y) \cdot d\delta^x d\delta^y \quad (3.5)$$

où $c(x, y)$ est la couleur du point de coordonnées (x, y) , $f()$ est la valeur du noyau de convolution, et (δ^x, δ^y) est le décalage de coordonnées le long des deux axes du plan visuel. En pratique, les images sont des ensembles discrets de pixels ; l'intégrale de convolution de l'équation 3.5 est donc approchée sous la forme de la somme pondérée des couleurs des pixels voisins.

De nombreux noyaux de convolution peuvent être utilisés pour traiter une image. Dans ce travail, nous avons adopté un filtre de convolution Gaussien pour générer des

¹⁰Le flou peut être appliqué directement dans l'espace de l'image, ou dans l'espace de l'objet sous la forme de textures qui sont ensuite plaquées sur des surfaces paramétrisées.

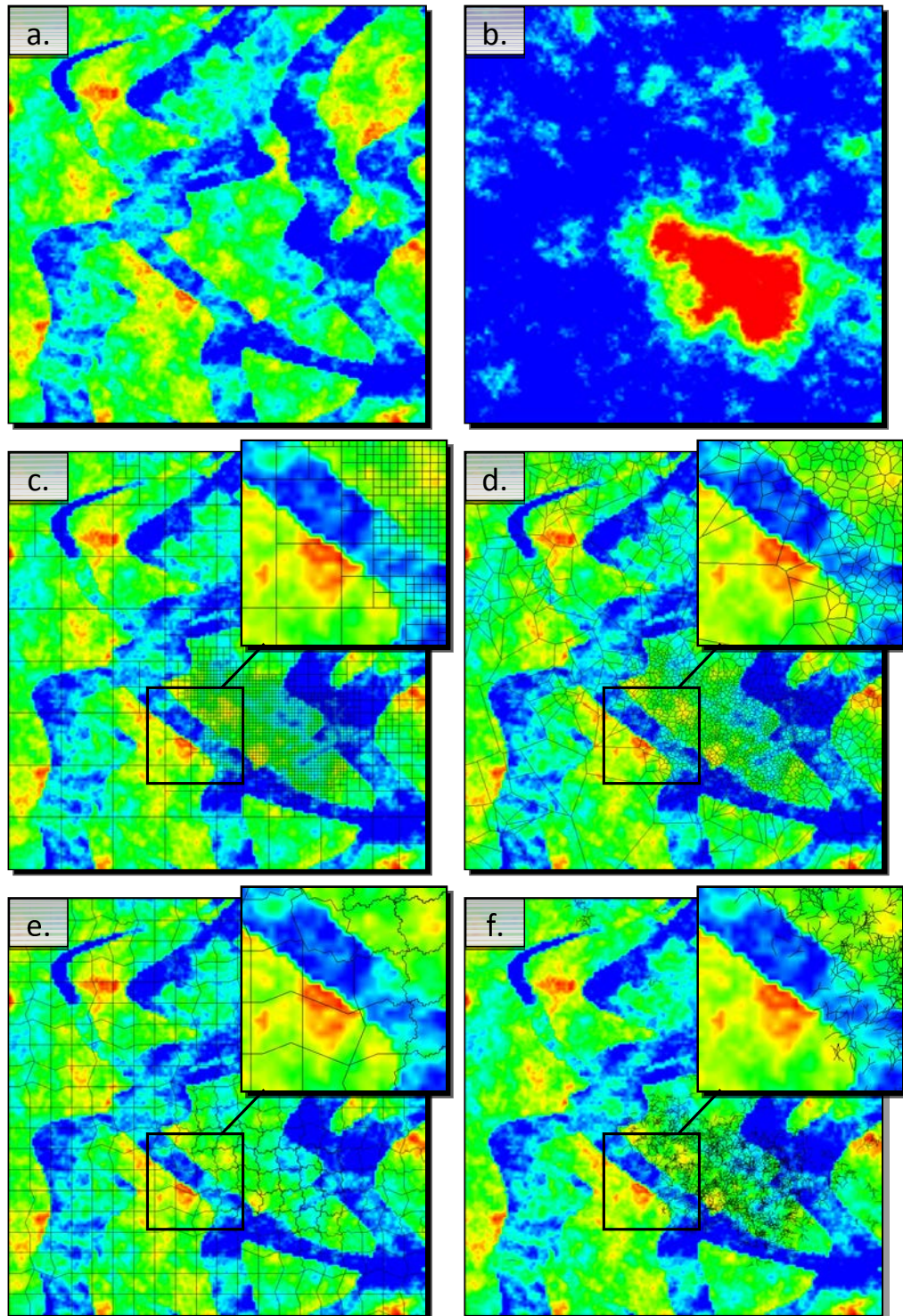


Fig. 3.18 – Méthodes de visualisation des incertitudes par résolution variable. (a) Carte de porosité. (b) Incertitudes associées. (c) Visualisation par résolution d'un maillage régulier. (d) Visualisation par résolution d'un maillage de Voronoi. (e) Visualisation par distorsion du maillage à l'aide de la fractale de Koch. (f) Visualisation par taux de développement d'une population de dendrites. Les images (c), (d), (e) et (f) présentent des zooms dans leur partie supérieure droite.

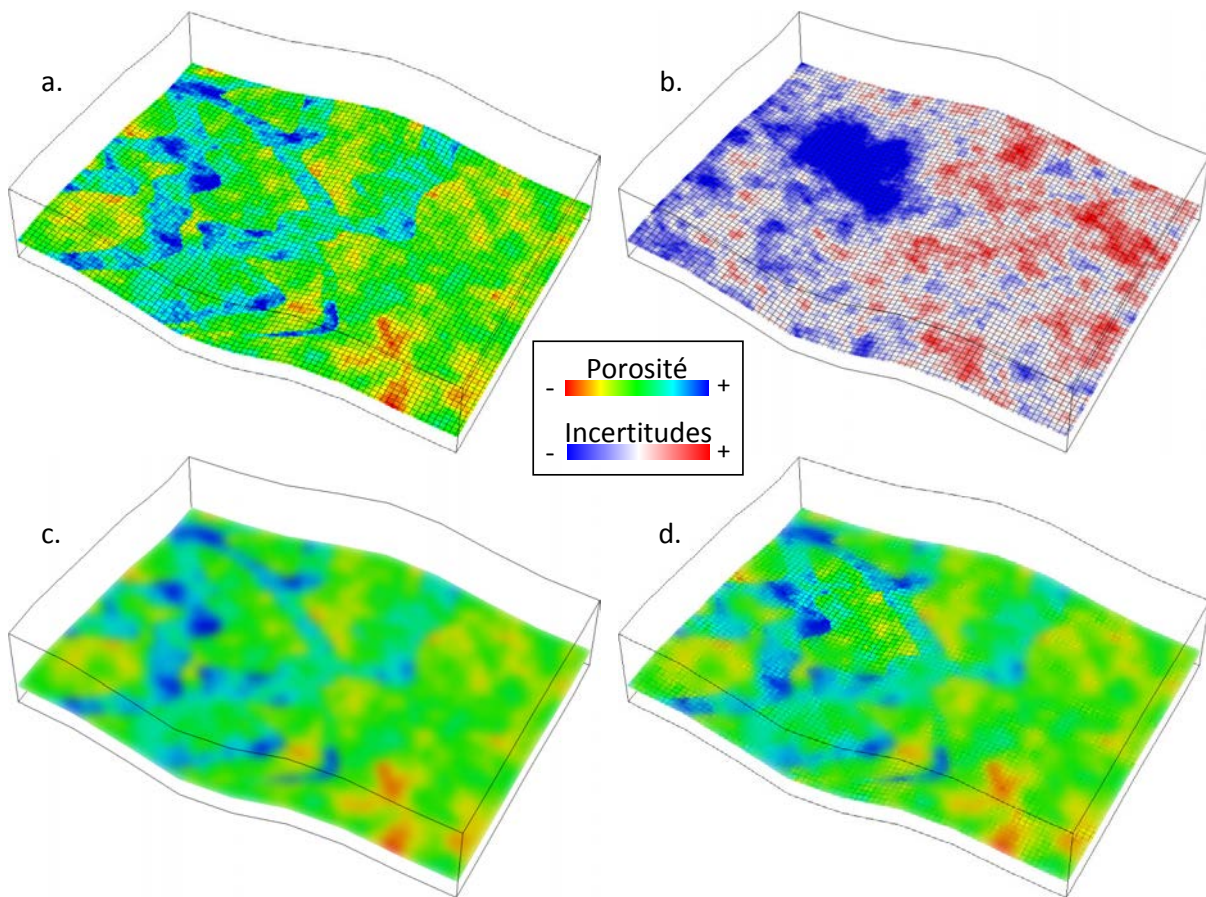


Fig. 3.19 – Visualisation des incertitudes par intensité de flou. (a) Image nette représentant une réalisation plausible de porosité. (b) Incertitudes associées. (c) Image floue générée à partir de l'image (a), en utilisant un filtre Gaussien d'un rayon de dix pixels. (d) Composition des images (a) et (c) en fonction du degré local d'incertitude.

images floues. Outre la qualité de rendu, ce filtre présente plusieurs optimisations possibles (Waltz et Miller, 1998) : (i) il est linéairement séparable, c'est-à-dire que l'application d'un filtre Gaussien à deux dimensions peut être remplacée par celle de deux filtres Gaussiens à une dimension selon deux axes perpendiculaires de l'image,¹¹ et (ii) l'application d'un noyau Gaussien de grande taille peut être remplacée par l'application successive de noyaux Gaussiens plus petits. Par ailleurs, la convolution des couleurs en un point (x, y) est indépendante des convolutions réalisées en tout autre point de l'espace, ce qui facilite la mise en place d'une implémentation parallèle du filtre Gaussien, notamment sur carte graphique (Jargstorff, 2004; James et O'Rorke, 2004).

¹¹Ce procédé permet de substituer une opération de complexité algorithmique $o(n^2)$ par deux opérations de complexité algorithmique $o(n)$, où n est la taille du noyau Gaussien.

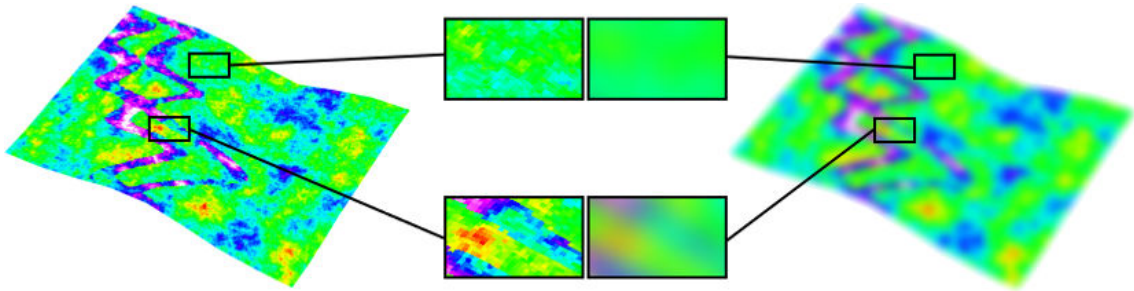


Fig. 3.20 – Visualisation d'une carte de porosité nette (gauche) et avec flou (droite). Les zooms permettent de comparer l'effet du flou sur deux zones ayant des fréquences spatiales différentes. Le flou n'a que peu d'effet sur la zone présentant une fréquence spatiale faible (zoom supérieur), mais est bien plus perceptible dans la zone à haute fréquence spatiale (zoom inférieur).

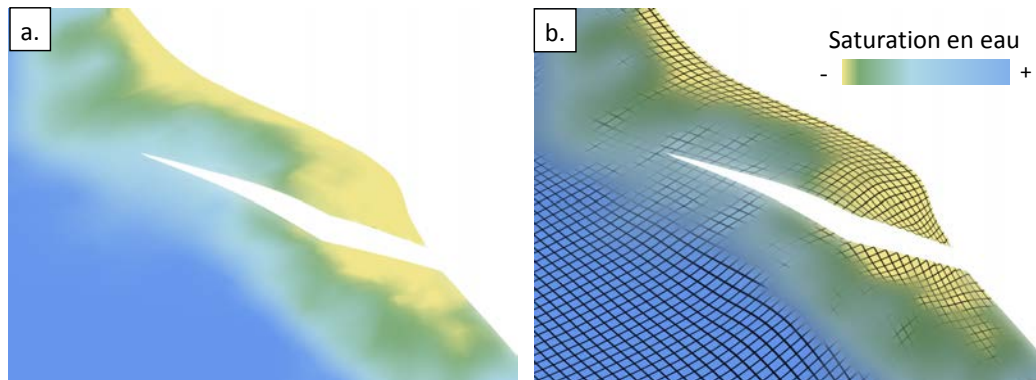


Fig. 3.21 – Visualisation des incertitudes par flou avec et sans motif répétitif. (a) Incertitudes représentées par l'intensité du flou sans motif répétitif. (b) Incertitudes représentées par l'intensité du flou avec motif répétitif. La couleur représente la saturation en eau la plus probable, et le motif répétitif reproduit une grille régulière. Données TNO, modèle fourni par Paradigm.

Le filtre Gaussien provoque une diminution du niveau de détail présent au sein de l'image, de manière similaire à un filtre passe-bas (réduction des variations spatiales à haute fréquence). De ce fait, si l'image initiale présente de faibles variations spatiales, l'action du flou est généralement difficilement discernable dans l'image traitée (Figure 3.20 Viard *et al.*, 2009d). Ce phénomène peut produire des ambiguïtés en termes de perception, l'observateur étant incapable de savoir si la zone est incertaine ou non, qui peuvent être nuisibles à la compréhension du modèle.

Pour pallier à ce défaut, nous proposons de superposer un motif répétitif à l'image avant d'appliquer le filtre Gaussien : si la couleur du motif répétitif contraste avec les couleurs codant pour la propriété d'intérêt, cette superposition augmente la variabilité spatiale de l'image et améliore donc la perception du flou (Figure 3.21).

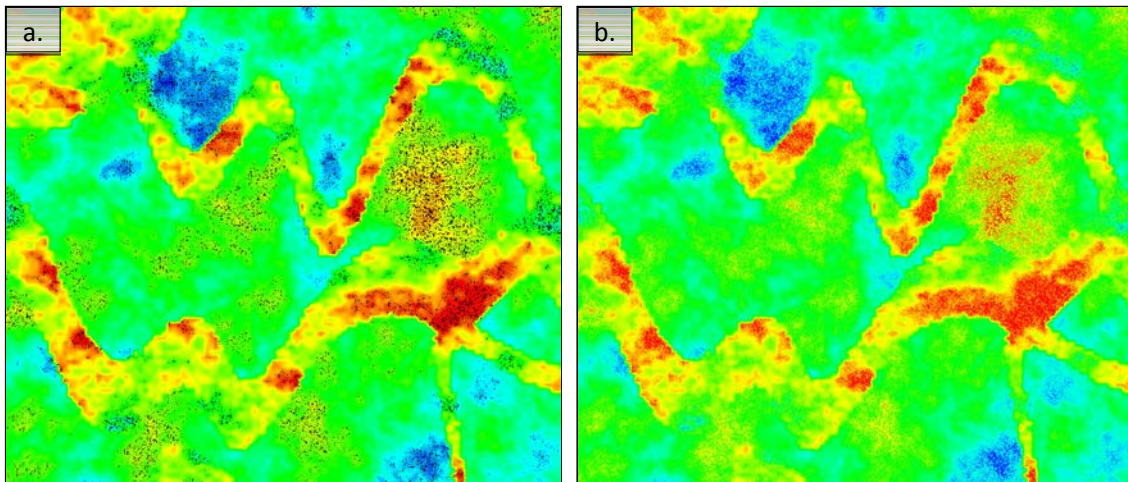


Fig. 3.22 – Visualisation des incertitudes par ajout de bruit. (a) Bruit ajouté dans l'espace des couleurs, sous la forme de taches noires. (b) Bruit ajouté dans l'espace des valeurs. Des bruits de même intensité ne sont pas perçus de la même manière en différents endroits, du fait du caractère non linéaire de la carte de couleurs utilisée. Les couleurs représentent une réalisation plausible de porosité.

Bruit Les méthodes par ajout de bruit consistent à ajouter à la propriété d'intérêt un champ scalaire généré aléatoirement présentant une forte variabilité spatiale, et dont l'intensité est fonction du degré local d'incertitude. L'ajout de bruit peut se faire dans l'espace des couleurs ou dans l'espace des valeurs (Figure 3.22).

Si le bruit est ajouté aux valeurs de la propriété d'intérêt, la perception des zones incertaines peut être biaisée par le code couleur utilisé. En effet, une carte de couleurs à basse fréquence a un effet lissant sur les variations à haute-fréquence associées au bruit et interfère donc avec la perception de l'incertitude locale. Par ailleurs, les variations haute-fréquence du bruit dans l'espace des valeurs peuvent être confondues avec les variations de la propriété d'intérêt elle-même. Il est donc généralement préférable d'utiliser le bruit dans l'espace des couleurs afin d'éviter toute ambiguïté entre propriété d'intérêt et incertitudes locales.

Nous avons également étudié une implémentation du bruit où le champ scalaire aléatoire est animé dans le temps. Nous proposons deux méthodes d'animation. La première est basée sur une animation linéaire du bruit, représenté sous la forme de champs scalaires à haute fréquence utilisés pour moduler la propriété d'intérêt ; cependant, la variabilité spatiale des champs scalaires intermédiaires est typiquement plus faible que celle des champs scalaires initiaux, provoquant des artefacts visuels au cours du temps. La seconde méthode repose sur un générateur congruentiel linéaire de nombres aléatoires pour générer analytiquement le bruit ; les paramètres du générateur congruentiel linéaire varient dans le temps pour animer le bruit. Ces deux méthodes provoquent un déplacement dans les

zones de forte incertitude, de manière similaire aux signaux parasites pouvant être reçus par un téléviseur, qui attire davantage l'attention de l'observateur en ces points.¹²

Distorsion Les méthodes par distorsion agissent sur la paramétrisation entre l'espace physique et l'espace des valeurs (Heckbert, 1986; Haker *et al.*, 2000). Cette paramétrisation est déformée le long d'un champ de déplacement prédéfini, selon une intensité de déplacement qui est fonction du degré local d'incertitude. De la sorte, les zones fortement déformées présentent une incertitude élevée, tandis que les zones fiables ne sont pas affectées.

Nous proposons quatre champs de déformation selon des motifs de damier, de vagues, de cercles concentriques et de spirale (Figure 3.23). Dans chacun de ces cas, les déformations sont évaluées de manière procédurale d'après une équation analytique. Il est possible de modifier l'amplitude et la fréquence de répétition de ces motifs à l'aide de paramètres fixés par l'utilisateur.

Les techniques par distorsion déplacent des ensembles de valeurs en des positions qui diffèrent de leur localisation réelle. De ce fait, ces techniques doivent être utilisées avec prudence : si un déplacement modéré des valeurs peut être acceptable dans certains cas, un déplacement significatif peut nuire à la compréhension des structures de la propriété d'intérêt, voire biaiser leur interprétation. Ces techniques sont donc généralement peu recommandées pour des problématiques géologiques.

Pixellisation Les méthodes par pixellisation sont fréquemment utilisées dans le domaine audiovisuel, afin de masquer des visages, des logos publicitaires ou des informations privées qui font partie intégrante de l'image, mais ne doivent pas être accessibles aux téléspectateurs (Figure 3.24). De manière similaire, nous proposons d'utiliser les méthodes de pixellisation pour masquer les zones incertaines de l'image.

La pixellisation fonctionne par combinaison des couleurs d'une image, généralement regroupées sous forme de carrés dont la couleur finale est la moyenne des couleurs que le carré recouvre. Traditionnellement, le caractère "pixellisé" d'une zone de l'image est donc booléen (la zone est pixellisée, ou elle ne l'est pas). De ce fait, une implémentation naïve de la pixellisation ne permettrait que la sélection d'un unique seuil d'incertitude, toutes les informations d'incertitude intermédiaires étant perdues. Pour éviter ce phénomène, nous proposons une extension de la pixellisation traditionnelle ; dans notre implémentation, une image entièrement pixellisée est composée avec une image non pixellisée selon un coefficient de composition qui est fonction du degré d'incertitude local. Cette approche

¹²Une vidéo illustrant le bruit animé dans l'espace des valeurs peut être accédée à l'adresse http://www.gocad.org/www/research/index_gallery.php?image=animated_noise.gif. Cette animation utilise une définition analytique du bruit à l'aide d'un générateur de nombres pseudo-aléatoires congruentiel linéaire.

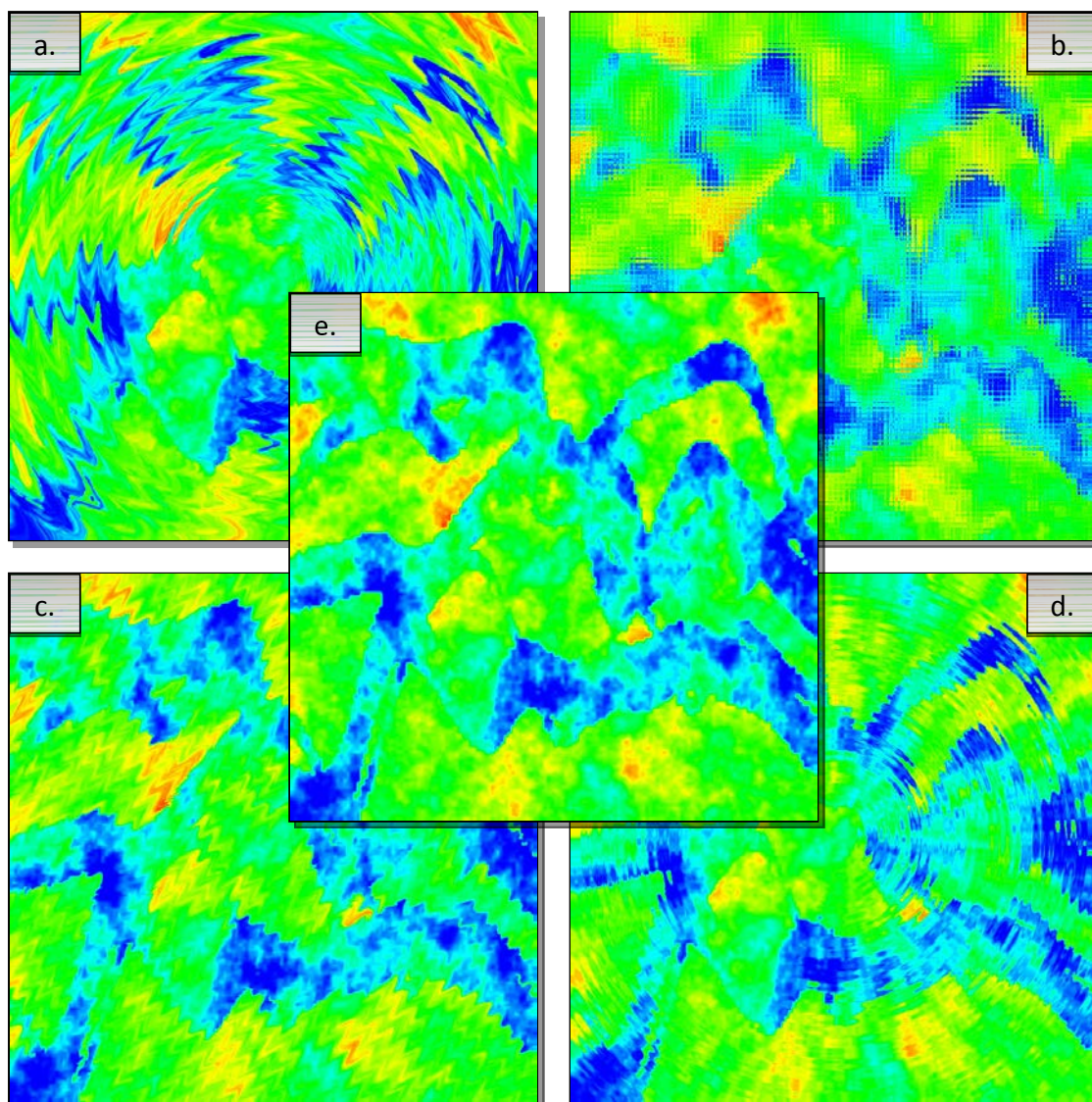


Fig. 3.23 – Visualisation des incertitudes par distorsion. (a) Distorsion en spirale. (b) Distorsion en damier. (c) Distorsion en vagues. (d) Distorsion en cercles concentriques. (e) Carte de porosité non distordue.

permet de percevoir une évolution plus continue des degrés locaux d'incertitude, tout en restituant une impression générale de pixellisation dans les zones incertaines (Figure 3.25).

3.1.5 Modification de la géométrie

Les techniques par modification de la géométrie transforment les cellules du maillage représentant le modèle afin de refléter le degré d'incertitude local. Ces approches sont généralement peu adaptées à des problématiques géologiques, car les déformations qui leur



Fig. 3.24 – Pixellisation d’une image afin de garantir l’anonymat de la personne photographiée. (a) Image initiale. (b) Image pixelisée.

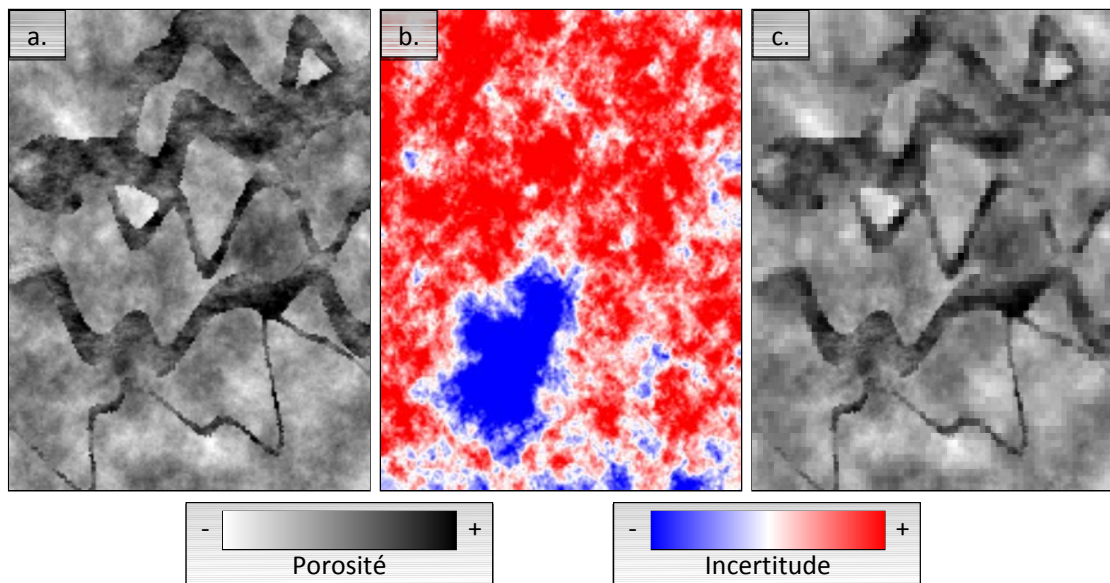


Fig. 3.25 – Pixellisation d’une carte de porosité en fonction du degré d’incertitude local. (a) Image initiale. (b) Incertitudes associées. (c) Image pixelisée.

sont associées peuvent obscurcir la perception des structures géologiques. Elles peuvent cependant s'avérer utiles avec certains types de grilles, notamment les maillages Cartésiens pour lesquels la cellule de référence avant transformation est connue de l'observateur.

Anamorphose L'anamorphose est un procédé de déformation de l'image, visant à accroître ou décroître les dimensions de certains éléments en fonction d'un critère objectif. Par exemple, sur une carte du monde visant à représenter le produit intérieur brut de chaque pays, les états de l'union européenne apparaîtraient sur une surface de la carte bien plus grande que les états du continent africain, bien que leur superficie réelle soit nettement inférieure. Dans le cadre de cette thèse, nous avons adopté le degré d'incertitude local comme critère objectif de déformation. Cette technique fait apparaître les zones incertaines avec de plus petites dimensions que les zones fiables.

L'anamorphose est *sensu-stricto* une déformation continue de l'espace et se rapproche donc des méthodes de distorsion (Section 3.1.4). Cependant, les limites des corps géologiques sont rarement discernables avec précision ; une implémentation continue de l'anamorphose produirait donc une visualisation de faible qualité dans laquelle la déformation locale ne serait que peu perceptible. Pour éviter cette limitation, nous proposons une anamorphose par morceaux où la déformation s'applique à chacune des cellules du modèle plutôt qu'au modèle dans sa totalité (Figure 3.26). Cette approche fournit à l'observateur une référence – la dimension caractéristique d'une cellule – qui améliore la perception du degré de déformation local. Cependant, les cellules peuvent être de dimensions très faibles dans les zones incertaines, jusqu'à être réduites à un point dans les zones les moins connues. Ce phénomène rend ces cellules ininterprétables en pratique. Il est donc recommandé de toujours conserver un seuil minimal pour la dimension des cellules.

Barres d'erreur La visualisation des incertitudes par barres d'erreur est une extension en deux dimensions de la technique fréquemment utilisée sur des graphes pour indiquer des intervalles de confiance (Olston et Mackinlay, 2002). Elle consiste à relier l'épaisseur de chaque cellule au degré d'incertitude local, les cellules les plus épaisses présentant une incertitude élevée (Figure 3.27).

Bien que cette technique ait démontré son efficacité en une dimension, son extension en deux dimensions se révèle de bien moindre qualité en termes de perception (Sanyal *et al.*, 2009), probablement du fait des occlusions qui se produisent entre les barres d'erreur. De ce fait, les barres d'erreur permettent d'évaluer aisément la différence d'incertitude locale entre deux cellules voisines, mais sont bien plus délicates à interpréter en termes de valeur absolue de l'incertitude. Deux stratégies peuvent être adoptées pour limiter ce phénomène : (i) n'inclure des barres d'erreur que sur un échantillonnage restreint des données, ou (ii) limiter la largeur des barres d'erreur pour leur faire occuper un espace inférieur à une cellule. Appliquées à des modèles géologiques réels, ces stratégies se révèlent cependant difficilement utilisables dans la majorité des cas.

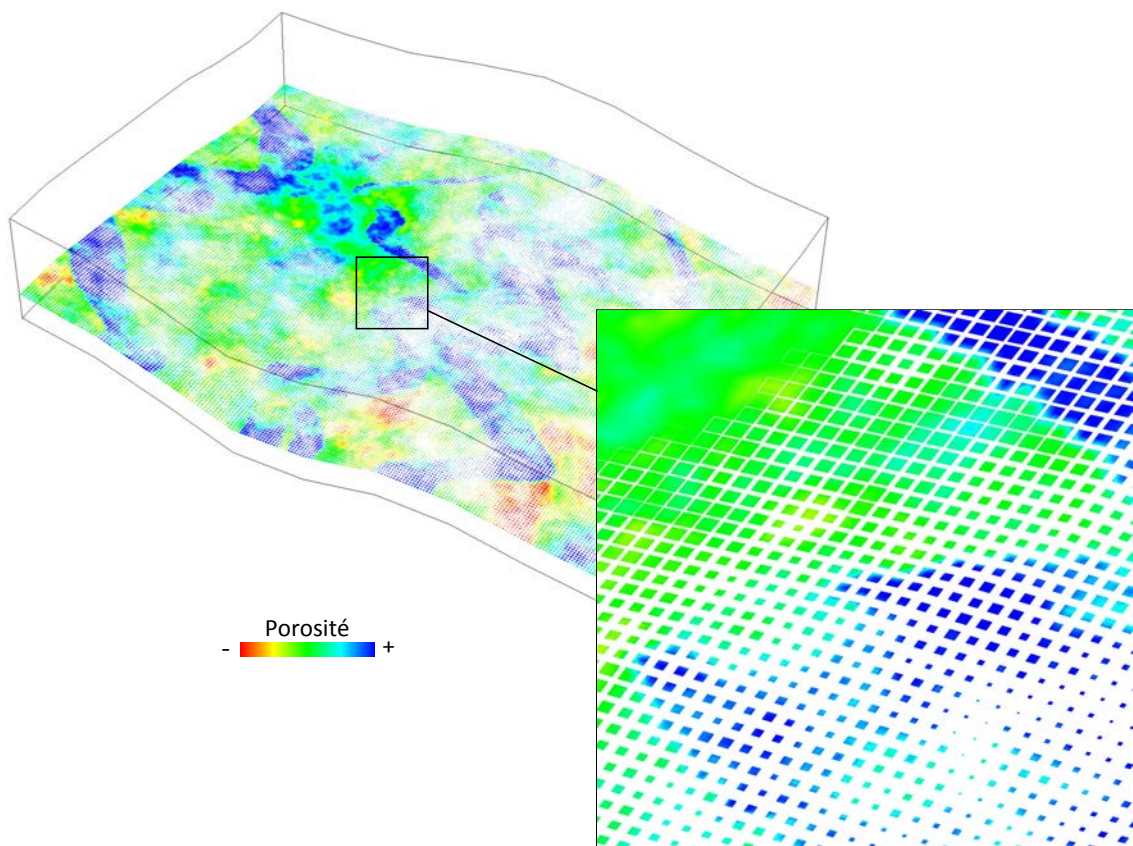


Fig. 3.26 – Visualisation des incertitudes par anamorphose discontinue. Les cellules de petite taille présentent une incertitude élevée, et la couleur représente une réalisation plausible de porosité.

3.2 Validation des méthodes statiques

Les méthodes statiques couvrent un large spectre de techniques et de variables visuelles associées – de ce fait, il est délicat de toutes les valider de manière empirique. Dans nos travaux, nous nous sommes donc concentrés sur une méthode particulière : la visualisation des incertitudes par motifs répétitifs (Viard *et al.*, 2010b).

3.2.1 Participants

Cette étude sur utilisateur a été réalisée au sein d'un panel de 123 participants, possédant tous des connaissances suffisantes en géosciences pour appréhender les tâches demandées. Au sein de la population des participants, un était professionnel du domaine de l'industrie pétrolière, cinq étaient doctorants au CRPG-CNRS de Nancy, et cent dix-sept étaient étudiants de seconde année à l'Ecole Nationale Supérieure de Géologie de Nancy (niveau Master 1). Le professionnel de l'industrie pétrolière était habitué à la prise de décision

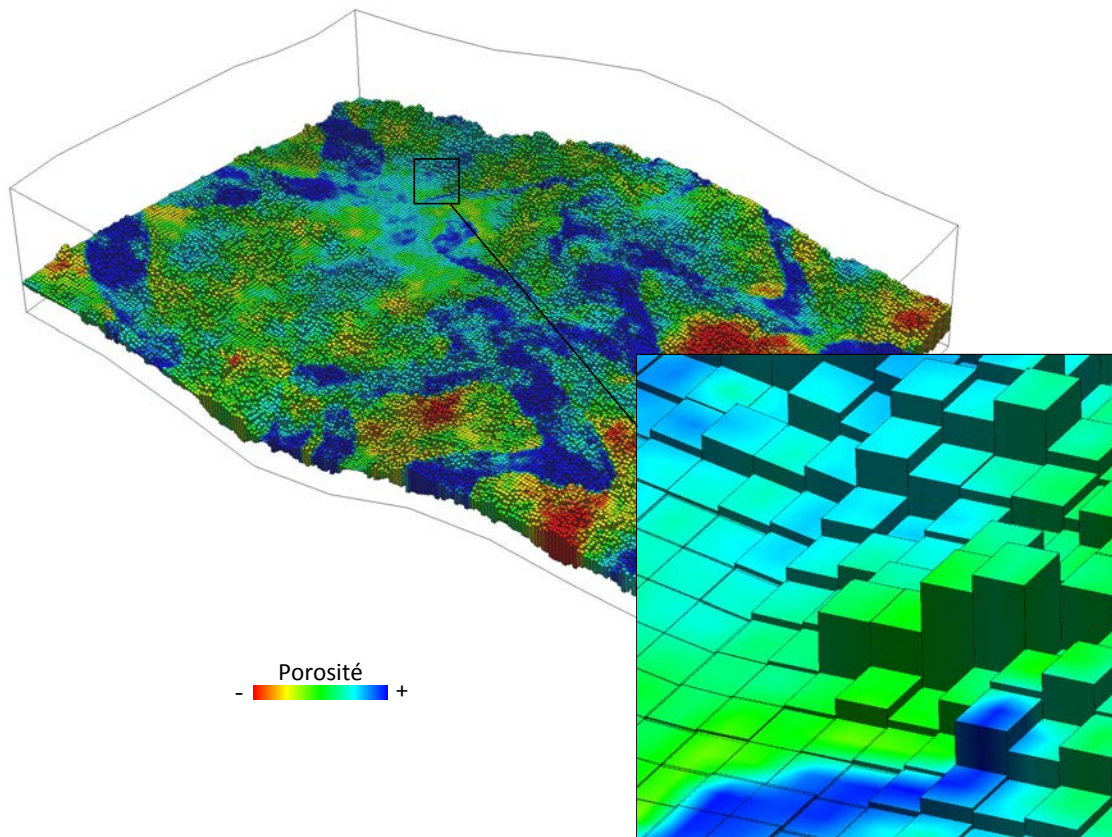


Fig. 3.27 – Visualisation des incertitudes par barres d'erreur. Les cellules les plus épaisses présentent une incertitude élevée, et la couleur représente une réalisation plausible de porosité.

en présence d'incertitudes ; son niveau peut être considéré comme excellent. Les cinq doctorants présentaient des spécialisations variées ; leur niveau varie de correct à bon. Les cent dix-sept étudiants de seconde année recevaient un enseignement dans diverses disciplines en géosciences ; leur niveau varie de faible à correct. Il convient de noter que le choix de ces participants a été motivé avant tout par des raisons logistiques, plutôt que par leur degré d'expertise.

D'après Roth (2009), la complexité de la tâche renforce l'influence du degré d'expertise des participants. Or, notre étude mettait en jeu des tâches complexes – en l'occurrence, le design de puits injecteurs. Il est donc raisonnable de s'attendre à un biais d'échantillonnage des participants dans notre étude, bien que cette hypothèse ne puisse être testée statistiquement du fait du manque d'experts disponibles.

3.2.2 Description des données

Les tâches proposées aux participants reposaient sur un jeu de données de pression, porté par le modèle Cloudspin.¹³ Cloudspin est un réservoir d'hydrocarbures présentant à la fois du pétrole et du gaz, localisés dans une formation supposée en surpression. La pression au puits étant plus faible que la pression du réservoir, les hydrocarbures doivent passivement s'écouler vers le puits producteur. L'orientation exacte et la vitesse de l'écoulement des hydrocarbures en réponse à un plan de production sont cependant incertaines, du fait d'incertitudes sur le comportement fluide des hydrocarbures eux-mêmes (tables Pression-Volume-Température) et d'incertitudes sur la perméabilité des roches du réservoir. Pour échantillonner les incertitudes associées à l'écoulement, un ensemble de 27 simulations d'écoulement ont été générées. A chaque pas de temps des simulations, les incertitudes sur la pression locale sont quantifiées à l'aide du coefficient de variation de l'ensemble des pressions. Les zones de forte incertitude mettent ainsi en valeur les frontières possibles de la zone où la plupart des hydrocarbures ont été extraits.

Pression et design de puits Dans cette étude, il était demandé aux participants de se pencher sur le positionnement de puits injecteurs, dont le but est d'injecter de l'eau dans le milieu poreux afin de déplacer les hydrocarbures vers les puits producteurs. La surpression locale doit cependant être étudiée préalablement à tout nouveau forage, afin que la colonne de boue du forage compense en permanence les différences de pression. Une sous-estimation de la pression pourrait ainsi provoquer une explosion de la tête de puits, mettant en danger des vies humaines et détruisant des équipements coûteux. A l'inverse, une sur-estimation peut conduire à une invasion de boue dans le voisinage du puits, réduisant ainsi la perméabilité locale.

3.2.3 Protocole

Dans cette étude sur des utilisateurs, nous avons séparé les participants en trois groupes suivant un schéma de répartition aléatoire. L'ensemble des participants devaient répondre aux mêmes questions ; cependant, les ressources disponibles pour l'exercice variaient d'un groupe à l'autre (Figure 3.28), comme suggéré par Deitrick et Edsall (2006). Le premier groupe ne disposait que d'informations sur la pression moyenne. Le second groupe disposait d'informations sur la pression moyenne et d'une carte des incertitudes associées, présentée sur une image séparée. Le troisième groupe disposait d'informations sur la pression moyenne et les incertitudes associées, présentées sur une même image à l'aide de techniques de visualisation des incertitudes par motif répétitif. Les images ont été capturées dans les mêmes conditions pour l'ensemble des groupes (point de vue, carte de couleurs, *etc.*). Tous ces documents ont été fournis aux participants sous la forme d'im-

¹³Données fournies par Schlumberger, modèle créé par Paradigm.

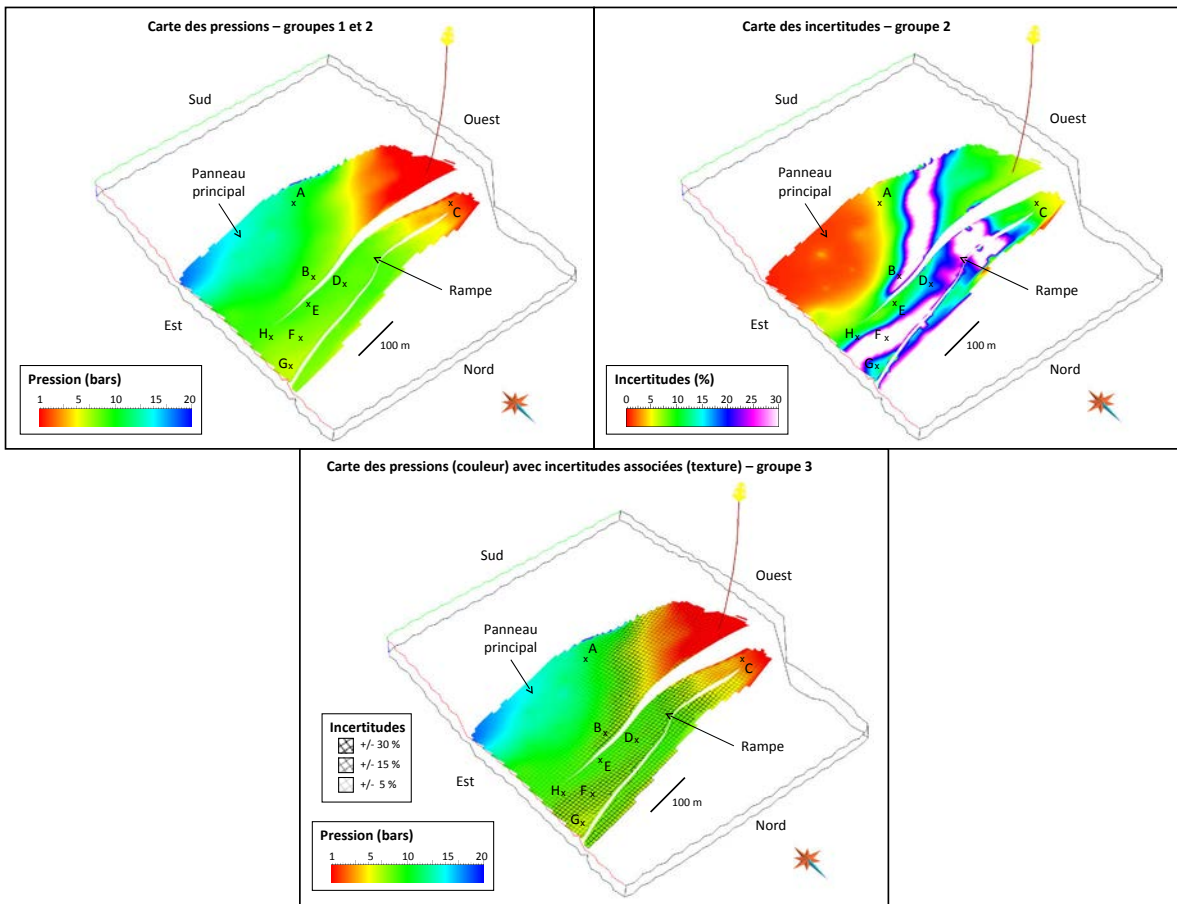


Fig. 3.28 – Ressources disponibles pour remplir les tâches de l'étude sur des utilisateurs. Les images représentent différentes visualisations du modèle Cloudspin.

pressions en couleur ; nous nous sommes assurés que les images étaient imprimées en haute qualité, et que les couleurs étaient rendues de manière satisfaisante par l'imprimante.

Design Les mêmes questions ont été posées à l'ensemble des participants ; de ce fait, nous considérons que toute différence statistique entre les réponses associées à chaque groupe provient des ressources mises à leur disposition. Pour être valable, ce postulat nécessite que (i) les groupes soient composés de participants ayant un niveau de compétence comparable, et (ii) une quantité significative de réponses soit collectée. Afin de répondre à ces pré-requis, nous avons mis en place une question dont le seul but était de comparer les compétences des participants. Par ailleurs, nous nous sommes efforcés de toucher un nombre de participants aussi large que possible et les avons assignés à chaque groupe de manière aléatoire, de manière à créer des groupes de dimensions similaires.

Description des questions Trois questions de complexité croissante ont été posées aux participants. Dans la première question, les participants devaient indiquer dans quelle zone du réservoir la pression était la plus élevée, parmi trois choix possibles - est, ouest ou centre du panneau principal. Le but de cette question était de vérifier les compétences des participants en matière de lecture de carte, afin de s'assurer que les trois groupes présentaient un niveau d'expertise similaire.

Dans la seconde question, deux emplacements possibles (A et B) étaient proposés pour le forage de puits injecteur ; les participants devaient comparer ces emplacements en termes de pression maximale possible. L'emplacement A présentait une pression moyenne plus élevée que l'emplacement B, mais son incertitude associée était plus faible. Cette question a été créée dans le but d'étudier si l'incertitude affectait les décisions dans le cadre de choix binaires, et le cas échéant, de quelle manière.

La troisième question proposait une tâche similaire, c'est-à-dire le tri de cinq emplacements de puits injecteurs possibles (D, E, F, G et H) en termes de pression maximale possible. De nouveau, les objectifs de cette question sont similaires aux objectifs de la question 2 ; cependant, les schémas de pensée mis en jeu pour résoudre cet exercice comportent des choix qualitatifs. En effet, la quantification de la pire pression possible d'un large ensemble de puits est nettement plus délicate à effectuer en un laps de temps restreint.

3.2.4 Analyse des résultats

Cette section présente dans un premier temps les outils statistiques utilisés pour analyser les réponses des participants, avant de commenter les résultats eux-mêmes.

Outils d'analyses statistiques : le test de la somme des rangs de Wilcoxon Le *test de la somme des rangs de Wilcoxon*, également connu sous le nom de *test de Mann-Whitney-Wilcoxon* (Wilcoxon, 1945; Mann et Whitney, 1947), est un test d'hypothèse statistique permettant de comparer deux séries de mesures ω_1 et ω_2 , sans présupposer de leur distribution réelle (test non paramétrique). Le test est basé sur le calcul d'une métrique U , sous la forme :

$$U_i = \sum_{k=0}^{N_i} \text{rang}(x_i^k) - \frac{N_i \cdot (N_i + 1)}{2} \quad (3.6)$$

Où U_i est la valeur de la métrique associée à la $i^{\text{ème}}$ série de mesures, N_i est le nombre d'éléments dans la $i^{\text{ème}}$ série, et $\text{rang}(x_i^k)$ est le rang de la $k^{\text{ème}}$ mesure de la $i^{\text{ème}}$ série, mesuré au sein de l'ensemble $\omega = \omega_1 \cup \omega_2$.

Le test de Wilcoxon part de l'*hypothèse nulle*, également appelée *hypothèse H0*, selon laquelle les deux séries de mesures échantillonnent un phénomène identique et doivent

donc présenter des distributions similaires. La probabilité conditionnelle d'apparition des valeurs U_1 et U_2 sachant H_0 a une loi de comportement connue, qui peut être approchée par une loi normale pour des séries suffisamment importantes. De la sorte, les valeurs prises par U_1 et U_2 permettent de déduire la probabilité de commettre une erreur de type I, c'est-à-dire la probabilité de considérer que les deux séries de mesures échantillonnent deux phénomènes différents alors qu'elles sont en réalité apparentées. Si la probabilité d'erreur de type I est inférieure à un seuil arbitraire α , l'hypothèse H_0 est acceptée ; dans le cas contraire, elle est rejetée.

Question 1 La première question a été analysée à l'aide d'un simple ratio entre réponses correctes et réponses incorrectes. Bien que cette approche ne permette pas d'étude plus avancée des distributions de réponses, l'écrasante majorité de réponses correctes a rendu toute autre analyse superflue.

Question 2 Les réponses apportées à la question 2 ont été analysées à l'aide du test de la somme des rangs de Wilcoxon, permettant la comparaison deux à deux des distributions de réponses associées à chaque groupe. Dans cette analyse, nous avons arbitrairement choisi un seuil α de 0,01 afin de limiter au maximum le risque d'erreur de type I.

Question 3 Dans le cadre de l'analyse de la question 3, nous avons calculé l'erreur E entre l'ordre choisi par les participants et l'ordre optimal des puits sous la forme :

$$E = \sum_{i \in \Omega} \sum_{j \in \Omega, j \neq i} \begin{cases} 0 & \text{si les puits } i \text{ et } j \text{ sont correctement ordonnés} \\ \text{Abs}(P_i - P_j) & \text{sinon} \end{cases} \quad (3.7)$$

Où Ω est l'ensemble des emplacements potentiels $\{\text{puits}_1 = D, \text{puits}_2 = E, \text{puits}_3 = F, \text{puits}_4 = G, \text{puits}_5 = H\}$, et P_i est la pression maximale possible au $i^{\text{ème}}$ puits. L'erreur E peut être considérée comme l'écart perceptuel par rapport au message porté par la visualisation.

Par ailleurs, nous avons appliqué le test de la somme des rangs de Wilcoxon à la distribution des erreurs E afin de comparer les groupes deux à deux. De nouveau, un seuil α de 0,01 a été choisi.

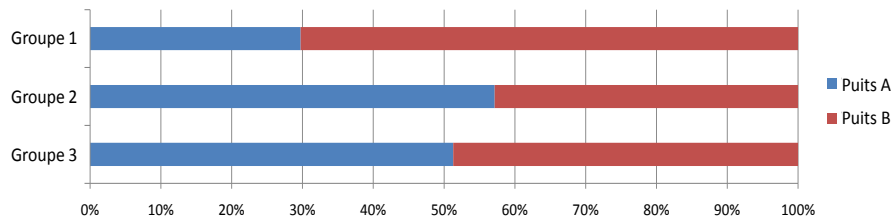
3.2.5 Résultats

Question 1 Parmi les 123 participants, plus de 96% ont répondu correctement à la première question. La majorité des participants étaient donc aptes à comprendre les données qui leurs étaient présentées. Les erreurs étaient toutes du fait d'étudiants de seconde année à l'ENSG.

En comparant les taux de réponses erronées par groupe, on constate que les erreurs sont

Tab. 3.2 – Pourcentage de réponses correctes à la question 1.

	Réponses correctes	Nombre de réponses	Nombre d'erreurs
Groupe 1	95,1%	41	2
Groupe 2	95,0%	40	2
Groupe 3	100,0%	42	0

**Fig. 3.29** – Proportions des réponses à la question 2.

réparties entre les groupes 1 et 2 (Table 3.2). Cependant, du fait de la large proportion de réponses correctes, nous considérons que les trois groupes sont équilibrés en termes de compétences ; les quatre participants aux réponses erronées ont été rejetés lors de l'analyse des questions suivantes pour plus de sécurité.

Question 2 On constate que la proportion de réponses en faveur du puits A est bien plus élevée lorsque les participants étaient informés de l'incertitude associée que lorsqu'ils l'ignoraient (Figure 3.29). Il apparaît donc que la présence de l'incertitude influence la prise de décision, en accord avec les résultats des études de [Leitner et Buttenfield \(2000\)](#) et [Deitrick et Edsall \(2006\)](#). Il convient cependant de noter que dans cet exemple, la pression maximale possible atteint des niveaux comparables aux deux emplacements (10,50 bars pour l'emplacement A et 10,42 bars pour l'emplacement B) ; de ce fait, la qualité de la prise de décision ne peut être évaluée.

Afin de comparer les réponses de manière plus poussée, nous avons utilisé le test de la somme des rangs de Wilcoxon (Table 3.3). L'analyse confirme que les groupes 2 et 3 ont des distributions de réponse similaires, tandis que le groupe 1 en diffère significativement. La manière selon laquelle le degré d'incertitude local est présenté n'a donc aucune influence sensible sur cet exercice. Nous interprétons ce phénomène comme une cause de la faible complexité de la tâche proposée ; en effet, il est aisé de quantifier la pression maximale possible en deux points, indépendamment du moyen utilisé pour présenter les incertitudes.

Question 3 L'analyse des erreurs d'ordre commises à la question 3 montre que le groupe 1 était plus précis que le groupe 3, ce dernier étant à son tour plus précis que le groupe 2 (Table 3.4). Ces résultats suggèrent qu'une visualisation coïncidente de l'incertitude est

Tab. 3.3 – Comparaison entre les distributions des réponses à la question 2.

	p(erreur de type I)	Conclusion
Groupes 1 et 2	$p_{1-2} = 0,0007$	Distributions différentes
Groupes 1 et 3	$p_{1-3} = 0,0053$	Distributions différentes
Groupes 2 et 3	$p_{2-3} = 0,3501$	Distributions similaires

Tab. 3.4 – Ecart de perception moyen pour chaque groupe à la question 3.

	Ecart moyen	Ecart type
Groupe 1	0,59 Kbars	$\pm 0,23$ bars
Groupe 2	2,60 Kbars	$\pm 0,31$ bars
Groupe 3	1,00 Kbars	$\pm 0,35$ bars

plus claire qu’une visualisation séparée, mais masque néanmoins une partie des informations initiales.

Cependant, l’écart-type laisse apparaître une intersection non-nulle entre les écarts de perception des groupes 1 et 3. Les deux distributions ont donc été comparées à l’aide du test de la somme des rangs de Wilcoxon (Table 3.5). Il apparaît alors que les distributions des réponses sont similaires entre les groupes 1 et 3, ce qui suggère que les deux visualisations présentent un niveau de clarté des informations initiales comparable.

3.2.6 Discussion

Cette étude avait pour objectif de répondre aux questions suivantes :

La visualisation des incertitudes influe-t-elle sur la prise de décision ? Cette question a d’ores et déjà reçu une réponse positive dans la littérature ; nous considérons donc ce test comme une vérification de la validité de notre étude plutôt que comme un nouveau résultat. La seconde question de notre étude a révélé des différences significatives dans les réponses des participants, selon qu’ils étaient ou non conscients de la présence d’incertitudes. Nos résultats sont donc en accord avec les découvertes des études précédentes.

Existe-t-il une différence perceptuelle entre les représentations coïncidente et adjacente de l’incertitude ? Notre étude présente des résultats contradictoires quant à l’influence de la manière selon laquelle l’incertitude est présentée : la seconde question ne montre aucune différence notable entre les deux modes de visualisation, tandis que la troisième question met en lumière une différence de qualité en termes de perception. Nous interprétons ces résultats comme la conséquence de la complexité croissante des tâches demandées aux participants. En effet, la seconde question était suffisamment simple pour

Tab. 3.5 – Comparaison entre les distributions des réponses à la question 3.

	p(erreur de type I)	Conclusion
Groupes 1 et 2	$p_{1-2} = 0,0002$	Distributions différentes
Groupes 1 et 3	$p_{1-3} = 0,9957$	Distributions similaires
Groupes 2 et 3	$p_{2-3} = 0,0043$	Distributions différentes

être résolue par des méthodes adjacentes ; cependant, la troisième question requérait une estimation plus complexe associée à un plus grand nombre d’emplacements, pour lequel les méthodes adjacentes introduisaient une surcharge perceptuelle et cognitive (Harrower, 2003). Ces résultats suggèrent que la manière selon laquelle l’incertitude est présentée influence sa perception. De ce fait, les techniques de visualisation des incertitudes devraient utiliser des représentations aussi compactes que possible afin de minimiser la charge cognitive imposée à l’observateur.

La visualisation des incertitudes clarifie-t-elle les informations présentées ? La troisième question étudiait la précision des réponses vis-à-vis de la vraie valeur portée par la visualisation. Son analyse a montré que les informations présentées sans incertitudes associées et avec visualisation coïncidente de l’incertitude présentaient le même degré de clarté, et étaient plus aisées à appréhender que les informations avec visualisation adjacente de l’incertitude. Ce résultat suggère que la visualisation de l’incertitude permet de présenter un ensemble dense de données sans pour autant obscurcir leur signification. Cette assertion doit cependant être contrastée au vu des méthodes de visualisation adjacentes de l’incertitude, pour lesquelles les décisions étaient nettement moins précises. De nouveau, ces résultats illustrent que seul un design soigneux des visualisations des incertitudes permet d’accroître la valeur de la prise de décision.

3.3 Méthodes dynamiques dédiées

Les méthodes dynamiques dédiées à la visualisation des incertitudes font intervenir une dimension temporelle dans l’affichage du modèle. Le temps peut être considéré de deux manières distinctes, (i) en le reliant directement à l’évolution temporelle de propriétés pétrophysiques dynamiques telles que la pression ou la saturation en fluides au sein d’un réservoir pétrolier, ou (ii) en l’utilisant pour faire vibrer l’affichage entre plusieurs réalisations possibles d’une unique propriété pétrophysique.

3.3.1 Animation par images clefs

L’animation par images clefs est une technique d’animation où les états intermédiaires du modèle sont reconstruits à partir des états connus. Cette technique permet de ne stocker

en mémoire qu'un nombre discret d'états \mathcal{E}_i du modèle, également nommés *images clefs*. Les états intermédiaires $\mathcal{E}(t)$ sont alors générés par interpolation entre les images clefs \mathcal{E}_i et \mathcal{E}_{i+1} à l'aide d'une fonction d'interpolation $f()$:

$$\forall t \in [t_i, t_{i+1}], \mathcal{E}(t) \leftarrow f(\mathcal{E}_i, \mathcal{E}_{i+1}, t) \quad (3.8)$$

où $[t_i, t_{i+1}]$ définit l'intervalle de temps pour lequel l'état intermédiaire $\mathcal{E}(t)$ est interpolé d'après les images clefs \mathcal{E}_i et \mathcal{E}_{i+1} .

La méthodologie de l'animation par images clefs est très générique et peut s'appliquer à de nombreuses entités, par exemple les valeurs de champs scalaires, les positions géométriques des nœuds d'un maillage ou les paramètres d'entrée d'une fonction analytique (Hanrahan et Sturman, 1985; Lasseter, 1987; Witkin et Popovic, 1995). Elle est en particulier à la base de nombreux formats de compression vidéo, par exemple MPEG4. Malgré la diversité des applications possibles, il est possible de dégager un ensemble de critères communs nécessaires à la mise en place de l'animation par images clefs :

1. Il est possible de définir une fonction d'interpolation entre les états \mathcal{E}_i et \mathcal{E}_{i+1} du modèle, ce qui présuppose généralement qu'une bijection entre \mathcal{E}_i et \mathcal{E}_{i+1} est connue.
2. Les images clefs \mathcal{E}_i sont échantillonnées à une résolution suffisante pour reconstruire l'évolution temporelle du modèle.
3. La fonction d'interpolation $f()$ produit des résultats cohérents avec le modèle présenté. Par exemple, soit un modèle de faciès géologiques en domaine fluvial dans lequel trois faciès sont encodés à l'aide de nombres entiers : 0 pour la plaine d'inondation, 1 pour les levées, et 2 pour les chenaux. Pour animer entre les faciès "plaine d'inondation" et "chenal", il est possible de proposer une interpolation par moyenne ; cependant, une telle interpolation fera apparaître le faciès "levée" à certains pas de temps, de manière incohérente par rapport aux données initiales.

Il est fréquent que l'ensemble des images clefs \mathcal{E}_i présente des caractéristiques communes. Chaque état \mathcal{E}_i peut alors être décomposé en une partie \mathcal{E}^C commune à l'ensemble des images et une partie spécifique \mathcal{E}_i^S qui lui est propre.¹⁴ Bien que cette décomposition soit mathématiquement équivalente aux états \mathcal{E}_i , elle permet de limiter l'espace de stockage nécessaire pour conserver les images clefs et de minimiser les quantités de données transférées lors de l'animation elle-même (Viard *et al.*, 2010a).

3.3.2 Types d'interpolations temporelles continues

Idéalement, les images intermédiaires $\mathcal{E}(t)$ devraient être générées avec les règles génériques qui ont servi lors de la création des images clefs \mathcal{E}_i , de telle sorte que leurs caractéristiques

¹⁴Par exemple, lors de l'animation de réalisations de porosité le long d'une surface stratigraphique, les états \mathcal{E}_i sont composés de la géométrie de la surface stratigraphique et du champ scalaire de porosité. Dans ce cas, la géométrie est identique pour toutes les images clefs et est donc stockée dans la composante \mathcal{E}^C des images clefs, tandis que les champs de porosité varient d'une image à l'autre et sont stockés dans les composantes individuelles \mathcal{E}_i^S .

téristiques respectent les paramètres physiques du phénomène étudié. Cependant, il n'est pas rare que les règles selon lesquelles les images clefs \mathcal{E}_i ont été générées ne permettent pas de produire des états intermédiaires évoluant de manière continue de \mathcal{E}_i à \mathcal{E}_{i+1} . Par exemple, [Srivastava \(1995\)](#) note qu'une faible variation de la graine utilisée lors de la génération de séquences de nombres pseudo-aléatoires pour des simulations géostatistiques produit des réalisations entièrement différentes *in-fine*.

En outre, même lorsqu'une évolution naturelle existe entre deux images clefs (par exemple lors de la modélisation de propriétés pétrophysiques dynamiques), l'évaluation de ces états intermédiaires à l'aide de règles génétiques est généralement coûteux en temps de calcul ([Scheidt et al., 2007](#); [Fetel et Caumon, 2008](#)). Or, l'animation est présentée en temps réel avec des états intermédiaires $\mathcal{E}(t)$ calculés à la volée ; pour que l'animation soit perçue de manière continue par un observateur, il est nécessaire de pouvoir reconstruire les états intermédiaires à une vitesse supérieure à vingt images par seconde ([Bederson et Boltman, 2003](#)).¹⁵

Les états intermédiaires sont donc généralement reconstruits à l'aide de fonctions d'approximation, choisies pour offrir un compromis satisfaisant entre erreur d'estimation et vitesse d'estimation. Dans la plupart des méthodes d'animation par images clefs, l'interpolation est évaluée sous la forme d'une combinaison linéaire des images clefs les plus proches, pour laquelle les coefficients de la combinaison linéaire sont des fonctions du temps :

$$\forall t \in [t_i, t_{i+1}], \mathcal{E}(t) \leftarrow \alpha(t) \cdot \mathcal{E}_i + \beta(t) \cdot \mathcal{E}_{i+1} \quad (3.9)$$

Dans le cadre de cette thèse, nous avons expérimenté trois fonctions d'approximation pouvant être exprimées sous la forme de combinaisons linéaires des images clefs : l'interpolation linéaire, l'interpolation d'Hermite par splines cubiques et la déformation graduelle. L'interpolation linéaire est équivalente à l'interpolation barycentrique dans cette application ; elle définit les coefficients $\alpha(t)$ et $\beta(t)$ comme des fonctions linéaires du temps, de telle sorte que $\alpha(t_i) = \beta(t_{i+1}) = 0$ et $\alpha(t_{i+1}) = \beta(t_i) = 1$:

$$\forall t \in [t_i, t_{i+1}], \quad \alpha(t) = \frac{t-t_i}{t_{i+1}-t_i} \quad \text{et} \quad \beta(t) = \frac{t_{i+1}-t}{t_{i+1}-t_i} \quad (3.10)$$

Cette fonction d'approximation est de loin la plus efficace, mais produit généralement des approximations grossières des états intermédiaires lorsque l'animation met en jeu des phénomènes hautement non-linéaires. Par ailleurs, l'interpolation linéaire de deux images clefs produit généralement des états intermédiaires présentant une plus faible variabilité spatiale, par effet de moyenne.

Pour éviter cet effet, [Ehlschlaeger et al. \(1996\)](#) proposent d'utiliser une interpolation temporelle basée sur des fonctions trigonométriques (sinus et cosinus) de manière

¹⁵La plupart des appareils de projection présentent les vidéos à une vitesse de 24 images par seconde, afin de garantir que le mouvement sera perçu de manière continue.

similaire aux techniques de déformation graduelle (Hu, 2000) :

$$\forall t \in [t_i, t_{i+1}], \quad \alpha(t) = \cos\left(\frac{t-t_i}{t_{i+1}-t_i} \cdot \frac{\pi}{2}\right) \quad \text{et} \quad \beta(t) = \sin\left(\frac{t-t_i}{t_{i+1}-t_i} \cdot \frac{\pi}{2}\right) \quad (3.11)$$

Cette méthode d'interpolation garantit que la variabilité spatiale des états interpolés soit identique à celle des images clefs, à condition que les images clefs présentent une distribution Gaussienne de moyenne nulle. Si les images clefs présentent une distribution Gaussienne de moyenne $\mu_{\mathcal{E}}$ non-nulle, il est cependant possible d'utiliser ce mode d'interpolation en décalant le champ scalaire des image clefs par la valeur moyenne ; la fonction d'interpolation devient alors :

$$\forall t \in [t_i, t_{i+1}], \quad \mathcal{E}(t) \leftarrow \cos\left(\frac{t-t_i}{t_{i+1}-t_i} \cdot \frac{\pi}{2}\right) \cdot (\mathcal{E}_i - \mu_{\mathcal{E}}) + \sin\left(\frac{t-t_i}{t_{i+1}-t_i} \cdot \frac{\pi}{2}\right) \cdot (\mathcal{E}_{i+1} - \mu_{\mathcal{E}}) + \mu_{\mathcal{E}} \quad (3.12)$$

Cette approche n'est cependant valable que pour des images clefs \mathcal{E}_i présentant une distribution Gaussienne ou sub-Gaussienne. Pour les images clefs ne respectant pas cette condition, il est possible d'utiliser des interpolations polynomiales de plus haut degré que l'interpolation linéaire. Parmi ces fonctions polynomiales, l'interpolation d'Hermitte par splines cubiques, notée $h(t)$, est assez populaire ; elle permet de respecter les dérivées entre les champs scalaires, tout en étant généralement plus précise que l'interpolation linéaire de par son degré plus élevé. Les coefficients $\alpha(t)$ et $\beta(t)$ sont alors calculés sous la forme :

$$\begin{aligned} \forall t \in [t_i, t_{i+1}], \quad \alpha(t) = h(t) \quad \text{et} \quad \beta(t) = 1 - h(t) \\ \text{avec} \quad h(t) = \sum_{j=0}^3 c_j \cdot \left(\frac{t-t_i}{t_{i+1}-t_i}\right)^j \end{aligned} \quad (3.13)$$

Les coefficients c_j du polynôme d'Hermitte sont inconnus, mais peuvent être calculés d'après les valeurs et les dérivées des images clefs S_i grâce au système d'équations :

$$\begin{aligned} h(t_i) = \mathcal{E}_i \quad \frac{\delta h}{\delta t}(t_i) = \delta \mathcal{E}_i \\ h(t_{i+1}) = \mathcal{E}_{i+1} \quad \frac{\delta h}{\delta t}(t_{i+1}) = \delta \mathcal{E}_{i+1} \\ \text{avec} \quad \frac{\delta h}{\delta t}(t) = \sum_{j=1}^3 j \cdot c_j \cdot \left(\frac{t-t_i}{t_{i+1}-t_i}\right)^{j-1} \end{aligned} \quad (3.14)$$

où $\delta \mathcal{E}_i$ et $\delta \mathcal{E}_{i+1}$ sont les dérivées des images clefs \mathcal{E}_i et \mathcal{E}_{i+1} . Cependant, ces dérivées ne sont pas toujours connues. Il est alors possible de les calculer d'après les images clefs voisines \mathcal{E}_{i-1} et \mathcal{E}_{i+2} , sous la forme :

$$\delta \mathcal{E}_i = \frac{\mathcal{E}_{i+1} - \mathcal{E}_{i-1}}{2} \quad \text{et} \quad \delta \mathcal{E}_{i+1} = \frac{\mathcal{E}_{i+2} - \mathcal{E}_i}{2} \quad (3.15)$$

Dans ce cas, le système d'équations présenté précédemment (Equation 3.14) peut être inversé pour obtenir la solution suivante :

$$\begin{pmatrix} c_3 \\ c_2 \\ c_1 \\ c_0 \end{pmatrix} = \frac{1}{2} \cdot \begin{pmatrix} 1 & 5 & -5 & -1 \\ -2 & -7 & 8 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 2 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} \mathcal{E}_{i+2} \\ \mathcal{E}_{i+1} \\ \mathcal{E}_i \\ \mathcal{E}_{i-1} \end{pmatrix} \quad (3.16)$$

L'approche par interpolation d'Hermite par splines cubiques nécessite que les images clefs \mathcal{E}_{i-1} et \mathcal{E}_{i+2} existent, ce qui n'est pas le cas pour la première et la dernière image clef de la séquence. Pour ces cas particuliers, il est possible de substituer \mathcal{E}_i à \mathcal{E}_{i-1} ou \mathcal{E}_{i+1} à \mathcal{E}_{i+2} . Cette technique produit toutefois des effets de bord lors de l'interpolation, réduisant la qualité d'approximation de la spline cubique.

3.3.3 Animation de réalisations non-équiprobables

Lorsque l'animation est utilisée pour faire vibrer le modèle entre plusieurs réalisations possibles prises dans un ensemble Ω , il est fréquent que chaque réalisation soit considérée comme aussi probable que toute autre réalisation de Ω (on parle de *réalisations équiprobables*). Cependant, il peut arriver que des informations complémentaires permettent d'assigner une probabilité globale à chaque réalisation ; ces informations peuvent par exemple correspondre à des jugements subjectifs d'experts géologues, à des données qui n'ont pas été utilisées lors de la création des réalisations (nouveaux puits, informations sismiques, etc.), ou à une qualité de "prédiction" du comportement d'un réservoir dans le cadre d'opérations de calage à l'historique de la production. Dans ce cas, l'animation "standard" ne permet pas de présenter la différence de probabilité entre deux réalisations ; nous proposons dans cette thèse une extension à l'animation standard, nommée *animation pondérée*, où la durée selon laquelle une réalisation est présentée à l'écran varie en fonction de la probabilité globale de la réalisation (Viard *et al.*, 2009f, 2010d). Cette technique peut être vue comme une extension des travaux de Lundström *et al.* (2007), qui permet d'appliquer les méthodes d'animation probabiliste à tout type d'animation plutôt qu'aux seules animations discontinues.

Lorsque l'animation est discontinue, c'est-à-dire que les images clefs sont affichées successivement sans calcul d'images intermédiaires, le lien entre la durée d'affichage d'une réalisation et la probabilité associée à cette réalisation est trivial. Cependant, dans le cadre d'une animation continue, l'évaluation de la durée d'affichage devient plus complexe : en effet, l'état du modèle est interpolé entre deux images clefs dans la quasi-totalité de l'animation, ce qui signifie que deux probabilités différentes doivent être prises en compte. Pour résoudre ce problème, l'animation pondérée procède par découplage entre le *temps réel* t , qui correspond à la durée pouvant être chronométrée par un observateur, et le *temps d'affichage* τ , qui correspond à la durée utilisée en interne dans la fonction d'interpolation

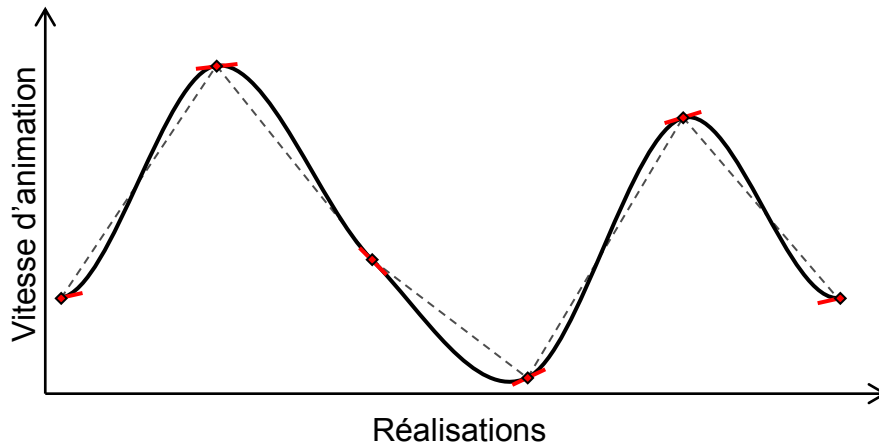


Fig. 3.30 – Vitesse d’affichage \mathcal{V} en fonction des réalisations \mathcal{E}_i . Le trait noir présente l’interpolation d’Hermite et le trait gris en pointillés l’interpolation linéaire de la vitesse \mathcal{V} . Les diamants rouges correspondent aux instants où une réalisation \mathcal{E}_i est affichée (pas d’interpolation) et les traits rouges à la dérivée de la vitesse en ces points, calculée d’après les interpolations linéaires à gauche et à droite.

$f()$. Ces deux temps sont reliés par une fonction de vitesse d’affichage \mathcal{V} , construite d’après les probabilités p_i et p_{i+1} associées aux réalisations \mathcal{E}_i et \mathcal{E}_{i+1} , de telle sorte que :

$$\tau = \int_0^t \mathcal{V}(r) dr \quad (3.17)$$

La fonction de vitesse d’affichage \mathcal{V} est construite à l’aide d’une interpolation par morceaux entre les probabilités p_i et p_{i+1} associées aux réalisations \mathcal{E}_i et \mathcal{E}_{i+1} . Les probabilités p_i et p_{i+1} sont inversées¹⁶ pour que les réalisations ayant une probabilité élevée présentent une vitesse plus faible, et soient donc affichées pour une durée réelle plus longue. Lorsque l’animation présente une image clef \mathcal{E}_i , la vitesse \mathcal{V} est donc égale à $\frac{1}{p_i}$. Entre les images clefs, nous utilisons une interpolation d’Hermite par splines cubiques afin d’éviter des variations abruptes de la fonction de vitesse (Figure 3.30). Les dérivées utilisées par l’interpolation d’Hermite sont calculées comme la moyenne des dérivées à droite et à gauche, calculées à l’aide d’une interpolation linéaire.

Bien que la fonction de vitesse \mathcal{V} soit le plus souvent strictement positive, il peut arriver qu’elle soit nulle ou négative si les dérivées varient de manière importante entre deux images clefs successives. En ce cas, ces valeurs incohérentes de vitesse produisent respectivement un arrêt de l’animation ou une oscillation des temps d’affichage dans un intervalle de temps restreint, convergeant vers le temps d’affichage où la vitesse est nulle. Pour éviter cet effet indésirable, nous imposons simplement un seuil à la vitesse \mathcal{V} .

¹⁶Les réalisations présentant une probabilité nulle sont éliminées avant la construction de la vitesse \mathcal{V} ; il est donc toujours possible de calculer l’inverse de la probabilité p_i .

3.4 Validation des méthodes dynamiques

Les techniques par animation ont été évaluées par le biais d'une étude sur des utilisateurs, parmi un panel de 29 participants au sein d'ExxonMobil Upstream Research Company domiciliés à Houston, Texas. L'ensemble de ces participants sont des experts dans le domaine du pétrole, dont les spécialisations sont pour la plupart l'interprétation sismique, l'ingénierie, le développement de logiciels ou la géomodélisation.

L'étude sur des utilisateurs consistait en un ensemble de dix questions, comportant quatre exercices distincts. Les exercices étaient supportés par une série d'animations présentées dans un format de vidéo numérique, accessibles via un réseau intranet à ExxonMobil.¹⁷ Pour comparer les techniques par animation pondérée et par animation standard, nous avons séparé les participants en deux groupes. Chaque groupe s'est vu poser le même jeu de questions; cependant, les animations supportant les questions différaient d'un groupe à l'autre.

3.4.1 Description des données

Les données utilisées dans le cadre de cette étude sur des utilisateurs sont un jeu synthétique représentant des faciès sableux et argileux. Le faciès de type argile est considéré comme entièrement imperméable, et le faciès de type sable comme perméable à l'écoulement de l'eau et des hydrocarbures. Les faciès sont générés à l'aide de simulations séquentielles Gaussiennes, dont les valeurs sont tronquées afin de produire des données catégoriques.

Le jeu de données présente un puits producteur, foré au centre d'un corps sableux dont les frontières sont incertaines (Figure 3.31). Pour accélérer la production, il est prévu de forer un puits injecteur, dont le but est d'injecter de l'eau dans le milieu poreux afin de balayer les hydrocarbures vers le puits producteur. Pour maximiser la quantité d'hydrocarbures balayés vers le puits producteur, il est nécessaire de forer le puits injecteur aux frontières du corps sableux. Cependant, ces frontières sont incertaines; or, si le puits injecteur est foré en domaine argileux, l'eau ne pourra s'immiscer dans le milieu poreux car la matrice rocheuse est considérée imperméable. L'exercice de localisation du puits injecteur consiste donc à maximiser le profit réalisé tout en minimisant le risque de forer dans les argiles.

Sur l'ensemble des questions de l'étude sur des utilisateurs, seule une section du modèle était présentée. Les questions étaient donc basées sur le présupposé que les structures

¹⁷Une copie de la version du questionnaire par animation pondérée est également accessible à l'adresse suivante : <http://www.gocad.org/www/people/files/SurveyWeighted.html>.

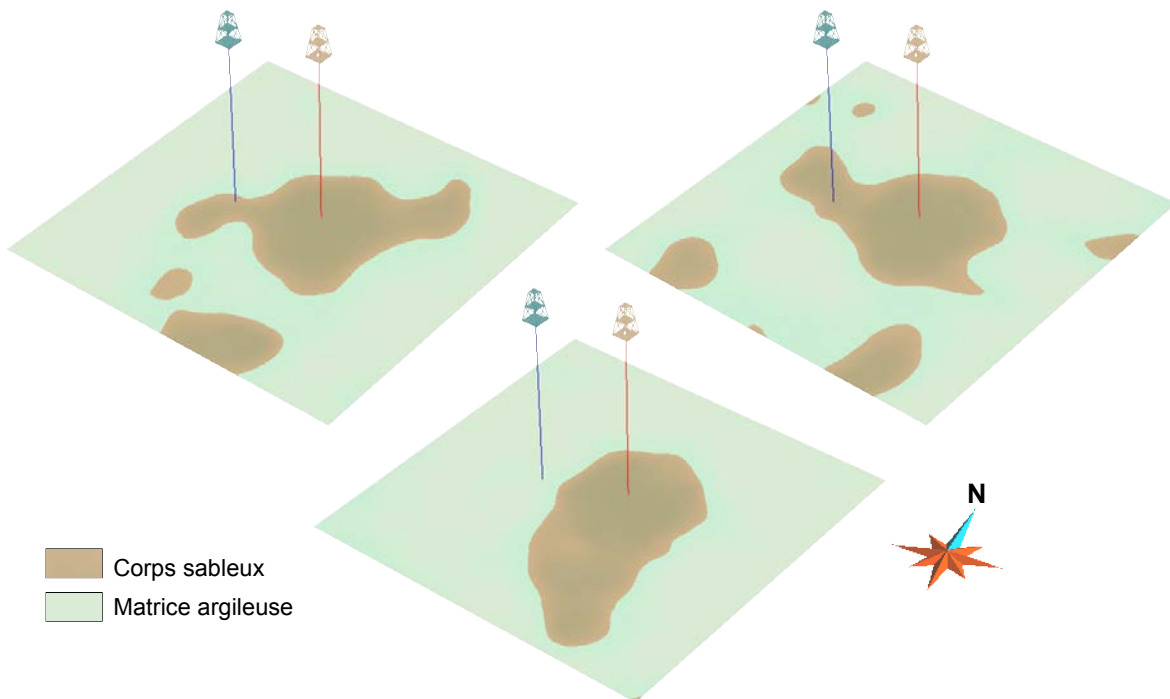


Fig. 3.31 – Jeu de données synthétique utilisé dans l’étude sur des utilisateurs des méthodes dynamiques. Plusieurs réalisations de faciès sont présentées. Les puits rouge et bleu correspondent respectivement au puits producteur et à une position proposée pour le puits injecteur.

visibles sur la section étaient révélatrices de la structure des faciès sur l’ensemble du volume d’intérêt.¹⁸

3.4.2 Description des questions

Question 1 La question 1 demandait aux participants de se positionner en terme d’expertise en visualisation des incertitudes sur une échelle allant de “novice” à “expert”. Cette question avait pour but de détecter des biais possibles liés à des différences d’expertise entre les deux groupes, qui rendraient la comparaison entre les groupes caduque.

Questions 2, 4 et 6 Dans le cadre de cette série de questions, les participants devaient visualiser une série de vidéos d’animations continues entre différentes réalisations de faciès, dont certaines prédisaient que le puits injecteur soit connecté au corps sableux et d’autres que le puits injecteur soit connecté à la matrice argileuse. Le but de l’exercice était d’estimer la probabilité que le puits producteur soit connecté au corps sableux, cette

¹⁸Dans le cas général, ce présupposé est faux, et une interprétation du modèle doit se faire sur l’ensemble du volume.

probabilité étant proportionnelle à l'intervalle de temps pour lequel une telle connexion apparaît sur l'animation.

Pour toutes ces questions, les réalisations utilisées dans l'animation étaient identiques ; cependant, elles n'étaient pas organisées dans le même ordre d'une animation à l'autre. Les questions 2 et 6 présentaient des réalisations en alternant connexion du puits injecteur avec l'argile et avec le sable. Inversement, la question 4 présentait l'ensemble des réalisations dont le puits injecteur était connecté à l'argile, puis l'ensemble des réalisations dont le puits injecteur était connecté au sable. Afin d'éviter tout biais du à des comparaisons avec les réponses précédentes, une notice explicative prétendait (à tort) que les réalisations utilisées variaient d'une question à l'autre.

Cette série de questions était posée afin de répondre à plusieurs interrogations :

1. En comparant les deux groupes, les réponses ont-elles un degré de précision comparable, ou l'une des deux méthodes d'animation permet-elle une meilleure perception des informations présentées ?
2. Si l'une ou l'autre des méthodes d'animation ne permet pas de percevoir les probabilités de manière correcte, existe-t-il un biais d'estimation significatif et répétable conduisant à une surestimation ou une sous-estimation du résultat ?
3. L'ordre dans lequel les réalisations sont présentées influe-t-il sur la perception du phénomène étudié ? Si oui, quel est l'ordre qui offre le meilleur résultat ?

Question 8 La question 8 demandait aux participants d'évaluer la probabilité de connexion entre le puits injecteur et le corps sableux, de manière similaire aux questions 2, 4 et 6. Cependant, cette question était supportée par une animation *discontinue* des réalisations de faciès en lieu et place des animations continues par images clefs présentées dans les questions précédentes. Cet exercice visait à déterminer si l'animation continue permet d'améliorer l'estimation de la probabilité de connexion au corps sableux ou, au contraire, complique la tâche des participants en augmentant le fardeau cognitif qui leur était imposé.

Questions 3, 5, 7 et 9 Après chaque animation, les participants étaient interrogés sur la confiance qu'ils avaient en leur estimation de la probabilité de connexion entre le puits injecteur et le corps sableux. Ces questions avaient pour objectif de déterminer si les utilisateurs étaient conscients de leur erreur d'estimation, ou s'ils présentaient un excès ou un défaut de confiance en leurs réponses.

Question 10 La dernière question permettait aux participants d'exprimer de manière ouverte leurs commentaires sur les techniques d'animation ou sur l'étude elle-même. Cette question était mise en place pour collecter des informations subjectives que les questions

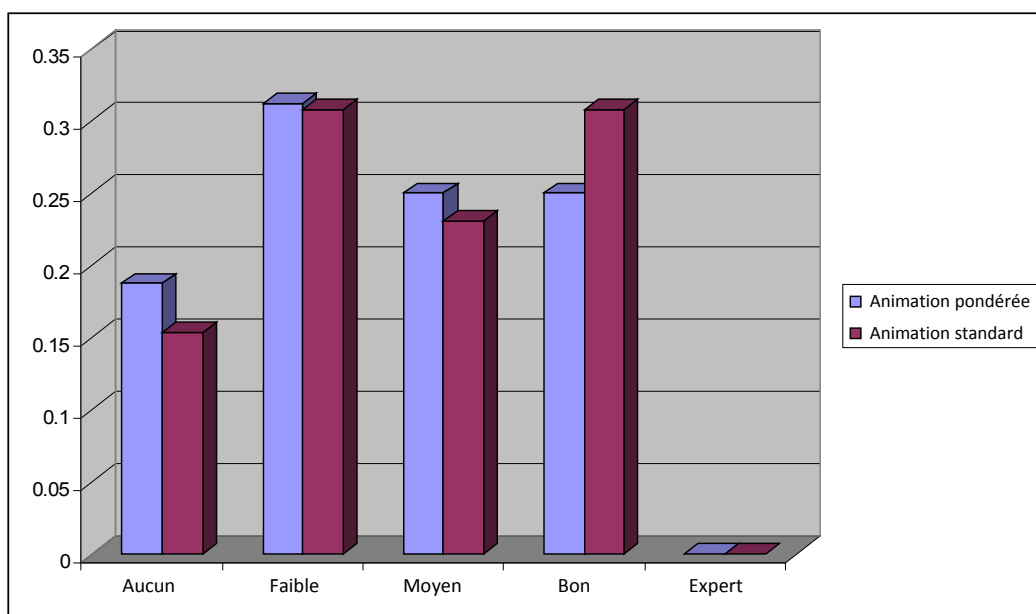


Fig. 3.32 – Distribution des degrés d'expertise des participants en visualisation des incertitudes, présentés séparément pour chaque groupe (animation standard et animation pondérée).

précédentes ne pouvaient capturer, ainsi que d'éventuelles suggestions d'amélioration ou d'application des méthodes par animation.

3.4.3 Résultats

Equilibre entre les groupes Le nombre de participants a été réparti de manière homogène entre les deux groupes (13 participants pour l'animation standard et 16 participants pour l'animation pondérée). De manière générale, la question 1 montre que le degré d'expertise des participants est équilibré d'un groupe à l'autre (Figure 3.32). Un léger biais est cependant détectable en faveur de l'animation standard, qui compte plus de participants ayant un bon degré d'expertise que le groupe associé à l'animation pondérée. L'on peut également remarquer qu'aucun participant ne s'est défini comme un expert en visualisation des incertitudes.

Perception quantitative On constate une sous-estimation significative de la probabilité de connexion entre le puits injecteur et le corps sableux au sein des deux groupes (Table 3.6); ces erreurs d'estimation illustrent la faible efficacité des méthodes par animation pondérée dans le cadre d'interprétations quantitatives des données. Lowe (1999) fournit une explication possible à ce phénomène : l'animation du modèle provoquerait une surcharge cognitive, résultant en une diminution de la qualité des informations pouvant être extraites mentalement par les participants. Néanmoins, on peut noter que la percep-

Tab. 3.6 – Estimation moyenne de la probabilité de connexion entre le puits injecteur et le corps sableux.

	Animation standard	Animation pondérée	Organisation
Question 2	25,4%	17,5%	En alternance
Question 4	43,8%	26,3%	Par famille
Question 6	29,2%	20,0%	En alternance
Question 8	44,6%	28,1%	En alternance
Valeur réelle	48,0%	39,7%	-

tion *qualitative* des données est généralement correcte ; en effet, la tendance du puits à être moins fréquemment connecté à du sable qu'à de l'argile a été correctement observée par les participants des deux groupes.

Dans cette étude, la probabilité de connexion entre le puits injecteur et le corps sableux est systématiquement sous-estimée pour les deux groupes sur l'ensemble des questions. La raison d'une telle répétabilité du biais d'estimation est présentement inconnue ; il est possible qu'elle soit liée à la méthode d'animation elle-même, mais il est plus probable que ses causes proviennent d'un paramètre spécifique à notre étude, par exemple la fonction de transfert utilisée pour représenter les faciès.

Influence de l'ordre des images clef En comparant les résultats de la question 4 aux résultats des questions 2 et 6, il apparaît que les réponses sont bien plus précises lorsque les réalisations sont organisées par famille que lorsque l'animation alterne les familles de réalisations. Cette découverte contraste avec les travaux de [Ehlschlaeger \(1998\)](#), qui n'observe pas d'écart perceptuel significatif entre les animations utilisant des réalisations ordonnées et en alternance. Ces résultats contradictoires peuvent être expliqués par des variabilités différentes au sein des données ; en effet, d'après les commentaires d'un des participants, il semble que les variations de couleurs à haute-fréquence aient eu un effet négatif sur ses capacités d'interprétation de la durée de connexion. Il est également possible que la plus haute dimension des données dans notre étude ait révélé des tendances non perceptibles dans les travaux de [Ehlschlaeger \(1998\)](#), du fait de données plus aisées à interpréter.

Les réalisations associées à l'animation discontinue étaient organisées en alternant réalisations dont le puits injecteur est connecté au sable et réalisations dont le puits injecteur est connecté à l'argile. Bien qu'un résultat similaire aux méthodes par animation continue puisse être attendu, aucune comparaison n'est possible dans ce cas ; la question de l'organisation optimale des images clefs dans le cadre d'une animation discontinue reste donc ouverte.

Dans l'exemple proposé aux participants, le classement des réalisations par famille était trivial (la réalisation est connectée, ou elle ne l'est pas). Cependant, dans le cas général,

les critères définissant une famille de réalisations peuvent être bien plus complexes à formaliser ; il peut même arriver que ces critères soient inconnus *a priori* et doivent être reconstruits d'après les réalisations elles-mêmes. Pour répondre à cette problématique, [Panhaleux et al. \(2010\)](#) ont mis en place un environnement de classification d'images bidimensionnelles ou tridimensionnelles basés sur la distance de Hausdorff ([Huttenlocher et al., 1993](#)), permettant de séparer des classes entre les réalisations. Cette approche présente un degré de complexité supérieur aux approches de segmentation classiques, les réalisations étant des entités de dimensionnalité très élevée.

Comparaison entre animation standard et animation pondérée On constate que les réponses associées à l'animation standard discontinue sont sensiblement plus précises que les réponses associées à l'animation pondérée discontinue. Cette observation semble de nouveau indiquer que l'animation pondérée obscurcit la perception des données. Cependant, la précision des réponses peut également s'expliquer par la simplicité de cet exercice. En effet, les participants du groupe de l'animation standard pouvaient se contenter de compter le nombre d'occurrences où le puits injecteur était connecté au corps sableux ; les durées d'affichage étant identiques pour chaque réalisation, un simple rapport vis-à-vis du nombre total de réalisations donnait immédiatement le résultat correct.

De manière similaire, les réponses observées par animation continue sont plus précises dans le cadre de l'animation standard qu'avec les méthodes par animation pondérée. D'après [Lowe \(1999\)](#), les méthodes par animation peuvent créer une surcharge perceptuelle et cognitive qui affecte la compréhension des données. Par comparaison avec l'animation standard, les méthodes par animation pondérée accroissent davantage la complexité de la visualisation. De ce fait, il est légitime de s'attendre à une surcharge plus importante encore.

Comparaison entre animation continue et animation discontinue Les résultats associés à l'animation discontinue sont significativement plus proches des valeurs réelles de probabilité de connexion que ceux obtenus par des méthodes d'animation continue, faisant de l'animation discontinue une méthode potentiellement utilisable pour l'interprétation quantitative des données.

Biais d'apprentissage En moyenne, les réponses de la question 6 sont plus précises que celles de la question 2. Bien que la différence de précision soit peu significative, cet effet pourrait correspondre à un biais d'apprentissage, les participants ayant amélioré leur perception des données et de la visualisation par animation grâce à leur expérience acquise en répondant aux questions précédentes.

Confiance dans les réponses En moyenne, les participants ont exprimé une confiance en leurs résultats respectivement de $\pm 16,6\%$ pour le groupe associé à l'animation standard

Tab. 3.7 – Comparaison de la confiance moyenne des animations standard et pondérée.

	Moyenne	Déviatiion standard
Animation standard	$\pm 16,6\%$	12,1%
Animation pondérée	$\pm 19,0\%$	12,7%

Tab. 3.8 – Comparaison entre les distributions de confiance des animations standard et pondérée.

	Question 3	Question 5	Question 7	Question 9	Seuil
p(erreur)	0,9281	0,6641	0,8576	0,3366	0,01
Conclusion	H0 acceptée	H0 acceptée	H0 acceptée	H0 acceptée	-

et de $\pm 19,0\%$ pour le groupe associé à l’animation pondérée (Table 3.7). Dans les deux cas, la valeur réelle de probabilité se situe dans la limite supérieure de l’intervalle de confiance ainsi défini. Par ailleurs, la confiance des participants en leur interprétation est généralement assez faible, illustrant le faible potentiel des méthodes par animation pour l’interprétation quantitative des données.

Cependant, le degré de confiance varie significativement d’un participant à l’autre, résultant en une valeur de déviation standard élevée. La moyenne n’est donc pas un indicateur fiable de la répartition des confiances ; pour une analyse plus complète des résultats, nous avons comparé les distributions des réponses à l’aide du test de somme des rangs de Wilcoxon (Table 3.8). Sur l’ensemble des questions, l’hypothèse nulle selon laquelle les réponses ont des distributions similaires est acceptée. De manière similaire, nous avons pu étudier les différences de distribution entre les questions, tous types d’animation confondus (Table 3.9) ; de nouveau, l’hypothèse nulle est systématiquement acceptée. Ces résultats suggèrent que les participants n’étaient pas conscients de leur erreur d’estimation : en effet, la confiance associée à leurs réponses ne varie pas en fonction du degré réel de précision des réponses.

Influence de l’expertise des participants Afin d’évaluer les effets de l’expertise sur les réponses des participants, nous avons transformé les réponses apportées à la question

Tab. 3.9 – Comparaison entre les distributions de confiance par question, toutes méthodes d’animation confondues.

	Question 3	Question 5	Question 7	Question 9
Question 3	1,0	-	-	-
Question 5	0,4840	1,0	-	-
Question 7	0,6785	0,8161	1,0	-
Question 9	0,9937	0,5197	0,7037	1,0

Tab. 3.10 – Corrélation des réponses avec le degré d’expertise des participants. Les tirets indiquent qu’aucune interprétation pertinente ne peut être apportée.

	Animation standard		Animation pondérée	
	Corrélation	Les experts sont...	Corrélation	Les experts sont...
Question 1	1,0	-	1,0	-
Question 2	-0,196	Plus précis	-0,255	Plus précis
Question 4	-0,202	Plus précis	0,062	-
Question 6	-0,427	Plus précis	-0,230	Plus précis
Question 8	0,329	Moins précis	0,351	Moins précis
Question 3	0,173	Plus confiants	-0,682	Moins confiants
Question 5	0,300	Plus confiants	-0,465	Moins confiants
Question 7	0,270	Plus confiants	-0,502	Moins confiants
Question 9	-0,004	-	-0,384	Moins confiants

1 en valeurs numériques – de 1 pour aucune expertise à 4 pour une bonne expertise en visualisation des incertitudes. Cette technique nous a permis d’étudier le coefficient de corrélation entre la variable aléatoire “expertise du participant” et les informations collectées dans les questions suivantes (Table 3.10). Bien qu’aucune relation linéaire clairement définie n’apparaisse (la plupart des coefficients de corrélation sont inférieurs à 0,5), il est possible de dégager un certain nombre de tendances d’après cette analyse.

La comparaison entre animation continue et discontinue a révélé que les experts étaient systématiquement plus précis avec l’animation continue, mais *moins* précis avec l’animation discontinue. Ce résultat est inattendu ; en effet, nous supposons que les effets de l’expertise seraient moins sensibles avec l’animation discontinue, et non pas que les novices seraient plus précis que les experts. La cause de ce phénomène est présentement inconnue, cependant, deux hypothèses peuvent être proposées : (i) l’évaluation de l’expertise en visualisation des incertitudes reposait sur un jugement subjectif des participants, qui peut donc être biaisé, ou (ii) les experts étaient habitués aux techniques par animation continue et ont tenté d’appliquer ce savoir à l’animation discontinue, plutôt que d’évaluer l’animation discontinue de manière indépendante.

La comparaison entre l’animation standard et l’animation pondérée montre que les experts étaient majoritairement plus confiants que les novices avec l’animation standard, mais moins confiants avec l’animation pondérée. Ce résultat suggère que les experts étaient plus sensibles à l’accroissement de difficulté que les novices, et par extension, plus conscients de leur possible erreur d’estimation.

Enfin, nous avons découvert que les effets de l’expertise sont généralement moins sensibles lorsque les réalisations sont organisées par famille ; de manière similaire, les effets de l’expertise apparaissent plus clairement avec l’animation pondérée qu’avec l’animation standard. Ces résultats étoffent les théories de Roth (2009), selon lesquelles les effets

de l'expertise sont nettement marqués lors de tâches complexes – en l'occurrence, les variations de couleur à haute fréquence et la vitesse non-uniforme de l'animation – mais n'ont que peu d'influence sur des tâches plus simples.

Commentaires des participants La question 10 nous a permis de collecter les commentaires de nombreux participants sur les techniques par animation ; ce paragraphe résume les informations les plus importantes :

1. Un des participants a exprimé la possibilité d'un biais d'apprentissage acquis au fil des questions, améliorant ses capacités à évaluer la probabilité de connexion au cours du temps.
2. Plusieurs participants ont indiqué qu'ils étaient tentés de comparer les durées ressenties d'une question à l'autre plutôt que d'évaluer chaque question de manière entièrement indépendante.
3. Une participante a signalé ses difficultés à évaluer la probabilité de connexion sur le seul critère de la durée d'affichage, sans y incorporer de connaissances externes telles que le réalisme de la forme des corps sableux, la possibilité de connexion même quand le puits injecteur n'était pas directement relié au corps sableux (les argiles n'étant jamais *complètement* imperméables), *etc.*
4. Plusieurs participants ont exprimé leur préférence pour l'animation discontinue plutôt que pour l'animation continue. Certains ont affirmé que l'animation continue manquait de repères temporels, l'interpolation entre images clefs masquant les images clefs elles-mêmes.
5. Plusieurs participants ont affirmé qu'ils interprèteraient des visualisations des incertitudes statiques plus rapidement et plus précisément que des méthodes par animation. La plupart ont indiqué que les méthodes par animation étaient utiles pour détecter des événements possibles au sein des données, mais peu efficaces pour quantifier ou interpréter des données.

Les participants ont également proposé plusieurs améliorations possibles pour les méthodes d'animation :

1. Incorporer à la visualisation un marqueur indiquant le changement d'images clefs, par exemple une image entièrement noire visible quelques fractions de secondes.
2. Intégrer dans l'interface une horloge ou un curseur temporel, permettant de suivre avec plus grande précision l'évolution du temps réel.

3.4.4 Limites

Quelle que soit la planification avant d'effectivement mettre en place une étude sur des utilisateurs, des améliorations possibles dans le format de l'étude apparaissent généralement *a posteriori*. Cette section répertorie les limites observées lors de notre étude :

1. Les questions 3, 5, 7 et 9 proposent un ensemble de réponses à caractère objectif (pourcentages) alors que les questions elles-mêmes sont de nature hautement subjective. Un ensemble de choix plus qualitatifs, du même type que ceux proposés en question 1, aurait probablement été plus cohérent sur ce groupe de questions.
2. Le format de l'étude ne nous a pas permis de mesurer les temps de réponse au questionnaire. Une analyse du temps moyen de réponse nous aurait permis de collecter plus d'indices sur un possible biais d'apprentissage (le temps de réponse devant alors décroître à chaque question), ainsi que sur le niveau de difficulté de chaque exercice.
3. Les exercices des questions 2 à 9 ont été proposés dans un ordre prédéfini, identique pour l'ensemble des participants. Pour limiter les effets du biais d'apprentissage, il aurait été intéressant de proposer une randomisation de l'ordre de ces exercices, dont l'action aurait permis de moyenniser les résultats du biais d'apprentissage sur chaque question.
4. Certains participants ont exprimé une comparaison avec des méthodes statiques de visualisation des incertitudes, affirmant qu'ils préféreraient utiliser des méthodes statiques que des méthodes par animation. Il aurait été intéressant de comparer ces deux approches de manière plus rigoureuse au sein de l'étude sur des utilisateurs, afin de confirmer ou d'infirmer les intuitions de ces participants.
5. Chaque groupe présentait un faible nombre de participants, ce qui accroît la variabilité des résultats obtenus et affecte les conclusions de cette étude, les rendant moins significatives.

3.4.5 Conclusions

L'étude sur des utilisateurs des méthodes par animation met en lumière une faible aptitude à communiquer des informations *quantitatives* sur les états possibles d'un modèle, particulièrement en ce qui concerne les techniques par animation continue. En revanche, les méthodes par animation semblent mettre en valeur des informations *qualitatives* de manière bien plus efficace.

De manière générale, les techniques d'animation pondérée semblent obscurcir le message porté par l'animation. Cette observation doit cependant être contrastée par les conditions de l'étude, qui limitent la confiance pouvant être accordée à ce résultat du fait d'importantes erreurs d'estimation dans le cas de l'animation continue et d'une sur-simplification de l'exercice dans le cas de l'animation standard discontinue.

Enfin, l'ordre dans lequel sont organisées les réalisations paraît avoir une influence significative sur la qualité de la perception des informations. L'étude a montré qu'il est préférable de classer les réalisations par familles, des variations de l'information à basse fréquence étant plus aisément interprétables que des variations à haute fréquence.

3.5 Application aux incertitudes pétrophysiques

Cette section présente deux exemples d'applications d'algorithmes de visualisation des incertitudes, respectivement dédiés aux incertitudes statiques (porosité, perméabilité, propriétés géomécaniques, *etc*) et aux incertitudes dynamiques (pression, saturation fluide, *etc*), afin d'illustrer des utilisations typiques de ces algorithmes en géosciences.

3.5.1 Propriétés statiques

Nous illustrons dans cette section une d'application basée les méthodes statiques par motifs répétitifs, dont le but est de fournir un support à la recherche de la position optimale d'un nouveau puits d'évaluation au sein d'un réservoir d'hydrocarbures. Les puits d'évaluation visent essentiellement à collecter davantage d'informations sur une zone méconnue du sous-sol, à un stade de développement immature du réservoir.

Le modèle Nancy1 Nancy1 est un champ de pétrole et de gaz localisé au Moyen-Orient, dont les dimensions couvrent approximativement vingt kilomètres carrés. Les hydrocarbures se sont accumulés dans un piège structural, composé d'un ensemble de couches plissées et bordé par des failles à faible perméabilité. Le réservoir est composé de deux formations distinctes nommées B et W.

La formation B présente un environnement de dépôt de chenaux deltaïques. Elle est composée de nombreux corps sableux empilés dont le contenu en hydrocarbures est important. De par la forte densité des chenaux, la formation B présente une faible hétérogénéité, et par conséquent, un degré d'incertitude spatial limité. La formation W correspond à un environnement de dépôt fluvial. La densité des corps sableux y est bien moindre et les hétérogénéités latérales sont donc élevées, résultant en un degré d'incertitude spatial significatif. Le contenu en hydrocarbures de la formation W est néanmoins estimé entre 10 et 20% des réserves de Nancy1 ; la maîtrise des incertitudes associées au volume poreux de cette formation est donc essentielle pour maximiser la production du réservoir.

Caractérisation des incertitudes Dans cette étude, nous nous concentrons sur la formation W du fait de sa sensibilité aux incertitudes spatiales. La formation W présentant un environnement de dépôt fluvial, nous modélisons les faciès à l'aide de la méthode de simulation objet FluvSim (Deutsch et Tran, 2002), qui est apte à reproduire la géométrie associée aux chenaux. Nous représentons deux faciès différents, les chenaux et la plaine d'inondation. Nous conditionnons la simulation à l'aide des données collectées par trois puits d'évaluation forés préalablement, ainsi que de la connaissance fournie par des analogues pétroliers (densité, sinuosité, hauteur, largeur des chenaux, *etc*).

Nous modélisons ensuite les propriétés pétrophysiques conditionnellement à chaque faciès, en nous basons sur les informations fournies par les échantillons prélevés *in-situ*, notam-

ment en termes de variabilité et distribution des propriétés. La porosité est ainsi considérée comme constante dans la plaine d'inondation ; cependant, les chenaux présentent une variabilité sensiblement plus élevée et leur porosité est donc estimée à l'aide de simulations séquentielles Gaussiennes. Le nombre de réalisations est un compromis entre qualité de l'échantillonnage des incertitudes et coût en puissance de calcul – dans cette application, nous plaçons le curseur à cent réalisations de porosité.

Enfin, nous générons un champ scalaire représentant les incertitudes associées à la porosité, dans le but de simplifier la population des réalisations. Cette approche utilise des métriques d'incertitude adaptées à la population étudiée, en l'occurrence un coefficient de variation.

Recherche de la position optimale d'un nouveau puits d'évaluation Pour maximiser la valeur des informations obtenues, un puits d'évaluation doit idéalement atteindre une zone à fort potentiel en hydrocarbures et dont l'incertitude est élevée. Ces deux caractéristiques peuvent s'avérer conflictuelles – en effet, une zone supposée fortement perméable avec une incertitude élevée peut se révéler être en réalité une zone peu perméable une fois couverte par un puits d'évaluation. Les géologues doivent donc décider s'ils préfèrent forer dans des zones très incertaines ou dans des zones à fort potentiel et modérément incertaines ; le premier choix est généralement intéressant dans les premières étapes du cycle de développement d'un réservoir, où l'acquisition de plus d'informations est critique, tandis que le second choix est préférable lorsque le réservoir devient mature. Nancy1 présente des structures hautement non-linéaires, telles que des chenaux, des failles et des plis, qui ne peuvent être adéquatement analysées à l'aide de trois puits d'évaluation. Il est donc plus important de minimiser l'incertitude à cette étape du développement que de minimiser le risque de forer dans une zone à faible potentiel.

Outre la réduction des incertitudes spatiales, un certain nombre d'objectifs secondaires peuvent être également définis. Par exemple, la localisation du contact entre les hydrocarbures et l'aquifère fournit des indications importantes sur le volume d'hydrocarbures présents dans le réservoir. Par ailleurs, il peut s'avérer intéressant de convertir un puits d'évaluation en puits producteur ou injecteur lors des développements ultérieurs du réservoir.

Nous nous appuyons sur une visualisation du réservoir par motifs répétitifs (Figure 3.33) pour rechercher des positions sub-optimales¹⁹ où forer un nouveau puits d'évaluation. En utilisant les critères définis ci-dessus, nous proposons deux cibles possibles, notées A et B. La cible A se trouve aux frontières d'une zone présentant une estimation moyenne de porosité élevée, de telle sorte que le puits d'évaluation puisse être converti en puits injecteur par la suite. Cette cible est cependant assez profonde, ce qui accroît le risque de

¹⁹Cette étude est réalisée le long d'une surface stratigraphique du modèle et ne garantit donc pas que la position sélectionnée soit optimale, car le comportement du réservoir le long de surfaces stratigraphiques déposées à des temps différents est négligé.

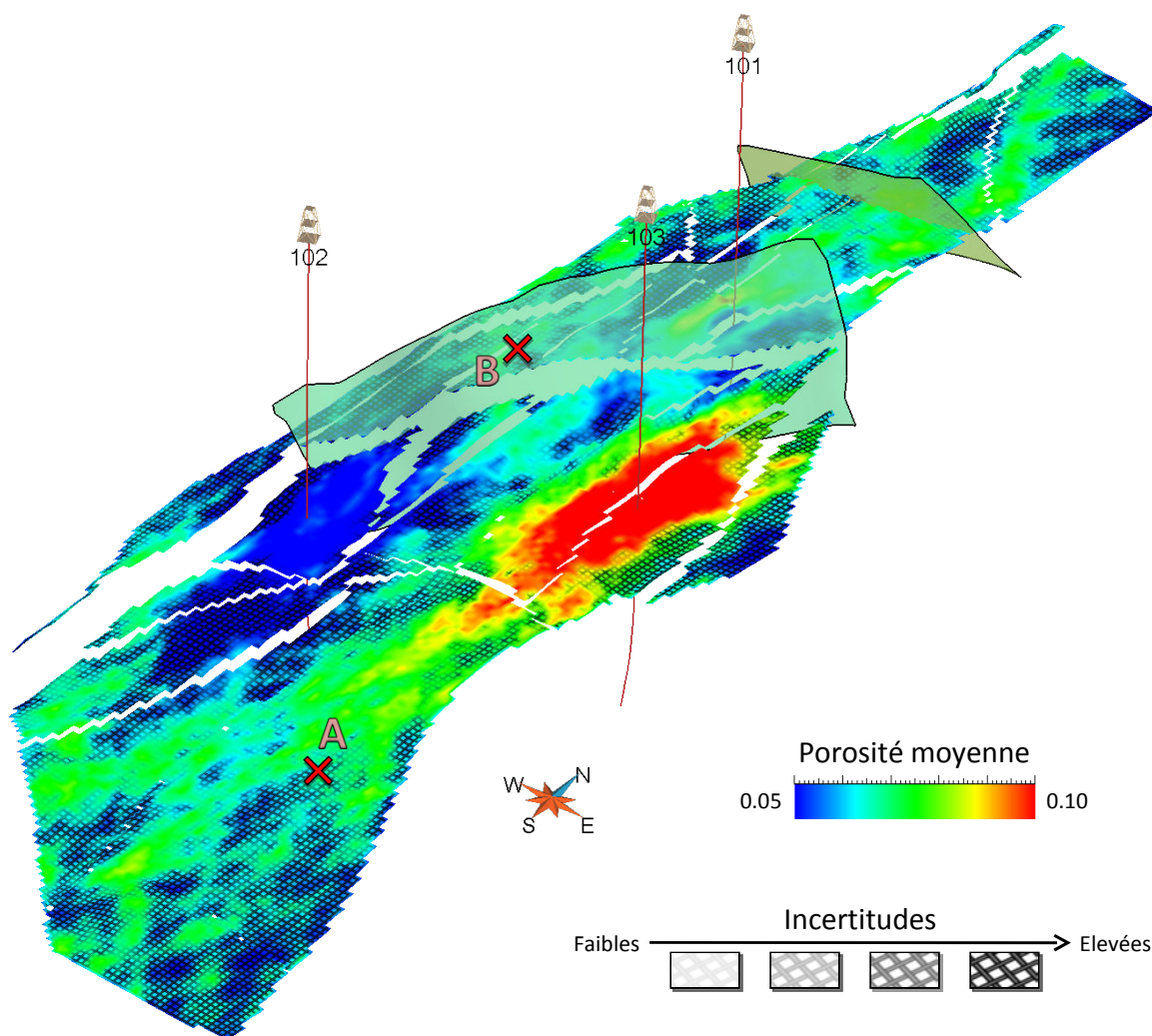


Fig. 3.33 – Visualisation d’un horizon faillé du réservoir Nancy1 présentant à la fois la porosité (couleur) et les incertitudes associées (motif répétitif). Les zones où le motif répétitif est visible présentent une incertitude élevée. Les puits représentés correspondent aux trois puits d’évaluation déjà forés, et les plans semi-transparents aux principales failles internes au réservoir (les failles mineures ne sont pas représentées). Données Total.

rater la zone saturée en hydrocarbures. La cible B est plus élevée et est donc plus à même de se trouver dans la zone saturée en hydrocarbures. Elle présente en outre une estimation moyenne de porosité élevée, de telle sorte que le puits d’évaluation puisse être converti en puits de production par la suite. La production d’un tel puits pourrait cependant être affectée par le comportement de la faille sud, qui est présentement inconnu.

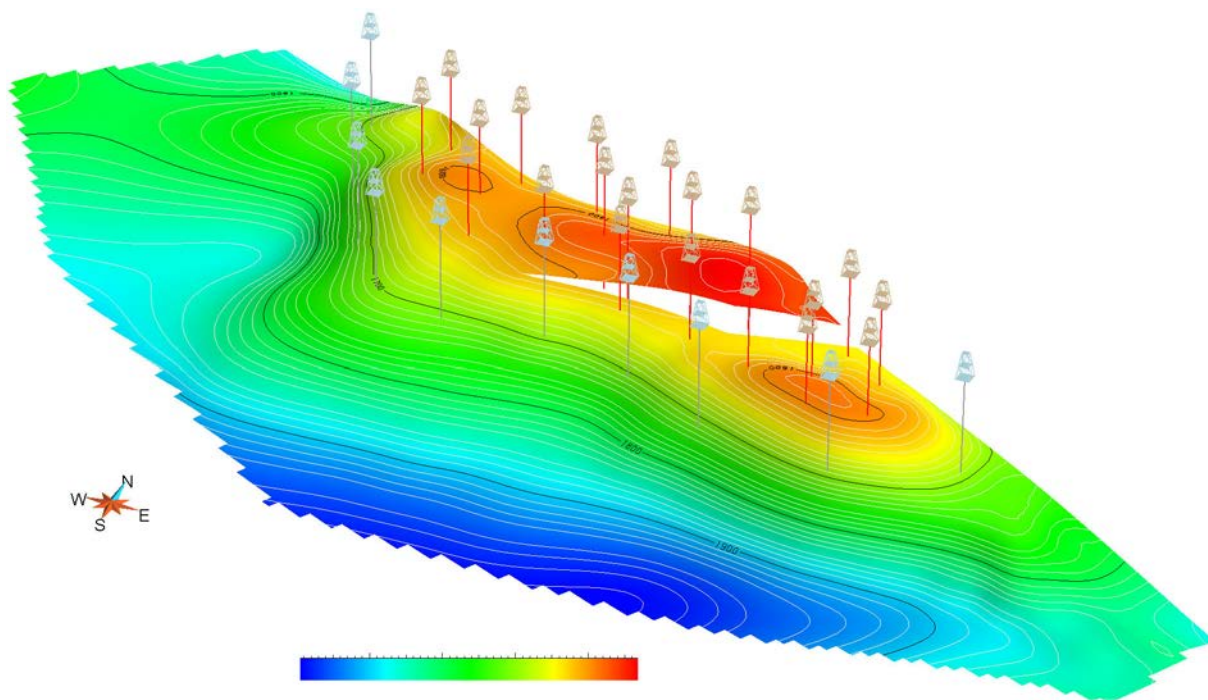


Fig. 3.34 – Visualisation de la structure du modèle Brugge. La couleur de la surface stratigraphique indique l'élévation locale. Les puits producteurs et injecteurs sont respectivement représentés en rouge et en bleu. Données TNO.

3.5.2 Propriétés dynamiques

Nous proposons dans cette section un exemple d'application utilisant conjointement les méthodes statiques par motifs répétitifs et les méthodes par animation, pour dénoter de l'évolution temporelle de propriétés dynamiques conjointement à l'évolution de leur incertitude associée.

Le modèle Brugge Nous avons appliqué nos algorithmes de visualisation aux données de saturation en eau dans le modèle de réservoir Brugge, qui a initialement été développé par TNO à l'occasion du SPE Applied Technology Workshop tenu en 2008 dans la ville de Bruges, Belgique.

Brugge est un modèle synthétique reproduisant la stratigraphie typiquement associée aux champs pétroliers de la mer du Nord. Les hydrocarbures sont situés dans un piège structural mesurant approximativement 10×3 kilomètres, compris entre une structure en dôme allongé et une faille majeure sub-verticale (Figure 3.34, Chen *et al.*, 2009; Denney, 2009). Une autre faille est présente au sein du réservoir et agit comme un drain pour l'écoulement des hydrocarbures.

Le réservoir est couvert par un ensemble de trente puits, dont vingt sont des puits producteurs situés au sommet de la structure en dôme allongé, et dix sont des puits injecteurs situés aux frontières du dôme. Les hydrocarbures présents dans le réservoir sont exclusivement du pétrole ; des simulations d'écoulement ont étudié l'évolution du réservoir sur une période de dix ans, sans qu'aucune apparition de bulles de gaz ne soit observée.

Caractérisation des incertitudes Les incertitudes dynamiques associées à la saturation en eau dans le réservoir ont été échantillonnées à l'aide de trente simulations d'écoulement, chacune utilisant différents paramètres de champs de perméabilité et de contact entre l'aquifère et le pétrole. Nous avons ensuite calculé des métriques d'incertitude à chaque pas de temps de la simulation d'écoulement pour chacune des positions du réservoir, sous la forme du coefficient de variation des saturations en eau simulées. Cette métrique n'est valable que si la saturation en eau moyenne est non-nulle ; cependant, même dans la zone saturée en pétrole, une saturation incompressible de 25% d'eau est observée, garantissant ainsi la validité du coefficient de variation.

Le faible nombre de simulations d'écoulement ne permet qu'un échantillonnage limité de l'espace des incertitudes, et ne peut vraisemblablement pas capturer l'ensemble des évolutions possibles de la saturation dans le réservoir. Cette limitation de l'échantillonnage est due au coût prohibitif des simulations d'écoulement en termes de puissance de calcul, particulièrement lorsque ces dernières sont appliquées à des modèles complexes.²⁰

Résultats La visualisation par motifs répétitifs révèle que la position du front de l'aquifère présente une incertitude élevée au temps initial.²¹ Cette incertitude sur la position du front est directement reliée à l'incertitude sur la profondeur du contact entre l'aquifère et le pétrole et était donc prévisible. Par ailleurs, l'animation de la saturation en eau dans le temps met en lumière l'accélération des digitations du front de l'aquifère dès que ces dernières entrent en contact avec la faille interne. De nouveau, cet effet est une conséquence directe de l'action drainante de la faille.

En étudiant l'animation jointe de la saturation en eau et de l'incertitude associée, un effet de dissipation des incertitudes peut être observé après que les digitations soient entrées en contact avec la faille interne (Figure 3.35). Nous interprétons la diminution locale du degré d'incertitudes comme un effet associé aux failles drainantes en général : dans la mesure où l'action drainante accélère l'écoulement, l'ensemble des simulations d'écoulement

²⁰Dans cette application, les simulations ont été appliquées à une grille stratigraphique comportant 60048 cellules. Approximativement 25% de ces cellules étaient cependant inactives, soit parce qu'elles présentaient un volume nul (cellules dites *mortes*), soit parce qu'elles étaient situées hors du domaine d'intérêt.

²¹Une vidéo illustrant l'évolution temporelle de la saturation en eau conjointement à ses incertitudes associées peut être accédée à l'adresse http://www.gocad.org/www/research/index_gallery.php?image=animated_texture_brugge.gif.

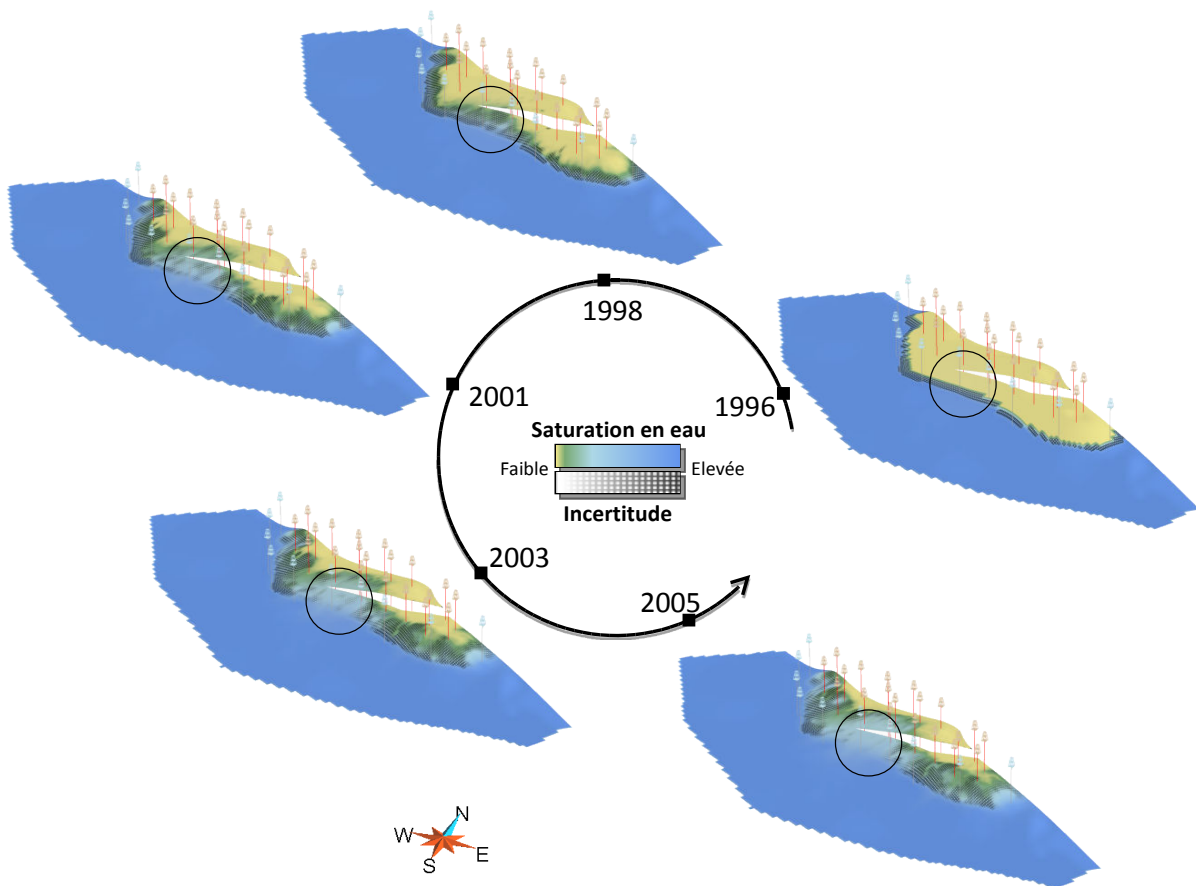


Fig. 3.35 – Evolution conjointe de la saturation en eau et des incertitudes associées dans le modèle Brugge. Le degré d'incertitude local est représenté par l'intensité d'un motif répétitif. L'animation met en valeur la dissipation des incertitudes après que le front de l'aquifère entre en contact avec la faille interne (c'est-à-dire après 1998). Données TNO.

présentent une augmentation locale de la saturation en eau, résultant en une plus faible dispersion locale des valeurs de saturation.

Les informations portées par ce mode de visualisation peuvent s'avérer utiles en première approximation lors d'opérations d'optimisation de la production d'un réservoir, telles que le placement de nouveaux puits producteurs. En effet, elles permettent de percevoir quelles zones sont les plus susceptibles d'être rapidement envahies par l'aquifère (c'est-à-dire peu productives en volume d'hydrocarbures), et quelle confiance peut être accordée à l'arrivée du front à une date donnée. La position d'un puits producteur doit cependant toujours être confirmée par des simulations d'écoulement, car nos algorithmes de visualisation ne peuvent indiquer dans quelle mesure le nouveau puits producteur affectera la production des puits producteurs déjà en place.

Tab. 3.11 – Résumé des méthodes de visualisation implémentées en fonction du support. ✓ : méthode implémentée. ✗ : méthode non implémentée. ✗ : méthode incompatible avec le support.

Méthodes		Sections		Rendu volumique	
		Statiques	Animées	Statique	Animé
Attributs de couleur	Opacité	✓	✓	✓	✓
	Saturation	✓	✓	✓	✓
	Valeur	✓	✓	✓	✓
Fonction de transfert 2D		✓	✗	✓	✗
Surcharge d'informations	Glyphes	✓	✓	✗	✗
	Motifs répétitifs	✓	✓	✓	✓
	Isovaleurs	✓	✗	✓	✗
	Résolution	✓	✗	✗	✗
Ambiguation	Flou	✓	✓	✗	✗
	Bruit	✓	✓	✓	✓
	Distorsion	✓	✓	✗	✗
	Pixellisation	✓	✗	✗	✗
Modification de géométrie	Anamorphose	✓	✗	✗	✗
	Barres d'erreur	✓	✗	✗	✗

3.6 Conclusion

La visualisation des incertitudes pétrophysiques est assimilable à une visualisation des incertitudes sur les attributs d'un modèle, et est donc relativement bien traitée dans la littérature. Cependant, la validation des différentes approches reste encore immature; nous avons adopté dans cette thèse une approche duale de création / validation de nos méthodes de visualisation des incertitudes.

Nous avons dans un premier temps présenté des méthodes de visualisation dites statiques, car elles n'utilisent pas explicitement le temps pour présenter le degré d'incertitude spatial. Nous introduisons dans cette section un certain nombre de visualisations des incertitudes, organisées selon une taxonomie proche de Pang *et al.* (1997). Parmi l'ensemble des techniques implémentées dans notre travail, trois catégories peuvent être dégagées : (i) certaines méthodes sont des contributions nouvelles qui n'ont, à notre connaissance, jamais été présentées dans la littérature, (ii) certaines traduisent des concepts existants appliqués dans un contexte de géomodélisation, (iii) certaines s'appuient sur des méthodes préexistantes et présentent des améliorations favorisant leur utilisation en géomodélisation, et (iv) certaines, enfin, sont une simple application de méthodes préexistantes pour tester leur adéquation à des problématiques de géomodélisation (Tables 3.11 et 3.12).

Une des méthodes de visualisation statique, la visualisation par motifs répétitifs, a fait l'objet d'une étude sur des utilisateurs afin de confirmer sa validité en géomodélisation.

Tab. 3.12 – Résumé des principales contributions pour chacune des méthodes de visualisation implémentées.

Famille	Méthode	Contributions
Attributs de couleur <i>Section 3.1.1</i>	Transparence	Contours suggestifs, rendu volumique
	Saturation	-
	Valeur	-
Fonction de transfert 2D <i>Section 3.1.2</i>		Définition de la fonction de transfert Classification automatique Rendu volumique
Surcharge d'informations <i>Section 3.1.3</i>	Glyphes	Stratégie d'échantillonnage
	Motifs répétitifs	Méthodes par épaisseur et fréquence
		Contrôle des interférences
		Rendu volumique
	Isovaleurs	∅ (réimplémentation)
	Résolution	Maillages adaptatifs Fractale de Koch Dendrites
Ambiguation <i>Section 3.1.4</i>	Flou	Combinaison aux motifs répétitifs
	Bruit	Animation du bruit
	Distorsion	Méthodologie de décalage des valeurs
	Pixellisation	Application aux incertitudes
Modification de géométrie <i>Section 3.1.5</i>	Anamorphose	∅ (réimplémentation)
	Barres d'erreur	∅ (réimplémentation)
Animation <i>Section 3.3</i>		Extension de l'animation probabiliste

Cette étude a révélé que non seulement la présence des incertitudes, mais également la manière selon laquelle les incertitudes sont représentées, pouvait affecter la prise de décision. De manière générale, les cartes adjacentes d'incertitude semblent être de mauvais vecteurs de la notion d'incertitude spatiale ; les cartes coïncidentes, dont le design a été étudié pour répondre au problème spécifique, semblent en revanche présenter un niveau de clarté comparable aux cartes ne présentant pas les incertitudes spatiales, tout en étant plus riches en information.

Nous nous sommes ensuite intéressés aux techniques d'animation, qui peuvent être utilisées (i) pour présenter la variabilité spatiale au sein d'un ensemble de réalisations grâce à la fréquence et à l'intensité des oscillations de la grandeur d'intérêt, ou (ii) pour présenter l'évolution temporelle de propriétés dynamiques conjointement à l'évolution temporelle des incertitudes associées, en combinant l'animation aux méthodes de visualisation statiques. Nous avons notamment présenté la technique d'animation par images clefs, qui

est adaptée à ces deux types d'application. Nous avons en outre proposé une extension aux travaux de [Lundström *et al.* \(2007\)](#), permettant d'étendre la technique de l'animation probabiliste aux animations continues.

Les méthodes par animation ont été validées dans le cadre d'une étude sur des utilisateurs, menée au sein d'un environnement professionnel. Cette étude a montré que l'ordre dans lequel les réalisations sont agencées influence fortement la qualité de la perception ; les réalisations doivent être ordonnées par famille afin de maximiser la compréhension des incertitudes. Le problème du tri par famille reste ouvert dans le cas général ; en effet, le tri à partir d'une fonction objective ayant un sens géologique, telle que la quantité d'hydrocarbures présents dans un réservoir pétrolier, ne se traduit pas nécessairement en une perception similaire de réalisations voisines.

Nous avons enfin proposé une illustration des méthodes de visualisation statique et par animation à des propriétés pétrophysiques statique (porosité) et dynamique (saturation en eau) sur deux géomodèles différents. Ces applications sont présentées comme un exemple d'utilisation des méthodes de visualisation des incertitudes présentées précédemment dans un contexte de géomodélisation.

Chapitre 4

Visualisation des incertitudes structurales en géomodélisation

Sommaire

4.1 Méthodes statiques dédiées - les surfaces fantômes	122
4.1.1 Génération d'un volume enveloppe	123
4.1.2 Population par une métrique d'incertitude	125
4.1.3 Visualisation de la surface fantôme	125
4.1.4 Résultats et discussion	126
4.2 Méthodes dynamiques dédiées - isosurfaces animées	127
4.2.1 Méthodes d'extraction	127
4.2.2 Modélisation de surfaces mourantes	128
4.2.3 Résultats et discussion	131
4.3 Conclusion	132

Le domaine des incertitudes structurales traite des incertitudes sur la géométrie ou la topologie d'un ensemble de surfaces géologiques. Leur visualisation est généralement plus complexe que celle d'incertitudes portant sur les attributs, car la théorie des variables régionalisées ne s'applique que de manière limitée aux incertitudes structurales ; en effet, une même entité (par exemple, un point d'une surface géologique) peut se voir proposer différentes positions spatiales, tandis qu'une variable régionalisée n'est associée qu'à une unique position de l'espace. Pour éviter ce problème, la plupart des approches existantes considèrent la notion de distance à une surface géologique dans l'ensemble du volume d'intérêt plutôt que la position de la surface. De fait, la distance peut être traitée comme un attribut du modèle et supporte donc l'ensemble des opérations mathématiques classiquement utilisées en géostatistiques. Dans ce chapitre, nous présentons deux techniques de visualisation des incertitudes structurales reposant sur la notion de distance : les surfaces fantômes proposent une visualisation statique de l'incertitude, tandis que les isosurfaces animées utilisent des méthodes dynamiques pour présenter le degré d'incertitude spatial.

Il convient de noter qu'en géomodélisation, les incertitudes pétrophysiques sont conditionnées aux structures géologiques ; de ce fait, les incertitudes structurales sont traitées en amont dans la chaîne de modélisation. Nous choisissons néanmoins de présenter la visualisation des incertitudes structurales après la visualisation des incertitudes pétrophysiques dans cette thèse, du fait de la plus grande complexité des algorithmes de visualisation mis en jeu.

Terminologie Différents concepts liés aux incertitudes sur la position d'une surface géologique peuvent aisément être confondus. Ce paragraphe décrit brièvement l'ensemble de ces concepts, ainsi que les notations qui leurs sont associées.

La *surface réelle* \mathcal{S} fait référence à l'état réel de la surface géologique. Cet état ne peut cependant jamais être caractérisé avec certitude du fait d'observations indirectes, imprécises ou parcellaires. Il est donc approché par un ensemble de *surfaces possibles* \mathcal{S}_p , dont les caractéristiques reproduisent aussi fidèlement que possible les propriétés de la surface réelle \mathcal{S} au vu des observations collectées.

L'ensemble de ces surfaces définissent un *modèle de la surface réelle*, noté \mathcal{S}_M . Le modèle peut être approché par un ensemble discret de surfaces \mathcal{S}_p , ou être représenté par un ensemble infini de surfaces \mathcal{S}_p qui définissent alors une fonction de densité de probabilité $P_{\mathcal{S}}(x)$, $x \in \mathbb{R}^3$, le long de trajectoires orientées par la géométrie la plus probable de la surface géologique.

Cette *surface la plus probable* ou *surface initiale*, notée \mathcal{S}_i , correspond au sous-ensemble de \mathbb{R}^3 où la fonction de densité de probabilité $P_{\mathcal{S}}(x)$ présente un maximum local, c'est-à-dire à l'ensemble des positions où la surface réelle \mathcal{S} a la plus grande probabilité de se trouver. La surface \mathcal{S}_i a un sens probabiliste, mais pas nécessairement un sens géologique – en effet, elle permet de minimiser l'écart à la surface réelle \mathcal{S} de manière locale, mais ses caractéristiques d'ensemble (rugosité, forme, dimensions, *etc.*) peuvent être significativement différentes des caractéristiques de \mathcal{S} .

Il est possible de ne pas être certain de l'existence d'une surface géologique \mathcal{S} , notamment lors de l'interprétation de données sismiques bruitées. En ce cas, une *probabilité d'existence* $p(\mathcal{S})$ est introduite ; cette probabilité est propagée à l'ensemble du modèle \mathcal{S}_M afin de refléter les incertitudes sur l'existence de \mathcal{S} .

4.1 Méthodes statiques dédiées - les surfaces fantômes

Les techniques par visualisation de surfaces fantômes nécessitent de connaître la valeur de $P_{\mathcal{S}}(x)$ sur l'ensemble de domaine d'étude. Or, le modèle \mathcal{S}_M est le plus souvent constitué d'un ensemble discret de surfaces possibles \mathcal{S}_p , auxquelles est parfois associé un degré de confiance spatiale sous la forme d'une grandeur scalaire. Nous proposons deux stratégies différentes permettant de convertir ces données 2,5D en une fonction de densité

de probabilité volumique, basées (i) sur une extrusion de surface ou (ii) sur une discrétisation du volume d'intérêt sous forme de grille Cartésienne (Viard *et al.*, 2010c).

Les techniques par visualisation de surfaces fantômes s'inspirent des approches par visualisation d'enveloppes d'incertitude présentées par Johnson et Sanderson (2003). Leur but est de présenter directement la fonction de densité de probabilité $P_S(x)$ sous la forme d'un halo entourant la surface la plus probable \mathcal{S}_i . Johnson et Sanderson (2003) présentent un halo de densité constante ; dans notre approche, en revanche, la densité du halo est fonction de $P_S(x)$. Ces stratégies reposent sur les mêmes étapes clef : un maillage volumique est généré sur l'ensemble du volume d'intérêt, puis peuplé par la fonction de probabilité de densité $P_S(x)$ avant d'être enfin visualisé.

4.1.1 Génération d'un volume enveloppe

Afin de générer la fonction de densité de probabilité $P_S(x)$, nous construisons dans un premier temps un maillage tridimensionnel dont les nœuds serviront de support à $P_S(x)$. Parmi les différents types de support utilisés en géosciences (Caumon *et al.*, 2005), nous générons un maillage non structuré par extrusion de surface, et un maillage structuré par discrétisation en grille Cartésienne.

Approche par extrusion L'extrusion de surfaces génère un maillage tridimensionnel par balayage de la surface initiale \mathcal{S}_i le long de vecteurs d'extrusion. Une cellule prismatique est alors créée pour chacune des facettes de la surface initiale, de telle sorte que la base et l'extrémité du prisme correspondent respectivement à la facette avant et après extrusion (Figure 4.1). Les vecteurs d'extrusion doivent être choisis avec soin, afin d'éviter des phénomènes d'auto-intersection ou de prismes dégénérés par inversion de polarité (Bombarde, 1997; Cognot *et al.*, 2001). En appliquant l'extrusion de part et d'autre de la surface initiale \mathcal{S}_i , un volume enveloppe est ainsi généré.

Approche par discrétisation en grille Cartésienne Dans l'approche par discrétisation en grille Cartésienne, le domaine d'intérêt est représenté par un parallélépipède englobant l'ensemble des positions de \mathbb{R}^3 où la fonction de densité de probabilité $P_S(x)$ n'est pas nulle (Figure 4.2). Des nœuds sont alors générés à intervalles réguliers dans la base définie par ce parallélogramme, et stockés sous la forme d'une grille Cartésienne. La simplicité de cette structure facilite les accès aux données, car il existe une fonction linéaire bijective permettant de passer d'une coordonnée spatiale $x \in \mathbb{R}^3$ à un des nœuds de la grille Cartésienne. Cependant, la résolution fixe de l'échantillonnage peut s'avérer néfaste. En effet, un faible nombre d'échantillons peut produire des approximations grossières de la fonction de densité de probabilité, tandis qu'un nombre élevé d'échantillons est extrêmement coûteux en termes d'espace mémoire. Par ailleurs, de nombreux échantillons sur les bords de la grille couvrent des positions de l'espace où $P_S(x)$ prend des valeurs

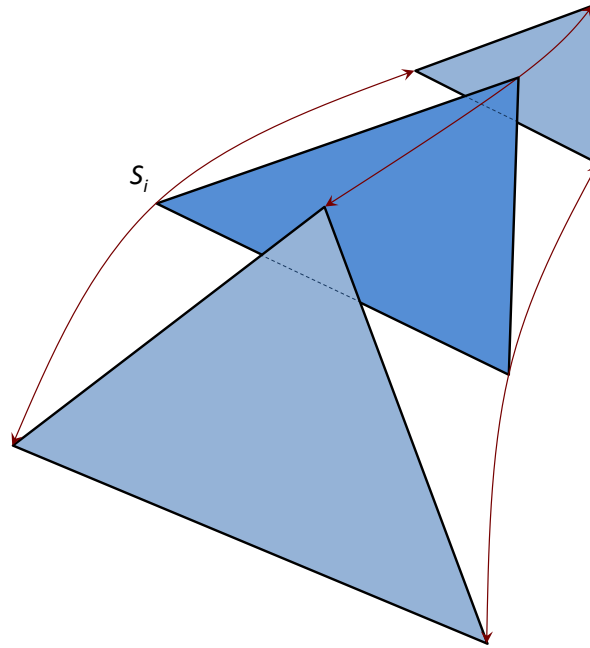


Fig. 4.1 – Extrusion d’une des facettes de la surface initiale S_i . Le triangle bleu sombre correspond à la facette initiale de S_i , et les triangles bleu clair aux facettes extrudées. Les vecteurs rouge sombre représentent les vecteurs d’extrusion.

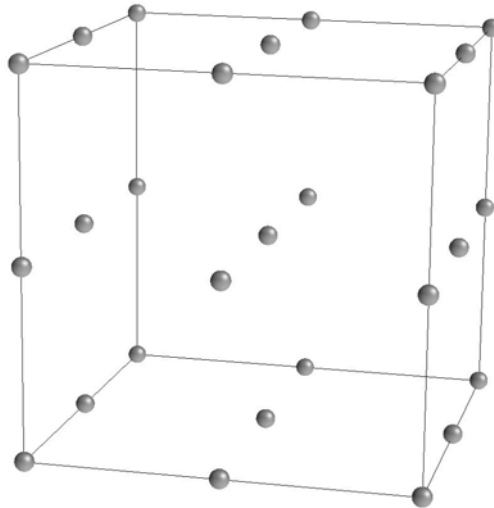


Fig. 4.2 – Discrétisation du volume d’intérêt à l’aide d’une grille Cartésienne. Les sphères grises représentent les points d’échantillonnage.

nulles. Ces échantillons pourraient être ignorés à l’aide de techniques d’élimination des espaces vides (*empty space skipping* en anglais, [Levoy, 1990](#); [Boada et al., 2001](#); [Li et al., 2003](#)).

4.1.2 Population par une métrique d'incertitude

L'extrusion ou la discrétisation en grille Cartésienne permettent de créer un support volumique, qui doit ensuite être peuplé par les valeurs de la fonction de densité de probabilité $P_S(x)$. Si un modèle mathématique représentant la probabilité d'apparition de la surface réelle \mathcal{S} est disponible, la population du maillage est triviale ; cependant, le comportement spatial de $P_S(x)$ est rarement connu analytiquement. Nous reconstruisons donc le champ de probabilités sous la forme d'une fonction du degré de confiance local si ce dernier est connu (dans le cas contraire, nous supposons le degré de confiance local uniforme), ainsi que de la distance $\delta(x)$ entre le point de coordonnées $x \in \mathbb{R}^3$ et la surface initiale \mathcal{S}_i .

Approche par extrusion Lorsque la surface est extrudée, δ correspond à une distance curvilinéaire mesurée le long du champ des vecteurs d'extrusion. Si les surfaces extrudées définissent un volume enveloppe au sein duquel la probabilité d'apparition de \mathcal{S} est égale à $P_{enveloppe}$, alors la fonction de densité de probabilité doit garantir que la surface réelle a une probabilité $P_{enveloppe}$ d'intersecter la trajectoire du vecteur d'extrusion :

$$\int_{\lambda_1}^{\lambda_2} P_S(\mathcal{V}(\delta)) \cdot d\delta = P_{enveloppe} \quad (4.1)$$

Où $\mathcal{V}(\delta)$ correspond à la trajectoire paramétrisée de la ligne tangente au champs des vecteurs d'extrusion, et λ_1 et λ_2 correspondent aux distances le long de la ligne paramétrisée $\mathcal{V}(\delta)$ où se trouvent les deux surfaces extrudées.

Approche par discrétisation en grille Cartésienne Lorsque l'espace est discrétisé sous la forme d'une grille Cartésienne, les vecteurs d'extrusion ne sont pas disponibles. δ peut alors correspondre à une distance euclidienne entre le point x et la surface initiale \mathcal{S}_i . Alternativement, un champ d'orientations peut être défini à l'aide du gradient de distance à la surface initiale \mathcal{S}_i ; δ correspond alors à la distance curvilinéaire mesurée le long de ce champ d'orientations. De manière similaire à l'approche par extrusion (Equation 4.1), la fonction de densité de probabilité doit alors garantir que la surface réelle est rencontrée en parcourant la trajectoire $\nabla(\delta)$ tangente à ce champ d'orientations.

4.1.3 Visualisation de la surface fantôme

De nombreuses techniques de visualisation par rendu volumique sont présentées dans la littérature (Section 2.1). Dans ce travail, nous calculons le rendu de la fonction de densité de probabilité à l'aide d'une approche par tranchage incrémental (Yagel *et al.*, 1996; Westermann et Ertl, 1998), qui permet un affichage de qualité et peut s'adapter à différents types de grilles sans altérations fondamentales de l'algorithme.

La seule différence algorithmique entre les approches par extrusion et par discrétisation en grille Cartésienne repose sur les tables d'index utilisées lors du tranchage des cellules. En effet, les cellules prismatiques et hexaédriques sont topologiquement différentes, notamment du fait d'une dissimilarité du nombre de sommets associé à chacune. Dans les deux cas, les cellules sont tranchées perpendiculairement au point de vue de l'observateur en extrayant les isosurface de distance au point de vue à des pas de distance constants. Cependant, les cellules prismatiques sont tranchées à l'aide une table d'index générée par la technique des *Marching Cells* (Caumon *et al.*, 2005; Buatois, 2008), alors que les cellules hexaédriques utilisent la table d'index associée à la technique des *Marching Cubes* (Lorenson et Cline, 1987).

Pour chaque tranche, l'opacité locale est déterminée en fonction de la valeur de $P_S(x)$. Les tranches sont alors composées entre elles de la tranche la plus lointaine à la tranche la plus proche de l'observateur, de manière à restituer une impression de rendu volumique.

4.1.4 Résultats et discussion

La visualisation des incertitudes structurales par surfaces fantômes présente de manière efficace le sous-domaine de \mathbb{R}^3 dans lequel la surface réelle \mathcal{S} a une probabilité élevée de se trouver (Figure 4.3). Par ailleurs, elle permet aisément de séparer visuellement les zones où la position de \mathcal{S} est relativement bien contrainte, qui présentent une forte opacité et une faible épaisseur, des zones où la position de \mathcal{S} est fortement incertaine, pour lesquelles le halo est plus diffus.

Cependant, cette approche présente un certain nombre de limitations. Tout comme les approches statiques en modélisation d'incertitudes pétrophysiques, le calcul d'une métrique d'incertitude (en l'occurrence, d'une fonction de densité de probabilité) induit une simplification de l'espace des incertitudes. Bien que cette simplification permette d'appréhender plus facilement l'ensemble des configurations possibles, elle peut également provoquer des erreurs de jugement. En effet, si le sous-ensemble de \mathbb{R}^3 où le halo est visible peut accueillir la surface \mathcal{S} , la géométrie de \mathcal{S} au sein de ce volume doit respecter certains critères afin d'être géologiquement réaliste – or, la visualisation du halo masque complètement ce type d'information.

Par ailleurs, la visualisation par surfaces fantômes présente essentiellement les incertitudes *géométriques* associées au modèle, mais ignore une partie des incertitudes sur la *topologie* des surfaces. En effet, seule la probabilité d'existence peut être représentée par le biais d'une transparence accrue du halo. Les incertitudes sur les relations d'intersection entre les surfaces géologiques, notamment dans le cadre de réseaux de failles où les branchements entre failles sont parfois peu visibles sur les images sismiques, ne peuvent cependant être prises en compte par cette technique.

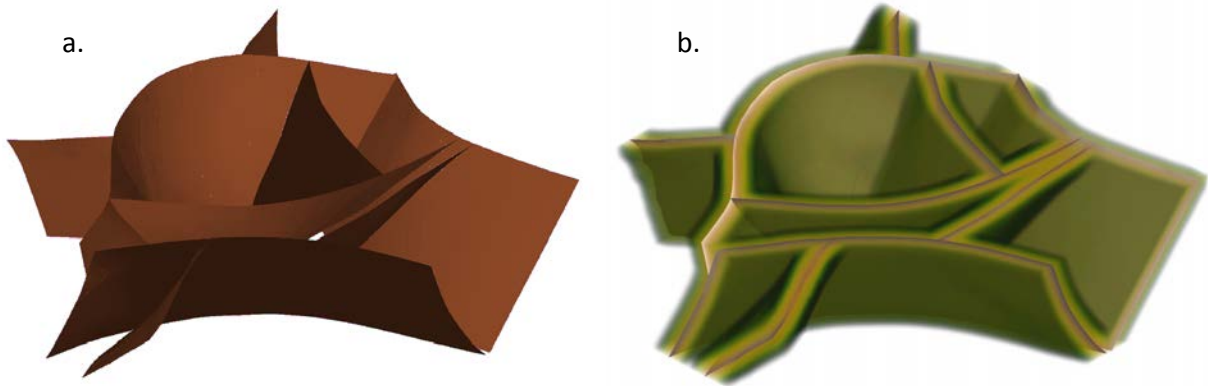


Fig. 4.3 – Visualisation des incertitudes structurales associées au réservoir Mandaros par la technique des surfaces fantômes. (a) Position la plus probable du réseau de failles. (b) Visualisation par surfaces fantômes des positions possibles des failles. Données Total.

4.2 Méthodes dynamiques dédiées - isosurfaces animées

Les méthodes de visualisation des incertitudes structurales par isosurfaces animées s’inspirent des techniques par animation de géométrie (Ehlschlaeger *et al.*, 1996; Brown, 2004) et des techniques de morphing (Buatois et Caumon, 2005). Leur objectif est de présenter de manière séquentielle l’ensemble des surfaces possibles \mathcal{S}_p , avec une animation continue par images clefs entre les réalisations (Viard *et al.*, 2010c).

4.2.1 Méthodes d’extraction

Dans notre approche, la géométrie des surfaces géologiques est représentée de manière implicite sous la forme d’une propriété de distance signée tridimensionnelle (Ledez, 2002; Moyen *et al.*, 2004; Frank *et al.*, 2007). Cette représentation permet de stocker des géométries différentes sur un unique support volumique (typiquement, un maillage tétraédrique), ce qui simplifie grandement les opérations de comparaison et de combinaison entre différentes réalisations. Notamment, l’unicité du support permet de définir une fonction d’interpolation $f()$ agissant indépendamment sur chaque nœud de la grille lors de l’animation par images clefs (Section 3.3.1). L’animation fait ainsi vibrer les différents champs de distances signées représentant les géométries possibles de la surface \mathcal{S} .

La conversion de la représentation implicite à la représentation explicite de la géométrie peut se faire à l’aide de techniques d’extraction d’isosurfaces (Section 2.1) pour chaque image. Un certain nombre d’optimisations possibles sont décrites par Buatois (2008) pour accélérer l’extraction des isosurfaces, généralement basées sur un catalogue des cellules in-

tersectées par l'isosurface à chaque pas de temps. Notre implémentation utilise cependant un algorithme à force brute appliqué directement sur la carte graphique, afin de bénéficier pleinement de la puissance de calcul des GPUs et notamment des shaders de géométrie. En effet, la présélection des cellules requiert des structures de données dynamiques qui sont difficilement implémentables sur GPU à l'heure actuelle. L'alternative serait de proposer une méthode de présélection des cellules sur CPU – l'extraction restant à la charge de la carte graphique – ou un algorithme en deux passes sur le GPU. Dans les deux cas, l'exécution séquentielle du CPU ou des deux passes sur GPU provoquerait une surcharge de calcul lors du changement d'images clef, qui résulterait en un temps de latence sensible lors de l'animation.

4.2.2 Modélisation de surfaces mourantes

Les maillages sur lesquels sont stockés les champs de distances signées sont généralement représentés sous la forme d'un ensemble discret de points ayant des relations de voisinage entre eux. La continuité de ces champs scalaires est reconstruite à l'aide d'une fonction d'interpolation spatiale adaptée au sein de chaque cellule. De ce fait, les isosurfaces sont intrinsèquement continues, à l'exception des zones où le maillage lui-même présente une discontinuité (bord ou discontinuité interne); elles peuvent ainsi traverser l'ensemble du domaine d'intérêt ou former des surfaces fermées. Cependant, cette caractéristique n'est pas adaptée à la représentation de certains objets géologiques; par exemple, les failles sont des surfaces géologiques d'extension finie, de même que les horizons stratigraphiques peuvent être limités par des discordances.

Critère de validité implicite Pour représenter des isosurfaces d'extension finie, nous introduisons un critère de validité spatial indiquant l'existence locale de la surface. Ce critère de validité est ainsi défini sous la forme d'un champ scalaire booléen; il définit une région de \mathbb{R}^3 dans laquelle la surface est physiquement présente. Ce critère permet de respecter les règles d'intersections spécifiques aux objets géologiques, selon lesquelles (i) seules les failles peuvent avoir des bords libres correspondant à des zones de rejet nul et (ii) les horizons sont des surfaces orientées qui ne peuvent se croiser (Figure 4.4, Caumon *et al.*, 2004).

Cascade de conditions Notre implémentation la plus simple du critère de validité spatial définit le domaine d'existence de la surface géologique par demi-plans de valeurs d'une surface implicite limités par une valeur seuil ξ . Cette définition permet de respecter la seconde règle d'intersection de surfaces géologiques selon laquelle un horizon est une surface orientée, la polarité de l'horizon étant définie de manière relative au seuil ξ (Figure 4.5, Algorithme 1).

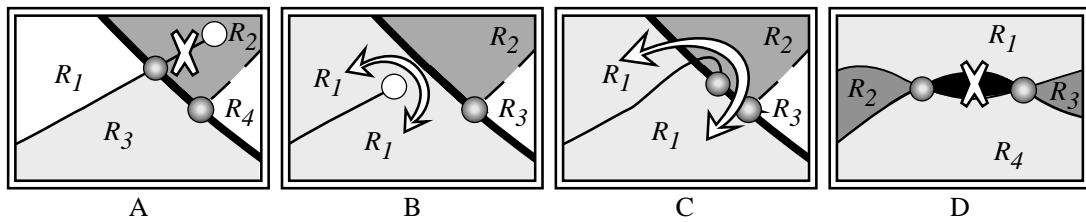


Fig. 4.4 – Exemples d’intersections de surfaces géologiquement incohérentes (d’après Caumon et al. (2004)). (A) Extrémité d’horizon libre traversant une surface de faille. (B) Extrémité d’horizon libre se terminant avant de rencontrer une surface de faille. (C) Croisement entre l’horizon et la surface de faille inversant la polarité de l’horizon. (D) Croisement entre deux horizons.

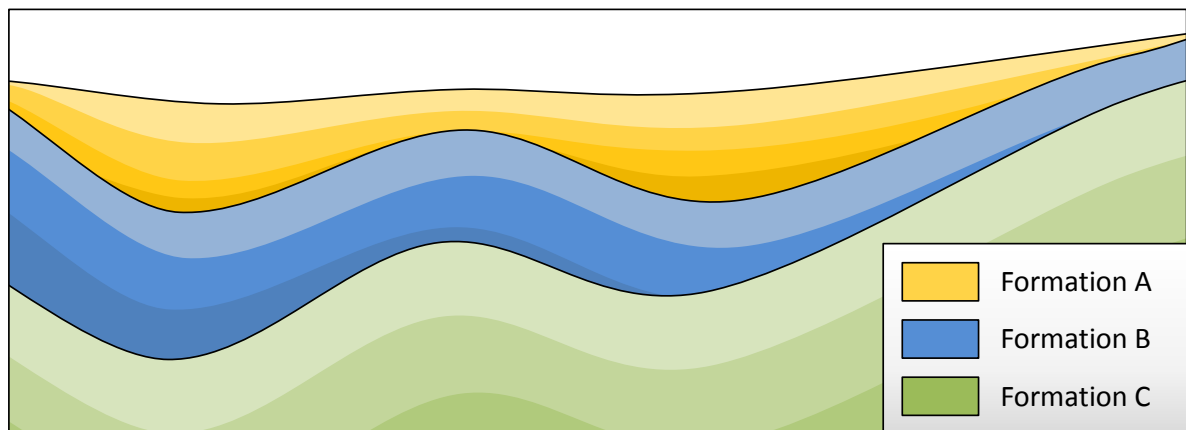


Fig. 4.5 – Illustration du critère de validité spatial par cascade de conditions implicites. Chaque formation est discordante avec les formations supérieure et inférieure. Les couches représentées peuvent être modélisées par la cascade de conditions présentée dans l’algorithme 1.

Cependant, seules des intersections entre horizons peuvent être modélisées de la sorte. Des intersections mettant en jeu une ou plusieurs failles ne peuvent être représentées par cascade de conditions, car la surface géologique affectée par le jeu de la faille principale existe de part et d’autre du plan de la faille (Figure 4.6) – or, chaque niveau de la cascade de conditions exclut un des demi-plans de valeurs de la surface implicite.

Arbre binaire Tertois et Mallet (2006) et Cherpeau et al. (2010) proposent d’utiliser un arbre binaire afin de représenter les intersections entre failles de manière implicite.²² Dans cette approche, chaque nœud de l’arbre représente une faille, définie par un champ scalaire et une isovaleur. Les premier et second descendants de la faille représentent une

²²La définition par cascade de conditions peut être vue comme un arbre binaire dont chaque nœud présente au plus un descendant.

Algorithme 1 Cascade de conditions associées à la figure 4.5.

Requiert: Les champs scalaires $\Psi_A(x)$, $\Psi_B(x)$ et $\Psi_C(x)$ ainsi que les seuils ξ_A , ξ_B et ξ_C permettant de définir leur polarité.

Si $\Psi_C(x) < \xi_C$ **Alors**

Formation = C.

Sinon Si $\Psi_B(x) < \xi_B$ **Alors**

Formation = B.

Sinon Si $\Psi_A(x) < \xi_A$ **Alors**

Formation = A.

Sinon

Formation = \emptyset .

Fin Si

Retourner Formation.

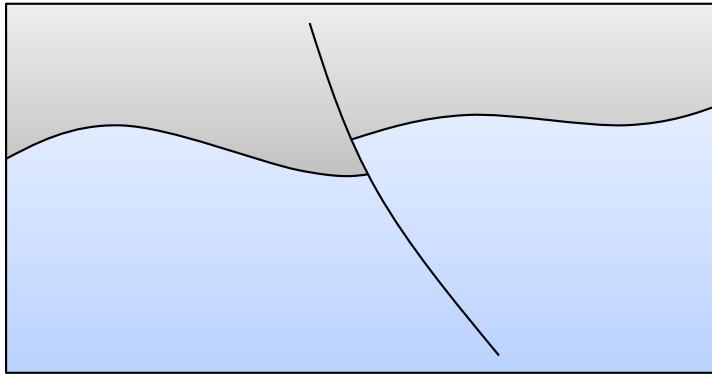


Fig. 4.6 – Exemple de faille inverse simple ne pouvant être représentée par une cascade de conditions implicites.

faille secondaire, qui se branche respectivement sur le côté positif ou sur le côté négatif de la faille principale. Les feuilles de l'arbre (nœuds n'ayant aucun descendant) représentent un bloc de faille au sein du modèle (Figure 4.7).

De manière similaire à la cascade de conditions, l'arbre binaire peut alors être représenté algorithmiquement par une série de tests logiques de type “*Si/Sinon*” (Algorithme 2) qui peut aisément se traduire en code interprétable par la carte graphique.

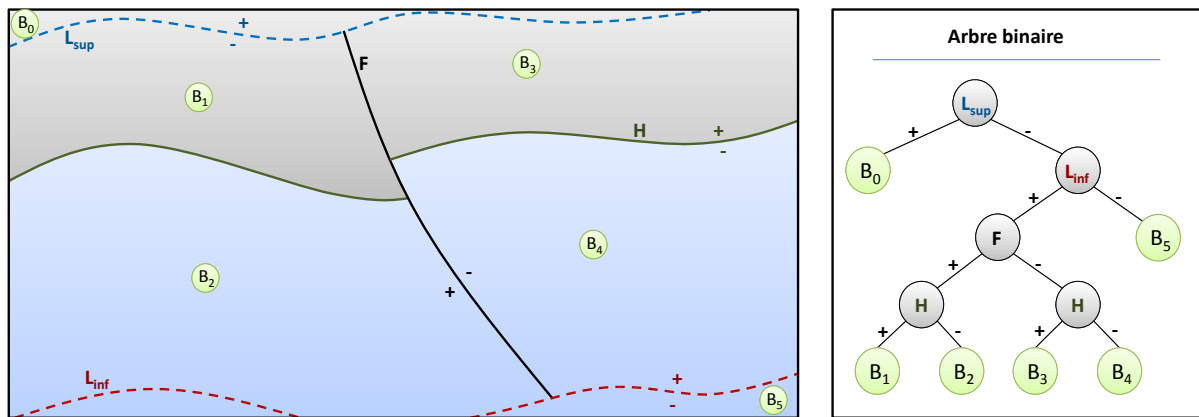


Fig. 4.7 – Représentation d’une faille inverse à l’aide d’un arbre binaire. L’horizon H est affecté par le jeu de la faille F (le rejet est ici représenté en utilisant deux isovaleurs ξ_{H0} et ξ_{H1} de part et d’autre de F , ce qui suppose que l’extension latérale de la faille est infinie). Les surfaces limites L_{sup} et L_{inf} sont purement virtuelles et définissent la limite au niveau de laquelle la faille F présente un rejet nul. L’ensemble de ces surfaces définit six blocs logiques notés B_i , $i \in [0, 5]$.

4.2.3 Résultats et discussion

La visualisation par isosurfaces animées permet de présenter une évolution continue de différentes géométries des surfaces géologiques, ainsi que de visualiser de manière discontinue différentes topologies en passant d’une définition par arbre binaire à une autre.²³

Cependant, cette approche présente un certain nombre de limitations. Tout d’abord, les surfaces géologiques doivent être stockées sur des maillages volumiques, plutôt que sous la forme d’entités surfaciques comme en modélisation structurale traditionnelle. Ce type de support numérique requiert davantage d’espace mémoire du fait de sa plus grande dimension.

Par ailleurs, la position des surfaces géologiques doit être extraite d’après les champs de distance signée à chaque image, ce qui peut être coûteux si un grand nombre de surfaces géologiques doit être représenté. L’extraction de ces surfaces utilise actuellement un algorithme en force brute, parallélisé sur la carte graphique à l’aide d’un shader de géométrie. Buatois (2008) propose un certain nombre de techniques d’optimisation permettant d’accélérer le processus d’extraction, notamment à l’aide de structures de données de type *octree* (espace physique) ou *bucket search* et *Kd-tree* (espace des valeurs), qui pourraient être utilisées pour accroître les performances.

²³Une vidéo illustrant la visualisation des incertitudes structurales par isosurfaces animées peut être accédée à l’adresse http://www.gocad.org/www/research/index_gallery.php?image=animated_level_sets.gif.

Algorithme 2 Traduction algorithmique de l'arbre binaire associé à la figure 4.7.

Requiert: Les champs scalaires $\Psi_H(x)$, $\Psi_F(x)$, $\Psi_{Lsup}(x)$ et $\Psi_{Linf}(x)$ ainsi que les seuils ξ_{H0} , ξ_{H1} , ξ_F , ξ_{Lsup} et ξ_{Linf} permettant de définir leur polarité.

Si $\Psi_{Lsup}(x) > \xi_{Lsup}$ **Alors** {Au dessus de la limite supérieure}

Bloc = B_0 .

Sinon {Sous la limite supérieure}

Si $\Psi_{Linf}(x) > \xi_{Linf}$ **Alors** {Au dessus de la limite inférieure}

Si $\Psi_F(x) > \xi_F$ **Alors** {Du côté positif de la faille}

Si $\Psi_H(x) > \xi_{H0}$ **Alors** {Au dessus de l'horizon}

Bloc = B_1 .

Sinon {Sous l'horizon}

Bloc = B_2 .

Fin Si

Sinon {Du côté négatif de la faille}

Si $\Psi_H(x) > \xi_{H1}$ **Alors** {Au dessus de l'horizon}

Bloc = B_3 .

Sinon {Sous l'horizon}

Bloc = B_4 .

Fin Si

Fin Si

Sinon {Sous la limite inférieure}

Bloc = B_5 .

Fin Si

Fin Si

Retourner Bloc.

4.3 Conclusion

La caractérisation des incertitudes structurales est généralement plus complexe que la caractérisation des incertitudes pétrophysiques, car la théorie des variables régionalisées ne s'applique que de manière restreinte à ce domaine. Nous avons utilisé dans cette thèse un attribut de distance signée indiquant indirectement la présence de surfaces géologiques, en nous basant sur des travaux récents dans ce domaine.

Nous avons dans un premier temps introduit une méthode de visualisation statique des incertitudes, les surfaces fantômes, représentant les incertitudes géométriques associées à la position de la surface sous la forme d'un halo dont l'épaisseur et la densité sont des fonctions du degré d'incertitude local. Cette technique est également à même de représenter les incertitudes topologiques associées à l'existence même de la surface, mais ne peuvent représenter les incertitudes topologiques associées aux interactions avec d'autres surfaces géologiques.

Pour répondre à ce problème, nous avons introduit une seconde méthode de visualisation utilisant une approche par animation : les isosurfaces animées. Cette méthode utilise une représentation de la topologie par arbre binaire, comme présenté dans les travaux de [Tertois et Mallet \(2006\)](#) et [Cherpeau *et al.* \(2010\)](#) ; il est ainsi possible d'exprimer visuellement les incertitudes topologiques en modifiant la définition de la topologie dans le temps, à l'aide d'une animation discontinue. Par ailleurs, il est possible d'animer continument les champs de distances signés afin de représenter l'ensemble des géométries possibles des surfaces géologiques.

Conclusions

Du fait d'observations indirectes, parcellaires ou imprécises, l'état exact du sous-sol ne peut jamais être connu avec certitude (Figure 1). La connaissance du degré d'incertitude local associé aux interprétations des données est donc cruciale lors du processus de prise de décision, car elle permet de déduire la probabilité associée à l'ensemble des cas de figure possibles. Cette information a ainsi un impact significatif sur de nombreuses applications pratiques, telles que l'emplacement où forer un puits producteur en géosciences pétrolières, la validation de l'implantation d'une usine rejetant des polluants en étude de risques, ou l'analyse de la stabilité de terrains en fonction des propriétés géomécaniques des roches dans le génie civil.

En géosciences, de nombreux outils géostatistiques ont été mis en place pour caractériser les incertitudes locales. Le manque d'informations sur le sous-sol permet rarement l'utilisation d'un modèle déterministe de manière fiable; la plupart des approches retenues s'inscrivent donc dans le cadre de la théorie des problèmes inverses, qui consiste à proposer plusieurs états du sous-sol honorant au mieux les observations collectées *in-situ*. Plus précisément, de nombreux algorithmes de simulation stochastique ont été présentés (par exemple, Srivastava, 1995; Goovaerts, 1997; Journel *et al.*, 1998; Caers, 2000; Deutsch et Tran, 2002; Chambers et Yarus, 2006), permettant de reproduire différents critères géologiques (répartition et variabilité des données, organisation des faciès, *etc.*) tout en respectant les observations de terrain à l'aide de méthodes de conditionnement aux données.

Ces algorithmes peuvent être utilisés pour échantillonner l'espace des incertitudes en générant un grand nombre de réalisations. Cependant, la manipulation d'un tel volume de données s'avère en pratique complexe : (i) dans le cadre de processus automatisés, la puissance de calcul nécessaire au traitement des réalisations augmente linéairement avec leur nombre, et (ii) l'interprétation d'un grand nombre de réalisations peut s'avérer impossible pour un géologue du fait de la surcharge d'informations associée aux réalisations multiples. Dans le cadre de cette thèse, nous avons présenté un ensemble de techniques de visualisation des incertitudes permettant de manipuler de tels ensembles de réalisations géostatistiques. Ces techniques s'appuient sur les routines de modélisation d'objets géologiques présentes dans le géomodeleur Gocad; de par leur intégration au sein de ce

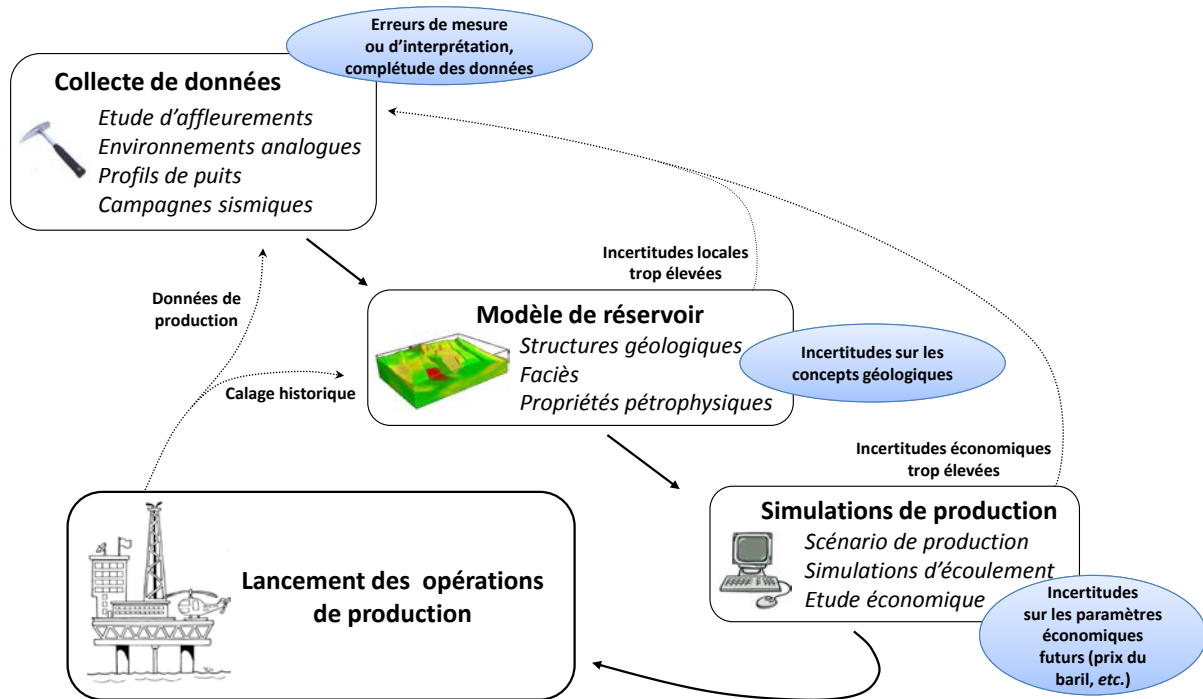


Fig. 1 – Chaîne simplifiée de modélisation d’un réservoir. Les flèches en trait plein indiquent la succession logique des étapes de modélisation. Les flèches en pointillés indiquent un retour possible à une étape précédente. Les bulles bleues indiquent, de manière non exhaustive, les principales sources d’incertitude affectant chaque étape.

géomodeleur, les méthodes de visualisation présentées dans cette thèse sont directement applicables à des problèmes concrets en géosciences.

Lors de la création d’un modèle géologique, les géologues s’intéressent initialement aux structures géologiques rencontrées dans le domaine d’intérêt. Nous avons proposé dans le chapitre 4 deux techniques de visualisation des incertitudes structurales, les isosurfaces animées et les surfaces fantômes. Les surfaces fantômes permettent de visualiser la fonction de densité de probabilité associée à la position d’une surface \mathcal{S} , et représentent ainsi les incertitudes géométriques associées aux surfaces géologiques.

Des travaux de recherche récents ont proposé des approches pour caractériser les incertitudes sur la topologie des surfaces géologiques en plus des incertitudes géométriques (Tertois et Mallet, 2006; Cherpeau *et al.*, 2010). Nous avons présenté des approches par isosurfaces animées, permettant de visualiser de telles incertitudes à l’aide d’une représentation en arbre binaire.

Lorsque les structures géologiques sont correctement caractérisées, les géologues s’intéressent aux faciès et aux propriétés pétrophysiques rencontrés dans les différentes couches du modèle. Les incertitudes associées à ces grandeurs peuvent être considérées comme des attributs du modèle ; de ce fait, de nombreuses techniques de visualisation présentées dans

la littérature s'appliquent à ces types d'incertitudes. Nous avons étudié un certain nombre de techniques dans le chapitre 3. Notre effort de développement s'est ainsi organisé autour de trois axes : (i) l'implémentation de méthodes préexistantes afin d'étudier leur intérêt dans le cadre d'une problématique en géosciences (par exemple, la visualisation par transparence ou par saturation), (ii) le développement d'améliorations ou d'extensions à des méthodes proposées dans la littérature (par exemple, les motifs répétitifs ou le flou), et (iii) le prototypage, développement et test d'idées novatrices telles que les approches par fractales probabilistes, les techniques de pixellisation ou les méthodes de distorsion de texture. L'ensemble de ces outils a été développé directement sur GPU à l'aide d'un langage de shaders ; l'infrastructure logicielle que nous avons adoptée, présentée dans l'annexe A, fournit un cadre unifié à l'ensemble de ces techniques de visualisation, et permet ainsi de les combiner pour visualiser plus de deux grandeurs dans une problématique de corendu.

Par ailleurs, nous nous sommes également intéressés aux mécanismes perceptuels régissant la manière dont une visualisation des incertitudes sera reçue par un observateur. Cet axe de recherche a fait l'objet de deux études sur utilisateurs, portant respectivement sur les techniques de visualisation statiques et les techniques de visualisation dynamiques des incertitudes. L'étude des méthodes statiques a révélé que la manière selon laquelle le degré d'incertitude est présenté influence fortement la perception de l'utilisateur ; les visualisations adjacentes semblent plus aisément interprétables que les visualisations coïncidentes de l'incertitude, car leur format compact minimise la charge perceptuelle et cognitive imposée à un observateur. L'étude des méthodes dynamiques a mis en valeur l'importance de l'ordre selon lequel les réalisations sont séquentiellement agencées lors de l'animation : un agencement par famille minimise les variations temporelles à haute fréquence lors de l'animation et favorise ainsi la perception des données. La notion de famille reste cependant à préciser ; en effet, des critères objectifs de classification reposant sur des grandeurs géologiques, par exemple le volume de sable connecté à un puits producteur, ne se traduisent pas nécessairement en réalisations perçues de manière similaire.

Perspectives

Dans le cadre de notre travail, le degré d'incertitude local est généralement quantifié à l'aide de métriques d'incertitude ; seule la grandeur scalaire produite par cette métrique est alors présentée lors de la visualisation proprement dite. Bien que cette simplification permette de présenter plus aisément les informations d'intérêt à l'observateur, elle peut induire une perte d'information significative sur la distribution locale des données, notamment lorsque cette distribution est fortement non-Gaussienne (par exemple, des distributions de type multimodales). De fait, l'incertitude est *sensus-stricto* portée par la distribution locale des données – le rendu de l'incertitude devrait donc se baser sur le champ de distributions plutôt que sur un champ scalaire. La visualisation de champs de distributions est un problème complexe, actuellement sous-traité dans la littérature. Les

approches existantes incluent l'utilisation d'opérateurs unaires appliqués à l'ensemble des coordonnées spatiales du domaine d'intérêt (Luo *et al.*, 2003), dans une approche similaire à nos métriques d'incertitude, ainsi que l'utilisation de fonctions de distance entre la distribution locale et une distribution de référence (Anderson *et al.*, 2007). Cette dernière approche n'a cependant qu'un potentiel limité en visualisation des incertitudes ; en effet, deux points ayant des distributions fortement différentes peuvent cependant présenter des degrés d'incertitude similaires.

Par ailleurs, notre approche portant sur les incertitudes pétrophysiques se limite à des propriétés d'intérêt représentées sous la forme d'une grandeur scalaire. Or, certaines grandeurs pétrophysiques ou géomécaniques, telles que la perméabilité ou la déformation, sont représentées sous la forme d'un tenseur comportant neuf composantes scalaires. Bien que nous nous soyons penchés sur les problèmes de rendu de tenseurs (Piquet *et al.*, 2010), la visualisation des incertitudes associées à de telles entités multidimensionnelles reste largement inexploree. Les méthodes actuellement disponibles reposent sur des méthodes de visualisation de champs de vecteur ou de tenseurs à l'aide de glyphes, dont la forme permet d'encoder plusieurs caractéristiques (Wittenbrink *et al.*, 1995b, 1996; Cronin, 2004; Harvard CMT, 2005). Ces techniques sont hautement dépendantes de la nature des entités à visualiser, de telle sorte que les glyphes associés à ces entités soient intuitivement interprétables (glyphes en "flèches" pour représenter des incertitudes sur des courants marins, et dièdres droits pour représenter des incertitudes sur un état de contraintes).

Enfin, l'étude empirique des mécanismes de perception permettant d'appréhender l'incertitude est encore immature et demande à être développée davantage. Les domaines d'investigation potentiels incluent notamment (i) l'étude de l'adéquation entre les différentes techniques existantes et les tâches fréquemment effectuées en géosciences – en particulier, la comparaison entre les méthodes statiques et les méthodes par animation lors d'évaluations quantitative et qualitative d'une modèle – (ii) la validation des méthodes d'incertitude dans un contexte pleinement tridimensionnel (Riveiro, 2007), et plus spécifiquement dans le cas de couplages avec un moteur de rendu volumique, et (iii) la validation des méthodes de visualisation des incertitudes structurales (la majeure partie des études existantes se focalisant sur des incertitudes portant sur les attributs du modèle et non sa structure).

Annexe A

Visualisation des incertitudes et programmation sur GPU

Sommaire

A.1 Outils utilisés	140
A.1.1 Shader par sommets	141
A.1.2 Shader de géométrie	141
A.1.3 Shader par fragments	141
A.2 Structure logicielle	142
A.2.1 Organisation d'un shader	142
A.2.2 Exemple	143
A.3 Prototypage de shaders	144

Depuis une dizaine d'années, les cartes graphiques sont devenues programmables grâce à l'arrivée des shaders. Un shader est une série d'instructions exécutable directement sur GPU,²⁴ remplaçant certaines des étapes du pipeline graphique standard dans le but d'apporter une flexibilité de rendu supérieure aux fonctionnalités fixes de ce pipeline (Figure A.1, Macedonia, 2003; Rost, 2006). A l'exception notable de Botchen *et al.* (2005), Zuk *et al.* (2005) et Zuk (2008), peu de travaux se sont explicitement intéressés aux méthodes de programmation sur GPU comme un support à la visualisation des incertitudes. Cette annexe présente nos stratégies pour utiliser de manière optimale les GPUs dans ce but.

²⁴Par abus de langage, le terme "shader" est fréquemment utilisé pour désigner le programme compilé à partir de ces instructions; dans cette thèse, nous parlons de "programme de shader" pour garder une distinction claire entre code source et programme compilé.

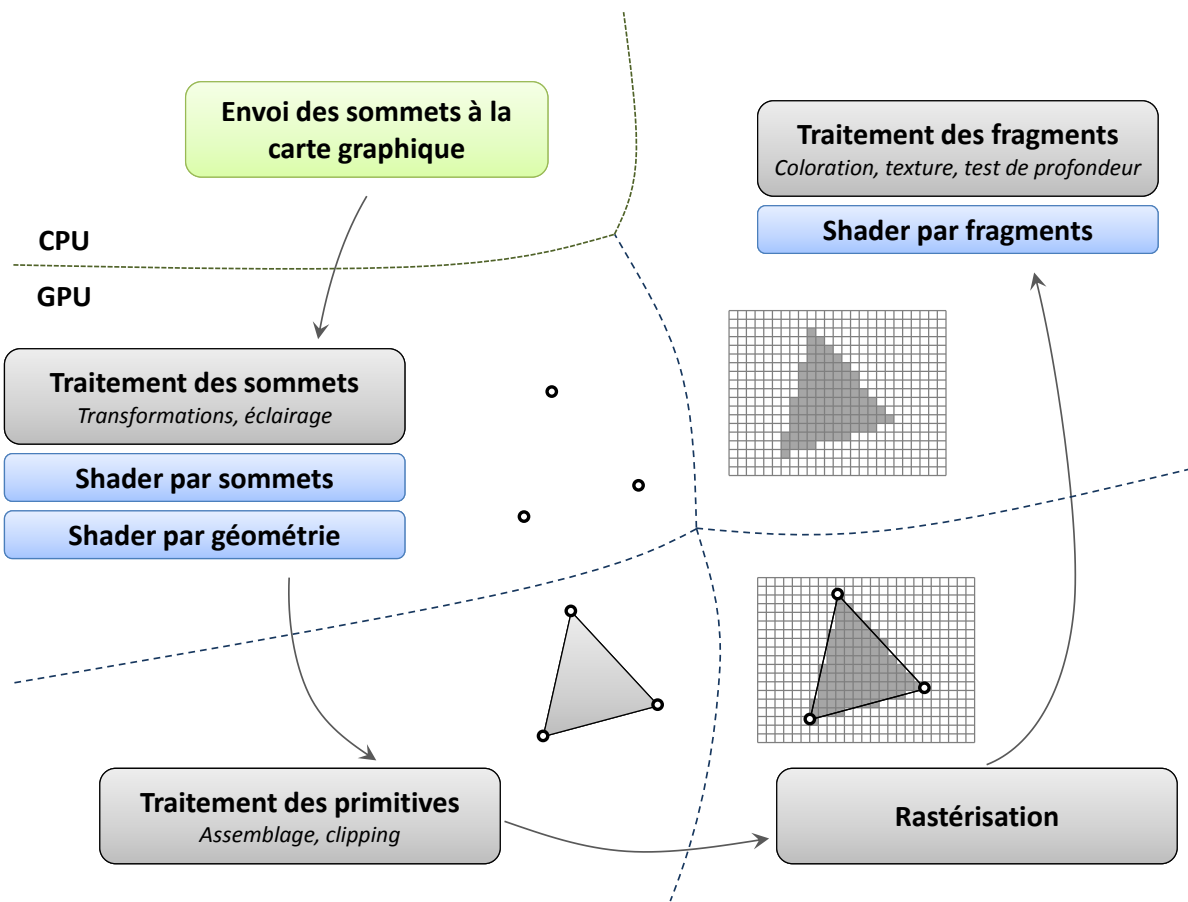


Fig. A.1 – Organisation simplifiée du pipeline graphique. Les encadrés verts représentent les étapes prises en charge par le CPU, et les encadrés gris et bleus représentent respectivement les étapes prises en charge par le GPU à l'aide des fonctionnalités fixes ou à l'aide des unités programmables (c'est-à-dire les shaders).

A.1 Outils utilisés

Il existe différents types de shaders, dont les principaux sont les shaders par sommets et les shaders par fragments. D'autres shaders ont été introduits par la suite – le shader de géométrie, et plus récemment le shader de calcul et les shaders de tessellation (enveloppe et domaine) introduits avec l'API DirectX 11 de Microsoft (Gee, 2008). Bien qu'ils gèrent des opérations très différentes, ces shaders reposent tous sur une architecture commune intimement liée au fonctionnement des cartes graphiques : leur code est conçu pour une exécution en parallèle sur le modèle Simple Instruction Multiples Données (SIMD, Flynn et Rudd, 1996). Cette architecture permet l'écriture de code générique appliqué de la même manière à l'ensemble des données, tout en offrant des capacités de calcul haute performance (Figure A.2, Buatois, 2008). Nous décrivons ci-dessous les trois types de shaders utilisés dans cette thèse : les shader par sommets, par fragments et de géométrie.

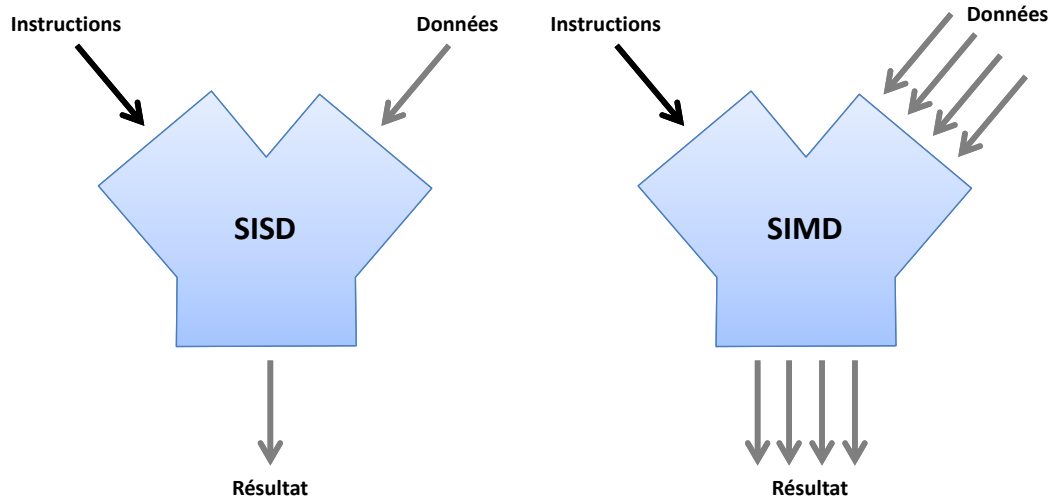


Fig. A.2 – Comparaison entre les modèles Simple Instruction Simple Donnée (SISD) et Simple Instruction Multiples Données (SIMD).

A.1.1 Shader par sommets

Le shader par sommets se substitue aux étapes de transformations de coordonnées et d'éclairage dans le pipeline graphique (Figure A.1). Il permet une gestion fine de la géométrie depuis le GPU, par exemple pour animer géométriquement des modèles (Beeson, 2004), ainsi que des calculs d'attributs par sommet tels que la couleur du sommet ou ses coordonnées de texture, mais ne peut pas générer de nouveaux sommets.

A.1.2 Shader de géométrie

Le shader de géométrie intervient immédiatement après le shader par sommets dans le pipeline graphique (Figure A.1); contrairement au shader par sommets, il permet la génération de nouvelles primitives graphiques directement sur GPU. Le shader de géométrie a des applications intéressantes dans les cas où le nombre de primitives graphiques n'est pas fixé, par exemple lors de l'extraction d'isosurfaces (Tatarchuk *et al.*, 2007), ou dans le cadre de tessellations adaptatives selon le niveau de détail voulu (Boubekeur et Schlick, 2007; Lorenz et Döllner, 2008). Il est également possible de l'utiliser de manière récursive grâce au mécanisme de feedback : les primitives générées par le shader de géométrie peuvent être renvoyées au shader par sommets pour une nouvelle itération (Brown, 2008).

A.1.3 Shader par fragments

Le shader par fragments, également appelé shader par pixels, intervient lors du calcul de la couleur et de l'application des coordonnées de texture des fragments dans le pipeline

graphique (Figure A.1). Cette étape permet notamment d’accepter ou de rejeter un fragment, ainsi que de moduler ou d’imposer sa couleur selon des règles personnalisées. Cette dernière fonctionnalité a été largement utilisée lors de cette thèse, la modulation de la couleur se faisant selon le degré d’incertitude locale.

A.2 Structure logicielle

Outre leur exécution sur le modèle SIMD, les shaders se distinguent également par leur flexibilité de compilation et d’édition de liens : ces deux opérations peuvent être menées à tout moment lors de l’exécution d’un programme sur CPU, permettant ainsi l’évolution dynamique du rendu selon les choix de l’utilisateur. En pratique, la compilation et l’édition de liens sont faites à la volée par le pilote de la carte graphique (Rost, 2006). Ce mécanisme offre des possibilités similaires à celles d’un langage interprété, sans nécessiter de compromis en termes de rapidité d’exécution. De ce fait, il est également aisé d’utiliser des méthodes de méta-génération du code, c’est-à-dire que le shader peut être créé de toutes pièces pendant l’exécution d’une application sur CPU (Folkegård et Wesslén, 2004; Coleman *et al.*, 2007). Cette propriété a été largement utilisée dans le cadre de cette thèse afin de fournir un environnement structuré de programmation sur GPU, permettant ainsi de combiner entre elles l’ensemble des techniques de visualisation présentées dans le chapitre 3. Cette section présente la hiérarchisation générale d’un shader telle que définie dans nos travaux.

A.2.1 Organisation d’un shader

Dans nos travaux, les shaders sont générés par concaténation d’un ensemble d’*éléments de code source*, eux-mêmes composés par un ensemble de *composantes*. Un élément de code source correspond typiquement à un concept graphique, tandis qu’une composante correspond à une action spécifique nécessaire à la traduction de ce concept à l’écran, encodée à l’aide d’une ou plusieurs instructions. Par exemple, l’affichage d’un motif répétitif avec une intensité variable (voir le paragraphe 3.1.3) sera gérée à l’aide d’un unique élément de code source. Cet élément contiendra lui-même plusieurs composantes, gérant notamment les accès à la couleur locale du motif répétitif et au degré d’incertitude locale, ainsi que le calcul de la modulation de la couleur en fonction de ces deux paramètres.

Ordre des composantes Les composantes d’un élément de code source doivent naturellement être effectuées dans une séquence bien précise pour parvenir au résultat désiré. Nous avons initialement envisagé un système de priorité par composante permettant de gérer l’ordre dans lequel ces dernières s’agencent. Ce système a cependant été rejeté car le mécanisme de priorité ne reflétait pas adéquatement la structure même d’un code de

shader et s'avérait donc peu intuitif à utiliser. Pour pallier à cette limitation, nous proposons de lier l'ordre d'exécution des composantes au type d'action qui lui est associée. Nous avons ainsi défini une structure fixe déterminant l'ordre des composantes de la manière suivante :

1. Déclaration et définition des fonctions globales (potentiellement partagées avec d'autres éléments de code source) ;
2. Déclaration et définition des fonctions locales²⁵ ;
3. Déclaration et définition des variables globales (potentiellement partagées avec d'autres éléments de code source) ;
4. Traitement des coordonnées de texture ;
5. Déclaration et définition des variables locales ;
6. Traitement de la couleur associée à un fragment.

Ce cadre fixe offre une plus grande simplicité d'utilisation que le système de priorité, mais peut s'avérer limitant en termes de flexibilité de programmation.

L'assemblage final des composantes implémentées par l'ensemble des éléments de code source est alors fait sous la forme :

```

1 1. Déclaration et définition des fonctions globales.
2 2. Déclaration et définition des fonctions locales.
3 3. Déclaration et définition des variables globales.
4 void main() { // Fonction principale du shader.
5     4. Traitement des coordonnées de texture.
6     5. Déclaration et définition des variables locales.
7     6. Traitement de la couleur.
8 }
```

Dans le cas où plusieurs éléments de code source implémentent des composantes de même type, donc potentiellement concurrentes les unes avec les autres, ces composantes sont concaténées selon l'ordre dans lequel les éléments de code source sont présentés.

A.2.2 Exemple

Ce paragraphe présente un exemple de méta-génération de shader selon la structure proposée dans la section A.2.1. Supposons que deux méthodes de visualisation doivent être utilisées simultanément, par exemple la visualisation des incertitudes par transparence et par saturation (voir le paragraphe 3.1.1). Chaque méthode sera représentée par un élément de code source qui lui est propre (Table A.1).

Lors de l'assemblage final du shader, on détecte que ces deux éléments présentent une variable globale partagée (le degré d'incertitude locale) et deux composantes concurrentes

²⁵La notion de fonction locale n'étant pas nativement supportée par le langage GLSL, chaque fonction locale doit avoir un nom unique qui n'est redéfini par aucun autre élément de code source.

Tab. A.1 – *Éléments de code source associés à la visualisation des incertitudes par transparence et par saturation. Les composantes sont présentées en pseudo-code plutôt que dans le langage GLSL.*

Transparence	
Fonctions globales	\emptyset
Fonctions locales	\emptyset
Variables globales	Incertitude
Coordonnées de texture	\emptyset
Variables locales	\emptyset
Couleur	Modulation de la transparence

Saturation	
Fonctions globales	Routines de conversion de la couleur des espaces RGB à HSV
Fonctions locales	\emptyset
Variables globales	Incertitude
Coordonnées de texture	\emptyset
Variables locales	\emptyset
Couleur	Transformation de la couleur de RGB vers HSV Modulation de la saturation Transformation de la couleur de HSV vers RGB

de type “Traitement de la couleur”. Supposons que la visualisation par saturation soit présentée avant la visualisation par transparence. Le traitement de la couleur sera alors composé des instructions de la composante associée à la saturation, puis des instructions de la composante associée à la transparence :

```

1 Déclaration et définition des routines de conversion de la couleur entre les espaces ↔
  RGB et HSV.
2 Déclaration et définition de la variable globale "Incertitudes".
3 void main() { // Fonction principale du shader.
4     Transformation de la couleur de RGB vers HSV.
5     Modulation de la saturation.
6     Transformation de la couleur de HSV vers RGB.
7     Modulation de la transparence.
8 }
```

A.3 Prototypage de shaders

Les possibilités de compilation et d'édition de liens des shaders facilitent grandement le prototypage rapide de nouvelles méthodes de visualisation, typiquement au sein d'environnements nommés designers de shader. Les designers de shader sont des programmes

CPU qui prennent en charge le plus gros du travail graphique – chargement des éléments de la scène graphique, envoi des primitives au GPU, création d’une fenêtre d’affichage, *etc.* – laissant ainsi le développeur se concentrer uniquement sur l’écriture des shaders. Dans le cadre de cette thèse, nous avons développé un environnement de designers de shaders au sein du géomodeleur Gocad. Ce designer de shader rend les propriétés pé-trophysiques ou géométriques d’un modèle géologique, ainsi que les champs d’incertitude associés, accessibles aux programmes de shaders en cours d’exécution. Notre designer de shaders supporte les shaders par sommets, les shaders de géométrie et les shaders par fragments sur la plupart des objets géologiques utilisés en géosciences (maillages tétra-édriques, maillages structurés Cartésiens ou curvilinéaires, surfaces triangulées, courbes et nuages de points).

Cet environnement de développement permet de prototyper extrêmement rapidement des algorithmes de visualisation, sans être contraint par le formalisme de la structure de données en éléments de code source et en composantes associées. Il nous a ainsi été possible de recoder un certain nombre de méthodes présentées dans la littérature (la plupart des images présentées dans la section 2.3 ont été générées de la sorte) pour tester leur potentiel dans des problématiques en géosciences, sans pour autant en proposer une implémentation complète avec une interface dédiée.

Annexe B

Espaces de couleur

La couleur peut être représentée sous différents formats de données. La représentation numérique la plus couramment utilisée stocke la couleur sous le format RGBA, c'est-à-dire comme l'ensemble de quatre composantes : l'intensité de rouge (R pour *red* en anglais), de vert (G pour *green* en anglais), de bleu (B pour *blue* en anglais) et de transparence (A pour α).

Cependant, la modification des attributs de la couleur est peu intuitive dans cet espace ; les artistes utilisent plus fréquemment les notions de saturation et de valeur de la couleur pour obtenir des nuances adaptées. La valeur correspond à la quantité de noir présente dans la couleur, et la saturation correspond à la pureté de la couleur qui est renvoyée par le matériau : une couleur saturée présente une forte chromaticité, tandis qu'une couleur sous-saturée est plus terne – jusqu'aux niveaux de gris dont la saturation est nulle. Ces notions peuvent être représentées numériquement à l'aide des modèles de couleurs HSV, HSL ou HSI,²⁶ inventés par Smith (1978). Dans ses travaux, Smith (1978) présente des algorithmes permettant de convertir la couleur de l'espace de couleurs RGB à l'espace de couleurs HSV (Algorithme 3) et inversement (Algorithme 4).

D'autres modèles de couleurs ont été définis par la suite, tels que les systèmes $L^*a^*b^*$ ou $L^*u^*v^*$ définis par la commission internationale de l'éclairage (CIE), ou encore le système de Munsell (Burt *et al.*, 2005). Ces systèmes reposent généralement sur des estimations empiriques des intensités de saturation et de valeur perçues par le système visuel humain.

²⁶HSV, HSL et HSI sont trois modèles de couleur voisins. Ils définissent la couleur comme la résultante de trois composantes : la *teinte* (*hue* en anglais), la *saturation*, et une métrique de la “brillance” de la couleur, qui est la *valeur* dans le modèle HSV, la *luminance* dans le modèle HSL, et l'*intensité* dans le modèle HSI.

Algorithme 3 Conversion de l'espace de couleurs RGB à l'espace HSV.

Requiert: Les composantes R (rouge), G (vert) et B (bleu) de la couleur

$V \leftarrow \text{maximum}(R, G, B)$

$\Delta \leftarrow V - \text{minimum}(R, G, B)$

Si $\Delta = 0$ **Alors** {Couleur en niveau de gris}

$H \leftarrow 0, S \leftarrow 0$

Sinon {Couleur saturée}

$S \leftarrow \Delta / V$

Si $R = \text{maximum}(R, G, B)$ **Alors**

$H \leftarrow (G - B) / (6 \cdot \Delta)$

Sinon Si $G = \text{maximum}(R, G, B)$ **Alors**

$H \leftarrow (B - R) / (6 \cdot \Delta) + 1/3$

Sinon { $B = \text{maximum}(R, G, B)$ }

$H \leftarrow (R - G) / (6 \cdot \Delta) + 2/3$

Fin Si

Fin Si

Retourner Les composantes H (teinte), S (saturation) et V (valeur) de la couleur

Algorithme 4 Conversion de l'espace de couleurs HSV à l'espace RGB.

Requiert: Les composantes H (teinte), S (saturation) et V (valeur) de la couleur

Si $S = 0$ **Alors** {Couleur en niveau de gris}

$R \leftarrow V, G \leftarrow V, B \leftarrow V$

Sinon {Couleur saturée}

$\lambda \leftarrow \text{partie entière}(6 \cdot H), \lambda' \leftarrow \text{modulus}(\lambda + 4, 6), \lambda'' \leftarrow \text{modulus}(\lambda + 8, 6)$

$\alpha \leftarrow V \cdot (1 - S), \beta \leftarrow V \cdot (1 - 6 \cdot H \cdot S + \lambda \cdot S), \gamma \leftarrow V \cdot (1 + 6 \cdot H \cdot S - (1 + \lambda) \cdot S)$

$\text{Vec} \leftarrow (V, \beta, \alpha, \alpha, \gamma, V)$

$R \leftarrow \text{Vec}(\lambda), G \leftarrow \text{Vec}(\lambda'), B \leftarrow \text{Vec}(\lambda'')$

Fin Si

Retourner Les composantes R (rouge), G (vert) et B (bleu) de la couleur

Annexe C

Publications

Cette annexe présente un pré tirage de l'article accepté dans *Computers and Geosciences*. Les autres publications effectuées lors de cette thèse ne peuvent être présentées dans ce manuscrit du fait de problèmes de confidentialité (articles publiés à l'occasion des Gocad Meetings de 2008, 2009 et 2010) ou de détention des droits (article soumis dans *IEEE Transactions on Visualization and Computer Graphics*).

Adjacent versus coincident representations of geospatial uncertainty: Which promote better decisions?[☆]

Thomas Viard^{a,*}, Guillaume Caumon^a, Bruno Lévy^b

^a*Gocad Research Group, CRPG-CNRS, Nancy Université, Vandoeuvre-les-Nancy, France*

^b*INRIA Lorraine, ALICE team, Villers-les-Nancy, France*

Abstract

3D geological models commonly built to manage natural resources are much affected by uncertainty because most of the subsurface is inaccessible to direct observation. Appropriate ways to intuitively visualize uncertainties are therefore critical to draw appropriate decisions. However, empirical assessments of uncertainty visualization for decision making are currently limited to two-dimensional map data, while most geological entities are either surfaces embedded in a 3D space or volumes.

This paper first reviews a typical example of decision making under uncertainty, where uncertainty visualization methods can actually make a difference. This issue is illustrated on a real Middle East oil and gas reservoir, looking for the optimal location of a new appraisal well. In a second step, we propose a user study that goes beyond traditional 2D map data, using 2.5D pressure data for the purposes of well design. Our experiments study the quality of adjacent versus coincident representations of spatial uncertainty as compared to the presentation of data without uncertainty; the representations' quality is assessed in terms of decision accuracy. Our study was conducted within a group of 123 graduate students specialized in geology.

Keywords:

Uncertainty visualization, user study, perception, transparency

1. Introduction

Spatial uncertainty is present at multiple levels in subsurface studies, from structural interpretation to dynamic flow simulation. In geological modeling, a significant endeavor has been made to sample this uncertainty by producing several possible 3D geological realizations instead of one best – and probably flawed – deterministic model (e.g., Arpat and Caers, 2007; Chambers and Yarus, 2006; Deutsch and Tran, 2002; Hu, 2000). Visualizing such a population of 3D model realizations is paramount for some applications such as assessing risks of water table contamination due to anthropic pollution, studying mechanical rock properties prior to the construction of large buildings (e.g., arch-gravity dams), or targeting of new drillholes to produce/discover natural resources in potentially high-pay areas.

While a large body of work proposes methods to visualize spatial uncertainty, we need to improve our understanding on how uncertainty is perceived and how it affects human decisions (Harrower, 2003). An increasing number of studies address this issue (e.g., Evans, 1997; Leitner and Bittenfield, 2000; Edwards and Nelson, 2001; Aerts et al., 2003; Deitrick and Edsall, 2006); however, these experiments are currently limited to two-dimensional data (e.g. chloropleth maps) and disregard 2.5D and 3D data that are routinely used in geological applications.

Contributions

This work discusses the applications of two different visualization methods for geological issues, i.e., adjacent versus coincident representations of geospatial uncertainty, and evaluates their relative merits through an empirical user study. In this paper:

[☆]Code available from server at <http://www.gocad.org/www/research/freesoftware.php>.

*Corresponding author. Gocad Research Group, CRPG-CNRS, Nancy Université, Vandoeuvre-les-Nancy, France. Tel.: (+33) 03.83.59.64.37; Fax: (+33) 03.83.59.64.60. E-mail address: viard@gocad.org

Table 1: Uncertainty typology. Adapted from Thomson et al. (2005).

Uncertainty type	Description
Accuracy / error	Deviation from the actual value
Precision	Resolution of the measurement
Completeness	Amount of missing information
Consistency	Coherency between pieces of information
Lineage	Information history, from collection to presentation
Currency / timing	Extent to which pieces of information are up-to-date
Credibility	Reliability of the source
Subjectivity	Amount of human judgments in the information
Interrelatedness	Independence between pieces of information

1. We review a taxonomy of the multiple concepts encapsulated in the term “uncertainty” (Section 2.1), and discuss which strategies have been adopted so far to represent these concepts in a meaningful way (Section 2.2). We further review studies about the effects of uncertainty visualization on decision making (Section 2.3);
2. We present an example of uncertainty visualization method (Section 3) with its use for a geological issue where uncertainty visualization can add value to decision making, i.e., the quest for the optimal location of a new appraisal well in a hydrocarbon reservoir (Section 4). This issue is illustrated on a real middle-east oil and gas field. The source code and executables can be freely downloaded at our website;
3. 123 graduate students in geoscience participated in a formal study to evaluate adjacent versus coincident uncertainty representations of 2.5D pressure data. The protocol of our study is presented in Section 5, and results and implications are discussed in Section 6.

2. Background

2.1. Preliminary notions

What is uncertainty? Uncertainty is a complex, multi-faceted concept that can affect different parameters of a study for different reasons. Most authors agree to see uncertainty as a metadata representing a lack of knowledge about a model or a piece of information, but they encompass this characteristic into several different notions such as error, accuracy, precision, completeness, volume support, *etc.* Some authors even consider uncertainty as a component of a broader concept, termed “data imperfection” according to Gershon (1998).

Past research has proposed several formalizations of the notions related to uncertainty. The US Geological Survey (USGS) propose a Spatial Data Transfer Standard (SDTS) in which they include notions of positional accuracy, attribute accuracy, logical consistency, completeness and lineage (USGS, 1977). Thomson et al. (2005) extend the SDTS with a typology of uncertainty that also integrates the concepts of currency, credibility, subjectivity and interrelatedness (see Table 1 for a description of their typology) – the appropriateness of their typology, revised after MacEachren et al. (2005), has been empirically assessed and refined by Roth (2009a) in the context of floodplain mapping under uncertainty. After discussions with experts involved in various scientific domains, Skeels et al. (2008) propose a hierarchical typology of uncertainty, with measure precision on the lowest level, completeness on the middle level and inferences on the highest level; they find that notions of credibility and disagreement are transversal to these three levels.

Uncertainty quantification. Even if uncertainty were a clearly defined concept, the problem of its characterization would stand still. Several impediments limit uncertainty characterization (Buttenfield, 1993); for example, one participant interviewed by Skeels et al. (2008) highlighted that he may not be aware of the presence of uncertainty (which he called *unknown unknowns*). Being aware of the presence of uncertainty is often not enough – in many scientific applications, there is a need to know the amount of uncertainty associated to the data. This is frequently related to the notion of error; however, the error compares a measured value to the actual value, which is often unknown (termed as *known unknowns* in Skeels et al. (2008), i.e. uncertainty you are aware of, but that can not be measured).

Table 2: Visual variables proposed by Bertin (1983), MacEachren (1992) and McGranaghan (1993).

Visual variable	Description
Location	(x,y) position of an element on the visual plane
Size	Dimensions of an element
Value	Local amount of black that is perceived
Grain	Local variation in the scale of the elements
Color	Local hue and saturation
Orientation	Local angle of the elements
Shape	Combination of size and orientation
Focus	Power of attraction of an element to the eye
Realism	Perceptual similarity of an element to a real-world object

2.2. Uncertainty visualization

Theoretical framework

In uncertainty visualization, data are at least bivariate – consisting of a primary attribute and its associated uncertainty degree – and may be multivariate if either primary data or uncertainty are not defined as scalar values; such a high data-dimensionality calls for appropriate visualization methods (Rheingans and Landreth, 1995). A wide variety of visualization algorithms have been used to depict local uncertainty together with the value of interest in various fields, including GIS, meteorology, oceanography or medical research (Griethe and Schumann, 2006; Pang, 2006; Johnson and Sanderson, 2003; Pang et al., 1997; MacEachren, 1992). Some of these research include *adjacent representations* of uncertainty, i.e., uncertainty map being displayed separately from the primary data, but most of them focus on *coincident representations* of uncertainty, i.e., representations where uncertainty is integrated into the same display as the primary data (MacEachren et al., 1998). Based on the work of Shortridge (1982), MacEachren (1995) further distinguishes between *coincident integral representations* of uncertainty, which directly alter the display of the primary data, and *coincident separable representations* of uncertainty, which allow for selective attention to either variable.

Zuk and Carpendale (2006) evaluate the quality of several uncertainty visualization methods based on the perception criteria developed by Bertin (1983), Tufte (2001) and Ware (2004). They notably work upon Bertin’s visual variables, which are *location*, *size*, *value*, *grain*, *color*, *orientation* and *shape* (Table 2). MacEachren (1992) adds an eighth visual variable which he calls *focus*. He proposes different ways of manipulating focus through contour crispness, fill clarity, fog and resolution. McGranaghan (1993) further expands this set of perceptual variables with the notion of *realism*, which is a significant concern in computer graphics.

Some other authors apply criteria from the perception community to assess the quality of visualization methods – for example, the uncertainty display should not obscure the display of the primary variable; also, visualization should use preattentive processing¹ to be understood at a glance (Healey et al., 1996; Tory and Möller, 2004; Ware, 2004).

Dedicated visualization methods

This section provides a quick review of some existing uncertainty visualization methods. It is not intended as a comprehensive state of the art, but rather as a set of examples of different separable versus integral methods.

Visually separable methods. Cedilnik and Rheingans (2000) overlay distorted wire-frame lines to the visualization as an annotation of the uncertainty degree (Figure 1, a). Their work can be seen as a variation of Wright (1942), who suggests changing the value and grain of contour lines to display contour uncertainty. Using lines results in little of the primary data hidden by the uncertainty display. Another method is the use of patterns, typically with a variable grain or intensity (Figure 1,b), as done by Interrante (2000) along maps, Rhodes et al. (2003) along isosurfaces and Djurcilov et al. (2001) as a post-processing step of a volume rendering algorithm. Their approaches subdivide uncertainty space

¹Preattentive processing is a perception mechanism inherent to our visual system that allows humans to perceive some features of a scene very quickly (less than 200ms according to Healey et al. (1996), which is the time required for eye movement) and without any need to focus on the scene.

into a set of smaller ranges, each of which is assigned a different pattern. This results in a categorization of uncertainty levels.

Visually integral methods. MacEachren (1992) uses focus in geographical information systems (GIS), e.g., by changing the fill clarity of land-cover patterns based on their uncertainty degree (Figure 1, c). Kosara et al. (2001) extend this notion to what they call semantic depth of field (SDOF). While classical image processing blurs objects based on their depth of field, SDOF blurs objects according to their relevance. Unlike most highlighting methods (Robinson, 2006), this de-emphasizes irrelevant objects in order to put the focus on the most important ones.

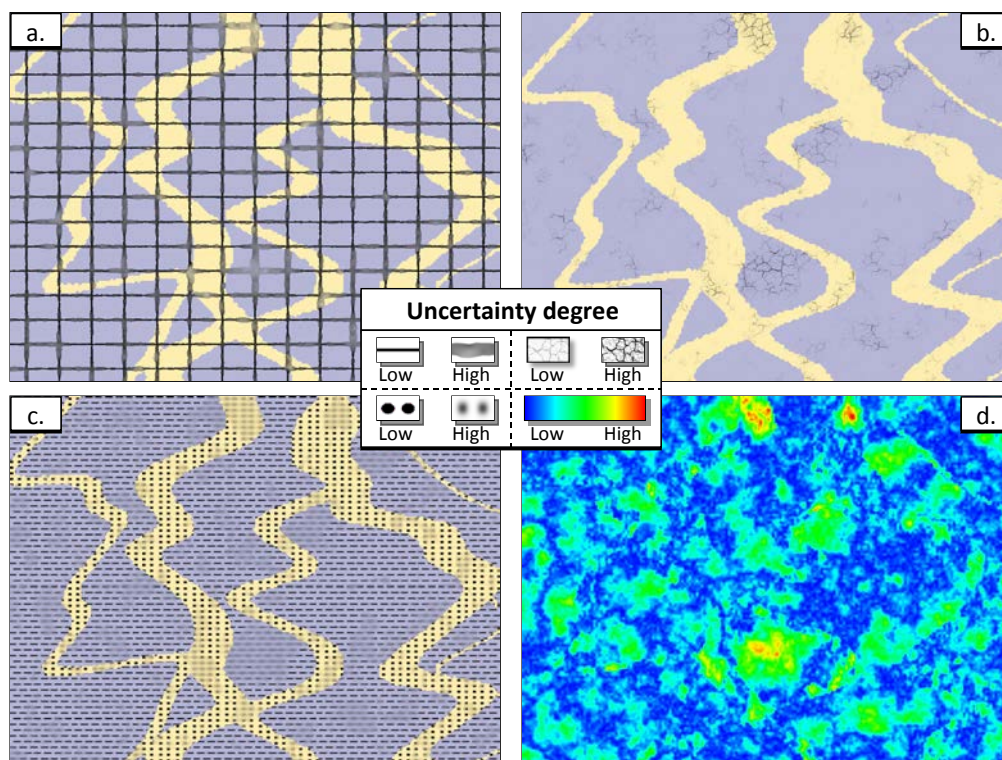


Figure 1: Visualizations of facies uncertainty on a channel dataset. (a) Uncertainty visualization with procedural annotation lines. The line width denotes of the local uncertainty degree. (b) Pattern-based uncertainty visualization with variable intensity. High-intensity areas have higher facies uncertainty. (c) Fill-clarity uncertainty visualization. Blurred patterns indicate high facies uncertainty. (d) Color-coded facies uncertainty map. For figures (a), (b) and (c), yellow denotes sandy areas and blue denotes shaly areas.

Animation methods. Animation methods use time to convey a sense of uncertainty.² Fisher (1993) introduces an animation method to convey a sense of uncertainty in soil maps; he selects patches at random map locations through time and re-evaluates the patch type based on the local soil types distribution. Srivastava (1994a) proposes animation to display a sequence of possible models, using a modified p-field algorithm. The frames of this animation can be ordered according to a relevant objective function (Srivastava, 1994b). Ehlschlaeger et al. (1996) improve the interpolation schemes between keyframes as in the gradual deformation algorithm (Hu, 2000). This preserves the spatial variability of Gaussian random fields in interpolated frames. Davis and Keller (1997) create a patch-based animation of slope stability data. They especially focus on the integration of several sources of uncertainty such as soil type, soil characteristics or elevation values, using fuzzy logic and Monte-Carlo sampling methods. Dooley and

²Note that although animation is an integral method *sensus-stricto*, as the uncertainty display is fully integrated into the display of the primary data, we treat animation methods separately because they typically do not need any explicit measure of uncertainty.

Lavin (2007) propose an animated visualization of plant hardiness zone maps, in which they consider uncertainties due to the data sampling density and to the selected interpolation method. Their approach results in maps being displayed sequentially.

2.3. Decision making under uncertainty

Representation of spatial uncertainty has little interest if it does not affect the way people think about their problems to improve their decisions (Deitrick and Edsall, 2006). Improvements can relate to accuracy, decision speed and confidence in the results (Harrower, 2003). Uncertainty was initially believed to clutter data display, because of the burden introduced by additional information (McGranaghan, 1993). However, studies performed by Leitner and Battenfield (2000) and Aerts et al. (2003) conclude that embedding primary map data with uncertainty tends to actually clarify the display. Their results are consistent with the findings of Evans (1997) and Edwards and Nelson (2001), whose studies showed that users perform better with integrated uncertainty displays than with separate maps. Sanyal et al. (2009) also found that uncertainty perception is not uniform, i.e., that the perception of low uncertainty locations differs from the perception of high uncertainty locations.

Most of the experiments discuss the effects of user expertise on decision-making under uncertainty. There is no clear consensus in the literature whether the level of expertise may bias decision-making under uncertainty; results from Kobus et al. (2001) and Hope and Hunter (2007) suggest that participants' expertise influences their decisions, whereas Evans (1997) and Aerts et al. (2003) find either little or no significant difference between novices and experts. Roth (2009b) argues that these contradictory findings may come from the task complexity – the more complex, the greater the influence of the participant's expertise.

An increasing number of studies focus on the effects of uncertainty visualization for decision-making, but most experiments have been performed with two-dimensional data, resulting in poor knowledge about perception of 2.5D/3D data that are widely used in geology. This work aims at bridging the gaps that exist with such higher dimensional data.

3. Uncertainty visualization design

Our tests on participants and our case study were performed using an implementation of the uncertainty visualization method based on pattern transparency, as described below. Whereas a full-featured software with 3D support has been implemented in a commercial geomodeling software, we provide an open-source package containing the source code, binaries and a technical documentation of our GPU-accelerated implementation at <http://www.gocad.org/www/research/freesoftware.php>.

Uncertainty visualization using pattern transparency. In our experiments, we use transparency to convey a sense of uncertainty, by compositing both the *value* and the *color* visual variables (Zuk and Carpendale, 2006). We blend a predefined pattern to the color-coded data, with a blending ratio expressed as a function of uncertainty. This results in the pattern being visible at some interesting areas; for our problem, we chose to make the pattern visible at high uncertainty areas to bring the attention of participants to these locations, because Sanyal et al. (2009) suggests that uncertainty perception is non-uniform.

The blending ratio can be seen as a measure of the pattern intensity, noted $I(u)$ where u corresponds to some local scalar uncertainty measure. For every single pixel, the blending results in a final RGB color C_f calculated as:

$$C_f = (1 - \alpha_p \cdot I(u)) \cdot C_i + \alpha_p \cdot I(u) \cdot C_p \quad (1)$$

where C_i is the RGB color before applying the pattern, C_p is the RGB color of the pattern and α_p is the pattern opacity at the current pixel.

Pattern definition. Our pattern-based uncertainty visualization method is a variation of Rhodes et al. (2003), for our pattern uses transparency with continuous uncertainty values rather than density with discrete uncertainty levels. Hence, the method removes the burden of defining a discrete set of patterns, i.e., any relevant 2D pattern can be used in our implementation. We have tested several types of pattern empirically, e.g., fabric, grid, circles, chessboard, etc. We generally use “fabric” patterns because they reproduce natural textures (Interrante, 2000) that make the pattern easy to recognize for most users. However, our tests were performed with participants experienced in geomodeling. We therefore used a grid pattern in the user study, which is similar to grid wire-frames that participants were familiar with.

Control on the interference. The degree of interference between the pattern and the background color-coded property can be tuned by replacing α_p in color equation (1) with $\alpha'_p = \omega \cdot \alpha_p$, where $\omega \in [0, 1]$ is a constant user-defined value. If ω is close to 0, the pattern should be hardly visible even at locations of maximum intensity $I(u)$, whereas it should be clearly visible if ω is close to 1 (Figure 2). In our experiments, we use an ω value of 0.8 to prevent the pattern from being over-perceived in the display.

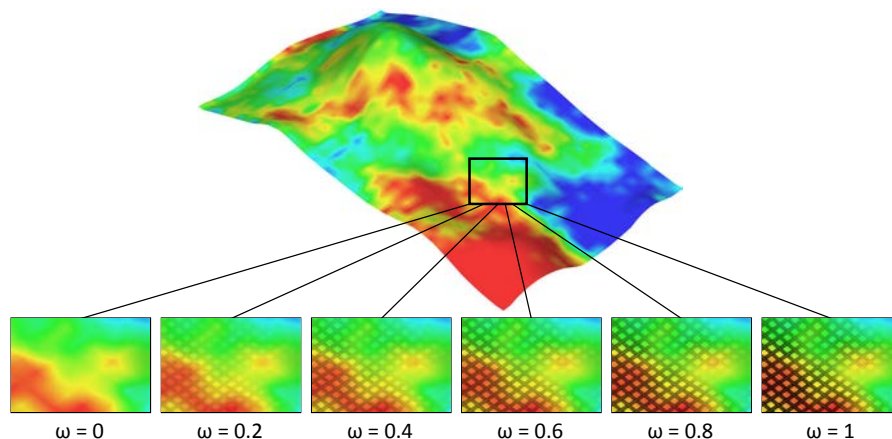


Figure 2: Average porosity map on a stratigraphic layer, blended with “fabric” pattern with several values of ω . Top: whole layer without pattern. Bottom: zooms with different values of ω .

Perception quality. We chose pattern-based uncertainty visualization over other existing methods because of its high quality in terms of perception:

- **Interferences** – Pattern-based uncertainty visualization has the advantage of being perceptually separable from the primary geological property; furthermore, the choice of adapted patterns minimizes the hidden areas of the model, which is likely to result in lower interferences with the underlying color-coded data.
- **Intuitive display** – Transparency has proven to be highly connected to the notion of uncertainty, resulting in an intuitive sign-vehicle for uncertainty (Griethe and Schumann, 2006; Grigoryan and Rheingans, 2004; Johnson and Sanderson, 2003; MacEachren, 1992).
- **Preattention** – Because blending the pattern affects the value of the image, the local levels of uncertainty are likely to be processed preattentively.

4. Uncertainty visualization in geology: an example scenario

Since geoscientists typically face various types of uncertainty, uncertainty visualization algorithms can be valuable in many different issues. In this section, we review one specific issue that frequently happens in the oil and gas industry – i.e., the quest for the optimal location of a new appraisal well (Lafont, 2007) – in order to provide a typical example of decision-making in geosciences. Appraisal wells primarily aim at gathering information about an unknown area of the subsurface, and are typically drilled at the early steps of the assessment of a reservoir. This example is illustrated on the Nan1 reservoir.³ We first describe the geological context of the Nan1 field (Section 4.1) and the uncertainty characterization workflow (Section 4.2), before discussing the well locations which are geologically relevant (Section 4.3).

³Courtesy of Total. For confidentiality reasons, all names and scales related to the reservoir have been modified.

4.1. The Nan1 field

Nan1 is a Middle East onshore oil and gas field, covering approximately twenty square kilometers. Oil is accumulated in a structural trap made of low permeability faults and folded strata. The reservoir is located in two stratigraphic formations, named B and W.

The B formation was formed in a channeled deltaic environment. It shows a lot of sand bodies stacked on each other, with a good potential in oil and gas content. The B formation is quite homogeneous due to the high density of sand bodies (Figure 3), resulting in a low spatial uncertainty degree on the location of sand-rich areas.

The W formation was formed in a fluvial environment. The density of sand bodies is much less favorable and lateral heterogeneity is significant. There is thus a higher uncertainty degree on the location of reservoir rocks. The W formation is estimated to contain between 10 and 20% of the oil in place. A better estimation of the net-to-gross is therefore a worthy challenge, and may help improving the oil production.

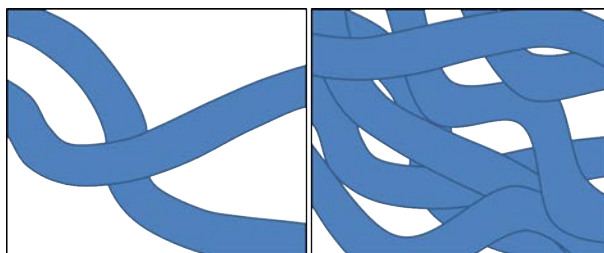


Figure 3: Two sections of a synthetic geological model showing stacked paleochannels with different channel densities. Left: the low channel density results in significant facies heterogeneity as in the W formation. Right: the high channel density results in an almost homogeneously sandy medium as in the B formation.

4.2. Uncertainty characterization

In this study, we focus on the W formation because its evaluation is very sensitive to spatial uncertainty. Because the W formation was deposited in a clastic fluvial environment, we model the facies with the Fluvsim object-based stochastic method (Deutsch and Tran, 2002), which reproduces the geometry of complex channelizing sand bodies (Journel et al., 1998). We represent two different facies, the channels and the flood plain. These sedimentary bodies are conditioned to the observations collected along three existing appraisal wells, and honor analog geological observations (sinuosity, height, width, etc.).

For each rock type simulation, we then simulate the petrophysical properties using statistics inferred from core samples. The floodplain porosity is considered constant, but the channels show a higher variability and are therefore simulated using a stochastic algorithm. In this study, we use sequential Gaussian simulation conditioned to well data. The number of realizations is a trade-off between the accuracy of the uncertainty sampling and the CPU cost of the realizations' generation and storage. In this example, we set the cursor to one hundred realizations.

After the generation of the set of realizations, it is possible to compute some uncertainty metrics at every single spatial location of the model. We use a normalized standard deviation metric on the set of porosity realizations.

4.3. Optimal well location

Appraisal well criteria. To bring as much value as possible, an appraisal well should ideally reach high-pay areas with a high uncertainty. These requirements conflict with each other, since a high-pay area with high uncertainty could actually turn out to be a low-pay area when drilled. Hence, geoscientists have to decide between drilling in an area with a high uncertainty degree where most is unknown, which is typically valuable at the early steps of reservoir characterization, or drilling in an expectedly high-pay zone with moderate uncertainty, which is often preferred when the reservoir is about to enter the production phase. Our visualization methods may be valuable for both approaches, as both require to know about the local uncertainty degree.

The Nan1 reservoir shows complex and highly non-linear structures, such as channels, faults and folds. Only three appraisal wells were drilled so far, which is clearly not sufficient to highlight the properties of the reservoir. Minimizing reservoir uncertainty is thus more valuable than minimizing the risk of drilling in a low-pay area at this step.

Reservoir engineers may be interested in a better estimation of the hydrocarbon reserves, typically by finding the oil-water contact. To avoid future overcosts, they may also consider conversion of appraisal wells later in the development of the field, which requires specific care about the placement of the well. For instance, a water-injection well should be placed so that water could sweep oil toward a production well, which means that there should not be any low-permeability zone in-between. Special attention must be paid to faults, which can act either as drains or seals, and to the elevation of the reservoir top. Injectors should also be located at the boundaries of high-pay areas, so that no oil is lost being swept in areas without production wells.

Uncertainty visualization. We applied our pattern-based uncertainty visualization (Section 3) to a layer of the Nan1 field (Figure 4) in order to determine the location of a new appraisal well. Note that for a truly optimal well location, this study should be performed on all of the layers of the reservoir.

Using the criteria defined earlier, we were able to define two possible well locations in the Nan1 model. Location A is on the boundary of an area with reasonably high expected porosity, so that the appraisal well may finally be converted into a water-injector. This location is quite deep, however, which means there is a high risk to miss the oil-saturated zone. Location B has a higher position, and is therefore more likely to reach hydrocarbons. Its expected porosity is high, so that the appraisal well could be converted into a producing well on further developments. It may nevertheless be affected by the southern fault permeability, i.e., turning it into a producer would result in low hydrocarbon production if the fault has a sealing effect.

4.4. Discussion

On this simple example, uncertainty visualization algorithms were used to support the choice of a new appraisal well location, as they could carry information about the porosity and its associated uncertainty at once. Secondary objectives of the appraisal well should however still be assessed by an expert geologist, as the location of the well is not solely based on the degree of uncertainty. Furthermore, the benefits of uncertainty visualization are not limited to data exploration: uncertainty visualization should also be seen as a way to communicate results to a non-expert audience, in a clear and intuitive manner.

5. Methodology

This section describes our methodology during the user study design. It is organized around two main axes: (i) the protocol of the user study (Section 5.1) and (ii) the methods used for the analysis of the answers (Section 5.2). Results are discussed in Section 6.

5.1. Protocol

Our user study was designed to determine whether uncertainty visualization tools affect decision-making, and if they do, to determine whether adjacent or coincident displays perform better than the other.

5.1.1. Participants

Our study involved 123 participants with a background in geology, including one domain expert, 5 PhD students from the CRPG-CNRS laboratory and 117 MSc students from the Ecole Nationale Supérieure de Géologie. The domain expert was a professional of the oil and gas industry; his level of expertise can be considered as excellent. The 5 PhD students had a variety of different backgrounds; their level of expertise in over-pressured reservoirs ranges from fair to good. The 117 MSc students were involved in geological thematics, but could not be considered as domain experts at this step of their training; their level of expertise ranges from low to fair. Note that participants were selected for logistical reasons rather than because of their expertise.

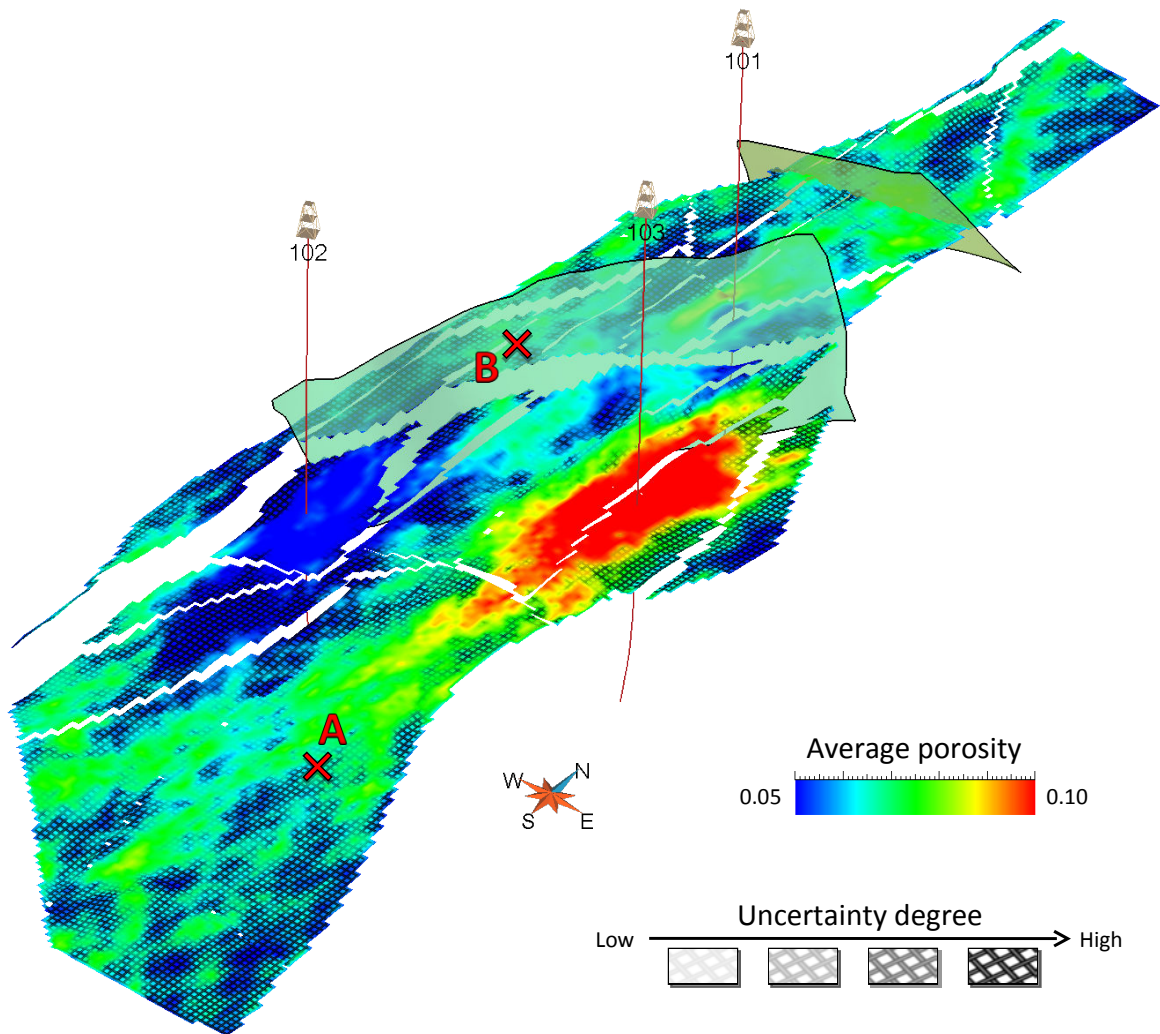


Figure 4: Visualization of a faulted geological layer of the Nan1 field showing both the average porosity (color) and its associated uncertainty (pattern intensity). Areas where pattern is visible have high uncertainty degree. The paths of the first appraisal wells are represented as red lines, with a derrick glyph showing their head location. The gray translucent planes correspond to important faults which split the model into several compartments (minor fault planes are not represented). Locations A and B are two possible candidates for the drilling of a new appraisal well, assessed using local uncertainty behavior and domain-dependent knowledge (possible well reconversion in the production phase).

5.1.2. Model and technical background

The Cloudspin model. The study used a pressure data set generated in the Cloudspin reservoir.⁴ The Cloudspin reservoir is an oil and gas field, whose expected production results were simulated with 27 flow simulations – the low number of simulations is due to the large CPU power requirements of flow simulations. As the local pressure around the production well is lower than the reservoir pressure, the hydrocarbons passively flow toward the production well. However, the simulations were performed with several realizations of permeability and various Pressure-Volume-Temperature tables as input parameters. As the viscosity of the hydrocarbons and the rocks permeability varies, so will the local speed of the hydrocarbons' flow, resulting in different reservoir production outcomes. Uncertainty was sampled using all simulated pressure fields taken at a specific time-step of the simulation, then quantified with a normalized standard deviation metric (also termed *variation coefficient*). The area of high uncertainty highlights the possible boundaries of the depleted area at this time-step.

Pressure in well design. In this user study, participants were required to consider the development of water-injection wells. The goal of such wells is to sweep oil toward the production well, resulting in improved production results. However, the possible local overpressure must be studied prior to drilling a new well, so that the well head counterbalances overpressure at any time. Incorrect estimation of the overpressure may result in a blow up of the well head, putting human lives at risk and destroying costly equipment.

5.1.3. User study description

We split the participants into three groups. All participants were asked the same questions, but were provided with different pictures of the reservoir depending on the group to which they belonged (Figure 5), as suggested in Deitrick and Edsall (2006). The first group was given pressure data only, the second group pressure and uncertainty data displayed separately, and the third group pressure and uncertainty data in a single display using our pattern transparency visualization method. The images were taken in the same conditions for all three groups; all were given short written explanations about the geological background and the risks associated to incorrect overpressure estimation (Section 5.1.2). All materials were provided as colored paper prints; we ensured that images were printed with high quality and that colors were correctly approximated by the printer.

Question design. Because all participants were asked the same questions, we assume that any statistical difference in the answers between the groups would come from the data that were provided. To be valid, this assumption requires that (i) all groups have participants with comparable skills and (ii) a significant amount of answers is collected for each group, so that the statistics are not biased by random behaviors. To meet these requirements, (i) we designed a question meant to compare the groups abilities and (ii) we tried to reach as many participants as possible and assigned them to either group with equal probability, so that the groups were approximately the same size.

Description and goal of the questions. Participants were asked three different questions with increasing complexity. In the first question, participants were asked to indicate in which area of the reservoir pressure was the highest, given three possible choices – east, center or west of the reservoir main panel. This question was intended as a simple map reading test, in order to assess and compare the groups' abilities.

The second question required participants to compare well locations A and B in terms of worst possible overpressure. Location A showed higher pressure than location B, but associated uncertainty was lower. We designed this question to study how uncertainty could affect decision in binary choices, and whether the way uncertainty is presented could influence decisions.

In the third question, participants were asked to rank well locations D, E, F, G and H from lowest to highest possible overpressure. The goal of this question is quite similar to question 2, but involves much more qualitative decision making, as quantifying the worst possible pressure for each of the five wells may be difficult to achieve in a limited amount of time.

5.2. Analysis

The analysis of the answers was performed using different statistical tools. The summary of the tools we used is listed below on a question-per-question basis.

⁴Data courtesy of Schlumberger and model courtesy of Paradigm.

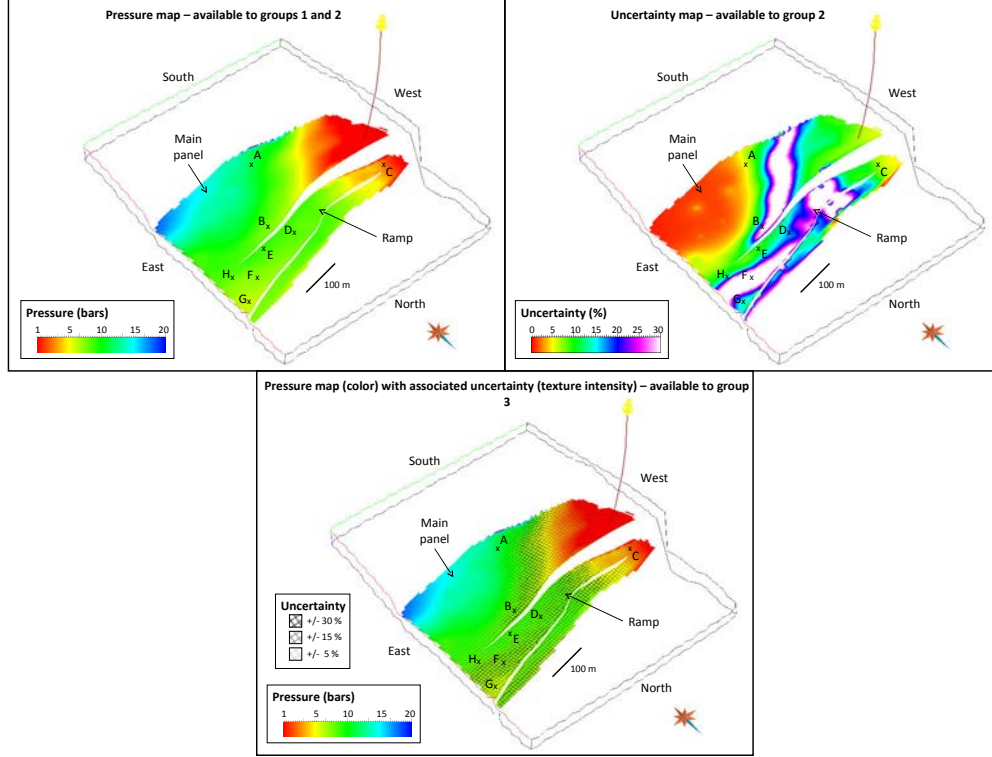


Figure 5: Pictures of the Cloudspin reservoir provided with the user study. Top left: average pressure map (available to groups 1 and 2). Top right: uncertainty map (available to group 2). Bottom: average pressure map with uncertainty (available to group 3).

Question 1. Question 1 was analyzed using a simple percentage of correct vs. incorrect answers. Although this method does not allow for advanced statistical testing of the distribution of the answers, the large number of correct answers made more sophisticated testing methods useless (Section 6).

Question 2. The results of question 2 were analyzed with the Wilcoxon signed-rank test (Wilcoxon, 1945), which was chosen because it does not assume any type of distribution for the answers (non-parametric test). The Wilcoxon test compares two sets of samples, assuming that they come from the same distribution (H_0 hypothesis). We then compute the conditional probability p_{i-j} of the answers of groups i and j under H_0 . If p_{i-j} is below a threshold α , the H_0 hypothesis is rejected, i.e. we make the choice to consider that groups i and j have statistically different answers. We compared the answers of groups 1, 2 and 3 with a threshold α of 0.01.

Question 3. For the analysis of question 3, we have computed the error between the answer and the actual order. This error is computed as the sum of the errors for each well, as detailed in equation (2). It can be seen as the total pressure mismatch between the answer and the actual maximum possible pressure of the reservoir, i.e. the error acts as a measure of the perception mismatch.

$$E = \sum_{i \in \Omega} \sum_{j \in \Omega, j \neq i} \begin{cases} 0 & \text{if wells } i \text{ and } j \text{ are correctly ordered} \\ \text{Abs}(P_i - P_j) & \text{otherwise} \end{cases} \quad (2)$$

where E is the error, Ω is the set of well locations [well₁ = D, well₂ = E, well₃ = F, well₄ = G, well₅ = H] and P_i is the maximum possible pressure at well i .

The average error of a group allows to determine whether the selected visualization method improves or clutters perception. Furthermore, we applied the Wilcoxon signed-rank test to the distribution of the errors, in order to highlight common answer patterns between groups.

Table 3: Correct answer rate in question 1.

Group	Correct answer rate	Answer count	Incorrect answer count
Group 1	95.1%	41	2
Group 2	95.0%	40	2
Group 3	100.0%	42	0

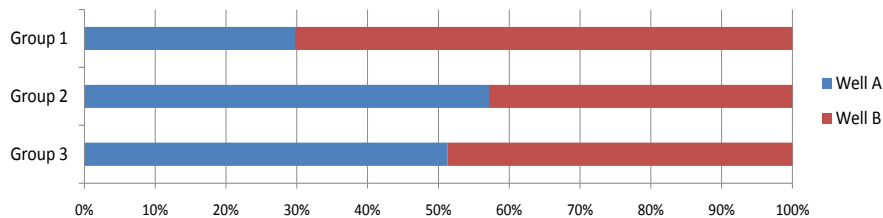


Figure 6: Answer rates for question 2.

6. Results and discussion

6.1. Interpretation of the results

This section provides detailed description of the results we obtained and their implications for each question.

Question 1

Results. Within the 123 participants to the study, more than 96% answered correctly to the first question. We analyzed the correct answer rates for each group; results are reported in Table 3. Participants with incorrect answers were all MSc students.

Interpretation. The large majority of the participants were able to read the data they were provided with. We found that all three groups had an equivalent ability to read the pressure map, which guarantees that our group assignment process did not introduce any bias in terms of map-reading skills. The four participants with incorrect answers were discarded in questions 2 and 3 for additional safety.

Question 2

Results. Distribution of the answers in question 2 were compared using statistical hypothesis testing; answer rates are reported in Figure 6, and probabilities that answers are sampled from similar distributions under the H0 hypothesis are reported in Table 4. We found that groups 2 and 3 have statistically similar answers, while group 1 differs from both group 2 and group 3.

Interpretation. Group 1 had no information about the local uncertainty, while groups 2 and 3 were provided with uncertainty maps. Hence, we interpret these results as the effects of uncertainty on decision-making. This conclusion is similar to the studies performed by Leitner and Buttenfield (2000) and Deitrick and Edsall (2006). Note that all groups have the same level of accuracy for that question since wells A and B have very similar worst possible pressures, i.e., 10.50 bars for well A and 10.42 bars for well B.

The way uncertainty is presented has no clear effect on the answers of this question. We believe this is connected to the question complexity; indeed, it is quite easy to quantify pressure and uncertainty at only two well locations, no matter whether uncertainty is presented jointly or separately.

Question 3

Results. The results of question 3 are reported in Table 5; they show that group 1 was closer to the actual pressure ordering than group 3, who in turn performed better than group 2. The Wilcoxon test showed that answers of groups 1 and 3 have similar distribution patterns under the H0 hypothesis, while answers of group 2 differ from the two other groups (Table 6).

Table 4: Comparison between the distribution of the answers in question 2.

Groups compared	p(similar distributions H0)	Conclusion
Groups 1 & 2	$p_{1-2} = 0.0007$	Different distributions
Groups 1 & 3	$p_{1-3} = 0.0053$	Different distributions
Groups 2 & 3	$p_{2-3} = 0.3501$	Similar distributions

Table 5: Average perception mismatch for each group in question 3.

Group	Average mismatch	Deviation
Group 1	0.59 Kbars	+/- 0.23 Kbars
Group 2	2.60 Kbars	+/- 0.31 Kbars
Group 3	1.00 Kbars	+/- 0.35 Kbars

Interpretation. The average error results suggest that a visualization which integrates a joint display of uncertainty is clearer than two separate visualizations for data and uncertainty for qualitative choices, but that it still occludes part of the primary information present in the map. However, looking at the confidence interval about the average error for groups 1 and 3, we found a non-empty intersection, i.e. both groups may actually be at the same level of perception quality.

This hypothesis was checked using the Wilcoxon signed-rank test; we found that answers of groups 1 and 3 had similar distribution patterns under the H0 hypothesis, i.e., that they are actually sampled from the same initial distribution. This result implies that the average error difference between groups 1 and 3 may actually be an artifact, possibly coming from a heterogeneous sampling of the participants, but more likely coming from random behaviors. This suggests that the joint display of uncertainty presented in this paper has the same level of clarity than the map by itself, i.e., that this joint display of uncertainty does not clutter the perception.

Conversely, the average error obtained for group 2 suggests that a separate display of the data and the uncertainty obscures perception when compared to the other methods; the Wilcoxon test confirms that group 2 does not share common distribution patterns with any other group.

6.2. Discussion

While the setup of the user study allowed us to answer some of the questions we were investigating, some technical, material or organizational issues prevented us from gathering some interesting pieces of information. This part discusses the limitations of our methodology.

Decision-making speed. Our study was practically distributed as a paper form that could be filled in a limited amount of time. Unfortunately, those conditions did not allow us to measure the time participants took on each question, since the form was distributed to large groups of participants at once. We were only able to limit the total amount of time available to answer all the questions. Time measures would have allowed statistical testing of the decision-making speed between groups, which could have been seen as a first approximation of the intuitiveness of each approach.

Participants' representativity. More than 95% of the participants were MSc students; hence, the population used in our experiment clearly suffered from a lack of domain experts. According to Roth (2009b) the task complexity strengthens the influence of the participant's expertise. Our experiments involved well targeting, which is indeed a challenging task. Hence, it is reasonable to expect a bias in the results of our user study, even though the limited number of experts does not allow statistical testing of this hypothesis.

Statistical relevance. In our experiments, participants were evenly distributed into three groups who were asked the same set of three questions, each group having access to a different visualization method. This methodology avoids learning bias, but limits the number of questions that can be tested, hence decreases the statistical relevance of the study. We could have provided the three visualization methods in a random order to each participant, which would have increased our pool of answers while minimizing the effects of learning bias.

Table 6: Comparison between the distribution of the answers in question 3.

Groups compared	p(similar distributions H0)	Conclusion
Groups 1 & 2	$p_{1-2} = 0.0002$	Different distributions
Groups 1 & 3	$p_{1-3} = 0.9957$	Similar distributions
Groups 2 & 3	$p_{2-3} = 0.0043$	Different distributions

7. Conclusion

This paper aimed at answering the following questions:

Does uncertainty visualization influence decision making? Since this question has already received a positive answer in the literature, we see this test as a cross-check of the consistency of our study, rather than as a new result. The second question of our study revealed different answer patterns whether participants were aware of uncertainty or not, which suggests that uncertainty visualization does have an influence on decision making; hence, our results agree with the findings of previous studies.

Is there a difference in terms of decision-making accuracy between coincident and adjacent displays of spatial uncertainty? Our study resulted in contradictory findings about the influence of the way uncertainty is displayed: our second question resulted in similar results for adjacent and coincident displays, while our third question highlighted significant differences between the two methods. We interpret these different results as a consequence of the increasing complexity of the tasks participants had to perform; the second question was simple enough to be carried out with adjacent displays, while the third question involved complex multi-location assessments, for which adjacent maps introduced a perceptual and cognitive overload (Harrower, 2003). This suggests that the way uncertainty is presented actually influences user perception; hence, visualizations dedicated to real-world applications should aim at using compact representations of the information in order to minimize cognitive burdens for the users.

Does uncertainty visualization act to clarify or to clutter the display of uncertain data? The third question studied the accuracy of the answers with respect to the actual value presented on the display; it has shown that data presented without uncertainty and coincident displays of uncertainty shared the same answer patterns and were close to the best decision, i.e., that uncertainty visualization can present a large amount of information in compact displays without obscuring the important structures in the data. This assertion has to be contrasted with the results of adjacent uncertainty displays, where decisions were much less accurate. This once again points out that only careful design of uncertainty visualizations, with minimal burden on the user, can actually add value to the decisions.

Is the perception of 2.5D/3D data different from the perception of 2D map data in presence of uncertainty? The results of our study are generally consistent with the previous research on this topic, e.g., Leitner and Battenfield (2000) and Deitrick and Edsall (2006), which suggests there is little difference in the perception of 2.5D data as compared to traditional map data. However, our experiments did not include any testing of actual 3D data. Exploration of 3D data substantially differs from exploration of 2D/2.5D data, because it typically involves more user interactions, e.g., moving the view point, zooming some particular areas or changing the lights. While such interactions increase the complexity of the logistics of user study design, they may influence the way uncertainty is perceived in the data. Furthermore, results from Sanyal et al. (2009) show that it was significantly harder for users to assess uncertainty on 2D data than on 1D data; a similar ascending complexity can thus be expected when jumping from 2D/2.5D data to 3D data. Hence, we consider that the question of user perception on volumes is open; understanding how 3D uncertainties can be seen is a fundamental issue that still needs to be answered in order to produce efficient uncertainty visualizations for geological applications, but also for other scientific domains such as medical research, meteorology or oceanography.

Acknowledgments

We want to express our acknowledgment to all the participants of the user study, and to Irina Panfilova for her useful pieces of advice when designing the user study. We also thank the anonymous reviewers for their suggestions

to improve this paper.

We thank Total for providing the Nan1 model used in the case study, and Paradigm for providing the Cloudspin uncertainty model used in the user study – many thanks to Alexandre Hugot and Emmanuel Gringarten for their work and explanations on the Cloudspin model. We also acknowledge Paradigm for providing the Gocad software and developer API. This research is part of a PhD thesis funded by the Gocad consortium. All the members of the consortium are hereby acknowledged for their support.

This is CRPG-CNRS contribution n°2044.

- Aerts, J., Clarke, K., Keuper, A., 2003. Testing popular visualization techniques for representing model uncertainty. *Cartography and Geographic Information Science* 30 (3), 249–262.
- Arpat, G. B., Caers, J., 2007. Conditional simulation with patterns. *Mathematical Geology* 39 (2), 177–203.
- Bertin, J., 1983. *Semiology of Graphics*. The University of Wisconsin Press.
- Buttenfield, B., 1993. Representing data quality. *Cartographica: The International Journal for Geographic Information and Geovisualization* 30 (2), 1–7.
- Cedilnik, A., Rheingans, P., 2000. Procedural annotation of uncertain information. *Proceedings of the 11th IEEE Visualization 2000 Conference (VIS 2000)*.
- Chambers, R., Yarus, J., 2006. Practical geostatistics - an armchair overview for petroleum reservoir engineers. *Journal of Petroleum Technology, Distinguished Author Series* 11.
- Davis, T., Keller, C., 1997. Modelling and visualizing multiple spatial uncertainties. *Computers and Geosciences* 23 (4), 397–408.
- Deitrick, S., Edsall, R., 2006. The influence of uncertainty visualization on decision making: An empirical evaluation. *Progress in Spatial Data Handling, 12th International Symposium on Spatial Data Handling*, 719–738.
- Deutsch, C., Tran, T., 2002. Fluvsim: a program for object-based stochastic modeling of fluvial depositional systems. *Computers and Geosciences* 28 (4), 525–535.
- Djurcilov, S., Kim, K., Lermusiaux, P., Pang, A., 2001. Volume rendering data with uncertainty information. *Proceedings of the EG+IEEE VisSym on Data Visualization*, 243–52.
- Dooley, M., Lavin, S., 2007. Visualizing method-produced uncertainty in isometric mapping. *Cartographic Perspectives* 56, 17–36.
- Edwards, L., Nelson, E., 2001. Visualizing data certainty: A case study using graduated circle maps. *Cartographic Perspectives* 38, 19–36.
- Ehlschlaeger, C., Shortridge, A., Goodchild, M., 1996. Visualizing spatial data uncertainty using animation. *Computers & Geosciences* 23 (4), 387–395.
- Evans, B., 1997. Dynamic display of spatial data-reliability: does it benefit the map user? *Computers and Geosciences* 23 (4), 409–422.
- Fisher, P., 1993. Visualizing uncertainty in soil maps by animation. *Cartographica: The International Journal for Geographic Information and Geovisualization* 30 (2), 20–27.
- Gershon, N., 1998. Visualization of an imperfect world. *IEEE Computer Graphics and Applications* 18 (4), 43–45.
- Griethe, H., Schumann, H., 2006. The visualization of uncertain data: Methods and problems. *Proceedings of SimVis 2006*, SCS Publishing House, 143–156.
- Grigoryan, G., Rheingans, P., 2004. Point-based probabilistic surfaces to show surface uncertainty. *IEEE Transactions on Visualization and Computer Graphics*, 546–573.
- Harrower, M., 2003. Representing uncertainty: Does it help people make better decisions? In: *UCGIS Workshop: Geospatial Visualization and Knowledge Discovery Workshop*, National Conference Center, Landsdowne, VA., Nov. pp. 18–20.
- Healey, C., Booth, K., Enns, J., 1996. High-speed visual estimation using preattentive processing. *Transactions on Computer-Human Interaction* 3 (2), 107–135.
- Hope, S., Hunter, G., 2007. Testing the effects of thematic uncertainty on spatial decision-making. *Cartography and Geographic Information Science* 34 (3), 199–214.
- Hu, L., 2000. Gradual deformation and iterative calibration of gaussian-related stochastic models. *Mathematical Geology* 32 (1), 87–108.
- Interrante, V., 2000. Harnessing natural textures for multivariate visualization. *IEEE Computer Graphics and Applications* 20 (6), 6–11.
- Johnson, C., Sanderson, A., 2003. A next step: Visualizing errors and uncertainty. *IEEE Computer Graphics and Applications* 23 (5), 6–10.
- Journel, A., Gunderso, R., Gringarten, E., Yao, T., 1998. Stochastic modelling of a fluvial reservoir: a comparative review of algorithms. *Journal of Petroleum Science and Engineering* 21 (1), 95–121.
- Kobus, D., Proctor, S., Holste, S., 2001. Effects of experience and uncertainty during dynamic decision making. *International Journal of Industrial Ergonomics* 28 (5), 275–290.
- Kosara, R., Miksch, S., Hauser, H., 2001. Semantic depth of field. *Proceedings of the IEEE symposium on Information Visualization*, 97–104.
- Lafont, F., 2007. Pers. com. Jaca field trip, Total.
- Leitner, M., Buttenfield, B., 2000. Guidelines for the display of attribute certainty. *Cartography and Geographic Information Science* 27 (1), 3–14.
- MacEachren, A., 1992. Visualizing uncertain information. *Cartographic Perspective* 13, 10–19.
- MacEachren, A., 1995. *How maps work*. Guilford Press New York.
- MacEachren, A., Brewer, C., Pickle, L., 1998. Visualizing georeferenced data: representing reliability of health statistics. *Environment and Planning A* 30, 1547–1562.
- MacEachren, A., Robinson, A., Hopper, S., Gardner, S., Murray, R., Gahegan, M., Hetzler, E., 2005. Visualizing Geospatial Information Uncertainty: What We Know and What We Need to Know. *Cartography and Geographic Information Science* 32 (3), 139–161.
- McGranaghan, M., 1993. A cartographic view of spatial data quality. *Cartographica: The International Journal for Geographic Information and Geovisualization* 30 (2), 8–19.
- Pang, A., 2006. Visualizing uncertainty in natural hazards. *Risk Assessment, Modeling and Decision Support* 14, 261–294.
- Pang, A., Wittenbrink, C., Lodha, S., 1997. Approaches to uncertainty visualization. *The Visual Computer* 13 (8), 370–390.
- Rheingans, P., Landreth, C., 1995. Perceptual principles for effective visualizations. In: Grinstein, G., Levkowitz, H. (Eds.), *Perceptual Issues in*

- Visualization. Springer-Verlag, pp. 59–73.
- Rhodes, P., Laramee, R., Bergeron, R., Sparr, T., 2003. Uncertainty visualization methods in isosurface rendering. *EUROGRAPHICS 2003 Short Papers*, 83–88.
- Robinson, A., 2006. Highlighting techniques to support geovisualization. In: *Proceedings of the ICA Workshop on Geovisualization and Visual Analytics*.
- Roth, R., 2009a. A qualitative approach to understanding the role of geographic information uncertainty during decision making. *Cartography and Geographic Information Science* 36 (4), 315–330.
- Roth, R., 2009b. The impact of user expertise on geographic risk assessment under uncertain conditions. *Cartography and Geographic Information Science* 36 (1), 29–43.
- Sanyal, J., Zhang, S., Bhattacharya, G., Amburn, P., Moorhead, R., 2009. A User Study to Compare Four Uncertainty Visualization Methods for 1D and 2D Datasets. *IEEE Transactions on Visualization and Computer Graphics* 15 (6), 1209–1218.
- Shortridge, B., 1982. Stimulus processing models from psychology: can we use them in cartography? *Cartography and Geographic Information Science* 9 (2), 155–167.
- Skeels, M., Lee, B., Smith, G., Robertson, G., 2008. Revealing uncertainty for information visualization. In: *Proceedings of the working conference on Advanced visual interfaces*. ACM, pp. 376–379.
- Srivastava, R., 1994a. The visualization of spatial uncertainty. *The Visualization of Spatial Uncertainty, Stochastic Modeling and Geostatistics: Principles, Methods, and Case Studies*, J.M Yarus and R.L. Chambers, eds., American Assoc. of Petroleum Geologists.
- Srivastava, R. M., 1994b. The interactive visualization of spatial uncertainty. *SPE* 27965.
- Thomson, J., Hetzler, E., MacEachren, A., Gahegan, M., Pavel, M., 2005. A typology for visualizing uncertainty. In: *Proc. SPIE*. Vol. 5669. pp. 146–157.
- Tory, M., Möller, T., 2004. Human factors in visualization research. *IEEE Transactions on Visualization and Computer Graphics* 10 (1).
- Tufte, E., 2001. *The Visual Display of Quantitative Information*, 2nd ed. Graphics Press, Cheshire.
- USGS, 1977. *Spatial data transfer standard (sdts): Logical specifications*.
- Ware, C., 2004. *Information Visualization: Perception for Design*, 2nd ed. Morgan Kaufmann Publishers.
- Wilcoxon, F., 1945. Individual comparisons by ranking methods. *Biometrics Bulletin* 1 (6), 80–83.
- Wright, J., 1942. Map Makers Are Human: Comments on the Subjective in Maps. *Geographical Review*, 527–544.
- Zuk, T., Carpendale, S., 2006. Theoretical analysis of uncertainty visualization. *Proceedings of SPIE*. Vol. SPIE-6060, 66–79.

Bibliographie

- AERTS, J., CLARKE, K. et KEUPER, A. *Testing popular visualization techniques for representing model uncertainty*. Cartography and Geographic Information Science, 2003, vol. 30, n°3, pp. 249–262.
- ALBERT, M. *Lightness and perceptual transparency*. Perception, 2006, vol. 35, n°4, pp. 433.
- ANDERSON, J., GOSINK, L., DUCHAINEAU, M. et JOY, K. *Feature identification and extraction in function fields*. In Proceedings of the Eurographics/IEEE-VGTC Symposium on Visualisation, 2007. Eurographics Association.
- ARTHUR, D. et VASSILVITSKII, S. *How slow is the k-means method ?* In Proceedings of the twenty-second annual symposium on Computational geometry, 2006. ACM, pp. 153.
- AURENHAMMER, F. et KLEIN, R. Voronoi diagrams. In Handbook of Computational Geometry. Elsevier, 2000. 5, pp. 201–290.
- BARBIER, S. et BONNEAU, G. *GPU Improvements on the sorting and projection of tetrahedral meshes for direct volume rendering*. IEEE Pacific Visualization Symposium 2008, Poster session, 2008.
- BARDOSSY, G. et FODOR, J. *Traditional and new ways to handle uncertainty in geology*. Natural Resources Research, 2001, vol. 10, n°3, pp. 179–187.
- BARRAS, C. *Pointillism shows the way for computer graphics*. The New Scientist, 2010, vol. 206, n°2754, pp. 18.
- BASSER, P. et LE BIHAN, D. *Fiber orientation mapping in an anisotropic medium with NMR diffusion spectroscopy*. In Proceedings of the 11th Annual Meeting SMRM, Berlin, 1992. pp. 1221.
- BEDERSON, B. et BOLTMAN, A. *Does animation help users build mental maps of spatial information ?* The craft of information visualization : readings and reflections, 2003, pp. 87.
- BEESON, C. *Animation in the “Dawn” demo*. In GPU Gems, Programming Techniques, Tips, and Tricks for Real-Time Graphics. Wiley, Chichester, UK, 2004. pp. 63–72.

- BERGER, M. et OLIGER, J. *Adaptive mesh refinement for hyperbolic partial differential equations*. Journal of computational Physics, 1984, vol. 53, n°3, pp. 484–512.
- BERTIN, J. *Semiology of graphics*. The University of Wisconsin Press, 1983.
- BLINN, J. *Simulation of wrinkled surfaces*. In Proceedings of the 5th annual conference on Computer graphics and interactive techniques, 1978. ACM New York, NY, USA, pp. 286–292.
- BLOOMENTHAL, J. *Polygonization of implicit surfaces*. Computer Aided Geometric Design, 1988, vol. 5, n°4, pp. 53–60.
- BOADA, I., NAVAZO, I. et SCOPIGNO, R. *Multiresolution volume visualization with a texture-based octree*. The Visual Computer, 2001, vol. 17, n°3, pp. 185–197.
- BOLLER, R., BRAUN, S., MILES, J. et LAIDLAW, D. *Application of uncertainty visualization methods to meteorological trajectories*. Earth Science Informatics, 2010, pp. 1–8.
- BOMBARDE, S. *Building stratigraphic surfaces tied to well markers using a morphing-like approach - application to stratigraphic grid building*. Proceedings of the 15th Gocad Meeting, 1997.
- BOND, C., GIBBS, A., SHIPTON, Z., JONES, S. et STRACHER, G. *What do you think this is? “Conceptual uncertainty” in geoscience interpretation*. GSA Today, 2007, vol. 17, n°11, pp. 4.
- BOTCHEN, R., WEISKOPF, D. et ERTL, T. *Texture-based visualization of uncertainty in flow fields*. IEEE Visualization (VIS 05), 2005, pp. 647–654.
- BOUBEKEUR, T. et SCHLICK, C. *Approximation of subdivision surfaces for interactive applications*. ACM SIGGRAPH Sketch Program, 2007.
- BROWN, P. *NV_transform_feedback2 extension specifications*, 2008. URL http://www.opengl.org/registry/specs/NV/transform_feedback2.txt.
- BROWN, R. *Animated visual vibrations as an uncertainty visualisation technique*. In Proceedings of the 2nd international conference on Computer graphics and interactive techniques in Australasia and South East Asia, 2004. ACM, pp. 84–89.
- BUATOIS, L. *Algorithmes sur GPU de visualisation et de calcul pour des maillages non-structurés*. Th. Doct.. Institut National Polytechnique de Lorraine, 2008. URL http://www.gocad.org/www/people/files/thesis_luc_buatois.pdf.
- BUATOIS, L. et CAUMON, G. *4D Morph : Dynamic Visualization of 4D Reservoir Data With Continuous Transitions Between Time Steps*. In Proceedings of the 25th Gocad Meeting, 2005.

-
- BURNS, M., KLAWE, J., RUSINKIEWICZ, S., FINKELSTEIN, A. et DECARLO, D. *Line drawings from volume data*. In ACM SIGGRAPH 2005 Papers, 2005. ACM, pp. 518.
- BURT, J., ZHU, A. et HARROWER, M. *Depicting fuzzy soil class uncertainty using perception-based color models*. Proceedings of the 11th World Congress of International Fuzzy Systems Association (IFSA2005) : Fuzzy Logic, Soft Computing and Computational Intelligence, 2005, pp. 112–117.
- BUTTENFIELD, B. *Representing data quality*. Cartographica : The International Journal for Geographic Information and Geovisualization, 1993, vol. 30, n°2, pp. 1–7.
- BUTTENFIELD, B. et WEIBEL, R. *Visualizing the quality of cartographic data*. In Third International Geographic Information Systems Symposium (GIS/LIS 88), San Antonio, Texas, 1988.
- CAERS, J. *Direct sequential indicator simulation*. Geostats, 2000, pp. 39–48.
- CAERS, J. *Geostatistical reservoir modelling using statistical pattern recognition*. Journal of Petroleum Science and Engineering, 2001, vol. 29, n°3-4, pp. 177–188.
- CAERS, J. et ZHANG, T. *Multiple-point geostatistics : a quantitative vehicle for integrating geologic analogs into multiple reservoir models*. AAPG Memoir : “Integration of outcrop and modern analog data in reservoir models”, (eds) Grammer, GM et al, 2002.
- CAUMON, G., LEPAGE, F., SWORD, C. et MALLET, J. *Building and editing a sealed geological model*. Mathematical Geology, 2004, vol. 36, n°4, pp. 405–424.
- CAUMON, G., LÉVY, B., CASTANIÉ, L. et PAUL, J. *Visualization of grids conforming to geological structures : a topological approach*. Computers and Geosciences, 2005, vol. 31, n°6, pp. 671–680.
- CEDILNIK, A. et RHEINGANS, P. *Procedural annotation of uncertain information*. Proceedings of the 11th IEEE Visualization 2000 Conference (VIS 2000), 2000, pp. 77–84.
- CHAMBERS, R. et YARUS, J. *Practical geostatistics - An armchair overview for petroleum reservoir engineers*. Journal of Petroleum Technology, Distinguished Author Series, 2006, vol. 11, pp. 78–87.
- CHARLES, T., GUÉMÉNÉ, J., CORRE, B., VINCENT, G. et DUBRULE, O. *Experience with the Quantification of Subsurface Uncertainties*. SPE 68703, 2001.
- CHEN, C., WANG, Y., LI, G. et REYNOLDS, A. *Closed-loop reservoir management on the Brugge test case*. Computational Geosciences, 2009, pp. 1–13.
- CHEN, J., VAN EEDEN, C. et ZIDEK, J. *Uncertainty and the conditional variance*. Statistics & Probability Letters, 2010.

- CHERPEAU, N., CAUMON, G. et LÉVY, B. *Stochastic simulations of fault networks in 3D structural modeling*. Comptes Rendus Geoscience, 2010.
- COGNOT, R., MALLET, J.-L., SOUCHE, L., MASSOT, J. et DENY, L. *Fiber based sgrid construction*. Proceedings of the 21st Gocad Meeting, 2001.
- COLEMAN, C., HARRIS, K., HINTON, A., EPHANOV, A. et RICHARDSON, T. *Automated shader generation using a shader infrastructure*. In Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC), 2007. NTSA.
- COMBA, J., KLOSOWSK, J., MAX, N., MITCHELL, J., SILVA, C. et WILLIAMS, P. *Fast polyhedral cell sorting for interactive rendering of unstructured grids*. Computer Graphics Forum, 1999, vol. 18, n°3, pp. 369–376.
- CORRE, B., THORE, P., DE FERAUDRY, V. et VINCENT, G. *Integrated Uncertainty Assessment For Project Evaluation and Risk Analysis*. SPE 65205, 2000.
- CRONIN, V. *A draft primer on focal mechanism solutions for geologists : Teaching Quantitative Skills in the Geosciences*, 2004. URL http://serc.carleton.edu/files/NAGTWorkshops/structure04/Focal_mechanism_primer.pdf.
- DANIELSSON, P. et SEGER, O. *Generalized and separable Sobel operators*. Machine vision for three-dimensional scenes, 1990, pp. 347–379.
- DAVIS, T. et KELLER, C. *Modelling and visualizing multiple spatial uncertainties*. Computers and Geosciences, 1997, vol. 23, n°4, pp. 397–408.
- DECARLO, D., FINKELSTEIN, A., RUSINKIEWICZ, S. et SANTELLA, A. *Suggestive contours for conveying shape*. ACM Transactions on Graphics (TOG), 2003, vol. 22, n°3, pp. 848–855.
- DEITRICK, S. et EDSALL, R. *The influence of uncertainty visualization on decision making : An empirical evaluation*. Progress in Spatial Data Handling, 12th International Symposium on Spatial Data Handling, 2006, pp. 719–738.
- DENNEY, D. *Results of the Brugge Benchmark Study for Flooding Optimization and History Matching*. Journal of Petroleum Technology, 2009, pp. 48–49.
- DEUTSCH, C. et TRAN, T. *FLUVSIM : a program for object-based stochastic modeling of fluvial depositional systems*. Computers and Geosciences, 2002, vol. 28, n°4, pp. 525–535.
- DEUTSCH, C. et WANG, L. *Hierarchical object-based stochastic modeling of fluvial reservoirs*. Mathematical Geology, 1996, vol. 28, n°7, pp. 857–880.
- DJURCILOV, S., KIM, K., LERMUSIAUX, P. et PANG, A. *Volume rendering data with uncertainty information*. Proceedings of the EG+IEEE VisSym on Data Visualization, 2001, pp. 243–52.

-
- DJURCILOV, S. et PANG, A. *Visualizing gridded datasets with large number of missing values (case study)*. In Proceedings of the conference on Visualization'99 : celebrating ten years, 1999. IEEE Computer Society Press, pp. 408.
- DJURCILOV, S. et PANG, A. *Visualizing sparse gridded data sets*. IEEE Computer Graphics and Applications, 2000, pp. 52–57.
- DOOLEY, M. et LAVIN, S. *Visualizing method-produced uncertainty in isometric mapping*. Cartographic Perspectives, 2007, vol. 56, pp. 17–36.
- EDWARDS, L. et NELSON, E. *Visualizing data certainty : A case study using graduated circle maps*. Cartographic Perspectives, 2001, vol. 38, pp. 19–36.
- EHLSCHLAEGER, C. *Exploring Temporal Effects in Animations Depicting Spatial Data Uncertainty*, 1998. URL <http://chuck.ehlschlaeger.info/older/AAG98/>.
- EHLSCHLAEGER, C., SHORTRIDGE, A. et GOODCHILD, M. *Visualizing spatial data uncertainty using animation*. Computers & Geosciences, 1996, vol. 23, n°4, pp. 387–395.
- ENGEL, K., KRAUS, M. et ERTL, T. *High-quality pre-integrated volume rendering using hardware-accelerated pixel shading*. Proceedings of Eurographics/SIGGRAPH Workshop on Graphics Hardware (2001), 2001, pp. 9–16.
- EVANS, B. *Dynamic display of spatial data-reliability : does it benefit the map user?* Computers and Geosciences, 1997, vol. 23, n°4, pp. 409–422.
- FETEL, E. *Reservoir Flow Uncertainty in Presence of Stochastic Parameters*. Proceedings of the 26th Gocad Meeting, 2006.
- FETEL, E. et CAUMON, G. *Reservoir flow uncertainty assessment using response surface constrained by secondary information*. Journal of Petroleum Science and Engineering, 2008, vol. 60, n°3-4, pp. 170–182.
- FISHER, P. *Visualizing uncertainty in soil maps by animation*. Cartographica : The International Journal for Geographic Information and Geovisualization, 1993, vol. 30, n°2, pp. 20–27.
- FLYNN, M. et RUDD, K. *Parallel architectures*. ACM Computing Surveys (CSUR), 1996, vol. 28, n°1, pp. 70.
- FOLKEGÅRD, N. et WESSLÉN, D. *Dynamic code generation for realtime shaders*. In Proceedings of SIGRAD 2004, 2004. pp. 11–15.
- FRANK, T., TERTOIS, A. et MALLET, J. *3D-reconstruction of complex geological interfaces from irregularly distributed and noisy point data*. Computers & Geosciences, 2007, vol. 33, n°7, pp. 932–943.

- GAHEGAN, M. et EHLERS, M. *A framework for the modelling of uncertainty between remote sensing and geographic information systems*. ISPRS Journal of Photogrammetry and Remote Sensing, 2000, vol. 55, n°3, pp. 176–188.
- GARCIA, A., BELL, J., CRUTCHFIELD, W. et ALDER, B. *Adaptive Mesh and Algorithm Refinement Using Direct Simulation Monte Carlo*. Journal of Computational Physics, 1999, vol. 154, pp. 134–155.
- GEE, K. *Direct3D 11 tessellation*. In Gamefest 2008 graphics track, 2008. URL <http://www.microsoftpost.com/microsoft-download/direct3d-11-tessellation/>.
- GERSHON, N. *Visualization of an imperfect world*. IEEE Computer Graphics and Applications, 1998, vol. 18, n°4, pp. 43–45.
- GIBOU, F. et FEDKIW, R. *A fast hybrid k-means level set algorithm for segmentation*. In 4th Annual Hawaii International Conference on Statistics and Mathematics, 2005. pp. 281–291.
- GOOVAERTS, P. *Geostatistics for natural resources evaluation*. Oxford University Press, 1997. ISBN 0-19-511538-4.
- GOOVAERTS, P. *Geostatistical analysis of disease data : visualization and propagation of spatial uncertainty in cancer mortality risk using Poisson kriging and p-field simulation*. International Journal of Health Geographics, 2006, vol. 5, n°1, pp. 7. URL <http://www.ij-healthgeographics.com/content/5/1/7>.
- GOTWAY, C., FERGUSON, R., HERGERT, G. et PETERSON, T. *Comparison of kriging and inverse-distance methods for mapping soil parameters*. Soil Science Society of America Journal, 1996, vol. 60, n°4, pp. 1237–1247.
- GRIETHE, H. et SCHUMANN, H. *The visualization of uncertain data : Methods and problems*. Proceedings of SimVis 2006, SCS Publishing House, 2006, pp. 143–156.
- GRIGORYAN, G. et RHEINGANS, P. *Point-based probabilistic surfaces to show surface uncertainty*. IEEE Transactions on Visualization and Computer Graphics, 2004, pp. 546–573.
- HAKER, S., ANGENENT, S., TANNENBAUM, A., KIKINIS, R., SAPIRO, G. et HALLE, M. *Conformal surface parameterization for texture mapping*. IEEE Transactions on Visualization and Computer Graphics, 2000, vol. 6, n°2, pp. 181–189.
- HANRAHAN, P. et STURMAN, D. *Interactive animation of parametric models*. The Visual Computer, 1985, vol. 1, n°4, pp. 260–266.
- HAR-PELED, S. et SADRI, B. *How fast is the k-means method?* Algorithmica, 2005, vol. 41, n°3, pp. 185–202.

-
- HARROWER, M. *Representing uncertainty : Does it help people make better decisions ?* In UCGIS Workshop : Geospatial Visualization and Knowledge Discovery Workshop, National Conference Center, Landsdowne, VA., Nov, 2003. pp. 18–20.
- HARVARD CMT. *Centroid Moment Tensor Project*, 2005. URL <http://www.globalcmt.org/CMTsearch.html>.
- HASTINGS, W. *Monte Carlo sampling methods using Markov chains and their applications*. Biometrika, 1970, pp. 97–109.
- HEALEY, C., BOOTH, K. et ENNS, J. *High-speed visual estimation using preattentive processing*. Transactions on Computer-Human Interaction, 1996, vol. 3, n°2, pp. 107–135.
- HECKBERT, P. *Survey of texture mapping*. IEEE Computer Graphics and Applications, 1986, vol. 6, n°11, pp. 56–67.
- HENGL, T. *Visualisation of uncertainty using the HSI colour model : computations with colours*. Proceedings of the 7th International Conference on GeoComputation, 2003.
- HENGL, T. et TOOMANIAN, N. *Maps are not what they seem : representing uncertainty in soil-property maps*. In Proceedings of the 7th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, 2006. pp. 5–7.
- HENGL, T., WALVOORT, D. et BROWN, A. *Pixel and Colour Mixture : GIS techniques for visualisation of fuzzyness and uncertainty of natural resource inventories*. In Proceedings of the 5th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, 2002.
- HENRION, V., CAUMON, G. et CHERPEAU, N. *ODSIM : An object-distance simulation method for conditioning complex natural structures*. Mathematical Geosciences, in press.
- HENRION, V., PELLERIN, J. et CAUMON, G. *A Stochastic Methodology for 3D Cave System Modeling*. Proceedings of the 8th International Geostatistics Congress, 2008, vol. 1, pp. 525–533.
- HOFF, K. E., KEYSER, J., LIN, M., MANOCHA, D. et CULVER, T. *Fast computation of generalized Voronoi diagrams using graphics hardware*. Proceedings of the 26th annual conference on Computer graphics and interactive techniques, 1999, pp. 277–286.
- HOLDEN, L., HAUGE, R., SKARE, O. et SKORSTAD, A. *Modeling of fluvial reservoirs with object models*. Mathematical Geology, 1998, vol. 30, n°5, pp. 473–496.
- HOPE, S. *Decision-making under spatial uncertainty*. Th. Doct.. The University of Melbourne, Department of Geomatics, 2005. URL <http://eprints.unimelb.edu.au/archive/00001041/>.

- HOPE, S. et HUNTER, G. *Testing the effects of thematic uncertainty on spatial decision-making*. Cartography and Geographic Information Science, 2007, vol. 34, n°3, pp. 199–214.
- HU, L. *Gradual deformation and iterative calibration of Gaussian-related stochastic models*. Mathematical Geology, 2000, vol. 32, n°1, pp. 87–108.
- HUTTENLOCHER, D., KLANDERMAN, G. et RUCKLIDGE, W. *Comparing images using the Hausdorff distance*. IEEE Transactions on pattern analysis and machine intelligence, 1993, pp. 850–863.
- IKITS, M., KNISS, J., LEFOHN, A. et HANSEN, C. *Volume rendering techniques*. GPU Gems. Addison-Wesley, 2004, pp. 667–692.
- INTERRANTE, V. *Harnessing natural textures for multivariate visualization*. IEEE Computer Graphics and Applications, 2000, vol. 20, n°6, pp. 6–11.
- JACKSON, C. *Displaying uncertainty with shading*. The American Statistician, 2008, vol. 62, n°4, pp. 340–347.
- JAMES, G. et O’RORKE, J. *Real-Time Glow*. GPU Gems, 2004, pp. 343–361.
- JARGSTORFF, F. *A Framework for Image Processing*. GPU Gems, 2004, pp. 445–467.
- JOHNSON, C. et SANDERSON, A. *A next step : Visualizing errors and uncertainty*. IEEE Computer Graphics and Applications, 2003, vol. 23, n°5, pp. 6–10.
- JONES, D. *Determining and visualizing uncertainty in estimates of fiber orientation from diffusion tensor MRI*. Magnetic Resonance in Medicine, 2003, vol. 49, n°1, pp. 7–12.
- JOSPEH, A., LODHA, S., RENTERIA, J. et PANG, A. *UISURF : Visualizing uncertainty in iso-surfaces*. In Proceedings of Computer Graphics and Imaging, 1999. pp. 184–191.
- JOURNAL, A. *The deterministic side of geostatistics*. Mathematical Geology, 1985, vol. 17, n°1, pp. 1–15.
- JOURNAL, A., GUNDESO, R., GRINGARTEN, E. et YAO, T. *Stochastic modelling of a fluvial reservoir : a comparative review of algorithms*. Journal of Petroleum Science and Engineering, 1998, vol. 21, n°1, pp. 95–121.
- JOURNAL, A. et ISAAKS, E. *Conditional indicator simulation : application to a Saskatchewan uranium deposit*. Mathematical Geology, 1984, vol. 16, n°7, pp. 685–718.
- KAJIYA, J. et VON HERZEN, B. *Ray tracing volume densities*. ACM SIGGRAPH Computer Graphics, 1984, vol. 18, n°3, pp. 165–174.

-
- KANUNGO, T., MOUNT, D., NETANYAHU, N., PIATKO, C., SILVERMAN, R. et WU, A. *An efficient k-means clustering algorithm : Analysis and implementation*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, vol. 24, n°7, pp. 881–892.
- KAY, T. et KAJIYA, J. *Ray tracing complex scenes*. ACM SIGGRAPH Computer Graphics, 1986, vol. 20, n°4, pp. 269–278.
- KINDLMANN, G. et DURKIN, J. W. *Semi-automatic generation of transfer functions for direct volume rendering*. IEEE Symposium on Volume Visualization, 1998, pp. 79–86.
- KNISS, J., KINDLMANN, G. et HANSEN, C. *Interactive volume rendering using multi-dimensional transfer functions and direct manipulation widgets*. Proceedings IEEE Visualization 2001, 2001, pp. 255–262.
- KNISS, J., PREMOZE, S., HANSEN, C., SHIRLEY, P. et MCPHERSON, A. *A model for volume lighting and modeling*. IEEE Transactions on Visualization and Computer Graphics, 2003, pp. 150–162.
- KNISS, J., UITERT, R. V., STEPHENS, A., LI, G.-S., TASDIZEN, T. et HANSEN, C. *Statistically quantitative volume visualization*. IEEE Visualization 2005 Proceedings, 2005.
- KOBUS, D., PROCTOR, S. et HOLSTE, S. *Effects of experience and uncertainty during dynamic decision making*. International Journal of Industrial Ergonomics, 2001, vol. 28, n°5, pp. 275–290.
- KOSARA, R., HEALEY, C., INTERRANTE, V., LAIDLAW, D. et WARE, C. *Thoughts on user studies : Why, how, and when*. IEEE Computer Graphics and Applications, 2003, vol. 23, n°4, pp. 20–25.
- KOSARA, R., MIKSCH, S. et HAUSER, H. *Semantic depth of field*. Proceedings of the IEEE symposium on Information Visualization, 2001, pp. 97–104.
- KRIGE, D. *Two-dimensional weighted moving average trend surfaces for ore-evaluation*. Journal of the South African Institute of Mining and Metallurgy, 1966, vol. 66, pp. 13–38.
- LASSETER, J. *Principles of traditional animation applied to 3D computer animation*. In Proceedings of the 14th annual conference on Computer graphics and interactive techniques, 1987. ACM, pp. 35–44.
- LAURENT, G., VIARD, T. et CAUMON, G. *Hardware-Accelerated Isosurface-Based Volume Rendering of Stratigraphic Grids*. In Proceedings of the 29th Gocad Meeting, 2009.
- LEDEZ, D. *Modélisation d’objets naturels par formulation implicite*. Th. Doct.. Institut National Polytechnique de Lorraine, 2002. URL http://www.gocad.org/www/people/files/thesis_david_ledez.pdf.

- LEE, C. et VARSHNEY, A. *Representing thermal vibrations and uncertainty in molecular surfaces*. In SPIE Conference on Visualization and Data Analysis, 2002. pp. 80–90.
- LEE, Y., MARKOSIAN, L., LEE, S. et HUGHES, J. *Line drawings via abstracted shading*. ACM Transactions on Graphics (TOG), 2007, vol. 26, n°3, pp. 18.
- LEITNER, M. et BUTTENFIELD, B. *Guidelines for the display of attribute certainty*. Cartography and Geographic Information Science, 2000, vol. 27, n°1, pp. 3–14.
- LEVOY, M. *Display of surfaces from volume data*. IEEE Computer Graphics and Applications, 1988, vol. 8, n°3, pp. 29–37.
- LEVOY, M. *Efficient ray tracing of volume data*. ACM Transactions on Graphics (TOG), 1990, vol. 9, n°3, pp. 261.
- LÉVY, B., CAUMON, G., CONREAUX, S. et CAVIN, X. *Circular incident edge lists : a data structure for rendering complex unstructured grids*. In ERTL, T., JOY, K. et VARSHNEY, A. (Eds), Proceedings of IEEE Visualization, 2001. pp. 191–198.
- LEWINER, T., LOPES, H., VIEIRA, A. et TAVARES, G. *Efficient implementation of marching cubes' cases with topological guarantees*. Journal of Graphics Tools, 2003, vol. 8, n°2, pp. 1–16.
- LI, W., MUELLER, K. et KAUFMAN, A. *Empty space skipping and occlusion clipping for texture-based volume rendering*. In Proceedings of the 14th IEEE Visualization 2003 (VIS'03), 2003. IEEE Computer Society, pp. 317–325.
- LODHA, S., CHARANIYA, A., FAALAND, N. et RAMALINGAM, S. *Visualization of spatio-temporal GPS uncertainty within a GIS environment*. In Proceedings of SPIE, 2002. pp. 216.
- LODHA, S., FAALAND, N., CHARANIYA, A., VARSHNEY, P., MEHROTRA, K. et MOHAN, C. *Visualization of uncertain particle movement*. In Computer Graphics and Imaging conference, 2002.
- LODHA, S., PANG, A., SHEEHAN, R. et WITTENBRINK, C. *UFLOW : Visualizing uncertainty in fluid flow*. In Proceedings of the 7th conference on Visualization'96, 1996. IEEE Computer Society Press Los Alamitos, CA, USA, pp. 249–254.
- LODHA, S., SHEEHAN, B., PANG, A. et WITTENBRINK, C. *Visualizing geometric uncertainty of surface interpolants*. In Graphics Interface, 1996. pp. 238–245.
- LODHA, S., WILSON, C. et SHEEHAN, R. *LISTEN : sounding uncertainty visualization*. In Proceedings of the 7th conference on Visualization'96, 1996. IEEE Computer Society Press, pp. 189–195.

-
- LORENSEN, W. et CLINE, H. *Marching cubes : A high resolution 3D surface construction algorithm*. ACM SIGGRAPH'87, 1987, pp. 163–169.
- LORENZ, H. et DÖLLNER, J. *Dynamic mesh refinement on GPU using geometry shaders*. Proceedings of the 16-th WCSG, 2008, pp. 97–104.
- LOWE, R. *Extracting information from an animation during complex visual learning*. European journal of psychology of education, 1999, vol. 14, n°2, pp. 225–244.
- LUNDSTRÖM, C., LJUNG, P., PERSSON, A. et YNNERMAN, A. *Uncertainty visualization in medical volume rendering using probabilistic animation*. IEEE Transactions on Visualization and Computer Graphics, 2007, vol. 13, n°6, pp. 1648–1655.
- LUO, A., KAO, D. et PANG, A. *Visualizing spatial distribution data sets*. In Proceedings of the symposium on Data visualisation 2003, 2003. Eurographics Association, pp. 38.
- MAC EACHREN, A. *Visualizing uncertain information*. Cartographic Perspective, 1992, vol. 13, pp. 10–19.
- MAC EACHREN, A. *How maps work*. Guilford Press New York, 1995.
- MAC EACHREN, A., BREWER, C. et PICKLE, L. *Visualizing georeferenced data : representing reliability of health statistics*. Environment and Planning A, 1998, vol. 30, pp. 1547–1562.
- MAC EACHREN, A., HOWARD, D., VON WYSS, M., ASKOV, D. et TAORMINO, T. *Visualizing the health of Chesapeake Bay : An uncertain endeavor*. In GIS LIS, 1993. American Society for Photogrammetry and Remote Sensing, pp. 449–449.
- MACEDONIA, M. *The GPU enters computing's mainstream*. Computer, 2003, vol. 36, n°10, pp. 106–108.
- MACQUEEN, J. *Some methods for classification and analysis of multivariate observations*, 1966.
- MANDELBROT, B. *The fractal geometry of nature*. Wh Freeman, 1982.
- MANN, C. *Uncertainty in geology*. In DAVIS, J. et HERZFELD, U. (Eds), Computers in geology—25 years of progress, 1993. Oxford University Press, Inc., pp. 241–254.
- MANN, H. et WHITNEY, D. *On a test of whether one of two random variables is stochastically larger than the other*. The Annals of Mathematical Statistics, 1947, vol. 18, n°1, pp. 50–60.
- MAO, X., HONG, L. et KAUFMAN, A. *Splatting of curvilinear volumes*. In Proceedings of the 6th conference on Visualization'95, 1995. IEEE Computer Society Washington, DC, USA, pp. 61–68.

- MASIN, S. *Test of models of achromatic transparency*. Perception, 2006, vol. 35, n°12, pp. 1611–1624.
- MASSONNAT, G. *Can We Sample the Complete Geological Uncertainty Space in Reservoir-Modeling Uncertainty Estimates ?* SPE Journal, 2000, vol. 5, n°1, pp. 46–59.
- MATHERON, G. *Le krigeage universel*. Ecole nationale supérieure des mines de Paris, 1969.
- MATTIS, P. et KIMBALL, S. *Gimp, the GNU image manipulation program*, 2008. URL <http://www.gimp.org/>.
- MAX, N. *Optical models for direct volume rendering*. IEEE Transactions on Visualization and Computer Graphics, 1995, vol. 1, n°2, pp. 99–108.
- MCGRANAGHAN, M. *A cartographic view of spatial data quality*. Cartographica : The International Journal for Geographic Information and Geovisualization, 1993, vol. 30, n°2, pp. 8–19.
- METROPOLIS, N. et ULAM, S. *The Monte Carlo method*. Journal of the American Statistical Association, 1949, vol. 44, n°247, pp. 335–341.
- MOSEGAARD, K. et TARANTOLA, A. *Monte Carlo sampling of solutions to inverse problems*. Journal of Geophysical Research, 1995, vol. 100, n°B7, pp. 12–431.
- MOYEN, R., MALLET, J., FRANK, T., LEFLON, B. et ROYER, J. *3D-parameterization of the 3D geological space : the GeoChron model*. In Proceedings of the European Conference on the Mathematics of Oil Recovery IX, Cannes, France, 2004.
- MUELLER, K., MÖLLER, T. et CRAWFIS, R. *Splatting without the blur*. In Proceedings of the conference on Visualization'99 : celebrating ten years, 1999. IEEE Computer Society Press Los Alamitos, CA, USA, pp. 363–370.
- NEWMAN, T. et LEE, W. *On visualizing uncertainty in volumetric data : techniques and their evaluation*. Journal of Visual Languages & Computing, 2004, vol. 15, n°6, pp. 463–491.
- OLSTON, C. et MACKINLAY, J. *Visualizing data with bounded uncertainty*. Information Visualization, IEEE Symposium, 2002, pp. 37–40. ISSN 1522-404X.
- PANG, A. *Visualizing uncertainty in natural hazards*. Risk Assessment, Modeling and Decision Support, 2006, vol. 14, pp. 261–294.
- PANG, A. et ALPER, N. *Bump mapped vector fields*. In SPIE & IS&T Conference Proceedings on Electronic Imaging, 1995. pp. 78–86.

-
- PANG, A., WITTENBRINK, C. et LODHA, S. *Approaches to uncertainty visualization*. The Visual Computer, 1997, vol. 13, n°8, pp. 370–390.
- PANHALEUX, S., VIARD, T. et CAUMON, G. *An image clustering framework to simplify multi-realization output*. In Proceedings of the 30th Gocad Meeting, 2010.
- PHONG, B. *Illumination for computer-generated pictures*. Communications of the ACM, 1975, vol. 18, n°6, pp. 311–317.
- PIQUET, G., VIARD, T. et CAUMON, G. *Particle based tensor field visualization*. In Proceedings of the 30th Gocad Meeting, 2010.
- PLEWA, T., LINDE, T., WEIRS, G. et OTHERS. *Adaptive mesh refinement-theory and applications*. Springer-Verlag Berlin Heidelberg, 2005.
- PORTER, T. et DUFF, T. *Compositing digital images*. ACM SIGGRAPH Computer Graphics, 1984, vol. 18, n°3, pp. 253–259.
- PROKHOROV, Y. *Convergence of random processes and limit theorems in probability theory*. Theory of Probability and its Applications, 1956, vol. 1, pp. 157.
- PRUSINKIEWCZ, P. et SANDNESS, G. *Koch curves as attractors and repellers*. IEEE Computer Graphics and applications, 1988, vol. 8, n°6, pp. 26–40.
- RHEINGANS, P. et JOSHI, S. *Visualization of molecules with positional uncertainty*. In GRÖLLER, E., LÖFFELMANN, H. et RIBARSKY, W. (Eds), Data Visualization '99. Springer-Verlag Wien, 1999. pp. 299–306. ISBN 3-211-83344-7.
- RHEINGANS, P. et LANDRETH, C. *Perceptual principles for effective visualizations*. In GRINSTEIN, G. et LEVKOWITZ, H. (Eds), Perceptual Issues in Visualization. Springer-Verlag, 1995. pp. 59–73.
- RHODES, P., LARAMEE, R., BERGERON, R. et SPARR, T. *Uncertainty visualization methods in isosurface rendering*. EUROGRAPHICS 2003 Short Papers, 2003, pp. 83–88.
- RIVEIRO, M. *Cognitive evaluation of uncertainty visualization methods for decision making*. In Proceedings of the 4th symposium on Applied perception in graphics and visualization, 2007. ACM, pp. 133.
- ROBINSON, A. *Highlighting techniques to support geovisualization*. In Proceedings of the ICA Workshop on Geovisualization and Visual Analytics, 2006.
- ROETTGER, S., BAUER, M. et STAMMINGER, M. *Spatialized transfer functions*. In Proc. of IEEE/Eurographics Symposium on Visualization (EuroVis), 2005. pp. 271–278.
- ROST, R. *OpenGL(R) shading language*. Addison-Wesley, 2006. ISBN 0321334892/9780321334893.

- ROTH, R. *A qualitative approach to understanding the role of geographic information uncertainty during decision making*. Cartography and Geographic Information Science, 2009, vol. 36, n°4, pp. 315–330.
- SABELLA, P. *A rendering algorithm for visualizing 3D scalar fields*. In Proceedings of the 15th annual conference on Computer graphics and interactive techniques, 1988. ACM, pp. 51–58.
- SAMBRIDGE, M., BEGHEIN, C., SIMONS, F. et SNIEDER, R. *How do we understand and visualize uncertainty?* The Leading Edge, 2006, vol. 25, n°5, pp. 542–546.
- SANYAL, J., ZHANG, S., BHATTACHARYA, G., AMBURN, P. et MOORHEAD, R. *A user study to compare four uncertainty visualization methods for 1D and 2D datasets*. IEEE Transactions on Visualization and Computer Graphics, 2009, vol. 15, n°6, pp. 1209–1218.
- SCHEIDT, C., ZABALZA-MEZGHANI, I., FERAILLE, M. et COLLOMBIER, D. *Toward a Reliable Quantification of Uncertainty on Production Forecasts : Adaptive Experimental Designs*. Oil & Gas Science and Technology - Revue de l'IFP, 2007, vol. 62, n°2, pp. 207–224.
- SCHMIDT, G., CHEN, S., BRYDEN, A., LIVINGSTON, M., OSBORN, B. et ROSENBLUM, L. *Multi-dimensional visual representations for underwater environmental uncertainty*. IEEE Computer Graphics & Applications, 2004, pp. 56–65.
- SHEPARD, D. *A two-dimensional interpolation function for irregularly-spaced data*. Proceedings of the 1968 23rd ACM National Conference, 1968, pp. 517–524.
- SHIRLEY, P. et TUCHMAN, A. *A polygonal approximation to direct scalar volume rendering*. ACM SIGGRAPH Computer Graphics, 1990, vol. 25, n°5, pp. 63–70.
- SHORTRIDGE, B. *Stimulus processing models from psychology : can we use them in cartography?* Cartography and Geographic Information Science, 1982, vol. 9, n°2, pp. 155–167.
- SINGH, M. et ANDERSON, B. *Photometric determinants of perceived transparency*. Vision research, 2006, vol. 46, n°6-7, pp. 879–894.
- SKEELS, M., LEE, B., SMITH, G. et ROBERTSON, G. *Revealing uncertainty for information visualization*. In Proceedings of the working conference on Advanced visual interfaces, 2008. ACM, pp. 376–379.
- SMITH, A. *Color gamut transform pairs*. Proceedings of the 5th annual conference on Computer graphics and interactive techniques, 1978, pp. 12–19.

-
- SRIVASTAVA, R. *Reservoir characterization with probability field simulation*. In SPE Annual Technical Conference and Exhibition, 1992.
- SRIVASTAVA, R. *The visualization of spatial uncertainty*. The Visualization of Spatial Uncertainty, Stochastic Modeling and Geostatistics : Principles, Methods, and Case Studies, J.M Yarus and R.L. Chambers, eds., American Assoc. of Petroleum Geologists, 1994, pp. 339–345.
- SRIVASTAVA, R. *An overview of stochastic methods for reservoir characterization*. AAPG Computer Applications in Geology, Tulsa, Yarus J. and Chambers R. eds., 1995. Tome 3.
- SRIVASTAVA, R. M. *The interactive visualization of spatial uncertainty*. SPE 27965, 1994.
- STEIN, M. *Interpolation of Spatial Data : some theory for kriging*. Springer Verlag, 1999.
- STROTHOTTE, T., PUHLE, M., MASUCH, M., FREUDENBERG, B., KREIKER, S., LUDOWICI, B. et ZENTRUM, G. *Visualizing uncertainty in virtual reconstructions*, 1999.
- TARANTOLA, A. *Inverse problem theory*. Elsevier, 1987.
- TATARCHUK, N., SHOPF, J. et DECORO, C. *Real-time isosurface extraction using the GPU programmable geometry pipeline*. In ACM SIGGRAPH 2007 courses, 2007. ACM, pp. 137.
- TERTOIS, A. et MALLET, J. *Distance maps and virtual fault blocks in tetrahedral models*. In Proceedings of the 26th Gocad Meeting, 2006.
- THOMSON, J., HETZLER, E., MACEACHREN, A., GAHEGAN, M. et PAVEL, M. *A typology for visualizing uncertainty*. In Proc. SPIE, 2005. pp. 146–157.
- TORY, M. et MÖLLER, T. *Human factors in visualization research*. IEEE Transactions on Visualization and Computer Graphics, 2004, vol. 10, n°1.
- TREECE, G., PRAGER, R. et GEE, A. *Regularised marching tetrahedra : improved isosurface extraction*. Computers & Graphics, 1999, vol. 23, n°4, pp. 583–598.
- TUFTE, E. *The visual display of quantitative information*. Graphics Press, Cheshire, 1983.
- TUFTE, E. *The visual display of quantitative information, 2nd ed*. Graphics Press, Cheshire, 2001.
- UPTON, G. et COOK, I. *Understanding statistics*. Oxford University Press, 1996.
- USGS. *Spatial data transfer standard (SDTS) : Logical specifications*, 1977.
- VAN DER VAART, A. et WELLNER, J. *Weak convergence and empirical processes*. Springer Verlag, 1996.

- VIARD, T., CAUMON, G., GRINGARTEN, E., HUGOT, A. et LÉVY, B. *3D Visualization of Dynamic Reservoir Uncertainty*. In Proceedings of the 2009 Winter Gocad Meeting, Doha, 2009.
- VIARD, T., CAUMON, G., GRINGARTEN, E., HUGOT, A. et LÉVY, B. *New perspectives for 3D visualization of dynamic uncertainties*. In International Petroleum and Technology Conference (IPTC), Doha, 2009.
- VIARD, T., CAUMON, G., GRINGARTEN, E., HUGOT, A. et LÉVY, B. *Visualization of dynamic uncertainties*. In Proceedings of the 30th Gocad Meeting, 2010.
- VIARD, T., CAUMON, G. et LÉVY, B. *Uncertainty Visualization in Geological Grids*. In Proceedings of the 28th Gocad Meeting, 2008.
- VIARD, T., CAUMON, G. et LÉVY, B. *An implementation of Projected Tetrahedra for Volume Rendering with Uncertainty*. In Proceedings of the 29th Gocad Meeting, 2009.
- VIARD, T., CAUMON, G. et LÉVY, B. *Visualization of Uncertainty on 3D geological models using Blur and Textures*. In Proceedings of the 29th Gocad Meeting, 2009.
- VIARD, T., CAUMON, G. et LÉVY, B. *Adjacent versus coincident representations of geospatial uncertainty : Which promote better decisions ?* Accepté dans Computers and Geosciences, 2010.
- VIARD, T., CHERPEAU, N., CAUMON, G. et LÉVY, B. *Structural uncertainty visualization : ghost surfaces and animation of level sets*. In Proceedings of the 30th Gocad Meeting, 2010.
- VIARD, T., CZERNUSZENKO, M. et ADAIR, N. *Weighted Animation for Uncertainty Visualization*. In Proceedings of the 2009 Fall Gocad Meeting, Houston, 2009.
- VIARD, T., CZERNUSZENKO, M. et ADAIR, N. *Weighted animation for uncertainty visualization*. ExxonMobil, 2009.
- VIARD, T., CZERNUSZENKO, M. et ADAIR, N. *Animation methods for uncertainty visualization : a user study*. Soumis à IEEE Transactions on Visualization and Computer Graphics, 2010.
- VIARD, T., ROYER, J., CAUMON, G. et LÉVY, B. *Visualisation des Incertitudes Associées à un Modèle Géologique*. In Réunion des Sciences de la Terre (RST), Nancy, 2008.
- VIARD, T., ROYER, J., CAUMON, G. et LÉVY, B. *Visualization of Geological Uncertainties along Cartesian and Stratigraphic Grids*. In International Geological Congress (IGC), Oslo, 2008.

-
- VON KOCH, H. *Une méthode géométrique élémentaire pour l'étude de certaines questions de la théorie des courbes planes*. Acta mathematica, 1906, vol. 30, n°1, pp. 145–174.
- WALTZ, F. et MILLER, J. *An efficient algorithm for Gaussian blur using finite-state machines*. In SPIE Conf. on Machine Vision Systems for Inspection and Metrology VII, 1998.
- WARE, C. *Information visualization : Perception for design, 2nd ed.* Morgan Kaufmann Publishers, 2004.
- WESTERMANN, R. et ERTL, T. *Efficiently using graphics hardware in volume rendering applications*. In Proc. of SIGGRAPH, 1998. pp. 169–178.
- WILCOXON, F. *Individual comparisons by ranking methods*. Biometrics Bulletin, 1945, vol. 1, n°6, pp. 80–83.
- WILLIAMS, P. *Visibility-ordering meshed polyhedra*. ACM Transactions on Graphics (TOG), 1992, vol. 11, n°2, pp. 103–126.
- WILLIAMS, P. et MAX, N. *A volume density optical model*. Proceedings of the 1992 workshop on Volume visualization, 1992, pp. 61–68.
- WITKIN, A. et POPOVIC, Z. *Motion warping*. In Proceedings of the 22nd annual conference on Computer graphics and interactive techniques, 1995. ACM, pp. 105–108.
- WITTENBRINK, C. *IFS fractal interpolation for 2D and 3D visualization*. In Proceedings of the 6th conference on Visualization'95, 1995. IEEE Computer Society Washington, DC, USA.
- WITTENBRINK, C. *CellFast : Interactive unstructured volume rendering*. IEEE Conference on Visualization (Vis'99), Late Breaking Hot Topics, 1999.
- WITTENBRINK, C., PANG, A. et LODHA, S. *Verity visualization : Visual mappings*. Tech. Report, UCSC-CRL-95-48, 1995.
- WITTENBRINK, C., SAXON, E., FURMAN, J., PANG, A. et LODHA, S. *Glyphs for visualizing uncertainty in environmental vector fields*. SPIE and IS and T Conference Proceedings on Electronic Imaging : Visual Data Exploration and Analysis, 1995, pp. 87–100.
- WITTENBRINK, C., SAXON, E., FURMAN, J., PANG, A. et LODHA, S. *Glyphs for visualizing uncertainty in environmental vector fields*. IEEE Transactions on Visualization and Computer Graphics, 1996, vol. 2, n°3, pp. 266–279.
- WRIGHT, J. *Map makers are human : Comments on the subjective in maps*. Geographical Review, 1942, pp. 527–544.

- WYLIE, B., MORELAND, K., FISK, L. et CROSSNO, P. *Tetrahedral projection using vertex shaders*. Proceedings of IEEE Volume Visualization and Graphics Symposium 2002, 2002, pp. 7–12.
- YAGEL, R., REED, D., LAW, A., SHIH, P. et SHAREEF, N. *Hardware assisted volume rendering of unstructured grids by incremental slicing*. In ACM Symposium on Volume Visualization, 1996. pp. 55–63.
- ZABALZA-MEZGHANI, I., MANCEAU, E., FERAILLE, M. et JOURDAN, A. *Uncertainty management : From geological scenarios to production scheme optimization*. Journal of Petroleum Science and Engineering, 2004, vol. 44, n°1-2, pp. 11–25.
- ZEHNER, B., WATANABE, N. et KOLDITZ, O. *Visualization of gridded scalar data with uncertainty in geosciences*. Computers & Geosciences, 2010.
- ZUK, T. *Visualizing Uncertainty*. Th. Doct.. Department of Computer Science, University of Calgary, Calgary, 2008. URL http://innovis.cpsc.ucalgary.ca/innovis/uploads/Publications/Publications/Zuk_2008_Visualizing_Uncertainty.pdf.
- ZUK, T. et CARPENDALE, S. *Theoretical analysis of uncertainty visualization*. Proceedings of SPIE. Vol. SPIE-6060, 2006, pp. 66–79.
- ZUK, T., CARPENDALE, S. et GLANZMAN, W. *Visualizing temporal uncertainty in 3D virtual reconstructions*. In VAST, 2005. pp. 6th.
- ZUK, T., DOWNTON, J., GRAY, D., CARPENDALE, S. et LIANG, J. *Exploration of uncertainty in bidirectional vector fields*. In KATY BÖRNER, Matti T. Gröhn, J. P. et ROBERTS, J. C. (Eds), Proceedings of the SPIE & IS&T Conference on Electronic Imaging, 2008.
- ZWICKER, M., VAN BAAR, J. et GROSS, M. *EWA volume splatting*. In Proceedings of the conference on Visualization'01, 2001. IEEE Computer Society Washington, DC, USA, pp. 29–36.

AUTORISATION DE SOUTENANCE DE THESE
DU DOCTORAT DE L'INSTITUT NATIONAL
POLYTECHNIQUE DE LORRAINE

o0o

VU LES RAPPORTS ETABLIS PAR :

Monsieur Georges-Pierre BONNEAU, Professeur, Université Joseph Fourier, LJK, INRIA Rhône-Alpes, Montbonnot

Monsieur Klaus MOSEGAARD, Professeur, Technical University of Denmark, Denmark

Le Président de l'Institut National Polytechnique de Lorraine, autorise :

Monsieur VIARD Thomas

à soutenir devant un jury de l'INSTITUT NATIONAL POLYTECHNIQUE DE LORRAINE,
une thèse intitulée :

"Algorithmes de visualisation des incertitudes en géomodélisation sur GPU"

en vue de l'obtention du titre de :

DOCTEUR DE L'INSTITUT NATIONAL POLYTECHNIQUE DE LORRAINE

Spécialité : « **Géosciences** »

Fait à Vandoeuvre, le 23 septembre 2010

Le Président de l'I.N.P.L.,

F. LAURENT



NANCY BRABOIS
2, AVENUE DE LA
FORET-DE-HAYE
BOITE POSTALE 3
F - 5 4 5 0 1
VANDŒUVRE CEDEX

Résumé

En géosciences, la majeure partie du sous-sol est inaccessible à toute observation directe. Seules des informations parcellaires ou imprécises sont donc disponibles lors de la construction ou de la mise à jour de modèles géologiques ; de ce fait, les incertitudes jouent un rôle fondamental en géomodélisation. La théorie des problèmes inverses et les méthodes de simulations stochastiques fournissent un cadre théorique permettant de générer un ensemble de représentations plausibles du sous-sol, également nommées réalisations. En pratique, la forte cardinalité de l'ensemble des réalisations limite significativement tout traitement ou interprétation sur le modèle géologique.

L'objectif de cette thèse est de fournir au géologue des algorithmes de visualisation permettant d'explorer, d'analyser et de communiquer les incertitudes spatiales associées à de larges ensembles de réalisations. Nos contributions sont les suivantes : (1) Nous proposons un ensemble de techniques dédiées à la visualisation des incertitudes pétrophysiques. Ces techniques reposent sur une programmation sur carte graphique (GPU) et utilisent une architecture garantissant leur interopérabilité ; (2) Nous proposons deux techniques dédiées à la visualisation des incertitudes structurales, traitant aussi bien les incertitudes géométriques que les incertitudes topologiques (existence de la surface ou interactions avec d'autres surfaces) ; (3) Nous évaluons la qualité des algorithmes de visualisation des incertitudes par le biais de deux études sur utilisateurs, portant respectivement sur la perception des méthodes statiques et par animation. Ces études apportent un éclairage nouveau sur la manière selon laquelle l'incertitude doit être représentée.

Mots-clés: Incertitude, géomodélisation, corendu, perception, GPU

Abstract

Most of the subsurface is inaccessible to direct observation in geosciences. Consequently, only local or imprecise data are available when building or updating a geological model ; uncertainties are therefore central to geomodeling. The inverse problem theory and the stochastic simulation methods provide a framework for the generation of large sets of likely representations of the subsurface, also termed realizations. In practice, however, the size of the set of realizations severely impacts further interpretation or processing of the geological model.

This thesis aims at providing visualization algorithms to expert geologists that allow them to explore, analyze and communicate on spatial uncertainties associated to large sets of realizations. Our contributions are : (1) We propose a set of techniques dedicated to petrophysical uncertainty visualization, based on a GPU programming approach that maintains their interoperability ; (2) We propose two techniques dedicated to structural uncertainty visualization that can handle both geometrical and topological uncertainties (e.g., the existence of the surface or its relationships with other surfaces) ; (3) We assess the quality of our uncertainty visualization algorithms through two user studies, which respectively focus on the perception of static and animated methods. These studies bring new elements on how uncertainty should be represented.

Keywords: Uncertainty, geomodeling, corendering, perception, GPU

