



HAL
open science

Modélisation de l'incertitude géologique par simulation stochastique de cubes de proportions de faciès - Application aux réservoirs pétroliers de type carbonaté ou silico-clastique

Abbas Zerkoune

► **To cite this version:**

Abbas Zerkoune. Modélisation de l'incertitude géologique par simulation stochastique de cubes de proportions de faciès - Application aux réservoirs pétroliers de type carbonaté ou silico-clastique. Modélisation et simulation. Université Joseph-Fourier - Grenoble I, 2009. Français. NNT: . tel-00410136

HAL Id: tel-00410136

<https://theses.hal.science/tel-00410136>

Submitted on 18 Aug 2009

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.



École Doctorale Terre, Univers, Environnement
Laboratoire de Géodynamique des Chaînes Alpines



Thèse de doctorat de l'université Joseph Fourier, Grenoble I

Modélisation de l'incertitude géologique par simulation stochastique de cubes de proportions de faciès : Application aux réservoirs pétroliers de type carbonaté ou silico-clastique

Abbas ZERKOUNE

Thèse soutenue le lundi 08 juin 2009

Directeurs de la thèse : MM. É. LEWIN, Physicien adjoint d'observatoire, Univ. J Fourier

Rapporteurs : MM. C. ROBIN, Maître de Conférence, Géosciences Rennes
MM. A. MAS, Professeur d'Université, Univ. Montpellier II

Président du jury : MM. É. JAILLARD, DR2 de l'IRD, Univ. J Fourier

Examineurs : MM. F. ROGGERO, Senior IFP Research Engineer, IFP
MM. H. WACKERNAGEL, Maître de Recherche, Mines Paris
MM. G. MASSONNAT, Expert Géologie de Gisement, Total

Remerciements

Voilà c'est ma thèse, une parmi d'autre avec son lot de remerciements. Réalisée entre Grenoble au LGCA, Pau au CSTJF et Saint Ouen l'Aumône à la base. Merci à toutes les personnes que j'ai fréquentées là-bas, de près ou de loin depuis ces temps lointain de janvier 2004.

Je remercie tout particulièrement Gérard Massonnat, panier d'idées à toute heure et directeur officieux de cette thèse. Merci à lui de m'avoir accueilli dans son équipe, de la confiance accordée et de cette touche d'autonomie dirigée qu'il a su insuffler à mon travail.

Je remercie aussi Annie Arnaud-Vanneau et Éric Lewin, mes directeurs académiques dans ce projet, qui ont fait énormément pour qu'il aille au bout, qui ont encadré en partie ces travaux. Un merci tout particulier pour leur patience aux responsables l'école doctorale TUE que j'ai usés, ainsi qu'à celle qui jouait pour moi la figure de l'administration, Céline Legrenzi qui a réussi à m'inscrire un bon nombre de fois, celle-là, c'était la dernière et ça valait le coup.

Ce projet est le fruit de l'expérience tirée des discussions et des problématiques soulevées par Fanny dont certains des modèles ont servi à faire une partie des illustrations de cette thèse, Bruno et Pierre Biver. Il a pris la forme d'une application utilisable par tous (normalement) grâce à Jean-Pull et Dibi qui ont eu le courage de les programmer. Ces travaux ont fait l'objet d'un dépôt de brevet dont Emmanuelle s'est chargée, merci. Comment ne pas remercier Myriam, qui a su gérer la partie logistique de mes séjours à Grenoble, même au dernier moment, bravo !

Un merci gros comme la mer où se jette le Nil à ma camarade préférée de bureau, Fanny, pour sa bonne humeur et toutes les discussions professionnelles ou surtout pas que l'on a eues. Au passage, merci à mes partenaires d'exploits sportifs, camarade Flonflon, Shah Mahmoud et Éric, et de souvenirs festifs : en particulier le Basque, Julien, Tibo, Pitou. Merci à ceux qui ont osé me prêter, une voiture ou un tracteur tondeuse, Enrico et Toto, au seul à connaître Yvon Leroux ou le nom de la femme de l'acteur qui joue dans CHIPS : Laurent et sa culture alternative. Bref, merci à tous ceux qui ont partagé mes journées de travail et pas que. Sans oublier toutes les brigades de stagiaires, apprentis et contractants qui se sont succédées, du BB au BA, et qui ont animés la vie au CSTJF et à Pau. Je cite au risque d'en oublier le père David, la Stache, la confrérie des Mathieu, le roi des Mouches, Lyes, Vahiiiiid, Mehdi, Gabi, Florence, Laurence... Merci également à tous les footeux du stade Blanchard, témoins privilégiés de mes occasions d'exceptions. À Grenoble, merci à tous ceux que j'ai plus ou moins côtoyés, en particulier le clan nantais : Marion et François, qui m'a fait apprécié mes séjours là-bas. Merci aussi au RU de Grenoble ? Bah non, pas le RU, plutôt la Gentille dame à la caisse du RU de Total.

Et maintenant la séquence dédicaces : une pensée à ma famille que j'ai un peu moins vue toutes ces années : mes frères Hassan, Samir et son épouse Sonia, mes sœurs Sabrina et Leïla, et bien évidemment mon père, Monsieur Zerkoune le vrai, et Maman Zerkoune.

à mon père, ma mère, mes frères et mes soeurs

PEANUTS



Résumé

Après sa découverte, les choix relatifs au développement d'un gisement se prennent sur la base de représentations incertaines du champ. Sa caractérisation utilise en effet des modèles numériques spatiaux du réservoir. Cependant, s'ils traduisent au mieux les hétérogénéités du sous-sol, ils restent porteurs de l'incertitude liée à la complexité du milieu souterrain. D'ordinaire, les méthodes de simulations stochastiques, qui génèrent un ensemble de modèles équiprobables du sous-sol, sont supposées les quantifier. Néanmoins, ces images alternatives du champ renvoient à des tirages au sein d'un modèle probabiliste unique. Elles oublient l'incertitude relative au choix du modèle probabiliste sous-jacent, et tendent à la sous-estimer. Ce travail de recherche vise à améliorer la quantification de ce type d'incertitude durant la modélisation des réservoirs. Elle retranscrit la part de doute relative à la compréhension des propriétés du milieu sur les modèles probabilistes, et propose de l'intégrer sur ces modèles.

Cette thèse précise d'abord la notion d'incertitude en modélisation pétrolière, en particulier sur les modèles géologiques 3D comprenant différents faciès. Leur construction demande avant toute simulation de définir en tout point de l'espace leur probabilité d'existence (ou proportion) : c'est le cube de proportions. Généralement, bien que ces probabilités soient peu connues, les méthodes actuelles d'évaluation de l'incertitude sédimentaire les gardent figées. De fait, elles oublient l'influence du caractère incertain du scénario géologique sur le choix du cube de proportions. Deux méthodes stochastiques de simulation ont été développées afin de générer des modèles équiprobables en termes de cubes de proportions. Elles permettent d'intégrer la variabilité liée aux proportions de faciès, et ainsi d'explorer dans son ensemble un tel domaine d'incertitude.

La première reste relativement attachée à la géologie. Elle intègre directement l'incertitude liée aux paramètres qui composent le scénario géologique. Elle développe une approche multi-réalisation sur les divers paramètres du scénario géologique, qu'ils prennent la forme de signaux aux puits, de cartes ou d'hypothèses plus globales à l'échelle du réservoir. Une démarche de type Monte-Carlo échantillonne ainsi les composantes du schéma sédimentaire. Chaque tirage permet de construire un cube de proportions par l'intermédiaire d'un géomodeleur qui intègre de façon plus ou moins explicite les paramètres du scénario géologique. La méthodologie est illustrée et appliquée à un processus inédit de modélisation des dépôts carbonatés en milieu marin.

La seconde revêt un caractère plus géostatistique en se concentrant davantage sur le cube de proportions. Elle vise plutôt à réconcilier les différents modèles sédimentaires possibles. Dans le modèle maillé représentant le réservoir, elle estime la loi de distribution des proportions de faciès cellule par cellule - supposées suivent une loi de Dirichlet, à partir de quelques modèles, construits sur la base de scénarios géologiques distincts. Elle simule alors les proportions de façon séquentielle, maille après maille, en introduisant une corrélation spatiale (variogramme) qui peut être déterministe ou probabiliste. Divers cas pratiques, composés de réservoirs synthétiques ou de champs réels, illustrent et précisent les différentes étapes de la méthode proposée.

Abstract

After finding out a potential oil field, development decisions are based on uncertain representations of the petroleum reservoir. Indeed, its characterisation uses numerical, spatial models of the reservoir. However, if they are representative of subsoil heterogeneities, the uncertainty linked to subsoil complexity remain. Usually, uncertainty is supposed to be assessed using many equiprobable models, which represent the heterogeneities expected into the reservoir. Nevertheless, those alternative images of the underground correspond to multiple realizations of a given and a single stochastic model. Those methods ignore the uncertainty related to the choice of the underlying probabilistic model. This work aims at improving that kind of uncertainty assessment when modelling petroleum reservoir. It conveys the doubt linked with our subsoil properties understanding on probabilistic models, and proposes to integrate it on them.

This thesis first defines uncertainty in the context of oil industry modelling, particularly on 3D geological models comprising several litho-types or facies. To build them, we need, before any simulations, to estimate for every point in the space the probability of occurring for each facies : this is the proportions cube. Even though those probabilities are often poorly known, they are frozen while using current methods of uncertainty assessment. So, the impact of an uncertain geological scenario on the definition of a proportion cube is forgotten. Two methods based on stochastic simulations of alternative, equiprobable proportion cubes have been developed to sample the complete geological uncertainty space.

The first one is closely linked to geology. It integrates directly uncertainty related to the parameters composing the geological scenario. Based on a multi-realisation approach, it describes its implementation on every parameters of geological scenario from information at wells to maps or global hypothesis at reservoir scale resolution. A Monte Carlo approach samples the components of the sedimentary scheme. Each drawing enables to build a proportion cube using modelling tools which integrates more or less explicitly parameters of geological scenario. That methodology is illustrated and applied to a modelling process which is used to model marine carbonate deposits.

The second method appears to be more ‘geostatistics’ focussing on proportion cubes. It rather aims at reconcile distinct eventual sedimentary models. In the meshed model symbolising the reservoir, it assesses the probabilistic law of facies proportion in each cells – they are supposed to follow Dirichlet’s probabilistic law. That assessment is done from some models inferred from different geological scenarios. Facies proportions are sequentially simulated, cell after cell, introducing a spatial correlation model (variogram), which could be deterministic as probabilistic. Various practical cases, comprising synthetic reservoirs or real field, illustrates and specifies the different steps of the proposed method.

Table des matières

Introduction.....	1
1. Vers une plus grande compréhension du sous-sol	1
2. Objectifs et organisation générale	5
Partie I : Incertitude géologique et modélisation du sous-sol.....	9
Chapitre 1 L’incertitude géologique	13
1.1 Un milieu à l’histoire géologique complexe	13
1.2 Des données incomplètes	14
1.3 Un modèle géologique incertain.....	16
Chapitre 2 Le traitement de l’incertitude géologique	21
2.1 Le cadre probabiliste	21
2.2 Les méthodes statistiques	24
2.2.1 Les simulations par cubature.....	24
2.2.2 Un exemple illustratif.....	25
2.3 Les méthodes statistiques spatiales	27
2.4 Les études à scénarios multiples	32

Partie II : Multi-réalisation sur les paramètres du scénario géologique....37

Chapitre 3 Incertitude et Neptune	41
3.1 Neptune, bases et concepts	41
3.2 Neptune, les paramètres incertains	45
3.3 Neptune, hiérarchie des sources d'incertitudes	48
3.4 Neptune, la multi-réalisation des scénarios géologiques	51
Chapitre 4 Incertitude aux puits (1D)	55
4.1 La bathymétrie aux puits.....	55
4.1.1 La relation bathymétrie/faciès	55
4.1.2 La multi-réalisation du modèle de faciès.....	57
4.2 Les corrélations aux puits	60
Chapitre 5 Incertitude sur les cartes (2D).....	63
5.1 Les signaux sur les surfaces isochrones.....	63
5.1.1 Extrapolation des épaisseurs.....	63
5.1.2 Calcul du potentiel d'accommodation	64
5.1.3 Définition de la carte de bathymétrie de référence	65
5.2 La multi-réalisation des cartes de bathymétrie	67
5.2.1 Une nécessaire adaptation des cartes de base	67
5.2.2 Une approche à scénarios multiples.....	68
5.2.3 Une approche selon les modèles de mélange.....	69
Chapitre 6 Incertitude sur le réservoir (3D)	77
6.1 Modélisation des proportions de faciès.....	77
6.2 Modélisation de la diagenèse précoce.....	81

Conclusions et Perspectives (1)	85
Partie III : Simulations stochastiques de cubes de proportions	89
Chapitre 7 Panorama d'une méthode de simulation des cubes de proportions ...	93
7.1 Introduction	93
7.2 Principe général de la méthode	95
Chapitre 8 Simulation régionalisée des proportions de faciès	99
8.1 Hypothèses de travail	99
8.1.1 Rappel : le cube de proportions de faciès.....	99
8.1.2 Des lois de probabilité ponctuelles	100
8.1.3 Un support géométrique commun	102
8.2 Choix des lois de probabilité ponctuelles.....	104
8.2.1 Des lois paramétriques	104
8.2.2 Contraintes de probabilité	104
8.2.3 Petit historique des lois testées.....	106
8.3 Caractérisation des distributions ponctuelles	112
8.3.1 Estimation des paramètres des lois $P(\alpha)$	112
8.3.2 Retour sur les scénarios de base.....	114
8.3.3 Simulations de vecteurs de proportions régionalisées	115
Chapitre 9 Simulation globale des cubes de proportions	119
9.1 Conservation de la cohérence spatiale.....	119
9.1.1 Où introduire la continuité spatiale ?	119
9.1.2 Recours au cadre géostatistique	121
9.2 En domaine de plate-forme	124
9.2.1 Le modèle de référence	124

9.2.2 Une corrélation spatiale inférée	125
9.2.3 Une corrélation spatiale infinie.....	128
9.3 En domaine chenalisé	130
9.3.1 Le modèle de référence.....	130
9.3.2 Cas pratique n° 1 : Simulation de zones préférentielles pour les chenaux.....	132
9.3.3 Cas pratique n° 2 : Simulation de cartes de proportions via la sismique.....	139
9.3.4 Cas pratique n° 3 : Intégration de plusieurs modèles de dérives.....	141
Chapitre 10 Mise en cohérence des cubes de proportions	149
10.1 Intégration des données de puits	149
10.1.1 Nature des contraintes.....	149
10.1.2 Mise en œuvre sur un exemple	152
10.1.3 Des données dures incertaines !.....	154
10.2 Intégration des cartes sismiques.....	156
10.2.1 Nature des contraintes.....	156
10.2.2 Assimilation des contraintes sismiques	158
10.2.3 Des cartes sismiques incertaines !	159
Chapitre 11 Combinaison de cubes de proportions	161
11.1 Des cubes de proportions d'origines diverses.....	161
11.2 Couplage des lois de Dirichlet.....	162
11.2.1 Objectif	162
11.2.2 Mélange par combinaisons linéaires.....	163
11.2.3 Mélange bayésien	164
Chapitre 12 Régionalisation de l'incertitude	167
12.1 Un cas synthétique	167

12.2 En amont des simulations.....	168
12.3 En aval des simulations	170
12.4 Tri des cubes de proportions	174
Conclusions et Perspectives (2)	177
Annexe A Notations utilisées.....	181
A.1 Abréviations	181
A.2 Typographie :	183
Références bibliographiques	187

Table des Figures

FIG. 1 : La construction d'un géomodèle.	3
FIG. 1.1: La représentativité des données en question.	15
FIG. 1.2 : Scénario géologique et modèle de continuité spatiale.	18
FIG. 2.1 : (a) Notions de variable et (b) de fonction aléatoire.	22
FIG. 2.2 : Cadres déterministe et probabiliste.	23
FIG. 2.3 : Les méthodes par cubature.	26
Fig. 2.4 : Définition du cube de proportions de faciès.	28
FIG. 2.5 : Simulations géostatistiques de la porosité.	29
FIG. 2.6 : Effet de l'incertitude géologique sur le cube de proportions de faciès.	30
FIG. 2.7 : Étude classique des incertitudes sur le réservoir Doushka.	31
FIG. 2.8 : Étude à scénarios multiples du réservoir Doushka.	33
FIG. 2.9 : Intégration des deux niveaux d'incertitude géologique.	34
FIG. 3.1 : Évolution dans le temps des variables associées à l'accommodation.	43
FIG. 3.2 : Les étapes du processus Neptune et les paramètres d'entrée associés.	46
FIG. 3.3 : Principe général de quantification des incertitudes.	49
FIG. 3.4 : Effet des principaux paramètres du scénario géologique sur le cube de proportions.	51
FIG. 3.5 : Principe des réalisations multiples en terme de scénario géologique.	53
FIG. 4.1 : La relation faciès/bathymétrie.	56

Table des figures

FIG. 4.2 : Intégration de l'incertitude sur le contexte de dépôt.....	58
FIG. 4.3 : Intégration de l'incertitude sur l'extension des plages de dépôts.....	59
FIG. 5.1 : Calcul des signaux connus aux puits sur l'ensemble des surfaces isochrones. ...	64
FIG. 5.2 : Calcul de la bathymétrie $\mathbf{b}(\mathbf{x},\mathbf{y},\mathbf{t})$ le long des surfaces isochrones.	65
FIG. 5.3 : Incertitude sur la paléogéographie.....	66
FIG. 5.4 : Contrainte des cartes de référence aux données de puits.....	67
FIG. 5.5 : Une approche discrète pour intégrer l'incertitude sur les modèles de cartes.	69
FIG. 5.6 : Les divers blocs diagrammes envisagés et compatibles avec les données.	70
FIG. 5.7 : Construction de cartes par métamorphose entre couples de cartes.	71
FIG. 5.8 : Combinaison linéaire des cartes de paléogéographie.	73
FIG. 5.9 : Vue en coupe des cartes issues de la combinaison « gaussienne» des cartes de bathymétrie.....	75
FIG. 6.1 : L'incertitude sur les probabilités a priori.	78
FIG. 6.2 : Un mode alternatif de calcul des poids a priori.....	80
FIG. 6.3 : Simulation des faciès.....	81
FIG. 6.4 : Régions soumises à la diagenèse précoce et incertitude sur leur extension.	82
FIG. 6.5 : Principe de l'estimation des zones émergées dans Neptune.....	83
FIG. 6.6 : Intégration de l'incertitude sur le volume de la zone soumise à la diagenèse précoce.	84
FIG. 7.1 : Étapes de la méthode de simulation stochastique des cubes de proportions.	97
FIG. 8.1 : Définition du cube de proportions.....	100
FIG. 8.2 : Définition de la notion de cube de proportions aléatoire.....	101
FIG. 8.3 : Un support géométrique commun à l'ensemble des $n_m(G)$ modèles de départ.	103
FIG. 8.4 : Exemples de simplexes sur lesquels sont simulés les vecteurs de proportions.	105
FIG. 8.5 : Simulation d'un vecteur de proportions $\mathbf{p}(\alpha)$ par ACP (ou vecteur gaussien).	107
FIG. 8.6 : Simulation d'un vecteur de proportions $\mathbf{p}(\alpha)$ avec transformation préalable des données.	109

FIG. 8.7 : Loi de Dirichlet et proportions.....	111
FIG. 8.8 : Estimation des paramètres de la loi de probabilité au point α	113
FIG. 8.9 : Simulation d'un vecteur de proportions au point α	115
FIG. 8.10 : Illustration de la perte de cohérence spatiale.....	117
FIG. 9.1 : La continuité spatiale des cubes de proportions simulées en question.....	120
FIG. 9.2 : Représentation de la fonction variogramme.....	122
FIG. 9.3 : Les cubes de proportions initiaux.....	124
FIG. 9.4 : Inférence de la continuité spatiale à partir des modèles de bases.....	126
FIG. 9.4 : Inférence de la continuité spatiale à partir des modèles de bases.....	127
FIG. 9.6 : Inférence de la continuité spatiale à partir des modèles de bases.....	129
FIG. 9.7 : Définition d'un cube de proportions aléatoires.....	129
FIG. 9.8 : Définition d'un cube de proportions aléatoires.....	131
FIG. 9.9 : Cas pratique n°1 : Incertitude sur les zones préférentielles de chenaux.....	132
FIG. 10.1 : Deux jeux de contraintes possibles aux puits pour les proportions.....	151
FIG. 10.2 : Simulations de quantiles conditionnés aux puits.....	153
FIG. 10.3 : Intégration de l'incertitude sur les données de puits.....	155
FIG. 10.4 : Définition des contraintes sismiques relativement au cube de proportions. ..	157
FIG. 10.5 : Intégration de l'incertitude sur les cartes tirées de la sismique.....	158
FIG. 10.6 : Intégration de l'incertitude sur les cartes tirées de la sismique.....	160
FIG. 11.1 : Combinaison des lois de Dirichlet.....	163
FIG. 12.1 : Présentation des possibilités de régionalisation de l'incertitude.....	168
FIG. 12.2 : Régionalisation de l'incertitude en amont des simulations.....	170
FIG. 12.3 : Régionalisation de l'incertitude en aval des simulations.....	173

Introduction

1. Vers une plus grande compréhension du sous-sol

Pendant la majeure partie de son histoire avec « l'or noir », l'homme ne cherche pas à comprendre les objets naturels que sont les réservoirs pétroliers. Tout simplement parce que ce n'est pas nécessaire : les ressources apparaissant à la surface (émanations de gaz, sable bitumineux, etc.) suffisent à subvenir aux besoins (éclairage, matériaux de construction...). Le XX^{ème} siècle, suite à l'essor de l'automobile, marque les débuts d'une accélération de la demande en produits dérivés du pétrole [Perrodon, 1990 ; Leflon, 2005]. Aujourd'hui, plus que jamais, l'activité humaine repose sur cette matière première. Les besoins s'intensifient avec une consommation en énergie toujours croissante¹, amplifiée depuis quelques années par les pays émergents (Inde, Chine, etc.). En parallèle, les ressources se raréfient, exacerbant d'autant plus la concurrence entre opérateurs [Perrodon, 1999]. Dans un contexte si dur, l'industrie n'a de cesse de produire davantage. A présent, il ne s'agit plus seulement de récupérer les matières premières à la surface de la Terre mais d'en extraire les ressources souterraines. Elles se situent plus précisément dans des roches suffisamment poreuses pour pouvoir contenir les fluides comme l'eau, l'huile ou le gaz, et, qui sont recouvertes par une couche imperméable interdisant à ces mêmes fluides de s'échapper [Foucault et Raoult, 1995].

Finis donc, les temps bénis du pétrole abondant, facile à pomper et bon marché. Les opérateurs du secteur s'intéressent maintenant à des zones de plus en plus difficiles d'accès et complexes. C'est notamment le cas des gisements « très enfouis » : à terre, à plus de 6000 mètres de profondeur, ou en mer, au-delà de 2000 mètres de fond. Désormais, qu'ils soient techniques, stratégiques ou financiers, les choix relatifs à l'exploration puis l'exploitation éventuelle d'un champ requièrent le souci constant d'une meilleure compréhension du milieu qui renferme ces ressources. En effet, une fois la zone intéres-

¹ Depuis une trentaine d'années, la consommation mondiale par habitant en hydrocarbures a augmenté de 11 %, les prévisionnistes tablent par contre sur une croissance de l'ordre de 27 % d'ici à 2030. [AIE/BP Stat review].

sante repérée par la géologie ou l'imagerie sismique¹, l'industrie cherche à apprécier les qualités qui en conditionnent le rendement potentiel. Pendant la phase d'exploration, il faut d'abord évaluer la capacité du champ en matières premières susceptibles d'être produites : accumulation / réserve². Au stade de l'exploitation, il est alors nécessaire de prévoir, à plus ou moins long terme, l'évolution de la production en matières premières : c'est l'établissement des profils de production.

De telles grandeurs dépendent fortement de la distribution spatiale du volume (enveloppe du gisement), de la porosité (volume occupé par les fluides), de la saturation en hydrocarbures (fraction de la quantité de fluide en huile ou en gaz), et de la perméabilité (aptitude du milieu à se laisser traverser par un fluide) [Chapellier *et al.*, 2004]. Ces propriétés, dites réservoir, présentent effectivement de larges degrés d'hétérogénéités, à l'origine d'une répartition inégale des différents types de milieux poreux comme de l'écoulement des fluides. C'est précisément cette part de variabilité pétrophysique, au sein même des couches susceptibles de contenir les hydrocarbures, que l'industrie se doit de connaître, si elle veut caractériser au mieux les performances du réservoir.

De fait, l'appréciation de ces paramètres clés pour l'industrie passe par la traduction des propriétés topologiques, géométriques et physiques des hétérogénéités géologiques les plus actives vis-à-vis de l'écoulement des fluides. Cette étape fait classiquement intervenir la modélisation numérique de réservoir. Elle consiste à construire un modèle géologique – généralement en 3 dimensions – du milieu, ou géomodèle (Fig. 1). Celui-ci intègre les caractéristiques pétrophysiques – pour les calculs en accumulations – à partir desquelles les simulateurs d'écoulements reproduisent le comportement dynamique de la zone – pour l'établissement des profils de production et des réserves. Plus exactement, ce géomodèle reproduit dans un maillage à trois dimensions :

- (1) la structure du gisement, c'est-à-dire les toits, murs, limites latérales, et isopaques, et isobathes³ à partir des horizons et des failles pointés sur la sismique 3D et les marqueurs aux puits ;
- (2) l'architecture interne du réservoir définie par exemple, à l'aide des surfaces géologiques intermédiaires non visibles sur la sismique, elles sont corrélées et interpolées entre les puits en tenant compte de l'histoire géologique de la zone; en ce qui concerne les environnements de plates-formes carbonatées par exemple, le recours aux principes de la stratigraphie séquentielle [Vail, 1977 ; Homewood *et al.*, 2000] permet d'obtenir ces limites en rendant compte de l'organisation séquentielle des dépôts ;

¹ **Sismique** : Technique d'investigation du sous-sol par l'étude d'ondes de compression (ondes P) et de cisaillement (ondes S) émises et enregistrées en surface à l'issue d'un parcours souterrain comportant des réflexions ou des réfractions sur des couches de caractéristiques différentes.

² **Accumulation / Réserve** : Volume d'hydrocarbures présent dans le réservoir pouvant être produit/Volume d'hydrocarbures considéré comme récupérable pendant la phase de production, sous un certain plan de développement.

³ **Isobathe** : Ligne joignant les points qui sont situés à la même profondeur sous l'eau ou sous la terre.

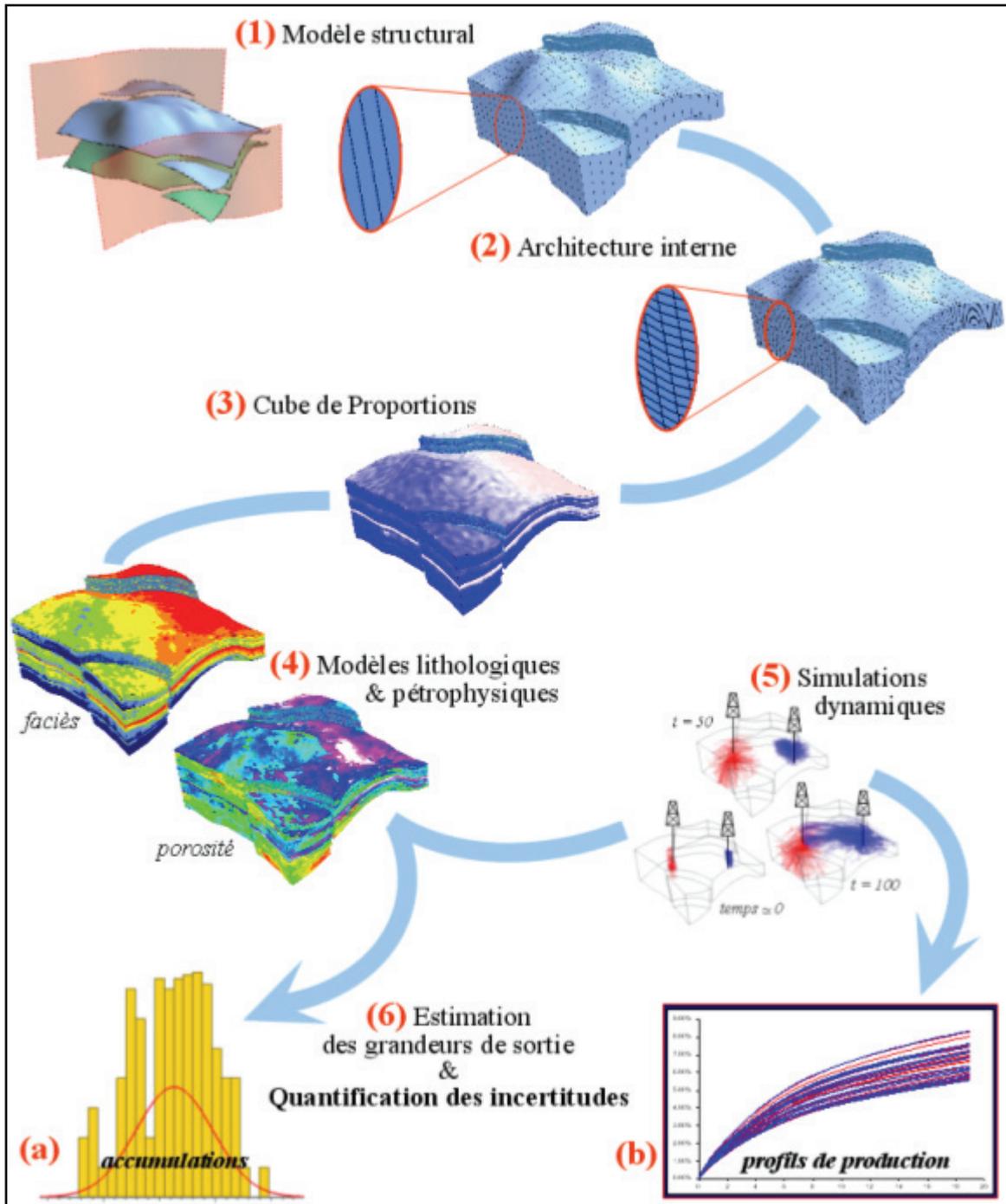


FIG. 1 : La construction d'un géomodèle. À partir de la structure générale de la zone (1), est défini le maillage interne du modèle (2). Le réservoir maillé est renseigné en probabilité d'apparition de faciès (3), nécessaire au remplissage en propriétés lithologiques puis pétrophysiques (4). Ces réalisations permettent d'estimer les accumulations (6-a) et, à travers un modèle du comportement dynamique du réservoir (5), l'établissement des profils de production (6-b) et d'en quantifier les incertitudes.

- (3) les propriétés relatives à la lithologie avec en particulier, (a) les proportions de faciès¹, on parle alors de **cube de proportions**. Issu de l'interprétation géologique des données et des statistiques concernant la répartition des faciès dans l'espace, il représente en tout point du modèle la probabilité d'occurrence des différentes catégories de sédiments ; et (b) les faciès eux-mêmes déterminés à partir de cette loi de probabilité ;
- (4) les évolutions pétrophysiques pour chaque faciès en tout point du réservoir, comme la porosité, la hauteur utile², la saturation en hydrocarbures ou les perméabilités absolues et relatives, etc.
- (5) les paramètres nécessaires à toutes simulations dynamiques, ils viennent compléter les données statiques (porosité, perméabilités) précédemment générées ; on recense notamment les propriétés thermodynamiques des fluides (compressibilité, diagramme de phase, facteur de volume) ainsi que le plan de développement qui précise l'implantation des puits et la stratégie de production ;
- (6) et bien entendu, les **incertitudes** associées et, inhérentes à toute prédiction, avec par exemple, une quantification autour de la capacité ou du profil médian, ces incertitudes se répercutant sur les estimations d'accumulations ou de productions du réservoir.

Car une telle description du milieu souterrain et de son comportement dynamique constitue un exercice ardu et particulièrement incertain. Or, d'importantes décisions liées à des investissements lourds, comme l'achat de licences d'exploitation ou le développement de projets en haute mer, se prennent sur la base de ces représentations incertaines du champ. Ainsi, une incertitude négligée, ou qui échappe à l'identification, conduit souvent à la ruine d'un projet de développement prometteur. Le passage d'un bon candidat à un champ économiquement viable impose donc de savoir évaluer et limiter ces incertitudes. Gérer celles qui ne peuvent être suffisamment réduites s'avère également capital. On distingue classiquement trois grandes familles d'incertitudes selon l'étape de construction du modèle où elles interviennent et le type d'information utilisé. On parle d'incertitudes structurales lors de la définition de la géométrie du réservoir (étape 1), géologiques pour la construction du maillage interne, le remplissage en propriétés lithologiques et pétrophysiques (étapes (2), (3) et (4)), et dynamiques pour la conception du modèle en écoulement – perméabilités relatives, données thermodynamiques, compressibilité, etc. – (étape (5)). Il est donc essentiel d'évaluer leurs rôles et de les hiérarchiser, afin de mieux les appréhender. Nous nous intéresserons plus particulièrement aux incertitudes géologiques qui ne sont pas complètement intégrées dans les chaînes d'étude actuelles [Massonnat, 2000 ; Poirot, 2002 ; Sheppers, 2003].

¹ **Faciès** : Catégorie dans laquelle on peut ranger une roche ou un terrain, et qui est déterminée par un ou plusieurs caractères lithologiques (lithofaciès) ou paléontologiques (biofaciès). Ce terme est également employé pour désigner une catégorie correspondant à un milieu ou à un domaine de sédimentation.

² **Hauteur utile** : Fraction du volume réservoir considérée comme économiquement et/ou techniquement récupérable.

L'incertitude géologique s'explique, outre la complexité du milieu lui-même, par le caractère imprécis, partiel et subjectif, (a) de l'information disponible et, (b) de la connaissance géologique du sous-sol. La modélisation d'un milieu aussi hétérogène ne repose finalement que sur un nombre limité de mesures dispersées dans l'espace, à partir desquelles a été élaboré un scénario géologique. Celui-ci sert de canevas à l'extrapolation des caractéristiques du sous-sol observées au niveau des puits. Ce scénario, issu de l'interprétation géologique des données, traduit différents choix et hypothèses de modélisation (paléogéographie, tailles des hétérogénéités, etc.), qui par essence sont subjectifs et malheureusement impossibles à vérifier sur le terrain [Caumon *et al.*, 2004]. Il reste par conséquent une grande part d'inconnue, principale préoccupation des acteurs de l'industrie. Actuellement, l'appréciation de cette famille d'incertitude consiste principalement à reproduire l'ensemble des états possibles du milieu compatibles avec les observations. On procède classiquement en faisant varier, dans un cadre probabiliste et à partir d'un cube de proportions (chapitre 2), les paramètres lithologiques (faciès) et pétrophysiques (porosité, perméabilité, saturations, etc.), dont dépendent les volumes récupérables et les profils de production. Grâce aux développements des moyens informatiques, il est désormais possible de générer rapidement, non plus un géomodèle unique (structure et caractéristiques des roches), mais une série de réalisations alternatives des propriétés lithologiques et pétrophysiques du réservoir, qui toutes vont respecter les mêmes données de départ (interprétation géologique, puits, information sismique, etc.) [Corre *et al.*, 2000]. L'examen de toutes les simulations permet de délimiter un intervalle dans lequel fluctue le volume total d'hydrocarbure ou le profil de production, et par la même de quantifier l'incertitude associée à ces grandeurs.

Pourtant, en règle générale, ce pool de réalisations équiprobables est généré sur la base d'un scénario géologique unique. On utilise effectivement un seul modèle probabiliste spatial des propriétés ou cube de proportions. En procédant de la sorte, on représente la variabilité statistique de l'état du milieu autour du scénario choisi. Or, la qualité et la quantité des données recueillies laisse une part essentielle à l'interprétation géologique qui peut se traduire par une profusion de scénarios, tous compatibles avec l'information disponible. Une telle approche ne suffit donc pas à décrire l'espace d'incertitude géologique dans son intégralité puisqu'elle ignore la variabilité que porte le scénario sédimentaire. Pour préciser au mieux ce domaine, il semble par conséquent primordial de revenir à la source des incertitudes qui pèsent sur la définition du schéma sédimentaire ; autrement dit, la géologie. Car cette discipline conduit la compréhension de l'architecture interne du milieu souterrain. En définitive, on en revient au constat de Gonod (1996) pour qui traiter de l'incertitude, c'est finalement aller au cœur de la compréhension du phénomène étudié.

2. Objectifs et organisation générale

Cette thèse, réalisée au Laboratoire de Géodynamique des Chaînes Alpines (LGCA), en partenariat avec le groupe Total, s'inscrit dans le programme de quantification de l'incertitude géologique rencontrée durant la phase de caractérisation des réservoirs. La

première partie (chapitres 1 et 2) précise ainsi la notion d'incertitude dans un tel contexte, et témoigne du caractère hautement incertain du scénario géologique. Elle se veut également une synthèse des méthodes actuelles d'évaluation de l'incertitude sédimentaire, et pose le problème d'une description incomplète de la variabilité liée au scénario géologique.

Nous nous sommes donc intéressés à la manière d'explorer dans son ensemble un tel domaine d'incertitude. L'objectif consiste à retranscrire la part de doute que peut porter le scénario géologique, sur les modèles probabilistes¹ (cube de proportions), puis sur les grandeurs clés pour l'industrie (accumulations, productions, etc.), et ainsi d'aider aux choix techniques et/ou stratégiques. De fait, cette thèse revêt essentiellement des aspects méthodologiques en proposant divers protocoles appliqués à la réalisation multiple en termes de cubes de proportions.

La **deuxième partie** expose une première méthode de simulation des cubes de proportions de faciès en accord avec la variabilité relative au scénario géologique. Elle participe ainsi à la gestion des incertitudes sur le scénario rencontrées dans un processus inédit de modélisation sédimentaire en milieu marin développé au sein du groupe Total : Neptune (chapitre 3). L'accent est mis sur une intégration de l'aléa² lié aux paramètres qui composent le scénario géologique (chapitres 4 à 6).

La **troisième partie** présente d'abord une méthodologie alternative d'intégration de l'incertitude du schéma sédimentaire au niveau du cube de proportions. Elle se veut plus intégrative et synthétique de l'information disponible. Elle s'appuie ainsi sur la construction, à partir de lois probabilités adaptées aux proportions (lois de Dirichlet), de modèles stochastiques en proportions de faciès, modèles préalables à toutes simulations pétrophysiques (chapitres 7 à 9).

Néanmoins, quelle que soit la méthode utilisée, elles ont tendance à élargir l'espace des modèles possibles. Les **chapitres 10 et 11** proposent de fait des techniques de mise en cohérence du pool de cubes de proportions simulés à des signaux secondaires du type sismique, ou historique de production³. Il s'agit par là de restreindre l'espace d'incertitude à un éventail de modèles compatibles avec toutes les sources d'informations disponibles.

Ce mémoire a également pour objet de développer d'une part, une méthode de régionalisation de l'incertitude (**chapitre 12**). L'idée consiste à identifier les zones plus ou moins (in)certaines du modèle maillé en terme de géologie. D'un autre côté, il aborde la question du classement des cubes de proportions de façon à établir la fonction de répartition des modèles simulés. Il s'agit ainsi de fournir des modèles clés – via les quantiles – utilisables dans les simulations dynamiques en écoulement.

¹ utilisés précisément pour décrire ces incertitudes.

² **Aléa** : Ce terme est uniquement considéré dans son sens le plus général, il se rapporte à un événement incertain ; dans ce mémoire, il ne renvoie en aucune manière à l'expression d'un risque.

³ **Historique de production** : Il s'agit de l'enregistrement des données enregistrées aux puits (évolution de la pression, des volumes extraits, ...) la production cumulée réalisée au cours de la production du champ.

Première partie

Incertitude géologique et modélisation du sous-sol

La partie suivante revient d'abord sur les différentes sources d'incertitudes géologiques rencontrées pendant la modélisation des propriétés du milieu souterrain. Elle souligne en particulier, l'ampleur de la variabilité que revêt la définition du scénario géologique, définition issue de l'interprétation d'un nombre somme toute limité de données. Le second chapitre s'intéresse aux techniques utilisées aujourd'hui pour quantifier cette forme d'incertitude. Pour l'essentiel, celles-ci reposent sur le cadre probabiliste, car on ne peut atteindre qu'une représentation probable de la réalité. L'examen de plusieurs modèles illustre l'incapacité de telles méthodes à explorer l'ensemble de l'espace d'incertitude. Ces dernières n'intègrent effectivement pas, ou uniquement de manière partielle, la variabilité du scénario géologique.

Chapitre 1

L'incertitude géologique

La quantification des incertitudes se rencontre aux différentes étapes de la vie d'un champ pétrolier, de son exploration à sa mise en production. Il s'agit par contre d'un passage particulièrement délicat au moment de l'estimation des accumulations, des réserves ou des profils de production à partir de modèles numériques. De telles représentations de la réalité, produits des données et de leur interprétation à travers le prisme de la géologie, demeurent largement entachées incertitudes. Ces dernières se répercutent en particulier au moment de la construction du maillage interne et du remplissage en propriétés lithologiques et pétrophysiques. Elles ont pour origine une connaissance incomplète de la distribution spatiale des propriétés géologiques du réservoir. Cette part d'ignorance s'explique par la complexité du milieu, le manque d'information disponible et la subjectivité des interprétations géologiques. Le chapitre suivant précise les sources d'incertitudes relatives à la connaissance géologique du milieu souterrain.

1.1 Un milieu à l'histoire géologique complexe

La prévision exacte des performances d'un réservoir impose de connaître de manière exhaustive la distribution spatiale des caractéristiques du sous-sol. Cependant, les propriétés pétrophysiques du milieu montrent de larges degrés d'hétérogénéités dans l'espace. Ces caractéristiques, inhérentes à la roche, correspondent en effet au produit de l'histoire globale du milieu. Celle-ci met par contre en jeu une succession de phénomènes physico-chimiques complexes, souvent couplés entre eux : sédimentation, diagenèse¹, compaction, fracturation, etc. –, dont certains paramètres demeurent intrinsèquement entachés d'incertitudes. De tels mécanismes s'avèrent ainsi être une source de difficultés quant à la compréhension du milieu souterrain puis de sa modélisation dans l'espace 3D. L'hétérogénéité naturelle des objets géologiques ainsi produite se répercute de fait sur la

¹**Diagenèse** : Ensemble de tous les processus qui affectent un dépôt sédimentaire et le transforment progressivement en roche sédimentaire solide.

distribution spatiale des propriétés pétrophysiques ; ce qui rend toute prédiction des ressources disponibles et de leurs écoulements hautement incertaine. Les réservoirs carbonatés, par exemple, à l'histoire géologique souvent compliquée, illustre les obstacles auxquels on peut être confronté au cours de leur modélisation [Massonnat, 2004] :

- « ces formations tirent en partie leur origine sédimentaire d'une activité biologique dont la sensibilité aux paramètres écologiques et climatiques évolue plus ou moins rapidement dans l'espace et le temps avec les espèces mises en jeu, les processus de transport et/ou de remobilisation, qui peuvent exister, compliquent encore davantage l'organisation des dépôts ;
- l'extrême diversité des textures sédimentaires, liée notamment à la densité, la taille et l'organisation des grains, ainsi qu'au type de ciment, s'avère une source de difficulté pour la compréhension et la prédiction des types de milieu poreux dans l'espace 3D ;
- cette difficulté est renforcée par la sensibilité très forte des sédiments originels aux divers types de diagenèse : dissolution, recristallisation, cimentation, dolomitisation, etc., qui constituent autant de phénomènes venant perturber les propriétés initiales, et en particulier la perméabilité du milieu ;
- les carbonates sont fréquemment affectés d'une fracturation intense précoce ou tardive favorisée par leurs propriétés mécaniques ; cette fracturation, dont une caractéristique principale est l'aspect multi échelles, constitue un élément clé si on considère les formations carbonatées d'un point de vue réservoir, aquifère ou à hydrocarbures. »

La complexité de tels milieux empêche par conséquent de déterminer parfaitement la véritable distribution spatiale des propriétés pétrophysiques. Seul un échantillonnage exhaustif de la zone d'étude, en pratique impossible, permettrait de connaître exactement l'agencement réel des caractéristiques du sous-sol ainsi que sa capacité, ou son futur comportement dynamique.

1.2 Des données incomplètes

Il importe alors d'exploiter autant que possible l'information tirée des mesures collectées au cours des campagnes de terrain (prélèvements sur carottes, diagraphies¹, imagerie sismique, essai de puits², etc.) Cette source de connaissance contribue à préciser la valeur des propriétés réservoir en tout point du milieu où aucune donnée n'a été prélevée. Néanmoins, en plus de l'existence inexorable d'erreurs de mesure ou de calibration des

¹ **Diagraphie** : Enregistrement en continu dans un sondage de paramètres physiques (résistivité électrique, la radioactivité naturelle, etc.) afin de connaître les propriétés des sédiments.

² **Essai de puits** : Technique qui consiste à perturber le régime de pression régnant dans le milieu en faisant varier le débit d'un puits afin de caractériser le comportement dynamiques du réservoir (directions et axes d'écoulement, anisotropies, etc.)

appareils, les données récoltées se révèlent plus ou moins fiables. Les puits, par exemple, se situent en général dans les zones supposées riches en ressources. L'information que fournit l'examen des carottes extraites de ces derniers, introduit le risque d'une estimation biaisée des propriétés. Un puits foré à un autre emplacement, pourrait engendrer une image dissemblable du réservoir. C'est notamment le cas en environnement fluviatile, où les faciès peuvent évoluer latéralement de manière très rapide, pour un paramètre tel que la proportion entre chenaux et encaissant argileux. Ce ratio, qui gouverne le nombre de chenaux, – et bien entendu, les volumes d'huile en place –, se calcule directement à partir des puits, dont la représentativité, dans des formations aussi variables, n'est pas forcément assurée (Fig. 1.1).

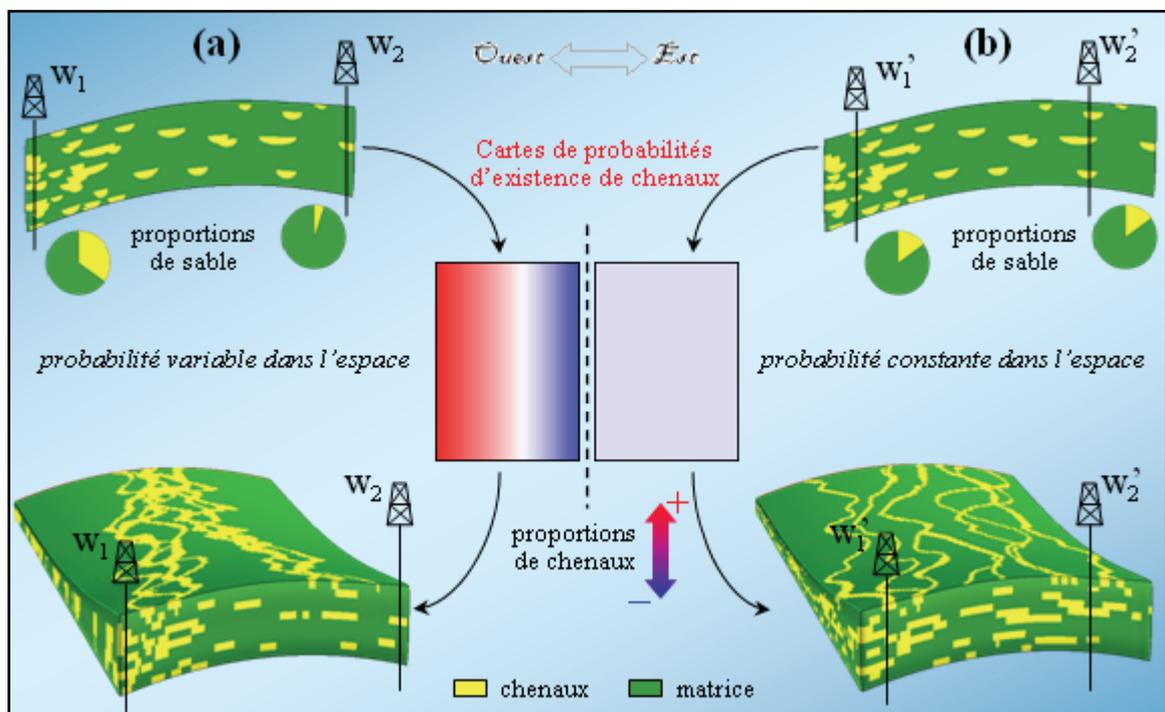


FIG. 1.1: *La représentativité des données en question. Dans un milieu fluviatile, une première configuration de forages, en (a), laisse croire que la proportion de sable décroît d'ouest en est; conduisant à construire un modèle où la probabilité d'apparition des chenaux varie dans l'espace. En (b), un positionnement alternatif des puits conduit plutôt à considérer une carte dont la proportion en chenaux reste stationnaire dans l'espace, aboutissant à une image complètement différente de la zone d'étude.*

À contrario, les techniques d'imagerie sismique développées au milieu du XX^{ème} siècle, peuvent permettre d'affiner le calcul de ce rapport entre sable et argile ou même, de délivrer une estimation de la porosité dans l'espace. Elles fournissent en effet dans certains cas, une image en deux dimensions du réservoir, rendant possible l'appréhension de l'évolution latérale des faciès [Doyen, 1988 ; Behrens *et al.*, 1998]. Cependant, les mesures effectuées n'apportent qu'une information indirecte sur des propriétés comme la

proportion de sables ou la porosité. Les données doivent d'abord être **interprétées** puis transformées afin d'établir un lien avec les caractéristiques pétrophysiques du milieu. Il s'agit entre autres :

- de l'établissement d'une relation fonctionnelle entre l'impédance acoustique et la porosité ;
- du choix des seuils utilisés pour attribuer une nature de faciès à un intervalle d'amplitudes de signaux sismiques.

Il existe par conséquent une réelle part d'incertitude quant au caractère subjectif du choix de la relation utilisée pour la conversion des attributs sismiques vers les propriétés pétrophysiques. L'information tirée des données reste finalement imprécise, partielle, subjective et de ce fait, participe au caractère incertain de la représentation du milieu souterrain.

1.3 Un modèle géologique incertain

Pourtant, de toutes les sources de variabilité rencontrées en modélisation de réservoir, la géologie joue un rôle majeur car elle conduit la compréhension de l'architecture interne du sous-sol [Massonnat, 2000]. Malgré l'arsenal d'outils disponible, malgré la variété d'échelle que balayent les différents procédés de mesures, le nombre de données reste limité en raison notamment des délais et des coûts. Le volume exploré par les puits ne représente, somme toute, qu'une fraction infime, de l'ordre de 10^{-11} , du volume occupé par la zone étudiée. Pour se fixer un ordre de grandeur, toutes proportions gardées, modéliser le réservoir avec les données disponibles, reviendrait à reproduire l'ensemble des hétérogénéités de la planète à partir seulement, de l'information qu'apporterait l'exploration d'une île. D'ailleurs, même si les données recueillies le sont en nombre suffisant, elles ne permettent pas toujours de traduire les caractéristiques du milieu dans un modèle. Par exemple, au sein des environnements de dépôts marins, principalement dans les carbonates, la sismique fournit rarement des renseignements sur la distribution spatiale des propriétés pétrophysiques. Dans de telles formations, le pouvoir de résolution sismique, de l'ordre de quelques dizaines de mètres [Fournier, 1990], interdit en général de connaître l'architecture interne des réservoirs et donne un aperçu très limité des hétérogénéités lithologiques. La construction de cartes de tendances à partir des observations réalisées aux puits se révèle être un exercice difficile pour contraindre correctement les modèles en raison de la non stationnarité des propriétés étudiées (les caractéristiques statistiques des variables fluctuent dans l'espace¹).

Il est par conséquent nécessaire, si l'on veut définir un géomodèle cohérent et prédictif, de dépasser l'incomplétude rencontrée au niveau de l'information tirée des données. De

¹ Un exemple typique de variable non stationnaire correspond à la température sur Terre dont les caractéristiques – la moyenne à l'année par exemple – varie selon la latitude. Un autre modèle correspondrait à la nature du faciès rencontré dans l'exemple de la figure 1.1 en (a), où la proportion de sable varie dans d'est en ouest.

manière générale, ce passage repose sur l'intégration dans les modèles de certaines hypothèses et concepts de la géologie, qui permettent d'établir un lien entre les données et les points où aucune mesure n'a été prise. En effet, bien qu'une importante variabilité caractérise le milieu souterrain, les structures géologiques qui le composent font qu'il n'est pas non plus complètement chaotique. Il présente, au contraire, une certaine cohérence géologique ou continuité spatiale, que le modéleur, guidé par la connaissance apportée par la géologie, cherche à traduire dans son interprétation des observations disponibles. Cette source d'information est introduite par l'intermédiaire de ce que l'on appelle le scénario géologique. Issu de l'interprétation par le sédimentologue des données disponibles (sismique, carottes, diagraphies, etc.), ce schéma sédimentaire spécifie les paramètres gouvernant la position des différents dépôts, comme par exemple, le modèle en faciès, la direction des apports sableux, la paléogéographie, la taille des hétérogénéités, les évolutions latérales de faciès, la position de la ligne de rivage, l'énergie des vagues, les corrélations, etc. Il sert ainsi de canevas à la modélisation et à l'extrapolation des caractéristiques du sous-sol. Le scénario règle, en quelque sorte, « l'image géologique » du réservoir [Poirot, 2002], puisqu'il traduit un choix de continuité géologique du milieu souterrain.

La figure 1.2 illustre la manière dont se traduisent à la fois, cette connaissance et le choix de scénario. Cet exemple s'intéresse au cas du réservoir LILOUCHE, de type turbiditique¹, au large de l'Afrique de l'Ouest. Il s'agit d'en modéliser la porosité à partir des observations qu'apportent quatre puits d'investigations. La zone étudiée correspond à un piège stratigraphique² créé suite au comblement, par les turbidites, d'un canyon situé sur le talus continental. L'utilisation seule des mesures de porosité réalisées aux puits conduit en général au tracé de lignes « iso-porosités » parallèles d'est en ouest (Fig 1.2-a). Ce dessin constitue en soi un choix de scénario géologique, car il impose un modèle certain de continuité de la porosité dans l'espace (polarité vers le Sud). Par contre, l'examen des forages, via la géologie, apporte un supplément d'information. Ils révèlent effectivement des corps sédimentaires typiques des turbidites : chenaux, lobes, etc. Or, le fait d'interpréter ces observations comme liées à cette famille de dépôts amène à privilégier des hétérogénéités plutôt anisotropes. Par conséquent, ce choix de scénario change complètement l'image du réservoir puisqu'il modifie la continuité des corps modélisés (Fig. 1.2-b). D'ailleurs, au sein même du scénario sédimentaire, il demeure, dans cet exemple, une part d'incertitude sur l'orientation de la zone d'apport des sables. L'interprétation de dépôts turbiditiques dirigées d'axe nord-sud ou encore d'axe est-ouest bouleverse une nouvelle fois l'image du modèle (Fig. 1.2-b' et b''). Ces variations se répercutent bien entendu sur les accumulations, les écoulements et modifient les profils de productions.

¹ **Turbidite** : Sédiment d'origine détritique déposé par des écoulements gravitaires (courants de turbidité) le long d'un talus continental sous-marin.

² **Piège stratigraphique** : Piège créé par le recouvrement des roches réservoirs par des couches imperméables ; on peut entre autres citer des sables d'origine fluviatile recouverts par les argiles de plaine d'inondation.

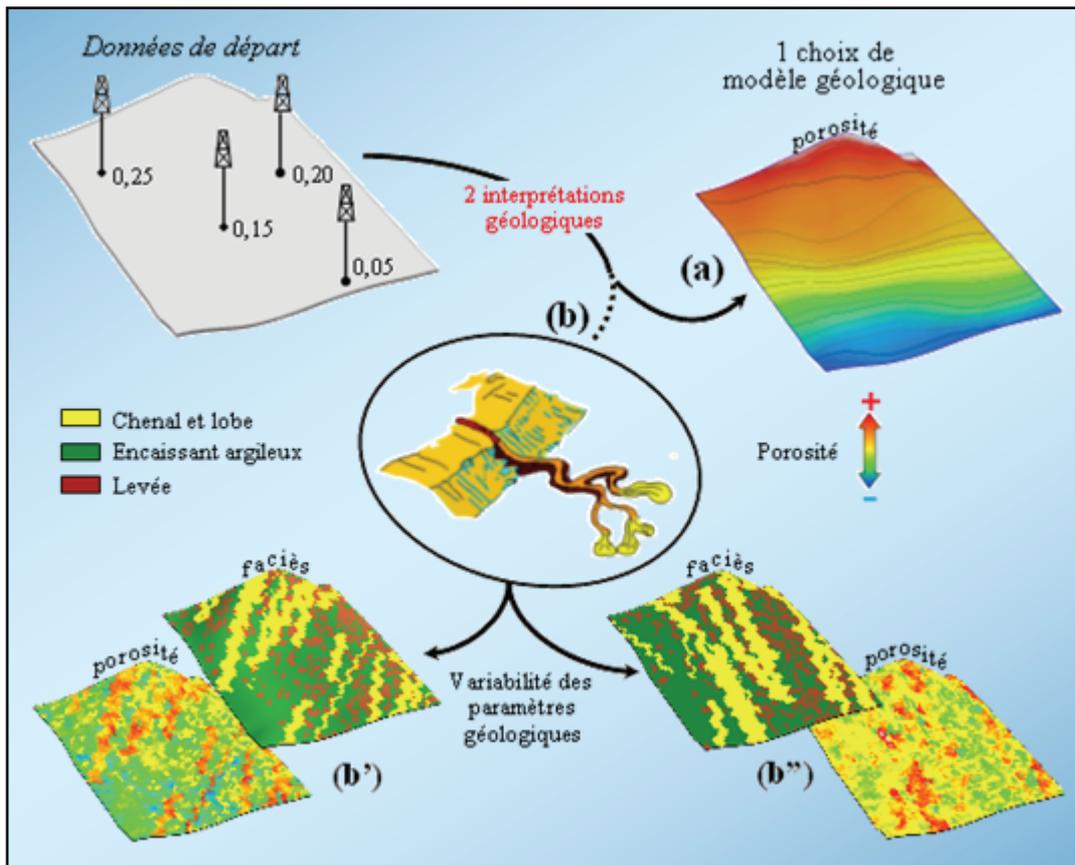


FIG. 1.2 : Scénario géologique et modèle de continuité spatiale. À partir d'un jeu de données aux puits identique, on peut réaliser différentes interprétations concernant la cartographie de la porosité. En (a), un modèle construit en interpolant les données de puits sans apporter explicitement de connaissance géologique. Cependant, implicitement, il traduit un choix de modélisation. En (b), on impose explicitement un scénario géologique et un modèle de continuité, lui-même variable selon la direction d'anisotropie privilégiée, modèles (b') et (b'').

Ce cas d'étude souligne ainsi, que la définition d'un schéma particulier ne correspond qu'à l'une des nombreuses interprétations concevables en ce qui concerne la cohérence sédimentaire du réservoir. Le manque de données comme la complexité des objets géologiques mis en jeu font que, certaines composantes du scénario restent difficilement accessibles et quantifiables. C'est le cas, par exemple, pour l'évaluation de l'orientation ou de la sinuosité des chenaux. Il s'agit de paramètres, dont l'impact se traduit en deux dimensions, et que les géologues doivent appréhender à partir d'observations réalisées essentiellement en une dimension. Il leur faut donc faire appel à leur expérience qui varie naturellement d'un expert à l'autre suivant son parcours, ses connaissances. Un cas analogue à celui-ci, présenté à une classe d'étudiants en géologie/géophysique de l'Institut Français du Pétrole a laissé la place à tout un éventail de modèles géologiques possibles en terme de scénario géologique (orientation, anisotropie des corps,...). La place allouée à

l'interprétation reste par conséquent importante, conduisant ainsi à la possibilité d'un large éventail de valeurs pour les caractéristiques du scénario, qui tous concourent à la diversité de ce dernier.

Le caractère imprécis, partiel et subjectif de la connaissance géologique du milieu souterrain confère de fait, à ce scénario ainsi qu'aux modèles qui en découlent, un caractère hautement incertain [Tamhane *et al.*, 1999]. Dans la suite de ce mémoire, on s'intéresse essentiellement aux incertitudes sédimentaires liées au choix du scénario géologique. Actuellement, leur prise en compte reste incomplète et repose essentiellement sur la multi-réalisation des propriétés pétrophysiques (Chapitre 2).

Chapitre 2

Le traitement de l'incertitude géologique

Aujourd'hui, les efforts de l'industrie portent principalement sur le traitement des incertitudes géologiques inhérentes à toute prédiction sur les grandeurs clés qu'elle s'est fixées : accumulation et production. L'objectif consiste à en fournir une estimation la plus « juste » et précise possible, mais surtout d'en quantifier la variabilité autour d'une capacité et d'un profil médian. Le chapitre suivant présente les méthodes couramment employées pour rendre compte de cette source d'incertitude. En règle générale, elles estiment ces quantités à partir de modèles numériques spatiaux du réservoir (ou géomodèles) pour l'ensemble des cas alternatifs envisageables et compatibles avec les données. Pourtant, quelle que soit la technique utilisée, elle ne permet pas d'explorer l'intégralité de l'espace d'incertitude, c'est-à-dire l'ensemble des modèles possibles, puisqu'elle n'intègre pas, ou seulement de façon discrète, la variabilité propre au scénario géologique.

2.1 Le cadre probabiliste

Les caractéristiques du milieu, comme les accumulations qui en découlent, bien que produits d'une histoire géologique complexe, restent complètement déterministes. Pourtant, compte tenu du manque de connaissance et de la complexité de la réalité géologique (chapitre 1), on ne peut construire un modèle détaillé qu'à condition de consentir une part d'incertitude sur celui-ci. Aussi, se place-t-on le plus souvent dans un cadre probabiliste. Plutôt que d'établir un modèle unique, on génère un très grand nombre de modèles « équiprobables »¹, chacun d'eux respectant a priori les données à disposition. Concrètement, quand on définit un géomodèle en faciès ou en porosité, on s'efforce de représenter sa vision de la réalité géologique la plus en accord avec l'information disponible. Cepen-

¹ En fait, ces modèles correspondent à des réalisations tirées indépendamment dans la distribution probabiliste des propriétés définie à l'aide de la connaissance et/ou de l'information disponible.

dant, si un tel modèle est cohérent au regard du scénario géologique¹, l'agencement des hétérogénéités pétrophysiques ne peut constituer une représentation précise de la réalité en raison des nombreuses incertitudes. Il paraît donc primordial de disposer des différents modèles alternatifs envisageables, dont l'étude et la comparaison permettent d'apprécier l'impact des incertitudes :

- d'abord, sur les faciès et les paramètres pétrophysiques, c'est-à-dire sur les accumulations en hydrocarbures ;
- puis, à travers un modèle de comportement dynamique de la zone réservoir, sur les réserves et la production.

La modélisation stochastique reproduit de telles variantes en simulant les réalisations de variables ou de fonctions aléatoires (Fig. 2.1). Une variable aléatoire est une fonction dont

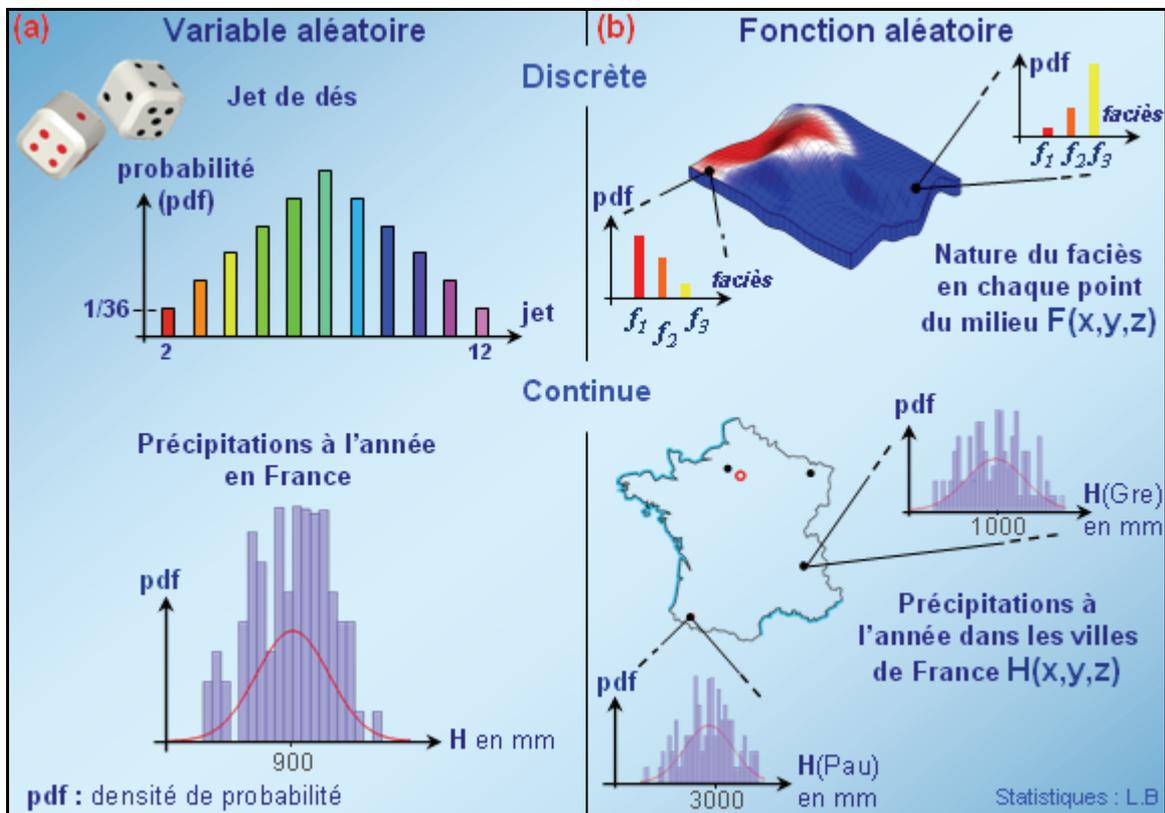


FIG. 2.1 : **(a)** Notions de variable et **(b)** de fonction aléatoire. La première décrit les valeurs observées comme le résultat d'un mécanisme aléatoire, la seconde comme celles d'une fonction définie dans l'espace et réalisation particulière d'un processus aléatoire. Toutes deux renvoient soit à des variables discrètes : variable catégorielle correspondant souvent à des éléments de classification, ou bien continues.

¹ En termes de direction d'apports, formes, tailles, et orientation des corps sédimentaires, distribution des phases de diagenèse, de paléogéographie, etc.

les valeurs possibles sont connues mais dont le résultat final reste a priori imprévisible¹ [Bouleau, 1986 ; Govaert *et al.*, 2004]. Une fonction aléatoire² se définit comme l'association à tout point d'un espace, aussi bien spatial que temporel, d'une variable aléatoire [Matheron, 1962, 1965]. Cette approche permet ainsi, plutôt que de traduire la réalité dans un modèle unique, d'atteindre une représentation probable de la réalité à travers des modèles statistiques, dont une ou plusieurs composantes sont aléatoires voire inconnues.

L'exemple suivant, inspiré de Chilès et Delfiner (1999), et Marcotte (2000) illustre l'intérêt de travailler dans un tel cadre en ce qui concerne l'évaluation des incertitudes. Il est ici question de déposer un câble sous-marin dont on cherche à estimer la longueur néces-

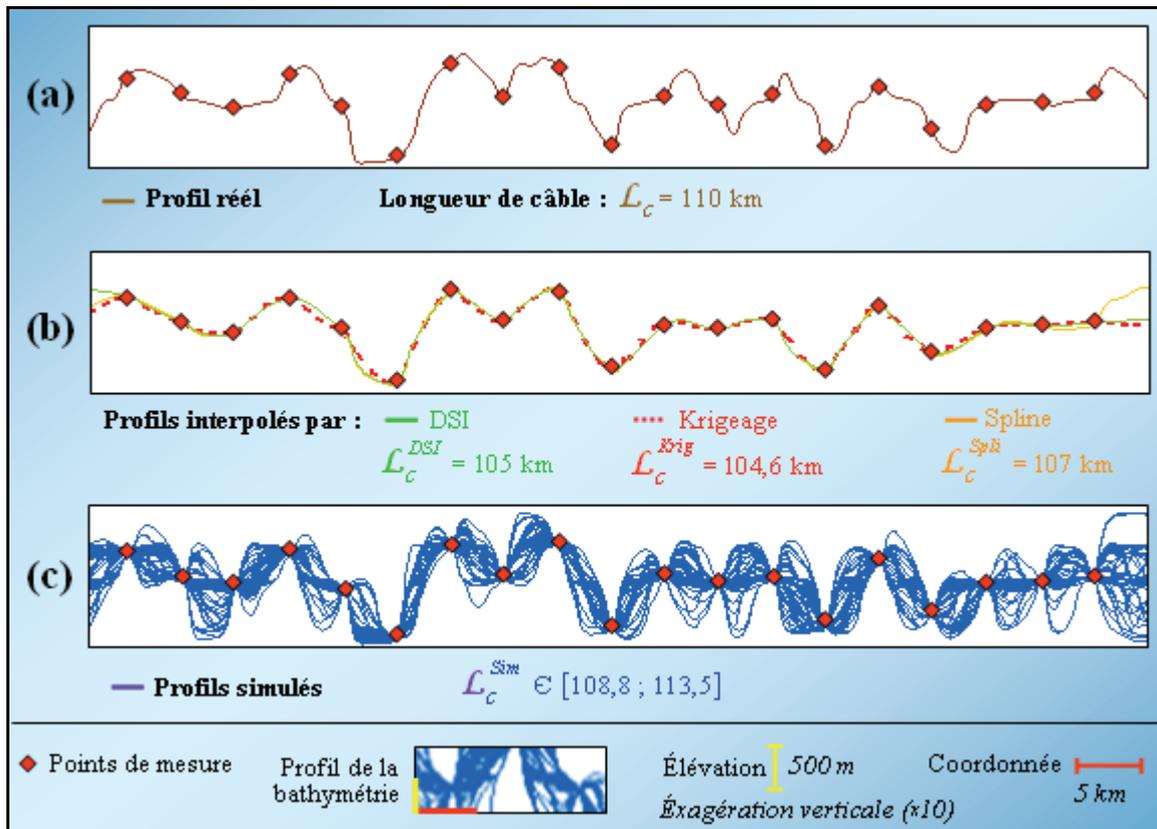


FIG. 2.2 : Cadres déterministe et probabiliste. En (a) sont représentés le profil réel du fond marin et les points d'échantillonnage. En (b) les techniques d'interpolation – DSI, Spline et Krigeage – donnent une image unique et lissée de la réalité. En (c) les simulations stochastiques reproduisent la variabilité de la longueur de câble et son incertitude.

¹ La variable aléatoire la plus simple est donnée par le jet d'une pièce, dont la réalisation vaut *pile* ou *face*. Un autre exemple est donné par le résultat d'un jet de dé, pour lequel les valeurs possibles sont 1, 2, 3, 4, 5, 6 (si le dé comprend six faces).

² Une fonction aléatoire simple serait par exemple le niveau d'eau dans une piscine, ou bien la nature du faciès dans une région, en chaque point de l'espace, le niveau de précipitations en France (Fig. 2.1)

saire au raccordement des deux côtes. La profondeur exacte (topographie) est connue uniquement en quelques points d'observations ; ce sont les équivalents des puits en modélisation de réservoirs. Dans un premier temps, on estime la longueur de câble de façon déterministe suivant quelques-unes des nombreuses méthodes d'interpolations. On recourt aux algorithmes de DSI – compromis entre la « rugosité » de la solution et le respect des données [Mallet, 1992, 2002] –, de Splines (ajustement d'une fonction définie par morceaux par des polynômes) [De Boor, 1978] puis, de krigeage (estimation des points en pondérant les données voisines). Les résultats de l'interpolation sont respectivement égaux à 105 km, 107 km et 104,6 km alors que la longueur réelle mesure 110 km. On court par conséquent le risque de manquer de câble car ces méthodes déterministes fournissent des images lissées de la réalité. A contrario, si on recourt à des simulations conditionnelles de la longueur (section 2.2) selon des méthodes géostatistiques (section 2.3), chaque réalisation reproduit une image possible de la réalité en traduisant sa variabilité. Après 1000 simulations de la longueur, l'histogramme obtenu (Fig. 2.1-c) permet d'atteindre sa valeur moyenne, égale ici à la longueur réelle (110 km). Plus important encore, il est désormais possible d'en appréhender les caractéristiques de dispersion par l'intermédiaire d'un intervalle de confiance [Verdel, 2000 ; Brethon *et al.*, 2002]. Le câble sous-marin présente ici 95 chances sur 100 de mesurer entre 108,8 km et 113,5 km de long¹. C'est dans ce cadre que sont étudiées les incertitudes géologiques.

2.2 Les méthodes statistiques

2.2.1 Les simulations par cubature

Afin de déterminer l'impact des incertitudes sur les accumulations, il est possible, dans un premier temps, de considérer le réservoir dans sa globalité. On parle dans ce cas de méthode par cubature ; le réservoir étant assimilé à une cellule (ou cube) unique [Biver, 1998 ; Cossé, 1993]. On procède selon une démarche purement probabiliste, dite de Monte-Carlo. Elle est couramment utilisée en analyse des risques et fait appel aux principes d'échantillonnage aléatoire [Robert & Casella, 1999]. L'information disponible sur les données d'entrée du modèle (volume, proportions de faciès, porosité, perméabilité,...) est ainsi supposée de nature aléatoire du fait de la variabilité naturelle des caractéristiques du milieu. Les grandeurs incertaines – porosité, perméabilité, etc. – sont représentées par des distributions de probabilité plus ou moins complexes que divers procédés statistiques permettent d'apprécier à l'aide des données recueillies. Ces méthodes statistiques recourent classiquement à des lois analytiques, en général, gaussiennes pour la porosité ou lognormales pour la perméabilité, dont les paramètres sont déterminés à partir des observations [Bouleau, 1986]. L'incertitude sur les lois peut, dans certains cas, être intégrée en procédant selon une démarche bayésienne. Cette dernière permet d'actualiser les estimations d'une probabilité ou d'un paramètre à partir des observations et de leur loi de proba-

¹ Il est intéressant de noter que les longueurs estimées par interpolation ne sont même pas comprises dans cet intervalle de confiance à 95%.

bilité a priori [Verdel, 2000 ; Bouleau, 1986]. D'autres approches reposent plutôt sur des techniques de ré-échantillonnage de l'information disponible (ou bootstrapping) [Efron *et al.*, 1993]. Une fois précisées les différentes distributions de probabilité, l'incertitude relative aux paramètres du modèle se traduit en terme de variabilité sur les accumulations (OOIP), considérées comme le produit de ces grandeurs :

$$OOIP = \sum_f^{\text{faciès}} RV \times p_f \times NTG_f \times \Phi_f \times (1 - S_{w_f}) \quad (\text{Éq. 2.1})$$

Dans l'équation 2.1, le terme RV représente le volume de roche définie par la structure, les contacts entre fluides et l'architecture interne. L'indice f se rapporte aux différents faciès. Les quantités p_f , NTG_f et Φ_f correspondent respectivement à la proportion, hauteur utile et porosité relatives au faciès f . S_w renvoie à la saturation en eau, elle est d'ordinaire, fonction de la perméabilité et de la porosité (fonction analytique J de Leverett par exemple [Leverett, 1941]).

En pratique, une combinaison des valeurs d'entrée est tirée aléatoirement suivant leur loi de probabilité respective, que l'on affecte à l'ensemble du volume de roche. Ce qui permet la perturbation simultanée de tous les paramètres. Une fois la capacité du modèle calculée, on réitère le procédé avec une nouvelle combinaison de données extraite des lois de distributions. Il est alors possible, avec un nombre suffisant de simulations, d'obtenir une densité de probabilité des volumes en place, utile pour déterminer l'impact des incertitudes : écart-type, quantiles, etc.

2.2.2 Un exemple illustratif

L'exemple suivant (Fig. 2.2) décrit le traitement statistique des incertitudes qui peuvent peser sur les proportions de faciès présents ainsi que sur leur porosité. On s'intéresse au cas d'un réservoir formé dans un environnement fluviatile et que renseignent quatre puits. L'examen des carottes a permis de distinguer deux familles de faciès : des sables d'origines fluviatiles et des argiles de plaine d'inondation. Le calcul des accumulations s'effectue suivant les techniques de Monte-Carlo. Les distributions de probabilité, préalables à toutes simulations, sont déterminées suivant les diverses méthodes précédemment citées (section 2.2.1) :

- (a) à l'aide des données observées aux puits, on détermine d'abord le rapport entre sables ("bons faciès") et argiles ("mauvais faciès"), c'est notre cube de proportions ; les grandeurs incertaines (ici la porosité), se traduisent par une loi de probabilité (gaussienne) dont les deux paramètres (moyenne et écart-type), sont eux-mêmes évalués grâce aux données¹ ;

¹ Dans le cas présenté ici, les simulations de Monte Carlo seraient inutiles, en effet, seule la porosité est incertaine et est amenée à varier, les autres paramètres étant figées dans l'équation 2.1 ; au final, l'incertitude sur les accumulations correspond à celle de la porosité pondérée par le volume du modèle, la proportion des sables...

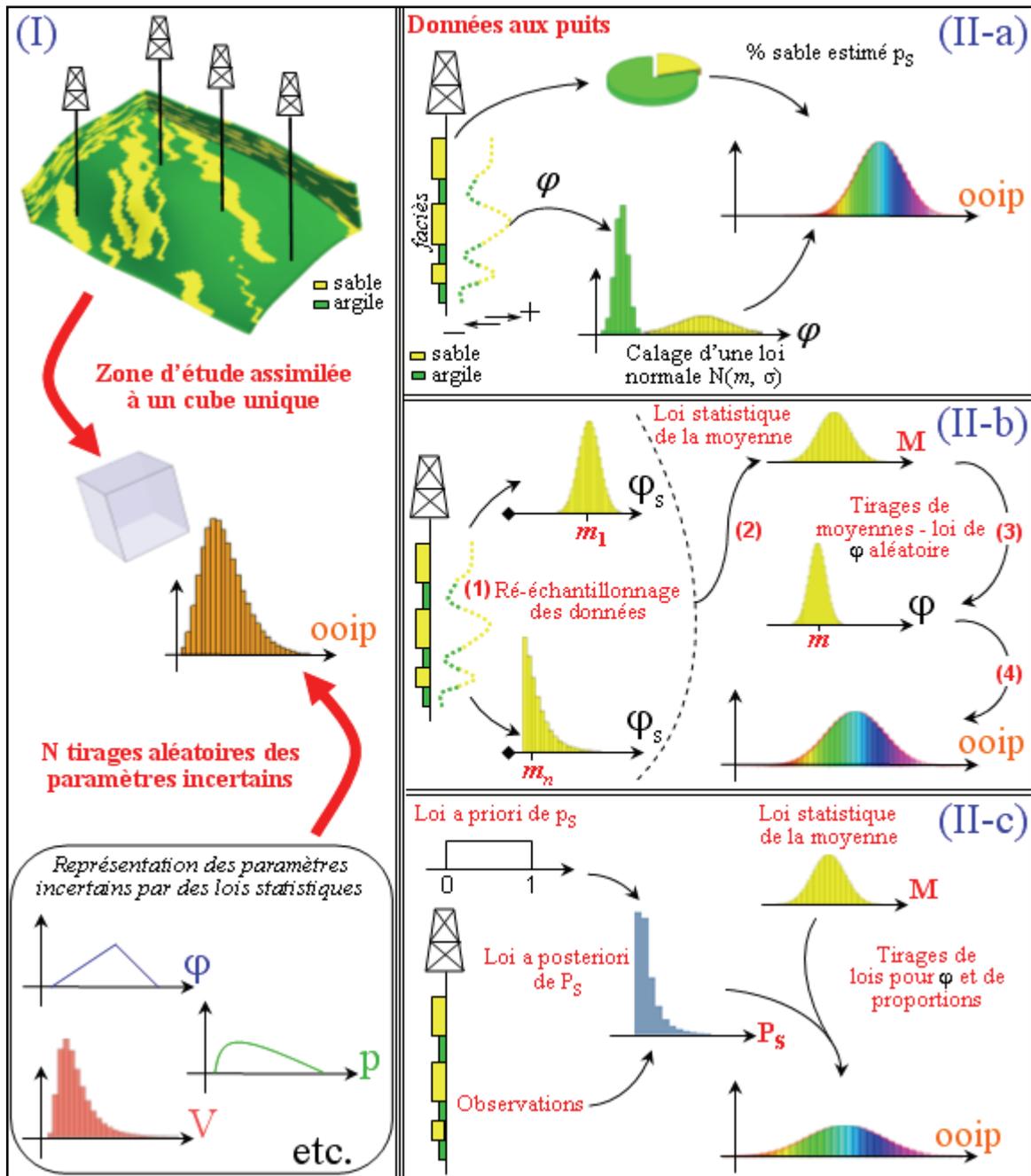


FIG. 2.3 : Les méthodes par cubature. En (I) la zone d'étude est assimilée à un cube au sein duquel varient les paramètres incertains contrôlant les grandeurs de sortie. Trois techniques sont à l'étude en (II) suivant lesquelles 200 réalisations sont calculées pour chacune. En (II-a) la porosité suit une loi gaussienne dont les paramètres sont déterminés à partir des données. En (II-b) les techniques de ré-échantillonnage permettent d'intégrer l'incertitude sur les éléments de la variable aléatoire, ici la moyenne. (II-c) l'approche bayésienne estime la distribution de la proportion en sable et complète le ré-échantillonnage de la moyenne de porosité.

- (b) le cas précédent ne tient pas compte de la variabilité qui pèse sur les paramètres de la loi de distribution (ici la moyenne) ; les techniques de ré-échantillonnage permettent de simuler la loi de probabilité de cette moyenne pour en évaluer le biais, la variance ou un intervalle de confiance même si la loi théorique est inconnue ; le principe consiste à estimer itérativement cette moyenne sur des réalisations différentes de l'échantillon obtenus aléatoirement par n tirages avec remise dans l'échantillon initial de données que fournissent les puits; ainsi, à chaque réalisation générée, la moyenne de la distribution en porosité varie suivant la loi estimée par ré-échantillonnage, ce qui permet de « tirer » une valeur de porosité dans cette nouvelle distribution, valeur qui est ensuite utilisée dans le calcul de la capacité du réservoir ;
- (c) afin d'introduire l'incertitude sur le ratio sable - argile, on procède selon une approche de type bayésien : l'examen des faciès aux puits (qui sont supposés suivre une loi multinomiale¹), est complété par l'inversion de la densité de probabilité a priori des proportions (sensée être une loi uniforme), ce qui conduit classiquement à une loi a posteriori des proportions de faciès correspondant à une distribution de Dirichlet (section 5.2.3).

Pour les différents cas présentés ici, on génère à chaque fois 200 réalisations. Leur examen permet de définir l'histogramme des accumulations et d'en apprécier l'incertitude. Cependant, ces approches statistiques s'affranchissent de la structure spatiale des variables étudiées. La continuité des formations géologiques impose en effet la corrélation dans l'espace des données. Or, les puits se trouvent dans les zones supposées riches en hydrocarbures. Il se pose donc toujours la question de l'indépendance et de la représentativité des observations aux puits [Caumon *et al.*, 2004]. L'incertitude évaluée ici est de nature statistique sur les données disponibles (puits) et ne porte pas véritablement sur le scénario géologique. En outre, elle est représentée sous sa forme globale (au sens de l'ensemble du modèle) sans qu'il soit vraiment possible de discuter de la répartition régionale des accumulations calculées, ou même d'en évaluer les réserves à travers un modèle dynamique. C'est ce que tentent dorénavant de palier certaines méthodes statistiques appliquées dans l'espace en s'intéressant aux différentes cellules du modèle de réservoir.

2.3 Les méthodes statistiques spatiales

Aujourd'hui, pour évaluer au mieux cette famille d'incertitude, on recourt à des modèles numériques spatiaux symbolisant le réservoir dans un maillage 3D. Leurs cellules intègrent ainsi les caractéristiques lithologiques et pétrophysiques du milieu. En effet, suite au développement de l'informatique, les méthodes géostatistiques – statistiques appliquées dans l'espace géométrique [Journel, 1989 ; Deutsch *et al.*, 1998 ; Goovaerts, 1997] –, permettent de générer plusieurs images équiprobables d'un gisement en présen-

¹ **Loi multinomiale** : La variable aléatoire prend au plus n_f valeurs (par exemple, la nature du faciès ou la couleur d'une boule tirée dans une urne) avec respectivement les probabilités p_1, \dots, p_{n_f} .

vant la structure spatiale des hétérogénéités. La connaissance du processus génétique à l'origine des dépôts n'est pas obligatoire. Par contre, il est nécessaire de connaître les dépôts, c'est-à-dire le résultat des processus géologiques, pour l'inférence du modèle statistique. Les techniques géostatistiques considèrent ainsi les dépôts sédimentaires comme la réalisation d'un processus aléatoire défini par sa loi spatiale de probabilité. Pour les faciès, c'est le **cube de proportions** qu'il faut fournir ou construire avant toute simulation stochastique. Celui-ci correspond à l'ensemble des distributions ponctuelles de probabilité d'occurrence des faciès qui sont définies en tout point α (ou cellule uvw) du modèle maillé (Fig. 2.4). Leur évaluation dépend principalement du scénario géologique [Poirot, 2002 ; Bergamo, 2002 ; Sheppers, 2003], dont l'incertitude se répercute au niveau du modèle en proportions qui en découle.

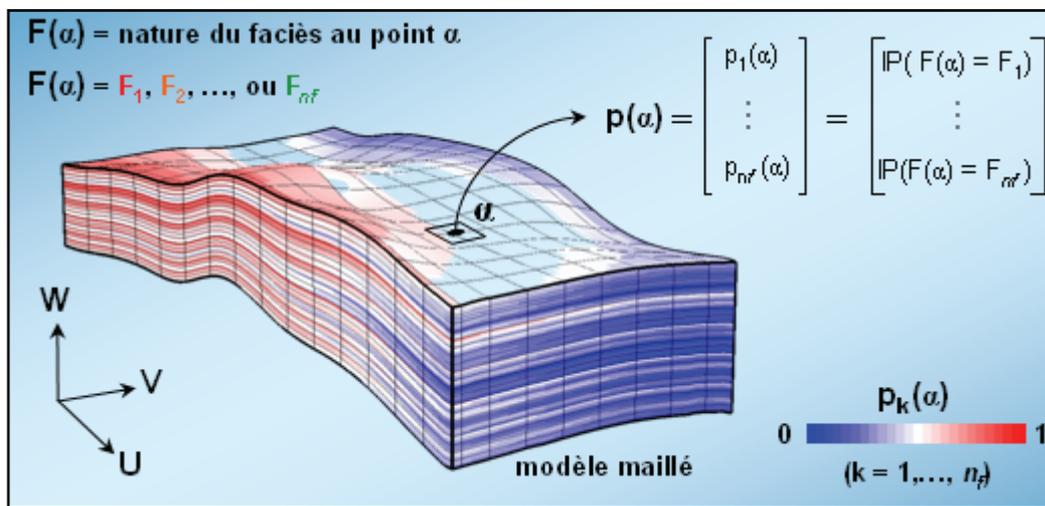


Fig. 2.4 : Définition du cube de proportions de faciès. C'est un modèle discrétisé de l'espace qui correspond en chaque point α (ou cellule d'indice uvw) à une loi multinomiale de paramètres le vecteur de probabilité à n_f composantes $\mathbf{p}(\alpha) = [p_1(\alpha), p_2(\alpha), \dots, p_{n_f}(\alpha)]^T$. n_f désigne le nombre de faciès considérés et $p_k(\alpha)$ la probabilité d'occurrence du faciès F_k . Ce cube de proportions permet de réaliser les simulations stochastiques en faciès.

Grâce aux simulations stochastiques, on obtient plusieurs réalisations d'un même objet. Elles simulent les distributions statistiques des propriétés sédimentaires et pétrophysiques en chaque point du modèle maillé de réservoir (Fig. 2.5), tout en respectant la cohérence spatiale de l'information disponible (puits, proportions de faciès, ...). À l'instar des méthodes de Monte-Carlo qui « tirent » un nombre selon une certaine distribution, ici, on « tire » un modèle 3D : en faciès, en porosité, etc., dans l'ensemble des modèles statistiques spatiaux possibles, sous la contrainte de respecter les observations aux puits, l'information sismique disponible, les relations entre différentes variables mais surtout le scénario géologique. L'ensemble des simulations permet d'apprécier l'impact des incertitudes sur les paramètres pétrophysiques, sur les volumes disponibles puis sur la capacité à les exploiter. Comme pour les tirages de Monte-Carlo, il est possible de combiner cette

approche avec les méthodes bayésiennes ou de ré-échantillonnage [Journel, 1994] afin de contraindre au mieux les lois statistiques des variables lithologiques et pétrophysiques.

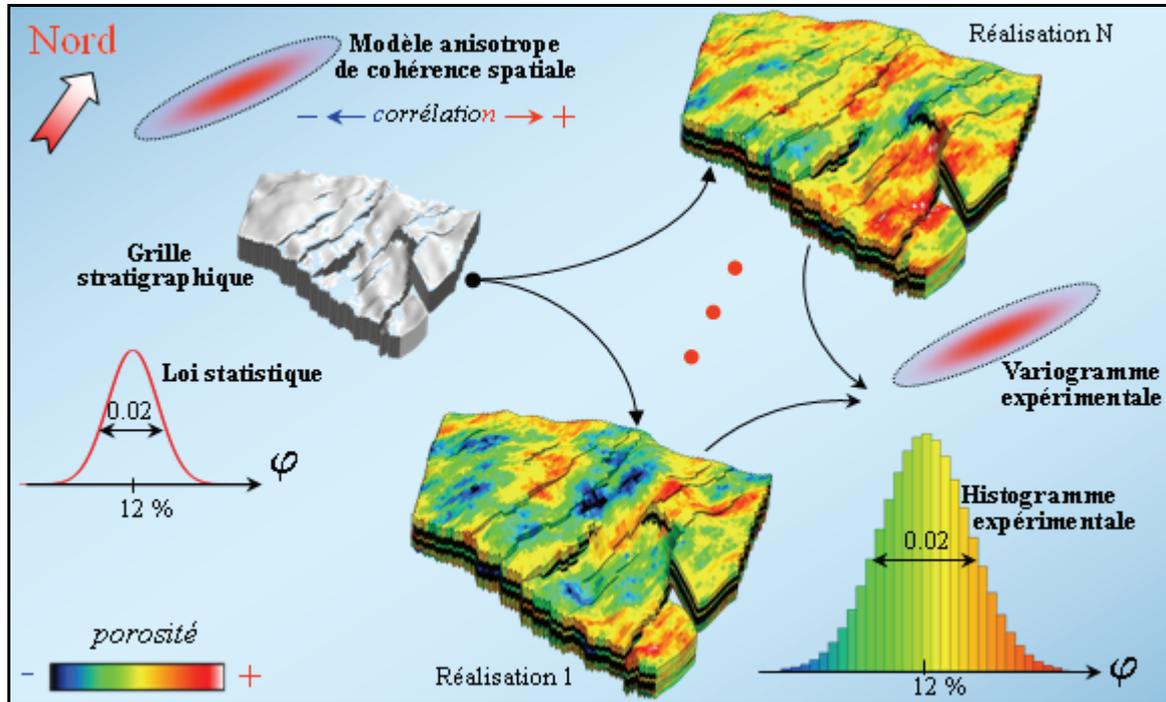


FIG. 2.5 : Simulations géostatistiques de la porosité. À partir d'une loi statistique et d'un modèle de continuité spatiale (a), la porosité est simulée en tout point de l'espace (b). Chaque réalisation présente un histogramme des valeurs simulées calculé sur l'ensemble de la grille correspondant à la loi statistique initiale, et qui traduit l'anisotropie de la structure spatiale imposée (c). Remarquons que les résultats des simulations restent comparables. Il faut donc réaliser l'étude des incertitudes en amont des modèles géostatistiques (i.e déterminer celles sur les modèles de lois et de continuité).

Ces images alternatives du champ sont sensées reproduire son hétérogénéité et rendre compte de l'aléa géologique. Toutefois, les différences entre réalisations ne reflètent pas la variabilité globale du réservoir car le scénario géologique comme les proportions de faciès qui en découlent, restent figés [Caumon *et al.*, 2004]. Les véritables sources de l'incertitude géologique, comprises dans la définition même du scénario et qui se traduisent en partie sur la loi spatiale (modèle en proportions), semblent donc oubliées. La multi-réalisation, quelle soit lithologique et/ou pétrophysique, représente en fait, des variations aléatoires d'ordre secondaire au regard de l'imprécision que portent le schéma sédimentaire et le cube de proportions qui en résulte (section 2.4). La multi-réalisation pétrophysique s'apparente alors à un bruit autour de l'image précise que donne le scénario sédimentaire à travers le cube de proportions [Poirot, 2002 ; Sheppers, 2003]. En revanche, si l'on veut retranscrire l'incertitude propre au scénario sédimentaire, il reste possible de rendre stochastique certains paramètres du modèle géologique comme la direction des

apports sableux, la forme des objets géologiques, etc [Norris *et al.*, 1993 ; Corre *et al.*, 2000]. Ces méthodes oublient par contre une part essentielle du scénario géologique : le cube de proportions. Ce dernier, qui régit entre autres l'hypothèse de stationnarité du modèle, est généralement considéré comme déterministe. Or, le choix d'un modèle de probabilité (cube de proportions) unique peut avoir une incidence très importante sur les grandeurs clés de l'industrie, et donc sur les décisions qui pourraient en découler.

L'exemple suivant souligne les limites actuelles de la multi-réalisation en propriétés lithologiques puis pétrophysiques. Il revient sur le réservoir carbonaté DOUSHKA, reconnu par trois puits et une campagne sismique. Ce champ présente la morphologie d'une plate-forme carbonatée¹. Les milieux de dépôts s'organisent ainsi en zones pseudo parallèles depuis les milieux fortement agités jusqu'aux secteurs les plus confinés du domaine interne (à proximité du continent) [Arnaud, 2004a]. Les géologues invités à modéliser la distribution spatiale des faciès, distinguent six faciès sédimentaires discriminants en terme de pétrophysiques. Par contre, si, à partir des données disponibles (puits et sismique) et de leur connaissance, les experts sont d'accord sur un environnement général de dépôt de type plate-forme carbonatée, la spécification de certains aspects du scénario géologique, comme les caractéristiques de la plate-forme (morphologie, taille, orientation, pente, position, intensité et position d'une dérive éventuelle, etc.), se révèlent incertaines. Une telle incomplétude conduit naturellement à la possibilité de plusieurs schémas sédimentaires (Fig. 2.6), chacun d'eux se traduisant par un cube de proportions distinct.

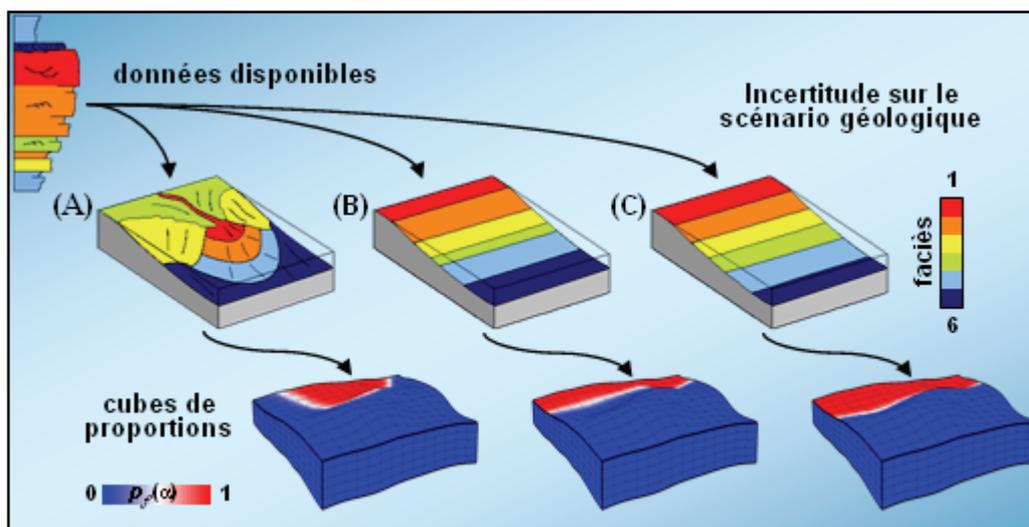


FIG. 2.6 : Effet de l'incertitude géologique sur le cube de proportions de faciès. Un même jeu de données risque de s'interpréter selon différents scénarios sédimentaires : ici notés (A), (B) ou (C) par exemple. Ceux-ci peuvent d'ailleurs présenter des morphologies plus ou moins différentes. L'incertitude sur le choix du modèle géologique se répercute donc sur le cube de proportions de faciès : autant de cubes sont envisageables.

¹ **Plate-forme carbonatée** : elle correspond à un domaine marin peu profond localisé dans les mers chaudes et caractérisée par des dépôts constitués uniquement de particules carbonatées.

Pourtant, en règle générale, par manque de temps et/ou de moyens humains, un seul scénario sédimentaire (cube de proportions) est envisagé lors de la construction du modèle de réservoir – ici, le modèle noté (B) (Fig. 2.7). L'impact des incertitudes géologiques sur les performances du gisement est étudié en faisant varier, à partir du cube de proportions choisi, les paramètres lithologiques (nature des faciès) et pétrophysiques : porosité, perméabilité et saturation en hydrocarbures. En pratique, des simulations géostatistiques génèrent, à partir d'un schéma sédimentaire et du cube de proportions qui en dérive, une multitude de réalisations équiprobables en propriétés lithologiques et pétrophysiques. Cependant, l'ensemble de ces réalisations ne reflètent pas la variabilité globale du réservoir (scénario (A) et (C)). En effet, seul le scénario géologique (B) est investigué, les proportions de faciès qui s'ensuivent, restant figées.

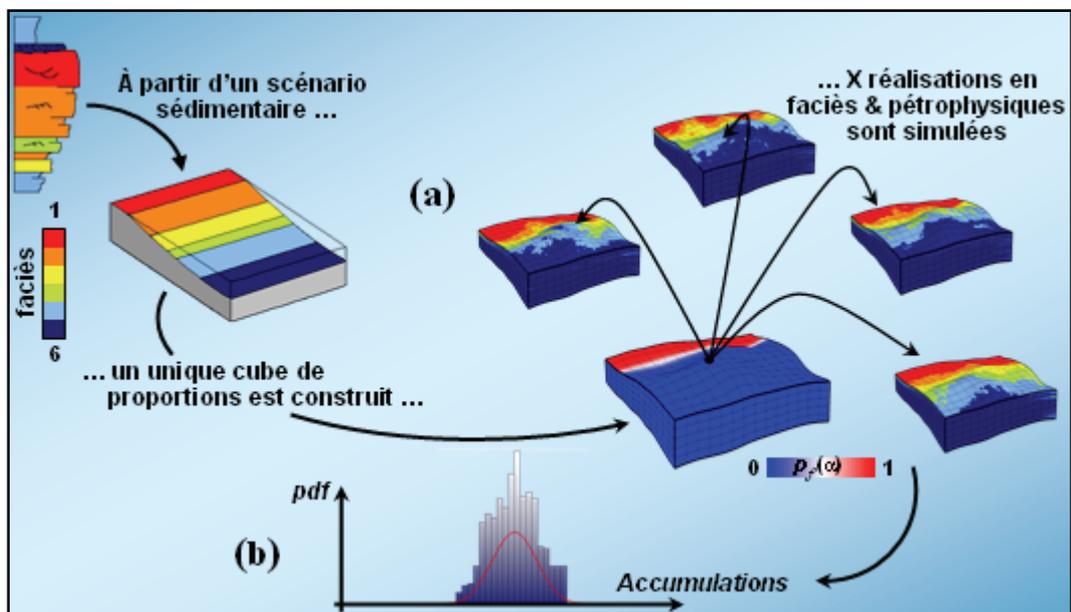


FIG. 2.7 : Étude classique des incertitudes sur le réservoir Doushka. De manière générale, on envisage un modèle géologique unique, ici le scénario (B). Ce choix revient ainsi à « figer » le cube de proportions, c'est-à-dire le modèle probabiliste sous-jacent. À partir de celui-ci, des images alternatives du champ en termes de faciès, puis de pétrophysiques sont simulées (a). Pour chaque réalisation, on calcule l'accumulation en hydrocarbure correspondante. Il est possible, avec un nombre suffisant de simulations, d'obtenir une densité de probabilité des volumes en place (b), et d'évaluer en partie l'impact de la variabilité des propriétés lithologiques et pétrophysiques. Pourtant, leurs variations d'une réalisation à l'autre ne balayent pas complètement l'incertitude sur le scénario géologique (cube de proportions), celui-ci restant figé.

Les simulations géostatistiques construites autour d'un cube de proportions unique ne suffisent donc pas à décrire la variabilité des modèles en proportions liée à l'incertitude sur le scénario sédimentaire. Son principal écueil réside dans la confusion de deux niveaux distincts de l'incertitude géologique. Le premier, aux effets majeurs, est d'abord lié

au caractère imprécis et incomplet de l'information en raison d'un manque de connaissance. Cette incertitude se traduit sur le scénario géologique, et se répercute au final sur le cube de proportions en faciès. Le second, de moindre importance, provient du caractère aléatoire de l'information relatif à la variabilité naturelle des propriétés lithologiques et pétrophysiques. Si ce dernier peut être appréhendé dans un cadre probabiliste en ayant recours à la multi-réalisation, le premier degré d'incertitude géologique, lui par contre, ne l'est pas. Certaines méthodes cherchent alors à d'abord faire varier les proportions de faciès (chapitre 9) [Haas *et al.*, 2002 ; Biver *et al.*, 2002]. Cependant, celles-ci expliquent uniquement la variabilité d'origine statistique sur les proportions (échantillonnage aux puits et/ou sismique). Elles négligent par conséquent la part d'incertitude qu'introduit l'interprétation géologique sur ces mêmes probabilités d'occurrence de faciès. Ces méthodes font en plus, l'hypothèse souvent abusive, de stationnarité du premier ordre : la probabilité a priori de trouver un faciès donné est supposée constante quelle que soit la position. Elles simulent ainsi pour chaque intervalle stratigraphique issu des corrélations, une proportion constante de faciès. Une telle hypothèse peut se trouver complètement erronée dans certains cas. Pour les dépôts sur une rampe carbonatée par exemple, la nature des sédiments varie dans l'espace avec la bathymétrie. D'autres procédés perturbent plutôt le modèle initial de probabilités¹, et les réalisations géostatistiques des propriétés du sous-sol en découlant. L'idée consiste en effet à optimiser par itérations successives le calage du profil de production qui en résulte à un historique de production [Hu *et al.*, 2001 ; Caers, 2003]. Toutefois, il faut disposer d'un tel historique ; ce qui n'est pas toujours le cas, notamment en phase d'exploration. D'ailleurs, la morphologie des cubes de proportions ne change pas véritablement. Seul est modifié l'ordre de grandeur global des probabilités d'apparition des faciès. De fait, cette approche, comme les précédentes, se limite à l'étude d'un scénario géologique unique, et n'évalue donc pas l'incertitude sur ce paramètre (Fig. 2.7). Il semble par conséquent naturel de faire varier ce dernier, ou autrement dit, de véritablement intégrer la variabilité du cube de proportions qui en dérive. À ce jour, l'étude de différents scénarios s'avère nécessaire afin de dépasser les caractères déterministes et plus ou moins figés des proportions de faciès propres à chacun des modèles géologiques.

2.4 Les études à scénarios multiples

L'étude de plusieurs scénarios peut compléter l'approche précédente pour en dépasser le caractère déterministe sur les proportions de faciès (Fig. 2.8). À partir de quelques cubes de proportions, construits sur la base de scénarios géologiques distincts, des simulations géostatistiques font varier les paramètres pétrophysiques (porosité, perméabilité, saturation,...). La démarche demeure néanmoins limitée par son caractère discret. Seul l'examen de quelques cas reste envisageable en raison de la lourdeur des simulations géostatistiques mises en œuvre. D'ailleurs, de manière générale, on ne travaille pas vraiment sur des scénarios géologiques différents. On part typiquement d'un « cas de base »

¹ c'est-à-dire le cube de proportions de faciès

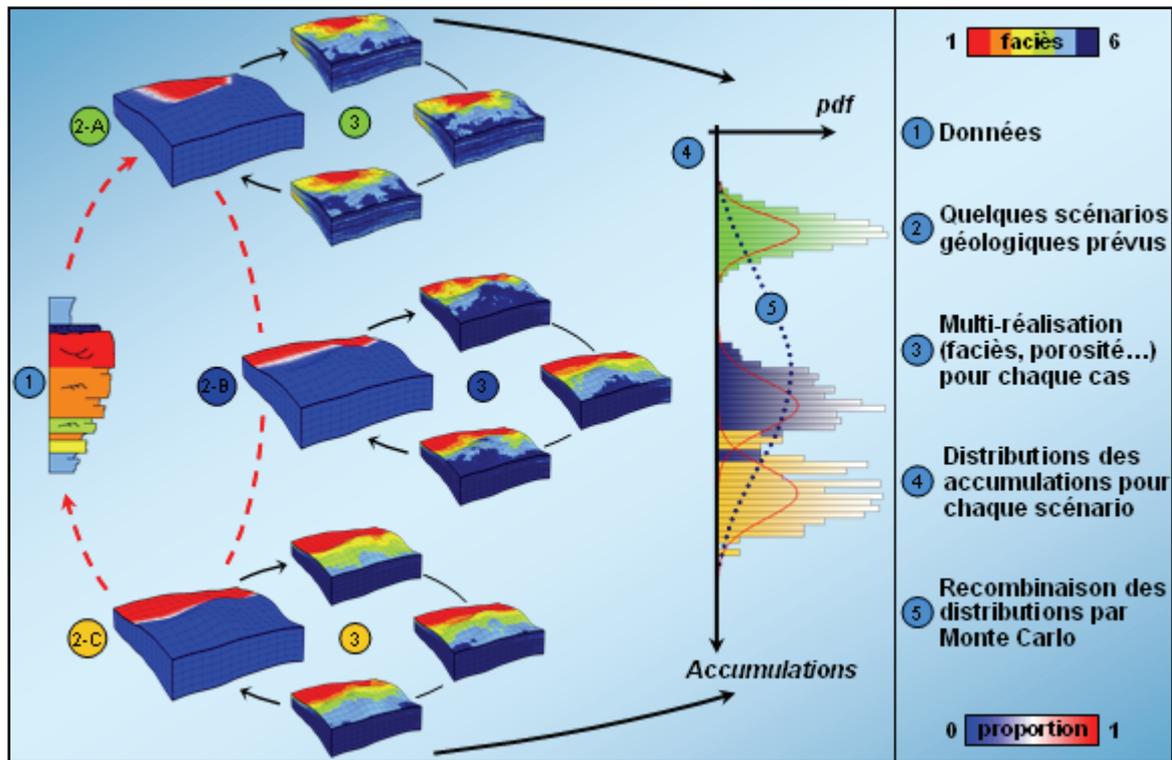


FIG. 2.8 : Étude à scénarios multiples du réservoir Doushka. L'interprétation d'un même jeu de données (1) conduit à envisager trois scénarios géologiques distincts dont dérivent trois cubes de proportions (2-A, B, C). Pour chaque cube de proportions, on procède à de la multi-réalisation sur les propriétés lithologiques et pétrophysiques (3), fournissant respectivement les distributions en accumulations pour les différents cas envisagés (4). Une démarche de type Monte Carlo permet de recombinaison ces distributions par combinaison linéaire, d'avoir une distribution continue.

que l'on dégrade en un cas pessimiste et/ou que l'on améliore en un « cas optimiste ». Le cas pessimiste minimise l'importance des bons faciès réservoirs, en ne retenant que les bons faciès prouvés aux puits et en limitant leur extension latérale. Cependant, on ne modifie pas véritablement la carte de base (la morphologie). Aussi ne rend-t-on pas véritablement compte de l'incertitude sur la disposition latérale des proportions de faciès. Pour dépasser cet aspect discret, il est possible de combiner les accumulations calculées sur les quelques cubes de proportions issus des scénarios géologiques concevables à travers une analyse de Monte-Carlo (Fig. 2.8-5). On obtient une distribution des accumulations qui tient compte de l'incertitude relative au choix du scénario géologique. Toutefois, la méthode ne balaye qu'une partie de l'espace d'incertitude. De telles combinaisons empêchant toute nouveauté, les modèles fournis au départ deviennent en quelque sorte des cas "extrêmes".

Les approches actuelles ne suffisent donc pas à décrire la variabilité des modèles en proportions, et qui est liée à l'incertitude sur le scénario sédimentaire. Il faudrait pour cela

pallier les caractères discrets et limités d'une étude à plusieurs scénarios ou de combinaisons d'accumulations après multi-réalisation pétrophysique. Ces méthodes se limitent en effet à des modèles en proportions plus ou moins figés. À contrario, les méthodes proposées dans les parties suivantes s'intéressent plus précisément au cube de proportions afin de quantifier l'incertitude sur les paramètres du schéma sédimentaire. Elles visent à améliorer l'évaluation de la variabilité sur les proportions de faciès et sur les grandeurs clés d'un champ - accumulation en hydrocarbure par exemple. Dans la partie à venir, on adopte une démarche de type Monte-Carlo, mais cette fois sur les paramètres qui composent le schéma sédimentaire (Fig. 2.9-I.a). Chaque tirage permet de construire un cube de proportions par l'intermédiaire d'un géomodeleur qui intègre de façon plus ou moins explicite les paramètres du scénario géologique. Ici, il s'agit de Neptune. Dans la troisième partie, la méthode présentée considère les cubes de proportions dans leur globalité. Elle

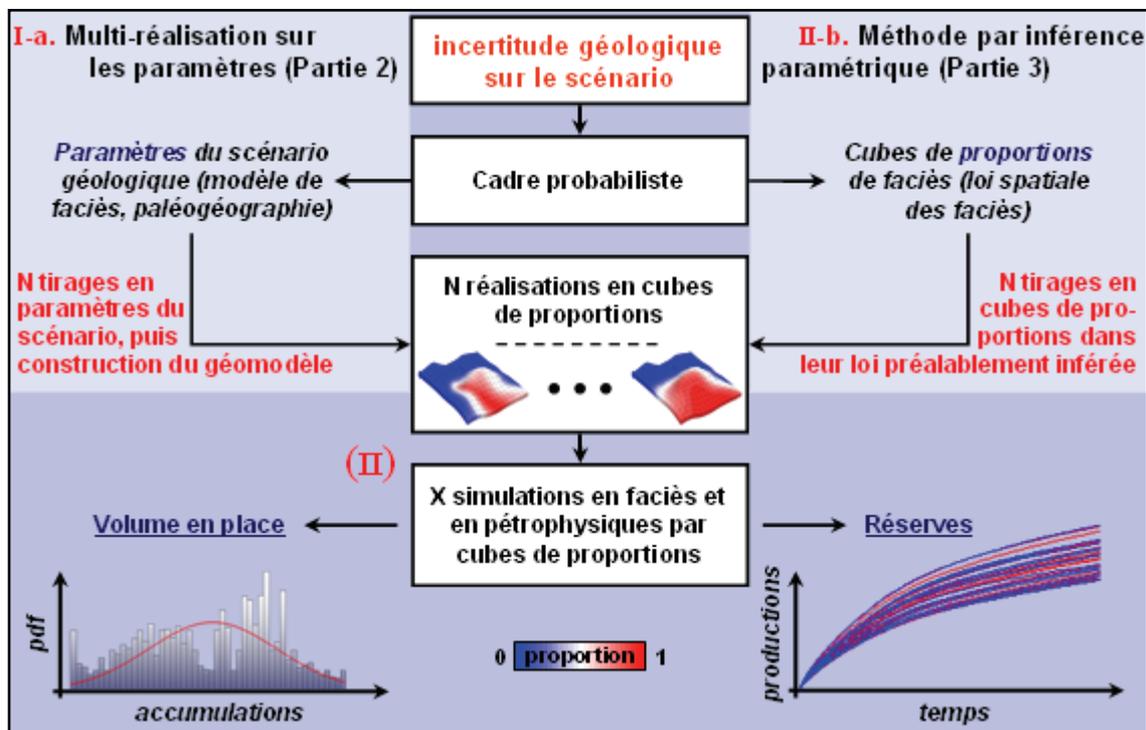


FIG. 2.9 : Intégration des deux niveaux d'incertitude géologique. (I-a) La méthode décompose le scénario géologique en ses principaux paramètres (paléogéographie, modèle de faciès, etc.), ceux-ci étant rendus probabilistes. Chaque tirage d'un ensemble de paramètres aboutit à la construction d'un scénario géologique, puis du cube de proportions qui en découle. (I-b) Une méthode alternative s'intéresse au cube de proportions dans sa globalité. À partir de quelques modèles de base, elle infère leur loi de distribution, dont les réalisations correspondent à des cubes de proportions. Quelle que soit la méthode utilisée (I-a) ou (I-b), les variations d'un cube de proportions à l'autre intègrent l'incertitude sur le scénario géologique. (II) Pour chaque cube de proportions ainsi simulé, on procède à de la multi-réalisation sur les propriétés lithologiques et pétrophysiques en vue d'intégrer l'incertitude de nature aléatoire sur la pétrophysique.

génère, à partir d'un échantillon de quelques cubes de proportions issus des différentes interprétations géologiques envisageables, l'ensemble des images alternatives du réservoir (Fig. 2.9-I.b). Le recours à la notion de variables aléatoires au niveau des proportions de faciès permet ce passage et l'intégration de l'incertitude géologique (chapitre 8). Au sortir de ces deux approches, on dispose de multiples réalisations en terme de cubes de proportions – incertitude sur le schéma sédimentaire. Celles-ci traduisent l'incertitude sur la variabilité des proportions de faciès. Ainsi, l'intérêt des deux techniques repose sur le passage d'hypothèses discrètes – les scénarios géologiques, à un phénomène continu. D'ailleurs, des simulations géostatistiques autour de chacun des modèles en proportions préalablement construits – incertitude pétrophysique, peuvent compléter ces approches (Fig. 2.9-II).

Deuxième partie

Multi-réalisation sur les paramètres du scénario géologique

Cette partie est dédiée aux solutions apportées à l'identification et la quantification des incertitudes géologiques dans Neptune. Il s'agit d'un processus inédit de modélisation des faciès en milieu marin élaboré au sein du groupe TOTAL, et implémenté sous la forme d'une extension de gOcad¹ [Massonnat, 1999 ; Moyen, 2001 ; Gaziglia, 2001 ; Leflon, 2005]. Une particularité essentielle de Neptune réside dans l'utilisation de la paléobathymétrie afin (a) de définir une grille de réservoir dont l'architecture interne est cohérente avec la stratigraphie réelle et, (b) de générer des probabilités d'apparition de faciès sur l'ensemble de la zone. Nos travaux se situent en parallèle des derniers développements réalisés par Leflon sur la méthodologie. Il était question pour lui de perfectionner la méthode afin de l'adapter aux cas complexes où le réservoir est plissé ou faillé, et où les puits sont déviés [Leflon, 2005]. Quant à nous, l'objectif est d'intégrer les incertitudes relatives aux paramètres qui définissent un scénario géologique, et qui représentent des grandeurs d'entrée explicites pour Neptune. Dans un premier temps, on présente les concepts attachés à cette méthodologie. L'accent est mis sur les paramètres géologiques d'entrée clés qui potentiellement sont entachés d'incertitude : paléogéographie, modèle de faciès, ... (**chapitre 3**). Pour cela, on revient sur les différentes analyses de sensibilité qui ont été réalisées pour étudier l'influence des divers facteurs incertains sur la variabilité de la réponse du modèle en termes d'accumulations et/ou de productions. Les chapitres suivants développent l'approche multi-réalisation et sa mise en oeuvre sur les paramètres du scénario géologique entrés dans Neptune, qu'ils soient en une, deux ou trois dimensions (respectivement **chapitres 4, 5 et 6**).

¹ **gOcad** est un outil de modélisation des objets géologiques du sous-sol quelle que soit l'échelle d'étude, en particulier à l'échelle réservoir où de puissants outils géophysiques et géostatiques ont été développés.

Chapitre 3

Incertitude et Neptune

Ce chapitre revient d'abord sur les principes que la méthodologie Neptune met en œuvre, avant d'en présenter les étapes importantes. Il s'agit d'un modèleur qui permet la construction de modèles pétrophysiques 3D en milieu marin. En terme d'incertitude, l'intérêt du processus réside dans la possibilité d'intégrer de façon explicite les paramètres du scénario géologique. Aussi, mettra-t-on l'accent sur les facteurs du scénario jugés a priori les plus influents sur la qualité et la cohérence du modèle final. Car, l'intégration de l'incertitude géologique oblige à savoir où et comment elles interviennent au cours du processus. On développe par la suite l'approche multi-réalisation utilisée dans cette partie sur les données géologiques d'entrée afin de traduire leur variabilité. Elle diffère des méthodes classiques dans le sens où elle intervient directement sur les paramètres à l'origine des cubes de proportions. Elle autorise ainsi la description de l'ensemble des scénarios géologiques alternatifs, et permet de traduire l'incertitude que portent ces scénarios sur les accumulations et/ou productions.

3.1 Neptune, bases et concepts

L'application Neptune, conçue par Gérard Massonnat (1999, 2002b), est une méthodologie de modélisation du sous-sol qui combine à la fois les concepts de la géostatistique (chapitres 2 et 9), et de la stratigraphie séquentielle¹ [Vail *et al.*, 1977]. Cette méthode permet entre autres :

- la construction d'une grille de réservoir compatible avec la stratigraphie séquentielle ; elle reproduit, à l'échelle du réservoir, les cortèges transgressifs ou régressifs enregistrés aux puits ;

¹ **Stratigraphie séquentielle** : c'est l'étude des relations notamment génétiques qui existent entre les roches se trouvant à l'intérieur de successions sédimentaires de même âge stratigraphique, bornées par des surfaces d'érosion ou l'absence de dépôts. Elle sert à prévoir le caractère lithologique de la roche ou à ordonner les étapes de remblaiement. À haute résolution, elle permet la corrélation entre des gisements de pétrole malgré les variations latérales de faciès.

- la modélisation en tout point du réservoir maillé de proportions de faciès cohérentes avec les données dures, la connaissance géologique et les hypothèses du scénario sédimentaire.

Pour ce faire, Neptune s'appuie sur l'hypothèse qu'en environnement marin, une relation forte existe entre d'un côté, la bathymétrie¹, l'énergie de dépôt², le contexte stratigraphique³ et de l'autre, la nature du faciès résultant. Elle cherche donc à modéliser la bathymétrie afin d'en déduire les probabilités d'apparition de faciès. À partir des carottes, il est possible d'estimer la bathymétrie le long des puits en observant les faciès. Par contre, son extrapolation se révèle aussi difficile et incertaine que pour les autres paramètres connus aux puits, qu'ils soient géologiques (faciès, hauteur utile, marqueurs de diagenèse,...) ou pétrophysiques (porosité, perméabilité,...). En effet, la bathymétrie, notamment dans le cas des plates-formes carbonatées, est une propriété non stationnaire tant dans l'espace que dans le temps. La faible quantité et/ou la représentativité discutable des observations aux puits comme la résolution limitée des cartes que fournit la sismique empêchent généralement de reproduire son évolution spatiale et temporelle (chapitre 1). Aussi, extrapole-t-on la bathymétrie par le biais d'un signal plus commode à manipuler, qui décrit une histoire commune à l'ensemble du réservoir, et qui offre ainsi des résultats moins bruités et dépendants des données : *l'accommodation*.

Cette propriété, notée $A(x,y,t)$ sur la figure 3.1, et encore appelée *niveau marin relatif*, correspond à l'instant t , pour le point M de coordonnées (x,y) , à la distance comprise entre la surface de l'eau et l'horizon de référence. Concrètement, elle représente l'espace total disponible pour accueillir les sédiments à ce moment précis. Elle équivaut donc à la somme de :

- (1) la bathymétrie $b(x,y,t)$, qui est l'espace libre pour la sédimentation entre la surface de l'eau et l'interface eau/sédiment à l'instant t ;
- (2) l'épaisseur de sédiments déposée $e(x,y,t)$, qui représente la distance entre l'interface eau/sédiment et le substratum⁴ à l'instant t ,

selon la relation :

$$A(x,y,t) = b(x,y,t) + e(x,y,t) \quad (\text{Éq. 3.1})$$

¹ **Bathymétrie** : Elle mesure la profondeur d'eau ; c'est le niveau de la mer mesuré à partir de la surface supérieure du sédiment.

² **L'énergie de dépôt** renvoie à la limite d'action des vagues, elle dépend de l'exposition du bassin aux vents dominants, de sa protection par une éventuelle barrière récifale (corallienne)

³ **Contexte stratigraphique** : Il renvoie au type de cortège de dépôt dans lequel on se trouve : est-on en système progradant ou transgressif ?

⁴ Par **substratum**, on entend un horizon de référence choisi arbitrairement : on considère en général la surface temps correspondant au début de la sédimentation.

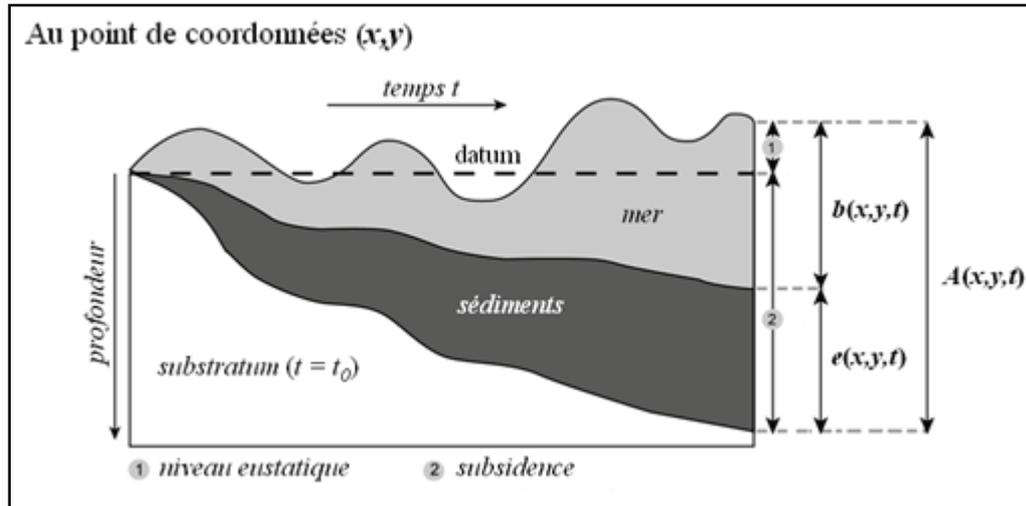


FIG. 3.1 : Évolution dans le temps des variables associées à l'accommodation au point de coordonnées (x,y) (d'après Homewood et al., 2000).

L'observation des faciès aux puits rend possible le calcul de la bathymétrie de dépôt $b(x,y,t)$ et de l'épaisseur de sédiments $e(x,y,t)$ le long des puits. À partir de l'équation 3.1, on évalue l'accommodation $A(x,y,t)$ en tout point le long des puits.

Cependant, l'information intéressante pour conduire l'extrapolation des données ne passe pas seulement par l'accommodation. Elle ressort plutôt de l'examen de son évolution dans le temps qui, à l'échelle du réservoir, peut être considérée comme un phénomène global [Homewood et al., 2000]. Un événement géologique comme une baisse du niveau marin par exemple, peut produire des transitions différentes de faciès selon sa position sur la plate-forme (proximale ou distale). Par contre, cette diminution se traduit à cet instant précis par une signature commune à l'ensemble du réservoir : un pic négatif sur la courbe d'évolution de l'accommodation, caractéristique d'une destruction d'espace disponible pour sédimenter.

On appellera alors *potentiel d'accommodation* ou *subsidence relative* $s(x,y,t)$, la variation d'accommodation entre le temps de référence t_0 et le temps t , pour un point M de paléo-coordonnées (x,y) . En s'appuyant sur l'équation 3.1, ce potentiel se définit par la formule :

$$s(x,y,t) = A(x,y,t) - A(x,y,t_0) = b(x,y,t) - b(x,y,t_0) + e(x,y,t) - e(x,y,t_0) \quad (\text{Éq. 3.2})$$

Dès lors, grâce aux concepts de la stratigraphie séquentielle, nous pouvons interpréter les variations d'accommodation observées aux puits à l'échelle du réservoir. Il devient alors possible, par le biais des corrélations entre puits, de relier les points de même âge stratigraphique, et de construire une grille qui reproduit les caractères géologiques. On obtient un paramétrage du temps géologique t , nécessaire au recalage au niveau des puits des signaux $b(x,y,t)$, $e(x,y,t)$, $A(x,y,t)$ et $s(x,y,t)$, et à la définition d'un espace « pseudo-

temps » (x,y,t) . C'est l'équivalent d'un diagramme de Wheeler [Wheeler, 1958] dans lequel tous les dépôts, non-dépôts ou érosions contemporains sont placés sur un même plan horizontal (à temps t constant). Les corrélations définissent au final un découpage vertical dans lequel les différentes couches sont cohérentes avec la stratigraphie réelle : une strate est plus ancienne que toute autre strate placée au-dessus d'elle, mais plus récente que celles situées au-dessous.

Il s'agit désormais d'extrapoler spatialement la subsidence relative $s(x,y,t)$, connue uniquement aux puits, le long de ces surfaces isochrones. Or, cette propriété correspond également à la somme de l'eustatisme¹ et des mouvements tectoniques du bassin² (Fig. 3.1), et représente de fait, un phénomène global à l'échelle du réservoir. Le potentiel d'accommodation devrait donc être le même sur tous les puits à chaque instant t fixé. Toutefois, ce n'est que très rarement le cas, des bruits apparaissant souvent. On les doit aux approximations de modélisation : corrélations, calcul de la courbe de bathymétrie et de l'épaisseur ; ou du point de vue géologique à des phénomènes de subsidence différentielle. Le potentiel d'accommodation est alors séparé en deux composantes : le facteur commun, correspondant à la moyenne du potentiel d'accommodation calculée sur chaque horizon, et les résidus, qui sont la différence entre le facteur commun et le potentiel d'accommodation. Le facteur commun et les résidus, connus seulement aux puits, sont extrapolés indépendamment sur chaque strate. On les additionne, pour d'abord retrouver la subsidence relative $s(x,y,t)$, et enfin l'accommodation. La bathymétrie découle directement de l'inversion de l'équation 3.1.

Une fois la bathymétrie connue dans l'espace et le temps, les probabilités d'occurrence (ou "proportions") de chaque faciès sont alors estimées en chaque point du modèle en se servant de la relation bathymétrie/faciès. Un cube de proportions cohérent géologiquement est obtenu, qui permet de contraindre les simulations stochastiques en faciès, puis de modéliser les propriétés pétrophysiques indépendamment pour chaque type de roche. Néanmoins, pour les carbonates en particulier, la porosité ou la perméabilité actuelle ne dépendent pas uniquement de la nature du faciès déposé. Il faut aussi compter avec les phénomènes de diagenèse précoce (dissolution, cimentation, etc.) liés aux roches de surface³, qui en général, bouleversent fortement les qualités pétrophysiques originelles des sédiments. Or, l'histoire bathymétrique de la zone permet d'identifier les régions émergées, et qui sont susceptibles de connaître une diagenèse pendant la formation du réservoir. Un même faciès se verra alors affecter des porosités et perméabilités différentes selon qu'il se trouve ou non dans une région ayant été émergée. De fait, le modèle est renseigné de propriétés pétrophysiques qui traduisent à la fois la nature du faciès déposé et son comportement diagénétique.

¹ **Eustatisme** : Il s'agit des variations du niveau marin absolu, qui représente la distance entre la surface de l'eau et un point de référence fixe (le centre de la Terre par exemple)

² On parle de **subsidence** quand le bassin s'enfonce au cours du temps; et de **surrection** lorsqu'il remonte.

³ Nous ne distinguons pas les sédiments proches de la surface et non encore enfouis, de ceux qui ont subi une ou plusieurs émergences au cours de leur diagenèse.

Au final, l'approche Neptune renvoie une image réaliste et cohérente de l'hétérogénéité du réservoir que ce soit sur le plan :

- de son architecture interne, avec un maillage vertical que définissent des surfaces isochrones ;
- des proportions de faciès, dont le calcul intègre toute la connaissance qu'apporte le scénario géologique ;
- ou de la pétrophysique, dont l'extrapolation respecte l'histoire diagénétique du milieu.

Toutefois, si le modèle construit est géologiquement cohérent, il n'est pas nécessairement « juste ». Il peut ne traduire qu'une vision erronée de la véritable géologie du réservoir, par essence incertaine. Plus exactement, Neptune s'appuie sur l'hypothèse d'un scénario géologique « juste » ; or, un scénario faux n'aboutit pas forcément à un modèle incohérent. La quantification de l'incertitude sur les accumulations ou productions oblige donc à intégrer la variabilité des paramètres du scénario sédimentaire, paramètres jugés incertains.

3.2 Neptune, les paramètres incertains

La figure 3.2 revient sur les principales étapes de construction d'un modèle Neptune : de la description des carottes disponibles aux puits à la simulation stochastique des faciès dans l'espace discrétisé. Pour plus de détails sur les divers modes de calcul, nous invitons le lecteur à lire les travaux des programmeurs ayant développé la méthodologie (Moyen, 2001 ; Gaziglia, 2001), en particulier, la thèse de Bruno Leflon (2005) à l'origine de l'implémentation actuelle de Neptune. Au-delà des modalités de calculs, cette figure précise plutôt les étapes et les choix de modélisation qui incluent la connaissance géologique comme donnée d'entrée. Elle recense les paramètres que nous supposons a priori les plus influents. Le déroulement linéaire du protocole Neptune peut se diviser en trois grandes étapes selon la nature et l'interprétation de l'information introduite, et qui feront l'objet d'une étude plus approfondie aux chapitres 4, 5 et 6 :

- (1) **le long des puits – en 1D** – à partir de sa description des faciès aux puits et de sa connaissance, le sédimentologiste propose une relation entre la profondeur de dépôt et la nature du faciès : c'est le **modèle de faciès** (Fig. 4.1), qui comme toute interprétation est incertain (section 4.1). Cette correspondance admise, on estime la bathymétrie grâce à l'interpolateur DSI [Mallet, 1992, 2002 ; Leflon, 2002]. À ce stade, apparaît une incertitude de deuxième ordre comparée au choix du modèle de faciès : l'aspect déterministe du **mode de construction** de la courbe de bathymétrie. Les données disponibles permettent également de calculer les épaisseurs, l'accommodation et la subsidence relative $s(x,y,t)$ le long des puits. Les **corrélations** entre puits, qui relient les points de même âge géologique, ont permis de définir le temps t et de construire un espace « pseudo-temps » qui facilite l'extrapolation des données. La définition de ces corrélations porte une part a priori



FIG. 3.2 : Les étapes du processus Neptune et les paramètres d'entrée associés. À gauche, les données, au centre la méthodologie en elle-même, à droite, les paramètres liés à la connaissance et au scénario géologique, potentiellement entachés d'incertitude.

importante d'incertitude. Elle traduit l'analyse des données à travers les concepts de la stratigraphie séquentielle, et contrôle en grande partie l'architecture interne du réservoir (section 4.2) ;

- (2) **en termes de surfaces et de cartes – en 2D** – connaissant les différents signaux $b(x,y,t)$, $e(x,y,t)$, $A(x,y,t)$ et $s(x,y,t)$ aux puits, il s'agit de les extrapoler le long des surfaces isochrones définies précédemment. L'extrapolation des épaisseurs et des composantes de la subsidence relative passe par les techniques géostatistiques (krigeage, simulations, etc.). Néanmoins, le **modèle de continuité géologique** (choix des variogrammes, chapitre 1) semble difficile à définir précisément en raison du manque de données disponibles (nombre de puits limité). Les valeurs d'entrée des paramètres du variogramme sont donc entachées d'incertitudes. Quant à la bathymétrie, on l'estime en utilisant l'épaisseur des horizons et l'accommodation calculées précédemment, ainsi qu'une **carte de paléo-bathymétrie** de référence à l'instant t_0 (Éq. 3.2). Cette carte correspond à la paléogéographie au temps t_0 . En règle générale, elle est évaluée par un sédimentologue (dessin, croquis, etc.) à l'aide de sa connaissance du bassin et des données sismiques. Elle revêt naturellement un caractère hautement incertain (chapitre 5) ;
- (3) **dans le réservoir – en 3D** – une fois la bathymétrie définie dans l'ensemble du réservoir maillé, il est possible, grâce au modèle en faciès, de déterminer les probabilités d'apparition de chaque faciès connaissant la bathymétrie. Ce calcul repose sur la formule de Bayes relatives aux probabilités conditionnelles (sections 2.2.1 et 6.1). Celle-ci confronte le sédimentologue à la tâche délicate de fournir la **probabilité a priori d'apparition** (appelée aussi poids) de chaque faciès. Il s'agit pour lui d'estimer la probabilité d'observer un faciès donné sans posséder la moindre information sur la bathymétrie, l'énergie de dépôts, etc. Plusieurs méthodes existent pour évaluer le poids des faciès, mais aucune n'efface complètement l'incertitude que porte son estimation (section 6.1). Cependant, une fois le cube de proportions construit, il peut être utilisé afin de contraindre les simulations en faciès. Par ailleurs, sachant la bathymétrie, il est également possible d'identifier les zones susceptibles d'avoir subi une **diagenèse précoce**. Cette dernière affecte les sédiments lorsqu'ils sont encore proches de la surface (jusqu'à quelques mètres de profondeur), ou bien émergés. À ce stade, une incertitude existe sur le volume des régions exposées compte tenu de la difficulté d'estimer l'extension de la zone vadose¹ : degré d'ouverture du biseau de diagenèse notamment (section 6.2). Par contre, une fois ces régions déterminées, les propriétés pétrophysiques sont simulées indépendamment selon que ces régions ont connu une diagenèse précoce ou non, afin de reproduire le comportement diagénétique des sédiments.

Dans les chapitres à venir – 4 à 6 – nous reviendrons de manière plus approfondie sur la description des principaux paramètres incertains. Car, toutes les sources de variabilité interne au scénario géologique n'ont pas la même incidence sur les proportions des faciès,

¹ **Vadose** : Se dit de l'eau qui percole dans le sous-sol, sans faire partie d'une nappe d'eau souterraine

les accumulations en huile, les profils de productions. Parmi les facteurs ainsi recensés, il s'agit désormais de déterminer ceux qui ont une influence non négligeable sur la variation des réponses. En effet, si on souhaite intégrer aux modèles Neptune les incertitudes relatives aux paramètres les plus influents, il faut établir une hiérarchie au sein de ces variables d'entrée. Le paragraphe suivant revient sur les résultats et conclusions de plusieurs études de sensibilité aux paramètres d'entrée de la méthodologie Neptune.

3.3 Neptune, hiérarchie des sources d'incertitudes

Quatre études méthodologiques [Poirot, 2002 ; Bergamo, 2002 ; Sheppers, 2003, Zerroune, 2003], distinguent et évaluent l'influence de ces données d'entrée inconnues sur la variabilité de la réponse d'un modèle Neptune, que ce soit sur le plan statique (proportions, accumulations), ou dynamique (productions). Elles mettent en œuvre sur des cas opérationnels ou synthétiques certaines méthodes d'analyse de sensibilité décrites par Saltelli *et al.* (1999, 2000a). Ces dernières étudient comment les variables de sortie d'un modèle réagissent à des perturbations de leurs entrées. Généralement, elles font varier les paramètres du modèle autour d'une valeur connue tout en gardant fixes les autres facteurs au cours de la modélisation [Félix *et al.*, 2005]. Ici, ces quatre études auront souvent fait appel à la technique des plans d'expérience, dont le domaine d'application se révèle finalement très large : marketing, publicité, robotique [Vivier, 2002], biologie, agriculture, pharmacologie, gestion de production, etc. De tels procédés modifient de manière simultanée les paramètres du modèle selon une série d'expériences déterminées a priori, ces paramètres ne pouvant prendre qu'un nombre fini de valeurs. Outre l'identification des facteurs sensibles, c'est-à-dire ceux pour lesquels une simple variation peut entraîner des changements considérables dans les sorties du modèle [Saltelli *et al.*, 2000b], ils permettent de déterminer les interactions possibles entre variables. Il existe différents types de plans d'expérience qui se distinguent suivant le nombre et le choix des combinaisons entre paramètres [Hines *et al.*, 2003]. Par exemple, les plans factoriels complets qui, pour P facteurs incertains variant continûment dans un intervalle de confiance, les fixent successivement à N niveaux de valeur, et décrivent l'ensemble des N^P configurations possibles. Le nombre d'expérience peut en outre être réduit : on parle alors de plans factoriels fractionnaires¹.

Dans ces quatre études, on considère ainsi, pour tous les paramètres d'entrée², deux niveaux de valeurs pris aux extrémités de leur intervalle de définition, afin de les faire varier avec le maximum d'amplitude. Chacune des combinaisons de facteurs étudiés,

¹ Les plans d'expérience les plus connus sont les plans de Taguchi à l'origine développé dans une optique d'utilisation industrielle, et qui ne prennent en compte que les interactions entre paramètres jugés importantes et en abandonnant d'autres ; on citera également les plans de Plackett-Burman qui omettent les interactions entre les facteurs les plus sensibles [Vivier, 2002].

² Ces études ne se sont pas intéressées à faire varier les corrélations ainsi que le poids des faciès. Elles ont étudié l'influence des autres paramètres, par groupe de six en général, en intégrant toujours la variabilité sur les cartes de paléo-bathymétrie.

fixés successivement à leurs deux valeurs possibles, définit un scénario géologique, lui-même à l'origine d'un modèle Neptune. L'effet de chaque modification est analysé sur trois sorties de Neptune, d'abord sur les probabilités d'occurrence des faciès dont la sensibilité est signifiée par le calcul d'un indice appelé "Indice de Shannon". Il s'agit d'un coefficient tiré de la théorie de l'information qui se révèle utile à l'étude de la biodiversité des espèces. Il traduit une entropie, autrement dit, le désordre [Shannon, 1948]. Les propriétés statiques (faciès, porosité, perméabilité, etc.) sont ensuite affectées à l'ensemble du

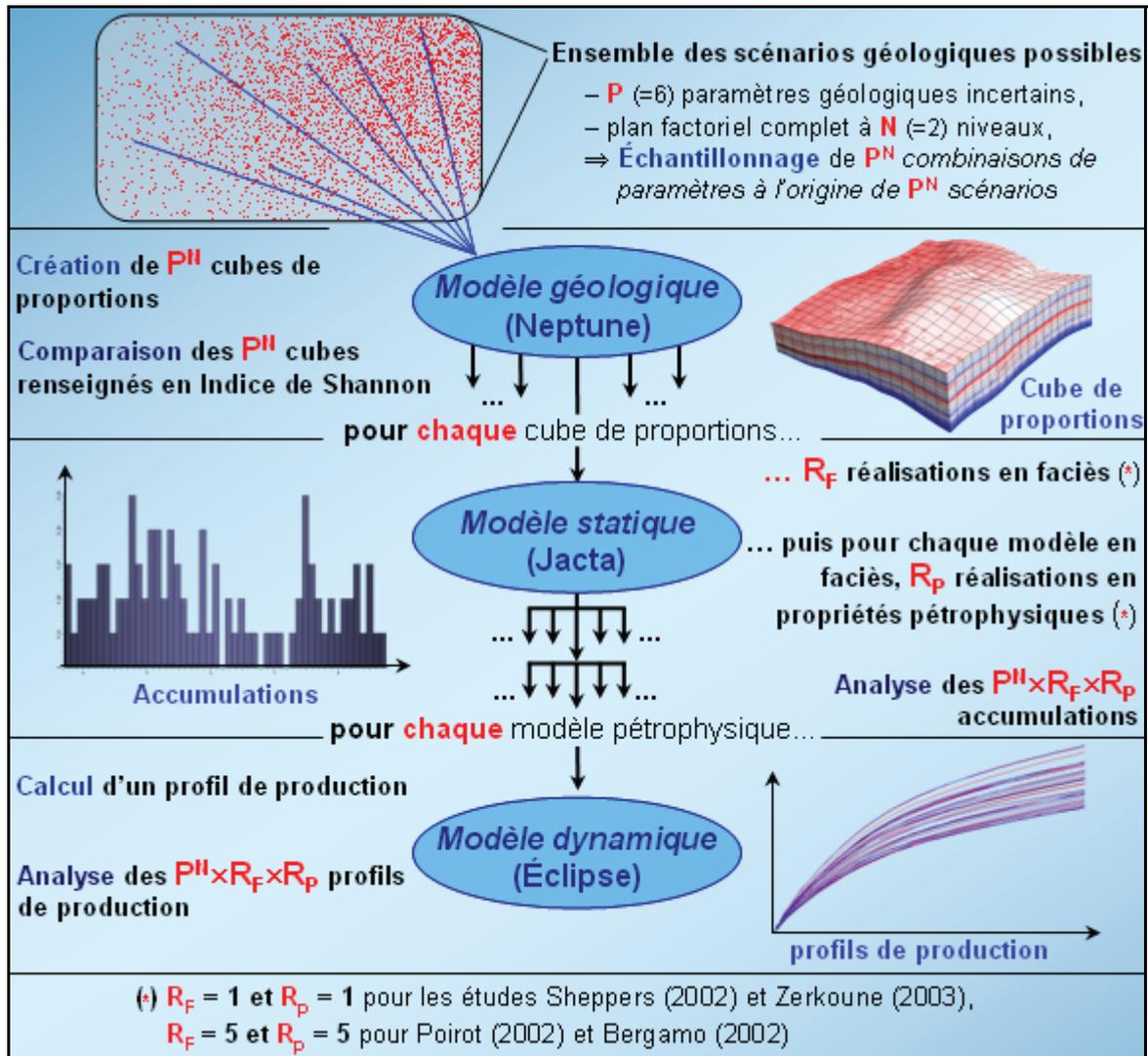


FIG. 3.3 : Principe général de quantification des incertitudes mis en œuvre dans les études de hiérarchisation des incertitudes. Les combinaisons de paramètres géologiques permettent de faire varier le modèle géologique construit sous Neptune. Les réalisations multiples en propriétés lithologiques et pétrophysiques sous Jacta, extension de gOcad quantifient le poids de la variabilité aléatoire sur les accumulations et sur les résultats dynamiques après simulation sous Éclipse, un simulateur d'écoulement.

modèle par simulation(s)¹ géostatistiques. On s'intéresse alors aux accumulations en hydrocarbures ; puis, à travers un modèle de comportement dynamique du réservoir, aux profils de production, dont les parts de variance dues à chaque variable d'entrée quantifient la sensibilité à ces paramètres.

Les réponses envisagées se montrent particulièrement sensibles aux paramètres qui définissent le modèle en faciès, les zones susceptibles de subir la diagenèse, mais surtout la paléogéographie. Quelle que soit l'étude, une modification de la carte de paléobathymétrie de référence provoque les changements les plus conséquents : en moyenne sur les six paramètres considérés, supérieurs au tiers de la dispersion totale pour les trois sorties en question. Une telle importance s'explique du fait que la carte de paléobathymétrie conditionne l'ensemble de la bathymétrie sur le modèle, et indirectement la répartition spatiale des hétérogénéités : elle règle l'image géologique du réservoir (Fig. 3.4). Si le modèle en faciès ou la « diagenèse » apparaissent en retrait, leur effet combiné peut s'avérer important. La comparaison des cubes de proportions accorde ainsi un rôle primordial au premier (plus de 20 % de leur dispersion [Zerkoune, 2003]). On doit de telles variations à l'influence du modèle en faciès sur la largeur des zones de transitions entre faciès (Fig. 3.4). Quant aux modifications des zones susceptibles de subir la diagenèse, elles jouent directement sur les volumes en place et la production selon qu'elles améliorent ou non les qualités réservoirs des faciès qu'elles affectent (environ 10% de leur dispersion [Zerkoune, 2003]). Ces quatre études montrent que des incertitudes importantes sur ces paramètres, en particulier sur la morphologie de base, peuvent notablement disperser les résultats obtenus pour les grandeurs de sortie. Par contre, d'autres facteurs du scénario géologique comme le mode de construction des courbes de bathymétrie, ou le choix des variogrammes induisent des changements de second ordre par rapport aux éléments précédents (moins de 5 % de la dispersion totale [Zerkoune, 2003]).

On notera également que la variabilité introduite par les simulations géostatistiques reste négligeable comparée à l'influence du scénario géologique. Elle représente en moyenne 2% de la dispersion totale en termes d'accumulations et de productions [Poirot, 2002 ; Bergamo, 2002]. Comme nous l'avons déjà évoqué, la multi-réalisation géostatistique peut être assimilée à un bruit autour de l'image qu'impose le schéma géologique. Nous rappelons qu'à ce jour, l'estimation de la variabilité des grandeurs de sortie se limite à des réalisations multiples en propriétés lithologiques et pétrophysiques sur un seul cube de proportions : un scénario géologique unique, dont les paramètres sont figés. Les sources majeures de dispersion, comprises dans la définition hautement incertaine du scénario sédimentaire, semblent par conséquent négligées. Pour pallier cette insuffisance, il importe d'appréhender la variabilité du schéma géologique au niveau des facteurs qui le composent. C'est pourquoi nous proposons d'adopter une approche multi-réalisation,

¹ Selon les études, le remplissage de la grille en propriétés statiques s'est fait avec une ou plusieurs simulations géostatistiques. Les premières études effectuent de la multi-réalisation pour tester l'effet de la variabilité liée à la génération d'images équiprobables. Les suivantes ne font plus, pour chaque modèle Neptune créé, qu'une seule série de simulations (une en faciès, une en porosité, etc.), la variabilité liée aux simulations étant négligeable par rapport à celle des paramètres du scénario géologique : plus de 95% de la part de variance sur les résultats est lié au choix de scénario [Poirot, 2002 ; Bergamo, 2002].

mais cette fois sur les paramètres du scénario géologique. On privilégiera bien entendu les plus incertains et les plus sensibles : le modèle de faciès, la diagenèse, et la paléogéographie. On reviendra en outre sur les corrélations (section 4.2), et le poids des faciès (section 6.1). Bien qu'ils n'aient fait l'objet d'aucune analyse de sensibilité (facteurs discrets, temps de calculs trop importants), ces paramètres gardent a priori une influence essentielle. Le premier contrôle en partie l'architecture interne du modèle ; le second, la prédominance des faciès les uns sur les autres à l'échelle du réservoir.

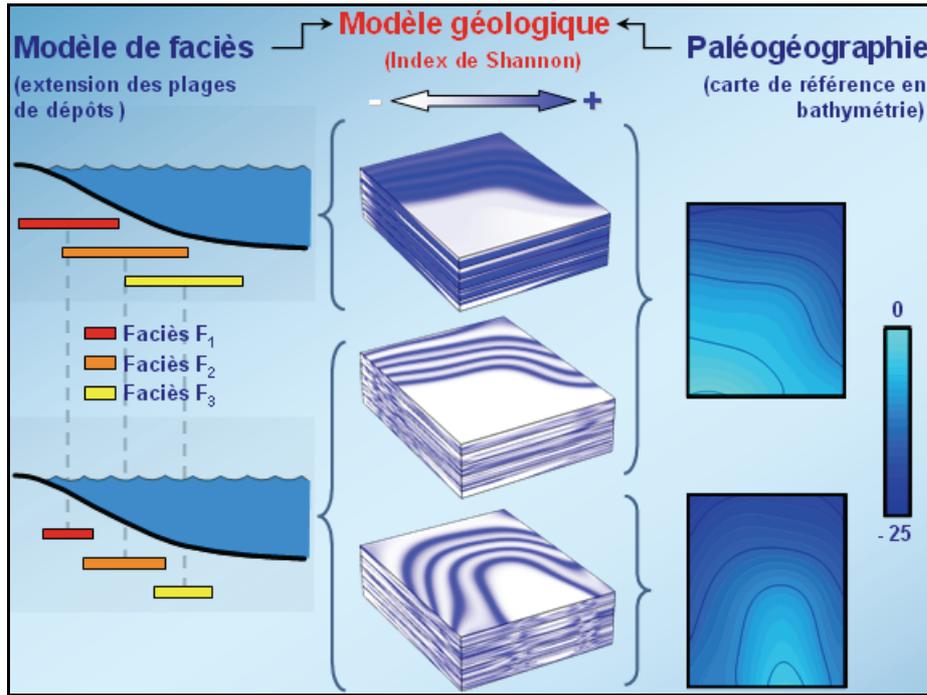


FIG. 3.4 : Effet des principaux paramètres du scénario géologique sur le cube de proportions. Le passage d'un modèle de faciès à l'autre élargie plus ou moins les zones de variabilités des proportions. La modification des cartes bouleverse le profil des lignes de niveaux.

3.4 Neptune, la multi-réalisation des scénarios géologiques

On l'a vu, il ne s'agit plus de traiter les incertitudes en aval des procédés de modélisation par de la multi-réalisation géostatistique, mais plutôt de les introduire au sein même de la démarche (Fig.3.3). L'idée consiste à balayer l'ensemble des scénarios géologiques alternatifs et par là même, de traduire l'incertitude qui les affecte, sur les accumulations et productions. La méthodologie présentée ici offre la possibilité d'inclure directement cette part de variabilité aux composantes du scénario. On associe la méthodologie Neptune à un processus de réalisations multiples sur les paramètres géologiques d'entrée. Il est alors possible, en faisant varier ces facteurs, de parcourir toutes les configurations géologiques

envisageables, et cohérentes avec d'une part les données disponibles, et avec la connaissance du sédimentologiste de l'autre.

Une démarche probabiliste de type Monte-Carlo permet de régir la combinaison des différents éléments. On recourt notamment aux méthodes de Monte Carlo avec échantillonnage aléatoire. Elles représentent les procédés de base pour décrire l'ensemble des combinaisons envisageables. Des techniques alternatives ont été proposées pour balayer plus efficacement l'espace des possibles, parmi lesquelles les procédés de simulation pseudo-probabilistes, comme l'échantillonnage stratifié ou par hypercube latin [McKay, 1995], les méthodes de Quasi-Monte Carlo [Niederreiter, 1992], ou encore les méthodes de Quasi-Monte Carlo randomisé [Owen, 1998]. Cependant, le choix du procédé d'échantillonnage importe peu, seule nous intéresse ici la façon probabiliste de décrire les différents scénarios. Nous supposons toute l'information concernant les paramètres de nature aléatoire et nous propageons celle-ci à travers le modèle considéré en utilisant la technique de Monte-Carlo. Les grandeurs incertaines du schéma géologique : modèle de faciès, cartes de bathymétrie, poids des faciès, etc., sont alors représentées par des variables aléatoires (section 2.1), que caractérisent leurs distributions de probabilité. Des tirages aléatoires, qui respectent ces lois, fournissent les combinaisons de paramètres étudiés, autorisant ainsi leur perturbation simultanée. Toutes ces configurations définissent une distribution de scénarios géologiques, eux-mêmes à l'origine d'autant de cubes de proportions alternatifs (Fig. 3.5). Il est alors possible, avec un nombre suffisant de simulations, de quantifier l'impact de l'incertitude affectant le scénario géologique au niveau des volumes en place : densité de probabilité, intervalles de confiance, etc., ou encore, sur le plan des profils de production en balayant les différentes alternatives réalisables.

Cette approche repose en premier lieu sur la définition des lois de probabilités régissant les composantes du scénario géologique. Toutefois, il faut souligner que la difficulté relève justement de leur détermination. C'est effectivement à travers ces distributions que l'expert traduit son incertitude sur le scénario, et la quantifie en terme de degré de variabilité sur les paramètres du schéma géologique. Les chapitres à venir proposent plusieurs méthodes pour traduire en termes de distributions, certaines situations d'incertitude sur le modèle géologique auxquelles peut être confronté le sédimentologiste. Classiquement, on dispose d'hypothèses discrètes sur un paramètre du scénario, qu'il s'agit de réconcilier quand c'est possible afin de balayer l'ensemble des modèles intermédiaires. Par la suite, on revient ainsi sur la quantification des sources d'incertitudes selon le niveau où elles interviennent dans le processus Neptune :

- en 1D, le long des puits (chapitre 4) ;
- en 2D, à travers les surfaces isochrones (chapitre 5) ;
- en 3D, dans le réservoir (chapitre 6).

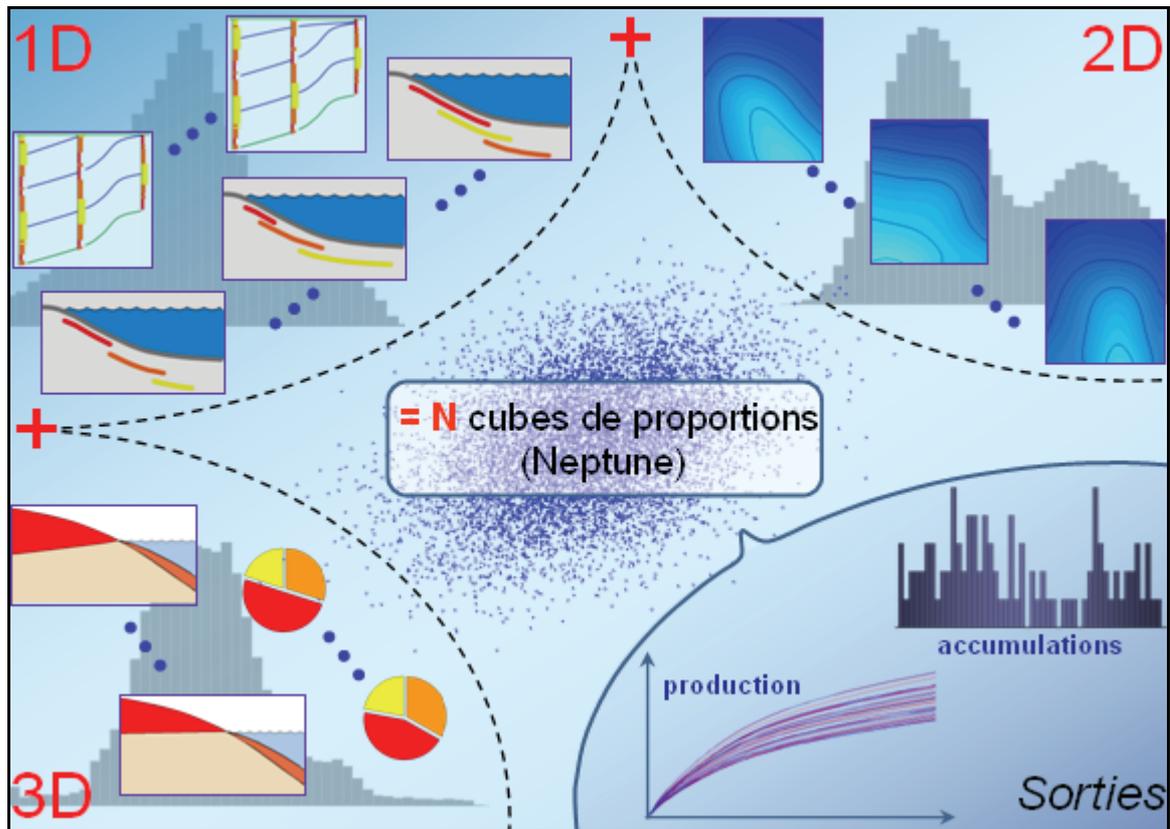


FIG. 3.5 : Principe des réalisations multiples en terme de scénario géologique. La description probabiliste des paramètres d'entrée à une (1D), deux (2D) et trois (3D) dimensions génère un ensemble de cubes de proportions (chacun d'eux correspondant à un point) compatibles avec l'incertitude géologique dont il est possible d'extraire accumulations et profils de production potentiels (Sorties).

Chapitre 4

Incertitude aux puits (1D)

En sortie de processus, sont créées le long des puits, les courbes de bathymétrie, d'accommodation et de subsidence relative dans un espace « pseudo-temps », comparable à l'espace de Wheeler. Leur construction requiert l'introduction d'informations géologiques : (1) le modèle de faciès, qui relie la bathymétrie et la nature des sédiments, et (2) les corrélations entre puits, qui autorisent le passage à l'espace « pseudo-temps ». Il s'agit de deux paramètres essentiels du scénario géologique, qui demeurent incertains, car liés à l'interprétation des données par le sédimentologiste. Ce chapitre cherche à intégrer ces sources de variabilité affectant ce scénario. Nous procédons selon une démarche multi-réalisation sur ces paramètres. L'objectif consiste à en décrire l'ensemble des états possibles, de façon à représenter l'espace réel d'incertitude qu'induisent ces éléments.

4.1 La bathymétrie aux puits

4.1.1 La relation bathymétrie/faciès

Le calcul des courbes de bathymétrie le long des puits constitue une étape importante de la méthodologie Neptune : elle conditionne la qualité et la cohérence du modèle final. Une telle opération repose sur l'hypothèse qu'un lien existe entre la bathymétrie de dépôt et la nature des faciès (section 3.1). Pour établir cette correspondance, le sédimentologiste s'appuie sur sa description des carottes disponibles aux puits. Elle lui fournit des indications sur l'environnement, la bathymétrie et l'énergie de dépôt grâce notamment à l'examen et l'interprétation [Boulvain, 2005] :

- des structures (laminations, stratification entrecroisée, etc.), et textures sédimentaires (granulométrie, composition, surfaces d'émersion, etc.) [Dunham, 1962 ; Arnaud-Vanneau, 2004] ;
- des caractères paléontologiques (contenu fossilifère, associations en faciès, microfaciès, etc.) [Wilson, 1975] ;

- des signatures se rapportant à la diagenèse précoce, et révélatrices d'éventuelles émersion [Carrio-Schaffhauser, 2004].

Par exemple, le type de communauté biologique renseigne sur la bathymétrie. La présence ou l'absence d'organismes photosynthétiques permet d'estimer la limite de la zone photique, que l'on situe la plupart du temps, à quelques dizaines de mètres de profondeur. La préservation de formes robustes ou délicates fournit éventuellement des indications sur la limite d'action des vagues, généralement à une dizaine de mètres sous la surface... Pourtant, une telle correspondance se révèle loin d'être évidente à établir. Même en domaine côtier, où certains auteurs proposent une série de « microfaciès standards » localisés dans un modèle général de plate-forme carbonatée [Wilson, 1975 ; Arnaud-Vanneau, 2004], il est en pratique quasiment impossible de définir une valeur précise de la paléo-bathymétrie de dépôt pour chaque faciès. Dans la majorité des cas, la description des carottes permet seulement d'obtenir des intervalles de profondeur d'eau : un minimum et un maximum de bathymétrie entre lesquels se dépose chaque type de sédiment.

Pourtant, ils représentent une information suffisante pour construire un cube de proportions Neptune. L'essentiel dans un modèle de faciès réside en effet dans leur ordre¹ de dépôt, le contexte – superficiel ou profond – de sédimentation, et la tranche d'eau qu'ils peuvent occuper : largeur de la plage de dépôt (Fig. 4.1–a). Pour traduire la relation entre bathymétrie et faciès, la méthodologie Neptune interprète ces intervalles en termes de lois de répartition probabilistes : loi uniforme, gaussienne, ou triangulaire, etc. (Fig. 4.1–b). Connaissant la distribution de bathymétrie en fonction du faciès, c'est-à-dire la probabilité $IP(B \leq b | F=f)$, il est possible de construire le long de chaque puits, la courbe de bathymétrie la « plus probable » par interpolation des données (Fig. 4.1–c). Néanmoins,

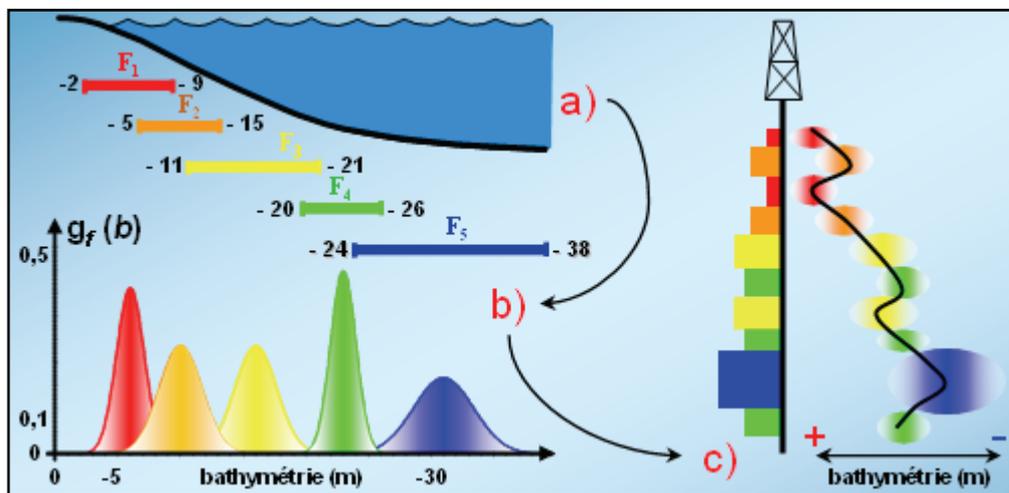


FIG. 4.1 : La relation faciès/bathymétrie. L'interprétation du modèle de faciès (a) en termes de lois probabilistes de répartition de la bathymétrie en fonction de chaque faciès – ici gaussiennes – (b) permet de construire les courbes de bathymétrie au puits (c).

¹ Par ordre, on entend selon la profondeur.

une telle qualité de la bathymétrie est valable uniquement pour le modèle de faciès considéré ; or, son choix, on l'a vu, reste toujours incertain : le sédimentologiste peut hésiter entre plusieurs (n_{mf}) modèles de faciès. Afin d'appréhender cette variabilité, on cherchera, dans le cadre d'un processus multi-réalisation, à construire les divers modèles de faciès compatibles et cohérents avec l'environnement sédimentaire défini à l'aide des données.

4.1.2 La multi-réalisation du modèle de faciès

La définition de l'ordre de dépôt des faciès représente l'étape préliminaire à l'établissement d'un modèle de faciès. Cependant, malgré son influence sur le résultat final, elle reste assez peu entachée d'incertitude. En général, le sédimentologiste, s'il dispose d'un nombre suffisant de carottes et de logs, décrit de manière précise l'agencement des faciès les uns par rapport aux autres. Toutefois, dans l'hypothèse d'une incertitude sur cette organisation, il est possible de procéder à de la multi-réalisation. Celle-ci permet de tester les différents (n_o) modèles d'ordonnement de faciès possibles. Elle se fait de manière discrète en attribuant une probabilité a priori à chacun d'eux. À chaque réalisation du processus, on sélectionne le paramètre « ordre des faciès » parmi tous les modèles candidats possibles selon leur probabilité respective. En fait, l'incertitude sur le modèle de faciès porte plutôt sur (a) l'enchaînement des faciès, et (b) l'extension des plages de bathymétrie. Elle se traduit alors sur le choix des distributions utilisées dans la méthodologie Neptune en termes, (a) de profondeurs moyennes des dépôts, et (b) d'écart types des lois de probabilité. L'approche multi-réalisation permet d'intégrer cette variabilité en simulant aléatoirement les paramètres : moyennes et écarts-types, des lois de probabilité de chaque faciès. Pour illustrer la méthode, on considère l'exemple de trois types de sédiments F_A , F_B et F_C (du plus superficiel au plus profond), dont il s'agit de traduire l'incertitude sur le modèle de faciès.

Incertitude sur le contexte de dépôt

Si l'ordonnement des n_f faciès est connu, il est souvent plus ardu de déterminer l'ordre de grandeur de leur profondeur moyenne de dépôt. Certains faciès se révèlent de bons indicateurs de la bathymétrie, d'autres peuvent être plus difficiles à interpréter à l'instar des faciès présentant des brachiopodes. En effet, c'est une espèce qui vit dans des profondeurs très variables, allant de la zone littorale à plusieurs milliers de mètres, en zone abyssale. Pourtant, le modèle finale reste relativement sensible aux hypothèses choisies sur les moyennes : des interactions existent entre le contexte de dépôt et le degré d'émersion de certaines régions du modèle [Zerkoune, 2003]. Moins il est profond, plus celles-ci sont susceptibles d'être émergées, donc de connaître la diagenèse précoce. Aussi, le sédimentologiste peut-il traduire son ignorance en fixant deux contextes de sédimentation qu'il estime être « extrêmes » :

- d'une part, un cortège superficiel, qui géométriquement correspond à un point S de coordonnées $[m_1(S), m_2(S), \dots, m_{n_f}(S)]^T$; $m_f(S)$ étant la moyenne de dépôt pour le faciès f ;

– et profond d’autre part, représenté géométriquement par le point P de coordonnées $[m_1(P), m_2(P), \dots, m_{nf}(P)]^T$; le terme $m_f(P)$ se rapportant à la moyenne de dépôt pour le faciès f .

Dès lors, pour régir l’incertitude sur la moyenne, il faut balayer l’ensemble des cortèges de dépôts intermédiaires. Une méthode simple consiste à décrire le segment SP , de longueur L . Il suffit alors de simuler un paramètre λ selon une loi uniforme dont le support est l’intervalle $]0 ; L[$. Le contexte intermédiaire correspondant est déterminé par la multiplication par ce facteur de toutes les moyennes, et de tous les écarts types du cortège superficiel S . Cette approche permet de faire varier les profondeurs moyennes entre les deux contextes en préservant la cohérence du modèle : l’ordre des faciès, et l’écart entre les zones d’influence. La figure 4.2 illustre cette approche sur l’exemple des trois faciès considérés ici.

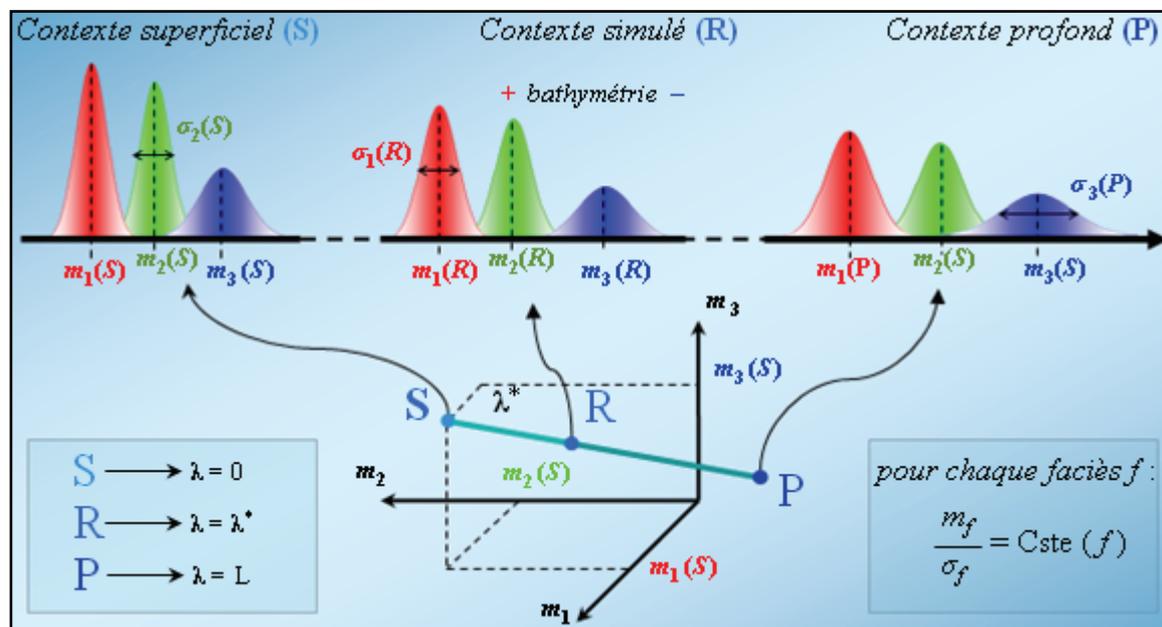


FIG. 4.2 : Intégration de l’incertitude sur le contexte de dépôt. À partir de deux contextes de dépôt, on le fait varier en simulant la bathymétrie moyenne de sédimentation pour chaque faciès en respectant leur ordonnancement et le rapport moyenne/écart-type.

Incertitude sur l’étendue des plages de dépôts

Elle se manifeste directement sur l’écart type des lois de répartition de chaque faciès. Le sédimentologue répercute cette variabilité dans un terme d’incertitude, qui se rapporte au manque d’information. Un moyen commode d’appréhender cette grandeur consiste à spécifier pour chaque faciès, les deux plages de dépôts que l’on juge « extrêmes » : un contexte extensif d’un côté, et restrictif de l’autre (Fig. 4.3). Selon le principe du maximum d’entropie [Deutsch, 1992], chaque écart type est supposée suivre une loi uniforme, dont les bornes se déduisent du terme d’incertitude. Il suffit ensuite de faire

respirer les plages de dépôts en ayant soin de respecter leur chevauchement mutuel. L'approche suivante simule des modèles de faciès cohérents avec ces contraintes :

1. pour chaque faciès f , calcul des bornes min_f et max_f de la loi de probabilité de l'écart type (σ_f) traduisant l'extension de sa plage de dépôt ;
2. pour f variant de 1 à n_f – nombre de faciès – simulation de σ_f selon une loi uniforme de support] min_f ; max_f [;
3. pour chaque paire de faciès contiguës f et f' ; f étant plus profond que f' , de bathymétrie moyenne de dépôt m_f et $m_{f'}$, test sur la cohérence du modèle simulé :
 - (a) chevauchement des plages de bathymétrie :
 si $m_f + 3.\sigma_f \geq m_{f'} - 3.\sigma_{f'}$ alors $e_f = 1$ sinon $e_f = 0$
 - (b) ordre des faciès,
 si $m_f + 3.\sigma_f \geq m_{f'}$ et $m_{f'} - 3.\sigma_{f'} \leq m_f$ alors $e_f = 1$ sinon $e_f = 0$
4. test sur la validité du modèle :
 si $e_1 \times e_2 \times \dots \times e_{n_f} = 1$, on accepte le modèle ; on le refuse sinon.

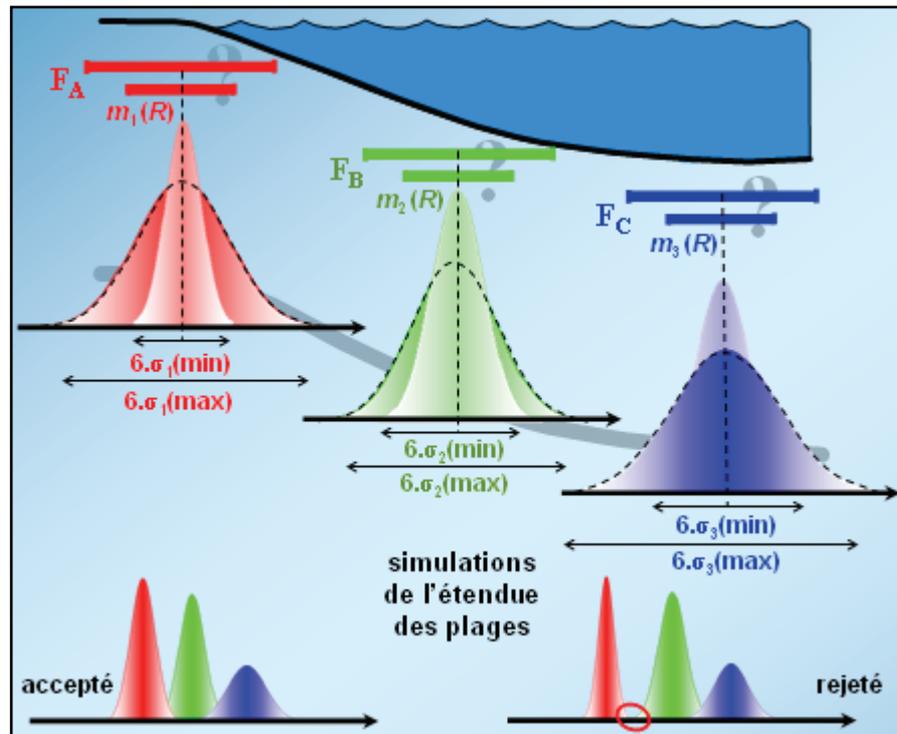


FIG. 4.3 : Intégration de l'incertitude sur l'extension des plages de dépôts. À partir de deux scénarios en termes de largeur, on fait varier de façon probabiliste la plage d'influence des différents faciès de manière à respecter leur chevauchement deux à deux.

On remarquera que l'incertitude sur l'extension des plages de dépôt peut se faire à deux niveaux : d'abord, pour établir la relation bathymétrie/faciès utilisée pour estimer les courbes de bathymétrie le long de chaque puits, puis inverser le modèle 3D de bathymétrie en cube de proportions. Sinon, il paraît néanmoins plus approprié de construire les courbes à partir d'un cas médian, et de recourir au processus multi-réalisation sur l'étendue des plages uniquement pour l'inversion du modèle de bathymétrie en cube de proportions. Des études de sensibilité montrent en effet que la dimension des plages de faciès a peu d'influence sur le calcul des courbes de bathymétrie si la condition de chevauchement est respectée.

À chaque simulation du processus multi-réalisation, on simule un contexte de dépôt (moyennes), puis des extensions de plages de dépôt (écarts types). Ces paramètres, cohérents avec l'information disponible, sont à l'origine d'une relation bathymétrie/faciès. Celle-ci est utilisée par la suite pour construire les courbes de bathymétrie le long de chaque puits, mais aussi pour évaluer les proportions de faciès partout dans le réservoir. Auparavant, on aura estimé les signaux d'accommodation le long des puits. Un tel calcul nécessite le recalage des puits en temps à l'aide des corrélations.

4.2 Les corrélations aux puits

Le tracé des corrélations – passage obligé en stratigraphie – établit les correspondances entre les points de chaque puits qui ont le même âge géologique. Il s'agit d'une étape majeure puisqu'elle contrôle en partie le découpage stratigraphique du modèle final selon des surfaces isochrones. Pour élaborer sa planche de corrélations, le sédimentologiste identifie des repères stratigraphiques contemporains, que ce soit à travers les données biostratigraphiques, lithologiques, sismiques ou diagraphiques. Leur interprétation via les concepts de la stratigraphie séquentielle met en évidence l'enchaînement de séquences élémentaires de dépôt¹ au cours du temps. Chacune d'elles, dont les limites ont une valeur chronostratigraphique, peut d'ailleurs être enregistrée sous la forme d'une surface (hiatus², érosion), et/ou d'un volume de sédiments [Arnaud, 2004b]. L'identification de ces corrélations, qui procèdent d'interprétations porte donc, par essence, une part d'incertitude.

Toutefois, la méthodologie Neptune offre la possibilité de tester toutes les corrélations envisageables, même en cas d'ignorance complète sur leur agencement. Les puits sont rassemblés dans un espace paramétrique dans lequel chaque isochrone est horizontale. La coordonnée verticale est un « temps » discrétisé en plusieurs intervalles, qui correspondent au nombre n_c de couches du modèle final. Chaque intervalle est délimité par deux corrélations au niveau des puits, et représente une strate dans la grille finale. À chaque fois qu'une corrélation est définie, l'échelle verticale des puits est modifiée afin que les points corrélés se retrouvent sur la même ligne horizontale (isochrone). Les signaux enregistrés aux puits (bathymétrie, accommodation, etc.) sont également recalculés.

¹ On parle plus souvent de paraséquence.

² **Hiatus** : Ce terme renvoie à une lacune de sédimentation.

Maintenant, dans l'hypothèse d'une d'incertitude totale sur la planche de corrélations, il suffit de simuler, sur chaque puits w :

- (1) la côte $T_w(1)$ des points à corrélérer les plus anciens selon une loi uniforme sur l'intervalle $[0, n_c]$;
- (2) la côte $T_w(2)$ des points correspondant à la deuxième corrélation la plus ancienne selon une loi uniforme dont le support correspond à l'intervalle $[T_w(2), n_c]$;
- (…)
- (3) la côte $T_w(n)$ des points correspondant à la corrélation la plus récente selon une loi uniforme sur l'intervalle $[T_w(n-1), n_c]$.

Un tel procédé, on l'admet, produit beaucoup de « déchets ». La majorité des modèles construits n'est en effet pas cohérente en terme de géologie. Un moyen de faire un tri des résultats serait de ne conserver que les modèles dont les Résidus, qui traduisent entre autres les erreurs de corrélations, sont globalement les plus faibles. On rappelle que les Résidus représentent l'écart entre la valeur d'accommodation calculée sur les puits et sa moyenne estimée indépendamment sur chaque couche à partir des données de puits (Facteur Commun).

Fort heureusement, le sédimentologiste fournit une planche de corrélations, et apporte donc de la connaissance géologique. L'incertitude porte alors ici sur la position précise des corrélations le long des puits. Il faut donc plutôt voir Neptune comme un outil d'aide à leur tracé : les variations d'accommodation devraient être les mêmes pour des points contemporains. En effet, un événement géologique comme une montée du niveau marin par exemple, se traduit, à cet instant précis, par une signature commune à l'ensemble du réservoir : un pic positif sur la courbe d'évolution de l'accommodation, caractéristique d'une création d'espace pour sédimenter. Au moment des corrélations, il suffit de connecter les pics semblables observés sur les courbes d'accommodation. Dans ce cas, la comparaison des courbes sur tous les puits permettrait de réduire l'incertitude sur le pointé des corrélations ; par exemple en minimisant la part des Résidus.

Dans certains cas, plusieurs (n_c) schémas de corrélations sont vraisemblables et cohérents avec les données (un par sédimentologiste en général). Pour appréhender cette incertitude, nous envisageons une stratégie de sélection probabiliste du paramètre « planche de corrélations » lors de la construction des scénarios géologiques à chaque simulation du processus multi-réalisation. À chaque itération, on tire une planche de corrélations parmi tous les modèles candidats possibles selon leur probabilité respective. La difficulté repose sur l'estimation de ces probabilités. Diverses méthodes existent pour affecter cette dernière à chaque modèle et sont proposés ici :

– **Schémas équiprobables,**

Ce cas traduit le fait qu'aucun modèle, indexé ici de 1 à n_c , n'est meilleur parmi les autres modèles possibles. Chaque scénario est supposé avoir la même probabilité no-

tée IP de se révéler « vrai » que les autres,

$$\forall i \in \{1, 2, \dots, n_c\}, IP(\text{planche} = i) = \frac{1}{n_c} \quad (\text{Éq. 4.1})$$

– **Définition par avis d'expert(s),**

$$\forall i \in \{1, 2, \dots, n_c\}, IP(\text{planche} = i) = \mathbf{F}(\text{experts}) \quad (\text{Éq. 4.2})$$

où F désigne une fonction de l'expert.

Il s'agit d'une méthode simple qui présente l'avantage de ne nécessiter aucune donnée. Elle révèle d'une certaine manière la préférence de l'expert pour certains modèles. Cette approche souffre néanmoins de subjectivité ; mais, l'appréciation de l'incertitude, par nature, ne peut être complètement objective. Vu les données à disposition, chaque modèle résulte au départ d'un choix de paramètres consécutif à une interprétation du phénomène réel. C'est pourquoi un traitement approprié de l'incertitude repose en grande partie sur l'expertise, seule à même de justifier les paramètres et hypothèses de modélisation.

À ce stade, on dispose pour chaque réalisation, d'une relation entre bathymétrie et faciès ainsi que d'une planche de corrélation. Ces données rendent alors possibles la construction des courbes intervenant dans le procédé Neptune au niveau des puits, et ce dans un espace « pseudo-temps », comparable à l'espace de Wheeler. Il s'agit désormais d'extrapoler ces signaux à travers les surfaces isochrones ainsi définies.

Chapitre 5

Incertitude sur les cartes (2D)

Cette étape correspond à l'extrapolation des signaux connus uniquement aux puits ($b(x,y,t)$, $s(x,y,t)$, $e(x,y,t)$, etc.), à travers les surfaces isochrones déterminées grâce aux corrélations. Si l'extrapolation des composantes de la subsidence relative (Facteur Commun et Résidus), ou des épaisseurs s'accommodent de l'hypothèse de stationnarité (la moyenne reste constante partout dans l'espace), le calcul de la bathymétrie reste quant à lui plus délicat. Il nécessite de fournir une carte de paléo-bathymétrie de référence. Classiquement, le sédimentologue l'estime à partir des données sismiques et de sa connaissance du bassin ; elle n'en demeure pas moins très incertaine. Ce chapitre propose plusieurs méthodes pour simuler ce paramètre inconnu du scénario géologique, afin de l'inclure au processus de multi-réalisation. L'objectif, on le rappelle, consiste à décrire l'ensemble des états possibles, pour ainsi représenter l'espace réel d'incertitude qu'induisent de tels éléments.

5.1 Les signaux sur les surfaces isochrones

Le calcul de la bathymétrie sur l'ensemble des surfaces isochrones contrôle l'image géologique du modèle : elle conditionne indirectement la répartition des divers corps sédimentaires. C'est en effet sur cette propriété qu'est introduite la non stationnarité qui affecte les caractéristiques finales du réservoir, notamment les proportions de faciès. Son estimation réclame l'extrapolation préalable, à travers les surfaces isochrones, des autres signaux acquis aux puits : épaisseur des couches et subsidence relative (somme du Facteur Commun et des Résidus). Elle suppose en plus l'introduction d'une carte de bathymétrie de référence.

5.1.1 Extrapolation des épaisseurs

Après les corrélations, chaque couche représente un intervalle de temps durant lequel l'épaisseur des sédiments varie d'un puits à l'autre. Cette épaisseur doit être extrapolée sur chacune des surfaces isochrones. Généralement, pour déterminer l'épaisseur des cou-

ches dans la totalité du réservoir, on procède selon les techniques géostatistiques de krigage ou de simulation. Elles sont en outre contraintes à l'épaisseur totale du réservoir. Ici, l'incertitude sur le modèle porte sur le choix du variogramme utilisé (chapitre 3). En pratique, il est préférable de travailler avec des variogrammes de type gaussien afin d'obtenir des variations d'épaisseurs relativement lisses dans l'espace. En ce qui concerne les autres paramètres (portée, orientation, etc.), certains auteurs [Norris *et al.*, 1993 ; Corre *et al.*, 2000] préconisent leur simulation préalable selon des distributions fournies par l'utilisateur. Cependant, dans notre processus multi-réalisation, nous ne tiendrons pas compte de cette incertitude, puisqu'elle est de second ordre par rapport aux autres.

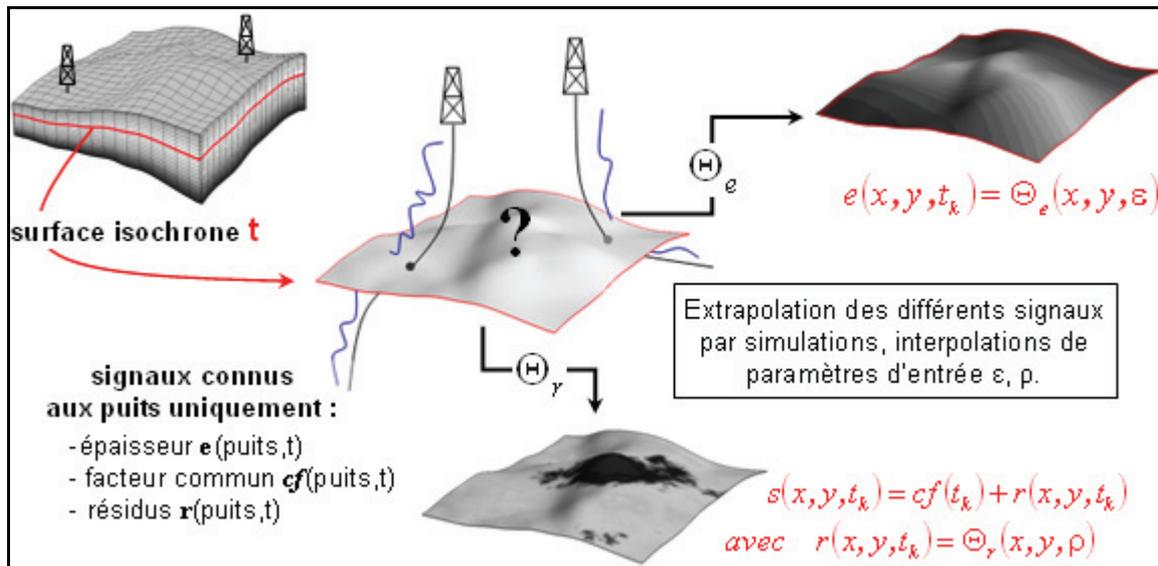


FIG. 5.1 : Calcul des signaux connus aux puits sur l'ensemble des surfaces isochrones. Ceux-ci sont extrapolés ou interpolés en général au moyen de techniques géostatistiques dont les paramètres d'entrée ε et ρ (variogrammes par exemple) ont une faible influence sur les résultats en sortie.

5.1.2 Calcul du potentiel d'accommodation

Le potentiel d'accommodation $s(x, y, t)$, représente un phénomène global à l'échelle du réservoir (chapitre 3). Il est donc supposé constant le long d'une surface isochrone, et peut être directement extrapolé couche par couche. Toutefois, ce n'est que très rarement le cas en raison des bruits parasites. Le signal $s(x, y, t)$ est décomposée en deux parties, le Facteur Commun $cf(t)$, et les Résidus $r(x, y, t)$, qui sont extrapolés séparément :

$$s(x, y, t) = cf(t) + r(x, y, t) \quad (5.1)$$

Le Facteur Commun correspond à ce que serait le potentiel d'accommodation s'il n'y avait aucun bruit parasite (erreurs de mesure, de calcul ou phénomènes géologiques locaux). Il représente la moyenne des données de puits estimée indépendamment sur chaque

couche du modèle. Il est constant sur chacune d'elle, et son extrapolation s'effectue directement (Fig. 5.1). Les Résidus correspondent à l'écart entre la valeur du potentiel d'accommodation calculée sur les puits et le Facteur Commun. Ils représentent l'ensemble des bruits dus à la géologie, mais surtout dus aux erreurs de mesure et de calcul. À l'instar des épaisseurs, leur extrapolation s'appuie sur les techniques géostatistiques ; la méthode la plus fréquemment utilisée étant la Simulation Gaussienne Séquentielle (SGS). Bien entendu, on retrouve l'incertitude liée au choix du modèle de continuité géologique, qui de la même manière ne sera pas prise en compte dans notre processus multi-réalisation. On prêtera plutôt attention à la portée du variogramme utilisé : une portée relativement grande sera utilisée si on suppose que les Résidus sont dus à des mouvements tectoniques, une très faible portée sera utilisée s'ils sont dus à des erreurs de mesure.

5.1.3 Définition de la carte de bathymétrie de référence

Une fois le potentiel d'accommodation et l'épaisseur des horizons déterminés dans la totalité du modèle, il est possible de remonter à la bathymétrie en inversant l'équation 3.2. Cependant, cette relation apporte simplement des informations sur son évolution entre les limites inférieures et supérieures des différentes couches qui composent le réservoir. Le calcul de la bathymétrie nécessite donc de fournir une carte de paléo-bathymétrie de référence sur au moins une surface isochrone du réservoir, d'ordinaire la mieux connue¹. Cette carte, qui représente la bathymétrie au temps t_0 dans l'équation 3.2, reproduit la morphologie de la plate-forme. Il s'agit donc d'extrapoler les données observées aux puits sur l'ensemble de la surface correspondant à l'isochrone t_0 .

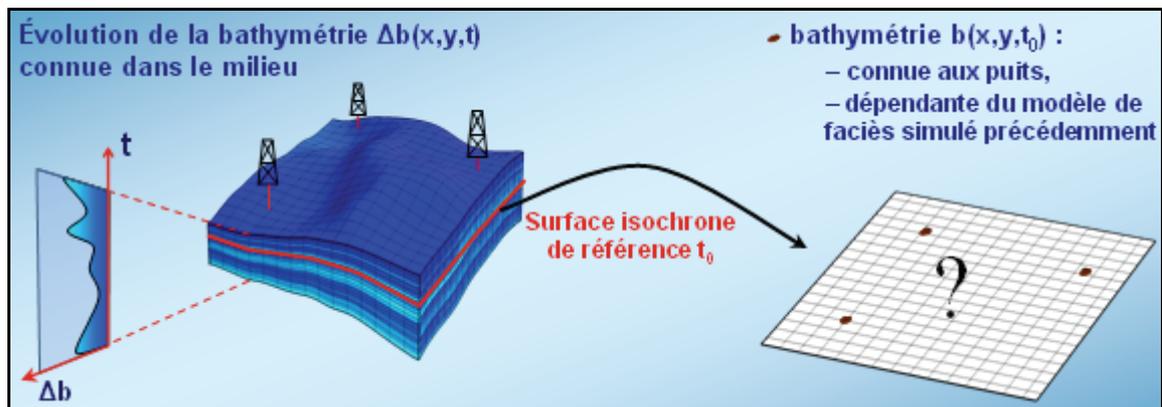


FIG. 5.2 : Calcul de la bathymétrie $b(x,y,t)$ le long des surfaces isochrones. Seule l'évolution $\Delta b(x,y,t)$ de la bathymétrie est connue, le calcul de la bathymétrie $b(x,y,t)$ en elle-même nécessite de fournir une carte de paléo-bathymétrie de référence à l'isochrone t_0 . La construction de cette carte doit respecter les valeurs de bathymétrie qui sont connues aux puits et qui dépendent du modèle de faciès calculé précédemment.

¹ En général, au toit ou au mur du réservoir.

La carte est dessinée grâce aux données de puits et au modèle sédimentaire. Dans certains cas, elle peut être élaborée à l'aide des attributs sismiques par cosimulation ou cokrigage¹ [Pernarcic, 2002]. Cependant, le nombre de données contraignantes (puits) reste limité, ou sont à une résolution trop faible (sismique) pour contraindre de manière satisfaisante le dessin de la carte. En effet, quelle que soit la méthode employée : krigage, simulation géostatistique, interpolation par DSI [Leflon *et al.*, 2004], etc., il faut considérer la non stationnarité de la bathymétrie caractéristique des rampes marines. Car, plus que la valeur précise de la bathymétrie en chaque point, c'est la forme générale de la topographie au temps t_0 qui prévaut, et affecte le plus les résultats finaux (section 3.3). Par conséquent, la paléogéographie de référence repose en grande partie sur l'interprétation des données aux puits (1D). L'analyse de leur évolution permet de définir l'environnement sédimentaire avec ses milieux et ses mécanismes de dépôt ; et par là même rend leur cartographie (2D) possible. En pratique, cette connaissance se traduit directement sur le tracé des courbes isobathes définies à partir du modèle sédimentaire. Elles sont alors utilisées comme données dures pour le krigage, ou encore comme contraintes pour l'interpolation DSI. Néanmoins, leur dessin revêt un caractère hautement incertain. Un large ensemble de paléogéographies – donc de courbes isobathes – est en réalité compatible avec la description, même précise, de l'environnement sédimentaire [Massonnat, 1999]. On retrouve finalement, à l'échelle de la carte, l'incertitude de modèle abordée au premier chapitre. C'est pourquoi, il est nécessaire de parcourir l'ensemble des cartes possibles durant la multi-réalisation des paramètres du scénario géologique.

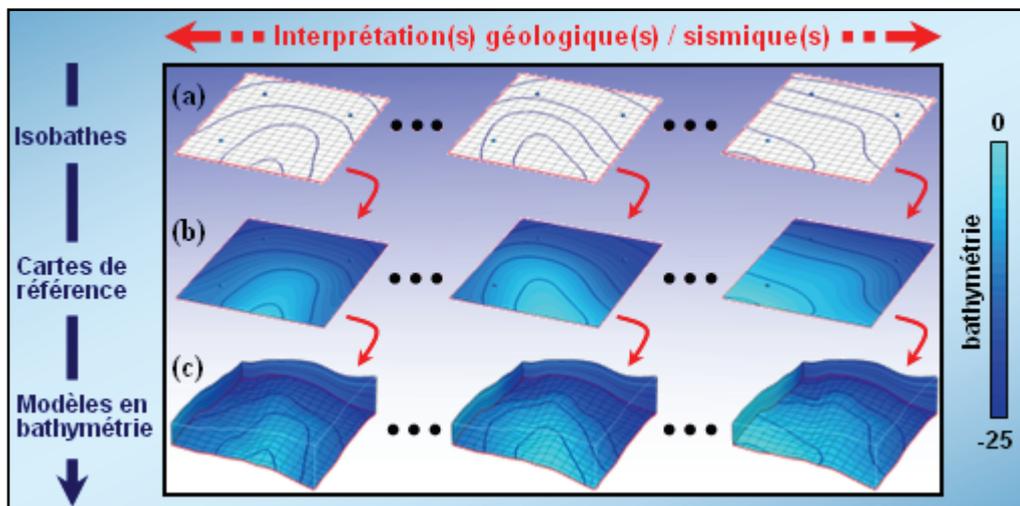


FIG. 5.3 : Incertitude sur la paléogéographie. La bathymétrie de référence, connue seulement aux puits est extrapolée en intégrant les tendances incertaines tirées de l'interprétation géologique ou sismique. (a) Les différentes dérives sont introduites sous forme d'isobathes, (b) qui permettent de construire par extrapolation des données les cartes de référence, et (c) d'en déduire les modèles 3D en bathymétrie correspondants selon l'équation 3.2.

¹ Ce sont des méthodes géostatistiques qui, outre les données relatives à la propriété à extrapoler, utilisent l'information qu'apporte une variable secondaire.

5.2 La multi-réalisation des cartes de bathymétrie

À ce stade, si on dresse un bilan de l'information disponible, plusieurs (n_b) hypothèses touchant aux cartes de référence – allure générale de la bathymétrie – restent encore envisageables. Toutes ont été élaborées à partir des observations aux puits et/ou de la sismique. Dès lors, il convient d'inclure la simulation des cartes de référence à la suite du processus multi-réalisation. Plusieurs méthodes sont ici proposées afin d'appréhender l'incertitude sur le modèle de paléogéographie de référence. Au préalable, il faut cependant contraindre les cartes aux observations faites aux puits après la simulation du modèle de faciès.

5.2.1 Une nécessaire adaptation des cartes de base

Chaque simulation du processus multi-réalisation modifie d'abord le modèle de faciès (chapitre 4). Elle affecte en particulier la bathymétrie attribuée aux faciès observés sur les puits au temps t_0 . Par conséquent, la valeur des isobathes, qui traduit les hypothèses sur la topographie générale, est décalée par rapport aux données de puits (Fig. 5.4). Dans la perspective d'une démarche multi-réalisation, il s'agit donc d'ajuster la bathymétrie inter-

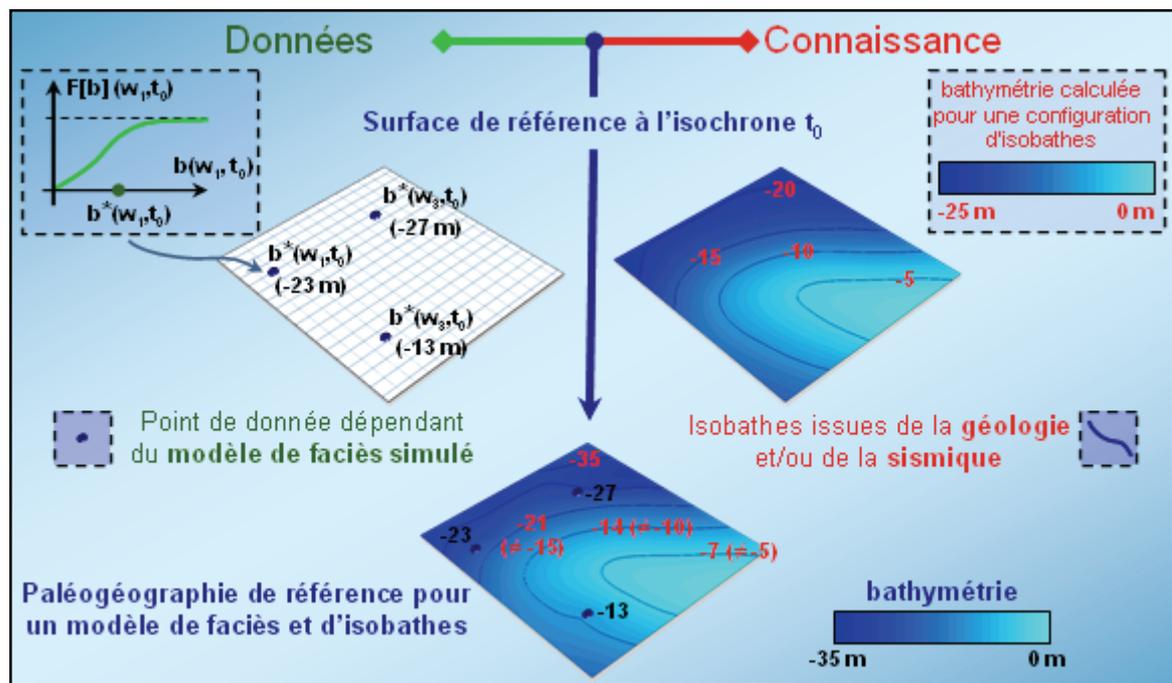


FIG. 5.4 : Contrainte des cartes de référence aux données de puits. À droite, les courbes isobathes sont créées une seule fois pour décrire une éventuelle topographie. À gauche, la simulation du modèle de faciès modifie les valeurs en bathymétrie au niveau des puits sur la surface isochrone t_0 . Celles-ci servent alors de données contraignantes à l'extrapolation de la bathymétrie, décalant ainsi les courbes isobathes (carte du bas).

polée aux puits sans avoir à recréer manuellement une nouvelle carte de même allure générale à chaque réalisation. En premier lieu, pour chaque carte de base, les contraintes de forme – isobathes – sont stockées sur des courbes annexes ; et ne sont finalement définies qu’une seule fois pour l’ensemble des réalisations. L’ajustement de chacune des cartes aux observations réalisées sur les puits peut se faire en deux étapes :

- (1) la valeur affectée à chaque isobathe est déterminée par le biais d’une fonction de régression ; cette dernière traduit la relation à l’emplacement des puits, entre les valeurs observées sur les cartes et celles simulées aux puits à partir du modèle de faciès ; Leflon (2005) recourt à un tel procédé pour corriger les erreurs d’échelle qui peuvent s’immiscer pendant la construction des cartes : par exemple, si la bathymétrie est comprise entre 5 et 30 mètres sur la carte de bathymétrie alors qu’en réalité, sur les puits elle doit être entre 0 et 60 m ;
- (2) sur la carte de base, la bathymétrie est interpolée en intégrant dans le processus de construction, krigeage ou interpolateur DSI, les modifications accomplies antérieurement au niveau des puits et des courbes isobathes (données dures ou nœuds de contrôle).

Les cartes de départ respectent les valeurs observées aux puits à chaque simulation du modèle de faciès. En outre, on conserve les n_b hypothèses touchant la forme générale de la topographie. Elles peuvent maintenant être utilisées dans la suite du processus multi-réalisation afin d’assimiler l’incertitude sur le modèle de paléogéographie.

5.2.2 Une approche à scénarios multiples

Cette méthode, déjà exposée au chapitre précédent (section 4.2), repose sur une démarche à scénarios multiples. À l’instar des corrélations, on envisage un mode de sélection probabiliste du paramètre « carte » lors de la construction des scénarios géologiques. À chaque simulation du processus multi-réalisation, on sélectionne une carte de référence au sein des n_b cartes admissibles suivant leur probabilité respective. L’estimation de la probabilité IP de chaque paléogéographie de référence (indexée de 1 à n_b) s’appuie toujours :

– soit sur l’équiprobabilité des cartes – cas d’une ignorance totale,

$$\forall i \in \{1, \dots, n_b\}, \mathbb{P}(\text{carte} = i) = \frac{1}{n_b} \quad (\text{Éq. 5.1})$$

– ou bien sur l’avis d’expert(s) – cas d’une ignorance partielle,

$$\forall i \in \{1, \dots, n_b\}, \mathbb{P}(\text{carte} = i) = \mathbf{Fonction}(\text{experts}) \quad (\text{Éq. 5.2})$$

Malgré sa simplicité, cette méthode permet de tester rapidement différentes hypothèses, et de traduire en partie l’incertitude sur le modèle géologique. Elle demeure néanmoins limitée par son caractère discret en balayant partiellement l’espace d’incertitude. La question se pose évidemment de savoir « ce qu’il se passe entre ces scénarios extrêmes » ; la relation associant paléogéographie, accumulations et productions étant tout sauf monotone.

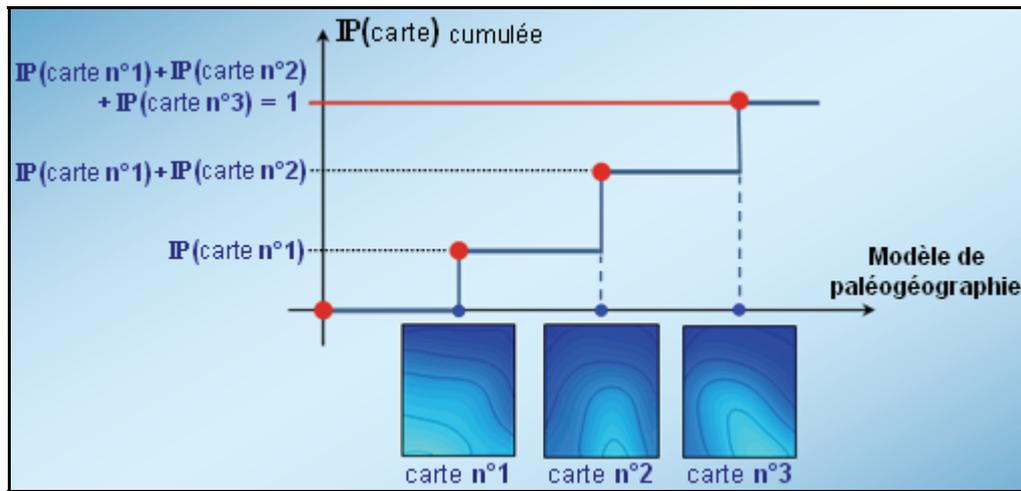


FIG. 5.5 : Une approche discrète pour intégrer l'incertitude sur les modèles de cartes. Une probabilité est attribuée à chacun des modèles envisagés. La carte utilisée dans le processus de multi-réalisation est simulée selon cette distribution discrète.

5.2.3 Une approche selon les modèles de mélange

Une idée naturelle consiste à considérer des cartes de référence construites à partir des différentes (n_b) hypothèses de paléogéographie possibles. Une telle approche peut s'apparenter aux modèles de mélange [McLachlan *et al.*, 2000]. Pour calculer les valeurs d'une variable aléatoire d'intérêt Y , ceux-ci recourent à un modèle défini comme une moyenne pondérée des divers modèles admissibles. On présente certaines techniques de combinaisons des cartes, utilisées en vue d'appréhender l'incertitude sur la forme générale de la paléogéographie. À titre d'exemple, l'application au réservoir synthétique TATA illustre les méthodes proposées (Fig. 5.6).

Il s'agit d'une plate-forme carbonatée que renseignent trois puits. L'incertitude qui touche le scénario géologique porte sur la morphologie de la rampe. La description des données aux puits permet d'envisager trois interprétations distinctes, (a) un profil de plage (modèle \mathbf{M}_A), (b) un promontoire d'axe nord-sud (\mathbf{M}_B) ou (c) d'orientation nord-ouest / sud-est (\mathbf{M}_C). On dispose de n_b modèles de paléogéographie – ici $n_b = 3$, dont les profils de bathymétrie diffèrent, traduisant ainsi l'incertitude sur le scénario géologique à ce niveau. Cette dernière se répercute sur le degré d'ouverture du promontoire et sur l'orientation de son axe. En fait, il est important de traduire l'incertitude sur les cartes à travers des paramètres continus dont la combinaison établit la topographie¹ finale. En effet, par la suite, on s'attache à faire des simulations de la paléogéographie à partir de ces cartes de base. Afin de conserver un contrôle sur la consistance géologique des modèles créés, il est nécessaire de travailler avec des configurations qui restent compatibles entre

¹ Outre les grandeurs précédemment citées, on mentionnera également le pourcentage de la pente, le degré d'ouverture d'un éventuel sillon, la largeur d'une barrière, etc.

elles. Dans ce cas, on parlera de « méta-scénario » dans le sens où il englobe les différents scénarios géologiques considérés traditionnellement, et non balayés par les méthodes conventionnelles (chapitre 2). Dans le cas contraire, on privilégiera une approche discrète appliquée sur chacun des « méta-scénarios » envisagés.

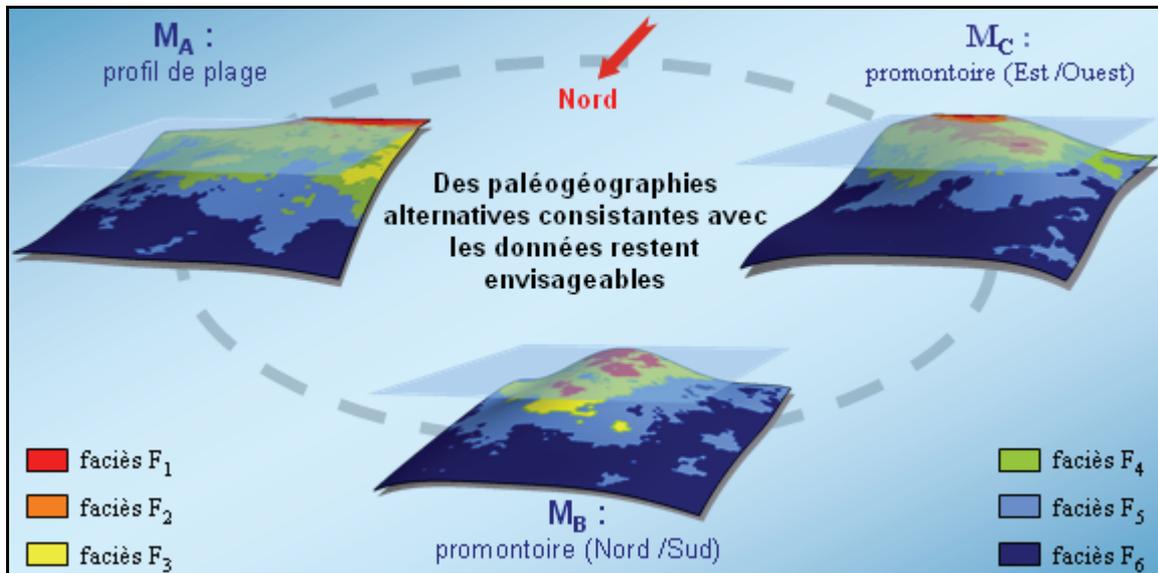


FIG. 5.6 : Les divers blocs diagrammes envisagés et compatibles avec les données. À partir d'un même jeu de données, différents paysages peuvent être envisagés. Ici, on s'intéresse à un profil de plage, ou de promontoire soit axé plein nord, soit dirigé vers le Nord-Est.

Métamorphose – *morphing* en anglais – entre couples de cartes

Le problème consiste à décrire les paléogéographies intermédiaires à partir de deux cartes de bathymétrie initiales, que l'expert considère comme « extrêmes ». L'idée qui vient naturellement repose sur le recours aux techniques de métamorphose, souvent utilisées pour les animations en infographie. Il s'agit de procédés qui transforment de façon continue et progressive un modèle – courbe, image, surface ou volume – en un autre par interpolation. Traditionnellement, le processus est décomposé en deux phases distinctes :

- (1) le couplage des sommets, il établit une correspondance entre les caractéristiques géométriques des formes de départ et d'arrivée ; la plupart des algorithmes imposent manuellement la mise en relation des différents points choisis [Alexa *et al.*, 2000] ;
- (2) l'interpolation, qui définit la déformation des formes de la première image vers celles de la seconde trajectoire afin de superposer leurs traits caractéristiques ; citons pour exemple, les méthodes de fondu, qui font progressivement disparaître et apparaître, respectivement les formes initiales et finales, les procédés par grilles ou segments d'influence, ou encore les méthodes de transformation à l'aide de la

transformée de Fourier [Gomes *et al.*, 1997] ; Lazarus (1995) et Alexa (2002) proposent chacun un état de l'art des techniques de métamorphose surfaciques et volumiques.

Ces deux problèmes suscitent d'ailleurs toujours un vif intérêt, puisque aucune définition formelle d'une solution satisfaisante n'existe à ce jour. Certaines méthodes sont a priori plus générales que d'autres, mais montrent un manque de contrôle sur la transformation. D'autres au contraire reposent sur une très grande interaction avec l'utilisateur [Lazarus, 1995].

S'agissant des cartes de bathymétrie, la difficulté relève de la création de paléographies intermédiaires qui préservent l'apparence, et les propriétés des cartes fournies en entrée. C'est pourquoi il faut porter une attention particulière à l'étape de couplage des points ; elle améliore grandement la qualité de la transformation et la cohérence des images construites [Surazhsky *et al.*, 2003]. Cependant, l'incertitude, transposée aux paramètres continus qui composent la topographie, simplifie de façon significative le couplage des sommets. Dans notre exemple, concernant l'indétermination relative à l'orientation du promontoire, il s'agit tout bonnement d'en coupler les axes (Fig. 5.3). Plus généralement, le problème sera uniquement de faire correspondre les isobathes de valeur identique d'une carte à l'autre, et de laisser « fixes » les points de données aux puits. Rien n'empêche d'ailleurs un contrôle de l'animation obtenue, afin de préserver la cohérence des modèles construits à partir des paléogéographies simulées.

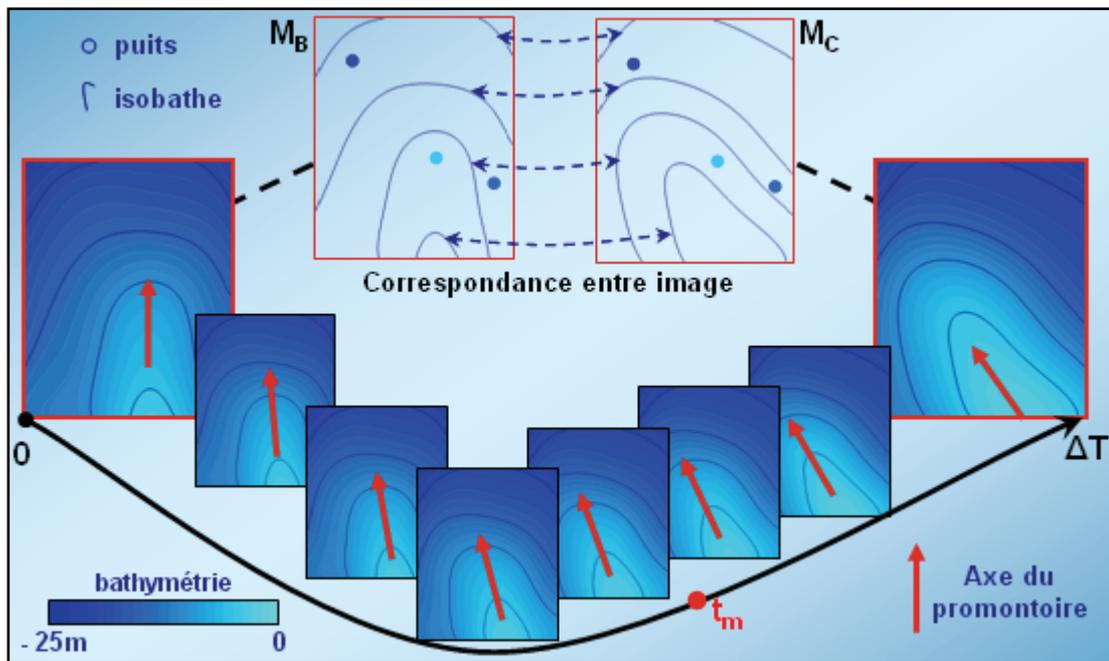


FIG. 5.7 : Construction de cartes par métamorphose entre couples de cartes. La correspondance entre cartes établie, le passage de l'une à l'autre se fait de continûment en parcourant le « temps » t_m ; ce qui permet de décrire les paléogéographies intermédiaires.

La carte utilisée dans le processus multi-réalisation correspond à l'image obtenue au bout d'un temps t_m après le début de l'animation. En effet, un seul paramètre régit l'évolution de la transformation : le temps t_m qui s'écoule entre le passage de la carte initiale (0) à la carte intermédiaire (t_m) ; la carte finale correspondant à la durée ΔT . Quant à l'échantillonnage des cartes utilisées dans le processus multi-réalisation, il dépend bien entendu de la connaissance du sédimentologiste. S'il n'a aucun a priori sur son choix, le paramètre t_m peut être simulé selon une loi uniforme de support $[0 ; \Delta T]$ pour chaque réalisation. Cette technique permet d'élargir le spectre de cartes possibles, en respectant les observations aux puits (points « fixes ») et en préservant la cohérence des modèles (Fig. 5.3). Cependant, elle travaille seulement par couples de cartes. L'idéal serait de la généraliser à plus de deux modèles, et d'extraire ainsi l'information issue de l'ensemble des cartes.

Combinaison de plusieurs cartes

Si plusieurs modèles (n_b) de paléogéographie (M_i) _{$1 \leq i \leq n_b$} sont définis, il est toujours possible de les combiner de façon linéaire en associant simplement un poids $p(M_i)$ à chacun :

$$\text{À } t_0, \text{ pour tout } (x, y), b(x, y, t_0) = \sum_{i=1}^{n_b} p(M_i) \cdot b(x, y, t_0 | M_i) \quad (\text{Éq. 5.4})$$

Dans l'équation 5.3, $b(x, y, t_0)$ représente la bathymétrie sur la surface isochrone t_0 , au point de coordonnées (x, y) ; $b(x, y, t_0 | M_i)$ désigne la bathymétrie correspondante pour le modèle M_i (i est l'index du modèle de carte). Le poids $p(M_i)$ renvoie à la probabilité pour que le modèle de paléogéographie M_i soit choisi, et vérifient la relation :

$$\sum_{i=1}^{n_b} p(M_i) = 1 \quad (\text{Éq. 5.5})$$

Quelle que soit la valeur des paramètres $p(M_i)$, la cohérence sur le scénario de dépôt reste préservée puisque les cartes ainsi calculées se situent toujours dans le domaine d'hypothèses fixé par le sédimentologiste (Fig. 5.4). Par conséquent, la multi-réalisation en termes de cartes devient envisageable par échantillonnage des coefficients $p(M_i)$.

Toutefois, la manière de sélectionner ces poids traduit des situations différentes d'incertitude. Par exemple, en cas d'ignorance complète sur leur valeur, on peut les simuler selon une loi uniforme en vertu du principe du maximum d'entropie. Cette approche présente l'avantage de balayer la totalité des cartes possibles connaissant les modèles de base. Pour le cas considéré ici, cela revient à tirer un point au « hasard » sur le triangle dont les sommets correspondent aux trois topographies de base (Fig. 5.5-a). Par extension, pour n_b modèles, il s'agit d'échantillonner des points sur le simplexe définis les n_b modèles de base. Il existe d'ailleurs plusieurs algorithmes de simulation d'une loi uniforme sur un simplexe de dimensions n_b .

Maintenant, si l'expert possède un a priori $\{ \hat{p}(M_i) - i = 1, \dots, n_b \}$ sur ces grandeurs, il est possible d'affecter des poids selon une loi probabiliste : elle traduit alors son incertitude autour d'une telle configuration. Sa moyenne correspond à cet a priori $\{ \hat{p}(M_i) \}$, et

sa variance à la confiance que place l'expert dans les différents modèles. Par exemple, les lois de Dirichlet, que l'on définira plus précisément dans la troisième partie, sont souvent utilisées dans ce sens. Elles présentent l'avantage de fournir pour chaque réalisation une série de nombres dont la somme vaut un sans avoir à les normaliser. On peut également citer les lois d'Aitchinson (1980, 1986) dont il use pour les analyses de compositions géochimiques (Fig. 5.5– b).

L'avantage de ces méthodes, outre leur simplicité, réside dans la possibilité de facilement les généraliser à un nombre important de paléogéographie de référence. Elle permet d'élargir le pool de cartes possibles tout en restant dans le domaine d'hypothèses de l'utilisateur. On espère ainsi prendre en compte les incertitudes liées à la paléogéographie. Différents paysages sont ainsi générés tout en restant dans le domaine défini par l'utilisateur, ce qui préserve une certaine cohérence géologique. Elle présente aussi l'avantage de fournir des réalisations qui sont toutes, par construction, contraintes au niveau des puits.

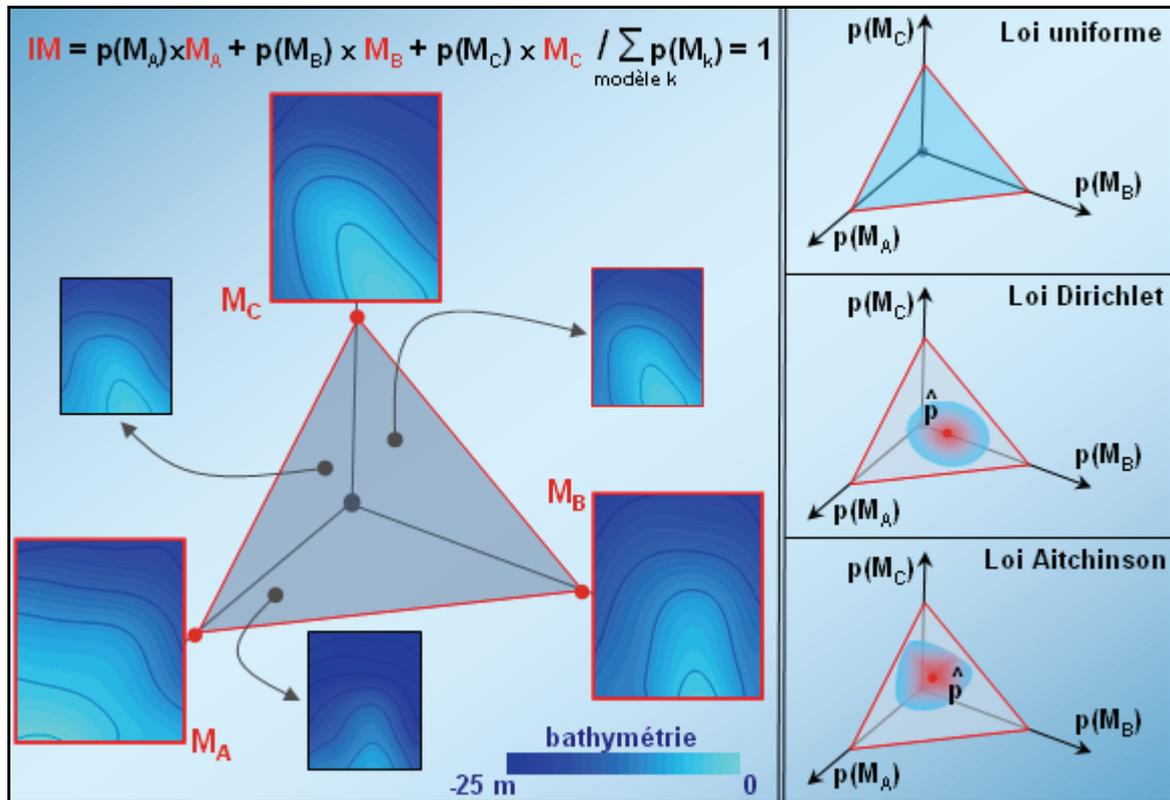


FIG. 5.8 : Combinaison linéaire des cartes de paléogéographie. À gauche, les cartes qui sont utilisées dans le processus de multi-réalisation sont issues de la combinaison linéaire des n_b (ici $n_b = 3$) cartes de base (ici notées M_A , M_B et M_C). Les poids qui leur sont respectivement attribués peuvent être simulés à partir de lois de probabilités uniforme, de Dirichlet ou d'Aitchinson (à droite).

Simulation « gaussienne » de cartes

Ce paragraphe revient sur une méthodologie de simulation des cartes qui repose en partie sur les outils géostatistiques. Elle reprend le principe des méthodes utilisées pour intégrer l'incertitude structurale en sismique [Gueméné *et al.*, 1998]. Leur intérêt réside dans le fait qu'une telle approche permet d'obtenir des cartes de bathymétrie difficilement reproductibles par l'intermédiaire des seules combinaisons linéaires des modèles initiaux. Elle travaille à l'échelle du point pour estimer la loi de la bathymétrie et à celle de la carte pour les combinaisons. Le protocole se divise en trois grandes étapes (Fig. 5 .6) :

1. Analyse des données

Le même point (x,y) , dans les différents modèles M_i , est affecté de valeurs de bathymétrie distinctes. Les données à traiter correspondent à un ensemble de n_b individus (les modèles), dont on connaît une variable quantitative (la bathymétrie). Cette série de données s'interprète comme la réalisation d'un échantillon de taille n_b (nombre de modèles) de la variable aléatoire bathymétrie $\mathbf{B}(x,y)$.

2. Loi jointe de la variable aléatoire

En chaque point de coordonnées (x,y) , on suppose que la réalisation de la fonction aléatoire $\mathbf{B}(x,y)$ correspond à une loi normale de moyenne $\bar{b}(x,y)$, et d'écart-type $\sigma(x,y)$. Considérer de telles distributions est un choix de modélisation qui donne généralement de bons résultats. D'ailleurs, il est toujours possible de se ramener à des lois normales, qui restent plus pratiques à manipuler pour les outils géostatistiques. Les « n_b » cas de base servent à définir les paramètres des lois normales régionalisées. On les considère comme n_b réalisations de ces lois. Cet échantillon permet l'estimation de leurs premiers moments statistiques $\bar{b}(x,y)$ et $\sigma(x,y)$. Ces paramètres définissent entièrement les lois régionalisées, et permettent de prendre en compte l'incertitude locale sur la bathymétrie.

3. Simulation des cartes

L'objectif consiste à faire respirer les cartes de bathymétrie. Ainsi, pour construire une nouvelle carte, on simule une réalisation des variables $\mathbf{B}(x,y)$ pour l'ensemble des points (x,y) de la carte. Il est par contre nécessaire de contraindre les simulations entre elles en chaque point afin de conserver la cohérence spatiale du modèle simulé. Une approche simple consiste, pour chaque réalisation de carte, à extraire aléatoirement, en chacun des points (x,y) , la bathymétrie correspondant à un même quantile u , et ce pour chaque loi $\mathbf{B}(x,y)$.

On souligne que toutes les cartes simulées sont contraintes aux puits. Au niveau de ceux-ci, la bathymétrie simulée correspond toujours à la bathymétrie observée : les cartes de base étant toutes contraintes aux puits, $\bar{b}(puits)$ est égale à la bathymétrie observée aux puits et $\sigma(puits)$ est nul. Appliquée au cas étudié ici, cette méthodologie décrit des modèles difficilement reproductibles par l'intermédiaire de combinaisons linéaires des modèles

initiaux. La figure 5.9 montre quelques réalisations en termes de cartes de bathymétrie simulées selon ce protocole.

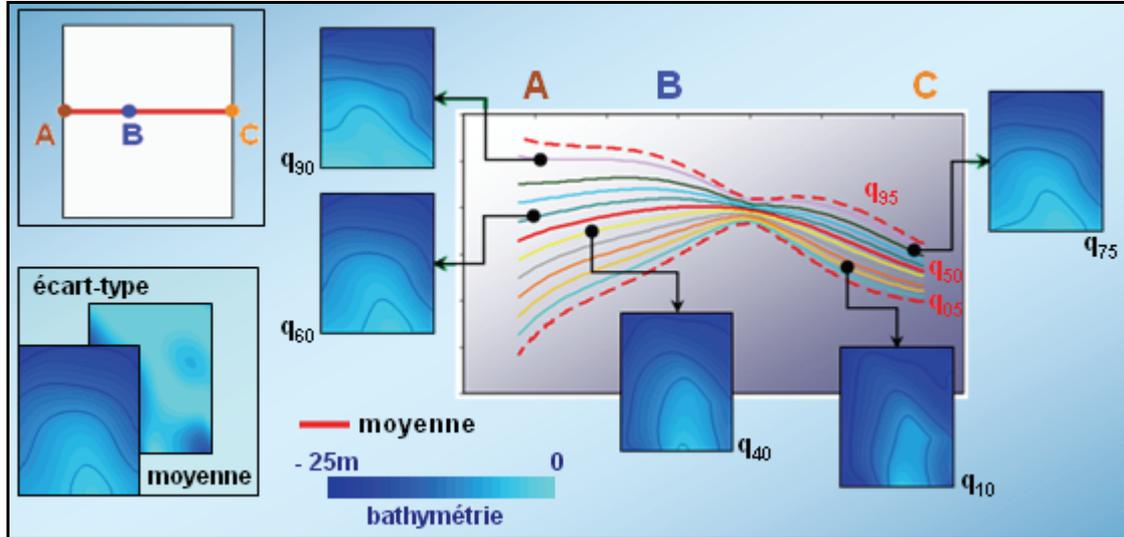


FIG. 5.9 : Vue en coupe des cartes issues de la combinaison « gaussienne » des cartes de bathymétrie de base. À partir de ces dernières, on estime en chaque point (x,y,t_0) , les paramètres d'une loi gaussienne (moyenne et écart-type). Les cartes utilisées pour le processus de multi-réalisation correspondant au même quantile pour les différentes lois gaussiennes estimées.

Pour chaque réalisation, les signaux observés aux puits autres que la bathymétrie, sont extrapolés le long des surfaces isochrones de manière déterministe en terme de modèle de continuité. Par contre, si la paléographie de référence demeure incertaine, le sédimentologue traduit son indétermination à travers plusieurs hypothèses de cartes. À partir de celles-ci, il est possible d'adopter une approche multi-réalisation sur ce paramètre, soit de manière discrète, ou en réconciliant les différents modèles. La carte ainsi générée est alors utilisée pour calculer la bathymétrie dans tout le réservoir. Il est maintenant question d'évaluer le cube de proportions correspondant ainsi que les zones susceptibles de connaître la diagenèse précoce.

Chapitre 6

Incertitude sur le réservoir (3D)

Dans un premier temps, connaissant la bathymétrie partout dans l'espace, il s'agit d'évaluer la loi de probabilité de chaque faciès en tout point du réservoir. Cette estimation repose sur une approche bayésienne qui inverse la relation entre faciès et bathymétrie proposée au départ. Classiquement, elle nécessite de fournir une série de poids notés $\mathbb{P}(F=f)$, qui correspondent aux probabilités a priori d'obtenir chacun des faciès f . Ils traduisent la prédominance des sédiments les uns par rapport aux autres dans l'ensemble du réservoir. Malgré l'abondance des techniques d'estimation, ces derniers portent toujours une part d'incertitude liée au caractère subjectif de leur définition. D'autre part, la bathymétrie permet de distinguer les environnements de dépôt : l'émergence du sol favorise souvent les phénomènes de diagenèse précoce, qui modifient les qualités pétrophysiques des faciès affectés. Néanmoins, une incertitude persiste quant au volume des régions exposées compte tenu de la difficulté à estimer précisément leur extension. Ce chapitre propose d'abord une méthode alternative pour estimer les poids des différents faciès. Il revient également sur un moyen de tester différentes hypothèses sur l'extension de la zone vadose en vue de l'inclure dans un processus multi-réalisation.

6.1 Modélisation des proportions de faciès

Le chapitre 4 décrit la création des courbes de bathymétrie le long des puits. Leur construction repose sur la correspondance entre faciès et bathymétrie, que l'on interprète en termes de lois probabilistes. Elles définissent la distribution de bathymétrie en fonction du faciès, c'est-à-dire la probabilité $\mathbb{P}(B = b|F = f)$ ¹. La bathymétrie, calculée dans tout le réservoir (chapitre 5), peut être utilisée pour inverser cette relation, et en déduire en tout point de l'espace géologique, la probabilité d'occurrence de chaque faciès sachant la bathymétrie, c'est-à-dire la probabilité $\mathbb{P}(F = f|B = b)$. L'inversion s'appuie sur le

¹ $\mathbb{P}(B = b|F = f)$ renvoie à la probabilité que la bathymétrie B au point (x, y, t) soit égale à une valeur b sachant que le sédiment F déposé en ce point correspond au faciès f .

formalisme bayésien, et nécessite d'ordinaire la définition préalable de la probabilité a priori d'apparition des n_f faciès $\mathbb{P}(F = f)$ dans l'intégralité de la zone d'étude :

pour chaque faciès f ,

$$\mathbb{P}(F = f | B = b) = \frac{\mathbb{P}(B = b | F = f) \times \mathbb{P}(F = f)}{\sum_{f=1}^{n_f} \mathbb{P}(B = b | F = f) \times \mathbb{P}(F = f)} \quad (\text{Éq. 6.1})$$

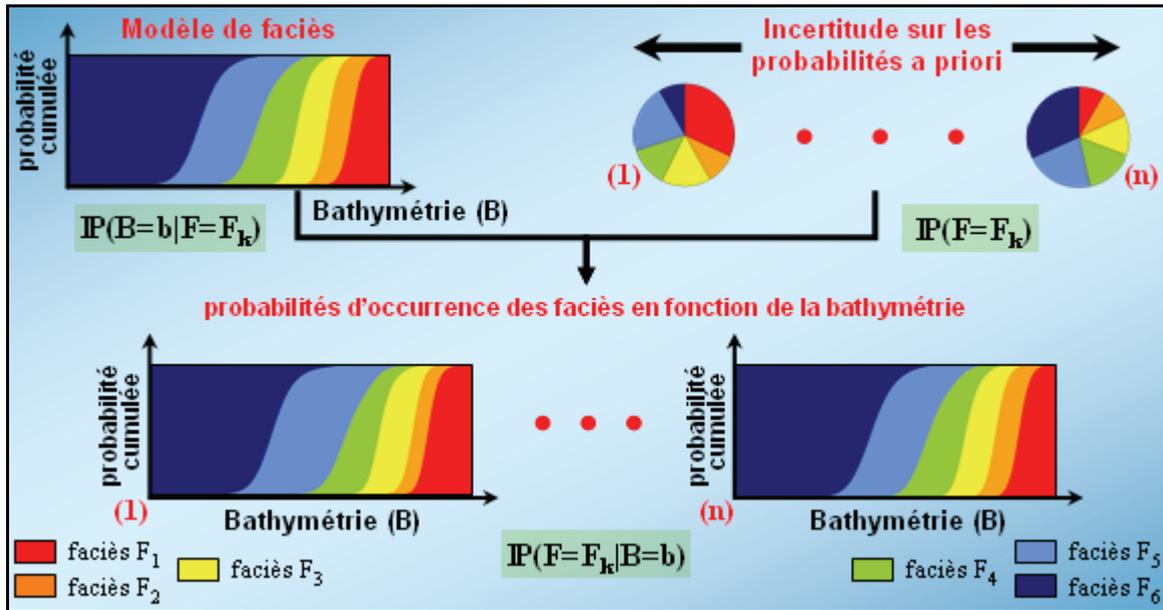


FIG. 6.1 : L'incertitude sur les probabilités a priori. La définition des probabilités permet de définir la relation permettant l'inversion de la relation faciès bathymétrie.

Ces probabilités a priori, aussi appelées poids, déterminent la prédominance des faciès les uns sur les autres. Elles représentent la proportion globale de chacun d'eux dans le réservoir, et sont postulées sans la moindre information locale. En pratique, l'estimation de ces termes $\mathbb{P}(F = f)$ n'a rien d'évident : un calcul direct à partir des données de puits introduit certainement un biais sur les résultats. Les observations aux puits sont rarement représentatives de l'ensemble du réservoir, car ces derniers sont habituellement forés dans les zones sensées être les plus prometteuses. Si par exemple on dispose d'une imagerie sismique de la région, on privilégiera les secteurs à faible impédance, potentiellement meilleures zones réservoirs. En conséquence, les statistiques traditionnelles [Biver *et al.*, 1998], qui supposent les observations aux puits indépendantes, ignorent les structures géologiques, sources de dépendances spatiales et de redondance. On citera notamment les techniques de rééchantillonnage, ou *bootstrap* en anglais [Efron *et al.*, 1983, 1993]. Il s'agit d'un ensemble de méthodes qui rééchantillonnent un grand nombre de fois l'échantillon réellement prélevé dans la population : ici, les données observées aux puits (chapitre 2). L'inférence statistique des poids $\mathbb{P}(F = f)$ s'appuie sur les résultats des

échantillons ainsi obtenus. Cependant, outre l'oubli de la corrélation spatiale entre les données, ces procédés supposent finalement les puits représentatifs du réservoir. C'est pourquoi plusieurs méthodes ont été proposées afin de lever ce biais.

Une première approche s'affranchie des données de puits en ce sens qu'elle repose davantage sur leur interprétation au sein d'un modèle sédimentaire que sur un calcul direct (ou non) des poids à l'aide des observations. Une solution commode consiste ainsi à faire l'hypothèse d'équiprobabilité entre tous les faciès ou encore, de s'en remettre à l'avis d'experts. Ces derniers définissent les poids respectifs des sédiments sur la base de réservoirs analogues présentant la même géologie. À l'instar des cartes de référence (section 5.2), si l'expert est capable de quantifier son ignorance autour de son a priori sur les poids, que l'on note $\{ \hat{p}(F = f_i) - i = 1, \dots, n_f \}$, il est possible d'intégrer la variabilité sur ces grandeurs à travers une loi de probabilité de type Dirichlet par exemple (partie 3). En effet, s'agissant des statistiques globales – les proportions de chaque faciès ici – il est en général demandé aux sédimentologues d'envisager au moins trois cas : pessimiste, médian et optimiste. Ces scénarios peuvent être utilisés pour paramétrer une distribution à partir de laquelle les proportions sont simulées. Par contre, en cas d'ignorance totale sur leur valeur, on peut toujours simuler cette série de proportions selon une loi uniforme¹, et balayer de la sorte toutes les configurations de poids sans exception.

D'autres méthodes préfèrent remédier à la non-représentativité des observations réalisées le long des puits. Journel (1994) ; Norris *et al.* (1993) proposent des techniques de rééchantillonnage spatial afin d'appréhender la variabilité associée à la localisation des données. Des puits forés à un autre emplacement fourniraient une estimation différente de la moyenne globale des faciès dans le réservoir. Une succession d'échantillons fictifs est alors définie : on déplace la configuration réelle des puits à travers une réalisation en faciès du réservoir obtenue préalablement par simulation. L'analyse des différents jeux de puits – au moyen des techniques classiques de rééchantillonnage – définit la distribution des fréquences moyennes de faciès dans la zone d'étude. Cependant, ces méthodes supposent une répartition stationnaire des faciès ; leur probabilité d'occurrence est constante quelle que soit la position. Or, sur une rampe carbonatée, la nature des sédiments varie dans l'espace. À première vue, cette non-stationnarité interdit le recours direct à de tels procédés dans les cas que traite Neptune par exemple.

Leflon (2005) évalue les poids $\mathbb{P}(F = f)$ en analysant la distribution de faciès relativement à la distribution de bathymétrie observée sur les puits. En notant $\{B_{\text{puits}}\}$ l'échantillon des bathymétries estimées le long des puits, la proportion de chaque faciès sur les puits renvoie à la probabilité d'un faciès conditionnée à un tel échantillonnage $\{B_{\text{puits}}\}$ de la bathymétrie, c'est-à-dire la probabilité $\mathbb{P}(F = f | B_{\text{puits}})$ ². Les distributions de

¹ Il s'agit d'échantillonner des points sur le simplexe défini par : $p(F = f_1) + \dots + p(F = f_{n_f}) = 1$, et tels que les poids $p(F = f_i)$ soient tous positifs. Il existe plusieurs algorithmes de simulation d'une loi uniforme sur un simplexe de dimensions n_f quelconque.

² La notation $\mathbb{P}(F = f | B_p)$ insiste sur le fait que la distribution de bathymétrie correspond à celle estimée au niveau des puits, elle ne correspond pas forcément à la distribution de la bathymétrie dans le réservoir.

bathymétrie définies pour chaque faciès¹ donne la probabilité de définir un tel échantillon de bathymétrie selon la nature du faciès f , soit $\mathbb{P}(B_{puits}|F=f)$. La formule de Bayes et les probabilités conditionnelles permettent le calcul des probabilités a priori de chaque faciès (Éq. 6.2). La probabilité $\mathbb{P}(B_{puits})$ n'a pas besoin d'être calculée, puisqu'elle correspond au terme normalisateur de la somme des probabilités a priori des différents faciès :

pour chaque faciès f ,

$$\mathbb{P}(F=f) = \frac{\mathbb{P}(F=f|B_{puits}) \times \mathbb{P}(B_{puits})}{\mathbb{P}(B_{puits}|F=f)} \quad (\text{Éq. 6.2})$$

Cette approche, si elle introduit l'information qu'apporte la bathymétrie sur le calcul du poids des faciès, ne permet pas véritablement d'apprécier la représentativité des puits. Elle suppose la bathymétrie observée aux puits $\{B_{puits}\}$ caractéristique de la distribution de bathymétrie considérée sur tout le réservoir, et que l'on désignera par la suite $\{B_{rés.}\}$. Pourtant à ce stade, on connaît la profondeur d'eau en tout point du réservoir, autrement dit la distribution globale $\{B_{rés.}\}$. Par conséquent, il est possible, dans un premier temps, de quantifier la représentativité de l'échantillon $\{B_{puits}\}$ à l'égard de la distribution de bathymétrie sur l'ensemble du réservoir. Pour cela, on recourt aux distances de Kolmogorov ou du χ^2 , couramment utilisées par les statisticiens [Verdel, 2000] pour mesurer l'écart entre distributions théoriques et observées. On vérifie alors le degré de légitimité du rac-

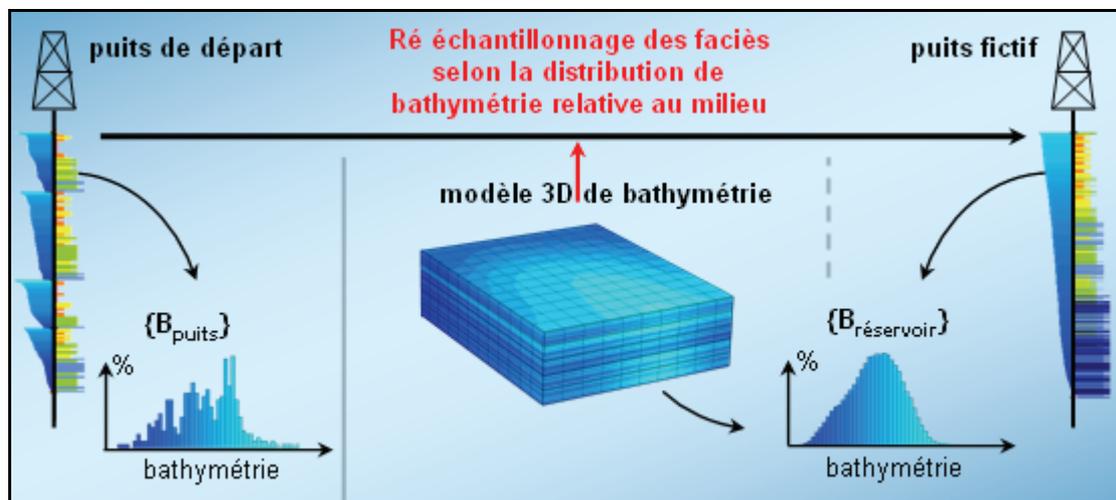


FIG. 6.2 : Un mode alternatif de calcul des poids a priori. Connaissant la distribution de la bathymétrie dans le réservoir, on échantillonne les couples (bathymétrie, faciès) aux puits selon la loi de répartition associée à la bathymétrie à l'échelle du réservoir. Le poids a priori des faciès considérés correspond à leur proportion respective dans l'échantillon ainsi construit.

¹ Autrement dit, le modèle de faciès.

cordement de la loi associée à $\{B_{puits}\}$, vis-à-vis de la distribution globale $\{B_{rés.}\}$. D'autre part, on propose de construire, à partir des puits, un échantillon de faciès représentatif de la bathymétrie globale. On sélectionne, le long des puits, une série de couples "bathymétrie/faciès" selon la loi de répartition associée à la bathymétrie dans le réservoir. L'échantillon ainsi obtenu est alors utilisé pour calculer la proportion globale de chacun des faciès. En fait, il s'agit d'une démarche similaire aux méthodes employées dans les sondages d'opinions. On réalise un choix raisonné d'individus selon un système de quota associé à une variable secondaire, la bathymétrie, sensée reproduire les caractéristiques d'une population, les faciès.

Au final, plusieurs alternatives existent pour quantifier l'incertitude sur le poids des faciès. On préférera néanmoins, les méthodes qui s'affranchissent de la non-représentativité des puits. Pour cela, elles introduisent soit une connaissance supplémentaire issue de l'avis d'experts, ou bien une information complémentaire, tirée d'un signal secondaire comme la bathymétrie. Quelle que soit la méthode employée, la définition de ces poids permet d'établir en tout point du réservoir, la probabilité d'occurrence de chaque faciès selon l'équation 6.1.

6.2 Modélisation de la diagenèse précoce

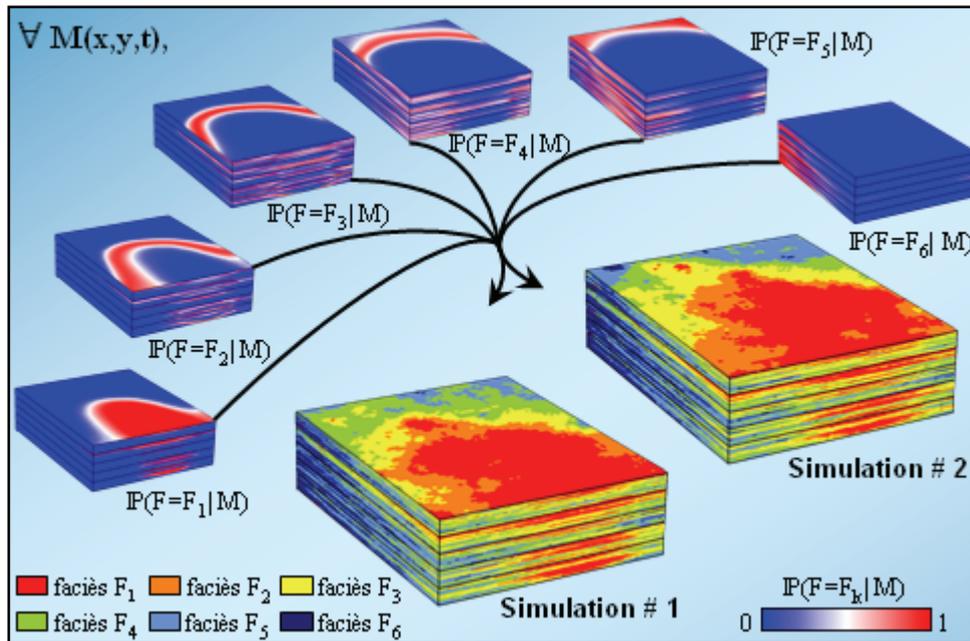


FIG. 6.3 : Simulation des faciès. À partir des probabilités d'apparition des faciès (grilles bleues à rouges), des réalisations en faciès sont générées (ici deux). Remarquons que les résultats des simulations sont très semblables. C'est pourquoi il est nécessaire de réaliser l'étude des incertitudes en amont de la construction du cube de proportions (ensemble des grilles rouges à bleues), en intégrant les incertitudes des modèles sédimentologiques et stratigraphiques jouant sur le cube de proportions (d'après Leflon, 2005).

Les probabilités $\mathbb{P}(F=f|B=b)$, appelées aussi proportions, sont utilisées pour contraindre les simulations en faciès. Plusieurs méthodes existent pour construire ces modèles stochastiques connaissant la loi spatiale des faciès. Labat (2004) revient sur les différentes techniques de simulations de faciès existantes. Chaque réalisation définit alors, en tout point du modèle, la nature des sédiments tels qu'ils se sont formés au moment de leur dépôt (Fig. 6.3). Dès lors, ils subissent différentes transformations liées à la diagenèse : actions bactériennes, dissolution, cimentation, recristallisation, compaction, etc. [Barthurst, 1975 ; Moore, 2001], qu'ils soient encore proches de la surface, c'est la diagenèse précoce, ou enfouis profondément. De tels phénomènes bouleversent naturellement les propriétés initiales du milieu, en particulier pétrophysiques. Aussi, les mécanismes de dissolution expliquent-ils généralement la présence de drains à forte perméabilité dans les réservoirs carbonatés. La dissolution affecte en partie les sédiments quand ils se trouvent en environnement diagénétique météorique¹ [Carrio-Shaffhauser, 2004], c'est-à-dire à proximité de la surface. Au cours de son histoire, la plate-forme carbonatée peut effectivement connaître des émergences durant lesquelles les eaux météoriques, très agressives à l'encontre des minéraux du domaine carbonaté, modifient fortement la porosité et la perméabilité de départ. Ces épisodes d'émergences améliorent souvent de façon considérable les qualités réservoirs des sédiments². Ces changements ne touchent d'ailleurs pas uniquement les parties en contact avec l'air. Les eaux météoriques s'infiltrent de la surface du paléosol au toit du domaine phréatique, c'est-à-dire à travers la zone vadose (Fig. 6.4). Celle-ci devient alors le siège d'intenses percolations des fluides à l'origine de la dissolution des roches carbonatées. Un tel contexte favorise d'ordinaire le développement

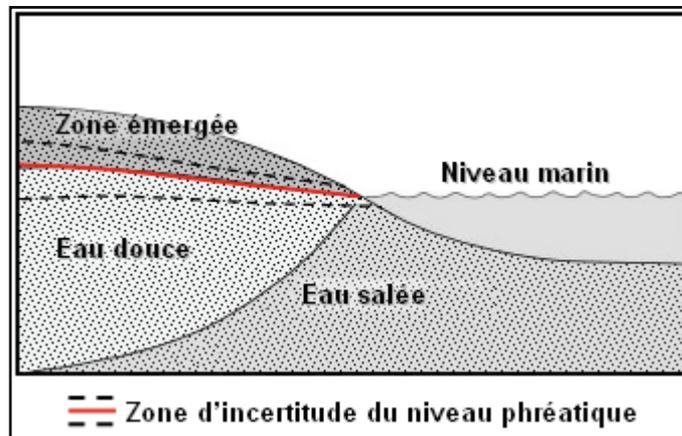


FIG. 6.4 : Régions soumises à la diagenèse précoce et incertitude sur leur extension. Les eaux météoriques s'infiltrent des zones émergées jusqu'au toit du domaine phréatique. Cette limite supérieure reste difficile à appréhender, et joue directement sur les volumes.

¹ D'autres phénomènes de dissolutions peuvent se produire pendant, et après l'enfouissement des sédiments. On ne s'intéresse dans ce mémoire qu'aux phénomènes liés à la diagenèse précoce, c'est-à-dire, avant l'enfouissement.

² Ces dissolutions peuvent être compensées par la cimentation, la recristallisation ou la précipitation.

de karsts¹, témoins de la transformation du réseau poreux comme de la perméabilité originelle. Par conséquent, un même sédiment présente souvent des propriétés pétrophysiques différentes selon s'il s'est trouvé ou non, durant son histoire, dans la zone vadose créée pendant les phases d'émersions.

La méthodologie Neptune reconstitue l'histoire bathymétrique de la zone d'étude. Combinée à l'évolution de l'épaisseur des couches, elle identifie les régions ayant été émergées pendant la formation de la séquence carbonatée (Fig. 6.5). Celles-ci, on l'a dit, sont susceptibles de connaître la diagenèse météorique. Néanmoins, si les régions émergées sont relativement simples à déterminer, la limite inférieure de la zone vadose, elle, ne l'est pas forcément. Elle correspond à la surface piézométrique, et varie au cours de la période d'émersion avec les battements de la zone phréatique, dont elle marque l'interface avec le domaine vadose. Il apparaît une incertitude sur l'extension de la zone vadose, et par là même sur le volume des régions exposées, compte tenu de la difficulté d'estimer l'ampleur de tels battements (Fig. 6.4).

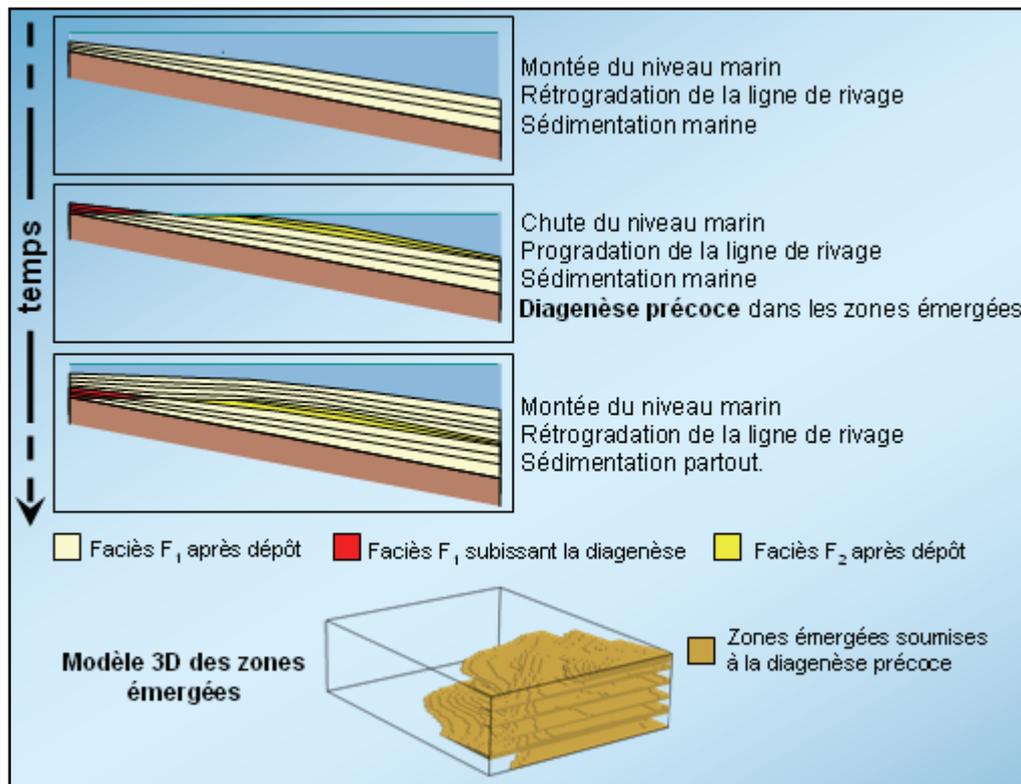


FIG. 6.5 : Principe de l'estimation des zones émergées dans Neptune. Connaissant l'histoire de la bathymétrie comme l'épaisseur des sédiments au cours du « temps », il est possible de reconstituer les régions émergées et susceptibles de subir une diagenèse précoce.

¹ **Karst** : il désigne une géomorphologie désordonnée et irrégulière montrant des formes particulières résultant de l'action dissolvante de l'eau sur une roche soluble, le plus souvent carbonatée.

En premier lieu, il est possible de reproduire l'équivalent de ces fluctuations en modifiant la bathymétrie $b(x,y,t)$ sur tout le modèle. La figure 6.6 montre en effet que l'augmentation « artificielle » de la tranche d'eau fait mécaniquement remonter le niveau de la surface piézométrique. Il en ressort une réduction du volume des régions qui subissent la diagenèse météorique. Une diminution de la bathymétrie aura bien entendu l'effet inverse. En pratique, on introduit un facteur multiplicatif qui porte sur l'ensemble de la bathymétrie, et dont le sédimentologue fixe le domaine de variation. Il peut alors tester différentes alternatives sur l'extension de la zone vadose, soit de manière discrète en pondérant ses hypothèses, ou bien de façon continue en échantillonnant ce paramètre selon une loi de probabilité qu'il aura préalablement établie.

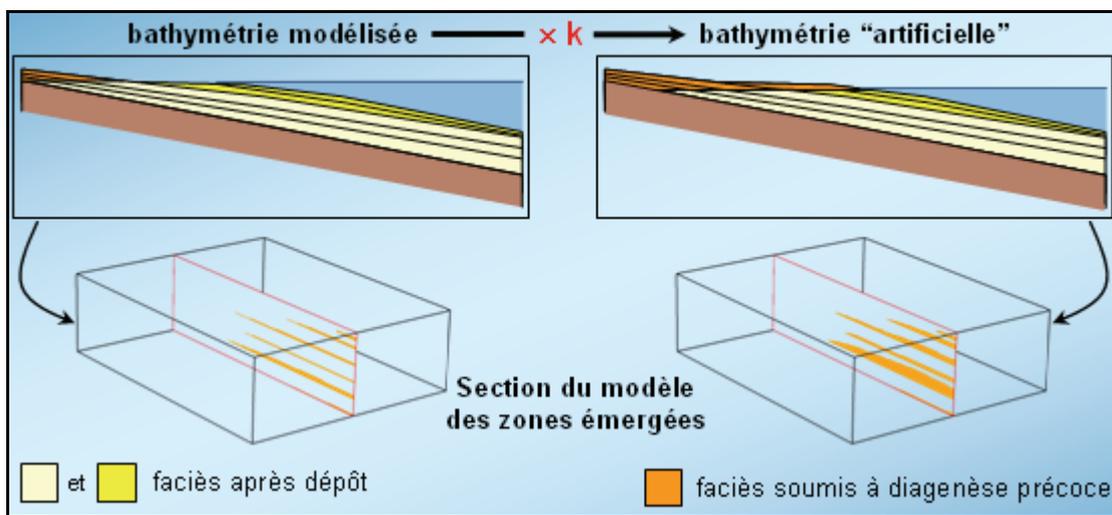


FIG. 6.6 : Intégration de l'incertitude sur le volume de la zone soumise à la diagenèse précoce. La bathymétrie modélisée initialement est multipliée par un coefficient multiplicateur, qui permet de faire varier le volume des zones émergées entre des scénarios optimiste et pessimiste.

Par conséquent, on identifie les régions qui connaissent au cours de leur histoire une diagenèse de type météorique en intégrant la variabilité sur leur volume. Ainsi, des porosités et perméabilités distinctes peuvent être affectées à un même faciès, selon qu'il se situe ou non dans de telles régions. Pour cela, la simulation des propriétés pétrophysiques dans ces régions sera faite indépendamment des simulations dans les zones qui n'ont pas subi ces transformations [Massonnat *et al.*, 2002a ; Leflon, 2005]. Au final, les propriétés réservoirs du modèle traduisent à la fois la nature du faciès déposé et son comportement diagénétique.

Conclusions et Perspectives (1)

Les travaux menés dans cette partie proposent une méthode de simulation des cubes de proportions en accord avec l'incertitude propre au scénario géologique dans des environnements sédimentaires complexes tels que les plates-formes carbonatée ou mixte. Ils participent notamment à la gestion des incertitudes sur le scénario géologique rencontrées dans le processus de modélisation sédimentaire en milieu marin, Neptune. L'approche qui a été adoptée représente alors de façon probabiliste les paramètres qui composent le scénario géologique, et qui correspondent aux grandeurs d'entrée de Neptune : modèle de faciès, cartes de paléogéographies, biseau de diagenèse... Une analyse de Monte Carlo conduite sur les composantes du schéma géologique traduit de manière explicite, à travers les réalisations stochastiques multiples en cubes de proportions qui en résultent, l'incertitude sur le scénario géologique.

La méthode proposée remédie de cette manière aux caractères discrets et limités des études à plusieurs scénarios ou de combinaisons d'accumulations, qui généralement oublient l'incertitude sur le modèle probabiliste (le cube de proportions) utilisé pour effectuer les simulations pétrophysiques. Elle évite par conséquent l'écueil de l'utilisation de distributions de probabilité uniques pour représenter notre état de connaissance sur les paramètres du modèle. Elle améliore l'évaluation de la variabilité (a) des proportions de faciès en chaque point du milieu, et (b) des grandeurs clés d'un champ qui en dépendent : accumulations, profils de production, réserves...

En définitive, les apports essentiels des travaux réalisés dans cette partie comprennent :

- l'interprétation dans un cadre probabiliste, indépendamment de leur dimension, d'un large panel de paramètres qui définissent le scénario géologique comme le modèle de faciès, les cartes de bathymétrie, etc. ;
- la génération rapide de multiples cubes de proportions cohérents en terme de géologie, et qui intègrent, avec leur incertitude, des interprétations et des données de nature ou de résolution variables ;
- le recentrage de la quantification de l'incertitude géologique autour des modèles de probabilités¹ utilisés pour accomplir les traditionnelles simulations géostatistiques des propriétés pétrophysiques du milieu ;

¹ c'est-à-dire les cubes de proportions de faciès.

- la correspondance avec les travaux de Leflon (2005). Il propose des incertitudes sur le scénario rencontrées dans le processus de modélisation sédimentaire en milieu marin, Neptune et la définition d'un Workflow.

Dans le prolongement de ces travaux de thèse, il serait néanmoins intéressant de revenir sur :

- l'application de l'approche développée à des procédés de modélisation autre que Neptune. Cette partie revêt en effet essentiellement des aspects méthodologiques : elle propose divers protocoles, qui tous visent à faire varier de manière stochastique les paramètres d'entrée d'un processus de modélisation sédimentaire particulier (Neptune). Pourtant, les méthodes présentées ici n'en perdent pas pour autant leur généralité, car les paramètres d'entrée utilisés dans Neptune, du modèle de faciès à leur poids global, sont (1) d'un emploi courant dans la plupart des techniques de modélisation du milieu souterrain, ou (2) posent, de façon générale, le problème d'intégrer la variabilité de grandeurs multidimensionnelles. Par exemple, malgré leur différence avec les cartes de bathymétrie employées dans Neptune, les cartes auxquelles recourent les méthodes géostatiques, des cartes de tendance¹ aux images d'entraînement², portent une incertitude qui semble similaire³, et à ce titre peut être appréhendée de la même manière ;
- la connexion avec la future version industrielle de Neptune. Dans la pratique, un certain nombre de paramètres porteurs d'incertitude sont amenés à être simulés. Il s'agit donc de les entrer ainsi que leur variabilité de manière efficace. Cela passera par un *workflow* qui est un interface graphique permettant un déroulement logique et rapide d'actions successives pour la création d'un modèle. Un premier *workflow* a été partiellement développé [Bergamo-Néel, 2006], au sein duquel il suffit d'intégrer les algorithmes de simulations sur les données d'entrée proposés ici. Certains d'entre eux ont d'ailleurs déjà été développés, notamment les fonctions associées aux réalisations multiples en termes de cartes de bathymétrie. Cependant, ceux-ci restent malgré tout des produits de recherche, qu'il est encore possible d'améliorer ;
- la limitation de l'incertitude sur les paramètres d'entrée, qui peut passer par un rapprochement des méthodes de multi-réalisation avec les techniques de recuit-simulé ;
- le développement d'outils « pratiques » qui faciliteraient la création des données d'entrée de Neptune et de leurs différentes alternatives. Un premier outil a été déve-

¹ Les cartes de tendance sont utilisées dans le but d'introduire des variations horizontales dans les modèles probabilistes utilisées dans les simulations géostatiques. Ces modèles probabilistes sont en effet issus de la combinaison entre justement cette carte de tendance horizontale (tirée de la sismique ou de l'interprétation du géologue) et une dérive dite verticale inférée à partir des puits (chapitres 2 et 9).

² Le rôle des images d'entraînement se rapproche davantage de celui du variogramme (chapitre 9) : ces cartes définissent le modèle de continuité des objets géologiques rencontrés que les méthodes de simulations géostatiques [Caers et Zang, 2004] doivent respecter.

³ Les diverses cartes de dérives ou de bathymétrie procèdent des différentes interprétations que proposent les géologues afin de tenir compte de l'incertitude sur ce paramètre du scénario sédimentaire.

loppé dans ce sens par Leflon (2005) au cours de sa thèse, pour une création rapide et simplifiée des cartes de bathymétrie par interpolation DSI. D'un autre côté, il doit être possible de créer une carte de bathymétrie en combinant une carte de faciès, couramment fournie par le géologue et un modèle de faciès selon une démarche similaire à la création des courbes de bathymétrie aux puits. Dans le même ordre d'idée, il serait possible de proposer un panel de « paysages » possibles sous la forme de cartes de tendances (avec des profils distincts). Cet ensemble peut être utilisé comme dérive pour l'extrapolation des données de bathymétrie aux puits par exemple ;

- le lien avec les travaux menés dans la partie à venir. Ceux-ci sont à l'origine d'une méthodologie alternative de simulation des cubes de proportions, qui repose sur des lois probabilistes dites de Dirichlet (chapitre 8). Cette approche nécessite un échantillon de quelques cubes de proportions. Les procédés proposés dans cette deuxième partie permettraient de construire rapidement les modèles de cet échantillon, et ainsi d'alimenter cette autre méthode.

Troisième partie

Simulations stochastiques de cubes de proportions

Cette partie présente une méthodologie alternative d'intégration de l'incertitude géologique au niveau du modèle sédimentaire. L'objectif consiste en effet à déterminer les lois de distribution régionalisées¹ de ces proportions dans l'ensemble de la zone d'étude. Pour cela, elle cherche à générer l'ensemble des images alternatives du réservoir à partir d'un échantillon de quelques cubes de proportions, tous issus de l'une des interprétations géologiques envisageables. Des réalisations multiples en terme de cubes de proportions (incertitude, dite épistémique, sur le schéma sédimentaire) remplacent les simulations géostatistiques sur la pétrophysique à partir d'un modèle statistique unique (incertitude de variabilité). L'ensemble de ces simulations traduit ainsi la variabilité des proportions de faciès. L'intérêt d'une telle technique réside dans le passage d'hypothèses discrètes – les scénarios géologiques, à un phénomène continu sur les proportions de faciès.

Les principes généraux de la méthode seront présentés dans un premier temps. Le **chapitre 8** s'attache à introduire les lois de Dirichlet utilisées ici pour la simulation des proportions de faciès, l'estimation de ces lois en tout point du milieu souterrain, ainsi que leur simulation ponctuelle. Le **chapitre 9** présente les méthodes, notamment géostatistiques, utilisées dans le but de conserver la continuité des formations géologiques au niveau des cubes de proportions quand on simule les lois de Dirichlet partout dans le modèle.

De telles simulations en termes de cubes de proportions tendent à grandement élargir l'espace des modèles construits. Le **chapitre 10** est donc dédié aux différentes méthodes utilisées dans le but de limiter l'espace d'incertitude en contraignant les simulations à d'autres signaux (sismique, données de puits, etc). D'un autre côté, cette méthode se montre plus intégrative et synthétique de l'information disponible (sismique, géologie, etc...) que la précédente (Deuxième partie). Le **chapitre 11** revient de fait sur la généralisation du formalisme des lois de Dirichlet à la combinaison de cubes de proportions issues de sources d'informations différentes : géologique, sismique et enfin statistique à partir des données de puits.

Enfin, le **chapitre 12** revient sur un moyen de régionaliser l'incertitude et ainsi, de définir les zones plus ou moins (in)certaines du modèle. Cette méthode peut d'ailleurs s'appliquer aux cubes de proportions construits selon une approche multi-réalisation (à partir de Neptune par exemple)

¹ c'est-à-dire en chaque point de l'espace.

Chapitre 7

Panorama d'une méthode de simulation des cubes de proportions

La simulation d'un modèle géologique nécessite l'estimation préalable des proportions de faciès dans l'espace. L'incertitude qui peut peser sur le scénario géologique affecte directement ces grandeurs. La méthode proposée dans cette partie, cherche donc à intégrer cette source d'incertitude directement au niveau du cube de proportions, qui décrit la loi de probabilité des faciès dans l'espace. C'est une méthode stochastique de simulation des cubes de proportions, qui à la différence d'un processus multi-réalisation, ne nécessite pas l'échantillonnage explicite des composantes du scénario géologique. Elle cherche plutôt à réconcilier différents modèles sédimentaires concurrents¹ que peut proposer le géologue. Elle en extrait un modèle synthétique, lui-même à l'origine des simulations en termes de cubes de proportions, que des sources d'informations autres que la géologie peuvent d'ailleurs contraindre.

7.1 Introduction

Le premier chapitre souligne que la définition du scénario géologique gouverne la distribution dans l'espace des probabilités d'existence, ou proportions, de chaque faciès. Par conséquent, le caractère incertain du scénario géologique : modèle de faciès, paléogéographie, etc., se répercute directement sur la répartition spatiale des probabilités d'occurrence de sédiments, c'est-à-dire le cube de proportions de faciès (chapitre 2). Pour quantifier cette incertitude, il est possible, comme exposé dans la deuxième partie,

¹ Il est important de préciser que ces scénarios géologiques font partie d'un même méta-scénario. Ce dernier distingue des environnements sédimentaires dont le passage de l'un à l'autre n'est pas continu dans le sens où il est difficile (intéressant ?) de combiner un modèle de réservoir de type clastique et carbonaté par exemple.

d'intégrer la variabilité des composantes du schéma sédimentaire par le biais d'une approche à réalisations multiples. Dans ce cas, il s'agit d'échantillonner les divers éléments incertains du scénario, dont les combinaisons génèrent une distribution de scénarios géologiques, elle-même à l'origine des différentes alternatives en termes de cubes de proportions. D'un autre côté, la question se pose malgré tout de savoir comment appréhender cette part d'incertitude, non pas simplement au niveau des paramètres du scénario, mais plutôt de manière directe sur les cubes de proportions. Car, dans la pratique, l'essentiel des méthodes de modélisation en faciès, notamment géostatistiques, n'introduisent pas de façon aussi explicite que Neptune les éléments nécessaires à la définition du scénario géologique.

Dès lors, certains auteurs [Haas *et al*, 2002 ; Biver *et al*, 2002] s'attachent à appréhender l'incertitude liée à l'estimation de ces probabilités. Toutefois, l'incertitude qu'ils évaluent est uniquement de nature statistique sur les données disponibles, en particulier aux puits (section 7.2). Ils s'intéressent principalement au caractère aléatoire de l'information, qu'explique la variabilité naturelle de toute grandeur mesurée à travers le milieu souterrain¹. Du reste, leur approche s'appuie d'ordinaire sur un modèle conceptuel ou statistique unique ; c'est l'équivalent d'un scénario géologique estimé de façon déterministe. Par exemple, en cas de non stationnarité de la fonction aléatoire associée aux faciès, on considère souvent une seule dérive horizontale afin de décrire l'évolution de leur proportion (soit leur moyenne) dans l'espace. Pourtant, cette dérive se révèle d'une importance comparable à la carte de référence introduite dans la méthodologie Neptune, et porte naturellement une part d'incertitude similaire. En conséquence, ces procédés n'évaluent pas l'incertitude liée à la répartition latérale des faciès, car celle-ci relève davantage du caractère imprécis et incomplet de l'information lié au manque de connaissance, que de sources de variabilités aléatoires.

Pour y remédier, nous proposons une méthode qui simule en tout point du réservoir, la probabilité d'occurrence de chaque faciès, et ce en accord avec l'incertitude liée au scénario géologique, sans toutefois recourir à la multi-réalisation de ses composantes. Plus précisément, c'est un processus stochastique qui permet d'aller au-delà des incertitudes de nature aléatoire, et qui cherche finalement à répondre aux objectifs suivants :

- quantifier l'incertitude sur le scénario géologique à travers la simulation directe des différents cubes de proportions de faciès alternatifs, pour décrire l'ensemble des états possibles, et ainsi représenter l'espace réel d'incertitude géologique : en d'autres termes, il s'agit de reconstruire la « population » de modèles compatibles avec cette incertitude ;
- être plus intégrative que les techniques de multi-réalisation sur les paramètres du scénario géologique, afin d'assimiler au mieux l'ensemble de l'information (sismique, puits) et de la connaissance (géologie) disponibles ainsi que leur incertitude ;

¹ On décrit souvent cette variabilité comme l'incertitude responsable de l'obtention de résultats différents lorsque l'on répète plusieurs fois dans des conditions identiques une expérience.

- construire une loi de distribution des grandeurs clés (accumulations en hydrocarbures), et décrire un faisceau de profils de productions, qui tiennent véritablement compte de l'incertitude géologique dans son intégralité ;
- régionaliser l'incertitude, c'est-à-dire définir les zones plus ou moins (in)certaines du modèle en terme de géologie.

7.2 Principe général de la méthode

Considérons un réservoir pétrolier, ou de façon plus générale une zone du sous-sol, dont il est question d'estimer les volumes d'hydrocarbures, la production potentielle, ou d'en évaluer les propriétés. À partir des mesures collectées¹, et de la connaissance géologique de la région, il est possible de construire un modèle unique du milieu souterrain, que traduit un cube de proportions. Celui-ci est censé synthétiser au mieux l'ensemble du savoir et des informations disponibles. Cependant, cette approche déterministe se révèle souvent insuffisante, car plusieurs modèles géologiques sont à même de respecter à la fois, les données et la connaissance à disposition. Aussi, l'expert transpose-t-il l'incertitude associée au scénario géologique, à travers la définition de plusieurs schémas sédimentaires concurrents les uns des autres. Par la suite, on notera $n_m(\mathbf{G})$ le nombre de scénarios géologiques pris en compte ; ceux-ci conduisent du reste à définir autant de cubes de proportions de faciès. Il est important de remarquer que la manière dont sont construits ces modèles de base ne remet pas en cause la généralité de la méthode présentée ici. Elle utilise pour données de départ un échantillon de cubes de proportions issus des différents scénarios sédimentaires concevables. Peu importe alors le procédé de construction, l'essentiel reposant en définitive sur le fait que, chaque cube de proportions reflète un scénario géologique différent.

La méthode que nous proposons repose sur l'idée naturelle qui consiste à réconcilier, quand c'est réalisable, les $n_m(\mathbf{G})$ scénarios géologiques admissibles. Elle diffère ainsi des stratégies classiques de sélection de modèles cherchant à ne conserver qu'un seul modèle suivant des critères d'optimisation en général, et à l'utiliser pour l'inférence statistique. Ces approches, comme le remarque Raftery (1993), ignorent l'incertitude issue du processus de sélection, et sous-estime l'incertitude totale existant sur la variable considérée. En fait, notre méthode s'apparente davantage aux techniques de mélange, que l'on rencontre fréquemment sur le thème de l'incertitude de modèle. Pour calculer les valeurs d'une variable aléatoire d'intérêt, elles recourent à un modèle défini comme une moyenne pondérée des différents modèles concurrents. L'utilisation de cette moyenne fournit en général de meilleures prédictions que l'emploi de n'importe quel modèle seul [Jacques, 2005].

Dans notre cas, il s'agit d'interpréter les $n_m(\mathbf{G})$ cubes de proportions initiaux comme un échantillon d'une population plus large de modèles. Celle-ci représente naturellement l'ensemble des cubes de proportions alternatifs et compatibles avec l'incertitude sur le

¹ Données de puits, campagne sismique, essais de puits, etc.

scénario géologique. En conséquence, la méthode présentée ici vise à reconstruire cette population en supposant qu'elle obéit à une loi probabiliste, dont on détermine les caractéristiques au moyen des cubes de proportions correspondant aux $n_m(G)$ scénarios géologiques. Elle présente ainsi une certaine similarité avec les techniques utilisées pour apprécier les incertitudes de mesures, responsables de l'obtention de résultats différents quand on répète plusieurs fois une expérience dans des conditions identiques. De la même manière, ces procédés estiment, à partir des mesures, la loi probabiliste qui représente la quantité physique étudiée et son incertitude. C'est le cas par exemple, avec les mesures de porosité utilisées pour inférer la distribution, en général gaussienne, de cette grandeur.

Nous insistons ici sur le fait que cette démarche constitue un choix méthodologique. Une telle décision permet le recours aux outils probabilistes et géostatistiques pour passer d'hypothèses discrètes : les $n_m(G)$ scénarios géologiques, à un phénomène continu. De la sorte, il devient possible de définir des modèles intermédiaires, et d'expérimenter des situations nouvelles, mais cohérentes en terme de scénario géologique. Autrement dit, elle balaie, à partir des $n_m(G)$ scénarios de base, l'ensemble des modèles statistiquement compatibles avec ces derniers tout en préservant une certaine cohérence géologique pour refléter l'incertitude globale sur le schéma sédimentaire. La figure 7.1 illustre les deux principales étapes de la méthodologie que l'on détaillera dans les chapitres à venir :

- (1) **la simulation des cubes de proportions alternatifs** : elle est d'abord contrainte à l'échelle de chaque point du réservoir par une loi de probabilité relative aux proportions de faciès, mais aussi au niveau du modèle afin de préserver la cohérence géologique. Plus précisément, ces simulations nécessitent l'estimation préalable de la distribution des proportions de faciès en chaque point de l'espace (chapitre 8), ces proportions étant supposées suivre une loi paramétrique. L'estimation de telles lois se fait à partir de quelques modèles construits sur la base des $n_m(G)$ scénarios distincts, issus de l'interprétation géologique des données. Ils permettent de définir les incertitudes sur les proportions qui en découlent dans l'espace. La simulation des proportions proprement dite, est réalisée en chaque point du milieu souterrain en ayant recours aux techniques géostatistiques. Selon l'environnement sédimentaire étudié, ces simulations s'effectuent, soit en introduisant une corrélation spatiale infinie, ou bien en déterminant cette cohérence par apprentissage sur les $n_m(G)$ cubes déjà existants (chapitre 9) ;
- (2) **la mise en cohérence des réalisations** : dans certains cas, il est possible de bénéficier d'informations supplémentaires autres que la géologie. Par exemple, l'imagerie sismique peut très bien fournir, à la résolution verticale du réservoir, une cartographie de la probabilité d'occurrence des faciès. L'objectif consiste donc à mettre en cohérence cette information avec les réalisations en cube de proportions décrivant l'espace des possibles en terme de géologie. Cette étape peut se faire, soit en amont de chaque simulation en contraignant la loi de probabilité des proportions avec cette information ou connaissance additionnelle, soit en contraignant à celles-ci les différentes réalisations (chapitre 10).

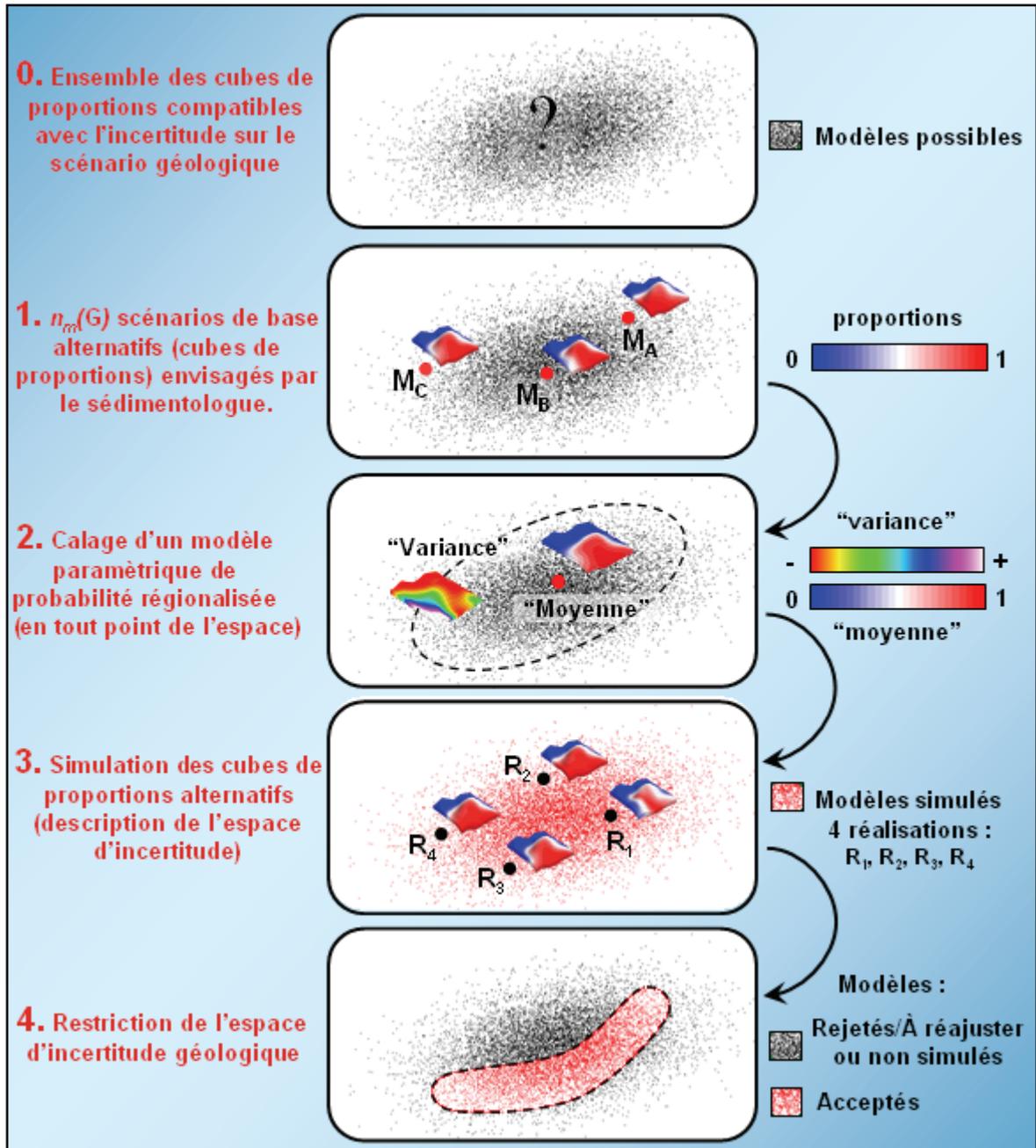


FIG. 7.1 : Étapes de la méthode de simulation stochastique des cubes de proportions. L'objectif consiste à décrire l'ensemble des cubes de proportions compatibles avec l'incertitude sur le scénario géologique : un point correspondant à un cube de proportions (0). À partir de quelques scénarios (1), on cale une loi de probabilité en chaque point de l'espace (2). Ces lois permettent de simuler des cubes de proportions alternatifs (3) que des données autres que la géologie peuvent contraindre (4), soit en amont des simulations en modifiant les lois de départ, ou en aval des simulations en recalant les cubes de proportions simulés.

Chapitre 8

Simulation régionalisée des proportions de faciès

Pour rappel, la méthode que nous proposons consiste d'abord à simuler des images alternatives en termes de proportions de faciès et ce, en tout point du réservoir. Celles-ci respectent l'incertitude géologique qui est évaluée au moyen d'un nombre $n_m(G)$ de cubes de proportions, eux-mêmes issus des diverses interprétations géologiques envisageables faites à partir des données. Par conséquent, des réalisations multiples en termes de cubes de proportions (incertitude sur le schéma sédimentaire) remplacent les simulations géostatistiques classiques initiées à partir d'un modèle statistique, ou loi de répartition spatiale des faciès, unique (incertitude de nature aléatoire). Le chapitre à venir détaille les étapes nécessaires à l'accomplissement de telles simulations, non pas à l'échelle du réservoir mais d'abord, à celle du point : nature des lois probabilistes utilisées (en particulier les lois de Dirichlet), leur paramétrage, la simulation des proportions à proprement parler, etc. Nous reviendrons auparavant sur les principales hypothèses que réclament celles-ci, avec d'une part, le recours au cadre probabiliste sur le plan des cubes de proportions, et d'autre part, la définition préalable du support géométrique indispensable à leur exécution.

8.1 Hypothèses de travail

8.1.1 Rappel : le cube de proportions de faciès

En premier lieu, il est important de revenir à la définition du cube de proportions afin de préciser les notations et le formalisme employés par la suite. Nous l'avons évoqué au chapitre 2, le cube de proportions, que Mallet et Shtuka (1997) appellent « membership function », porte deux types d'informations (Fig. 8.1) :

- d'un côté, il réalise une partition du milieu souterrain en un modèle discrétisé, dont les noeuds (ou cellules) indexés par un entier uvw , forment une grille recouvrant le

domaine géologique pour servir de support aux propriétés à modéliser (faciès, porosité, perméabilité, etc.) ;

- de l'autre, il exprime, en tout point α de l'espace géologique – en fait, sur chaque nœud uvw –, la probabilité d'occurrence de chaque faciès. Plus exactement, pour chaque cellule, il définit un vecteur¹ de probabilité d'appartenance à l'un des n_f lithofaciès pris en compte. Ce vecteur $\mathbf{p}(\alpha) = [p_1(\alpha), p_2(\alpha), \dots, p_{n_f}(\alpha)]^T$ comprend donc n_f composantes, qui renvoient respectivement à la probabilité $p_f(\alpha)$ du point situé en α (ou nœud uvw) d'appartenir au faciès f . Ces probabilités, encore appelées proportions, sont alors utilisées pour contraindre la simulation des faciès.

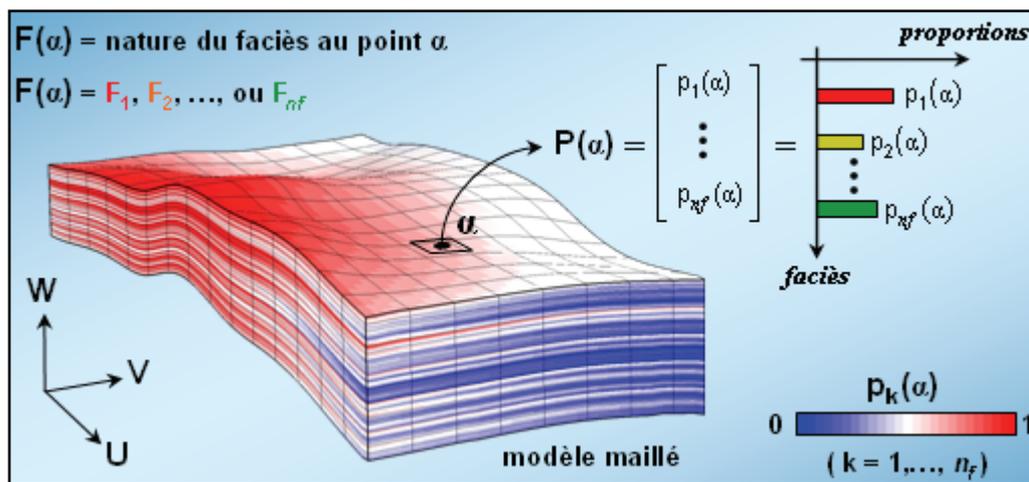


FIG. 8.1 : Définition du cube de proportions. C'est un modèle discrétisé de l'espace qui décrit en chaque point α (ou cellule d'indice uvw) une loi vectorielle de probabilité $\mathbf{p}(\alpha) = [p_1(\alpha), p_2(\alpha), \dots, p_{n_f}(\alpha)]^T$ à n_f composantes où n_f désigne le nombre de faciès considérés et $p_k(\alpha)$ la probabilité d'occurrence du faciès F_k .

La valeur de ces vecteurs de proportions est connue a priori en chaque point du sous-sol qui a été échantillonné au cours de la phase de prospection. Pour tous les autres, notre méthode propose de les simuler en accord avec l'incertitude sur le scénario géologique préalablement évaluée au moyen des $n_m(G)$ cas de base.

8.1.2 Des lois de probabilité ponctuelles

La méthode présentée ici repose sur l'hypothèse de l'existence d'une loi de probabilité à même de régir la population des cubes de proportions, une telle loi devant être compatible avec les données disponibles, et l'interprétation de l'expert. Si on adopte le

¹ Par abus de langage, nous parlerons même de « vecteur de proportions », on entend par là, un n_f -uplet dont les éléments correspondent à la collection des probabilités d'appartenance à chaque faciès. Car, on ne peut pas véritablement définir un espace vectorielle des proportions de faciès : la condition de sommation à l'unité de ses composantes empêche en effet la stabilité par un scalaire.

formalisme probabiliste, cela revient à interpréter chaque cube de proportions, noté \mathbf{CP} comme la réalisation d'un vecteur aléatoire dont les composantes se rapportent aux proportions de faciès en chaque point du modèle : $\mathbf{CP} = \{ \mathbf{p}(\alpha) / \alpha \in \Omega \}$, où Ω désigne l'ensemble des points du réservoir (Fig. 8.2). C'est un vecteur dont la simulation fait intervenir plusieurs variables aléatoires (ses composantes) considérées dans leur ensemble et non chacune isolément. Il est alors question d'identifier cette distribution en s'appuyant sur les modèles initiaux qui tous, découlent de l'un des $n_m(G)$ scénarios géologiques envisagés. Une fois cette loi déterminée, on génère un grand nombre de réalisations alternatives, à travers lesquelles il est possible de traduire l'incertitude, évaluée à l'aide des modèles de base, sur les proportions de faciès.

À première vue, il paraît laborieux de considérer dans son intégralité le cube de proportions comme un vecteur aléatoire. Son nombre de composantes en fait rapidement un objet de taille souvent rédhibitoire. Pour fixer un ordre de grandeur, un modèle géologique dépasse généralement le million de cellules, chiffre qu'il faut multiplier par le nombre de faciès pris en compte. Aussi, sommes-nous confrontés à des espaces de grande dimension où, même si théoriquement les calculs restent comparables à ceux effectués dans le plan, ils deviennent en pratique difficilement réalisables. L'estimation de la loi dont dériverait un tel vecteur \mathbf{CP} comme sa simulation, s'avèrent donc délicate.

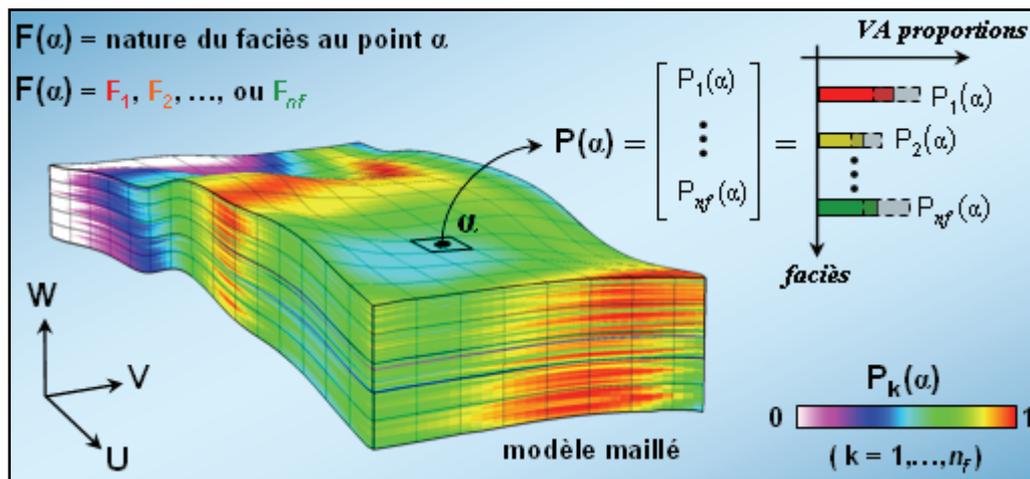


FIG. 8.2 : Définition de la notion de cube de proportions aléatoire. C'est un modèle discrétisé de l'espace représentant en chaque point α (ou cellule d'indice uvw) un vecteur aléatoire $\mathbf{P}(\alpha) = [P_1(\alpha), P_2(\alpha), \dots, P_{n_f}(\alpha)]^T$ à n_f composantes. n_f désigne le nombre de faciès considérés et $P_k(\alpha)$ la variable aléatoire qui décrit le comportement de la proportion de faciès F_k au point α .

C'est pourquoi, il est préférable d'étudier les proportions de faciès en chaque point – en fait, dans chacune des cellules d'indice uvw – du modèle maillé de réservoir. Ainsi, le même point α , considéré dans les $n_m(G)$ modèles de départ, est affecté de valeurs de proportion distinctes pour chacun des n_f faciès. Au point α , les données à traiter, regroupées dans un tableau numérique de dimension $n_m(G) \times n_f$ correspondent à un ensemble de $n_m(G)$

individus : les modèles, dont on connaît n_f variables quantitatives : les proportions de faciès. Cette matrice s'interprète comme la réalisation d'un échantillon de taille $n_m(G)$ (nombre de modèles) d'un vecteur aléatoire de dimension n_f (nombre de faciès). Par la suite, on suppose que le vecteur de proportions $\mathbf{p}(\alpha)$ attaché à chaque point α^1 du milieu souterrain, et qui définit sa probabilité d'appartenir à l'un des n_f faciès considéré, représente une réalisation de ce vecteur aléatoire $\mathbf{P}(\alpha) = [P_1(\alpha), P_2(\alpha), \dots, P_{n_f}(\alpha)]^\top$. Plus précisément, ses composantes $p_i(\alpha)$ – avec $i = 1, \dots, n_f$ – renvoient respectivement à une réalisation des n_f variables aléatoires $P_1(\alpha), \dots, P_{n_f}(\alpha)$. C'est en quelque sorte, l'équivalent stochastique de la fonction vectorielle de probabilité, ou « membership function », calculée de façon déterministe par Mallet et Shtuka (1997).

8.1.3 Un support géométrique commun

La simulation des vecteurs de proportions $\mathbf{P}(\alpha)$ en chaque point α nécessite, comme les techniques géostatistiques qu'elle met en œuvre (section 9.1), la définition préalable d'un support géométrique. En pratique, il s'agit d'une grille stratigraphique, ou grille curviligne, dont l'architecture interne, c'est-à-dire le maillage « vertical », est censé suivre la stratigraphie du milieu. Elle se compose de cellules déformées, qui sont indexées en utilisant un système de coordonnées curvilignes² (u, v, w) . Le paramétrage horizontal (u, v) décrit la surface des couches qui composent le réservoir. La coordonnée verticale (w) , normale aux deux premières, parcourt la succession des divers niveaux sédimentaires.

Au départ, nous disposons de $n_m(G)$ cubes de proportions établis selon les méthodes traditionnelles de géomodélisation : géostatistiques, procédé du type de Neptune, etc. À première vue, rien n'indique un paramétrage (u, v, w) commun de l'espace géologique entre les modèles : chacun présente a priori un support géométrique (maillage) qui lui est propre. Malgré tout, la majorité des techniques actuelles sépare la modélisation, ou l'extrapolation, des propriétés de la création du support de ces mêmes propriétés. En effet, la plupart des grilles sont construites en utilisant les surfaces, horizons ou failles, fournies par l'imagerie sismique. La stratification interne est alors générée indépendamment des propriétés à calculer en créant des surfaces intermédiaires entre les horizons observés de manière relativement simple. Par exemple, on peut définir des couches d'épaisseurs proportionnelles à l'épaisseur totale parallèles au toit ou au mur, ou bien calculées à partir d'une moyenne pondérée entre le toit et le mur. C'est pourquoi les $n_m(G)$ modèles de base réalisent d'ordinaire une partition du milieu similaire ; seule diffère la valeur des vecteurs de proportions modélisés, synonyme de l'incertitude sur le scénario géologique. C'est ainsi le cas des techniques géostatistiques à l'instar du krigeage ou de la simulation d'indicatrices de faciès, ou encore des procédés qui s'appuient sur l'interpolateur DSI, comme celui proposé par Labat (2004).

¹ Indépendamment du système de coordonnées, qu'il soit géographique en (x, y, z) , ou à caractère chronostratigraphique en (x_t, y_t, t) comme l'espace de Wheeler, l'espace paramétrique évoqué dans Neptune, ou bien l'espace Geochron proposé par Mallet (2004) et décrit par Moyen (2005).

² Elles sont aussi appelées coordonnées stratigraphiques.

En fait, la difficulté découle plutôt des modèles construits selon les méthodologies du genre de Neptune. Outre la détermination des propriétés, elles visent aussi à produire une véritable grille stratigraphique, dans laquelle les différentes couches de cellules sont cohérentes avec la stratigraphie réelle. Certaines des propriétés modélisées jouent alors un rôle majeure dans la définition de l'architecture interne du modèle. Par exemple, dans Neptune, la bathymétrie sert à contraindre l'épaisseur des différentes cellules : il y aura peu ou pas de dépôt dans les zones émergées, là où la bathymétrie est négative. Dans ce cas, l'épaisseur de sédiments simulée devrait être nulle (chapitre 3), ce qui oblige à une transformation de l'épaisseur, qui modifie la géométrie de la grille stratigraphique. Par conséquent, l'emploi de telles méthodologies conduit généralement à définir une architecture interne distincte, respectivement pour les $n_m(G)$ cubes de proportions initiaux.

Dans un premier temps, les modèles concurrents seront supposés avoir le même support géométrique : des nœuds ou des cellules communs à tous. Il s'agit par là de se placer dans le cas le plus fréquemment rencontré, car la plupart des modèles sont aujourd'hui construits selon les méthodes géostatistiques. Cependant, nous verrons au chapitre *Conclusions et Perspectives (2)*, qu'il est quand même possible de généraliser cette approche à des modèles de base – issus de Neptune par exemple –, à l'architecture interne différente. Nous pourrions alors intégrer aux simulations l'incertitude sur le pointé des corrélations, l'épaisseur des couches, etc. Dans les deux cas, avant de combiner les $n_m(G)$ modèles de départ, il faut d'abord préciser, en chaque noeud α du support géométrique, la loi probabiliste $\mathbf{P}(\alpha)$, dont dérivent les vecteurs de proportions $\mathbf{p}(\alpha)$, sans se préoccuper encore de la cohérence spatiale induite par les objets géologiques.

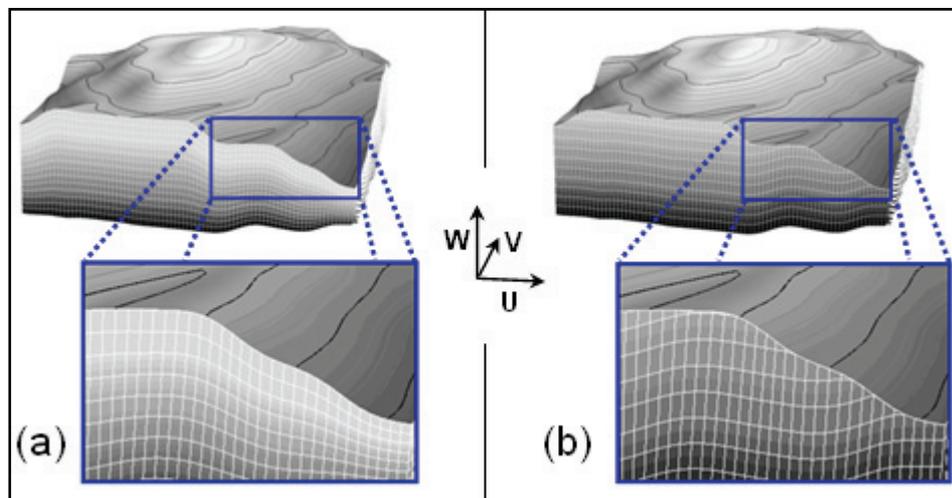


FIG. 8.3 : Un support géométrique commun à l'ensemble des $n_m(G)$ modèles de départ. Les modèles (a) et (b) présentent une architecture interne différente. La première grille définit un maillage vertical pseudo parallèle au toit du réservoir. La seconde, au mur. Elles ne définissent donc pas un paramétrage de l'espace géologique (u,v,w) commun, empêchant toute combinaison de leurs vecteurs de proportions respectifs. Ceux-ci correspondent à des propriétés discrétisés dans l'espace qui ont pour support les nœuds des grilles (a) et (b), ceux-ci coïncidant uniquement de manière singulière.

8.2 Choix des lois de probabilité ponctuelles

8.2.1 Des lois paramétriques

Connaissant, en chaque point α , un certain nombre de réalisations du vecteur aléatoire $\mathbf{P}(\alpha)$, il s'agit d'en inférer la distribution à partir de celles-ci. Pour cela, nous préférons caler les observations à une loi paramétrique prédéfinie plutôt que d'utiliser d'autres outils comme les histogrammes. Il faut en effet composer avec une situation paradoxale. D'une part, bien que l'on raisonne à l'échelle du point, nous restons face à des espaces de dimension élevée ; en général, on envisage de 3 à 10 faciès par modèles. Nous voilà alors confrontés à l'augmentation explosive du volume de données consécutif à l'ajout de dimensions supplémentaires dans un espace mathématique. Cet obstacle, connu sous le nom de « fléau de la dimension » [Bellman, 1961], ou *curse of dimensionality* en anglais, correspond au fait que les espaces de grande dimension sont vides. Par exemple, la sphère de rayon 0,75 de \mathbb{R}^{10} contient seulement 5 % des points du cube encadrant celle-ci, parallèle aux axes, et que l'on aurait rempli uniformément. Bellman donne ainsi l'exemple d'une centaine d'observations recouvrant l'intervalle $[0,1]$ à partir desquelles, il dresse un histogramme précisant ainsi leur loi expérimentale. Par contre, dans un espace de dimension 10 comme $[0,1]^{10}$, ces 100 observations représentent des points isolés dans un espace vide ; ils ne permettent donc aucune analyse statistique. Pour réaliser dans $[0,1]^{10}$ une couverture équivalente à celle des 100 points dans $[0,1]$, pas moins de 1020 observations se révèlent nécessaires, nombre concrètement impossible à atteindre dans notre cas. D'autre part, on dispose de $n_m(G)$ réalisations du vecteur $\mathbf{P}(\alpha)$ qui d'ordinaire, sont en nombre relativement faible. La pratique montre effectivement qu'en général très peu, voire un seul modèle géologique sont considérés : on est donc confronté à la difficulté d'évaluer des lois avec peu de répétition. Par conséquent, il faut tirer des inférences d'un nombre réduit d'expériences dans un espace de possibilités de dimension élevée, ce qui proscrit l'estimation d'une loi expérimentale au moyen d'un histogramme. En revanche, il est toujours envisageable d'estimer les paramètres d'un modèle de loi prédéfinie. On cherchera tout de même à rendre cet échantillon représentatif de la connaissance de l'expert, soit en affectant des poids aux $n_m(G)$ modèles de base, soit en évaluant des paramètres statistiques robustes (section 8.3.2). Il s'agit maintenant de déterminer, en chaque point α , une loi de probabilité qui vérifie un certain nombre de contraintes liées à la nature des vecteurs de proportions.

8.2.2 Contraintes de probabilité

En tout point α du milieu, chaque vecteur de proportions $\mathbf{p}(\alpha) = [p_1(\alpha), \dots, p_{nf}(\alpha)]^T$ est une réalisation d'un vecteur aléatoire $\mathbf{P}(\alpha)$. Ce dernier suit une loi de probabilité paramétrique. L'objectif consiste alors à déterminer celle-ci en chaque point α de telle sorte que ses réalisations $\mathbf{p}(\alpha)$ soient comparables à des distributions de probabilité : (a) chacune de leurs composantes doit par conséquent être comprise entre zéro et un (Éq. 8.1), et (b) la somme de leurs composantes respectives doit être égale à l'unité (Éq. 8.2) :

$$\forall \alpha, \forall f \in \{1, \dots, n_f\}, 0 \leq p_f(\alpha) \leq 1 \quad (\text{Éq. 8.1})$$

$$\forall \alpha, p_1(\alpha) + \dots + p_f(\alpha) + \dots + p_{n_f}(\alpha) = 1 \quad (\text{Éq. 8.2})$$

Géométriquement, pour tout nœud α du réservoir, les $n_m(G)$ vecteurs $\mathbf{p}(\alpha)$ déduits de l'ensemble des modèles initiaux, peuvent être représentés par un nuage de points dans l'espace à n_f dimensions \mathbb{R}^{n_f} ; leurs composantes sont les proportions de faciès¹. Les contraintes (8.1) et (8.2) s'interprètent donc comme la nécessité de simuler des points sur le simplexe $S(n_f) = \{(x_1, \dots, x_{n_f}) / x_i \geq 0 (i = 1, \dots, n_f), x_1 + \dots + x_{n_f} = 1\}$. Comme le montre la figure 8.4, pour deux faciès ($n_f = 2$), ce simplexe correspond au segment reliant les points I(1,0) et J(0,1)¹ ; pour trois faciès ($n_f = 3$), c'est le triangle IJK où I (1,0,0), J (0,1,0) et K (0,0,1)¹. Sachant ces contraintes et leur éventuelle interprétation géométrique, il est à présent envisageable de définir une loi $\mathbf{P}(\alpha)$ en chaque point du réservoir. Pour cela, nous recourons à un certain nombre de distributions vectorielles, qui produisent des résultats plus ou moins satisfaisants s'agissant du respect des relations (8.1) et (8.2). La difficulté majeure de leur modélisation réside en effet dans la non indépendance des composantes des vecteurs de proportions due à la relation (8.2). Le problème a été considéré sous deux angles différents : d'abord, de manière géométrique, en se basant sur les méthodes exploratoires utilisées en analyse de données, puis selon une approche plus fonctionnelle, en examinant les lois qui satisfont simplement à ces contraintes.

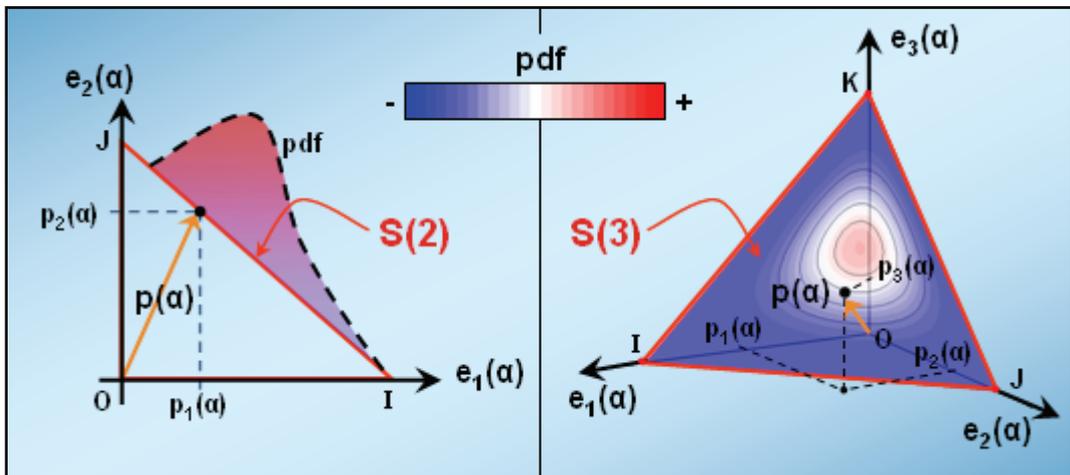


FIG. 8.4 : Exemples de simplexes sur lesquels sont simulés les vecteurs de proportions. À gauche, en deux dimensions, il faut simuler des vecteurs de proportions $\mathbf{p}(\alpha) = [p_1(\alpha), p_2(\alpha)]^T$ dont les extrémités se situent sur le segment [IJ]. À droite, en trois dimensions, les vecteurs de proportion ayant pour origine O et notés $\mathbf{p}(\alpha) = [p_1(\alpha), p_2(\alpha), p_3(\alpha)]^T$, pointent sur le triangle défini par les points I, J et K. Dans les deux cas, les vecteurs sont simulés selon des lois de probabilité qui restent à définir (section 8.2.3). Celles-ci sont aussi représentées sur les deux simplexes (pdf de bleue à rouge).

¹ Dans la base canonique respectivement de \mathbb{R}^{n_f} , \mathbb{R}^2 et \mathbb{R}^3

8.2.3 Petit historique des lois testées

Analyse en composantes principales

L'analyse en composantes principales (ACP), d'abord développée par Pearson (1901) pour deux variables, a été étendue par Hotelling (1933), à un nombre quelconque de variables. Les ouvrages de Jackson (1991) ou de Lebart et *al.* (1995) offrent un panorama détaillé de la méthode. À l'instar d'autres techniques d'analyse de données, elle vise entre autres à remplacer les données initiales par un nombre réduit de nouvelles variables sans perte notable d'information. Elle présente la particularité de définir les nouvelles variables comme des combinaisons linéaires de celles de départ, et qui en plus, sont non corrélées linéairement. Si l'on représente les données initiales à l'aide d'un nuage de points, ce problème revient à chercher les droites, les plans et de manière plus générale les variétés linéaires proches du nuage initial.

Comme évoqué précédemment (section 8.2.2), en chaque nœud α du milieu, les $n_m(G)$ vecteurs $\mathbf{p}(\alpha)$ déduits des modèles initiaux, dessinent un nuage de points dans \mathbb{R}^{n_f} . Pour plus de clarté, on notera par $\mathbf{p}(k, \alpha) = [p_1(k, \alpha), \dots, p_{n_f}(k, \alpha)]^T$ le vecteur de proportions au point α relatif au modèle de base d'indice k . Dans notre cas, l'ACP permet de construire une nouvelle base orthonormée $[e_1^*(\alpha), \dots, e_{n_f}^*(\alpha)]$, obtenue par rotation de la base canonique de \mathbb{R}^{n_f} autour de son origine O , et qui passe au plus près du nuage initial (Fig. 8.5-a).

La projection des différents points du nuage sur cette nouvelle base définit, pour chaque composante sur l'axe $(O, e_i^*(\alpha)) - i = 1, \dots, n_f -$, un échantillon d'une variable aléatoire $V_i(\alpha)$. Par construction, celles-ci ne sont pas corrélées entre elles. Dès lors, on suppose qu'il s'agit de lois normales en vue de se ramener à des variables indépendantes¹. Leurs paramètres – moyenne et écart type – sont estimés à partir des composantes des vecteurs de proportions $\mathbf{p}(k, \alpha)$, où $k = 1, \dots, n_m(G)$, sur $[e_1^*(\alpha), \dots, e_{n_f}^*(\alpha)]$ (Fig. 8.5-b).

À présent, connaissant les lois $V_i(\alpha) - i = 1, \dots, n_f -$, autrement dit les composantes du vecteur probabiliste $\mathbf{P}(\alpha)$ dans la $[e_1^*(\alpha), \dots, e_{n_f}^*(\alpha)]$, il suffit de tirer un nombre aléatoire $v_i(\alpha)$, avec $i = 1, \dots, n_f$, dans chacune de ces n_f distributions normales indépendantes (voir Fig. 8.3-c). La réalisation $\mathbf{p}(\alpha)$ du vecteur de proportions correspondant s'obtient alors en projetant le vecteur $[v_1(\alpha), \dots, v_{n_f}(\alpha)]^T$ dans la base canonique de \mathbb{R}^{n_f} (Fig. 8.5-d).

On peut remarquer que l'ensemble du procédé correspond à un algorithme connu [Govaert *et al.*, 2004], utilisé pour simuler un vecteur gaussien, c'est-à-dire une variable aléatoire multidimensionnelle, dont les composantes sont des lois normales. En effet, certains voient dans l'ACP l'équivalent de la recherche des axes principaux de l'ellipsoïde qui caractérise les distributions normales multidimensionnelles. En d'autres termes, cela revient à supposer dès le départ que $\mathbf{P}(\alpha)$ est un vecteur aléatoire gaussien. Il s'agit alors d'en estimer, à partir des données initiales, les paramètres – vecteur moyen et matrice de variance –, et d'en réaliser la simulation selon cette distribution.

¹ Considérons un ensemble de n variables aléatoires (X_1, X_2, \dots, X_n) non corrélées, en faisant l'hypothèse supplémentaire qu'elles suivent toutes des lois normales, celles-ci sont indépendantes.

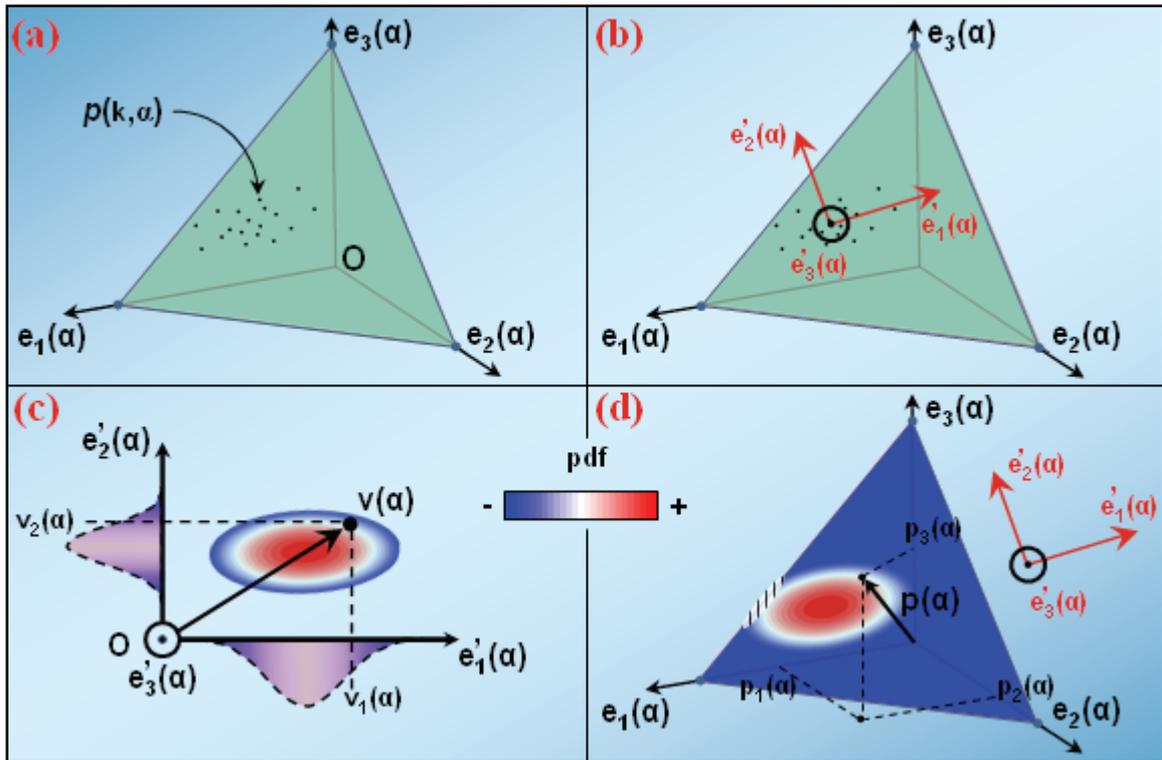


FIG. 8.5 : Simulation d'un vecteur de proportions $\mathbf{p}(\alpha)$ par ACP (ou vecteur gaussien). En (a) le nuage de point correspond aux vecteurs de proportions des modèles initiaux d'indice k au point α . En (b), l'ACP détermine les axes principaux $[\mathbf{e}'_1(\alpha), \dots, \mathbf{e}'_3(\alpha)]$ du nuage. Les points $\mathbf{p}(k, \alpha)$ sont projetés sur cette base, permettant ainsi d'estimer les paramètres des lois normales supposées régir leurs composantes $V_i(\alpha)$, puis d'en simuler une réalisation $\mathbf{v}(\alpha)$ en (c). Puis, ce vecteur est à nouveau projeté en (d) dans le repère initial $[O, \mathbf{e}_1(\alpha), \dots, \mathbf{e}_3(\alpha)]$ pour définir une réalisation $\mathbf{p}(\alpha)$ du vecteur de proportions.

Au final, les résultats obtenus sont globalement bons, même si des corrections restent nécessaires. Si la relation (8.2) est toujours respectée, on risque par contre dans certaines configurations, de simuler des vecteurs dont quelques composantes sont négatives ou supérieures à un (Fig. 8.5-d, partie hachurée). Pour y remédier, une première approche consiste à rejeter les réalisations qui ne satisfont pas aux contraintes de probabilité jusqu'à obtenir un vecteur qui les respectent. Bien que localement, ce soit une solution convenable, elle n'est plus tenable à l'échelle du réservoir. Vu le nombre de points, ou mailles, du modèle, il est presque sûr de simuler au moins un vecteur de proportions incorrect. En supposant que c'est au point α_i , on ne peut pas rejeter $\mathbf{p}(\alpha_i)$ pour une nouvelle simulation sans remettre en question les vecteurs simulés aux points entourant α_i . Le risque est de perdre la continuité spatiale des modèles, car les vecteurs de proportions ne sont pas indépendants les uns des autres (section 8.3.3 et chapitre 9). Par conséquent, le tirage d'une nouvelle réalisation de $\mathbf{P}(\alpha_i)$ oblige à simuler le cube de proportions dans son intégralité, simulation qui génère presque sûrement un vecteur de proportions incorrect.

Une solution alternative cherche plutôt à corriger les réalisations qui, comme $p(\alpha_i)$, ne respectent pas la relation (8.1). Elle ramène d'abord les composantes négatives ou positives respectivement à zéro ou à un. Elle applique alors un coefficient multiplicateur λ à chaque proportion de sorte que leur somme soit égale à l'unité :

1. pour tout faciès f ,
 - si $p_f(\alpha_i) < 0$ alors $p_f(\alpha_i) = 0$
 - si $p_f(\alpha_i) > 1$ alors $p_f(\alpha_i) = 1$
2. calcul du terme correctif, $\lambda^{-1} = p_1(\alpha_i) + \dots + p_{n_f}(\alpha_i)$
3. pour tout faciès f , $p_f(\alpha_i) = \lambda \times p_f(\alpha_i)$

Géométriquement, une telle correction a tendance à ramener les vecteurs de proportions sur les bords du simplexe (Fig. 8.5-d, partie hachurée), risquant par là même de surestimer la probabilité de cette zone de points. En outre, comme les modifications sont locales, on retrouve les problèmes liés au respect de la cohérence spatiale des cubes de proportions générés.

Transformation préalable des données

Pour y remédier, une autre approche consiste à passer d'abord, des valeurs de type « probabilité » $p_f(\alpha)$ à des nombres réels de manière à pouvoir utiliser les lois normales, qui ne sont pas forcément adaptées aux intervalles bornés. Ces transformations sont d'un usage courant en statistiques, notamment lorsque l'on fait de la régression logistique¹. L'idée est de faire intervenir une fonction réelle monotone « g », dite de lien, qui modifie le domaine de variation [0,1] des variables aléatoires $P_f(\alpha)$ en l'intervalle $]-\infty ; +\infty[$. De nombreuses fonctions existent pour remplir ce rôle, les plus couramment utilisées [Besse, 2003] étant définie en tout point α , pour chaque faciès f par les applications :

- (1) **probit**, g est la fonction inverse de la fonction de répartition Φ d'une loi normale centrée et réduite, $g[p_f(\alpha)] = \text{probit}[p_f(\alpha)] = \Phi^{-1}[p_f(\alpha)]$;
- (2) **log-log**, g est définie $g[p_f(\alpha)] = \ln[-\ln(1-p_f(\alpha))]$, où $x \rightarrow \ln(x)$ correspond au logarithme népérien ;
- (3) **logit**, c'est l'application $g[p_f(\alpha)] = \text{logit}[p_f(\alpha)] = \ln\left[\frac{p_f(\alpha)}{1-p_f(\alpha)}\right]$;

¹ **Régression logistique** : C'est le pendant de la régression linéaire pour les variables catégorielles. Il s'agit d'une technique statistique dont l'objectif est de construire, à partir d'une série de variables explicatives continues, un modèle permettant de prédire les valeurs prises par une variable catégorielle. Elle est largement répandue dans de nombreux domaines : épidémiologie, biologie, assurance, banque, etc.

Ces transformations sont assez semblables, en particulier les applications probit et logit. Toutefois, pour des commodités de calcul, l'expression du probit n'étant pas explicite, on préfère souvent le logit. Une raison supplémentaire à son utilisation est l'interprétation du rapport $p_f(\alpha)/(1-p_f(\alpha))$ en termes de chances relatives – *odds* en anglais – de simuler le faciès f plutôt que l'un des (n_f-1) autres. En chaque nœud α , pour effectuer la simulation d'un vecteur de proportions, il est possible de recourir à un vecteur gaussien $\mathbf{V}(\alpha)$ obtenu de la manière suivante :

1. on transforme les données issues des modèles initiaux (indexés par $k, k=1, \dots, n_m(G)$) $\mathbf{p}(k, \alpha) = [p_1(k, \alpha), \dots, p_{n_f}(k, \alpha)]^T$ en une série de vecteurs que l'on notera ici abusivement, $\text{logit}[\mathbf{p}(k, \alpha)] = [\text{logit}[p_1(k, \alpha)], \dots, \text{logit}[p_{n_f}(k, \alpha)]]^T$;
2. on estime, à partir de ce nouveau nuage de points, les paramètres : moyenne et matrice de variance, du vecteur gaussien $\mathbf{V}(\alpha) = [\mathbf{V}_1(\alpha), \dots, \mathbf{V}_{n_f}(\alpha)]^T$ qui correspond à ces données.

Pour générer un vecteur de proportions $\mathbf{p}(\alpha)$, on simule d'abord une réalisation du vecteur aléatoire $\mathbf{V}(\alpha)$ estimé précédemment : on la note $\mathbf{v}(\alpha) = [v_1(\alpha), \dots, v_{n_f}(\alpha)]^T$. Le vecteur de proportions correspondant s'obtient naturellement en appliquant à chaque composante du vecteur $\mathbf{v}(\alpha)$, la fonction inverse du logit qui se définit pour tout réel x par la relation $g^{-1}[x] = e^x/(1+e^x)$, $x \rightarrow e^x$ étant la fonction exponentielle (notée aussi $\exp(x)$) :

$$\text{Ainsi, pour tout faciès } f, p_f(\alpha) = g^{-1}[v_f(\alpha)] = \frac{\exp[v_f(\alpha)]}{\exp[v_f(\alpha)] + 1}$$

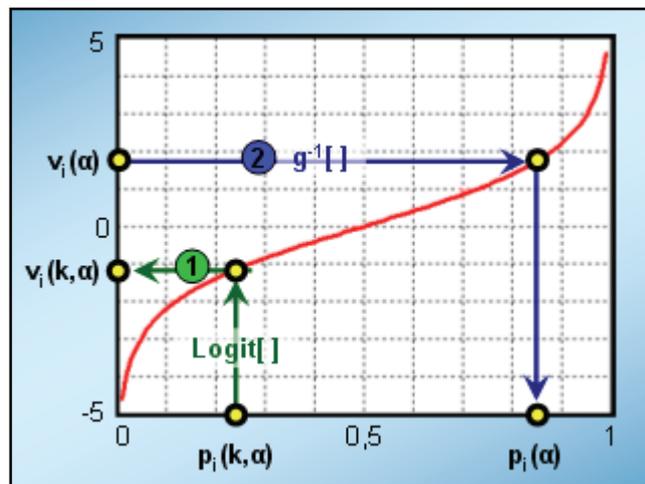


FIG. 8.6 : Simulation d'un vecteur de proportions $\mathbf{p}(\alpha)$ avec transformation préalable des données. (1) Les données de départ $p_i(k, \alpha)$ – i et k renvoyant respectivement au faciès et au modèle de base – sont transformées selon une fonction $\text{Logit}[\cdot]$. Le calcul des paramètres des lois auxquelles obéit le vecteur $\mathbf{V}(\alpha)$ est identique à celui évoqué pour les vecteurs gaussiens. Seules les données de départ changent : ce sont les $v_i(k, \alpha)$. (2) La transformation selon $g^{-1}[\cdot]$ des composantes $v_i(\alpha)$ d'une réalisation $\mathbf{v}(\alpha)$ de $\mathbf{V}(\alpha)$ correspond à $p_i(\alpha)$, composante d'une réalisation d'un vecteur de proportions $\mathbf{p}(\alpha)$.

Les résultats obtenus respectent ainsi à la fois les contraintes de probabilités, et les paramètres de la loi estimée sur la base de l'échantillon de vecteurs de proportions.

Lois intrinsèques de $S(n_f)$

Ces procédés d'analyse de données, qui en somme reposent sur des lois normales multidimensionnelles, nécessitent relativement peu de temps de calculs. Elles offrent des résultats corrects en ce qui concerne le respect des contraintes de probabilité, même s'ils ne sont pas aussi bons que ceux obtenus à partir d'autres distributions dont dérive le vecteur aléatoire $\mathbf{P}(\alpha)$. Car d'un autre côté, il existe en effet des lois probabilistes, que l'on notera $\pi : \mathbb{R}^{n_f} \rightarrow [0,1]$, ayant pour support en dimension n_f , le simplexe $S(n_f)$, et qui intrinsèquement respectent les relations (8.1) et (8.2) :

- $\oint_{S(n_f)} \pi(\mathbf{p}).d\mathbf{p} = 1$, où $\mathbf{p} = [p_1, \dots, p_{n_f}]^T$ est un vecteur de \mathbb{R}^{n_f} ;
- $\pi(\mathbf{p}) = 0$ si $\mathbf{p} \notin S(n_f)$

Leurs paramètres sont déterminés à partir des données extraites des différents scénarios géologiques envisagés. Parmi ces lois de probabilité, nous pouvons citer les distributions « logistiques normales » proposées par Aitchinson (1997), ou encore les lois de Dirichlet – extension des lois Bêta à plus d'une dimension –, auxquelles nous prêterons particulièrement attention. Nous ne prétendons aucunement à l'exhaustivité, même si Aitchinson et Shen (1980) dressent le constat que, pour la majorité des statisticiens, ce sont là, les seules distributions à même de respecter les contraintes (8.1) et (8.2). Aussi, un nombre important de procédés statistiques, qui utilisent des variables de type « probabilité », se réfèrent-ils à ces lois. Par exemple, pour concilier des évaluations subjectives d'un modèle de probabilité, Lindley *et al.* (1979), appliquent les lois logistiques normales sur les chances relatives de succès pour décrire ces estimations. Lindley (1964), Bloch et Watson (1967) ou Leonard (1973) se servent de ces deux familles de distributions en tant que lois a priori pour l'inférence bayésienne de variables multinomiales.

Aitchinson et Begg (1976) suggèrent que les lois logistiques normales sont plus riches que celles de Dirichlet, car elles offrent davantage de degré de liberté (paramètres). Selon eux, il est ainsi possible d'approcher, à partir d'une loi logistique normale, n'importe quelle loi de Dirichlet. Cependant, plusieurs raisons, tant théoriques que pratiques, font préférer cette dernière famille de distributions pour la simulation des vecteurs de proportions en chaque point α .

Retour sur les lois de Dirichlet

Le recours à cette famille de lois paraît naturel pour modéliser l'incertitude sur les proportions, car d'une part ses réalisations respectent les contraintes de probabilité (8.1) et (8.2). D'autre part, ce formalisme est parfaitement adapté à la modélisation des incertitudes sur les proportions de faciès : ces lois ont été utilisées dans ce sens dans les travaux de

Haas et al. (2002), et de Biver et al. (2002). Ils proposent d'inférer une loi de probabilité relative aux proportions de chacun des n_f faciès estimée à partir des données puits. Pour cela, ils supposent que le vecteur des occurrences de faciès $N_b = [N_b(1), \dots, N_b(n_f)]^T$ observées le long des puits, pour un vecteur donné de proportions de faciès $\mathbf{p} = [p_1, \dots, p_{n_f}]^T$ suit une loi multinomiale $\mathbb{P}(N_b/\mathbf{p})$. En pratique, les proportions de chaque faciès, c'est-à-dire le vecteur \mathbf{p} , ne sont pas connues. Cependant, il est possible de déterminer leur loi de probabilité conditionnelle $\mathbb{P}(\mathbf{p}/N_b)$, connaissant le vecteur N_b d'occurrence de faciès observé aux puits, par inversion bayésienne (Fig. 8.7) :

$$\mathbb{P}(\mathbf{p}/N_b) = \frac{\mathbb{P}(N_b/\mathbf{p}) \times \mathbb{P}(\mathbf{p})}{\int_{S(n_f)} \mathbb{P}(N_b/\mathbf{p}) \times \mathbb{P}(\mathbf{p}) . d\mathbf{p}} \quad (\text{Éq. 8.3})$$

En l'absence d'information supplémentaire, la loi a priori des proportions de faciès $\mathbb{P}(\mathbf{p})$ est une loi uniforme. Ce qui conduit naturellement à inférer pour loi a posteriori des proportions de faciès, une distribution de Dirichlet, qui de façon générale, s'écrit dans un espace à n_f dimensions avec n_f paramètres $[c_1, \dots, c_{n_f}]$:

pour tout vecteur $\mathbf{p} = [p_1, \dots, p_{n_f}]^T$ de $S(n_f)$,

$$\pi(\mathbf{p}) = \frac{\Gamma(\mathbf{c})}{\Gamma(c_1) \times \dots \times \Gamma(c_{n_f})} p_1^{c_1-1} \times \dots \times p_{n_f}^{c_{n_f}-1} \quad (\text{Éq. 8.4})$$

avec $c = c_1 + c_2 + \dots + c_{n_f}$, et $x \rightarrow \Gamma(x)$ est la fonction gamma¹ (section 9.1)

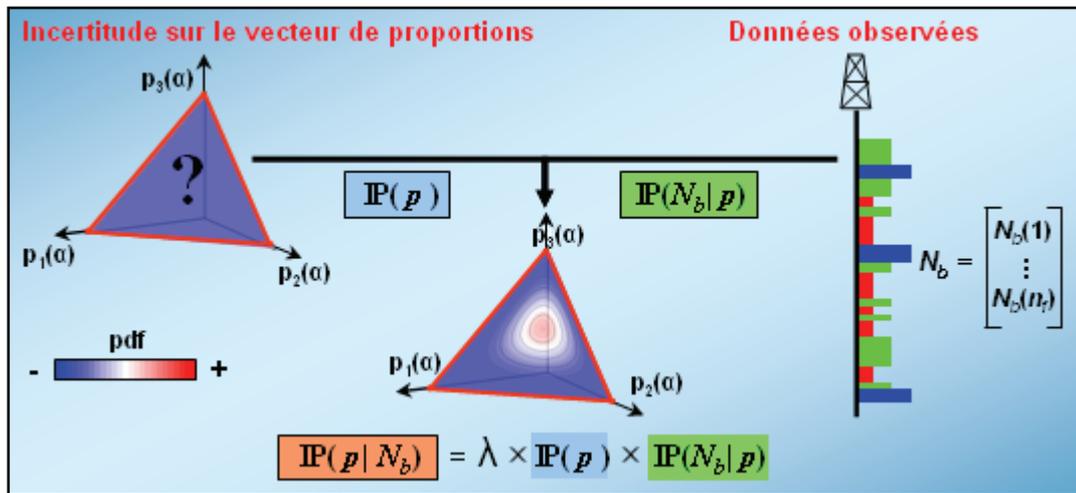


FIG. 8.7 : Loi de Dirichlet et proportions. La combinaison, dans un cadre bayésien, d'une loi a priori uniforme pour les proportions (à gauche) et des occurrences de faciès aux puits (à droite), conduit à inférer pour loi a posteriori des proportions une loi de Dirichlet.

¹ Bien connue des mathématiciens, la fonction gamma peut être considérée comme une extension au nombre réels de la factorielle, c'est-à-dire par abus de notations $\Gamma(x) = \Gamma(x-1) !$ pour tout réel positif x .

Remarquons au passage que l'incertitude évaluée précédemment est exclusivement de nature statistique sur les données disponibles (puits). Ces auteurs abordent la question de l'indépendance et de la représentativité des observations aux puits. Leur approche considère d'ailleurs un scénario géologique unique, et n'évalue donc pas l'incertitude sur ce paramètre. De façon complémentaire, la méthode que nous proposons, se sert des modèles de Dirichlet pour combiner des cubes de proportions correspondants à des scénarios géologiques différents, et définit ainsi l'incertitude géologique, qui en dérive, sur les proportions. Un tel formalisme permet en plus de combiner ces deux lois de Dirichlet issues respectivement d'analyses différentes (chapitre 11). Dans le reste du mémoire, nous ne considérerons plus que les lois de Dirichlet pour décrire les autres étapes de la méthode de simulation proposée. Néanmoins, il est important de remarquer que la démarche appliquée par la suite avec des lois de Dirichlet, peut très bien s'effectuer en utilisant l'une ou l'autre des lois précédemment citées (section 8.2.3). C'est simplement que leurs résultats sont jugés moins bons. Par conséquent, on suppose qu'en chaque noeud α du réservoir, le vecteur de proportions $\mathbf{P}(\alpha)$ suit sa propre loi vectorielle de Dirichlet, loi qu'il s'agit désormais de caractériser.

8.3 Caractérisation des distributions ponctuelles

8.3.1 Estimation des paramètres des lois $\mathbf{P}(\alpha)$

A ce stade, le bilan de l'information et de la connaissance disponible peut être le suivant, en chaque point α du réservoir maillé :

- (1) le vecteur de proportions $\mathbf{P}(\alpha)$ est supposé suivre une loi de Dirichlet définie par les n_f paramètres inconnus $[c_1(\alpha), \dots, c_{n_f}(\alpha)]$ et $c(\alpha) = c_1(\alpha) + \dots + c_{n_f}(\alpha)$;
- (2) les $n_m(G)$ modèles initiaux fournissent un échantillon de taille $n_m(G)$ de cette distribution vectorielle ; si on considère le même point α dans les différents modèles de base, il est affecté de valeurs de proportion distinctes pour le i ème faciès.

De fait, en chaque point, l'évaluation des paramètres de $\mathbf{P}(\alpha)$ passe d'abord, par l'estimation de ses premiers moments statistiques : moyenne, variance et covariance de ses composantes, sur la base des valeurs de proportions prises en ce point par les $n_m(G)$ modèles initiaux. Il devient dès lors possible d'utiliser les relations qui relient ces moments aux paramètres $[c_1(\alpha), \dots, c_{n_f}(\alpha)]$ de la loi (Fig. 8.8) :

quel que soit le faciès f ,

$$E[\mathbf{P}_f(\alpha)] = \frac{c_f(\alpha)}{c(\alpha)} \quad (\text{Éq. 8.5})$$

$$\sigma^2[\mathbf{P}_f(\alpha)] = \frac{c_f(\alpha) \times (c_f(\alpha) + 1)}{c(\alpha) \times (c(\alpha) + 1)} \quad (\text{Éq. 8.6})$$

et pour tout faciès f' différent de f ,

$$\sigma [P_f(\alpha) ; P_{f'}(\alpha)] = \frac{c_f(\alpha) \times c_{f'}(\alpha)}{c(\alpha) \times (c(\alpha) + 1)} \quad (\text{Éq. 8.7})$$

$E[Z]$ et $\sigma^2[Z]$ définissant respectivement l'espérance mathématique (la moyenne arithmétique), et la variance de la variable aléatoire Z ; $\sigma[Z;Z']$ désignant sa covariance avec la variable aléatoire Z' .

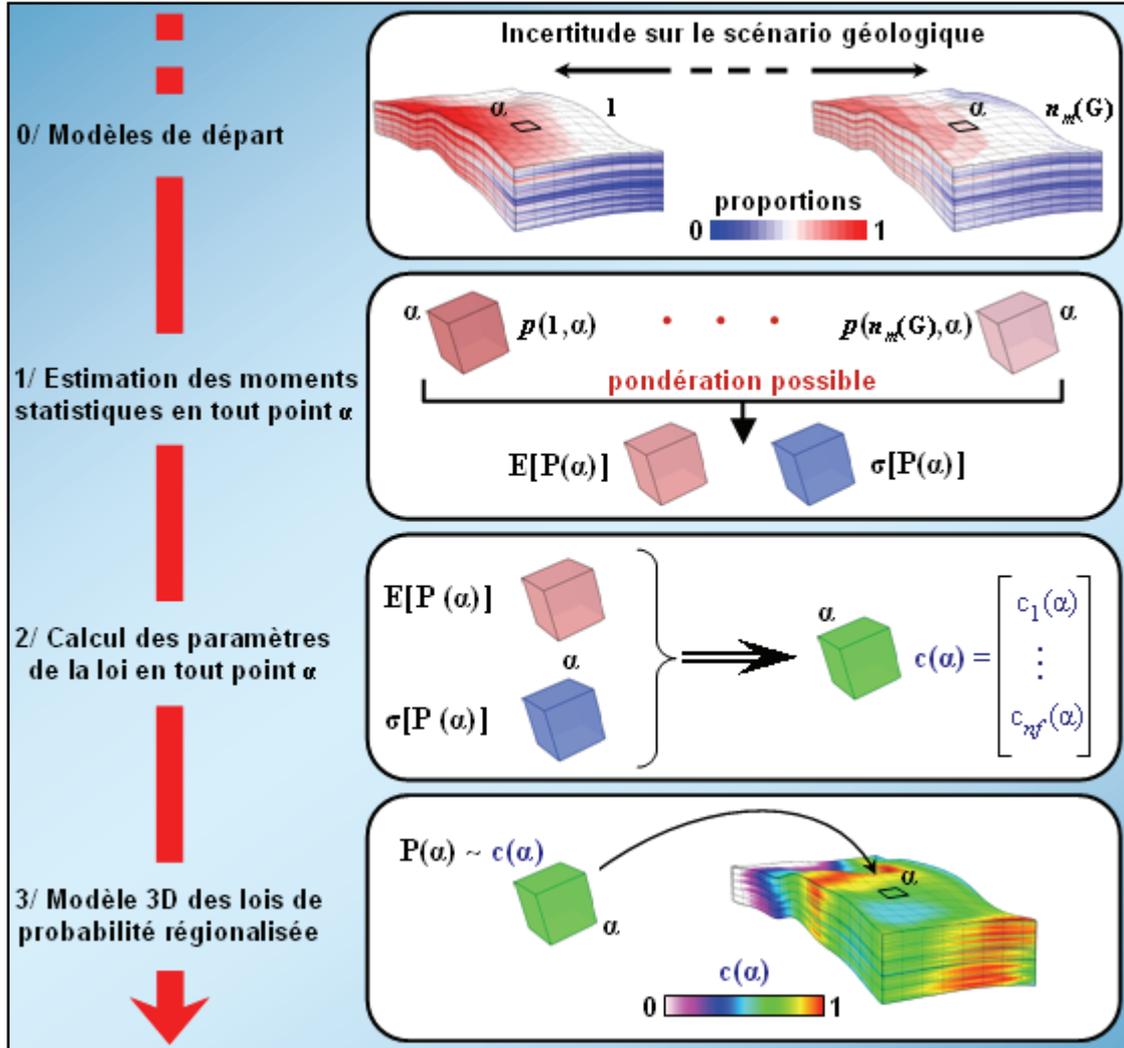


FIG. 8.8 : Estimation des paramètres de la loi de probabilité au point α . (0) $n_m(G)$ modèles de base traduisent l'incertitude sur le scénario. (1) En tout point α , les premiers moments sont estimés avec les vecteurs de proportions $p(k, \alpha)$ – avec k indice des modèles de base. (2) De la combinaison de ces moments dérivent le vecteur de paramètres $c(\alpha)$ de la loi de Dirichlet régissant la distribution des vecteurs de proportions. (3) Ces paramètres $c(\alpha)$ déterminés en tout point α du milieu définissent l'ensemble des lois de Dirichlet régionalisées dans tout le modèle.

8.3.2 Retour sur les scénarios de base

L'évaluation des paramètres $[c_1(\alpha), \dots, c_{nf}(\alpha)]$ des lois de Dirichlet régionalisées nécessite de fournir au préalable une série de $n_m(G)$ modèles initiaux. Se pose quand même la question de leur représentativité. En pratique, les experts cherchent plutôt à fournir des cas « extrêmes ». En conséquence, on aura tendance à surestimer l'incertitude ; celle-ci se traduit en effet sur la variance des proportions. D'un autre côté, le nombre limité de modèles disponibles oblige souvent à apprécier les distributions avec peu de répétition. C'est l'exercice qui l'impose, car d'ordinaire en analyse de risques, seul trois cas sont envisagés : pessimiste, vraisemblable et optimiste. Néanmoins, il ne faut pas oublier l'origine des modèles de départ. Ils décrivent la connaissance et les hypothèses du sédimentologue établies à partir des données disponibles. L'écart entre scénarios peut donc s'interpréter comme son ignorance à l'égard de celui-ci.

Des alternatives existent pour préciser le calcul des paramètres de chaque distribution en tout point α . Une première solution consiste à pondérer les scénarios géologiques dans le calcul des premiers moments statistiques de la loi. En affectant un poids ω_k à chacun des modèles de base, k renvoyant à l'indice de celui-ci, si on considère par exemple, la moyenne des proportions, son estimation devient pour chaque faciès f :

$$E[P_f(\alpha)] = \sum_{k=1}^{n_m(G)} \omega_k \times p_f(k, \alpha) \quad (\text{Éq. 8.8})$$

où pour le modèle d'indice k , $p_f(k, \alpha)$ désigne la proportion du faciès f au point α , c'est-à-dire $p_f(\alpha) = p_f(k, \alpha)$; la somme des $n_m(G)$ poids étant égale à un.

Là encore, la difficulté ressort de la définition des poids notés ω_k , pour lesquels diverses interprétations et évaluations restent envisageables [Buslik, 1993 ; Winkler, 1993] notamment par avis d'expert(s) (section 4.1.2). Ces derniers traduisent de cette façon la confiance ou la vraisemblance qu'ils accordent à chacun des scénarios initiaux : en s'appuyant sur leur analyse, ils affectent un poids d'autant plus faible aux scénarios qu'ils jugent possibles mais peu probables. Il s'agit bien entendu d'un exercice délicat, qui demande un recul supplémentaire sur les hypothèses pour rendre l'échantillon plus représentatif. Car, le choix d'un modèle géologique reste un exercice fondamentalement subjectif. Cependant, les études et techniques actuelles conduisent surtout à négliger l'incertitude sur la géologie. Au final, ce protocole rend sa place à une part d'incertitude non appréhendée jusqu'ici.

Une solution alternative s'appuie sur l'évaluation probabiliste du meilleur estimateur des paramètres des lois en chaque point. Jusqu'à présent, on utilise la moyenne et l'écart-type. Aussi, en ce qui concerne la question des modèles « extrêmes » évoquée précédemment, un meilleur estimateur de la variabilité serait la distance interquartile plutôt que l'écart-type. D'autre part, à ce stade, il n'est pas certain que la loi estimée soit la bonne ni même qu'il existe une « bonne loi ». Il n'y a d'ailleurs pas d'autres moyens de l'appréhender. On se retrouve finalement dans la même situation qu'en géostatistique, où

pour Matheron, la « bonne loi » est celle qui donne de bons résultats ». En fait, il faut plutôt voir cette méthode comme une technique de mélange qui utilise les notions de probabilités pour aller chercher des modèles intermédiaires. On insiste sur le fait que cette démarche constitue « un choix méthodologique ». Cette décision nous permet d'utiliser les outils géostatistiques pour passer d'hypothèses discrètes – les scénarios géologiques, à un phénomène continu, et expérimenter des situations nouvelles à travers la simulation des vecteurs de proportions régionalisés $\mathbf{P}(\alpha)$.

8.3.3 Simulations de vecteurs de proportions régionalisées

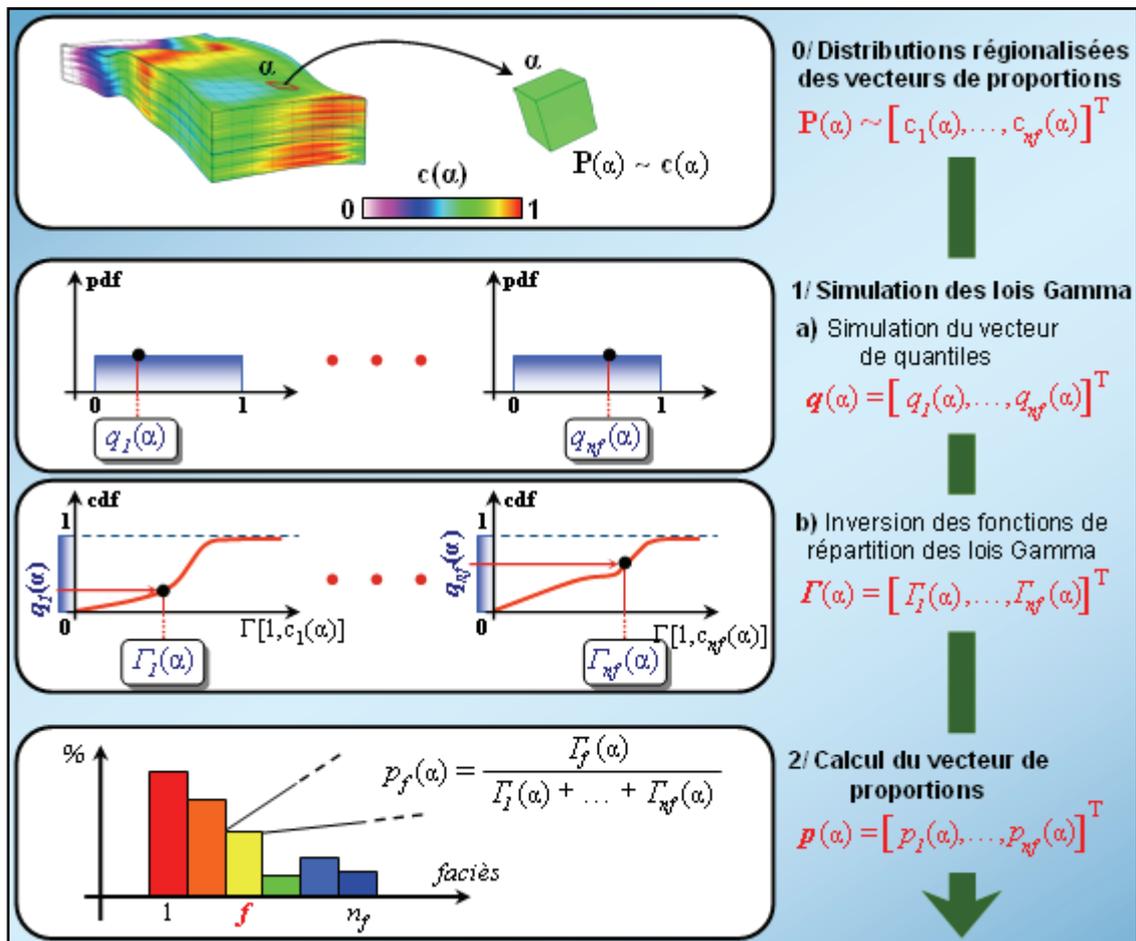


FIG. 8.9 : Simulation d'un vecteur de proportions au point α . (0) Au point α , les vecteurs probabilistes $\mathbf{P}(\alpha)$ suivent une loi de Dirichlet de paramètres $\mathbf{c}(\alpha)$. (1) Une réalisation $\mathbf{p}(\alpha)$ de $\mathbf{P}(\alpha)$ requiert de simuler un jeu de n_f variables intermédiaires selon des lois Gamma $\Gamma_1(\alpha), \dots, \Gamma_{n_f}(\alpha)$, dont les paramètres procèdent de la loi de Dirichlet. (a) Pour cela, n_f valeurs sont tirées indépendamment dans une loi uniforme de support $[0;1]$. (b) Ces nombres, appelés quantiles, permettent l'inversion des fonctions de répartitions des lois Gamma, dont les réalisations correspondent aux composantes d'un vecteur $\mathbf{\Gamma}(\alpha)$. (2) Le vecteur de proportions $\mathbf{p}(\alpha)$ est proportionnelle à ce dernier.

La création d'un nouveau cube de proportions nécessite de tirer aléatoirement un vecteur de proportions dans chaque distribution de Dirichlet régionalisée qui a préalablement été estimée. Cette section revient sur la méthode employée pour effectuer les tirages de manière à respecter la loi de Dirichlet régionalisée, dont dérive chaque vecteur aléatoire (Fig. 8.9).

Au point α , la génération du vecteur probabiliste $\mathbf{P}(\alpha)$ selon une distribution de Dirichlet de paramètres $[c_1(\alpha), \dots, c_{n_f}(\alpha)]$ peut se faire sur la base de distributions Gamma [Devroye, 1986]. Il s'agit en effet d'effectuer la simulation préalable d'un jeu de variables aléatoires intermédiaires indépendantes $\Gamma_1(\alpha), \dots, \Gamma_{n_f}(\alpha)$ qui, respectivement suivent une loi Gamma définie par les couples de paramètres $[1, c_1(\alpha)], \dots, [1, c_{n_f}(\alpha)]$. Les composantes du vecteur simulé $\mathbf{p}(\alpha) = [p_1(\alpha), \dots, p_{n_f}(\alpha)]^T$ correspondant, coïncident avec le rapport respectif des réalisations $\Gamma_1(\alpha), \dots, \Gamma_{n_f}(\alpha)$ par leur somme. Il est d'ailleurs possible d'utiliser une méthode alternative à la précédente pour réaliser la simulation d'une loi de Dirichlet. Elle recourt à des variables intermédiaires qui suivent des lois Bêta, car il existe également un lien fort entre ces deux familles de distributions. La première définit en effet des lois de Dirichlet unidimensionnelles.

Quel que soit le jeu de variables intermédiaires utilisé, la méthode privilégiée pour générer des nombres aléatoires suivant leur loi de probabilité respective (Gamma ou Bêta), s'effectue par anamorphose [Mallet, 2002 ; Sucquet, 2003]. Cette technique repose sur l'inversion de la fonction de répartition, dont les valeurs correspondent aux probabilités cumulées de la variable aléatoire à simuler. En pratique, comme l'illustre la figure 8.6, on effectue la correspondance entre un nombre u , préalablement simulé selon une loi uniforme sur l'intervalle $[0;1]^1$, et la courbe de probabilité cumulée. On peut alors en déduire la valeur simulée $\Pi^{-1}(u)$, où Π^{-1} désigne la fonction inverse de la fonction de répartition Π de la variable aléatoire considérée. Dans notre cas, au point α , la génération d'un vecteur de proportions $\mathbf{p}(\alpha) = [p_1(\alpha), \dots, p_{n_f}(\alpha)]^T$ dans la distribution de Dirichlet de paramètres $[c_1(\alpha), \dots, c_{n_f}(\alpha)]$ s'effectue ainsi selon l'algorithme :

1. pour tout entier f de $\{1, \dots, n_f\}$,
 - tirer de façon indépendante un nombre aléatoire $q_f(\alpha)$ selon une loi uniforme sur l'intervalle $[0;1]$;
 - effectuer l'inversion de la fonction de répartition $\Pi_f[\alpha]$ de la loi Gamma $\Gamma_f(\alpha)$ de paramètres $[1, c_f(\alpha)]$, pour en calculer la réalisation correspondante $\Gamma_f(\alpha) = \Pi_f[\alpha]^{-1}(q_f(\alpha))$;
2. calculer la somme $\Gamma(\alpha) = \Gamma_1(\alpha) + \dots + \Gamma_{n_f}(\alpha)$;
3. calculer les proportions des différents faciès f , $p_f(\alpha) = \Gamma_f(\alpha) / \Gamma(\alpha)$.

¹ La simulation d'un nombre « au hasard » compris entre 0 et 1 se fait à partir de générateurs de nombre aléatoire ; les plus usités appartiennent à la classe des générateurs aléatoires congruentiels [Delahaye, 2007].

En chaque point α , la quantité $q_f(\alpha)$ renvoie à la probabilité cumulée $\mathbb{P}(F_f(\alpha) \leq q_f(\alpha))$, c'est pourquoi on appellera par la suite le n_f -uplet $\mathbf{q}(\alpha) = [q_1(\alpha), \dots, q_{n_f}(\alpha)]^T$, le vecteur de quantiles. Il correspond à la réalisation d'un vecteur aléatoire $\mathbf{Q}(\alpha)$ dont les n_f composantes respectives $Q_1(\alpha), \dots, Q_{n_f}(\alpha)$ suivent des lois uniformes indépendantes sur l'intervalle $[0;1]$.

Pour établir un nouveau cube de proportions, il suffit donc de simuler, en chaque noeud α du réservoir, une réalisation $\mathbf{p}(\alpha)$ du vecteur de proportions aléatoire $\mathbf{P}(\alpha)$ selon cet algorithme. Néanmoins, cette étape impose auparavant de contraindre les simulations de chaque noeud entre elles, si on veut conserver la cohérence spatiale du modèle simulé. Plus exactement, on ne peut pas affecter en chaque point un vecteur de proportions simulé indépendamment de ceux qui l'entourent au risque de perdre, comme le montre la figure 8.10, la continuité spatiale du milieu souterrain. En effet, malgré sa variabilité dans l'espace, de par les structures géologiques qui le composent, il n'est en général pas non plus complètement chaotique. L'objectif consiste par conséquent à préserver cette cohérence spatiale au niveau des cubes de proportions simulés. Il s'agit de passer de la simulation régionalisée des vecteurs aléatoires $\mathbf{P}(\alpha)$ à leur simulation globale.

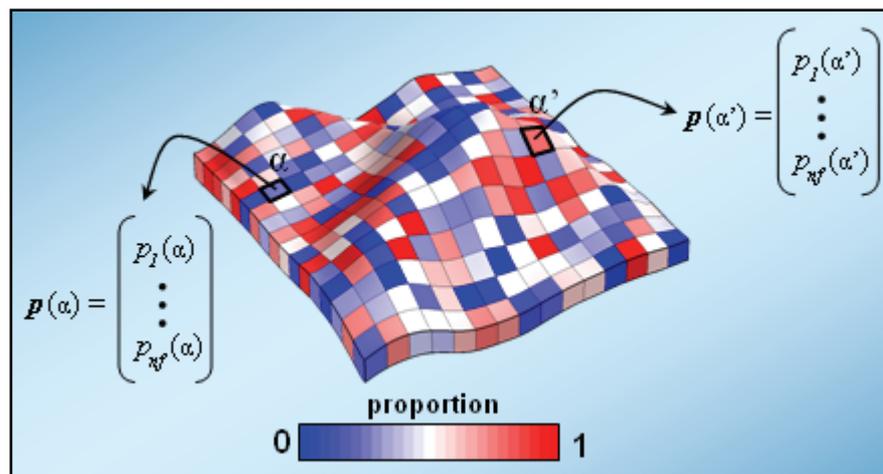


FIG. 8.10 : Illustration de la perte de cohérence spatiale. En chaque point du milieu (par exemple au point α), les vecteur de proportions de faciès $\mathbf{p}(\alpha)$ est simulé indépendamment des vecteurs de proportions aux autres points du milieu (par exemple, $\mathbf{p}(\alpha')$ au point α').

Chapitre 9

Simulation globale des cubes de proportions

La continuité des formations géologiques empêche la simulation indépendante des vecteurs de proportions en chaque point α du sous-sol. Il convient alors de conserver les relations dans l'espace entre vecteurs de proportions. Aussi, propose-t-on de corrélérer leur simulation, cette corrélation étant introduite lors de la construction des vecteurs de quantiles. Ceux-ci sont en effet simulés dans un cadre géostatistique au moyen de méthodes stochastiques basées sur la notion de variogramme. Ce dernier mesure la continuité régionale des quantiles, et par là même des proportions. Son choix dépend naturellement de l'environnement de dépôt, selon qu'il s'agit de sédimentation carbonatée ou silico-clastique, comme il dépend de la nature et du degré d'information ou de connaissance que l'on a du milieu. Divers cas pratiques, composés de réservoirs synthétiques ou de champs réels, illustrent et précisent ces différentes situations. Toutes conduisent à la construction de cubes de proportions cohérents.

9.1 Conservation de la cohérence spatiale

9.1.1 Où introduire la continuité spatiale ?

Le milieu souterrain présente une structure spatiale qu'il s'agit de conserver dans chaque cube de proportions simulé. Plus exactement, il importe de traduire l'idée fondamentale que deux vecteurs de proportions relatifs à des points situés l'un près de l'autre doivent, en moyenne, se ressembler davantage que deux vecteurs correspondant à des points éloignés. Pour cela, il est nécessaire de contraindre, en tout point α , la simulation du vecteur de proportions $\mathbf{P}(\alpha)$ avec celles de son voisinage.

L'algorithme utilisé pour leur simulation (section 8.3.3) montre que cette cohérence spatiale peut être introduite lors du tirage des variables intermédiaires indépendantes $\Gamma_1(\alpha), \dots$, et $\Gamma_{n_f}(\alpha)$. Pour rappel, ces variables sont nécessaires à la réalisation des lois de Dirichlet préalablement estimées en tout point du milieu. Leur simulation s'effectue par

inversion de leur fonction de répartition respective : en chaque point α du réservoir, nous tirons un vecteur de quantiles $\mathbf{q}(\alpha) = [q_1(\alpha), \dots, q_{n_f}(\alpha)]^T$ dans une loi uniforme multidimensionnelle sur l'ensemble $[0;1]^{n_f}$. Ses composantes déterminent les valeurs simulées $\Gamma_1(\alpha), \dots, \Gamma_{n_f}(\alpha)$, puis le vecteur de proportions $\mathbf{p}(\alpha)$. En conséquence, comme le montre la figure 9.1, la manière de simuler dans l'espace, les vecteurs de quantiles, d'un point α à un point contigu α' , influe sur la cohérence spatiale des réalisations. Leurs vecteurs de quantiles respectifs $\mathbf{q}(\alpha)$ et $\mathbf{q}(\alpha')$, tirés de façon indépendante, produiront généralement des résultats incohérents en termes de proportions. Par contre, l'introduction d'une corrélation entre leurs composantes respectives est à l'origine de vecteurs de proportions d'autant plus « proches » que cette corrélation est forte.

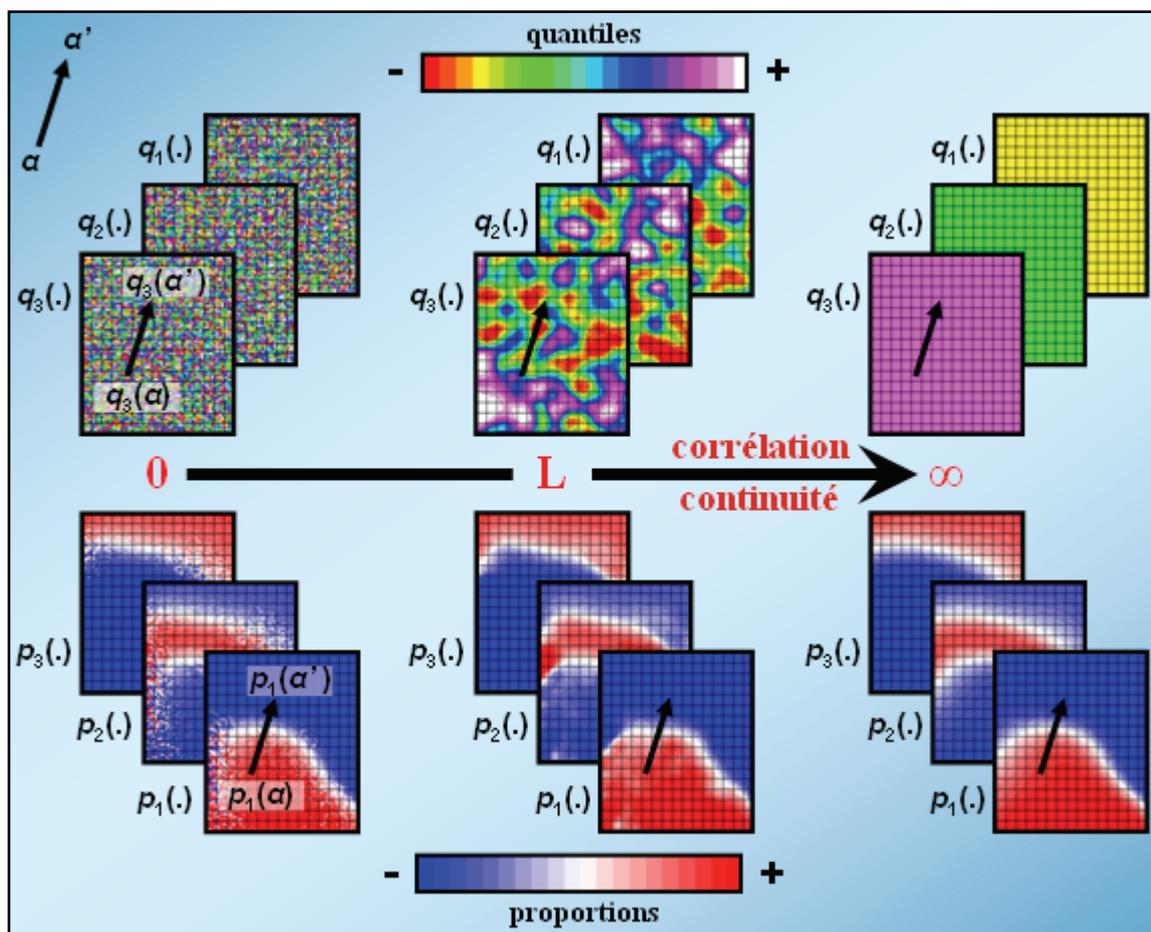


FIG. 9.1 : La continuité spatiale des cubes de proportions simulés en question. De gauche à droite, on augmente (au-dessus de la flèche) la corrélation spatiale des quantiles simulés $q_1(\cdot)$, $q_2(\cdot)$ et $q_3(\cdot)$ de zéro (pas de corrélation) à l'infini (quantiles constants dans l'espace). De cette manière, on accroît d'autant plus la continuité dans l'espace des n_f (ici $n_f = 3$) composantes $p_1(\cdot)$, $p_2(\cdot)$ et $p_3(\cdot)$ des vecteurs de proportions correspondants (sous la flèche).

La création d'un modèle en proportions consistant, qui passe par une réalisation cohérente des variables secondaires $\Gamma_f(\alpha) - f = 1, \dots, n_f$ – entre les divers points α du milieu, réclame finalement des simulations corrélées entre les vecteurs de quantiles $\mathbf{Q}(\alpha)$. Leurs composantes $Q_f(\alpha)$ ($f = 1, \dots, n_f$) étant indépendantes entre elles, le problème se réduit à générer pour chacune d'elle, plusieurs images alternatives et équiprobables, qui préservent la structure spatiale des hétérogénéités. De telles simulations, qui sont spatialement corrélées, entrent naturellement dans le cadre des méthodes géostatistiques.

9.1.2 Recours au cadre géostatistique

L'adoption du formalisme géostatistique permet de modéliser la distribution spatiale des composantes des vecteurs de quantiles tout en conservant leur structure respective. Pour chaque composante d'indice f , on considère que, lorsque le point α décrit le milieu, l'ensemble des valeurs $q_f(\alpha)$ prises par la réalisation respective des quantiles $Q_f(\alpha)$, correspond à celles d'une fonction $q_f(\cdot) : \alpha \mapsto q_f(\alpha)$ définie sur le domaine d'étude Ω . Cette fonction, appelée aussi variable régionalisée, est elle-même supposée représenter une réalisation particulière d'une fonction aléatoire $Q_f(\cdot) : \alpha \mapsto Q_f(\alpha)$, définie sur Ω et obéissant à une loi spatiale. Cette dernière décrit la fonction aléatoire $Q_f(\cdot)$ en terme de probabilité, et caractérise sa variabilité dans l'espace, d'ordinaire à travers la définition d'un variogramme. Ce dernier est l'outil fondamental de la géostatistique, car il détermine la forme des corps et des hétérogénéités que l'on cherche à modéliser. Il quantifie d'un point de vue statistique le degré de corrélation entre paires de points en fonction de leur configuration. Celui-ci dépend à la fois de la distance entre les deux points : plus ceux-ci sont éloignés, moins la corrélation entre eux est forte, et de la direction qu'ils définissent, puisque rien n'assure que la continuité soit identique dans toutes les directions. Par définition, pour la fonction aléatoire $Q_f(\cdot)$ par exemple, le variogramme $\gamma_f(\mathbf{h})$ mesure, l'écart quadratique observé en moyenne entre les valeurs que prend la fonction $Q_f(\cdot)$ sur tous les couples de points séparés d'un vecteur \mathbf{h} :

$$\gamma_f(\mathbf{h}) = \mathbb{E} \left[\frac{(Q_f(\alpha) - Q_f(\alpha + \mathbf{h}))^2}{2} \right] \quad (\text{Éq. 9.1})$$

$\mathbb{E}[Z]$ définissant l'espérance mathématique de la variable aléatoire Z .

Différents modèles de variogramme peuvent être pris en compte selon le degré de continuité à traduire ; les plus courants étant les modèles sphérique, exponentiel, gaussien, etc. Les paramètres nécessaires à leur définition sont essentiellement la *portée* R_f (*range* en anglais) et le *palier*¹ (seuil ou *sill*). Le premier définit la distance maximale au-delà de laquelle il n'y a plus de corrélation entre deux points. Le second représente la variance maximale entre deux points. Le variogramme dépend aussi de l'*azimut* θ_f des éventuelles directions d'anisotropie de la continuité spatiale, par exemple, lorsque le corps étudié montre des allongements préférentiels (lentilles, chenaux). En pratique, on représente le

¹ Si ce palier existe. Dans notre étude, ce sera toujours le cas.

variogramme par une ellipse. Elle permet de visualiser directement les directions principales et la portée du variogramme.

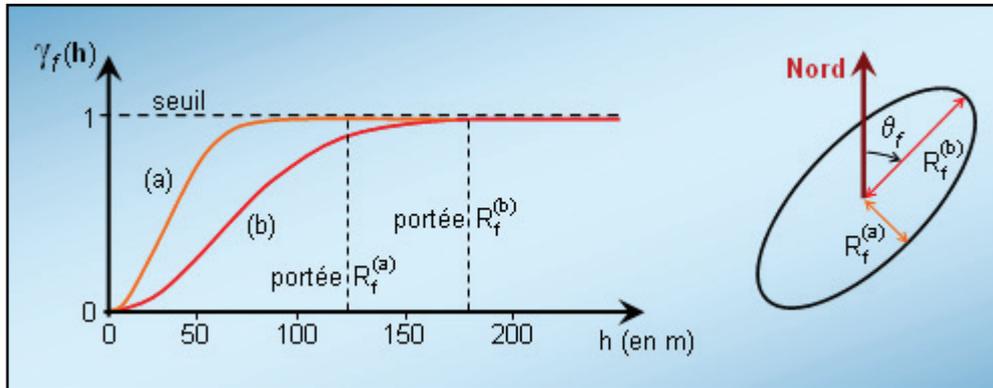


FIG. 9.2 : Représentation de la fonction variogramme. À gauche, la courbe (a) correspond au variogramme de type gaussien dans une direction donnée, la courbe (b) au variogramme dans la direction orthogonale à la première. Le paramètre h renvoie quant à lui à la distance entre deux points pour laquelle le variogramme est calculé. Le seuil ($s=1$) et la portée ($R_f^{(b)} = 180$ m) sont représentés pour la courbe (b). À droite, l'ellipse est une représentation commode du variogramme. Elle permet de visualiser directement les directions principales et la portée du variogramme (d'après Leflon, 2005).

Le cadre géostatistique permet de recourir aux techniques classiques de simulations stochastiques en approche dite *pixel* afin de générer des réalisations alternatives en quantiles $Q_f(\cdot)$ dans le réservoir. Ces procédés appréhendent l'espace comme un ensemble de points – d'où le terme de *pixel* –, ces points servant d'ailleurs de support aux propriétés à modéliser. Parmi les nombreuses méthodes en approche *pixel* qui peuvent exister, avec entre autres les gaussiennes tronquées [Matheron *et al.*, 1987], les plurigaussiennes tronquées [Loc'h *et al.*, 1996], ou la géostatistique multipoints proposée par Caers (2001) et Strebelle *et al.* (2001), on privilégiera les méthodes de simulation gaussienne séquentielle (SGS) [Journel, 1989 ; Deutsch *et al.*, 1998]. Elles représentent l'une des techniques les plus couramment utilisées pour simuler des variables continues en géostatistique. Outre le conditionnement direct aux données observées, elles reproduisent des champs de valeurs qui révèlent une structure spatiale (variogramme) et un histogramme identique à ceux inférés au moyen des données, ou bien supposés dans le milieu. En pratique, les méthodes de simulations gaussiennes séquentielles génèrent d'abord un chemin aléatoire qui permet de visiter successivement chacun des nœuds du domaine d'étude préalablement discrétisé. Pour chaque point, à tour de rôle, elles simulent sa valeur en se servant (a) de l'ensemble des points connus, incluant ainsi les points conditionnants¹ comme ceux déjà simulés, et (b) d'un variogramme.

¹ S'il y en a, auquel cas on parlera de simulations conditionnées, sinon on parlera plutôt de simulations non conditionnées.

Ces méthodes sont appliquées à la simulation indépendante de chacun des quantiles $Q_f(\cdot)$. Elles distribuent dans l'espace les valeurs extraites d'une loi uniforme sur $[0;1]$ de manière à honorer le variogramme $\gamma_f(\mathbf{h})$, caractéristique de la structure spatiale du quantile $Q_f(\cdot)$. Ces champs de valeurs correspondent aux composantes de vecteurs de quantiles spatialement corrélés, dont vont dériver des vecteurs de proportions qui le sont également, et un cube de proportions qui présente une certaine continuité géologique. D'un point de vue pratique, si le support géométrique (modèle maillé) décrit un cadre chronostratigraphique intégrant l'organisation séquentielle des faciès issue des corrélations, la simulation du quantile $Q_f(\cdot) - f = 1, \dots, n_f -$, s'effectue indépendamment d'un intervalle stratigraphique à l'autre. En effet, dans ce contexte, nous considérons que les propriétés n'offrent aucune corrélation verticale entre les intervalles ainsi définis. Leur architecture interne (maillage vertical) présente un caractère chronostratigraphique, qui rend chaque couche indépendante des couches du dessus ou du dessous, contraignant à accomplir les simulations de quantiles $Q_f(\cdot)$ en deux dimensions, horizon par horizon.

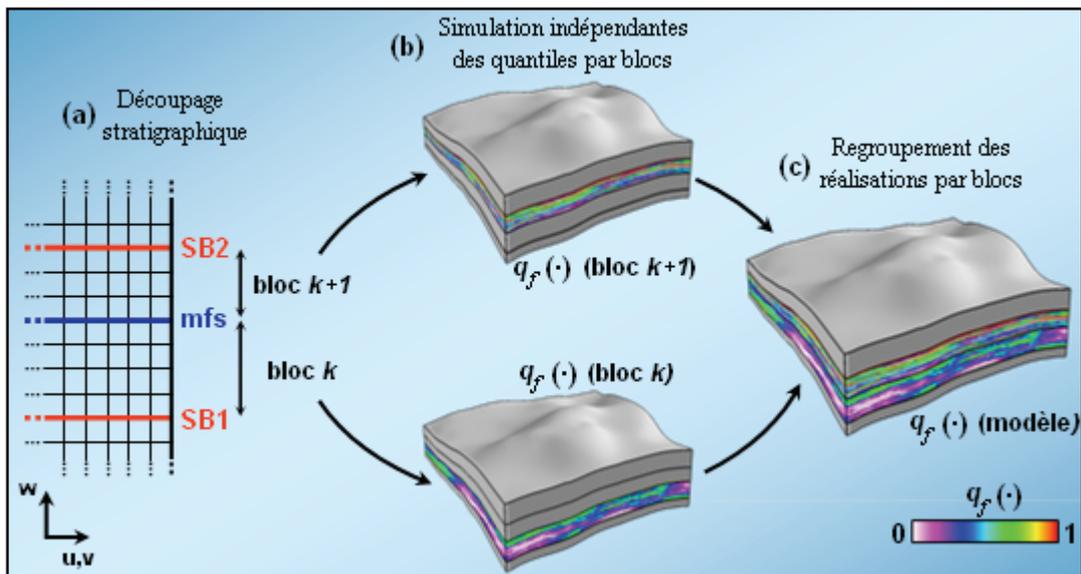


FIG. 9.3 : Intégration du contexte chronostratigraphique. **(a)** Le support géométrique décrit un cadre chronostratigraphique délimité par les limites de séquences (SB) et une surface d'ennoyage maximum (mfs). **(b)** Les simulations des quantiles (donc des proportions de faciès) s'effectue indépendamment d'un bloc à l'autre, puis sont regroupées dans le même modèle **(c)**.

Du reste, ces simulations nécessitent en plus la détermination préalable des variogrammes $\gamma_f(\mathbf{h})$. Le choix de leurs paramètres (portées, seuils, azimuth) dépend fortement de l'environnement de dépôt considéré, selon qu'il s'agit de sédimentation (1) carbonatée, plutôt continue avec des changements graduels de faciès, ou (2) silico-clastique, plus « factuelle » avec des variations brutales de faciès. D'un milieu à l'autre, la continuité spatiale ne se traduit donc pas de la même manière, ce qui oblige à un traitement particulier de chacun de ces environnements.

9.2 En domaine de plate-forme

Dans ce type d'environnement, la nature des sédiments suit en général une certaine polarité, et varie de façon graduelle le long de la rampe carbonatée (section 9.2.1). Une telle répartition se caractérise bien entendu par une variation spatiale relativement lisse des proportions de faciès et ce, selon une certaine « carte » de tendance (paléogéographie). Dans notre approche, ces caractéristiques sont en grande partie portées par les lois de Dirichlet régionalisées, elles-mêmes extraites des modèles de base. Par conséquent, il est possible de déterminer les variogrammes à employer pour la simulation des quantiles $Q_f(\cdot)$ en analysant la continuité spatiale des modèles initiaux (section 9.2.2). Les variogrammes ainsi inférés reflètent fréquemment une corrélation spatiale importante des propriétés dans l'espace. C'est pourquoi on envisage par la suite, le cas limite d'une corrélation spatiale infinie (section 9.2.3), où tous les points sont corrélés entre eux. Quelle que soit la démarche utilisée, chacune permet de conserver une continuité géologique forte, continuité que l'on observe ordinairement sur les cubes de proportions modélisant ce type d'environnement. L'application à un cas synthétique de sédimentation en milieu carbonaté illustre et précise les deux approches.

9.2.1 Le modèle de référence

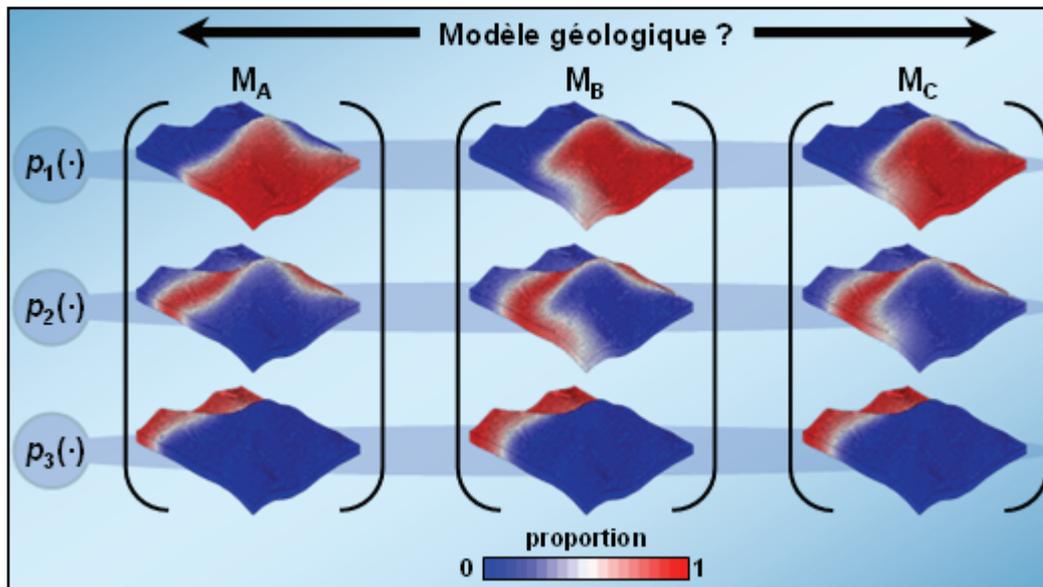


FIG. 9.4 : Présentation des cubes de proportions initiaux. Les $n_m(G)$ modèles à partir desquels sont calculées les lois de Dirichlet décrivent différentes hypothèses liées à la paléogéographie : un profil de plage (M_A), un promontoire d'axe N-S (M_B), SO-NE (M_C).

Le modèle étudié recense trois catégories de faciès ($n_f = 3$), que l'on désigne respectivement par F_1 , F_2 et F_3 , de la meilleure à la moins bonne roche réservoir. Il revient en

outre sur le cas synthétique TATA déjà présenté au chapitre 5 (section 5.2.3). Ce dernier illustre les effets de l'incertitude géologique sur la morphologie du réservoir exprimée sous la forme de trois cartes de bathymétrie. On reprend successivement ces hypothèses pour construire, à l'aide d'une relation probabiliste entre nature de faciès et bathymétrie semblable à celle employée dans Neptune (deuxième partie), trois cubes de proportions correspondant (a) à un profil de plage (modèle (\mathbf{M}_A)), (b) un promontoire d'axe nord-sud (\mathbf{M}_B), ou (c) d'orientation sud-ouest / nord-est (\mathbf{M}_C). En conséquence, l'incertitude géologique se répercute sur l'extension latérale des faciès et de leurs proportions, qui diffèrent d'un modèle à l'autre (Fig. 9.4). Pour simplifier, le modèle comprend une seule unité stratigraphique, l'aléa portant essentiellement sur la disposition horizontale des proportions de faciès. Une grille de type « boîte à sucres » symbolise le domaine d'étude, et le discrétise en $70 \times 60 \times 1$ cellules à section carrée de 100 mètres de côté. Chaque point α du modèle maillé renvoie à un vecteur de proportions $\mathbf{P}(\alpha)$, qui suit une loi de Dirichlet de paramètres $\mathbf{c}(\alpha)$, celle-ci ayant été préalablement estimée grâce aux modèles (\mathbf{M}_A), (\mathbf{M}_B) et (\mathbf{M}_C), pondérés de façon identique (chapitre 8). Il s'agit à présent de simuler les quantiles $Q_f(\cdot)$ par les méthodes SGS grâce à leur variogramme défini selon l'une des deux approches citées précédemment.

9.2.2 Une corrélation spatiale inférée

Une première méthode s'appuie sur les $n_m(G)$ modèles de base, c'est-à-dire les cas (\mathbf{M}_A), (\mathbf{M}_B) et (\mathbf{M}_C) dans cet exemple, afin d'étudier la variabilité spatiale de chaque quantile $Q_f(\cdot)$. Après tout, chacun des modèles initiaux représente, à une transformation¹ près, une réalisation particulière des différentes fonctions aléatoires $Q_f(\cdot)$. L'idée consiste donc à inférer, à partir de ces réalisations, les caractéristiques (type, seuil, portée) d'un variogramme expérimental $\gamma_f(\mathbf{h})$, que ces réalisations sont sensées respecter. Cette approche, appliquée respectivement à chacun des modèles de départ, se décompose en deux étapes (Fig. 9.5) :

- en tout point α du modèle considéré, connaissant la loi de Dirichlet qui lui est attachée, on détermine les composantes du vecteur de quantiles qui, une fois transformées¹, fournissent les proportions de faciès en ce point ;
- chaque composante étant désormais connue partout dans l'espace, on réalise de façon traditionnelle son analyse variographique ; d'une part, on calcule pour des distances croissantes et ce, dans un certain nombre de directions, les valeurs d'un variogramme expérimental ; puis, on cherche à ajuster à ces séries de points expérimentaux, un modèle analytique de variogramme parmi les modèles gaussiens, sphériques et exponentiels. Ceux-ci permettent de déduire la covariance entre deux points quelconques en fonction de leur espacement et, éventuellement, de la direction qu'ils définissent.

¹ D'après l'algorithme décrit à la section 8.8.8 utilisé pour simuler une loi de Dirichlet, la transformation, qui permet le passage de quantiles à des proportions, correspond à la combinaison d'une anamorphose et d'un quotient.

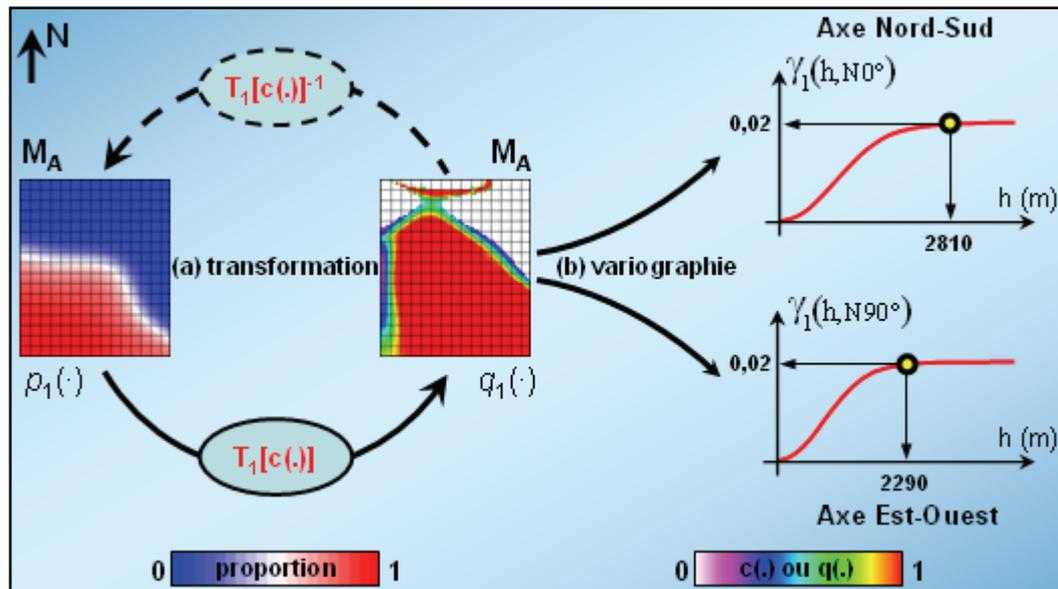


FIG. 9.5 : Inférence de la continuité spatiale des quantiles à partir des modèles initiaux (1). Pour chaque modèle de base (ici \mathbf{M}_A), les proportions dans l'espace de chaque faciès (ici $p_1(\mathbf{M}_A, \cdot)$) sont converties en quantiles $q_1(\mathbf{M}_A, \cdot)$ par la transformation $\mathbf{T}_1[\mathbf{c}(\cdot)]$. Il s'agit des quantiles, qui par la transformation inverse $\mathbf{T}_1[\mathbf{c}(\cdot)]^{-1}$ fournissent les proportions $p_1(\mathbf{M}_A, \cdot)$. Une variographie de ces quantiles (b) détermine les paramètres du variogramme associé selon deux directions principales. Ce travail doit être réalisé pour tous les faciès et tous les modèles.

Pour chaque quantile $Q_f(\cdot)$ – ici $f = 1, 2$ ou 3 –, on dispose donc de $n_m(G)$ expressions analytiques du variogramme $\gamma_f(\mathbf{h})$ qui lui est associé, chacune étant issue des cas de base. Par exemple, l'étude successive des modèles (\mathbf{M}_A), (\mathbf{M}_B) et (\mathbf{M}_C) définit pour la composante $Q_f(\cdot)$, un modèle théorique de variogramme pour chaque cas initial. Ceux-ci s'ajustent visuellement le mieux aux variogrammes expérimentaux calculés selon les directions nord-sud (N-S) et est-ouest (E-O). Dans tous les cas, ils correspondent à des variogrammes gaussiens :

- de portée $a = 2810$ mètres selon l'axe (N-S) (2290 m pour la direction (E-O)) et de palier $S = 0,02$ (respectivement $S = 0,02$) pour le modèle (\mathbf{M}_A) ;
- pour le cas (\mathbf{M}_B), il présente une portée de 2030 mètres et de palier de 0,01 selon l'orientation (N-S) ; quant à l'axe (E-O), il révèle une portée de 1650 m et un seuil de 0,03 ;
- pour le modèle (\mathbf{M}_C), selon la direction (N-S) et (E-O), sa portée est respectivement de 2860 m et 2320 m pour des paliers de 0,02 et 0,02.

Pour ne privilégier aucun modèle de corrélation, on propose de combiner, pour chaque quantile $Q_f(\cdot)$, les $n_m(G)$ variogrammes tirés de l'analyse variographique des scénarios de base. Une solution admissible consiste à utiliser un modèle qui correspond à une moyenne

pondérée des divers variogrammes possibles (Fig. 9.6). Dans le cas général, on affecte un poids positif ω_k à chacun des modèles de base, k renvoyant à son indice. Le variogramme $\gamma_f(\mathbf{h})$ utilisé pour rendre compte de la corrélation spatiale du quantile $Q_f(\cdot)$ correspond ainsi à :

$$\gamma_f(\mathbf{h}) = \sum_{k=1}^{n_m(G)} \omega_k \times \gamma_f[k](\mathbf{h}) \quad (\text{Éq. 9.2})$$

où $\gamma_f[k](\mathbf{h})$ désigne pour tout couple de points séparés d'un vecteur \mathbf{h} , la valeur du variogramme du quantile $Q_f(\cdot)$ estimé à partir du scénario d'indice k ; la somme des $n_m(G)$ poids ω_k est égale à un.

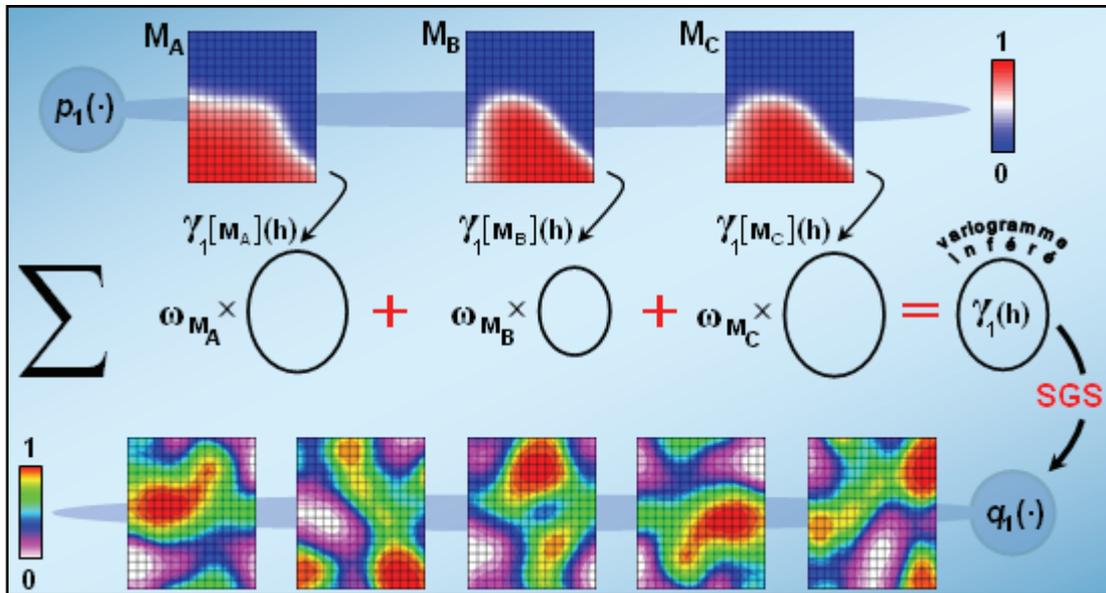


FIG. 9.6 : Inférence de la continuité spatiale à partir des modèles de bases (2). Pour tous les faciès (ici F_1), chaque modèle de base d'indice k fournit un variogramme $\gamma_1[k](h)$ associé au quantile correspondant $q_1[k](\cdot)$ (Fig. 9.5). Un poids w_k est affecté à chacun des variogrammes obtenus à l'aide des modèles initiaux. Leur combinaison linéaire permet de déterminer un variogramme $\gamma(h)$. Ce dernier est utilisé pour réaliser les simulations SGS donnant les différentes réalisations en termes de composantes $q_1(\cdot)$ du vecteur de quantiles. Ce travail doit être réalisé pour l'ensemble des faciès.

Dans cet exemple, on affecte aux scénarios de départ (M_A), (M_B) et (M_C) un poids identique, égal à un tiers. Les variogrammes obtenus sont alors utilisés pour la simulation des quantiles nécessaires au calcul des proportions de faciès. La figure 9.7 présente ainsi trois simulations en termes de cubes de proportions. On obtient des images alternatives du réservoir plus bruitées que les modèles de départ, même si on conserve leur cohérence et structure géologique.

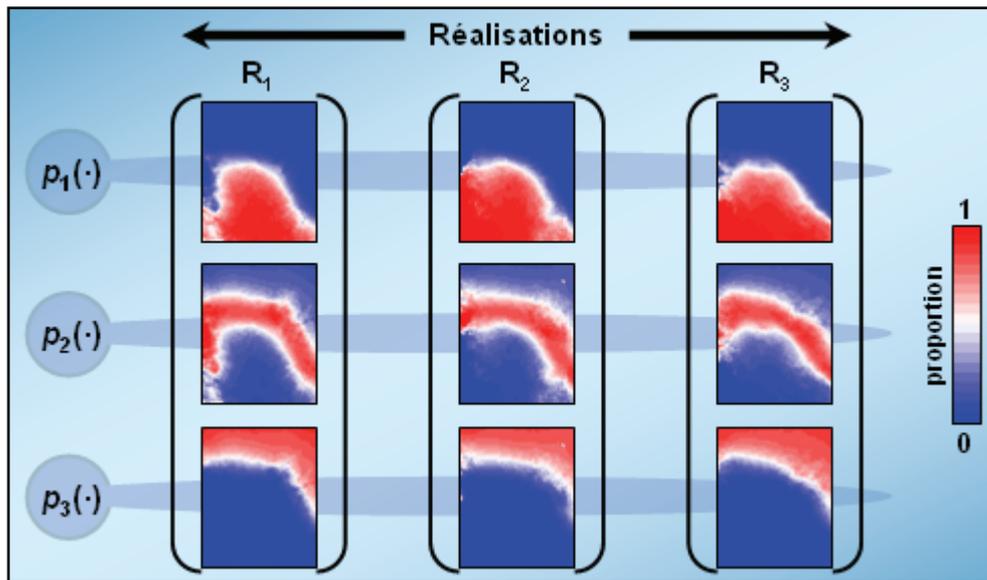


FIG. 9.7 : Cubes de proportions simulés avec un vecteur de quantiles simulés par SGS. Pour chacune des réalisations, les composantes des vecteurs de quantiles sont simulées par simulations gaussiennes séquentielles, avec respectivement les variogrammes inférés à partir des modèles de base.

9.2.3 Une corrélation spatiale infinie

Une alternative possible suppose une corrélation spatiale infinie afin de préserver la continuité géologique : c'est comme si tous les points du modèle étaient spatialement corrélés. En fait, cette hypothèse revient à simuler les cubes de proportions qui, pour chaque faciès f , renvoient respectivement, en tout point α , une proportion correspondant au même quantile, et ce pour toutes les lois de probabilité $P_f(\alpha)$. On conserve ainsi la cohérence des différents cas de base : on décrit dans l'espace les mêmes quantiles pour chacune des lois de Dirichlet régionalisées, dont les paramètres – la moyenne surtout –, portent finalement la non stationnarité des faciès.

Pour cela, il suffit d'utiliser dans les méthodes SGS un variogramme de portée infinie et de seuil nul lors de la simulation des quantiles $Q_f(\cdot)$. Néanmoins, en pratique, on génère respectivement, pour chaque réalisation de cubes de proportions, un vecteur de quantiles global, $\mathbf{q} = [q_1, \dots, q_{n_f}]^T$ ($n_f = 3$), le même pour l'ensemble des nœuds du modèle. Quel que soit le point α du réservoir, le vecteur de quantiles $\mathbf{q}(\alpha) = [q_1(\alpha), \dots, q_{n_f}(\alpha)]^T$ qui lui est affecté vérifie :

$$\mathbf{q}(\alpha) = \mathbf{q} \quad (\text{Éq. 9.3})$$

soit pour chaque composante d'indice f (ici égal à 1, 2 ou 3) :

$$q_f(\alpha) = q_f \quad (\text{Éq. 9.4})$$

q_f correspondant à un nombre aléatoire tirée dans une loi uniforme sur $[0,1]$.

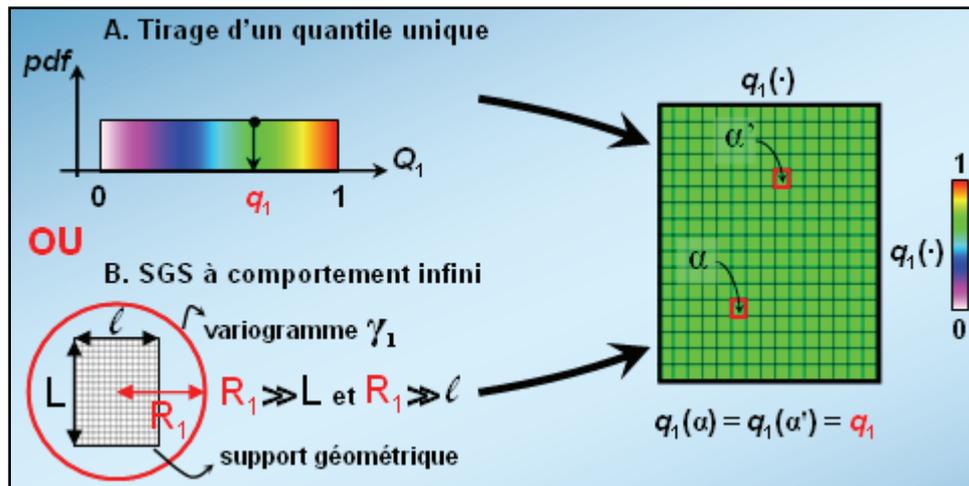


FIG. 9.8 : Deux procédés de simulation d'un vecteur de quantiles unique. Pour chaque quantile (ici $q_1(\cdot)$), (A) on tire dans une loi uniforme une valeur q_f (ici q_1) qui est affectée à l'ensemble des points α du modèle ; ou (B) on recourt à des simulations SGS au comportement « infini ». Les variogrammes ont une portée très grande devant les caractéristiques du modèle, fournissant ainsi un quantile unique pour chaque point α .

Dans ce cas d'étude, on effectue quatre tirages indépendants de vecteurs quantiles fournissant ainsi quatre réalisations en termes de cubes de proportions. On obtient des images géologiques, qui présentent un degré de continuité équivalent à celui des modèles de base. Elles sont d'ailleurs difficilement reproductibles par l'intermédiaire d'une simple combinaison linéaire des cubes initiaux.

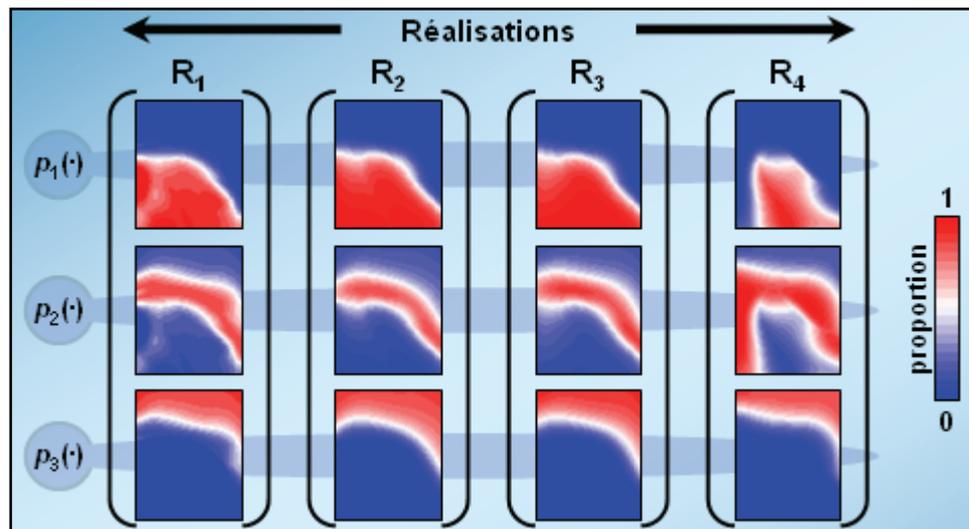


FIG. 9.9 : Cubes de proportions simulés avec un vecteur de quantiles stationnaire. Pour chaque réalisation R_k , un unique vecteur de quantiles (un jeu de quantiles $q_f(\cdot)$) est simulé, puis est affecté à l'ensemble des points α du modèle.

En milieu carbonaté, la cohérence spatiale peut être inférée à partir des cubes de proportions initiaux, ou bien supposée infinie. De telles approches induisent une continuité majeure sur les cubes de proportions. Elles paraissent de ce fait difficiles à transposer aux environnements silico-clastiques, que caractérisent des changements abrupts de faciès.

9.3 En domaine chenalisé

Ce type d'environnement favorise la sédimentation d'objets géologiques ayant une forme bien définie (comme les chenaux par exemple), et selon une direction privilégiée. Il se caractérise par une évolution latérale des faciès très rapide, ce qui conduit en général à faire l'hypothèse de stationnarité. Néanmoins, dans certains cas (section 9.3.1), les apports sableux se font selon des axes préférentiels, que Leflon (2005) appelle ceintures chenalisantes, ou *fairway* en anglais. Physiquement, cela signifie par exemple, que les chenaux respectent le tracé d'éventuelles vallées. Dans ces régions, la proportion des faciès de chenaux est globalement plus élevée qu'ailleurs ; la probabilité a priori de trouver un faciès donné n'est donc pas indépendante de la position dans le milieu. D'ailleurs, ces axes préférentiels imposent souvent une anisotropie importante au cube de proportions : les proportions de faciès montrent en général une meilleure continuité parallèlement à la direction des apports que perpendiculairement à celle-ci. Dans notre approche, ces caractéristiques sont en grande partie régies par le choix des variogrammes utilisés lors de la simulation des quantiles $Q_f(\cdot)$ (section 9.3.2). D'un autre côté, on peut observer, à une autre échelle, une évolution latérale relativement continue du nombre (proportion) de chenaux (section 9.3.3). À l'intérieur d'un *fairway* par exemple, ces variations se font perpendiculairement à la direction des apports. Là, ce sont plutôt les lois de Dirichlet extraites des modèles de base qui portent la continuité géologique (section 9.3.4). Dans tous les cas, l'objectif consiste à exprimer l'anisotropie observée d'ordinaire sur les cubes de proportions modélisant ce type d'environnement. L'application à un cas synthétique de sédimentation en milieu silico-clastique illustre et précise le choix du modèle de continuité spatiale utilisé pour la simulation des quantiles $Q_f(\cdot)$ et ce, pour divers types et degrés d'information ou de connaissance du milieu.

9.3.1 Le modèle de référence

Pour illustrer les différents cas proposés, nous allons utiliser tout au long de cette section le modèle de référence ZOUINE. Il se rapporte à un cas synthétique de sédimentation en milieu fluvial. Le réservoir est orienté selon un axe est-ouest, et délimite une zone de 10 kilomètres de long pour 5 km de largeur. Par commodité, les calculs sur les propriétés sont réalisés dans une grille cartésienne¹. Les résultats peuvent être transférés dans une grille stratigraphique, qui rend compte de la véritable géométrie du milieu, et ce grâce à une transformation qui assure le passage d'un support à l'autre [Mallet, 2004 ; Moyen, 2005]. Cette grille comprend une unité stratigraphique ; elle est discrétisée horizontale-

¹ De type « boîte à sucres ».

ment en 100×50 cellules de section carrée de 100 mètres de côté (Fig. 9.10). Le modèle sédimentaire distingue trois variétés de faciès typiques des dépôts fluviaux : des sables grossiers de chenaux (F_1), des silts de levées (F_2), et pour finir des argiles (F_3) de plaines d'inondation. Bien entendu, les réservoirs principaux se situent dans les chenaux sableux.

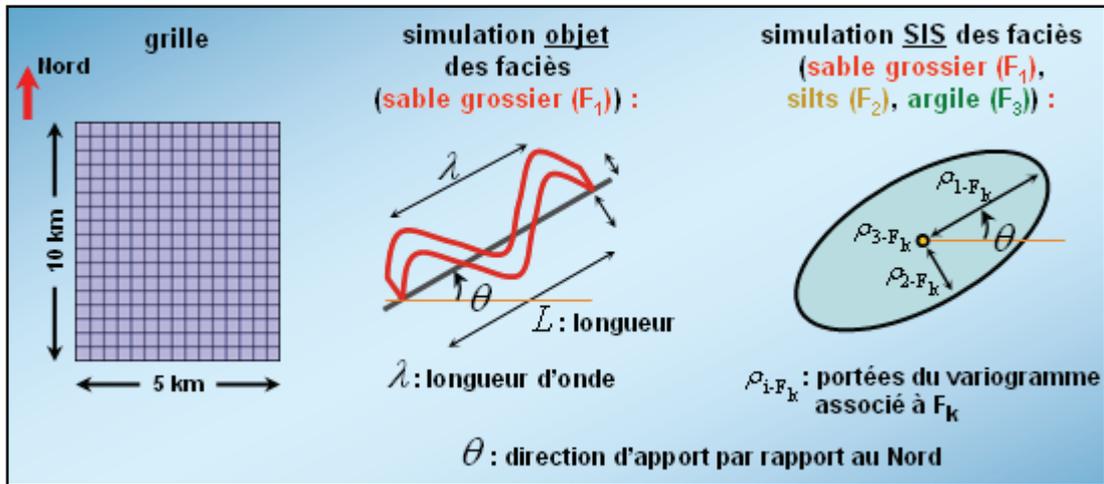


FIG. 9.10 : Le modèle de référence utilisé (ZOUINE). C'est un cas synthétique où les dépôts sont d'origine fluviale. Les faciès sont simulés à partir d'un cube de proportions à l'aide des méthodes classiques de simulation « objet » (F_1) ou d'indicatrices (F_1, F_2, F_3).

Dans les cas pratiques envisagés par la suite, la simulation des faciès F_2 et F_3 passe par les techniques *basées pixels* de simulations séquentielles par indicatrices (SIS) [Journel, 1989 ; Goovaerts, 1997]. Elles correspondent à l'application des méthodes SGS aux variables discrètes qui sont préalablement transformées. Chaque faciès est codé par une Indicatrice¹ binaire qui, en tout point, caractérise sa présence ou son absence. Un variogramme propre à chaque faciès² rend compte des corrélations spatiales et des directions d'anisotropie. D'ailleurs, on suppose habituellement que les portées correspondent à la taille des corps modélisés. Les réalisations obtenues reflètent ainsi les différences de continuité spatiale des divers faciès tout en respectant leur proportion spatiale. La figure 9.10 décrit le paramétrage des variogrammes utilisés pour simuler les silts (F_2). Quant à la simulation des faciès de chenaux, elle s'effectue d'une part, selon une approche SIS, dont le variogramme associé est fourni à la figure 9.10. D'autre part, on envisage la simulation des chenaux au moyen de méthodes *basée objets*, aussi appelées méthodes booléennes [Viseur, 2001 ; Lopez, 2003]. Elles assimilent les corps sédimentaires – comme par exemple les chenaux – à des formes paramétriques prédéfinies. Elles distribuent alors

¹ Il s'agit d'une fonction aléatoire associée à un faciès qui, partout où on l'observe vaut un, zéro sinon.

² En pratique, il est toujours un faciès correspondant au remplissage de la grille ; il n'est donc pas généré en utilisant une simulation. Aussi, pour un cas à n faciès, seuls $n-1$ variogrammes sont nécessaires. Une approche plus élaborée définit pour chaque faciès une indicatrice, un variogramme, et un ordre de simulation relatif à un ordre d'érosion.

dans l'espace des objets géométriques simples – comme des sinusoïdes, des disques –, que décrivent des paramètres (orientation, largeur, épaisseur...) aléatoires ou non. Les faciès de chenaux (F_1) sont ici représentés par des bandes sinusoïdales dont la figure 9.10 détaille le paramétrage. Nous remarquerons que, dans les deux approches, le modèle de faciès impose une anisotropie forte en faveur de la direction d'apports sédimentaires.

À partir du modèle de référence, nous envisageons successivement trois cas pratiques que distinguent la nature comme le degré d'information et de connaissance du sous-sol. Ces différences se répercutent sur le choix des variogrammes à utiliser pour la simulation des quantiles $Q_f(\cdot)$. Certaines situations ont d'ailleurs été déclinées sur des cas opérationnels ; ces derniers ont ainsi permis de tester la méthodologie mais aussi d'intégrer des contraintes supplémentaires.

9.3.2 Cas pratique n° 1 : Simulation de zones préférentielles pour les chenaux

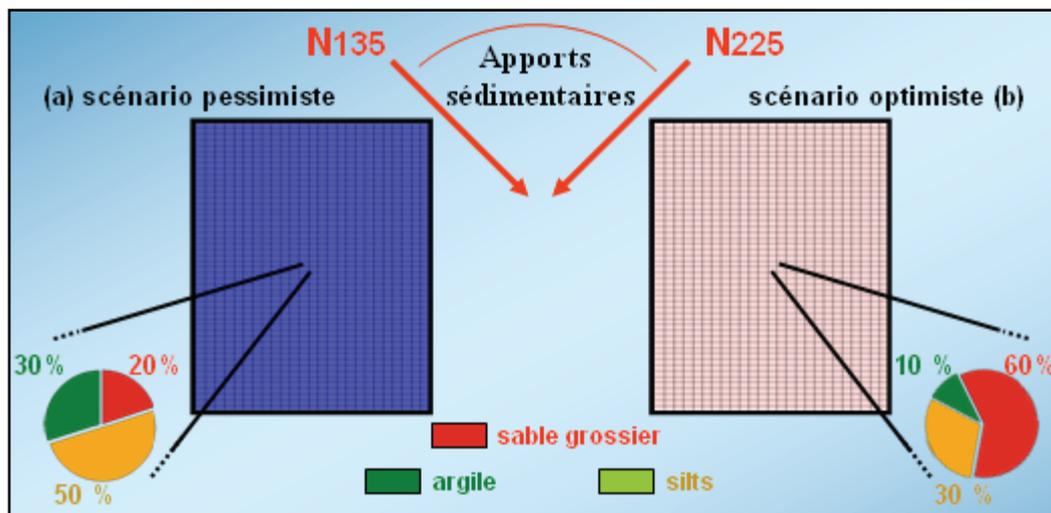


FIG. 9.11 : Cas pratique n°1 : Incertitude sur les zones préférentielles de chenaux. L'incertitude sur le scénario porte sur la direction des apports sableux, les proportions de faciès et les dimensions caractéristiques des zones préférentielles.

Les apports sableux peuvent se faire selon des axes préférentiels, où la proportion des faciès de chenaux est plus élevée qu'ailleurs. Néanmoins, il arrive que leurs positions comme leurs caractéristiques, naturellement liées à celles des chenaux, soient incertaines. Dans le cas présenté ici, l'incertitude touche précisément (Fig. 9.11) :

- (1) la valeur globale des proportions dans chaque couche du réservoir, que l'on appréhende à partir d'un cas de base, (a) en le dégradant en un scénario pessimiste, avec 20% de sables grossiers, 50 % de silts et 30 % d'argiles, ou (b) en l'améliorant en un modèle optimiste : sables grossiers (60%), silts (30%) et argiles (10%). Dans les deux cas, la proportion des divers faciès est constante sur chaque horizon ;

- (2) la direction des apports sableux, qui en l'absence d'azimut préférentiel, peut définir un angle avec le Nord compris entre 135 et 225 degrés ;
- (3) l'extension des zones préférentielles, dont les caractéristiques (largeur, longueur) sont naturellement liées à celles des faciès, mais restent difficiles à évaluer. Ici, seule la largeur est indéterminée, mais reste comprise entre 800 m et 2 km.

La question est donc de transcrire le caractère non stationnaire et incertain de la fonction aléatoire associée à chaque faciès : la première incertitude porte sur la valeur globale des proportions, les deux dernières sur la localisation des zones préférentielles.

Certains auteurs contournent la difficulté en simulant des ceintures chenalisantes dans le domaine d'étude avant de modéliser les faciès identifiés aux puits. Ils interprètent la présence ou l'absence de ces ceintures comme les valeurs prises par l'Indicatrice qui leur est associée. Sa loi de probabilité dans l'espace, entendez son cube de proportions, est souvent supposée constante sur chaque couche du modèle. En général, on l'estime à partir des statistiques réalisées le long des puits en fonction de l'épaisseur stratigraphique, représentant ainsi l'évolution verticale des ceintures chenalisantes. Les méthodes géostatistiques de simulations des variables catégorielles¹ permettent alors de simuler, en tout point de l'espace, ces *fairways*. Chaque réalisation définit une partition du milieu souterrain à partir de laquelle, il devient possible de simuler des chenaux exclusivement dans les régions délimitées par les ceintures chenalisantes préalablement simulées. Si cette approche appréhende la non stationnarité des faciès, l'estimation des proportions de *fairways* à partir des puits s'avère néanmoins délicate, puisque ces ceintures ne renvoient pas forcément à un faciès observé le long des puits. D'ailleurs, ils n'intègrent en général pas la variabilité des proportions de faciès : ils ne considèrent qu'un seul cube de proportions pour les *fairways*.

Plutôt que de construire ces objets intermédiaires, nous proposons d'interpréter directement la non stationnarité des faciès à travers leur probabilité d'apparition dans l'espace. Pour cela, on simule, grâce au formalisme des lois de Dirichlet, l'ensemble des modèles en proportions de faciès, qui présentent des zones préférentielles où la probabilité de simuler des chenaux sera plus forte, et qui sont compatibles avec l'incertitude sur le scénario géologique. En adoptant ce formalisme, la combinaison² des scénarios pessimiste et optimiste, qui présentent des proportions de faciès constantes sur chaque horizon, définit une loi de Dirichlet stationnaire sur chaque couche : en tout point α du milieu, les vecteurs de proportions $\mathbf{P}(\alpha)$ suivent la même loi statistique. En conséquence, la non stationnarité des faciès, disons les degrés de continuité et d'anisotropie de leur proportion, s'explique par le choix des modèles de variogramme $\gamma_f(\mathbf{h})$ utilisé pour simuler les quantiles $Q_f(\cdot)$.

¹ En général, en environnement de chenaux, ces simulations sont réalisées à partir de méthodes *basées « objets »*, et ce dans l'idée d'assurer la continuité des régions qui correspondent aux ceintures chenalisantes.

² Dans le calcul des paramètres des lois de Dirichlet en chaque point, on attribue un poids identique aux deux scénarios de base.

Ce choix est guidé à la fois, par les caractéristiques des faciès étudiés, et par la nature des propriétés à modéliser, en l'occurrence des proportions. Aussi, rapproche-t-on l'anisotropie du variogramme $\gamma_f(\mathbf{h})$ associé au quantile $Q_f(\cdot)$ de celle du faciès correspondant. Pour chaque faciès f , la direction principale θ_f du variogramme $\gamma_f(\mathbf{h})$ coïncide naturellement avec la direction selon laquelle le faciès montre des allongements préférentiels, et offre donc le plus de continuité¹. Dans le cas présenté ici, cela revient à considérer l'azimut des apports sédimentaires. On tire par conséquent la direction principale θ_f dans une distribution uniforme de support $[135^\circ; 225^\circ]$, puisque l'orientation des apports est inconnue. Rappelons que cette dernière peut définir un angle avec le Nord compris entre 135 et 225 degrés.

En revanche, pour les autres paramètres du variogramme : type, portées et seuils, le lien avec la structure spatiale des faciès ne se révèle pas aussi directe. Comme le rappelle la figure 9.12, l'information portée par les proportions représente une information à grande échelle. Par conséquent, les variations spatiales propres aux proportions s'effectuent de façon plus régulière que celles des faciès correspondants². À ce titre, pour chaque faciès f , le variogramme $\gamma_f(\mathbf{h})$ associé au quantile $Q_f(\cdot)$ ne peut être identique au modèle de continuité (variogramme ou paramètres géométriques des objets) du faciès f . Par exemple, dans le cas d'une simulation des faciès de chenaux suivant les méthodes SIS, le variogramme $\gamma_f(\mathbf{h})$ doit avoir un seuil plus bas, une pente à l'origine plus faible et une portée plus longue que le variogramme associé aux chenaux (Fig. 9.12).

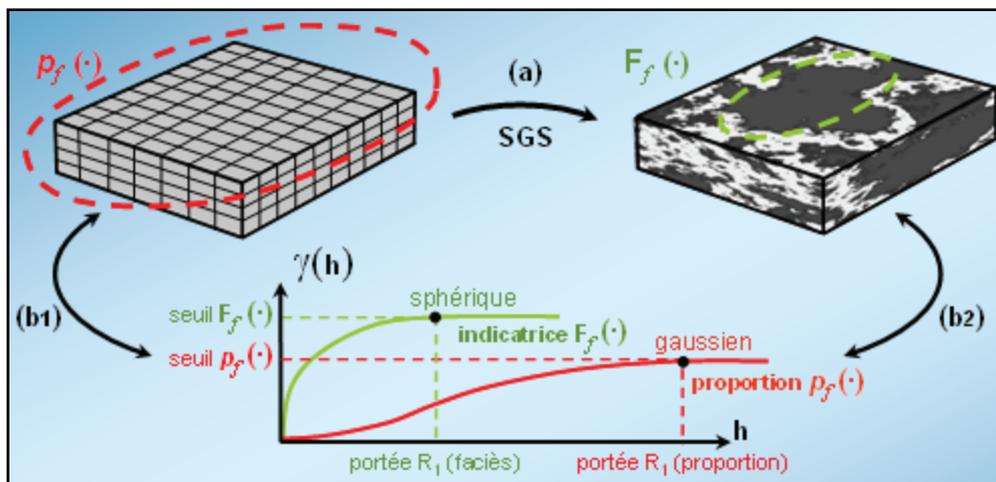


FIG. 9.12 : Comparaison entre variogrammes de proportions et de faciès. La proportion $p_f(\cdot)$ décrit des variations à plus grande échelle (valeur uniforme sur le modèle) que celle du faciès F_f simulé (a) à partir de celle-ci. Le variogramme relatif à $p_f(\cdot)$ (b1) a ainsi un seuil plus bas, une portée plus longue que celui associé à F_f (b2).

¹ Par exemple, pour les silts de levées, c'est la direction du variogramme utilisé pour leur simulation selon les méthodes SIS.

² La proportion de chaque faciès peut être interprétée comme la moyenne de l'Indicatrice qui code le faciès, ce qui explique le caractère lissé des cubes de proportions par rapport aux modèles en faciès.

Ainsi, pour chaque variogramme, en ce qui concerne :

- (a) **le type** : on privilégie l'emploi de variogrammes gaussiens, car ils expriment une très grande continuité, typique de variables comme les proportions. Ils permettent par conséquent d'obtenir des variations de quantiles, donc de proportions de faciès, relativement lisses dans l'espace ;
- (b) **le seuil** : rappelons qu'il représente la variance maximale entre deux points. Dans notre cas, il correspond à la variance des composantes de la loi de Dirichlet estimée à partir des scénarios initiaux dont dérive les vecteurs de proportions ;
- (c) **la portée** : pour chaque direction d'anisotropie, elle traduit la distance maximale au-delà de laquelle il n'y a plus de corrélation entre deux points. Selon l'état des connaissances, on peut relier les portées aux caractéristiques des fairways, ou bien considérer des longueurs de corrélation plus grandes que la taille des corps modélisés (en pratique deux à trois fois la portée des variogrammes, ou les caractéristiques géométriques des faciès). Dans le cas présenté ici, la portée R_{1f} selon la direction principale du variogramme est fixée à une valeur plus grande que la longueur du modèle. On modélise en effet des corps qui traversent le modèle. Sa portée R_{2f} selon la deuxième direction du variogramme, perpendiculaire à la première, exprime l'incertitude sur la largeur des fairways. Elle est donc simulée dans une distribution triangulaire de support [800 m ; 2 km], et de mode 1400 m.

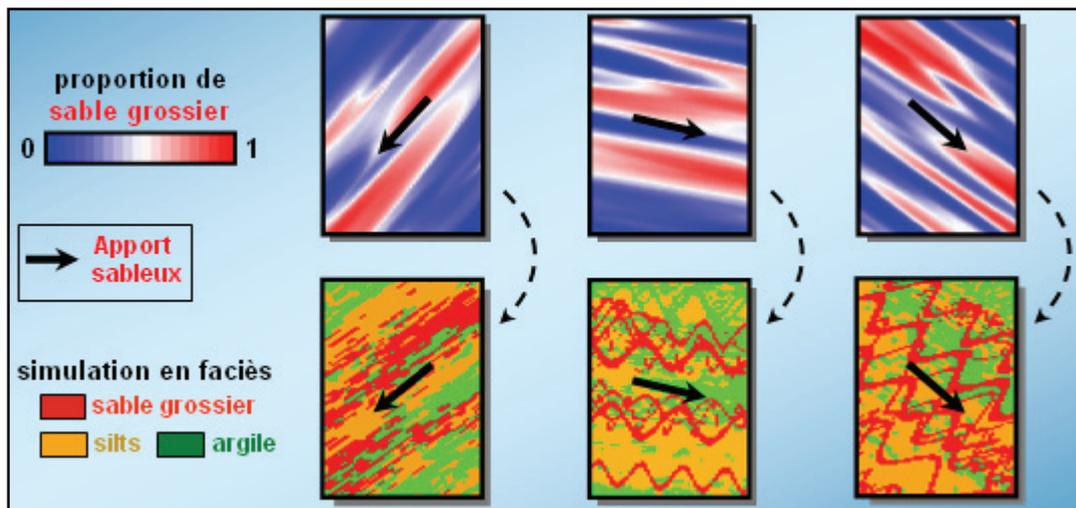


FIG. 9.13 : Simulations des zones préférentielles pour les chenaux en termes de proportions. Les paramètres des variogrammes sont tirés dans une distribution, permettant la simulation des quantiles, puis le calcul des proportions de faciès correspondantes. Pour chacune d'elle, une simulation en faciès est réalisée. Celle-ci conserve la direction des apports sédimentaires utilisées pour la simulation en quantiles.

La figure 9.13 présente les résultats obtenus en termes de cubes de proportions, puis sur les simulations en faciès basées pixels d'une part, et objets de l'autre. Sur toutes les ima-

ges concernant ce jeu de données et illustrant ce chapitre, seule la composante relative aux faciès de chenaux sera présentée. Nous aurions pu faire de même avec n'importe quelle autre composante, les conclusions restant identiques. Ces résultats expriment le caractère non stationnaire et incertain de la fonction aléatoire associée à chaque faciès. On peut ainsi simuler les faciès de chenaux dans ces zones préférentielles.

Un autre cas d'application

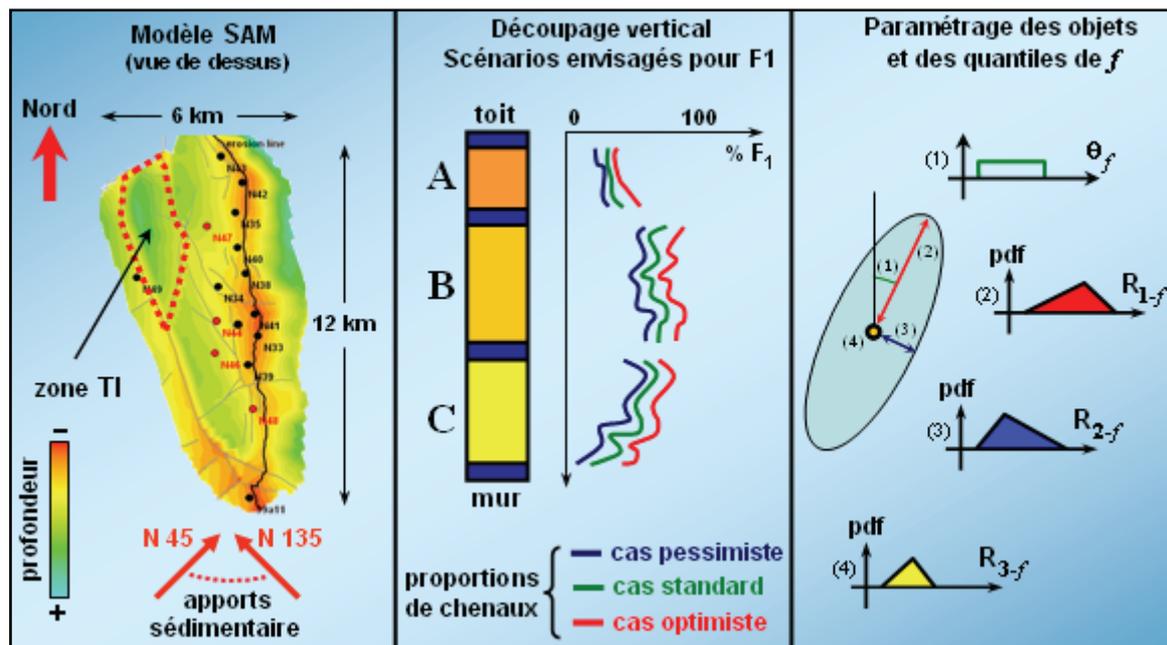


FIG. 9.14 : Présentation du cas d'application SAM. (1) Vue de dessus de la région d'étude, en particulier la zone à investiguer TI. L'incertitude du scénario géologique porte sur la direction des apports et la largeur des zones préférentielles. (2) Le découpage vertical du modèle en trois zones réservoir indépendantes A, B et C. Deux cas, pessimiste et optimiste, sont envisagés pour les proportions de faciès : ils permettent le calcul d'une loi de Dirichlet stationnaire sur chaque horizon. (3) Paramétrage des quantiles dont les grandeurs du variogramme associé sont tirées dans des distributions triangulaires déduites de celles des objets correspondants (direction, portées multipliées par 3).

Ce protocole a en outre été testé sur un cas réel, que l'on appellera SAM, dont le jeu de données est constitué de (Fig. 9.14) :

- une grille stratigraphique faillée qui discrétise une zone orientée selon un axe nord/sud en plus d'un million de cellules ($60 \times 100 \times 536$). Elle se divise verticalement en trois réservoirs indépendants notés (A), (B) et (C), eux-mêmes composés de plusieurs unités. Toutes correspondent à un intervalle stratigraphique indépendant des autres : la simulation des quantiles se fera donc indépendamment d'une unité à l'autre ;

- une quinzaine de puits, échantillonnés à l'échelle de la grille stratigraphique, et répartis pour la plupart à l'est du domaine le long d'un axe nord/sud ;
- un modèle sédimentaire qui distingue jusqu'à huit faciès discriminants en terme de pétrophysique, tous caractéristiques d'un milieu fluvatile. On retrouve, avec des qualités réservoirs décroissantes, des sables de chenaux imbriqués¹ ($F_1 \subset F_2 \subset F_3$), des silts de levées imbriqués également ($F_4 \subset F_5 \subset F_6$), ainsi que des argiles de lobes de débordement (F_7) ou bien de plaines d'inondation (F_8). Tous sont simulés au moyen d'une méthode *basée objets* ;
- un modèle pétrophysique qui simule indépendamment dans chaque faciès, la porosité selon les méthodes SGS avec un modèle de distribution paramétrique et un variogramme propre à chaque faciès. La perméabilité est cosimulée avec la porosité indépendamment par faciès en utilisant les histogrammes expérimentaux des données aux puits. Un paramètre multiplicatif corrige la perméabilité pour tenir compte de l'enfouissement. La hauteur utile (NTG) est égale à un, la saturation en huile est une fonction paramétrique de la perméabilité, de la porosité et de la profondeur ;
- un cube de proportions de base, qui représente la proportion des différents faciès, estimée le long des puits, en fonction de l'épaisseur stratigraphique ; la proportion des divers faciès est constante sur chaque couche.

La finalité de l'étude porte sur l'intérêt économique du forage d'un puits dans un panneau jusqu'ici inexploré (TI^2). Celui-ci se trouve au nord-ouest du champ, sachant que l'ensemble des puits est situé le long d'un axe nord-sud à l'est du champ. Il s'agit donc d'évaluer, en faisant l'hypothèse de non stationnarité des faciès, les accumulations dans le réservoir, notamment dans le domaine (TI), accumulations qui doivent être compatibles avec les incertitudes relevées sur :

- (1) la valeur globale des proportions de faciès dans chaque couche du réservoir, que l'on appréhende à partir du cube de proportions de base, (a) en l'améliorant en un scénario optimiste, et (b) en le dégradant en un modèle pessimiste. Ces deux cas permettent de définir les lois de Dirichlet en chaque point du modèle discrétisé. Cette étude présente d'ailleurs une particularité : tous les faciès ne sont pas forcément observés aux puits et modélisés dans chaque unité. Par exemple, la zone (A_1) compte quatre faciès : F_1 , F_2 , F_3 et F_8 . Les faciès F_4 , F_5 , F_6 et F_7 , absents des scénarios optimiste et pessimiste, conduisent par construction à des lois de Dirichlet qui présentent une composante nulle³ pour ces modalités. La simulation de telles lois attribuent naturellement aux faciès F_4 , F_5 , F_6 et F_7 une proportion nulle, prévenant de fait leur apparition dans cette région.

¹ Afin de prendre en compte des variations internes de lithologies.

² Comprenez « Terra Incognita ».

³ Plus exactement, l'estimation de la composante aléatoire de la loi de Dirichlet, qui correspond à chacun des faciès absents, conduit à une moyenne et une variance nulle ; elle peut s'interpréter comme un Dirac centré sur 0, dont la réalisation sera toujours nulle.

- (2) la provenance des apports sableux, dont l'azimut définit un angle avec le Nord compris entre 160 et 200 degrés. Par conséquent, on tire la direction principale θ_f des variogrammes, utilisés pour simuler les quantiles $Q_f(\cdot)$, dans une loi uniforme, qui décrit cette plage angulaire ;
- (3) l'extension des zones préférentielles dont on évalue les caractéristiques à partir de celles des faciès à simuler. Le modèle proposé distribue dans l'espace des objets géométriques dont les paramètres de taille sont stochastiques (distributions triangulaires). On utilise ces paramètres multipliés par trois pour contraindre les portées (R_{1f} et R_{2f}) des variogrammes utilisés dans la simulation des quantiles comme illustré à la figure 9.14 (à droite).

Une particularité de ce modèle repose sur l'imbrication des faciès à l'instar des sables de chenaux ($F_1 \subset F_2 \subset F_3$). En pratique, cela se traduit d'abord par la simulation des proportions du regroupement de faciès $F_c = F_1 + F_2 + F_3$, puis du ratio du faciès F_1 dans F_c , et de la proportion de F_1 dans F_c . Pour conserver la cohérence des objets imbriqués, les quantiles sont simulés à partir d'un germe identique. Le chemin aléatoire qu'il définit est le même pour tous les quantiles. D'un faciès imbriqué à l'autre (F_1, F_2, F_3), on échantillonne de cette manière des quantiles voisins pour chaque variable régionalisée Gamma associée aux différents faciès.

Une centaine de cubes de proportions ont été simulés (Fig. 9.15-a), à partir desquels on réalise une simulation en faciès, puis en propriétés pétrophysiques suivant les hypothèses décrites dans le modèle pétrophysique (Fig. 9.15-b).

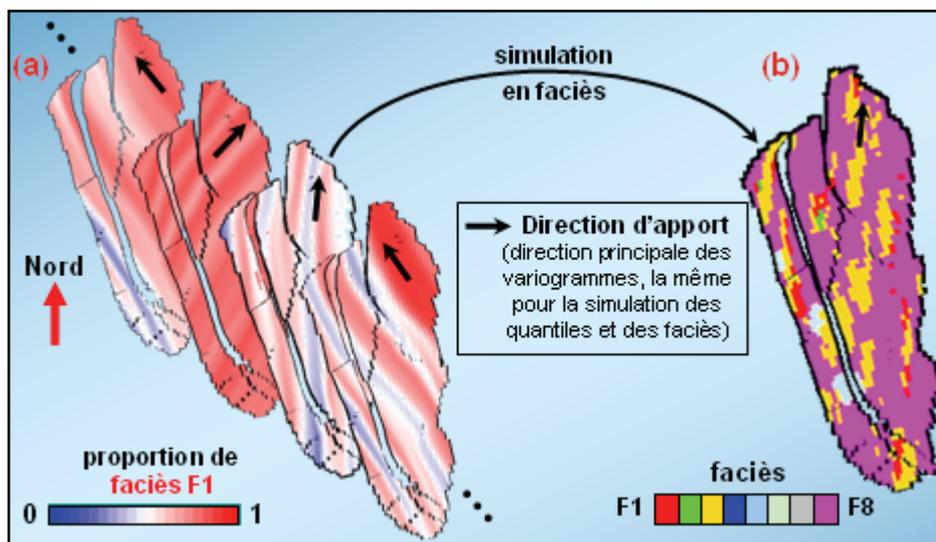


FIG. 9.15 : Description de l'incertitude liée au scénario géologique pour le modèle SAM. (a) Plusieurs réalisations en cubes de proportions (seule celle du faciès F_1 est représentée). (b) Pour une des réalisations (azimut du variogramme θ fixée), on simule les faciès selon une méthode basée « objets », leur orientation étant la même (égale à θ) pour préserver la cohérence du scénario géologique préalablement « simulé ».

Le calcul des accumulations montre l'incertitude liée aux proportions. Remarquons que la dispersion relative n'est pas si importante (Fig. 9.16). En fait, on la doit au modèle de faciès qui n'est pas discriminant en terme de pétrophysique. Les lois associées à la pétrophysique sont quasiment identiques pour tous les faciès. La variabilité introduite au niveau des proportions est en quelque sorte homogénéisée par le modèle pétrophysique.

	Zone B3		Zone C3	
	Totalité	Panneau TI	Totalité	Panneau TI
q_{10}	93,76	81,89	92,33	75,25
moyenne	100	100	100	100
q_{90}	105,29	118,77	108,97	123,49
σ/m	0,05	0,15	0,06	0,19

Fig. 9.16 : Volumes en place pour les différents cas. Les accumulations sont calculées après une unique simulation en faciès puis en pétrophysique.

9.3.3 Cas pratique n° 2 :

Simulation de cartes de proportions via la sismique

Dans certains cas, l'imagerie sismique permet de déterminer le contour des ceintures chenalisantes. La technique consiste à réaliser des vibrations en un point du sous-sol, puis d'en étudier la réponse en différents points de la zone. Le signal émis par la source se réfléchissant sur les différentes interfaces en fonction de l'angle d'incidence, il est alors possible d'imager le milieu souterrain en mesurant les amplitudes reçues sur les différents récepteurs. Leur traitement permet de relier l'amplitude sismique aux propriétés pétrophysiques des roches qui sont alors interprétées en termes de faciès (section 10.2). La cartographie issue de cette conversion reste néanmoins entachée d'incertitude, et ne peut être considérée aussi fiable que les données de puits.

Dans l'exemple suivant (Fig. 9.17), qui reprend le modèle de référence ZOUINE, une campagne sismique définit deux zones contrastées qui partagent le modèle du nord au sud. À l'ouest, un domaine (**Fw**) présente des proportions de chenaux élevées : il correspond à la ceinture chenalisante ; à l'est, un domaine (**Ovb**) aux proportions de chenaux plus faibles : ce sont les débordements. Bien que la direction des apports sédimentaires soit déterminée – en provenance du nord –, le modèle demeure toujours entaché d'incertitudes qui touchent entre autres :

- (1) la valeur globale des proportions de faciès dans chaque couche du réservoir, liée à l'analyse des amplitudes sismiques. On appréhende celle-ci à partir d'un cas de base (a) en le dégradant en un modèle pessimiste, avec dans le domaine (**Fw**), 20% de sables grossiers, 50 % de silts et 30 % d'argiles, tandis que le domaine (**Ovb**) présente 0% de sables, 20% de silts et 80% d'argiles, ou bien (b) en l'améliorant

en un scénario optimiste : dans la zone (**Fw**), sables grossiers (60%), silts (30%) et argiles (10%), et dans la région (**Ovb**), sables grossiers (10%), silts (50%) et argiles (40%) ;

- (2) la répartition spatiale des proportions de faciès à l'intérieur de la ceinture chenalisante, car on observe en général des changements de lithologies à mesure que l'on s'éloigne de l'axe de la ceinture, ce qui impose une certaine variabilité des proportions de faciès à l'intérieur de l'axe chenalisant ;
- (3) les limites latérales de la ceinture chenalisante, puisque la transition vers la zone de débordements (**Ovb**) ne peut être aussi nette, c'est-à-dire déterministe, en raison des incertitudes liées à la conversion de l'amplitude sismique en termes de proportions de faciès.

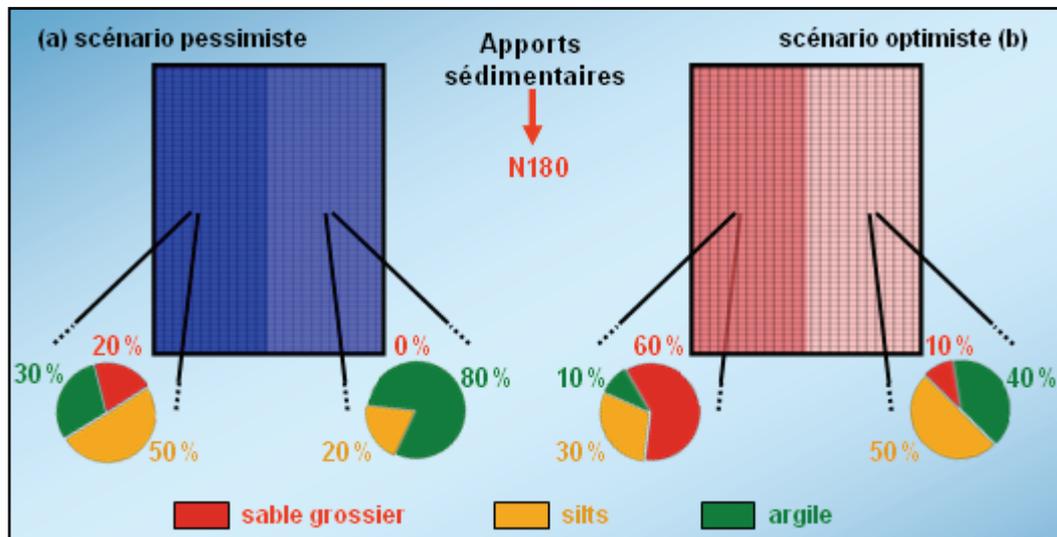


FIG. 9.17 : Cas pratique n° 2 : Incertitude sur les proportions au sein des fairways. La direction des apports sableux est connue, l'imagerie sismique fournit les contours des zones préférentielles. L'incertitude porte sur la valeur globale des proportions, la répartition spatiale des proportions à l'intérieur des ceintures chenalisantes.

La question consiste alors à décrire l'ensemble des modèles en proportions de faciès, qui se révèlent compatibles avec les incertitudes décrites précédemment. Pour cela, on combine les scénarios pessimiste et optimiste en leur attribuant un poids similaire. On définit alors des lois de Dirichlet stationnaires, respectivement dans les régions (**Fw**) et (**Ovb**). La variabilité des proportions de faciès est donc portée, d'une part par les lois de Dirichlet, puisqu'elles délimitent les régions : elles traduisent l'incertitude globale sur les proportions (1). D'autre part, elle peut s'expliquer à l'aide des variogrammes $\gamma_f(\mathbf{h})$, ils introduisent la non stationnarité des faciès à l'intérieur et entre les deux domaines : ils traduisent les incertitudes sur la variabilité interne des proportions (2) et sur la limite (**Fw**)-(Ovb) (3).

Ainsi, pour chaque faciès f , le variogramme $\gamma_f(\mathbf{h})$ utilisés est de type gaussien afin d'obtenir des variations de proportions lisses dans l'espace. Dans le cas proposé ici, l'orientation des apports étant connue, la direction principale du variogramme $\gamma_f(\mathbf{h})$ correspond à cet axe. Sa portée R_{1f} selon cette direction est déterministe et égale à 15 km. Elle dépasse la longueur du modèle car on simule des objets traversants. Sa portée R_{2f} selon la deuxième direction du variogramme, perpendiculaire à la première, est simulée dans une distribution triangulaire de support [800 m ; 2 km], et de mode 1400 m.

La figure 9.18 présente les résultats obtenus en termes de cubes de proportions, puis sur les simulations en faciès basées *pixels* d'une part, et *objets* de l'autre. On remarquera la définition d'une limite latérale de fairway relativement floue et variable d'une réalisation à l'autre en proportions. Ce qui traduit l'incertitude sur la façon de passer d'une zone (Fw) à l'autre (Ovb). Ces simulations permettent également d'appréhender la variabilité des proportions à l'intérieur de l'axe chenalisant, comme l'incertitude sur la valeur globale des proportions.

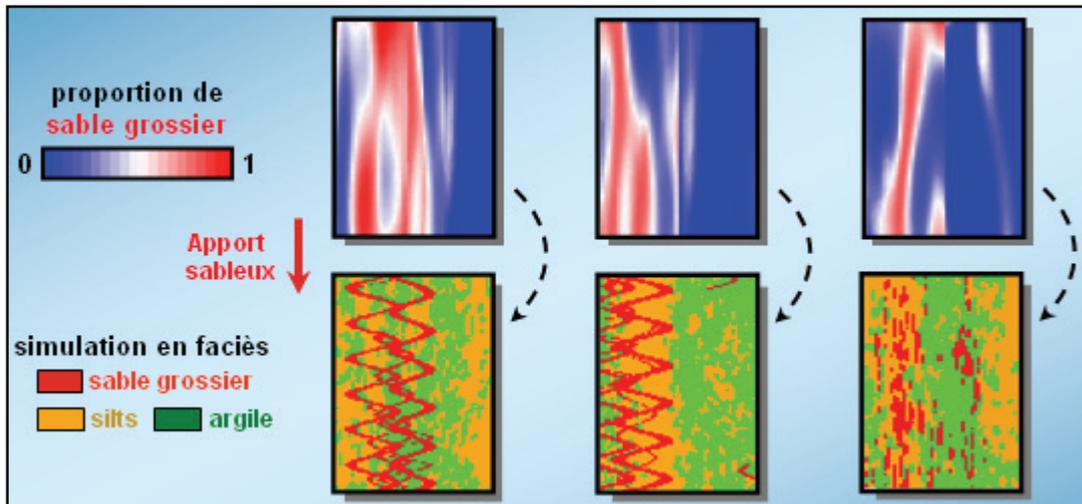


FIG. 9.18 : Incertitude sur la distribution des proportions au sein des axes chenalisants. Les paramètres des variogrammes (portée R_2 seulement) sont tirés dans une distribution probabiliste. Ce qui permet la simulation des quantiles selon la direction des apports sableux, puis le calcul des proportions de faciès correspondant. Pour chaque réalisation, une simulation en faciès est effectuée selon la direction d'apport.

9.3.4 Cas pratique n° 3 :

Intégration de plusieurs modèles de dérives

Dans certaines situations, le milieu peut présenter une tendance verticale ou horizontale, avec des variations progressives de faciès. Souvent, à l'intérieur d'une ceinture chenalisante par exemple, le milieu passe graduellement d'une prédominance de sédiments sableux à une prépondérance des dépôts argileux à mesure que l'on s'éloigne de son axe. Pour rendre compte de ces tendances, il est possible d'introduire des données

qualitatives telles que la connaissance du contexte géologique de la zone, du bloc voisin ou d'un champ analogue sous la forme de cartes de dérives en termes de proportions de faciès. Ces tendances peuvent également être appréhendées sur la base d'un modèle statistique inféré à partir des données. On l'obtient en général par régression linéaire, voisinage glissant¹, ou encore par krigeage de la moyenne. Néanmoins, le dessin précis de ces dérives reste entachée d'incertitudes.

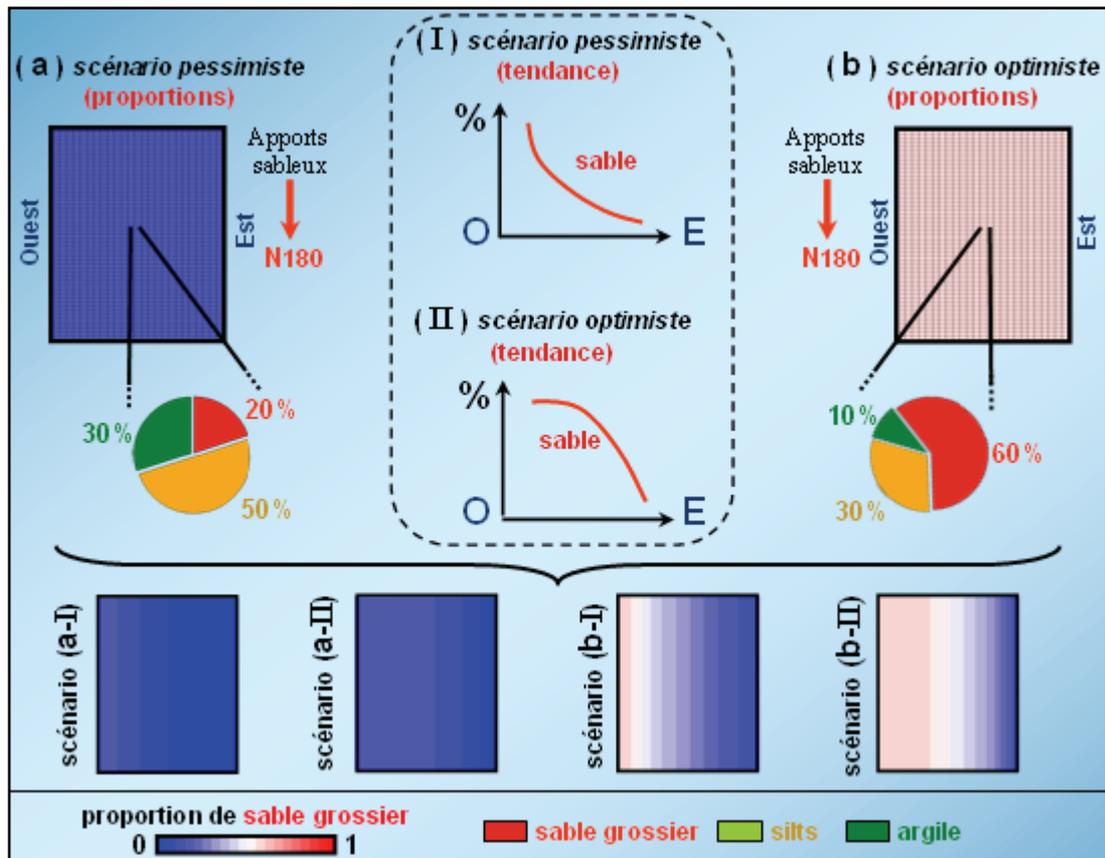


FIG. 9.19 : Cas pratique n° 3 : Incertitude sur le modèle de dérive en proportions. La direction des apports sableux est connue, comme la « dégradation » des faciès d'ouest en est. L'incertitude porte sur la valeur globale des proportions (scénarios (a) et (b)), ainsi que sur la forme de la dérive (hypothèses (I) et (II)). Leur combinaison permet la création de quatre modèles de base utilisé pour l'estimation des lois de Dirichlet.

Dans cet exemple qui reprend le modèle de référence ZOUINE, on connaît l'origine des apports sédimentaires : ils proviennent du nord. De plus, la géologie informe d'une dégradation des faciès réservoir d'ouest en est. Néanmoins, leur proportion comme leur évolution sont quant à elles moins bien connues. L'incertitude affecte précisément :

¹ **Voisinage glissant** : encore appelée moyenne mobile. C'est une moyenne qui au lieu d'être calculée sur n valeurs fixes, est calculée sur n valeurs consécutives « glissantes ».

- (1) la valeur globale des proportions de faciès dans chaque couche du réservoir, que l'on capture à partir d'un cas de base, (a) en l'améliorant en un scénario optimiste, avec une évolution des proportions d'ouest en est de 60 à 10% pour les sables grossiers, de 30 à 50% pour les silts et de 10 à 40% pour les argiles, ou bien (b) en le dégradant en un modèle pessimiste avec une évolution d'ouest en est, de 20 à 0% pour les sables, de 50 à 20% pour les silts et de 30 à 80% pour les argiles (Fig. 9.19-a,b) ;
- (2) l'évolution latérale des proportions sur chaque horizon, que l'on appréhende au moyen de deux hypothèses, respectivement pessimiste ou optimiste selon la décroissance plus ou moins « rapide » des proportions de faciès de chenaux d'ouest en est (Fig. 9.19-I,II).

La combinaison de ces différentes hypothèses permet de définir quatre scénarios de base (Fig. 9.19-a,b-I,II), qu'il est possible d'intégrer dans le formalisme de Dirichlet. En fait, cette situation se rapproche des cas pratiques rencontrés en environnements carbonatés. À l'instar de ceux-ci, ce sont plutôt les lois de Dirichlet régionalisées, extraites de scénarios de base relativement continus, qui portent la corrélation spatiale. De la même manière, on recourt soit (a) à un modèle de variogramme inféré à partir des modèles initiaux, ou (b) à une corrélation spatiale infinie qui définit un vecteur de quantile stationnaire pour chaque réalisation. La figure 9.20 présente les résultats obtenus en termes de cubes de proportions puis sur les simulations en faciès *basées pixels* puis *objets*. Ces simulations permettent d'appréhender l'incertitude sur la valeur globale des proportions et sur les modèles de dérive : elles définissent l'ensemble des cartes de tendances compatibles avec ces sources de variabilité.

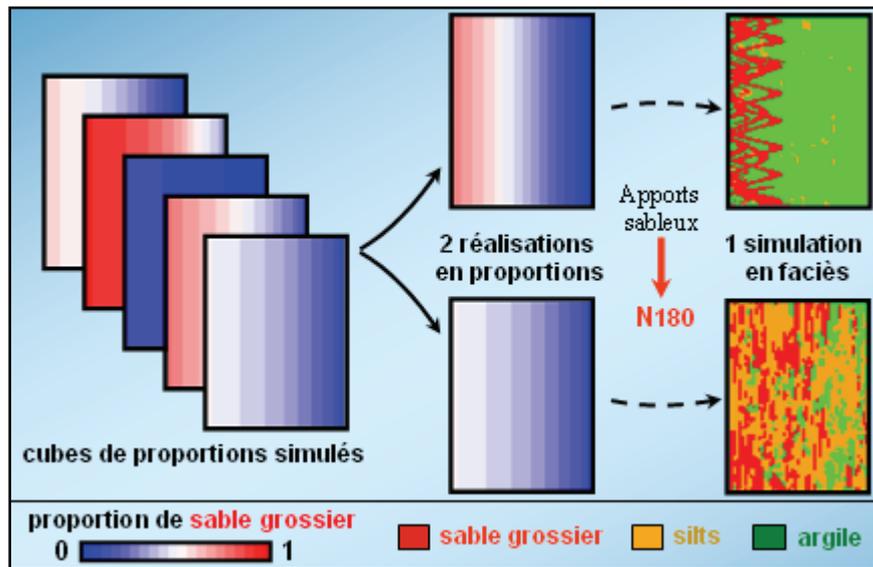


FIG. 9.20 : Incertitude sur le modèle de dérive horizontale. Pour réalisation et chaque faciès, un unique quantile est simulé permettant le calcul des proportions. Des simulations en faciès sont effectuées selon la direction des apports sableux (N180°).

Un autre cas d'application

Cette méthodologie a également été appliquée à un cas plus opérationnel, appelé LAKH'D, et qui se compose de :

- une grille stratigraphique faillée délimitant une zone carrée de 20 kilomètres de côté, qui comprend plus de 400 000 cellules ($105 \times 120 \times 33$), et divisée en trois réservoirs indépendants notés (R_1), (R_2) et (R_3) du plus profond au plus superficiel ;
- un ensemble de 19 puits, rééchantillonnés à l'échelle de la grille stratigraphique, et répartis au centre du domaine suivant un axe nord-est/sud-ouest ;
- un ensemble de trois faciès interprétés et échantillonnés le long des puits. Là encore, ces dépôts sont caractéristiques d'une sédimentation en milieu fluvatile, avec un encaissant argileux (F_0) et des sables déposés par des chenaux plus (F_2) ou moins (F_1) sinueux ;
- un modèle sédimentologique qui suppose les apports sableux en provenance du sud (vers l'azimut Nord 0°) et un modèle pétrophysique ;

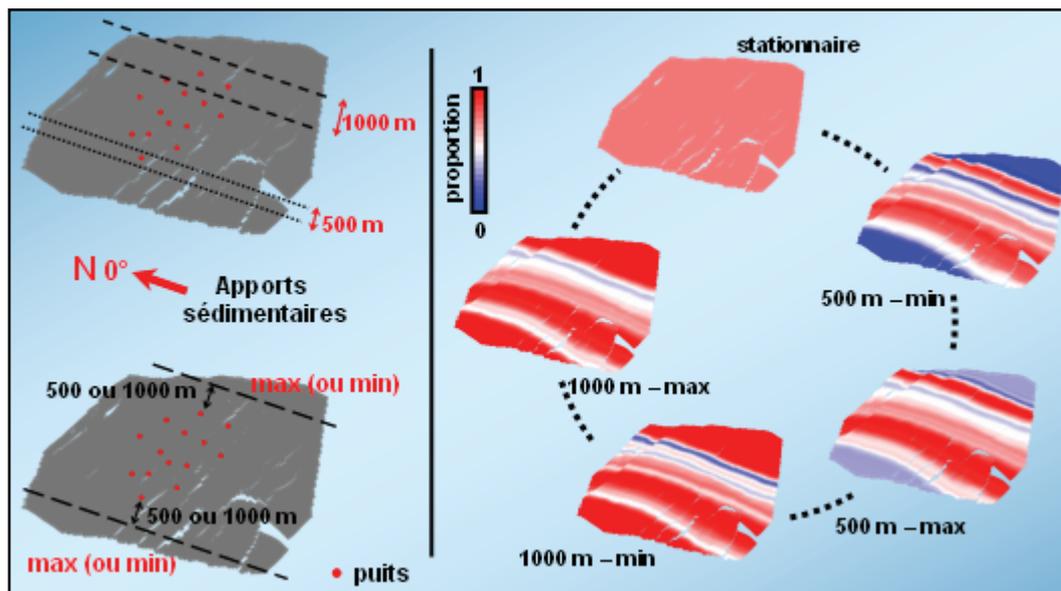


FIG. 9.21 : Cas LAKH'D : Construction des modèles de base. Les sédiments proviennent du sud. (a) Plusieurs dérives sont créées par voisinage glissant sur les données de puits. Une alternative porte sur la largeur des fenêtres (500 m ou 1000 m), une autre sur la proportion de sable dans les zones extérieures (maximum ou minimum des proportions de sable). (b) La combinaison de ces hypothèses décrit quatre modèles auxquels s'ajoute un cas stationnaire (moyenne des proportions aux puits).

- une série de cartes de tendance (Fig. 9.21) créée par voisinage glissant à partir des puits construites selon différentes hypothèses réalisées (1) sur la largeur de la fenêtre, 500 m ou 1000 m, et (2) sur la façon d'extrapoler les proportions dans les zones

non échantillonnées, pessimiste ou optimiste. La combinaison de ces hypothèses définit quatre scénarios possibles. On considère en plus un modèle stationnaire dans chaque couche correspondant aux proportions des différents faciès observés aux puits horizons par horizons ;

- les lois de Dirichlet dont dérivent les vecteurs de proportions $\mathbf{P}(\alpha)$ estimées en chaque point α à partir des cinq modèles de base cités précédemment, et pondérés de manière identique.

Une centaine de cubes de proportions (Fig. 9.22) ont ainsi été simulés à partir desquels on réalise une simulation en faciès, puis en propriétés pétrophysiques selon les hypothèses décrites dans le modèle pétrophysique. Pour chaque faciès, on simule indépendamment la porosité selon les méthodes SGS avec un modèle de distribution paramétrique et un variogramme propre à chaque faciès. La perméabilité est cosimulée par faciès avec la porosité en utilisant les histogrammes expérimentaux des données aux puits. Un paramètre multiplicatif corrige la perméabilité pour tenir compte de l'enfouissement. La hauteur utile (NTG) est égale à un, la saturation en huile est une fonction paramétrique de la perméabilité, de la porosité et de la profondeur

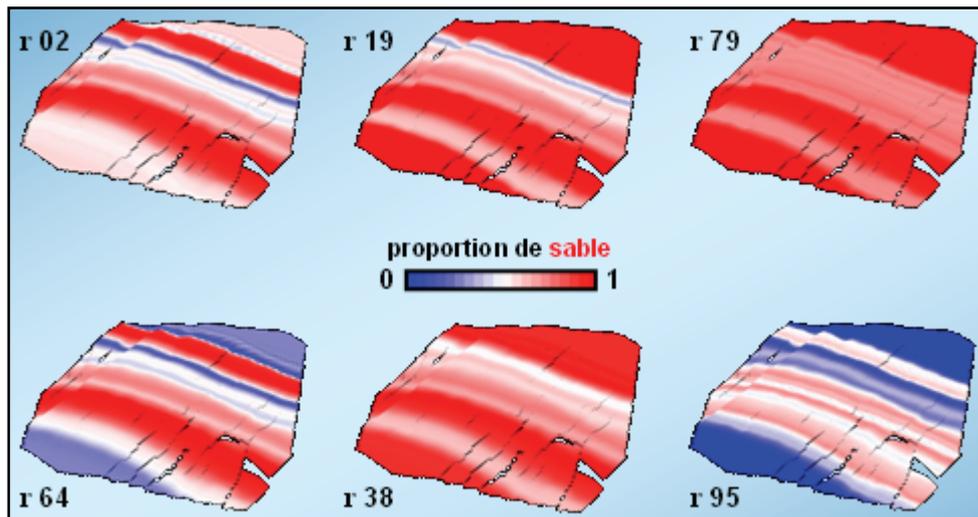


FIG. 9.22 : Cas LAKH'D : Six réalisations en proportions de faciès sableux. La construction des différentes réalisations passe par la simulation, respectivement pour chaque faciès, d'un quantile unique. Ceux-ci permettent alors le calcul des proportions, elles-mêmes utilisés pour simuler les images en faciès.

Les volumes moyens obtenus d'un scénario de base à l'autre (Fig. 9.23) affichent des différences significatives. On retrouve la variabilité liée au choix du modèle. D'un autre côté, la dispersion des accumulations reste relativement faible au sein d'un même scénario. Il s'agit de l'incertitude de nature aléatoire. Ainsi, la variabilité introduite par le choix du type de dérive, c'est-à-dire le cube de proportions, constitue bien la principale source d'incertitude sur les volumes d'hydrocarbures et leur distribution. Malgré tout, les résul-

tats acquis par combinaisons des scénarios montrent que le protocole proposé permet d’appréhender cette variabilité. Notons d’ailleurs que les dispersions relatives (rapport σ/m) des distributions définies respectivement selon les modes de combinaisons classiques (chapitre 2) et de Dirichlet restent proches l’une de l’autre. La seconde approche permet tout de même de balayer une plus grande plage de variabilité.

	Scénarios de base					Combinaison	
	<i>500 min</i>	<i>500 max</i>	<i>1000 max</i>	<i>1000 min</i>	<i>Statio.</i>	<i>Classique</i>	<i>Dirichlet</i>
q_{10}	91	74	77	123	92	78	69
m	100	82	85	131	100	100	101.9
q_{90}	109	90	92	139	109	131	134
σ/m	0.08	0.08	0.08	0.05	0.08	0.18	0.25

Fig. 9.23 : Cas LAKH'D : Comparaisons des volumes en place pour les différents scénarios. Pour ceux de base, les accumulations en hydrocarbures sont calculées par multi-réalisation pétrophysique. Les accumulations sont normalisées par rapport à la moyenne(m) observée sur le modèle stationnaire (Statio).

Remarque : Ce cas d’étude présente en plus une incertitude sur la direction des apports qui n’a pas été prise en compte ici. Néanmoins, initialement, pour intégrer cette incertitude, nous avons appliqué cette méthodologie en faisant varier de façon discrète la direction des apports : tous les dix degrés. Pour chaque direction, indépendamment des autres, on construit les cinq scénarios de base, puis on réalise les simulations de Dirichlet. Ce qui rend cette approche relativement lourde. C’est d’ailleurs la raison qui a conduit à appréhender l’anisotropie des modèles à travers les variogrammes (cas pratique n° 1). Ceux-ci permettent en effet d’intégrer de manière continue ce type d’hypothèses dans le choix de son azimut.

Le recours au cadre géostatistique permet de conserver la cohérence spatiale qui définit les relations dans l’espace entre vecteurs de proportions. Combinée aux lois de Dirichlet préalablement estimées, cette approche simule des vecteurs de proportions spatialement corrélés. La corrélation est introduite lors de la construction des vecteurs de quantiles, qui sont alors simulés au moyen de méthodes stochastiques basées sur la notion de variogramme. Son choix dépend naturellement de l’environnement de dépôt, selon qu’il s’agit de sédimentation carbonatée ou silico-clastique, comme il dépend de la nature et du degré d’information ou de connaissance du milieu. Une telle approche nécessite la définition des variogrammes qui traduisent la continuité des propriétés étudiées. On construit ainsi des cubes de proportions cohérents en terme de géologie puisqu’ils respectent la continuité des formations géologiques sous-jacentes, et compatibles avec l’incertitude géologique. Cependant, à l’instar des méthodes de multi-réalisation, elle tend à élargir l’espace des modèles possibles en termes de cubes de proportions. Or, il existe des sources

d'information alternatives qui peuvent servir à préciser l'espace d'incertitude lié à la définition du scénario géologique (**chapitre 10**).

Chapitre 10

Mise en cohérence des cubes de proportions

Les simulations décrites aux chapitres précédents définissent l'ensemble des cubes de proportions compatibles avec l'incertitude géologique. Néanmoins, on bénéficie aussi de données autres que la géologie : puits, imagerie sismique, historique de production, etc. L'information supplémentaire qui en est déduite peut servir à mettre en cohérence les divers cubes de proportions générés. Elle permet effectivement (a) de contraindre les réalisations en amont du processus de simulation, et/ou (b) de corriger les réalisations en aval du procédé. Ce chapitre revient sur la manière d'appréhender et d'intégrer à la méthodologie proposée des sources d'information additionnelles, – qui elles-mêmes peuvent être incertaines. L'objectif consiste à préciser l'espace d'incertitude lié à la définition du scénario géologique. Il s'agit ainsi de contraindre les cubes de proportions simulés grâce au formalisme de Dirichlet à l'information complémentaire extraite des données de puits ou de l'imagerie sismique. Les informations supplémentaires sont dans un premier temps supposées « exactes » avant d'être considérées avec leur incertitude.

10.1 Intégration des données de puits

10.1.1 Nature des contraintes

Les données de puits (carottes, diagraphies, etc.) apportent une information à haute résolution verticale, bien que localisée sur les zones de forages et un volume rocheux limité. Elles correspondent malgré tout à des observations directes du milieu souterrain et à ce titre, sont généralement supposées exactes. Elles sont utilisées comme données dures, *hard data* en anglais, afin de contraindre la modélisation du sous-sol. C'est en particulier le cas des faciès observés ou interprétés le long des puits. Ils servent d'ordinaire de données de conditionnement lors de la simulation des faciès dans le reste du réservoir : pour chaque réalisation, le faciès simulé en un point qui coïncide avec une observation corres-

pond au faciès réellement observé. De la même façon, il est possible de construire un modèle en proportions de faciès qui intègre de manière cohérente l'information disponible aux puits, en particulier les logs de faciès. Pour cela, on s'intéresse, parmi les cellules qui composent le modèle discrétisé, à celles qui sont traversées par les puits. Pour chacune de ces mailles notées α_w , on peut conditionner les vecteurs de proportions $\mathbf{p}(\alpha_w)$ simulés selon l'interprétation des données de puits privilégiée :

- (1) soit au faciès supposé présent dans la cellule, déterminé classiquement par les méthodes de changement d'échelle¹ – *upscaling* en anglais –, des données de faciès observés aux puits dans le volume définie par la cellule ;
- (2) soit aux proportions de faciès observées aux puits, et calculées à la résolution de la cellule considérée.

Dans les deux cas, la démarche adoptée pour faire respecter de telles contraintes dans les cellules α_w traversées par les puits, reste identique. Seul diffère le vecteur de conditionnement² $\mathbf{p}^*(\alpha_w) = [p_1^*(\alpha_w), \dots, p_{nf}^*(\alpha_w)]^T$ affecté à chaque maille α_w , et qui représente le vecteur de proportions supposé dans celle-ci.

La première situation réconcilie directement les données dures de faciès avec les vecteurs de proportions des cellules intersectées par les puits. Pour chacune des mailles α_w , on renseigne les valeurs des composantes du vecteur de conditionnement $\mathbf{p}^*(\alpha_w)$ de telle façon que pour chaque faciès f :

$$p_f^*(\alpha_w) = \begin{cases} 1 & \text{si le faciès } f \text{ est rencontré au noeud } \alpha_w \\ 0 & \text{sinon} \end{cases} \quad (\text{Éq. 10.1})$$

En d'autres termes, en chaque cellule investiguée par un puits où l'on observe le faciès d'indice f , le vecteur de proportions traduit une probabilité de 100% de rencontrer ce faciès. Il ramène au contraire cette probabilité à 0% dans les autres cellules « puits » où ce faciès n'a pas été observé. Par exemple, la figure 10.1, qui décrit un contexte avec trois faciès, s'intéresse plus particulièrement à la cellule α_w . Elle est traversée par un puits, et on considère le faciès d'indice 3 comme donnée dure (Fig. 10.1-a). Dans ce cas, le vecteur de conditionnement $\mathbf{p}^*(\alpha_w)$ correspondant est défini par :

$$\mathbf{p}^*(\alpha_w) = \begin{bmatrix} p_1^*(\alpha_w) \\ p_2^*(\alpha_w) \\ p_3^*(\alpha_w) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad (\text{Éq. 10.2})$$

¹ Il s'agit de savoir comment représenter une propriété variable dans un volume fini par une valeur unique attachée à ce volume. Diverses méthodes existent : pour le faciès, le plus fréquent peut être choisi, mais il est aussi possible de choisir le faciès observé le plus proche du barycentre de la cellule. C'est d'ailleurs ces faciès qui servent de données de conditionnement lors de la simulation des propriétés lithologiques.

² Par analogie avec les méthodes géostatistiques, on pourrait parler de vecteur de proportions « dur ».

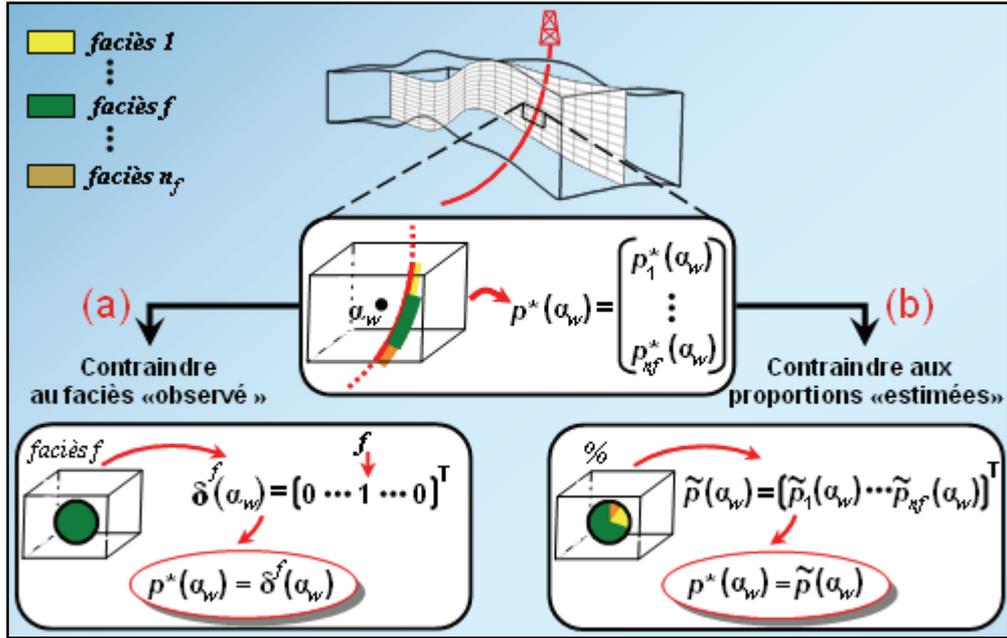


FIG. 10.1 : Deux jeux de contraintes possibles aux puits pour les proportions. Le vecteur de proportions $\mathbf{p}(\alpha_w)$ simulé dans la cellule α_w traversée par un puits doit correspondre à un vecteur $\mathbf{p}^*(\alpha_w)$, qui en (a) traduit le fait que l'on observe le faciès f , et qui en (b) renvoie à un estimateur des proportions de faciès.

La seconde approche privilégie plutôt les proportions supposées aux puits. Pour chaque cellule α_w traversée par les puits, l'examen des points de données localisés dans celle-ci peut fournir un estimateur $\tilde{\mathbf{p}}(\alpha_w) = [\tilde{p}_1(\alpha_w), \dots, \tilde{p}_{n_f}(\alpha_w)]^T$ du vecteur de proportions dans ce volume. Dans cette cellule, on renseigne les valeurs des composantes du vecteur de conditionnement $\mathbf{p}^*(\alpha_w)$ de telle façon que :

pour chaque faciès f ,

$$p_f^*(\alpha_w) = \tilde{p}_f(\alpha_w) \quad (\text{Éq. 10.3})$$

Si on reprend l'exemple de la figure 10.1, la cellule α_w contient plusieurs points de données à partir desquels on calcule la proportion des différents faciès, qui par faciès d'indice croissant vaut respectivement 10, 20 et 70 % (Fig. 10.1–b). Le vecteur de conditionnement $\mathbf{p}^*(\alpha_w)$ correspondant est donc défini par :

$$\mathbf{p}^*(\alpha_w) = \begin{bmatrix} p_1^*(\alpha_w) \\ p_2^*(\alpha_w) \\ p_3^*(\alpha_w) \end{bmatrix} = \begin{bmatrix} 0,1 \\ 0,2 \\ 0,7 \end{bmatrix} = \begin{bmatrix} \tilde{p}_1(\alpha_w) \\ \tilde{p}_2(\alpha_w) \\ \tilde{p}_3(\alpha_w) \end{bmatrix} = \tilde{\mathbf{p}}(\alpha_w) \quad (\text{Éq. 10.4})$$

En définitive, l'objectif commun à ces deux interprétations consiste à faire en sorte que pour chaque cellule α_w traversée par un puits, les réalisations $\mathbf{p}(\alpha_w)$ du vecteur de propor-

tion $\mathbf{P}(\alpha_w)$ correspondent toujours au vecteur de conditionnement $\mathbf{p}^*(\alpha_w)$, c'est-à-dire :

$$\mathbf{p}(\alpha_w) = \begin{bmatrix} p_1(\alpha_w) \\ \vdots \\ p_{nf}(\alpha_w) \end{bmatrix} = \begin{bmatrix} p_1^*(\alpha_w) \\ \vdots \\ p_3^*(\alpha_w) \end{bmatrix} = \mathbf{p}^*(\alpha_w) \quad (\text{Éq. 10.5})$$

Pour cela, il suffit de considérer comme fixé le vecteur de proportions en chaque cellule α_w quand on simule les cubes de proportions alternatifs. Or, l'algorithme utilisé pour leur construction (chapitres 8 et 9) montrent que le respect d'un vecteur de proportions figé $\mathbf{p}^*(\alpha_w)$ peut se faire à l'échelle : soit des vecteurs de quantiles, ou bien des lois de Dirichlet. Dans les deux cas, ce conditionnement, dont un exemple précise la mise en œuvre, s'effectue en amont du processus de simulation des cubes de proportions.

10.1.2 Mise en œuvre sur un exemple

Considérons par exemple le modèle de référence ZOUINE déjà décrit au chapitre précédent (section 9.3). On réalise l'hypothèse supplémentaire de l'existence d'un puits au centre du domaine. Il investigate les cellules notées α_w sur lesquelles sont déterminés les vecteurs de conditionnement $\mathbf{p}^*(\alpha_w)$ qu'il s'agit d'honorer. Cependant, la mise en cohérence de chaque réalisation aux données de puits diffère selon la manière de simuler les cubes de proportions. Nous distinguerons par conséquent le cas (1) des simulations géostatistiques des quantiles de celui (2) d'une corrélation spatiale infinie de ces mêmes quantiles.

Cas des simulations géostatistiques de quantiles

Pour rappel, la première situation correspond à celles rencontrées dans les cas pratiques n° 1 et 2 (sections 9.3.2 et 9.3.3). Plus généralement, elle renvoie à des modèles où sont définis des vecteurs de quantiles qui varient au sein d'une couche¹, la simulation de leurs composantes supposant une corrélation spatiale finie. Les méthodes géostatistiques SGS (section 9.1) utilisées pour cela, peuvent être conditionnées, c'est-à-dire assurer que chaque réalisation honore d'éventuels points de données. Or, chacun des vecteurs de conditionnement $\mathbf{p}^*(\alpha_w) = [p_1^*(\alpha_w), \dots, p_{nf}^*(\alpha_w)]^T$ représente une réalisation particulière du processus de Dirichlet. En effet, un vecteur de conditionnement est issu de l'anamorphose (section 8.3.3) des composantes d'un vecteur de quantiles particulier que l'on notera $\mathbf{q}^*(\alpha_w) = [q_1^*(\alpha_w), \dots, q_{nf}^*(\alpha_w)]^T$. Par conséquent, le respect des vecteurs $\mathbf{p}^*(\alpha_w)$ peut se faire en contraignant pour chaque faciès f , la simulation SGS du quantiles $Q_f(\cdot)$ aux données dures $q_f^*(\alpha_w)$, ces données dures étant préalablement obtenues à partir des composantes $p_f^*(\alpha_w)$. Cette approche se décompose en deux étapes. En reprenant les notations du chapitre 8, elles se résument à l'algorithme suivant (Fig. 10.2) :

¹ C'est également le cas des réservoirs carbonatés dont on modélise la continuité spatiale à partir d'un variogramme inféré (paragraphe 9.2.2).

- (1) dans chaque cellule α_w , on détermine la valeur des réalisations respectives des lois Gamma, $\Gamma_1(\alpha_w), \dots, \Gamma_{nf}(\alpha_w)$, dont les rapports renvoient aux composantes du vecteur de conditionnement. En posant $\Gamma(\alpha_w) = \Gamma_1(\alpha_w) + \dots + \Gamma_{nf}(\alpha_w)$, qui en première approximation peut être considéré comme la somme des moyennes des lois gamma selon la relation $\Gamma(\alpha_w) = E[\Gamma_1(\alpha_w)] + \dots + E[\Gamma_{nf}(\alpha_w)]$ où $E[\cdot]$ est l'espérance, on réalise pour chaque faciès f :
 - (2) le calcul du produit, $\Gamma_f(\alpha_w) = p_f^*(\alpha_w) \times \Gamma(\alpha_w)$;
 - (3) l'application de la fonction de répartition de la loi Gamma $\Gamma_f(\alpha_w)$ de paramètres $[1, c_f(\alpha_w)]$ à l'abscisse $\Gamma_{nf}(\alpha_w)$, $q_f^*(\alpha_w) = \Pi_f[\alpha_w](\Gamma_{nf}(\alpha_w))$;
 - (4) pour chaque quantile d'indice f , on simule une réalisation du quantile $Q_f(\cdot)$ au moyen des méthodes SGS, celles-ci étant conditionnées aux points de données α_w par les valeurs $q_f^*(\alpha_w)$.

À partir de cette réalisation conditionnée des vecteur de quantiles dans l'espace, il est possible de déterminer le cube de proportions qui en découle. Ce modèle étant alors lui-même contraint aux données de puits.

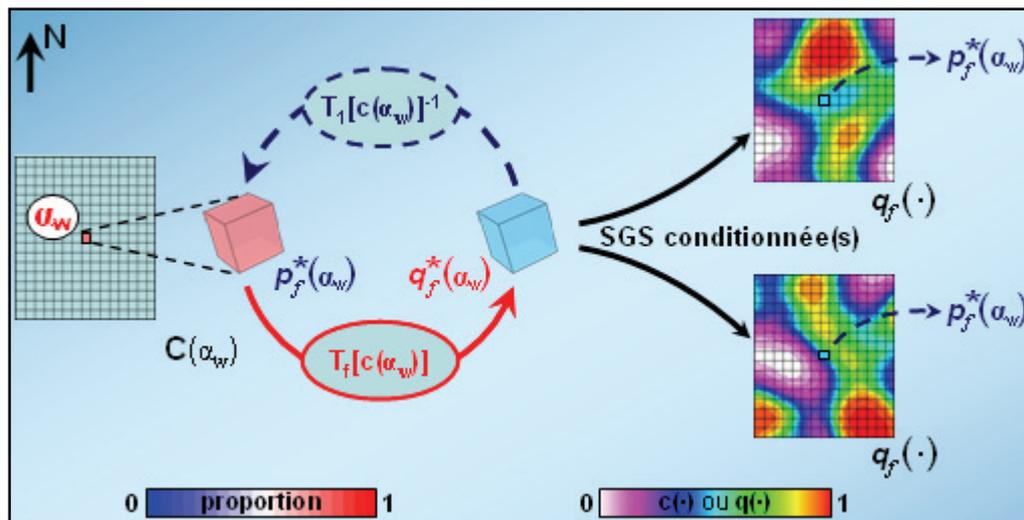


FIG. 10.2 : Simulations de quantiles conditionnés aux puits. Au point α_w traversé par un puits, pour chaque faciès f , la composante $p_f(\alpha_w)$ du vecteur de conditionnement $\mathbf{p}^*(\alpha_w)$ est convertie en quantile $q_f^*(\alpha_w)$ par la transformation $\mathbf{T}_f[\mathbf{c}(\cdot)]$ décrite précédemment. C'est le quantile, qui par la transformation inverse $\mathbf{T}_f[\mathbf{c}(\cdot)]^{-1}$ renvoie à $p_f(\alpha_w)$. Le respect des vecteurs $\mathbf{p}^*(\alpha_w)$ s'effectue en contraignant pour chaque faciès f , la simulation SGS du quantiles $Q_f(\cdot)$ aux données dures $q_f^*(\alpha_w)$.

Cas d'une corrélation infinie des quantiles

La seconde situation correspond davantage au cas pratique n° 3 décrit à la section 9.3.4. Elle renvoie à des modèles où sont définis des vecteurs de quantiles constants au

sein d'une couche¹, la simulation de leurs composantes supposant une corrélation spatiale infinie. De ce fait, le conditionnement aux vecteurs $\mathbf{p}^*(\alpha_w)$ passe par les lois de Dirichlet. Il s'agit de s'assurer que par construction, dans toute cellule α_w traversée par un puits, le vecteur de proportions $\mathbf{P}(\alpha_w)$ suit une loi de Dirichlet de moyenne $\mathbf{p}^*(\alpha_w)$ et de variance « nulle ». Pour cela, il faut que le paramètre $\mathbf{c}(\alpha_w) = [c_1(\alpha_w), \dots, c_{nf}(\alpha_w)]^T$ de la distribution corresponde au vecteur $[p_1^*(\alpha_w) \times N, \dots, p_{nf}^*(\alpha_w) \times N]^T$, où N est un entier qui tend vers l'infini. Concrètement, les $n_m(G)$ modèles de base à partir desquels sont inférés les lois de Dirichlet doivent tous être contraints aux cellules « puits ». Pour chacun de ces modèles d'indice k , en notant $\mathbf{p}(k, \alpha_w) = [p_1(k, \alpha_w), \dots, p_{nf}(k, \alpha_w)]^T$ le vecteur de proportions au point α_w , celui-ci vérifie la relation :

$$\mathbf{p}(k, \alpha_w) = \begin{bmatrix} p_1^*(k, \alpha_w) \\ \vdots \\ p_3^*(k, \alpha_w) \end{bmatrix} = \begin{bmatrix} p_1^*(\alpha_w) \\ \vdots \\ p_3^*(\alpha_w) \end{bmatrix} = \mathbf{p}^*(\alpha_w) \quad (\text{Éq. 10.6})$$

Dès lors, quel que soit le vecteur de quantiles \mathbf{q} simulé², le vecteur de proportions $\mathbf{p}(\alpha_w)$ qui en dérive coïncide avec $\mathbf{p}^*(\alpha_w)$. Le conditionnement des cubes de proportions aux données de puits s'effectue directement sur les modèles de base, qui naturellement doivent être cohérents avec cette information. Cette approche s'applique d'ailleurs au cas des simulations géostatistiques de quantiles évoqué précédemment. En revanche, le recours aux simulations conditionnelles s'avère inutile, car les lois de Dirichlet étant déjà contraintes, les techniques SGS non conditionnées suffisent.

10.1.3 Des données dures incertaines !

Dans ce qui précède, on suppose exactes les données de puits. Néanmoins, elles portent naturellement une part d'incertitude qui s'explique entre autres par :

- la subjectivité implicite des données, car elles résultent d'une interprétation ou d'un étalonnage des observations ou des mesures effectuées aux puits. Par exemple, on utilise les logs de puits pour interpréter la succession des faciès le long des forages, et ce sont alors ces faciès qui font office de données dures ;
- la différence de résolution entre d'un côté les mesures accomplies le long des puits et le support géométrique, c'est-à-dire les cellules qui composent le modèle discrétisé. Les puits investiguent un volume bien inférieur à celui des cellules qu'ils traversent. On effectue donc un changement d'échelle implicite en supposant la mesure réalisée aux puits représentative du volume délimité par la cellule où elle se situe ;

¹ C'est également le cas des réservoirs carbonatés dont on modélise la continuité spatiale à partir de variogrammes de portées « infinies » (paragraphe 9.2.3).

² On rappelle en effet que dans le cas d'une corrélation infinie, en tout point α , pour chaque réalisation, on ne simule qu'un vecteur de quantiles \mathbf{q} , le même pour toutes les cellules α .

- l'existence de plusieurs estimateurs $\tilde{p}(\alpha_w) = [\tilde{p}_1(\alpha_w), \dots, \tilde{p}_{nf}(\alpha_w)]^T$ du vecteur de proportions dans la cellule α_w traversée par les puits, allant de la moyenne arithmétique des données observées à un krigeage ordinaire de la moyenne des indicatrices représentatives des différents faciès.

D'ailleurs dans le cas d'un calage des modèles de réservoir à un historique de production, celui-ci nécessite souvent d'introduire une certaine variabilité, ou degré de liberté, sur les données dures. Car, le fait de fixer les données de puits complique, ou empêche généralement tout ajustement [Hoffman, 2005]. Akamine et Caers (2006) proposent donc de simuler les données dures de faciès en considérant leur incertitude, et en respectant leur continuité spatiale. Par conséquent, pour tenir compte de la variabilité des contraintes aux puits, on honore, au moyen des méthodes décrites précédemment (section 10.1.2), chaque simulation de cube de proportions à un jeu de données dures, qui aura été préalablement simulé selon une technique analogue à celle d'Akamine et Caers (Fig. 10.3).

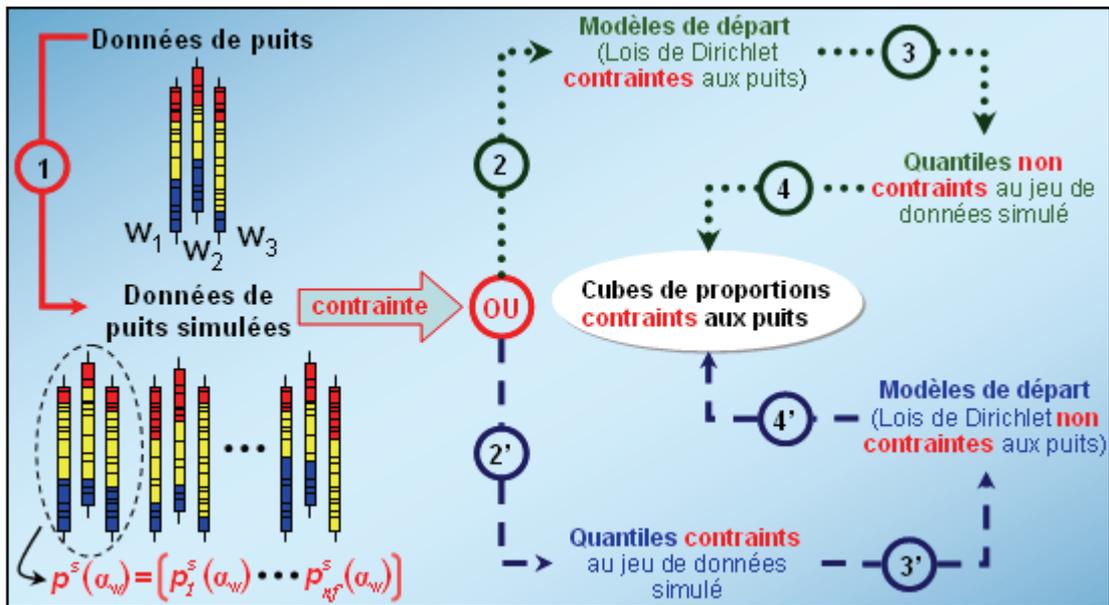


FIG. 10.3 : Intégration de l'incertitude sur les données de puits. Pour toute réalisation en cubes de proportions, on simule d'abord (1) un jeu de données aux puits (vecteur noté $p^s(\alpha_w)$) selon la méthode proposée par Akamine et Caers (2006). Le respect de ces données alternatives de puits imposent :

... soit de construire $n_m(G)$ modèles de base qui coïncident avec les valeurs simulées aux puits (i.e $p^s(\alpha_w)$). On définit ainsi (2) des lois de Dirichlet contraintes aux puits (moyenne égale à $p^s(\alpha_w)$ et variance nulle). Ces lois combinées (3) à des simulations SGS de quantiles non conditionnées fournissent (4) des cubes de proportions correspondants au jeu de données simulé aux puits ;

-- ou bien de simuler des quantiles contraints (2') à ces données de puits – i.e $p^s(\alpha_w)$. Ces quantiles, combinés (3') aux lois de Dirichlet issues des $n_m(G)$ modèles de base non contraints aux puits, définissent (4') des cubes de proportions qui respectent le jeu de données simulé aux puits.

10.2 Intégration des cartes sismiques

10.2.1 Nature des contraintes

Les données sismiques sont des mesures de variables secondaires (attributs ou impédances), qui peuvent être reliées à la propriété considérée. Elles apportent une information indirecte sur les propriétés du milieu souterrain, information qui, contrairement aux données de puits, offre une résolution bien plus faible. Ces données proposent cependant une couverture extensive du domaine étudié. Elles sont ainsi couramment utilisées en tant que contraintes complémentaires (données *soft* en anglais) pour caractériser le sous-sol [Xu, 1992 ; Johann et al., 1996]. C'est notamment le cas des cartes d'attributs ou d'impédance acoustique définies par l'imagerie sismique à la résolution verticale du réservoir. L'examen des faciès et des données sismiques localisés uniquement le long des puits permet en effet de construire une relation probabiliste entre les occurrences de faciès et les grandeurs sismiques (\mathbf{D}_S) :

$$\mathbb{P}(F = f | \mathbf{D}_S = \mathbf{d}_S) = \psi_f(\mathbf{d}_S) \quad (\text{Éq. 10.7})$$

$\mathbb{P}(F = f | \mathbf{D}_S)$ étant la probabilité conditionnelle d'observer le faciès f connaissant les données sismiques \mathbf{D}_S ;

Avec pour toute valeur ou gamme de grandeurs sismiques \mathbf{d}_S , $\psi_1(\mathbf{d}_S) + \dots + \psi_{n_f}(\mathbf{d}_S) = 1$ et pour tout indice f , $0 \leq \psi_f(\mathbf{d}_S) \leq 1$.

La définition des fonctions ψ_f peut se faire selon diverses méthodes plus ou moins élaborées : classification sur diagrammes, analyse en composantes principales, régression multivariée, etc. Une approche classique consiste ainsi à regrouper les observations aux puits en classes d'éléments ayant des caractéristiques communes en termes d'attributs sismiques. Pour cela, on utilise les méthodes traditionnelles de classification [Doyen *et al.*, 1989 ; Pairazian, 1998] à l'instar des nuées dynamiques [Saporta, 1990]. Pour chaque groupe $\{S_k\}$ ainsi défini, on calcule la fréquence de tous les faciès f . Ce qui fournit une estimation de la distribution conditionnelle des faciès pour la classe d'attributs $\{S_k\}$, c'est-à-dire la valeur de la fonction ψ_f dans la gamme de grandeurs délimitée par $\{S_k\}$ ¹.

La relation probabiliste obtenue est alors utilisée pour spécifier la cartographie des variations latérales de la probabilité d'apparition des faciès [Mao et al., 1999], simplement en appliquant les fonctions ψ_f à l'ensemble de la carte sismique. Dans l'exemple précédent, chaque point de la carte sismique, où l'on dispose des mesures sismiques, est classé dans un des groupes $\{S_k\}$, et se voit attribuer la distribution conditionnelle du groupe auquel il appartient.

À présent, si on considère une grille stratigraphique de taille $(n_u \times n_v \times n_w)$ (Fig. 10.4), celle-ci se décompose en $(n_u \times n_v)$ blocs verticaux, ou colonnes stratigraphiques, de coordonnées (u, v) . Dans la pratique, on suppose souvent que les cartes de probabilités

¹ Pour toutes gammes d'attributs sismiques \mathbf{d}_S appartenant à la classe $\{S_k\}$, $\psi_f(\mathbf{d}_S) = \mathbb{P}(F = f | \mathbf{D}_S \in S_k)$

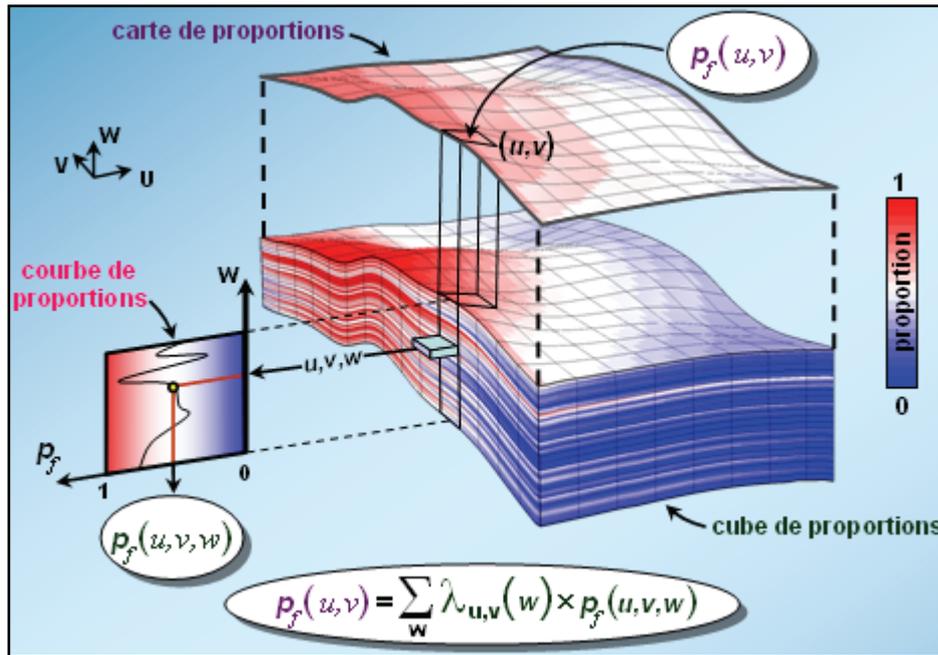


FIG. 10.4 : Définition des contraintes sismiques relativement au cube de proportions. La sismique fournit pour chaque faciès f une carte correspondant à sa proportion $p_f(u, v)$ pour toutes les colonnes stratigraphiques de coordonnées (u, v) du cube de proportions, celui-ci définissant en chaque cellule (u, v, w) la proportion de faciès.

fournies par la sismique représentent respectivement les composantes de vecteurs de proportions $\mathbf{p}(u, v) = [p_1(u, v), \dots, p_{n_f}(u, v)]^T$ pour l'ensemble des colonnes de coordonnées (u, v) de la grille¹. Pour chaque faciès f , la quantité $p_f(u, v)$ correspond à sa proportion dans le volume défini par la colonne (u, v) . L'ensemble de ces cartes – que nous appellerons cartes de proportions² –, offrent ainsi la possibilité de contraindre les cubes de proportions simulés selon le formalisme de Dirichlet. En effet, les vecteurs de proportions $\mathbf{p}(u, v)$ portés par cette carte peuvent être exprimés en fonction des vecteurs de proportions $\mathbf{p}(a)$ issus du processus de Dirichlet proposé dans ce mémoire de telle manière que :

pour chaque colonne stratigraphique (u, v) ,

$$\mathbf{p}(u, v) = \sum_{w=1}^{n_w} \lambda_{u,v}(w) \cdot \mathbf{p}(u, v, w) \quad (\text{Éq. 10.8})$$

où $\lambda_{u,v}(w)$ est un poids associé à la cellule de coordonnées (u, v, w) : volume, épaisseur de la maille par exemple ;

¹ Au point de coordonnées (u, v) de la carte où les attributs sismiques correspondent à \mathbf{d}_s , pour chaque faciès f , on peut écrire la relation $p_f(u, v) = \psi_f(\mathbf{d}_s)$.

² Ces cartes de proportions ne sont pas uniquement fournies par la sismique, elles peuvent être construite par le géologue sur la base de sa connaissance du contexte géologique de la région, du bloc voisin ou d'un réservoir analogue.

c'est-à-dire que pour chaque faciès f ,

$$p_f(u,v) = \sum_{w=1}^{n_w} \lambda_{u,v}(w) \times p_f(u,v,w) \quad (\text{Éq. 10.9})$$

Par conséquent, pour conserver la cohérence des modèles simulés avec l'information sismique définie à l'échelle des cartes, il est nécessaire de construire des cubes de proportions qui respectent les contraintes (10.8) et (10.9).

10.2.2 Assimilation des contraintes sismiques

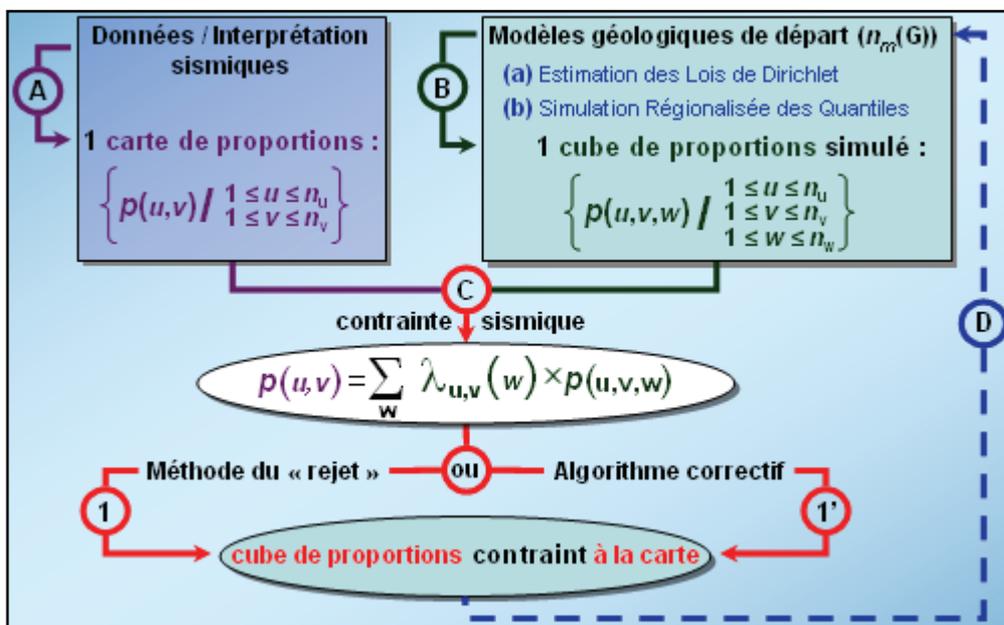


FIG. 10.5 : Intégration de l'incertitude sur les cartes tirées de la sismique. **(A)** L'imagerie sismique fournit une carte des vecteurs de proportions $p(u,v)$ sur l'ensemble des colonnes stratigraphiques (u,v) . **(B)** Le formalisme de Dirichlet simule un cube de vecteurs de proportions $p(u,v,w)$ sur l'ensemble des cellules (u,v,w) . **(C)** Le respect de la contrainte sismique passe **(1)** soit par le rejet de toute réalisation n'honorant pas ces conditions ; **(1')** soit par une méthode de combinaisons de dérivées. Dans les deux cas, on obtient un cube de proportions contraint à la carte sismique. **(D)** Pour améliorer l'estimation des lois de Dirichlet initiales, il est en outre possible de les actualiser en ajoutant le cube de proportions simulé, puis accepté ou corrigé, au pool de modèles de base qui servent à estimer les lois de Dirichlet.

Une première possibilité consiste à rejeter toute réalisation en terme de cube de proportions n'honorant pas les relations (10.8) et (10.9) (Fig. 10.5-C1). Cependant, dans la pratique, en raison de la dimension des objets simulés ($n_u \times n_v \times n_w \times n_f$), cette approche produit inévitablement une majorité de réalisations non contraintes. C'est pourquoi on préfère caler les modèles simulés aux cartes de proportions en aval du processus de cons-

truction. Pour cela, si on considère les colonnes stratigraphiques (u,v) individuellement, celles-ci s'interprètent comme une série de courbes verticales de proportions¹. Or, diverses techniques existent pour combiner carte et courbe verticale de proportions. Certaines sont aussi simples qu'une combinaison linéaire ; d'autres plus complexes comme les méthodes bayésiennes. Labat (2004) propose de prendre en compte ces données sous la forme de contraintes en utilisant l'interpolateur Discrete Smooth Interpolation (DSI). Aucune d'entre elle n'est plus satisfaisante que les autres en toutes circonstances. Dans ce mémoire, on recourt à un algorithme développé par Bouzid (2003) sur une idée originale de Massonnat. La méthode distribue de façon itérative les proportions de faciès par colonne stratigraphique de façon à respecter les conditions (10.8) et (10.9). Nous ne reviendrons pas davantage sur la méthode utilisée, aucun changement n'ayant été apporté au cours de cette thèse. L'essentiel réside dans la possibilité de construire un cube de probabilités d'apparition de chaque faciès qui respecte les cartes sismiques et le scénario géologique (Fig. 10.5-C1'). Pour aller plus loin, il est possible d'actualiser les lois de Dirichlet en recalculant le vecteur de paramètres $\mathbf{c}(\alpha) = [c_1(\alpha), \dots, c_{nf}(\alpha)]^T$ en chaque nœud α . Il suffit d'ajouter chaque cube de proportions simulé puis corrigé, au pool de modèles de base qui servent à estimer les lois de Dirichlet (Fig. 10.5-D).

10.2.3 Des cartes sismiques incertaines !

Dans ce qui précède, nous ne tenons absolument pas compte de l'incertitude qui pèse sur les cartes de proportions tirées de la sismique. Cette variabilité est due en partie au nombre limité de données utilisées pour établir la relation probabiliste entre la nature des faciès et l'information sismique. Reprenons par exemple, le cas de la relation décrite à la section 10.2.1. Elle se définit en partitionnant les observations aux puits en plusieurs groupes homogènes $\{S_k\}$, à partir desquels sont estimées les probabilités conditionnelles d'occurrence des faciès pour chaque classe $\{S_k\}$. La fiabilité de ces distributions dépend naturellement de la quantité de données utilisées pour estimer celles-ci. Elle sera d'autant plus grande que le nombre de données indépendantes dans les groupes $\{S_k\}$ est important. Biver *et al.* (2002) modélise ainsi l'incertitude relative à l'estimation de ces distributions, en l'occurrence les proportions de faciès dans l'espace, au moyen des lois de Dirichlet. Leurs paramètres sont alors calculés à partir du nombre de données indépendantes utilisées pour réaliser ces estimations. D'ailleurs, il est possible de manière générale de se ramener au formalisme des lois de Dirichlet afin d'exprimer l'incertitude sur des cartes de proportions déterminées à partir d'une autre relation probabiliste.

En conséquence, pour honorer les contraintes qu'imposent ces cartes, on peut dans un premier temps, introduire un certain degré de tolérance au niveau des conditions (10.8) et (10.9). Celui-ci doit traduire l'incertitude sur les cartes, et permet d'accepter tout cube de proportions simulé qui respectent ces contraintes assouplies. D'un autre côté, une alternative possible reposerait plutôt sur (1) la simulation d'une carte de proportions au moyen

¹ Dans chaque bloc (u,v), la courbe verticale de proportions donnent pour chaque couche de la grille la proportion attendue des divers faciès (Fig. 10.3).

du formalisme de Dirichlet appliquée aux données issues de la sismique, puis (2) celle d'un cube de proportions géologique selon l'approche exposée dans ce mémoire, et enfin (3) la mise en cohérence de ce cube de proportions à cette carte simulée à l'aide des méthodes décrites à la section 10.2.2 (Fig. 10.6).

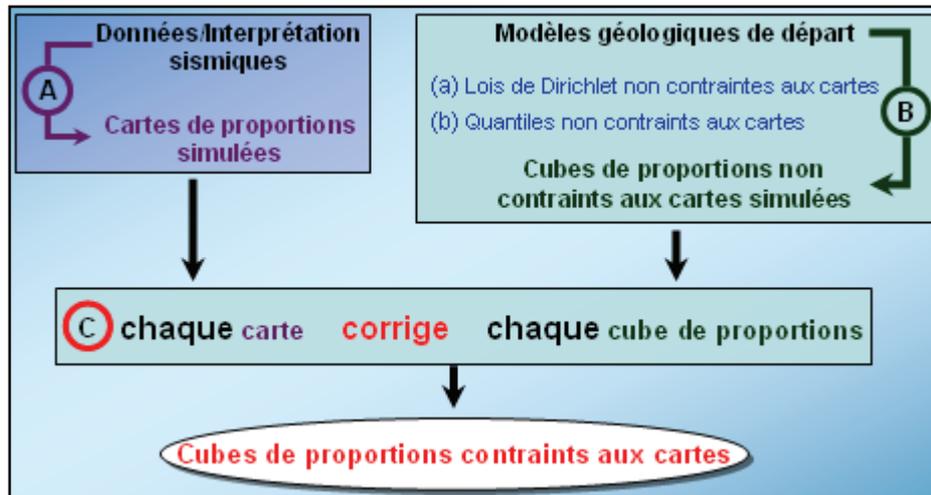


FIG. 10.6 : Intégration de l'incertitude sur les cartes tirées de la sismique. Avant toute réalisation en cube de proportions, on simule (1) une carte de proportions selon le formalisme de Dirichlet appliqué aux proportions issues de la sismique. On simule alors selon le protocole de Dirichlet (2) un cube de proportions, qu'une méthode de combinaisons de dérivées (3) contraint à la carte préalablement simulée.

Il est donc possible de simuler des cubes de proportions qui intègrent l'incertitude sur le scénario géologique, et respectent l'information disponible additionnelle : en particulier les données de puits ou les cartes tirées de la sismique. De plus, les méthodes proposées dans ce chapitre permettent également d'assimiler la variabilité des données complémentaires introduites. Il suffit pour cela de simuler les contraintes (valeurs aux puits ou cartes) à respecter. Remarquons néanmoins que la mise en cohérence des cubes de proportions issus de la géologie s'effectue en aval de leur processus de simulation. Se pose alors la question de contraindre directement les lois de Dirichlet avec les sources d'informations alternatives : sismique, puits (**chapitre 11**).

Chapitre 11

Combinaison de cubes de proportions

Jusqu'à présent, nous nous sommes exclusivement intéressés à combiner des cubes de proportions issus de divers scénarios géologiques, puis à introduire une, voire plusieurs sources d'information complémentaires extraites de catégories alternatives de données (sismiques, puits, etc). Dans certains cas pourtant, l'utilisation séparée de ces diverses sources d'informations et/ou de connaissance permet d'obtenir autant de cubes de proportions distincts. Pour chacun, l'incertitude sur les proportions en tout point du milieu peut être modélisée à travers le formalisme des lois de Dirichlet. Aussi, cherche-t-on, afin de valoriser ces sources variées d'information à combiner de manière cohérente ces modèles issus d'analyses distinctes comme leur incertitude. Ce chapitre s'inscrit ainsi dans la continuité des travaux de Biver (2006) en proposant leur application à la combinaison des lois de Dirichlet, et ce dans un cadre bayésien. Le recours à ce formalisme se solde par la construction d'une loi de Dirichlet généralisée qui tient compte de l'incertitude propre à chaque méthodologie, ici géologique, statistique et sismique.

11.1 Des cubes de proportions d'origines diverses

Dans ce mémoire, le recours aux distributions de Dirichlet permet la modélisation des incertitudes géologiques sur les proportions de faciès, l'estimation de ces lois passant par la combinaison des quelques scénarios géologiques (notés par l'indice \mathbf{Ge}) dont on dispose. En tout point α , les vecteurs de proportions $\mathbf{p}_{\mathbf{Ge}}(\alpha) = [p_{l,\mathbf{Ge}}(\alpha), \dots, p_{nf,\mathbf{Ge}}(\alpha)]^T$ suivent une loi de Dirichlet $\mathbf{P}_{\mathbf{Ge}}(\alpha)$ ayant pour paramètres $\mathbf{c}_{\mathbf{Ge}}(\alpha) = [c_{1,\mathbf{Ge}}(\alpha), \dots, c_{nf,\mathbf{Ge}}(\alpha)]^T$ (chapters 8 et 9). Cependant, le formalisme des lois de Dirichlet se prête également à des modèles issus de données différentes, ou d'une analyse distincte de ces mêmes données, c'est-à-dire :

- un cube de proportions dérivé de statistiques réalisées à partir des observations faites aux puits : une courbe verticale de proportions par exemple ;

- un cube de proportions construit sur la base d’une analyse des signaux sismiques comme l’impédance acoustique.

L’incertitude propre à chaque cube de proportions peut être modélisée par une loi de Dirichlet. La première situation correspond au cas où l’on bénéficie d’un nombre suffisant de puits. On construit un modèle dont le vecteur de proportions au point α est issu d’une estimation statistique (notée par l’indice **St**), $\mathbf{p}_{\text{St}}(\alpha) = [p_{l,\text{St}}(\alpha), \dots, p_{nf,\text{St}}(\alpha)]^T$ basée sur les données de puits. L’incertitude autour de ce vecteur dépend naturellement de la densité locale des données. Elle sera d’autant plus importante que la zone est moins bien échantillonnée. Haas *et al.* (2002) modélise cette variabilité par une loi de Dirichlet $\mathbf{P}_{\text{St}}(\alpha)$ ayant pour paramètres $\mathbf{c}_{\text{St}}(\alpha) = [c_{1,\text{St}}(\alpha), \dots, c_{nf,\text{St}}(\alpha)]^T$. La quantité “ $c_{f,\text{St}}(\alpha)-1$ ” est le nombre de données aux puits utilisées pour estimer la proportion du faciès f au point α .

La seconde situation renvoie au cas où l’on dispose d’un modèle 3D de données sismiques, qui peuvent être corrélés à la géologie. Car, de la même manière qu’il est possible de bâtir une carte de proportions de faciès (section 10.2), on peut parfois définir un cube de proportions à partir d’une relation probabiliste. En tout point α du milieu, le vecteur de proportions est issu d’une estimation basée sur les mesures sismiques (noté par l’indice **Si**) $\mathbf{p}_{\text{Si}}(\alpha) = [p_{l,\text{Si}}(\alpha), \dots, p_{nf,\text{Si}}(\alpha)]^T$. Biver *et al.* (2002) modélise la variabilité autour de ce vecteur par une loi de Dirichlet $\mathbf{P}_{\text{Si}}(\alpha)$ de paramètres $\mathbf{c}_{\text{Si}}(\alpha) = [c_{1,\text{Si}}(\alpha), \dots, c_{nf,\text{Si}}(\alpha)]^T$. Le terme “ $c_{f,\text{Si}}(\alpha)-1$ ” représente le nombre de données sismiques utilisées pour estimer la proportion du faciès f au point α .

En chaque point α du modèle maillé, on dispose donc de plusieurs lois de Dirichlet qui traduisent l’incertitude sur les proportions de faciès, chacune privilégiant une source d’information distincte. Pour valoriser celles-ci, il s’agit désormais d’en combiner les lois de Dirichlet résultantes, c’est-à-dire géologique (**Ge**), statistique (**St**) et sismique (**Si**). L’utilisation de diverses sources de données peut effectivement réduire l’incertitude sur l’estimation de la distribution $\mathbf{P}(\alpha)$ des vecteurs de proportions en ce point. De fait, une fois les lois $\mathbf{P}(\alpha)$ déterminées en chaque point α , la simulation des cubes de proportions s’effectue avec ces lois selon les méthodes exposées au chapitre 9.

11.2 Couplage des lois de Dirichlet

11.2.1 Objectif

En chaque point α du sous-sol, l’ensemble de l’information disponible, exprimée sous la forme de lois de Dirichlet $\mathbf{P}_{\text{Ge}}(\alpha)$, $\mathbf{P}_{\text{St}}(\alpha)$ et $\mathbf{P}_{\text{Si}}(\alpha)$, doit être intégrée au sein d’un modèle cohérent du vecteur de proportions $\mathbf{P}(\alpha)$. Il s’agit en fait de combiner au mieux les distributions disponibles au moyen d’une relation Ψ à définir (Fig. 11.1) :

en tout point α du modèle,

$$\mathbf{P}(\alpha) = \Psi(\mathbf{P}_{\text{Ge}}(\alpha), \mathbf{P}_{\text{St}}(\alpha), \mathbf{P}_{\text{Si}}(\alpha)) \quad (\text{Éq. 11.1})$$

Journel (1998) propose une méthode basée sur les probabilités conditionnelles. Elle utilise une fonction Ψ qui fait intervenir une série de paramètres appelés *tau*, ceux-ci quantifiant la redondance entre les différentes sources de données. De manière plus générale, Clemen et Winkler (1999) détaillent les méthodes disponibles pour combiner des distributions de probabilités. Parmi celles-ci, on s'intéresse plus particulièrement aux techniques qui se prêtent aisément aux lois de Dirichlet à l'instar des combinaisons linéaires et de l'approche bayésienne.

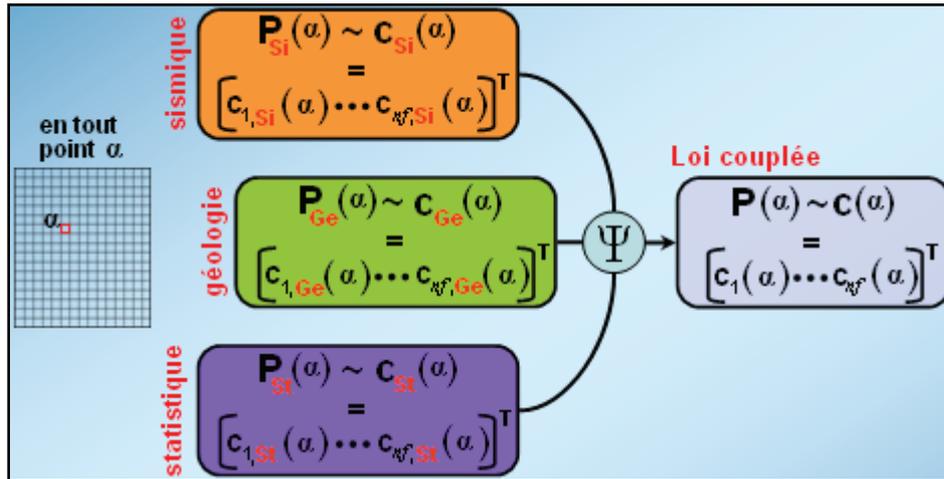


FIG. 11.1 : Combinaison des lois de Dirichlet. En chaque point α du modèle maillé, les mesures aux puits (statistique), la sismique et la géologie définissent respectivement une loi de Dirichlet notée $\mathbf{P}_{St}(\alpha)$, $\mathbf{P}_{Si}(\alpha)$ et $\mathbf{P}_{Ge}(\alpha)$. La relation Ψ permet de combiner ces distributions issues de sources d'informations distinctes en une loi couplée $\mathbf{P}(\alpha)$, qui est toujours une loi de Dirichlet (sections 11.2.2 et 11.2.3).

11.2.2 Mélange par combinaisons linéaires

C'est la méthode la plus simple pour combiner des lois de probabilités. Elle consiste à attribuer un poids ω_k à chaque distribution :

en tout point α du modèle,

$$\mathbf{P}(\alpha) = \omega_{Ge} \cdot \mathbf{P}_{Ge}(\alpha) + \omega_{St} \cdot \mathbf{P}_{St}(\alpha) + \omega_{Si} \cdot \mathbf{P}_{Si}(\alpha) \quad (\text{Éq. 11.2})$$

en confondant les distributions avec leur densité de probabilité ; la somme des poids étant égale à 1, $\omega_{Ge} + \omega_{St} + \omega_{Si} = 1$.

La difficulté repose sur le choix des poids, car cela reste un exercice subjectif. Ils peuvent être définis de façon à refléter la représentativité ou la qualité relative de chaque distribution, soit la confiance relative que l'on a en chacune d'entre elle. Une alternative consiste à utiliser les paramètres $\mathbf{c}_K(\alpha) = [c_{1,K}(\alpha), \dots, c_{nf,K}(\alpha)]^T$, qui définissent les lois de Dirichlet à combiner. Ces derniers peuvent en effet être reliés aux occurrences de faciès dans un échantillon fictif de taille " $c_{1,K}(\alpha) + \dots + c_{nf,K}(\alpha) - n_f$ ". Par exemple, pour la distribution $\mathbf{P}_{St}(\alpha)$

définie par les paramètres $\mathbf{c}_{St}(\alpha) = [c_{1,St}(\alpha), \dots, c_{nf,St}(\alpha)]^T$, c'est comme si on observait " $c_{f,St}(\alpha)-1$ " occurrences du faciès f^l . Par conséquent, il paraît naturel d'utiliser ces tailles d'échantillons fictifs pour traduire la confiance que l'on peut avoir en ces distributions. Plus le nombre de données utilisées pour conduire leur estimation est important, plus celle-ci est précise. Ainsi, on attribue respectivement à chacune des lois à combiner un poids égal à la moyenne pondérée des tailles d'échantillons fictifs qui leur sont associés :

en tout point α ,

$$\omega_K = \frac{\sum_K (c_{f,K}(\alpha) - 1)}{\sum_{f=1}^{n_f} (c_{f,Ge}(\alpha) - 1) + \sum_{f=1}^{n_f} (c_{f,St}(\alpha) - 1) + \sum_{f=1}^{n_f} (c_{f,Si}(\alpha) - 1)} \quad (\text{Éq. 11.3})$$

où l'indice $\mathbf{K} = \mathbf{Ge}, \mathbf{St}$ ou \mathbf{Si} , et renvoie ainsi aux diverses sources d'informations.

11.2.3 Mélange bayésien

Les méthodes bayésiennes offrent la possibilité de mettre à jour une distribution a priori en y incorporant des informations supplémentaires. Haas et Formery (2002) proposent une solution dans un tel cadre, que Biver *et al.* (2002) mettent en application. Elle consiste à combiner deux distributions en actualisant l'une des lois, considérée comme un a priori, à l'aide de l'autre transformée en un échantillon fictif. Ils utilisent la particularité des lois de Dirichlet d'être des lois conjuguées, c'est-à-dire que si la loi a priori est supposée suivre une distribution de Dirichlet, la loi a posteriori déduite des données se trouve toujours être une loi de Dirichlet.

Considérons le point α du réservoir maillé. On détermine d'abord la distribution $\mathbf{P}_c(\alpha)$ résultant de la combinaison des lois $\mathbf{P}_{Ge}(\alpha)$ et $\mathbf{P}_{St}(\alpha)$. La loi a priori de Dirichlet $\mathbf{P}_{Ge}(\alpha)$ issue de la géologie a pour paramètres $\mathbf{c}_{Ge}(\alpha) = [c_{1,Ge}(\alpha), \dots, c_{nf,Ge}(\alpha)]^T$. La loi $\mathbf{P}_{St}(\alpha)$ dérivée d'une analyse statistique, peut être associée à un échantillon fictif où l'on observerait " $c_{f,St}(\alpha)-1$ " occurrences du faciès d'indice f . La loi a posteriori se définit donc par une loi de Dirichlet $\mathbf{P}_{co}(\alpha)$ de paramètres $\mathbf{c}_{co}(\alpha) = [c_{1,co}(\alpha), \dots, c_{nf,co}(\alpha)]^T$ de sorte que :

pour chaque faciès f ,

$$c_{f,co}(\alpha) = c_{f,Ge}(\alpha) + (c_{f,St}(\alpha) - 1) \quad (\text{Éq. 11.4})$$

Il est important de noter que l'on obtiendrait le même résultat en utilisant la loi $\mathbf{P}_{St}(\alpha)$ comme a priori. Dans ce cas, il suffirait de transposer les paramètres de la loi $\mathbf{P}_{Ge}(\alpha)$ en un échantillon équivalent où on observerait " $c_{f,Ge}(\alpha) - 1$ " occurrences du faciès f . De la même manière, il est possible d'ajouter l'information apportée par les données sismiques à travers la loi de Dirichlet $\mathbf{P}_{Si}(\alpha)$. On considère alors comme a priori la loi $\mathbf{P}_c(\alpha)$, qui résulte

¹ Ces occurrences sont utilisées pour inférer la loi de Dirichlet.

de la combinaison de $\mathbf{P}_{\text{Ge}}(\alpha)$ et de $\mathbf{P}_{\text{St}}(\alpha)$. La loi a posteriori $\mathbf{P}(\alpha)$ résultant de cette combinaison est une loi de Dirichlet de paramètres $\mathbf{c}(\alpha) = [c_1(\alpha), \dots, c_{n_f}(\alpha)]^T$ définis par la relation :

pour chaque faciès f ,

$$c_f(\alpha) = c_{f,\text{co}}(\alpha) + (c_{f,\text{si}}(\alpha) - 1) = c_{f,\text{Ge}}(\alpha) + (c_{f,\text{st}}(\alpha) - 1) + (c_{f,\text{si}}(\alpha) - 1) \quad (\text{Éq. 11.5})$$

Remarquons d'ailleurs que l'ordre des combinaisons n'a aucun effet sur la distribution résultante.

Ainsi, la loi de Dirichlet inférée dans cette thèse à partir des différents scénarios géologiques dont on dispose entre dans le cadre des combinaisons de distributions, qu'elles soient linéaires ou bayésiennes comme proposé par Biver. Ces combinaisons de lois de Dirichlet issues de différentes sources d'information permettent donc d'affiner la quantification de l'incertitude sur les proportions de faciès. Un tel formalisme permet de fait de combiner les différentes sources d'informations (sismique, statistique et géologique) de façon optimale.

Chapitre 12

Régionalisation de l'incertitude

Le formalisme des lois de Dirichlet, ou encore l'approche multi-réalisation simulent un ensemble cohérent de cubes de proportions alternatifs. Ces procédés permettent ainsi de balayer l'espace d'incertitude lié à la définition du scénario géologique. Se pose néanmoins toujours la question d'une quantification, dite régionalisée, de cette incertitude en chaque point α du milieu. Autrement dit, il s'agit de définir les zones plus ou moins (in)certaines du modèle en terme de géologie. Ce chapitre revient donc sur divers index utilisés afin de régionaliser l'incertitude géologique. Tous sont appliqués à un même cas synthétique, en amont ou en aval des simulations de cubes de proportions. D'un autre côté, ce chapitre aborde également la question du classement des cubes de proportions de façon à établir la fonction de répartition des modèles simulés. L'objectif consiste une nouvelle fois à quantifier l'incertitude géologique, non pas au niveau des points du milieu, mais sur le cube de proportions considéré dans sa globalité. Il s'agit par là de fournir des modèles clés – via les quantiles – utilisables dans les simulations dynamiques en écoulement.

12.1 Un cas synthétique

L'objectif consiste à déterminer un index $I_R(\alpha)$, calculé en chaque point α du modèle maillé, qui identifie les zones plus ou moins (in)certaines en terme de géologie. On considère pour cela le modèle de plate-forme carbonatée *Tata* déjà évoqué au chapitre 5 (section 5.2.3) puis repris au chapitre 9 (section 9.2.1). Pour rappel, celui-ci compte trois catégories de faciès ($n_f = 3$), notées respectivement F_1 , F_2 et F_3 , de la meilleure à la moins bonne roche réservoir. Le domaine d'étude comporte une unité stratigraphique que discrétise une grille de type « boîte à sucres » de dimensions $70 \times 60 \times 1$ cellules. Ce cas, outre sa simplicité, offre effectivement l'avantage d'avoir été étudié selon les deux méthodologies proposées dans ce mémoire : les formalismes (1) de multi-réalisation sur les paramètres du scénario géologique, et (2) des lois probabilistes de Dirichlet. Le modèle pose le problème d'une incertitude sur l'extension latérale des faciès et de leurs proportions, incertitude exprimée sous la forme, soit (1) de cartes de bathymétrie alternatives, ou bien

(2) de plusieurs cubes de proportions de base (Fig. 12.1-a et b). Cependant, les méthodes évoquées précédemment ne quantifient pas directement en tout point α , l'incertitude régionalisée sur les proportions de faciès. Les deux prochaines sections proposent de fait divers indices $I_R(\alpha)$ permettant une telle régionalisation. Celle-ci s'effectue soit amont (Fig. 12.1-b), soit en aval (Fig. 12.1-c) des simulations multiples de cubes de proportions. Tous ces indices sont appliqués au modèle *Tata*, plus exactement en chacun des points α du modèle indépendamment des autres points.

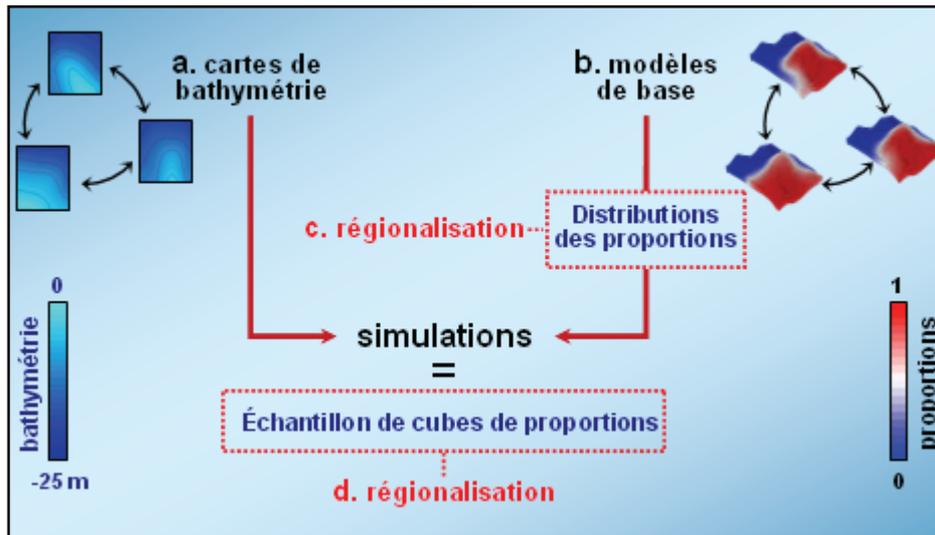


FIG. 12.1 : Présentation des possibilités de régionalisation de l'incertitude. L'incertitude relative à la paléogéographie, exprimée sous la forme (a) de cartes de bathymétrie ou (b) de cubes de proportions peut être régionalisée en amont des simulations (c) à partir des distributions (Dirichlet) régionalisées des proportions. La multi-réalisation sur les paramètres ou le formalisme de Dirichlet simulent un échantillon de cubes de proportions, à partir duquel peuvent être définis (d) divers indices régionaux quantifiant l'incertitude sur les proportions de faciès.

12.2 En amont des simulations

Ce cas renvoie à la définition, en tout point α , d'un indice $I_R(\alpha)$, qui caractériserait, avant de lancer le processus de simulation des cubes de proportions, les zones plus ou moins (in)certaines du modèle. Cette approche présente l'avantage d'être directe car elle évite la simulation d'un grand nombre de cubes de proportions. Néanmoins, seul le formalisme des lois de Dirichlet s'y prête véritablement. Dans ce cas en effet, l'incertitude sur les proportions, matérialisée en un point α du modèle par une loi de Dirichlet¹ de paramètres $\mathbf{c}(\alpha) = [c_1(\alpha), \dots, c_{n_f}(\alpha)]^T$, peut « se mesurer » par la variance de la distribution de chacune des proportions autour de sa moyenne. La somme de ces variances représente

¹ la loi de Dirichlet a été estimée à l'aide des $n_m(G)$ modèles de base.

ainsi un indice simple de régionalisation de l'incertitude :

$$I_R^{(1)}(\alpha) = \sum_{f=1}^{n_f} (\sigma[P_f(\alpha)])^2 \quad (\text{Éq. 12.1})$$

où n_f désigne le nombre de faciès considérés, et $\sigma^2[P_f(\alpha)]$ la variance de la variable aléatoire associée aux proportions, au point α , du faciès d'indice f . Cette variance se déduit simplement du jeu de paramètres $\mathbf{c}(\alpha) = [c_1(\alpha), \dots, c_{n_f}(\alpha)]^T$ (Éq. 8.6).

Cependant, pour comparer les indices d'incertitude d'un point à un autre, on utilisera de préférence un coefficient normalisé. Plutôt que de considérer la somme des variances, on privilégiera ainsi la somme des coefficients de variation respectifs de chacune des distributions attachées aux proportions de faciès, c'est-à-dire la somme des rapports entre leurs écart-type et leur moyenne :

$$I_R^{(2)}(\alpha) = \sum_{f=1}^{n_f} \left(\frac{\sigma[P_f(\alpha)]}{E[P_f(\alpha)]} \right) \quad (\text{Éq. 12.2})$$

où n_f désigne le nombre de faciès considérés ; $\sigma[P_f(\alpha)]$ et $E[P_f(\alpha)]$ renvoient respectivement à l'écart-type et l'espérance de la variable aléatoire associée aux proportions du faciès d'indice f au point α . Ils se déduisent des paramètres $\mathbf{c}(\alpha) = [c_1(\alpha), \dots, c_{n_f}(\alpha)]^T$ (Éq. 8.5 et Éq. 8.6).

Une alternative consiste à utiliser la somme des paramètres $c_1(\alpha), \dots, c_{n_f}(\alpha)$ ¹ qui définissent la loi de Dirichlet inférée au point α . Les sections 11.2.2 et 11.2.3 soulignent que par analogie, la valeur " $c_f(\alpha)-1$ " représente le nombre d'occurrences du faciès f dans un échantillon fictif de taille " $c_1(\alpha)+\dots+c_{n_f}(\alpha)-n_f$ ". Ces observations servant à estimer la moyenne des lois de Dirichlet, la confiance que l'on peut avoir en ces distributions s'exprime à travers cette taille d'échantillon fictif. Plus le nombre de données utilisées pour conduire leur estimation est important, plus celle-ci devrait se révéler précise :

$$I_R^{(3)}(\alpha) = c_1(\alpha) + \dots + c_{n_f}(\alpha) - n_f \quad (\text{Éq. 12.3})$$

Cet indice entre dans le cadre plus général des combinaisons de cubes de proportions issus de plusieurs types de données (chapitre 11). Il s'applique ainsi aux diverses lois de Dirichlet estimées à partir des différentes sources d'information, régionalise l'incertitude propre à chacune, et permet leur comparaison en tout point α du modèle maillé.

Les indices déterminés en aval des simulations offrent l'avantage de nécessiter peu de calculs, ceux-ci se limitant à l'estimation des lois de Dirichlet (Fig. 12.2-I). Ces indices semblent équivalents (Fig. 12.2-II). Par contre, aucun d'eux ne se prêtent, en amont des simulations, au procédé de multi-réalisation sur les éléments du scénario géologique (partie 2). Leur calcul explicite suppose en effet de connaître la loi de probabilité des proportions de faciès. Or, si la distribution des paramètres incertains est spécifiée, la définition précise de celle des proportions semble difficile. Néanmoins, les techniques de

¹ La valeur n_f du dernier indice correspond au nombre de faciès étudiés.

simulations telles que les méthodes de Monte Carlo permettent d'approcher numériquement ces indices. Cela passe bien entendu par la simulation préalable des cubes de proportions. En conséquence, dans le cadre d'une approche multi-réalisation, les indices de régionalisation évoqués précédemment appartiendront à la famille des index déterminés en aval des simulations de cubes de proportions.

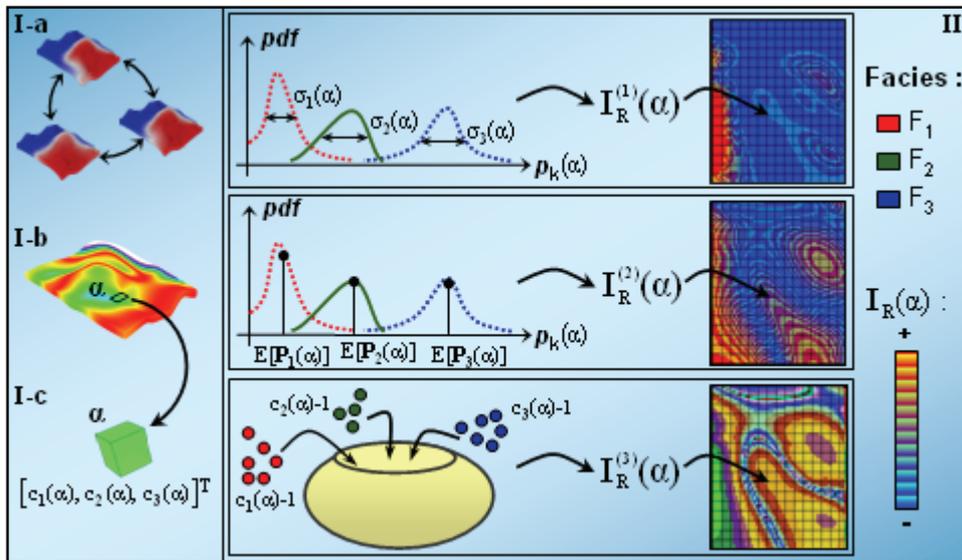


FIG. 12.2 : Régionalisation de l'incertitude en amont des simulations. Les modèles de base (I-a) déterminent en tout point du modèle maillé les distributions des vecteurs de proportions de faciès (I-b). En un point α du géomodèle (I-c), la connaissance des lois de probabilités permet le calcul des trois indices de régionalisation présentés dans cette section. Ceux-ci donnent des résultats équivalents (II).

12.3 En aval des simulations

Ce cas renvoie à la définition d'un indice $I_R(\alpha)$, qui caractériserait, après la simulation des cubes de proportions, les zones plus ou moins (in)certaines du modèle. Cette approche, bien que plus coûteuse en temps de calcul, est applicable quel que soit le mode de simulation des cubes de proportions : formalismes de multi-réalisation sur les paramètres ou des lois de Dirichlet. Elle permet (1) de se rapporter aux indices présentés à la section précédente, ou plus généralement à des indices dont l'évaluation nécessite le recours à des méthodes de calculs probabilistes, et (2) d'offrir un éventail de mesures alternatives de la dispersion.

D'une part, les indices $I_R^{(1)}$, $I_R^{(2)}$ et $I_R^{(3)}$ (section 12.2) se révèlent difficiles à calculer de manière explicite quand on privilégie une approche du type multi-réalisation. Néanmoins, la simulation des cubes de proportions revient en particulier, à échantillonner au point α , la distribution du vecteur de proportions aléatoire $\mathbf{P}(\alpha)$, ainsi que celles de ses n_f composantes aléatoires $P_1(\alpha), \dots, P_{n_f}(\alpha)$. Un tel échantillon permet l'estimation des premiers

moments¹ (espérance et variance) de chacune des lois de probabilité de $P_1(\alpha), \dots, P_{n_f}(\alpha)$, dont se déduisent directement les indices $I_R^{(1)}$ et $I_R^{(2)}$, puis l'index $I_R^{(3)}$ une fois évalués les paramètres $c_1(\alpha), \dots, c_{n_f}(\alpha)$ de la loi de Dirichlet modélisant la distribution². Dans le même ordre d'idée, l'entropie [Shannon, 1948], dont la mise en œuvre pratique passe par les techniques probabilistes, peut alors servir d'indice de régionalisation. Cette notion traduit en effet l'incertitude contenue dans une distribution de probabilité : c'est l'incertitude sur la réalisation d'un événement³. Par exemple, l'incertitude sur la distribution des faciès correspond au caractère plus ou moins bien réparti des probabilités sur les différents faciès. Au point α , une distribution de trois faciès avec le vecteurs de proportions $[1, 0, 0]$ est moins incertaine qu'une autre distribution avec le vecteur de probabilités $[\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]$. Pour des lois de probabilité discrètes comme celles-ci, l'entropie se définit par :

$$H = -\sum_{f=1}^{n_f} p_f \times \log(p_f) \quad (\text{Éq. 12.4})$$

où H représente l'entropie de la distribution discrète, n_f le nombre de faciès, p_f la probabilité d'occurrence du faciès d'indice f , et $x \rightarrow \log(x)$, le logarithme népérien.

L'entropie est minimale pour les distributions présentant un état avec une probabilité de un, et pour lesquelles elle est nulle (car la fonction $x \rightarrow x \log x$ est nulle en 0 et 1), et maximale pour les systèmes équiprobables pour lesquels elle vaut $\log n_f$. À l'instar des distributions de faciès précédentes, l'incertitude régionalisée sur les proportions de faciès peut être exprimée sous la forme d'une entropie, qui dans le cas continu prend une forme intégrale :

$$I_R^{(4)}(\alpha) = -\oint_{S(n_f)} \pi_\alpha(\mathbf{p}) \times \log[\pi_\alpha(\mathbf{p})] \cdot d\mathbf{p} \quad (\text{Éq. 12.5})$$

où n_f désigne le nombre de faciès considéré, $S(n_f)$ le simplexe en dimension n_f . La fonction $\pi_\alpha : x \rightarrow \pi_\alpha(x)$, définie sur \mathbb{R}^{n_f} , représente la densité de probabilité de la distribution des proportions de faciès au point α , et $x \rightarrow \log(x)$ le logarithme népérien.

Néanmoins, un calcul direct de cette intégrale, compte tenu des dimensions en jeu, devient rapidement délicat. Nous recourons plutôt aux techniques de Monte Carlo qui écrivent l'expression intégrale sous la forme d'une espérance mathématique :

$$\frac{1}{n_s} \times \sum_r^{\text{échantillon}} (-\log[\pi_\alpha(\mathbf{p}^r)]) \longrightarrow E[-\log \circ \pi_\alpha](\mathbf{P}(\alpha)) = I_R^{(4)}(\alpha) \quad (\text{Éq. 12.6})$$

¹ L'estimation des moments en chaque point α , peut être réalisée de façon itérative après la simulation de chaque cube de proportions. Ce qui évite de conserver l'ensemble des réalisations, et prévient donc d'éventuels problèmes d'espace mémoire liés à des capacités informatiques limitées.

² Les paramètres $c_1(\alpha), \dots, c_{n_f}(\alpha)$ sont calculés à l'aide des espérances et des variances préalablement estimées à l'aide de l'échantillon de vecteurs de proportions simulé.

³ Cette entropie est inspirée de l'entropie thermodynamique de Boltzmann en théorie cinétique des gaz. De la même manière, elle se révèle être une mesure du désordre.

Avec \mathbf{p}^r la réalisation d'indice \mathbf{r} dans un échantillon de n_s vecteurs de proportions au point α du modèle maillé. Ils sont tirés selon la distribution $\mathbf{P}(\alpha)$ de densité de probabilité π_α . $E[Z]$ renvoie à l'espérance mathématique de la variable aléatoire Z , et la fonction $x \rightarrow \log(x)$ renvoie au logarithme népérien.

Cette espérance est déterminée au moyen de l'échantillon de cubes de proportions simulés selon le formalisme des lois de Dirichlet. Là encore, si ce dernier s'y prête facilement¹, un tel mode de calcul paraît difficilement applicable à une approche comme la multi-réalisation. Dans ce cas, nous proposons un moyen alternatif de calcul. L'idée repose sur la discrétisation du simplexe $S(n_f)$ où sont simulés les vecteurs de proportions (section 8.2.2) en un nombre n_c de classes $\{C_k\}_{1 \leq k \leq n_c}$. La simulation indépendante de n_s cubes de proportions revient en particulier, à simuler au point α un nombre n_s de vecteurs de proportions $\mathbf{p}(\alpha)$ indépendants dans la distribution $\mathbf{P}(\alpha)$. Il suffit de les affecter à chacune des classes $\{C_k\}_{1 \leq k \leq n_c}$; l'étude des effectifs fournit une première approximation de l'entropie, et par conséquent de l'indice de régionalisation (Fig. 12.3) :

$$\frac{1}{n_s} \times \sum_{\text{échantillon}} (-\log[\pi_\alpha(\mathbf{p}(\alpha))]) \approx -\sum_{k=1}^{n_c} \frac{1}{N(C_k)} \times \log\left(\frac{1}{N(C_k)}\right) \quad (\text{Éq. 12.7})$$

avec $N(C_k)$ l'effectif de la classe d'indice k : C_k ; les fonctions π_α et $x \rightarrow \log(x)$ renvoient respectivement à la densité de probabilité des vecteurs de proportions et au logarithme népérien.

D'autre part, la mise en oeuvre des simulations permet d'utiliser d'autres types de mesures de la dispersion, notamment géométrique. À ce stade en effet, on dispose pour tout point α du modèle, d'un échantillon de vecteurs de proportions. Chaque réalisation $\mathbf{p}(\alpha)$ renvoie à un point dans le simplexe $S(n_f)$, dont les coordonnées sont les proportions de faciès. La disposition dans l'espace de ces points exprime la variabilité des réalisations en réponse aux incertitudes sur les proportions. On définit alors un nouvel indice de l'incertitude régionalisée qui s'appuie sur la mesure de l'écartement des points au centre de gravité du nuage qu'ils composent (Fig. 12.3). Pour cela, on utilise les distances classiques dans un espace de dimension finie, comme la distance euclidienne ou de Manhattan, qui ne sont en fait que des cas particuliers de la distance de Minkowski² :

$$d_n(\mathbf{p}(\alpha), \mathbf{p}'(\alpha)) = \left(\sum_{f=1}^{n_f} w_f \times |p_f(\alpha) - p'_f(\alpha)|^n \right)^{\frac{1}{n}} \quad (\text{Éq. 12.8})$$

où $d_n(\mathbf{p}(\alpha), \mathbf{p}'(\alpha))$ représente la distance entre les vecteurs de proportions $\mathbf{p}(\alpha)$ et $\mathbf{p}'(\alpha)$, qui ont pour composantes respectives $p_f(\alpha)$ et $p'_f(\alpha)$ – f l'indice des faciès, $f = 1, \dots, n_f$.

¹ Dans le cadre du formalisme des lois de Dirichlet, l'expression de la densité de probabilité est connue de manière explicite. Le calcul de l'expression intégrale de l'entropie correspond à la moyenne des valeurs prises par cette densité aux points échantillonnés. D'ailleurs, la nécessité de simuler cet échantillon justifie la présentation de cet indice pour les lois de Dirichlet dans cette section.

² Toutes ces distances sont équivalentes au sens mathématique du terme, et conduisent à des résultats comparables.

On peut affecter un poids w_f selon la nature du faciès : ceux-ci peuvent, par exemple, être reliés à leurs propriétés pétrophysiques. En fonction du paramètre n , on parlera des distances de Manhattan ($n = 1$), euclidienne ($n = 2$) ou de Chebychev ($n = \infty$).

À la position α , le calcul de la distance de l'ensemble des points – ou extrémité des vecteurs $\mathbf{p}(\alpha)$ – à leur centre de gravité (extrémité du vecteur moyen $\bar{\mathbf{p}}(\alpha)$) définit une distribution de distance, dont sont extraits en plus de la médiane (q_{50}), les premiers (q_{10}) et derniers déciles (q_{90}). Ce qui nous affranchie des points aberrants. Une mesure normalisée de la dispersion des points, donc de l'incertitude régionalisée, peut donc s'écrire :

$$I_R^{(5)}(\alpha) = \frac{q_{90} - q_{10}}{q_{50}} \quad (\text{Éq. 12.9})$$

D'autres indices au caractère géométrique ont été envisagés durant cette thèse. On citera par exemple la mesure de la forme contenant les points, ou une certaine proportion d'entres eux pour éliminer les points aberrants. Il suffit pour cela d'extraire l'enveloppe convexe des points et d'en approximer l'hypersurface au moyen des méthodes classiquement utilisées en géométrie (Fig. 12.3) :

$$I_R^{(6)}(\alpha) = S[\{\mathbf{p}^1(\alpha), \dots, \mathbf{p}^{n_S}(\alpha)\}] \quad (\text{Éq. 12.10})$$

avec $\mathbf{p}^r(\alpha)$ la réalisation d'indice r dans un échantillon de n_S vecteurs de proportions au point α du modèle maillé, tirés selon la distribution $\mathbf{P}(\alpha)$. $S[\cdot]$ est un algorithme permettant le calcul de la surface de l'enveloppe convexe. Notons que cette dernière approche n'a été mise en œuvre qu'à de faibles dimensions (deux à trois faciès). Au-delà, elle devient relativement lourde à utiliser.

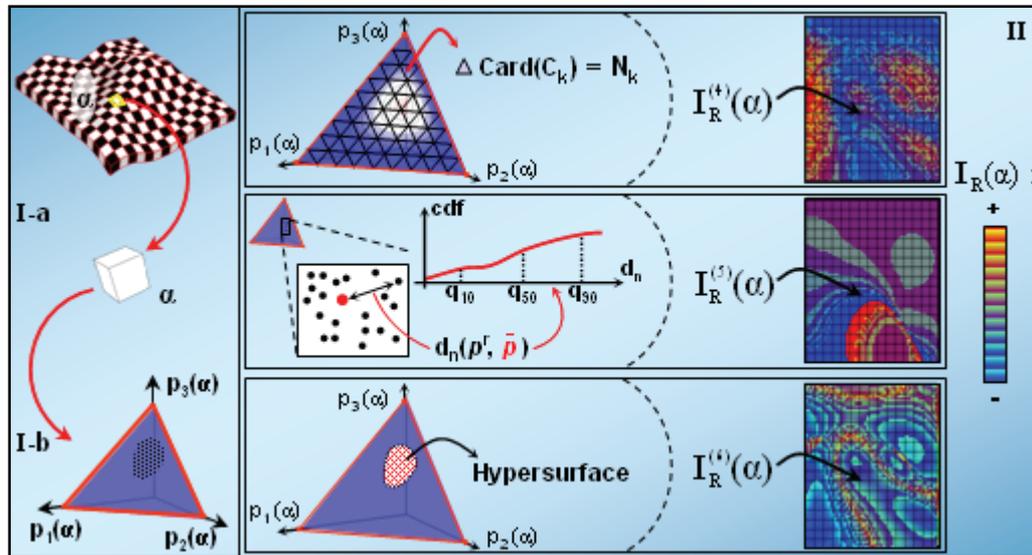


FIG. 12.3 : Régionalisation de l'incertitude en aval des simulations. Après simulations, en tous points α (I-a), on dispose d'un échantillon de vecteurs de proportions (I-b), qui permet de définir les trois indices de régionalisation présentés dans cette section (II).

Plusieurs indices sont ainsi susceptibles de rendre compte de la variabilité, ou dispersion, des proportions de faciès, qu'elles aient été simulées ou non. Tous sont d'autant plus faibles que les réalisations sont proches les unes des autres et s'appliquent pour la plupart, aux formalismes de simulations des cubes de proportions présentés dans ce mémoire. Néanmoins, s'ils quantifient l'incertitude des proportions sur le plan régional : en chaque point du milieu, il serait intéressant, une fois les cubes de proportions simulés, de la quantifier sur un plan global : pour l'ensemble des cubes de proportions.

12.4 Tri des cubes de proportions

Cette section est consacrée au classement des cubes de proportions simulés de façon à établir en quelque sorte, leur fonction de répartition selon une grandeur monodimensionnelle à préciser, notée R^K . Une telle distribution des cubes de proportions offre la possibilité :

- (a) sur un plan pratique, d'éviter de conserver l'ensemble des modèles en proportions simulés. Leur sauvegarde s'avèrerait effectivement lourde en termes de capacités informatiques si le nombre de simulations devient important. Un cube de proportions dépasse souvent le million de cellules, chiffre qu'il faut multiplier par les nombres de faciès pris en compte, et de simulations effectuées. Au contraire, la définition d'une telle fonction de répartition, peu consommatrice de mémoire, permet de relier chaque cube de proportions simulé à l'un de ses quantiles. À partir d'un quantile, il est simple de construire à nouveau le ou les(s) cubes de proportions lui correspondant ;
- (b) de procurer les modèles en proportions clés relatifs à des quantiles particuliers, et utilisables par la suite dans la simulation stochastique des propriétés pétrophysiques, et surtout les simulations dynamiques en écoulement. Leur mise en œuvre sur l'intégralité des cubes de proportions se révèle bien souvent lourde et coûteuse en temps de calculs. D'ailleurs, les approches actuelles, qui considèrent un cube de proportions unique, privilégie l'étude du comportement dynamique de quelques modèles sur l'ensemble des réalisations issues de la multi-réalisation pétrophysique. En général, ils correspondent aux premiers (q_{10}) et derniers (q_{90}) déciles ainsi qu'à la médiane (q_{50}) des accumulations.

Les indices de régionalisation présentés dans les sections précédentes (12.2 et 12.3) représentent une première étape vers un tri des cubes de proportions simulés. C'est en particulier le cas de l'indice $I_R^{(5)}(\alpha)$ qui met en œuvre la distance à leur moyenne des réalisations en vecteurs de proportions au point α . De la même manière, on définit une classification des réalisations en cubes de proportions, notés CP^r avec " r " son indice dans l'échantillon, en fonction de leur distance respective au cube moyen \overline{CP}^1 . En fait, cette

¹ Le cube de proportions moyen renvoie, en chacun de ses points α , à la moyenne des vecteurs de proportions respectifs en ce point, des diverses réalisations en cubes de proportions.

distance au cube moyen correspond à la somme des distances sur l'ensemble de ses points α , entre sa valeur $\mathbf{p}^r(\alpha)$ et le vecteur de proportions moyen $\bar{\mathbf{p}}(\alpha)$ en ce point :

$$R^K[CP^r] = D_n(CP^r, \bar{CP}) = \sum_{\alpha} (\lambda(\alpha) \times d_n[\mathbf{p}^r(\alpha), \bar{\mathbf{p}}(\alpha)]) \quad (\text{Éq. 12.11})$$

avec $R^K[CP^r]$ la distance de la réalisation en cubes de proportions d'indice "r" au cube moyen, et $\lambda(\alpha)$ un poids affecté à chaque point (ou cellule) α , comme le volume de la cellule associée à ce point par exemple. $d_n[\cdot, \cdot]$ est une des distances évoqués à la section 12.3.

Le calcul de la distance $D_n[\cdot, \cdot]$ pour l'ensemble des cubes de proportions simulés à leur centre de gravité définit une distribution de distance, dont les quantiles correspondent au rang des différents cubes de proportions simulés.

Une autre solution consiste à classer les cubes de proportions selon les accumulations en hydrocarbures. Néanmoins, un cube de proportions renvoie généralement, non pas à une accumulation, mais à une distribution d'accumulations¹. Une alternative repose donc sur l'utilisation de l'espérance des accumulations en hydrocarbures relative à chaque cube de proportions. En effet, les travaux de Massonnat (2000), Bergamo (2002), Poirot (2002), Sheppers (2003) et Zerkoune (2003) soulignent que le passage d'un modèle de probabilité à l'autre (ou d'un cube de proportions à l'autre) induit une dispersion plus importante sur les accumulations que les simulations stochastiques de la lithologie et de la pétrophysique. C'est pourquoi on propose ici de ne retenir que la simulation des proportions de faciès, et de ne pas réaliser les simulations liées à la pétrophysique. Cela constitue un moyen simple et rapide de trier les cubes de proportions avant toute simulation des propriétés pétrophysiques :

$$R^K[CP^r] = E[ooip(CP^r)] = \sum_{\alpha} RV(\alpha) \times \left(\sum_{f=1}^{n_f} p_f^r(\alpha) \times \bar{\phi}_f \times (1 - \bar{S}w_f) \right) \quad (\text{Éq. 12.12})$$

avec $R^K[CP^r]$ la grandeur monodimensionnelle associée à la réalisation en cubes de proportions d'indice "r". $E[ooip(CP^r)]$ renvoie à l'espérance des accumulations en hydrocarbures relatives au cube de proportions CP^r . $RV(\alpha)$ et $p_f^r(\alpha)$ représentent respectivement le volume de la cellule associée au point α et la proportion du faciès d'indice "f" en ce point. Les grandeurs $\bar{\phi}_f$ et $\bar{S}w_f$ sont les porosités et saturations en eau moyennes relatives au faciès d'indice f.

Le calcul de l'espérance $E[ooip(CP^r)]$ pour l'ensemble des cubes de proportions simulés définit une distribution de distance, dont les quantiles correspondent au rang des différents cubes de proportions simulés. Remarquons que cette approche nécessite la définition préalable d'un modèle pétrophysique ($\bar{\phi}_f$ et $\bar{S}w_f$). Cependant, plus généralement, on peut

¹ Pour chaque cube de proportions, il est possible de réaliser des simulations stochastiques en faciès, puis en propriétés pétrophysiques (porosité, perméabilité, saturation en eau). L'ensemble de ces réalisations définit autant d'accumulations en hydrocarbures (section 2.3).

s'affranchir de ce modèle en faciès en pondérant les différents faciès, ou en utilisant pour grandeur de classement $R^k[CP^r]$ l'espérance des faciès dans chaque cellule.

Conclusions et Perspectives (2)

La méthode développée ici se révèle être une alternative, avec un champ d'application plus large, à la méthodologie proposée dans la deuxième partie. Elle traite les problèmes en domaine de plate-forme carbonatée ou silico-clastique. Elle intègre à partir d'un certain nombre de scénarios de base l'incertitude relative au schéma sédimentaire au niveau des cubes de proportions, qui décrit les probabilités d'occurrence de faciès dans l'espace.

Cette approche considère pour cela le cube de proportions – par extension, le scénario géologique sous-jacent – comme une fonction spatiale aléatoire et permet, par l'utilisation des méthodes probabilistes et géostatistiques, de caractériser sa variabilité à travers la simulation stochastique d'un ensemble de cubes de proportions alternatifs. La construction d'un tel pool de modèles passe d'abord par l'estimation, en chaque point du milieu, des distributions des proportions de faciès. Celles-ci sont supposées suivre des lois de Dirichlet, qui représentent une famille de lois de probabilités particulièrement adaptées aux proportions. L'estimation des lois se fait à partir de quelques modèles construits sur la base de scénarios géologiques distincts, et issus des interprétations alternatives des données. Les techniques géostatistiques participent à la simulation des proportions de faciès en tout point du milieu souterrain. Elles permettent de conserver la continuité régionale du sous-sol par l'intermédiaire d'un modèle de corrélation spatiale, dont le choix dépend bien entendu de l'environnement de dépôt comme du degré de connaissance disponible, et peut à ce titre :

- (a) soit être supposé « infini » ;
- (b) soit déterminé par apprentissage à partir des cubes de proportions de base, qui sont utilisés pour estimer les lois de Dirichlet régionalisées ;
- (c) ou alternativement imposé en fonction de variogrammes fournis par le géologue, selon sa connaissance du milieu ;
- (d) ou encore être introduite de façon stochastique en simulant les paramètres du modèle de corrélation spatiale (variogramme) dans des distributions probabilistes.

De la même façon que les méthodes présentées à la partie 2, la méthodologie proposée remédie de cette manière aux caractères discrets et limités des études à plusieurs scénarios ou de combinaisons d'accumulations, qui généralement oublient l'incertitude sur le mo-

dèle probabiliste utilisé pour effectuer les simulations pétrophysiques. Elle évite par conséquent l'écueil de l'utilisation de distributions de probabilité uniques pour représenter notre état de connaissance sur les paramètres du modèle. Elle améliore alors l'évaluation de la variabilité (a) des proportions de faciès en chaque point du milieu, et (b) des grandeurs clés d'un champ qui en dépendent : accumulations, profils de production, réserves.

En définitive, les apports essentiels des travaux réalisés dans cette partie comprennent :

- le recentrage de la quantification de l'incertitude géologique directement au niveau des modèles de probabilités utilisés pour accomplir les traditionnelles simulations géostatistiques des propriétés pétrophysiques du milieu. C'est une méthode stochastique de simulation des cubes de proportions, qui au contraire d'un processus multi-réalisation, ne nécessite pas l'échantillonnage explicite des composantes du scénario géologique ;
- l'interprétation dans un cadre probabiliste du cube de proportions de faciès, et ce à deux niveaux : sur un plan régional d'une part, avec la construction des lois de probabilités, et sur un plan global d'autre part, lors de la simulation cubes de proportions avec l'utilisation de modèles de continuité pouvant être tirés dans une distribution ;
- la réconciliation des différents modèles sédimentaires concurrents que peut proposer le géologue. La méthode proposée extrait de ceux-ci un modèle synthétique qui est à l'origine des simulations en termes de cubes de proportions, que des sources d'informations autres que la géologie peuvent d'ailleurs contraindre ;
- la génération rapide de multiples cubes de proportions, tous cohérents en terme de géologie, et qui intègrent, avec leur incertitude, des interprétations et des données de nature ou de résolution variables ;

Dans le prolongement de ces travaux de thèse, il serait néanmoins intéressant de revenir entre autres sur :

- la généralisation de la méthodologie proposée à des modèles de base utilisés pour l'estimation des lois de Dirichlet régionalisées, qui présenteraient une architecture interne différente. Jusqu'à présent, seule la modélisation des proportions de faciès est réalisée ; un support géométrique commun à tous les modèles de base est adopté. L'idée consisterait à regrouper la construction du maillage vertical et le remplissage des propriétés du géomodèle au moyen des lois de Dirichlet. Un tel formalisme offre d'abord, la possibilité de simuler l'épaisseur des cellules le long de chaque paléo-verticale. L'épaisseur d'une cellule représente en définitive, une proportion de l'épaisseur stratigraphique totale. Ce formalisme peut en suivant, toujours être employé à la simulation des proportions de faciès. Par contre, l'estimation des lois de Dirichlet régionalisées réclamerait un paramétrage commun à l'ensemble du modèle de base. Une première voie consiste à travailler avec le paramétrage tiré de la simulation des épaisseurs. Un premier algorithme a été partiellement développé, qu'il est encore possible d'améliorer. En outre, une telle approche permettra la connexion avec les cubes de proportions issus d'un processus comme Neptune par exemple ;

-
- l'application des lois de Dirichlet aux procédés de changement d'échelle, ou *downscaling* en anglais. Ceux-ci visent à générer des modèles à haute résolution à partir de données plus grossières. L'imagerie sismique, par exemple, fournit bien souvent un aperçu limité des hétérogénéités lithologiques. Cependant, combinée aux techniques de changement d'échelle, elle peut permettre de connaître l'architecture interne des réservoirs. D'autres [Tran et al., 1999] réalisent un calage des historiques de production avec une étape de *downscaling* entre les modèles dynamique et statique¹. Ici, les lois de Dirichlet permettraient de redistribuer les proportions de faciès, ou toute autre propriété additive à l'instar de la porosité, qui sont connues à l'échelle de l'épaisseur du réservoir, le long des paléo-verticales sur les différentes couches qui composent le modèle. En effet, la proportion de faciès dans chaque cellule selon une paléo-verticale représente une fraction de la proportion globale de faciès ;
 - l'intégration des contraintes dynamiques. Les cubes de proportions simulés ici peuvent être amenés à respecter un large éventail de données. On citera les données issues de la sismique et les données de puits, qui se présentent respectivement sous la forme de cartes et courbes de proportions. D'autres contraintes peuvent encore restreindre l'espace des cubes de proportions simulés : les données dynamiques avec en particulier, les historiques de production. Pour cela, une piste envisageable recourrait aux procédés qui perturbent les réalisations géostatistiques des propriétés du sous-sol (faciès, porosité) [Hu et al., 2001 ; Caers, 2003]. Leur objectif consiste à optimiser, par perturbations successives du géomodèle, le calage du profil de production qui en résulte à un historique de production donné. Dans le cas du processus de Dirichlet, de telles perturbations s'effectueraient, non plus au niveau du remplissage en propriétés lithologiques et pétrophysiques, mais à l'étape de simulation des quantiles ;
 - le développement d'outils « pratiques » qui faciliteraient la création des scénarios de base afin d'alimenter le formalisme de Dirichlet. Si un procédé comme Neptune permet de générer des scénarios par modification de ses paramètres d'entrée, d'autres méthodes peuvent encore être développées dans ce sens dans gOcad : calculs des courbes verticales de proportions, méthodes de combinaisons des dérives horizontale et verticale de proportions ;
 - l'utilisation de méthodes alternatives aux simulations gaussiennes séquentielles pour la construction des quantiles, comme par exemple, les techniques de simulations séquentielles multipoints basées sur les images d'entraînement [Caers, 2001] ;
 - l'industrialisation de la méthodologie, un brevet a d'ailleurs été déposé dans ce sens. Sa mise en oeuvre passerait par deux voies : d'une part, sous la forme d'un *workflow*

¹ Le modèle **statique** renvoie au géomodèle construit pour la modélisation des propriétés pétrophysiques (Fig. 1-4). Il est souvent à une résolution plus fine que le modèle utilisé pour les étudier les écoulements au sein du réservoir, ce dernier correspondant au modèle dit **dynamique**. Il est construit sur la base du modèle statique au moyen d'un changement d'échelle : un ensemble de cellules qui prennent différentes valeurs sont regroupées dans une seule maille, dont la valeur est unique. Dans ce cas, on parlera d'*upscaling* en anglais.

indépendant. Un prototype est en cours de développement sur la plate-forme gOcad. D'autre part, il est possible d'implémenter ce formalisme dans Jacta, un *workflow* sous gOcad. C'est l'outil utilisé chez Total pour réaliser les simulations stochastiques des propriétés du réservoir. Les simulations en proportions seraient ainsi réalisées avant la simulation des propriétés pétrophysiques dans le déroulement du processus Jacta. Quelle que soit sa forme, le lien avec les travaux de Biver (2006), qui recourt à ce formalisme pour intégrer la variabilité statistique des proportions, est conservé, comme est conservé la possibilité de combiner des lois de Dirichlet issues de sources distinctes. D'ailleurs, ces travaux de thèse et ceux de Biver sont complémentaires ; une publication commune serait tout à fait indiquée pour conclure cette collaboration.

Annexe A

Notations utilisées

A.1 Abréviations

Voici quelques abréviations ou acronymes couramment employés dans ce mémoire, et qui se rencontrent également dans l'ensemble de la littérature spécialisée, qu'elle soit pétrolière, pétrophysique ou mathématique.

NTG	hauteur utile (de l'anglais Net To Gross) ; suivi de l'indice f , elle renvoie à la saturation en eau du faciès d'index f
OOIP	accumulations en hydrocarbures (traduit de l'anglais Original Oil In Place)
Φ , φ	porosité ; suivi de l'indice f , elle renvoie à la porosité du faciès d'index f
RV	volume du modèle (de l'anglais Rock Volume)
S_w	saturation en eau d'une roche ; suivi de l'indice f , elle renvoie à la saturation en eau du faciès d'index f
ACP	pour Analyse en Composantes Principales
cdf, Π	fonction de répartition d'une loi de probabilité (de l'anglais Cumulative Density Function)
exp	fonction exponentielle
$E[Z]$	espérance mathématique de la variable aléatoire Z
$\mathbf{h}(\cdot)$	correspond à une fonction vectorielle
$h(\cdot)$	renvoie à une fonction scalaire h
ln	fonction logarithme népérien

\log	fonction logarithme
$IP(A)$	probabilité de l'événement A
$IP(A B)$	probabilité de l'événement B conditionnellement à la réalisation de l'événement A
pdf	densité de probabilité d'une loi probabiliste (de l'anglais P robability D ensity F unction)
Π^{-1}	Π^{-1} désigne la fonction inverse de la fonction de répartition Π de la variable aléatoire considérée
SGS	S imulations G aussiennes S équentielles
SIS	Simulation séquentielle d'indicatrices (de l'anglais S equential I ndicator S imulation)
$\sigma^2[Z]$	variance de la variable aléatoire Z
$\sigma[Z;Z']$	covariance entre les variables aléatoires Z et Z'
\sum	opérateur de sommation
$[x_1, x_2, \dots, x_n]^T$	désigne un vecteur de composantes x_1, x_2, \dots , et x_n dans la base canonique de \mathbb{R}^n par défaut
$[x_1, x_2, \dots, x_n]$	désigne un n-uplet, n étant un entier
\tilde{z}	estimateur de la valeur (ou du vecteur) de conditionnement z^*
\bar{z}	valeur moyenne de la variable aléatoire Z
z^*	valeur (ou vecteur) de conditionnement de la variable (ou vecteur) aléatoire Z.
α ou uvw	marquent respectivement le nœud de la grille et l'indice de la cellule du modèle discrétisé de l'espace
α_w	nœud de la grille stratigraphique intersecté par la trajectoire d'un puits
$[c_1, c_2, \dots, c_n]$	paramètres de la loi de Dirichlet
$[c_1(\alpha), \dots, c_n(\alpha)]$	paramètres de la loi de Dirichlet dont dérive le vecteur proportion au nœud α
F	variable aléatoire discrète décrivant la nature du faciès, $f = 1, 2, \dots$ ou n_f
f	indice du faciès
$\Gamma(\cdot)$	fonction gamma, qui peut être considérée comme une extension aux nombres réels de la factorielle, c'est-à-dire par abus de notations $\Gamma(x) = \Gamma(x-1) !$ pour tout réel positif x

$\gamma_f(\cdot)$	variogramme du quantile associé au faciès d'indice f
n_f	nombre de faciès considéré dans les modèles
$n_m(G)$	nombre de modèles géologiques de base utilisés pour estimer les lois de Dirichlet
n_u, n_v, n_w	dimensions de la grille stratigraphique, les deux premières se rapportant à des plans stratigraphiques, la dernière à la verticale
p_f	probabilité d'occurrence (abusivement appelée proportion) du faciès d'indice f
q_{10}, q_{90} et q_{50}	désignent respectivement les premiers et derniers déciles ainsi que la médiane d'une population
$S(n)$	simplexe de \mathbb{R}^n défini par l'équation : $x_1 + \dots + x_n = 1$ où les x_i sont les coordonnées du point M dans la base canonique

A.2 Typographie :

A.2.1 Polices majuscule ou minuscule :

En ce qui concerne la modélisation des phénomènes aléatoires, régionalisés ou non, les lettres minuscules sont réservées à une valeur prise par la variable ou la fonction aléatoire, les lettres majuscules au modèle probabiliste sous-jacent. Le passage d'une majuscule à une minuscule correspond à une opération de réalisation de la variable ou de la fonction aléatoire.

Z	définit une Variable Aléatoire
z	correspond à une réalisation de la Variable Aléatoire Z
$Z(\cdot)$	désigne une Fonction Aléatoire Scalaire, fonction définie à la fois sur l'espace « géographique » et sur un espace probabilisé, dont les réalisations correspondent à des fonctions scalaires.
$z(\cdot)$	renvoie la Variable Régionalisée $z(x,y,z)$. C'est une fonction déterministe de l'espace à valeurs réelles

A.2.2 Formats gras ou normal :

Les caractères en gras renvoient à des variables ou des fonctions vectorielles, c'est-à-dire de dimension supérieure ou égale à deux. Les lettres au format normal correspondent à des scalaires. Ainsi, avec les notations habituelles :

- \mathbf{v} désigne un vecteur. Par exemple, $\mathbf{p}(\alpha)$ désigne le vecteur de proportions au point α
- v désigne un scalaire. Exemple : $p(\alpha)$ désignerait une proportion
- $\mathbf{v}(\cdot)$ désigne une fonction à valeurs vectorielles ; $\mathbf{p}(\cdot)$ désigne la fonction qui à tout point α de l'espace géographique associe un vecteur de proportions
- $v(\cdot)$ renvoie à une fonction à valeurs réelles. Ainsi, $p(\cdot)$ désignerait la fonction qui à tout point α de l'espace géographique associe une proportion

A.2.3 Combinaisons majuscule/minuscule et gras/normal :

Compte tenu des conventions précédentes, les différentes combinaisons possibles désignent respectivement :

- $\mathbf{Z}(\cdot)$ désigne une Fonction Aléatoire Vectorielle. C'est une fonction définie à la fois sur l'espace « géographique » et sur un espace probabilisé, dont les réalisations correspondent à des fonctions vectorielles
- $\mathbf{z}(\cdot)$ correspond à une Variable Régionalisée vectorielle $\mathbf{z}(x,y,z)$. Il s'agit d'une fonction déterministe de l'espace
- $z(\cdot)$ renvoie à la Variable Régionalisée $z(x,y,z)$. C'est une fonction déterministe de l'espace à valeurs réelles
- $Z(\cdot)$ désigne une Fonction Aléatoire Scalaire, fonction définie à la fois sur l'espace « géographique » et sur un espace probabilisé, dont les réalisations correspondent à des fonctions scalaires

Compte tenu de ces conventions, les notions rencontrées au cours de ces travaux de thèse peuvent être récapitulées dans le tableau suivant :

- $\mathbf{P}(\cdot) = [P_1(\cdot), \dots, P_{n_f}(\cdot)]^T$ désigne la Fonction Aléatoire Vectorielle associée aux vecteurs proportions de faciès dans l'espace. Chacune des fonctions $P_f(\cdot)$, $f=1, \dots, n_f$, renvoie à la Fonction Aléatoire relative à la proportions de faciès d'indice f
- $\mathbf{p}(\cdot) = [p_1(\cdot), \dots, p_{n_f}(\cdot)]^T$ définit une Variable Régionalisée vectorielle relative aux vecteurs de proportions sur l'ensemble de la grille stratigraphique : c'est une réalisation de $\mathbf{P}(\cdot)$. $p_f(\cdot)$, $f=1, \dots, n_f$, renvoie à une Variable Régionalisée scalaire associée à la proportion de faciès f en tout point du modèle

$\mathbf{P}(\alpha) = [P_1(\alpha), \dots, P_{n_f}(\alpha)]^T$	correspond au Vecteur Aléatoire associé aux vecteurs de proportions au point α . $P_f(\alpha), f=1, \dots, n_f$, renvoie à la variable Aléatoire relative à la proportion du faciès d'indice f au point α
$\mathbf{p}(\alpha) = [p_1(\alpha), \dots, p_{n_f}(\alpha)]^T$	renvoie au vecteur de proportions au point α . C'est la valeur de $\mathbf{p}(\alpha)$ au noeud α . Chacune des composantes $p_f(\alpha)$, où $f=1, \dots, n_f$, correspond à la proportion de faciès f au point α
$\mathbf{Q}(\cdot) = [Q_1(\cdot), \dots, Q_n(\cdot)]^T$	désigne la Fonction Aléatoire Vectorielle associée aux vecteurs de quantiles. Les fonctions $Q_f(\cdot), f=1, \dots, n_f$, renvoient respectivement à la Fonction Aléatoire propre à la proportions de faciès d'indice f
$\mathbf{q}(\cdot) = [q_1(\cdot), \dots, q_n(\cdot)]^T$	définit une réalisation de $\mathbf{Q}(\cdot)$ sur l'ensemble de la grille stratigraphique. $q_f(\cdot), f=1, \dots, n_f$, renvoie à une Variable Régionalisée scalaire associée au quantile de la proportion du faciès f sur en tout point du modèle
$\mathbf{Q}(\alpha) = [Q_1(\alpha), \dots, Q_{n_f}(\alpha)]^T$	correspond au Vecteur Aléatoire associé aux quantiles des proportions de faciès au point α . $Q_f(\alpha), f=1, \dots, n_f$, renvoie à la variable Aléatoire relative au quantile de la proportion du faciès d'indice f au point α
$\mathbf{q}(\alpha) = [q_1(\alpha), \dots, q_{n_f}(\alpha)]^T$	renvoie au vecteur de quantiles au point α . Les composantes $q_f(\alpha)$, où $f=1, \dots, n_f$, correspondent respectivement au quantile associé à la proportion du faciès d'indice f en α
$\mathbf{\Gamma}(\cdot) = [\Gamma_1(\cdot), \dots, \Gamma_{n_f}(\cdot)]^T$	désigne la Fonction Aléatoire Vectorielle dont les composantes renvoient aux Fonctions Aléatoires intermédiaires associées à des lois Gamma, et utilisées pour la simulation des lois de Dirichlet relatives à $\mathbf{P}(\cdot)$
$\mathbf{\Gamma}(\cdot) = [\Gamma_1(\cdot), \dots, \Gamma_{n_f}(\cdot)]^T$	définit une réalisation de $\mathbf{\Gamma}(\cdot)$ sur l'ensemble de la grille stratigraphique. $\Gamma_f(\cdot), f=1, \dots, n_f$, renvoie à une Variable Régionalisée scalaire ayant pour support la loi de probabilité Gamma associée à la composante f de la loi de Dirichlet de $\mathbf{P}(\cdot)$
$\mathbf{\Gamma}(\alpha) = [\Gamma_1(\alpha), \dots, \Gamma_{n_f}(\alpha)]^T$	désigne le Vecteur Aléatoire « Gamma » dont les composantes correspondent aux lois de probabilité Gamma utilisées pour la simulation de la loi de Dirichlet en α
$\mathbf{\Gamma}(\cdot) = [\Gamma_1(\alpha), \dots, \Gamma_{n_f}(\alpha)]^T$	renvoie à une réalisation au point α du vecteur « Gamma ». Ses composantes sont utilisés pour le calcul des proportions de faciès en ce point

Références bibliographiques

- AITCHINSON, J. & BEGG, C.B. (1976) – Statistical diagnosis when cases are not classified with certainty. *Biometrika* 63, pp. 1-12.
- AITCHINSON, J. & SHEN, S.M. (1980) – Logistic-normal distributions : Some properties and uses. *Biometrika* 67, pp. 261-72.
- AITCHINSON, J. (1986) – *The Statistical Analysis of Compositional Data*. London : Chapman and Hall.
- AITCHINSON, J. (1997) – The one-hour course in compositionnal data analysis or composition data analysis is simple. In Pawlowsky-Gahn (Ed.), *Proceedings of IAMG'97*. Barcelona : V.CIMNE, 35 p.
- AKAMINE, J. & CAERS, J. (2006) – Conditioning stochastic realizations to hard data with varying reliability. In *Stanford Center for Reservoir Forecasting Annual Affiliates Meeting*. Stanford, California.
- ALEXA M., COHEN-OR D. & LEVIN, D. (2000) – As-rigid-as-possible shape interpolation. In *SIGGRAPH '00: Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, New York, NY, USA. ACM Press/Addison-Wesley Publishing Co., pp. 157–164.
- ALEXA, M. (2002) – Recent advances in mesh morphing. *Comput. Graph. Forum* 21, 2, 173–196.
- ARNAUD, H. (2004a) – Typologie sommaire des plates-formes carbonatées. In PREAT, A., ARNAUD-VANNEAU, A., CARRIO, E., ARNAUD, H. & FERRY, S. (Ed.), *École d'été, Carbonates récifaux et de plate-forme*. Fontainebleau : Association des Sédimentologues Français.
- ARNAUD, H. (2004b) – Stratigraphie séquentielle appliquée aux plates-formes carbonatées. In PREAT, A., ARNAUD-VANNEAU, A., CARRIO, E., ARNAUD, H. & FERRY, S. (Ed.), *École d'été, Carbonates récifaux et de plate-forme*. Fontainebleau : Association des Sédimentologues Français.

- ARNAUD-VANNEAU, A. (2004) – Les microfaciès des plates-formes carbonatées. In PREAT, A., ARNAUD-VANNEAU, A., CARRIO, E., ARNAUD, H. & FERRY, S. (Ed.), *École d'été, Carbonates récifaux et de plate-forme*. Fontainebleau : Association des Sédimentologues Français.
- BARTHURST, R.G.C. (1975) – *Carbonate sediments and their diagenesis*. New-York : Elsevier Publ., 2nd Ed., 658 p.
- BAUER, W.F. (1958) – The Monte Carlo Method. *Journal of the Society for Industrial and Applied Mathematics*, 6(4): 438-451.
- BEHRENS, R.A., MACLEOD, M.K., TRAN, T.T. & ALIMI, A.O. (1996) – *Incorporating seismic attribute maps in 3-D reservoir models*. Annual Meeting Abstracts, Society Of Exploration Geophysicists.
- BELLMAN, R.E. (1961) – *Adaptive Control Processes*. Princeton, NJ : Princeton University Press.
- BERGAMO, L. (2002) – Évaluation des incertitudes géologiques dans les modélisations de réservoirs carbonatés : impact sur les profils de production. Rapport de stage industriel de Mastère II : Université Joseph Fourier, Grenoble, France.
- BESSE, P. (2003) – *Data mining : Modélisation Statistique & Apprentissage de nombres aléatoires*. Publication du Laboratoire de Statistiques et Probabilités, Université Paul Sabatier, Toulouse, France. Disponible en ligne en décembre 2004 sur www.lsp.ups-tlse.fr/Besse.
- BIVER, P. (1998) – *Conseils méthodologiques pour l'élaboration de calculs d'accumulations probabilisées*. Note interne Total S.A, Pau.
- BIVER, P., HAAS, A. & BACQUET, C. (2002) – Uncertainties in facies proportion estimation II. Application to geostatistical simulation of facies and assessment of volumetric uncertainties. *Mathematical Geology*, Volume 34, No 6.
- BIVER, P. (2006) – Combinaison de cubes de proportions issus de scénarios géologiques différents. Note Interne Total S.A, Pau.
- BLOCH, D.A. & WATSON, G.S. (1967) – A Bayesian study of the multinomial distribution. *Ann. Math. Statist.* 38, pp. 1423-35.
- BOULEAU, N. (1986) – *Probabilité de l'Ingénieur*. Hermann.
- BOULVAIN, F. (2005) – *Cours en ligne de sédimentologie, de pétrologie sédimentaire et sur les processus sédimentaire*. Ouvrage collectif du laboratoire de Pétrologie Sédimentaire, Université de Liège, Belgique. Disponible en ligne en avril 2005 sur <http://www.ulg.ac.be/geolsed>.
- BOUZID, I. (2003) – *Développement de NeptuneWorkflow, outil de modélisation et simulation de réservoirs carbonatés en environnement marin*. Rapport de stage industriel de 3^{ème} année : École Nationale Supérieure de Géologie, Nancy, France.

-
- BRETHON A., CARAUX, G. & SAPORTA, G. (2002) – *Cours en ligne de statistique*. Techniques de la statistique. Ouvrage collectif CNAM/CRA Languedoc-Roussillon / AGRO de Montpellier/Université Montpellier II. Disponible en ligne sur <http://www.agro-montpellier.fr/cnam-lr/statnet>.
- BUSLIK, A. (1993) – Bayesian approach to model uncertainty. In *Workshop Model Uncertainty : its Characterization and Quantification*, Annapolis (Maryland, USA).
- CAERS, J. (2001) – Geostatistical reservoir modelling using statistical pattern recognition. *Journal of Petroleum Science and Engineering*, 29 : pp. 177-188.
- CAERS, J. (2003) – History matching under training image based geological model constraints. *SPE Journal*, 218.
- CARRIO-SCHAFFHAUSER, E. (2004) – La diagenèse en milieu carbonaté. In PREAT, A., ARNAUD-VANNEAU, A., CARRIO, E., ARNAUD, H. & FERRY, S. (Ed.), *École d'été, Carbonates récifaux et de plate-forme*. Fontainebleau : Association des Sédimentologues Français.
- CAUMON, G., STREBELLE, S., CAERS, J. K. & JOURNEL, A.G. (2004) – Assessment of Global Uncertainty for Early Appraisal of Hydrocarbon Fields. In *SPE Annual Technical Conference and Exhibition*, Houston, Texas. Paper SPE 89943.
- CHAPELLIER, D., BARON, L., MARESCOT, M., MARI, L., MONNET, R., ROCHAT, S. & TACCHINI, G. (2004) – *Cours en ligne de géophysique* sur <http://www-ig.unil.ch/cours>, Université de Lausanne, Institut Français du pétrole.
- CHILÈS, J.P. & DELFINER, P. (1999) – *Geostatistics : Modeling Spatial Uncertainty*. Wiley éd., Wiley Series in Probability and Statistics: Applied Probability and Statistics Section.
- CLEMEN, R.T. & WINKLER, R.L. (1999) – Combining probability distributions from experts in risk analysis. *Risk Analysis*, 19, pp. 187–203.
- CORRE, B., THORE, P., DE FERAUDY, V., VINCENT, G. (2000) – Integrated Uncertainty Assessment For Project Evaluation and Risk Analysis. In *SPE European Petroleum Conference*, Paris, France. Paper SPE 65205.
- COSSÉ, R. (1993) – *Basics of reservoir engineering*. Paris : Éditions Technip, 376 p.
- DE BOOR, C. (1978) – *A Practical Guide to Splines*. New York : Springer-Verlag, cop. 2001, coll. Applied Mathematical Sciences, Vol. 27, 1^{ère} éd., XVIII, 372 p., 32 illus.
- DELAHAYE, J.P. (2007) – *Générateurs de nombres aléatoires*. Disponible en ligne en août 2007 sur www.univ-irem.fr/commissions/ci3m/CDci3m/caen/nbaleat/nbalea.htm.
- DEUTSCH, C.V. & JOURNEL, A.G. (1998) – *GSLIB : Geostatistical Software Library and User's Guide*. New York : Oxford University Press, 340 p.

- DEUTSCH, C.V. (1992) – *Annealing techniques applied to reservoir modelling and the integration of geological and engineering (well test) data*. PhD Thesis of Applied Earth Sciences submitted to the Committee on Graduate Studies of Stanford University, USA.
- DEVROYE, L. (1986) – *Non-Uniform Random Variate Generation*. New York : Springer-Verlag.
- DOYEN P., GUIDISH, T. & DE BUYL, M. (1989) – Seismic Discrimination of lithology in sand/shale reservoir - A Bayesian approach. In *SEG Annual Meeting*, Dallas, SEG edition.
- DOYEN, P.M. (1988) – Porosity from seismic data: A geostatistical approach. *Geophysics*, Volume 53, Issue 10, pp. 1263-1275.
- DUNHAM, R. L. (1962) – Classification of carbonate rock according to depositional texture. In HAM, W. G. (Ed), *Classification of carbonate rocks*, American Association of Petroleum Geologists Memoir, 1, pp. 108–121.
- EFRON B. & GONG G. (1983) – A leisurely look at the bootstrap, the jackknife and cross-validation. *The American statistician*, vol.37, No 1, pp 36-48.
- EFRON, B. et TIBSHIRANI, R.J. (1993) – *An Introduction to the bootstrap*. New York, N.Y., Chapman and Hall Monographs on Statistics and applied probability, vol. 57, 436 p.
- FÉLIX, R & XANTHOULIS, D. (2005) – Analyse de sensibilité du modèle mathématique “Erosion Productivity Impact Calculator” (EPIC) par l’approche One-Factor-At-A-Time (OAT). *Biotechnol. Agron. Soc. Environ.* 9 (3), 179-190.
- FOUCAULT, A. & RAOULT, J-F. (1995) – *Dictionnaire de Géologie*. Paris : Masson, 4^e Éd.
- FOURNIER, F. (1990) – *Application de l’analyse des données à l’exploration pétrolière – Caractérisation des faciès sismiques par analyses statistiques multivariées*. Thèse Doct. Géosciences, Institut National Polytechnique de Lorraine, Nancy, France.
- GAZIGLIA, G. (2001) – *Développement d’outils utilisant la bathymétrie pour une modélisation 3D par le logiciel Neptune*. Rapport de stage industriel de 2^{ème} année : École Nationale Supérieure de Géologie, Nancy, France.
- GOMES, J., VELHO, L, COSTA, B. & DARSA, L. (1997) – *Course on Warping and Morphing of graphical objects*, Los Angeles, California. Disponible en ligne en mars 2005 sur <http://w3.impa.br/~morph/sig-course/about.html#Introduction>.
- GONOD, P. (1996) – *Dynamique des systèmes et méthodes prospectives*. TRP NO2, Futuribles International.
- GOOVAERTS, P. (1997) – *Geostatistics for natural resources evaluation*. Applied Geostatistics Series, New-York : Oxford University Press, 483 p.
- GOVAERT, G. & AMBROISE, C. (2004) – *Analyse de Données et Data Mining*. Disponible en ligne en février 2005 sur www.etu.utc.fr/rentree/catuv/pdf/SY09.pdf.

-
- GUÉMÉNÉ, J.M. & SCHNAPPER, D. (1998) – *Les Incertitudes Structurales : Approche Pratique*. Rapport interne Elf EP/T/ER/T3G NO JMG-98-038, Pau.
- HASS, A. & FORMERY, P. (2002) – Uncertainties in facies proportion estimation I. Theoretical framework: the Dirichlet distribution. *Mathematical Geology*, Volume 34, No 6.
- HINES W.W., MONTGOMERY, D.C., GOLDSMAN, D.M. & BORROR, C.M. (2003) – *Probability and statistics in engineering*. Phoenix : John Wiley and Sons Inc.
- HOFFMAN, T.B. (2005) – History matching while perturbing facies. PhD Thesis, Stanford University.
- HOMWOOD, P.W., MAURIAUD, P. & LAFONT, F. (2000) – *Best Practices in Sequence Stratigraphy for explorationists and reservoir engineers*. [Vade-mecum de Stratigraphie Séquentielle pour géologues, géophysiciens et ingénieurs réservoir]. Bull centre Recherche Elf Explor. Prod., Mem. 25, 81 pp. 50 fig.
- HOTELLING, H. (1933) – Analysis of a complex of a statistical variables into principal components. *J. Educ. Psy.* 24, p. 417-441, p. 498-520.
- HU, L-Y., BLANC, G. & NOETINGER, B. (2001) – Gradual deformation and iterative calibration of sequential stochastic simulations. *Mathematical Geology* 33, No 4, 475.
- JACKSON, J.E. (1991) – *A User's Guide to Principal Components*. New York, N.Y : Wiley.
- JACQUES, J. (2005) - *Contributions à l'analyse de sensibilité et à l'analyse discriminante généralisée*. Thèse Doct. Mathématiques Appliquées, Université Joseph Fourier. Grenoble, France, 209 p.
- JOHANN, P., FOURNIER, F., SOUZA, O.Jr., ESCHARD, R. & BEUCHER, H. (1996) – 3D stochastic reservoir modeling constrained by well and seismic data on a turbidite field. In *SPE Annual Technical Conference and Exhibition*. Paper SPE 36501.
- JOURNEL, A.G. (1989) – *Fundamentals of Geostatistics in Five Lessons*. Short course presented at the 28th International Geological Congress, Washington D.C., American Geophysical Union.
- JOURNEL, A.G. (1994) – Resampling from stochastic simulations. *Environmental and Ecological Statistics* 1, No. 1, 63.
- JOURNEL, A.G. (1998) – Combining knowledge from diverse data sources : an alternative to traditional data independence hypothesis. *Mathematical Geology*, v. 34 573-596.
- LABAT, L. (2004) - *Simulations stochastiques de faciès par la méthode des membership functions*. Thèse Doct. Géosciences, Institut National Polytechnique de Lorraine, Nancy, 105 p.
- LAZARUS, F. (1995) – *Courbes, Cylindres et Métamorphoses pour l'image de synthèse*. Thèse Doct. Informatique, Université de Paris VII, France, 393 p.

- LE LOC'H, G. & GALLI, A. (1999) – *Truncated Plurigaussian Method: Theoretical and Practical Points of View*. In *Geostatistics Wollongong 96 - Proceedings of the Fifth International Geostatistics Congress*, Wollongong, Australia, p.211-222.
- LEBART, L., MORINEAU, A. & PIRON, M. (1995) – *Statistique exploratoire multidimensionnelle*. Paris : Dunod, 439 p.
- LEFLON, B. (2002) – *Développement de Neptune, outil de modélisation stochastique de réservoirs en environnement marin*. Rapport de stage industriel de 3^{ème} année : École Nationale Supérieure de Géologie, Nancy, France.
- LEFLON, B. (2005) – *Modélisation des hétérogénéités lithologiques à l'échelle du réservoir pétrolier en milieu marin et fluvial*. Thèse Doct. Géosciences, Institut National Polytechnique de Lorraine, Nancy, France.
- LEFLON, B., MALLET, J.-L. & MASSONNAT, G.J. (2004) – Neptune project : use of discrete smooth interpolation in an accommodation-based simulator. In *The 24th Gocad Meeting Proceedings*, Nancy, France. Association Scientifique pour la Géologie et ses Applications.
- LEONARD, T. (1973) – A Bayesian method for histograms. *Biometrika* 60, pp. 287-308.
- LEVERETT, M.C. (1941) – *Capillary behavior in porous solids*. AIME Trans. 142, pp. 152–168.
- LINDLEY, D.V. (1964) – The Bayesian analysis of contingency tables. *Ann. Math. Statist.* 35, pp. 1622-43.
- LINDLEY, D.V., TVERSKY, A. & BROWN, R.V. (1979) – On the reconciliation of probability assessments. *J. R. Statist. Soc. A* **142**, pp. 146-80.
- LOPEZ, S. (2003) – *Modélisation stochastique et générique de réservoirs pétroliers canalisés et méandriques*. Thèse Doct. Géostatistiques/Géologie des bassins sédimentaires, Centre de géostatistiques, École Nationale Supérieure des Mines, Paris, France.
- MALLET, J.L. (2002) – *Geomodeling*. Oxford Press.
- MALLET, J.L. & SHTUKA, A. (1997) – Stratigraphic Inversion Integrating Wells, Seismic and Sedimentology Data. In *The 67th Annual International Soc. of Expl. Geoph.*, pp. 1632–1635.
- MALLET, J.L. (1992) – *Discrete Smooth Interpolation in Geometric Modeling*. *Computer-Aided Design*, 24(4):177–191.
- MALLET, J.-L. (2004) – Space–time mathematical framework for sedimentary geology. *Mathematical Geology*, vol. 36, no1, pp. 1–32.
- MAO, S. & JOURNEL, A.G. (1999) – Conditional 3D simulation of lithofacies with 2D seismic data. *Computers and Geosciences*, 25 : pp. 845–862.

-
- MARCOTTE, D. (2004) – *Cours en ligne de Géologie et géostatistique minières*. École Polytechnique de Montréal. Disponible en avril 2005 sur <http://geo.polymtl.ca/~marcotte/g1q3401geo.html>.
- MASSONNAT, G.J. & PERNARCIC, E. (2002a) – *Assessment and Modeling of High Permeability Areas in Carbonate Reservoirs*. In *SPE Annual Technical Conference and Exhibition*, San Antonio, Texas. Paper SPE 77591.
- MASSONNAT, G.J. & PERNARCIC, E. (2002b) – NEPTUNE : An Innovative Approach to significantly Improve Reservoir Modeling in Carbonate reservoirs. In *Adipec Conference*, Abu Dhabi, EAU. Paper SPE 78528.
- MASSONNAT, G.J. (1999) – Breaking of a paradigm : geology can provide 3D complex probability fields for stochastic facies modelling. In *SPE Annual Technical Conference and Exhibition*, Houston, Texas. Paper SPE 56652.
- MASSONNAT, G.J. (2000) – Can we sample the complete geological uncertainty space in reservoir-modeling uncertainty estimates? *SPE Journal* 5(1), Paper SPE 59801, pp. 46–59.
- MASSONNAT, G.J. (2004) – Modélisation géologique des sédiments carbonatés. In PREAT, A., ARNAUD-VANNEAU, A., CARRIO, E., ARNAUD, H. & FERRY, S. (Ed.), *École d'été, Carbonates récifaux et de plate-forme*. Fontainebleau : Association des Sédimentologues Français.
- MATHERON G., BEUCHER, H., DE FOUQUET C. & GALLI, A. (1987) – Conditional simulation of the geometry of fluvio-deltaic reservoirs. In *The 62nd Conference of the Society of Petroleum Engineers*. Paper SPE 16753.
- MATHERON, G. (1962) – *Traité de Géostatistique Appliquée*. Paris : Éd. Technip, Tome 1.
- MATHERON, G. (1965) – *Les variables régionalisées et leur estimation*. Paris : Masson.
- MCKAY, M.D. (1995) – Evaluating prediction uncertainty. Technical Report NUREG/CR-6311, US Nuclear Regulatory Commission and Los Alamos National Laboratory, 1995.
- MCLACHLAN, G.J. & Peel, D. (2000) – *Finite Mixture Model*. New York : Wiley.
- MOORE, C.H. (2001) – Carbonate reservoirs – Porosity evolution and diagenesis in a sequence stratigraphic framework. *Developments in Sedimentology*, 55, Elsevier.
- MOYEN, R. (2001) – NEPTUNE : *Modélisation de faciès contrôlée par la bathymétrie*. Rapport de stage industriel de 3^{ème} année : École Nationale Supérieure de Géologie, Nancy, France.
- MOYEN, R. (2005) – *Paramétrisation 3D de l'espace en géologie sédimentaire : le modèle GeoChron*. Thèse Doct. Géosciences, Institut National Polytechnique de Lorraine, Nancy, France.
- NIEDERREITER, H. (1992) – *Random Number Generation and Quasi-Monte Carlo Methods*. Philadelphia : SIAM.

- NORRIS, R., MASSONNAT, G.J. & ALABERT, F. (1993) – *Early quantification of uncertainty in the estimation of oil in place in a turbidite reservoir*. In SPE Annual Technical Conference and Exhibition, Houston, Texas. Paper SPE 26490
- OWEN, A. (1998) – Monte carlo extension of quasi-monte carlo. In *1998 Winter Simulation Conference*, Washington (DC, USA).
- PAIRAZIAN, K. (1998) – *Modélisation 3D des Réservoirs Pétroliers dans l'Intégration des Données Sismiques et Géologiques : Approches Quantitatives Multivariées*. Thèse Doct. Géosciences, Institut National Polytechnique de Lorraine, Nancy, France.
- PEARSON, K. (1901) – On lines and planes of closest fit to systems of points in space. *Phil Mag.* 2, No11, pp. 559-572.
- PERNARCIC, E. (2002) – *Neptune, Principes et Exemples d'applications*. Rapport interne Total S.A : DGEP/GSR/THEM/CARB, Pau.
- PERRODON, A. (1990) – *L'homme et le pétrole*. Lyon: Muséum de Lyon, 1 vol. 84 p.
- PERRODON, A. (1999) – *Quel pétrole demain ?* Paris : Éd. Technip, 1 vol. 95 p.
- POIROT, N. (2002) – *Évaluation des incertitudes géologiques dans les modélisations de réservoirs carbonatés*. Rapport de stage industriel de 3^{ème} année : École Nationale Supérieure des Mines, Nancy, France.
- RAFTERY, A.E. Bayesian model selection in social research. *Sociological Methodology 1995*. Oxford, U.K : Blackwells, Peter V. Marsden Edition, 1995, pp. 111-196.
- ROBERT, C. & CASELLA, G. (1999) – *Monte Carlo Statistical Methods*. New York : Springer.
- SALTELLI, A., CHAN, K. & SCOTT, E.M. (2000b) – Sensitivity Analysis. *Probability and Statistics series*. New York: John Wiley, 475 p.
- SALTELLI, A., TARANTOLA, S. & CHAN KPS. (1999). A quantitative model-independent method for global sensitivity analysis of model-output. *Technometrics* 41 (1), p. 39–49.
- SALTELLI, A., Tarantola, S. & Compolongo, F. (2000a) – Sensitivity analysis as an ingredient of modelling. *Stat. Sci.* 15 (4), p. 377–395.
- SAPORTA, G. (1990) – *Probabilités, Analyse des données et Statistiques*. Paris : Éditions Technip.
- SHANNON, C. (1948) – A mathematical theory of communication. *Bell System Technical Journal*, 27, pp. 379–623.
- SHEPPERS, K.C. (2003) – *Quantification of uncertainty in reservoir simulations influenced by varying input parameters, Maria reservoir, Cahu field*. PhD thesis of Geology submitted to the Office of Graduate Studies of Texas A&M University, 155 p.

-
- STREBELLE, S., JOURNEL, A.G. & CAERS, J. (2001) – Data integration with multiplepoint geostatistics. In *The Third IMA Conference on Modeling Permeable Rocks*, Cambridge.
- SUCQUET, C. (2003) – Introduction au calcul des probabilités. Université des Sciences et Technologies de Lille, France, U.F.R. de Mathématiques Pures et Appliquées. Disponible en ligne en mars 2004 sur <http://math.univ-lille1.fr/~sucket>.
- SURAZHISKY, V. & GOTSMAN, C. (2003) – Intrinsic morphing of compatible triangulations. *International Journal of Shape Modeling* 9, 2, pp. 191–201.
- TAMHANE, D., WANG, L. & WONG, P.M. (1999) – The role of geology in stochastic reservoir modelling: The future trends. In *SPE Asia Pacific Oil and Gas Conference and Exhibition*, Jakarta, Indonesia. Paper SPE 54307, 5 pp.
- TRAN, T.T., WEN, X.-H. & BEHRENS, R.A. (1999) – Efficient Conditioning of 3D Fine-Scale Reservoir Model to Multiphase Production Data Using Streamline-Based Coarse-Scale Inversion and Geostatistical Downscaling. In *SPE Annual Technical Conference and Exhibition*, Houston, Texas. Paper SPE 56518.
- VAIL, P.R., MITCHUM, R.M.JR., TODD, R.G., WIDMERI, J.W., THOMSON, S., SANGREE, J.B., BUBB, J.N. & HATELID, W.G. (1977) – Seismic Stratigraphy and global changes of Sea Level. In PAYTON, C. E. (Ed), *Seismic Stratigraphy – Application to Hydrocarbon Exploration*. Tulsa : American Association of Petroleum Geologists Memoir, 26, pp. 49-219.
- VERDEL, T. (2000) – *Décisions et Prévisions Statistiques*. Cours de statistiques et d'analyse des risques de l'École des Mines, Nancy, France.
- VISEUR, S. (2001) – *Simulation stochastique basée-objet de chenaux*. Thèse Doct. Géosciences, Institut National Polytechnique de Lorraine, Nancy, France.
- VIVIER, S. (2002) – *Stratégies d'optimisation par la méthode des plans d'expériences et Application aux dispositifs électrotechniques modélisés par éléments finis*. Thèse Doct. Génie Électrique, Ecole Centrale de Lille et Université des Sciences et technologie de Lille, France.
- WHEELER, H.E. (1958) – *Time-stratigraphy*. Bulletin of the American Association of Petroleum Geologists, vol. 42, No 5, pp. 1047–1063.
- WILSON, J.L. (1975) – *Carbonate facies in geologic history*. New-York : Springer Verlag, 471 pp.
- WINKLER, L. (1993) – Model uncertainty : Probabilities for models ? In *Workshop "Model Uncertainty : its Characterization and Quantification"*, Annapolis (Maryland, USA).
- XU, W. (1992) – Integrating seismic data in reservoir modeling : the collocated cokriging alternative. In *SPE Annual Technical Conference & Exhibition*, Washington, D.C. Paper SPE 24742.

Références Bibliographiques

ZERKOUNE, A. (2003) – *Évaluation des incertitudes géologiques dans la modélisation Neptune*. Rapport de stage industriel de 3^{ème} année : École Nationale Supérieure des Mines de Nancy, France.

Résumé

Modélisation de l'incertitude géologique par simulation stochastique de cubes de proportions de faciès - Application aux réservoirs pétroliers de type carbonaté ou silico-clastique

Après sa découverte, les choix relatifs au développement d'un gisement se prennent sur la base de représentations incertaines du champ. En effet, sa caractérisation utilise des modèles numériques spatiaux porteurs de l'incertitude liée à la complexité du milieu souterrain. D'ordinaire, les méthodes de simulations stochastiques, qui génèrent des modèles équiprobables du sous-sol, sont supposées les quantifier. Néanmoins, ces images alternatives du champ renvoient à des tirages au sein d'un modèle probabiliste unique. Elles oublient l'incertitude relative au choix du modèle probabiliste sous-jacent, et tendent à la sous-estimer. Ce travail de recherche vise à améliorer la quantification de cette incertitude. Elle retranscrit la part de doute relative à la compréhension des propriétés du milieu sur les modèles probabilistes, et propose de l'intégrer à ce niveau. Cette thèse précise d'abord la notion d'incertitude en modélisation pétrolière, en particulier sur les modèles géologiques 3D comprenant différents faciès. Leur construction demande au préalable de définir en tout point de l'espace leur probabilité d'existence : c'est le cube de proportions. Généralement, bien que ces probabilités soient peu connues, les méthodes actuelles d'évaluation de l'incertitude sédimentaire les gardent figées. De fait, elles oublient le caractère incertain du scénario géologique et son impact sur le cube de proportions. Deux méthodes stochastiques de simulation ont été développées afin de générer des modèles équiprobables en termes de cubes de proportions. Elles intègrent la variabilité liée aux proportions de faciès, et explorent dans son ensemble un tel domaine d'incertitude. La première reste relativement attachée à la géologie. Elle intègre directement l'incertitude liée aux paramètres qui composent le scénario géologique. Elle décrit sa mise en oeuvre sur les divers paramètres du scénario géologique, qu'ils prennent la forme de signaux aux puits, de cartes ou d'hypothèses plus globales à l'échelle du réservoir. Une démarche de type Monte-Carlo échantillonne les composantes du schéma sédimentaire. Chaque tirage permet de construire un cube de proportions par l'intermédiaire d'un géomodeleur qui intègre de façon plus ou moins explicite les paramètres du scénario géologique. La méthodologie est illustrée et appliquée à un processus inédit de modélisation des dépôts carbonatés en milieu marin. La seconde revêt un caractère plus géostatistique en se concentrant davantage sur le cube de proportions. Elle vise plutôt à réconcilier les différents modèles sédimentaires possibles. Dans le modèle maillé de réservoir, elle estime la loi de distribution des proportions de faciès cellule par cellule - supposées suivre une loi de Dirichlet, à partir de quelques modèles, construits sur la base de scénarios géologiques distincts. Elle simule alors les proportions de façon séquentielle, maille après maille, en introduisant une corrélation spatiale (variogramme) qui peut être déterministe ou probabiliste. Divers cas pratiques, composés de réservoirs synthétiques ou de champs réels, illustrent et précisent les différentes étapes de la méthode proposée.

Mots clés : cube de proportions de faciès, géostatistique, incertitude, simulation, loi de Dirichlet

Abstract

Geological uncertainty modelling using stochastic simulations of facies proportion cubes – Application to hydrocarbon reservoirs : carbonates and silici-clastics

After finding out a potential oil field, development decisions are based on uncertain representations of the reservoir. Indeed, its characterisation uses numerical, spatial models of the reservoir. However, if they are representative of subsoil heterogeneities, the uncertainty linked to subsoil complexity remain. Usually, uncertainty is supposed to be assessed using many equiprobable models, which represent the heterogeneities expected into the reservoir. Nevertheless, those alternative images of the underground correspond to multiple realizations of a given and a single stochastic model. Those methods ignore the uncertainty related to the choice of the underlying probabilistic model. This work aims at improving that kind of uncertainty assessment when modelling petroleum reservoir. It conveys the doubt linked with our subsoil properties understanding on probabilistic models, and proposes to integrate it on them. This thesis first defines uncertainty in the context of oil industry modelling, particularly on 3D geological models comprising several litho-types or facies. To build them, we need, before any simulations, to estimate for every point in the space the probability of occurring for each facies : this is the proportions cube. Even though those probabilities are often poorly known, they are frozen while using current methods of uncertainty assessment. So, the impact of an uncertain geological scenario on the definition of a proportion cube is forgotten. Two methods based on stochastic simulations of alternative, equiprobable proportion cubes have been developed to sample the complete geological uncertainty space. The first one is closely linked to geology. It integrates directly uncertainty related to the parameters composing the geological scenario. Based on a multi-realisation approach, it describes its implementation on every parameters of geological scenario from information at wells to maps or global hypothesis at reservoir scale resolution. A Monte Carlo approach samples the components of the sedimentary scheme. Each drawing enables to build a proportion cube using modelling tools which integrates more or less explicitly parameters of geological scenario. That methodology is illustrated and applied to a modelling process which is used to model marine carbonate deposits. The second method appears to be more geostatistics focussing on proportion cubes. It rather aims at reconcile distinct eventual sedimentary models. In the meshed model symbolising the reservoir, it assesses the probabilistic law of facies proportion in each cells – they are supposed to follow Dirichlet's probabilistic law. That assessment is done from some models inferred from different geological scenarios. Facies proportions are sequentially simulated, cell after cell, introducing a spatial correlation model (variogram), which could be deterministic as probabilistic. Various practical cases, comprising synthetic reservoirs or real field, illustrates and specifies the different steps of the proposed method.

Keywords : facies proportion cube, geostatistics, uncertainty, simulation, Dirichlet's law