Connectivity of channelized sedimentary bodies: analysis and simulation strategies in subsurface modeling

Connectivité de corps sédimentaires chenalisés : stratégies d'analyse et de simulation en modélisation de subsurface

THÈSE

présentée et soutenue publiquement le 15 mars 2016
pour l'obtention du grade de

Docteur de l'Université de Lorraine et de l'Université de Neuchâtel

Spécialité Géosciences

par

Guillaume RONGIER

Composition du jury:

Rapporteurs:	Prof. Klaus MOSEGAARD	University of Copenhagen
	Prof. Thierry MULDER	Université de Bordeaux
	Dr. Michael PYRCZ	Chevron
Examinateurs:	Dr. Julien STRAUBHAAR	Université de Neuchâtel
	Dr. Vanessa TELES	IFP Energies nouvelles
Directeur de thèse:	Prof. Philippe RENARD	Université de Neuchâtel
Directrice de thèse:	Prof. Judith SAUSSE	Université de Lorraine
Co-directrice de thèse:	Dr. Pauline COLLON	Université de Lorraine
Invités:	Prof. Gregoire MARIETHOZ	Université de Lausanne
	Dr. Sebastien STREBELLE	Chevron

GeoRessources – UMR 7359
Université de Lorraine – ENSG / CNRS / CREGU
2 rue du Doyen Marcel Roubault, TSA 70605
54518 Vandœuvre-lès-Nancy, France

Institut d'Hydrogéologie et de Géothermie (CHYN)
Université de Neuchâtel – Faculté des sciences
Bâtiment UniMail, Rue Emile-Argand 11
2000 Neuchâtel, Suisse

UNIVERSITÉ DE NEUCHÂTEL

Faculté des sciences
Secrétariat-décanat de Faculté
Rue Emile-Argand 11
2000 Neuchâtel - Suisse
Tél: + 41 (0)32 718 2100
E-mail: secretariat.sciences@unine.ch

IMPRIMATUR POUR THESE DE DOCTORAT

La Faculté des sciences de l'Université de Neuchâtel autorise l'impression de la présente thèse soutenue par

Monsieur Guillaume RONGIER

Réalisée en cotutelle avec

L'Université de Lorraine, France

Titre :

« Connectivity of channelized sedimentary bodies: analysis and simulation strategies in subsurface modeling »

sur le rapport des membres du jury:

- Prof. Philippe Renard, Université de Neuchâtel, directeur de thèse
- Prof. Judith Sausse, Université de Lorraine, directrice de thèse
- Dr Pauline Collon, Université de Lorraine, co-directrice de thèse
- Prof. Klaus Mosegaard, Université de Copenhague, Danemark
- Prof. Thierry Mulder, Université de Bordeaux, France
- Dr Michael Pyrcz, Chevron, USA
- Dr Julien Straubhaar, Université de Neuchâtel
- Dr Vanessa Teles, IFPEN, France

Neuchâtel, le 1er avril 2016

Le Doyen, Prof. B. Colbois

« Son petit visage en forme de cœur paraissait dur sous le voile de deuil que formait ses cheveux noirs. »

Philip K. Dick
Le Bal des schizos, 1972

Abstract

Keywords turbiditic channel connectivity connected component stochastic simulation Lindenmayer system channel migration

Channels are the main sedimentary structures for sediment transportation and deposition from the continents to the ocean floor. The significant permeability of their deposits enables fluid circulation and storage. As illustrated with turbiditic systems, those channel fill is highly heterogeneous. Combined with the spatial organization of the channels, this impacts significantly the connectivity between the permeable deposits. The scarcity of the field data involves an incomplete knowledge of these subsurface reservoir architectures. In such environments, stochastic simulations are used to estimate the resources and give an insight on the associated uncertainties. Several methods have been developed to reproduce these complex environments. They raise two main concerns: how to analyze and compare the connectivity of a set of stochastic simulations? How to improve the representation of the connectivity within stochastic simulations of channels and reduce the uncertainties?

The first concern leads to the development of a method to objectively compare realizations in terms of connectivity. The proposed approach relies on the connected components of the simulations, on which several indicators are computed. A muldimensional scaling (MDS) representation facilitates the realization comparison. The observations on the MDS are then validated by the analysis of the heatmap and the indicators. The application to a synthetic case study highlights the difference of connectivity between several methods and parameter values to model channel/levee complexes. In particular, some methods are far from representing channel-shaped bodies.

Two main issues derive from the second concern. The first issue is the difficulty to simulate a highly elongated object, here a channel, conditioned to well or seismic-derived data. We rely on a formal grammar, the Lindenmayer system, to stochastically simulate conditional channel objects. Predefined growth

rules control the channel morphology to simulate straight to sinuous channels. That morphology conditions the data during its development thanks to attractive and repulsive constraints. Such constraints ensure the conditioning while preserving at best the channel morphology. The second issue arises from the limited control on the channel organization. This aspect is addressed by taking into account the evolution processes underlying channel organization. An initial channel is simulated with a Lindenmayer system. Then that channel migrates using sequential Gaussian simulation or multiple-point simulation. This process reproduces the complex relationships between successive channels without relying on partially validated physical models with an often constraining parameterization.

The applications of those various works to synthetic cases highlight the potentiality of the approaches. They open up interesting prospects to better take into account the connectivity when stochastically simulating channels.

Résumé

Mots-clefs chenal turbiditique connectivité composante connexe simulation stochastique système de Lindenmayer migration de chenaux

Les chenaux sont des structures sédimentaires clefs dans le transport et le dépôt de sédiments depuis les continents jusqu'aux planchers océaniques. Leurs dépôts perméables permettent la circulation et le stockage de fluides. Comme illustré avec les systèmes turbiditiques, le remplissage de ces chenaux est très hétérogène. Son impact sur la connectivité des dépôts perméables est amplifié par les variations d'organisation spatiale des chenaux. Mais du fait de l'aspect lacunaire des données, l'architecture de ces structures souterraines n'est que partiellement connue. Dans ce cas, les simulations stochastiques permettent d'estimer les ressources et les incertitudes associées. De nombreuses méthodes ont été développées pour reproduire ces environnements. Elles soulèvent deux questions capitales : comment analyser et comparer la connectivité de simulations stochastiques ? Comment améliorer la représentation de la connectivité dans les simulations stochastiques de chenaux et réduire les incertitudes ?

La première question nous a conduits à développer une méthode pour comparer objectivement des réalisations en se concentrant sur la connectivité. L'approche proposée s'appuie sur les composantes connexes des simulations, sur lesquelles sont calculés plusieurs indicateurs. Une représentation par positionnement multidimensionnel (MDS) facilite la comparaison des réalisations. Les observations faites grâce au MDS sont ensuite validées par une carte de chaleur et les indicateurs. L'application à un cas synthétique de complexes chenaux/levées montre les différences de connectivité entre des méthodes et des valeurs de paramètres différentes. En particulier, certaines méthodes sont loin de reproduire des objets avec une forme de chenaux.

La seconde question amène deux principaux problèmes. Premièrement, il apparaît difficile de conditionner des objets très allongés, comme des chenaux, à des données de puits ou dérivées de sismiques. Nous nous appuyons sur

une grammaire formelle, le système de Lindenmayer, pour simuler stochastiquement des objets chenaux conditionnés. Des règles de croissance prédéfinies contrôlent la morphologie du chenal, de rectiligne à sinueuse. Cette morphologie conditionne les données au fur et à mesure de son développement grâce à des contraintes attractives ou répulsives. Ces contraintes assurent le conditionnement tout en préservant au mieux la morphologie. Deuxièmement, l'organisation spatiale des chenaux apparaît peu contrôlable. Nous proposons de traiter ce problème en intégrant les processus qui déterminent l'organisation des chenaux. Un premier chenal est simulé avec un système de Lindenmayer. Puis ce chenal migre à l'aide d'une simulation gaussienne séquentielle ou d'une simulation multipoints. Cette approche reproduit les relations complexes entre des chenaux successifs sans s'appuyer sur des modèles physiques partiellement validés et au paramétrage souvent contraignant.

L'application de ces travaux à des cas synthétiques démontre le potentiel de ces approches. Elles ouvrent des perspectives intéressantes pour mieux prendre en compte la connectivité dans les simulations stochastiques de chenaux.

Contents

List of Figures

List of Tables

Preface

Remerciements

Commençons par une petite citation attribuée à M. Winston Churchill : « Ce n'est que quand il fait nuit que les étoiles brillent. » Peut-être pas la meilleure, mais elle prend un sens tout particulier au moment d'écrire ces lignes. Parce que je suis loin d'avoir été le seul acteur de mon parcours jusqu'à cette thèse. Parce que j'ai eu la chance d'avoir été accompagné par une ribambelle de gens sans qui rien n'aurait été possible. Trop de gens pour pouvoir tous les remercier nommément, même si je n'en pense pas moins. Ainsi ce qui va suivre tiens plus de l'énumération focalisée sur mes années de thèse et n'évoquera pas l'ensemble des possibles, malgré quelques débordements.

Cela va tellement de soi et pourtant... Un incommensurable merci à Pauline pour m'avoir fait confiance durant toutes ces années, comprenant mes années de thèse mais également mes années de master. Suite logique, un immense merci à Philippe pour sa confiance, mais aussi pour tous les agréables moments et discussions associées à Neuchâtel. Un sincère merci à Judith, qui a toujours répondu présente quand j'en avais besoin. Enfin un ample merci à Julien et Sébastien pour des discussions toujours pertinentes durant ces trois ans (et demi).

Un auguste merci à Guillaume pour m'avoir laissé l'opportunité de poursuivre en thèse, ainsi qu'à Fatima puis Sophie pour m'avoir libéré des tâches administratives avec une efficacité sans égale. Merci également à Pierre et Christophe pour avoir fait de même avec les aspects informatiques.

Un colossal merci à tous les thésards dont j'ai croisé la route, aussi bien à Nancy qu'à Neuchâtel. Vous êtes tous pour beaucoup dans le merveilleux souvenir que je conserverai de cette expérience. Une petite mention particulière aux divers membres du bureau F-221 pour leur résistance contre vents et marées face à la disparition du respect. Je ne me lancerai pas plus avant dans

une énumération qui serait déjà bien fournie avec les seuls thésards, et qui s'élargirait plus encore en ajoutant les nombreuses autres personnes qui m'ont accompagné de plus ou moins loin dans mon périple.

J'ai bien sûr une pensée pour ma famille, en commençant par la famille proche – en particulier mes sœurs – puis en poursuivant avec le sens le plus large que le mot « famille » puisse prendre : oncles, tantes, cousins, cousines, nourrice, amis et bien d'autres. J'extrais plus particulièrement de cet agrégat mon colloc', avec qui cette aventure fut un monumental plaisir. Malheureusement, toutes les bonnes choses ont une fin.

Naturellement mes parents n'étaient pas directement mentionnés au paragraphe précédent pour leur réserver le mot de la fin : je vous aime.

Acknowledgements

This work was performed in the frame of the RING project at the Université de Lorraine in co-supervision with the CHYN at the Université de Neuchâtel. I would like to thank the industrial and academic sponsors of the Gocad Research Consortium managed by ASGA for their support and Paradigm for providing the SKUA-GOCAD software and API. I also thank the thesis committee members for assessing this work and for their helpful remarks.

Résumé étendu

Les fluides et leur circulation sont des éléments fondamentaux en géologie. D'une part ils sont liés à la formation et l'évolution de nombreuses structures géologiques. D'autre part ces fluides peuvent être eux-mêmes des ressources exploitables ou aboutir à la formation de ressources exploitables. De ce point de vue les systèmes chenalisés fluviatiles et turbiditiques sont des structures majeures. Parmi les dépôts liés à ces systèmes, certains ont une perméabilité élevée autorisant la circulation et le stockage de fluides, comme de l'eau, du pétrole et du gaz [p. ex. Fitterman et al., 1991, Mayall et al., 2006, Braathen et al., 2012]. Un élément architectural fondamental de ces systèmes est le chenal.

Dans le cas des systèmes turbiditiques, les chenaux assurent le transport des sédiments depuis la marge continentale jusqu'au bassin océanique. Mais ces chenaux sont eux-mêmes le siège de dépôts de sédiments, dont des dépôts perméables à fort potentiel réservoir. Ceci explique le fort intérêt qui leur est porté. Le remplissage de ces chenaux n'est pas homogène [p. ex. Labourdette, 2007, McHargue et al., 2011]. Ils contiennent également des dépôts plus argileux qui peuvent bloquer la circulation de fluides entre deux chenaux, donc compartimenter le réservoir et compliquer son exploitation. Cet aspect est encore amplifié par sa combinaison à l'organisation spatiale des chenaux [p. ex. Labourdette, 2007, Mayall et al., 2006]. En effet les chenaux résultent d'une longue évolution composée de migrations continues ou abruptes et d'avulsions [p. ex. Abreu et al., 2003, Kolla, 2007]. L'architecture qui en résulte est donc particulièrement complexe, avec un impact significatif sur la connectivité des dépôts perméables et la compartimentation du réservoir (figure 1).

Ces structures étant enfouies au moment de leur exploitation potentielle, leur architecture ne peut être connue complètement par des observations directes. Différents types de données – données de puits ou données sismiques notamment – donnent une image extrêmement lacunaire de cette architecture et de

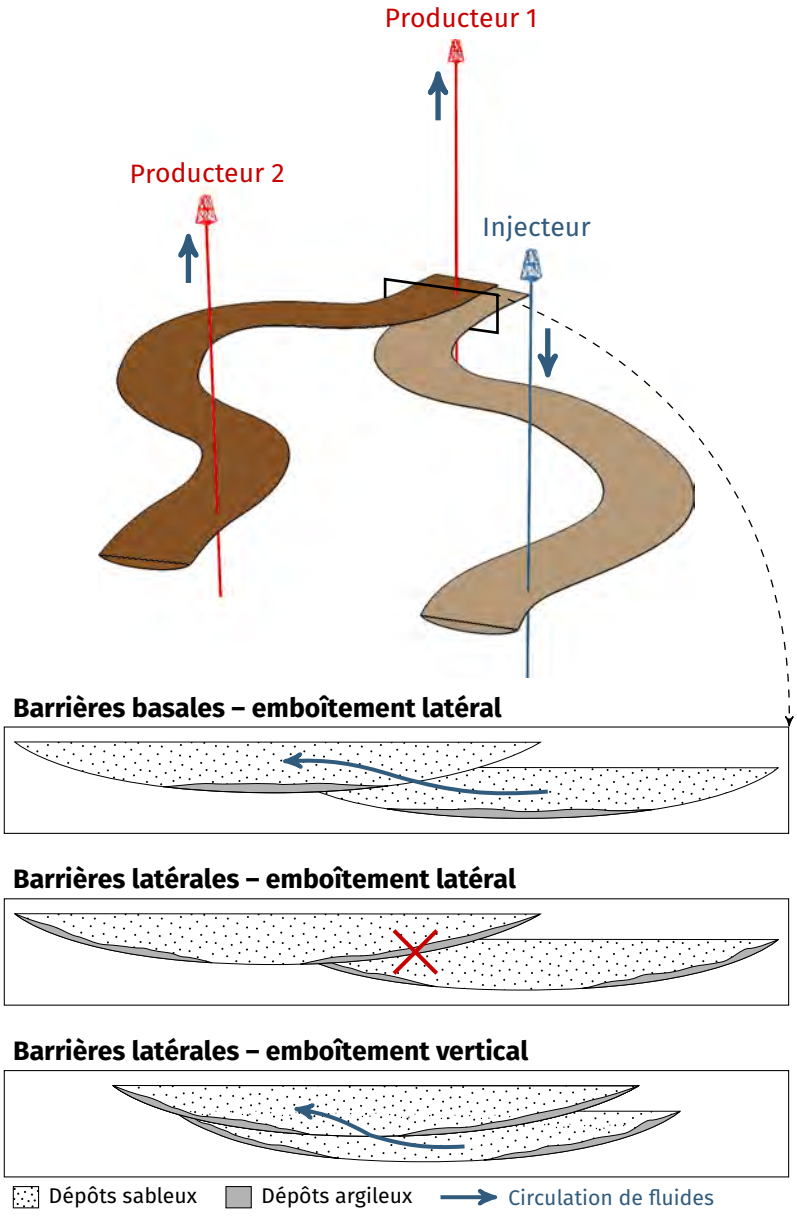


FIGURE 1 Effet du remplissage et de l'organisation spatiale de chenaux, plus ou moins emboîtés les uns dans les autres, sur la circulation des fluides. Lors de l'implantation d'un puits producteur, le problème est d'estimer si celui-ci sera bien connecté au puits injecteur. Ici suivant le type de barrières argileuses et l'emboîtement des chenaux, le puits producteur 2 peut ne pas être connecté à l'injecteur. Le producteur 1 est certain d'être connecté à l'injecteur, mais son rendement concernant la récupération de pétrole risque d'être moins intéressant.

la connectivité associée. De ce point de vue, les simulations stochastiques permettent de reconstituer un ensemble d'architectures probables en prenant en compte les données disponibles. Elles donnent ainsi accès à l'incertitude autour des modèles simulés. Plusieurs méthodes se sont attaquées à la simulation de chenaux [p. ex. Viseur, 2001, Strebelle, 2002, Pyrcz et al., 2009]. Ces méthodes se sont retrouvées confrontées à un choix cornélien :

– les méthodes qui conditionnent facilement aux données, basées sur le remplissage des cellules d'une grille, ont des difficultés à reproduire des corps sédimentaires avec une forme de chenaux ;

– les méthodes qui reproduisent facilement des chenaux, basées sur la définition d'une forme géométrique de chenal, ont des difficultés à conditionner aux données.

Ceci implique que souvent la reproduction des corps sédimentaires et la connectivité sous-jacente sont sacrifiées au profit du conditionnement. Deux interrogations ressortent de ces observations :

– comment analyser la connectivité de simulations stochastiques afin de comparer ces simulations ?

– comment améliorer la reproduction de la connectivité dans ces simulations ?

Analyse de la connectivité de simulations stochastiques

Comme de nombreuses simulations se basent sur la reproduction de statistiques d'un modèle a priori, la comparaison des réalisations se fait souvent d'un point de vue statistique [p. ex. Boisvert et al., 2010, Tan et al., 2014]. Or cette approche néglige la connectivité, qui est pourtant un aspect clef dans les utilisations des simulations stochastiques. D'autres méthodes se focalisent sur les composantes connexes des réalisations [p. ex. Deutsch, 1998, De Iaco, 2013]. Une composante connexe est ici un ensemble de cellules, de telle façon que pour deux cellules il existe un chemin de cellules voisines de même faciès reliant l'une à l'autre. Sur chaque composante connexe constituant une réalisation, il est possible de mesurer divers indicateurs, comme leur volume, puis de comparer ces indicateurs entre les réalisations. Cela permet d'avoir une idée de la connectivité des corps sédimentaires.

Nous proposons ici un ensemble d'indicateurs basés sur ces composantes connexes pour aller plus loin dans l'analyse de la connectivité (figure 2). Ces

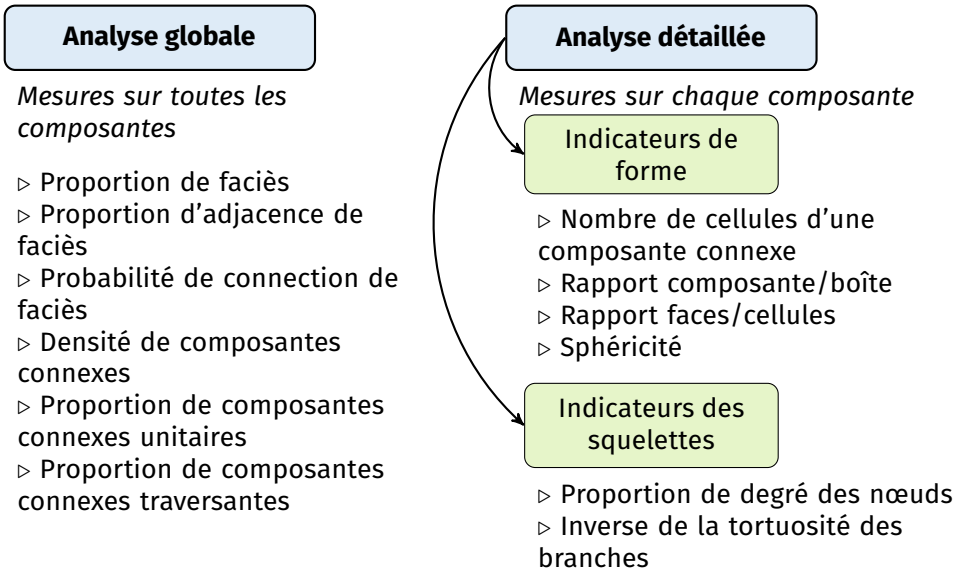


FIGURE 2 Indicateurs utilisés pour la comparaison de réalisations.

indicateurs sont ensuite utilisés pour comparer des réalisations. Cette comparaison est simplifiée par le calcul d'une dissemblance entre les réalisations à l'aide des indicateurs. Cette dissemblance se base sur une combinaison de distance euclidienne et de divergence de Jensen-Shannon. L'analyse des dissemblances s'avère longue et fastidieuse face à plusieurs centaines de réalisations, ce qui est commun en simulation stochastique. Le positionnement multidimensionnel (*MultiDimensional Scaling* ou MDS) [pour une revue de la méthode voir Cox and Cox, 1994] représente chaque réalisation par un point, avec les distances entre les points les plus proches possible des dissemblances entre les réalisations (figure 3). Le principal avantage de cette méthode est de pouvoir placer les points dans un espace de faible dimension, idéalement deux dimensions pour faciliter l'analyse. Comme la représentation des dissemblances est biaisée par le faible nombre de dimensions, il est indispensable de retourner aux valeurs de dissemblances, sous la forme d'une carte de chaleur, puis aux indicateurs pour valider et parfaire l'analyse.

L'application à un cas synthétique montre que le classement des réalisations diffère de ce qui est attendu suite à une analyse visuelle. L'étude plus précise des indicateurs révèle néanmoins la cohérence des indicateurs, qui s'avèrent rapide à mettre en œuvre et simples à analyser.

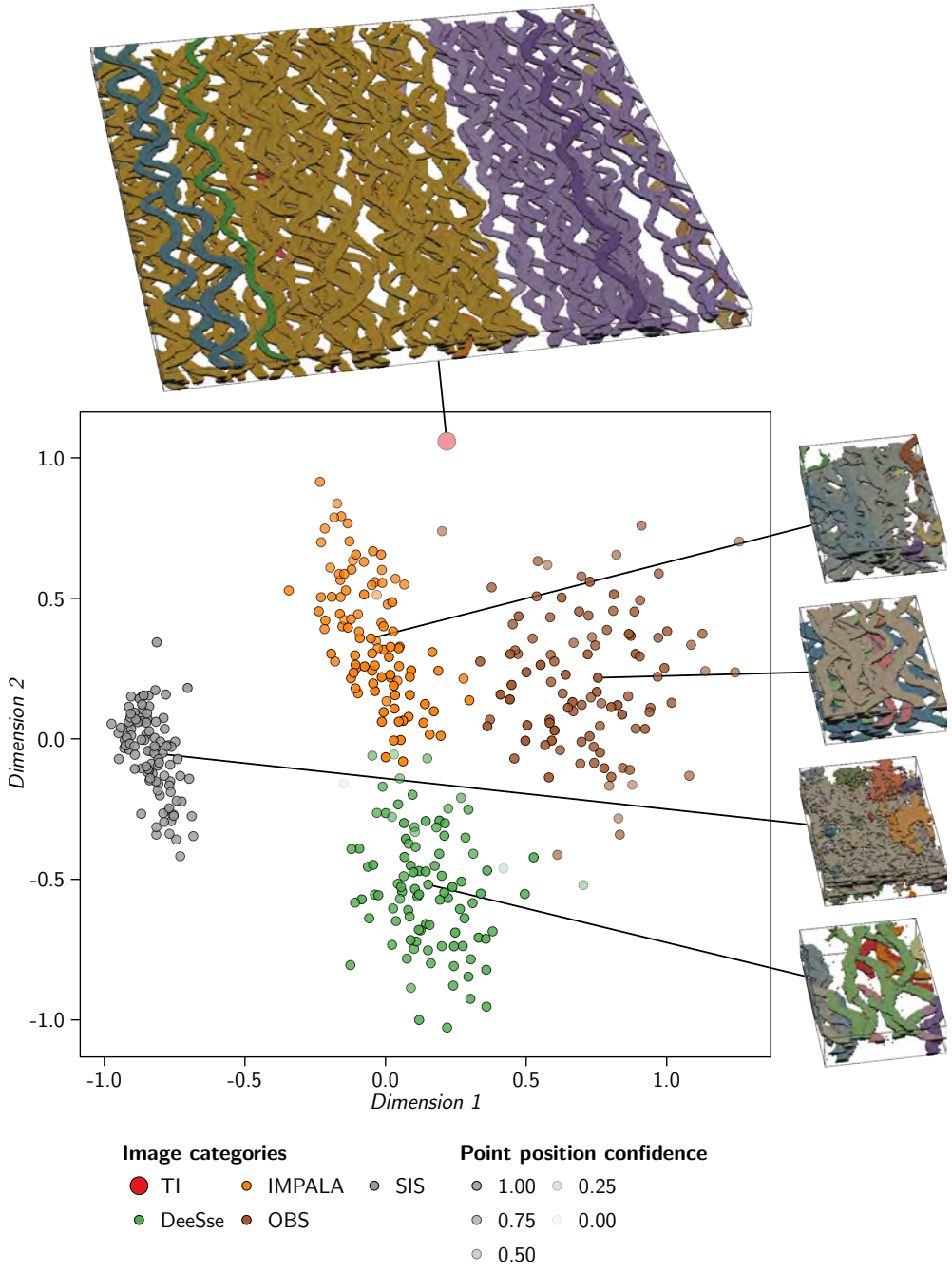


FIGURE 3 Exemple d'application du MDS à un ensemble de réalisations provenant de différentes méthodes (d'après Rongier et al. [2015]). Les réalisations de chacune des méthodes forment des groupes clairement distincts les uns des autres qui sont reliés à des différences de connectivité.

Préservation de la morphologie des chenaux dans les simulations stochastiques conditionnées

Un des éléments ayant un impact significatif sur la connectivité et les simulations d'écoulements est la préservation de la continuité des chenaux. Mais concilier cet élément avec l'ensemble des données disponibles est une tâche délicate. La méthode proposée ici se base sur une grammaire formelle, le système de Lindenmayer ou L-système [Lindenmayer, 1968]. Les L-systèmes sont particulièrement employés en biologie afin de simuler des plantes et des arbres. La simulation d'arbres implique notamment la prise en compte d'informations environnementales pour intégrer par exemple l'effet de la gravité ou la compétition pour la lumière.

Cette approche a été adaptée pour simuler stochastiquement des chenaux conditionnés à des données de puits ou dérivées d'images sismiques. Des règles prédéfinies gèrent la simulation de la morphologie des chenaux, permettant de simuler aussi bien des chenaux rectilignes que sinueux. Au cours de son développement, cette morphologie est influencée par des contraintes attractives ou répulsives. Ces contraintes permettent de prendre en compte divers types de données (figure 4) : interprétation de corps sédimentaires au niveau des puits, connectivité entre les puits, cube de proportion de faciès ou confinement des chenaux.

L'application à plusieurs cas synthétiques intégrant plus ou moins de données illustre la capacité de la méthode à prendre en compte de nombreuses données. Les contraintes induisent une déformation de la morphologie du chenal, même si tout est fait pour garder cette déformation la plus faible possible. En contrepartie, elles permettent d'assurer le conditionnement aux données et la continuité du chenal est toujours préservée. Or cette continuité a plus d'impact sur les écoulements que la sinuosité des chenaux.

Prise en compte des relations spatiales entre chenaux dans les simulations stochastiques

Si les L-systèmes permettent de prendre en compte de nombreuses données, reproduire précisément l'organisation des chenaux reste compliqué, voir impossible. Cette organisation découle des processus d'évolution des chenaux et contrôle l'emboîtement de ces chenaux, donc la connectivité. Dans un système turbiditique, un chenal est dévié au fur et à mesure de son développement de manière plus ou moins abrupte par migration et avulsion. Peu de méthodes s'at-

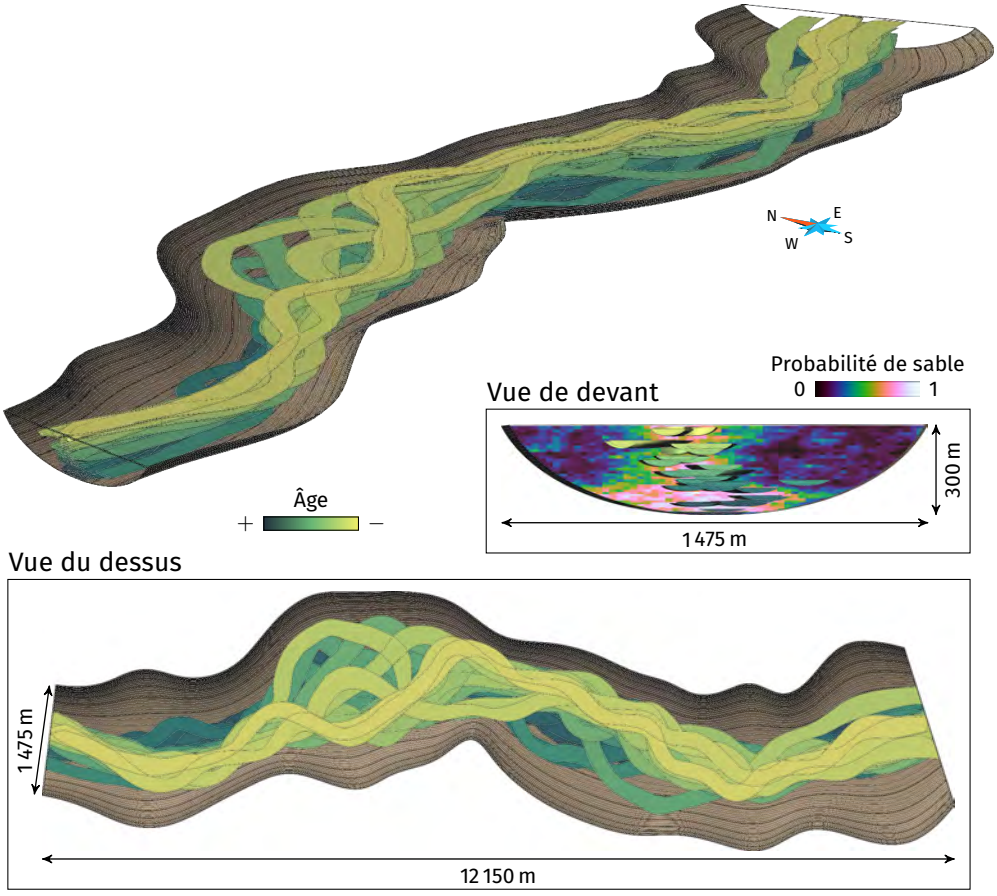


FIGURE 4 Exemple de réalisation de chenaux turbiditiques dans un chenal principal. Les chenaux sont conditionnés à un cube de proportion de sable et par le chenal principal, dont ils ne doivent pas dépasser les bords.

tachent à reproduire précisément l'organisation des chenaux. Les plus connues se basent sur des modèles physiques de migration [Lopez, 2003, Pyrcz et al., 2009]. Ces modèles ont comme inconvénients un paramétrage souvent contraignant et une validité loin d'être établie [Camporeale et al., 2007]. Peu de méthodes essayent d'être plus descriptives pour passer outre les limitations des modèles physiques [p. ex. Teles et al., 1998, Viseur, 2001, Labourdette, 2008].

La méthode proposée ici est similaire dans son principe aux méthodes basées sur des modèles physiques. Elle part d'un chenal initial, soit simulé avec un L-système, soit interprété à partir de données sismiques. Puis la migration de

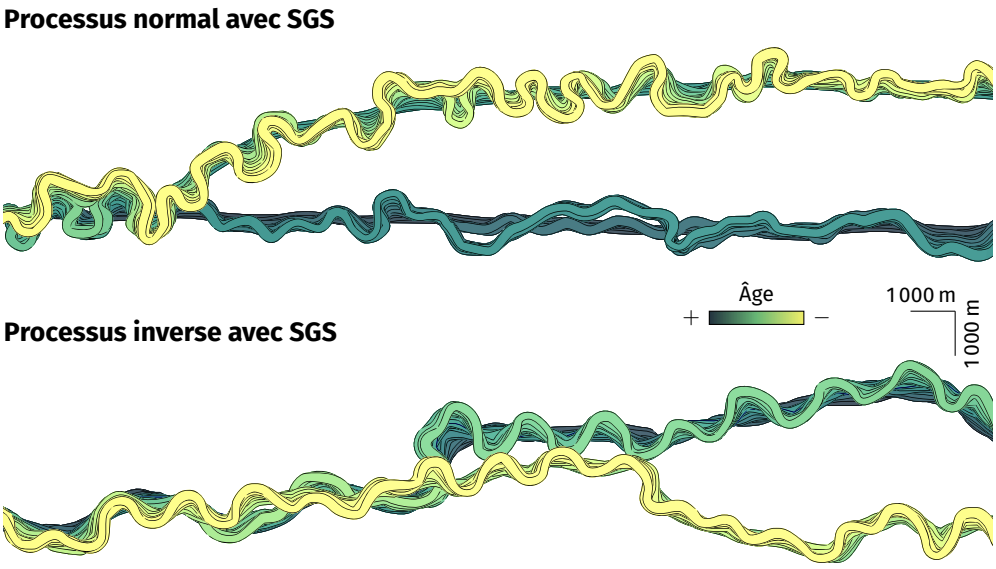


FIGURE 5 Exemples de réalisations avec une migration basée sur des simulations gaussiennes séquentielles. Le processus normal implique de commencer la simulation de la migration par le plus jeune des chenaux, c-à-d celui en bas de séquence. Le processus inverse implique de commencer la simulation de la migration par le plus vieux des chenaux, c-à-d celui en haut de séquence. Cette dernière se justifie par le fait que ce chenal en haut de séquence est souvent observable sur les données sismiques.

ce chenal est simulée soit par simulation gaussienne séquentielle (*Sequential Gaussian Simulation* ou SGS) soit par simulation multipoints (*Multiple-Point Simulation* ou MPS) si un modèle d'entrainement est disponible. L'avulsion est gérée aléatoirement : un emplacement est tiré aléatoirement le long du chenal et un nouveau chenal est simulé par L-système à partir de ce point.

L'ensemble permet de simuler les motifs de migrations observés de manière récurrente dans les systèmes turbiditiques anciens comme modernes. De plus, cette méthode permet de simuler la migration soit à partir du premier chenal d'une séquence, selon un processus normal, soit à partir du dernier, selon un processus inverse (figure 5). Dans le processus inverse les chenaux rétro-migrent depuis le dernier jusqu'au premier. Cette méthode est particulièrement pertinente car sur des données sismiques le dernier chenal d'un complexe peut souvent être interprété. Cela permet ainsi d'initier le processus à partir des données.

Conclusion

L'ensemble de ces nouvelles approches permettent de mieux appréhender la connectivité dans les simulations stochastiques. L'analyse de la connectivité est un élément primordial afin d'étudier sa reproduction et les impacts potentiels en ce qui concerne l'estimation des incertitudes et la simulation des écoulements. Mieux représenter les corps sédimentaires liés aux dépôts perméables permet une meilleure estimation de la connectivité. De ce point de vue, être capable de représenter des chenaux conditionnés aux données, comme le font les L-systèmes, puis de reproduire précisément leur organisation, comme le fait la migration, est primordial. Le conditionnement aux données de la migration doit encore être exploré afin de parfaire le processus de simulation.

Introduction

1 Motivations

Fluids and their circulation play a fundamental role in geological structure formation, evolution and alteration. This begins at the surface, from mountains to the sea-floor, with the erosion and deposition of sediments due to fluid circulations. These fluids often circulate within channels: the resulting sedimentary deposits are so concentrated in space following given patterns. This affects many different and essential resources. Concentrations of organic matter within the deposits can latter produce coal, oil and/or gas [e.g., McCabe et al., 1993]. The channels themselves are often filled for a large part with sand. The preserved spaces between the sand particles enable underground fluid circulations. These circulations may lead to the deposition of metals such as gold or uranium with a significant economic interest [e.g., Buck, 1983, Youngson and Craw, 1999]. Moreover, the fluids contained within the rocks are also interesting for themselves. Channels form reservoirs when they store oil and/or gas after their migration [e.g., Martin, 1993, Mayall et al., 2006, Pranter and Sommer, 2011]. They can also constitute aquifers, containing potential drinkable water resources or geothermal energy sources [e.g., Fitterman et al., 1991, Jussel et al., 1994, García-Gil et al., 2015]. And fluid storage within those rocks is more and more investigated for economic or climatic concerns [e.g., Evans and Chadwick, 2009, Braathen et al., 2012].

In those prospects, the remaining spaces within the rock, i.e., the rock porosity (figure 6), are fundamental. The connections between these spaces, i.e., their connectivity, determine the fluid ability to circulate within the domain of interest. The rock permeability measures this ability. The permeability is then a major property to understand flow circulations, so a major element required for flow simulations. Estimating the permeability calls for an estimation of the connected porosity. And estimating the porosity calls for a representation of

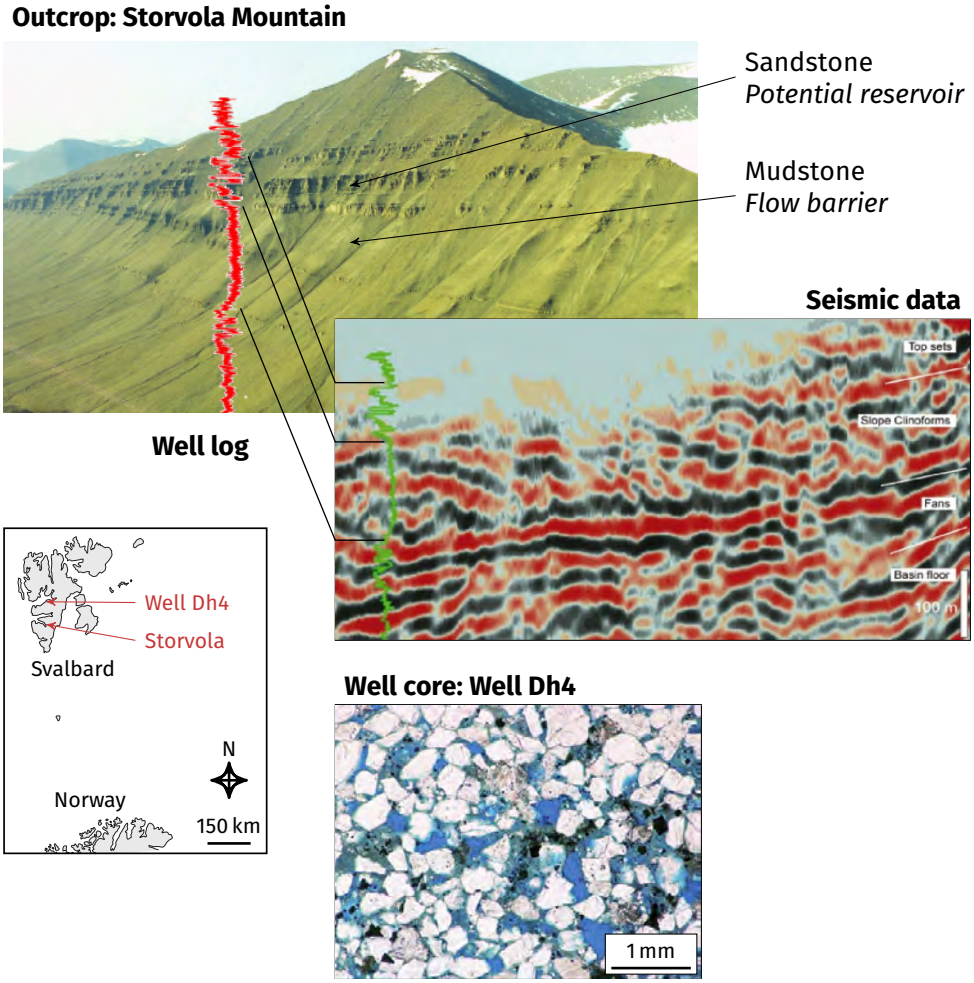


Figure 6 Example of an outcrop analog of an oil field: the Storvola Mountain, Svalbard, Norway (modified from Johannessen and Steel [2005], Johannessen et al. [2011], Grundvåg et al. [2014]). The Storvola mountain is about 7 km long and 900 m high. The well log is a gamma ray log coming from the Sysselmannbreen well, drilled north-west of the Storvola mountain. It is 450 m long and is approximately placed along the mountain stratigraphy [Luthi and Flint, 2014]. The seismic comes from a 2D seismic line passing through the Sysselmannbreen well. The microscopic image shows a sandstone sampled from the well Dh4 located farthest north in Svalbard (modified from Braathen et al. [2012]). The white grains are quartz and the blue areas are the rock pores. These pores store fluids.

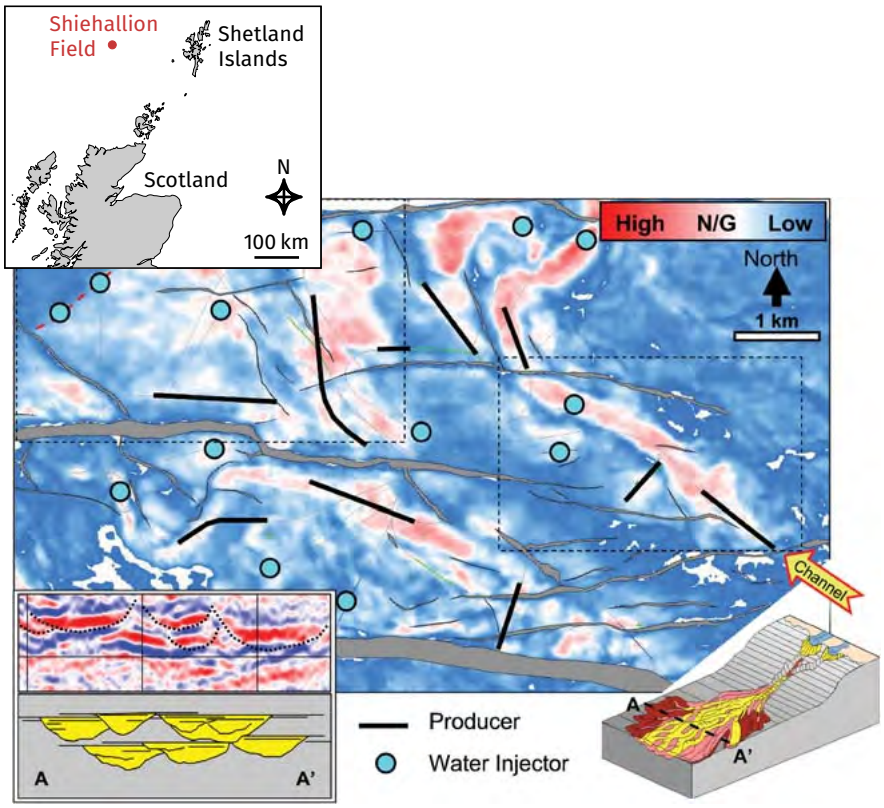


Figure 7 Schiehallion field main sand reservoir (modified from Gainski et al. [2010]). High net-to-gross (N/G) areas corresponds to submarine channels whose depositional environment is illustrated in the bottom right figure. The grey lines indicate faults.

the rock type within the domain of interest. As stated previously, sandy facies (figure 6) often have a significant connected porosity and form interesting reservoirs. On the other hand shales (figure 6) have a very low permeability: fluids can not circulate in such facies, which can act as flow barriers and may compartmentalize a reservoir.

The Shiehallion field is an interesting case to illustrate those concepts of reservoir and barrier features. It is an oil field located north of Scotland (figure 7), in the North Atlantic Ocean. Its reservoir compartments are composed of sandy rocks deposited by submarine channels with a south to north trend [Gainski et al., 2010]. Figure 7 shows the net-to-gross ratio of the main reservoir part. The net-to-gross ratio represents the ratio between the volume of actual

reservoir facies over the total volume of rock within the field. A high ratio, the red areas on the map, means a significant sand content, a low ratio, the blue areas on the map, a low sand content. Thus, the red areas roughly represent the channels and the related reservoir compartments. Those compartments have a very elongated shape characteristic of the channels.

The field production started in 1998. But soon the reservoir compartments appeared less inter-connected than initially estimated. This came from a wrong estimation of the potential flow barriers. A first and significant type of barriers are large faults that cut the field in a west-east trend. A fault results from a failure within the rock. However, such failure does not necessarily increase the porosity: some faults are sealed and act as flow barriers. This happens to be the case for most of the faults within the field. And their orientation almost perpendicular to the channels reduces even more the connectivity. Another type of barriers was observed: shales at the channel interfaces inducing a loss of connectivity between channels. These sedimentary barriers were to small to appear on the available data, making them difficult to infer.

At the end, the number of wells to maximize the oil recovery had to be more than doubled compared to the development plan. All this led to direct economic consequences, for instance an increase of the exploitation costs or a possible inability to maximize the recovery in some reservoir areas. But more generally it induces a miss-management of the resource. If some strategies exist to manage those cases at best [e.g., Gainski et al., 2010], the Shiehallion case illustrates why better assessing the connectivity from the start is essential. Here representing the connectivity calls for representations of reservoir and barrier features. Faults act as barriers. Channels are not just filled with sandy material and can contain flow barriers. The filling defines the intra-connectivity of the channels but also the inter-connectivity between different channels. Delimiting the channels is a first step to help estimate their filling.

This delimitation transfers the notion of connectivity from the voids within the channel filling to the channels themselves. The channel connectivity refers to two aspects:

- Representing the connectivity along a channel. This requires in particular to be able to maintain the continuity of the channelized body.

- Representing the connectivity between the channels. Channels are not just individual objects but result from a temporal and spatial evolution related to flow processes. Thus, their spatial arrangement reflects this history.

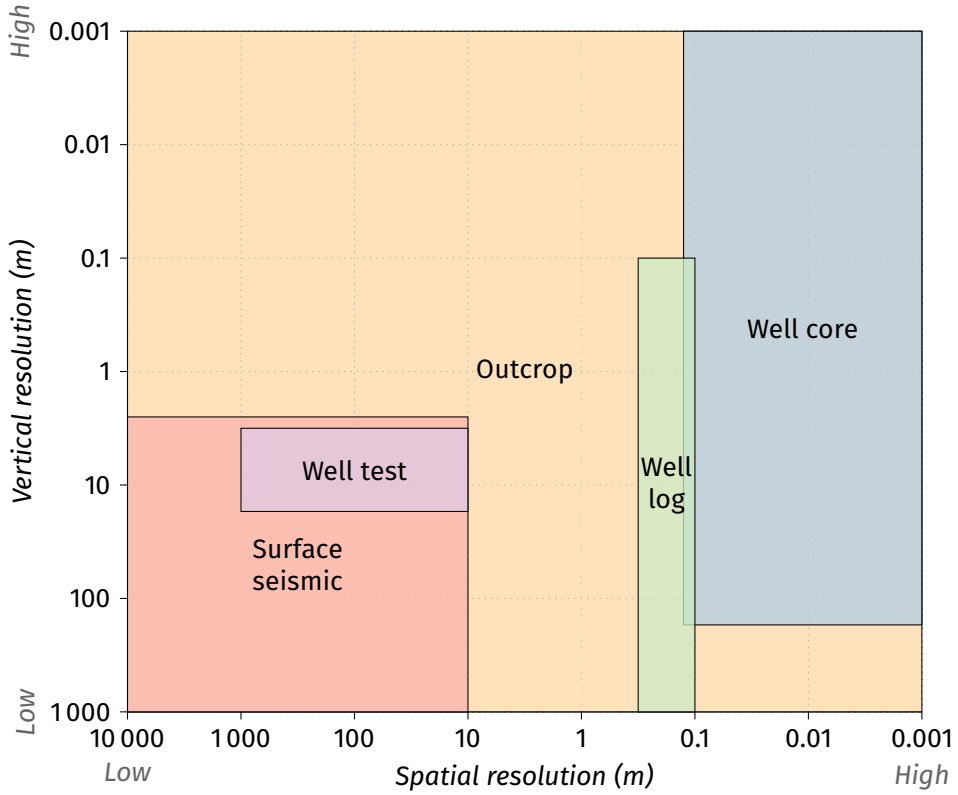


Figure 8 Examples of data used to characterize a reservoir (modified from Hanna et al. [2015]). The outcrop is a surface analog of a reservoir and helps to better understand potential reservoir characteristics. Figure 6 gives an illustration of these elements.

Those aspects are fundamental to represent the reservoir parts and barriers within a channel and so to reproduce at best the permeability of the reservoir.

However, our view of the geological structures in the underground is always partial. The Shiehallion field illustrates also this aspect: the net-to-gross was deduced from seismic data. A seismic image is similar to an ultrasound of the underground. This gives a view of the whole field, but only at a low resolution (figures 6 and 8). Indeed no precise channels are visible on figure 7, and some flow barriers were missed because their size is under the seismic resolution. In the best cases, it remains possible to interpret sedimentary bodies on a seismic if the resolution is high enough. Well data give a preciser view of the rock prop-

erties but only at a local scale (figure 8). Tracer tests between wells and well tests can help to identify larger scale information on the structure connectivity. In the Shiehallion field such tracer tests helped to determine more precisely the reservoir compartmentalization. For a geological study, all those data provide a precise view of the underground at only some locations. Aside from those locations, they give at best a glimpse of the geological structures. Being able to cross those views requires building a three-dimensional representation of the rock properties.

These representations, called geomodels, integrate as much information as possible from the well and/or seismic data and look at completing the remaining blanks. The usual workflow (figure 9) consists in modeling the horizons delimiting the geological layers and the faults. The resulting model is called a structural model [Caumon et al., 2009]. This structural model serves as a basis to build a geological grid. The geological grid is populated with properties representing the lithology, then the porosity and finally the permeability. For computational reasons, this grid is usually upscaled to a flow simulation grid. Flow simulations then help to better estimate the flow behavior within the domain of interest.

When building a geomodel, there is a need to complete some blanks due to:

- Some lacks in the data, as illustrated previously. These lacks come from the data resolution and the spatial coverage, and also from potential measurement errors.

- The lack of knowledge about the geological processes that have shaped the domain of interest. This embraces two questions: what are the involved processes? And how do they work?

It is then impossible to completely describe the geology of a given domain of interest. Thus, our vision of that domain is uncertain. A solution is to introduce some interpretation linking the data to geological concepts coming from the regional context or analog cases. This interpretation is often human-driven and subjective, and leads to a single deterministic model. Depending on the interpreter and its experience, the resulting models can be radically different [Bond et al., 2007]. This leads to the necessity of building several geomodels for a given case. However, it requires a significant amount of work. To overcome this, the idea is to simulate several equiprobable models considering all the available data. Those equiprobable models come from some randomness added to the workflow. Several steps, such as the horizon and fault modeling

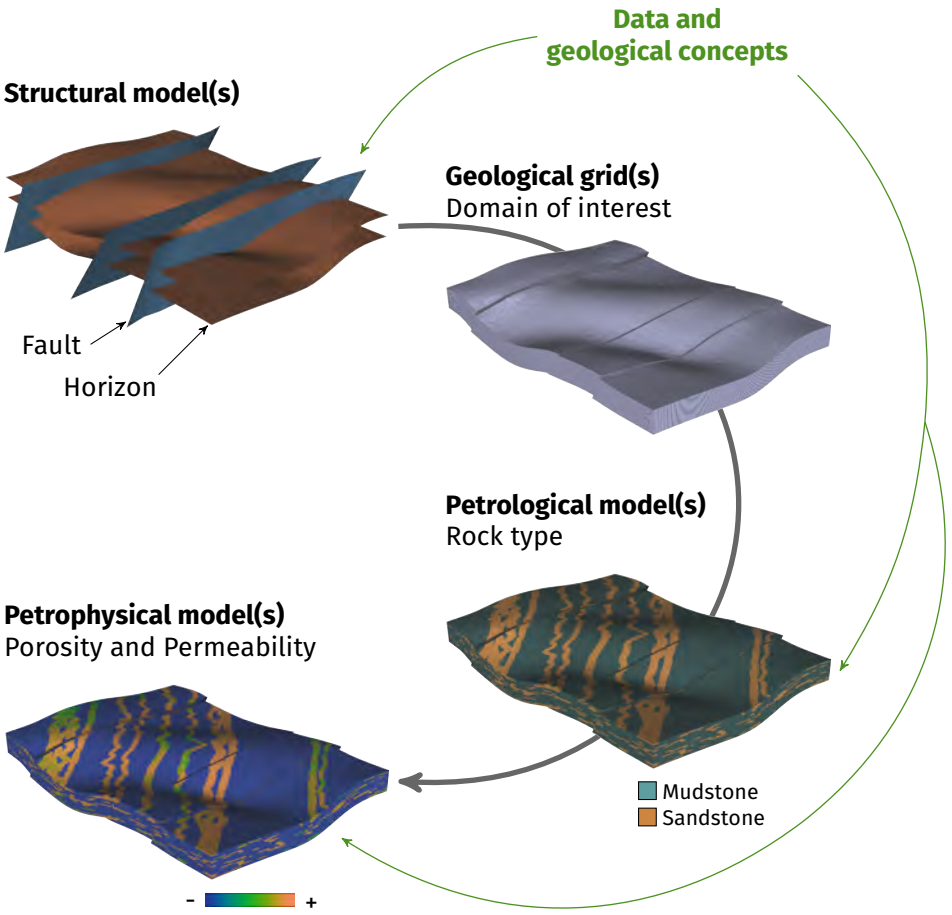


Figure 9 Usual workflow to build a geological model for flow simulation.

or the rock property simulations, are then stochastic. It gives access to a potentially infinite number of models for a given set of data. And depending on the differences between the models, it gives an idea of the uncertainties attached to the modeling process.

Studying these uncertainties is a prime concern when coming to the petrophysical modeling. Indeed, the related models integrate the smaller-scale geological bodies on which the lack of data is the most striking: well data do not inform much on the body extension and seismic data resolution is often too small to completely identify these bodies. Stochastic models are then es-

sential. Simulating petrophysical models begins with the simulation of the lithology. This step includes in particular the simulation of the channelized bodies, whose connectivity is fundamental. Different approaches exist to simulate such channels [e.g., Deutsch and Journel, 1992, Mariethoz and Caers, 2014, Pyrcz and Deutsch, 2014], with a struggle between data and geological concept integration:

- Some methods easily take into account the data but have trouble reproducing the channel bodies and their connectivity.
- Other methods provide far better representation of the channelized bodies but have trouble taking into account the data.

2 Objectives

This work aims at exploring the integration of the connectivity in the stochastic simulation of geological structures, focusing on channelized environments. This is done through two main themes of research.

The first theme is the analysis of the connectivity of realizations. Most of the stochastic simulation methods are based on reproducing statistics and not the connectivity. In that perspective, being able to compare realizations from the viewpoint of the connectivity is essential considering its impact on the flow. The proposed method further explores the use of connected components to analyze realization connectivity and to compare those realizations.

The second theme aims at improving the simulation of channelized sedimentary bodies in terms of connectivity. One difficulty in stochastic simulation is to preserve the significant continuity of channels while conditioning to various data. The proposal here is to rely on a formal grammar, the Lindenmayer system, to simulate the channel morphologies while conditioning to the available data. Then it serves as a basis to stochastically simulate the channel evolution. It improves the connectivity by introducing more concepts derived from geological observations in the simulation process.

3 Outline

The thesis is divided in four chapters developing the two previous scopes. If the connectivity is a unifying concern throughout the manuscript, each chapter can be read independently from the other.

Chapter 1 further develops the notion of connectivity in channelized systems and their modeling. In that perspective, it focuses on a particular channelized environment, the turbiditic systems. After describing the system architecture and its impact on the connectivity, the processes controlling this architecture are detailed. The integration of the connectivity within the stochastic simulation of such environments is then discussed.

Chapter 2 looks into the first scope. It first exposes a set of indicators to compare realizations with a focus on the connectivity of the geological structures. These indicators rely on the connected components of realizations or conceptual models. The computation of dissimilarities with the indicators facilitates the comparison. The application on a synthetic case study highlights the solutions to make the dissimilarity analysis easier.

Chapter 3 is a first look at the second scope: it focuses on the preservation of the connectivity along a channel while assuring data conditioning. It details a method of channel simulation based on a formal grammar system, the Lindenmayer system. The method relates on rules to develop channelized morphologies. Constraints related to data conditioning influence this development by attraction or repulsion. A synthetic case illustrates the modeling possibilities offered by this method.

Chapter 4 further explores the second scope to better preserve the connectivity between channels. As developed in chapter 1, this connectivity results from the channel evolution. The proposed method stochastically simulates channel evolution with a descriptive method based on geostatistical approaches. It introduces more geological concepts within the simulations while keeping a descriptive approach. A synthetic case illustrates the impact of such method on the channel connectivity.

A last chapter summarizes the results and further develops some perspectives.

Some of the parameters used in the application cases of the previous chapters are given in the appendices A, B and C.

This work is completed by appendices describing a parallel work about karsts.

Appendix C summarizes the results of a master project supervised during the doctoral work. It applies the Lindenmayer system to the stochastic simulation of karst networks. The connectivity of the network was a main concern during the method development. It has the ability to simulate both branchwork and anastomotic patterns.

Appendix D is an article published in Geomorphology during the doctoral work. It presents the work initiated during the M.Sc. internship about the stochastic simulation of karst conduits, and thus completes the work presented

in appendix A.

4 Working context

This work was funded by the Gocad Research Consortium managed by ASGA. It was performed in the frame of the RING project in the GeoRessources laboratory at Université de Lorraine and of the CHYN at Université de Neuchâtel.

Its development was mostly integrated and put into practice through the plug-in *ConnectO* of the SKUA-GOCAD software provided by Paradigm. The software environment for statistical computing R was used for the dissimilarity analysis in chapter 3.

It leads to the following contributions:

$^{(r)}$ *peer-reviewed article,* $^{(o)}$ *oral presentation,* $^{(p)}$ *poster,* $^{(a)}$ *proceeding article, * presenter*

Chapter 2 $^{(r)}$ **G. Rongier**, P. Collon, P. Renard, J. Straubhaar, J. Sausse *Comparing connected structures in ensemble of random fields*, submitted to Advances in Water Resources

$^{(p,a)}$ **G. Rongier***, P. Collon, P. Renard, J. Straubhaar, J. Sausse (2015) *Quality Analysis of Geostatistical Simulations through their Connected Structures*, EAGE Petroleum Geostatistics 2015, Biarritz, France

$^{(p)}$ **G. Rongier***, P. Collon, P. Renard, J. Straubhaar, J. Sausse (2014) *Assessment of multiple point simulation quality focusing on connected geobodies*, GeoEnv 2014 - 10$^{\text{th}}$ conference on Geostatistics for Environmental Applications, Paris, France

$^{(p)}$ **G. Rongier***, P. Collon, P. Renard, J. Straubhaar, J. Sausse (2014) *Exploring multiple-point realization quality through connected geobodies*, AAPG 2014 Annual Convention & Exhibition, Houston, Texas, USA

$^{(o,a)}$ **G. Rongier***, P. Collon, P. Renard, J. Straubhaar, J. Sausse (2013) *Reproduction assessment of connected geobodies in multiple-point simulation*, 33$^{\text{rd}}$ Gocad Meeting, Nancy, France

Chapter 3 $^{(p,a)}$ **G. Rongier***, P. Collon, P. Renard, J. Ruiu (2015) *Channel Simulation Using L-system, Potential Fields and NURBS*, EAGE Petroleum Geostatistics 2015, Biarritz, France

$^{(o,a)}$ **G. Rongier***, P. Collon, P. Renard (2014) *Channel axis simulation based on Lindenmayer systems*, 34$^{\text{th}}$ Gocad Meeting, Nancy, France

Chapter 4 $^{(o,a)}$ **G. Rongier***, P. Collon, P. Renard (2015) *A new application of L-systems to model channel system architecture and connectivity*, 35$^{\text{th}}$ Gocad Meeting, Nancy, France

Other contributions result from parallel works:

$^{(o,a)}$ P. Anquez*, **G. Rongier**, P. Collon (2015) *Stochastic simulations of karst networks with Lindenmayer systems*, 35th Gocad Meeting, Nancy, France

$^{(p,a)}$ D. Lavandier*, **G. Rongier**, J. Ruiu, P. Collon (2015) *3D reconstruction of karstic conduit from sparse laser data*, 35th Gocad Meeting, Nancy, France

$^{(o,a)}$ M. Parquer*, J. Ruiu, G. Caumon, P. Collon, **G. Rongier** (2015) *Towards more genetics concepts and data integration in channel simulation*, 35th Gocad Meeting, Nancy, France

$^{(r)}$ **G. Rongier**, P. Collon, M. Filipponi (2014) *Simulation of 3D karst conduits with an object-distance based method integrating geological knowledge*, Geomorphology 217, 152–164

$^{(r)}$ **G. Rongier**, P. Collon (2014) *Conduits Karstiques en 3D : une méthode de simulation stochastique*, Karstologia, 63, 58–60

$^{(o,a)}$ P. Collon*, **G. Rongier**, M. Le Mesnil, L. Dewaide, V. Hallet (2014) *Modélisation géométrique des systèmes karstiques : des réseaux aux conduits*, 24^{e} Réunion des Sciences de la Terre, Pau, France

$^{(p,a)}$ M. Le Mesnil*, P. Collon, L. Dewaide, P. Mejía-Herrera, **G. Rongier**, V. Hallet (2014) *Karst conduit generation: the Furfooz area case study (Belgium)*, 34th Gocad Meeting, Nancy, France

$^{(o)}$ **G. Rongier***, P. Collon, M. Filipponi (2014) *Incorporation of geomorphological information into the stochastic simulation of karst conduits*, GeoEnv 2014 - 10th conference on Geostatistics for Environmental Applications, Paris, France

All this includes the supervision of three master students at the National School of Geology (ENSG) in Nancy: P. Anquez (2014-2015), D. Lavandier (2014-2015) and P. Mourlanette (2015-2016).

Chapter 1

Channelized systems and connectivity: a review through the example of turbiditic systems

Abstract Channelized systems develop in various environments. Among them turbiditic systems rise a significant interest due to their ability to store fluids, such as oil or gas. Analyzing the sedimentary structures within turbiditic systems reveals a complex and heterogeneous organization. The fundamental architectural element of turbiditic systems, the channel, is often filled by sand-rich deposits, potential reservoir units, often framed by mud-rich deposits that constitute potential flow barriers. The channel stacking impacts the spatial relationship between the sand-rich and the mud-rich deposits. Thus, it has a significant impact on the connectivity of the sand-rich deposits. This stacking comes from a long-term evolution of the channels through migration and avulsion. Reproducing such processes in stochastic simulations helps to get closer to the reservoir connectivity. Object-based simulation methods start to take into account migration and avulsion. Cell-based methods still have some trouble in reproducing such complex architectures. However, they better integrate data than object-based methods. Thus, a better reproduction of the complex architecture within turbiditic systems while conditioning all the available data remains an open topic. It is necessary to obtain a more accurate estimation of the reservoir connectivity.

1.1 Introduction

Channelized systems are complexes characterized by the development of channelized sedimentary bodies due to the circulation of fluids. Many environments develop channelized bodies: in fluvial and turbiditic environments, channelized systems lead to the erosion and deposition of sediments from the continents to the abyssal plain under the oceans. They form sand-rich bodies, including channels or lobes.

These sand-rich bodies have raised an increasing interest, especially because of their ability to store and let circulate fluids. Their economic impact is already undeniable as oil and/or gas reservoirs, regardless of their origin, fluvial [e.g., Tyler, 1988, Martin, 1993, Pranter and Sommer, 2011] or turbiditic [e.g., Abreu et al., 2003, Mayall et al., 2006, Snedden, 2013]. Besides, channels can constitute highly interesting aquifers [e.g., Fitterman et al., 1991, Jussel et al., 1994, Mulligan et al., 2007].

Channelized system study is also useful outside their ability to store fluids. Placers and paleoplacers represent economically significant deposits in terms of gold [e.g., Youngson and Craw, 1999], uranium [e.g., Buck, 1983], tin [e.g., Aleva, 1985], platinum-group elements [e.g., Konstantinovskii, 2005] or diamonds [e.g., Bluck et al., 2005]. In other domains, unconsolidated sediments related to channelized systems can have a huge impact on buildings [e.g., Hernández Del Pozo et al., 2005] and these systems can also be implicated in various environmental studies [e.g., Blum and Törnqvist, 2000, Moore et al., 2003].

The growing interest of the oil industry for deep-water reservoirs has further developed the concern about turbiditic systems [e.g., Mayall et al., 2006, Labourdette, 2007, Janocko et al., 2013b]. Turbiditic systems form the largest sediment accumulation on Earth [Talling et al., 2015], with related structures spreading over tens to hundreds of kilometers. These structures and the related heterogeneity have a significant impact on the reservoir connectivity. However, they are thousands of meters under the oceans, making them difficult to study. The processes responsible for their formation remain less well known than in fluvial context.

In such context, the uncertainties due to the lack of knowledge about these processes and the data scarcity make the use of stochastic simulation even more valuable. It implies to be able to simulate channels, structures with a highly elongated shape, that condition to the available data.

This chapter aims at developing the significant impact of channels on the reservoir connectivity, with the particular interest of reservoir modeling. To do

so, it relies on the example of the turbiditic environment (section 1.2). The next sections illustrate the connectivity in such systems (section 1.3) and the processes leading to that connectivity (section 1.4). The last section describes the types of methods that exist to stochastically simulate channels and the options to deal with the connectivity (section 1.5).

1.2 An example of channelized environment: the turbiditic systems

Turbiditic systems are transitional systems transporting sediments from the continental margin to the oceanic basin floor. Turbidity currents play a major role in this transport.

1.2.1 About turbidity currents

Turbiditic systems are the results of density current circulations on the sea floor, usually in water depth from 100 m to 4 000 m. A density current is a current denser than its surrounding fluids and kept in motion by gravity in relation to the difference of density. Concerning turbiditic systems, the current is subaqueous and turbid, meaning that it contains suspended sediments. The higher density of the current originates from these sediments. More strictly, turbidity currents are a type of sediment gravity flows with a Newtonian rheology and a turbulent state. The upward component of the flow turbulence mainly supports the sediments [e.g., Middleton and Hampton, 1973, Mulder and Alexander, 2001, Shanmugam, 2002]. However, this definition is not universally accepted [Mulder and Alexander, 2001, Shanmugam, 2002] and some authors use a less strict definition that includes other types of density currents [e.g., Kneller and Buckee, 2000].

Turbidity currents have large dimensions. They can be from several kilometers to more than 200 km wide [e.g., Pirmez and Imran, 2003, Talling, 2014], even if they become thinner as the distance from the current axis increases. They vertically measure from a few meters to hundreds of meters [e.g., Komar, 1969, Stow and Bowen, 1980, Pirmez and Imran, 2003]. Considering the large dimensions of the resulting deposits, turbidity currents can travel over several hundreds of meters to more than 1 000 km [e.g., Salles, 2006, Talling et al., 2007]. Their speed varies from a few kilometers per hour to more than 70 $km.h^{-1}$ [e.g., Gennesseaux et al., 1980, Piper et al., 1999, Salles, 2006]. The major currents that build a turbiditic system can happen every year to every

thousands of years [e.g., Talling, 2014] and last from a few hours to several days.

1.2.2 About turbiditic systems

Turbiditic systems are classified according to two main parameters [Reading and Richards, 1994, Richards et al., 1998]:

- The number of system feeders and their characteristics.
- The size of the sediment particles, e.g., gravel-rich, sand-rich or mud-rich systems.

The feeders are the current suppliers for the system. They divide in three categories:

- One punctual source, forming submarine fans.
- Multiple punctual sources, forming submarine ramps.
- Multiple linear sources, forming slope aprons.

Submarine fans are the most famous turbiditic systems, and the term "turbiditic system" is often used instead of submarine fan. With submarine ramps, they rely on stable feeder, such as rivers or deltas. The turbidity currents tend to come from hyperpycnal flows, which form when the density of the river flow is higher than the density of the basin water, or from failures due to sediment accumulations at the river mouth, at the prodelta or at the shelf-edge [e.g., Piper and Normark, 2001, Wynn et al., 2007]. Submarine ramps and fans tend to have an organized architecture. On the contrary, slope aprons come from more ephemeral feeders, such as a slope failure. They do not have a clear organization, which reduces their interest as reservoirs. The following sections tend to focus on submarine fans due to an abundant literature, especially on modern analogs.

Besides the feeder and the sediment type, other features influence the development of a turbiditic system [e.g., Bouma, 2004]: the tectonic activity, the sea-level variations, the climate, the basin characteristics, the Coriolis force, etc. At the end, the system architecture results from the alternation of erosion and deposition events from the shore to the oceanic basin. Turbidity currents drive these events by transporting the sediments through the system. The system often begins with a canyon, i.e., an entrenchment along the continental slope. After the slope, the canyon ends and the system is far less confined. It is then

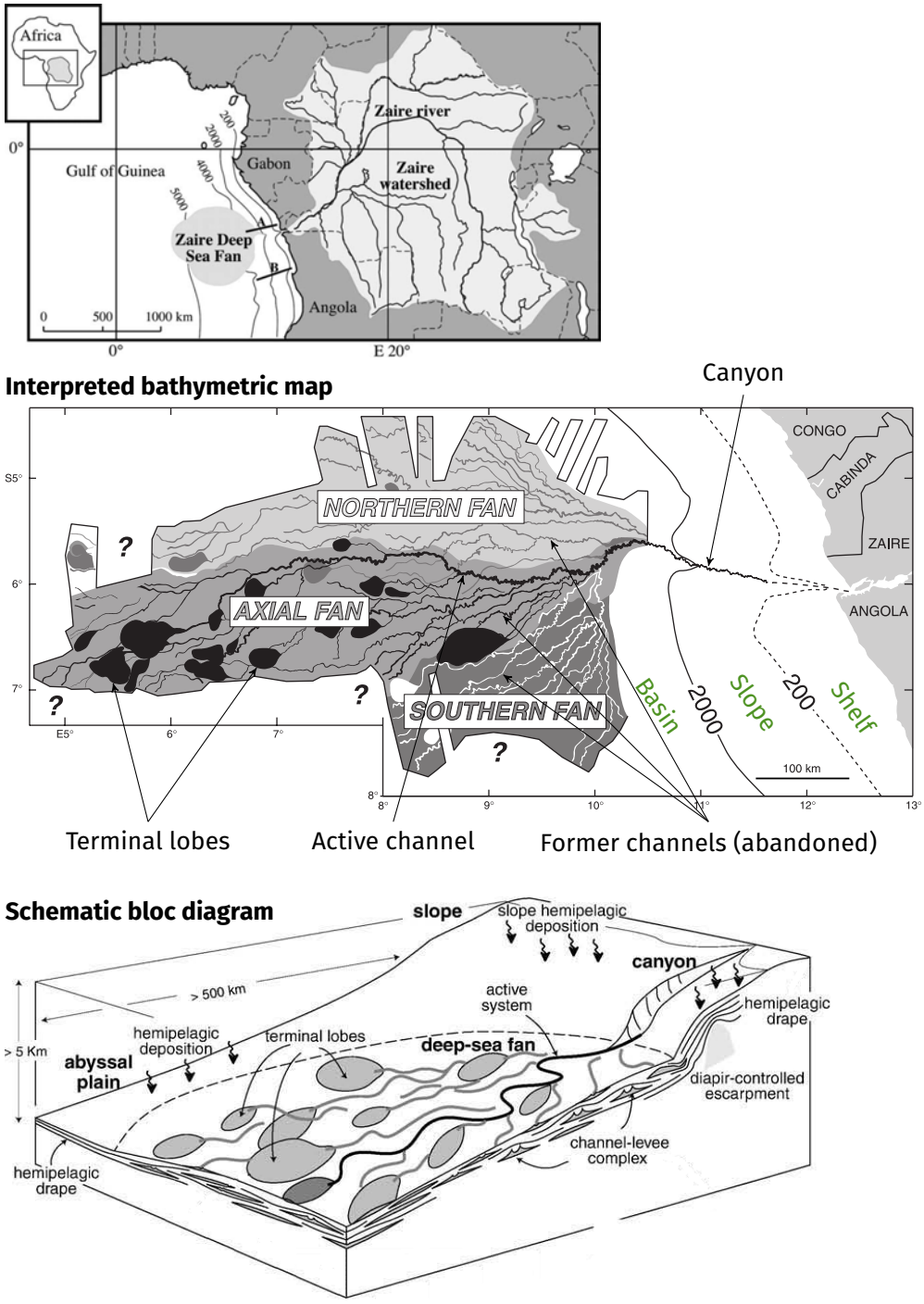


Figure 1.1 A modern submarine fan: the Congo (ex-Zaire) fan (modified from Babonneau et al. [2002], Droz et al. [2003] and Anka et al. [2009]).

free to evolve along the oceanic basin (figure 1.1). Channelized structures can develop from the canyon to the basin, but their development is more striking in the basin. Due to the lighter confinement than in the canyon, the channel path can vary a lot. Avulsions lead to the abandonment of a channel and to the formation of a new channel. It gives a fan architecture to the system. Only one channel path is active at a given time [Wynn et al., 2007]. The other observable channels are abandoned. From the slope to the basin, the channel widths tend to decrease [Flood and Damuth, 1987]. Channels often end with large ellipsoidal structures called terminal lobes.

1.2.3 Constituting elements

Turbidity currents usually start to erode the sea floor after their initiation, due to a high speed and a coarser load of sediments. In its advance, the current progressively loses its speed due to a decreasing slope and a loss of energy. Thus, it progressively moves from an erosive regime to a depositional regime. This depositional regime leads to the formation of several types of sedimentary bodies [e.g., Mutti and Normark, 1991, Normark et al., 1993].

Channel-levee systems

The main structures resulting from this erosion/deposition alternating are channels. Channels are highly elongated structures that are U or V-shaped in cross-sections. Turbidity currents initiate these structures during their erosive phase. Then the channels drive the currents and are progressively filled when the currents get into a depositional regime.

Due to the low difference of density between the turbidity current and the surrounding water, the current often overspills from the channel margins (figure 1.2). Such overspilling currents often have a mudier content than the current remaining in the channel axis. The deposition of their sediments form margin deposits called levees or overbanks.

Channels in turbiditic systems often develop in a nested structure: channel-like shapes can be identified within other channel-like shapes. Sometimes the limit between a channel and a confining structure is hard to determine [Wynn et al., 2007].

Architectural element hierarchy within channel-levee systems

Several classifications have been proposed to define relevant and recurring architectural elements related to turbiditic channels [e.g., Sprague, et al., 2002,

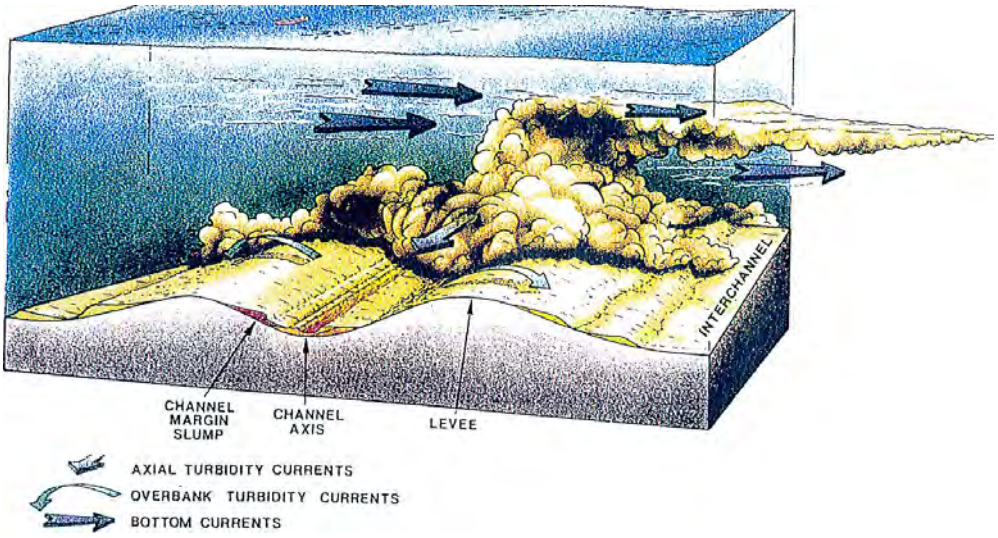


Figure 1.2 Conceptual model showing a turbidity current within a channel (from Shanmugam et al. [1993]). The channel constraints more or less the current. Thus some components of the current leave the axial current and form overbank currents.

Mayall et al., 2006, McHargue et al., 2011] (figure 1.3). The classifications differ in terms of approach and element denominations, but their characteristics are relatively similar.

The principal architectural element is the channel element of McHargue et al. [2011] or the individual channel of Mayall et al. [2006]. A channel comprises a bounding surface and all the sediments filling it. It is usually tens of meters thick and hundreds of meters wide. A channel contains several subchannel elements called stories [Sprague, et al., 2002, McHargue et al., 2011]. These elements have a channel-like shape and are composed of beds. Beds are the smallest architectural elements, with dimensions from the millimeter to the meter. They correspond to the deposits formed by a single turbiditic event [Macauley and Hubbard, 2013, Stright et al., 2014]. These deposits include divisions from the Bouma et al. [1962] and Lowe [1982] sequences. Along the channel margins, levees are formed in a depositional context. Terraces can appear in a more erosional context.

Channels usually stack to form larger-scale architectural elements, e.g., complexes and complex-sets [McHargue et al., 2011] or channel belt and channel belt complexes [Janocko et al., 2013b]. The stacking pattern and the genetic relationship between the channels define these larger-scale elements. Their

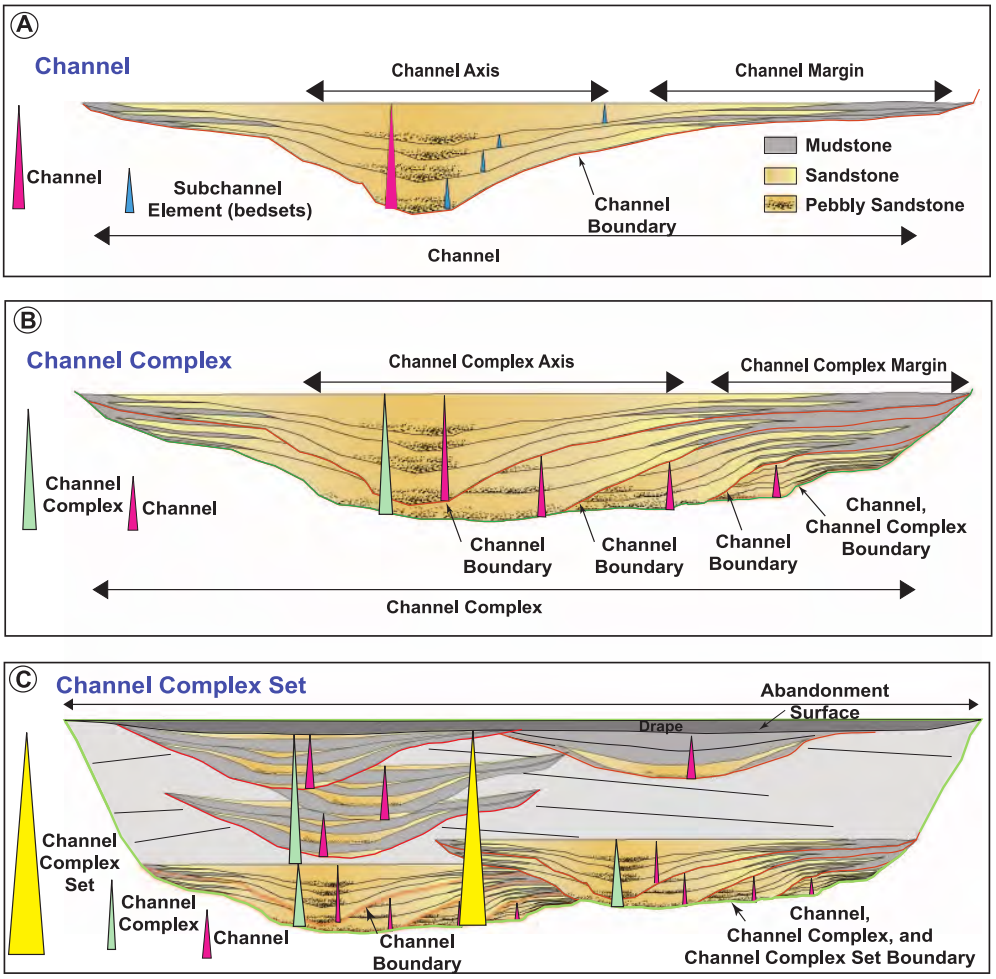


Figure 1.3 Example of hierarchy of turbiditic architectural elements, based on that of Campion et al. [2005] (from Stright et al. [2014]).

boundaries may be difficult to determine on the field. It leads to some difficulty in interpreting the complete hierarchy, with the use of the same denomination to characterize different elements at different scales [e.g., Wynn et al., 2007].

In any case, channels often form within a larger-scale confinement. Along the continental slope, the canyon is a first confinement. It is mainly an erosional structure, into which sediments circulate to the basin. On the basin, the erosive early phase of the turbidity events can create a new confinement. The resulting element has a channel-like shape (figure 1.4). It is usually hundreds of meters

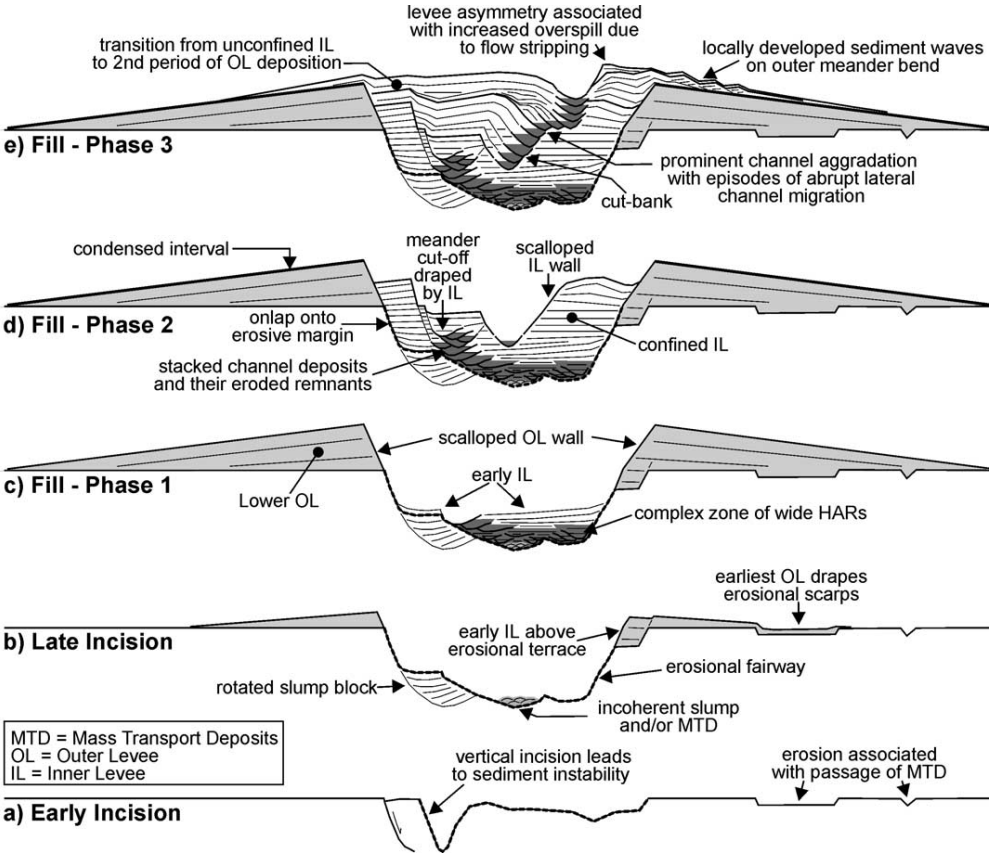


Figure 1.4 Conceptual model showing the incision and filling of a master channel (from Deptuck et al. [2003]). Channel initiation and filling are responsible for the filling of the master channel.

thick and several kilometers wide. This element corresponds to the channel as introduced by Mayall et al. [2006]: a large erosionally based channel delimited by a third order stratigraphic sequence boundary. It has various names, e.g., master channel [Kolla, 2007], submarine incised valley [Janocko et al., 2013b], erosional fairway [Deptuck et al., 2003]. The upper part of the currents may also overspill from the master channel margins. This leads to the deposition of sediments over the channel banks. These sediments form overbanks or levees. Thus, overbanks form at two scales [Kolla, 2007] (figure 1.4):

- Along the channel margins. They are called inner/secondary overbanks/levees.

- Along the master channel margins. They are called outer/master overbanks/levees.

When the outer overbanks are high enough, they confine the channel development. The overbank height depends on the mud content of the currents. A high mud-content increases the levee cohesion, leading to higher levees with a larger confining capacity [e.g., Galloway, 1998].

Lobes

Channelized structures act as transitional conduits for the sediments from the continental margin to the sea floor. Lobes result from the accumulation of sediments at a channel way out. During their circulation through the channels, the sediments are more or less sorted. The characteristics of the terminal lobes depend on the nature of the sediment input, e.g., mud-rich or sand-rich, and on the sorting power of the channels [Mulder and Etienne, 2010]. Terminal lobes are the most distal element of a turbiditic system. A distinction can be made between terminal lobes at the end of a channel and avulsion lobes. Avulsion lobes occur when an overbank breaches. It is a first step before the formation of a new channel path [e.g., Droz et al., 2007, Kolla et al., 2007] (for more details about avulsion, see section 1.4.3).

Globally a lobe has an ellipsoidal and convex-up morphology. The horizontal shape is more or less elongated depending on the confinement. It is usually several meters thick and tens of kilometers wide [e.g., Shanmugam and Moiola, 1988, Mulder and Etienne, 2010]. Internally, the lobe starts by a channeled part, corresponding to the continuation of the channels [Mulder and Etienne, 2010]. Overbanks can develop, but with a high sand content. Thus, avulsions are frequent, with the formation of new channel paths. The external parts of a lobe are composed of sheet-like structures. The first structure is deposited right after the channel mouth and contains the coarser sediments. The second structure contains the finer sediments. Similarly to channels, hierarchical models exist for lobes, which gather for instance into complexes (figure 1.5) [Prélat et al., 2009].

1.3 Connectivity in turbiditic systems and impact on fluid flows

Turbiditic systems contain various structures with complex internal facies heterogeneities. The main question is then the impact of such heterogeneities on

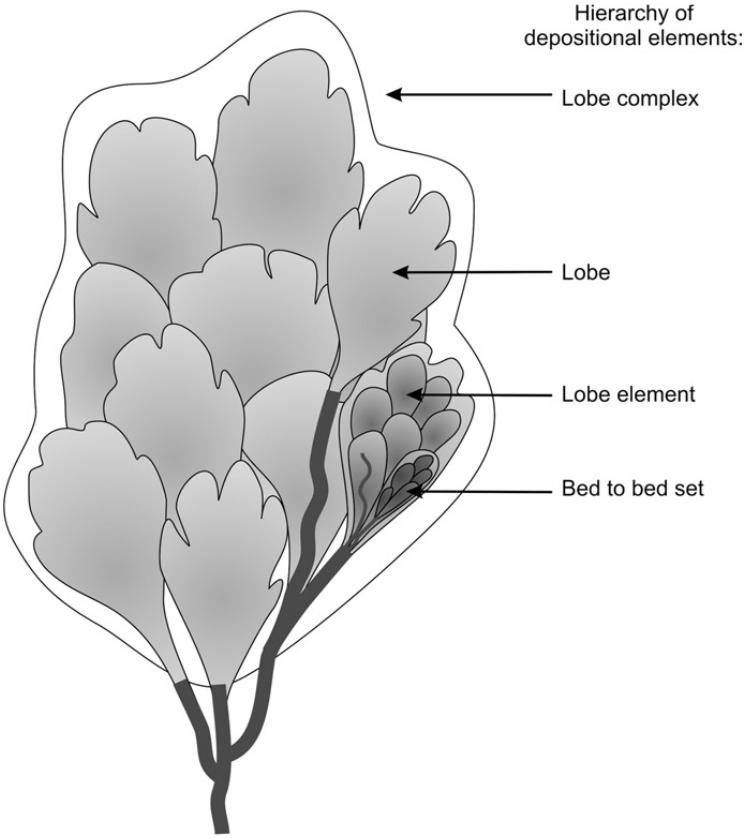


Figure 1.5 Example of lobe architectural element classification (from Prélat et al. [2010]). The feeder channel is at the bottom, in dark gray.

the flow.

1.3.1 Turbiditic system deposits and reservoir properties

The facies deposited by a turbiditic system can be roughly separated in two categories:

- The sand-rich deposits have a high sand content, so coarse-grained sediments, and a low mud content. The grain size lets plenty of space after deposition. Thus, such deposits often have a significant permeability and interesting reservoir properties.
- The mud-rich deposits have a high mud content, so fine-grained sedi-

ments. Such deposits have a far lower permeability than sand-rich deposits and can constitute flow barriers.

Among the architectural elements previously defined, channels and lobes are known to form interesting reservoirs due to their high sand-content [e.g., Shanmugam and Moiola, 1988, Mayall et al., 2006]. Overbanks can also form potential reservoirs depending on their sand content [Mayall and O'Byrne, 2002, Beaubouef, 2004]. The main questions concern the fluid circulation ability within and between all these structures. For instance fluid circulation within the overbanks may be difficult considering their heterogeneity and the poor lateral continuity of the thicker sand-rich beds. This question is even more complex because a channel and a levee can be spatially adjacent without a genetic relationship [Cronin et al., 2000]. The following sections illustrate the potential connectivity issues by focusing on the channels, as they are a fundamental structuring feature of turbiditic systems and have a significant reservoir potential.

1.3.2 Connectivity within a channel and between channels

A key aspect in a reservoir perspective is the connectivity between the sand-rich deposits. Channels are filled with such deposits, hence their reservoir potential. But the channel fill is usually not homogeneous: it depends on the characteristics of the turbiditic system, such as the mud-content of the sediment supply. The resulting architecture can be quite complex with combinations of sand-rich beds and mud-rich beds (figure 1.6) [e.g., Labourdette, 2007, McHargue et al., 2011, Alpak et al., 2013].

The first deposits result from the end of the erosive phase. They are composed of poorly sorted particles with mud clasts. This basal lag has a low permeability and constitutes a potential barrier for the flow. The channel axis is then filled with alternating sequences of coarse-grained sand to fine-grained sand. These sequences have usually a significant permeability and concentrate the essential part of the reservoir potential. The channel margins may contain mud drapes along the margins and mud-rich beds intercalated with sand-rich beds. These muddy elements have a low permeability, with two consequences:

- The fluid circulation between the channel axis and margins can be perturbed. Some fluids are then bypassed and remains in the margin
- The fluid circulations between the channels can be impossible.

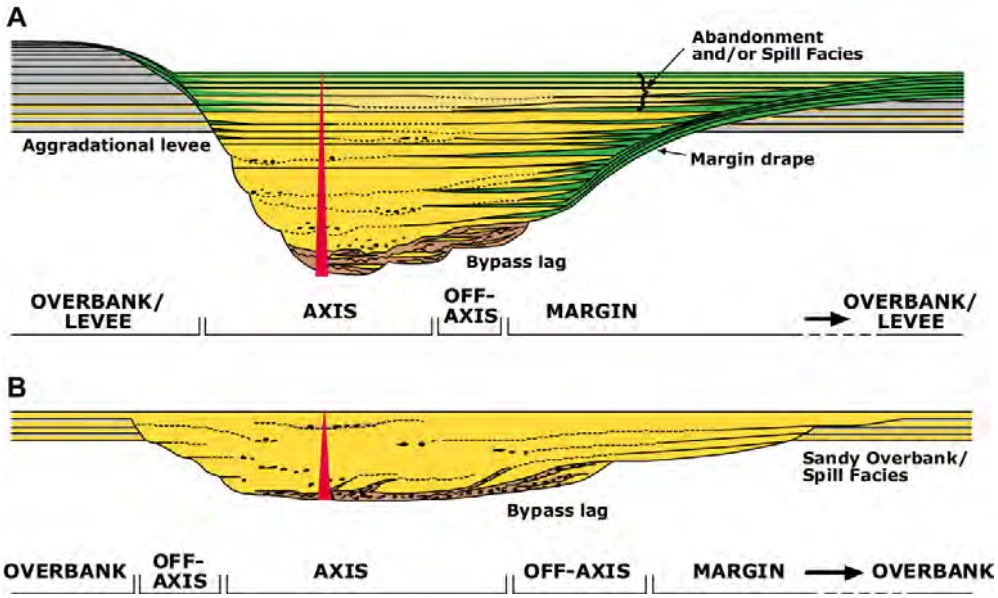


Figure 1.6 Conceptual models showing two different kinds of channel filling (from McHargue et al. [2011]). Model A is an under-filled channel with a moderate to high rate of overbank development. It contains several mud-rich deposits (in green, gray and brown), potential flow barriers. Model B is filled channel with a low rate of overbank development. Mud drapes are rare and the sand-content (in yellow) is important, with even potential sand deposits on the overbanks.

The channel margins and the related overbanks can also collapse before and/or during the channel filling. It forms slumps with a significant mud content, so a poor permeability. The upper part of the channel may be filled with abandonment deposits. Such deposits are mud-rich and constitute another potential flow barrier.

To summarize, margin drapes, margin collapse deposits, basal lags and abandonment deposits can all form flow barriers all around the channel. But they do not necessarily compartmentalize the reservoir. Indeed, the relative positioning of channels has a huge impact on the barrier influence [Labourdette et al., 2006, Labourdette, 2007]. Following their genetic relationships, channels are more or less stacked. The stacking can be described by a lateral and a vertical offset between two successive channels. The offset is not constant between two channels and can even vary significantly over short distances [Mayall et al., 2006, Deptuck et al., 2007]. It impacts the spatial relationships between the reservoir deposits and the flow barriers. The stacking also influ-

ences the preservation of the potential flow barrier by limiting their erosion [Labourdette, 2007]. Depending on the stacking, a mud-rich deposit prevents the flow circulation between two channels or not.

1.3.3 Connectivity influence on the reservoir and its exploitation

Two elements are essential to characterize the reservoir compartmentalization: the channel fill and the channel stacking. The channel fill determines the spatial distribution of the flow barriers and the channel stacking controls the impact of such barriers on the compartmentalization. Thus, their effect on the reservoir connectivity is significant [e.g., Labourdette et al., 2006, Larue and Hovadik, 2008, Alpak et al., 2013]. Moreover, while the channel stacking may be interpretable on seismic data, the channel fill is often subseismic, what complicates the estimation of the reservoir connectivity. It has been well illustrated by the Shiehallion field case [e.g., Gainski et al., 2010]: the reservoir compartments are less inter-connected than initially estimated. The first overestimation of the connectivity partly comes from unobserved margin collapses and drapes due to their subseismic scale. In practice, it led to double the well number to maximize the oil recovery.

Thus channel fill and stacking have a significant impact on the exploitation strategy, in particular concerning the well positioning [Mayall and O'Byrne, 2002]. Figure 1.7 illustrates this aspect: an injector was drilled and a location for a producer has to be chosen. The most permeable units are within the channels, and the inter-channel area is composed of mud-rich overbank deposits. Thus, the permeability is far higher along the channel direction than perpendicularly. The flow starts from the injector and reaches the producer 1 at first. Then it goes in the second channel and reaches the producer 2. From this point of view, the producer 1 risks a rapid water breakthrough and less oil recovery. The producer 2 seems far more interesting. However, this approach does not take into account the potential barriers between the channels. Basal barriers, such as basal lags, may let the flow circulate. Lateral barriers, such as margin drapes, may not, especially if the stacking is lateral. A vertical stacking at this location may let the flow circulate. Thus, the connectivity between the wells has to be carefully estimated, or the producer 2 will be useless due to the lack of connectivity between the two channels.

Channel fill and stacking also influence the well trajectory [Mayall and O'Byrne, 2002, Mayall et al., 2006]. A strategy when drilling a well is to go through as many barriers as possible. This allows bypassing the potential compartmentalization. The well geometry must be adapted to the channel

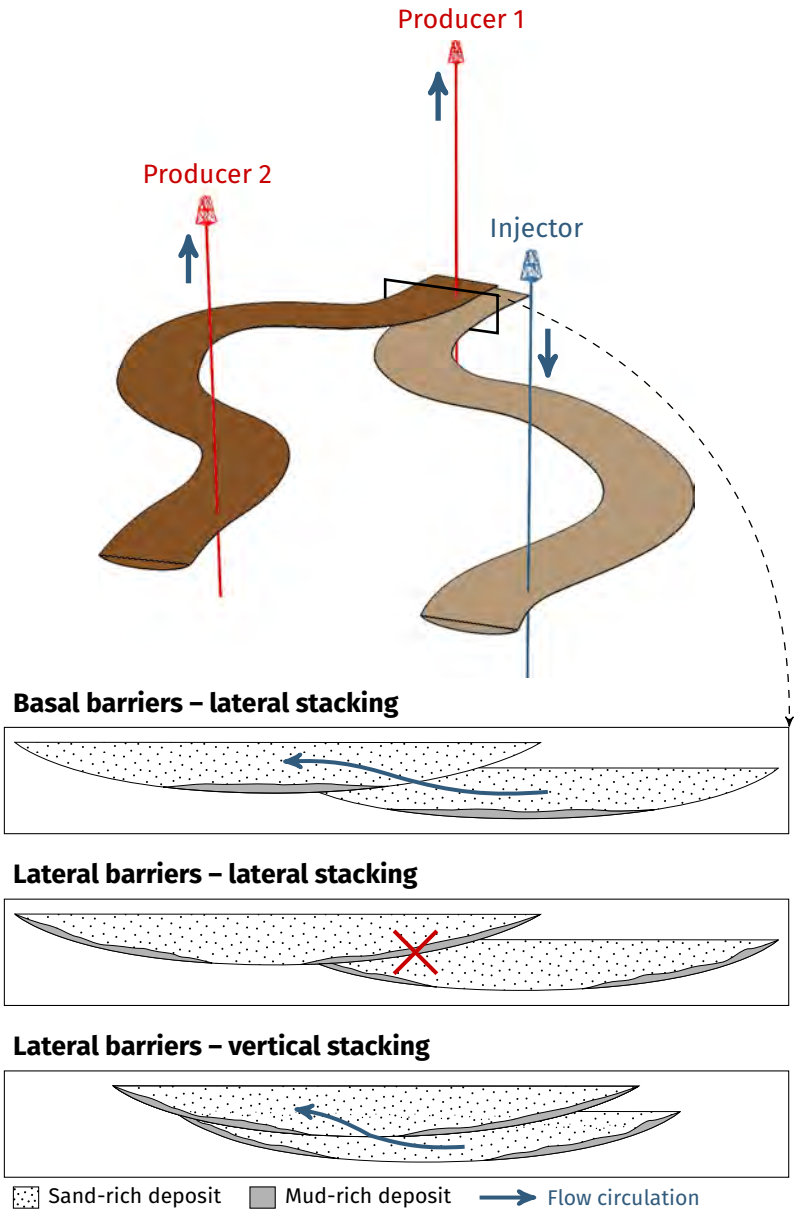


Figure 1.7 Impact of the channel fill and stacking on the connectivity. This impact has significant implications on the well locations (see the text for details).

stacking to cross has many channels as possible: vertical if the stacking is mostly vertical, horizontal if the stacking is mostly lateral. Thus, the channel stacking is a fundamental aspect to study because of its relationship with the flow barriers and the related impact on the connectivity.

1.4 Processes controlling channel stacking in turbiditic systems

The channel stacking comes from the lateral and vertical evolution of channels. This evolution depends on the turbidity current circulation within the channels and on the channel confinement, e.g., the presence of a master channel and overbanks.

1.4.1 Turbidity current circulation and secondary flow

Within a channel, a current follows a primary flow oriented along the channel direction. Perpendicularly to the channel orientation, the flow varies following the channel morphology, especially its curvature [e.g., Einstein, 1926, Rozovskiĭ, 1957, Imran et al., 1999]. The centrifugal force pushes the current toward the outer bank of the bend, elevating the current at the outer bend compared with the inner bend. This induces a pressure gradient in the opposite direction, so a force toward the inner bank of the channel. Contrary to the pressure force, the centrifugal force depends on the current velocity. And that velocity is not uniform. Vertically, the velocity is lower along the channel floor due to frictions (figure 1.8). Thus, the centrifugal force overcomes the pressure force at the top of the channel, and at the bottom, the pressure force overcomes the centrifugal force. This induces a rotational secondary flow. Combined with the primary flow, it gives a helical structure to the current. Besides the curvature, other features influence the development of a secondary flow, such as the channel topography or the current turbulence.

Secondary flows have been highly studied in the subaerial currents of rivers [e.g., Rozovskiĭ, 1957, Ikeda et al., 1981, Johannesson and Parker, 1989]. In such currents, the secondary flow goes from the outer bank to the inner bank along the bottom of the bend apex. It is directed inward, with the helix rising against the inner bend (figure 1.9). However, the direction of rotation of the secondary flow is more controversial in submarine currents. Both laboratory experiments and numerical simulations have shown inward-directed secondary flows in submarine channels, i.e., with the same direction than in rivers [e.g.,

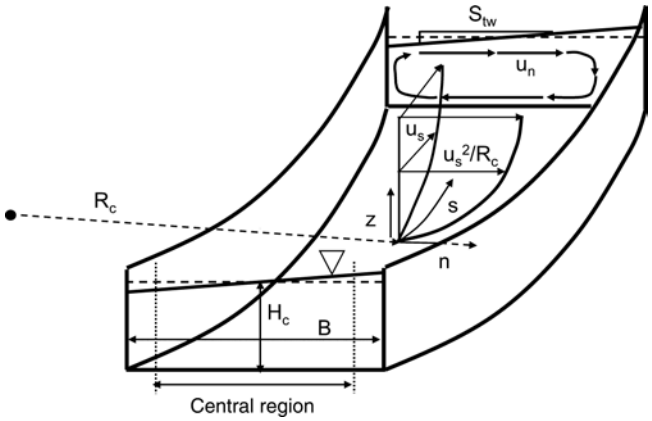


Figure 1.8 Diagram illustrating a secondary flow in a channel bend (from Lajeunesse et al. [2010]). u_s is the primary flow velocity, u_n is the secondary flow velocity and u_s^2/R_c represents the centrifugal force. As a result, the flow is super-elevated along the outer bend (on the right).

Imran et al., 2007, Islam et al., 2008, Imran et al., 2008]. But other laboratory experiments and numerical simulations show a reverse direction of rotation [e.g., Corney et al., 2006, Keevil et al., 2006, Corney et al., 2008]. The secondary flow then goes from the apex inner bank to the apex outer bank along the bend bottom (figure 1.9). It is directed outward, with the helix rising against the outer bank.

Such reversed secondary flow comes from the particular velocity profile of submarine flow. In subaerial flow, the difference of density between the current, i.e., the water in the river, and the surrounding fluid, i.e., the air, is such that the current does not mix with the surrounding fluid. Thus, the elevation of the current is limited and the maximum speed is at the top of the current. In submarine conditions, the difference of density between the current and the surrounding fluids is low. This enables a far higher current elevation, with the current going far more easily outside the channel. It induces a different velocity profile, with a maximum velocity in the lower part of the current. The maximum velocity can be closer to the channel bottom, with lower velocities at the channel top. Then the direction of rotation is the opposite of that in subaerial flow.

Recently, some authors point out that the direction of rotation of the secondary flow within submarine channel may not be uniform [Abad et al., 2011, Giorgio Serchi et al., 2011, Janocko et al., 2013a]. Indeed, the direction of

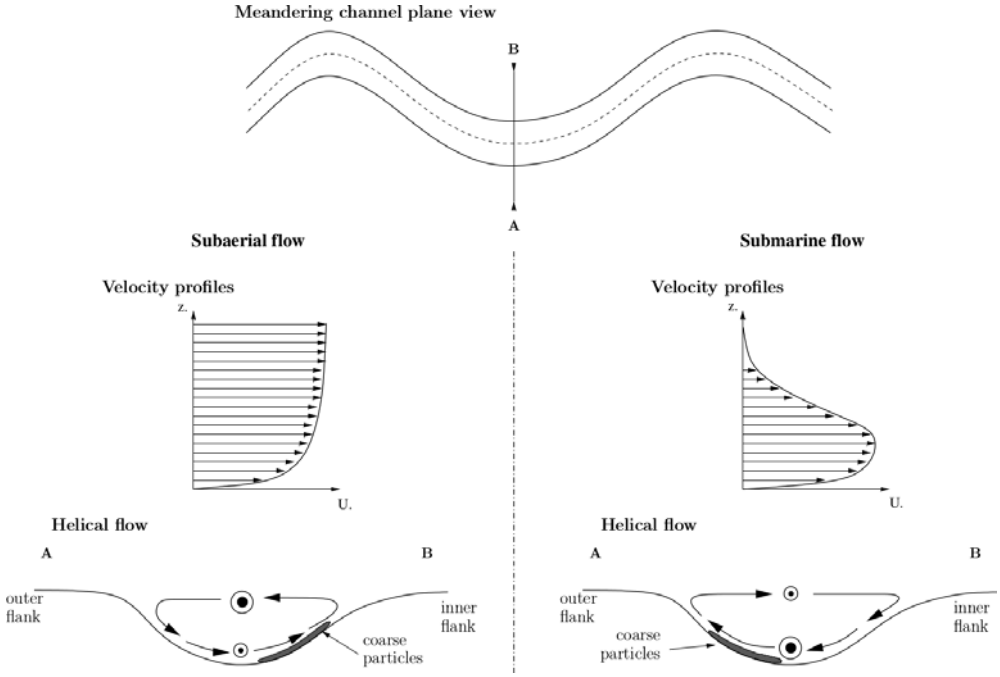


Figure 1.9 Comparison between the vertical velocity profile in subaerial and in submarine flows (modified from Salles [2006]). The location of the maximum velocity along the profile can give rise to a different direction of rotation of the resulting helical flow.

rotation depends on the velocity profile, and especially on the elevation of the maximum velocity of the current. If that elevation is close to the channel floor, the secondary flow is directed outward. If that elevation is close to the channel top, the secondary flow is directed inward. Thus, the channel morphology and the current characteristics both influence the rotation direction.

Under the action of the centrifugal, pressure and friction forces, the velocity profile of the current is non-uniform. Variations of the velocity profile appear vertically, but also transversally to the channel orientation (figure 1.10). The velocity is lower along the bank on which the helix is rising, and higher along the opposite bank. These velocity differences induce sediment deposition along the bank on which the helix is rising and erosion along the opposite bank. Such processes rule the channel evolution.

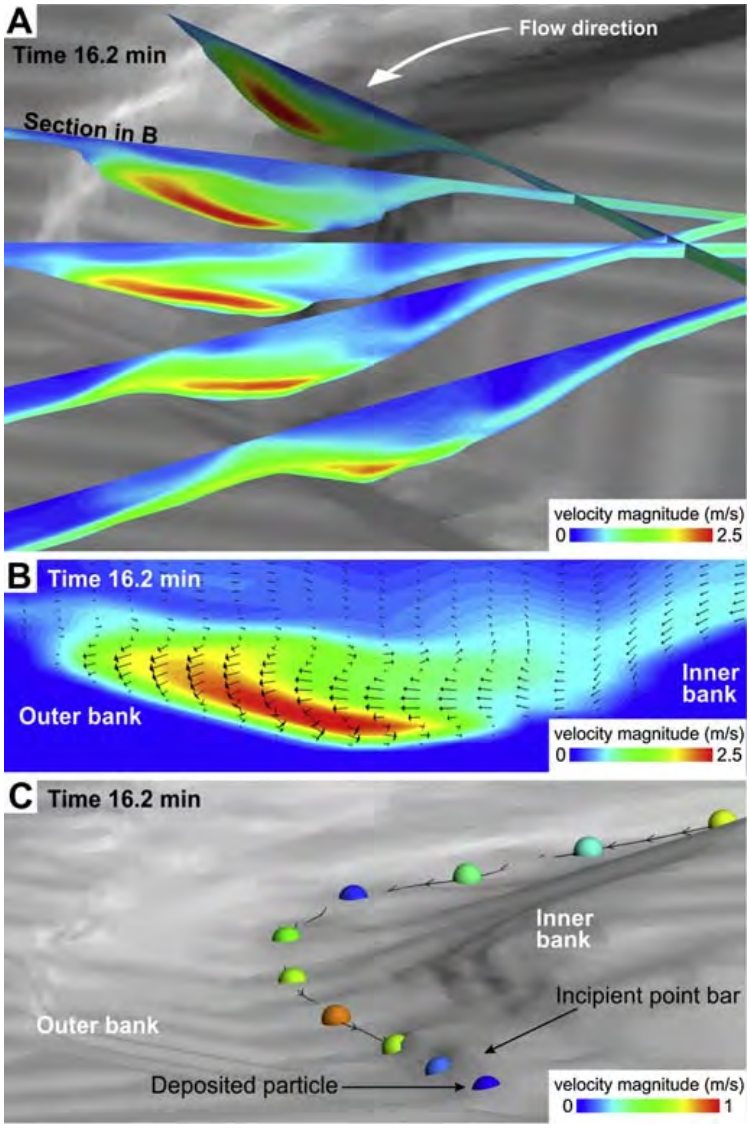


Figure 1.10 Example of numerical flow simulation results within the bend of a submarine channel (modified from Janocko et al. [2013a]): A. Velocity field of the current along the channel. The maximum velocity concentrates along the outer bank of the bend. B. Profile along the bend with the velocity field and the associated helix. C. Simulation of the transport of a particle. This particle is deposited along the inner bank due to the lower velocities.

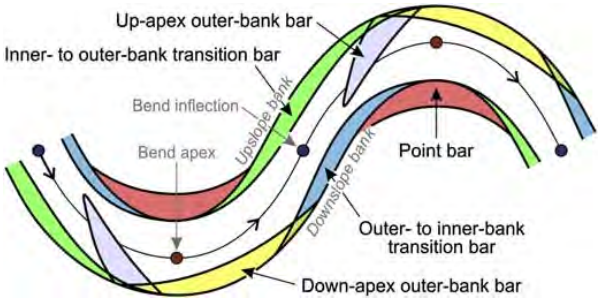


Figure 1.11 Potential bars that can form within a sinuous turbiditic channel (from Janocko et al. [2013a]).

1.4.2 Continuous channel migration

Channel evolution often refers to channel migration. Erosion and deposition processes along the two banks of the channel influence the migration, which is called continuous or gradual migration [e.g., Abreu et al., 2003, Kolla, 2007]. While one bank is eroded, the other is stretched by sediment deposition, leading to channel migration.

The migration deeply relies on the sediment deposition through the location of the depocentres, also called bars or accretion packages, within the channel (figure 1.11) [e.g., Abreu et al., 2003, Janocko et al., 2013a]. Three main types of bars are observable:

- Point-bars, also called lateral accretion packages (LAP), develop along the inner bank of a bend.
- Transition bars develop between the inner- to outer- or outer- to inner-bank transition.
- Outer-bank bars (OBB) develop along the outer bank of a bend.

Bars often have a significant sand content, with more or less muddy interbeds that can perturb the connectivity [Abreu et al., 2003]. Inner overbanks cover them. All these depocentres are associated to different trends in the development of the channel. Their deposition relates on the flow characteristics.

The factors controlling the flow circulation within channels are still subject to controversy [e.g., Dorrell et al., 2013, Sumner et al., 2014]. According to Janocko et al. [2013a], four key factors control the channel development:

- The deviation of the conduit profile from the equilibrium profile: for a flow with given characteristics, it is possible to define an equilibrium

profile along which the aggradation and the erosion are minimal. Thus, the channel-floor gradient tends to approximate the equilibrium profile. If so, the flow is in a slope-equilibrium state. If not, the flow tends to aggrade or to erode the channel.

- The deviation of the helix curvature from the channel curvature: the flow helix depends on the channel morphology. It also depends on the flow discharge, which impacts in particular the planform amplitude and wavelength of the helix. When these parameters evolve similarly to the amplitude and wavelength of the channel bends, the flow is "in phase" with the channel geometry. If the discharge changes abruptly, the consistency between the flow and the channel geometry can be lost. In such cases, the direction of secondary flow rotation in the bends can be reversed.

- The deviation of the current elevation from the channel elevation: the ratio of the flow elevation over the channel elevation determines the flow confinement. When this ratio is low, the flow is "under-sized" to "in-size" with the channel. When this ratio is high, the flow is "over-sized" compared to the channel. In such case, the current is poorly confined. The centrifugal force dominates and leads preferentially to the deposition of sediments at the outer bank. This classification concerns the flow entering the channel, independently of the downstream degree of flow overspill.

- The relationship between the flow power and the bank strength: it is fundamental for channel migration. If the flow power is too high and/or the bank strength too weak, the flow tends to leave the channel, either in an avulsion or in a braided pattern. If the flow power is too weak and/or the bank strength too high, the flow can not erode the bank and the channel do not migrate.

All these combined factors influence the sediment deposition and the erosion (figure 1.12). This gives rise to four major migration patterns:

- Lateral migration of the channel bends, also called swing, possibly associated to expansion, relies on point bar formation. The channel tends to migrate perpendicularly to the flow direction [Peakall et al., 2000, Posamentier, 2003] (figure 1.13), increasing the channel sinuosity.

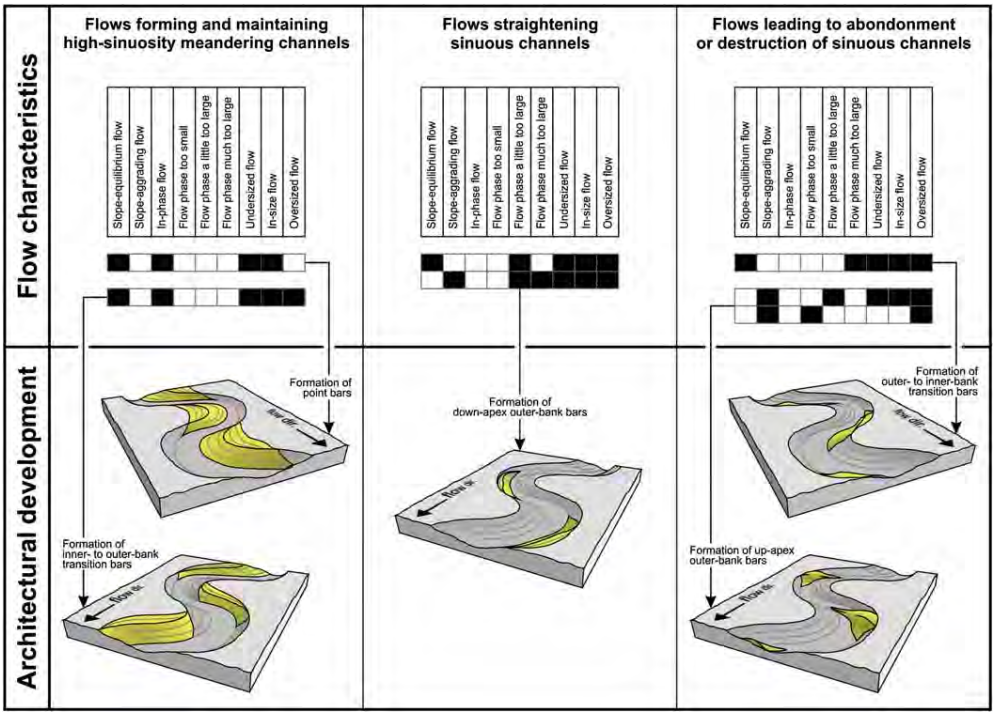


Figure 1.12 Influence of the flow characteristics on the formation of bars (from Janocko et al. [2013a]). The black squares indicate the development condition of the related bars.

- Downsystem migration or sweep relies on the development of inner- to outer-bank bars. The channel tends to migrate along the flow direction [Peakall et al., 2000, Posamentier, 2003] (figure 1.13).

- Retro-migration relies on the development of outer-bank bars. The migration is perpendicular to the flow direction, but with a straightening of the channel [Nakajima et al., 2009].

- Aggradation corresponds to the vertical migration of the channel. It is highly significant in turbiditic systems, with channels aggrading over hundreds of meters [Peakall et al., 2000].

All these patterns can combine depending on the flow properties.

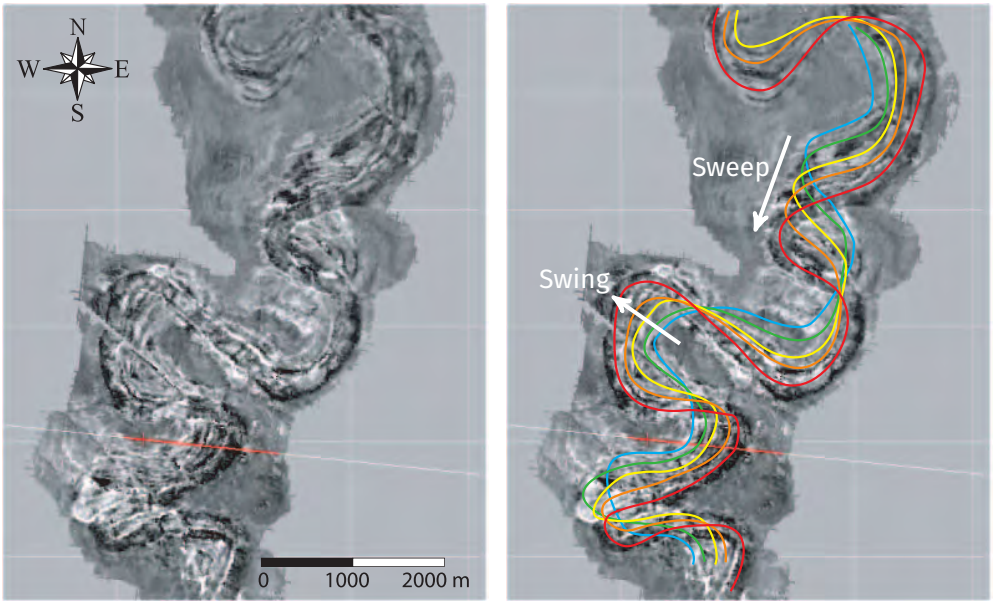


Figure 1.13 Example of channel migration patterns in a turbiditic system in West Africa (modified from Labourdette and Bez [2010]).

1.4.3 Discrete processes and avulsion

The flow deviation from the channel elevation and the bank strength are also influenced by the overbank deposits. These deposits increase the margin height and participate to the flow confinement. The development of the overbanks depends on the sand-content of the flow [e.g., McHargue et al., 2011], so on the sediment supply and on the energy of the flow. When the flow is sand-rich, the aggradation is low due to the low cohesion of the sand deposits. Thus, the overbanks are poorly developed, what significantly limits the confinement. When a new flow event begins, the previous channel has less influence on the development of the new channel, which can follow more or less different paths. The channel stacking is then disorganized.

When the flow is muddier, the cohesiveness of the deposits increases, and so does the aggradation. The overbanks are then more developed, with a better ability to confine the flow. The previous channel have then much more influence on the new one. The channel stacking is organized. The channel sinuosity also influences the stacking organization: when the sinuosity is high, the flow tends to have difficulty in following the channel geometry due to

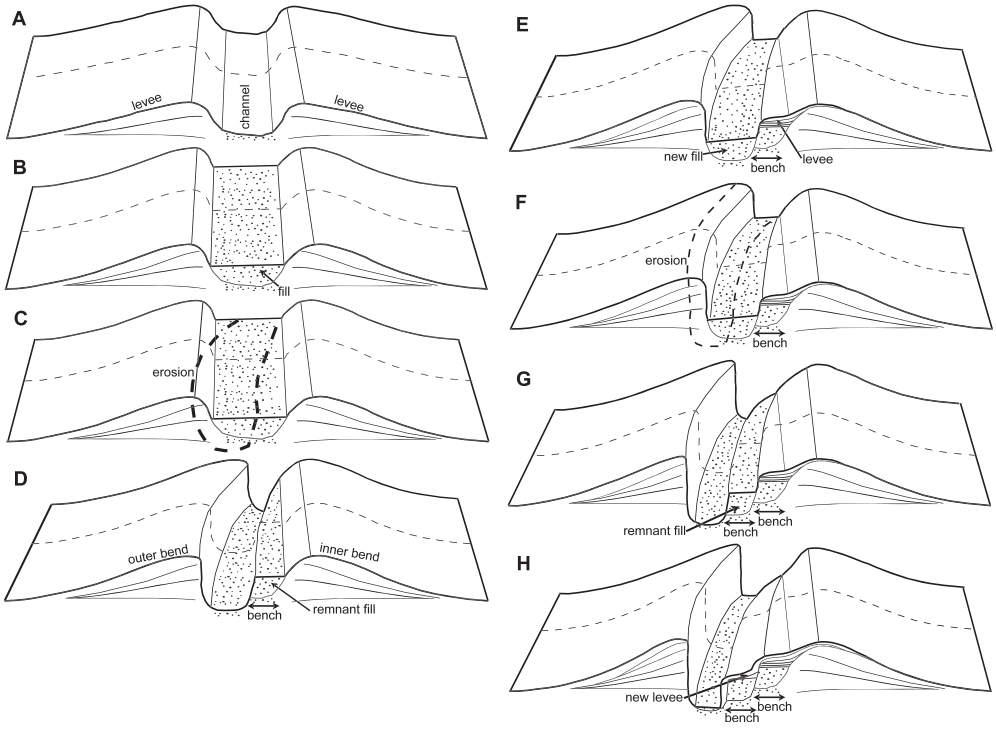


Figure 1.14 Conceptual model of punctuated migration (from Maier et al. [2012]). A channel forms (A, D, G) and is partially filled (B, E). A new channel is incised, with a lateral and eventually a negative vertical shift (C; F). The remaining part of the previous channel has a terrace-shape on which overbanks develop (E; H).

abrupt changes in the channel orientation. Thus, the new channel less likely follows the previous one, leading to a disorganized stacking.

The development of overbanks favors the flow confinement, leading to a less lateral or downsystem continuous migration. It is often associated with a significant aggradation. Less developed overbanks induces a disorganized stacking between the channels. This process relies on local changes of the channel path with the incision of a new channel. It does not relate on the deposition of bars combined with erosion, but still induces channel migration (figure 1.14). This process is called discrete, abrupt or punctuated migration [Abreu et al., 2003, Deptuck et al., 2003, Kolla, 2007, Maier et al., 2012]. In continuous migrations, the bars form along the channel margins, which influence the bar organization. In discrete migration, the previous channel is filled and then eroded. This influences the spatial relationship between the

heterogeneities of the channel fill. The resulting impact on the connectivity is then different than with a continuous migration process. Otherwise, the same migration patterns can be found with continuous and discrete migration [Kolla, 2007]. Continuous and discrete migrations do not seem exclusive, and both can occur within a single channel [Abreu et al., 2003].

Discrete migrations correspond to a local abandonment of the channel path, i.e., a local avulsion. The avulsion can be regional and leads to the complete abandonment of the channel path. Upstream to the avulsion location, the flow continues to circulate within the channel. Downstream to the avulsion location, a new channel initiates (figure 1.15). The process forms a new channel and possibly a new master channel. The conditions and development phases of an avulsion remain poorly known. As with discrete migration, the overbank confinement ability plays a significant role. Less developed overbanks favor avulsion. The progressive channel fill can also decrease the overbank confinement and favors avulsion [Armitage et al., 2012]. The avulsion initiation usually requires a levee breaching. Thus, the flow velocity plays a major role in the process [Kolla et al., 2007]. The channel sinuosity also influences the avulsion initiation: a higher sinuosity favors the possibility of an avulsion. Avulsion tends to occur most likely at the end of the canyon mouth, where the confinement is less significant [Armitage et al., 2012]. Once the overbank is breached, a lobe forms outside the channel. If the flow continues to go through the breach, it feeds the lobe, until the development of small channels and associated overbanks [Droz et al., 2007]. This last phase continues until one channel stabilizes and becomes the new channel.

Neck cutoffs are another kind of discrete process [Peakall et al., 2000, Posamentier, 2003]. Such structures come from the lateral migration. When a bend migrates with expansion, its extremities tend to get closer. This evolution comes with an increasing curvature along the bend. When the bend extremities are close enough, the high curvature makes it easier for the flow to bypass the bend. The bend is then abandoned and the flow circulates directly between the former bend extremities, along a straighter path (figure 1.16). Such process tend to limit the spatial expansion of the channels due to the migration. The abandoned bend is then filled by mud-rich sediments overspilling from the channel. This sedimentary body constitutes a potential flow barrier.

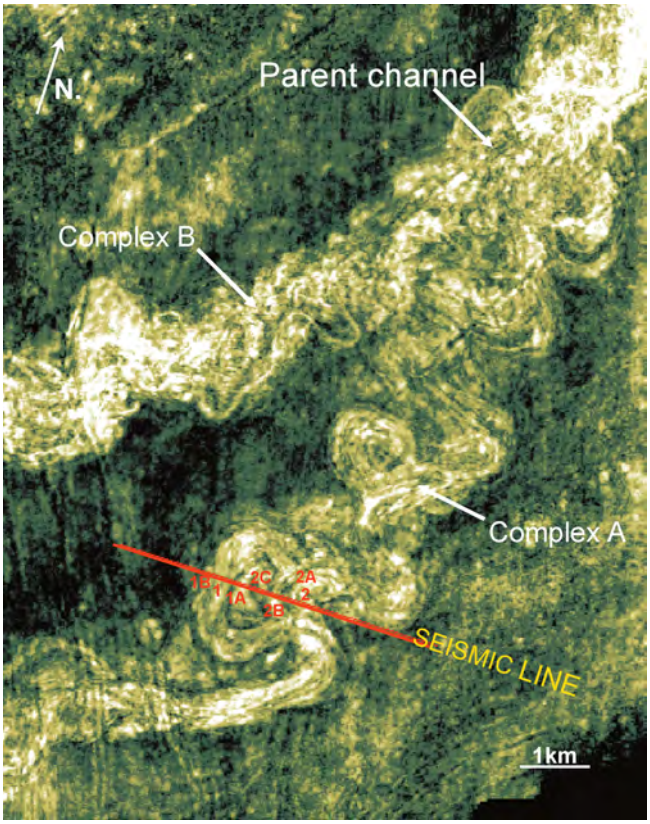


Figure 1.15 Seismic amplitude map showing an avulsion, Angola, West Africa (from Kolla et al. [2001]). This system started with the migration of a channel from the parent channel to the complex A. Then an avulsion occured, leading to the abandonment of the complex A. A new channel formed and has migrated to the current course, going from the parent channel to the complex B.

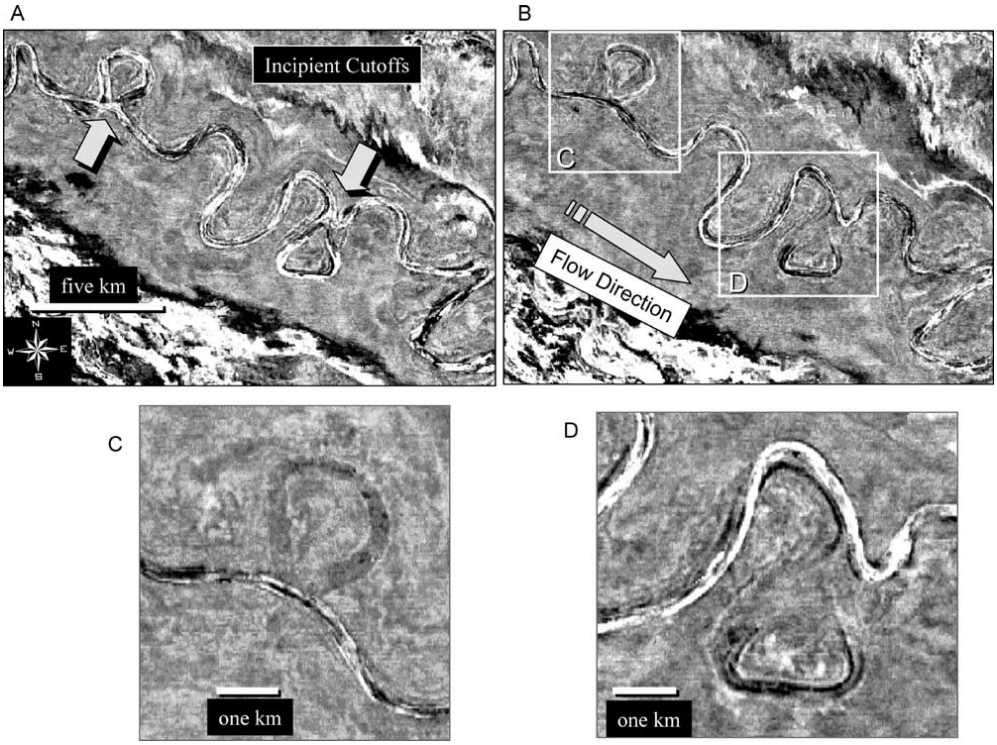


Figure 1.16 Neck cutoffs observed on seismic slices along the Joshua channel, Gulf of Mexico (from Posamentier [2003])

1.5 Stochastic simulation of channelized bodies and their connectivity

Turbiditic systems can form reservoirs either before or after the sediment consolidation. Before consolidation [e.g., Ostermeier, 2001], the system is close to its deposition state, thus potentially under thousands of meters of water. After consolidation [e.g., Braathen et al., 2012], the system is altered due to further geological processes, such as diagenesis, folding or faulting. Thus, the uncertainties about the reservoir structure and connectivity are often significant when facing a potential reservoir. Stochastic geomodeling is then especially useful to limit the biases in the interpretation.

1.5.1 Methods to stochastically simulate channelized bodies

Stochastic simulation methods can be divided following several criteria. A classical criterion is the support of the simulation results, i.e., a property within a grid or vectorial objects. In both cases, the methods simulate sedimentary bodies, such as channels, overbanks or lobes.

Cell-based methods

Cell-based methods regroup all the methods that simulate realizations by giving a value to each cell of a grid. The grid is a discretized representation of the domain of interest. The value gotten by each grid cell can be a sedimentary body type, for instance channel or levee (figure 1.17). These methods are mostly based on geostatistics. Their purpose is to reproduce in the whole grid the statistics obtained from a prior model. The model comes from the data or from a conceptual model. The data are most often punctual well data, interpreted in terms of sedimentary bodies. The conceptual model is a complete representation of the sedimentary objects of interest. Cell-based methods reproduce directly the result of the geological processes. Thus, they can also be qualified as structure-imitating methods.

Sequential indicator simulation (SIS) [Deutsch and Journel, 1992] is a well-known method for the simulation of discrete properties. The prior model comes from the data. Two elements compose this prior:

- The sedimentary object proportions, which indicate the importance of each sedimentary body.
- The variogram, which measures the correlation between two points depending on the distance between these points. For instance a channel leads to a high correlation along the channel direction and a lower correlation along the transversal direction (figure 1.17). The variogram captures the elongated shape of channel.

SIS relies on a sequential simulation process, i.e., the previously simulated values influence the next simulation steps, performed following a random path. At each simulation step, a value is simulated on the basis of a kriging system using the variogram. Such process ensures the conditioning to punctual data, as the data are attributed to cells at the beginning of the process. The plurigaussian simulation (PGS) [Galli et al., 1994] introduces a facies transition map to the prior. This map governs the relationships between the different sedimentary bodies, for instance if two bodies are in contact or not.

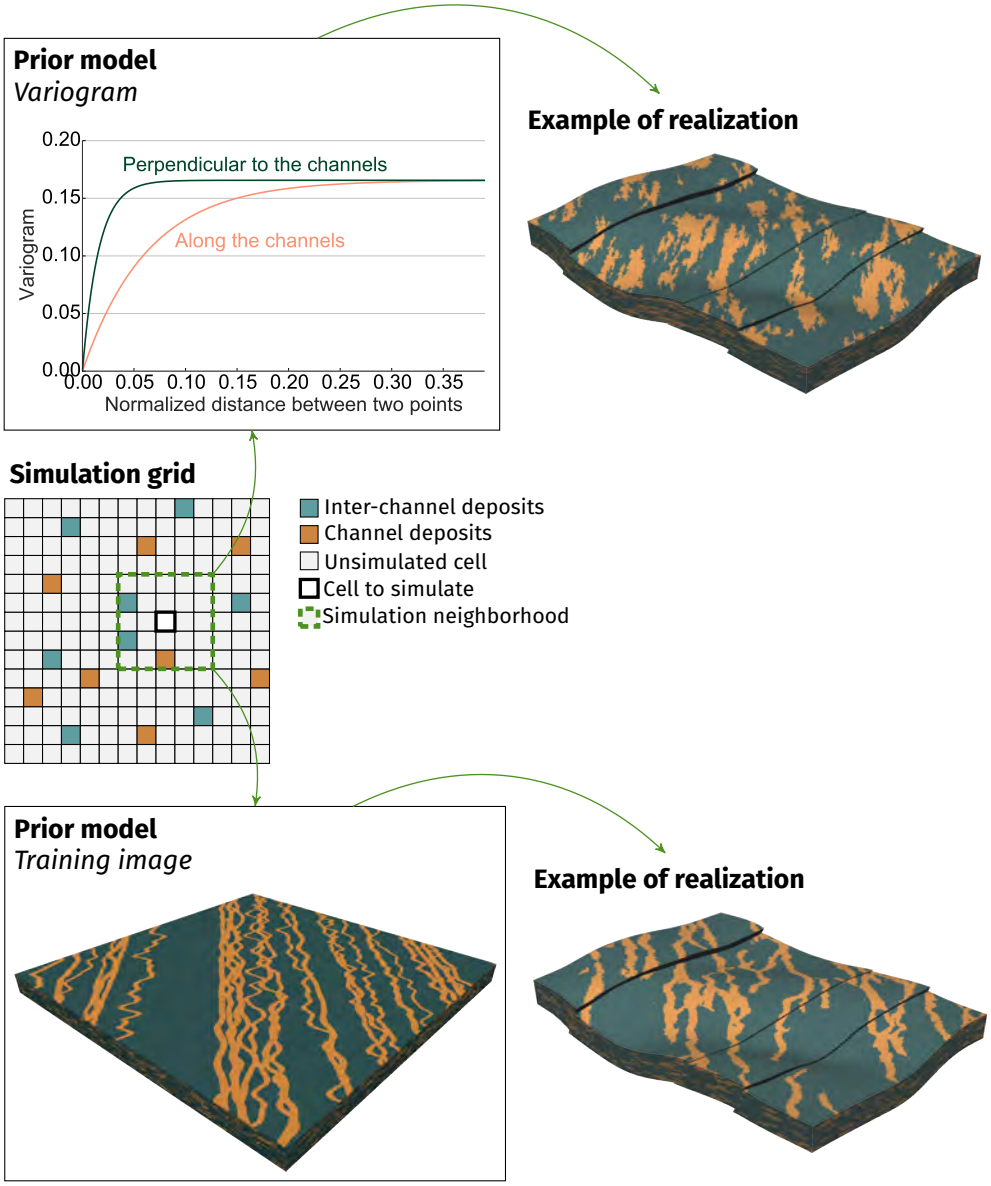


Figure 1.17 Principle of cell-based methods, with the example of sequential indicator simulation using a variogram [Deutsch and Journel, 1992] and multiple-point simulation, here the method IMPALA [Straubhaar et al., 2011], using a training image. The simulated values come from a prior model based on the known cell configuration in the neighborhood of the cell to simulate.

With multiple-point simulations (MPS), the prior no longer comes from the data. The principle is to borrow the higher-order statistics from a conceptual representation of the sedimentary objects to simulate [Guardiano and Srivastava, 1993]. This representation is called a training image and is the prior model (figure 1.17). The simulation process remains sequential. The purpose is to look at the configuration of cells with a known value around the cell to simulate. The training image is scanned to find the cells with the same neighboring configuration. This gives the probabilities of the possible discrete values considering the configuration of the neighborhood. The main issue with the original method is its computation time, as the training image is completely scanned for each cell to simulate. Two strategies were developed to improve the computation time:

- Scanning the training image once and for all before the simulation process [Strebelle, 2002, Straubhaar et al., 2011]. At the end of the scan, a data structure storage, either a tree or lists, conserves the probabilities of the possible discrete values considering the configuration of the neighborhood.
- Scanning the training image at each simulation step to compute a distance between neighborhood configurations [Mariethoz et al., 2010]. A distance threshold determines when the configuration with the lowest distance value is reached. The central cell of this configuration provides the simulated value. Thus, the scanning the whole training image is not always necessary.

Object-based methods

Object-based methods rely on the definition of geometric forms representing the sedimentary bodies of interest (figure 1.18). Each geometric form comes with its set of parameters. For instance, a radius and a thickness are enough to define a puck. It is a vectorial approach to the modeling of sedimentary objects. Thus, object-based methods do not necessarily require a grid in the simulation process, contrary to cell-based methods.

The classical approach for object-based modeling starts with the definition of a domain of interest. The objects are then randomly placed within the domain. The object parameters are drawn from statistical laws given in input.

The simulation of channels was at first introduced for a fluvial context [Deutsch and Wang, 1996]. But the similar morphologies between fluvial and turbiditic channels [e.g., Kolla, 2007] explain the use of the same object geometry in

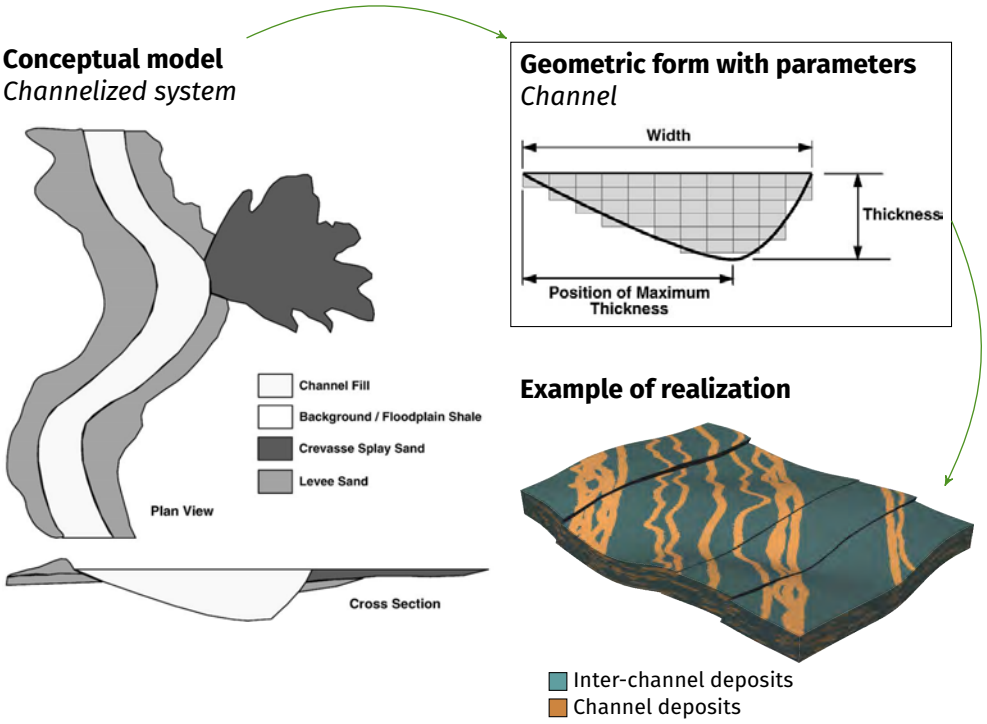


Figure 1.18 Principle of object-based methods, with the example of FLUVSIM for channel object simulation (modified from Deutsch and Tran [2002]).

turbiditic context. The classical parameterization of a channel relies on the channel centerline, also called backbone [e.g., Wietzerbin, 1994, Viseur, 2001, Deutsch and Tran, 2002]. The centerline defines the channel orientation and sinuosity. Then three parameters build the channel section (figure 1.18):

- The width at the top of the channel.
- The maximal channel thickness.
- The position of the channel maximal thickness along the section.

Other geometric definitions exist for levees, lobes and point bars [e.g., Viseur, 2001, Deutsch and Tran, 2002, Hassanpour et al., 2013, Ruiu et al., 2015b]. Some methods even rely on the simulation of point bars rather than on the simulation of a sinuous channel object [Hassanpour et al., 2013].

1.5.2 Controlling the connectivity in stochastic simulations of channelized bodies

The connectivity of the turbiditic deposits relies on the representation of channels, with their stacking and their fill. Possible strategies to integrate the connectivity within stochastic simulations are (i) better representing the channels and their stacking and (ii) conditioning to as much data as possible. In practice, this representation is in conflict with data conditioning in both cell-based and object-based methods. The data, especially well data, are yet the only information about the reservoir to have a low uncertainty. Thus their conditioning is essential.

Reproducing the channelized bodies while conditioning to the sedimentary data

Sedimentary data inform about the location of the channelized bodies. They come from punctual interpreted well data or from channel parts directly interpreted on seismic data if the resolution is high enough. These data are called hard data due to their low uncertainty.

Cell-based methods ensure hard data conditioning by filling the grid one cell at a time. However, they have difficulty in reproducing sedimentary objects. This aspect is well known with SIS. The variogram looks at the two-points statistics. So it ignores all the higher-order statistics and miss some information [Journel and Zhang, 2006]. Despite a better control on the object relationships, PGS faces the same issue. Higher-order statistics are necessary to build more consistent models. But they are difficult, if not impossible, to infer from the data. This explains the development of MPS.

MPS better reproduce sedimentary objects than SIS or PGS. Nevertheless, highly continuous structures like channels or small-scale and thin features like margin drapes are still far from being precisely reproduced. The preservation of their continuity is a major source of concern. It has given rise to many improvements [e.g., Liu and Journel, 2004, Strebelle and Remy, 2005] and many new methods, such as simulating all the cells within a neighborhood at once from the training image instead of a single cell [e.g., Arpat and Caers, 2004, Chatterjee and Dimitrakopoulos, 2012, Tahmasebi et al., 2012]. But this introduces conditioning difficulties.

On the other hand, the reproduction of sedimentary bodies by object-based simulations is far more satisfying from a geometrical point of view. The only requirement is to have a sedimentary body sufficiently known and described to

define a parameterization. Another interesting aspect is the possibility to build curvilinear grids within the objects [e.g., Deutsch and Wang, 1996, Shtuka et al., 1996, Ruiu et al., 2015a]. This facilitates the modeling of the channel fill and its heterogeneities, such as basal lags or margin drapes.

However, data conditioning of predefined objects is uneasy. The most straightforward approach is the rejection sampling: a simulated channel is kept if it fits some data [e.g., Deutsch and Wang, 1996]. Such method is highly inefficient when the number of data increases or with seismic-interpreted bodies. Other methods simulate a channel centerline conditioned to the data with a sequential Gaussian simulation [e.g., Shmaryan and Deutsch, 1999]. The simulation of high-sinuosity channels is not directly possible with such approach. Oliver [2002] proposes an iterative process to fit a whole channel centerline simulated with any method to well data. Other methods have a more local approach, either by deforming the channel [e.g., Viseur, 2001] or locally resimulating the channel [Mariethoz et al., 2014]. However, all these methods tackle the well data conditioning and not the conditioning of seismic-derived data such as sand probability cube. MPS can also be considered as a way to condition object-based methods. This is at the cost of the sedimentary object reproduction, so at the cost of their connectivity.

Reproducing the relationships between the channelized bodies

Simulating the relationship between sedimentary objects, in particular the channel stacking, is rather difficult with cell-based methods. Again, MPS are better able to handle the channel stacking than SIS or even PGS thanks to the training image. Similarly to the channelized body morphology, the stacking is yet poorly reproduced. Object-based methods *sensu stricto* randomly place the objects. While the placement of the previous objects can influence that of a new object, the resulting organization is far from the precise channel organization due to migration and avulsion.

Thus, event-based methods, also called pseudo-genetic or pseudo-process methods, manage to introduce genetic concepts in object-based methods [e.g., Pyrcz et al., 2012]. They intend to be process-imitating and not just structure-imitating. Pseudo-process-based methods were introduced for fluvial modeling. They rely on the work done on 2D physical simulation of river migration. 3D simulation of the flow circulating in a channel is highly time and resource consuming. 2D methods rely on a depth-averaged simulation and are more efficient. They link the migration to the asymmetry in the flow field induced by the channel curvature and responsible for the erosion/deposition processes

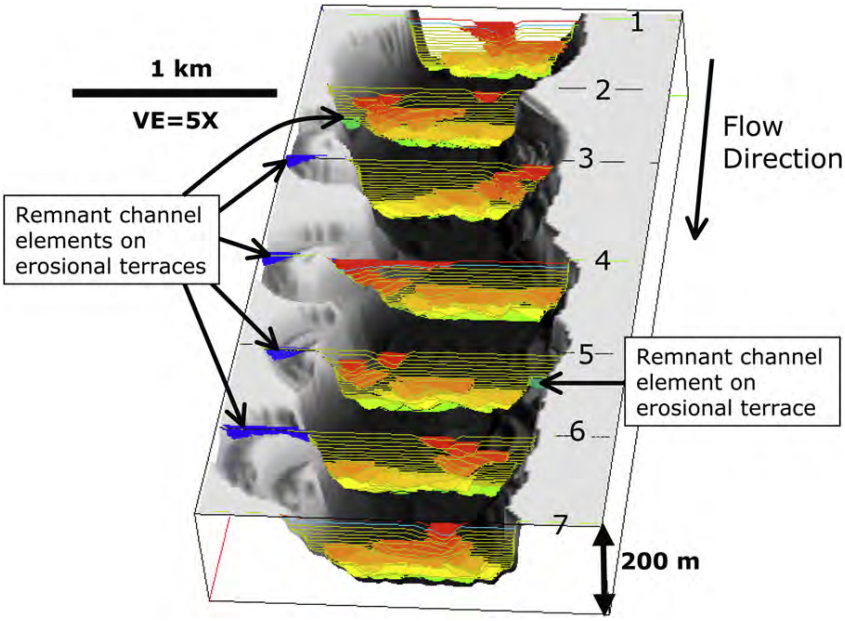


Figure 1.19 Example of realization simulated with a pseudo-process method (from McHargue et al. [2011]). The channels at the bottom of the master channel have a disorganized stacking, while the channel at the top have an organized stacking.

leading to the migration [Ikeda et al., 1981]. The simulation starts from an initial centerline which migrates during the process. These methods are used in a stochastic simulation process by Lopez [2003] and Pyrcz et al. [2009]. The channel parameterization remains the same as in object-based methods, with a centerline, a width, a thickness and a position of the thickness maximum. Each channel is related to the previous one through the centerline evolution.

Data conditioning is more problematic as the consistency between the channels has to be ensured. However, these methods significantly improve the possibility to simulate realistic-looking channel complexes (figure 1.19). They are primary designed for a fluvial context, but Imran et al. [1999] adapted the 2D genetic simulation of migration to turbiditic channels. Pyrcz et al. [2009] method was also directly applied to illustrate the development of turbiditic channels by McHargue et al. [2011]. However, the physical processes behind submarine channels are controversial [e.g., Corney et al., 2006, Imran et al., 2008, Corney et al., 2008]. Lately, Dorrell et al. [2013] argue that two-dimensional physical models are not accurate enough to capture the full three-dimensional structure of the flow field. This influences the possible migra-

tion patterns, so the channel stacking and the connectivity. Three-dimensional simulations are then required, but their validity remains controversial [e.g., Sumner et al., 2014] and they are computationally demanding. Their utility is then doubtful in a stochastic framework.

Conditioning to connectivity data

Seismic data are a first source of information about the connectivity. If the resolution is high enough, some channels can be directly interpreted and provide hard data. The channel stacking is then often observable [e.g., Mayall et al., 2006, Deptuck et al., 2007]. However, the channel fill is usually under the seismic resolution [e.g., Labourdette et al., 2006, Gainski et al., 2010]. If the seismic resolution is too low, some information can be extracted to condition the simulation, such as a facies proportion or probability cube [e.g., Strebelle et al., 2003]. Thanks to their spatial coverage, such data can help to constraint the simulation at a large scale. They are called soft data.

The wells give a more precise information about the location of the reservoir compartments and the flow barriers. However, this information is only punctual. The wells can only give large-scale connectivity data in terms of flow circulation. Tracer tests help to determine if two wells are connected. A tracer is introduced in an injector well. If the tracer is recovered on a producer well, the two wells are connected. Other information, such as the rate of recovery or the travel time, gives a first insight on the circulation path between the two wells. Production data also help to better characterize the flow behavior within the reservoir. Putting together these deductions with other data such as seismic data facilitates the understanding of the reservoir dynamic.

These data can also be introduced within the realizations. The usual approach consists in finding the model parameters such as the reservoir behavior of the resulting models matches that observed on the wells. This process, called history matching, is done through an inverse problem [e.g., Aanonsen et al., 2009, Oliver and Chen, 2010]. In a forward process, a model is simulated with given parameter values to predict the flow behavior. In an inverse process, the flow behavior observed on the wells is used to determine the possible model parameters. History matching relies on an optimization scheme to find the model parameters [Oliver and Chen, 2010], and is often computationally demanding.

Only few methods have been developed to constrain directly the realizations during their simulation. The principle consists in taking into account if two or several wells are connected, i.e., if a path of channel deposits link one well to the other(s). The productivity data themselves are not taken into account.

Allard and Group [1993] constrain a cell-based simulation to both the spatial structure of the prior and connectivity data. This approach simulates an initial field honoring the connectivity data and applies a Gibbs sampler to reproduce the spatial structure. Renard et al. [2011] constrain MPS realizations by extracting a continuous path between the wells from the training image and by using it as conditioning data. Thus, it can be used with all the MPS methods. This approach enables to condition object-based methods to connectivity data, by using the object-based realizations as training images for the MPS.

1.6 Conclusions

The interest for channelized system is due to the ability of their deposits to store fluids. From this point of view the connectivity between the reservoir facies, i.e., the sand-rich deposits, is fundamental. Turbiditic systems transport large quantities of sediments from the continental margin to the oceanic basin. These sediments circulate within channels through turbidity currents. The channels have nested structures, with channel evolving within master channels. Considering channels, the main features influencing the connectivity are their fill and their stacking. Channel fill is often heterogeneous, with mud-rich flow barriers. The channel stacking comes from channel migration and avulsion. These processes give rise to various genetic and spatial relationships between the channels, with a significant impact on the connectivity.

Such stacking is difficult to take into account in stochastic simulations. This is especially due to the difficulty to combine advanced geological representations with data conditioning. A better representation of the channel stacking is fundamental to ensure a correct connectivity estimation. It must be done jointly to a better integration of the available data, including the connectivity data. The study of the connectivity through channel heterogeneities and the related modeling aspects are here presented through turbiditic systems. Similar principles remains valid to reservoir/aquifer from fluvial systems.

Chapter 2

Analyzing stochastic simulations of sedimentary bodies using connected components

Abstract Many stochastic simulation methods are available to model sedimentary bodies. But realizations obtained with different methods and parameter values seem to differ visually. A major concern is the impact of those differences on the connectivity and then on fluid flow. The impartial analysis of the differences between realizations is so of prime interest. Here we propose to emphasize that analysis on the connectivity. The process relates on indicators computed from the connected components of the realizations. Some describe their spatial distribution, others their global shape or their topology. To analyze the realization differences, we compute dissimilarities between the images from the indicators. Heat maps and multidimensional scaling then facilitate the dissimilarity analysis. The application of this method to a synthetic case leads to some practical considerations. The multidimensional scaling is a powerful visualization tool, but it induces dissimilarity misrepresentations. The heat map displays the real dissimilarities and is more appropriate for a detailed analysis. The comparison with a multiple-point histogram method shows differences in the realization ranking. In any case, the indicators highlight differences between the realizations, whose connectivity actually varies. Further studying the indicator behavior could be beneficial to a better analysis of the connectivity within stochastic simulations.

Chapter submitted to Advances in Water Resources

2.1 Introduction

Connectivity is a key aspect of a geological study for its influence on fluid circulations. From a reservoir engineering perspective, it especially relates to sedimentary bodies with high and low permeabilities. It also relates to the spatial organization of these bodies and the resulting inter-connections. An incorrect body connection can bias the results of the flow simulations [Journel and Alabert, 1990, Gómez-Hernández and Wen, 1998, Labourdette et al., 2006]. Reproducing these sedimentary bodies together with their relations is so of prime importance [e.g., Deutsch and Hewett, 1996, King and Mark, 1999].

Stochastic simulations aim at generating possible representations of the sedimentary objects with respect to the available data. Several methods exist, with an usual separation in two main categories:

- Pixel-based methods simulate one cell at a time, based on a prior model describing the objects of interest. In sequential indicator simulation (SIS) [Deutsch and Journel, 1992], the prior is a variogram built upon the two-point statistics of the data. Hard data conditioning with such method is easy. But the simulated bodies do not look like geological objects. This is especially true for objects with curvilinear geometries such as channels, whose continuity is badly preserved. The plurigaussian simulation (PGS) [Galli et al., 1994] limits this difficulty by accounting for the facies relationships. Multiple-point simulations (MPS) go a step further by borrowing multiple-point statistics not from the data but from an external representation of the expected geology, the training image (TI) [Guardiano and Srivastava, 1993].

- Object-based methods rely on the definition of geometric forms and their associated parameters. Each form represents a particular sedimentary body [e.g., Viseur, 2001, Deutsch and Tran, 2002]. The objects are then randomly placed in the domain of interest with parameters drawn in statistical laws. More recent approaches introduce some genetic aspects to improve the object organization [e.g., Lopez, 2003, Pyrcz et al., 2009]. They provide more geologically consistent results. For instance channel continuity and relationships are better preserved than with pixel-based methods. But this is at the detriment of the ease of parametrization. And all those approaches have difficulty to condition the objects to data.

All those methods have advantages and drawbacks. This will influence the choice of a method and its parameters when dealing with a case study.

However few work aims at systematically analyzing the quality of a set of realizations. The quality control often consists in comparing the histogram and variogram of several realizations with those of the data, or of the training image if any [e.g., Strebelle, 2002, Mariethoz et al., 2010, Tahmasebi et al., 2012]. It is completed by a visual evaluation of the global structures, which is subjective and limited to a few realizations, often in two-dimensions [e.g., Tahmasebi et al., 2012, Yin et al., 2009].

Some authors provide some objective measures of the realization quality. Boisvert et al. [2010] and Tan et al. [2014] propose to analyze the multiple-point histogram. De Iaco and Maggio [2011] and De Iaco [2013] also explore the multiple-point statistics with high-order cumulants. Other authors focus on the further use of the realizations, often related to fluid circulations, and not on the imposed prior model. For instance Meerschman et al. [2012] use the connectivity function with the histogram and variogram to analyze the simulation parameter impact for the Direct Sampling MPS method [Mariethoz et al., 2010]. Deutsch [1998] uses directly the connected components determined from lithofacies, porosity and permeability models. He computes indicators such as the number of connected components or their sizes to rank the realizations. De Iaco and Maggio [2011] and De Iaco [2013] also use some statistics related to the connected components, such as their number or their mean surface and volume. These indicators on connected components have the advantage of being easy to compute and figure out. Comunian et al. [2012] rely on some of the previous indicators to analyze the quality of three-dimensional structures simulated from two-dimensional training-images. They also consider the equivalent hydraulic conductivity tensor as an indicator. However, this requires to have an idea of the hydraulic conductivities for the simulated facies.

In parallel, an important question is how to visualize the results. To compare the realizations authors often look at each quality indicator one by one. Scheidt and Caers [2009] and Tan et al. [2014] both rely on the computation of dissimilarity values between the realizations. Those dissimilarities are computed based on the quality indicators measured on each realization. They are then visualized based on a MultiDimensional Scaling (MDS) [e.g., Torgerson, 1952, Shepard, 1962a,b]. MDS represents the realizations as points, with the distance between the points as close as possible to the dissimilarities. The global analysis of the realization dissimilarities is so easier.

The present work aims at analyzing and discussing a set of statistical indicators to quantify the quality of stochastic simulations. This method performs on categorical three-dimensional images representing the facies constituting the sedimentary bodies of interest. It can be applied on realizations from one

or several stochastic simulation methods and/or parameter values. Conceptual images representing ideally the bodies to simulate can also be considered. Here the simulation quality is seen through the reproduction of the sedimentary body connectivity. The chosen set of indicators (section 2.2) relies on quantitative measurements on connected components and their skeletons. The indicators are used in dissimilarity computations to analyze the quality more directly (section 2.3). Several realizations obtained with different simulation methods (section 2.4.1) are then used to test the methodology (section 2.4). A comparison is made with the multiple-point histogram used as indicator.

2.2 Indicators to measure simulation quality

The indicators studied in this paper are based on the notion of connected components. Thus, they focus on the connectivity inside the three-dimensional image of the sedimentary bodies of interest. They are all quite simple and each one only gives a partial information about the connectivity and its structure. But their combination provides a more detailed characterization.

The indicators work with both regular and irregular structured grids. Thus the indicators based on volumes or surfaces are rather computed using number of cells and number of faces. This avoids biases related to different grid geometries, which give different indicator values even if the objects are the same when transferred in the same grid. The idea is to compare the components in the parametric space of the grids [e.g., Shtuka et al., 1996], as simulation methods often model geological objects in that space. The following sections describe the various indicators retained for this study. They can be separated in two groups: the first one contains global indicators related to the set of all connected components of a given facies, the second one contains indicators characterizing the connected components individually.

2.2.1 Basic element: the connected component

A simulation of geological objects is often done into a grid composed of cells. It is so a discretized representation of the sedimentary bodies of interest. Those cells are linked one to another by their faces, their edges and/or their corners [Renard et al., 2011]. In the case of the structured grids used for this work, one cell has three possible neighborhoods (figure 2.1):

1. One neighborhood composed of six face-connected cells.
2. One neighborhood composed of eighteen face- and edge-connected cells.

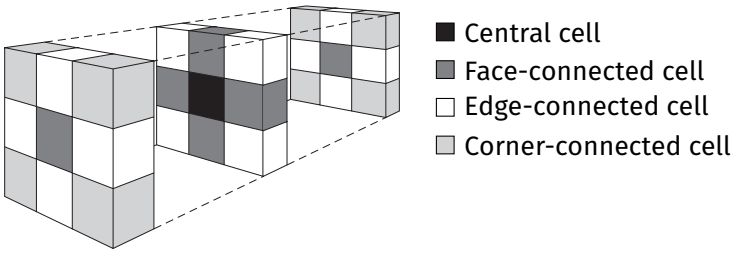


Figure 2.1 Possible neighborhoods for a given central cell in a regular grid (modified from Deutsch [1998]).

3. One neighborhood composed of twenty-six face-, edge- and corner-connected cells.

The connected components are defined using those neighborhoods: two cells belonging to the same facies are connected if a path of neighboring cells remaining within the same facies exists (figure 2.2).

This definition leads to a distinction between the sedimentary objects, such as a channel or a crevasse splay, and the connected components. Indeed, the sedimentary objects often tend to cross each others, giving one connected body where there is in fact several sedimentary bodies (figure 2.2). The range of possible shapes is larger for the connected components than for the individual object. This aspect complicates the comparison between images. But determining the connected components is far easier than trying to retrieve the geological objects. This is also close to the functioning of pixel-based methods, which do not try to reproduce geological objects but groups of cells, and therefore connected components.

2.2.2 Global analysis through all the components of a given facies

The global analysis gathers the indicators that do not characterize a particular connected component but an ensemble of connected components. They provide one value per facies.

Facies proportion

The facies proportion is of major importance in reservoir modeling considering its influence on porosity and permeability simulation. The proportion p of a facies f is defined as:

$$p_f = \frac{n_f}{n_t} \tag{2.1}$$

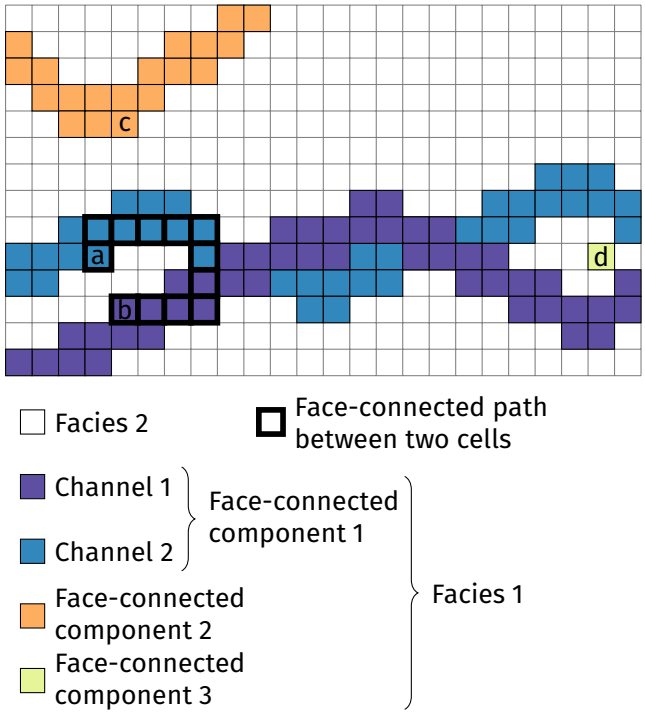


Figure 2.2 Connected components of a given facies in a two-dimensional structured grid. The cells *a* and *b* are connected and belong to the same connected body. There is no possible connected path between those cells and *c*, which belongs to another connected body. The cell *d* constitutes a third connected body in the case of a face-connected neighborhood. In the case of an edge- or corner-connected neighborhood, *d* belongs to the connected body 1.

with n_f the number of cells of facies f and n_t the total number of cells.

Facies adjacency proportion

The adjacency proportion between the facies quantifies the spatial relationships between those facies. The proportion p^a of a facies f adjacent to a facies j is defined as:

$$p^a_{f,j} = \frac{n_{f,j}}{n_{f,t}} \tag{2.2}$$

with $n_{f,j}$ the number of cells of facies f adjacent to the facies j and $n_{f,t}$ the total number of cells of f adjacent to a facies different from f.

Facies connection probability

The connection probability Γ quantifies the global connectivity of a given facies f. It is calculated as the proportion of pairs of connected cells among all the pairs of cells of the considered facies [Renard et al., 2011]:

$$\Gamma_f = \frac{1}{n_f^2} \sum_{i=1}^{N_f} (n_f^i)^2 \tag{2.3}$$

with n_f the total number of cells of facies f, N_f the number of connected components of the facies f and n_f^i the number of cells of the connected component i associated to facies f.

Connected component density

The density ϵ of connected components for a facies f expresses the number of components per image cells:

$$\epsilon_f = \frac{N_f}{n_t} \tag{2.4}$$

with N_f the number of connected components of the facies f and n_t the total number of cells.

Unit connected component proportion

The unit connected component proportion quantifies the number of connected components constituted by only one cell. The creation of such bodies may correspond to some erroneous and unwanted small scale noise. In that case, it is better to avoid them for the shape indicator computation (see section 2.2.3). The proportion p^u of unit volume connected components of a facies f is:

$$p_f^u = \frac{N_f^u}{N_f} \tag{2.5}$$

with N_f^u the number of unit volume connected components of the facies f and N_f the number of connected components for the same facies.

Traversing connected component proportion

Connected components can have various behaviors against the grid: they can be in the middle of the grid without contact with a border (figure 2.2, connected component 3), along one border, along two adjacent borders (figure

2.2, connected component 2) or traversing the grid completely from one border to the opposite border (figure 2.2, connected component 1). These last components having a main impact on the flow, their reproduction is of prime importance. The proportion p^c of traversing connected components of facies f is defined as the proportion of connected components that connect one border to the opposite one:

$$p_f^c = \frac{N_f^c}{N_f - N_f^u} \tag{2.6}$$

with N_f^c the number of traversing components for the facies f, N_f the total number of components for the facies f and N_f^u the number of unit components for the facies f.

2.2.3 Detailed analysis of the connected components

The following indicators are proposed to characterize any particular connected component. The values for different components are gathered to obtain one value per facies and make the indicator comparison easier.

Shape indicators

Global measures such as facies proportions are not sufficient to characterize precisely the impact of the related facies on the flow [e.g., Western et al., 2001, Mariethoz, 2009]. In particular, Oriani and Renard [2014] showed the influence of the connected component geometry – i.e., their shape – on the equivalent hydraulic conductivity, and therefore on the flow behavior. The following shape indicators characterize the connected component shape through simple surface and volumetric measures. They all give one value per component. The arithmetic mean of those values provides a value of the indicator for a given facies.

Number of connected component cells The number of cells n_f^g within a connected component g of a facies f represents the size of a connected component.

Box ratio The box ratio is based on the axis-aligned bounding boxes of the components (figure 2.3). It is related to the tortuosity of a component and to its orientation relative to the grid axis. The box ratio β of the connected component g of the facies f is expressed between 0 and 1 using the following

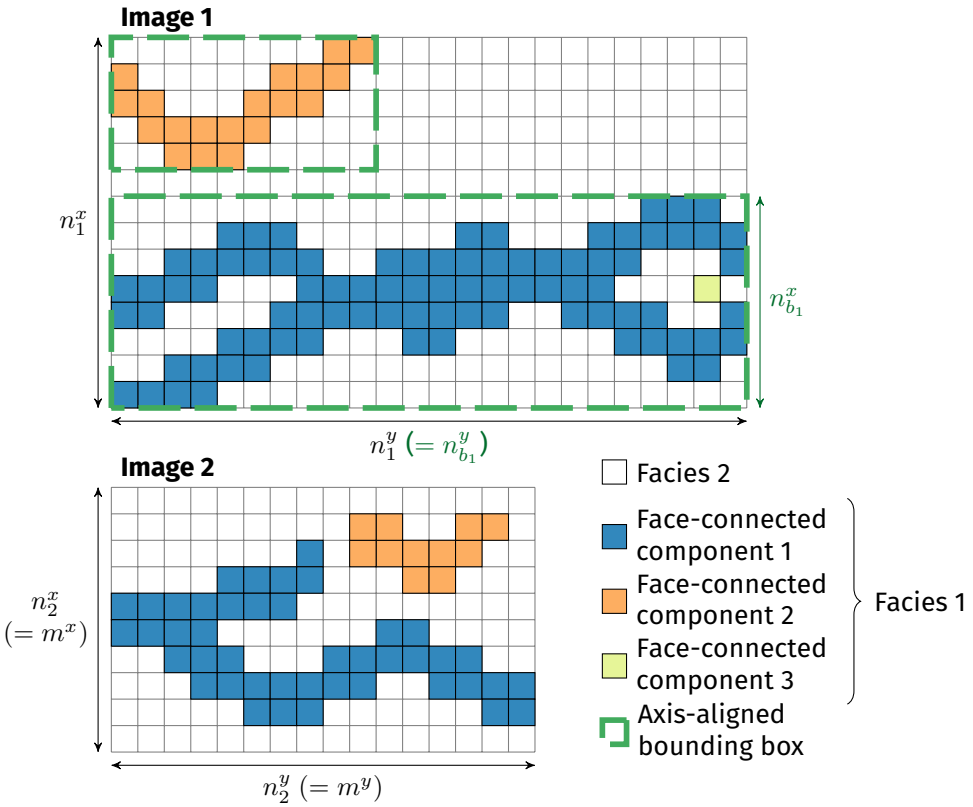


Figure 2.3 Connected components of figure 2.2 represented with their bounding boxes (image 1). Image 2 is another image in a different grid. The two grids have different dimensions along the x axis, which can influence both the number of connected components and the number of grid cells. But they also have different dimensions along the y axis, along which the channels are oriented. Changes along the y axis influence the number of grid cells, but not the number of channels. In that case, the comparison of the density of the two images is biased. The same principle stands for the number of cells of each components. Differences between the number of cells for the component 1 of each image mainly come from the difference in grid dimension along y, not from real differences of connectivity. n_1^x and n_1^y represent the number of cells along the axis x and y for image 1, n_2^x and n_2^y the number of cells along the axis x and y for image 2 and $n_{b_1}^x$ and $n_{b_1}^y$ the number of cells along the axis x and y for the axis-aligned bounding box b of the connected component 1. m^x and m^y are the lower sizes between the two grids for the axis x and y (see section 2.3.1 for more details).

formula:

$$\beta_f^g = \frac{n_f^g}{n_f^b} \tag{2.7}$$

with n_f^g the number of cells of the connected component g of the facies f and n_f^b the number of cells occupied by its axis-aligned bounding box.

Faces/cells ratio The faces/cells ratio is equivalent to the surface/volume ratio, which compares the surface area of a component with its volume. Deutsch [1998] uses it as a measure of the tortuosity of the components, but it is also affected by their roughness. The faces/cells ratio ζ of the connected component g of the facies f is expressed as:

$$\zeta_f^g = \frac{m_f^g}{n_f^g} \tag{2.8}$$

with m_f^g the number of faces along the surface of the connected component g of the facies f and n_f^g its number of cells.

Sphericity The sphericity ϕ compares the shape of a connected component g for a facies f with a sphere [Wadell, 1935]. It ranges between 0 and 1, where 1 corresponds to a sphere, and is expressed as:

$$\phi_f^g = 36\pi \frac{(n_f^g)^2}{(m_f^g)^3} \tag{2.9}$$

with n_f^g the number of cells of the connected component g of the facies f and m_f^g its number of faces along its surface.

Skeleton indicators

The skeletons are one-dimensional representations of the connected components. They are particularly used to assess the reproduction of the component topology.

Basic element: the skeleton A curve-skeleton is a thin one-dimensional representation of a three-dimensional shape. It is composed of nodes linked together by one or more segments (figure 2.4). The degree of a node is the number of segments connected to that node. Skeletons are often used to study some geometrical and topological features of a shape (figure 2.4).

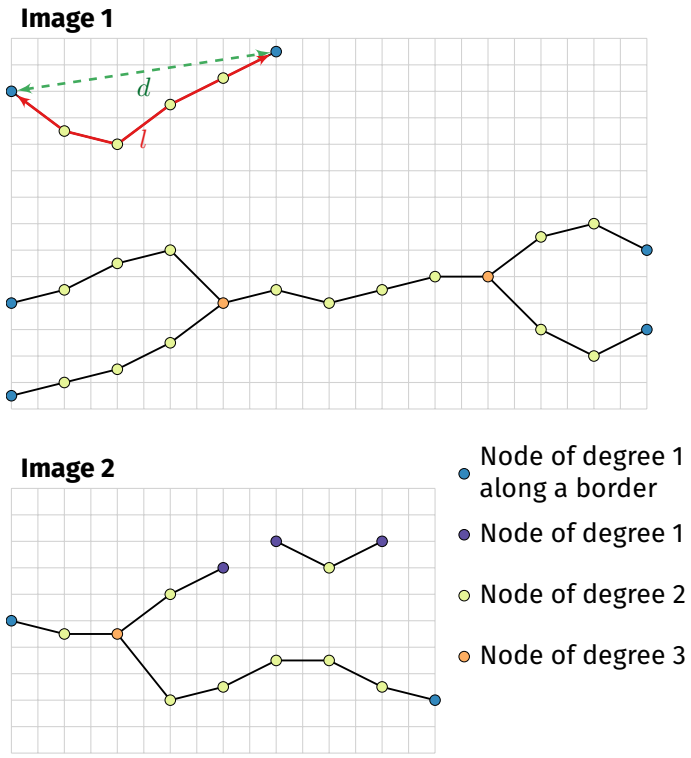


Figure 2.4 Example of skeletons for the connected components of the figure 2.3. If image 2 has no connection with a degree higher than 3, such as image 1, it displays some dead-ends highlighted by nodes with a degree one in the middle of the grid. This difference emphasize mis-reproductions of the channel connectivity. In image 1, l represents the curvilinear length of the branch and d the distance between its extremities. These measures are used to compute the tortuosity, which can also differ between the images.

There are several ways to compute skeletons [e.g., Serra, 1983, Jain, 1989, Brandt and Algazi, 1992]. Two methods have been considered for this work:

- The thinning-based method is distinguished by Cornea et al. [2007] as the methods with the best control on the skeleton connectivity. The algorithm retained here was defined by Lee et al. [1994] and implemented in Gocad by Barthélemy and Collon-Drouaillet [2013].

- A second method is based on slicing the grid along a given axis. The grid is subdivided into parallel slices of a given thickness. On each slice

the connected components are computed and one node is assigned to each component. The nodes are then linked by computing the connected components over two adjacent slices. If two components from two slices form one connected component when the slices are combined, their nodes are linked. If they form several components, their nodes are not linked.

The thinning-based method appears to perform better in two dimensions than the slicing-based method. But in three dimensions it tends to generate many small-scale loops which perturb both the topology and the geometry. This gives indicator behaviors that are hard to predict or explain. The slicing does not necessarily capture those small-scale elements depending on the slice size. While the geometry of the components is less well reproduced, the topology is better captured. It may not skeletonized the components smaller than the slice thickness, but this is compensated in some way by the other indicators, such as the number of component cells or the unit component proportion, which notice the presence of such small components. The thinning tends to generate skeletons with many nodes, which are heavy to manipulate. The slicing-based method has the same issues but far more limited when using quite high slice thicknesses. This aspect can be essential when dealing with several hundreds of images. Only the slicing method has been kept for this work, due to its ability to better capture the topology and to limit the skeleton size.

Node degree proportion The node degree proportion is based on the number of segments connected to a given node (figure 2.4). It is related to the topology of the skeleton and gives access to a deeper study of the connectivity of the sedimentary bodies of interest. The proportion p^n of nodes of degree n over all the skeletons for a given facies f is expressed as:

$$p_f^n = \frac{o_f^n}{o_f} \tag{2.10}$$

with o_f^n the number of nodes of degree n for all the skeletons of the connected components of the facies f and o_f the total number of nodes of the skeletons of the connected components for the facies f. Here we separate into two different classes the nodes of degree one along a border of an image and those in the middle of the image. The first ones relate to component terminations due to the limited size of the image. The other ones may relate to a bad body reproduction, for instance with channels that should be continuous and without dead-ends. The proportions for the different degrees give a histogram of node degrees.

Branch inverse tortuosity A branch is a part of a skeleton defined as an ensemble of segments that link nodes of degree 2 and delimited by two nodes of degree different from 2 (figure 2.4). The inverse tortuosity t quantifies how twisted a branch b is, with values ranged between 0 and 1. This measure is related to the geometry of the skeleton. It is expressed for a branch b belonging to a skeleton s of a component of a given facies as:

$$t_s^b = \frac{d_s^b}{l_s^b} \tag{2.11}$$

with l_s^b the curvilinear length of the branch b of the skeleton s and d_s^b the Euclidean distance between the two extremity nodes of the same branch. The values for all the branches of all the skeletons related to an image are averaged to obtain one indicator value for that image.

2.3 Quality analysis considerations

The final purpose of this work is to compare the quality of realizations. The indicators are thus computed on the realizations, which may come from different methods and/or parameters. Then dissimilarity values based on the indicators help to compare the realizations.

2.3.1 Influence of different grid dimensions

It is common to compare realizations having different grid dimensions. For instance in MPS, the training image is often chosen larger than the simulation grid to maximize pattern repeatability. Consequently, the grid dimensions influence the size of the traversing connected components. Channels are a perfect example of such continuous bodies traversing the grid completely. This impacts some indicators, such as the density and the number of component cells. A direct use of those indicators may be detrimental to the quality analysis and a workaround is required.

Correction for the connected component density

When some connected components of a facies are traversing (see section 2.2.2), reducing the grid dimension along the component traversing extension can change the component size but not necessarily their number. For instance

changing the grid size along the channel orientation does not change the number of channels. But it changes the grid volume and so does the density. Comparing two images of different sizes can affect the density whereas there is nothing wrong about the quality.

To overcome this aspect, the connected component density (formula 2.4) is replaced by:

$$\epsilon_f = \frac{N_f}{\vartheta_t} \tag{2.12}$$

with N_f the number of connected components of the facies f and ϑ_t is a corrected number of grid cells corresponding to a mix between the dimensions of each compared images:

$$\vartheta_t = \vartheta^x \times \vartheta^y \times \vartheta^z \tag{2.13}$$

with ϑ^x, ϑ^y and ϑ^z respectively the corrected number of cells along the axis x, y or z. These corrected numbers of cells are chosen depending on the smallest axis size among all the compared images and the size of the connected components. For an ensemble S of images to be compared, we note m^i the minimal size along the axis i over the images in S:

$$m^i = \min_{s \in S}(n_s^i) \;\; i \in \{x, y, z\} \tag{2.14}$$

with n_s^i the number of cells along the axis i for the image s. Determining the corrected numbers of axis cells for a given image r of S depends on the size of its connected components:

$$\vartheta^i = \begin{cases} m^i \text{ if } \exists\, g \in r,\ n_g^i \geq m^i \\ n_r^i \text{ otherwise} \end{cases} \quad i \in \{x, y, z\} \tag{2.15}$$

with g a connected component of r, n_r^i the number of cells along the axis i for the image r and n_g^i the number of cells along the axis i for the component g.

For instance in (figure 2.3), the component 1 of image 1 is larger than image 2 along the y axis. On the other hand, no component of that facies is larger than image 2 along the other axes. The corrected volumes for the image 1 $\vartheta_{t,1}$ and for the image 2 $\vartheta_{t,2}$ are so:

$$\begin{aligned} \vartheta_{t,1} &= n_1^x \times n_2^y \times n_1^z \\ \vartheta_{t,2} &= n_2^x \times n_2^y \times n_2^z \end{aligned} \tag{2.16}$$

with n_1^x, n_1^y and n_1^z the number of cells along the axis x, y and z for image 1 and n_2^x, n_2^y and n_2^z the number of cells along the axis x, y and z for image 2. Image 1 is so considered as truncated to limit the effect of the traversing bodies on the density.

Correction of the number of connected component cells

The objects are expected to have roughly the same range of sizes in all the images. But this expectation is no longer relevant in the case of traversing bodies (see section 2.2.2), whose dimensions must vary following the size of the supporting grid. As for the connected geobody density, the possible difference of image sizes imposes to rescale the number of cells of a connected component.

The rescaling is based on the estimation of the component bounding box size if this component was transfered within the grid of the smallest size. This bounding box size can be defined following a number of cells ϑ_b:

$$\vartheta_b = \vartheta^x \times \vartheta^y \times \vartheta^z \tag{2.17}$$

with ϑ^x, ϑ^y and ϑ^z respectively the corrected numbers of cells of the bounding box along the axis x, y or z. These corrected numbers of cells are following the same principles as to find the corrected number of cells for the density. Determining the corrected numbers of axis cells requires the same parameter m^i (formula 2.14):

$$\vartheta^i = \begin{cases} m^i \text{ if } n_b^i \geq m^i \\ n_b^i \text{ otherwise} \end{cases} \quad i \in \{x, y, z\} \tag{2.18}$$

with n_b^i the number of cells along the axis i for the bounding box of the component. Finally the corrected number of cells ϑ_f^g for the component g of the facies f is proportional to the ratio between the number of cells of its corrected bounding box ϑ_b and the number of cells of its actual bounding box n_b:

$$\vartheta_f^g = n_f^g \times \frac{\vartheta_b}{n_b} \tag{2.19}$$

For instance the component 1 on image 1 of figure 2.3 has a corrected number of cells ϑ_1 of:

$$\vartheta_1 = n_1 \times \frac{n_{b_1}^x \times n_2^y \times n_{b_1}^z}{n_{b_1}^x \times n_{b_1}^y \times n_{b_1}^z} = n_1 \times \frac{n_2^y}{n_{b_1}^y} \tag{2.20}$$

with n_1 the number of cells of component 1, $n_{b_1}^x$, $n_{b_1}^y$ and $n_{b_1}^z$ the numbers of cells of its axis-aligned bounding box along the axis x, y or z and n_2^x, n_2^y and n_2^z the numbers of cells of image 2 along the axis x, y or z. This reflects the impossibility to make that component fit into image 2. Its size needs so to be rescaled to be compared with objects of image 2.

The image dimensions has less impact on the other shape indicators. For instance when the size of a traversing object increases, so does its bounding box and the box ratio does not necessarily vary. Thus the corrected number of component cells is not used in the other shape indicator computation.

2.3.2 Indicator rescaling

The rescaling ensures that the differences between the ranges of indicator values will not affect the comparison. The histogram-based indicators – facies proportion, facies adjacency proportion and node-degree proportion – are not rescaled, to preserve their histogram behavior for the dissimilarity computation (section 2.3.3). Two methods can be used for rescaling: normalization and standardization.

The normalization method consists in rescaling linearly the indicators values between 0 and 1. The indicator I_i is the i$^{\text{th}}$ indicator of the set previously defined. When computed for the facies f of the realization r, we will denote the computed indicator I_{if}^r. The normalization is then obtained by rescaling it between its minimum and maximum values:

$$\text{norm}(I_{if}^r) = \frac{I_{if}^r - m_{if}}{M_{if} - m_{if}} \tag{2.21}$$

with M_{if} the maximum value for the same indicator and facies among all the images and m_{if} the minimal value for the same indicator and facies among all the images.

The standardization method consists in using reduced-centered indicator values. For an indicator i the standardized value for a facies f of a realization r is obtained using the following formula:

$$\text{stand}(I_{if}^r) = \frac{I_{if}^r - \mu_{if}}{\sigma_{if}} \tag{2.22}$$

with μ_{if} the mean for the same indicator and facies among all the images and σ_{if} the standard deviation for the same indicator and facies among all the images. The indicator normalization has been kept in this work to maximize the indicator variations between the images.

2.3.3 Dissimilarity calculation

The principle of comparing two images is to determine how dissimilar these images are. The indicators can be seen as coordinates of the compared images.

But these indicators are heterogeneous: they are either based on histograms or on continuous values. The computation of a dissimilarity value between two images requires a heterogeneous metric.

Following the example of Wilson and Martinez [1997], two different metrics are combined into a heterogeneous Euclidean/Jensen-Shannon metric. It uses the Jensen-Shannon distance, square root of the Jensen-Shannon divergence [Rao, 1987, Lin, 1991], for the histogram-based indicators – facies proportion, facies adjacency proportion and node-degree proportion – and the Euclidean distance for all the other indicators. The distance between two images r and s for a given indicator i of a given facies f is given by:

$$d(I^r_{if}, I^s_{if}) = \begin{cases} d_{JS}(I^r_{if}, I^s_{if}) \text{ if } I^r_{if} \text{ and } I^s_{if} \text{ are histograms} \\ d_E(I^r_{if}, I^s_{if}) \text{ if } I^r_{if} \text{ and } I^s_{if} \text{ are continuous values} \end{cases} \tag{2.23}$$

with I the indicator values. d_{JS} represents the Jensen-Shannon distance:

$$d_{JS}(H^r_i, H^s_i) = \sqrt{\frac{1}{2}\sum_{j=1}^{n}\left[H^r_{ij}\log\left(\frac{H^r_{ij}}{\frac{1}{2}(H^r_{ij}+H^s_{ij})}\right) + H^s_{ij}\log\left(\frac{H^s_{ij}}{\frac{1}{2}(H^r_{ij}+H^s_{ij})}\right)\right]} \tag{2.24}$$

with H^r_i and H^s_i the histograms of the indicator i for respectively the images r and s, n the number of classes for each histogram, H^r_{ij} and H^s_{ij} the proportions for the class j in the corresponding histograms. d_E represents the Euclidean distance used with rescaled indicators:

$$d_E(I^r_{if}, I^s_{if}) = \sqrt{(\text{resc}(I^r_{if}) - \text{resc}(I^s_{if}))^2} \tag{2.25}$$

with I^r_{if} and I^s_{if} the values of the indicator i for the facies f of respectively the images r and s and *resc* either *norm* (formula 2.21) or *stand* (formula 2.22). The final dissimilarity δ between two images r and s given their respective sets of indicators I^r and I^s is:

$$\delta(I^r, I^s, \omega, \nu) = \sqrt{\omega_1 d_{JS}(I^r_1, I^s_1, \nu)^2 + \sum_{i=2}^{12}\sum_{f=1}^{n}\omega_i\nu_f d(I^r_{if}, I^s_{if})^2} \tag{2.26}$$

with I^r_1 and I^s_1 the facies proportion histogram for the two images, I^r_{if} and I^s_{if} all the other indicator values depending on the indicator and the facies and

n the number of facies. ω represents the set of weights ω_i that control the impact of each indicator. ν represents the set of weights ν_f that control the impact of each facies. Note that the facies proportion histograms are the only indicators with one result for all the facies. Thus the facies proportions are treated differently from all the other indicators. The Jensen-Shannon distance used in that case is slightly modified:

$$d_{JS}(H_i^r, H_i^s, \nu) = \sqrt{\frac{1}{2}\sum_{f=1}^{n} \nu_f \left[H_{if}^r \log\left(\frac{H_{if}^r}{\frac{1}{2}(H_{if}^r + H_{if}^s)} \right) + H_{if}^s \log\left(\frac{H_{if}^s}{\frac{1}{2}(H_{if}^r + H_{if}^s)} \right) \right]} \quad (2.27)$$

The dissimilarity values computed by 2.26 between all the images constitute a non-negative symmetric matrix. This matrix has a zero diagonal corresponding to the dissimilarity between an image and itself. The dissimilarity matrix can be directly visualized with a heat map or treated by multidimensional scaling to get a more practical visualization.

2.3.4 Heat map

The heat map is a simple graphical representation of a matrix where the matrix values correspond to colors. In our case, the heat map is a two-dimensional representation. This colored representation highlights patterns in the dissimilarity matrix, either between realizations or between simulation methods. The main advantage of the heat map is to show the real dissimilarity values, contrary to the multidimensional scaling described in the next subsection.

The heat map also enables to classify the images and/or to apply clustering methods on it. A simple yet informative classification is the ranking according to the dissimilarities of the images toward one particular image. When using more advanced clustering methods, the matrix rows and columns are permuted to gather close values into the same cluster.

2.3.5 Multidimensional scaling

Multidimensional scaling (MDS) [e.g., Torgerson, 1952, 1958, see Cox and Cox, 1994 for a review] is a set of data visualization methods to explore dissimilarities between objects – represented by a dissimilarity matrix – through

a dimensionality reduction: it aims at producing a configuration of the objects as optimal as possible in a lower dimensional representation.

Principle and methods

Finding the configuration of the images in a k dimensional representation consists in locating a set of points representing the objects in a k-dimensional Euclidean space – with k being at most equal to the number of images minus one. The point localization is done so that the Euclidean distance d (2.28) between two points matches as closely as possible the dissimilarities between the images.

$$d_{r,s} = \sqrt{\sum_{i=1}^{k}(x_{ri} - x_{si})^2} \tag{2.28}$$

with r and s two images, k the dimension number of the Euclidean space, x_{ri} and x_{si} the coordinates of respectively r and s in the i-th dimension. Dissimilarities can be determined from data gathered on the objects.

Several multidimensional scaling methods have been proposed, depending on the type of dissimilarities and on the way to match the dissimilarities with the distances. Here, two metric multidimensional scaling methods are used: the classical scaling and the Scaling by MAjorizing a COmplicated Function (SMACOF).

Classical scaling Classical scaling [Torgerson, 1952, 1958, Gower, 1966] is a metric MDS method that considers the dissimilarities to be already Euclidean distances. The coordinates of the points corresponding to the images are so directly retrieved from that matrix. As the Euclidean assumption is very strict, it may be relaxed to a metric assumption. In that case, the dissimilarities must be distances. This is consistent with our dissimilarity based on a heterogeneous metric. To overcome the relaxation effects – in particular possible negative eigenvalues during the calculation of the point coordinates – two solutions can be considered [Cox and Cox, 1994]:

- Either choose the number of dimensions without taking into account those linked to negative eigenvalues.
- Either add a constant term to all the dissimilarity values except those between the same image. The purpose is to make the dissimilarities equal to Euclidean distances.

Those eigenvalues can also be used to assess the choice of the number of dimensions. The proportion of variation p^v of the k first eigenvalues – corresponding to the k chosen dimensions – compared with all the eigenvalues is given by [Cox and Cox, 1994]:

$$p^v = \frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{n} |\lambda_i|} \tag{2.29}$$

with λ_i the eigenvalues and n the total number of eigenvalues. It represents the weight of the k first dimensions on the distance values. The closest from 1 p^v is, the better is the chosen dimension number k. In the case of the metric assumption without using a constant term, the negative values may not be taken into account for the calculation of p^v.

Scaling by majorizing a complicated function (SMACOF) The SMACOF [De Leeuw, 1977, De Leeuw and Heiser, 1977, 1980] is also a metric MDS method, but it only considers that the dissimilarity matrix is metric and not Euclidean. Its goal is to get distances as close as possible from the dissimilarities using a majorization, i.e., the optimization of a given objective function called stress, through an iterative process.

Following the same principle as the impact of the eigenvalues with classical scaling, the final stress value helps to assess the choice of the number of dimensions: the lower the stress is, the better is that choice.

Validation of the number of dimensions

Following the chosen number of dimensions for the representation, the point configuration matches more or less the dissimilarity values. Verifying that the dimension number is enough for a good match between the dissimilarities and the distances is so of prime importance. Two approaches allow testing the chosen dimension number:

The scree plot: It represents the stress of the SMACOF against the dimension number. An equivalent representation for the classical scaling is the eigenvalue against the dimension number, besides looking at the proportion of variation. The stress and the eigenvalues both follow globally convex decreasing functions that tend toward 0 when the dimension number increases. A stress or an eigenvalue close or equal to zero means that the higher dimensions are unnecessary to represent the dissimilarities. The best number of dimensions is between the point with the highest flexion

of the curve and the beginning of the sill at zero. The dimension value right after the point with the highest flexion is generally enough for a decent representation.

The Shepard diagram: It represents the distances against the dissimilarities. The better the correlation, the better the choice of dimension number.

The dimension number k generally used is 2 or 3 for practical visualization purpose. Both MDS methods have been used on the following case study to analyze how good is the match between the distances and the dissimilarities for a two-dimensional representation and what is the impact of the MDS method on the resulting representation.

Estimation of the point position confidence

The point position confidence is another way to assess the MDS ability to represent the dissimilarities. For each point r, an error e highlights the mismatch between the dissimilarities δ and the distances d with all the other points s:

$$e_r = \sum_s |(a\delta_{r,s} + b) - d_{r,s}| \tag{2.30}$$

with a and b the linear regression coefficients found on the Shepard diagram. This measure gives a more local representation of the miss-representation than the scree plot or the Shepard diagram.

For visualization purpose, that error is then normalized, giving the confidence c for a given image r:

$$c_r = 1 - \frac{e_r - e_{min}}{e_{max} - e_{min}} \tag{2.31}$$

with e_{max} and e_{min} respectively the greatest and the lowest error values amongst the errors of all the images. This confidence can then be attributed to its corresponding point in the MDS representation through the point transparency: the less transparent the point is, the best the dissimilarities related to this point with all the other points are represented.

2.4 Application of the method and discussion

The method, as described in the previous sections, consists of three steps:

1. Indicator computation.

2. Dissimilarity computation in a matrix.

3. Dissimilarity visualization and analysis, especially with MDS.

The first two steps have been implemented in a C++ plugin for the SKUA-GOCAD geomodeling software [Paradigm, 2015]. The last step has been realized using the software environment for statistical computing R [R Core Team, 2012] with the addition of the R packages SMACOF [De Leeuw and Mair, 2009], NMF [Gaujoux and Seoighe, 2010] and ggplot2 [Wickham, 2009].

2.4.1 Dataset

The whole analysis process is applied to a synthetic dataset related to a channelized system. This case study aims to highlight some practical aspects of the method application. It uses some particular simulation methods but no general conclusion should be drawn on the simulation methods themselves based on this case.

The channelized system is composed of sandy channels with levees into a mudstone environment. This case falls within the scope of a MPS study: a training image (TI) (figure 2.5) is generated using the object-based method of the software Petrel [Schlumberger, 2015] (see A, table A.3, for the simulation parameters), and several simulation methods are used to try to reproduce the sedimentary bodies observed in the TI. The huge difference of grid size comes from the need of enough pattern repeatability for the MPS methods. The following realizations are simulated:

DeeSse: 100 MPS realizations generated with the DeeSse implementation [Straubhaar, 2011] of the direct sampling method [Mariethoz et al., 2010]. Contrary to more traditional MPS methods, the direct sampling bypasses the conditional probability computation and resamples randomly the training image. It relies on the compatibility – measured with a distance – between the conditioning data and the patterns scanned in the training image. The resampling step selects the first pattern with a distance lower than a given threshold. The training image is the TI and the set of parameters is given in table A.1) in the appendix A.

IMPALA: 100 MPS realizations generated with the method IMPALA [Straubhaar et al., 2011]. Contrary to the DeeSse, IMPALA still computes the conditional probabilities during the simulation. To improve the efficiency of this computation, the method stores the training image patterns in a list. The training image is scanned once at the beginning and the list is

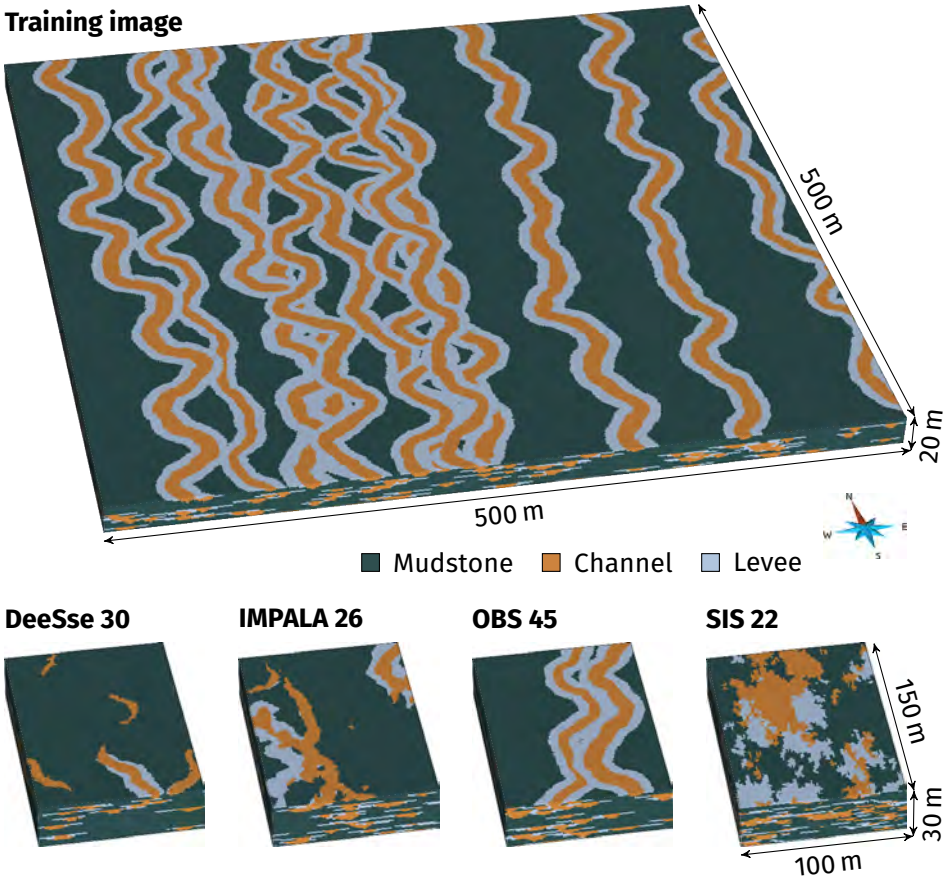


Figure 2.5 Training image and examples of realizations for each categories. These realizations are the closest to their category mean points in the MDS representation (figure 2.10).

used instead during the simulation. The training image is the TI and the set of parameters is given in table A.2) in the appendix A.

OBS: 100 object-based realizations generated with the same method and parameters used to generate the TI (A, table A.3).

SIS: 100 sequential indicator simulation realizations generated using variograms based on the facies in the TI (A, table A.4).

The purpose here is to compare the realizations with the TI. It should lead to retain the method and associated parameters that reproduce at best the

Table 2.1 Set of indicators used for the case study. The indicator descriptions are in section 2.2.

Set	Subset	Indicator	Symbol	Weight
Global analysis *(measures on all the components)*		Facies proportion	p	1
		Facies adjacency proportions	p^a	1
		Facies connection probability	Γ	1
		Corrected connected component density	ϵ	1
		Unit connected component proportion	p^u	1
		Traversing connected component proportion	p^c	1
Detailed analysis *(measures on each component)*	Shape indicators	Corrected number of connected component cells	n	1
		Box ratio	β	1
		Faces/cells ratio	ζ	1
		Sphericity	ϕ	1
	Skeleton indicators	Node degree proportions	p^n	1
		Branch inverse tortuosity	t	1

connectivity of the TI for the studied case. The indicators used in this case study (table 2.1) are all equally considered ($\omega_i = 1$ for all i in formula 2.26). This avoids any subjective bias that could arise from favoring a given indicator. The mudstone environment is the resultant of the channels and levees placement. It has so no precise shape by itself and may blur the analysis. It gets a weight of 0 while channels and levees both get each a weight of 1 ($\nu_{mudstone} = 0$, $\nu_{channel} = 1$ and $\nu_{levee} = 1$). Channels and levees are considered equally important to reproduce, but this aspect is related to the case study and could be further discussed. Slices of 17 cells along the grid axis with the same orientation than the channels are used for the skeletonization. The unit volume connected components (see section 2.2.2) give little information on the shape of the most important connected bodies considering flow circulation: the channels. But they can have a significant influence on the averaged shape indicators (section 2.2.3). They are so ignored during the computation of those shape indicators.

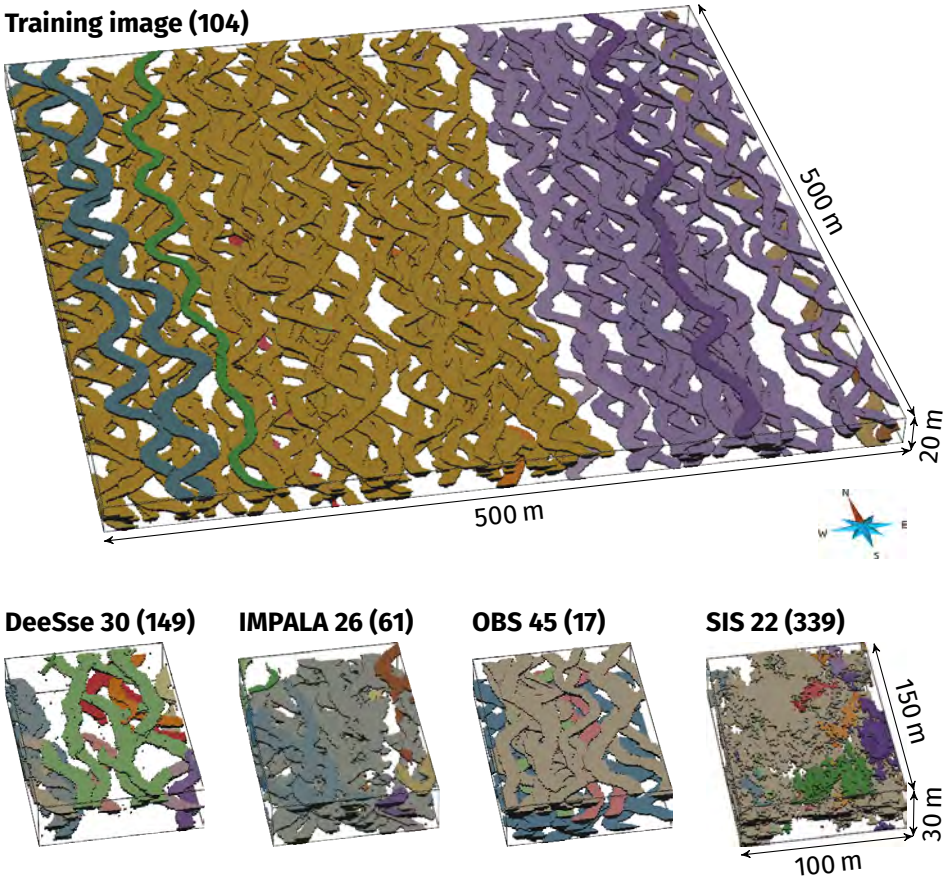


Figure 2.6 View of all the channel connected components within the reference image and examples of realizations for each categories. The number in parentheses are the number of connected components of each image.

2.4.2 Visual inspection of the realizations

Looking at the connected components (figure 2.6) highlights some expectations for the dissimilarity analysis. Two aspects must be analyzed: the reproduction of the channel shapes and the reproduction of the connectivity between the channels. In the studied case, the reproduction of the shape is pretty easy to analyze visually. The SIS realizations do not display any objects similar to channel/levee system and are so far dissimilar from the TI. The OBS realizations look similar to the TI, which is what is expected considering that they come from the same method and parameters. DeeSse realizations have objects

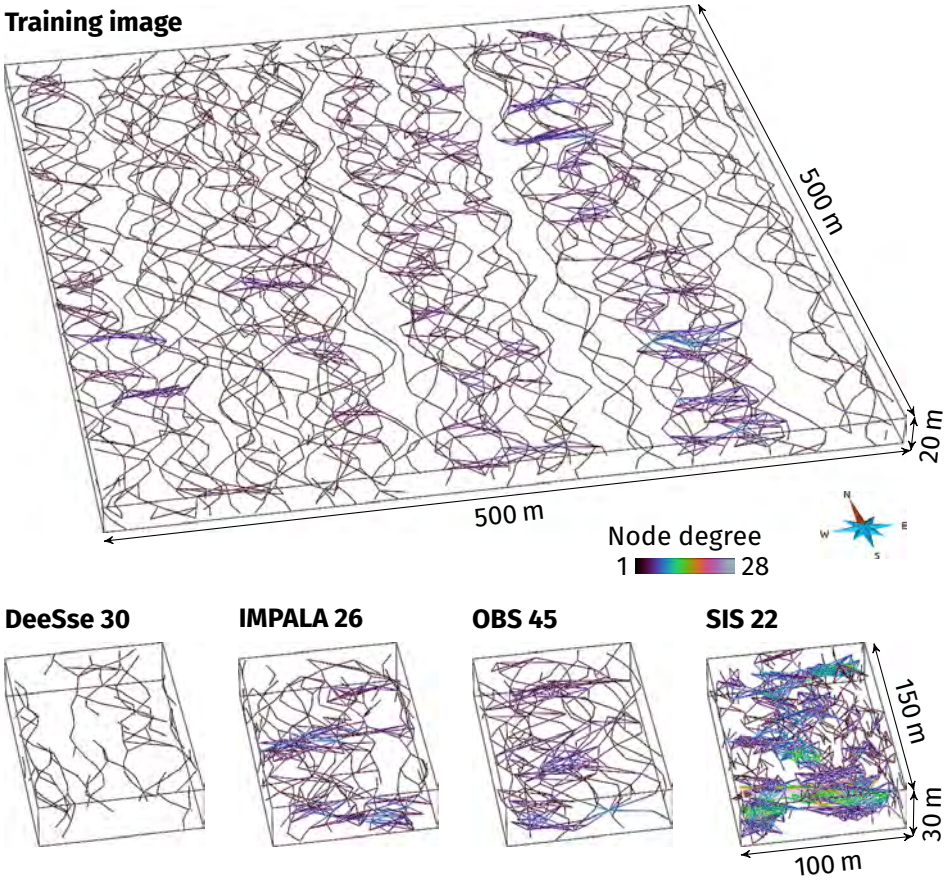


Figure 2.7 View of all the skeletons of the connected components for the TI and for a realization of each category.

similar to channels, even if some continuity issues appear. They also seem to have an insufficient number of channels. IMPALA realizations have quite linear objects but which poorly reproduce channel and levee shapes.

Estimating the connectivity of the objects in three-dimensional images is more difficult. The TI channels seem highly connected. The objects in the SIS realizations do not locally intersect like channels do and are far too connected. DeeSse realizations contain less objects and seem under-connected compared to the TI. The distinction between OBS and IMPALA realizations is difficult concerning the connectivity. Looking at the skeletons of the connected components (figure 2.7) corroborates those observations. DeeSse realizations

Table 2.2 Comparison of dissimilarities and distances between the TI and DeeSse realization 12 and 76 for to MDS methods. CS stands for classical scaling.

Compared images	Dissimilarity	CS distance	SMACOF distance
TI - DeeSse 12	1.773	0.661	1.577
TI - DeeSse 76	1.590	0.988	1.445
DeeSse 12 - DeeSse 76	0.962	0.329	0.213

are clearly under-connected compared with the other categories. SIS ones are over-connected. IMPALA realizations seem a bit more connected than OBS ones. The connectivity within the training image is clearly non-stationary.

2.4.3 About MDS methods and accuracy

Two MDS methods are tested with this case study: the classical scaling (figure 2.8) and the SMACOF (figure 2.9). The representations of each methods seem similar at first glance: the TI position is the only significant difference. But the point positions can vary significantly between the two representations. This is especially noticeable on the Shepard diagrams associated with both representations: if the correlation coefficients are rather good, the distances are sometimes far from being at least proportional to the dissimilarities. This can also lead to some misrepresentations of the dissimilarities. For instance, the realization DeeSse 76 should be closer to the TI than the realization DeeSse 12 (table 2.2). These two realizations are also more similar all together than with the TI. These dissimilarities are better reproduced by the SMACOF representation. The classical scaling even places the DeeSse realization 12 closer to the TI than the DeeSse realization 76. In this case using the classical scaling could lead to misleading interpretations.

The SMACOF method leads to a better representation as shown by the coefficient of determination r^2 of the Shepard diagram. This coefficient represents the goodness of fit of the linear regression. It measures so the correlation between the dissimilarities and the distances. The scree plot of the classical scaling also shows that a two-dimensional representation is not sufficient: a three-dimensional representation would perform better, and a four-dimensional representation would have been the best choice. For the SMACOF, a two dimensional representation seems enough. A three-dimensional one would have been better, but at the detriment of visibility. For visualization purpose a two-dimensional representation should always be favored. Several MDS methods

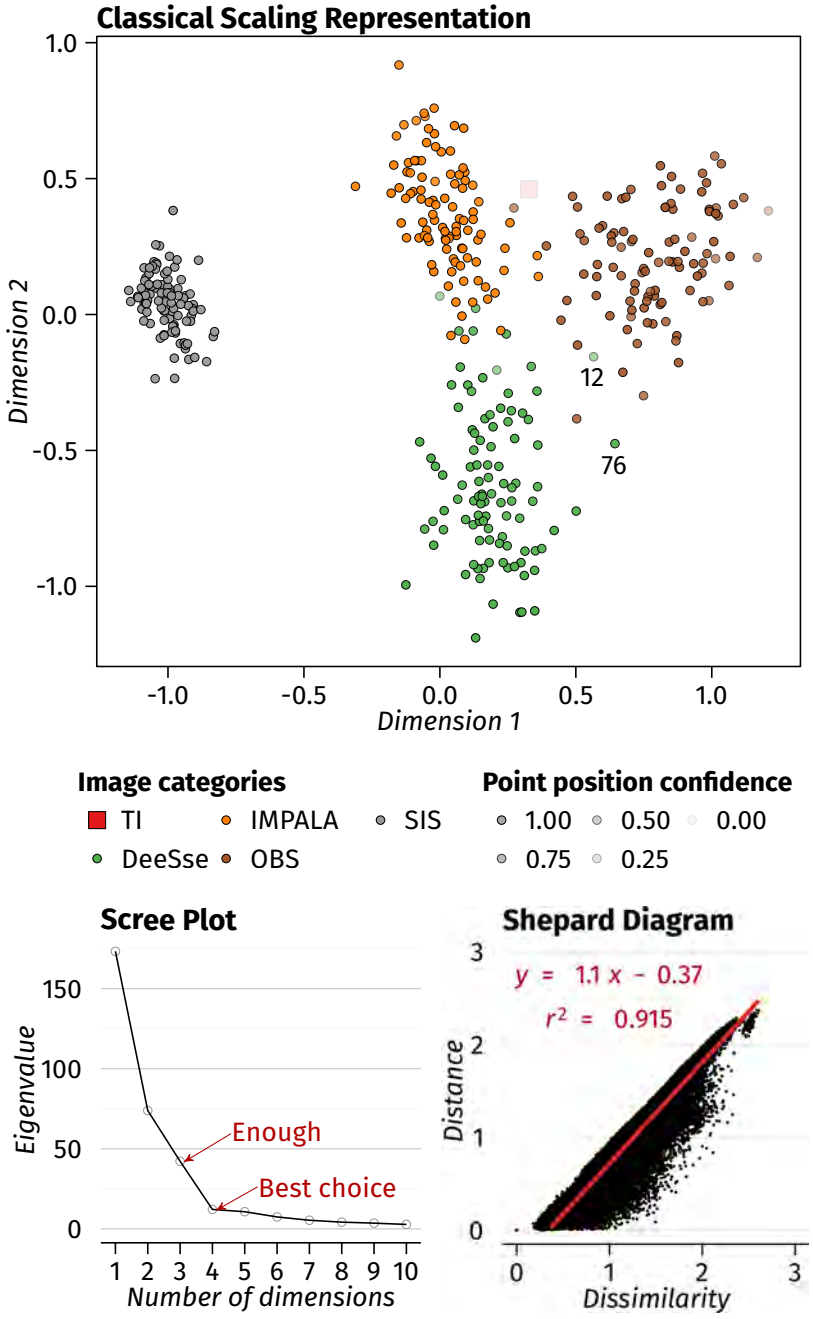


Figure 2.8 MDS representation of the dissimilarities between the images of the case study generated using classical scaling and validation graphs. The scree plot only displays the eigenvalues up to 10 dimensions on 400 possible.

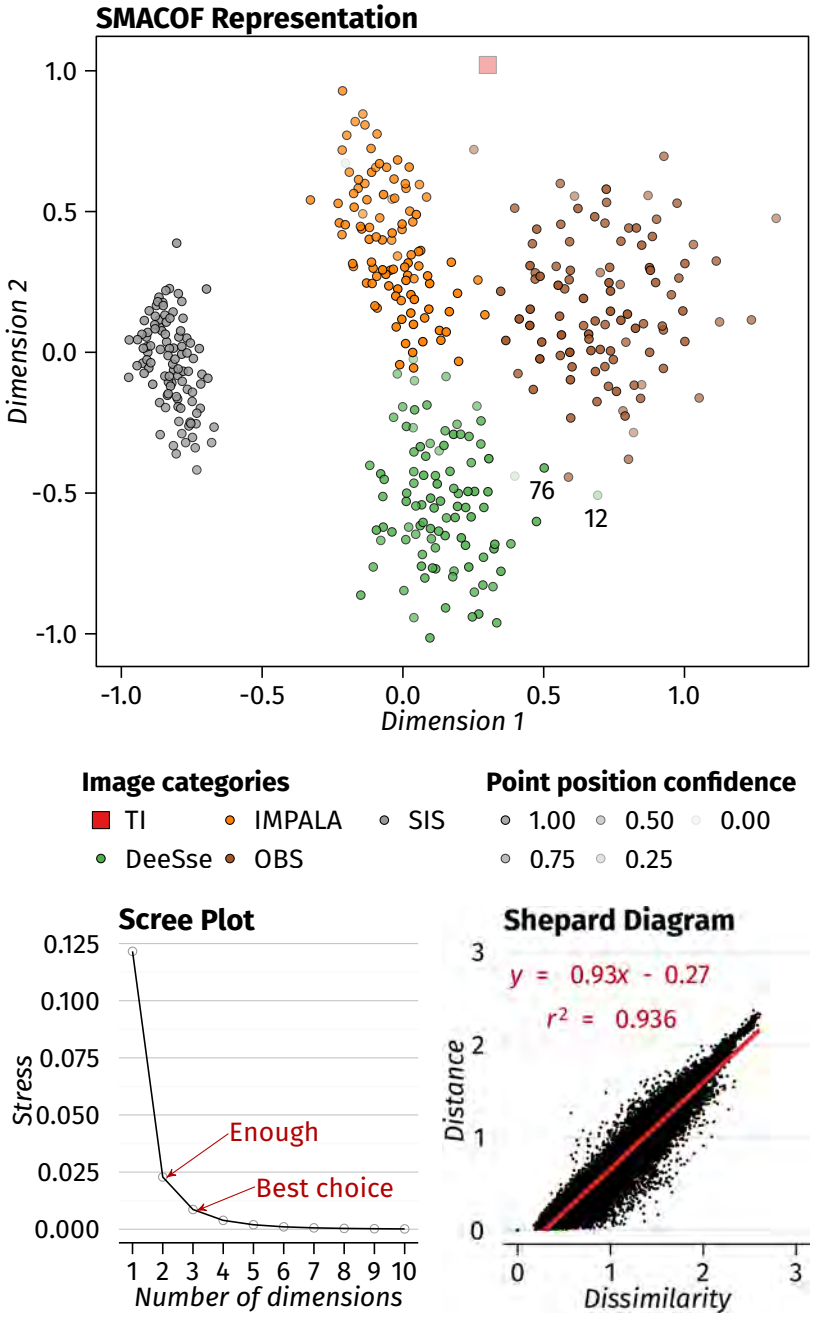


Figure 2.9 MDS representation of the dissimilarities between the images of the case study generated using SMACOF and validation graphs. The scree plot only displays the stress values up to 10 dimensions on 400 possible.

exist and they are not all equivalent for a given dissimilarity matrix. Testing different MDS methods could help to avoid using greater number of dimensions. In this case study, the two-dimensional representation given by the classical scaling is not sufficient. The SMACOF one is more precise, even if it does not represent perfectly the dissimilarities.

2.4.4 Analysis of the global dissimilarities

The first step of the dissimilarity analysis consists in looking at the global relationships between the realizations and the related methods. The MDS representation is ideal due to its clarity and ease of reading. But as highlighted in the previous section, caution should be exercised in its interpretation.

Inter-category dissimilarities

The MDS representation reveals a clear distinction between the different methods. The TI is quite far from the different categories, but seems closer to the IMPALA and OBS realizations. The SIS realizations are completely separated from the other categories.

A simplification of the MDS representation can also make the inter-category analysis clearer (figure 2.10). In that case, each category is represented by the point of mean coordinates and the convex hull of all its points. The proximity of the mean points and the eventual intersection of the convex hulls make the global analysis easier. Here the SIS realizations appear more clearly separated. DeeSse, IMPALA and OBS realizations are quite close. The overlapping between DeeSse and IMPALA seems significant, but it is mainly due to one realization. The IMPALA realizations are globally the closest from the TI, but some OBS realizations are closer. As the TI and the OBS realizations come from the same method and parameters, having the IMPALA realizations closer is quite surprising. This calls for a deeper analysis, especially considering the possible mismatches caused by the MDS.

Analyzing the heat map (figure 2.11) allows verifying the interpretations done on the MDS. The heat map is a more elegant way of analysis than examining a dissimilarity matrix which may contain several hundreds of values. It appears to be more difficult to read than a MDS representation. But the same conclusions can be made. The main point is to look at the squares formed by comparing two categories of images. When those squares are red, the categories are dissimilar. Each category is identifiable through the squares that appear on the colored matrix. The SIS realizations are clearly dissimilar to the

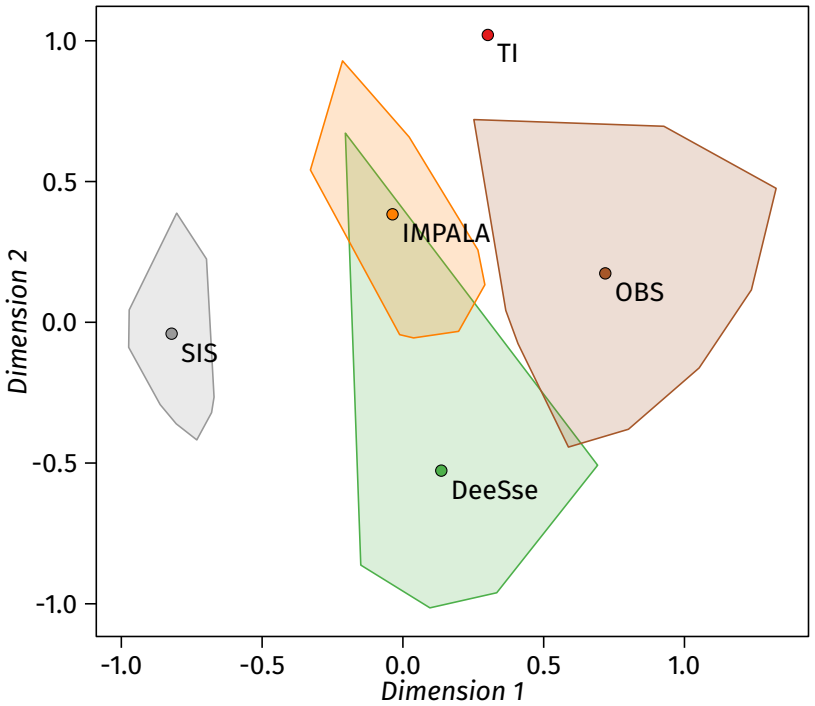


Figure 2.10 Simplification of the MDS representation of the dissimilarities between the images of the case study generated using SMACOF. The points are defined using the mean coordinates of all the MDS points of each category. The extension of each category is represented by the convex hull of all the MDS points of each category.

other categories, especially the OBS category. The OBS realizations are also distinguishable from the DeeSse and IMPALA realizations.

Intra-category dissimilarities

The same principles can be applied to compare the realizations within a category, so the realizations of a given method and a given set of parameters. Looking at the convex hull size provides a representation of how dissimilar the realizations of the category are, so how variable are the results of a given method. The realizations with less variability are the SIS ones, which all seem to have similar connectivity patterns. This is consistent with the great visual similarity of the SIS realizations. And the globular aspect of the SIS connected components enhances this similarity. Then come the IMPALA realizations, with a bit more of variability. The DeeSse and OBS categories have even larger

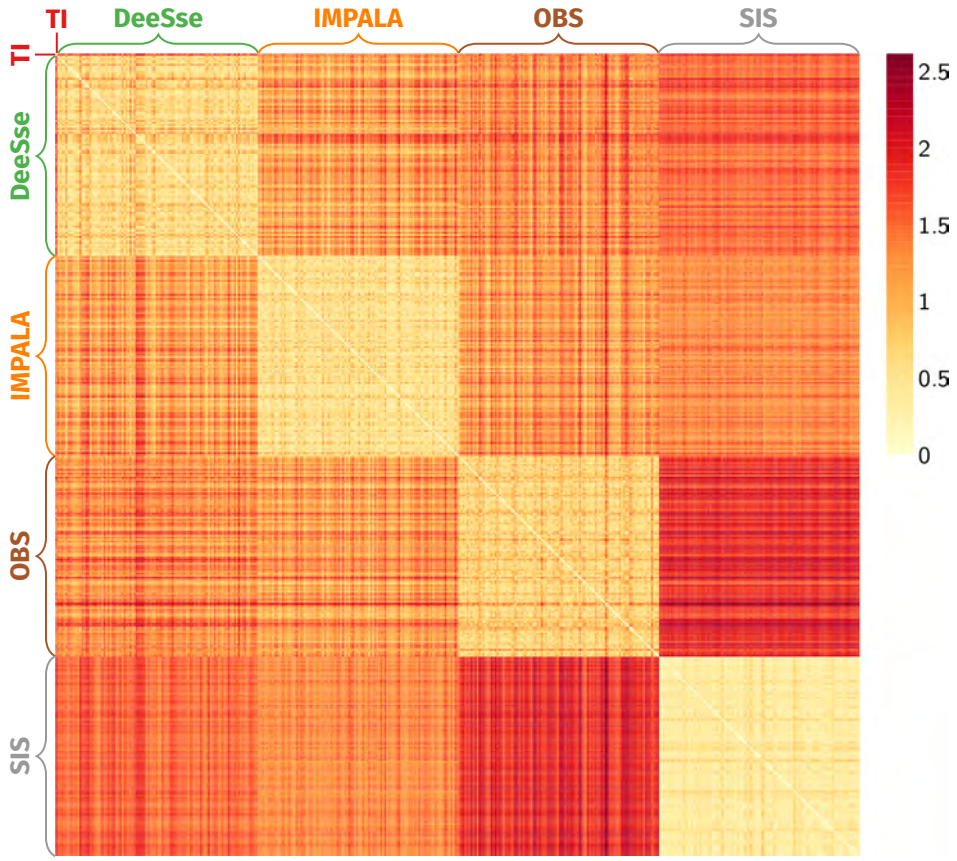


Figure 2.11 Heat map representation of the dissimilarity matrix computed based on the reference image and simulations of the case study. The color scale represents the dissimilarity values.

convex hulls, so a larger variability.

The same observations can be done on the heat map when looking at the color whiteness and uniformity of the squares along the diagonal. The SIS realizations appear highly similar all together and the IMPALA realizations more variable. The DeeSse and OBS realizations have again a greater variability, with quite high dissimilarity values.

About the difference in grid dimensions

One surprising aspect is the clear difference in the dissimilarities between the TI and the OBS realizations. All these images come yet from the same simulation method with the same parameter values, only the grid changes: the grid of the TI – the training grid – is larger than that of the OBS realizations – the simulation grid. The different grid dimensions could be an explanation for that clear difference:

- The TI displays a non-stationary connectivity pattern: some areas contain only one connected component as the channels are all connected, whereas other areas contain more connected bodies. This implies huge differences in terms of connectivity depending on the areas of the TI. When restrained to a smaller grid, this could give different results.

- The indicators most sensitive to variations of image size have already been adapted to take this aspect into account. However, this adaptation is rather an approximation and a bias may still exist in the comparison.

To check that the difference in the dissimilarities really comes from the difference of grid size, several TI analogs have been simulated. These analogs are realizations from the same method, parameter values (see A, table A.3) and grid than the TI. The application of the method highlights a significant difference between the realizations from different grids: the TI and analogs on one side, the OBS realizations on the other side (figure 2.12). Thus, the TI is not a realization with a particular connectivity pattern.

A closer look at the indicators is part of the answer. The TI, the analogs and the OBS realizations have all similar channel and levee proportions (figure 2.13). The channels and levees occupy the same volume inside the two grids. The facies connection probabilities for both channels and levees give a clearer difference between the realizations in the two grids (figure 2.13). The probability that two cells of the same facies belong to the same connected component is higher in the training grid than in the simulation grid. This is

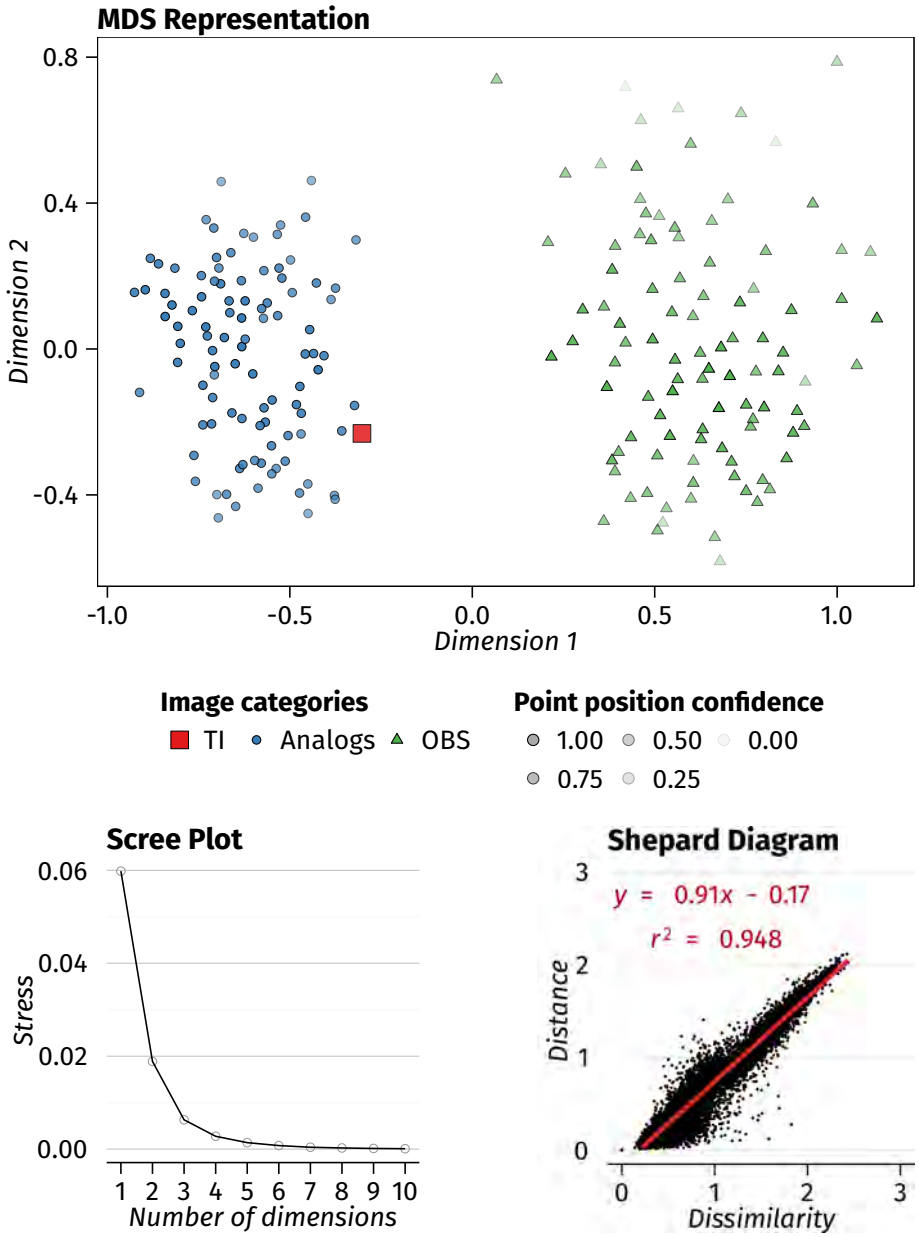


Figure 2.12 MDS representation of the dissimilarities between the TI, some TI analogs and the OBS realizations generated using SMACOF and validation graphs. The scree plot only displays the stress values up to 10 dimensions on 200 possible.

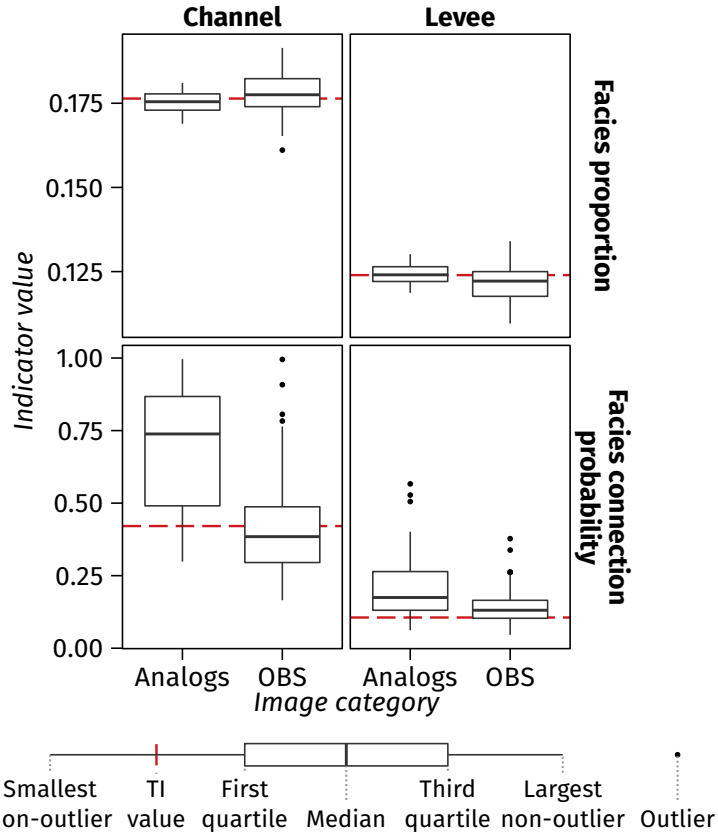


Figure 2.13 Box-plots comparing the facies proportions and facies connection probability for the TI, some TI analogs and the OBS realizations.

consistent with the difference of grid dimensions. When the grid dimension along the channel direction increases, the probability that two channels cross each other to form a single connected component increases too, especially here with sinuous channels. In such case, the characteristics of the connected components are impacted by the grid size.

Comparison with multiple-point histograms

Multiple-point histograms have made their way as indicators of a realization quality with MPS methods [e.g., Boisvert et al., 2010, Tan et al., 2014]. We propose here to compare the results obtained with those histograms to the previous results. The histograms are based on a $3 \times 3 \times 3$ pattern and are

computed on three levels of multi-grid [Tran, 1994], giving three histograms per image. The comparison between the images is based on the dissimilarity calculation presented in section 2.3.3. To compare two images, the Jensen-Shannon distance 2.24 is used between the histograms of each multi-grid levels. Then the Jensen-Shannon distances are assembled with a Euclidean norm, similarly to 2.26.

The SMACOF representation (figure 2.14) shows some similarities with the previous one (figure 2.9). The methods are all separated one from the other. The SIS realizations are all far from the other categories, even if their dissimilarities do not seem well represented. The OBS realizations do not mix much with the DeeSse and IMPALA realizations. Similarly, DeeSse and IMPALA realizations are quite close but do not really merge much.

Some differences appear in terms of variability. The SIS realizations seem far more variable and so does the IMPALA ones. But the most significant difference comes from the dissimilarity with the TI. The TI is inside the convex hull of the OBS realizations. These OBS realizations are so the closest to the TI. It is then hard to say which one of the DeeSse or IMPALA category is the closest.

All this highlights first differences between indicators. The indicator choice, if any, is not insignificant. It has an influence on the realization ranking. Further study of the indicator behavior is important to better characterize their measures.

2.4.5 Analysis of the dissimilarities with the conceptual model

Using MPS calls for a conceptual model of the bodies to simulate. The realizations have to reproduce these bodies and their characteristics, as the connectivity. This makes the quality analysis easier. Indeed, the analysis comes down to a single row of the dissimilarity matrix (or column due to the matrix symmetry). A MDS representation is not necessary any more and all the process can rest on the real dissimilarity values.

Here only the first row of the dissimilarity matrix is kept. It corresponds to the dissimilarities with the TI (figure 2.15, All the indicators). The IMPALA realizations are globally the closest to the TI; then come the OBS realizations that have a large dissimilarity range; follow the DeeSse realizations and finally the SIS realizations. This differs from the visual inspection of the component shapes: the OBS realizations should be the closest, followed by the MPS methods (DeeSse and IMPALA) and then by SIS realizations. When considering also connections between the objects in the visual inspection, having the DeeSse and SIS realizations as the furthest seems more consistent. This is coherent

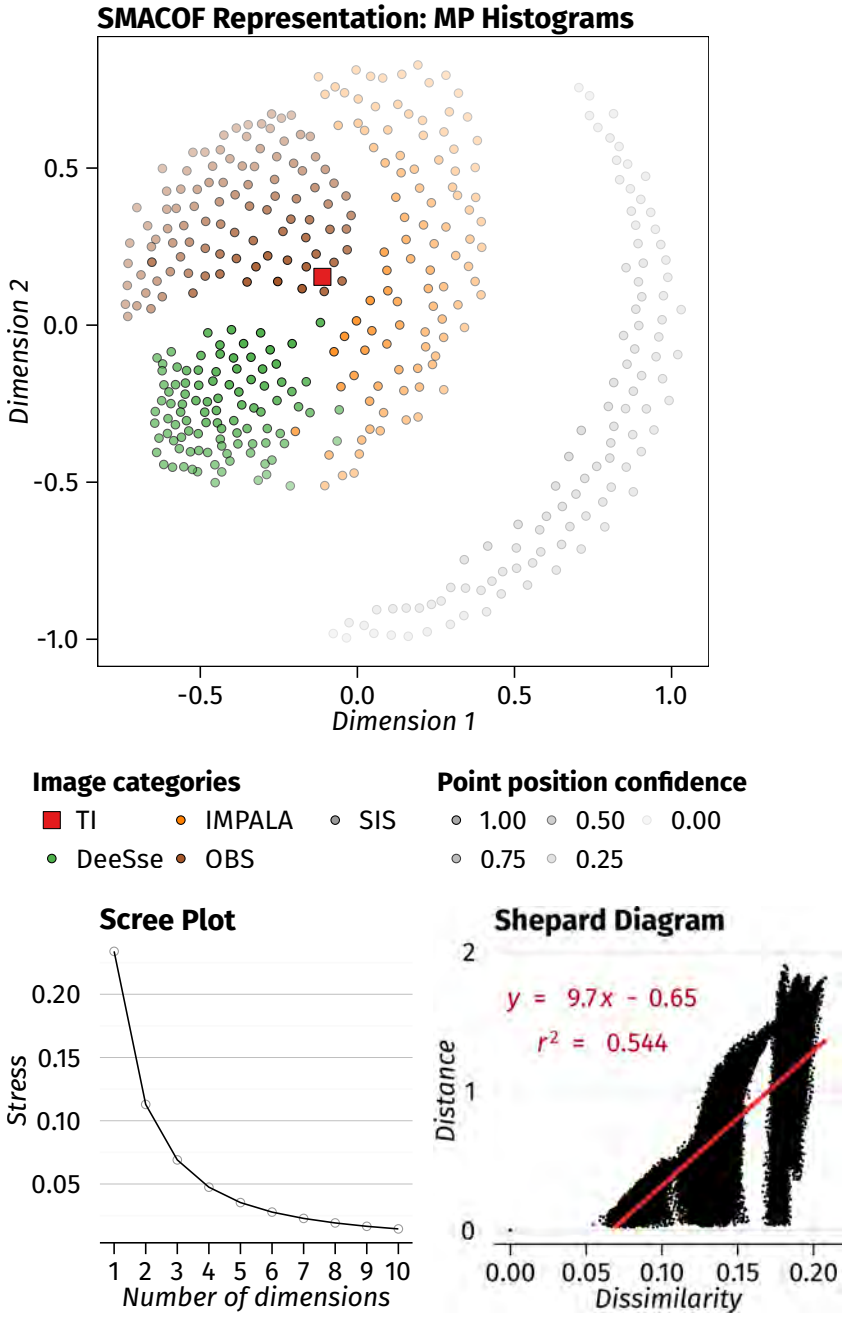


Figure 2.14 MDS representation of the dissimilarities between the images of the case study generated using SMACOF and validation graphs. The dissimilarities are based on the multiple-point (MP) histograms of the images. The scree plot only displays the stress values up to 10 dimensions on 400 possible.

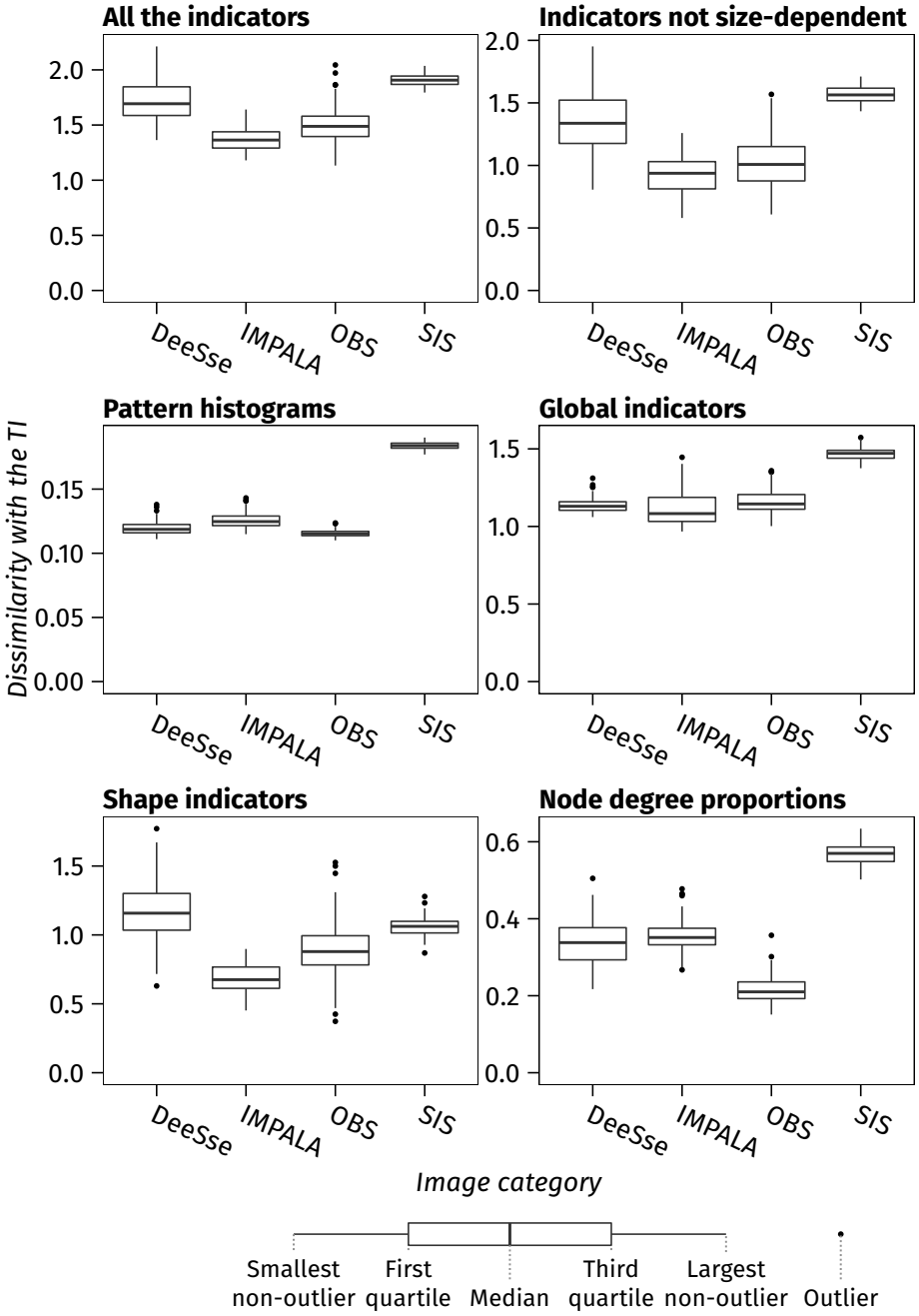


Figure 2.15 Box-plots comparing the realizations of each method with the TI. The dissimilarities depend on different indicators in each box-plot. The indicators not size-dependent are all the indicators except the component density (section 2.2.2) and the numbers of cells (section 2.2.3).

with the analysis previously done on the MDS representation. But a more detailed analysis is needed to understand the ranking.

The main advantage of such box-plot representation is to compare the ranking based on different indicators. The global (figure 2.15, Global indicators) and shape indicators (figure 2.15, Shape indicators) both have a ranking close to all the indicators. But the global indicators place the DeeSse realizations a bit closer to the TI than the OBS realizations. And the shape indicators place the SIS realizations closer to the TI than the DeeSse realizations. All those indicators measure globally the connected component distribution and shapes. On the other side, the node degree proportions (figure 2.15, Node degree proportions) have a ranking identical to the multiple-point histograms (figure 2.15, MP histograms): OBS realizations are the closest to the TI, followed by the DeeSse and the IMPALA realizations. The SIS realizations are the farthest and remain completely separated from the other categories. This ranking is more consistent with what can be observed visually. And those indicators are different from the other ones as they have a closer view at the connected components. The shape indicators analyze the components at a larger scale and do not see the shape variations in detail, contrary to the multiple-point histograms. On the other side the multiple-point histograms have no clear information on the connections between the patterns, so on the connectivity. The node degree proportions only look at the connectivity, not the shape. Their ranking is consistent with the one of the multiple-point histograms. But they separate more clearly the realizations than the multiple-point histograms.

The choice of the indicators and their combination have therefore a major influence on the analysis. Each indicator has its own way of seeing the sedimentary bodies in the realizations. Understanding the impact of various body organization on them is thus of prime importance, to see the kind of information they can provide, how they behave combined with other indicators and if this information and combinations are clearly relevant to analyze sedimentary bodies and their connectivity.

2.4.6 Indicator behavior

Indicators computed for the different methods are given in figures 2.16, 2.18 for the channels and in figures 2.17, 2.19 for the levees. For both channels and levees, the facies proportions and the facies adjacency proportions of IMPALA and OBS are closer to the TI than the other methods, with globally better results with the object-based method. OBS realizations are also interesting in terms of facies connection probability. Globally the connection probability is

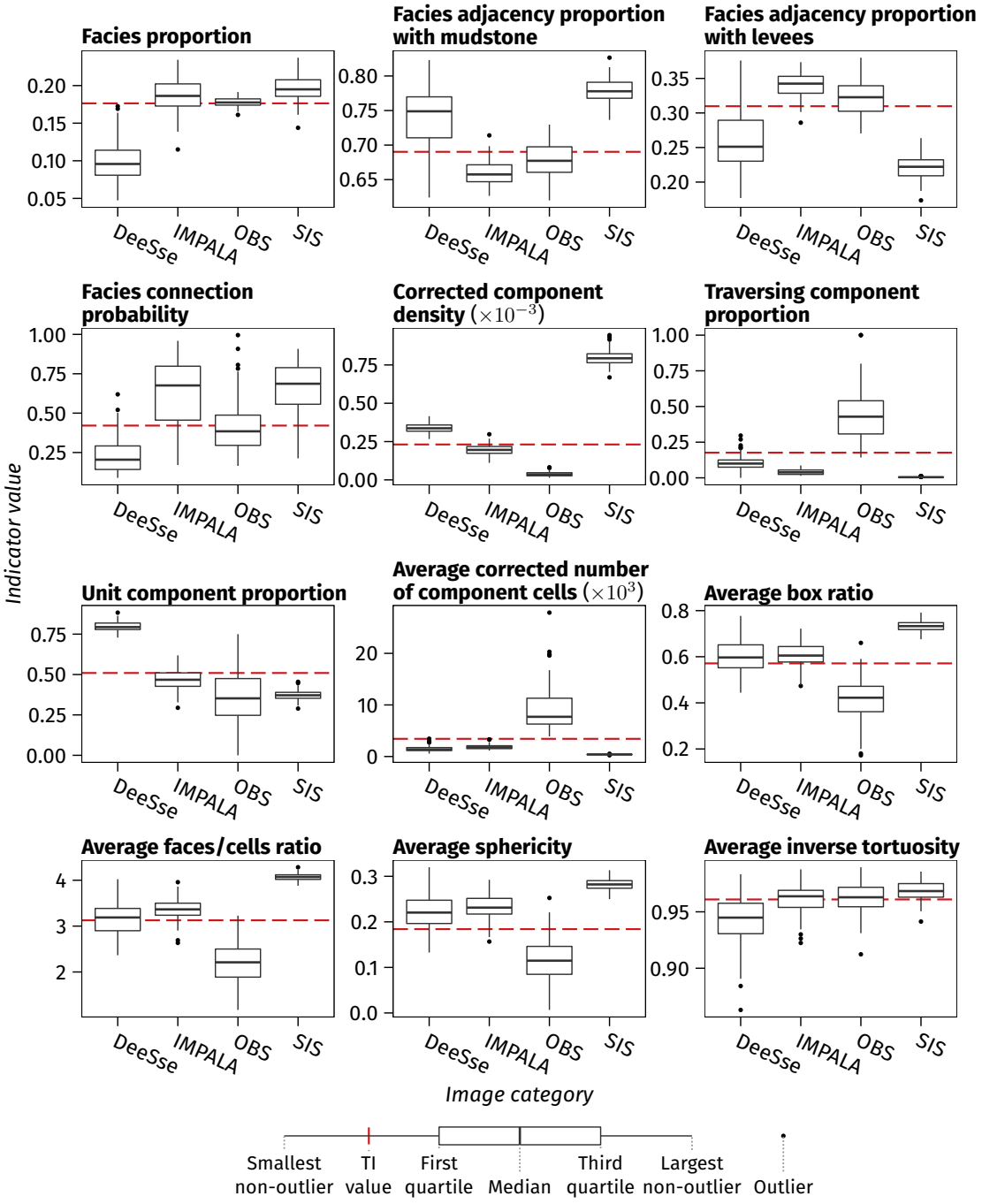


Figure 2.16 Box-plots comparing the range of indicators computed on the channels for the different methods, except the node degree proportions.

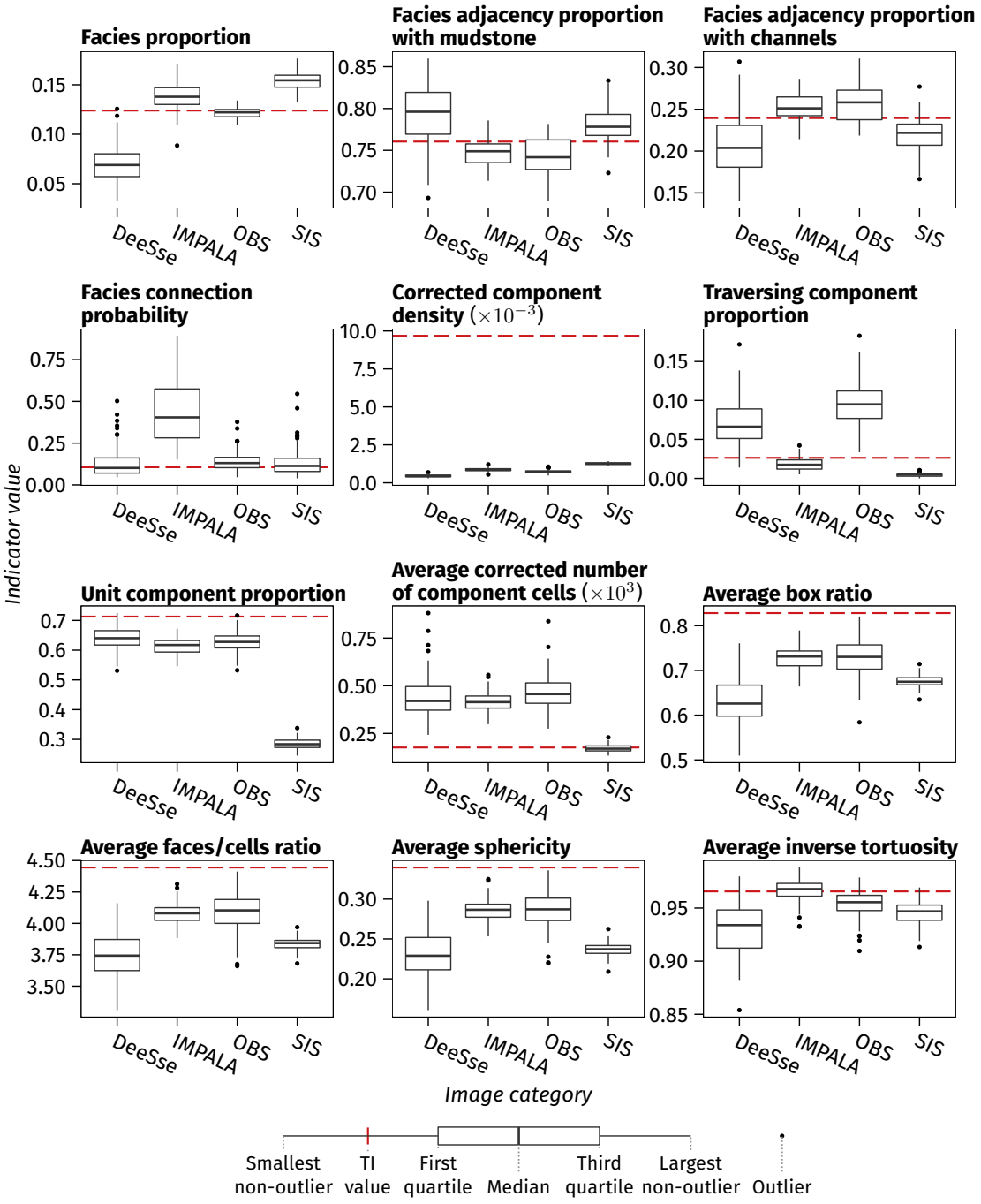


Figure 2.17 Box-plots comparing the range of indicators computed on the levees for the different methods, except the node degree proportions.

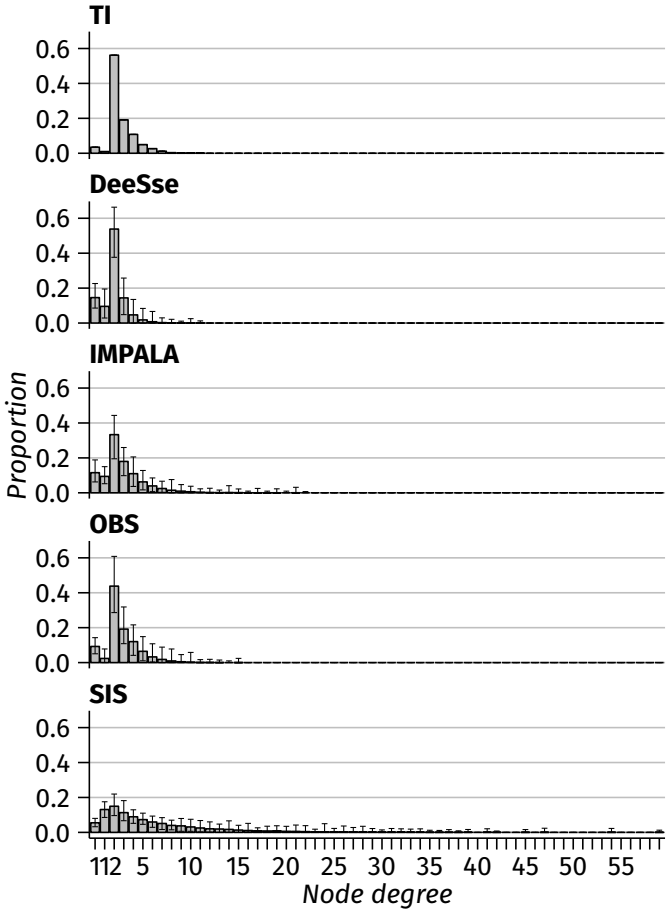


Figure 2.18 Mean node degree proportions of the channel skeletons for each category. The error bars display the minimum and maximum proportions. The first node degree 1 corresponds to the nodes of degree one along a grid border. The second node degree 1 corresponds to the nodes of degree one inside the grid.

consistent with the visual observations on the realizations and their skeletons: DeeSse realization components are clearly less connected than OBS ones, this is related to the poor reproduction of the channel proportion, and IMPALA and SIS realization components tend to be over-connected compared to the TI.

The other indicators related to the connected components display much contrasted behaviors. The corrected component density behavior differs following the facies. For the levees, all the methods are quite close all together but far from the TI. For the channels, the SIS realizations have a high density com-

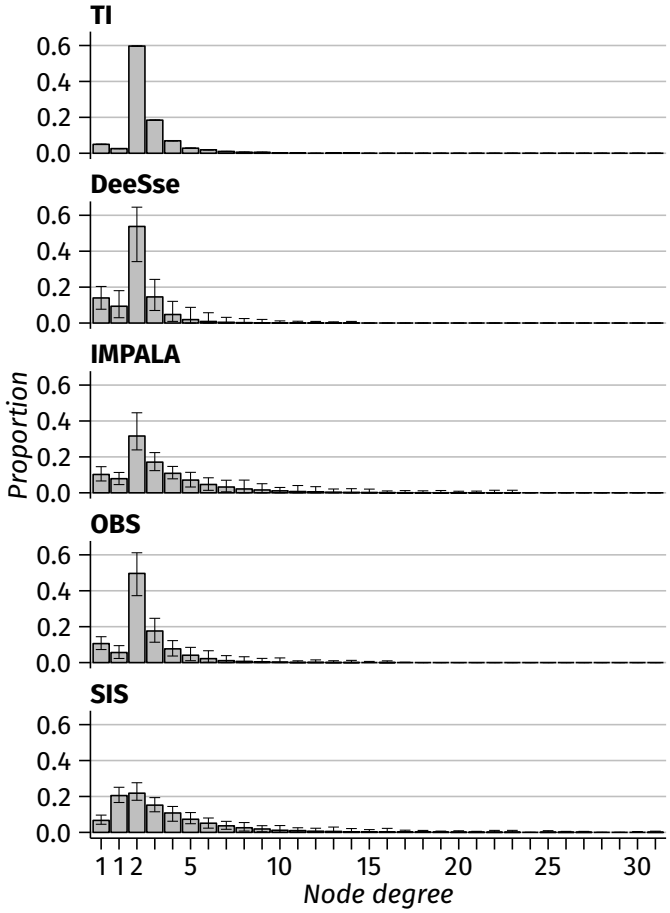


Figure 2.19 Mean node degree proportions of the levee skeletons for each category. The error bars display the minimum and maximum proportions. The first 1 node degree corresponds to the nodes of degree one along a grid border. The second 1 node degree corresponds to the nodes of degree one inside the grid.

pared with the other categories. DeeSse realizations follow them and OBS realizations have the smallest density. IMPALA realizations are the closest from the TI. This behavior can be linked with the number of component cells: the OBS realizations have the largest components, so at similar proportions a low density is consistent. SIS corrected number of component cells is quite low and their density is high, meaning that those realizations contain many small components. The high connection probability indicates in that case the presence of one huge component. This corresponds to the observations on the realizations.

The same observation can be made with IMPALA, but with a less clear behavior. The tendency of SIS and IMPALA realizations to generate small components also appears through the traversing component proportion: it is lower than for the other methods, so fewer components connecting two opposite borders of the image. OBS realizations are less dense, so fewer components relative to the occupied volume, and have relatively bigger components than the TI, with many traversing ones. Expecting a higher connection probability would be reasonable, but as said before the connection probability of OBS realizations is roughly similar to the TI value. This behavior can be related to the difference of image size in two ways:

- The density and the number of component cells for the TI are not the ones measured on the image: they are modified to take into account the difference of image size. This process is essential to be able to compare the indicators. However, it does not remove all the errors due to the difference of grid size. The indicator correction may be improved to lower the errors. But there will always be a doubt on the indicator accuracy.

- When the image size following the channel direction is increased, the probability to have channels hitting each other also increases, especially here with sinuous channels. The channels tend to form one connected component at the image scale, whereas at a smaller scale they can form different components. This aspect can be seen in the box ratio and in the sphericity. These two indicators are far less influenced by different grid dimensions than the density or the number of component cells. They are so not based on corrected values. However, they are smaller for the OBS realizations than for the TI. The sphericity highlights the more linear aspect of the OBS components, which are gathering fewer channels. The difference in box ratio must come from the fact that gathering two channels does not necessarily increase the bounding box compared to the channels taken individually. This tends to increase the box ratio, what explains the behavior of the TI compared to the OBS realizations.

One solution to limit the impact of different image size could be to extract one or several sub-images of the same size from each image. The comparison would then be made between the sub-images. However, it should be noticed that in this particular case the global ranking to the TI do not change much when not using the corrected density and numbers of component cells (figure 2.15, Indicators not size-dependent).

The unit component proportion for the channels appears to be high for all the

methods. The only clear distinction is from the DeeSse realizations, which have much more unit components than the other images. This is even clearer for the levees, with high proportions for all the categories except the SIS. It is useful to take into account that parameter in the calculation of the other indicators considering its potential influence. But it does not seem highly informative of an image quality. Using it as an indicator could be questioned considering the influence of such component on the fluid flow.

When choosing the skeletonization method and its parameters, we privileged the topology over the geometry. This means avoiding small-scale loops which perturb the node degree proportions. This choice has a direct impact on the inverse tortuosity values: if some variations appear between the categories, those variations are very low with values between 0.9 and 1. As 1 is a straight line, it appears that the skeleton branches are not very tortuous. This comes from the rather large slice thickness used to generate skeletons without small-scale loops. Using thiner slices could help to get more valuable information from the inverse tortuosity. But it could be at the detriment of the node degree proportions. Further work has to be done to improve the feature extraction by the skeletons.

Nevertheless, skeletons seem interesting objects to study the connectivity. Both facies have similar distributions of node degree proportions for all the categories (figure 2.18 and 2.19). SIS realizations have more connections than the other realizations, with far higher node degrees. In return this lowers the proportions of the smaller node degrees, especially the nodes of degree 2. The other categories have more similar distributions. The training image has numerous nodes of degree 2. DeeSse realizations have fewer connections than the TI, with less node of high degrees and more nodes of degree 1. IMPALA realizations have more connections, with fewer nodes of degree 2 and more high degree nodes. OBS realizations have the most similar distributions compared with the TI, as stated by the dissimilarity analysis (figure 2.15, Node degree proportions). This confirms the visual observations on the skeletons (figure 2.7). All the categories have fewer nodes of degree 2 than the TI and more nodes of degree one. The numbers of higher degree nodes vary following the category. At the end the connectivity seems indeed different from the TI for all the categories.

Compared to the previous indicators, the multiple-point histograms provide a more detailed analysis of the different object shapes in an image. Integrating it to the other indicators is so interesting. It would bring their ability to analyze object shape with more details than the shape indicators currently used and with less impact of different image sizes, at least from what can be seen on this

example. Node degree proportions could perfectly complete them to deeply analyze the connectivity and the reproduction at large-scale. Considering that, further studies are needed to understand the relationship between those histograms and the previous indicators, and in particular the shape indicators. This illustrates the difficulties to objectively analyze realization quality.

2.5 Conclusions and perspectives

This work develops an analysis process to compare several categorical images in terms of connectivity. The process relies on the computation of dissimilarities from a given set of indicators. These indicators are based on connected components. They cover various aspect of the realizations, from their facies proportions to the size of their components. Their analysis in the presented case highlights their ability to distinguish between realizations simulated with different methods.

The multidimensional scaling summarizes all this information through two-dimensional representations of the dissimilarities. It is a powerful and practical visualization tool to get a first idea of the relationships between the images. But it should be kept in mind that MDS representations provide only a partial view of the dissimilarities. Moreover, the representation and its quality are method-dependent. The MDS representation should so always be interpreted cautiously. Analyzing the dissimilarity matrix shall always be favored rather than analyzing the MDS representations. The heat map helps to have an easier check on the MDS interpretation. Then an analysis should be done on particular rows or column of the matrix depending on the studied cases. This gives access to simple but powerful visualization tools that work on the real dissimilarities.

The realization ranking when gathering the indicators is not what could be expected. But a detailed analysis shows the consistency of the indicator behaviors. Indicators of the connected component shape are interesting for their simplicity. But their sensitivity to the grid size is a real issue. A correction is proposed here, but it does not remove completely the error. It could be interesting to perform the analysis not on the images but on sub-images with the same dimensions. Multiple-point histograms also seem a valid option to analyze more locally the component shape. Skeleton-based indicators appear to be quite promising. The node degree proportions provide a simple way to compare connectivity structures. Further work should be done on the skeletons to better characterize both the connected component topology and geometry.

Thus, the present work shows that analyzing the behavior of various indica-

tors needs to be continued. More case studies should be considered. Channel are known to have an heterogeneous filling, with shale acting as a flow barrier. It could be interesting to model and work on the component formed by the reservoir facies rather than by the sedimentary objects they belong to. Starting with real data would also be beneficial.

Chapter 3

Stochastic simulation of channelized sedimentary bodies using a constrained L-system

Abstract Stochastically simulating realistic sedimentary bodies while conditioning all the available data is an important topic of research. In the particular case of channelized systems, that task is even more difficult due to the highly elongated shape of the channels. The classical dilemma is either simulating channel objects with a poor conditioning ability or simulating within a grid with a poor reproduction of the channel objects. We propose to focus the simulation of channel objects on data conditioning rather than reproducing perfectly their sinuosity. We rely on a formal grammar system, the Lindenmayer system or L-system, to stochastically simulate the channel morphologies resulting of the deposition processes. The L-system puts together channel segments based on user-defined rules and parameters, such as the element size or the angle between two consecutive elements. The succession of segments is then interpreted to generate non-rational uniform B-splines (NURBS) representing straight to highly meandering channels. Data are taken into account during the channel development using constraints. The constraints attract or repulse the channel from the data. They enable to condition various types of data, from well data to probability cubes or a confinement. The application to various synthetic cases highlights its ability to manage various data while preserving at best the channel morphology. Some improvements are still required to better condition the data at small-scale.

3.1 Introduction

Stochastic simulation methods rely on two main objectives: representing the geology as precisely as possible and honoring all the available data. Precisely representing the geological structures is essential to pursue the study of a reservoir. These structures may contain fluids or act as flow barriers. From this point of view their connectivity and its reproduction are fundamental. Channelized systems are a specific case of such structures. Their main constituting elements are highly continuous channelized sedimentary bodies – up to several hundreds of kilometers in turbiditic systems – often filled with sand. These channels are so potential reservoirs.

Representing channels causes several problems to stochastic simulation methods. Cell-based methods [e.g., Deutsch and Journel, 1992, Galli et al., 1994, Strebelle, 2002] attribute a category to each cell within a discretized domain, based on a prior model. Conditioning to well data is easy with such process, but most of the time the channel continuity is not preserved. This leads to a connectivity different from the prior model or from object-based methods (figure 3.1). These object-based methods [e.g., Viseur, 2001, Deutsch and Tran, 2002, Pyrcz et al., 2009] rely on defining a geometrical representation of the channels, with parameters such as their width or their wavelength. They preserve the channel continuity. However, data conditioning is difficult, contrary to cell-based methods. This is partly due to the highly elongated shape of channels and the poor flexibility of their representations.

Simulating channels requires a balance between some geological concepts, such as the channel elongated shape, and data conditioning (figure 3.2). Wells are a first and prime source of data. They help to locate the channels and may give some indication about the connectivity between the wells. Seismic data are less precise than well data, but they give an information at larger scales. They can provide sand probability cubes for instance. If the resolution is good enough, some objects can be directly interpreted on seismic data.

In numerical biology, tree and root simulation methods rely also on object-based approaches [e.g., Prusinkiewicz and Lindenmayer, 1996, Palubicki et al., 2009, Leitner et al., 2010]. One of their interesting developments is the consideration of the environment influence to improve the simulation realism: influence of gravity, influence of the sun light distribution or influence of the presence of other trees, etc. [e.g., Mvech and Prusinkiewicz, 1996, Streit et al., 2005, Taylor-Hell, 2005]. These methods are based on a formal grammar, the Lindenmayer system or L-system. It was first proposed as a formal description of the development of filamentous organisms [Lindenmayer, 1968]. Despite

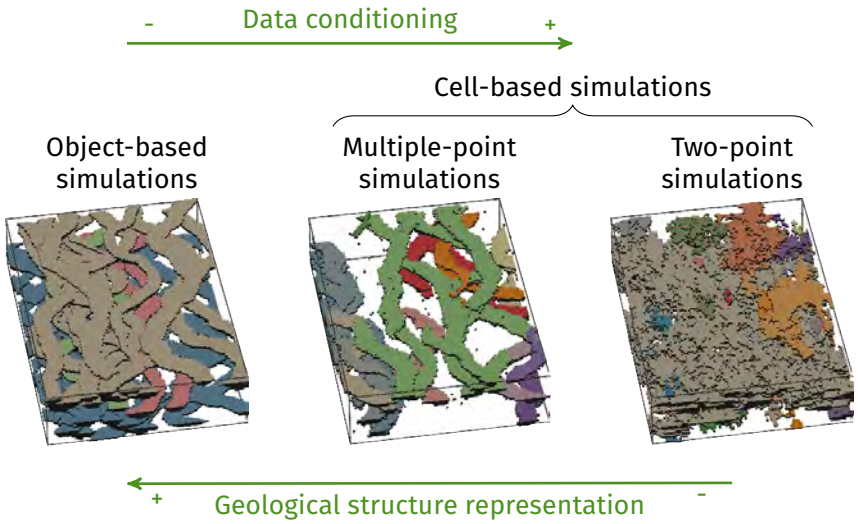


Figure 3.1 Types of stochastic simulation methods and relationship between geological structure representation and conditioning. The figures show the channels in realizations from different simulation methods, with the colors highlighting the connected components of the channels. These simulation methods have to model the same structures with priors from the same source. The resulting realizations differ significantly, in particular in term of connectivity.

their ability to integrate various data, very few works relate on formal grammars to simulate channels. Hill and Griffiths [2009] use an analog model to define some rules for a formal grammar system. These rules are the basis for stochastic simulation, in a process similar to multiple-point simulation. However data conditioning remains limited to sparse hard data, i.e., well data or channel segment interpreted from seismic data.

The present work aims at applying the principles of L-systems (section 3.2) to object-based approaches. The main idea is to simulate the morphology resulting from the flow processes forming channels. L-system rules simulate the channel morphology (figure 3.2), from straight to highly sinuous channels (section 3.3). During the application of the L-system rules, constraints influence the channel growth to condition to various data (section 3.4). The purpose of this research is to facilitate data conditioning as compared to other object-based approaches while keeping the main added-value of these approaches: the preservation of the channel shape. An illustration of the parameter impact and the addition of constraints to the system is based on simple cases (section 3.5). A more

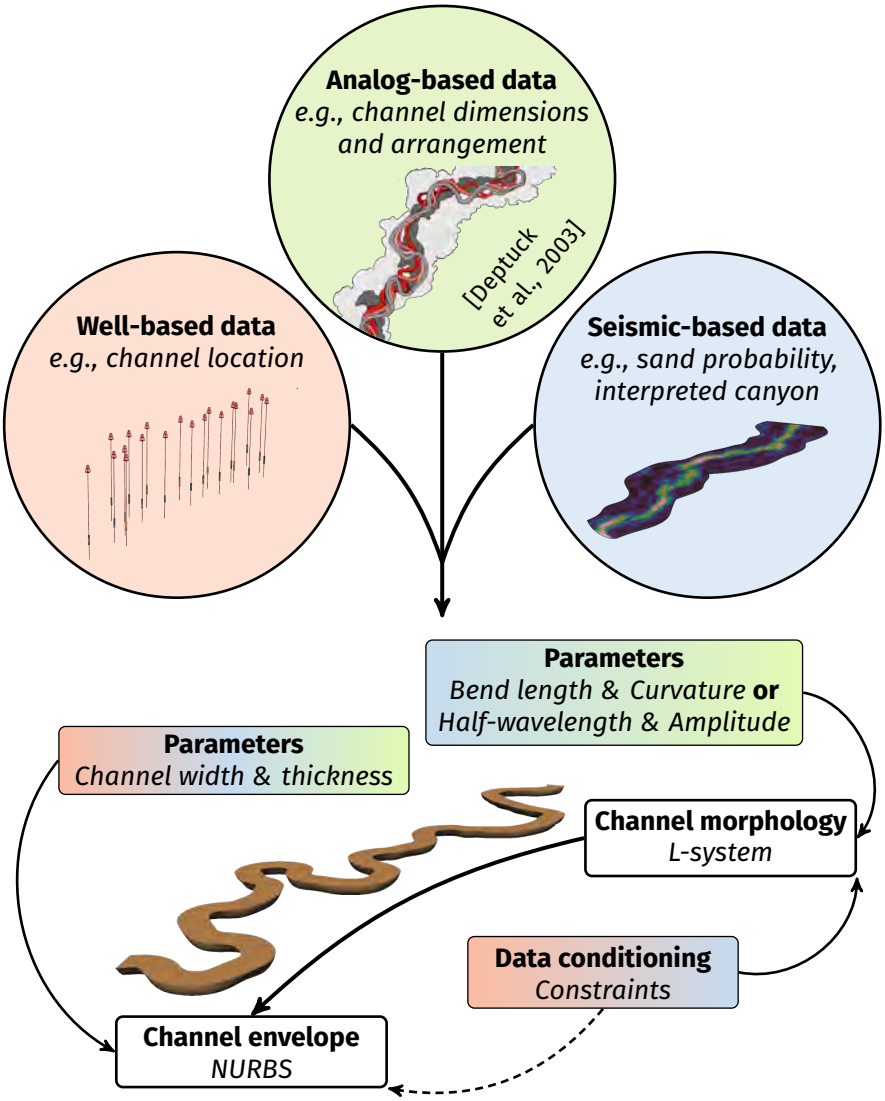


Figure 3.2 Workflow for channel simulation based on L-systems with data and parameter sources.

complete synthetic case study considers the simulation of turbiditic channels within a submarine canyon. It leads to a discussion on the integration of data of different types, such as well sedimentary data or facies probability cubes, in the simulation of turbiditic channels within a submarine canyon (section 3.6).

The whole method was implemented in C++ within the Gocad plug-in ConnectO. The channel envelopes with NURBS were implemented by Jérémy Ruiu within the Gocad plug-in GoNURBS [Ruiu et al., 2015b].

3.2 L-system principles

As formal grammar systems, L-systems rely on string rewriting. Letters from an alphabet are tied together by production rules to form and change strings. For an exhaustive review about L-systems, see Prusinkiewicz and Lindenmayer [1996].

3.2.1 Basic L-systems

A L-system starts with an initial string called the axiom (ω). A set of rules, called production rules ($p_i, i \in [1, n]$, n being the number of rules), controls the rewriting. The rules are first applied on the axiom and then on its substitutes, following a certain number of iterations, the order. In the most simple type of L-system, the rules are made of a predecessor and a successor:

$$predecessor \longrightarrow successor$$

The principle is to scan the axiom or its substitute. Each time a predecessor is found, its corresponding successor replaces it. For instance with the following system:

$$\begin{array}{lcll} \omega & : & b & \\ p_1 & : & a & \longrightarrow \quad ab \\ p_2 & : & b & \longrightarrow \quad a \end{array}$$

The evolution of the system for an order of five is:

$$\begin{array}{lcl} 0 & : & b \\ 1 & : & a \\ 2 & : & ab \\ 3 & : & aba \\ 4 & : & abaab \\ 5 & : & abaababa \end{array}$$

In the previous system, both rules are applied at each iteration until the order is reached. Another possibility is to express directly which rule can be used at

Table 3.1 Basic L-system alphabet and corresponding geometrical interpretation.

Letters	Interpretation
F	Move forward of a length l_s and draw a segment between the new position and the former one
f	Move forward of a length l_s without drawing a segment
$+$	Turn left by an angle α based on the rotation matrix $R(\hat{U}, \alpha)$
$-$	Turn right by an angle α based on the rotation matrix $R(\hat{U}, -\alpha)$

which iteration, giving here:

$$\begin{array}{lcll} \omega & : & b_0 & \\ p_1 & : & a_i \longrightarrow a_{i+1}b_{i+1} & 0 \leq i \leq 4 \\ p_2 & : & b_i \longrightarrow a_{i+1} & 0 \leq i \leq 4 \end{array}$$

Compared to other formal grammar systems, the L-system is parallel and not sequential: at a given iteration, the rules are applied simultaneously on all the letters.

3.2.2 Classical alphabet and geometrical interpretation

The purpose of the geometrical interpretation is to graphically represent the string resulting from the rewriting. At each letter of the alphabet corresponds an action that drives the string representation. Several interpretation methods exist, the most famous being the turtle interpretation [Prusinkiewicz, 1986].

The turtle is represented by a position vector $\vec{P}$ and an orthonormal coordinate system centered on the turtle position (figure 3.3). Three orientation vectors compose that system:

- $\hat{H}$ is the direction along which the turtle moves forward.
- $\hat{L}$ is the direction to the left for the turtle.
- $\hat{U}$ is the up direction for the turtle.

The position and the three vectors of the turtle's coordinate system form the turtle's state.

Two operations change the turtle's state:

- Moving forward of a given length l_s.

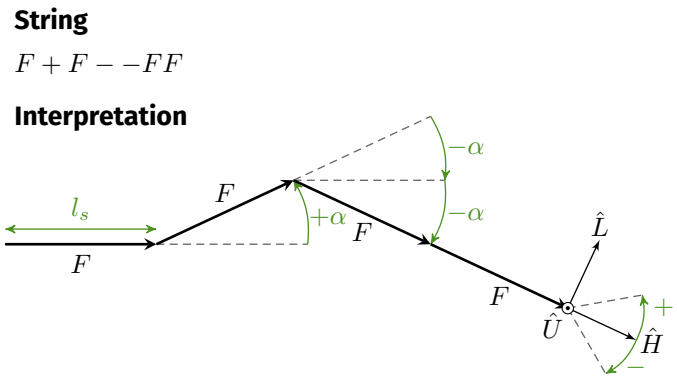


Figure 3.3 Example of string and its geometrical interpretation. l_s is a segment length and α the angle between to successive segments. The letter interpretations are given in table 3.1. The vectors $\hat{H}$, $\hat{L}$ and $\hat{U}$ are represented at the current turtle position $\vec{P}$. They control the further displacement of the turtle.

- Rotating of a given angle α following a given axis represented by $\hat{H}$, $\hat{L}$ or $\hat{U}$.

The rotation is based on the following update of the vectors:

$$\left[\hat{H}'\ \hat{L}'\ \hat{U}'\right] = \left[\hat{H}\ \hat{L}\ \hat{U}\right] R(\hat{V}, \alpha) \tag{3.1}$$

where $\hat{H}$, $\hat{L}$ and $\hat{U}$ are the current vectors, $\hat{H}'$, $\hat{L}'$ and $\hat{U}'$ are the updated vectors, $\hat{V}$ is $\hat{H}$, $\hat{L}$ or $\hat{U}$ depending on the wanted rotation axis, α is the rotation angle in radian and R is the rotation matrix:

$$R(\hat{V}, \alpha) = \begin{pmatrix} \cos\alpha + V_x^2(1-\cos\alpha) & V_xV_y(1-\cos\alpha) - V_z\sin\alpha & V_xV_z(1-\cos\alpha) + V_y\sin\alpha \\ V_yV_x(1-\cos\alpha) + V_z\sin\alpha & \cos\alpha + V_y^2(1-\cos\alpha) & V_yV_z(1-\cos\alpha) - V_x\sin\alpha \\ V_zV_x(1-\cos\alpha) - V_y\sin\alpha & V_zV_y(1-\cos\alpha) + V_x\sin\alpha & \cos\alpha + V_z^2(1-\cos\alpha) \end{pmatrix} \tag{3.2}$$

with $\hat{V} = (V_x, V_y, V_z)$. Then, the turtle position $\vec{P}$ is updated to its new position $\vec{P'}$ by taking into account the wanted length l_s:

$$\vec{P'} = \vec{P} + l_s \hat{H'} \tag{3.3}$$

A letter of the alphabet can then change the turtle's state based on those two operations, i.e., moving forward and rotating. It is also possible to give a tangible form to the interpretation by generating an object segment (figure 3.3). The main symbols of the classical L-system alphabet and their interpretation are given in table 3.1. Other symbols exist for pitching based on a rotation around $\hat{L}$ and rolling based on a rotation around $\hat{H}$.

3.2.3 Stochastic L-systems

Basic L-systems produce deterministic systems. There are two ways of adding stochasticity to obtain different strings and interpretation from one system:

- Randomizing the L-system parameters, especially those of the geometric interpretation, i.e., the angle α and the length l_s.
- Randomizing the production rule selection:

  $$predecessor \xrightarrow{p} successor$$

  with $p \in [0, 1]$ the rule probability for the rule drawing.

Whereas the first way affects only the geometry of the system, the second one can affect both the geometry and the topology. In the following example, the rules p_1 and p_2 are not both applied at each iteration. Only one of them is chosen each time following a random draw based on the given probabilities:

$$\begin{array}{lcll} \omega & : & b & \\ p_1 & : & a & \xrightarrow{0.75} \quad ab \\ p_2 & : & a & \xrightarrow{0.25} \quad bbb \\ p_3 & : & b & \longrightarrow \quad a \end{array}$$

3.2.4 Context-sensitive L-systems

Context-sensitive L-systems give the possibility to look at the context of the predecessor. When looking at a rule application, the string has not only to match the predecessor, but also its context. This context is separated in two

parts: the left context corresponds to the letters that should be just before the predecessor and the right context to the letters that should be just after:

$$left\ context < predecessor > right\ context \longrightarrow successor$$

The context can be one-sided if needed, i.e., with only a left context or only a right context. For instance, the rule p_1 can only be applied if the predecessor a is followed by b and the rule p_2 can only be applied if the predecessor b is preceded by aa:

$$\begin{array}{lcccccccc} \omega & : & b & & & & & & \\ p_1 & : & & & a & > & b & \longrightarrow & ab \\ p_2 & : & aa & < & b & & & \longrightarrow & a \end{array}$$

3.2.5 Parametric L-systems

Parametric L-systems were defined to improve the numerical power of L-systems and their ability to deal with continuous phenomena. One or several parameters can be assigned at each letter of the alphabet. They can be updated when rewriting a predecessor into its successor:

$$predecessor(parameters) : conditions \longrightarrow successor(parameters')$$

The use of parameters also enables the definition of conditions. The use of the rules is then subject to the condition validation. In this case, a string has not only to match the predecessor letters, but the number of parameters for each letters has also to be the same and the conditions, if any, have to be validated. Here, the letter b is associated to one parameter h that can represent anything from a branch length to a concentration of a certain element. The rule p_2 can only be applied to a letter b with one parameter and as long as that parameter is less than 1:

$$\begin{array}{lcccccc} h = 0 & & & & & & \\ \omega & : & b(h) & & & & \\ p_1 & : & a & & & \longrightarrow & ab(h + 0.1) \\ p_2 & : & b(h) & : & h \leq 1 & \longrightarrow & a \end{array}$$

h is global for the whole system and is not reset at the end of each iteration.

3.2.6 Environmentally-sensitive and open L-systems

In the previous L-system types, the turtle interpretation is a post-processing step to associate a geometry to a string. It has no effect on the string rewriting, which can be limiting in many applications. For instance the growth of a plant is influenced by its environment, such as the light exposure. In these cases the rewriting and the interpretation must take into account the influence of the environment.

In environmentally-sensitive L-systems, the turtle interpretation is done after each rewriting step [Prusinkiewicz et al., 1994]. Query modules enable to get some information from the interpretation and its relationship with its environment. For instance the rewriting process can get the current turtle state to influence the next rewriting steps.

Open L-systems generalize environmentally-sensitive L-systems [Mvech and Prusinkiewicz, 1996]. They have the same ability to receive information from the environment, but with the additional ability to update that environment following the system growth. For instance the growth of a root is influenced by the water repartition within the ground. Along its development, the root uses that water and so modify its distribution in the soil. In this case the root growth has also an impact on the environment, which then affects the root growth.

3.2.7 Bracketed L-systems

Trees are based on a branching structure. To simulate trees, L-systems have to be able to generate branches. Square brackets are used to handle branches: the opening bracket [records the current turtle state while the closing bracket returns to the saved state]. Between the brackets are mentioned the letters corresponding to the branch (figure 3.4).

The introduction of brackets is simple for all the previous L-system types, except the context-sensitive L-systems. When searching a left context and meeting a closing bracket, the branch has to be ignored and the search has to start again after the corresponding opening bracket.

All those systems can be combined together, giving the possibility to take into account various natural phenomena, for instance the propagation of a signal in a branch through auxin transport, and to generate various objects, from trees to fractals (for more details see Prusinkiewicz and Lindenmayer [1996]).

String

$ABC - [+DE][SG[+HI[-JK] + L]MNO]$

Interpretation

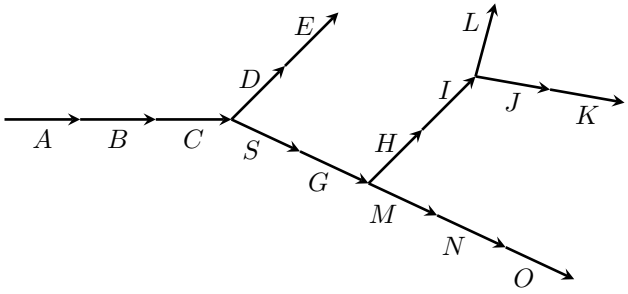


Figure 3.4 Example of string and geometrical interpretation (adapted from Prusinkiewicz and Lindenmayer [1996]). Letters A to O are different to facilitate the reading, but have all the same effect as a F from table 3.1.

3.3 L-system for channel stochastic simulation

Depending on their depositional environment, channels can display various morphologies, from straight to highly sinuous channels. In this work, we use L-system to simulate the channel morphology. The idea is to represent the channelized bodies resulting from the flow processes and not the flow processes themselves.

3.3.1 Simulation domain definition

Channelized systems that raise interest as reservoirs are not necessary in their depositional state. They have been affected by various geological events, possibly implying folding and faulting. It is difficult if not impossible to introduce all these events in the L-system.

The structural modeling workflow leads to the generation of a geological grid based on the horizons and the faults within the domain of interest (figure 9, introduction). This geological grid is often hexahedral and conformable to those horizons and faults. Its parametric space (u, v, t) can so be interpreted as a depositional space, with (u, v) the geographical coordinates and t the time (figure 3.5) [Mallet, 2004].

The channel simulation does not necessarily need to be applied in the Euclidean (x, y, z) space representing the current geological space. Any space can be used if a transformation exists to represent the system in the Euclidean

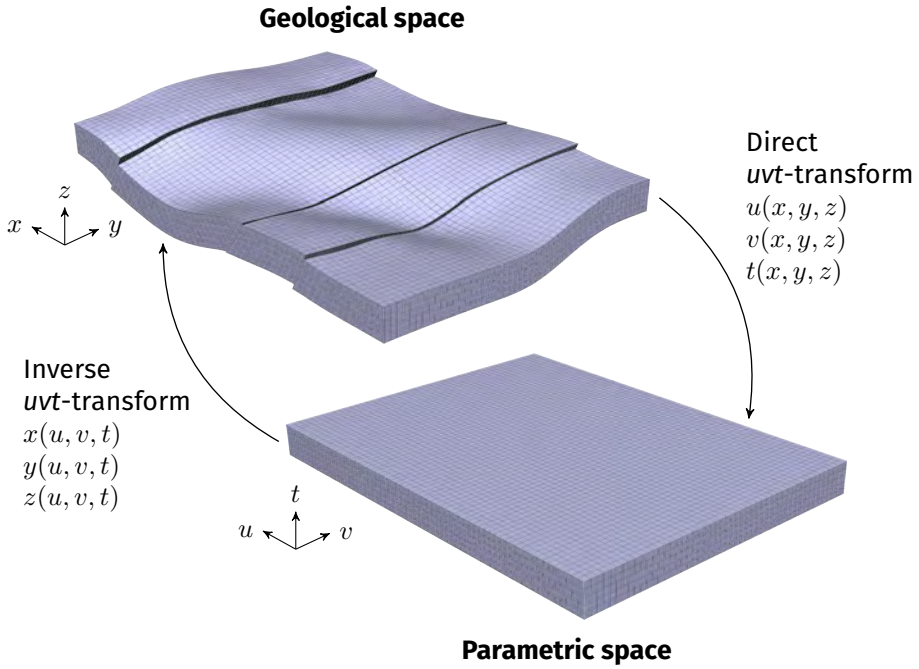


Figure 3.5 Example of curvilinear grid and representation of its parametric space (inspired by Mallet [2004]).

space at the end. Using the parametric space of a grid representing the domain of interest circumvents the issue of taking into account the evolution of the domain after the channel deposition. It implies that the grid is conformable to the horizons delimiting the formations associated with channelized systems.

In some cases the parametric space of a grid has some drawbacks, such as unwanted deformations of the grid cells. A more complex approach such as the Geo-Chronological model (GeoChron) [Mallet, 2004] can be used to parameterize a depositional space. GeoChron relies on a mathematical formulation of time-stratigraphy concepts defined by Wheeler [1958]. It provides a new space in which the horizons are flat and the faults have disappeared, similarly to the deposition space. From a simulation point of view the principle remains the same: simulating the channels in the depositional space and then transferring the resulting models into the Euclidean space.

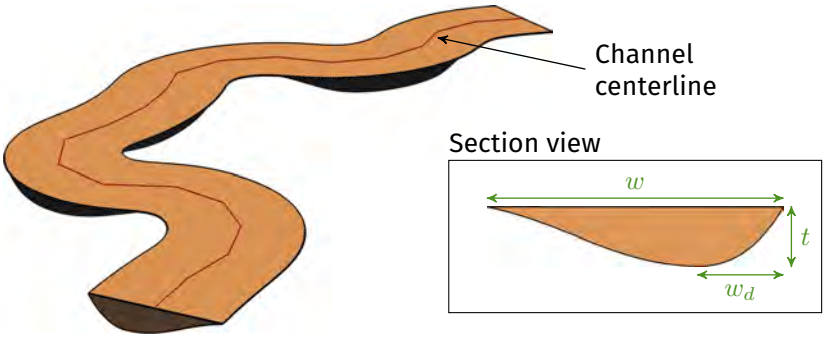


Figure 3.6 Channel envelope modeled with NURBS surfaces and required parameters for its simulation. w is the channel width and t is the channel thickness. w_d is the length between the outer border of the channel and the maximum depth. The asymmetry aspect ratio is equal to w_d/w.

3.3.2 Channel object parameterization

The channel object is based on the channel boundary model defined by Ruiu et al. [2015b]. It relies on Non-Uniform Rational B-Splines (NURBS), a generalization of the Bézier curves [Piegl and Tiller, 1995]. NURBS present the advantage of being easy to deform, which can be useful for data conditioning. Five NURBS surfaces are gathered to form a channel: one flat surface represents the top of the channel, two curved surfaces joining at the channel maximal depth represent the sides and bottom of the channel and two flat surfaces represent the channel extremities. Four parameters are necessary to build the NURBS model (figure 3.6):

- A discretized channel centerline, also called backbone [e.g., Wietzerbin, 1994, Viseur, 2001, Pyrcz et al., 2009]. This centerline is at the channel top and informs about the channel sinuosity.
- A set of channel widths along the centerline.
- A set of channel thicknesses along the centerline.
- An asymmetry aspect ratio.

The process simulates at first the L-system string, which contains all the parameter values required to generate the channel object. Thus, the channel representation is only created at the end of the simulation process.

3.3.3 Initial turtle state

The initial position of the system can come from different sources:

- Directly given by the user.
- Randomly drawn among the not yet conditioned well channel data. In this case, a data is first drawn, and the position is randomly perturbed both horizontally and vertically depending on the channel minimal width and minimal thickness. At the end, the channel still fits the data but at different locations at each simulation.
- Randomly drawn within the domain of interest.
- Randomly drawn within the domain of interest with the influence of a sand probability cube, if one is defined within the domain.

A channel global direction comes from an input azimuth and defines the turtle's initial coordinate system, with the turtle's head directly oriented along this direction. The global direction can be the same for the whole channel, or it can come from a property defined within the domain of interest. In this case, the direction is given by the property at the initial position of the channel.

3.3.4 Basic considerations and alphabet

After the parameterization, a channel can be seen as a discretized object composed of channel segments separated by transversal channel sections (figure 3.7). A section is at the end of its respective segment. Each section is aligned on the bisector between its segment and the next segment, as in the NURBS model [Ruiu et al., 2015b]. An orientation and a distance from the previous section characterize a given section. The L-system rewriting process determines the location of each section, which then gets a channel width and thickness at this location. Thus, each section knows all the parameters to build the channel object (see section 3.3.2) at that position. The geometrical interpretation consists in building the NURBS representation.

Five letters constitute the alphabet to model the succession of sections: $+$, $-$, C, T and I (table 3.2). The letters $+$ and $-$ modify the orientation of the system just as in the classical L-system. They have a main impact on the channel morphology. The letter C is comparable in its effect to the letter F defined in table 3.1. A channel section is symbolized by a $\pm C$ pair.

The letter I has no influence in the string rewriting process. It is only used during the geometrical interpretation to state the initial position of the channel.

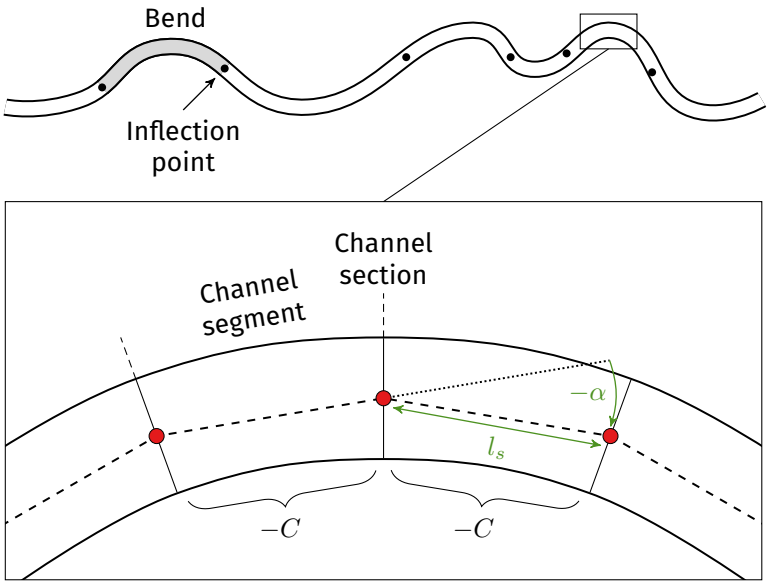


Figure 3.7 Principle of channel discretization for L-system simulation. The red dots correspond to the position of the channel sections, obtained from the turtle interpretation. l_s is both a channel element length and the distance between two successive sections. α is the angle between to successive sections, i.e., the change of orientation between two successive channel segments.

The letter T has no influence on the geometrical interpretation. It is only used during the string rewriting to assure the channel growth, as detailed in the next section.

3.3.5 Rule definition

A channel is decomposed into several bends separated by inflection points (figure 3.7). A bend is a succession of channel segments whose orientations change along the same direction. Thus, it is a succession of $\pm C$ letters, with $\pm$ being $+$ or $-$ but remaining the same along the bend. The number of channel segments within each bend is here an essential parameter. The L-system process for channel simulation calls for four parameters:

- A default segment length l_d.
- An angle between two successive segments α.

Table 3.2 Basic L-system alphabet and corresponding geometrical interpretation for channel axis simulation.

Letters(parameters)	Interpretation
I	First letter of a L-system string giving the initial position
T	Do nothing
$C(l_s)$	Move forward of a length l_s and draw a channel section between the new position and the former one
$+(\alpha)$	Turn left by an angle α based on the rotation matrix $R(\hat{U}, \alpha)$
$-(\alpha)$	Turn right by an angle α based on the rotation matrix $R(\hat{U}, -\alpha)$

- A bend length l_B, i.e., the curvilinear distance between two inflection points.
- A channel length l_C, which is not necessary with a simulation domain.

The L-system rules are separated in three steps: the initiation, the development and the closure. In each rule, l_B and l_d determine the number n_s of segments of a bend by:

$$n_s = \lceil l_B / l_d \rceil \tag{3.4}$$

where $\lceil x \rceil$ rounds x upward. The actual segment length l_s is recomputed as follows:

$$l_s = l_B / n_s \tag{3.5}$$

Initiation and development rules

The initial position of the channel does not need to be an extremity of the channel. Indeed, the initial position can be drawn in the middle of the simulation domain and not necessary on its border. Thus, a channel is divided into two branches: the first branch grows in one direction and the second branch in the opposite direction. The initiation rules start this branching structure (figure 3.8). It is symbolized by the brackets, with in between the letters corresponding to the first branch. The initiation rules simulate a single bend separated into two parts, one for each branch. The numbers of segments for each part, n_{s_1} and n_{s_2}, are computed by:

$$\begin{aligned} n_{s_1} &\sim \mathcal{U}(1, n_s - 1) \\ n_{s_2} &= n_s - n_{s_1} \end{aligned} \tag{3.6}$$

Axiom IT

Initiation

Rules

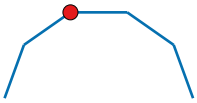
$$T \xrightarrow{0.5} \left[+(\alpha_1)C(l_s)\{+(\alpha)C(l_s)\}^{n_{s_1}-1}T\right] + (180) - (\alpha_2)C(l_s)\{-(\alpha)C(l_s)\}^{n_{s_2}}T$$

$$T \xrightarrow{0.5} \left[-(\alpha_1)C(l_s)\{-(\alpha)C(l_s)\}^{n_{s_1}-1}T\right] + (180) + (\alpha_2)C(l_s)\{+(\alpha)C(l_s)\}^{n_{s_2}}T$$

Final string *Iteration 1*

$I\,[-C-C-C\,T\,] + (180)+C+C\,T$

Interpretation



Development

Rules

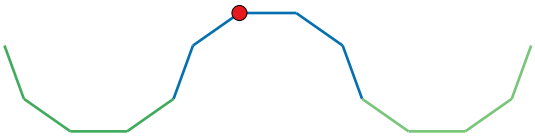
$$-(\alpha)C(l_s) < T \longrightarrow \{+(\alpha)C(l_s)\}^{n_s}T$$

$$+(\alpha)C(l_s) < T \longrightarrow \{-(\alpha)C(l_s)\}^{n_s}T$$

Final string *Iteration 2*

$I\,[-C-C-C+C+C+C+C\,T\,] + (180)+C+C-C-C-C-C\,T$

Interpretation



Final string *Iteration 3*

$I\,[-C-C-C+C+C+C+C-C-C-C-C\,T\,] + (180)+C+C-C-C-C-C+C+C+C+C\,T$

Interpretation

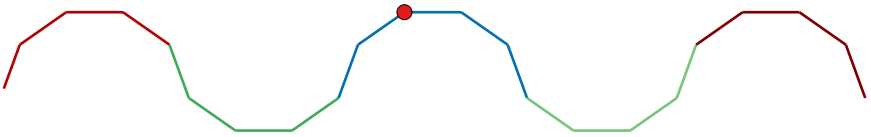


Figure 3.8 Simplified illustration of channel simulation based on a L-system. The red point is the initiation point for the interpretation. The first iteration initiates the channel structure, divided in two branches, based on stochastic rules. The following iterations handle channel growth depending on the context.

where n_s is the total number of segments for the bend and $\mathcal{U}$ is a uniform discrete distribution with its minimal and maximal values as parameters. n_{s_1} is randomly drawn from that distribution. This random draw implies: $n_s \geq 2$, so $l_B > l_d$ based on equation 3.4. The initiation rules are stochastic: they simulate a bend in one direction or the other depending on an equal probability of occurrence.

The channel must be oriented along the global direction, which defines the initial turtle's state. The initial turtle's state influences the orientation of the first segment of each branch. The angle of these segments must be modified, otherwise they would follow the global direction whereas the channel could follow a different direction. The angles α of these segments become:

$$\begin{aligned} \alpha_1 &= \alpha \times (\tfrac{n_s}{2} - n_{s_1} + 1) \\ \alpha_2 &= \alpha \times (\tfrac{n_s}{2} - n_{s_2}) \end{aligned} \tag{3.7}$$

A letter T is placed at each branch extremity to continue the growth.

This growth is assured by the development rules. Their purpose is to assure the channel growth by tying channel bends one after the other. These rules simply replace a T by a new bend, with a letter T at the end to pursue the channel development. They are context-sensitive. This enables to change the bend orientation from one bend to the other.

Stopping criteria and closure rules

For channel simulation the stopping criterion of the rewriting is not an order. When the L-system process occurs within a simulation domain, the process can stop as soon as the two branches go outside the grid. A second possibility is to stop the process as soon as a target channel length l_C is reached. A parameter l_c keeps track of the channel length during the growth. When that length is equal or larger than the target channel length l_C, the rewriting process stops.

Obtaining channels of a precise length is possible through closure rules. In this case, the development rules are called only if the new channel length does not exceed the target channel length:

$$l_c + l_B \leq l_C \tag{3.8}$$

The closing rules redefine the bend length as:

$$l_B = l_C - l_c \tag{3.9}$$

They are called only two times, one per branch, and are expressed as:

$$\begin{aligned} -(\alpha)C(l_s) < T &\longrightarrow \{+(\alpha)C(l_s)\}^{n_{s_n}} \\ +(\alpha)C(l_s) < T &\longrightarrow \{-(\alpha)C(l_s)\}^{n_{s_n}} \end{aligned} \tag{3.10}$$

The number n_{s_n} of segments in those rules is defined as:

$$n_{s_n} \sim \mathcal{U}(0, n_s) \times i + n_{s_o} \tag{3.11}$$

where $\mathcal{U}$ is a uniform discrete distribution with its minimal and maximal values as parameters. At the very beginning of the rewriting process, i is set to 1 and n_{s_o} to 0. Once a closure rule is called, i is set to 0 and n_{s_o} becomes:

$$n_{s_o} = n_s - n_{s_n} \tag{3.12}$$

At the second call, the number of segments of the bend becomes the n_{s_o} of the first call. This operation gives to the channel the target length l_C.

In the whole process, two parameters are drawn from some statistical distributions to add more stochasticity to the system:

- The channel length is drawn before the start of the rewriting process.
- A bend length is drawn each time a rule is called, so for each new bend.

The default segment length remains the same for all the segments. The angle between each segment varies for each segment. But this parameter is not directly asked in input, as detailed in the following section.

3.3.6 L-system parameters

The rules for channel simulation call for four parameters: the channel length, the bend length, the default segment length and the angle between two successive segments. Among those parameters, the bend length and the angle define the morphology of the bends and of the channel. From a user point of view, an angle is not a parameter easy to infer. Thus, it could be useful to replace it by other parameters more classically used in the geomorphological description of channel bends. The L-system rules are then modified to handle the conversion between the user input parameters and the L-system parameters. Two sets of exclusive parameters are proposed here: a bend length and a curvature or a bend amplitude and a bend half-wavelength (figure 3.9).

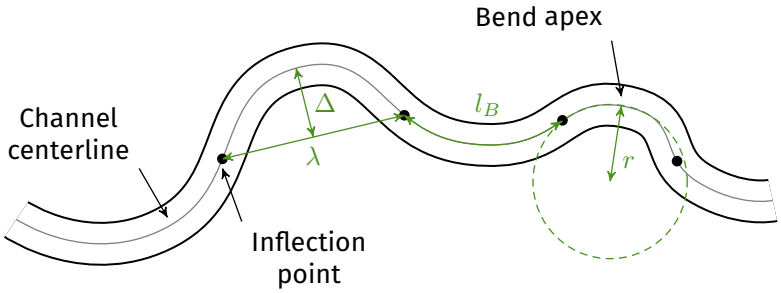


Figure 3.9 Parameters used to characterize an ideal channel bend. l_B is the bend length, λ is the bend half-wavelength, Δ is the bend amplitude and $r = 1/c$ is the radius of curvature, with c the curvature.

Bend length and curvature

The first set of parameters keeps the bend length and only changes the angle between successive channel segments. The idea is to rely on the radius of curvature of a bend, which is the inverse of the curvature. This radius corresponds to the radius of a circle fitted to the apex of the bend (figure 3.9).

For a section i, α is then expressed as:

$$\alpha_i = \frac{c(l_{s,i-1} + l_{s,i})}{2} \tag{3.13}$$

with c the curvature, $l_{s,i-1}$ and $l_{s,i}$ the distances to the previous and next sections, or equivalently the length of the current channel segment and the length of the next segment element. α is in radian. The curvature does not need to be signed. The sign is defined by the L-system itself by means of the $+$ and $-$ letters. From a user point of view, a curvature is more independent from the segment length than an angle.

While the draw of a bend length is done from a statistical distribution for each rule, the curvature is drawn from a statistical distribution for each segment. This leads to more various bend morphologies without introducing more parameters. Indeed, using a single radius of curvature for a whole bend is not enough to characterize the variety of bend shapes [Howard and Hemberger, 1991].

From a process perspective, both curvature or radius of curvature could be used. The radius of curvature is a length, so it may be a simpler parameter from a user perspective than a curvature. From a geomorphological point of view, a single value of radius of curvature is often used to characterize a given

bend. Here the radius of curvature changes along a bend. So we choose to keep the curvature, which could be less misinterpreted.

Table B.4 of appendix B gathers the complete L-system rules for channel simulation with a bend length and a curvature.

Bend half-wavelength and bend amplitude

The second set of parameters is a simplification of the first one. Indeed, both a curvature and a radius of curvature may be difficult to infer on a bend. This set relies on the bend half-wavelength, i.e., the Euclidean distance between two successive inflection points, and the bend amplitude, i.e., the maximal distance between the half-wavelength line and the bend centerline (figure 3.9).

Without stochasticity in the geometrical parameters, the L-system generates bends with the shape of a circular arc. It is thus possible to link both the bend length l_B and the curvature c to a half-wavelength λ and an amplitude Δ thanks to the intersecting chord and the law of cosines. The curvature c is expressed as:

$$c = \frac{8\Delta}{4\Delta^2 + \lambda^2} \tag{3.14}$$

And the bend length l_B becomes:

$$l_B = \frac{2\arccos(1 - \Delta c)}{c} \tag{3.15}$$

During the simulation, a half-wavelength and an amplitude are drawn for each bend. They can not be drawn for each segment, as the bend length would then vary for each segment. Thus, the curvature is always the same for each segment of a bend. This may limit the stochasticity of the system and limit the variety of bend morphologies. A deviation angle δ is added to the angle α to compensate that effect and to add some variability:

$$\alpha_i = \frac{c(l_{s,i-1} + l_{s,i})}{2} + \delta\frac{\pi}{180} \tag{3.16}$$

where α is in radian and δ is in degree.

3.3.7 Extending the influence of the global direction

The global direction is one of the input parameters. It is the direction along which the channel should develop. Only the initial turtle's state takes it into account. It does not influence the further developments of the system.

When the L-system parameter values have few variations from bend to bend, the channel continues to follow roughly the channel direction. When the L-system alternates straight and sinuous bends, the resulting channels often self-intersect. These self-intersections are an inconsistency. They result from the L-system interpretation principle: a new segment only knows its previous neighbor and ignores all about the previous segments and their locations.

Extending the influence of the global direction to the whole system helps to better control the channel orientation and limit self-intersections. The global direction $\hat{D}$ directly influences the turtle's head $\hat{H}$:

$$\begin{aligned} \hat{H'} &= \frac{\vec{K}}{\|\vec{K}\|} \\ \vec{K} &= \epsilon_h \hat{H} + \epsilon_d \hat{D} \end{aligned} \tag{3.17}$$

where $\hat{H'}$ is the new turtle's head, ϵ_h is a weight on the L-system, ϵ_d is a weight on the global direction. ϵ_h and ϵ_d are user-defined and allow controlling the influence of the L-system and of the global direction. This process is done by the letters that define the change of orientation between the channel sections, i.e., $+$ and $-$. It applies after the orientation change done by the L-system.

The input global direction comes from two possible sources:

- A direction directly defined for the whole system.
- A property cube defined over the whole domain of interest. In this case, the direction is updated when simulating a new section depending on the cell in which is the current section.

3.3.8 Simulating the channel width and thickness

The L-system in itself simulates the channel centerline required by the channel object. The channel width and thickness are simulated for each section at the end of the L-system rewriting and interpretation. Then, building the channel object means interpolating the channel sections based on their widths and thicknesses and interpolating the channel shape between the sections. The channel width and thickness are simulated at each channel section using a Sequential Gaussian Simulation (SGS).

To simplify the use, several SGS parameters are kept constant, with the idea of having smooth variations between the sections:

- The variogram nugget effect is set to 0.

- The variogram sill is set to 1, as the SGS calls for a normal score transform.
- The variogram model is defined as Gaussian.

Two additional parameters are let to the user: distributions of width or thickness values and variogram ranges for the width and thickness. The ranges define the scale of oscillation of the width and thickness. Both the width and the thickness may vary following the channel curvature. Thus, this curvature is used as a secondary variable of the SGS. This adds another parameter: the curvature weight. When the weight is equal to one, the spatial structure of the width or the thickness strictly follows the spatial structure of the curvature. When the weight is equal to zero, the width or the thickness are completely independent from the curvature.

Table 3.3 resumes all the parameters required to simulate a channel object. It contains the parameters related to the L-system and the parameters defining the section of the channel. The L-system rules do not appear in this list: as they are predefined, they are not strictly an input, but nothing prevents from modifying them.

3.4 Adding external constraints

In biology, constraining the development of L-systems manages the interaction between trees and their environment [e.g., Taylor-Hell, 2005, Longay et al., 2012]. External constraints are added to the L-system through an environmentally-sensitive process. After each rewriting step, an interpretation allows gathering some information about the channel and its environment, such as the channel position relative to a well data. Such information helps to build constraints that influence the further development of the system. The following sections provide a general definition of the constraints, before detailing the constraints predefined for channel simulation.

3.4.1 Formalization

A vector materializes a constraint and indicates where the system should preferentially go under the constraint influence. This vector decomposes in three elements:

- $\hat{\Lambda}$ is a normalized vector indicating the constraint direction.
- ρ is the vector magnitude which determines the force of the constraint.

Table 3.3 Input parameters to simulate a channel from L-system and NURBS. Only the L-system and envelope parameters are essential. An asterisk symbolizes a parameter not necessarily required with a simulation domain. The parameters in italic call for statistical distributions.

	Simulation input
Domain	Grid representing the domain of interest
L-system	* Initial position
	L-system weight
	Azimuth defining the global direction (in °)
	Global direction weight
	Default segment length (in m)
	* *Channel length* (in m)
	Bend length (in m) / *Curvature* (in m^{-1}) **or** *Bend half-wavelength* (in m) / *Bend amplitudes* (in m) / *Deviation angles* (in °)
Channel section	*Channel width* (in m)
	Channel width range (in m)
	Curvature weight for the width
	Channel thickness (in m)
	Channel thickness range (in m)
	Curvature weight for the thickness
	Asymmetry aspect ratio

- ϵ is a user defined weight to adjust the relative importance of each constraint.

Either the vector materializing a constraint is directly given as input, or it is defined from a vector between the turtle's head and a given constraining element. In the second case, determining a constraint means choosing a constraining element and determining the constraint vector (figure 3.10).

The constraints are separated into two categories:

- Relative constraints have their magnitude that depends on the distance to their constraining element. Thus, they are only defined relatively to a constraining element, for instance a well or a seismic-interpreted object.
- Global constraints have a magnitude independent from the distance to

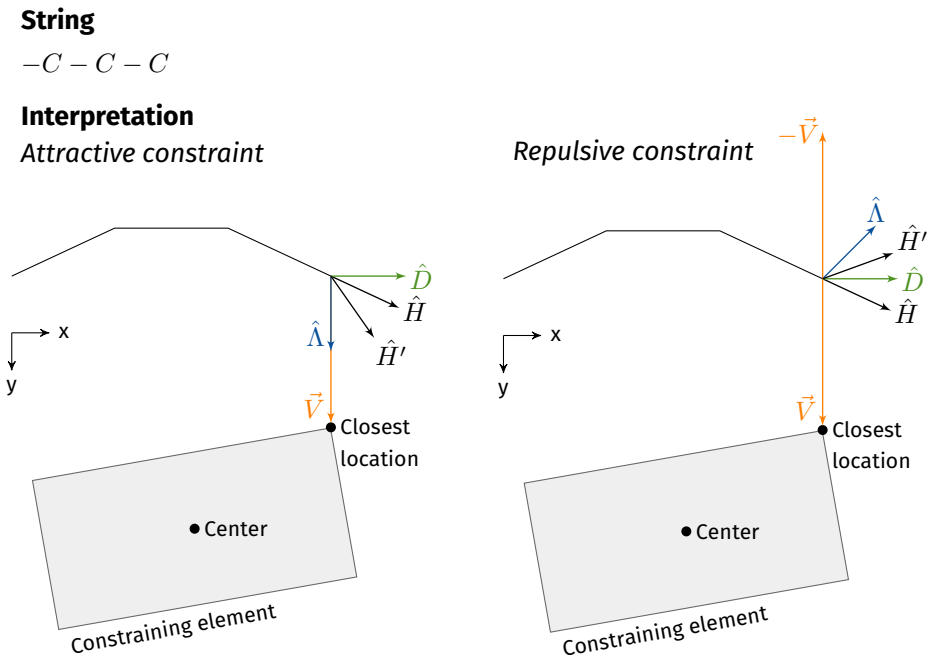


Figure 3.10 Principle of constraint application to a L-system. $\hat{H}$ is the current turtle's head and $\hat{H}'$ the new turtle's head that takes into account the constraint. $\hat{D}$ is the global direction of the channel. $\hat{\Lambda}$ is the constraint direction, deduced from $\vec{V}$. $\vec{V}$ is the vector between the turtle's head and a point on the constraining element, either its center or the closest location to the turtle's head. The next sections give further details about constraint application.

their constraining element. Thus, they do not necessarily require a constraining element. In this case, the constraint vector is given in input and is constant over the whole system. The channel global direction is equivalent to a global constraint without constraining element. A probability cube is another example of global constraint, with the constraining element depending on the local maximum probability value.

In addition, a constraint can be (figure 3.10):

- Attractive, i.e., it attracts the system toward the constraining element.
- Repulsive, i.e., it repulses the system from the constraining element.

3.4.2 Sources of constraining elements

In channel simulation, adding external constraints has two goals:

- Controlling the relationship between simulated channels and between different parts of a channel.
- Conditioning the channel to data. The two main sources of data are the wells and the seismic cubes.

The previous channels, channel parts, well or seismic data provide potential constraining elements for the system.

Previous channel sections

When simulating several channels, some can develop at the same depth and intersect each other. If required, this configuration is preventable by adding repulsion constraints between the previously simulated channels and the new channels. Such repulsive constraints are also useful to further prevent channel self-intersections when the channel sinuosity becomes locally too important. Indeed, the global direction is not always sufficient to prevent self-intersections.

Well sedimentary data

Wells are classical data in geology. They provide a local and often precise information about the facies and the associated sedimentary bodies. The simulation process supposes that the sedimentary objects are already interpreted along the wells.

We consider separately two main types of data:

- Channel data comes from channelized bodies, whatever their mud content. The simulated channels have to condition such data, i.e., each channel data must end up in at least one channel for all the realizations. Constraints create an attraction to the data to ensure the conditioning. But all the data should not attract a single channel at the same time.
- Inter-channel data correspond to all the other sedimentary bodies around channels, such as overbanks. They are often deposited alongside a channel and eroded by the younger channels. A channel should not condition such data, i.e., no inter-channel data must end up in a channel for all the realizations.

Well connectivity data

When studying or exploiting a reservoir, connectivity data between wells are sometimes available. If so, they can inform about the connection or disconnection of two wells. This forms sets of connected data, which knows their non-connected counterparts. When a channel conditions a data from such set, it has to condition the other connected data. It has also to avoid the entire non-connected data set.

Probability cube

Seismic data provide an overview of the underground in three-dimensions. Unfortunately their resolution is not always sufficient to identify geological structures. However, some information can still be extracted, such as the sand probability within the domain of interest [e.g., Strebelle et al., 2003]. As channels often have a sandy filling, those data influence the channel locations: a high channel probability is considered to be positively correlated to a high sand probability. A probability cube first helps to determine the channel initial position (see section 3.3.3) and then influences the channel development. The mean of many realizations, i.e., the E-map, should be as close as possible to the sand probability cube, without forgetting that simulating channels is not simulating sand distributions.

Seismic-interpreted objects

Sometimes seismic data have a sufficient quality for object interpretation. This interpretation is most of the time done manually. Ruiu et al. [2015a] developed such semi-automated approach with NURBS. The interpreter must digitize the channel centerline. A deformation process automatically fits the NURBS surfaces to the channel observed on the seismic data. The simulation process should take into account such interpreted sedimentary objects.

In turbiditic environments, channels often develop within a confining master channel. This confinement is a major source of uncertainty [Larue and Hovadik, 2008] and is one of the first features to determine in a reservoir study. When observable on seismic data, the confinement must be interpreted to influence the simulated channels. During the simulation within the confinement, constraints prevent the channels from going out of the master channel: its margins repulse the channel.

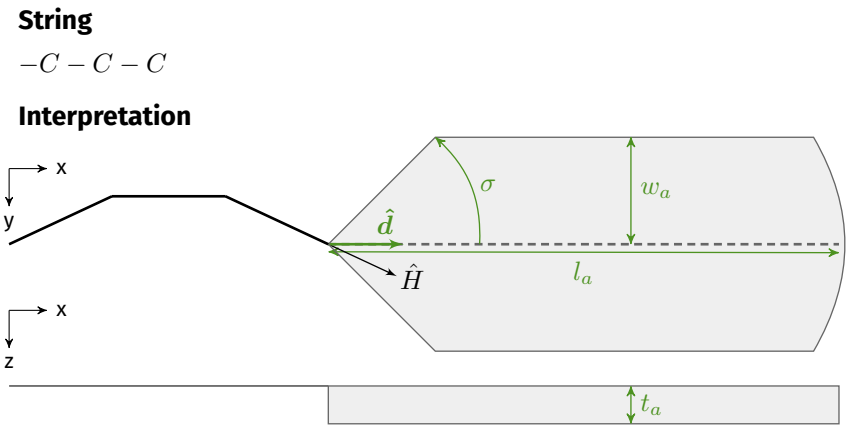


Figure 3.11 Parameters defining the area which select the constraining elements. $\hat{H}$ is the turtle's head. $\hat{d}$ is the direction of the area, σ is the direction tolerance, w_a is the maximal half-width of the area, l_a is the area length and t_a is the area thickness.

3.4.3 Constraining element selection

A constraining element influences the system development depending on two factors. The first factor is a perception area from the turtle's head: the constraining elements falling into that area may influence the system. The other elements that fall outside the area are ignored.

This perception area is a band starting from the current position of the system. Four parameters define it (figure 3.11):

- $\hat{d}$ is the center direction of the band.
- σ is the direction tolerance, defined as an angle from the center direction.
- w_a is the bandwidth, i.e., the maximal half-width of the area.
- l_a is the area length.
- t_a is the area thickness, defined below the channel segments because they are at the top of a channel.

w_a and l_a are used to reduce the perception area and limit the number of constraints.

The second factor is a selection criterion, which defines if a constraining element within the perception area induces a constraint or not. There are three possibilities:

- All the elements within the perception area influence the system.
- Only the closest element to the turtle's head influences the system.
- Only the element with the highest value of a given property influences the system.

Usually those factors are applied for each section during the interpretation. Thus the constraining elements may change between two successive sections: a section has no memory about the constraining elements of the previous sections. For some constraints, especially relative attractive constraints, the system can memorize the constraining elements until the conditioning, even if that element is not in the perception area any more.

3.4.4 Constraint definition

A vector characterizes a constraint through three components: a direction, a magnitude and a weight. The weight is a constant scalar value. The other two elements have more complex definition.

Constraint direction

The direction $\hat{\Lambda}$ of constraint application defines if a constraint is attractive or repulsive:

$$\hat{\Lambda} = \begin{cases} \frac{\vec{V}}{\|\vec{V}\|} \text{ if the constraint is attractive} \\ \frac{\hat{D}-\vec{V}}{\|\hat{D}-\vec{V}\|} \text{ if the constraint is repulsive} \end{cases} \tag{3.18}$$

where $\hat{D}$ is the global channel direction. Adding the global direction to repulsive constraints helps to limit abrupt direction changes due to a high repulsion.

$\vec{V}$ is the vector between the turtle's head and a particular point on the constraining element. This point can be:

- The center of the constraining element. The vertical component of $\vec{V}$ is set to zero to avoid any vertical perturbation of the channel centerline.
- The closest location along the element border at the same vertical coordinate.

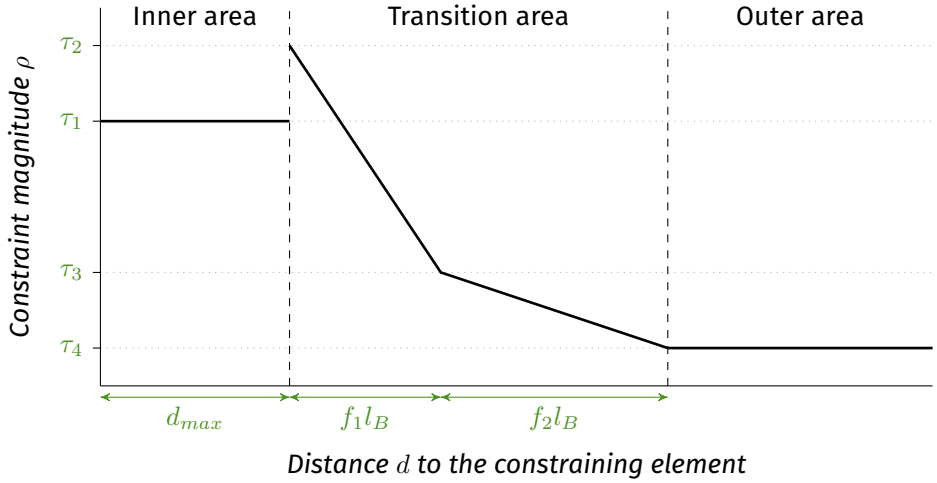


Figure 3.12 Parameters defining the magnitude of a relative constraint.

Constraint magnitude

The constraint magnitude ρ varies between 0 and 1 and depends on the constraint type. For a global constraint, the magnitude is expressed as:

$$\rho = 1 \tag{3.19}$$

By definition, the magnitude of a relative constraint varies following the distance d between the turtle's head and the constraining element:

$$\rho = \begin{cases} \tau_1 \text{ if } d < d_{max} \\ \tau_2 + (\tau_3 - \tau_2)\frac{d - d_{max}}{f_1 l_B} \text{ if } d_{max} \leq d \leq d_{max} + f_1 l_B \\ \tau_3 + (\tau_4 - \tau_3)\frac{d - d_{max} - f_1 l_B}{f_2 l_B} \text{ if } d_{max} + f_1 l_B < d \leq d_{max} + (f_1 + f_2) l_B \\ \tau_4 \text{ if } d > d_{max} + (f_1 + f_2) l_B \end{cases} \tag{3.20}$$

with $0 \leq \tau_1, \tau_2, \tau_3, \tau_4 \leq 1$.

Adding a constraint induces a deformation of the L-system. The purpose with a relative constraint is to disrupt as little as possible the morphology generated by the L-system. Thus the magnitude evolves following three areas (figure 3.12).

The inner area is defined by a distance d_{max} and a magnitude τ_1. This is the conditioning area toward which the channel evolves in case of an attrac-

tive constraint or the absolutely forbidden area for a repulsive constraint. For channels, d_{max} is equal to half the channel width.

The outer area is when the turtle's head is far away from the data. The magnitude value in this area τ_4 must be as low as possible to perturb at least the channel morphology.

In between, the magnitude evolves following the bend length l_B, because a larger bend length requires more anticipation to attract or repulse. This transition area gradually adjust the magnitude from the outer to the inner area. It has two parts:

- The first part is adjacent to the inner area and its length is controlled by the bend length and a factor f_1. It is the core part of the constraint influence, which ensures the conditioning.
- The second part is near the outer area and its length is controlled by the bend length and a factor f_2. Its magnitude variation is less abrupt than in the first part. Thus the constraint effect can begin early without deforming too much the channel morphology.

The magnitude goes from τ_4, the magnitude in the outer area, to τ_3 and then τ_2, the highest magnitude value of the constraint.

Representing the magnitude with a piecewise function requires many parameters, but this function is easier to define than a single function with similar variations. Moreover, adding other parameters that could influence the magnitude is far easier with this representation than with a single continuous function.

3.4.5 Adding the constraints to the L-system

Once defined, the constraints are added to the turtle head $\hat{H}$ by modifying the formula 3.17:

$$\begin{aligned}
\hat{H}' &= \frac{\vec{K}}{\|\vec{K}\|} \\
\vec{K} &= \epsilon_h \rho_r \hat{H} + \epsilon_d \rho_r \hat{D} + \sum_{j=1}^{n_g} \epsilon_{g,j} \rho_r \hat{\Lambda}_j + \sum_{i=1}^{n_r} \epsilon_{r,i} \rho_i \hat{\Lambda}_i \\
\rho_r &= \prod_{i=1}^{n_r} (1 - \epsilon_{r,i} \rho_i)
\end{aligned} \tag{3.21}$$

with $\hat{H}'$ the new turtle head, n_g the number of global constraints and n_r the number of relative constraints. $\epsilon_{g,j}$ is a user-defined weight for each global

Table 3.4 Constraints defined for channel simulation.

Constraint	Constraint type	Constraining element
Channel self-repulsion	Relative & repulsive	Channel section
Well channel data	Relative & attractive	Well point
Well inter-channel data	Relative & repulsive	Well point
Well connected channel data	Relative & attractive	Well point
Well non-connected channel data	Relative & repulsive	Well point
Probability cube	Global & attractive	Grid cell
Confinement	Relative & repulsive	Grid border

constraint j. $\epsilon_{r,i}$ is a user-defined weight for each relative constraint i. ϵ_h and ϵ_d are the user-defined weights for the L-system and the global direction. ρ_i is the magnitude of the constraint i and $\hat{\Lambda}_i$ the constraint direction. The magnitude ρ_r decreases the turtle's head magnitude, the global direction magnitude and the global constraint magnitudes when a relative constraint has a high magnitude. This improves the conditioning by limiting the influence of the other constraints when getting close the considered data.

3.4.6 Constraint setting

From all the sources of constraining elements defined in section 3.4.2, some have been retained for their interest in the stochastic simulation of channelized systems:

- Channel self-repulsion helps to prevent self-intersections when simulating highly sinuous channels, which are inconsistencies.
- Wells are classical data and the only data with a low uncertainty. Thus, their conditioning is essential. A well is here considered as a set of aligned points.
- A probability cube derived from seismic data, e.g., a sand probability cube, can give previous information about the sedimentary body continuity and spatial organization.
- The confinement, for instance a canyon or a master channel, is highly important due to its influence on the channel location. Its impact on the uncertainties of the models is significant [Larue and Hovadik, 2008].

Table 3.5 Parameters to determine a constraint.

Feature	Sub-feature	Type	Symbol	Definition
Weight		All	ϵ	Weight to control the relative strengths of the constraints
Constraining element selection	Perception area	All	$\hat{d}$	Center direction
			σ	Direction tolerance
			w_a	Bandwidth
			l_a	Length
			t_a	Thickness
	Selection	All	C	Selection criterion within the perception area
	Memory	All	M	Keep the constraining element over the next sections
Direction		All	L	Location on the constraining element from which comes the constraint
			$\hat{\Lambda}$	Constraint direction
Magnitude		Relative	τ_1, τ_2, τ_3, τ_4	Magnitude values delimiting the inner, transition and outer areas
			d_{max}	Size of the inner area
			f_1, f_2	Factors controlling the transition area
		Global		Always 1

Here the confining structure is defined as a hexahedral grid. The channels are simulated within the parametric space of that grid. The two opposite borders parallel to the channel global direction are the constraining elements.

All those elements define different constraints (table 3.4).

A relative constraint requires seventeen parameters. A global constraint requires ten parameters, as the magnitude is always one (table 3.5). Such a number of parameters is rather constraining from a user point of view. For the sources introduced previously, it is possible to define default values (tables 3.6, 3.7 and 3.8). These values come from different tests of the method. For most

Table 3.6 Parameters to define the perception area and to choose the constraining element(s) within that area. $\hat{H}$ is the turtle's head, $\hat{U}$ is the turtle's up direction, R is the rotation matrix from 3.2, $\hat{D}$ is the global direction of the channel, l_s is the channel segment length, $w_{a,CD}$ is the perception area bandwidth for the well channel data and t_{min} is the minimal channel thickness. The user input parameters are in bold.

	$\hat{d}$	σ	w_a	l_a	t_a	C	M
Channel self-repulsion	$\hat{H}$	90°	$+\infty$	$+\infty$	$\boldsymbol{t_{min}}$	Closest	–
Well channel data	$\hat{H}+\hat{\boldsymbol{D}}$	90°	$\boldsymbol{w_{a,CD}}$	$+\infty$	$\boldsymbol{t_{min}}$	Closest	No
Well inter-channel data	$\hat{H}$	125°	$+\infty$	$+\infty$	$\boldsymbol{t_{min}}$	All	–
Well connected data	$\hat{H}+\hat{\boldsymbol{D}}$	90°	$+\infty$	$+\infty$	$\boldsymbol{t_{min}}$	Closest	Yes
Well non-connected data	$\hat{H}$	125°	$+\infty$	$+\infty$	$\boldsymbol{t_{min}}$	All	–
Probability cube	$\hat{H}+\hat{\boldsymbol{D}}$	45°	$+\infty$	$5l_s$	$\boldsymbol{t_{min}}$	Max value	–
Confinement (border 1)	$\hat{\boldsymbol{D}}+R(\hat{U},\frac{\pi}{2})$	0°	$+\infty$	$+\infty$	0	Closest	–
Confinement (border 2)	$\hat{\boldsymbol{D}}+R(\hat{U},-\frac{\pi}{2})$	0°	$+\infty$	$+\infty$	0	Closest	–

Table 3.7 Parameters to define the constraint direction. $\vec{V}$ is a vector between the turtle's head and the constraining element location and $\hat{D}$ is the global direction of the channel. The user input parameters are in bold.

	L	$\hat{\Lambda}$
Channel self-repulsion	Section position	$\frac{\hat{\boldsymbol{D}}-\vec{V}}{\|\hat{\boldsymbol{D}}-\vec{V}\|}$
Well channel data	Well point	$\frac{\vec{V}}{\|\vec{V}\|}$
Well inter-channel data	Well point	$\frac{\hat{\boldsymbol{D}}-\vec{V}}{\|\hat{\boldsymbol{D}}-\vec{V}\|}$
Well connected data	Well point	$\frac{\vec{V}}{\|\vec{V}\|}$
Well non-connected data	Well point	$\frac{\hat{\boldsymbol{D}}-\vec{V}}{\|\hat{\boldsymbol{D}}-\vec{V}\|}$
Probability cube	Cell center	$\frac{\vec{V}}{\|\vec{V}\|}$
Confinement (border 1)	Border closest location	$\frac{\hat{\boldsymbol{D}}-\vec{V}}{\|\hat{\boldsymbol{D}}-\vec{V}\|}$
Confinement (border 2)	Border closest location	$\frac{\hat{\boldsymbol{D}}-\vec{V}}{\|\hat{\boldsymbol{D}}-\vec{V}\|}$

Table 3.8 Parameters to define the constraint magnitude. w_{min} and w_{max} are the minimal and maximal channel widths. The user input parameters are in bold.

	ρ	τ_1	τ_2	τ_3	τ_4	d_{max}	f_1	f_2
Channel self-repulsion	–	1	1	0.15	0	$\boldsymbol{w_{max}}$	0.25	0.25
Well channel data	–	0	1	0.15	0.15	$\boldsymbol{w_{min}}$	1	0
Well inter-channel data	–	1	1	0.15	0	$\boldsymbol{w_{max}}$	0.25	0.25
Well connected channel data	–	0	1	0.25	0.25	$\boldsymbol{w_{min}}$	1	0
Well non-connected channel data	–	1	1	0.15	0	$\boldsymbol{w_{max}}$	0.25	0.25
Probability cube	1	–	–	–	–	–	–	–
Confinement (border 1)	–	1	1	0.15	0	$\boldsymbol{w_{max}}$	0.25	0.25
Confinement (border 2)	–	1	1	0.15	0	$\boldsymbol{w_{max}}$	0.25	0.25

cases, they should be sufficient to ensure the conditioning. If not, the user can still adjust the values.

Among all the parameters, ten parameters remain required in input. Five of them are the weights ϵ associated to each constraint. They allow the user to adjust the relative importance of each constraints. Some weights are the same for different constraints, for instance within the different types of well data. The channel global direction is already required by the L-system. The channel minimal and maximal width and minimal thickness come from the channel width and thickness distributions, which are also required as input by the L-system. The bandwidth for the well channel data is the only parameter with the constraint weights that is not shared with the L-system. It controls on the lateral extension of the perception area, so it determines which data a channel can condition. Hence, this parameter influences the well correlation by the channels.

3.5 Illustration of channel simulation on simple cases

The following sections delve into some aspects of the channel simulation with L-systems. The main purpose is to illustrate the impact of the L-system parameters, of the L-system rules and of the constraints on the simulated channels.

Table 3.9 Input parameters for the user. ϵ_i are constraints weights, t_{min} is the minimal channel thickness, $\hat{D}$ is the global direction of the channel, $w_{a,CD}$ is the perception area bandwidth for the well channel data and w_{min} and w_{max} are the minimal and maximal channel widths.

	Input parameters				
Channel self-repulsion	ϵ_{SR}	$\hat{D}$	–	w_{max}	t_{min}
Well channel data	ϵ_{CD}	$\hat{D}$	$w_{a,CD}$	w_{min}	t_{min}
Well inter-channel data	ϵ_{ICD}	$\hat{D}$	–	w_{max}	t_{min}
Well connected channel data	ϵ_{CD}	$\hat{D}$	–	w_{min}	t_{min}
Well non-connected channel data	ϵ_{ICD}	$\hat{D}$	–	w_{max}	t_{min}
Probability cube	ϵ_{PC}	$\hat{D}$	–	–	t_{min}
Confinement (border 1)	ϵ_C	$\hat{D}$	–	w_{max}	–
Confinement (border 2)	ϵ_C	$\hat{D}$	–	w_{max}	–

3.5.1 L-system parameter impact

Among the L-system parameters, some are pretty straightforward, such as the initial position or the global direction. The default segment length determines the length of the channel segments. It influences bend shape: the larger the segment length, the rougher the bend representation. It also influences the simulation time: for a given channel length, using smaller segments calls for more segments.

Two sets of parameters define the channel morphology: the bend length and the curvature or the bend half-wavelength, the bend amplitude and the deviation angle. They are all given as statistical distributions.

The bend length defines the curvilinear distance along a single bend (figure 3.13). The curvature defines the channel sinuosity. Using different distribution types have less impact than changing the distribution parameters. All the channels of figure 3.13 have a uniform distribution defining the curvature, except the blue channel at the bottom, which has a triangular distribution. The distribution for the bend length remains uniform. The channel behind the blue channel is its equivalent with a uniform distribution for the curvature, so $\mathcal{U}(2 \times 10^{-3}, 4 \times 10^{-3})$ for the curvature and $\mathcal{U}(600, 1\,000)$ for the bend length. The triangular distribution leads to fewer variations of the curvature, as the extremums are less likely to be drawn, but it is quite subtle.

The second set of parameters has a mixed influence on the morphology (figure 3.14). Both the bend half-wavelength and amplitude influence the bend

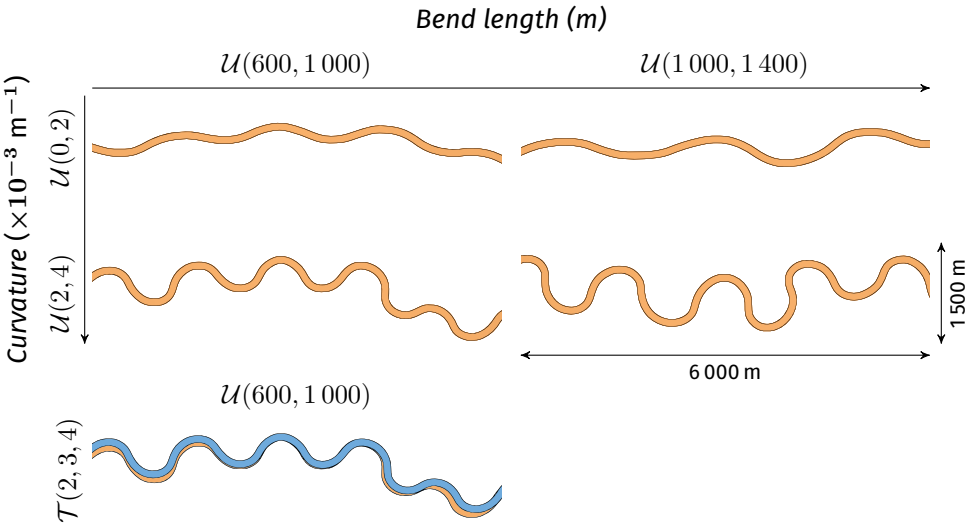


Figure 3.13 Variations of bend length and curvature and influence on the channel morphology. $\mathcal{U}$ is a uniform distribution with a minimum and a maximum value as parameters. $\mathcal{T}$ is a triangular distribution with a minimum, a mode and a maximum as parameters. See the text for more details.

length and the sinuosity. The higher the half-wavelength is, the longer are the bends, the lower is the sinuosity. The higher the amplitude is, the longer are the bends, the higher is the sinuosity. The deviation angle perturbs the resulting morphology. All the channels of figure 3.14 have a uniform distribution defining the deviation angle, with a minimal value of 0 and a maximal value of 0.057. The perturbation remains small. Only the blue channel at the bottom comes from a different uniform distribution, with a minimal value of -11.45 and a maximal value of 11.45. The channel behind the blue channel is its equivalent with the deviation angle distribution $\mathcal{U}(0, 0.057)$. The deviation angle principally perturbs both the bend half-wavelengths and amplitudes. It also perturbs the bend at a lower scale, with more sharp angles between some segments and more various bend shapes.

The global direction has a direct influence on the channel morphology. When its weight increases, the channel tends to follow more strictly the global direction and to be straighter (figure 3.15). The weight should be adjusted depending on how much the channel has to follow the global direction.

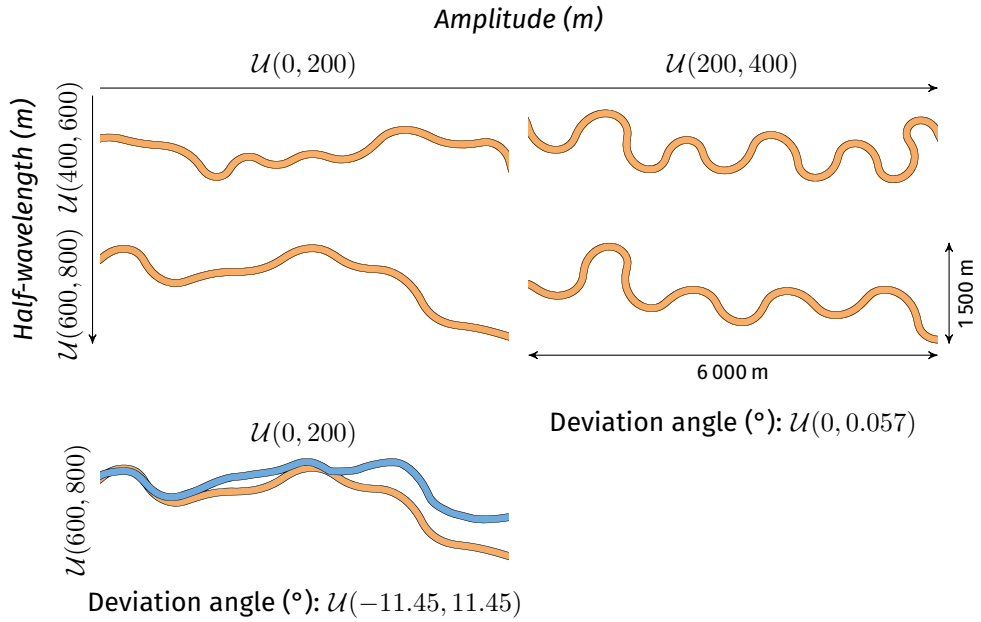


Figure 3.14 Variations of half-wavelength and amplitude and influence on the channel morphology. $\mathcal{U}$ is a uniform distribution with a minimum and a maximum value as parameters. See the text for more details.

3.5.2 L-system rule change

The process of channel simulation with a L-system integrates predefined L-system rules. Predefined rules have two main advantages:

- The method is more easily applied.
- The set of rules gives an example of how to simulate channels and facilitates the design of different rules if required.

This second point is especially important: in a L-system framework, the user is free to modify the pre-existing rules or even to choose its own rules. It gives the opportunity to adapt the simulation to particular cases.

This aspect is further illustrated with two examples of non-stationary channels. The first example relies on channels in submarine fans, whose sinuosity tends to decrease from the start of the fan to the abyssal plain [Babonneau et al., 2010]. This can be integrated in the rules by adding a multiplying factor

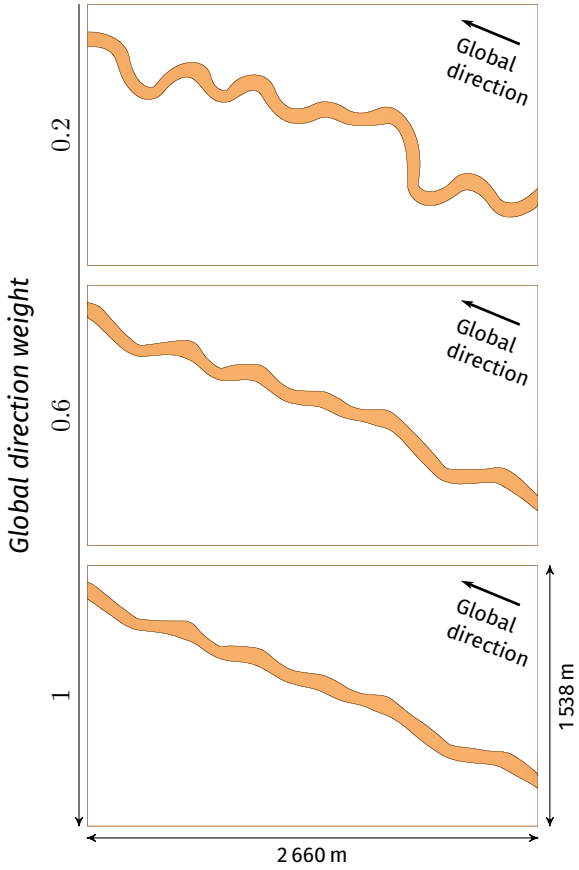


Figure 3.15 Impact of the global direction weight on the simulated channels. The L-system weight is set to one in each case.

to the curvature. It calls for a modification of the initiation rules:

$$p_0 \quad : \quad \begin{array}{lll} T & \xrightarrow{0.5} & \left[+C\{+C\}^{n_{s_1}-1}\,T\right] + (180) - C\{-C\}^{n_{s_2}}\,D \\ T & \xrightarrow{0.5} & \left[-C\{-C\}^{n_{s_1}-1}\,T\right] + (180) + C\{+C\}^{n_{s_2}}\,D \end{array}$$

The developments of each branch are independent by means of a letter D in addition to the letter T. Thus, the development rules are doubled: one set for the T growth, one set for the D. Each branch gets a different multiplication factor for the curvature. The factors introduced here are purely illustrative and do not come from any empirical law. In the branch with the T, the curvature c

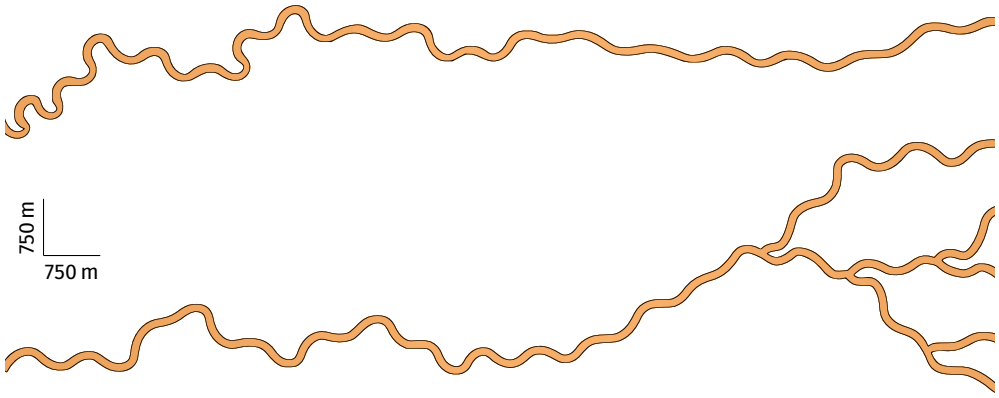


Figure 3.16 Non-stationary channels obtained by modifying the L-system rules. The input parameters are given in table B.1, appendix B.

becomes:

$$c' = c\frac{l_c + 0.5l_C}{2 - 0.5l_C} \tag{3.22}$$

with c' the new curvature value, l_c the current channel length and l_C the wanted channel length. Thus, the curvature decreases as the channel grows. In the branch with the D, the curvature c becomes:

$$c' = c\frac{2 - 0.5l_C}{l_c + 0.5l_C} \tag{3.23}$$

Thus the curvature increases as the channel grows. This way the channel development is non-stationary (figure 3.16, top channel).

The second example uses the same principle to simulate bifurcations along a channel. The initiation rules are modified as previously. The development rules of the T branch becomes:

$$\begin{array}{rcrcrcl}
p_1 & : & -C & < & T & \xrightarrow{0.95} & \{+C\}^{n_s}\, T \\
 & & -C & < & T & \xrightarrow{0.05} & \{+C\}^{n_{s_1}}\, [+C\, T]\, \{+C\}^{n_{s_2}}\, T \\
 & & +C & < & T & \xrightarrow{0.95} & \{-C\}^{n_s}\, T \\
 & & +C & < & T & \xrightarrow{0.05} & \{-C\}^{n_{s_1}}\, [-C\, T]\, \{-C\}^{n_{s_2}}\, T
\end{array}$$

with n_s the total number of segments for the bend and:

$$\begin{aligned}
n_{s_1} &\sim \mathcal{U}(1, n_s - 1) \\
n_{s_2} &= n_s - n_{s_1}
\end{aligned} \tag{3.24}$$

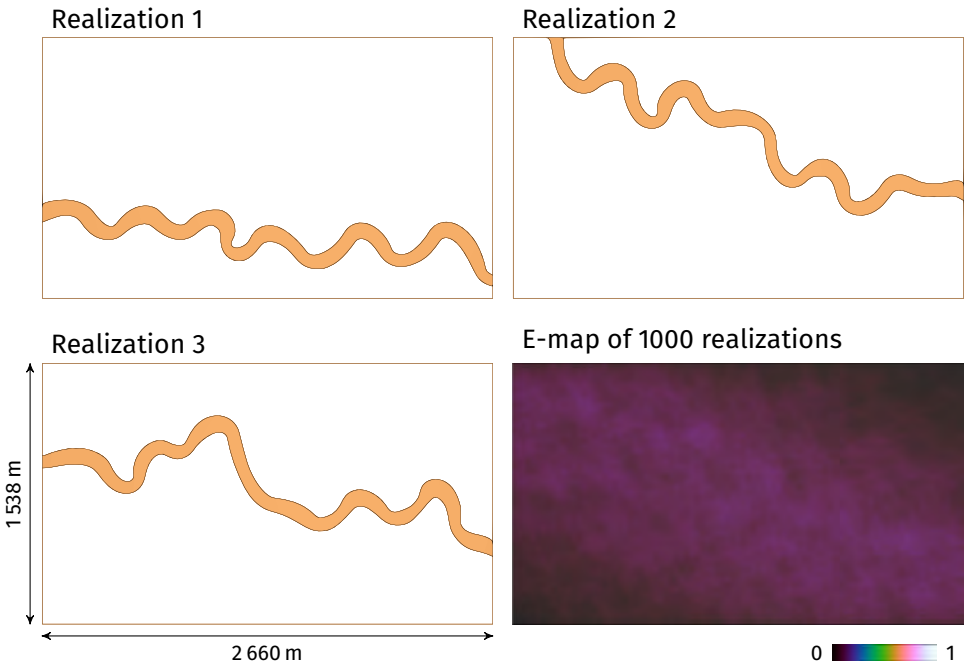


Figure 3.17 Examples of unconditional realizations and E-map of 1 000 realizations. L-system weight: 1; global direction weight: 0.2

with $\mathcal{U}$ a uniform discrete distribution with its minimal and maximal values as parameters and $n_s \geq 2$. This introduces bifurcations in the downstream part of the channel (figure 3.16, bottom channel). The channel self-repulsion limits intersections between the branches. However, when the number of branches becomes higher than four, intersections are difficult to avoid.

3.5.3 Conditioning impact

This section illustrates the impact of conditioning on the L-system and the channel morphology through several synthetic cases. In each of the following cases, the data contain only one channel within the domain. Thus, only one channel constitutes each realization which must condition to all the data. The channel development stops as soon as the two channel branches are outside the grid. The input parameters for the different cases are given in table B.2, appendix B.

Figure 3.17 presents some channels without data conditioning. The bottom

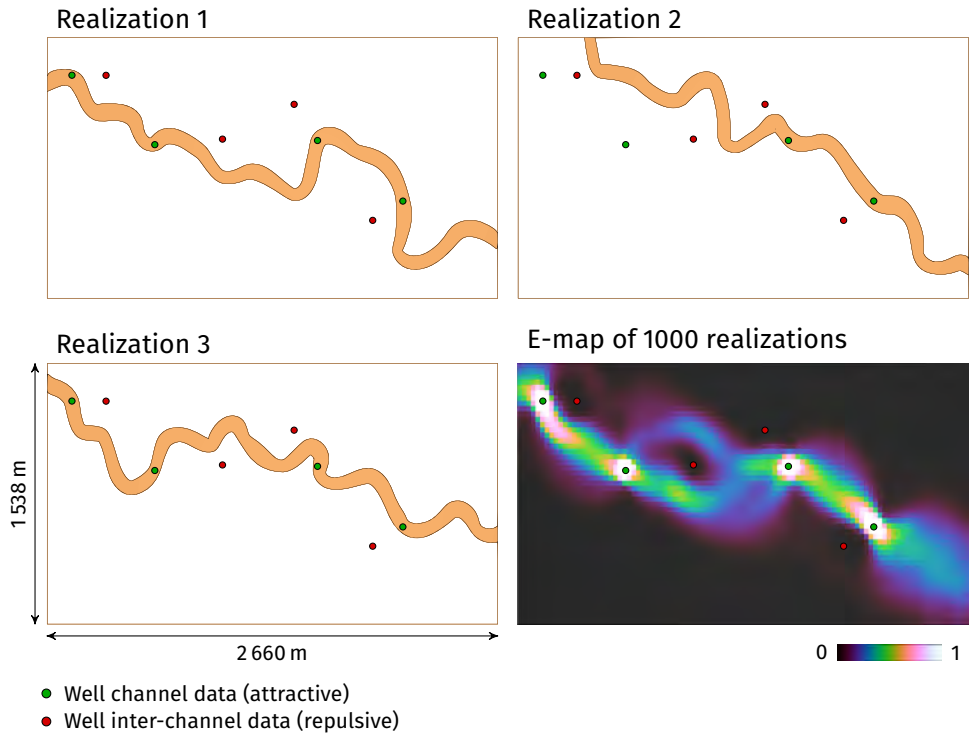


Figure 3.18 Examples of realizations conditioned to well sedimentary data and E-map of 1 000 realizations. L-system weight: 1; global direction weight: 0.2; channel data weight: 1; inter-channel data weight: 1.

right image is the channel count within each cell of a grid. Without particular data, the channels are distributed inside the whole grid.

Well sedimentary data

Figure 3.18 illustrates well sedimentary data conditioning. The channel distribution is not homogeneous any more: the data constraints concentrate the channels along one diagonal of the grid. The conditioning strategy implies a deformation of the channels. These channels tend to be straighter than those without data due to the channel data attraction. The perception area bandwidth is large enough so that this parameter has no influence any more. It favors the conditioning of all the data by a single channel. Figure 3.19 illustrates the influence of the bandwidth: when the bandwidth is small, the L-system has less conditioning possibilities. This gives a control on the well correlation.

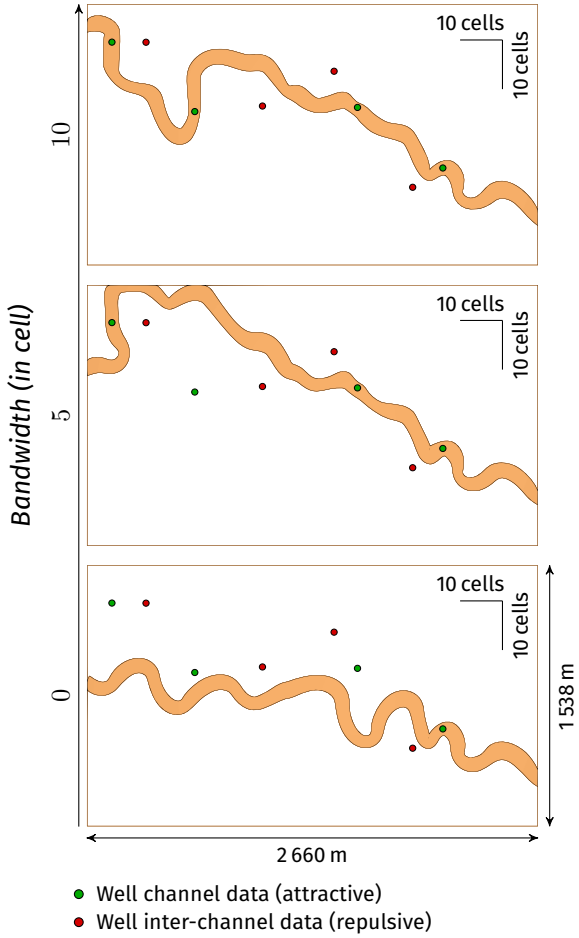


Figure 3.19 Impact of the bandwidth on channels conditioned to well sedimentary data. Global direction weight: 0.2; channel data weight: 1; inter-channel data weight: 1.

Not a single channel conditions a well inter-channel data (figure 3.18). The repulsion effect appears clearly on the E-map. The channels have even some difficulty in coming close to inter-channel data. The constraint magnitude has to be set high enough to ensure that all the channels avoid inter-channel data. In return, this introduces a bias in the conditioning process.

The well channel data are not conditioned by all 1 000 channels (figure 3.18), but by 925 channels at least and 985 channels at best. This reflects the incompatibility that may occur between the channel development and the

data location. Indeed, the channel data to condition may change between two successive channel sections, when the data goes outside the perception area of the second section. It reflects the stochastic well correlation done by the process: all the wells should not necessarily be conditioned by a single channel, and between two realizations the well correlation should be able to change. Depending on their location, inter-channel data can also prevent channel data conditioning (for instance realization 2 on figure 3.18).

The relative force of the L-system weight compared to the channel data weight also influences both the conditioning and the channel morphology. With a high weight for the data and a low weight for the L-system, the channel goes more directly to the data (figure 3.20). It impacts significantly the channel morphology, but it further ensures the conditioning.

The system is able to condition many well channel data (figure 3.21). Progressively decreasing the L-system weight as the number of data increases ensures the conditioning of all the data. Similarly, the weight associated to well inter-channel data is more or less decreased to ensure the channel data conditioning. In return, some channels condition some inter-channel data.

Well connectivity data

Figure 3.22 illustrates well connectivity data conditioning with a L-system. The data set is the same as that of the previous section. Among the four channel data, two have to be connected and one should not be connected to those two data. The last channel data does not have any connectivity information: the channels are free to condition it or not. Again, the perception area bandwidth is set large enough to favor the conditioning of all the data by the channels. The channel initial position is fixed near the most right green square for practical reasons: all the channels start from the connected set and not from the non-connected data.

Contrary to without connectivity data, the two channel data to be connected are always connected, with the 1 000 channels that condition them. The channel morphology tends to be straighter between those data due to the higher constraint magnitude. The non-connected channel data is never connected to the two previous data. The last channel data is conditioned by 999 channels over 1 000. Configuration such as in realization 2 on figure 3.18 do not occur much because the two green-square data are already conditioned. It facilitates the conditioning of the last data as the inter-channel are less bothering. The channels still do not condition the inter-channel data.

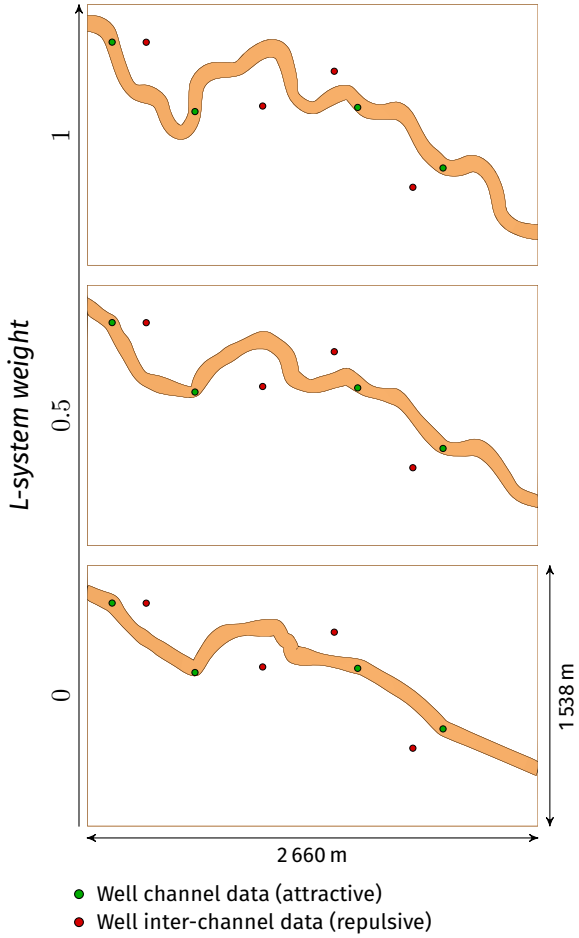


Figure 3.20 Impact of the L-system weight on channels conditioned to well sedimentary data. Global direction weight: 0.2; channel data weight: 1; inter-channel data weight: 1.

Sand probability cube

The sand probability cube is a property defined on each cell of a grid. This cube derives from seismic data for instance. Each grid cell is associated to the probability of having sand-rich deposits in it. Here sand-rich deposits are considered to be associated to a channel. Thus, the sand probability is used to influence the channel development.

The highest probability areas influence the initial position draw and attract

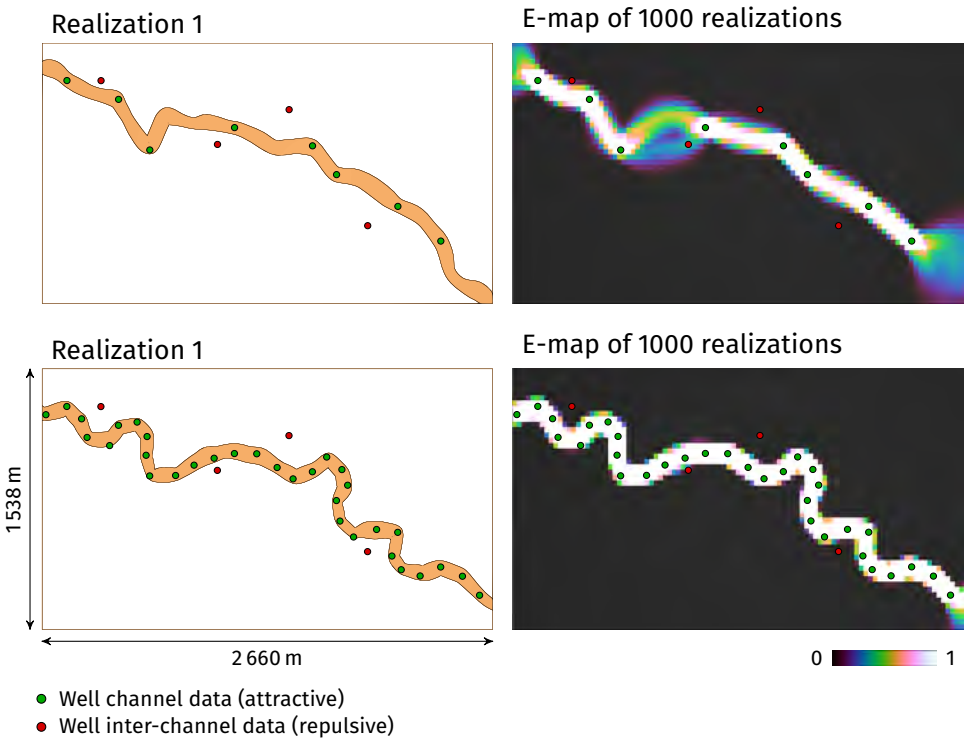


Figure 3.21 Examples of realizations conditioned to well sedimentary data and associated E-maps of 1 000 realizations. L-system weight: 0.4; global direction weight: 0.1; channel data weight: 1; inter-channel data weight: 0.15 (Top). L-system weight: 0.01; global direction weight: 0.002; channel data weight: 1; inter-channel data weight: 0.25 (Bottom).

the channels during their development. When simulating a high number of realizations, the resulting E-map should be similar to the sand probability: the purpose is to reproduce the probability cube.

Here the channels tend to follow the diagonal of high probabilities (figure 3.23). It appears clearly on the E-map, with a similar diagonal that concentrates more channels than the rest of the domain. However a significant difference remains between the probability values and the E-map values.

Similarly to well channel data, lowering the L-system weight compared to the probability cube weight lets the channels get closer to the high probability values (figure 3.24). This is still at the cost of the channel morphology. The resulting E-map is less different to the probability cube. Some differences are

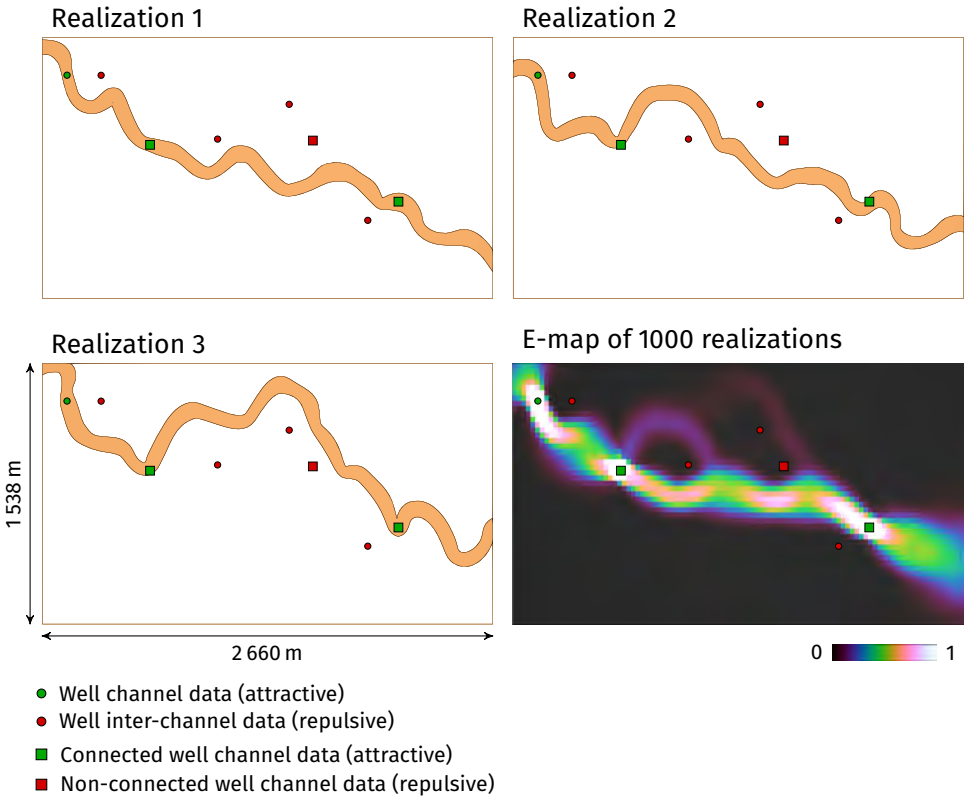


Figure 3.22 Examples of realizations conditioned to well sedimentary and connectivity data and E-map of 1 000 realizations. L-system weight: 1; global direction weight: 0.2; channel data weight: 1; inter-channel data weight: 1.

due to the discontinuity of the high probability areas, whereas the channels are necessarily continuous. It highlights a difference between a sand probability, which can vary within a channel, and a channel probability. When the probability cube weight is high relatively to the L-system weight, the influence of the global direction is fundamental. It ensures that the channels continue their development without going backward. Indeed, the highest sand probability areas are circular and can sometimes re-attract the channel after its first passage.

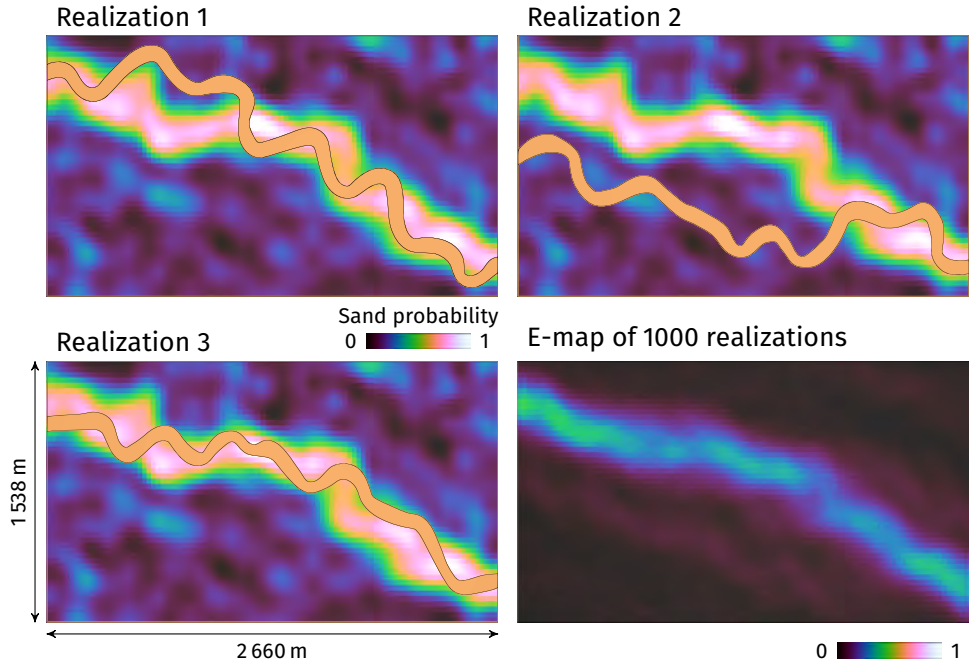


Figure 3.23 Examples of realizations conditioned to sand probability map and E-map of 1 000 realizations. L-system weight: 1; global direction weight: 0.2; probability cube weight: 0.2.

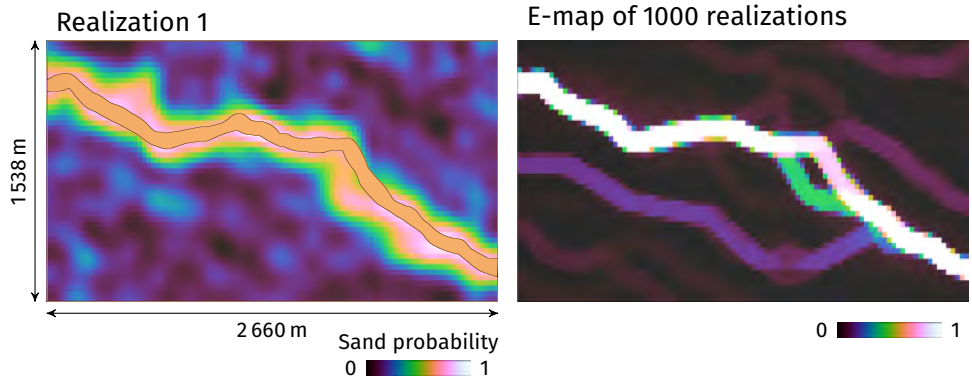


Figure 3.24 Example of realization conditioned to a sand probability map and E-map of 1 000 realizations. L-system weight: 0.2; global direction weight: 0.2; probability cube weight: 1.

3.6 Application

In this section, a synthetic test case with several data of different types illustrates the application of the method. It requires the combination of several constraints within a three-dimensional environment.

3.6.1 Dataset

This synthetic case is inspired by turbiditic systems. In such environment, channels evolve within a master channel. Continuous and discrete migrations control channel evolution. Channel evolution tends to depend on the energy of a turbiditic event [McHargue et al., 2011]. At the beginning of an event, the high energy mostly leads to erosion. As the energy decreases, deposition of sediment starts. Channels mostly migrate laterally, with very few aggradation and a stacking between channels that tend to be disorganized due to discrete migrations. When the energy is low enough, the sediment content of the turbidity currents become muddier. This muddier content favors the development of large overbanks. The flow is then more confined, with less lateral migrations. Aggradation dominates this phase, with a significant vertical migration. The purpose is to reproduce that evolution in two successive channel development trends, one horizontal at the bottom of the master channel and one vertical at the top, in the channel realizations.

The master channel is simulated using the channel simulation process. A L-system defines its morphology and NURBS surfaces represent its boundaries. The NURBS surfaces help to build a hexahedral grid aligned with the master channel margins [Ruiu et al., 2015b] (figure 3.25). This grid materializes the master channel and corresponds to the simulation domain. The channels are simulated within the parametric space of the grid. Having the grid aligned with the master channel margins helps the channels to follow the master channel orientation.

The information about the channel evolution comes from a sand probability cube:

- In the lower part of the master channel, the high sand probabilities occupy almost the whole master channel width due to the significant lateral migration.

- In the upper part, the high sand probabilities occupy only a small part of the master channel width due to the significant aggradation.

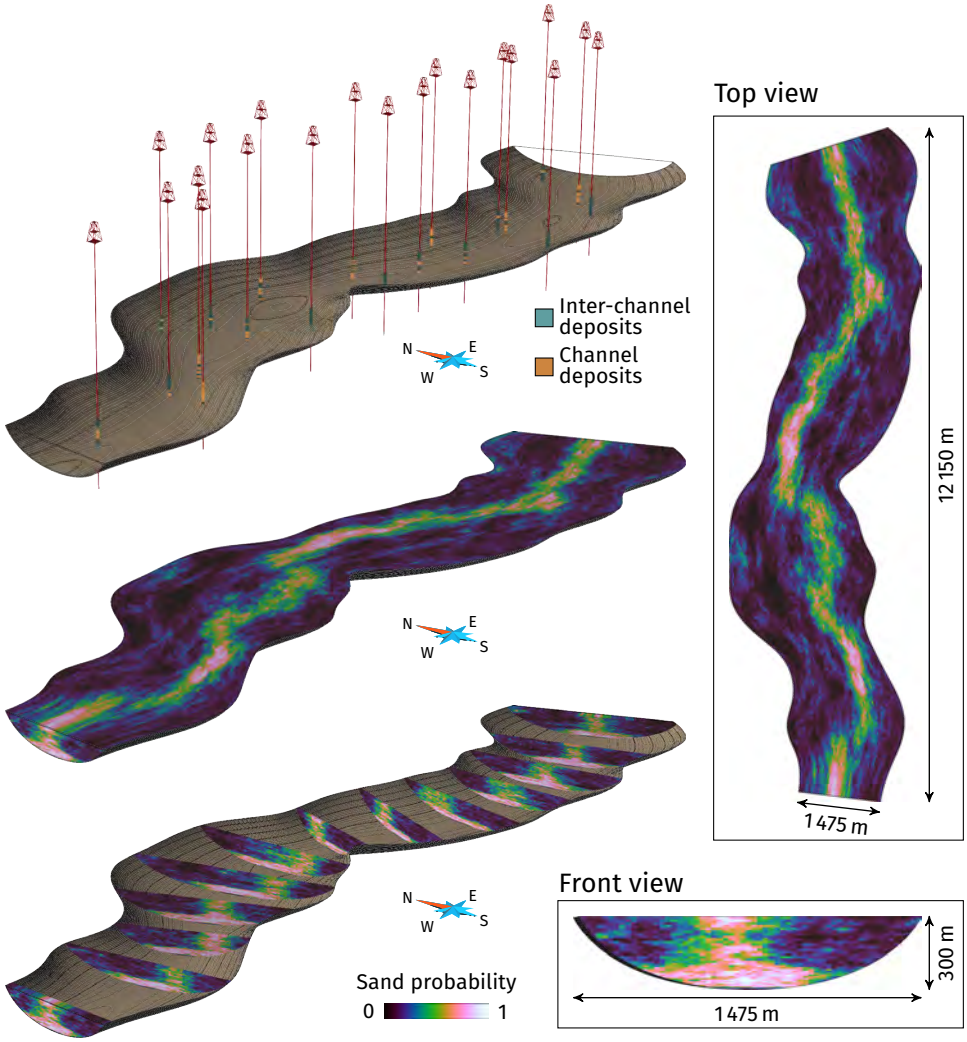


Figure 3.25 Dataset of the application: a curvilinear grid representing a master channel with a sand probability cube. Twenty wells are positioned along the canyon. Well data are extracted along each well from one simulation.

In a real case, the master channel could come from a seismic interpretation and the sand probability cube could derive from the seismic amplitude.

Three sets of channel realizations are simulated within the master channel:

- The first set does not have other constraint than the repulsion from the

confinement, i.e., the master channel.

- The second set must also condition the sand probability cube.

- The last set must condition the confinement, the sand probability cube and the well data. The well data are extracted from a realization of the second set along 20 well paths (figure 3.25). Among these data, well channel data attract the channels and well inter-channel data repulse the channels.

The input parameters for the different sets are given in table B.3, appendix B.

3.6.2 Simulation principle

The principle of the simulation process is simple: channels are simulated until a given number of channels is reached. Without data the channel initial position is randomly drawn inside the grid. The only constraint is that the channels can not go outside the master channel margins. This implies that the initial position can not be drawn at less than half a channel width from a canyon side. When adding the probability cube, the initial position is influenced by the sand probabilities.

When adding well data, the process is a bit different. It still calls for a number of channels. As long as there are non-conditioned channel data, channels are simulated. The initial position of those channels are drawn close to channel data, and they are then attracted or repulsed by the other data. Once all the well channel data are conditioned and if the channel number is lower than the target channel number, new channels are simulated until this target is reached.

3.6.3 Unconditional realizations

Some unconditional realizations have been simulated within the grid (figure 3.26). The only constraints are the two repulsive constraints from the master channel margins. The resulting channel distribution is quite uniform in the middle of the grid, but the probability to have a channel is lower close to the master channel margins (figure 3.27). However, the channels can still occupy the whole width of the master channel. And the channel morphology remains sinuous, without any large orientation change between two successive segments. It is only at the master channel bottom that the margin influence is such that the channels are straighter (figure 3.28).

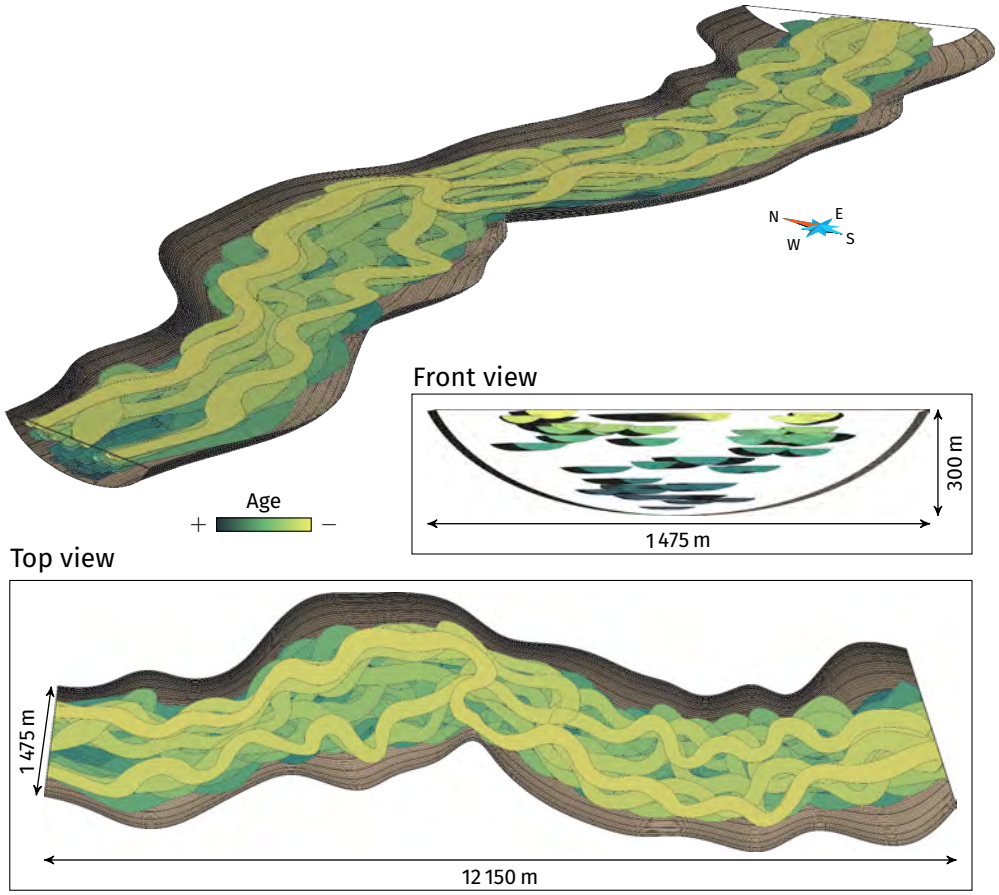


Figure 3.26 Example of a realization containing 40 channels within a master channel.

The constraints influence the channel section position. Here the efficiency of the repulsive constraints of the margins is such that no channel location ever goes outside those margins. All channels go through the whole master channel.

3.6.4 Realizations conditioned to the sand probability cube

Adding the sand probability cube as a constraint limits the horizontal dispersion of the channels (figure 3.29). The influence of the cube is significant in the upper part of the master channel (figure 3.30 and 3.31). The channels tend to gather along a more or less similar path along the high probability areas. The

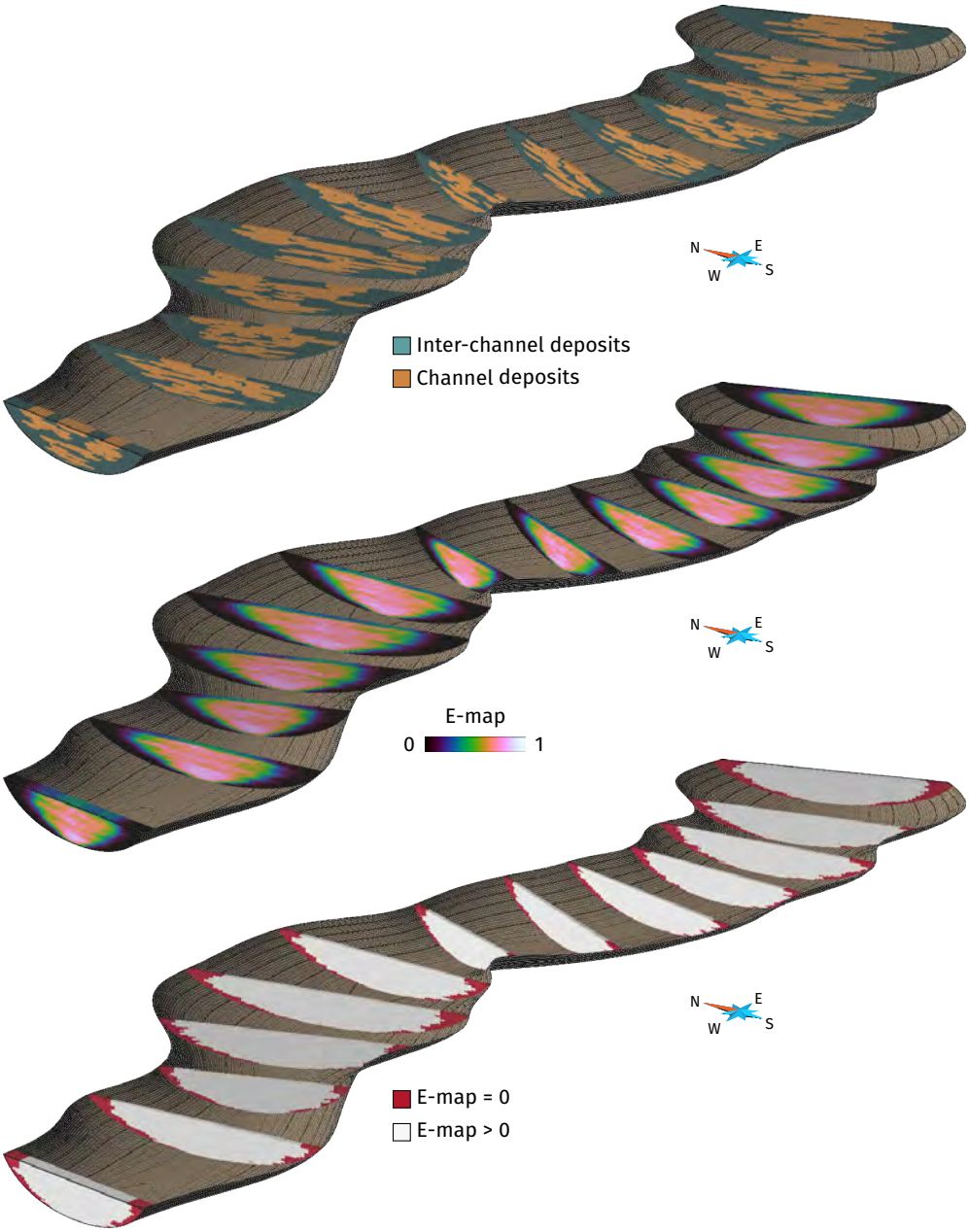


Figure 3.27 Details on the channel distribution inside the master channel for the realization in figure 3.26 rasterized in the master channel grid. The E-map illustrates the mean channel distribution over the 100 unconditional realizations.

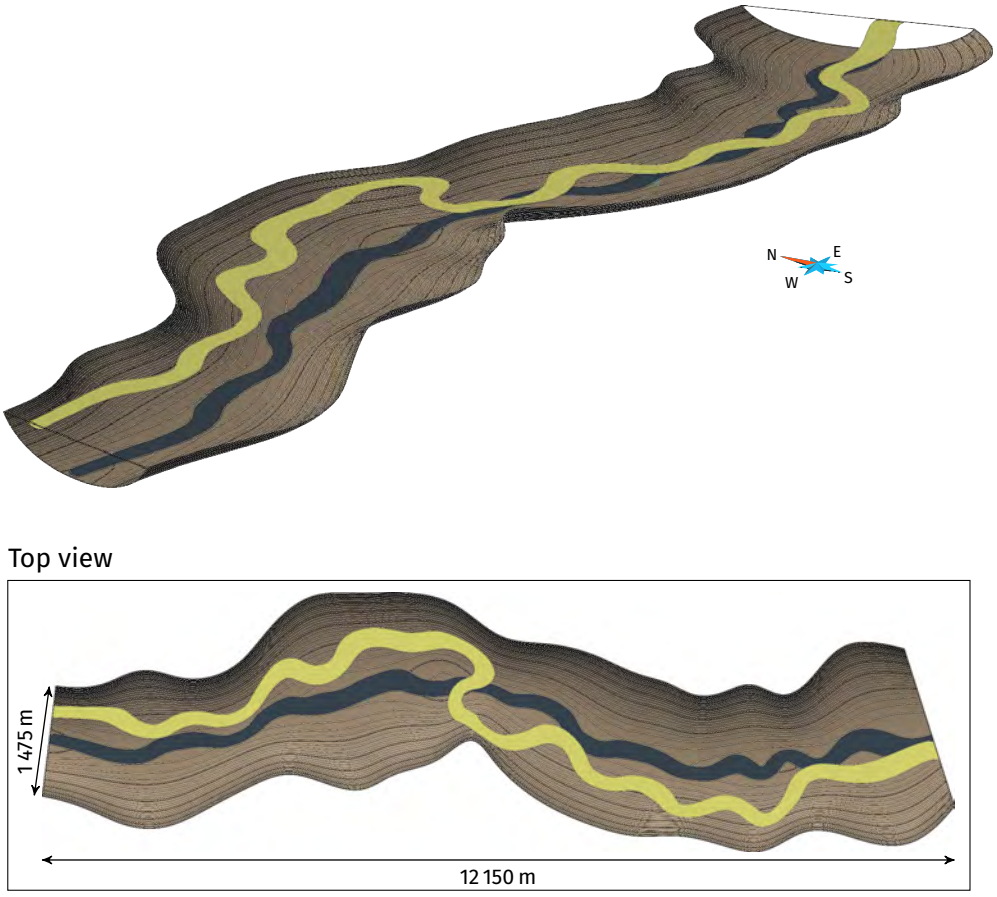


Figure 3.28 Channels at the top and at the bottom of the realization in figure 3.26.

channels in the upper part locally tend to move apart from the vertical high probability trend and to get close to the master channel margins. These shifts are often related to lower probability areas. In the lower part, the influence of the margins is such that the channels are often straighter, with less lateral variations. These aspects tend to blur the transition from a significant lateral development to a more vertical evolution (figure 3.31).

The duality of the channel evolution introduced in the probability cube is more or less clear depending on the realization and the location within the master channel. On average, the vertical trend seems to have a too huge influence on the channels (figure 3.31).

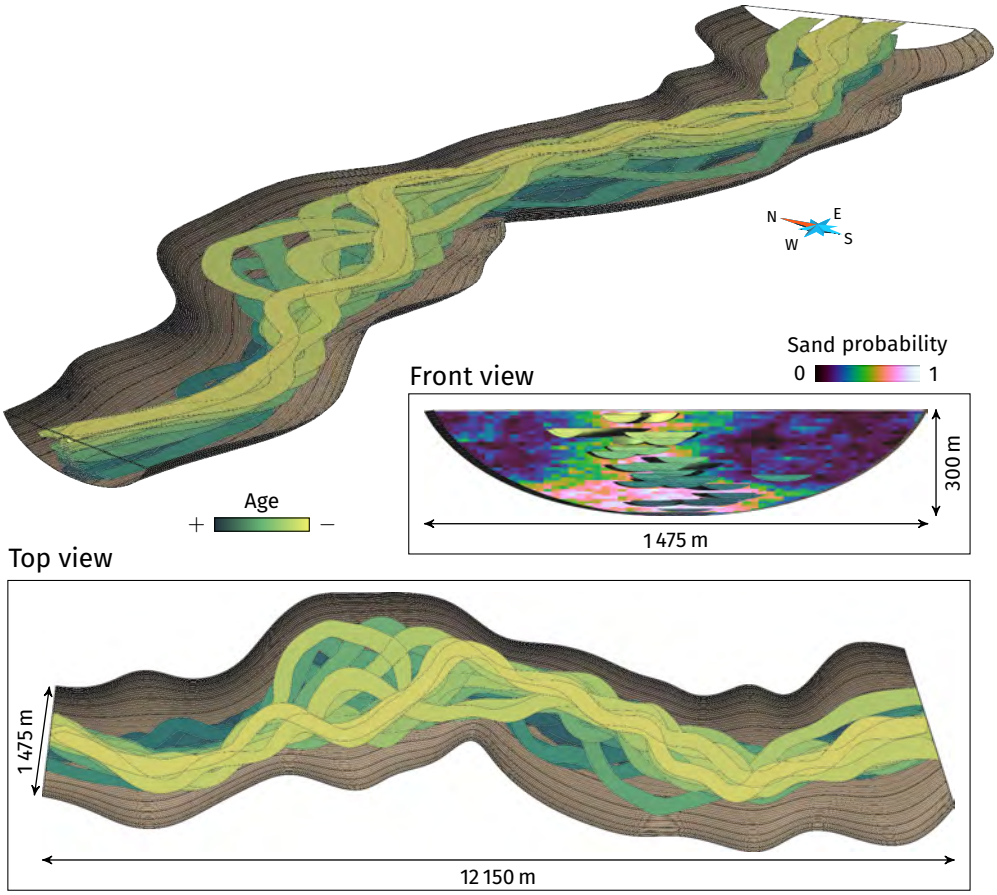


Figure 3.29 Example of a realization containing 40 channels conditioned to a sand probability cube within a master channel.

3.6.5 Realizations conditioned to the sand probability cube and the well sedimentary data

The last set of realizations takes into account the sand probability cube but also the well sedimentary data (figure 3.32). The well data are distributed along 20 wells and form a set of 625 points. The point spacing along a well depends on the master channel grid cells, with one point per cell. The channel and inter-channel data come from the realization presented in section 3.6.4. 234 points are channel data and 391 points are inter-channel data. The channels are free to evolve within the whole master channel and to condition channel

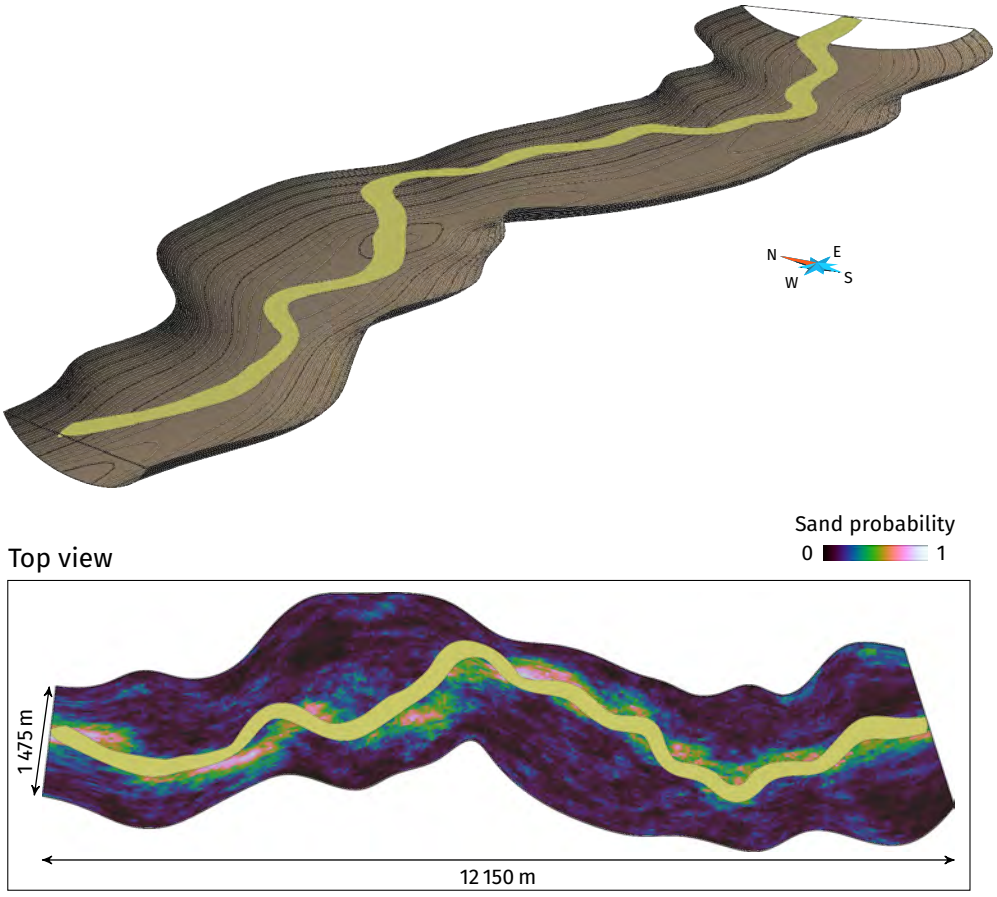


Figure 3.30 Channel at the top of the realization in figure 3.29.

data along the whole master channel width. Thus the channel data bandwidth for the perception area is set to half the master channel grid width.

This case introduces many constraints:

- The two repulsive constraints from the two master channel margins.
- The attractive constraint from the sand probability cube.
- Possibly one attractive constraint from a channel data, depending on the channel development and the perception area.
- Possibly several repulsive constraints from the inter-channel data, depending on the channel development and the perception area.