
Approche statistique multipoints

Dans ce chapitre nous présentons l'approche statistique multipoints et discutons ses avantages et ses inconvénients. Nous commençons par en donner le principe ainsi que l'algorithme initial proposés par Guardiano et Strivastava (1993). Nous décrivons ensuite la mise en œuvre de cette approche en détaillant la méthode de stockage des statistiques empiriques et de leur réutilisation proposée par Strebelle (2002). Une attention particulière sera portée aux conditions limites d'application de l'approche MP. Soulignons que tous les exemples présentés dans ce chapitre sont effectués par simulation stationnaire. En fin de chapitre, nous abordons le problème central de cette étude – la simulation non stationnaire.

2.1 Présentation générale

2.1.1 Principe

Le principe de base de l'approche MP est d'utiliser l'image d'apprentissage, supposée représentative de la structure géométrique de réservoir, comme la source des statistiques empiriques d'ordre multiple. À partir de ces statistiques, la loi spatiale d'une fonction aléatoire est construite. Cette loi permet de reproduire une structure géométrique complexe. L'algorithme de simulation MP s'appuie ainsi sur trois idées :

- 1) utilisation d'une image d'apprentissage ;
- 2) acquisition des statistiques MP à partir de cette image;
- 3) simulation séquentielle à partir des statistiques MP acquises.

L'algorithme ne mettant pas en jeu de forme géométrique de façon explicite, leurs caractéristiques (taille, orientation, longueur etc.) doivent être reproduites de façon implicite. Les exemples présentés (Fig. 4-7) ont été choisis de natures différentes (modèle booléen, modèle gaussien, réseau de fractures) pour illustrer les possibilités de l'approche multipoints.

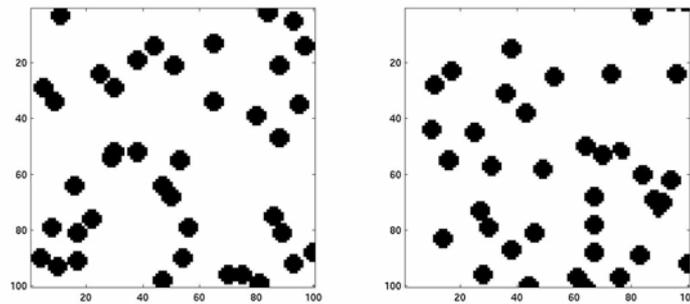


Fig. 4 Simulation MP stationnaire. Reproduction d'un modèle booléen avec des objets ronds : l'image d'apprentissage à gauche, la réalisation MP à droite.

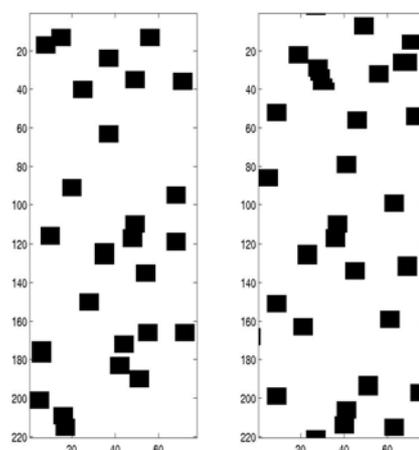


Fig. 5 Simulation MP stationnaire. Reproduction d'un modèle booléen avec des objets carrés : l'image d'apprentissage à gauche, la réalisation MP à droite.

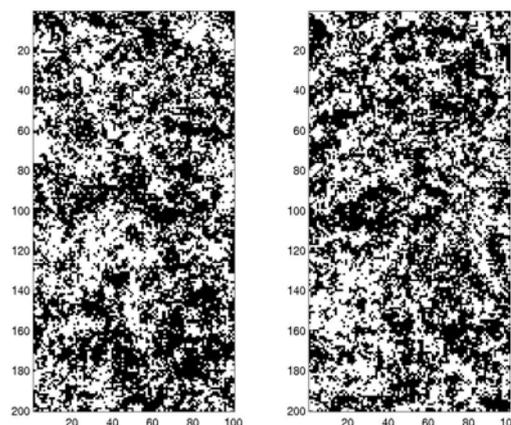


Fig. 6 Simulation MP stationnaire. Reproduction d'un modèle gaussien seuillé (variogramme hyperbolique avec une portée unité) : l'image d'apprentissage à gauche, la réalisation MP à droite.

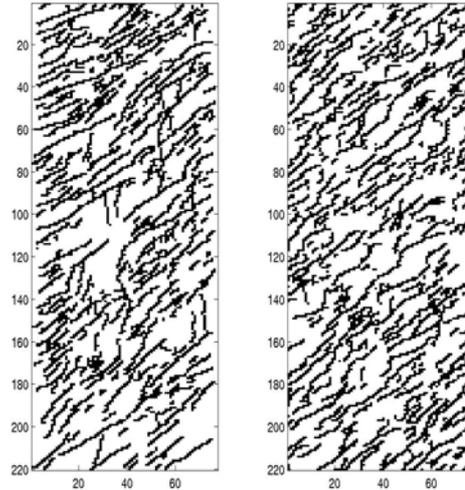


Fig. 7 Simulation MP stationnaire. Reproduction d'un réseau de fractures en 2D : l'image d'apprentissage à gauche, la réalisation MP à droite.

2.1.2 Algorithme initial

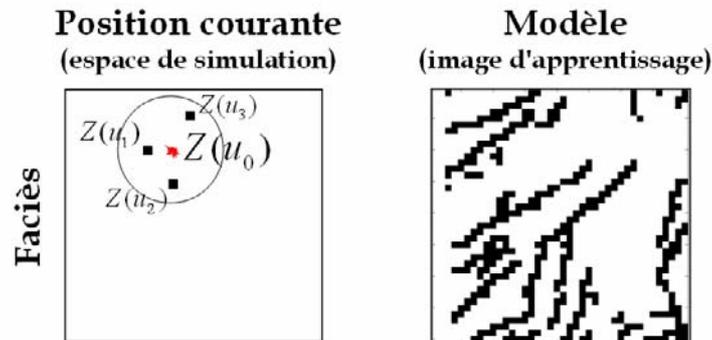


Fig. 8 L'espace de simulation à gauche ; l'image d'apprentissage à droite.

Ramenée à un cadre discret, l'approche MP se présente de la manière suivante :

Soit \mathbf{I} la grille de discrétisation de l'Image d'apprentissage, et soit \mathbf{E} la grille de l'Espace de simulation. Usuellement, \mathbf{I} et \mathbf{E} ont la même dimension physique (2 ou 3) et ont la même résolution spatiale. Par contre, leur taille peut être très différente. À chaque point u de l'une de deux grilles est attaché, ou bien doit être attaché, un attribut discret $Z(u)$ (fonction aléatoire) qui peut prendre K états correspondants aux K faciès différents. La simulation MP consiste à générer une valeur d'attribut en chaque point de \mathbf{E} .

Supposons que l'on veuille simuler Z au point $u_0 \in \mathbf{E}$ (Fig. 8). Soient $u_1, u_2, \dots, u_n \in \mathbf{E}$ un ensemble des points ordonnés composé par les points déjà simulés au tour de u_0 (Fig. 9).

Déf. On appelle voisinage de u_0 toute famille finie de points ordonnés $\{u_1, u_2, \dots, u_n\}$ ne contenant pas u_0 . Cet ensemble de points sera noté \mathbf{u} dans ce qui suit.

Déf. Le voisinage \mathbf{u} du point u_0 est équivalent au voisinage \mathbf{w} du point w_0 s'ils ont le même nombre de points et s'ils coïncident à une translation près :

$$\mathbf{u} \equiv \mathbf{w} \Leftrightarrow (u_1 - u_0, u_2 - u_0, \dots, u_n - u_0) = (w_1 - w_0, w_2 - w_0, \dots, w_n - w_0).$$

Soit $\mathbf{z} = (z_1, z_2, \dots, z_n)$ un *état* (une réalisation) de Z sur un voisinage \mathbf{u} . On écrit $Z(u_1, u_2, \dots, u_n) = (z_1, z_2, \dots, z_n)$ ou, simplement, $Z(\mathbf{u}) = \mathbf{z}$.

Déf. Dire que $Z(\mathbf{w}) = \mathbf{y}$ est une *réplique* de $Z(\mathbf{u}) = \mathbf{z}$ est signifie l'identité des voisinages $\mathbf{w} \equiv \mathbf{u}$ et l'égalité des valeurs $\mathbf{y} = \mathbf{z}$.

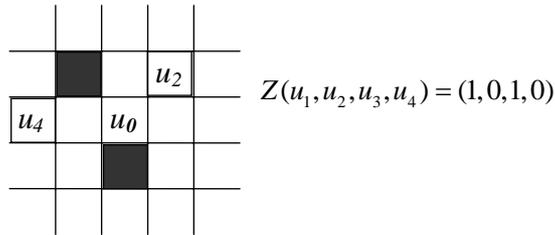


Fig. 9 Un voisinage simulé (*état*) du point u_0 . Cas 2D.

Les valeurs de $Z(\mathbf{I})$ étant structurées, la valeur d'un point $w_0 \in \mathbf{I}$ dépend des valeurs de ses points voisins $\mathbf{w} \subset \mathbf{I}$. Pour pouvoir reproduire la structure, la valeur en un point $u_0 \in \mathbf{E}$ est générée en fonction des points déjà simulés de son voisinage $\mathbf{u} \subset \mathbf{E}$. Ainsi, pour simuler la variable aléatoire $Z(u_0)$, il nous faut estimer la loi conditionnelle $P[Z(u_0) | Z(\mathbf{u}) = \mathbf{z}]$. Cette loi est extraite de l'image d'apprentissage en enregistrant l'ensemble des répliques de $Z(\mathbf{u}) = \mathbf{z}$. La valeur de u_0 est ensuite générée selon cette loi conditionnelle discrète (Fig. 10). La simulation s'arrête quand tous les points de l'espace de simulation ont été traités. L'algorithme de simulation MP s'écrit alors ainsi :

- (i) Choisir un point non encore simulé $u_0 \in \mathbf{E}$ et en extraire l'état de son voisinage $Z(\mathbf{u}) = \mathbf{z}$;
- (ii) Enregistrer toutes les répliques de $Z(\mathbf{u}) = \mathbf{z}$ dans l'image d'apprentissage :
 $W = \{w_0 \in \mathbf{I} : Z(\mathbf{w}) = \mathbf{z}, \mathbf{w} \equiv \mathbf{u}\}$;
- (iii) Si $W = \phi$, faire $\mathbf{u} = \mathbf{u} \setminus u_n$ et retourner en (ii) ;

(iv) Simuler la valeur de Z au point u_0 selon la loi conditionnelle :

$$P[Z(u_0) = k \mid Z(\mathbf{u}) = \mathbf{z}] = \frac{P[Z(u_0) = k, Z(\mathbf{u}) = \mathbf{z}]}{P[Z(\mathbf{u}) = \mathbf{z}]} .$$

$$= \frac{\#\{w_0 \in W : Z(w_0) = k\}}{\#\{w_0 \in W\}}, \quad k = 1, 2, \dots, K$$

où le symbole $\#$ désigne la cardinalité ensembliste.

(v) Retourner en (i).

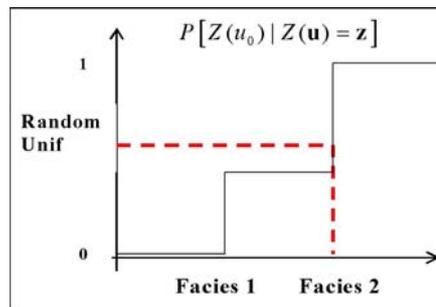


Fig. 10 Simulation d'une fonction aléatoire à partir d'une loi empirique pour deux faciès.

2.2 Mise en oeuvre

L'algorithme, initialement proposé par Guardiano et Strivastava (1993), présente l'inconvénient d'être particulièrement gourmand en mémoire et en temps calcul. En effet, le scanning de l'image d'apprentissage est nécessaire pour la simulation de chaque point. Strebelle (2000) a proposé un algorithme permettant de ne scanner qu'une seule fois l'image d'apprentissage, et de mémoriser les statistiques multipoints sous une forme arborescente (nommée ensuite *arbre de stockage*) afin de pouvoir en disposer plus rapidement.

Une fois l'image d'apprentissage scannée, la simulation peut débuter. Les points sont simulés selon un chemin prédéfini (aléatoire ou non) qui parcourt tout l'espace de simulation. Pour chaque point traité, une loi conditionnelle est extraite directement de l'arbre de stockage, correspondant à l'état de son voisinage. La simulation s'arrête quand tous les points du champ de simulation ont été traités. Néanmoins, il existe dans la littérature des algorithmes MP de post-simulation, qui consistent à resimuler des points afin de raffiner la reproduction de la géométrie. Nous discuterons de cette problématique au paragraphe 2.3.3. Dans le cadre de cette thèse, pour ne pas trop nous disperser, nous utilisons l'algorithme MP où chaque point n'est simulé qu'une seule fois.

Voici un diagramme de l'algorithme MP (Fig. 11).

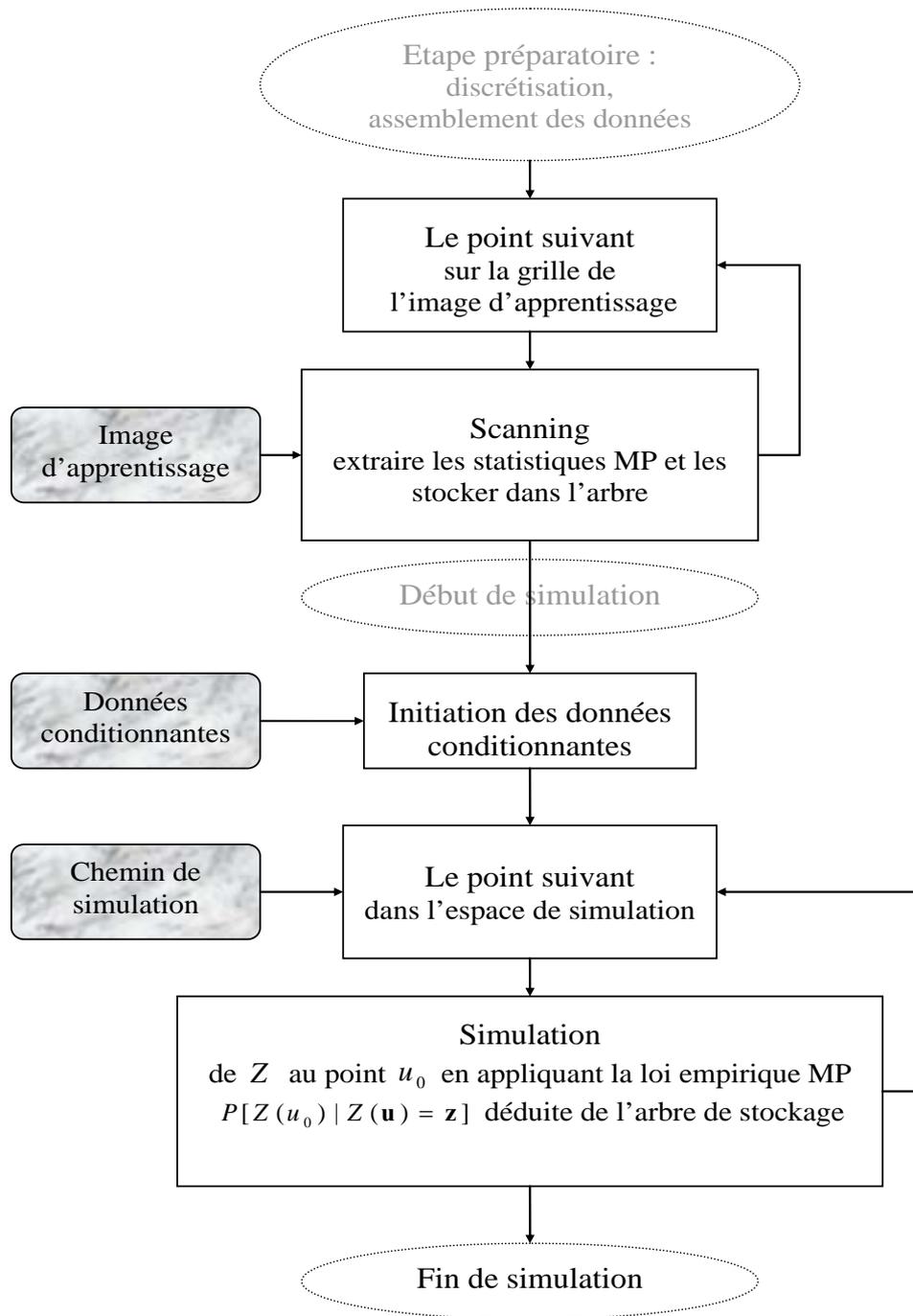


Fig. 11 Algorithme de simulation statistique MP.

2.2.1 Scanning de l'image d'apprentissage

L'étape de scanning de l'algorithme MP consiste à dériver et stocker les statistiques MP au moyen d'une fenêtre glissante de taille fixe (dite *template* dans la littérature anglophone), qui explore toute l'image d'apprentissage point après point (Fig. 12). À l'intérieur de cette fenêtre, chaque état sera enregistré dans l'arbre de stockage.

v_{11}	v_{12}	v_{13}	v_{14}	v_{15}
v_{10}	v_5	v_1	v_6	v_{16}
v_9	v_4	w_0	v_2	v_{17}
v_{24}	v_8	v_3	v_7	v_{18}
v_{23}	v_{22}	v_{21}	v_{20}	v_{19}

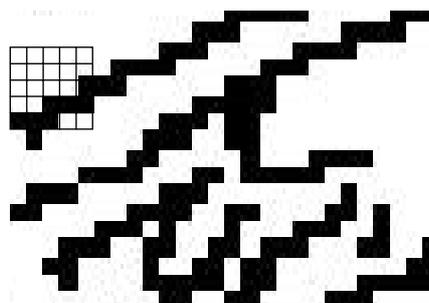


Fig. 12 Fenêtre d'exploration avec un point central w_0 . Cas 2D.

Soit $\mathbf{v}^l = (v_1, v_2, \dots, v_l)$; $l: 1, \dots, L$, un voisinage du point w_0 qui contient l'ensemble de points les plus proches de w_0 , L étant le nombre maximal de ces points. La configuration de cet ensemble définit la zone de conditionnement du point w_0 , c'est-à-dire que seules les valeurs de cet ensemble de points seront prises en compte pour dériver la loi conditionnelle de $Z(w_0)$. La numérotation de points dans un tel ensemble définit l'ordre de prise en compte des points voisins. Nous appelons un ensemble \mathbf{v}^L une *fenêtre*, des sous-ensembles $\mathbf{v}^l, l < L$ – des *sous-fenêtres* et w_0 – le *point central* de fenêtre. Les valeurs prises par la fonction aléatoire sur une fenêtre se dénotent par $Z(v_1, v_2, \dots, v_l) = (z_1, z_2, \dots, z_l)$ ou, simplement, $Z(\mathbf{v}^l) = \mathbf{z}^l$.

Pendant le scanning le point central se déplace sur toute l'image d'apprentissage et enregistre les états de la fenêtre et de ses sous fenêtres dans l'arbre de stockage. Cet arbre est constitué des noeuds de niveau entre 0 et L (la taille maximale d'une fenêtre d'exploration a été préalablement choisie pour scanner l'image d'apprentissage). Un niveau l correspond à la sous fenêtre \mathbf{v}^l et peut contenir autant de noeuds qu'il existe d'états différents de $Z(\mathbf{v}^l)$ dans l'image d'apprentissage. A chaque noeud de l'arbre de stockage sont associés 1) un état particulier $Z(\mathbf{v}^l) = \mathbf{z}^l$; 2) les nombres de répliques de cet état pour les différentes valeurs du point central. L'ensemble de cette information est équivalent à une loi empirique $P[Z(u_0) | Z(\mathbf{v}^l) = \mathbf{z}^l]$, $u_0 \in I$.

Le schéma de l'arbre de stockage est présenté dans le cas général de K faciès (Fig. 13). La Fig. 14 montre l'exemple d'un noeud avec quatre voisins définis dans le cas des deux faciès.

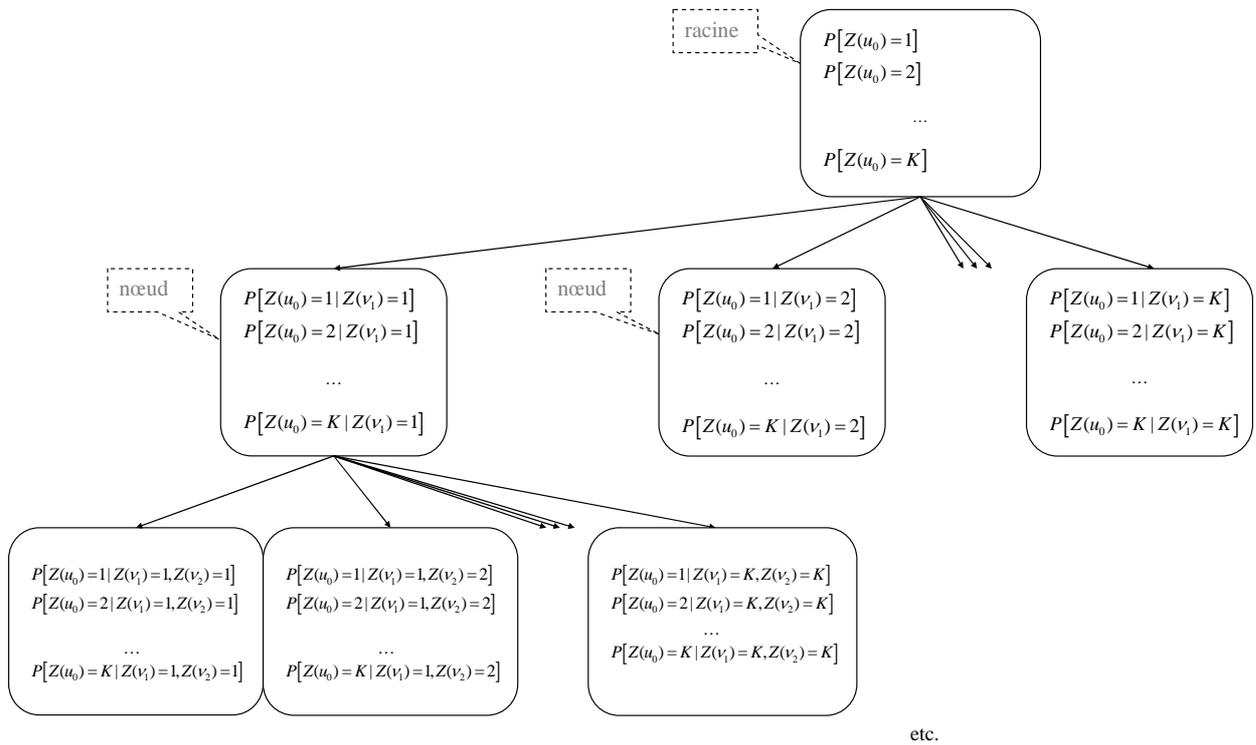


Fig. 13 L'arbre de stockage pour les statistiques MP extraites de l'image d'apprentissage. Cas général de K faciès.

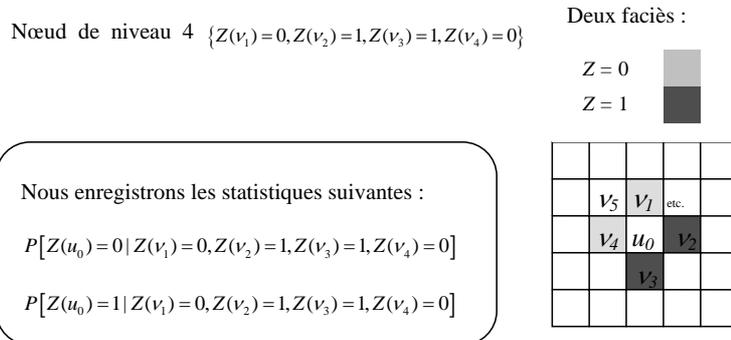


Fig. 14 Exemple d'un nœud de niveau $l = 4$ pour deux faciès.

2.2.2 Simulation

La simulation commence par le positionnement des données fixes et certaines dans l'espace de simulation. (En pratique, ce sont les données de puits dont la position et les faciès géologiques sont connus.) Au cours de la simulation, chaque point simulé devient lui-

même un point fixe. Les points déjà évalués (d'une manière ou d'une autre) sont considérés comme des points simulés.

Soit \mathbf{u} le voisinage simulé du point u_0 . Ce voisinage est un sous ensemble de fenêtre d'exploration, mais pas nécessairement une sous fenêtre $\mathbf{u} \subseteq \mathbf{v}^l$, $\mathbf{z} \subseteq \mathbf{z}^l$. Ainsi, le nombre maximal des points simulés qui seront prises en compte pendant la simulation est limité par la taille de fenêtre d'exploration. L'exemple d'un voisinage simulé de trois points est présenté sur la Fig. 15.

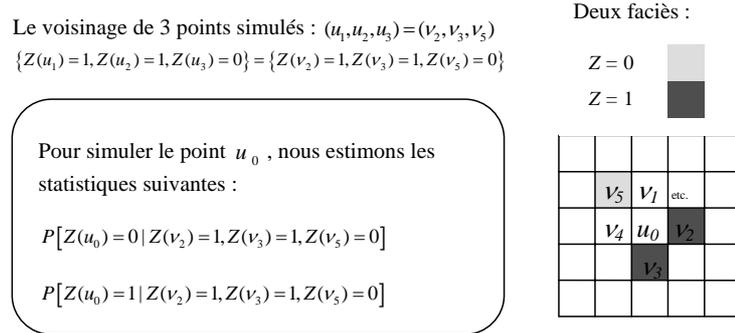


Fig. 15 Exemple d'un voisinage simulé pour deux faciès.

En chaque point de simulation $u_0 \in \mathbf{E}$, il est nécessaire de connaître la loi conditionnelle :

$$P[Z(u_0) | Z(\mathbf{u}) = \mathbf{z}] \tag{Eq 2.1}$$

Pour estimer cette loi, on considère l'algorithme suivant :

- (i) Choisir le point suivant $u_0 \in \mathbf{E}$ selon le chemin prédéfini et extraire l'état de son voisinage $Z(\mathbf{u}) = \mathbf{z}$
- (ii) Choisir le niveau d'arborescence qui est la taille de la plus petite sous fenêtre contenant le voisinage simulé :

$$l = \min_{0 \leq l \leq L} \{ \mathbf{u} \subseteq \mathbf{v}^l \}.$$

- (iii) Parmi l'ensemble des nœuds de ce niveau, retenir ceux qui coïncident avec l'état du voisinage $Z(\mathbf{u}) = \mathbf{z}$:

$$W = \{ w_0 \in \mathbf{I} : Z(\mathbf{v}^l) = \mathbf{z}^l, \mathbf{z} \subseteq \mathbf{z}^l \}.$$

- (iv) Si $W = \emptyset$, faire $\mathbf{u} = \mathbf{u} \setminus v_l$ et retourner en (ii) ;
- (v) Calculer la loi conditionnelle :

$$P[Z(u_0) = k | Z(\mathbf{u}) = \mathbf{z}] = \frac{\#\{w_0 \in W : Z(w_0) = k\}}{\#\{w_0 \in W\}}, \quad k = 1, 2, \dots, K \tag{Eq 2.2}$$

- (vi) Retourner en (i).

2.3 Conditions d'application

Les conditions d'application de l'approche MP sont dictées par des considérations théoriques, économique et pratiques.

2.3.1 Taille de l'image d'apprentissage

L'approche MP consiste à estimer la probabilité de $Z(u_0)$ conditionnée à l'état du voisinage $Z(\mathbf{u})$ par le nombre des répliques de cet état (Eq 2.2). D'après la loi forte des grands nombres, plus on retrouve ces répliques dans l'image d'apprentissage, plus notre estimation sera précise. Naturellement, la chance de trouver plusieurs répliques sur un voisinage de taille proche à celle de l'image d'apprentissage est très faible. Ainsi, nous supposons, que la taille de l'image d'apprentissage (nombre de points au total) est relativement grande par rapport à la taille de fenêtre d'exploration (nombre maximal de voisinage).

Étant donnée une image d'apprentissage de K faciès, nous voulons dériver les probabilités d'apparition d'une combinaison de n points voisins. Si l'image est complètement désordonnée (entropie forte), nous pouvons avoir K^n combinaisons différentes au maximum. Puisque nous espérons trouver plusieurs répliques de chaque combinaison dans l'image d'apprentissage, la taille de l'image doit être au moins supérieure au nombre de combinaisons possibles.

Prenons un simple exemple de structure 2D de deux faciès que nous voudrions explorer avec une fenêtre de $10 \times 10 = 100$ points. Au pire nous aurons $2^{100} \approx 10^{30}$ combinaisons différentes. Dans ce cas, il nous faut une image de la taille d'au moins de $10^{15} \times 10^{15}$ points (pour avoir une réplique pour chaque combinaison possible), ce qui n'est jamais le cas en pratique.

Dans les faits, nous disposons de moins de 10^{30} combinaisons différentes, car l'image d'apprentissage est structurée (entropie faible). Le travail de Arpat et al. (2002) montre un exemple de reconstruction d'un milieu fluvatile avec seulement 100 patterns (combinaisons de 40 points chacune) les plus fréquents. De plus, il faut tenir compte que la simulation MP n'assure pas l'utilisation de toutes les statistiques MP dérivées de l'image d'apprentissage. En effet, une des réalisations MP ne présente pas tous les patterns se trouvant dans l'image d'apprentissage.

2.3.2 Fenêtre d'exploration

Le choix d'une fenêtre d'exploration adaptée est essentiel pour la reproduction correcte de la structure géométrique. La taille de fenêtre d'exploration doit être suffisamment grande par rapport à la taille caractéristique de la texture de l'image d'apprentissage. De plus, pour saisir puis reproduire l'étendue et l'agencement des faciès, la fenêtre d'exploration doit pouvoir couvrir à la fois un faciès qui représente l'objet géologique et un autre ou plusieurs autres pour représenter son environnement (cf. Fig. 16 dans le cas de deux faciès).

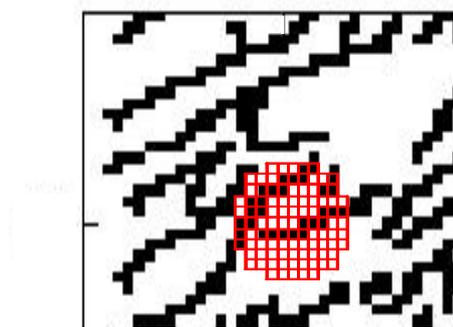


Fig. 16 Exemple d'une fenêtre d'exploration qui couvre le faciès noir entouré par de faciès blanc.

Dans le cadre de cette étude, les simulations MP ont été effectuées avec différentes tailles de fenêtre d'exploration. Il a été observé qu'au-delà d'une certaine taille du voisinage assez grande, le résultat de simulation n'est plus sensible à l'augmentation ultérieure de voisinage.

Il est possible d'adapter la forme d'une fenêtre d'exploration en fonction de l'orientation majeure des patterns. Pour étudier la sensibilité de la simulation MP vis-à-vis du choix de la forme de la fenêtre, Barrera et al. (2004) a effectué un test avec des fenêtres d'orientation différente. Ce test montre que les statistiques multipoints acquises par une fenêtre qui a la même orientation que les patterns sont plus robustes. Dans ce travail nous trouvons des indications sur le choix optimal de la forme d'une fenêtre. Le choix d'une fenêtre peut jouer un rôle important dans la reproduction de la structure si la taille de la fenêtre est optimale. Dès que la fenêtre dépasse une taille critique (dans toutes les dimensions de simulation), la forme n'est plus importante. Cela vient de ce que les statistiques MP acquises avec un voisinage assez grand sont capables de saisir toutes les caractéristiques géométriques de l'image, telles que la forme ou l'orientation.

Dans la suite de notre étude, nous supposerons la fenêtre d'exploration de taille suffisamment grande. Sauf précision contraire, la fenêtre utilisée dans les simulations en

2D sera de taille 100 points, ordonnée comme suit (Fig. 17) :

			100	92	77	85	93			
	84	76	68	60	45	53	61	69	81	
	75	52	44	36	21	29	37	49	70	
99	67	43	28	20	9	13	25	38	62	94
91	59	35	19	8	1	5	14	30	54	86
80	48	24	12	4	0	2	10	22	46	78
90	58	34	18	7	3	6	15	31	55	87
98	66	42	27	17	11	16	26	39	63	95
	74	51	41	33	23	32	40	50	71	
	83	73	65	57	47	56	64	72	82	
			97	89	79	88	96			

Fig. 17 Fenêtre d'exploration pour la simulation en 2D.

2.3.3 Chemin de simulation

Il y a plusieurs possibilités pour effectuer une simulation séquentielle.

Une des possibilités est de définir un **chemin aléatoire** qui parcourt tout l'espace de simulation indépendamment de la géométrie à simuler. Comme le nombre de points simulés augmente au fur et à mesure que la simulation se construit, les voisinages deviennent de plus en plus grands, si bien que les lois conditionnelles deviennent de plus en plus déterministes (Fig. 18). Ainsi, la probabilité conditionnelle à estimer dépend de toutes les estimations précédentes [Daly (2004)], et donc du chemin de simulation.

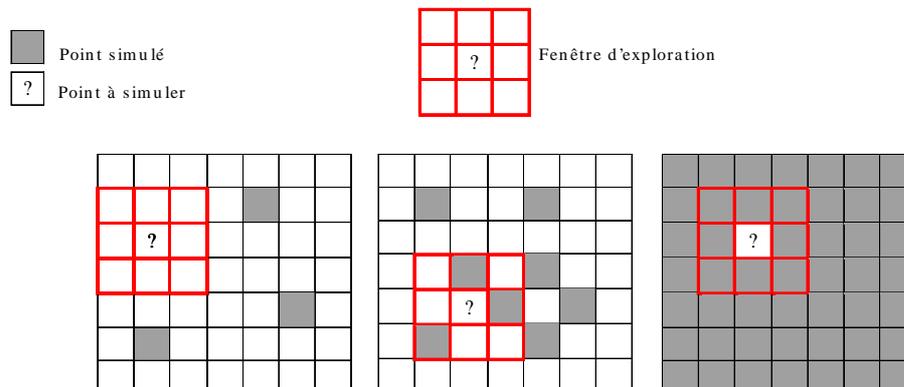


Fig. 18 Simulation par le chemin aléatoire (2D).

Soit $u_1, u_2, \dots, u_{n-1}, u_n$ un chemin aléatoire de simulation. Le voisinage évolue à chaque nouveau point simulé. Ainsi, la probabilité conditionnelle du point u_n dépend des

probabilités conditionnelles des points u_1, u_2, \dots, u_{n-1} de manière itérative.

Cet effet n'est pas négligeable dans le cas de la simulation d'objets continus (e.g. chenaux). Souvent, le manque de points conditionnants en début de simulation crée des configurations de points qui n'existent pas dans l'image d'apprentissage (le point blanc préalablement simulé sur le parcours d'un chenal), ce qui empêche la continuité des objets. Ici nous illustrons la non continuité causée par un cheminement aléatoire dans le cas de simulation MP de lignes parallèles (Fig. 19).

Pour une meilleure reproduction des objets de la taille supérieure à la taille de fenêtre d'exploration il est conseillé [Strebelle 2002] de simuler d'abord sur une grille grossière puis d'affiner la grille (technique **multigrille**). Ceci permet de reproduire la structure spatiale à une distance supérieure de celle de la fenêtre d'exploration (effet de nid, fractales). Nous proposons ici deux exemples de simulation qui montrent l'amélioration obtenue par des techniques multigrille d'ordre 3 et 5 (Fig. 20). Nous pouvons remarquer que la continuité des lignes restaurées est « proportionnelle » à l'ordre de multigrille.

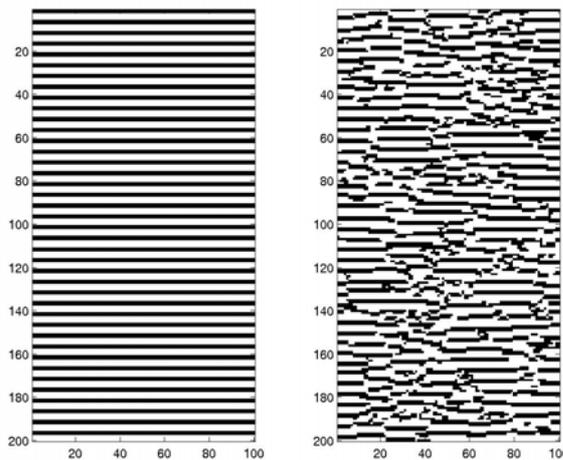


Fig. 19 Simulation MP par le chemin aléatoire 2D. Reproduction des lignes parallèles : l'image d'apprentissage à gauche, la réalisation MP à droite.

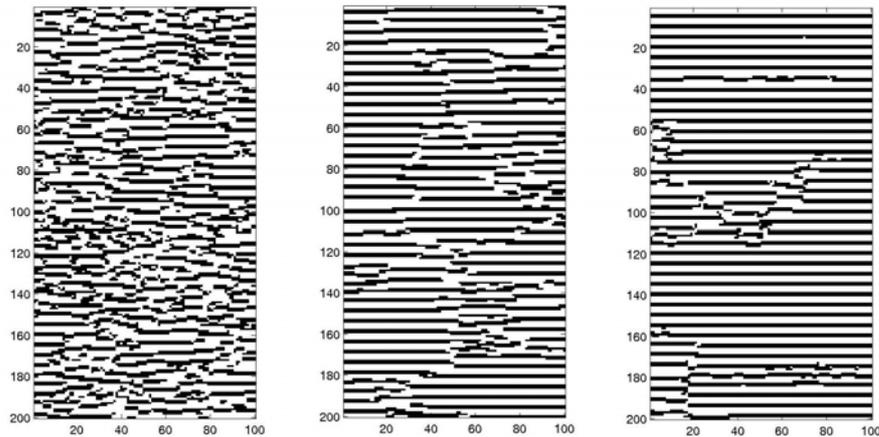


Fig. 20 Simulation MP par le chemin aléatoire 2D : grille fine à gauche ; multigrille d'ordre 3 au centre ; multigrille d'ordre 5 à droite.

Une autre façon de simuler séquentiellement est de définir un **chemin unilatéral** [Daly et Knudby (2007)]. Un exemple est un cheminement régulier selon l'ordre des points de la grille (Fig. 21). Dans ce cas, chaque point à simuler a son voisinage de points déjà simulé toujours de la même forme (sauf au bord). Il est alors simulé selon la même loi conditionnelle, au contraire des points d'un chemin aléatoire. Étant donnée une image d'apprentissage et une grille de simulation illimitées, nous nous trouvons dans le cadre du Modèle Markovien Unilatéral sans configurations inexistantes. En pratique toutefois, ce mode de simulation n'exclut pas des artefacts. Il s'avère néanmoins efficace pour simuler des structures continues présentant une orientation préférentielle et un support composé par des points déjà simulés de taille non négligeable (Fig. 22). Dès que la continuité des objets ne suit pas cette orientation ou que leur sinuosité est importante (chenaux méandriques), le chemin unilatéral perd en efficacité.

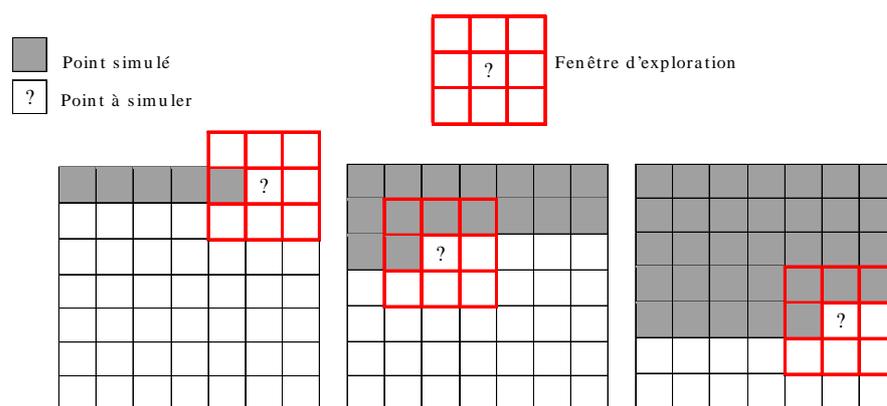


Fig. 21 Simulation par le chemin régulier (2D).

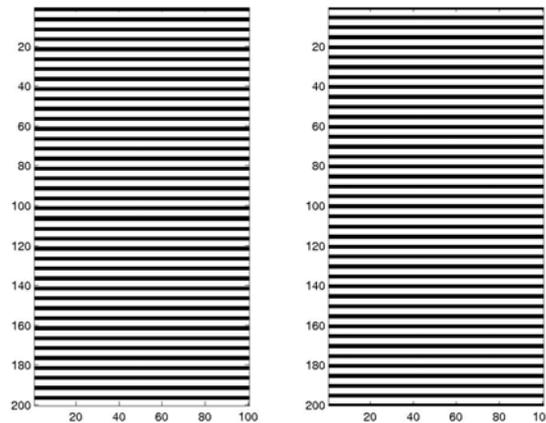


Fig. 22 Simulation MP par le chemin régulier 2D. Reproduction des lignes parallèles : l'image d'apprentissage à droite, la réalisation simulée à gauche.

Pour raffiner la reproduction de géométries complexes et, par conséquent, la reproduction des objets continus, Remy (2001), Liu et Journel (2004), Strebelle et Remy (2004) proposent d'utiliser le **chemin « post-processus »**. Autrement dit, l'algorithme MP «post-processus» consiste à revenir sur les points simulés avec peu de points voisins et à les resimuler avec plus de points voisins connus. Cette méthode peut être appliquée de manière itérative ou bien par une étape complémentaire en fin de simulation initiale. Nous proposons deux exemples de simulation post-processus d'abord avec 20% des points et puis avec 50% des points à resimuler (Fig. 23). Ces exemples montrent que cette méthode ne garantit pas la reproduction des objets très continus.

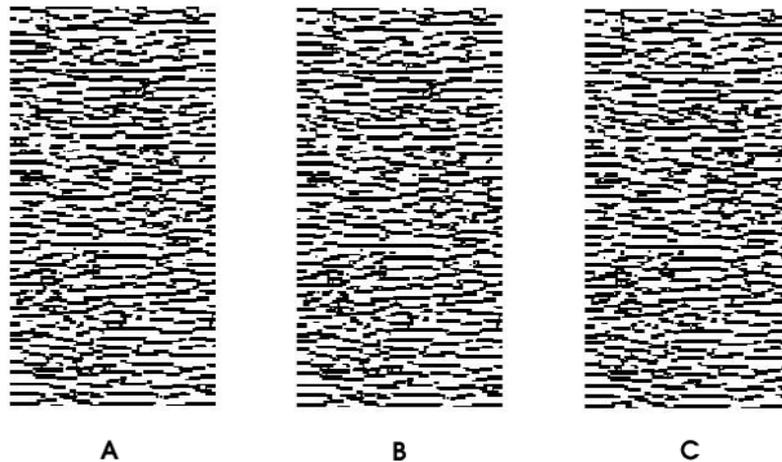


Fig. 23 Simulation MP par le chemin aléatoire post-processus 2D. (A) Reproduction des lignes parallèles : sans resimulation à gauche ; (B) 20% des points resimulés au centre ; (C) 50% des points resimulés à droite. (La simulation est obtenue en utilisant le module SGeMS développé à Stanford).

Par la suite, nous utiliserons un cheminement aléatoire pour obtenir la plupart des simulations MP sur la grille fine. Nous ne recourons à un cheminement régulier que dans

le cas où la reproduction des objets continus est essentielle.

2.3.4 Connexité

La connexité est une caractéristique, qui évalue la continuité (entre autre) des objets géologiques. C'est une caractéristique importante, car elle affecte l'écoulement de fluides dans un réservoir. Lors de la reproduction de la structure interne d'un réservoir, il est important donc de préserver sa connexité. Dans ce paragraphe, nous étudions la qualité de reproduction de la connexité par l'approche MP.

La littérature fournit diverses définitions de la connexité [Krishnan et Journel (2003)]. Dans le cadre de cette étude, nous nous appuyons sur la version classique de la connexité entre deux points [Allard (1994)] :

Déf. Deux points u et u' sont proches voisins si leur zone d'influence possède une frontière commune (1 dimension : la frontière est ponctuelle ; 2 dimensions - elle est linéaire ; 3 dimensions - elle est surfacique).

Déf. Deux point u et u' sont connectés $u \leftrightarrow u'$, s'ils sont reliés par une suite de proches voisins, tous appartenant au même faciès : u, u_1, \dots, u_n, u' .

Soient G l'espace discrétisé et A un ensemble de points dans G représentant un faciès : $Z(A) = 1$; $Z(A^c) = 0$. La fonction de connexité α entre deux points u, u' est la probabilité d'avoir ces deux points connectés :

$$\alpha(u, u') = P[u \leftrightarrow u' \mid u, u' \in A]$$

Cette fonction dépend non seulement de la distance entre les deux points u et u' , mais aussi de leur position relative. Dans un cadre stationnaire, la fonction de connexité $\alpha(u, u')$ ne dépend que du vecteur $u \leftrightarrow u'$, qui sépare les deux points.

A titre d'illustration, nous avons choisi un cas d'étude basé sur le modèle gaussien seuillé de variogramme hyperbolique. Les exemples sont bidimensionnels et de la taille 100x200. Pour observer les variations de continuité sur nos exemples, nous utilisons plusieurs portées horizontales de variogramme, tandis que la portée verticale reste unitaire. Pour estimer la qualité de reproduction de connexité et de continuité des *clusters* [définition en Annexe B] nous allons comparer les fonctions de covariance et de connexité horizontales dérivées de l'image d'apprentissage et de 10 de ses simulations MP.

La première image d'apprentissage est construite avec un variogramme horizontal de

portée unité (Fig. 24). Les fonctions de covariance régionale (Fig. 25) paraissent comparables, ce qui n'est pas vraiment le cas des fonctions de connexité (Fig. 26). Malgré le seuil 0 du modèle gaussien, qui fournit la même proportion entre les deux faciès, la présence de fluctuations statistiques fait que leurs connexités sur l'image d'apprentissage sont assez différentes. Puisque le modèle de simulation MP est fourni par l'image d'apprentissage, l'approche MP reproduit les caractéristiques de l'image d'apprentissage mais pas du modèle gaussien. Cela peut expliquer, pourquoi la connexité des clusters noirs est mieux reproduite.

La deuxième image d'apprentissage est construite avec un variogramme horizontal de portée 5 (Fig. 27). Même si l'accord entre les covariances reste satisfaisant (Fig. 28), la connexité des réalisations MP est inférieure à celle de l'image d'apprentissage (Fig. 29).

La troisième et dernière image d'apprentissage est construite avec un variogramme horizontal de portée 20 (Fig. 30). Les covariances montrent, que la portée de la réalisation n'atteint pas 20 (Fig. 31). Cet effet est causé par l'étendue des clusters qui est supérieure non seulement à la taille de fenêtre d'exploration mais aussi à la taille de l'image d'apprentissage. Plus l'étendue (continuité) des clusters augmente, plus les connexités de l'image d'apprentissage et celle des réalisations MP s'avèrent différentes (Fig. 32).

L'exemple précédent nous a montré que les multiples réalisations MP peuvent présenter des connexités assez variées. Ainsi, la connexité peut jouer un rôle important dans le choix d'une réalisation par rapport aux autres. Dans l'exemple suivant, nous allons montrer, que la fluctuation de connexité peut indiquer la qualité de reproduction de la géométrie et donc, doit être prise en compte comme caractéristique qualificative.

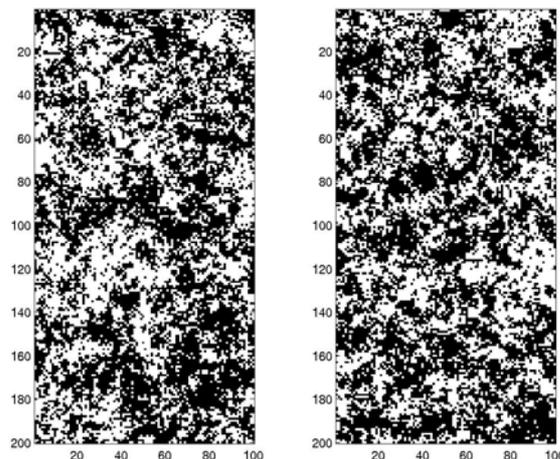


Fig. 24 Simulation stationnaire par l'approche MP. Reproduction du modèle gaussien hyperbolique seuillé (portée horizontale unité) : l'image d'apprentissage à gauche, l'une des 10 réalisations MP à droite.

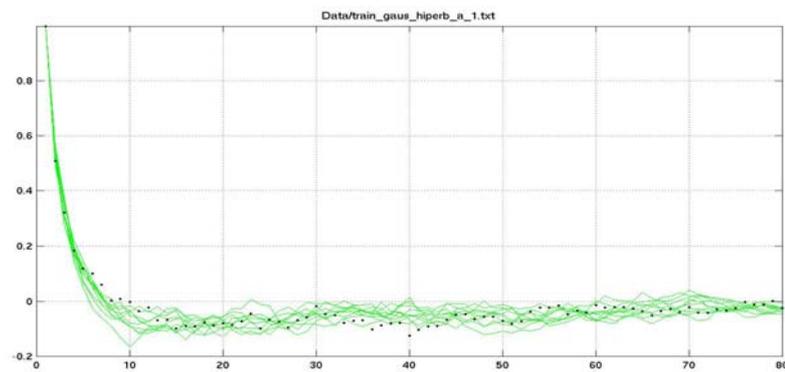


Fig. 25 Fonction de covariance horizontale du modèle gaussien hyperbolique seuillé (portée horizontale unité) : la ligne pointillée correspond à l'image d'apprentissage et les lignes pleines aux 10 réalisations MP.

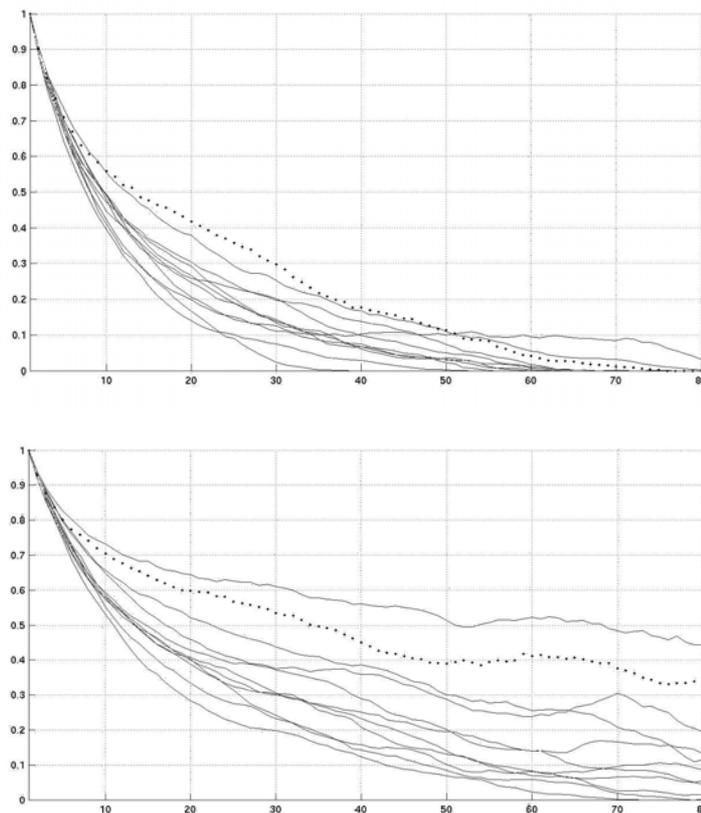


Fig. 26 Fonction de connexité horizontale du modèle gaussien hyperbolique seuillé (portée horizontale unité) : la ligne pointillée correspond à l'image d'apprentissage et les lignes pleines aux 10 réalisations MP ; en haut pour les clusters noirs, en bas pour les clusters blancs.

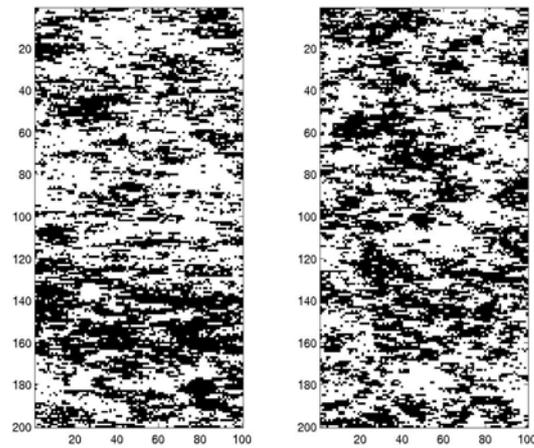


Fig. 27 Simulation stationnaire par l'approche MP. Reproduction du modèle gaussien hyperbolique seuillé (portée horizontale égale à 5) : l'image d'apprentissage à gauche, une des 10 réalisations MP à droite.

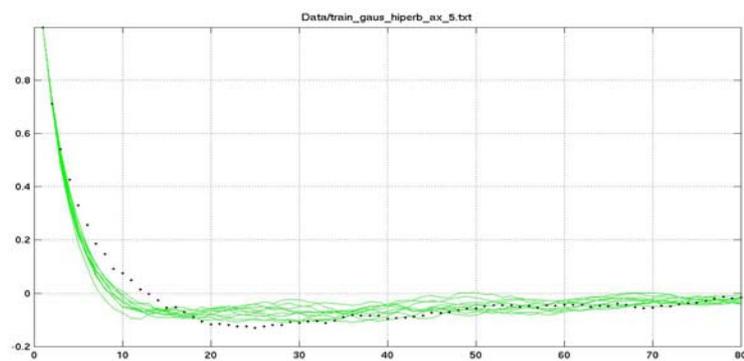


Fig. 28 Fonction de covariance horizontale du modèle gaussien hyperbolique seuillé (portée horizontale égale à 5) : la ligne pointillée correspond à l'image d'apprentissage et les lignes pleines aux 10 réalisations MP.

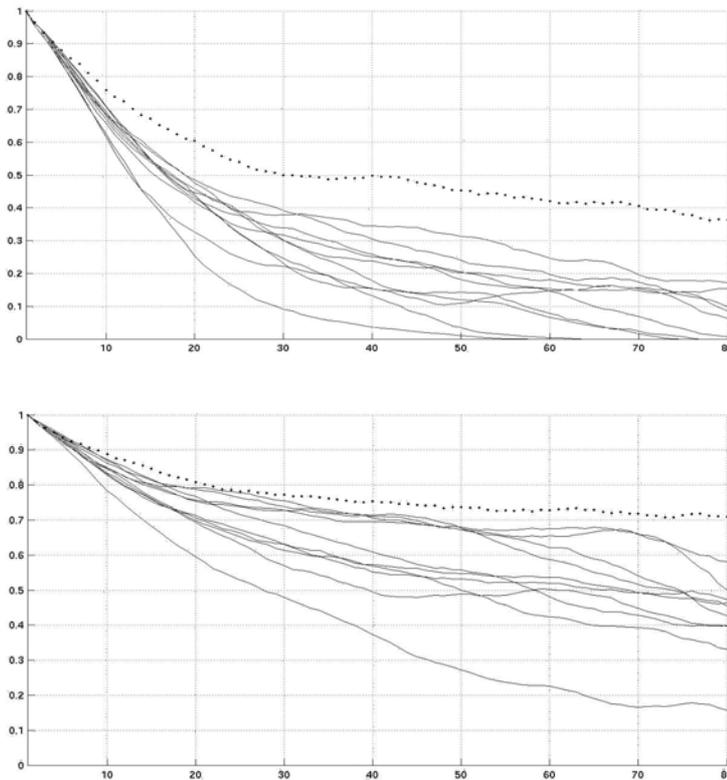


Fig. 29 Fonction de connexité horizontale du modèle gaussien hyperbolique seuillé (portée horizontale égale à 5) : la ligne pointillée correspond à l'image d'apprentissage et les lignes pleines aux 10 réalisations MP ; en haut pour les clusters noirs, en bas pour les clusters blancs.

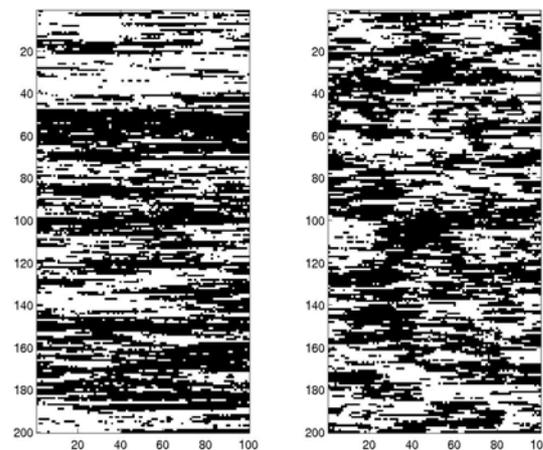


Fig. 30 Simulation stationnaire par l'approche MP. Reproduction du modèle gaussien hyperbolique seuillé (portée horizontale égale à 20) : l'image d'apprentissage à gauche, une des 10 réalisations MP à droite.

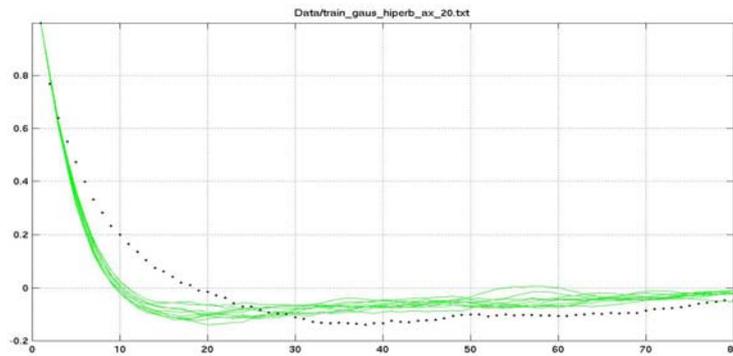


Fig. 31 Fonction de covariance du modèle gaussien hyperbolique seuillé (portée horizontale égale à 20) : la ligne pointillée correspond à l'image d'apprentissage et les lignes pleines aux 10 réalisations MP.

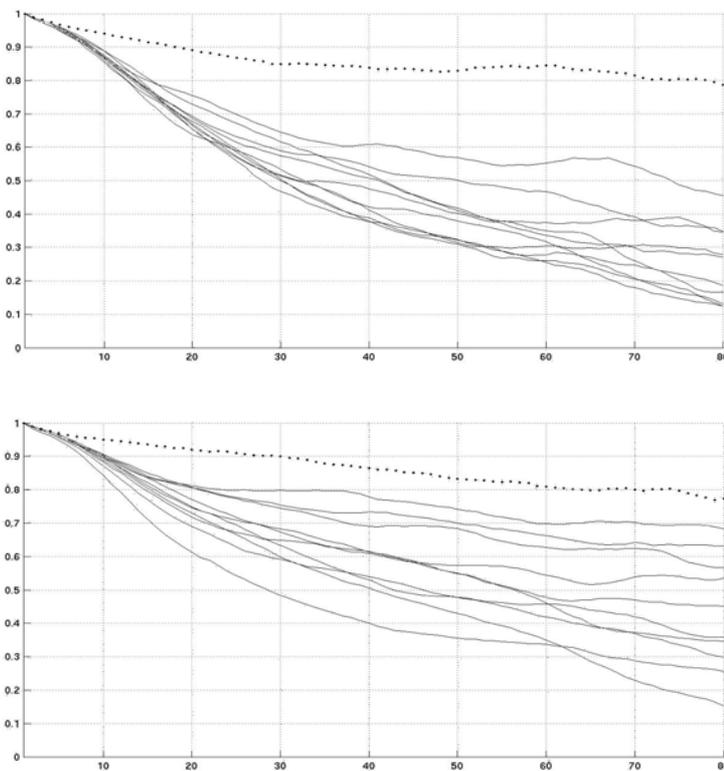


Fig. 32 Fonction de covariance du modèle gaussien hyperbolique seuillé (portée horizontale égale à 20) : la ligne pointillée correspond à l'image d'apprentissage et les lignes pleines aux 10 réalisations MP; en haut pour les clusters noirs, en bas pour les clusters blancs.

Le cas d'étude suivant est basé sur un réseau de fractures d'un gisement situé en Afrique du Nord et qui a déjà fait l'objet de plusieurs études [Gauthier et al. (2002)]. Nous disposons d'une image d'apprentissage en 2D de taille 100x200 représentant le réseau de fractures à grande échelle (Fig. 33). Dix réalisations sont simulées à partir de l'image d'apprentissage par approche MP (Fig. 34 - 35). Dix fonctions de connexité sont extraites de

ces réalisations et comparées avec celle de l'image d'apprentissage (Fig. 36).

Toutes les réalisations reproduisent assez bien la connexité verticale des fractures, pourtant elles ne parviennent pas à reproduire correctement celle de leur complémentaires (clusters blancs). Nous pouvons également observer que les fractures présentent une connexité très peu de fluctuante, au contraire de leur complémentaire. Cela résulte vraisemblablement de ce que les fractures sont mieux reproduites que leur complémentaire. En revenant au paragraphe 2.3.2, ce défaut s'explique par l'allongement des clusters blancs qui est comparable à celle de l'image d'apprentissage ; ils sont donc supérieurs à la taille de fenêtre d'exploration.

A partir des résultats de cette étude, nous tirons les conclusions suivantes :

- 1) La connexité peut être reproduite par l'approche MP si les paramètres de simulation sont adaptés à la géométrie.
- 2) La connexité peut être à la fois une caractéristique quantitative (continuité des faciès) et qualitative de la simulation MP :
 - faible fluctuation et l'accord avec la connexité de l'image d'apprentissage nous assure la bonne reproduction de la géométrie y compris la continuité des faciès ;
 - fluctuation importante et désaccord avec la connexité de référence nous signale que la reproduction de la géométrie est mauvaise et que les paramètres de simulation ne sont pas adaptés à la géométrie.

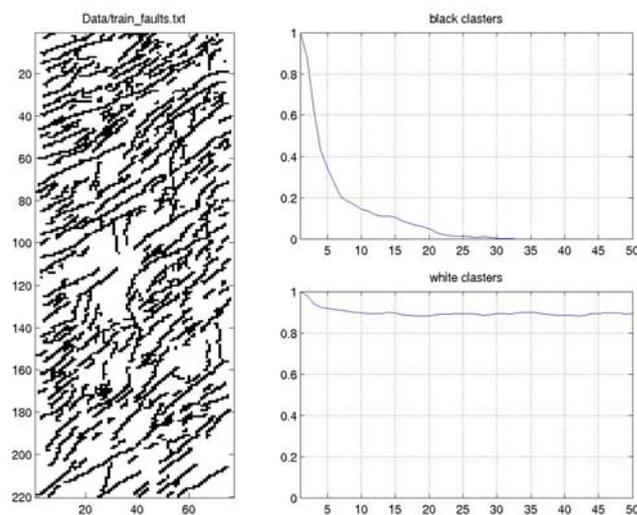


Fig. 33 L'image d'apprentissage du réseau de fractures. Fonction de connexité verticale : en haut pour les clusters noirs, en bas pour les clusters blancs.

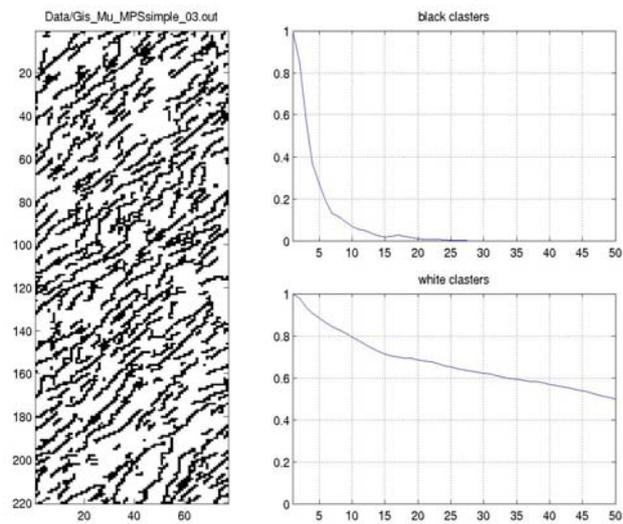


Fig. 34 Une des 10 réalisations MP. Fonction de connexité verticale : en haut pour les clusters noirs, en bas pour les clusters blancs.

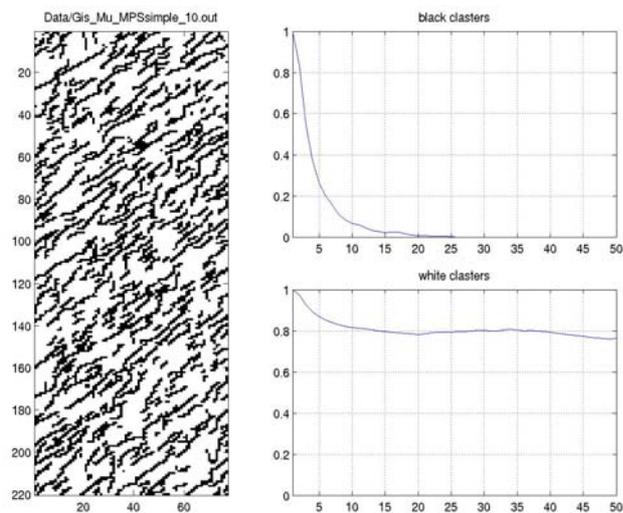


Fig. 35 Une des 10 réalisations MP. Fonction de connexité verticale : en haut pour les clusters noirs, en bas pour les clusters blancs.

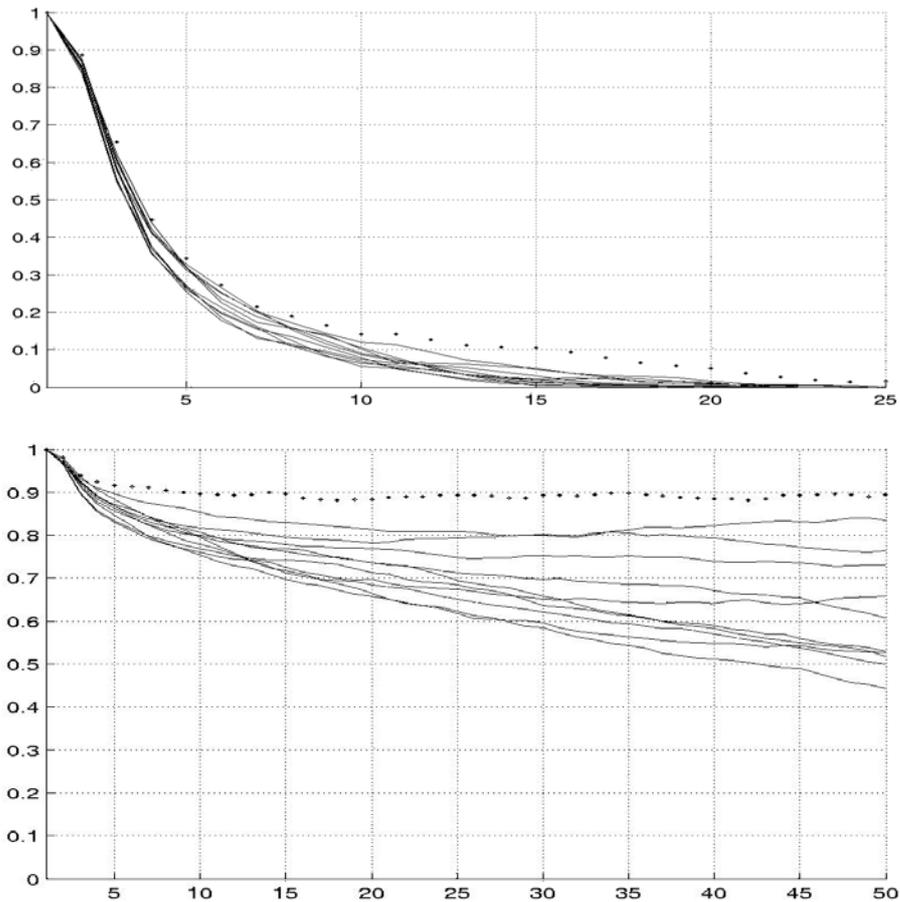


Fig. 36 Fonction de connectivité verticale du réseau de fractures et ses 10 réalisations MP : la ligne pointillée correspond à l'image d'apprentissage et les lignes pleines aux 10 réalisations MP.

2.3.5 Stationnarité

L'utilisation classique de l'approche MP nécessite la stationnarité spatiale pour pouvoir construire une loi empirique à partir des statistiques acquises de l'image d'apprentissage quelque soit leur localisation spatiale. Il existe plusieurs formes de stationnarité :

1) **stationnarité forte :**

Rappelons que la fonction aléatoire $Z(\cdot)$ est stationnaire si pour n'importe quel nombre de points u_1, u_2, \dots, u_n et n'importe quel h la distribution conjointe des fonctions aléatoires $Z(u_i)$, $i = 1, \dots, n$ est égale à la distribution conjointe des

fonctions aléatoires $Z(u_i + h)$, $i = 1, \dots, n$:

$$F\{Z(u_1), Z(u_2), \dots, Z(u_n)\} = F\{Z(u_1 + h), Z(u_2 + h), \dots, Z(u_n + h)\}.$$

Ce type de stationnarité est difficilement vérifiable et rarement utilisé dans la pratique géostatistique.

- 2) **stationnarité faible** ou stationnarité du second ordre. Toutes les variables ont la même moyenne. De plus, la covariance entre deux points d'appui ne dépend pas spécifiquement d'eux, mais seulement du vecteur qui les sépare. Cette forme de stationnarité beaucoup plus lâche n'est toutefois pas toujours satisfaite.
- 3) **hypothèse intrinsèque**. Elle ne nécessite que la stationnarité de la différence de premier ordre $Z(u_i + h) - Z(u_i)$, $i = 1, \dots, n$. Cette forme de stationnarité est très souvent utilisée en géostatistique et en analyse des séries temporelles.
- 4) **stationnarité faible avec dérive**. L'idée de cette forme de stationnarité est qu'on peut se débarrasser de la dérivée pour se retrouver dans un contexte de stationnarité faible.

L'approche MP a été développée sous l'hypothèse de stationnarité forte de la fonction aléatoire Z .

Précisons que la stationnarité est une caractéristique d'une fonction aléatoire $Z(\cdot)$ et non pas des données spatiales que l'on cherche à modéliser. La conséquence est que la forme de stationnarité de Z ne peut pas être strictement inférée à partir d'une ou de plusieurs réalisations de cette fonction aléatoire. Elle peut par contre être présumée. Il est possible de tester après coup la légitimité de notre hypothèse en étudiant la compatibilité des réalisations de $Z(\cdot)$ à nos données. Il est à noter que ce test ne permet que de réfuter l'hypothèse faite. La non réfutation de l'hypothèse ne garantit pas la forme de stationnarité retenue, mais laisse cette éventualité possible.

En pratique, le cas stationnaire correspond au cas où l'image d'apprentissage est suffisamment homogène pour que les statistiques que l'on en extrait, puissent être considérées comme représentatives de toute l'image. La vérification pratique de l'hypothèse de stationnarité à partir d'une image souligne l'importance d'un facteur d'échelle qui est totalement absent du formalisme probabiliste [Chauvet (1999)]. Un même objet géométrique peut être considéré comme constant ou variable, structuré ou non, selon l'échelle à laquelle il sera examiné.

Prenons l'exemple du réseau de fractures (Fig. 37). Supposons, que notre image d'apprentissage est une réalisation d'un processus stationnaire. Cette hypothèse est plausible si, par exemple, on travaille à l'échelle d'un groupe de fractures. Une simulation

MP a été réalisée sous hypothèse de stationnarité (Fig. 37). Il est intéressant de voir que la répartition des fractures de la réalisation est plus homogène que celle de l'image d'apprentissage. Si l'on travaille à l'échelle d'une fracture, la variabilité de taille, d'orientation et de présence de faciès est enregistrée.

Cet exemple illustre parfaitement le fait que la stationnarité est une notion dépendante de l'échelle de travail. À la question de savoir si la réalisation donnée peut être considérée comme la réalisation d'une fonction stationnaire, la réponse dépend finalement de l'objectif de modélisation (ce qui est d'ailleurs le cas pour toute modélisation géostatistique) et des caractéristiques que l'on veut préserver par cette modélisation.

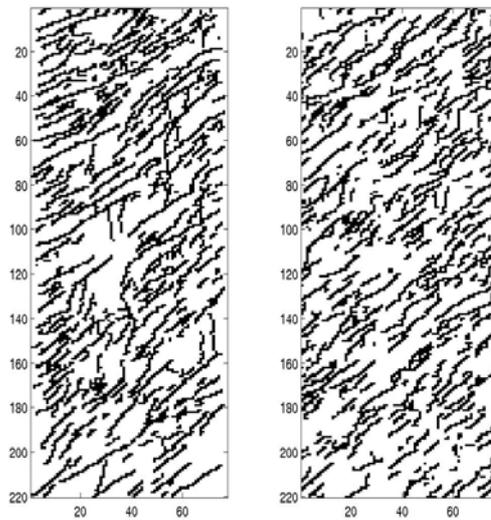


Fig. 37 Simulation stationnaire par l'approche MP. Reproduction d'un réseau de fractures : l'image d'apprentissage (à gauche) et la simulation MP (à droite) avec la proportion de faciès respectivement égale à 0.28 et 0.29.

2.3.6 Résumé

Pour conclure la partie des conditions d'application de l'approche MP, nous rappelons les 4 conditions nécessaires pour une bonne reproduction d'une structure géométrique :

- 1) L'image d'apprentissage doit contenir les patterns géométriques représentatifs de la structure et répétitifs dans l'espace (de faible entropie) ;
- 2) La taille de l'image d'apprentissage doit être suffisamment grande par rapport à la taille de fenêtre d'exploration ;
- 3) La taille de la fenêtre d'exploration doit être suffisamment grande par rapport à la taille des patterns répétitifs de l'image d'apprentissage ;

- 4) La fonction aléatoire $Z(\cdot)$ est présumée stationnaire. De plus, la variabilité spatiale de l'hétérogénéité de l'image d'apprentissage doit se stabiliser au-delà d'une certaine distance bien inférieure à la taille de l'image.

Si ces 4 conditions d'applications sont respectées, l'application de l'approche MP garantit une reproduction correcte de la géométrie en préservant ses caractéristiques (formes, orientation, continuité ...). Dans ce cas, le chemin de simulation peut être aléatoire, ce qui permet une meilleure reproduction de la diversité des patterns présentés dans l'image d'apprentissage. Dès que l'une de ces conditions n'est pas respectée, l'application de l'approche MP ne garantit pas une reproduction correcte de la géométrie. Dans ce cas, un ajustement des paramètres peut améliorer les résultats de simulation MP :

- 1) La configuration et la taille de la fenêtre d'exploration sont choisies en fonction de la géométrie de patterns répétitifs ;
- 2) Le cheminement régulier de simulation est défini de manière à pouvoir suivre l'orientation préférentielle des structures continues ;
- 3) Une simulation multigrille est appliquée pour reproduire les objets géométriques de taille supérieure à celle de fenêtre d'exploration.

Après avoir obtenu une réalisation MP, une vérification de sa qualité de reproduction est conseillée. Les caractéristiques d'image (moyenne, écart-type ...) peuvent faire apparaître les discordances majeures entre l'image d'apprentissage et sa réalisation MP. Les moyennes locales et les statistiques multipoints permettent de comparer plus finement leur géométrie. De son côté, la fonction de connexité aide à vérifier toute la conformité de la continuité et aussi de la géométrie dans les deux images.

2.4 Vers la simulation non stationnaire

Nous venons de voir le fonctionnement de l'approche multipoints moyennant l'introduction implicite d'une hypothèse de stationnarité (§2.3.5). En pratique toutefois, la réalité est souvent non stationnaire. Les structures géométriques formées par l'hétérogénéité du réservoir possèdent des caractéristiques (dites morphologiques) variables dans l'espace. Voici plusieurs caractéristiques d'un réservoir pétrolier susceptibles d'évoluer de façon non stationnaire :

- proportion locale (densité) de faciès ;
- taille des objets géologiques ;
- orientation régionale des objets ;
- connexité régionale ;
- etc.

Si la variabilité spatiale de l'hétérogénéité de l'image d'apprentissage ne se stabilise au-delà d'une certaine distance bien inférieure à la taille de cette image, nous la considérons comme non stationnaire. Dans ce cas, il est intéressant de préserver et de reproduire la non stationnarité spatiale par simulation MP.

Cette non stationnarité est souvent accessible via une information, issue, par exemple, de la sismique [Gauthier et al. (2002)]. La Fig. 38 montre un exemple dans un cadre non stationnaire spatiale sur le réseau de fractures qui l'on a déjà utilisé dans un cadre stationnaire (§§2.3.4-2.3.5). L'image d'apprentissage présente des zones de faible et de forte fracturation, ce qui est confirmé par la carte sismique.

Une autre source de données non stationnaires est l'image d'apprentissage elle-même. Dans ce cas, les caractéristiques géométriques de forte variabilité spatiale peuvent être extraites de l'image par des méthodes morphologiques tel que la moyenne mobile ou l'analyse de clusters ...).

Dans le chapitre suivant, nous allons montrer comment prendre en compte ces données dans une simulation multipoints.

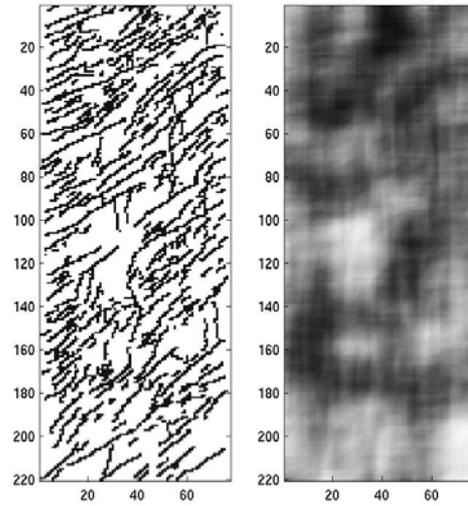


Fig. 38 Image d'apprentissage à gauche ; sa carte de densité de fracturation à droite.