HAL
open science

Morphologie mathématique, géostatistique et stéréologie : contributions à l'élaboration de modèles

Christian Lantuéjoul

To cite this version:

Christian Lantuéjoul. Morphologie mathématique, géostatistique et stéréologie : contributions à l'élaboration de modèles. Mathématiques [math]. Université Paris Sud - Paris XI, 2008. tel-00430389

HAL Id: tel-00430389

https://theses.hal.science/tel-00430389v1

Submitted on 6 Nov 2009

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

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

UNIVERSITÉ PARIS-SUD
FACULTÉ DES SCIENCES D'ORSAY

MÉMOIRE

Présenté pour obtenir

LE DIPLÔME D'HABILITATION À DIRIGER
DES RECHERCHES EN SCIENCE
DE L'UNIVERSITÉ PARIS XI

Spécialité : Mathématiques

par

Christian LANTUÉJOUL

MORPHOLOGIE MATHÉMATIQUE, GÉOSTATISTIQUE ET STÉRÉOLOGIE: CONTRIBUTIONS À L'ÉLABORATION DE MODÈLES

Soutenu le 12 février 2008 devant la commission d'examen :

M.	Jean-Paul CHILÈS	(Examinateur)
M.	Xavier GUYON	(Examinateur)
M.	Pascal MASSART	(Examinateur)
M.	Georges OPPENHEIM	(Rapporteur)
M.	Martin SCHLATHER	(Rapporteur)
M.	Michel SCHMITT	(Examinateur)

Remerciements

Après plus de trente ans d'activité à l'Ecole des Mines dans les domaines de la morphologie mathématique et de la géostatistique, il m'est agréable de remercier tout ceux qui de près et de loin ont contribué à l'aboutissement de cette thèse.

En premier lieu vient Georges Matheron, homme de culture et d'une exceptionnelle fécondité scientifique, qui m'a fait découvrir le monde de la recherche. Ce fut un privilège que de travailler dans son entourage. Je garde aussi un merveilleux souvenir de Philippe Formery qui m'a formé au calcul des probabilités et à l'enseignement par des classes de travaux dirigés portant sur des exercices particulièrement imaginatifs et imagés. Francis Maisonneuve, actuel professeur de mathématiques à l'Ecole des Mines m'a apporté en plus de son amitié un soutien scientifique constant tout au long de ces années.

Je souhaite exprimer ma reconnaissance à mes collègues, ou anciens collègues (Hélène, Pierre, Jean-Paul, Nathalie, François, Christian, Gaëlle, Didier, Françoise, Jacques, Isabelle, Yves et Hans), ainsi qu'aux thésards (Marco, Béatrice, Edwige, Marie... Claire ?) de l'équipe géostatistique pour l'atmosphère de travail studieuse et conviviale qui y règne. Les discussions ne manquent pas de se poursuivre lors de la sortie café qui suit le déjeuner, voire le long des pistes de Franchard (sur ce terrain, les participants sont un peu moins nombreux !).

Des remerciements particuliers vont à Michel Schmitt, actuel Directeur de la recherche à l'Ecole des Mines qui suit de près mon travail et qui m'encourage depuis un certain nombre d'années à soutenir une thèse d'habilitation. Je suis heureux qu'il ait accepté d'être examinateur de ce travail.

Lors d'un des séminaires de statistiques spatiales régulièrement organisés par Liliane Bel à AgroParisTech, cette dernière m'a proposé une soutenance commune en compagnie d'un troisième complice (Philippe Naveau). Nous sommes tous trois intéressés par les statistiques de valeurs extrêmes en contexte spatial. Un heureux évènement n'a finalement pas permis à Philippe de s'associer à nous jusqu'au bout, mais pour lui cela n'est certainement que partie remise.

Je tiens à remercier Georges Oppenheim de l'Université de Marne-la-Vallée qui a accepté d'être rapporteur de ce mémoire en plus de ses nombreuses activités. Ses remarques et suggestions m'ont été précieuses. De leur coté, Adrian Baddeley, Professeur à l'Unversité de Western Australia ainsi que Martin Schlather, Professeur à l'Université de Göttingen m'ont fait l'amitié d'être rapporteurs extérieurs. Je les en remercie d'autant plus chaleureusement que tous deux ont eu à surmonter les difficultés du français.

J'exprime toute ma gratitude à Pascal Massart, Directeur du Laboratoire de Mathématiques à l'Université d'Orsay Paris-Sud qui me fait l'honneur d'être examinateur de ce mémoire. Il en va de même pour Xavier Guyon, Professeur Emérite à l'Université Panthéon-Sorbonne que j'ai eu l'occasion de rencontrer à plusieurs reprises et dont j'ai pu apprécier les nombreux apports au développement des statistiques spatiales.

Pour finir, je voudrais remercier mon chef d'équipe Jean-Paul Chilès qui a aussi accepté de d'être examinateur. A sa grande compétence scientifique sont associées des qualités humaines rares.

Morphologie Mathématique, Géostatistique et Stéréologie : Contributions à l'élaboration de modèles

Christian LANTUÉJOUL

Le présent document donne un aperçu des sujets sur lesquels j'ai travaillé pendant un peu plus de trente ans à l'Ecole des Mines, dans les domaines de la morphologie mathématique et de la géostatistique. Issus tous deux de la nécessité d'examiner des jeux de données à des échelles différentes, ces deux domaines ont ensuite évolué dans des directions assez différentes, l'un vers l'analyse d'image l'autre vers les statistiques spatiales.

Mon séjour en morphologie mathématique de 1972 à 1984 fut pour moi l'occasion de participer à l'essor d'une discipline nouvelle, dans laquelle tout était ouvert. J'espère que l'application du chapitre 1 rend compte de cet état d'esprit.

En 1985, mon passage en géostatistique m'a ouvert de nouveaux horizons. Initiée en Afrique du Sud par les travaux précurseurs de H. Sichel et D. Krige, cette discipline a été complètement reformalisée par G. Matheron. De vocation au départ minière, elle s'est progressivement enrichie au contact d'applications nouvelles (pétrole, environnement, génie forestier, météo, halieutique, épidémiologie, science des matériaux...), faisant à leur tour surgir des problématiques nouvelles. Les chapitres 3 et 4 témoignent de cette évolution avec la construction de modèles stochastiques spatiaux et la conception d'algorithmes de simulation conditionnelle.

A la charnière des deux domaines figure la stéréologie. Elle s'avère incontournable dans la mesure où les données morphologiques ou géostatistiques sont souvent disponibles dans un espace de dimension inférieure à l'espace de travail (lames minces ou sections polies pour des images microscopiques, forages dans un réservoir pétrolier...). Quelques travaux de stéréologie sont présentés au chapitre 2.

Finalement une liste de quelques problèmes sur lesquels il conviendrait de faire porter nos efforts fait l'objet du dernier chapitre.

Table des matières

1 Morphologie mathématique **2**

2 Stéréologie **5**
2.1 Simulation de points uniformes dans un domaine de grande dimension 5
2.2 Relation de Cartier 6

3 Modèles stochastiques spatiaux **9**
3.1 Tempêtes poissonniennes 9
3.2 Fonction aléatoire de substitution 12

4 Simulation **14**
4.1 Fonction aléatoire gaussienne 14
4.1.1 Simulation des fonctions de base 14
4.1.2 Combien de fonctions de base ? 18
4.2 Schéma booléen 19

5 Quelques thèmes de recherche **23**
5.1 Choix de modèles 24
5.2 Simulation conditionnelle 24
5.3 Simulation exacte 26
5.4 Modèles statistiques spatiaux de données compositionnelles 27
5.5 Valeurs extrêmes en contexte spatial 28
5.6 Problèmes de voisinage 29

1 Morphologie mathématique

Considérons l'image de la figure 1. Il s'agit d'une section polie de fonte à graphite vermiculaire. On sait que plus les particules de graphite sont allongées, plus elles tendent à fragiliser la fonte. Comment mesurer leur allongement de façon automatique ?

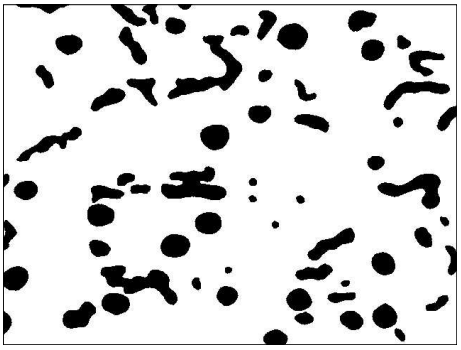

FIG. 1 – Section polie d'une fonte à graphite vermiculaire

Pour ce faire, il nous faut définir un indicateur d'allongement des particules.

Dans ce qui suit, chaque particule est modélisée par une partie du plan X, compacte et connexe par arcs. Si x et y sont deux points quelconques de X, on désigne par $d_X(x, y)$ la borne inférieure des longueurs des arcs reliant x et y tout en restant dans X. Pour éviter certaines pathologies, nous supposerons d_X finie[1]. Dans ce cas, d_X est une métrique sur X que l'on appelle *distance géodésique*. Nous la supposerons de plus comparable[2] à la métrique euclidienne induite sur X.

[1]Il peut exister des points x et y de X pour lesquels on a $d_X(x, y) = \infty$. Il suffit par exemple de considérer les deux extrémités de la spirale hyperbolique d'équation polaire $r(\theta) = \frac{1}{\theta}$ pour $\theta \in [1, +\infty]$.

[2]Il n'est pas difficile de construire des ensembles X qui sont compacts pour la métrique euclidienne et non compacts pour la métrique géodésique. Par exemple, un peigne avec un nombre dénombrable de dents.

A partir de la distance géodésique, il est utile de déterminer l'éloignement maximum attaché à chaque point de X. Pour ce faire, j'ai été amené à introduire la *fonction de propagation* :

$$T(x) = \max_{y \in X} d_X(x, y)$$

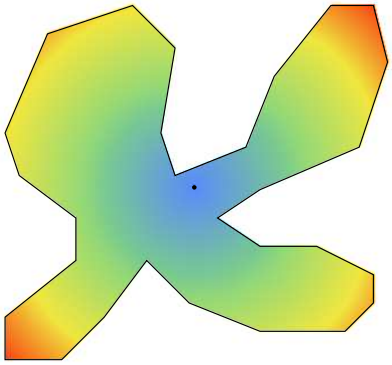

FIG. 2 – Une particule et sa fonction de propagation

Cette fonction a des propriétés tout à fait intéressantes. Tout d'abord, elle est continue et même lipschitzienne

$$|T(x) - T(y)| \leq d_X(x, y) \qquad x, y \in X$$

Comme X est compact, T atteint ses valeurs extrémales. Les points où T présente des maxima locaux sont appelés *extrémités* de X. Les valeurs en ces points n'étant pas forcément les mêmes, une hiérarchie entre extrémités est ainsi établie. On appelle *diamètre géodésique* la valeur $D(X) = \max_{x \in X} T(x)$.

Supposons maintenant X simplement connexe. Dans ce cas, on peut établir que deux points quelconques x et y de X sont reliés par un arc et un seul de longueur $d_X(x, y)$ (Lantuéjoul, 1982). Cet arc est appelé *géodésique* entre x et y, et noté G_{xy}. Il se trouve que la fonction de propagation se comporte comme une fonction convexe le long des géodésiques de X :

$$d_X(x, y)T(z) \leq d_X(z, y)T(x) + d_X(z, x)T(y) \qquad x, y \in X,\ z \in G_{xy}$$

Une conséquence est que T prend sa valeur minimale en un point unique c. Ce point est appelé *centre géodésique*, et la valeur $R(X) = T(c)$ *rayon géodésique*. Dans le cas où X est convexe, c et $R(X)$ ne sont rien d'autre que le centre et le rayon du cercle circonscrit à X.

Il a été montré par Maisonneuve que le diamètre géodésique et le rayon géodésique sont liés par les inégalités

$$\sqrt{3}R(X) \leq D(X) \leq\ 2R(X),$$

avec égalité à gauche ou à droite lorsque X est un triangle équilatéral ou bien un disque (Maisonneuve et Lantuéjoul, 1984). A deux dimensions, ces deux inégalités généralisent le résultat établi par Jung dans le cas où X est convexe (Berger, 1979).

A partir de ces deux facteurs, il est naturel d'envisager les deux indices d'allongement suivants :

$$\alpha(X) = \frac{\pi}{4}\frac{D^2(X)}{A(X)} \qquad \beta(X) = \pi\frac{R^2(X)}{A(X)}$$

Ces deux indices mettent en jeu l'aire $A(X)$ de X. Ils sont normés de façon à ne prendre que des valeurs ≥ 1, l'égalité ayant lieu dans le cas du disque.

Pour revenir à notre problème initial, l'indice β a été utilisé pour quantifier l'allongement des particules de graphite. La fonction de propagation a été calculée en parallèle sur toutes les particules par un algorithme fonctionnant en trame hexagonale (Lantuéjoul et Maisonneuve, 1984)[3]. La figure 3 montre les rayons géodésiques et les indices d'allongement obtenus. La taille des images est de 640×480. La valeur maximale prise par les rayons géodésiques est de 68. Quant aux indices d'allongement, ils varient entre 1.1 et 5.3.

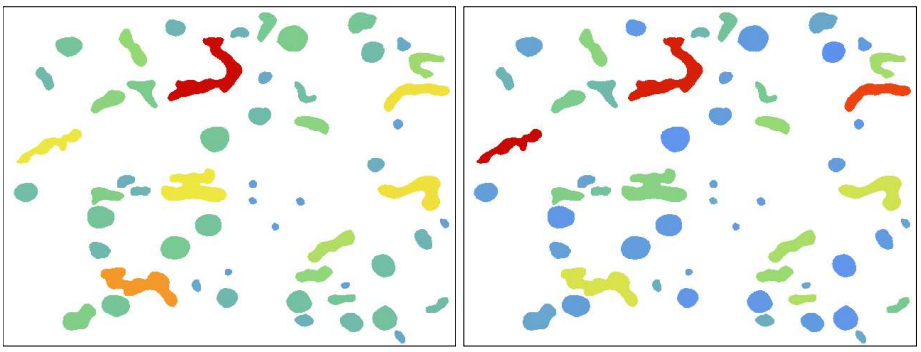

FIG. 3 – Représentation des rayons géodésiques (à gauche) et des indices d'allongement β (à droite) de particules de graphite à l'aide d'une palette de couleur arc-en-ciel. Certaines particules sont vues comme longues mais non allongées

Pour estimer un indice d'allongement moyen, on ne peut pas se contenter de prendre la moyenne des indices mesurés. En effet, les mesures ont été effectuées sur les particules visibles en totalité dans le champ de l'image, et il se trouve que ce sont les particules de plus grosse taille qui ont le plus de chance de recouper les bords du champ. C'est la raison pour laquelle chaque particule a son indice d'allongement pondéré par la chance qu'elle a de figurer à l'intérieur du champ. Cette chance est proportionnelle à la surface du champ érodé par la particule. On est ainsi amené à prendre un estimateur de la forme

$$\hat{\beta} = \sum_{i \in I} \frac{\beta(X_i)}{|D \ominus \check{X}_i|} \Big/ \sum_{i \in I} \frac{1}{|D \ominus \check{X}_i|}$$

Dans le cas d'une population stationnaire ergodique de particules, il a été établi par Miles (1974) que cet estimateur converge en probabilité vers l'indice d'allongement moyen d'une particule typique X dès que $P\{D \ominus \check{X} = \emptyset\} = 0$.

[3]Il existe d'autres algorithmes, tels que le suivi de contour sur trame hexagonale ou dodécagonale mis au point par Schmitt (1989), ou encore les approches de type géométrie par ordinateur sur des polygones (voir Mitchell (2004) et les références incluses).

2 Stéréologie

2.1 Simulation de points uniformes dans un domaine de grande dimension

Ce problème est d'importance pratique considérable, ne serait ce parce que générer un vecteur de densité donnée revient à générer un point uniforme dans son sous-graphe.

Considérons une partie ouverte A de $I\!R^d$, de volume fini. Il n'y a pas de difficulté particulière à générer un point uniforme dans A quand ce domaine est de forme simple (e.g. boule, parallélotope). Lorsque A est de forme plus compliquée, on peut recourir à la méthode de rejet de von Neumann (1951). Elle consiste à enclore A dans un domaine A_0 de forme simple, et à générer des points uniformes dans A_0 jusqu'à obtenir un point dans A qui y est alors implanté de façon uniforme.

L'efficacité de cette méthode se mesure par le nombre moyen de points générés dans A_0, qui n'est autre que le rapport $|A_0|/|A|$ entre le volume de A_0 et celui de A.

Le problème est que la méthode de rejet perd en efficacité au fur et à mesure que la dimension de l'espace de travail augmente (il devient de plus en plus difficile de trouver un domaine A_0 contenant A de près). Elle devient totalement inopérante aux grandes dimensions. Pour donner un ordre de grandeur, il faut en moyenne 10^{70} itérations pour parvenir à générer un point uniforme dans une boule unité à 100 dimensions que l'on a placée dans un hypercube de coté 2.

Plutôt qu'une méthode de rejet, qui n'apporte rien à chaque essai infructueux, nous allons tenter une approche stéréologique où l'on gagne un peu à chaque itération.

L'algorithme que j'ai imaginé est connu dans la littérature sous le nom de hit-and-run[4]

Algorithme *(Point uniforme – algorithme "hit-and-run")*

(i) Soit x un point quelconque de A ;
(ii) Générer une droite D de direction uniforme et passant par x ;
(iii) Générer un point uniforme y le long de $A \cap D$;
(iv) Poser $x = y$ et retourner en (ii).

Du point de vue probabiliste, la suite $(X_n, n \in I\!N)$ des points aléatoires successifs constitue une chaîne de Markov dans A. Son noyau de transition $P(x, B) = P\{X_{n+1} \in B \mid X_n = x\}$ admet la densité

$$f(y \mid x) = \frac{2}{d\omega_d} \frac{1}{\ell(x, y, A)\, |x - y|^{d-1}} \qquad y \neq x$$

où ω_d désigne le volume de la boule unité à d dimensions, et $\ell(x, y, A)$ représente la longueur de l'intersection entre A et la droite passant par x et y.
De l'écriture de cette densité, deux faits retiennent l'attention : d'une part la propriété de positivité $f(y \mid x) > 0$, qui garantit que la chaîne est à la fois irréductible et apériodique ; d'autre part la relation de symétrie $f(y \mid x) = f(x \mid y)$, qui implique que la chaîne laisse invariante la loi uniforme de A. Il s'ensuit que la loi uniforme de A est bien la loi ergodique de la chaîne.

[4] Après coup, J. Besag m'a signalé que cet algorithme avait déjà été proposé par R.L. Smith en 1984. Mais il figure déjà dans un article de V.F. Turčin (1971).

Cet algorithme a ceci de remarquable qu'il ne requiert aucune hypothèse sur le domaine A. Ce domaine n'a pas besoin d'être borné, ni convexe, ni connexe, ni même simplement connexe. La seule contrainte qu'impose l'algorithme est que l'intersection de A et de toute droite doit être de longueur finie. Cette condition se trouve presque sûrement satisfaite dès que A est de volume fini[5].

La vitesse de convergence s'établit simplement lorsque A est borné. Supposons qu'il existe $M > 0$ tel que $|x - y| < M$ pour toute paire de points x et y de A. Dans ce cas, on a aussi $\ell(x, y, A) \leq M$. Il s'ensuit que l'on a pour tout borélien B de A

$$P(x, B) = \frac{2}{d\omega_d} \int_B \frac{dy}{\ell(x, y, A)\, |x - y|^{d-1}} > \frac{2}{d\omega_d} \int_B \frac{dy}{M^d} = \frac{2}{d\omega_d} \frac{|A|}{M^d} \frac{|B|}{|A|}$$

Cela montre que le noyau de transition est minoré par la loi uniforme sur A. Il s'agit du critère de Doeblin qui, moyennant l'irréductibilité et l'ergodicité de la chaîne, garantit que la chaîne est uniformément ergodique (Meyn et Tweedie, 1995)

$$\lim_{n \longrightarrow \infty} \sup_{x \in A} \sup_{B \in \mathcal{B}(A)} \left| P^n(x, B) - \frac{|B|}{|A|} \right| = 0$$

La convergence de l'algorithme est illustrée Figure 4 en considérant un domaine A partagé en deux composantes connexes. Un point initial x a été choisi dans la composante la plus grande. A partir de ce point, on a fait tourner 1000 fois l'algorithme pendant 10 itérations, créant ainsi une loi empirique dont on peut suivre l'évolution au fil des itérations. Au bout de 5 itérations seulement, la loi empirique obtenue est déjà très proche d'une loi uniforme.

2.2 Relation de Cartier

Soit Z une fonction aléatoire stationnaire d'ordre 2. On désigne par $Z(V)$ la moyenne de Z dans le "support" V

$$Z(V) = \frac{1}{|V|} \int_V Z(x) dx$$

Supposons que V est la réunion de n domaines jointifs $v_1, ..., v_n$, tous identiques, à une translation près, à un support v. Soit $\dot{v}$ un ensemble sélectionné au hasard parmi les v_i. Entre les moyennes de $Z(\dot{v})$ et de $Z(V)$, on a ce que Matheron (1984) appelle la relation de Cartier

$$E\big\{Z(\dot{v}) \mid Z(V)\big\} = Z(V)$$

Cette relation est importante car combinée à l'inégalité de Jensen, elle montre que la loi F_v de v est plus dispersée que celle F_V de V, au sens où

$$\int \varphi(z) dF_v(dz) \geq \int \varphi(z) dF_V(dz)$$

pour toute fonction convexe φ. Lorsque $Z \geq 0$, un critère équivalent est l'egalité des moyennes et l'inégalité $C_v \leq C_V$ liant les courbes de concentration des deux lois.

[5] K.T. Smith et D. Solmon ont établi en 1975 que si A est de d-volume fini, alors la longueur de l'intersection de A avec presque toute droite de direction θ est finie, et cela pour presque toutes les directions θ.

FIG. 4 – 1000 simulations de l'algorithme hit-and-run observées aux instants $0, 1, 2, 3, 5$ et 10

A titre d'illustration, prenons pour Z une chaîne de Markov homogène à deux états 0 et 1. Les durées de séjour dans chaque état sont exponentielles de moyennes $1/a_0 = 2$ et $1/a_1 = 1$. La figure 5 ci-contre montre une simulation de ce processus.

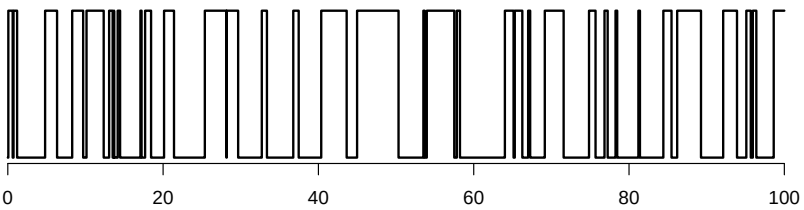

FIG. 5 – Simulation d'une chaîne de Markov homogène à deux états

Soit F_ℓ la loi de la moyenne de ce processus sur $[0, \ell]$. Cette loi comporte un atome en 0 et un atome en 1 de poids respectifs

$$w_0 = \frac{a_1}{a_0 + a_1} e^{-a_0 \ell} \qquad w_1 = \frac{a_0}{a_0 + a_1} e^{-a_1 \ell}$$

Elle prend aussi des valeurs intermédiaires entre 0 et 1 selon la densité

$$\begin{aligned} f_\ell(u) &= e^{-a_0(1-u)\ell - a_1 u\ell} \frac{\sqrt{a_0 a_1}\ell}{a_0+a_1} \Big(2\sqrt{a_0 a_1} I_0\big(2\ell\sqrt{a_0 a_1 u(1-u)}\big) + \\ &\qquad \big(a_0\sqrt{u/(1-u)} + a_1\sqrt{(1-u)/u}\big) I_1\big(2\ell\sqrt{a_0 a_1 u(1-u)}\big)\Big) \end{aligned}$$

qui met en jeu les fonctions de Bessel modifiées I_0 et I_1. La forme de cette densité est représentée figure 6 pour plusieurs valeurs de ℓ.

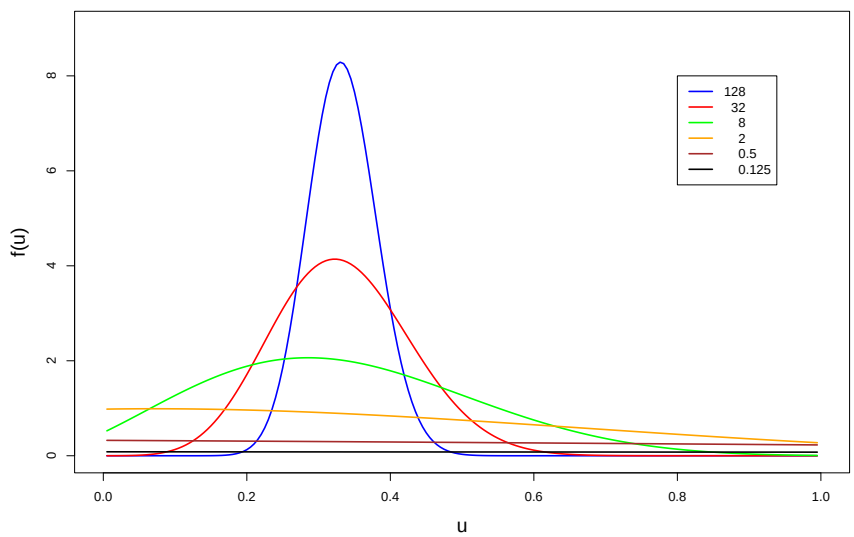

FIG. 6 – Représentation de la composante non atomique de la loi de $Z([0,\ell])$ à différents supports ($\ell = 0.125, 0.5, 2, 8, 32$ et 128)

La relation de Cartier permet aussi la comparaison d'estimateurs stéréologiques, dans l'esprit de Rao-Blackell.

Pour fixer les idées, soit K un compact convexe non vide de $I\!R^3$. La largeur moyenne $\bar{\ell}(K)$ de K est définie comme l'espérance de la longueur projetée $\ell(K \mid \dot{L})$ de K sur une droite $\dot{L}$ d'orientation uniforme. Cette définition suggère de prendre $\ell(K \mid \dot{L})$ comme estimateur sans biais de $\bar{\ell}(K)$. Mais d'autres d'estimateurs peuvent être envisagés. On sait par exemple que la projection de K sur un plan $\dot{P}$ d'orientation uniforme a pour périmètre moyen $\bar{\ell}(K)/\pi$ (formule de Cauchy). D'où un deuxième type d'estimateur sans biais, de la forme $p(K \mid \dot{P})\pi$.
Lequel de ces deux estimateurs vaut il mieux prendre ?
Pour le voir, considérons une droite $\ddot{L}$ d'orientation uniforme contenue dans $\dot{P}$. D'après la formule de Cauchy, on a

$$E\big\{\ell(K \mid \ddot{L}) \mid \dot{P}\big\} = p(K \mid \dot{P})/\pi,$$

ou encore

$$E\big\{\ell(K \mid \ddot{L}) \mid p(K \mid \dot{P})/\pi\big\} = p(K \mid \dot{P})/\pi$$

Ceci montre que les deux estimateurs $\ell(K \mid \ddot{L})$ et $p(K \mid \dot{P})/\pi$ sont en relation de Cartier. Comme $\ddot{L}$ a une orientation uniforme dans $I\!R^3$, il s'ensuit que la loi de $\ell(K \mid \dot{L})$ est plus dispersée que celle de $p(K \mid \dot{P})/\pi$.

Le cas du cube unité va nous servir à illustrer la différence de dispersion entre ces deux estimateurs. 1000 valeurs estimées ont été obtenues en simulant $\dot{P}$ puis $\ddot{L}$, fournissant le nuage bivariable que voici :

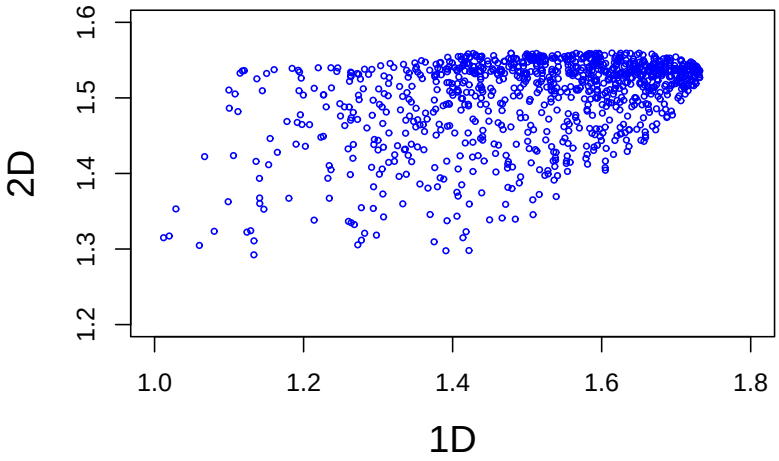

FIG. 7 – Comparaison de deux estimateurs de la largeur moyenne d'un cube unité

La forme assez affirmée du nuage est le signe d'une certaine corrélation les valeurs estimées 1D et 2D (0.424). Bien que cela ressorte assez mal sur la figure, les moyennes expérimentales sont comparables : 1.515 pour l'estimateur 1D et 1.501 pour l'estimateur 2D (la largeur moyenne d'un cube unité est de 1.5). L'écart type 2D est 3 fois plus faible que l'écart type 1D (0.056 au lieu de 0.153).

A trois dimensions, la largeur moyenne de K n'est que l'une de ses 4 fonctionnelles de Minkowski[6]. Le travail qui vient d'être présenté s'étend sans difficulté à l'estimation de toute fonctionnelle de Minkowski de tout compact convexe à d dimensions (Lantuéjoul, 1988).

3 Modèles stochastiques spatiaux

3.1 Tempêtes poissonniennes

Il s'agit d'un prototype de modèle spatial stationnaire max-stable, de loi Fréchet unité. Introduit par Smith (1990), ce modèle a été regardé de façon plus systématique par

[6]Toute fonction numérique définie sur la classe des compacts convexes non vides de $I\!R^d$ et qui est à la fois invariante par déplacement, continue pour la distance de Haussdorff et additive est nécessairement une combinaison de $d+1$ fonctionnelles de base, appelées fonctionnelles de Minkowski, et usuellement notées $W_0, W_1, ..., W_d$. La fonctionnelle W_i est homogène de degré $d-i$. A trois dimensions, les 4 fonctionnelles sont respectivement proportionnelles au volume, à la surface de la fromtière, à la largeur moyenne et à l'unité.

Schlather (2002).

Soit Π un processus de Poisson homogène de densité μ sur $I\!R^d \times I\!R_+$. Soit aussi $\big(Y_{y,t},\ (y,t) \in I\!R^d \times I\!R_+\big)$ une famille de copies indépendantes d'une fonction aléatoire Y définie sur $I\!R^d$, positive et intégrable. On pose ensuite

$$Z(x) = \sup_{(y,t)\in\Pi} \frac{Y_{y,t}(x-y)}{t} \qquad x \in I\!R^d \tag{1}$$

La loi spatiale de Z s'obtient sans grande difficulté. On a

$$P\left\{\forall i \in I\ Z(x_i) < z_i\right\} = \exp\left(-\mu \int_{I\!R^d} E\Big\{\max_{i\in I} \frac{Y(x_i+x)}{z_i}\Big\}\, dx\right) \tag{2}$$

pour toute famille finie $\big((x_i, z_i), i \in I\big)$ de $I\!R^d \times I\!R_+$. Cette formule montre que ce modèle est stationnaire, le membre de droite restant inchangé lorsque les x_i sont translatés selon le même vecteur h.

D'intéressantes simplifications surviennent dans le cas où Y est la fonction indicatrice d'un compact aléatoire X. On a alors

$$P\left\{\forall i \in I\ Z(x_i) < z_i\right\} = \exp\left(-\mu \sum_{\emptyset\neq J\subset I} (-1)^{|J|-1} \frac{E\{|X \ominus \check{C}_J|\}}{\max_{j\in J} z_j}\right) \tag{3}$$

où $C_J = \{x_j,\ j \in J\}$, et où $X \ominus \check{C}_J$ désigne l'érodé de X par C_J, c'est-à-dire l'ensemble des positions où les translatés de C_J sont totalement contenus dans X.

Pour certains ensembles X, le membre de droite pourra être évalué explicitement. Tel est le cas des polygones poissonniens que nous présentons maintenant :

Le point de départ de cet exemple est la paramétrisation des droites du plan $x\cos\alpha + y\sin\alpha = p$ à l'aide d'un paramètre de position $p \in I\!R$ et d'un paramètre d'orientation $\alpha \in [0, \pi[$. Une droite peut donc être vue comme un point dans son espace de paramètres $I\!R \times [0, \pi[$. On considère alors un processus de Poisson de densité λ dans $I\!R \times [0, \pi[$. La figure 8 donne un exemple du réseau de *droites poissonniennes* ainsi construit.

Les domaines délimités par ces droites sont appelés *polygones poissonniens*. Ces polygones possèdent la propriété d'être *conditionnellement invariants par érosion*. Etablie par Matheron (1975), cette propriété dit qu'un polygone érodé, conditionnellement au fait qu'il est non vide, a exactement les mêmes caractéristiques statistiques que le polygone initial[7]. En particulier, si X est un polygone poissonnien et C un ensemble quelconque, on a

$$E\big\{|X \ominus \check{C}| \ \big|\ X \ominus \check{C} \neq \emptyset\big\} = E\{|X|\}$$

On en déduit

$$E\{|X \ominus \check{C}|\} = E\{|X|\} P\{X \ominus \check{C} \neq \emptyset\}$$

[7]On sait qu'un réseau de droites poissonniennes dessine un processus de points poissonniens sur toute droite du plan. Les intervalles exponentiels délimités par ces points possèdent la propriété classique de non vieillissement. C'est précisément cette propriété qui est retranscrite dans un cadre bidimensionnel.

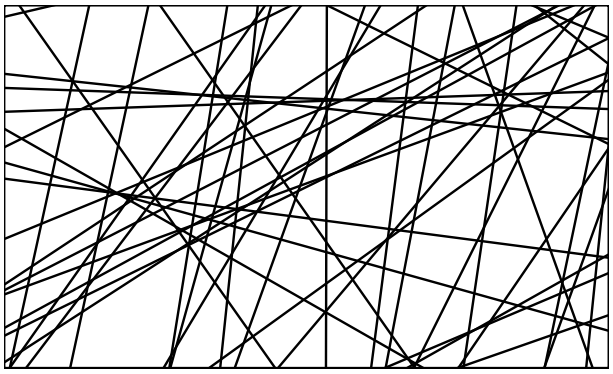

FIG. 8 – Réseau de droites poissonniennes. Les polygones poissonniens sont les domaines délimités par ces droites

La surface moyenne d'un polygone poissonnien vaut $E\{|X|\} = 1/\pi\lambda^2$ (Miles, 1969 ; Matheron, 1975). Par ailleurs, les polygones poissonniens étant convexes, il revient au même d'éroder X par C ou bien par son enveloppe convexe $\hat{C}$. La probabilité pour que cet érodé soit non vide est égale à la probabilité pour que $\hat{C}$ ne soit intersecté par aucune droite poissonnienne. Compte tenu de ce que le nombre de droites qui rencontrent $\hat{C}$ suit une loi de Poisson de moyenne proportionnelle à la densité λ du réseau et au périmètre $p(\hat{C})$ de $\hat{C}$ (Miles, 1969), on obtient

$$E\{|X \ominus \check{C}|\} = \frac{1}{\pi\lambda^2} \exp\bigl(-\lambda p(\hat{C})\bigr)$$

et finalement

$$P\left\{\forall i \in I \; Z(x_i) < z_i\right\} = \exp\left(-\frac{\mu}{\pi\lambda^2} \sum_{\emptyset \neq J \subset I} (-1)^{|J|-1} \frac{\exp\bigl(-\lambda p(\hat{C}_J)\bigr)}{\max_{j \in J} z_j}\right)$$

La figure 9 montre une simulation de tempête poissonnienne. Elle a été obtenue en utilisant un algorithme dû à Miles (1974) pour simuler exactement un polygone poissonnien sans passer par la simulation du réseau de droites. Une méthode de rejet est introduite pour éviter tout effet de bord de champ qui résulte du caractère non borné des polygones (Lantuéjoul, 2005).

Pour finir, mentionnons que la formule (2) peut être généralisée lorsque Z est s.c.s. Dans ce cas, son sous-graphe $S_Z = \{(x, z)\, z \leq Z(x)\}$ est un ensemble fermé aléatoire que l'on peut étudier à l'aide de sa fonction de répartition (Matheron, 1975). En particulier, on peut s'intéresser à la probabilité pour que le compact K implanté en x à l'altitude z ne rencontre pas le sous-graphe, ce qui n'est pas autre chose que la probabilité pour que le maximum $Z^K(x) = \max_{y \in K} Z(x + y)$ de Z sur le compact K translaté en x n'atteigne pas la cote z. Cette approche peut être étendue au cas de plusieurs compacts, et l'on aboutit au résultat suivant :

$$P\left\{\forall i \in I \; Z^{K_i}(x_i) < z_i\right\} = \exp\left(-\mu \int_{\mathbb{R}^d} E\Bigl\{\max_{i \in I} \frac{Y^{K_i}(x_i + x)}{z_i}\Bigr\}\, dx\right)$$

Cete formule est sans doute plus intéressante en pratique parce qu'aussi bien les échantillons que les quantités à estimer sont non ponctuels.

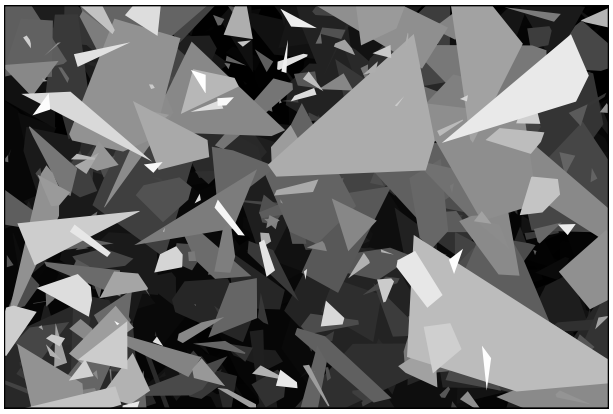

FIG. 9 – Simulation d'une tempête de polygones poissonniens. La teinte de gris des polygones indique leur temps d'arrivée (teinte claire pour ceux survenus tôt, teinte sombre pour les plus tardifs)

3.2 Fonction aléatoire de substitution

Quand on regarde une carte de géographie un peu complexe, on est frappé non seulement par l'information qu'elle contient mais aussi par le codage mis en oeuvre pour rendre cette information lisible. Le modèle de fonction aléatoire que nous avons proposé en 1993 et repris en 2002 est construit selon ce double principe. Une fonction aléatoire de substitution s'écrit

$$Z(x) = Y \circ T(x) \qquad x \in I\!R^d$$

où T est une fonction aléatoire (l'information) et Y est un processus (le codage) indépendant de T.

La grande latitude de choix pour T et pour Y fait qu'il n'est guère possible d'étudier ce modèle dans toute sa généralité. Pour le moment, nos investigations se sont limitées au cas stationnaire.

Un moyen simple de garantir cette propriété est de prendre Y stationnaire et T à accroissements stationnaires stricts. Dans ce cas, la loi ponctuelle de Z n'est autre que celle de Y. Sa fonction de covariance C_Z est liée à celle C_Y du processus de codage par la relation

$$C_Z(h) = E\big\{C_Y\big(T(h)\big)\big\} \qquad h \in I\!R^d$$

Notons en particulier que lorsque le processus de codage est périodique (e.g. $Y(t) = \sqrt{2}\cos(\omega t + \Phi)$ avec Φ uniforme sur $[0, 2\pi[$) et que la loi de $T(h)$ est paire, alors $C_Z(h)$ n'est autre que la transformée de Fourier de $T(h)$ à la fréquence considérée

$$C_Z(h) = E\Big\{e^{i\omega T(h)}\Big\}$$

Mentionnons aussi la propriété dite de *permanence isofactorielle* : si les lois bivariables du processus de codage (supposé ici à valeurs discrètes) admettent des déve-

loppements isofactoriels[8]

$$P\{Y(s)=i, Y(t)=j\} = \sum_{n=0}^{\infty} \lambda_n^{|s-t|}\, u_n(i)\, u_n(j)\, p(i)\, p(j) \qquad s,t \in I\!R,$$

alors il en est de même de celles de la fonction de substitution

$$P\{Z(x)=i, Z(y)=j\} = \sum_{n=0}^{\infty} E\big\{\lambda_n^{|T(x-y)|}\big\}\, u_n(i)\, u_n(j)\, p(i)\, p(j) \qquad x,y \in I\!R^d$$

Cette propriété pourra faciliter l'inférence des paramètres du modèle[9] .

Les deux exemples qui suivent visent à montrer qu'il est possible de contrôler la géométrie et la topologie des réalisations d'une fonction de substitution moyennant un choix judicieux de sa fonction d'information et de son processus de codage.

Pour ce premier exemple, la fonction d'information est construite selon le procédé de Chentsov (1957). Soit $\mathcal{D}$ un réseau de droites poissonniennes. Chaque droite partage $I\!R^2$ en deux demi-plans. Tous les points d'un demi-plan choisi au hasard sont mis à $+1$, tous les points de l'autre à -1. On définit ainsi une fonction aléatoire du plan V_D qui ne prend que les deux valeurs[10] ± 1. On procède ainsi pour toutes les droites du réseau. Finalement, on pose

$$T(x) = \sum_{\substack{D\in\mathcal{D} \\ D\cap]o,x[\neq\emptyset}} V_D(x) \qquad x \in I\!R^2$$

On vérifie bien que T est une fonction à accroissements stationnaires. Son variogramme est linéaire.
En ce qui concerne le processus de codage, on a pris une chaîne de Markov à trois états équiprobables : tous les états ont la même durée moyenne de séjour et les probabilités de transition d'un état à l'autre sont identiques.
Les trois faciès du modèle ainsi construit sont interchangeables et constituent ce que nous appelons une *isopartition*. Ce n'est pas ce que montre la simulation de la figure 10 où l'on voit clairement l'un des faciès occuper une place prépondérante. La dissymétrie apparente entre les trois faciès est le signe d'une absence marquée d'homogénéité spatiale. Celle-ci résulte de ce que la fonction de covariance du modèle est à décroissance lente : sa portée intégrale (Yaglom, 1987 ; Matheron, 1989 ; Lantuéjoul, 1991) est infinie.

Pour le second exemple, j'ai repris la construction de Chentsov, à ceci près que les valuations des demi-plans ne sont maintenant plus attribuées au hasard. On affecte la valeur $+1$ au demi-plan situé au dessus de chaque droite poissonnienne, et la valeur -1 à l'autre demi-plan. La fonction d'information obtenue est une fonction monotone de l'ordonnée des points le long de toute verticale. Combinée au même processus de codage que dans l'exemple précédent, elle fournit des réalisations dont la géométrie pourra évoquer celle d'un assemblage stratigraphique (cf. figure 11) :

[8] Il en est par exemple ainsi lorsque le processus de codage est une chaîne de Markov réversible.

[9] Elle a aussi l'avantage de fournir un procédé de construction simple de modèles spatiaux à lois bivariables isofactorielles, si utiles à la prédiction non linéaire par krigeage disjonctif (Matheron, 1976).

[10] Aucune valuation n'est attribuée aux points de la droite. Cela n'a pas d'importance car il s'agit d'un ensemble de mesure nulle.

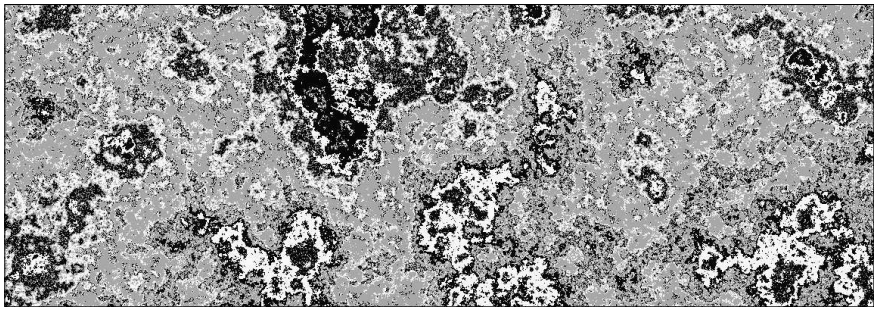

FIG. 10 – Fonction de substitution : simulation d'une isopartition

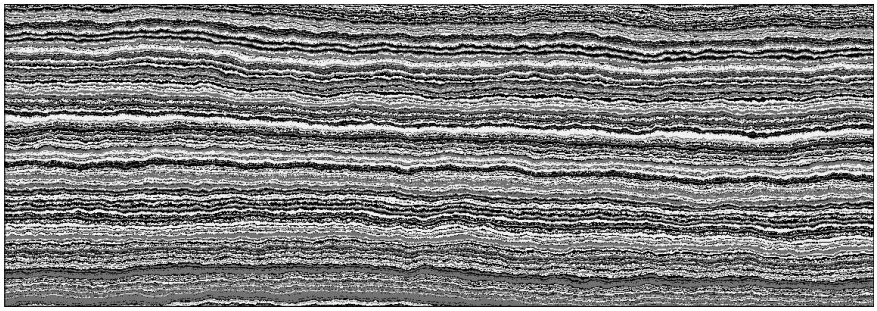

FIG. 11 – Fonction de substitution : simulation d'une stratification

4 Simulation

4.1 Fonction aléatoire gaussienne

Dans ce qui suit, je voudrais parler du travail accompli pour mettre en oeuvre un algorithme de simulation introduit par Matheron en 1973 et appelé *méthode des bandes tournantes*.

Pour présenter cet algorithme, considérons une suite de fonctions aléatoires $(Y_n, n \in \mathbb{N})$ de $\mathbb{R}^d$, mutuellement indépendantes et de même loi spatiale. Supposons-les stationnaires d'ordre 2 et centrées de fonction de covariance C. D'après le théorème central limite, la loi spatiale de

$$Y^{(n)} = \frac{Y_1 + \cdots + Y_n}{\sqrt{n}}$$

tend à devenir gaussienne de covariance C lorsque n tend vers l'infini.

Cette remarque est le point de départ d'un algorithme de simulation approchée d'une fonction gaussienne. Encore faut il savoir comment simuler les fonctions de base ainsi que de déterminer le nombre de fonctions qu'il convient de prendre pour obtenir une simulation gaussienne avec une précision spécifiée à l'avance.

4.1.1 Simulation des fonctions de base

L'approche la plus directe est la méthode spectrale (Shinozuka et Jan, 1972). Elle part du théorème de Bochner qui énonce qu'une fonction continue à l'origine est de type

positif (i.e. une fonction de covariance) si et seulement si elle est la transformée de Fourier d'une mesure positive sommable (mesure spectrale) :

$$C(h) = \int_{I\!R^d} \exp\big(i < u, h >\big)\, d\mathcal{X}(u) \qquad h \in I\!R^d$$

Si $C(0) = 1$, alors $\mathcal{X}$ est une mesure de probabilité. Un calcul immédiat montre que si Ω est une variable aléatoire de loi $\mathcal{X}$, et si U est une variable uniforme sur $]0, 1[$ indépendante de Ω, alors la fonction aléatoire définie par

$$Y(x) = \sqrt{2}\cos\big(< \Omega, x > +2\pi U\big) \qquad x \in I\!R^d$$

est centrée de covariance C.

Cette approche est efficace lorsque quelques vecteurs spectraux suffisent à fournir une approximation empirique raisonnable de la mesure spectrale. La transformée de Fourier échangeant les propriétés du 0 et de l'infini, cette situation se produit lorsque la covariance est régulière au voisinage de l'origine (covariances gaussienne, sinus cardinal...). Dans le cas contraire, cette approche manque d'efficacité et d'autres solutions doivent être envisagées.

C'est ici qu'entre en jeu la méthode des bandes tournantes. Par passage aux coordonnées polaires sous la forme $u = (r, \theta)$ avec $r \in I\!R$ et $\theta \in \frac{1}{2}S_d$ (demi sphère unité), la mesure spectrale $d\mathcal{X}(u)$ se factorise en une loi de direction $d\varpi(\theta)$ et une loi de module connaissant la direction $d\mathcal{X}_\theta(r)$. Moyennant quoi, le théorème de Bochner devient

$$\begin{aligned} C(h) &= \int_{\frac{1}{2}S_d} \int_{I\!R} \exp\big(ir < \theta, h >\big)\, d\mathcal{X}_\theta(r) d\varpi(\theta) \\ &= \int_{\frac{1}{2}S_d} C_\theta\big(< \theta, h >\big)\, d\varpi(\theta) \end{aligned}$$

où C_θ est une covariance sur $I\!R$.

Soit alors un processus centré X_θ de covariance C_θ. Soit aussi Θ une direction aléatoire de loi ϖ. La fonction aléatoire Y définie par

$$Y(x) = X_\Theta\big(< x, \Theta >\big) \qquad x \in I\!R^d$$

est centrée de covariance C.

Cette construction donne lieu à de remarquables simplifications dans le cas d'une fonction aléatoire gaussienne tridimensionnelle isotrope[11] (Matheron, 1973). Posons en effet $C(h) = C_3(|h|)$ pour tout $h \in I\!R^3$. Dans ce cas, toutes les covariances C_θ sont identiques à une fonction paire C_1, caractérisée par la relation

$$C_1(r) = \frac{d}{dr}\Big(r\, C_3(r)\Big) \qquad r \geq 0$$

et tout se ramène à savoir construire des processus centrés de covariance C_1.

[11] Le cas isotrope de dimension quelconque est traîté dans (Matheron, 1973). Le cas anisotrope figure dans (Lantuéjoul, 2002).

Pour ce faire, Matheron propose d'écrire C_1 comme le produit de convolution entre une fonction f et sa symétrisée $\check{f}$ par rapport à l'origine

$$C_1(r) = f * \check{f}(r) = \int_{I\!R} f(t)f(t+r)dt \qquad r \in I\!R$$

ce qui permet de l'interpréter comme la covariance d'un processus de dilution

$$X(t) = \sum_{p \in \mathcal{P}} \varepsilon_p f(t-p) \qquad t \in I\!R$$

où $\mathcal{P}$ désigne un processus de Poisson de densité unité sur $I\!R$, marqué par des valeurs $(\varepsilon_p, p \in \mathcal{P})$ uniformes sur $\{-1, +1\}$ et indépendantes.

En toute généralité, la simulation de X en un point s'obtient en prenant en compte l'ensemble – infini – des points poissonniens. Il y a là une difficulté qui ne disparait complètement que lorsque f est à support borné, ou bien, ce qui est équivalent, lorsque C_1 est de portée finie stricte. On est donc amené à distinguer plusieurs cas :

C est de portée finie stricte : Dans ce cas, C_1 est aussi de portée finie stricte, et l'approche de Matheron est directement applicable. Notons toutefois que son implémentation peut encore être simplifiée en remplaçant le processus de Poisson par un réseau périodique de période égale à la portée (avec une origine uniforme sur la période pour garantir la stationnarité).

A titre d'illustration, regardons le cas de la covariance sphérique de portée a :

$$C_3(r) = \left(1 - \frac{3}{2}\frac{r}{a} + \frac{1}{2}\frac{r^3}{a^3}\right) 1_{0 \le r \le a} \implies C_1(r) \;=\; \left(1 - 3\frac{r}{a} \;+ 2\frac{r^3}{a^3}\right) 1_{0 \le r \le a}$$

La covariance C_1 peut être vue comme l'autoconvoluée de la fonction

$$f(t) \;=\; t\, 1_{|t| \le \frac{a}{2}} \qquad t \in I\!R$$

La Figure 12 montre deux réalisations monodimensionnelles, selon que le processus ponctuel considéré est poissonnien ou périodique.

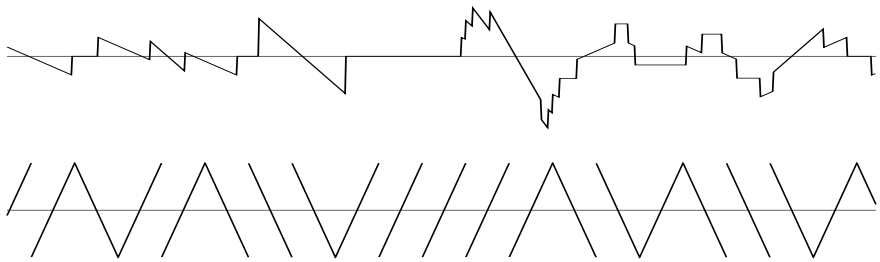

FIG. 12 – Simulation de processus de dilution poissonnien (en haut) et régulier (en bas). Ces deux processus peuvent être utilisés pour simuler une fonction aléatoire gaussienne tridimensionnelle de covariance sphérique par la méthode des bandes tournantes

C **n'est pas de portée finie stricte :** Dans ce cas, la fonction f est de support infini, et la construction précédente secrète un effet de bord qu'il est difficile de quantifier. Il est alors préférable de procéder autrement.

Evacuons d'emblée le cas où la covariance C est différentiable en 0. Il en est alors de même pour C_1, de sorte que la méthode spectrale s'applique efficacement au processus monodimensionnel. Mais cela revient à appliquer directement la méthode spectrale à la fonction tridimensionnelle elle-même. La simulation en bas à gauche de la figure 14 (covariance gaussienne) a été obtenue selon ce procédé.

Considérons maintenant le cas où C est non différentiable en 0, pour lequel la méthode spectrale est peu efficace. Le cas de la covariance exponentielle va servir à illustrer notre propos.

Une première idée est d'écrire C comme mélange d'une famille de covariances, toutes de portée finie stricte. C'est ainsi que X. Emery a établi qu'une covariance exponentielle est un mélange de covariances sphériques :

$$\exp\left(-\frac{|h|}{a}\right) = \int_0^\infty \left(1 - \frac{3}{2}\frac{|h|}{t} + \frac{1}{2}\frac{|h|^3}{t^3}\right) 1_{|h|\leq t}\, p(t)\, dt$$

où p est un mélange de deux densités gamma de paramètres 2 et 3, et de même indice $\frac{1}{a}$ (proportions respectives $\frac{1}{3}$ et $\frac{2}{3}$). Cela permet de simuler chaque processus X exactement comme dans le cas sphérique, à ceci près que sa portée doit être générée selon la densité p (Emery et Lantuéjoul, 2006).

Mais d'autres possibilités existent, basées sur la forme explicite de C_1 :

$$C_3(r) = \exp\left(-\frac{r}{a}\right) 1_{r\geq 0} \implies C_1(r) = \left(1 - \frac{r}{a}\right) \exp\left(-\frac{r}{a}\right) 1_{r\geq 0}$$

On peut recourir ici à un cas particulier d'un processus de migration (Haas et al., 1967). Cela consiste à prendre un processus de Poisson qui découpe la droite en intervalles exponentiels de longueur moyenne égale à $2a$. La première moitié de chaque intervalle est mise à -1, la seconde à $+1$ (cf. Figure 13).

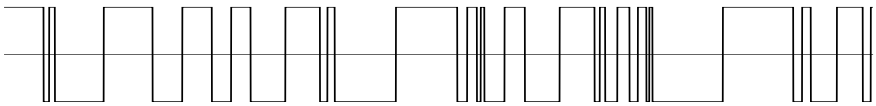

FIG. 13 – Simulation d'un processus de migration. Ce processus peut être utilisé pour simuler une fonction aléatoire gaussienne tridimensionnelle de covariance exponentielle par la méthode des bandes tournantes

La Figure 14 montre des réalisations de fonctions gaussiennes, de covariances sphérique, exponentielle, gaussienne et hyperbolique. La covariance hyperbolique peut être vue comme comme une covariance exponentielle de facteur d'échelle a/T, où T suit une loi exponentielle

$$\frac{1}{1 + |h|/a} = \int_0^\infty e^{-|h|t/a}\, e^{-t}\, dt \qquad h \in I\!R^3$$

Il s'ensuit que les processus associés peuvent être simulés comme des processus exponentiels, et donc aussi comme des processus sphériques.

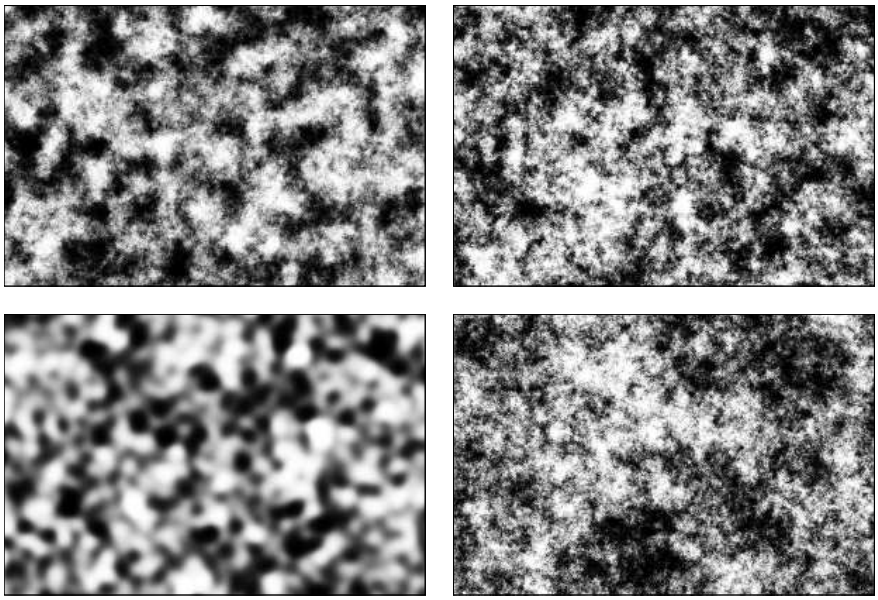

FIG. 14 – Simulation de fonctions aléatoires gaussiennes de covariance sphérique (en haut à gauche), exponentielle (en haut à droite), gaussienne (en bas à gauche) et hyperbolique (en bas à droite)

4.1.2 Combien de fonctions de base ?

Pour faciliter le choix de n, il faut chercher à quantifier l'écart séparant la loi spatiale de $Y^{(n)}$ de celle d'une fonction aléatoire gaussienne centrée normée de covariance C. Ceci peut être regardé sur les combinaisons linéaires de $Y^{(n)}$. De façon plus précise, soit $\lambda = \sum_{j=1}^{p} \lambda_j \delta(x_j)$ une mesure de $I\!R^d$ qui est une combinaison linéaire de masses de Dirac, et soit $Y^{(n)}(\lambda) = \sum_{j=1}^{p} \lambda_j Y^{(n)}(x_j)$ la variable aléatoire qui lui correspond. Il s'agit de voir en quoi la loi de $Y^{(n)}(\lambda)$ diffère d'une loi gaussienne centrée de variance $\sigma^2 = \sum_{j,k=1}^{p} \lambda_j \lambda_k C(x_j - x_k)$?

Pour regarder cela, nous avons essentiellement expérimenté deux approches :

– *L'approche des moments :* Pour $n = 1, 2, ..., \infty$, désignons par $\mu_k^{(n)}$ le moment absolu d'ordre k de $Y^{(n)}(\lambda)$. Compte tenu de $\mu_2^{(n)} = \sigma^2$, et de ce que les moments absolus d'ordre impair sont difficiles à calculer, le premier moment à considérer est celui d'ordre 4. Pour une variable gaussienne centrée, on sait que ce moment vaut 3 fois le carré de la variance, c'est-à-dire $\mu_4^{(\infty)} = 3\sigma^4$. On peut établir que la différence $\mu_4^{(n)} - \mu_4^{(\infty)}$ est inversement proportionnelle au nombre de fonctions de base :

$$\mu_4^{(n)} - 3\sigma^4 = \frac{1}{n}\left(\mu_4^{(1)} - 3\sigma^4\right)$$

– *L'approche distributionnelle*
Le théorème de Berry-Esséen (Feller, 1971) donne un majorant absolu de la différence entre la loi de $Y^{(n)}(\lambda)$ normée et une loi gaussienne standard :

$$\sup_{y \in I\!R} \left| P\left\{ \frac{Y^{(n)}(\lambda)}{\sigma/\sqrt{n}} < y \right\} - G(y) \right| < \frac{\alpha}{\sqrt{n}} \frac{\mu_3^{(1)}}{\sigma^3}$$

où α est une constante numérique inférieure à 1.32132 (Feller, 1971). Ce moment n'est pas toujours numériquement accessible, mais l'inégalité de Cauchy-Schwarz $\mu_3^{(1)} \leq \sigma\sqrt{\mu_4^{(1)}}$ permet de le remplacer par le moment d'ordre 4.

Des résultats numériques fournis par ces deux approches sont présentés dans (Lantuéjoul, 1994).

4.2 Schéma booléen

Ce modèle spatial, devenu classique, fait l'objet d'une littérature abondante (Matheron, 1967, 1975 ; Serra, 1982 ; Stoyan *et al.*, 1987 ; Hall, 1988 ; Molchanov, 1997). Sa construction repose sur les deux ingrédients suivants :

(i) un processus de Poisson ponctuel $\mathcal{P}$, de densité $\theta = (\theta(x), x \in I\!R^d)$;

(ii) une famille $(A(x), x \in I\!R^d)$ de compacts aléatoires non vides – appelés ici "objets" – et mutuellement indépendants de $I\!R^d$. Les propriétés statistiques de l'objet $A(x)$ peuvent être caractérisées par sa fonction de répartition[12] : il s'agit de l'application T_x qui associe à tout compact de $I\!R^d$ la probabilité d'être rencontré par $A(x)$:

$$T_x(K) = P\{A(x) \cap K \neq \emptyset\} \qquad K \in \mathcal{K}(I\!R^d)$$

Moyennant ces ingrédients, le schéma booléen est la réunion des objets implantés aux points du processus de Poisson :

$$X = \bigcup_{x \in \mathcal{P}} A(x)$$

On s'intéresse ici au problème de la simulation d'un schéma booléen dans un champ compact D sachant que deux ensembles finis de points C_1 et C_0 sont respectivement contenus dans X et dans son complémentaire X^c. Ce problème est posé par les ingénieurs pétroliers qui ont besoin de connaître la géométrie de leur réservoir à l'entrée d'un programme de simulation d'écoulement afin de prédire la quantité d'huile récupérable. Les ensembles C_0 et C_1 correspondent à des données lithologiques fournies par les forages[13].

[12] Cet outil a été mis en place par Matheron (1975) pour caractériser la loi des ensembles fermés aléatoires. Il a le statut mathématique d'une capacité de Choquet alternée d'ordre infini.

[13] Le problème est présenté ici dans sa version minimaliste car bien d'autres informations sont souvent disponibles. Par exemple, l'interprétation de données sismiques donne une idée de la proportion de faciès dans certaines parties du réservoir. Les essais de puits renseignent sur la topologie du faciès renfermant l'huile. Pour les réservoirs en exploitation, l'historique de production fournit aussi de précieuses informations...

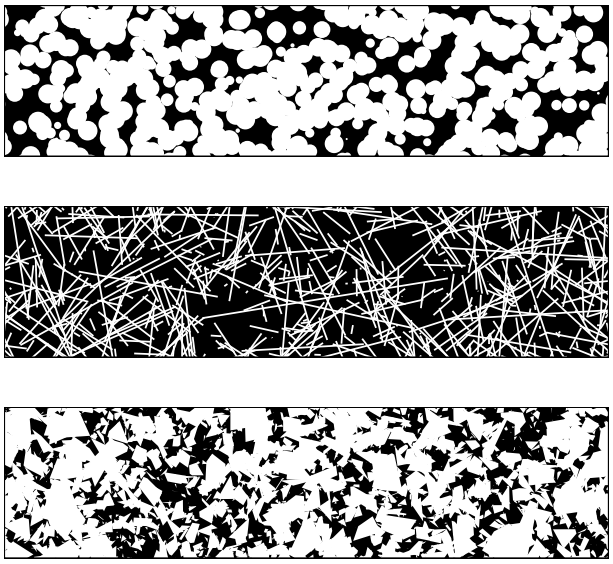

FIG. 15 – Schémas booléens de disques, de segments et de polygones poissonniens

Commençons par une première remarque : l'ensemble $X \cap D$ est encore un schéma booléen, mais ses paramètres ne sont pas ceux de X. Sa densité au point x dépend non seulement de la densité initiale $\theta(x)$, mais aussi de la chance qu'un objet implanté en ce point rencontre D. Explicitement elle vaut $\theta(x)T_x(D)$. De son côté la fonction de répartition d'un objet implanté en x n'est autre que $T_x(K)/T_x(D)$ pour tout compact K de D. C'est bien sûr avec ces nouveaux paramètres que doit s'effectuer la simulation[14].

Supposons ici la densité de $X \cap D$ intégrable :

$$\int_{I\!R^d} \theta(x)T_x(D)dx = \vartheta(D) < \infty$$

Dans ce cas, on peut montrer que $X \cap D$ a exactement la même fonction de répartition que la réunion d'un nombre poissonnien – de moyenne $\vartheta(D)$ – d'objets "typiques" (i.e. délocalisés) indépendants. La fonction de répartition de tels objets a pour expression

$$T(K) = \int_{I\!R^d} \frac{\theta(x)T_x(D)}{\vartheta(D)} \frac{T_x(K)}{T_x(D)} dx = \frac{\vartheta(K)}{\vartheta(D)} \qquad K \in \mathcal{K}(D)$$

De cette interprétation découle un algorithme itératif pour la simulation *non conditionnelle* de la population des objets typiques du schéma booléen. La simulation de la loi de Poisson est assurée par l'algorithme de Metropolis (1953) :

Algorithme *(Schéma booléen non conditionnel)*
(i) Poser $\Phi = \emptyset$ *;*

[14]Ce changement de paramètres n'est pas propre à la simulation. De la même façon, si D représente non pas le champ de simulation mais le support de l'union des échantillons, alors le modèle perçu à partir des échantillons est un schéma booléen dont la densité et la fonction de répartition des objets sont donnés par les formules du texte. Le biais qui en résulte doit être corrigé, ce qui rend l'inférence statistique particulièrement délicate.

(ii) Générer u qui vaut $+1, -1$ et 0 avec les probabilités respectives

$$p_{+1} = \frac{\vartheta(D)}{\vartheta(D) + \#\Phi + 1} \qquad p_{-1} = \frac{\#\Phi}{\vartheta(D) + \#\Phi} \qquad p_0 = 1 - p_{+1} - p_{-1};$$

(iii) De trois choses l'une :
– si $u = +1$, générer un objet typique et l'insérer à Φ ;
– si $u = -1$, sélectionner un objet typique au hasard et l'enlever de Φ ;
– si $u = 0$, ne rien faire ;
(iv) retourner en (ii).

Il suffit maintenant de prescrire le respect des conditions à chaque itération pour rendre cet algorithme conditionnel. La correction de cette approche est établie lorsque le noyau de transition qui régit la chaîne non conditionnelle est *réversible* et lorsque le noyau induit qui régit la chaîne conditionnelle est *irréductible* (Lantuéjoul, 2002). La réversibilité est automatiquement assurée par l'algorithme de Metropolis. L'irréductibilité résulte de la stabilité par union des populations respectant les conditions.

Algorithme *(Schéma booléen conditionnel)*
(i) Générer une population Φ d'objets typiques respectant les conditions C_1 et C_0 ;
(ii) Générer u qui vaut $+1, -1$ et 0 avec les probabilités respectives

$$p_{+1} = \frac{\vartheta(D)}{\vartheta(D) + \#\Phi + 1} \qquad p_{-1} = \frac{\#\Phi}{\vartheta(D) + \#\Phi} \qquad p_0 = 1 - p_{+1} - p_{-1};$$

(iii) De trois choses l'une :
– si $u = +1$, générer un objet typique A. Insérer A à Φ si $A \cap C_0 = \emptyset$;
– si $u = -1$, sélectionner un objet typique A au hasard. Enlever A de Φ si $C_1 \subset \bigcup_{B \in \Phi \setminus \{A\}} B$;
– si $u = 0$, ne rien faire ;
(iv) retourner en (ii).

L'étape d'initialisation (i) a pour fonction de générer une population compatible avec les conditions. Une façon simple de procéder est de simuler des objets typiques l'un après l'autre en rejetant ceux qui contiennent un point de C_0, et en continuant jusqu'à ce que tous les points conditionnants de C_1 aient été recouverts. Bien entendu, cette étape n'aboutit jamais lorsque les conditions ne sont pas compatibles avec le modèle.

Dans le cas $u = 0$ de l'étape (iii), il est préconisé de ne rien faire. Il est aussi possible de procéder au remplacement d'un objet typique, pourvu que le nouvel objet généré assure le maintien des conditions.

La figure 16 montre un exemple de simulation conditionnelle. Le modèle est un schéma booléen stationnaire de disques (rayon exponentiel de moyenne 5). La densité a été choisie de façon à ce qu'en moyenne 100 disques rencontrent le champ de simulation 100×100. De la simulation non conditionnelle (en haut à gauche), 100 points répartis uniformément dans le champ ont été prélevés (en bas à gauche) pour servir de points conditionnants. Au milieu et à droite figurent deux simulations conditionnelles, représentées avec et sans leurs points conditionnants.

L'utilisation d'un algorithme itératif de simulation pose immanquablement le problème de la détermination de sa vitesse de convergence.

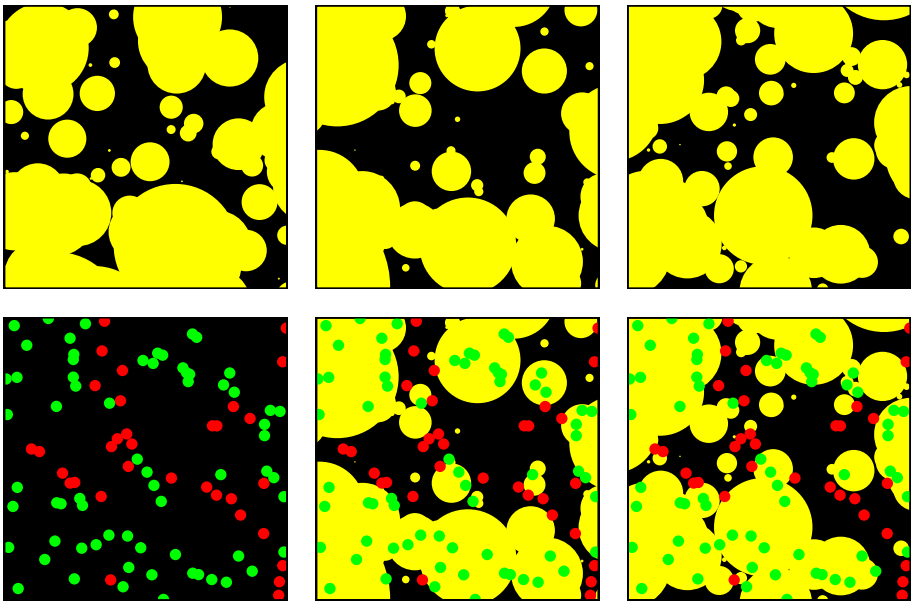

FIG. 16 – Simulation conditionnelle d'un schéma booléen. En haut à gauche une simulation non conditionnelle. En bas à gauche, un jeu de points conditionnants. Au milieu et à droite, deux simulations conditionnelles, représentées avec (en bas) et sans (en haut) leurs points conditionnants

A cette fin, on va comparer la fonction de répartition de l'ensemble aléatoire X_t obtenu au bout de t itérations et celle du schéma booléen conditionnel X_∞. Pour résumer simplement les choses, il peut être établi que

$$|P\{X_t \cap K \neq \emptyset\} - P\{X_\infty \cap K \neq \emptyset\}| = O(\lambda^t) \qquad K \in \mathcal{K}(D)$$

où λ désigne la seconde plus grande valeur propre du noyau de transition qui régit l'évolution du nombre d'objets de la simulation au cours des itérations (Lantuéjoul, 1998). Tout se ramène donc à l'estimation de λ.

Une façon naturelle de procéder est de lancer le programme de simulation et de le faire tourner pendant un certain nombre d'itérations. On enregistre le nombre d'objets présents à chaque itération pour construire un noyau de transition empirique dont on calcule ensuite les valeurs propres.

Cette méthode est très imprécise. Elle n'offre aucune garantie sur la qualité d'estimation du noyau de transition. Cela est d'autant plus préoccupant que λ doit être estimée avec précision : deux valeurs voisines telles que 0.99 et 0.999 correspondent à des vitesses de convergence qui diffèrent d'un ordre de grandeur.

Soit N_t le nombre aléatoire d'objets à l'itération t, et soit $\mathcal{X}$ un vecteur propre associé à λ. Des techniques de portée intégrale (Yaglom, 1987 ; Matheron, 1989 ; Lantuéjoul, 1991) montrent que l'on doit avoir

$$Var\left\{\frac{1}{t}\sum_{u=1}^{t}\mathcal{X}(N_u)\right\} \approx \frac{A}{t}$$

lorsque t est suffisamment grand. Par passage au logarithme, cette formule se récrit

$$\ln Var\left\{\frac{1}{t}\sum_{u=1}^{t}\mathcal{X}(N_u)\right\} \approx \ln A - \ln t$$

En traçant le logarithme de cette variance par rapport au logarithme de t, on doit donc s'attendre à voir apparaître une droite de pente -1, ce qui peut faire l'objet d'un contrôle expérimental.

Voici la procédure utilisée pour déterminer le nombre d'itérations à prendre en compte. On commence par fixer un nombre t_0. On fait tourner deux fois l'algorithme de simulation pendant t_0 itérations. La première simulation est utilisée pour construire un estimateur $\hat{\mathcal{X}}$ de $\mathcal{X}$ à partir d'un noyau de transition empirique[15]. Remplaçant $\mathcal{X}$ par son estimateur, la deuxième simulation est utilisée pour estimer $Var\left\{\frac{1}{t}\sum_{u=1}^{t}\mathcal{X}(N_u)\right\}$ pour plusieurs supports t diviseurs de t_0[16]. On stoppe la procédure si les points obtenus pour t suffisamment grand s'organisent selon une droite de pente -1 en coordonnées log-log. Sinon le nombre t_0 est incrémenté et la procédure est recommencée.

Une fois le nombre d'itérations évalué, on peut obtenir un estimateur $\hat{A}$ de A par simple lecture graphique. Il se trouve que l'on a

$$A = \frac{1+\lambda}{1-\lambda}$$

moyennant une normation convenable de $\mathcal{X}$. Il suffit alors de poser

$$\hat{\lambda} = \frac{\hat{A}-1}{\hat{A}+1}$$

comme valeur estimée de λ, le passage de A à λ se faisant dans le sens de la stabilité. A titre indicatif, la valeur propre λ associée au modèle de la Figure 16 a été estimée à 0.993, ce qui correspond à une portée intégrale d'environ 330 itérations.

5 Quelques thèmes de recherche

Dans ce qui suit, je dresse une liste de quelques thèmes de recherche que je souhaiterais approfondir. Cette liste n'est pas exhaustive, et bien d'autres thèmes auraient pu être mentionnés (échantillonnage, problèmes d'inférence, mesures de connexité ...). Les thèmes considérés ici ne semblent pas tous liés à une application immédiate. Et pourtant, tous ces thèmes trouvent leur origine dans le monde applicatif. Mais il n'est pas toujours nécessaire de les maintenir dans leur contexte pour les étudier ; il est même parfois avantageux de les en sortir pour mieux le faire. Par exemple, le conditionnement de simulations dans des modèles différents achoppe souvent sur des difficultés algorithmiques de même nature. Dans ce cas, il est préférable de traiter le problème de façon générique, c'est-à-dire libre de toute contrainte applicative explicite.

La liste proposée ci-desous n'est pas exhaustive. Même si elle mentionne à la fois des problèmes récurrents et des préoccupations du moment, elle ne fait que refléter mes propres goûts, et sans doute ma méconnaissance d'un certain nombre de sujets...

[15] La situation est ici avantageuse : les composantes des vecteurs propres associés à la seconde plus grande valeur propre d'un noyau de diffusion discret varient de façon monotone.

[16] Cela consiste en un calcul de variance expérimentale de dispersion.

5.1 Choix de modèles

Le lecteur est invité à regarder les trois images de la Figure 17 :

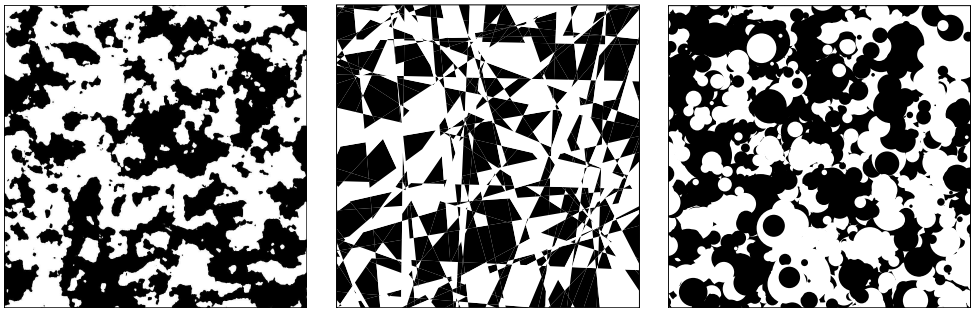

FIG. 17 – Réalisations d'une fonction aléatoire gaussienne seuillée (à gauche), d'un pavage poissonnien (au milieu) et d'un schéma des feuilles mortes (à droite). Ces trois ensembles aléatoires admettent exactement les mêmes lois trivariables.

Il s'agit de simulations de trois modèles d'ensembles aléatoires : de gauche à droite, une fonction aléatoire gaussienne seuillée, un pavage poissonnien et un schéma des feuilles mortes. Les paramètres de ces trois ensembles ont été conçus pour leur conférer la même covariance exponentielle (pour plus de détails, voir Chilès et Lantuéjoul (2005)). Etant autoduaux[17], ils ont les mêmes lois bivariables, et a fortiori les mêmes lois trivariables[18]

Quelles données (nombre, support, répartition spatiale) permettent de discriminer entre ces trois modèles ? Cette question est d'autant plus dérangeante qu'un coup d'oeil sur leur frontière suffit à le faire : les contours des feuilles mortes sont circulaires, ceux des polygones poissonniens rectilignes tandis que ceux de la fonction gaussienne seuillée ne sont que rectifiables.

Plus généralement, on est amené à se demander quelle est l'information que contient l'ensemble des lois multivariables à n points d'appui ? Que vaut cette information par rapport à celle contenue dans la loi spatiale ? Quel est l'apport d'un point d'appui supplémentaire ?

5.2 Simulation conditionnelle

La perspective d'avoir à simuler des modèles de plus en plus élaborés sous des contraintes sans cesse plus riches et plus sophistiquées suscite deux questions, duales l'une de l'autre :

– étant donné un modèle, quelles sont les contraintes que l'on sait prendre en compte lors d'une simulation ?

– étant données des contraintes, quels sont les modèles que l'on sait simuler conditionnellement ?

Dans l'état actuel des choses, notre opinion est que nous ne savons simuler conditionnellement que quelques modèles simples, et encore pourvu qu'ils soient faible-

[17]Un ensemble est dit *autodual* s'il admet la même loi spatiale que son complémentaire.

[18]Plus généralement, la formule d'inclusion-exclusion montre qu'un ensemble autodual a ses lois à $2n+1$ points d'appui complètement spécifiées par celles à $2n$ points d'appui.

ment contraints. Dès que le jeu des contraintes s'alourdit, la simulation conditionnelle des modèles même les plus rudimentaires semble soulever de redoutables difficultés d'ordre combinatoire.
Pour ne donner qu'un exemple, considérons le problème de la simulation d'un processus de Poisson dans $I\!R^d$ sous contrainte de la connaissance du nombre de points $(n_i, i \in I)$ tombés dans les domaines $(A_i, i \in I)$. Ce problème ne paraît pas hors de portée, car qu'il y a-t-il de plus simple qu'un processus constitué de points mutuellement indépendants ?

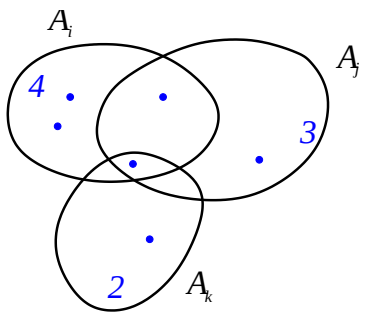

FIG. 18 – Comment simuler un processus de Poisson connaissant le nombre de points tombés dans une famille de domaines ?

Supposons le processus homogène de densité θ. Associons à toute partie non vide J de I l'ensemble A_J des points qui appartiennent à tous les domaines indexés par J et à aucun de ceux indexés par $I \backslash J$. Désignons aussi par N_J le nombre aléatoire de points que renferme chaque A_J. En raison de l'indépendance du nombre de points tombés dans des domaines deux à deux disjoints, la loi multivariable des N_J admet une expression de la forme

$$P\{\cap_J N_J = n_J \mid \cap_i N_i = n_i\} \propto \prod_J \frac{\theta_J^{n_J}}{n_J!} \prod_i 1_{\sum_{J \ni i} n_J = n_i}$$

où l'on a posé $\theta_J = \theta |A_J|$. Avec quelques domaines seulement, une méthode de rejet est tout à fait applicable[19]. Il en va tout autrement lorsque le nombre de domaines est grand. Cent domaines peuvent donner lieu à 2^{100} intersections...

En 2002, nous avons mis au point un algorithme itératif de simulation conditionnelle, reposant sur un noyau de Metropolis dans l'espace des populations de points autorisées. Son intérêt est toutefois très limité car il suppose que l'on puisse accéder à la valeur explicite de chaque θ_J à chaque fois que l'on en a besoin, ce qui est rarement le cas en pratique. Dans bon nombre de situations, les domaines A_i sont très enchevêtrés, de sorte que la seule information tangible sur laquelle on peut s'appuyer est le fait qu'un point quelconque de l'espace appartient à un domaine ou non. Cela permet de construire des estimateurs sans biais de θ_J, mais pas d'accéder à sa valeur effective.

[19]Elle revient à simuler non conditionnellement un processus de Poisson dans la réunion des A_i, et à ne garder que les simulations qui respectent les contraintes.

5.3 Simulation exacte

L'algorithme markovien de simulation conditionnelle du schéma booléen montre que l'établissement de la vitesse de convergence reste un problème ardu. Pour s'affranchir de ce problème, des méthodes de simulation "exacte" ou "parfaite" ont été développées. Elles permettent de générer un état selon la loi cible en un nombre fini d'itérations.
C'est dans ce cadre que figure le travail pionnier de Propp et Wilson (1996). Nous le reprenons ici, tout en adoptant la présentation plus synthétique de Diaconis et Freedman (1999).
Soit Ω un espace d'états (supposé ici discret), et soit p une loi qui charge tous les points de Ω. On suppose que p est la loi limite d'une chaîne de Markov irréductible et apériodique, de noyau de transition P :

$$\lim_{n \longrightarrow \infty} P^n(x,y) = p(y) \qquad x, y \in \Omega$$

On sait qu'il existe des procédés de construction d'une application aléatoire F de Ω dans Ω compatible avec le noyau de transition, c'est-à-dire telle que

$$P\{F(x) = y\} = P(x,y) \qquad x, y \in \Omega$$

Considérons une suite $(F_n, n \in I\!N)$ d'applications aléatoires indépendantes, de même loi que F. L'algorithme de simulation markovienne standard consiste à partir d'un état initial x puis de construire la suite des états itérés $x_n = F_n \circ \cdots \circ F_1(x)$. L'algorithme de simulation exacte procède en sens inverse. On part de Ω, et l'on construit la suite des ensembles itérés $\Omega_n = F_1 \circ \cdots \circ F_n(\Omega)$. Cette suite est décroissante :

$$\Omega_{n+1} = F_1 \circ \cdots \circ F_{n+1}(\Omega) = F_1 \circ \cdots \circ F_n \circ F_{n+1}(\Omega) \subset F_1 \circ \cdots \circ F_n(\Omega) = \Omega_n$$

Il existe des conditions sur la loi de F qui garantissent que $\Omega_\infty = \lim_{n \longrightarrow \infty} \Omega_n$ est presque sûrement un singleton[20] . Le singleton aléatoire ainsi obtenu suit la loi cible p[21]. Sous cette forme, l'algorithme de Propp et Wilson n'est pas utilisable pour de grands espaces d'états : son temps d'exécution est lent et surtout la capacité de mémoire qu'il requiert est importante. Une façon de s'en tirer est de structurer l'espace. Propp et Wilson ont examiné le cas où Ω est partiellement ordonné avec un état minimal et un état maximal. Si F est croissante pour l'ordre partiel, toutes les trajectoires sont prises en sandwich entre celles issues des deux états extrémaux. Il suffit alors de tester la rencontre (la "coalescence") des deux trajectoires extrémales pour savoir quand un singleton a été obtenu.

[20] L'application aléatoire F induit le noyau de transition $P_F(X,Y) = P\{F(X) = Y\}$ sur les parties non vides de Ω. Une CNS pour avoir $\#\Omega_\infty = 1$ p.s. est que les singletons de Ω soient les seuls états récurrents de P_F.

[21] Définissons $A_n(x)$ comme l'événement $\Omega_n = \{x\}$ pour tout entier n et tout état x. D'après le lemme de Fatou, on a

$$P\big\{\Omega_\infty = \{x\}\big\} = P\big\{\underline{\lim} A_n(x)\big\} \leq \underline{\lim} P\big\{A_n(x)\big\}$$

Il s'ensuit

$$P\big\{\Omega_\infty = \{x\}\big\} \leq \underline{\lim} P\{F_1 \circ \cdots \circ F_n(\Omega) = x\} \leq \underline{\lim} P\{F_1 \circ \cdots \circ F_n(x) = x\} = \underline{\lim} P^n(x,x) = p(x)$$

L'hypothèse que Ω_∞ est presque sûrement un singleton implique alors $P\big\{\Omega_\infty = \{x\}\big\} = p(x)$ pour tout état x de Ω.

Cette approche a été appliquée avec succès par Propp et Wilson pour simuler le modèle d'Ising. Plus récemment Kendall et Thönnes (1999) sont parvenus à mettre au point un algorithme exact de simulation conditionnelle du schéma booléen en supposant les compacts aléatoires de taille bornée. C'est là un exemple typique de simulation exacte dans un espace continu d'états. Mais d'autres exemples ont été considérés, tels que la simulation exacte d'un point uniforme dans un domaine ouvert A selon l'approche du hit-and-run (Tournois, 1998). Dans ce cas l'application aléatoire F est spécifiée par une direction Θ et une proportion U. Etant donné le point x de A, on désigne par $D(x, \Theta)$ la droite de direction Θ passant par x. $F(x)$ est alors défini comme le point qui partage $A \cap D(x, \Theta)$ en deux parties dans le rapport U. Tous les points de $A \cap D(x, \Theta)$ ont ainsi la même image par F. La Figure 19 montre les images de A par F lorsque la direction est de $45°$ et que la proportion prend les trois valeurs $0.25, 0.50$ et 0.75. Cet exercice ne peut toutefois pas être considéré comme concluant car les images ont une forte tendance à se morceler, ce qui est très difficile à gérer du point de vue informatique.

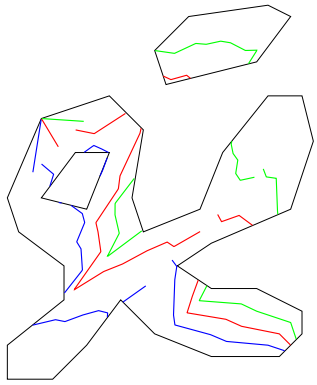

FIG. 19 – Exemples d'application aléatoire pour la simulation uniforme exacte selon l'approche du hit-and-run.

5.4 Modèles statistiques spatiaux de données compositionnelles

On appelle modèle spatial de données compositionnelles à d dimensions et à n états toute famille $Z = (Z_1, ..., Z_n)$ de n fonctions aléatoires positives et reliées par la condition $Z_1(x) + \cdots + Z_n(x) = 1$ pour tout point $x \in I\!R^d$.

Ce modèle reste compositionnel à tout support v. En effet, si l'on pose

$$Z_i(v) = \frac{1}{|v|} \int_v Z_i(x) dx \qquad 1 \leq i \leq n$$

on s'aperçoit que l'on a encore $Z_1(v) + \cdots + Z_n(v) = 1$.

Soit $Z(v) = \big(Z_1(v), ..., Z_n(v)\big)$. Il n'est pas difficile d'établir que si $\dot{x}$ désigne un point uniforme de v, alors $Z(\dot{x})$ et $Z(v)$ sont en relation de Cartier (Matheron, 1984)

$$E\{Z(\dot{x})|Z(v)\} = Z(v),$$

ce qui implique, par l'inégalité de Jensen, que l'on a

$$\int_{I\!R^d} \varphi(z) dF_v(z) \leq \int_{I\!R^d} \varphi(z) dF(z)$$

pour toute fonction numérique convexe φ de $I\!R^d$.

La construction de modèle spatial de données compositionnelles calculable n'est pas simple. Compte tenu de ce qu'en de nombreuses situations, les échantillons ne sont pas tous de même support ou bien que l'estimation doit être faite à un support qui n'est pas celui des échantillons, ce modèle doit pouvoir être accessible à des supports différents. De ce point de vue, le recours à la loi de Dirichlet

$$f(z_1, ..., z_n) = \frac{\Gamma(\alpha_1 + \cdots + \alpha_n)}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_n)} z_1^{\alpha_1 - 1} ... \, z_n^{\alpha_n - 1} 1_{z_1 + \cdots + z_n = 1} \qquad z_1, ..., z_n > 0$$

préconisé par Haas et Formery (2002) comme loi multivariable à un support de référence n'est pas totalement satisfaisant dans le domaine des sciences de la terre. Cette loi ne comporte pas de composante atomique (ce qui exclut la possiblité que le support soit totalement contenu dans un seul faciès). De plus, elle n'est en général pas préservée par changement de support[22].

Une façon de lever la difficulté est de partir d'une fonction aléatoire de $I\!R^d$ étagée sur n niveaux et de régulariser chacun de ses niveaux à l'ensemble des supports prescrits par l'étude. De tels modèles de fonctions aléatoires ne manquent pas (fonctions de dilution (Matérn, 1960 ; Serra, 1968), fonctions booléennes (Jeulin et Jeulin, 1981 ; Préteux et Schmitt, 1988), schéma des feuilles mortes (Jeulin, 1991), fonctions mosaïques (Lantuéjoul, 2002), fonctions plurigaussiennes (Armstrong *et al.*, 2003), mais ces modèles sont difficilement calculables à des supports différents ; ils doivent donc être simulés.

Une autre façon de procéder est de partir d'un modèle unidimensionnel simple mais calculable, tel qu'une chaîne de Markov à temps continu et à n états. On s'efforcera de déterminer la loi de la proportion de temps passé dans chaque état pendant certains intervalles de temps qui joueront alors le rôle de supports. Ce genre de modèle a l'avantage de ne dépendre que d'un petit nombre de paramètres (durées moyennes de séjour dans chaque état, matrice de transition entre états), tout particulièrement lorsque la chaîne est réversible.

5.5 Valeurs extrêmes en contexte spatial

Soient X et Y deux variables aléatoires positives de même loi F. Supposons F dans le domaine d'attraction de la loi extrême G : il existe alors deux suites (a_n) et (b_n), à termes respectivement positifs et quelconques, telles que

$$\lim_{n \longrightarrow \infty} P^n\{X < a_n x + b_n\} = \lim_{n \longrightarrow \infty} P^n\{Y < a_n x + b_n\} = G(x) \qquad x > 0$$

On dira alors que X et Y sont *asymptotiquement indépendantes* si la loi bivariable du couple (X, Y) est dans le domaine d'attraction d'une loi extrême bivariable factorisée, qui est alors nécessairement de la forme $G \otimes G$. Analytiquement parlant, cela s'écrit

$$\lim_{n \longrightarrow \infty} P^n\{X < a_n x + b_n, Y < a_n y + b_n\} = G(x)G(y) \qquad x, y > 0$$

[22]Une exception notable est celle où les Z_i sont déduites de fonctions aléatoires qui sont les régularisées de mesures aléatoires orthogonales gamma indépendantes.

Dans le cas contraire X et Y seront dits asymptotiquement dépendantes.

Pour caractériser les propriétés de dépendance-indépendance asymptotique du couple (X, Y), il est d'usage de recourir aux deux indices suivants (que nous introduisons sans recours aux copules) :

$$\begin{aligned}
\chi &= \lim_{z \longrightarrow \infty} 2\left[1 - \frac{\ln P\{X < z,\, Y < z\}}{\ln P\{X < z\} P\{Y < z\}}\right] \\
\bar{\chi} &= \lim_{z \longrightarrow \infty} \left[\frac{\ln P\{X > z\} P\{Y > z\}}{\ln P\{X > z\, Y > z\}} - 1\right]
\end{aligned}$$

On a $0 \leq \chi \leq 1$ et $-1 \leq \bar{\chi} \leq 1$. Selon Coles (2004), $\chi = 0$ si et seulement si X et Y sont asymptotiquement indépendantes, tandis que $\bar{\chi} = 1$ si et seulement si X et Y sont asymptotiquement dépendantes. Dans le cas $\chi = 0$, alors $\bar{\chi} < 1$ et cette valeur résume le degré de dépendance extrêmale dans la classe des lois asymptotiquement indépendantes. Dans le cas $\bar{\chi} = 1$, alors $\chi > 0$ et cette valeur résume le degré de dépendance extrêmale dans la classe des lois asymptotiquement dépendantes. De leur coté, Beirlant *et al.* (2004) ne sont pas aussi catégoriques. En particulier, ils n'excluent pas la possibilité d'avoir à la fois $\chi = 0$ et $\bar{\chi} = 1$.
Il a été clairement établi par Resnick (1987) que $\chi = 0$ est une condition nécessaire et suffisante pour l'indépendance asymptotique. Il est possible par ailleurs, d'exhiber des exemples pour lesquels on a $\chi = 0$ et $\bar{\chi} = 1$ simultanément[23].

Cette courte étude bibliographique suggère que $\bar{\chi}$ n'est peut être pas l'indice le plus approprié pour analyser la dépendance asymptotique de couples de variables. D'autres critères devraient sans doute être envisagés.

5.6 Problèmes de voisinage

Lors de la simulation conditionnelle d'un processus ponctuel (processus de Cox), nous nous sommes trouvés confrontés à un problème de mise en oeuvre de l'échantillonneur de Gibbs. Pour en faciliter l'exposition, nous le présentons dans le cadre simple de la simulation d'un vecteur gaussien. L'algorithme formellement envisagé est le suivant :

Algorithme *(Vecteur gaussien – échantillonneur de Gibbs)*
(i) Générer des valeurs gaussiennes indépendantes aux composantes du vecteur ;
(ii) Générer une permutation sur l'ensemble des composantes ;

[23] Soient U, V et W trois variables exponentielles indépendantes de moyenne unité. Posons $X = U + V$ et $Y = U + W$. On a

$$P\{X > z\} = P\{Y > z\} = e^{-z} + \int_0^z e^{-u} e^{-(z-u)} du = e^{-z}(z+1)$$

ainsi que

$$\begin{aligned}
P\{X > z, Y > z\} &= e^{-z} + \int_0^z e^{-u} e^{-2(z-u)} du = e^{-z}(2 - e^{-z}) \\
P\{X < z, Y < z\} &= \int_0^z e^{-u}(1 - e^{-(z-u)})^2 du = e^{-z}(e^z - e^{-z} - 2z)
\end{aligned}$$

d'où l'on déduit aisément $\chi = 0$ et $\bar{\chi} = 1$.

(iii) Considérer tour à tour chaque composante dans l'ordre spécifié par la permutation. Générer une valeur gaussienne pour cette composante connaissant l'ensemble des valeurs prises par les autres ;
(iv) retourner en (ii).

On saura faire tourner cet algorithme effectivement pourvu que le vecteur n'ait pas un nombre trop élevé de composantes. Dans le cas contraire, l'étape *(iii)* doit être modifiée en générant chaque composante non pas en fonction de toutes les autres, mais seulement de celles qui lui sont le plus corrélées.

Pour fixer les idées, supposons que l'on veuille simuler une fonction gaussienne normée de covariance sphérique de portée 10 aux 100×100 noeuds d'une grille carrée de maille unité. Telle quelle, l'étape *(iii)* n'est pas réalisable car une matrice de krigeage d'ordre 10000 ne peut être inversée. En contrepartie, en travaillant sur des voisinages de krigeage de rayon 15, on tombe sur des matrices de krigeage d'ordre 700 qui ne posent pas de problème d'inversion. Au terme de 1500 balayages, aucun test statistique n'invalide les simulations fournies par l'algorithme (cf. Figure 20) :

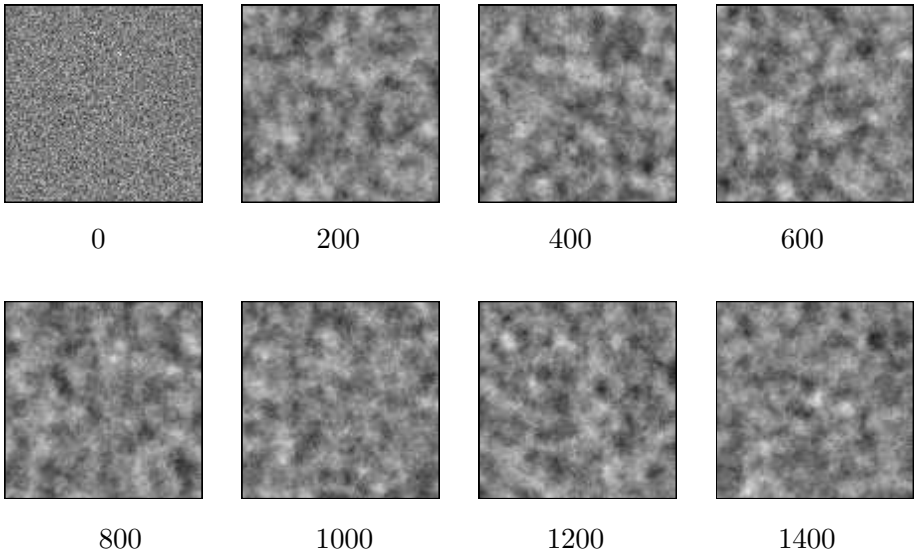

FIG. 20 – Simulation d'une fonction gaussienne par échantillonneur de Gibbs et voisinage réduit (1.5 fois la portée).

L'inversion de matrices d'ordre 700 prenant un temps non négligeable, nous avons cherché à réduire encore le voisinage de krigeage. La Figure 21 montre le résultat obtenu en prenant un rayon de voisinage de 5, soit la moitié de la portée. Au départ, la texture de la simulation obtenue est peu différente de celle obtenue avec un plus grand voisinage. Mais au fil des balayages, sa portée s'allonge tandis qu'elle devient de plus en plus contrastée :

Graduellement les valeurs intermédiaires s'amplifient pour laisser place à des gammes de valeurs sans rapport avec ce que l'on est en droit d'attendre d'une fonction gaussienne normée. Les deux courbes de la Figure 22 témoignent de cet état de fait :

Quelle interprétation donner à ce phénomène[24] ? Faut-il y voir la manifestation d'une

[24]Mentionnons que ce problème avait été observé par D. Allard en 1994 alors qu'il travaillait encore

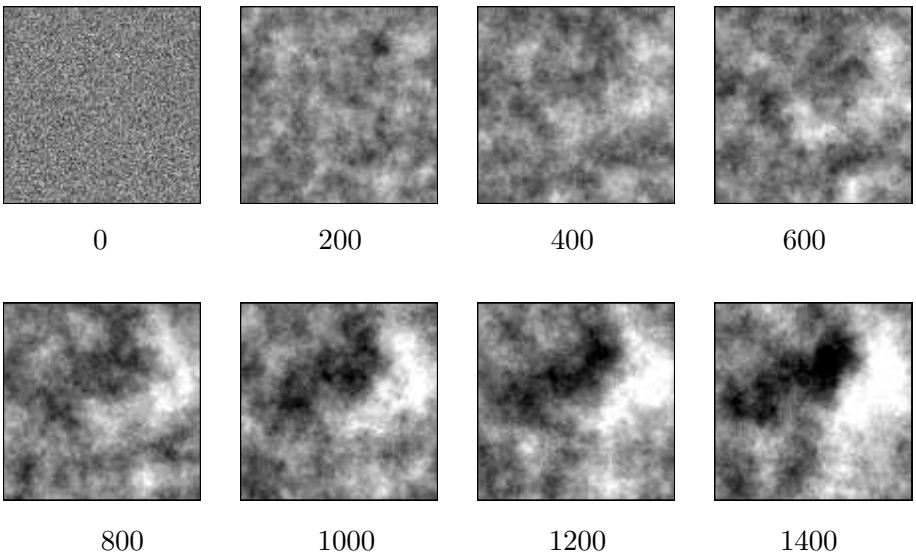

FIG. 21 – Simulation d'une fonction gaussienne par échantillonneur de Gibbs et voisinage très réduit (0.5 fois la portée).

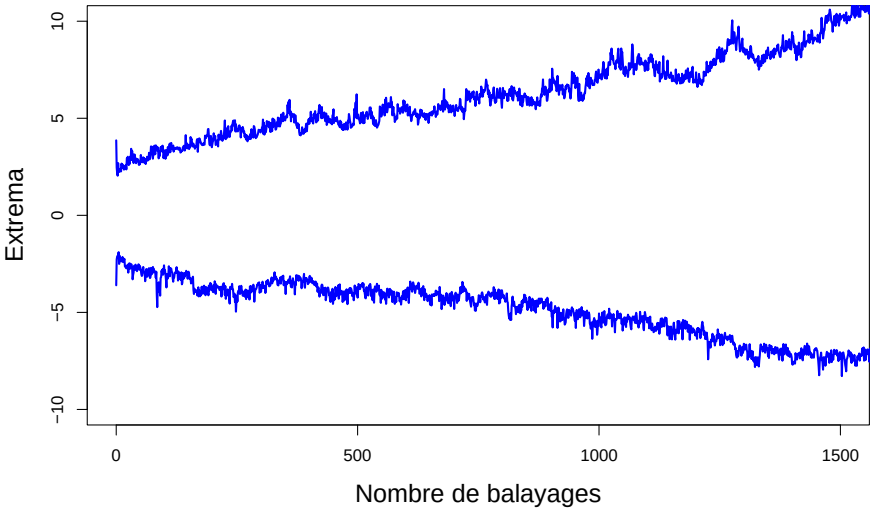

FIG. 22 – Minimum et maximum de la simulation en fonction du nombre de balayages.

transition de phase comme le suggère B. Noetinger de l'IFP ? En tout cas, il est important d'élucider les circonstances au cours desquelles il se produit ; cela permettrait de mettre en oeuvre l'échantillonneur de Gibbs sans encourir le risque d'obtenir des résultats aberrants.

à l'Ecole des Mines sur des applications pétrolières, et a été consigné dans un rapport resté confidentiel ("Quelques remarques sur la mise en oeuvre de l'échantillonneur de Gibbs"). Plus récemment, M. Woillez, actuellement doctorant à l'Ecole des Mines, a été à son tour confronté à ce problème dans le domaine de l'halieutique.

Bibliographie

- Armstrong M., Galli A., Le Loch G., Geffroy F. et Eschard R. (2003) – Plurigaussian simulations in geosciences. Springer, Berlin.
- Beirlant J., Goegebeur Y., Segers J. et Teugels J. (2004) – Statistics of extremes : Theory and applications. Wiley, Chichester.
- Berger M. (1978) – Géométrie : convexes et polytopes, polyèdres réguliers, aires et volumes. Nathan, Paris.
- Beucher S. et Lantuéjoul C. (1979) – "Use of the watersheds in contour detection". International workshop in image processing, real-time edge and motion detection/estimation. Rennes
- Chentsov N.N. (1957) – "Lévy brownian motion for several parameters and generalized white noise". Th. of Prob. and its Appl., Vol. 2, pp. 265-266.
- Chilès J.P. et Delfiner P. (1999) – Geostatistics : modeling spatial uncertainty. Wiley, New York.
- Chilès J.P. et Lantuéjoul C. (2005) – "Prediction by conditional simulation : models and algorithms". Space, structure and randomness (Bilodeau M., Meyer F. et Schmitt M. eds), Lecture notes in statistics N183. Springer, New York, pp. 39-68.
- Coles S. (2004) – An introduction to statistical modeling of extreme values. Springer, Londres.
- Diaconis P. et Freedman D. (1999) – "Iterated random functions". SIAM Review, Vol. 41-1, pp. 45-76.
- Emery X. et Lantuéjoul C. (2006) – "TBSIM : a computer program for conditional simulation of tri-dimensional Gaussian random fields via the turning bands method". Comp. and Geosc., Vol. 32, pp. 1615-1628.
- Feller W. (1971) – An introduction to Probability Theory and its Applications (Vol. 2). Wiley, New York.
- Haas A. et Formery P. (2002) – "Uncertainties in facies proportion estimation 1. Theoretical background. The Dirichlet distribution". Math. Geol., Vol. 34-6, pp. 679-706.
- Haas A., Matheron G. et Serra J. (1967) – "Morphologie mathématique et granulométrie en place". Annale des Mines, Vol. 11, pp. 736-753 et Vol. 12, pp. 767-782.
- Hall P. (1988) – Introduction to the theory of coverage processes. Wiley, New York.
- Jeulin D. (1991) – Modèles morphologiques de structures aléatoires et de changement d'échelle. Thèse de doctorat, université de Caen.
- Jeulin D. et Jeulin P. (1981) – "Synthesis of rough surfaces of random morphological functions". Stereo. Iugosl., Vol. 3-1, pp. 239-246.
- Kendall W.S. et Thönnes E. (1999) – "Perfect simulation in Stochastic Geometry". Pattern Recognition, Vol. 32, pp. 1569-1586.
- Lantuéjoul C. (1982) – "Existence et unicité géodésique". Rapport technique du CMM, 23 pages.
- Lantuéjoul, C. (1988) – "Some Stereological and statistical consequences derived from Cartier's formula", Journal of Microscopy, Vol. 151.3, pp. 265-276.
- Lantuéjoul C. (1991) – "Ergodicity and Integral Range". Journal of Microscopy, Vol. 161-3, pp. 387-404.
- Lantuéjoul C. (1993) – "Substitution Random Functions". Geostatistics Tróia'92 (Soares A. ed.). Kluwer, Dordrect, pp. 37-48.
- Lantuéjoul C. (1994) – "Non conditional simulation of stationary isotropic multigaussian Random Functions". Geostatistical Simulations (Armstrong M. et Dowd P.A. eds.), Kluwer, Dordrecht, pp. 147-177.
- Lantuéjoul, C. (1996) – "A stereological algorithm to generate uniform points over a domain in a high dimensional space". Acta Stereologica, Vol. 15-1, pp. 71-76.
- Lantuéjoul C. (1997) – "Conditional simulation of object-based models". Advances in theory and applications of random sets (Jeulin, D. (ed.). World Scientific Publishing Company, pp. 289-318.

- Lantuéjoul C. (2002) – Geostatistical simulation : models and algorithms. Springer, Berlin.
- Lantuéjoul C. (2005) – "Quelques commentaires sur deux modèles spatiaux max-stables". Note de recherche non publiée.
- Maisonneuve F. et Lantuéjoul C. (1984) – "Geodesic convexity". Acta Stereologica, Vol. 3-2, pp. 169-174.
- Matérn B. (1986) – Spatial Variation. Lecture Notes in Statistics N36. Springer, Berlin.
- Matheron G. (1967) – Éléments pour une théorie des milieux poreux. Masson, Paris.
- Matheron G. (1973) – "The intrinsic random functions and their applications". Adv. Appl. Prob., Vol. 5, pp. 439-468.
- Matheron G. (1975) – Random sets and integral geometry. Wiley, New York.
- Matheron G. (1976) – "A simple substitute for conditional expectation : the disjunctive kriging". Advanced geostatistics in the mining industry (Guarascio M., David M. et Huijbregts C. eds). Reidel, Dordrecht, pp. 221-236.
- Matheron G. (1984) – "The selectivity of the distributions and the second principle of geostatistics". Geostatistics for natural ressource characterization (Verly G. et al. eds.). Reidel, Dordrecht, pp. 421-433.
- Matheron G. (1989) – Estimating and choosing. An essay on probability in practice. Springer, Berlin.
- Miles R.E. (1969) – "Poisson flats in euclidean spaces". Adv. Appl. Prob., Vol. 1, pp. 1-43.
- Miles R.E. (1974) – "A synopsis of Poisson flats in Euclidean spaces". Stochastic Geometry (Harding E.F. et Kendall D.G. ed.). Wiley, London, pp. 202-227.
- Mitchell J.S.B. (2004) – "Shortest paths and networks". Handbook of discrete and computational geometry (Goodman J.E. et O'Rourke J. eds.). Chapman Hall, Boca Raton, pp. 607-641.
- Molchanov I.S. (1997) – Statistics of the boolean model for practitioners and mathematicians. Wiley, Chichester.
- Préteux F. et Schmitt M. (1988) – "Boolean texture analysis and synthesis". Image analysis and mathematical morphology, theoretical advances (Serra J. ed.). Academic Press, London, pp. 377-400.
- Propp J.G. et Wilson,D.B (1996) – "Exact sampling with coupled Markov chains and applications to statistical mechanics". Random structures and Algorithms, Vol. 9, pp. 223-252.
- Resnick S.I. (1987) – Extreme values, regular variation, and point processes. Springer, New York.
- Schlather M. (2002) – "Models for stationary max-stable random fields". Extremes, Vol. 5-1, pp. 33-44.
- Schmitt M. (1989) – Des algorithmes morphologiques à l'intelligence artificielle. Thèse, Ecole des Mines de Paris (disponible sur le site http ://pastel.paritech.org).
- Serra J. (1968) – "Fonctions aléatoires de Dilution". Note interne C-12. Centre de morphologie mathématique, Fontainebleau.
- Serra J. (1982) – Image analysis and mathematical morphology. Academic Press, Londres.
- Shinozuka M. et Jan C.M. (1972) – "Digital simulation of random processes and its applications". J. of Sound and Vib., Vol. 25-1, pp. 111-128.
- Smith R.L. (1984) – "Efficient Monte Carlo procedures for generating points uniformly distributed over bounded regions". Operations Research, Vol. 32, pp. 1296-1308.
- Smith R.L. (1990) – "Max-stable processes and spatial extremes". Manuscrit non publié.
- Smith K.T. et Solmon D.C. (1975) – "Lower dimensional integrability of L^2 functions". Journal of Math. Anal. and Appl., Vol. 51, pp. 539-549.
- Stoyan D., Kendall W.S. et Mecke J. (1995) – Stochastic geometry and its applications. Wiley, Chichester.
- Tournois F. (1998) – "Simulation parfaite de schéma booléen". Mémoire de DEA Algorithmique de Jussieu.
- Turčin V.F. (1971) – "On the computation of multidimensional integrals by the Monte Carlo method". Theory of Prob. and Appl., Vol. 16, pp. 720-4.

– Yaglom A.M. (1987) – Correlation Theory of Stationary and related Random Functions. Springer, New York.