

CONDITIONNEMENT DU MODÈLE GAUSSIEN PAR DES INÉGALITÉS OU DES RANDOMISÉES

THÈSE

présentée à

l'Ecole Nationale Supérieure des Mines de Paris

par

Xavier FREULON

pour obtenir le titre de
Docteur en Géostatistique

Soutenue le 1^{er} Juin 1992
devant le jury composé de :

M. G. MATHERON	<i>Président</i>
M. J.-P. CHILÈS	<i>Rapporteur</i>
M. P. FORMERY	<i>Rapporteur</i>
M ^{lle} C. de FOUQUET	<i>Examineur</i>
M. J. JACOD	<i>Examineur</i>
M. C. LANTUÉJOUL	<i>Examineur</i>

Remerciements

Je remercie les membres du jury qui se sont intéressés à mon travail et ont accepté de le juger.

Que toutes les personnes qui par leurs conseils, suggestions ou encouragements ont contribué à l'élaboration de ce mémoire soient ici remerciées.

Ce travail a été effectué au Centre de Géostatistique, laboratoire de l'École Nationale Supérieure des Mines de Paris. Les données expérimentales ont été fournies par la société Total Compagnie Minière et l'I.R.S.I.D.

Résumé

Pour simuler conditionnellement une variable, il est nécessaire d'explicitier le lien entre celle-ci et les observations. Deux types de relation ont été étudiés : une relation fonctionnelle, la variable observée est la transformée par une fonction déterministe du vecteur à simuler, et une randomisation, la variable observée est une variable aléatoire dont le paramètre est la variable à simuler. Le modèle de la variable à simuler étant le modèle gaussien anamorphosé, la simulation de la loi conditionnelle peut se ramener à la simulation d'un vecteur multigaussien dont chaque composante doit vérifier une ou plusieurs inégalités. La simulation d'un tel vecteur aléatoire s'effectue à l'aide d'une chaîne de Markov ergodique dont la loi stationnaire est la loi conditionnelle du vecteur gaussien. Ces techniques ont été utilisées pour modéliser la corégionalisation entre une teneur et un faciès lithologique d'un gisement d'uranium ainsi que pour filtrer la mesure entachée d'un bruit poissonien d'une image micro-sonde. Dans ce dernier cas, l'espérance conditionnelle ne pouvant être calculée directement, on a recours à des techniques de Monte-Carlo : une approximation de l'estimateur est construite sur des simulations de la loi conditionnée par les observations.

Abstract

A conditional simulation is based on an explicit model that links the measured variable to the variable to be simulated. Two different links have been studied :

- a deterministic link, i.e. the observed variable can be expressed as a deterministic function of the variable to be simulated,
- a probabilistic link, i.e. the distribution of the observed variable depends on a parameter which is the variable to be simulated.

If the variable to be simulated derives from a gaussian model, then its conditional distribution can be simulated in the same way as a multigaussian vector whose coordinates satisfy inequalities. These inequalities can be validated by sampling the stationary distribution of an ergodic Markov chain. This technique has been used to model the coregionalization between the grade and a geological indicator in an uranium orebody, and also to filter an image blurred by a Poisson noise. In this latter case, since a direct computation of the conditional expectation was impossible, a Monte-Carlo method has been used to build an estimator starting from several simulations of the conditional distribution.

Introduction

Les méthodes de simulations conditionnelles gaussiennes ont été développées il y a une vingtaine d'années [25] et sont maintenant d'usage courant en géostatistique. On oppose généralement simulations et estimation. Un estimateur construit sur un jeu de données expérimentales fournit en chaque point une valeur aussi proche que possible de la variable vraie inconnue. Une simulation conditionnelle reproduit seulement les principales caractéristiques de dispersion de la variable étudiée — on impose généralement que la loi bivariable et par conséquent l'histogramme et la covariance soient reproduits — et repasse par les données expérimentales. Cependant les simulations ne sont pas calculées pour elles-mêmes mais sont utilisées bien souvent pour estimer une fonctionnelle de la variable : son calcul sur les simulations fournit un échantillonnage de sa dispersion. On peut alors en déduire un estimateur. Un exemple classique est l'estimation de la surface d'une île à partir de mesures bathymétriques [7] : 100 bathymétries sont simulées conditionnellement et sur chaque simulation la surface de l'île — les points où la profondeur est négative — est calculée. L'histogramme des terres émergées est calculé. En sont déduits des estimateurs tels que la moyenne, l'écart-type et quelques quantiles. Les simulations conditionnelles apparaissent alors comme des intermédiaires de calcul pour construire un estimateur. Une telle démarche est lourde, elle est employée lorsque la construction directe d'un estimateur n'est pas possible — lorsqu'on ne dispose pas d'expression explicite de l'estimateur ou bien lorsque les calculs sont impraticables. De plus il convient de préciser le statut des estimateurs ainsi construits.

Pour cela on rappelle de manière succincte les techniques d'estimations utilisées en géostatistique. Il faut tout d'abord formaliser l'objet d'étude :

- Le phénomène étudié sur un domaine borné de l'espace, D , est représenté par une fonction y — scalaire ou vectorielle — définie sur D .
- On dispose de mesures liées au phénomène :

$$z = (z_1, z_2, \dots, z_n)$$

On recherche une estimation d'une fonctionnelle de la variable y , $\psi(y)$, qui peut être, par exemple, la valeur de y en un point de D ou une fonction de cette valeur, soit $y(u)$ ou $f \circ y(u)$, ou bien la valeur moyenne de $f \circ y$ sur un sous-ensemble V du domaine :

$$\frac{1}{|V|} \int_V f \circ y(u) du$$

L'estimateur est une fonction des données expérimentales, $\hat{\psi}(z)$. Le cadre de travail est fixé par l'hypothèse constitutive de la géostatistique [37]:

La variable étudiée y ¹ et les mesures sont interprétées comme une réalisation respectivement de la fonction aléatoire Y et de la variable aléatoire vectorielle Z (Y et Z ne sont pas indépendantes).

Dans ce cadre, un estimateur est défini comme la fonction des données expérimentales minimisant l'écart quadratique moyen :

$$\sigma_E^2 = \mathbf{E}\{(\psi(Y) - \hat{\psi}(Z))^2\}$$

La classe des fonctions des données sur laquelle la minimisation s'effectue définit le type d'estimation. Ainsi aux estimateurs couramment utilisés en géostatistique, le krigeage simple [33], le krigeage disjonctif [34, 36] et l'espérance conditionnelle, correspondent respectivement les classes de fonctions suivantes :

$$\begin{aligned} \mathcal{C}_{KS} &= \{\Sigma \lambda^i z_i\} \\ \mathcal{C}_{KD} &= \{\Sigma f_i(z_i)\} \\ \mathcal{C}_{EC} &= \{f(z_1, \dots, z_n)\} \end{aligned}$$

L'estimateur est alors la projection orthogonale de la variable à estimer, le modèle $\psi(Y)$, sur le sous-espace linéaire engendré par les données expérimentales et la classe des fonctions considérées. Mais comme $\mathcal{C}_{KS} \subset \mathcal{C}_{KD} \subset \mathcal{C}_{EC}$, il en est de même pour les espaces linéaires sur lesquels s'effectue la projection orthogonale. Plus les estimateurs sont précis — c'est-à-dire, plus la variance d'estimation est petite —, plus ils exigent en contrepartie de préciser le modèle de fonction aléatoire utilisé. Pour le krigeage simple seules sont nécessaires les covariances entre les données et la variable à estimer, pour le krigeage disjonctif les lois bivariées et pour l'espérance conditionnelle la loi du vecteur composé par les données expérimentales et la variable à estimer.

Dans la pratique, on retient le degré de spécification qui soit le plus faible possible et qui soit compatible avec la résolution effective du problème posé. En effet à un type d'estimateur sont associées les caractéristiques d'un modèle de fonction aléatoire. Cependant ces caractéristiques ne déterminent pas en général une fonction aléatoire unique, mais un ensemble de fonctions aléatoires — par exemple, la classe des fonctions aléatoires de covariance exponentielle ou la classe des fonctions aléatoires de lois bivariées gaussiennes. L'estimateur est optimal pour toutes les fonctions aléatoires appartenant à cet ensemble. Ainsi limiter le degré de spécification de la fonction aléatoire assure une certaine robustesse à l'estimateur. D'autre part, tant que les caractéristiques du modèle utilisées peuvent être estimées ou testées à partir des données expérimentales — bien souvent tel est le cas des covariances voire des lois bivariées —, l'estimation peut être considérée comme une procédure objective.

Cependant, en fonction du problème posé, il n'est pas toujours souhaitable de se limiter aux estimateurs linéaires. Par exemple, l'estimation de f o y à partir de z peut toujours être

¹Pour bien distinguer la réalité du modèle, y est appelée une variable régionalisée.

faite par krigeage linéaire et dans ce cas seules les covariances doivent être inférées. Pour assurer que le modèle des covariances simples et croisées soit un modèle autorisé [39], on doit bien souvent passer par la modélisation des lois bivariées. Ayant fait ces hypothèses, il faut utiliser l'estimateur qui les exploite au mieux, le krigeage disjonctif. Ce problème de la cohérence des modèles se rencontre par exemple lors de l'estimation d'une indicatrice : expérimentalement, on peut calculer les variogrammes simples et croisés sur les données, $(y_\alpha, \mathbb{1}_{y_\alpha \geq y_c})_{\alpha \in A}$. Mais pour l'ajustement de la covariance simple de l'indicatrice et de la covariance croisée entre la variable et l'indicatrice, on ne peut utiliser un modèle quelconque. Pour apporter une réponse cohérente à un problème d'estimation complexe, on peut donc être amené à utiliser un modèle qui n'a pu être testé exhaustivement sur les données expérimentales. Mais dans ce cas un estimateur ne s'appuyant pas sur un modèle cohérent et explicite ne garantirait à la procédure qu'une illusoire objectivité. On rencontre là les limites du jeu de données qui ne peut pas tout dire.

A cette présentation des estimateurs utilisés en géostatistique, on peut ajouter les estimateurs construits sur les simulations conditionnelles. Il faut tout d'abord se donner une définition formelle d'une simulation — la définition proposée au début de cette introduction est une définition des objectifs d'une simulation, reproduire la dispersion de la variable étudiée. La variable étudiée y est interprétée comme une réalisation d'une fonction aléatoire $Y(u)$. Si on se donne une famille finie de points de l'espace, $S = \{u_1, \dots, u_p\}$ — on considère généralement une grille régulière —, la restriction de la fonction aléatoire à ces points fournit le vecteur aléatoire $Y = (Y_\alpha)_{\alpha \in S}$ défini par sa loi μ^Y , appelée la loi *a priori*. Une simulation non conditionnelle est une réalisation d'un vecteur aléatoire de loi μ^Y . Pour conditionner une simulation par le vecteur des mesures z , on se donne un modèle du vecteur aléatoire (Z, Y) défini par sa loi μ^{ZY} tel que la loi marginale de Y soit la loi *a priori* de la simulation. Une simulation conditionnée par z est une réalisation de la loi conditionnelle de Y à z fixé, notée μ_z^Y . A l'aide des simulations conditionnelles $(\tilde{y}^{(i)})_{1 \leq i \leq \nu}$, on construit l'estimateur empirique de $\psi(y)$ par

$$\hat{\psi}^{(\nu)}(z) = \frac{1}{\nu} \sum_{i=1}^{\nu} \psi(\tilde{y}^{(i)})$$

Si on a la convergence,

$$\lim_{\nu \rightarrow \infty} \hat{\psi}^{(\nu)}(z) = E\{\psi(Y)|z\}$$

cet estimateur est une approximation de l'espérance conditionnelle calculée avec la loi μ^{ZY} . Comme pour l'espérance conditionnelle, les estimateurs construits à l'aide des simulations conditionnelles s'appuient sur des hypothèses qui ne peuvent pas toutes être contrôlées sur les données expérimentales. Mais procédant de la sorte, on peut intégrer dans un cadre cohérent des mesures de natures diverses : enrichissant le conditionnement on limite l'arbitraire du modèle. De plus pour choisir un modèle, on peut s'appuyer sur la physique du phénomène ou de sa mesure.

L'unique modèle utilisé tout au long de ce travail est le modèle gaussien anamorphosé. Ce modèle, très utilisé en pratique, a été l'objet de très nombreux travaux. Ici on présente

le conditionnement dans quelques nouveaux cas. En effet la méthode de simulation conditionnelle *classique* a des conditions d'utilisation très strictes qui sont parfois outrepassées au prix d'approximations peu justifiées. En particulier, la fonction d'anamorphose doit être bijective ce qui exclut la représentation par ces méthodes de distributions présentant des atomes (variables discrètes, effet zéro).

Dans le chapitre 1, on rappelle tout d'abord le modèle gaussien ainsi que la mise en œuvre des simulations conditionnelles *classiques*. Conditionner une simulation suppose un lien explicite entre la mesure et la variable simulée. Deux modèles sont présentés : la variable observée est soit la transformée par une fonction de carré sommable, soit la randomisée d'une variable gaussienne liée à la gaussienne à simuler. De cette modélisation, on déduit alors la loi conditionnelle à simuler. Pour chacun de ces modèles, un exemple d'application est présenté :

- La corégionalisation entre une teneur en uranium et l'indicatrice d'un faciès géologique est modélisée par des fonctions d'anamorphose d'une corégionalisation de facteurs gaussiens (chapitre 2).
- Les simulations conditionnelles sont utilisées pour filtrer la carte chimique d'un élément obtenue par une micro-sonde électronique (chapitre 4). Le bruit à éliminer est poissonien, c'est-à-dire que la variable observée est une variable de Poisson dont le paramètre est égal à la concentration à estimer.

Quelques algorithmes permettant la mise en œuvre des simulations conditionnelles dans le cas où la condition s'exprime sous la forme d'inégalités sont étudiés au chapitre 3. Un des algorithmes présentés est également utilisé, après quelques modifications mineures, pour conditionner une simulation par des randomisées.

Chapitre 1

Des exemples de conditionnement

Ce chapitre présente les différents types de conditionnement qui seront mis en œuvre par la suite. Le conditionnement permet de caler les simulations aux observations. Cependant la variable mesurée n'est pas toujours la variable à simuler. Aussi faut-il expliciter la contrainte sur la variable à simuler définie par l'observation. On recense les méthodes disponibles pour intégrer à un modèle gaussien anamorphosé des mesures de nature diverse. Cette modélisation permet de conditionner par des données expérimentales la simulation d'un tel modèle. Quatre types de relations ont été envisagés :

- les corrélations expérimentales entre plusieurs variables,
- des supports de mesure variables,
- des mesures entachées de bruit,
- les mesures sont exprimées par des inégalités.

Ces types d'observations sont modélisés par

- une transformation ou
- une randomisation

de variables gaussiennes. Ces modélisations déterminent la loi conditionnelle à simuler.

1.1 Le modèle gaussien anamorphosé

Dans le modèle gaussien anamorphosé, la variable régionalisée y , étudiée sur le domaine D (sous-ensemble de l'espace ou du plan), est interprétée comme une réalisation d'une fonction aléatoire stationnaire, $Y(\cdot)$, définie par

$$\forall u \in D, Y(u) = \phi(X(u))$$

où $X(\cdot)$ est une fonction aléatoire stationnaire multigaussienne centrée et normée et ϕ une fonction d'anamorphose.

1.1.1 Fonction aléatoire multigaussienne

Une fonction aléatoire $X(\cdot)$ définie sur le champ D est multigaussienne si et seulement si quel que soit le domaine fini ¹ $U \subset D$, le vecteur $(Y(u))_{u \in U}$ est un vecteur multigaussien.

Un vecteur aléatoire de dimension n , $X = (X_i)_{1 \leq i \leq n}$, est un vecteur multigaussien si et seulement si il vérifie une des conditions suivantes :

1. Toute combinaison linéaire des composantes du vecteur X est une variable aléatoire gaussienne.
2. La densité du vecteur X est

$$g(x) = \frac{1}{(2\pi)^{\frac{n}{2}}} \frac{1}{|\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x - m)' \Sigma^{-1}(x - m)\right)$$

où m est un vecteur réel et Σ la matrice carrée définie positive des covariances tels que :

$$\begin{aligned} \mathbf{E}\{X\} &= m \\ \mathbf{E}\{(X - m)(X - m)'\} &= \Sigma \end{aligned}$$

Dans le modèle gaussien anamorphosé, la fonction aléatoire multigaussienne et stationnaire $X(\cdot)$ est donc entièrement spécifiée par sa covariance

$$\text{Cov}(X_u, X_{u+h}) = \mathbf{E}\{X_u X_{u+h}\} = \rho(h) .$$

1.1.2 Anamorphose

Dans le cas général, l'histogramme de la variable régionalisée étudiée n'est pas gaussien. Par exemple, la variable est positive ou bien l'histogramme est fortement dissymétrique. Une fonction d'anamorphose, ϕ , permet de modéliser une variable d'histogramme quelconque comme la transformée d'une variable gaussienne. C'est une fonction non décroissante de $L^2(\mathbb{R}, g)$ telle que

$$F(\phi(x)) \leq G(x) \leq F(\phi(x)^+)$$

où F et G sont les fonctions de répartition respectivement de la loi marginale de Y et d'une gaussienne centrée normée [29].

¹Un domaine fini est une réunion finie de points.

1.2 Conditionnement

Une simulation conditionnelle de la fonction aléatoire $Y_{(\cdot)}$ sur un champ fini S est une réalisation de la loi du vecteur aléatoire $(Y_{(u)})_{u \in S}$ conditionnée par le vecteur des observations $Z = z$. Pour cela, on simule conditionnellement la variable gaussienne, puis on calcule la valeur de la variable: $u \in S$, $Y_u = \phi(X_u)$. On notera désormais X_S le vecteur $(X_{(u)})_{u \in S}$.

Il faut préciser tout d'abord la relation entre les observations et la variable à simuler. Deux types de conditionnement seront étudiés ici :

- 1 - le vecteur des observations, Z , est le transformé d'un vecteur multigaussien centré normé X_M , c'est-à-dire il existe une fonction ψ de \mathbb{R}^n dans \mathbb{R}^n telle que $Z = \psi(X_M)$,
- 2 - le vecteur des observations est le randomisé de X_M , c'est-à-dire il existe une famille de variables aléatoires, $(\Theta_x)_{x \in \mathbb{R}^n}$, telle que Z est la variable Θ de paramètre aléatoire X_M .

On suppose de plus que le vecteur (X_M, X_S) est multigaussien. Par la suite, les composantes de X_M seront appelées les gaussiennes associées aux mesures.

Il faut donc simuler la loi du vecteur X_S conditionnée par la valeur du vecteur Z . Cette loi sera notée $\mu_z^{X_S}$ —l'exposant rappelle la variable aléatoire dont on considère la loi et l'indice désigne la valeur de la variable conditionnante. L'introduction de la variable X_M permet de décomposer la simulation en deux étapes :

- (i) *validation des contraintes imposées par les observations, c'est-à-dire la simulation de la loi de X_M conditionnée par les observations, $d\mu_z^{X_M}(x)$,*
- (ii) *simulation conditionnelle de la loi du vecteur X_S conditionnée par la valeur prise par X_M , $g(x_S|x_M)dx_S$.*

Démonstration : On considère le triplet (Z, X_M, X_S) dont la loi est :

$$d\mu_{x_M, x_S}^Z(z) g(x_M, x_S) dx_M dx_S$$

Dans les modèles retenus, les observations ne dépendent que des valeurs gaussiennes qui leur sont associées. Les lois conditionnelles de Z ne dépendent alors que des valeurs prises par X_M . En particulier, on a

$$\mu_{x_M, x_S}^Z(z) = \mu_{x_M}^Z(z)$$

D'autre part, le couple (X_M, X_S) est bigaussien, sa loi est donc

$$g(x_M, x_S) dx_M dx_S = g(x_M) dx_M g(x_S|x_M) dx_S$$

La loi du triplet se réécrit donc sous la forme :

$$d\mu_{x_M}^Z(z) g(x_M) dx_M g(x_S|x_M) dx_S$$

Mais $d\mu_{x_M}^Z(z)g(x_M)dx_M$ est la loi du couple (Z, X_M) qui s'écrit aussi $d\mu^Z(z)d\mu_z^{X_M}(x_M)$. La loi du triplet se réécrit donc

$$d\mu^Z(z) d\mu_z^{X_M}(x_M) g(x_S|x_M) dx_S$$

et $d\mu_z^{X_M}(x_M) g(x_S|x_M) dx_S$ est l'expression de la loi conditionnelle de (X_M, X_S) lorsque la valeur de Z est fixée. De cette factorisation on déduit le procédé de simulation en deux étapes. Le vecteur ainsi simulé suit la loi

$$d\mu_z^{X_S}(x_S) = dx_S \int d\mu_z^{X_M}(x_M) g(x_S|x_M)$$

■

L'introduction de la variable auxiliaire X_M serait superflue si on pouvait directement calculer et simuler la loi $\mu_z^{X_S}(x_S)$. Elle est nécessaire ou simplifie grandement la méthode de simulation dans tous les exemples qui seront traités par la suite. Dans les paragraphes qui suivent, on présentera tout d'abord la mise en œuvre de l'étape (ii) — la simulation d'un vecteur multigaussien dont la valeur de plusieurs de ses composantes est fixée s'obtient par un très classique krigeage conditionnant — puis on ne parlera plus que de la validation des contraintes (i).

1.3 Simulations conditionnelles classiques

1.3.1 Rappels : le krigeage conditionnant

La simulation d'un vecteur multigaussien centré, $(X_i)_{1 \leq i \leq p}$, dont la valeur des n premières coordonnées est fixée s'appuie sur l'équivalence dans le modèle gaussien entre la non-corrélation et l'indépendance. En effet toute composante du vecteur dont la valeur n'est pas fixée peut se décomposer en une somme de deux termes indépendants,

$$n < i \leq p, \quad X_i = X_i^{KS} + (X_i - X_i^{KS})$$

où X_i^{KS} , le krigeage simple de X_i par les gaussiennes fixées, est défini par

$$X_i^{KS} = \sum_{1 \leq j \leq n} \lambda_i^j X_j$$

les $(\lambda_i^j)_{1 \leq j \leq n}$ étant les solutions du système,

$$1 \leq k \leq n, \quad \sum_{1 \leq j \leq n} \mathbf{E}\{X_k X_j\} \lambda_i^j = \mathbf{E}\{X_k X_i\}$$

et $X_i - X_i^{KS}$ l'erreur de krigeage. Cette erreur étant indépendante des $(X_j)_{1 \leq j \leq n}$, il suffit de simuler non conditionnellement un vecteur multigaussien centré, $(U_i)_{n < i \leq p}$, dont la matrice des covariances est $\Sigma_{22} - \Sigma'_{12} \Sigma_{11}^{-1} \Sigma_{12}$ avec

$$\Sigma_{11} = [\mathbf{E}\{X_i X_j\}]_{1 \leq i, j \leq n} \quad \Sigma_{12} = [\mathbf{E}\{X_i X_j\}]_{1 \leq i \leq n, n < j \leq p} \quad \Sigma_{22} = [\mathbf{E}\{X_i X_j\}]_{n < i, j \leq p}$$

La simulation conditionnelle du vecteur X est alors

$$\tilde{X} = (\tilde{X}_i)_{1 \leq i \leq p} \quad \text{où} \quad \tilde{X}_i = \begin{cases} X_i & 1 \leq i \leq n \\ X_i^{KS} + U_i & n < i \leq p \end{cases}$$

On vérifie aisément que le vecteur aléatoire ainsi construit vérifie les conditions :

1. $\text{Cov}(\tilde{X}_i, \tilde{X}_j) = \text{Cov}(X_i, X_j)$,
2. $1 \leq i \leq n, \tilde{X}_i = X_i$,
3. la loi de \tilde{X} est multigaussienne.

Le krigeage conditionnant[25] permet de simuler ce résidu : on simule non-conditionnellement² le vecteur aléatoire $T = (T_i)_{1 \leq i \leq p}$ dont la matrice des covariances est : $[\mathbf{E}\{X_i X_j\}]_{1 \leq i, j \leq p}$. Puis on calcule l'erreur de krigeage,

$$1 \leq i \leq p, \quad T_i - \sum_{1 \leq j \leq n} \lambda_i^j T_j$$

La simulation conditionnelle de la gaussienne est alors

$$1 \leq i \leq p, \quad \tilde{X}_i = \sum_{1 \leq j \leq n} \lambda_i^j X_j + (T_i - \sum_{1 \leq i \leq n} \lambda_i^j T_j) = T_i + \sum_{1 \leq j \leq n} \lambda_i^j (X_j - T_j)$$

²On dispose de très nombreuses méthodes pour simuler une fonction aléatoire stationnaire gaussienne (voir [6] ou annexes).

Si la fonction d'anamorphose est bijective, cette technique permet de simuler le vecteur $Y = (\phi(X_i))_{1 \leq i \leq p}$ sous la condition $1 \leq j \leq n$, $Y_j = y_j$. La simulation conditionnelle *classique* se décompose alors en trois étapes [25]:

- (i) calcul des gaussiennes "expérimentales", $1 \leq j \leq n$, $X_j = \phi^{-1}(y_j)$
- (ii) simulation conditionnelle des gaussiennes, $(\tilde{X}_i)_{1 \leq i \leq p}$
- (iii) transformation des gaussiennes, $1 \leq i \leq p$, $\tilde{Y}_i = \phi(\tilde{X}_i)$

1.3.2 Accélération des calculs par krigeage dual

Le krigeage conditionnant doit s'effectuer en voisinage unique³. Il est alors possible de réduire très sensiblement le temps de calcul en utilisant la représentation duale du krigeage:

$$n < i \leq p, \quad \tilde{X}_i = T_i + \sum_{1 \leq j \leq n} \mathbf{E}\{X_j X_i\} \tilde{b}^j$$

Les coefficients $(\tilde{b}^j)_{1 \leq j \leq n}$ sont la solution du système,

$$1 \leq k \leq n, \quad \sum_{1 \leq j \leq n} \mathbf{E}\{X_k X_j\} \tilde{b}^j = X_k - T_k$$

Démonstration: On utilise les notations matricielles suivantes:

- Σ est la matrice de dimension $n \times n$ des covariances entre les composantes *expérimentales*, $\Sigma = [\mathbf{E}\{X_i X_j\}]_{1 \leq i, j \leq n}$,
- Δ est le vecteur de dimension n des différences entre les composantes *expérimentales* mesurées et les composantes *expérimentales* simulées, $\Delta = (X_i - T_i)_{1 \leq i \leq n}$,
- $\sigma(i)$ est le vecteur de dimension n des covariances entre les points expérimentaux et la composante i ($n < i \leq p$), $\sigma(i) = (\mathbf{E}\{X_i X_j\})_{1 \leq j \leq n}$.

Utilisant l'associativité du produit matriciel, l'erreur de krigeage simulée peut donc se récrire sous la forme souhaitée:

$$\sum_{1 \leq j \leq n} \lambda_i^j (X_j - T_j) = (\sigma(i)' \Sigma^{-1}) \Delta = \sigma(i)' (\Sigma^{-1} \Delta) = \sum_{1 \leq j \leq n} \mathbf{E}\{X_j X_i\} \tilde{b}^j$$

■

³Lorsque le nombre des données conditionnantes est grand on procède généralement à l'aide d'un voisinage glissant.

Evaluons l'économie induite par cette représentation du krigeage conditionnant :

- Pour le krigeage direct, on calcule tout d'abord Σ^{-1} et Δ , puis pour chaque composante ($n < i \leq p$) $\sigma(i)$ et les produits matriciels $(\Sigma^{-1}\sigma(i))' \Delta$.
- Pour le krigeage dual, on calcule Σ^{-1} , Δ et le produit matriciel $\Sigma^{-1} \Delta$, puis pour chaque composante ($n < i \leq p$) $\sigma(i)$ et le produit matriciel $\sigma(i)' (\Sigma^{-1} \Delta)$.

Procédant par krigeage dual, on économise $p - n$ produits d'une matrice $n \times n$ avec un vecteur ; en contrepartie on ne dispose pas de la variance de krigeage⁴ qui dans le cas d'une simulation présente peu d'intérêt sinon pour évaluer l'intensité du conditionnement⁵.

1.4 La loi conditionnelle

L'objet de ce paragraphe est le calcul pour quelques exemples de la loi de X_M conditionnée par les observations z . Ceci est nécessaire car le modèle liant la mesure à la gaussienne ne précise que la loi conditionnelle de Z à x_M fixé, $\mu_{x_M}^Z$, alors que pour procéder à une simulation conditionnelle, l'expression de la loi conditionnelle de X_M à z fixé est nécessaire. On sait par ailleurs que X_M est un vecteur multigaussien. Sa loi marginale est donc $d\mu^{X_M}(x_M) = g(x_M)dx_M$. On en déduit la loi du couple (X_M, Z) :

$$d\mu_{x_M}^Z(z)g(x_M)dx_M$$

Disposant de la loi du couple (Z, X_M) , on peut calculer la loi de X_M à z fixé en utilisant les formules des probabilités totales et composées.

1.4.1 Formules de probabilités totales et composées

Ces formules bien connues permettent d'exprimer la probabilité d'événements en fonction de probabilités conditionnelles. On distingue le cas discret, élémentaire, du cas continu.

Le cas discret

Soient (Ω, \mathcal{B}, P) un espace probabilisé et $(A_i)_{i \in I}$ une partition dénombrable de Ω par des éléments de \mathcal{B} tels

- $\bigcup_{i \in I} A_i = \Omega$,
- $i \neq j, A_i \cap A_j = \emptyset$.

⁴La variance de krigeage est $\sigma_{(i)}^2 = \text{Var}(X_i - X_i^{KS})$ et vaut $\sigma_{(i)}^2 = \text{Var}(X_i) - \sum_{1 \leq j \leq n} \lambda_i^j \mathbf{E}\{X_i X_j\}$.

⁵A une variance de krigeage conditionnant proche de 1 correspond un conditionnement faible et à une variance quasi-nulle un très fort conditionnement.

La formule dite des *probabilités totales et composées* permet de calculer la probabilité de tout élément de \mathcal{B} en fonction des probabilités conditionnelles $(P(B|A_i))_{i \in I}$ et les probabilités des $(A_i)_{i \in I}$:

$$P(B) = \sum_{i \in I} P(B|A_i)P(A_i) \quad (1.1)$$

Cette formule est utilisée pour exprimer les probabilités conditionnelles des A_i connaissant B en fonction des probabilités conditionnelles de B connaissant A_i (formule de Bayes):

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(B|A_i) P(A_i)}{P(B)}$$

d'où d'après la formule 1.1

$$P(A_i|B) = \frac{P(B|A_i) P(A_i)}{\sum_{j \in I} P(B|A_j) P(A_j)}$$

Le cas continu

Soit (X, Y) un couple de variables aléatoires vectorielles à valeurs respectivement dans \mathbb{R}^n et \mathbb{R}^p . $\mathcal{B}_{\mathbb{R}^n}$ (resp. $\mathcal{B}_{\mathbb{R}^p}$) désigne la tribu des boréliens de \mathbb{R}^n (resp. \mathbb{R}^p).

Soit une application μ_x^Y de $\mathbb{R}^n \times \mathcal{B}_{\mathbb{R}^p}$ dans \mathbb{R}^+ telle que

- à x fixé, $x \in \mathbb{R}^n$, $\mu_x^X(\cdot)$ soit une probabilité sur \mathbb{R}^p ,
- à B fixé, $B \in \mathcal{B}_{\mathbb{R}^p}$, $\mu_x^X(B)$ soit une application mesurable.

μ_x^Y est une loi conditionnelle de Y à x fixé si et seulement si pour toute fonction ϕ mesurable et positive:

$$\mathbf{E}\{\phi(Y)|x\} = \int \phi(y) d\mu_x^Y(y).$$

Une telle loi conditionnelle existe[42] et est caractérisée par le théorème suivant:

Théorème: μ_x^Y est une probabilité conditionnelle de Y sachant x si et seulement si pour tous boréliens A et B respectivement de \mathbb{R}^n et \mathbb{R}^p

$$P\{X \in A, Y \in B\} = \int_A d\mu^X(x) \mu_x^Y(B) \quad (1.2)$$

où μ^X (resp. μ^Y) désigne la loi marginale de X (resp. Y). Cette relation définit une classe de lois conditionnelles qui sont équivalentes μ^X -presque partout.

Si $A = \mathbb{R}^n$, la formule 1.2 donne $P\{Y \in B\} = \int d\mu^X(x) \mu_x^Y(B)$ qui est donc la généralisation au cas continu de la formule de probabilités totales et composées. Si la loi du couple est absolument continue par rapport à un produit de mesures, on obtient très facilement l'expression des lois conditionnelles ainsi qu'une généralisation au cas continu de la formule de Bayes.

Théorème: Si le couple (X, Y) est absolument continu par rapport à la mesure $\lambda_1 \otimes \lambda_2$, sa densité étant notée $h(x, y)$, alors

- a - la loi marginale de X (resp. Y) est absolument continue par rapport à λ_1 (resp. λ_2) et sa densité est

$$h_1(x) = \int h(x, y) d\lambda_2(y) \quad (1.3)$$

(resp. $h_2(y) = \int h(x, y) d\lambda_1(x)$),

- b - la loi conditionnelle de Y à X fixé est absolument continue par rapport à λ_2 et sa densité est

$$h(y|x) = h(x, y)/h_1(x). \quad (1.4)$$

On a de même $h(x|y) = h(x, y)/h_2(y)$.

Démonstration :

a - $\forall A \in \mathcal{B}_{\mathbb{R}^n}, P\{X \in A\} = \int_A d\lambda_1(x) \int h(x, y) d\lambda_2(y) = \int_A d\lambda_1(x) h_1(x)$.

b - $\forall (A, B) \in \mathcal{B}_{\mathbb{R}^n} \times \mathcal{B}_{\mathbb{R}^p}$,

$$\begin{aligned} P\{X \in A, Y \in B\} &= \int_{A \times B} d\lambda_1(x) d\lambda_2(y) h(x, y) \\ &= \int_A d\lambda_1(x) h_1(x) \int_B d\lambda_2(y) h(x, y)/h_1(x). \end{aligned}$$

■

On en déduit immédiatement la généralisation au cas absolument continu de la formule de Bayes :

$$h(x|y) = \frac{h(x, y)}{h_2(y)} = \frac{h(h|x) h_1(x)}{\int h(h|u) h_1(u) d\lambda_1(u)} \quad (1.5)$$

qui permet donc d'exprimer la densité conditionnelle de X sachant y en fonction de la densité marginale de X et la densité conditionnelle de Y sachant x . Par la suite, la formule 1.5 sera très souvent utilisée pour établir la loi du vecteur des gaussiennes associées aux observations lorsque la loi est absolument continue. Cependant, dans quelques cas (relation fonctionnelle) il sera nécessaire de vérifier à l'aide de la formule 1.2 que la loi retenue est bien la loi conditionnelle.

1.4.2 Relation fonctionnelle

Supposons qu'il existe une fonction ψ de \mathbb{R}^n dans \mathbb{R}^n telle que $Z = \psi(X_M)$. La loi conditionnelle de Z à X_M fixé est donc $d\mu_{x_M}^Z(z) = \delta_{\psi(x_M)}(z)$ et la loi du couple (Z, X_M) est

$$\delta_{\psi(x_M)}(z) g(x_M) dx_M$$

où δ_{z_0} est la mesure de Dirac implantée au point z_0 .

On calcule la loi conditionnelle de X_M à Z fixé dans les cas particuliers suivants :

a - ψ est un difféomorphisme sur \mathbb{R}^n .

La loi conditionnelle de X_M à z fixé est $d\mu_z^{X_M}(x_M) = \delta_{\psi^{-1}(z)}(x)$.

Démonstration: $\forall (A, B) \in \mathcal{B}_{\mathbb{R}^n} \times \mathcal{B}_{\mathbb{R}^n}$,

$$\begin{aligned} P\{Z \in A, X_M \in B\} &= \int \mathbb{1}_A(z) \mathbb{1}_B(x) \delta_{\psi(x)}(z) g(x) dx \\ &= \int \mathbb{1}_A(z) \mathbb{1}_B(x) \delta_{\psi^{-1}(z)}(x) g(x) \frac{dz}{|J_\psi(\psi^{-1}(z))|} \\ &= \int_A g(\psi^{-1}(z)) \frac{dz}{|J_\psi(\psi^{-1}(z))|} \int_B \delta_{\psi^{-1}(z)}(x) \end{aligned}$$

En effet, ψ étant un difféomorphisme, on a $\delta_{\psi(x)}(z) dx = \delta_{\psi^{-1}(z)}(z) dz / |J_\psi(\psi^{-1}(z))|$ — J_ψ est le jacobien de ψ . Or $g(\psi^{-1}(z)) / |J_\psi(\psi^{-1}(z))|$ est la densité marginale de Z , donc $\delta_{\psi^{-1}(z)}(\cdot)$ est la loi conditionnelle de X_M à Z fixé. ■

La validation des contraintes se réduit donc au calcul de $\psi^{-1}(z)$. Une simulation *classique* correspond à ce cas lorsque $\psi(x) = (\phi(x_i))_{1 \leq i \leq n}$ où ϕ est un difféomorphisme sur \mathbb{R} .

b - ψ est une fonction étagée.

ψ est définie par des ouverts disjoints de \mathbb{R}^n , $(C_i)_{i \in I}$, tels que $\overline{\bigcup_{i \in I} C_i} = \mathbb{R}^n$ et des éléments de \mathbb{R}^n , $(\alpha_i)_{i \in I}$:

$$\psi(x) = \sum_{i \in I} \alpha_i \mathbb{1}_{C_i}(x) \text{ p.p.}$$

La loi du couple (Z, X_M) est absolument continue par rapport à la mesure $\sum_{i \in I} \delta_{\alpha_i}(z) \otimes \lambda(x)$ et sa densité est $g(x) \sum_{i \in I} \mathbb{1}_{\{\alpha_i\} \times C_i}(x, z)$ — λ désigne ici la mesure de Lebesgue sur \mathbb{R}^n .

Démonstration:

$$\delta_{\psi(x)}(z) g(x) dx = \sum_{i \in I} \mathbb{1}_{C_i}(x) g(x) \delta_{\alpha_i}(z) dx = g(x) \sum_{i \in I} \mathbb{1}_{\{\alpha_i\} \times C_i}(z, x) \left(\sum_{i \in I} \delta_{\alpha_i}(z) \otimes dx \right)$$

Des formules 1.3 et 1.4, on déduit: ■

- la loi marginale de Z , $d\mu^Z(z) = \sum_{i \in I} \delta_{\alpha_i}(z) \int_{C_i} g(u) du$,
- la loi conditionnelle de X_M à Z fixé, $d\mu_z^{X_M}(x) = \mathbb{1}_{C_i}(x) g(x) / \int_{C_i} g(u) du$.

Ceci correspond au modèle dit des gaussiennes tronquées ou seuillées.

c - Une forme de ψ un peu plus générale

$$\psi = \sum_{i \in I} \mathbb{1}_{B_i} \psi_i + \sum_{j \in J} \mathbb{1}_{C_j} \alpha_j \text{ p.p.}$$

où les $(B_i)_{i \in I}$ et les $(C_j)_{j \in J}$ sont des ouverts disjoints de \mathbb{R}^n , tels que $\overline{\bigcup_{i \in I} B_i} \cup \overline{\bigcup_{j \in J} C_j} = \mathbb{R}^n$, les $(\psi_i)_{i \in I}$ sont des difféomorphismes de B_i et les $(\alpha_j)_{j \in J}$ des éléments de \mathbb{R}^n . On retrouve les deux cas particuliers précédents lorsque $B_1 = \mathbb{R}^n$ ou lorsque $\overline{\bigcup_i B_i} = \emptyset$. La loi marginale de Z est alors :

$$d\mu^Z(z) = dz \sum_{i \in I} \mathbb{1}_{A_i}(z) \frac{g \circ \psi_i^{-1}(z)}{|J_\psi(\psi_i^{-1}(z))|} + \sum_{j \in J} \delta_{\alpha_j}(z) \int_{C_j} g(u) du$$

où $A_i = \psi_i(B_i)$ et J_ψ est le jacobien de ψ . Quant à la loi conditionnelle de X_M à z fixé, on a :

$$\left\{ \begin{array}{l} z \neq \alpha_j \quad d\mu_z^{X_M}(x) = \frac{\sum_{i \in I} \mathbb{1}_{A_i}(z) \frac{g \circ \psi_i^{-1}(z)}{|J_\psi(\psi_i^{-1}(z))|} \delta_{\psi_i^{-1}(z)}(x)}{\sum_{i \in I} \mathbb{1}_{A_i}(z) \frac{g \circ \psi_i^{-1}(z)}{|J_\psi(\psi_i^{-1}(z))|}} \\ d\mu_{\alpha_j}^{X_M}(x) = \frac{\mathbb{1}_{C_j}(x) g(x) dx}{\int_{C_j} g(u) du} \end{array} \right.$$

Cette formule s'interprète de la manière suivante :

— ou bien la valeur observée est un atome, $Z = \alpha_i$, on choisit alors la valeur de X_M selon la densité

$$\frac{g(x) \mathbb{1}_{C_i}(x)}{\int_{C_i} g(u) du}$$

— ou bien la valeur observée n'est pas celle d'un atome, on choisit une valeur dans $\psi^{-1}(\{z\}) = \{\psi_i^{-1}(z), z \in A_i\}$ selon la loi de probabilité

$$\frac{g \circ \psi_i^{-1}(z)}{|J_\psi(\psi_i^{-1}(z))|} / \sum_k \mathbb{1}_{A_k}(z) \frac{g \circ \psi_k^{-1}(z)}{|J_\psi(\psi_k^{-1}(z))|}$$

Par exemple, pour la fonction de transformation ($n = 1$),

$$\psi(x) = \mathbb{1}_{]-\infty, -1[}(x) - x \mathbb{1}_{]-1, 0[}(x) + x^2 \mathbb{1}_{]0, +\infty[}(x)$$

on a les cas suivants :

$$\begin{aligned} 0 < z < 1 \quad d\mu_z^{X_M}(x) &= \frac{\delta_{-z}(x)g(-z) + \delta_{\sqrt{z}}(x)g(\sqrt{z})/2\sqrt{z}}{g(-z) + g(\sqrt{z})/2\sqrt{z}} \\ z = 1 \quad d\mu_1^{X_M}(x) &= \frac{\mathbb{1}_{]-\infty, -1[}(x)g(x)dx}{\int_{]-\infty, -1[} g(u) du} \\ z > 1 \quad d\mu_z^{X_M}(x) &= \delta_{\sqrt{z}}(x) \end{aligned}$$

Il peut sembler surprenant que pour $z = 1$, X_M ne puisse prendre la valeur 1. Mais la probabilité de l'événement $\{X_M = 1\}$ est nulle, puisque la loi marginale de X_M ne charge

pas le point 1, alors que celle de l'événement $\{X_M < -1\}$ est strictement positive. Ceci implique $P\{X_M < -1|Z = 1\} = 1$. Pour une observation $z = 1$, on attribuera donc à X_M une valeur inférieure à -1 . Pour une valeur observée de z petite (< 1), l'influence du jacobien est très sensible : que x soit égale à \sqrt{z} est alors beaucoup plus probable que $-z$.

Dans la pratique, une telle situation apparaîtrait si l'observation était la transformée par une fonction non monotone de la variable à simuler. Dans les applications qui seront développées par la suite, il n'a pas été fait usage de cette formule.

d - ψ se décompose en une partie bijective et une autre qui ne l'est pas.

$$z = (z_1, z_2) = (\psi_1(x_1), \psi_2(x_2))$$

La loi conditionnelle se factorise :

$$d\mu_z^X(x) = \delta_{\psi_1^{-1}(z_1)} d\mu_{x_1, z_2}^{X_2}(x_2) .$$

La simulation de X_M conditionnée par z s'effectue en deux étapes. On calcule tout d'abord $x_1 = \psi_1^{-1}(z_1)$, puis on simule x_2 selon la loi $d\mu_{x_1, z_2}^{X_2}(x_2)$.

Des exemples de ces différents cas seront présentés aux chapitres 2 et 3.

1.4.3 Randomisation

Soit $(\mu_x)_{x \in \mathbb{R}^n}$, une famille de lois de probabilité dépendant du paramètre x et ν une loi de probabilité sur \mathbb{R}^n . On définit [18] $\tilde{\mu}$, la randomisation de μ_x par ν la loi de probabilité telle que :

$$\mathcal{A} \in \mathcal{B}_{\mathbb{R}^n} \quad \tilde{\mu}(\mathcal{A}) = \int \mu_x(\mathcal{A}) d\nu(x)$$

On considère un couple de variables aléatoires (Z, X_M) dont la loi est $d\mu_x(z) d\nu(x)$. Les lois marginales respectives sont $\tilde{\mu}$ et ν . La loi conditionnelle de Z connaissant x_M , est $\mu_{x_M}^Z(z)$. Comme il est toujours possible de trouver une variable aléatoire, Θ_x , de loi $\mu_{x_M}^Z$ (x_M est alors fixé), on définit intuitivement Z comme la variable Θ_{X_M} , c'est à dire Θ de paramètre aléatoire X_M .

Pour simuler X_M conditionnellement à $Z = z$, il nous faut connaître la loi conditionnelle de X_M connaissant la valeur de Z . Pour cela, on fait l'hypothèse que toutes les mesures $\mu_{x_M}^Z$ sont absolument continues par rapport à une même mesure. Par ailleurs, X_M est absolument continue par rapport à la mesure de Lebesgue. Soient $f(z|x)$ et $g(x)$ les densités respectives. La loi du couple étant absolument continue par rapport au produit des mesures, la formule de Bayes (page 13) donne alors la densité conditionnelle de X sachant z :

$$g(x|z) = \frac{f(z|x)g(x)}{\int f(z|u)g(u)du}$$

Les modèles utilisés en pratique supposent de plus que, connaissant la valeur de la gaussienne X_M , les composantes du vecteur des observations Z sont indépendantes et que leur loi ne

dépend que de la valeur de la gaussienne associée. La densité conditionnelle, $f(z|x)$, se factorise alors : $f(z|x) = \prod_{1 \leq i \leq n} f(z_i|x_i)$. La densité conditionnelle à simuler se réécrit donc :

$$g(x|z) = \frac{\prod_{1 \leq i \leq n} f(z_i|x_i)g(x_1, \dots, x_n)}{\int \prod_{1 \leq i \leq n} f(z_i|u_i)g(u_1, \dots, u_n) du_1 \cdots du_n} \quad (1.6)$$

1.5 Des observations

L'objet de cette partie est d'illustrer comment les modèles présentés plus haut peuvent être utilisés pour expliciter la relation liant les observations et la variable à simuler. Quatre exemples sont présentés :

- la corégionalisation de variables,
- le support de la mesure,
- un bruit affectant la mesure,
- l'observation est exprimée par des inégalités.

1.5.1 Corégionalisation linéaire de facteurs gaussiens

En un point expérimental on dispose de mesures de plusieurs variables corrélées. Ainsi, en mine, sur un échantillon on détermine différentes teneurs ou la nature lithologique. De même, lorsque l'on étudie une veine (considérée alors comme un plan), en un point il est possible de définir la puissance de la veine et son accumulation en métal. Les corrélations entre les variables sont mises en évidence par le calcul des variogrammes expérimentaux simples et croisés.

On impose au modèle de corégionalisation simulé de reproduire les histogrammes des variables et les covariances simples et croisées. Un modèle possible consiste à considérer les différentes variables comme les transformées de fonctions aléatoires multigaussiennes corrélées. Aux variables étudiées, $(Z^{(i)})_{1 \leq i \leq nv}$, sont alors associées les fonctions aléatoires gaussiennes, $(X^{(i)})_{1 \leq i \leq nv}$, à l'aide des fonctions d'anamorphose, $(\phi^{(i)})_{1 \leq i \leq nv}$. On a donc :

$$Z^{(i)} = \phi^{(i)}(X^{(i)})$$

Les fonctions d'anamorphose, $(\phi^{(i)})_{1 \leq i \leq nv}$, sont choisies de manière à modéliser correctement l'histogramme des variables. Mais ces fonctions de transformation, éléments de $L^2(\mathbb{R}, g)$, s'écrivent comme des combinaisons linéaires des polynômes d'Hermite (normalisés) :

$$\phi^{(i)}(x) = \sum_{n \geq 0} \phi_n^{(i)} \eta_n(x)$$

Les corrélations entre les variables étudiées, $C^{ij}(h) = \text{Cov}(Z_x^{(i)}, Z_{x+h}^{(j)})$, déterminent les corrélations entre les gaussiennes selon les relations :

$$1 \leq i, j \leq nv \quad C^{ij}(h) = \sum_{n > 0} \phi_n^{(i)} \phi_n^{(j)} (\rho^{ij}(h))^n$$

On recherche donc un modèle pour les corrélations entre les gaussiennes, $(\rho^{ij}(h))_{i,j}$, tel que les covariances $(C^{ij}(h))_{i,j}$ ajustent les covariances expérimentales simples et croisées. Un modèle couramment utilisé est celui d'une corégionalisation linéaire [47] de facteurs gaussiens (indépendants), $(U^k)_{1 \leq k \leq \nu}$. Ces facteurs sont définis par leur covariance $\tilde{\rho}^k(h) = \text{Cov}(U_x^k, U_{x+h}^k)$. Les gaussiennes s'écrivent donc :

$$X^{(i)} = \sum_{1 \leq k \leq \nu} a_k^i U^k$$

où les a_k^i sont choisis de sorte que :

$$\rho^{ij}(h) = \sum_{1 \leq k \leq \nu} a_k^i a_k^j \tilde{\rho}^k(h)$$

Une telle modélisation n'est pas toujours possible, car elle impose en particulier que les covariances croisées soient symétriques. Un contre-exemple simple est celui de la corégionalisation (X_1, X_2) où $X_2(u) = X_1(u + h_0)$. La covariance croisée entre ces deux variables est donc la covariance de X_1 translatée de h_0 , cette corégionalisation ne peut donc être modélisée par une corégionalisation linéaire.

Lorsque la corégionalisation étudiée peut être modélisée comme la transformée d'une corégionalisation de fonctions aléatoires gaussiennes, on construit le vecteur, X_M , des gaussiennes associées par une fonction de transformation à une observation. On peut alors récrire de manière synthétique le vecteur des observations sous la forme :

$$Z = \psi(X_M)$$

1.5.2 Changement de support

Dans de nombreuses applications, le support de la mesure et celui de la variable à simuler diffèrent. On doit simuler une régularisée de la variable mesurée ou, inversement, la variable mesurée est une régularisée de la variable à simuler. En mine, le volume des échantillons analysés et celui de l'unité de sélection diffèrent de plusieurs ordres de grandeur (respectivement des centimètres cubes et des décimètres cubes). La teneur des blocs est donc la régularisée de la teneur des échantillons. Par contre, les mesures indirectes sont souvent des régularisées de la variable d'intérêt. Ainsi de la radioactivité sont déduites des teneurs en uranium, et de la gravimétrie des densités. La régularisation d'une fonction aléatoire Y par la mesure finie ν est définie par :

$$Y_u^\nu = \int Y_{u+v} d\nu(v)$$

Dans le cas où la variable ponctuelle est gaussienne et la mesure de régularisation à support fini, le lien explicite entre X et X^ν est ainsi :

$$X_i^\nu = \sum_{\alpha \in V_i} a_i^\alpha X_\alpha$$

Le conditionnement de l'une des variables par l'autre s'effectue aisément par un co-krigeage conditionnant.

Dès que la fonction d'anamorphose n'est plus linéaire, la résolution du problème n'est pas chose aisée, aussi dans la pratique on a recours à des approximations. On s'intéressera ici à un cas simple de régularisation, le changement de support. La variable définie sur le volume V , Y_u^V , est déduite de la variable ponctuelle Y_u par :

$$Y_u^V = \frac{1}{|V|} \int_V Y_{u+v} dv$$

Deux modèles sont ici utilisés, le modèle gaussien discret et la randomisation. Tous deux s'appuient sur les hypothèses suivantes. Au lieu de la teneur ponctuelle en un point, Y_u , on considère la teneur ponctuelle d'un point aléatoire uniformément distribué dans V , $Y_{\underline{u}}$. Conditionnellement à la teneur Y_u^V du bloc, $Y_{\underline{u}}$ est indépendante de la teneur des autres blocs, Y_v^V , et de la teneur ponctuelle mesurée dans un autre bloc, Y_v . Il reste alors à déterminer la loi de changement de support, c'est-à-dire la loi du couple $(Y_{\underline{u}}, Y_u^V)$.

Le modèle gaussien discret

Dans le modèle gaussien discret [35], la teneur ponctuelle $Y_{\underline{u}}$ et la teneur de bloc Y_u^V sont toutes deux modélisées comme l'anamorphosée de variables gaussiennes :

$$\begin{aligned} Y_{\underline{u}} &= \phi(X_{\underline{u}}) \\ Y_u^V &= \phi^V(X_u^V) \end{aligned}$$

Le couple $(X_{\underline{u}}, X_u^V)$ est bigaussien avec $\text{Cov}(X_{\underline{u}}, X_u^V) = r$. Les paramètres de ce modèle sont ϕ^V et r . Ils sont déduits de ϕ , déterminée à partir des données expérimentales, et de la variance de la variable régularisée, calculée par géostatistique linéaire, en utilisant la relation de Cartier :

$$\mathbf{E}\{Y_{\underline{u}}|Y_u^V\} = Y_u^V$$

Dans ce modèle, la corégionalisation $(Y_{\underline{u}}, Y_u^V)$ est donc modélisée à l'aide d'une corégionalisation linéaire de facteurs gaussiens (ce modèle n'est utilisable que si on dispose d'un échantillon au plus par bloc) :

$$X_{\underline{u}} = rX_u^V + \sqrt{1-r^2} U_u$$

où U est un facteur pépétique. La relation entre l'observation et la variable à simuler est donc là encore de la forme :

$$Z = \psi(X_M)$$

Randomisation

Dans le modèle de changement de support par randomisation, les variables $Y_{\underline{u}}$ et Y_u^V sont modélisées comme les régularisées d'une même mesure orthogonale (c'est-à-dire un effet de

pépite):

$$\begin{aligned} Y_u &= \frac{1}{|v|} \int_{v_u} dM(w) \\ Y_u^V &= \frac{1}{|V|} \int_{V_u} dM(w) \end{aligned}$$

avec $v_u \subset V_u$. Des choix particuliers de la mesure M conduisent aux modèles suivants.

La proportion minéralisée

Le bloc V est constitué par la réunion de n_V petits volumes élémentaires. Ces petits volumes peuvent être ou non minéralisés et il n'y a pas de corrélation entre eux. On définit la proportion minéralisée du bloc V , Y^V , comme la proportion des volumes élémentaires minéralisés. Cette variable est donc comprise entre 0 et 1. La proportion minéralisée d'un petit bloc v , $v \subset V$, est donc la proportion de volumes élémentaires minéralisés contenus dans v . Comme il n'y a pas de corrélation entre la minéralisation des volumes élémentaires, un petit bloc est un tirage *sans remise* de n_v volumes élémentaires dans une urne en contenant n_V , dont $n_V Y^V$ sont minéralisés et $n_V(1 - Y^V)$ ne le sont pas. On a donc

$$Y^v \stackrel{\mathcal{L}}{\equiv} \frac{1}{n_v} \mathcal{H}(n_V Y^V, n_V(1 - Y^V), n_v)$$

où \mathcal{H} est une variable aléatoire hypergéométrique ⁶. On peut se ramener à des lois plus simples dans les cas particuliers suivants.

- Si $n_v \ll n_V$: $Y^v \stackrel{\mathcal{L}}{\equiv} \frac{1}{n_v} \mathcal{B}(n_v, Y^V)$ où \mathcal{B} est une variable aléatoire binomiale ⁷.
- Si Y^V est petit, n_v grand et n_V encore plus grand, $Y^v \stackrel{\mathcal{L}}{\equiv} \frac{1}{n_v} \mathcal{P}_{n_V Y^V}$ où \mathcal{P} est une variable aléatoire de Poisson ⁸.

En effet sous certaines conditions⁹, une variable aléatoire hypergéométrique converge en loi vers une variable aléatoire binomiale et une variable aléatoire binomiale vers une variable aléatoire de Poisson.

Le nombre de pépites

Y^V est le nombre de pépites dans V , rapporté au volume de V . Les pépites sont supposées réparties de manière uniforme dans V et leurs localisations sont mutuellement indépendantes.

⁶La loi d'une variable aléatoire hypergéométrique est : $0 \leq k \leq n$, $P\{\mathcal{H}(a, b, n) = k\} = \frac{C_a^k C_b^{n-k}}{C_{a+b}^n}$

⁷La loi d'une variable aléatoire binomiale est : $0 \leq k \leq n$, $P\{\mathcal{B}(n, p) = k\} = C_n^k p^k (1-p)^{n-k}$

⁸La loi d'une variable aléatoire de Poisson est : $0 \leq k$, $P\{\mathcal{P}_\theta = k\} = e^{-\theta} \frac{\theta^k}{k!}$

⁹

$$\begin{array}{ccc} \lim_{\substack{a \rightarrow \infty \\ b \rightarrow \infty \\ a/a+b \rightarrow p}} \mathcal{H}(a, b, n) = \mathcal{B}(n, p) & & \lim_{\substack{n \rightarrow \infty \\ np \rightarrow \theta}} \mathcal{B}(n, p) = \mathcal{P}_\theta \end{array}$$

Le nombre de pépites contenues dans le petit bloc v divisé par le volume de v est donc :

$$Y^v \stackrel{\mathcal{L}}{=} \frac{1}{|v|} \mathcal{B}(|V|Y^V, \frac{|v|}{|V|})$$

Si $|V|Y^V$ est grand et $|v|/|V|$ petit, on a l'approximation

$$Y^v \stackrel{\mathcal{L}}{=} \frac{1}{|v|} \mathcal{P}_{|v|Y^V}$$

Dans ces deux cas, la relation de Cartier est vérifiée, $\mathbf{E}\{Y_u^v | Y_u^V\} = Y_u^V$. La variable mesurée est modélisée comme la randomisée de la variable d'intérêt qui elle-même peut être modélisée comme la transformée d'une variable gaussienne.

1.5.3 Erreurs de mesures

Bien souvent les mesures dont on dispose sont entachées d'erreurs. Pour modéliser ces erreurs, on suppose généralement qu'elles sont d'une certaine manière indépendantes et homogènes (c'est-à-dire, d'une mesure à l'autre les caractéristiques du bruit ne varient pas trop).

Notons $Z = (Z_i)_{1 \leq i \leq p}$ le vecteur des observations bruitées et $Y = (Y_i)_{1 \leq i \leq p} = (\phi(X_i))_{1 \leq i \leq p}$ le vecteur de la variable non bruitée. La modélisation probabiliste de ces bruits [3] suppose généralement l'indépendance conditionnelle des composantes de Z à Y fixé et que la loi conditionnelle de Z_i ne dépend que de la valeur y_i de Y_i . Si $f(z_i | y_i)$ est la densité conditionnelle de Z_i sachant y_i , la densité conditionnelle de Z sachant x est donc : $f(z|x) = \prod_{1 \leq i \leq p} f(z_i | \phi(x_i))$ et la densité conditionnelle du vecteur des gaussiennes associées à la variable non bruitée est calculée à l'aide de la formule de Bayes (voir page 13) :

$$g(x|z) = \frac{\prod_{1 \leq i \leq p} f(z_i | \phi(x_i)) g(x_1, \dots, x_n)}{\int \prod_{1 \leq i \leq p} f(z_i | \phi(u_i)) g(u_1, \dots, u_n) du_1 \dots du_n}$$

De nombreux bruits peuvent être modélisés par une randomisation de la variable sous-jacente :

- Le modèle de bruit additif et indépendant : $Z_i = Y_i + \epsilon_i$ où les ϵ_i sont des variables aléatoires indépendantes et identiquement distribuées de densité h . La densité conditionnelle de Z_i à x_i fixé est tout simplement $h(z_i - \phi(x_i))$. Si les erreurs sont uniformément réparties sur $[-a, a]$,

$$g(x|z) = \frac{\prod_{1 \leq i \leq n} \mathbb{1}_{[-a, a]}(z_i - \phi(x_i)) g(x_1, \dots, x_n)}{\int \prod_{1 \leq i \leq n} \mathbb{1}_{[-a, a]}(z_i - \phi(u_i)) g(u_1, \dots, u_n) du_1 \dots du_n} = \frac{\mathbb{1}_B(x) g(x)}{\int_B g(u) du}$$

avec ici $B = [\phi^{-1}(z_1 - a), \phi^{-1}(z_1 + a)] \times \dots \times [\phi^{-1}(z_n - a), \phi^{-1}(z_n + a)]$. ϕ est ici supposée bijective et connue.

- Le bruit poissonien : ce bruit est rencontré en imagerie micro-sonde[10] : la variable observée est une variable aléatoire de Poisson dont le paramètre est la variable sous-jacente. Dans ce cas, les valeurs faibles sont relativement plus bruitées que les fortes valeurs.

1.5.4 Contraintes d'inégalités

Parfois, l'observation est exprimée par des inégalités : $1 \leq i \leq n$, $Y_i \in C_i$ où les $(C_i)_{1 \leq i \leq n}$ sont des sous-ensembles de \mathbb{R} . Si le vecteur à simuler est $Y = (\phi(X_i))_{1 \leq i \leq p}$, la contrainte sur la variable Y peut s'exprimer comme une contrainte sur la variable X :

$$1 \leq i \leq n, \quad Y_i \in C_i \quad \iff \quad X_i \in B_i = \phi^{-1}(C_i).$$

Les exemples suivants entrent dans ce cadre :

anamorphose non bijective

Des variables discrètes peuvent être modélisée par une fonction d'anamorphose¹⁰ de la forme, $\phi = \sum_i a_i \mathbb{1}_{B_i}$. L'observation " $z_\alpha = a_i$ " se traduit en termes du modèle par " $x_\alpha \in B_i$ ". De cette manière [40, 14], sont modélisés des faciès géologiques séquencés de réservoirs pétroliers. Au chapitre 2, cette technique a été utilisée pour représenter un faciès lithologique dans la corégionalisation géologie-teneur d'un gisement d'uranium.

On peut aussi modéliser une variable présentant des atomes. Ainsi en mine, les teneurs présentent bien souvent un fort atome en zéro. Cet *effet zéro* peut se modéliser à l'aide de la fonction d'anamorphose, $\phi(x) = \mathbb{1}_B(x)\psi(x)$ où l'intervalle B est de la forme $[x_c, +\infty[$ et ψ est une fonction strictement monotone sur B telle que $\psi(x_c) = 0$. La mesure de z se traduit alors en termes du modèle de la manière suivante : si $z_\alpha > 0$ alors $x_\alpha = \psi^{-1}(z_\alpha)$ sinon $z_\alpha = 0$ impose la contrainte $x_\alpha < x_c$.

appareil de mesure

La valeur de la mesure est fournie par un appareil. Elle n'est significative que sur un intervalle $[a, b]$. Cette limitation peut être due à l'étalonnage de l'instrument, réalisé sur cet intervalle, au bruit de fond — au dessous du seuil, la mesure est une mesure du bruit de fond et non du phénomène étudié — ou à la saturation de l'appareil de mesure. Aussi à une mesure z_α on fait correspondre dans le modèle la contrainte sur la variable y_α :

$$\alpha \in A \quad \begin{cases} Y_\alpha < a & \text{si} & z_\alpha < a \\ Y_\alpha = z_\alpha & \text{si} & a \leq z_\alpha \leq b \\ Y_\alpha > b & \text{si} & b < z_\alpha \end{cases}$$

V. Langlais[26] a étudié la prise en compte, dans une procédure d'estimation, d'informations données sous forme d'inégalités. Dans l'industrie pétrolière, ce type d'information apparaît par exemple lorsque un horizon géologique n'a pu être atteint par un sondage de taille limitée.

fonction aléatoire de substitution

La simulation d'un vecteur gaussien tronqué intervient dans la simulation conditionnelle d'un tout autre modèle de fonction aléatoire que le modèle gaussien anamorphosé : les fonctions aléatoires de substitution. Le modèle à simuler est $Z_{(u)} = X_{(T_{(u)})}$ où X est un processus

¹⁰Le terme d'anamorphose est ici abusif, car la fonction ainsi définie n'est pas strictement croissante.

markovien discret et T est une FAI-0 gaussienne. X et T sont indépendantes. Cette fonction aléatoire est observée sur le domaine $A: \forall \alpha \in A, Z_\alpha = z_\alpha$. L'algorithme de simulation conditionnelle proposé par Ch. Lantuéjoul[28] se décompose en quatre étapes :

(i) On simule le processus $\tilde{X}_{(t)}$ de même loi que X .

(ii) \tilde{X} ayant été simulé, $\tilde{X}_{(T_\alpha)} = z_\alpha$ définit aux points expérimentaux les contraintes,

$$\alpha \in A \quad T_\alpha \in \tilde{X}^{-1}(\{z_\alpha\})$$

$\tilde{X}^{-1}(\{z_\alpha\})$ étant une réunion d'intervalles, la simulation du vecteur $(T_\alpha)_{\alpha \in A}$ se ramène à la simulation d'un vecteur gaussien tronqué. Puis T est simulé sur tout le champ conditionnellement aux valeurs simulées aux points expérimentaux.

(iii) On simule X conditionnellement aux valeurs prises par le processus aux temps $(t_\alpha)_{\alpha \in A}$, c'est-à-dire $\alpha \in A, X_{t_\alpha} = z_\alpha$.

(iv) Enfin on calcule

$$u \in S \quad Z_u = X_{(T(u))}$$

En toute rigueur, $\tilde{X}^{-1}(\{z_\alpha\})$ est une réunion infinie d'intervalles. Cependant, \tilde{X} est simulé sur un domaine fini, ce qui limite le nombre d'intervalles. On a donc là un exemple où la contrainte sur chacune des composantes du vecteur gaussien tronqué est une réunion de très nombreux intervalles.

1.6 Conclusion

Pour simuler une variable régionalisée, il faut spécifier un modèle de fonction aléatoire. Or les données disponibles ne permettent pas toujours de tester rigoureusement cette hypothèse. C'est pourquoi il importe de conditionner le plus possible les simulations de sorte que l'importance du choix du modèle soit tempérée par le conditionnement. Les relations entre les variables mesurées et la variable à simuler sont modélisées à l'aide de deux outils :

- la transformation de variables par une fonction,
- la randomisation.

Dans le cadre de ces modèles, la simulation conditionnelle se décompose en deux étapes, la validation des contraintes suivie d'une simulation conditionnelle classique. La mise en œuvre de la validation des contraintes est présentée aux chapitres 2 et 3 pour le cas de gaussiennes tronquées et au chapitre 4 pour le cas d'une randomisation (poissonisation).

Chapitre 2

Le gisement d'uranium de Laporte

Dans de très nombreux gisements miniers, la teneur est une variable régionalisée liée à un guide géologique : deux types de contrôle, non exclusifs voire complémentaires, sont le contrôle lithologique et le contrôle structural. Ainsi le géologue pour décrire un gisement utilisera des propositions telles que : *“les minéralisations sont contenues par un type de roche”* ou *“les minéralisations ne peuvent être rencontrées dans le faciès X”*. Mais bien souvent il s'échappe de cette alternative, inclusion-exclusion, pour évoquer des notions plus complexes de voisinage ou de localisation préférentielle. A cette logique ensembliste — le contrôle lithologique — se superpose celle d'éléments structuraux tels que failles, fractures, diaclases etc. *“Les minéralisations sont portées par un réseau de fractures.”* Ces logiques géologiques ne sont pas impératives : un faciès favorable n'est pas entièrement minéralisé, pas plus que ne le sont les fractures porteuses. Elles ne sont pas indépendantes — la distinction entre le contrôle lithologique et le contrôle structural est une simplification propre au discours —, leur conjonction permet d'affiner le modèle géologique : ainsi *“les minéralisations sont localisées à l'intersection d'un filon et d'un réseau de fractures”*. Cependant, si le modèle géologique permettait une localisation des minéralisations — il faudrait connaître exactement la géologie, ce qui n'est pas possible en début d'un projet minier — il ne détermine ni les formes ni le métal contenu. Aussi est-il indispensable d'intégrer les caractéristiques principales du gîte au modèle géostatistique, ceci afin d'utiliser au mieux toute l'information disponible et d'autre part d'assurer la compatibilité entre le modèle numérique et le contexte géologique.

Au cours d'une étude géostatistique, l'étude géologique est implicitement prise en compte. Ainsi le champ étudié sera décomposé en sous-ensembles sur lesquels la teneur est homogène. Les paramètres du modèle structural — portée du variogramme, anisotropies, etc. — sont inférés sur les mesures. Cependant ils reflètent la géologie, sans que le lien soit explicite. Un deuxième exemple est donné par le modèle qui a été développé il y a quelques années pour rendre compte des contrastes des teneurs dans certains gisements d'uranium [24, 13] : l'enveloppe des fortes teneurs était tout d'abord déterminée et représentée par son indicatrice. Les teneurs étaient ensuite simulées séparément de l'enveloppe et seules les teneurs situées dans l'enveloppe étaient conservées. Une conditionnalisation approximative de la géométrie permettait localement le calage des teneurs fortes au cœur de l'enveloppe plutôt

qu'en bordure. Ce modèle présentait deux inconvénients : le modèle reliant la géométrie de l'enveloppe des teneurs fortes aux valeurs de la teneur n'est pas explicité et la conditionnalisation de l'indicatrice de la géométrie des teneurs riches était très approximative. Dans ces exemples, les inconvénients résultent de l'imprécision du modèle : quel est le statut des variables géologiques fixant la géométrie et par conséquent à quel modèle de corégionalisation géométrie-teneur fait-on appel ?

Le but de ce chapitre est donc de présenter un exemple de corégionalisation entre la teneur et une lithologie, établi sur le gisement de Laporte. Ce gisement ayant déjà été exploité, on disposait de sondages d'exploration, des résultats d'exploitation ainsi que d'un modèle géologique permettant de préciser la localisation des minéralisations. Ce gisement ainsi que les données disponibles sont présentés au paragraphe 2.1. A l'aide de covariances, il est possible d'étudier la relation entre la teneur et une géométrie : ceci est l'objet du paragraphe 2.2. Un modèle de corégionalisation entre la teneur et une variable géologique est spécifié. Ce modèle ainsi que son utilisation, c'est à dire les simulations, seront présentés au paragraphe 2.3.

2.1 Le gisement de Laporte

Le gisement d'uranium de Laporte est situé dans le Cantal (France) à proximité de Saint-Privat. Il a été exploité en carrière par la société Total Compagnie Minière (TCM) au cours des années 83 et 84. Le gîte est de type fissural encaissé dans un granite : les minéralisations se rencontrent à l'intersection de filons de lamprophyres subverticaux de direction Nord-Sud avec un réseau de diaclases quasi-perpendiculaires (N70) d'espacement décamétrique.

Le modèle géologique est illustré par le levé d'exploitation du niveau 508 de la carrière (figure 2.1). Les trois composantes du modèle sont représentées :

- 5 filons de lamprophyres sont visibles. Leur puissance varie d'une cinquantaine de centimètres à cinq mètres, leur espacement de 1.5 à 12 mètres. Leur tracé présente une bonne continuité Nord-Sud, avec des décalages dextres aux intersections avec les diaclases.
- Les fissures ouvertes et les diaclases sont reportées, principalement sur les pourtours du niveau, au pied du gradin.
- La radioactivité a été mesurée au sol. Le géologue en déduit une représentation approximative et lissée des courbes isoteneurs pour quatre coupures différentes. Les teneurs les plus fortes sont situées à l'intérieur des lamprophyres, elles ne présentent pas de continuité Nord-Sud. Les teneurs plus faibles diffusent en dehors de ce faciès, préférentiellement en suivant les diaclases ; elles réalisent localement des jonctions entre des filons voisins. Inversement, bien que les lamprophyres portent les teneurs les plus fortes, ils ne sont pas minéralisés intégralement.

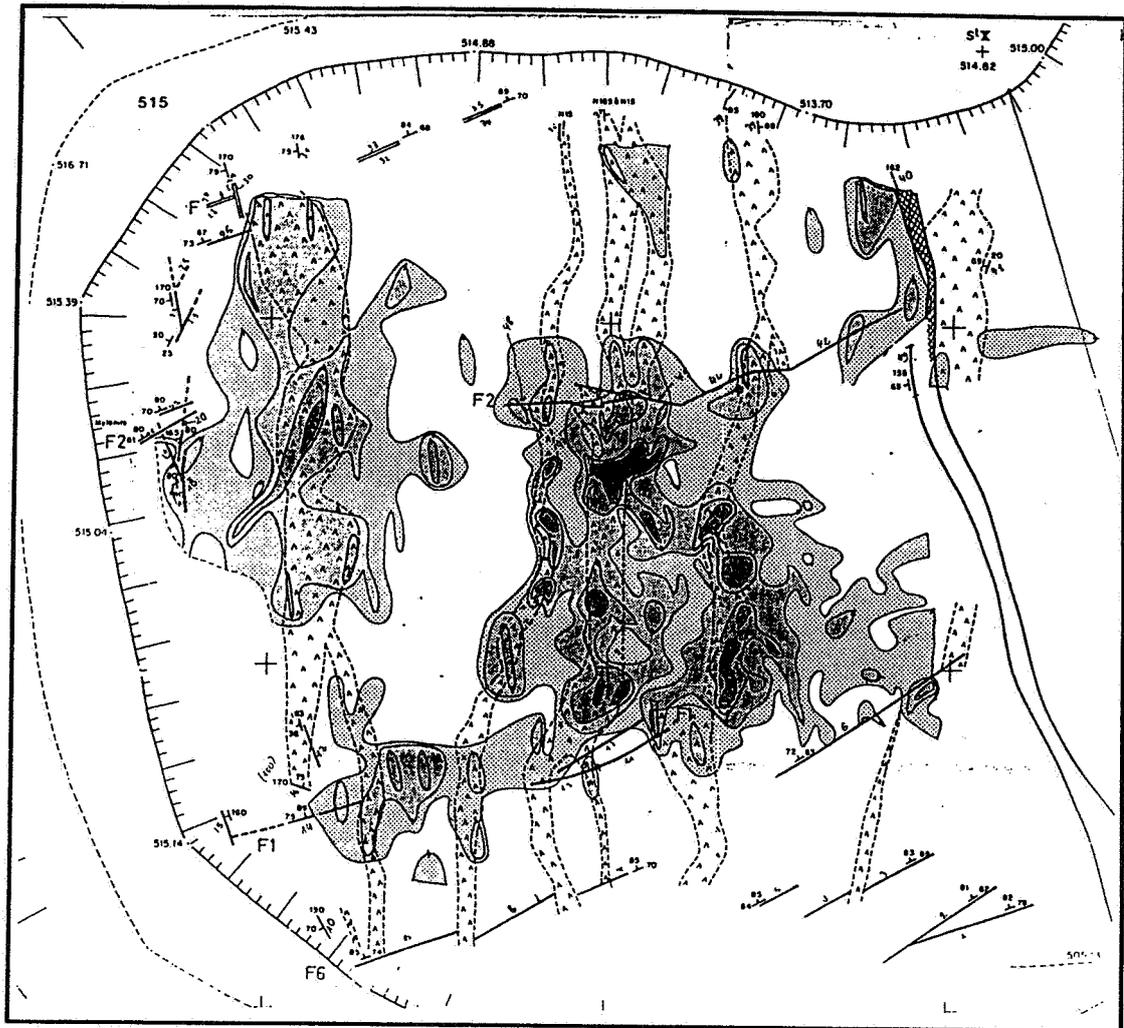
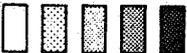


Figure 2.1 : Levé d'exploitation du niveau 508 de Laporte (d'après document SCUMRA,1984)

teneur :  lamprophyres : 

Pour mener l'étude géostatistique de ce gisement, on disposait de deux types de données : le levé d'exploitation présenté ci-dessus et les sondages de reconnaissance.

Le levé d'exploitation

Le levé d'exploitation mettant en évidence la relation entre la géométrie et la teneur servira au départ à la construction du modèle. Pour ce niveau, il offre une description exhaustive des filons. Par contre, le report des teneurs manque de fiabilité du point de vue numérique : les mesures étaient effectuées approximativement sur une grille, puis elles ont été lissées et interpolées de manière empirique. De plus le regroupement en cinq classes rend malaisée l'inférence du modèle. Enfin, fonder un modèle de gisement sur un levé obtenu à l'issue de l'exploitation ne correspondrait pas à une pratique minière puisque l'étude géostatistique précède généralement l'exploitation.

Les sondages d'exploration

Le gisement a été reconnu par des sondages radiométriques. Après corrélation, les échantillons sont informés en teneur à pas de un mètre. La lithologie dominante est déterminée à partir des *cuttings*. Dans le parallélépipède d'enveloppe de la carrière, les 73 sondages sont répartis suivant quatre directions : verticale, à 60 degrés Sud, c'est-à-dire dans le plan des lamprophyres, et les recoupant à 60 degrés Est et 45 degrés Ouest.

Les variables étudiées par la suite sont l'indicatrice des lamprophyres et leur teneur en uranium. On recherche un modèle de corégionalisation entre ces deux variables : pour le gîte de Laporte, le canevas géométrique sera donné par les filons de lamprophyres. Il n'a pas été possible d'intégrer explicitement une des composantes du modèle géologique, les failles, au modèle géostatistique. En effet les mesures des failles données par le levé d'exploitation sont localisées préférentiellement à la périphérie de la zone d'étude ; elles ne sont pas disponibles sur les sondages radiométriques.

2.2 Covariances et géométrie

Dans le modèle géologique, les minéralisations sont "*plutôt*" dans les lamprophyres. Cette affirmation du géologue peut-elle se traduire en termes quantitatifs ? Tout d'abord, on présente les outils structuraux utilisés pour préciser le lien entre une variable régionalisée et un ensemble. Puis, ce lien est étudié sur le gîte de Laporte.

2.2.1 Les covariances d'indicatrice, des outils structuraux

Sur \mathbb{R}^n , on considère une variable régionalisée z et un ensemble d . Sur le gisement de Laporte, z représente la teneur et d les filons de lamprophyres. Le support ¹ de ces variables est

¹Le support d'une variable régionalisée z , est l'ensemble sur lequel la variable est non nulle :

$$\text{supp}z = \{x \in \mathbb{R}^n, z(x) \neq 0\}$$

d'extension limitée. L'ensemble d est défini par son indicatrice $\mathbb{1}_d$:

$$x \in \mathbb{R}^n, \mathbb{1}_d(x) = \begin{cases} 1 & \text{si } x \in d \\ 0 & \text{sinon} \end{cases}$$

On convient d'un domaine d'étude, S , contenant les supports des deux variables étudiées, z et $\mathbb{1}_d$. Une première étape est le calcul de la valeur moyenne de z sur différents ensembles : S , d et $d^c \cap S$. On notera \bar{z}_E la valeur moyenne de z sur l'ensemble E , c'est-à-dire

$$\bar{z}_E = \frac{1}{|E|} \int_E z(u) du$$

De la valeur de ces moyennes, on peut déduire des relations simples sur la géométrie de z :

- si $\bar{z}_{d^c \cap S} = 0$ le support de z est contenu dans d ,
- si $\bar{z}_d > \bar{z}_S$ les valeurs de z sont en moyenne plus élevées dans d .

Elles ne permettent pas de préciser comment se répartissent les valeurs de z à l'intérieur ou au voisinage de d . A titre d'exemple, quatre cas sont considérés : z est constante sur d (I), les valeurs les plus fortes sont au cœur de d (II), les valeurs les plus fortes sont sur les bordures de l'ensemble (III) et les valeurs les plus fortes sont localisées sur un des bords de l'ensemble (IV). Ils sont schématisés sur la figure 2.2. Sur ces quatre exemples, l'épaisseur et la teneur moyenne de la traversée sont constantes. Dans les cas II, III ou IV, il y a une relation forte entre z et la géométrie d , qu'il conviendrait de modéliser.

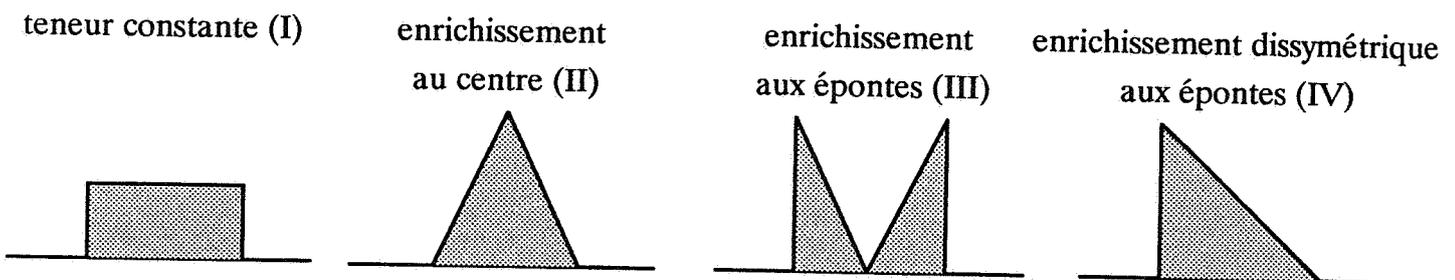


Figure 2.2: Quatre exemples de traversée minéralisée

Pour cela, on définit les bords de la manière suivante :

- le bord intérieur de d est $d \cap d_{-h}^c$,
- le bord extérieur de d est $d^c \cap d_{-h}$

où d_h est le translaté de d par le vecteur h et d^c le complémentaire de d (figure 2.3). La valeur moyenne sur les bords intérieurs calculée sur les exemples triviaux précédents est une fonction de h : figure 2.4. Dans ce cas, l'allure de ces courbes permet de mettre en évidence les *effets de bords*.

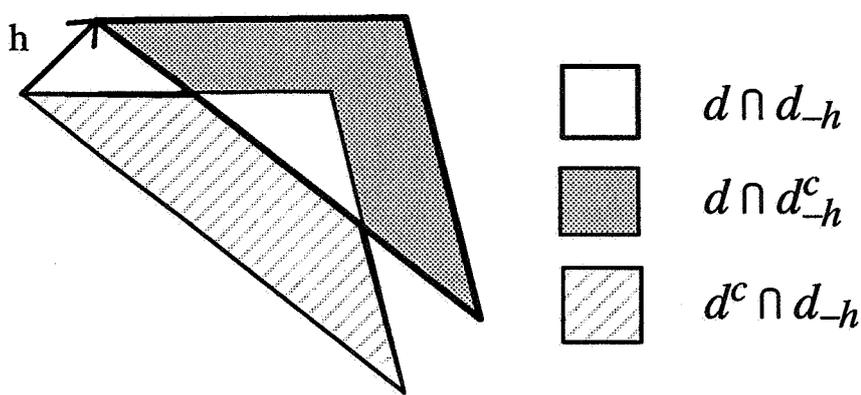
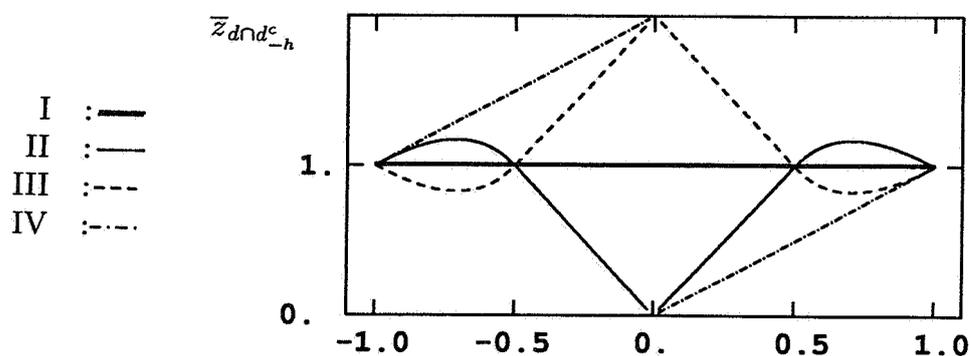


Figure 2.3 : Bord intérieur et bord extérieur

Figure 2.4 : Valeurs moyennes calculées sur le bord intérieur de I, II, III et IV



Il est possible d'une part de calculer une approximation de la valeur moyenne sur les bords à l'aide de covariances expérimentales et d'autre part d'interpréter les résultats expérimentaux en termes de modèle. En effet, sur un domaine S contenant d , on calcule la covariance croisée expérimentale des variables $z\mathbb{1}_d$ et $\mathbb{1}_{d^c}$:

$$\begin{aligned}\hat{\sigma}_{zd}(h) &= \frac{1}{|S \cap S_{-h}|} \int_{S \cap S_{-h}} z(x) \mathbb{1}_d(x) \mathbb{1}_{d^c}(x+h) dx \\ &= \frac{1}{|S \cap S_{-h}|} \int_{d \cap d_{-h}^c \cap S_{-h}} z(x) dx\end{aligned}$$

Dans le cas particulier où z est la fonction constante égale à 1, la covariance croisée est :

$$\hat{\sigma}_d(h) = \frac{|d \cap d_{-h}^c \cap S_{-h}|}{|S \cap S_{-h}|}$$

De même, on calcule les covariances expérimentales des variables $z\mathbb{1}_{d^c}$ et $\mathbb{1}_d$:

$$\begin{aligned}\hat{\sigma}_{zd^c}(h) &= \frac{1}{|S \cap S_{-h}|} \int_{d^c \cap d_{-h} \cap S} z(x) dx \\ \hat{\sigma}_{d^c}(h) &= \frac{|d^c \cap d_{-h} \cap S|}{|S \cap S_{-h}|}\end{aligned}$$

Le rapport des covariances expérimentales donne :

$$\begin{aligned}\frac{\hat{\sigma}_{zd}(h)}{\hat{\sigma}_d(h)} &= \frac{\int_{d \cap d_{-h}^c \cap S_{-h}} z(x) dx}{|d \cap d_{-h}^c \cap S_{-h}|} \\ \frac{\hat{\sigma}_{zd^c}(h)}{\hat{\sigma}_{d^c}(h)} &= \frac{\int_{d^c \cap d_{-h} \cap S} z(x) dx}{|d^c \cap d_{-h} \cap S|}\end{aligned}$$

Ce sont là presque les valeurs moyennes de z sur les bords, intérieurs et extérieurs, de largeur h . Ce mode de calcul introduit un biais dû au champ. Mais S peut être choisi tel que $d \subset S \cap S_{-h} \cap S_h$, on a alors

$$\bar{z}_{d \cap d_{-h}^c} = \frac{\hat{\sigma}_{zd}(h)}{\hat{\sigma}_d(h)} = \frac{\int_{d \cap d_{-h}^c} z(x) dx}{|d \cap d_{-h}^c|}$$

Cependant, ce choix est incompatible avec l'hypothèse de l'indépendance entre le champ S et les variables considérées, hypothèse sur laquelle s'appuie le formalisme intrinsèque.

Ces grandeurs se traduisent aisément en termes de modèle : on interprète alors z et d comme une réalisation respectivement d'une fonction aléatoire Z et d'un ensemble aléatoire stationnaire D — on fait l'hypothèse de l'indépendance entre la corégionalisation (Z, D) et le champ S . Les covariances expérimentales sont des estimateurs non biaisés des covariances théoriques :

$$\begin{aligned}\mathbf{E}\{\hat{\sigma}_{ZD}(h)\} &= \sigma_{ZD}(h) \\ &= \mathbf{E}\{Z(x) \mathbb{1}_D(x) \mathbb{1}_{D^c}(x+h)\} \\ &= \mathbf{E}\{Z(x) | x \in D \cap D_{-h}^c\} P\{x \in D \cap D_{-h}^c\}\end{aligned}$$

On a de même :

$$\mathbf{E}\{\hat{\sigma}_D(h)\} = \sigma_D(h) = P\{x \in D \cap D_{-h}^c\}$$

On prendra comme estimateur de $\mathbf{E}\{Z(x)|x \in D \cap D_h^c\}$, le rapport des covariances expérimentales : $\hat{\sigma}_{ZD}/\hat{\sigma}_D$. Rien ne nous assure cependant qu'il soit sans biais puisque :

$$\mathbf{E}\{\hat{\sigma}_{ZD}/\hat{\sigma}_D\} \neq \mathbf{E}\{\hat{\sigma}_{ZD}\}/\mathbf{E}\{\hat{\sigma}_D\}$$

Par contre, si (Z, D) est ergodique,

$$\begin{aligned} \lim_{|S| \rightarrow +\infty} \hat{\sigma}_{ZD} &= \sigma_{ZD} \quad \text{p.s.} \\ \lim_{|S| \rightarrow +\infty} \hat{\sigma}_D &= \sigma_D \quad \text{p.s.} \end{aligned}$$

Donc si $\sigma_D(h) \neq 0$,

$$\lim_{|S| \rightarrow +\infty} \frac{\hat{\sigma}_{ZD}(h)}{\hat{\sigma}_D(h)} = \frac{\sigma_{ZD}(h)}{\sigma_D(h)} = \mathbf{E}\{Z(x)|x \in D \cap D_h^c\}$$

Si les expressions précédentes sont symétriques en h , $\mathbf{E}\{Z(x)|x \in D \cap D_h^c\}$ peut s'exprimer aussi en termes de variogrammes simples et croisés :

$$\begin{aligned} \gamma_{ZD}(h) &= \frac{1}{2} \mathbf{E}\{(Z(x)\mathbb{1}_{D(x)} - Z(x+h)\mathbb{1}_{D(x+h)})(\mathbb{1}_{D(x)} - \mathbb{1}_{D(x+h)})\} \\ \gamma_D(h) &= \frac{1}{2} \mathbf{E}\{(\mathbb{1}_{D(x)} - \mathbb{1}_{D(x+h)})^2\} \end{aligned}$$

On a donc $\mathbf{E}\{Z(x)|x \in D \cap D_h^c\} = \gamma_{ZD}(h)/\gamma_D(h)$.

Les outils structuraux présentés ci-dessus précisent la répartition d'une variable régionalisée à l'intérieur d'une géométrie. Le statut de la géométrie par rapport à la variable étudiée permet de distinguer deux cas :

- la géométrie est *définie* par la variable z . Dans les exemples rencontrés, la géométrie est le champ sur lequel la variable est définie, ou bien la géométrie est définie par une coupure sur z .

– Relation entre une variable et son champ géométrique [32]

z est la variable étudiée, d désigne le champ sur lequel elle est définie. Le calcul de la valeur moyenne de z sur les bords (intérieurs) de d permet de tester l'indépendance entre la variable et son champ. Cette indépendance interne se traduit par l'égalité entre la valeur moyenne sur les bords et la valeur moyenne sur le champ. Elle est contenue dans l'hypothèse de base du formalisme intrinsèque dans lequel la variable régionalisée est interprétée comme une réalisation d'une fonction aléatoire intrinsèque.

– Relation entre les indicatrices d'une variable z

La suite des coupures $(z_i)_{1 \leq i \leq n}$, telle que $z_1 < z_2 < \dots < z_n$ définit une suite d'ensembles emboîtés : $(\mathbb{1}_{z(x) \geq z_i})_{1 \leq i \leq n}$. Pour un couple de coupures $z_i < z_j$, on peut s'intéresser à la répartition de l'ensemble $\mathbb{1}_{z(x) \geq z_j}$ à l'intérieur de l'ensemble $\mathbb{1}_{z(x) \geq z_i}$. Pour cela, on calcule le rapport de la covariance croisée des indicatrices et de la covariance de l'indicatrice de la coupure la plus basse. Ce rapport est un estimateur de

$$P\{Z_x \geq z_j | Z_x \geq z_i \text{ et } Z_{x+h} < z_i\}$$

En fonction des résultats expérimentaux, on choisit de modéliser la variable z avec un modèle sans *effet de bords*, appartenant à la famille des modèles à résidus d'indicatrices auto-krigeables développés par J. Rivoirard[43, 44], ou un modèle à *effet de bords*, par exemple le modèle multigaussien anamorphosé.

- la géométrie est une variable auxiliaire, qui est corrélée avec la variable étudiée sans être totalement définie par elle. Dans cette catégorie entrent les géométries dessinées par les faciès lithologiques.

2.2.2 Etude de Laporte

Les outils structuraux présentés aux paragraphes précédents ont été utilisés sur le gisement de Laporte pour préciser la localisation relative des fortes teneurs et des faibles teneurs et la relation entre la teneur et les filons de lamprophyres. L'étude a été faite sur le levé d'exploitation du niveau 508 (figure 2.1).

Ce levé est échantillonné sur une grille 351×351 . La teneur discrétisée sur cinq classes ($t \in \{0, 1, 2, 3, 4\}$) est représentée par les quatre indicatrices : $k \in \{1, 2, 3, 4\} \mathbb{1}_{t \geq k}$. Les filons de lamprophyre sont également représentés par leur indicatrice. Les valeurs moyennes des indicatrices des teneurs sur les bords des ensembles définis par une coupure plus basse et sur les bords des lamprophyres sont calculées par le rapport des covariances non centrées expérimentales. $\hat{\sigma}_{kk'}/\hat{\sigma}_k$ ($k < k'$) désigne donc le rapport des covariances croisées avec au numérateur les variables $\mathbb{1}_{t(x) \geq k'}$ et $\mathbb{1}_{t(x) < k}$ et au dénominateur les variables $\mathbb{1}_{t(x) \geq k}$ et $\mathbb{1}_{t(x) < k}$. De même, $\hat{\sigma}_{Lk}/\hat{\sigma}_L$ est le rapport des covariances croisées avec d'une part les variables $\mathbb{1}_{t(x) \geq k} \times \mathbb{1}_{x \in L}$ et $\mathbb{1}_{x \in L^c}$ et d'autre part les variables $\mathbb{1}_{x \in L}$ et $\mathbb{1}_{x \in L^c}$.

- A la figure 2.5 sont reportés les rapports $\hat{\sigma}_{23}/\hat{\sigma}_2$ et $\hat{\sigma}_{24}/\hat{\sigma}_2$ pour les directions Nord-Sud et Est-Ouest. Les courbes ont un comportement très semblable pour des décalages positifs ou négatifs; elles sont croissantes avec $|h|$. Au bord de l'ensemble défini par $t \geq 2$, les proportions des valeurs supérieures à 3 ou 4 sont inférieures à la proportion moyenne sur cet ensemble — 0.24 pour $k = 3$ et 0.05 pour $k = 4$. Les courbes $\hat{\sigma}_{23}/\hat{\sigma}_2$ et $\hat{\sigma}_{24}/\hat{\sigma}_2$ se stabilisent à ces valeurs pour de grands décalages. Les teneurs les plus fortes sont entourées par des teneurs moyennes, elles mêmes entourées par des teneurs faibles.
- A la figure 2.6 sont reportés les rapports $(\hat{\sigma}_{kL}/\hat{\sigma}_L)_k$ et $\hat{\sigma}_{kL^c}/\hat{\sigma}_{L^c}$ pour la direction Est-Ouest. On peut distinguer deux types de comportement. Pour les faibles et moyennes teneurs ($t = 1$ ou 2), les courbes sont approximativement symétriques et la proportion des teneurs supérieures à 1 ou 2 est légèrement supérieure à l'intérieur qu'à l'extérieur des lamprophyres. Par contre, les fortes teneurs ($t = 3$ ou 4) sont localisées sur le bord Ouest des lamprophyres (ce qui correspond à $h > 0$ pour le bord intérieur et $h < 0$ pour le bord extérieur) et de très fortes teneurs ne sont pas rencontrées sur le bord extérieur Est. En moyenne, les lamprophyres sont plus riches que leur complémentaire.

Dans la modélisation utilisée par la suite, il n'a pas été tenu compte du décalage des fortes teneurs sur les bords Ouest des lamprophyres. Seul a été retenu le caractère diffusif de la minéralisation et la localisation préférentielle des fortes teneurs dans les lamprophyres.

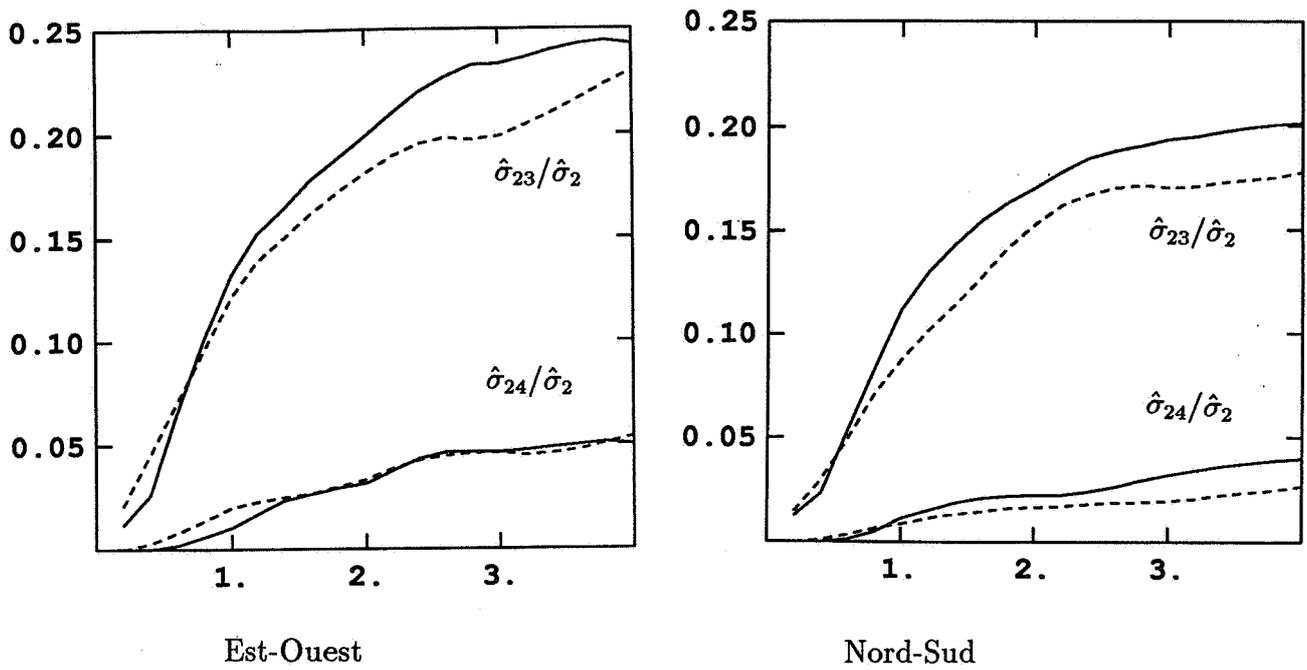


Figure 2.5: Rapports des covariances expérimentales $(\hat{\sigma}_{i2}/\hat{\sigma}_2)_{i=3,4}$ en fonction du décalage ($-h < 0$ et $---h > 0$)

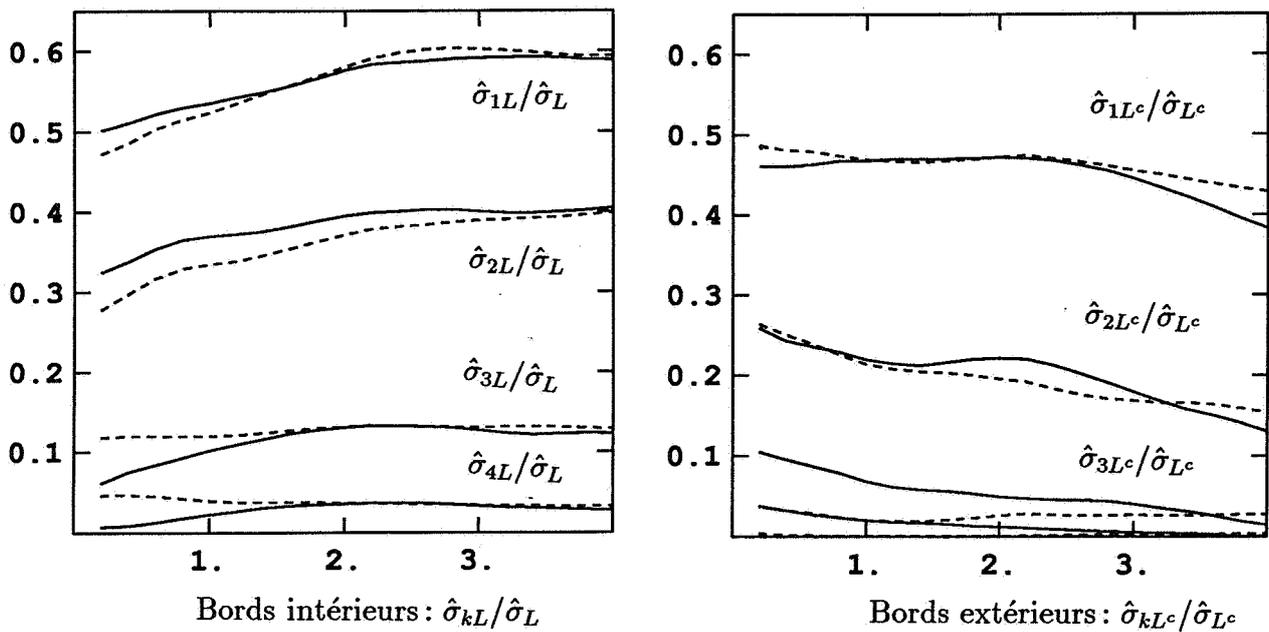


Figure 2.6: Rapports des covariances expérimentales $(\hat{\sigma}_{iL}/\hat{\sigma}_L)_i$ en fonction du décalage ($-h < 0$ et $---h > 0$)

2.3 Modélisation de Laporte

L'étude expérimentale a ainsi confirmé le modèle géologique : les fortes valeurs sont localisées de manière préférentielle au cœur des filons de lamprophyres et les teneurs faibles diffusent autour des plus fortes. Cependant les filons ne sont pas entièrement minéralisés et on rencontre des teneurs moyennes ou faibles dans les lamprophyres et en-dehors. Le but de l'étude étant une simulation du gisement, il nous faut disposer d'un modèle de corégionalisation qui rende compte des observations précédentes. Le modèle retenu est tout d'abord présenté, il est spécifié pour Laporte sur les sondages d'exploration dans la partie suivante. Enfin quelques exemples d'utilisation de ce modèle, des simulations conditionnelles, sont donnés.

2.3.1 Modèle de corégionalisation teneur-géologie

Le tracé continu mais irrégulier des lamprophyres suggère une représentation par une gaussienne tronquée. La teneur présentant des transitions continues des valeurs fortes vers des valeurs faibles est décrite par un modèle gaussien anamorphosé. La relation entre la géométrie des filons et la répartition des teneurs est modélisée par une corrélation entre les variables gaussiennes associées respectivement à la teneur et la géologie. Explicitement, le modèle de corégionalisation est le suivant :

- La teneur est obtenue par anamorphose d'une fonction aléatoire gaussienne centrée normée :

$$Z(x) = \phi(Y_1(x))$$

- L'indicatrice des lamprophyres est modélisée par la troncature d'une gaussienne :

$$\mathbb{1}_L(x) = \mathbb{1}_{Y_2 \geq y_c}(x)$$

Afin de modéliser les corrélations entre les deux gaussiennes, on décompose la gaussienne des teneurs, Y_1 , en un terme géométrique, linéaire en Y_2 , et un résidu indépendant également gaussien : $Y_1(x) = \alpha Y_2(x) + \beta U(x)$ α est donc le coefficient de corrélation entre les deux gaussiennes.

Ce modèle de corégionalisation permet de traduire et de quantifier les remarques du géologue en termes du modèle. Supposons α positif et appelons une forte teneur une teneur supérieure à $z^* = \phi(y^*)$. Vérifier si les fortes teneurs sont préférentiellement dans les lamprophyres nous amène à comparer les deux probabilités conditionnelles :

$$\begin{aligned} P\{Z \geq z^* | x \in L\} &= P\{Y_1 \geq y^* | Y_2 \geq y_c\} \\ P\{Z \geq z^* | x \in L^c\} &= P\{Y_1 \geq y^* | Y_2 < y_c\} \end{aligned}$$

A titre d'exemple, prenons $P\{Z \geq z^*\} = 0.05$ et $P\{x \in L\} = 0.2$. Pour différentes valeurs de α , les valeurs de $P\{Z \geq z^* | x \in L\}$ et $P\{Z \geq z^* | x \in L^c\}$ sont reportées au tableau 2.1 : dans le modèle, la concordance plus ou moins stricte entre les fortes teneurs et les filons de lamprophyres dépend de la valeur de α . D'autre part, la proportion des lamprophyres non

minéralisés est donnée par $P\{Z < z^* | x \in L\} = P\{Y_1 < y^* | Y_2 \geq y_c\}$. Ainsi la lithologie explique en partie la répartition de la minéralisation et joue un rôle prépondérant pour la génération des valeurs fortes de la teneur. La structure du résidu $U(x)$ explique des différences de structure des filons et des teneurs : différence d'anisotropie, répartition des classes de teneurs à l'intérieur des lamprophyres et en-dehors.

Tableau 2.1 : Relation théorique entre les fortes valeurs et l'indicatrice de lamprophyres

$$P\{Z \geq z^*\} = 0.05 \quad P\{x \in L\} = 0.20$$

α	$P\{Z \geq z^* x \in L\}$	$P\{Z \geq z^* x \in L^c\}$	$\frac{P\{Z \geq z^* x \in L\}}{P\{Z \geq z^* x \in L^c\}}$	$P\{Z < z^* x \in L\}$
0.5	0.1460	0.0260	5.6	0.85
0.6	0.1705	0.0199	8.5	0.83
0.7	0.1965	0.0134	14.7	0.80
0.8	0.2225	0.0069	32.4	0.77
0.9	0.2440	0.0015	162.0	0.75

Ce modèle comporte les paramètres suivants :

- la fonction d'anamorphose de la teneur ϕ ,
- la coupure définissant la géologie y_c ,
- les covariances simples et croisée des variables gaussiennes Y_1 et Y_2 (ou, ce qui est équivalent, la corrélation entre Y_1 et Y_2 , α , et les covariances simples des gaussiennes indépendantes Y_2 et U).

L'objet de l'analyse structurale est de vérifier ces hypothèses sur le gisement et de choisir une valeur pour chacun des paramètres.

2.3.2 Analyse structurale

Statistiques et histogrammes des échantillons de sondages donnent les résultats suivants :

- La forme de la distribution des teneurs est usuelle pour des données d'uranium. De très nombreuses faibles valeurs, voire strictement nulles pour 13 % d'entre elles, s'opposent à quelques très fortes valeurs (figure 2.7). A partir de cet histogramme on détermine une fonction d'anamorphose ϕ [29].
- La proportion des échantillons appartenant aux lamprophyres, p , détermine la coupure gaussienne associée, $y_c = G^{-1}(1 - p)$: $p = 0.23$ $y_c = 0.74$.
- La relation teneur-lithologie est précisée par le tableau 2.2.

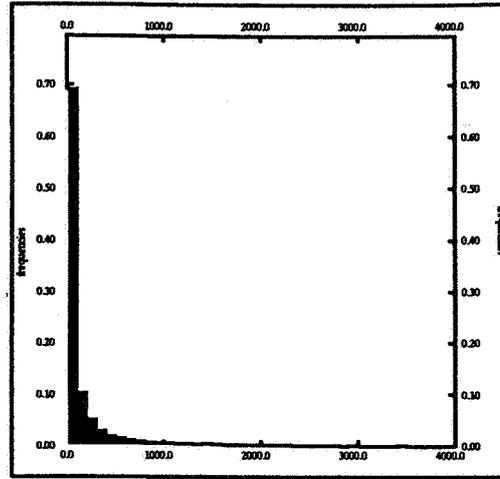


Figure 2.7 : Histogramme des teneurs du gisement de Laporte

Les variogrammes simples et croisés calculés le long des sondages pour la teneur et l'indicatrice des lamprophyres sont présentés sur la figure 2.8. Ils présentent des structures comprises entre trois et dix mètres. Afin d'assurer la cohérence du modèle[39], ces variogrammes expérimentaux ne sont pas ajustés directement. On a recherché un modèle de covariance pour les gaussiennes, modèle qui après transformation ajuste les variogrammes expérimentaux²: les deux fonctions de transformation, $\phi(\cdot)$ et $1_{\geq y_c}$, peuvent s'écrire comme des combinaisons linéaires des polynômes d'Hermite normés,

$$Z(x) = \phi(Y_1(x)) = \sum_{n \geq 0} \phi_n \eta_n(Y_1(x))$$

$$\mathbb{1}_L(x) = \mathbb{1}_{Y_2(x) \geq y_c} = \sum_{n \geq 0} \psi_n \eta_n(Y_2(x))$$

Les covariances simples et croisée entre $Z(x)$ et $\mathbb{1}_L(x)$ s'expriment donc de la façon suivante :

²Le variogramme croisé entre la teneur et la géologie est symétrique et ne peut rendre compte du léger décalage entre les très fortes teneurs et les lamprophyres

Tableau 2.2: Relation teneur-lithologie sur le gisement de Laporte

lithologie	nombre d'échantillons	moyenne des teneurs	variance
granite	1058	122	82000
lamprophyres	323	269	415000
total	1381	189	170000

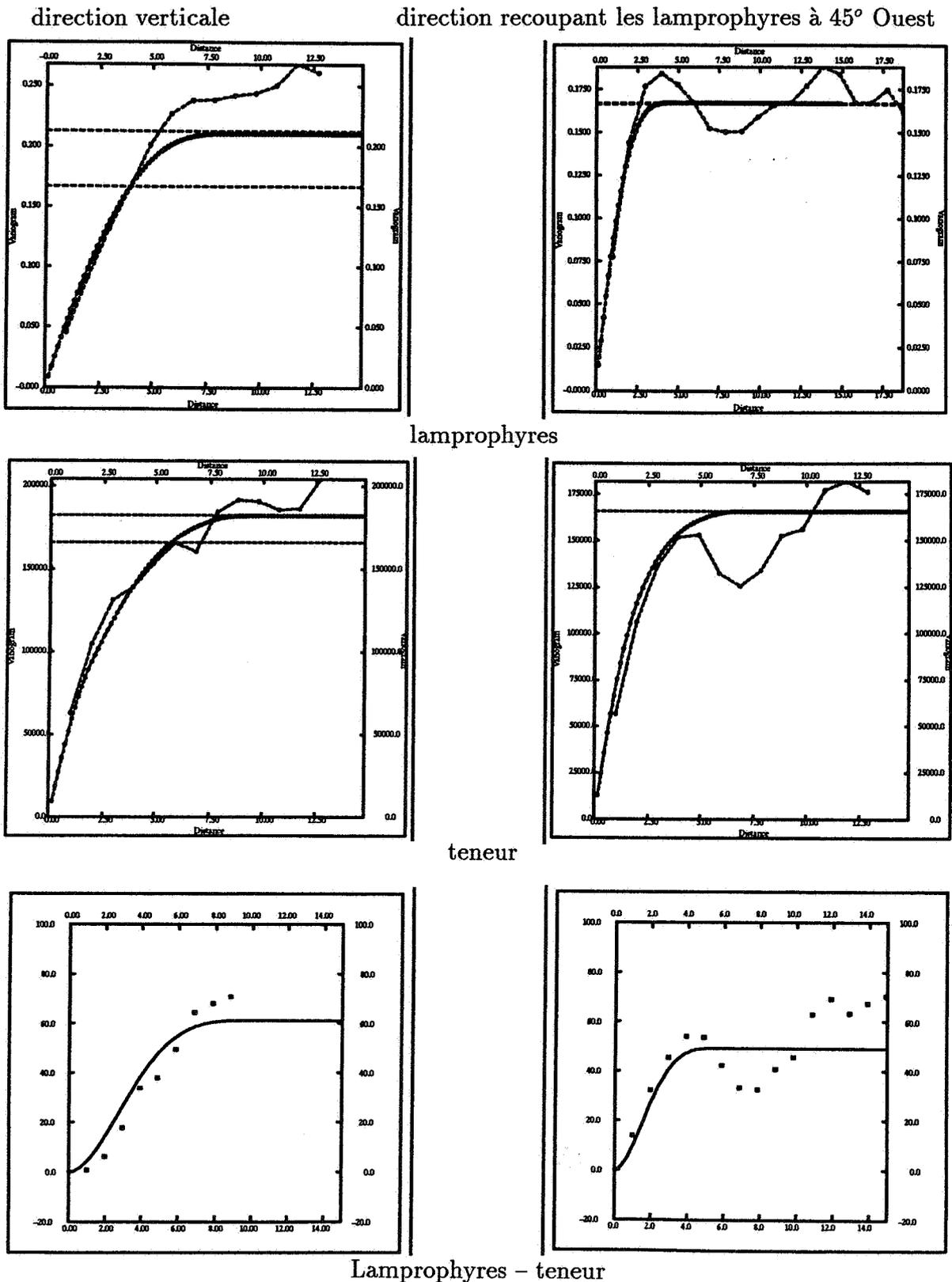


Figure 2.8: Variogrammes simples et croisés du gisement de Laporte

$$\begin{aligned}
C^{ZZ}(h) &= \text{Cov}(Z(x), Z(x+h)) = \sum_{n>0} \phi_n^2 [\text{Cov}(Y_1(x), Y_1(x+h))]^n \\
C^{ZL}(h) &= \text{Cov}(Z(x), \mathbb{1}_L(x+h)) = \sum_{n>0} \phi_n \psi_n [\text{Cov}(Y_1(x), Y_2(x+h))]^n \\
C^{LL}(h) &= \text{Cov}(\mathbb{1}_L(x), \mathbb{1}_L(x+h)) = \sum_{n>0} \psi_n^2 [\text{Cov}(Y_2(x), Y_2(x+h))]^n
\end{aligned}$$

Ces relations imposent un comportement parabolique au voisinage de l'origine pour la covariance de la gaussienne Y_2 , ceci afin que le comportement au voisinage de l'origine du variogramme de l'indicatrices des lamprophyres soit linéaire.

Les résultats sont les suivants :

- La gaussienne de la géologie est la somme de deux schémas cubiques anisotropes. Une des deux composantes présente une anisotropie zonale dans la direction Nord-Sud afin de rendre compte de l'aspect filonien.
- Le résidu est une somme de deux composantes sphériques anisotropes.
- Le coefficient α déterminé à partir de la corrélation expérimentale des variables est 0.5. On a donc,

$$Y_1(x) = 0.5Y_2(x) + 0.866U(x)$$

Les caractéristiques des facteurs gaussiens, Y_2 et $\mathbb{1}_L$, sont reportées au tableau 2.3. Ce modèle ajuste également les variogrammes expérimentaux de l'indicatrice des lamprophyres du levé d'exploitation : figure 2.9.

Tableau 2.3 : Détails des facteurs du modèle de Laporte

	gaussienne de la géologie : Y_1		résidu : U	
	$Y_1 = 0.58X_1 + 0.81X_2$		$U = 0.47X_3 + 0.88X_4$	
composante	X_1	X_2	X_3	X_4
type de covariance	cubique	cubique	sphérique	sphérique
portée	9 m	9 m	3 m	9 m
anisotropie selon Ox	3.2	2.5	1.0	3.2
anisotropie selon Oy	1.0	0.0	1.0	1.0
anisotropie selon Oz	1.0	1.0	1.0	1.0

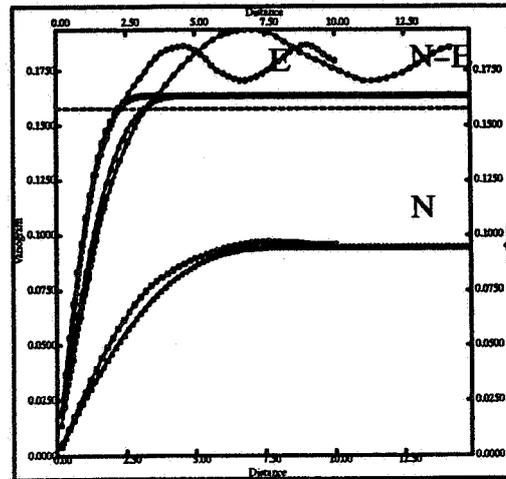


Figure 2.9: Variogrammes de l'indicatrice des lamprophyres du levé d'exploitation et ajustement (direction Est, Nord-Est et Nord)

2.3.3 Simulations

Le modèle structural présenté plus haut se prête aisément à une simulation conditionnelle de la corégionalisation géologie-teneur. La simulation du niveau 508 sera présentée pour deux types de conditionnement : simulation conditionnée par les sondages d'exploration (lithologie et teneur) et simulation conditionnée par la géologie du levé d'exploitation.

Simulation conditionnée par les sondages

En tout point conditionnant, on dispose de la valeur de la teneur et de l'indicatrice des lamprophyres. Cette information est équivalente dans le modèle à la donnée de la valeur de la gaussienne associée à la teneur — cette valeur est calculée par inversion³ de la fonction d'anamorphose $y_1(x_\alpha) = \phi^{-1}(z_1(x_\alpha))$ — et une inégalité sur la gaussienne associée à la géologie :

$$\begin{aligned} y_2(x_\alpha) < y_c & \text{ si } \mathbb{1}_L(x_\alpha) = 0 \\ y_2(x_\alpha) \geq y_c & \text{ si } \mathbb{1}_L(x_\alpha) = 1 \end{aligned}$$

La simulation conditionnelle se décompose en deux étapes :

- validation des inégalités, c'est-à-dire la simulation aux points expérimentaux de la variable gaussienne associée à la géologie et respectant les inégalités. Les méthodes pour réaliser cette simulation sont présentées dans le chapitre 3.
- simulation conditionnelle classique. En chaque point conditionnant la valeur de deux gaussiennes est connue ; on calcule donc les valeurs des deux facteurs indépendants. La simulation conditionnelle de la corégionalisation se ramène alors à la simulation conditionnelle classique de chacun des facteurs.

³On a négligé ici l'atome en zéro sur les teneurs.

A la figure 2.10, une simulation du niveau 508 est présentée. Pour ce niveau le nombre des données conditionnantes est faible (73). On ne pouvait espérer retrouver la géologie connue *a priori*. Cependant nous retrouvons les filons décrits par le géologue, par rapport auxquels les corps minéralisés s'organisent convenablement.

Simulation conditionnée par la géologie du levé d'exploitation

A l'issue de l'étude géologique, l'exploitant dispose d'une interprétation géologique du gisement. En pratique, un tel document peut être utilisé pour conditionner des modèles numériques des teneurs. Le levé des lamprophyres du niveau 508 est discrétisé sur une grille régulière de pas 0.6 mètre. La gaussienne de la géologie est simulée en respectant les inégalités, puis indépendamment le résidu. Par combinaison linéaire de ces deux facteurs et anamorphose on obtient les teneurs ponctuelles. Afin de les comparer à la cartographie lissée du levé, les teneurs simulées sont régularisées sur neuf points.

2.4 Conclusion

L'étude de Laporte a permis d'intégrer au modèle géostatistique du gisement une partie du modèle géologique : un faciès lithologique, les filons de lamprophyres, y est représenté par son indicatrice. Ainsi le recours à des variables auxiliaires permet-il d'affiner la représentation de la réalité étudiée ; il dépend de la nature des mesures disponibles. Ceci suppose une étude des relations entre les variables à incorporer au modèle. Ici les outils structuraux utilisés sont de covariances expérimentales. Le modèle retenu permet de ne rendre compte que d'une partie du phénomène étudié. La corégionalisation de facteurs gaussiens est une généralisation du modèle gaussien anamorphosé qui avait été étendue à la simulation d'ensembles aléatoires pour représenter des faciès lithologiques séquencés [40, 14].

Cette étude est un cas particulier — le lien particulier existant entre les variables sur le gisement de Laporte n'a pas de valeur générale pour tous les gisements. Mais intégrer dans un modèle cohérent toutes les mesures disponibles sur un phénomène est un des objets de la géostatistique. De plus, il est nécessaire d'introduire dans le modèle des caractéristiques du gîte pour assurer sa compatibilité avec le contexte géologique. Les contraintes apparaissent dans ce cas comme des éléments d'information qu'on ne peut négliger.

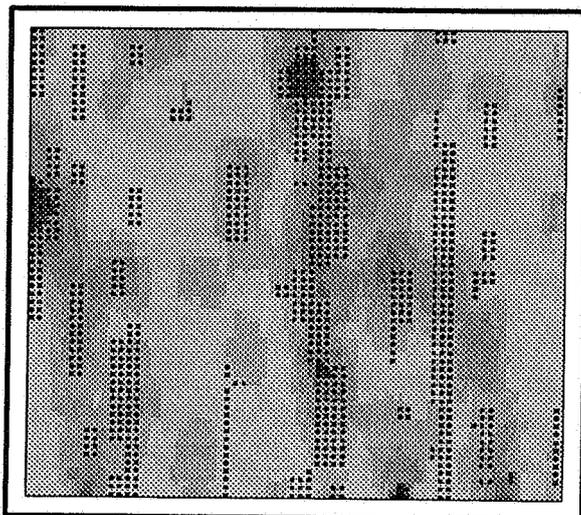


Figure 2.10 : Simulation de la corégionalisation teneur-lamprophyres (niveau 508) conditionnée par les sondages (les teneurs faibles sont représentées par un gris clair et les fortes teneurs un gris foncé ; la trace des filons de lamprophyres apparaît sous forme de points noirs)

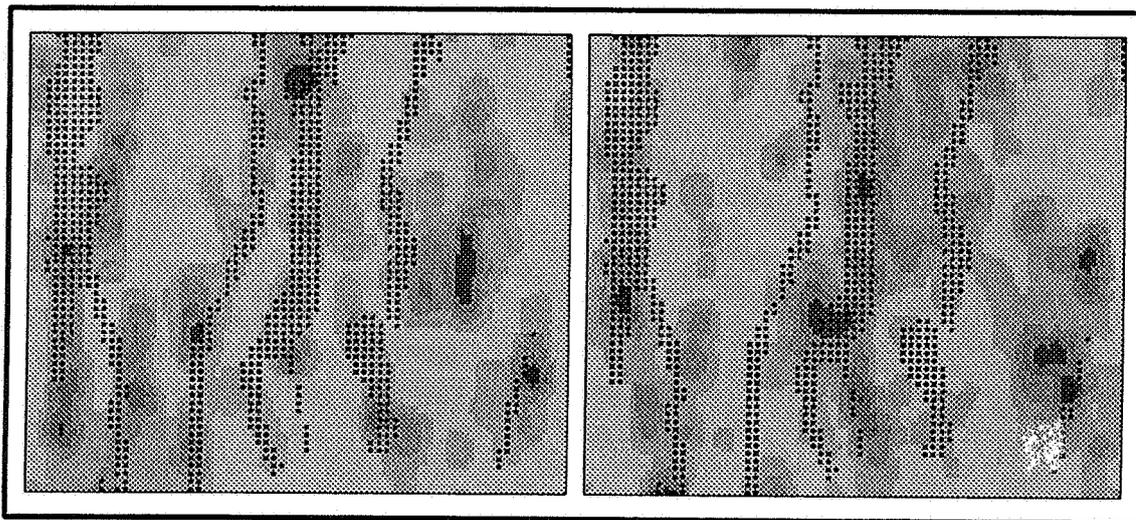


Figure 2.11 : Simulations de la teneur conditionnées par la géologie du niveau 508 (les teneurs faibles sont représentées par un gris clair et les fortes teneurs un gris foncé ; la trace des filons de lamprophyres apparaît sous forme de points noirs)

Chapitre 3

Simulation de gaussiennes tronquées

Dans le chapitre précédent, le modèle géostatistique utilisé pour décrire la corégionalisation d'une teneur et l'indicatrice d'un faciès associée à une mesure de la lithologie L une inégalité sur une variable gaussienne :

$$\begin{aligned} u \in L &\iff X_2(u) < x_c \\ u \in L^c &\iff X_2(u) \geq x_c \end{aligned}$$

Pour simuler conditionnellement cette corégionalisation, il faut donc simuler un vecteur multigaussien vérifiant des inégalités. L'objet de ce chapitre est de présenter des techniques pour réaliser cette simulation.

On précise tout d'abord la forme des inégalités que doit vérifier la simulation. Ceci définit ce qui est appelé par la suite un vecteur aléatoire gaussien tronqué. Puis deux types d'algorithmes sont présentés : ceux fournissant une simulation exacte et ceux simulant une approximation de la loi modélisée. Dans la première catégorie entrent deux algorithmes, la simulation par la méthode d'inversion et la simulation par acceptation et rejets. Tous deux ne s'appliquent qu'à des vecteurs de très petite dimension. Ils sont introduits ici car ils seront utilisés pour mettre en œuvre les algorithmes approximatifs. Dans cette dernière catégorie, sont successivement présentées la méthode incrémentale — toutes les composantes sont simulées l'une après l'autre — et la méthode itérative — la loi est simulée comme la loi stationnaire d'une chaîne de Markov.

Dans la définition d'une gaussienne tronquée deux éléments interviennent, d'une part la loi gaussienne et d'autre part la troncature. Un vecteur gaussien centré¹ de dimension n est défini par la matrice strictement² définie positive des covariances,

$$\Sigma = [\sigma_{ij}]_{1 \leq i, j \leq n} \text{ où } \sigma_{ij} = \text{Cov}(X_i, X_j) = \mathbf{E}\{X_i X_j\}$$

Sa densité est :

$$x \in \mathbb{R}^n \quad g(x) = \frac{1}{(2\pi)^{\frac{n}{2}}} \frac{1}{|\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}x'\Sigma^{-1}x\right)$$

On impose au vecteur la contrainte :

$$X \in B = B_1 \times B_2 \times \cdots \times B_n$$

Chacune des composantes doit appartenir à la réunion finie d'intervalles :

$$B_i = [a_i^1, b_i^1] \cup [a_i^2, b_i^2] \cup \cdots \cup [a_i^{n_i}, b_i^{n_i}]$$

Dans la plupart des applications, la contrainte sur chaque composante se réduit à un intervalle. Cependant, il arrive que B_i soit la réunion de très nombreux intervalles (un exemple est donné page 22). La densité conditionnelle du vecteur X sachant $X \in B$ est

$$x \in \mathbb{R}^n \quad f(x) = \frac{g(x)\mathbb{1}_B(x)}{\int_B g(u)du}$$

Un vecteur gaussien tronqué par B est un vecteur aléatoire réel dont la densité est f .

Le type d'inégalités retenu peut sembler restrictif. Il répond au problème qui nous était posé — la conditionnalisation de la corégionalisation teneur-géologie du gisement de Laporte. D'autre part, il apparaît dans de nombreuses applications, dont quelques exemples ont été présentés au paragraphe 1.5.4.

3.1 Simulations exactes

3.1.1 Simulation par inversion

La méthode de simulation par inversion est une méthode générale pour simuler une variable aléatoire scalaire X , dont la loi est donnée par la fonction de répartition :

$$x \in \mathbb{R} \quad F(x) = P\{X < x\} .$$

¹Si le vecteur n'est pas centré, on simule le vecteur centré $Y = (X_i - \mathbf{E}\{X_i\})_{1 \leq i \leq n}$: la matrice des covariances centrées n'est pas modifiée et seules les inégalités à vérifier sont translatées.

²La matrice des covariances est supposée strictement définie positive car si $r < n$ est le rang de la matrice, la simulation de X sous la contrainte $X \in B_1 \times \cdots \times B_n$ se ramène à la simulation d'un vecteur gaussien de dimension r sous une contrainte qui n'est plus un produit cartésien de contraintes sur chaque composante.

Comme F est une fonction croissante et continue à gauche, on peut lui associer un pseudo-inverse $F^{-1}(u) = \sup\{y, F(y) \leq u\}$ — si F est bijective, ce pseudo-inverse coïncide avec l'inverse. La méthode est déduite de la proposition suivante : si U est une variable aléatoire uniformément répartie sur $]0, 1[$ alors $F^{-1}(U)$ est une variable aléatoire dont la loi est F .

Il est possible d'utiliser cette méthode pour simuler une gaussienne tronquée scalaire car la fonction de répartition d'une gaussienne, G , et son inverse sont tabulées, et on dispose de méthodes pour simuler un nombre aléatoire uniformément réparti sur $]0, 1[$ — par exemple, la méthode de congruence de Lehmer [30]. Cette méthode présente deux limitations : la contrainte ne comporte qu'un intervalle — cette limitation est d'ordre pratique, les expressions de F et F^{-1} deviennent inextricables pour plusieurs intervalles —, et la contrainte doit être incluse dans le domaine sur lequel G et G^{-1} sont tabulées. L'expression de F et F^{-1} est alors

$$F(x) = \begin{cases} 0 & x \leq a \\ \frac{G(x)-G(a)}{G(b)-G(a)} & a < x < b \\ 1 & b \leq x \end{cases}$$

$$F^{-1}(u) = G^{-1}((1-u)G(a) + uG(b))$$

L'algorithme de simulation est donc

Algorithme 1 (i) simulation de U , variable aléatoire uniformément répartie sur $]0, 1[$,
(ii) calcul de $X = F^{-1}(U)$.

3.1.2 Simulation par acceptation et rejets

Cette méthode de simulation due à Von Neumann [46] permet de simuler une variable aléatoire réelle, scalaire ou vectorielle, de densité f . n désigne la dimension de la variable à simuler. Un élément de \mathbb{R}^{n+1} est noté (x, y) où $x \in \mathbb{R}^n$ et $y \in \mathbb{R}$. A_f désigne l'intersection du sous-graphe de f et du demi-espace $\mathbb{R}^n \times \mathbb{R}^+$:

$$A_f = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}, 0 \leq y \leq f(x)\}$$

La méthode d'acceptation et rejets s'appuie sur les deux théorèmes suivants :

Théorème 1 Si (X, Y) est un point uniformément réparti dans A_f , alors la loi de X admet f pour densité.

Théorème 2 Si X et U sont des variables aléatoires de densité respective f et $\mathbb{1}_{]0,1[}$, alors $(X, Uf(X))$ est un point uniforme dans A_f .

Les démonstrations détaillées de ces théorèmes peuvent être trouvées par exemple dans le cours de Ch. Lantuéjoul [27] que suit cette présentation de la méthode d'acceptation et rejets. La simulation d'une variable aléatoire X dont la densité est f suppose que f soit majorée par une densité h que l'on sait simuler (par exemple, par la méthode d'inversion si $n = 1$) :

$$x \in \mathbb{R}^n \quad f(x) \leq C h(x)$$

La simulation d'une variable aléatoire X dont la densité est f suppose que f soit majorée par une densité h que l'on sait simuler (par exemple, par la méthode d'inversion si $n = 1$):

$$x \in \mathbb{R}^n \quad f(x) \leq C h(x)$$

D'après le *théorème 1*, il suffit de simuler un point uniformément réparti dans A_f . Pour cela, on simule un point uniformément réparti sur A_{Ch} : on simule U et X de densité respective $\mathbb{1}_{]0,1[}$ et h . D'après le *théorème 2*, $(X, Cg(X)U)$ est uniformément réparti dans A_{Ch} . Ce point s'il tombe dans A_f est uniforme dans A_f puisque $A_f \subset A_{Ch}$ (figure 3.1).

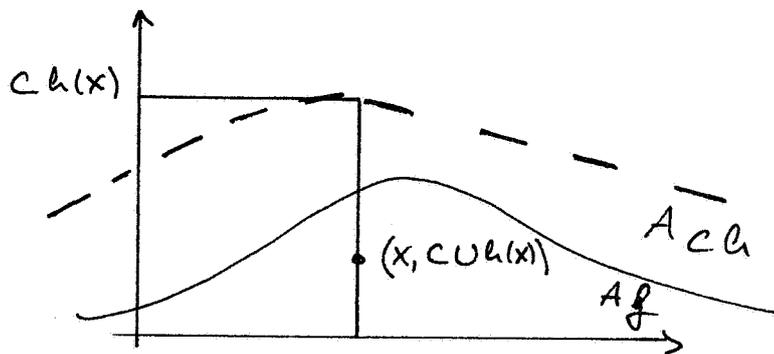


Figure 3.1 : La méthode d'acceptation et rejets

L'algorithme de la méthode de simulation par acceptation et rejets est le suivant :

- (i) *génération des variables X et U de densité respective h et $\mathbb{1}_{]0,1[}$,*
- (ii) *si $CUh(X) \leq f(X)$ le candidat est retenu, sinon retour en (i).*

La vitesse de cet algorithme est déterminée par le nombre de candidats qu'il faut simuler puis tester pour accepter une valeur. Ce nombre est une variable aléatoire de Pascal ³ de paramètre $1/C$, dont l'espérance est donc C . Cette méthode est applicable dès que l'on connaît une densité h que l'on sait simuler et pour laquelle la constante C n'est pas trop grande. Cette méthode est utilisée pour simuler une gaussienne tronquée scalaire — la contrainte peut être une réunion finie d'intervalles — et de petits vecteurs gaussiens tronqués — mais dans ce cas la contrainte sur chacune des composantes est un intervalle.

Une gaussienne tronquée scalaire

La contrainte est une réunion finie d'intervalles non vides disjoints, $B = \cup_i I_i$. Sur chacun des intervalles, on majore la densité g par une fonction positive et sommable h_i :

$$x \in I_i \quad g(x) \leq h_i(x)$$

³La loi d'une variable aléatoire de Pascal de paramètre p ($0 < p < 1$), \mathcal{N} est :

$$n \in \mathbb{N}^* \quad P\{\mathcal{N} = n\} = (1-p)^{n-1}p$$

On majore la densité de la gaussienne tronquée par

$$x \in \mathbb{R} \quad f(x) = \frac{g(x)\mathbb{1}_B(x)}{\int_B g(u)du} \leq C \sum_i \varpi_i \frac{h_i(x)\mathbb{1}_{I_i}(x)}{\int_{I_i} h_i(u)du} \quad \text{où} \quad \begin{cases} C = \frac{\sum_j \int_{I_j} h_j(u)du}{\int_B g(v)dv} \\ \varpi_i = \frac{\int_{I_i} h_i(u)du}{\sum_j \int_{I_j} h_j(u)du} \end{cases}$$

Les fonctions $(h_i)_i$ sont choisies de sorte que soit possible le calcul explicite du système de probabilité $(\varpi_i)_i$, ainsi que l'inverse des fonctions de répartition des densités construites avec les fonctions positives majorantes. La loi majorante est un mélange qui se simule aisément par la méthode d'inversion (cf. page 46). En pratique (figure 3.2), les inégalités utilisées pour majorer la densité gaussienne sur chacun des intervalles sont de deux sortes :

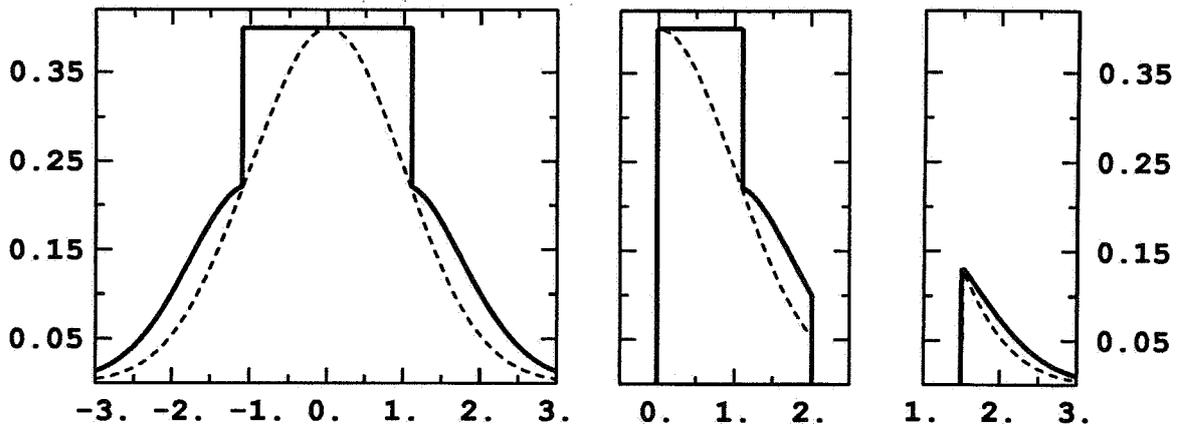
$$a \leq |x| \quad \begin{cases} e^{-\frac{x^2}{2}} \leq 1 \\ e^{-\frac{x^2}{2}} \leq \frac{|x|}{a} e^{-\frac{x^2}{2}} \end{cases}$$

Figure 3.2: Exemples de majoration de la densité de Gauss

a) $-\infty \leq X \leq +\infty$

b) $0 \leq X \leq 2$

c) $1.5 \leq X \leq +\infty$



--- densité de Gauss

— majoration

a) $-\infty \leq X \leq +\infty \quad f(x) = \frac{1}{\sqrt{2\pi}} (\mathbb{1}_{|x| < a} + \frac{|x|}{a} e^{-\frac{x^2}{2}} \mathbb{1}_{|x| > a}) \quad a = 1.09$

b) $0 \leq X \leq 2 \quad f(x) = \frac{1}{\sqrt{2\pi}} (\mathbb{1}_{0 < x < a} + \frac{|x|}{a} e^{-\frac{x^2}{2}} \mathbb{1}_{a < x < 2})$

c) $1.5 \leq X \leq +\infty \quad f(x) = \frac{1}{\sqrt{2\pi}} \frac{x}{1.5} e^{-\frac{x^2}{2}} \mathbb{1}_{1.5 < x}$

Il faut donc simuler une variable aléatoire X selon la densité majorante $h(\cdot)$ et U uniformément répartie sur $]0, 1[$. Le test $CUh(X) \leq f(X)$ se simplifie grandement et ne fait plus intervenir que des grandeurs aisément calculables. En effet, X appartenant à un intervalle, par exemple i_0 , l'inégalité à vérifier est équivalente à $Uf_{i_0}(X) \leq g(X)$. L'algorithme de simulation des gaussiennes tronquées scalaires par la méthode d'acceptation et rejets ⁴ est donc :

Algorithme 2 (i) *simulation de l'intervalle i , selon la loi $(\varpi_i)_i$,*

(ii) *simulation des variables aléatoires X et U de densité respective*

$$\frac{h_i \mathbb{1}_{I_i}(x)}{\int_{I_i} h_i(u) du} \quad \text{et} \quad \mathbb{1}_{]0,1[}(u),$$

(iii) *si $U h_i(X) \leq g(X)$, le candidat X est retenu. Sinon retour en (i).*

Comparaison pratique des méthodes de simulation d'une gaussienne tronquée scalaire

Les temps moyens d'exécution ont été calculés sur 10000 valeurs simulées à l'aide des méthodes par inversion et par acceptation et rejets⁵ (tableau 3.1) : lorsque la contrainte s'exprime à l'aide d'une seule inégalité, les deux méthodes de simulation ont des temps d'exécution comparables. Par contre si la contrainte est définie par deux inégalités, le temps d'exécution pour la méthode d'inversion est le double de celui de la méthode d'acceptation et rejets. Enfin, pour la méthode d'inversion, on rencontre des problèmes numériques — le calcul de $G^{-1}(0)$ ou $G^{-1}(1)$ — dès que l'on impose à la valeur absolue de la gaussienne d'être supérieure à 3. La méthode de simulation par inversion ne peut être utilisée dans des applications. Aussi, toutes les simulations de gaussiennes tronquées scalaires qui seront utilisées par la suite seront simulées par la méthode d'acceptation et rejets.

⁴Cet algorithme quelque peu modifié — les intervalles sont alors pondérés par les poids p_i , le système de probabilités $(\varpi_i)_i$ est remplacé par $(\varpi'_i)_i$ où $\varpi'_i = p_i \int_{I_i} f_i(u) du / \sum_j p_j \int_{I_j} f_j(v) dv$ — sera utilisé au chapitre 4 pour simuler une gaussienne conditionnée par sa poissonisée.

⁵Ces comparaisons ont été effectuées sur une station de travail en FORTRAN77 avec des réels déclarés en simple précision.

Tableau 3.1 : Temps d'exécution pour la simulation de 10000 gaussiennes scalaires tronquées
a) $X \geq x_c$

x_c	-2.5	-1.5	-0.5	0.0	0.5	1.5	2.5
acceptation et rejets	0.51	0.50	0.47	0.50	0.54	0.54	0.49
inversion	0.74	0.74	0.61	0.43	0.62	0.83	0.85

(t_{CPU} en secondes)

b) $a \leq X \leq b$

inégalités	$[-2, 2]$	$[-1, 1]$	$[0, 0.5]$	$[0, 1]$	$[0, 1.5]$
acceptation et rejets	0.56	0.49	0.46	0.49	0.53
inversion	1.26	1.02	0.83	0.83	0.96

(t_{CPU} en secondes)

Un vecteur gaussien tronqué

Pour simuler un vecteur gaussien tronqué par la méthode d'acceptation et rejets, une loi majorante acceptable est la loi gaussienne *a priori*. En effet on a trivialement,

$$\forall x \in \mathbb{R}^n \quad f(x) = \frac{g(x)\mathbb{1}_B(x)}{\int_B g(u)du} \leq \frac{1}{\int_B g(u)du} g(x)$$

Le test pour vérifier si le point généré $(X, C U g(X))$ appartient au sous-graphe de f est équivalent à $U \leq \mathbb{1}_B(X)$. Or U appartenant à $]0, 1[$ et $\mathbb{1}_B(X)$ ne prenant que les valeurs 0 ou 1, le résultat du test ne dépend que de la valeur de l'indicatrice : il n'est donc pas nécessaire de tirer la variable aléatoire uniformément répartie. L'algorithme consiste à tirer des vecteurs gaussiens indépendants jusqu'à ce que la valeur tirée vérifie la contrainte. Mais ceci peut être très long quand $\int_B g(u)du$ est très petit.

L'efficacité de cet algorithme peut être améliorée en utilisant pour loi majorante non pas celle de la gaussienne *a priori* mais la loi d'une gaussienne tronquée sur un ensemble A contenant B et que l'on sait simuler par une autre méthode. En effet si $B \subset A$,

$$\frac{g(x)\mathbb{1}_B(x)}{\int_B g(u)du} \leq \frac{\int_A g(v)dv}{\int_B g(u)du} \frac{g(x)\mathbb{1}_A(x)}{\int_A g(u)du}$$

Le nombre moyen d'essais est alors $\frac{\int_A g(v)dv}{\int_B g(u)du}$ qui est inférieur à $\frac{1}{\int_B g(u)du}$ si A est strictement inclus dans \mathbb{R}^n .

Le problème est la détermination de l'ensemble A , sur lequel il est possible de simuler un vecteur tronqué à l'aide des techniques présentées aux paragraphes précédents. Ces techniques ne s'appliquent qu'à des variables scalaires. Cependant, si la densité gaussienne ainsi que la contrainte se factorisent,

$$\begin{aligned} g(x) &= g(x_1) \quad g(x_2) \quad \cdots \quad g(x_n) \\ B &= B_1 \times B_2 \times \cdots \times B_n \end{aligned}$$

la densité du vecteur gaussien tronqué se factorise aussi

$$\frac{g(x)\mathbb{1}_B(x)}{\int_B g(u)du} = \frac{g(x_1)\mathbb{1}_{B_1}(x_1)}{\int_{B_1} g(u_1)du_1} \frac{g(x_2)\mathbb{1}_{B_2}(x_2)}{\int_{B_2} g(u_2)du_2} \cdots \frac{g(x_n)\mathbb{1}_{B_n}(x_n)}{\int_{B_n} g(u_n)du_n}$$

La simulation du vecteur se ramène alors à la simulation des composantes indépendamment les unes des autres. Ce cas très particulier correspond à la troncature d'un vecteur sans corrélation.

Par la suite, nous nous limiterons au cas où la contrainte est de la forme

$$B = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$$

Il faut donc rechercher un ensemble A sur lequel il sera possible de se ramener au cas des composantes indépendantes. Pour cela, on utilisera la représentation des composantes d'un vecteur gaussien comme des combinaisons linéaires de variables aléatoires gaussiennes indépendantes :

$$X \stackrel{L}{\equiv} LT$$

Σ , la matrice de covariance de X , étant strictement définie positive, les dimensions de X et T sont égales et la matrice carrée L peut être obtenue par factorisation de Σ sous la forme LL' ⁶. Deux factorisations classiques d'une matrice définie positive sont la décomposition de Cholesky et la décomposition par éléments propres⁷. Simuler X sous la contrainte $B = \{x \in \mathbb{R}^n, a_i \leq x_i \leq b_i\}$ est équivalent à simuler T sous la contrainte $B^* = \{t \in \mathbb{R}^n, a_i \leq \sum_j l_{ij}^2 t_j \leq b_i\}$. T tronqué sur B^* sera simulé par acceptation et rejets. La loi majorante est celle d'une gaussienne tronquée sur le plus petit hypercube contenant B^* :

$$A^* = [\alpha_1, \beta_1] \times [\alpha_2, \beta_2] \times \cdots \times [\alpha_n, \beta_n]$$

Les bornes de A^* sont déterminées par : $\alpha_i = \min_{(t_1, \dots, t_n) \in B^*} t_i$ et $\beta_i = \max_{(t_1, \dots, t_n) \in B^*} t_i$. Ces problèmes peuvent être résolus par la méthode du simplexe [8]. Mais dans le cas d'une décomposition de Cholesky, la géométrie très particulière du polyèdre A^* — la matrice L est triangulaire inférieure — peut être utilisée pour calculer de manière beaucoup plus efficace les bornes de A^* .

⁶Cette représentation est à la base des méthodes matricielles de simulation [15, 12, 11].

⁷ Σ étant une matrice symétrique et définie positive s'écrit sous la forme $\Sigma = P\Lambda P'$ où P est la matrice des vecteurs propres et Λ celle des valeurs propres. $L = P\Lambda^{\frac{1}{2}} = \Lambda^{\frac{1}{2}} = [\delta_{\alpha\beta}\sqrt{\lambda_\alpha}]_{(\alpha,\beta) \in A}$ — fournit donc la décomposition souhaitée.

L'algorithme de simulation du vecteur gaussien tronqué par la méthode d'acceptation et rejets est donc :

Algorithme 3 (i) Détermination de L et A^* .

(ii) Simulation de T tronqué sur A^* .

(iii) Si $X = LT \in B$, le candidat est retenu, sinon retour en (ii).

La simulation du couple bigaussien centré normé (X, Y) de covariance $\text{Cov}(X, Y) = 0.5$ sous la contrainte $B = [0, 1] \times [0, 1]$ permet d'illustrer cette méthode. Sur la figure 3.3, sont reportés les ensembles A^* et B^* obtenus par les deux types de factorisation de la matrice de covariance. Le nombre moyen de tirages en fonction de la loi majorante est 7.08 pour la loi gaussienne *a priori*, 1.77 pour la loi obtenue par factorisation de Cholesky et 1.42 pour la loi obtenue par la factorisation par éléments propres. Dans ce cas particulier, la décomposition par éléments propres conduit à un algorithme plus efficace.

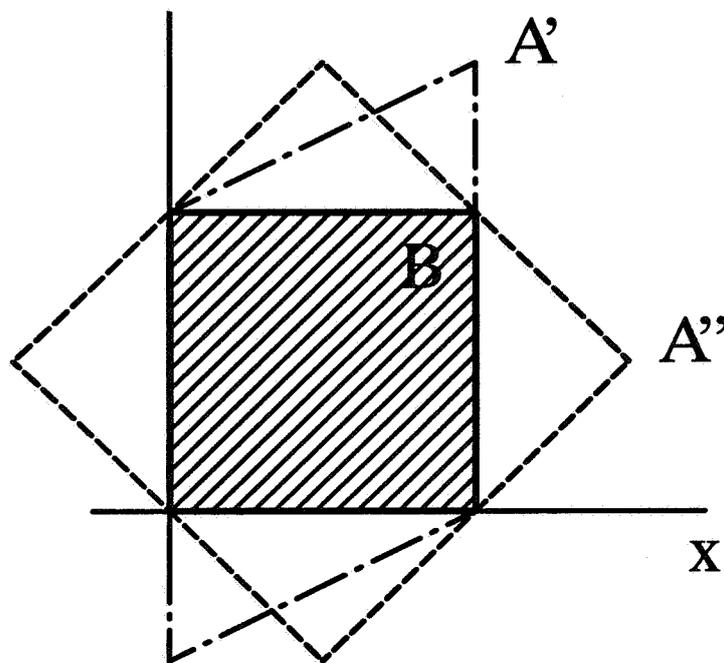


Figure 3.3 : Simulation d'un couple bigaussien tronqué par la méthode d'acceptation et rejets $(X, Y) \in [0, 1] \times [0, 1] \quad \text{Cov}(X, Y) = 0.5$

Les candidats sont simulés selon la densité $\frac{\mathbb{1}_A(x,y)g(x,y)}{\int_A g(u,v)dudv}$ où A est égal à

A' avec la décomposition de Cholesky et

A'' avec la décomposition par éléments propres.

$B \quad [0, 1] \times [0, 1]$

Dans la méthode par acceptation et rejets, la forme de la contrainte n'intervient que dans la détermination de A^* . On pourrait donc l'utiliser pour simuler des gaussiennes sous des contraintes plus générales, par exemple un polygone quelconque

$$B = \{x \in \mathbb{R}^n, k = 1, \dots, p \quad \sum_{1 \leq i \leq n} u_k^i x_i \leq v_k\}$$

Cependant, la dimension du vecteur à simuler limite le domaine d'utilisation de cette méthode, car le temps de calcul croît très rapidement. La simulation d'un vecteur de dimension n se décompose en trois étapes : détermination de la matrice L (I), détermination de la loi majorante (II), simulation et test des candidats (III). On leur associe les temps respectifs, t_I , t_{II} et t_{III} . Ces temps de calcul ont été estimés sur un exemple : le vecteur gaussien est un échantillon d'un processus stationnaire de covariance exponentielle. La matrice de covariance est donc,

$$1 \leq i, j \leq n \quad \text{Cov}(X_i, X_j) = \rho^{|i-j|}$$

Pour une coupure x_c , on simule tout d'abord l'indicatrice $(\mathbb{1}_{X_i \geq x_c})_{1 \leq i \leq n}$, puis le vecteur gaussien tronqué déterminé par cette indicatrice. Sur 10000 simulations, la moyenne des temps de calcul ainsi que le nombre d'essais ont été calculés pour les deux types de factorisation de la matrice de covariance ; les simulations des candidats ont été faites avec pour loi majorante ou bien la loi gaussienne *a priori* ou bien la loi d'une gaussienne tronquée (tableau 3.2).

Tableau 3.2: Temps d'exécution et nombre d'essais moyens pour 10000 simulations de vecteurs gaussiens tronqués par la méthode d'acceptation et rejets ($\rho = 0.7$ et $x_c = 0.0$)

temps de simulation : $t_I + t_{II} + t_{III}$

nombre d'essais : nt

n	Calcul de L		simulation			majoration			simulation			
	t_I		nt	t_{III}		prog. lin.		geom.	E.P.		Cholesky	
	E.P.	Cho.		E.P.	Cho.	E.P.	Cho.	Cho.	nt	t_{III}	nt	t_{III}
1	0.5	0.3	2	1.5	1.4	11.5	12.7	0.3	1.0	1.5	1.0	1.2
2	3.8	0.5	4	2.6	2.6	40.0	41.6	0.6	2.0	3.5	1.5	2.7
3	8.5	0.7	8	6.5	7.8	95.7	84.6	1.2	4.1	9.4	2.2	4.9
4	12.0	0.9	16	13.3	15.2	164.4	157.7	3.2	12.0	33.9	3.4	8.7
5	16.5	1.3	32	36.7	42.4	316.4	242.0	8.3	25.0	89.1	5.0	14.0
6	22.0	1.8	64	74.7	84.2	528.2	382.4	22.1	57.7	235.7	7.6	21.9

(t_{CPU} en secondes)

Sur l'exemple étudié ($\rho = 0.7$ et $x_c = 0.0$), la décomposition par éléments propres conduit à des temps de simulation beaucoup plus grands. Pour la décomposition de Cholesky, le temps moyen de simulation avec pour loi majorante la loi gaussienne *a priori* ou la loi d'une gaussienne tronquée est reporté à la figure 3.4. Dès que la taille du vecteur est supérieure à deux, la recherche d'une loi gaussienne tronquée permet d'économiser du temps de calcul. Cependant, l'allure exponentielle de cette courbe limite cette méthode à de petits vecteurs ($n \leq 10$).

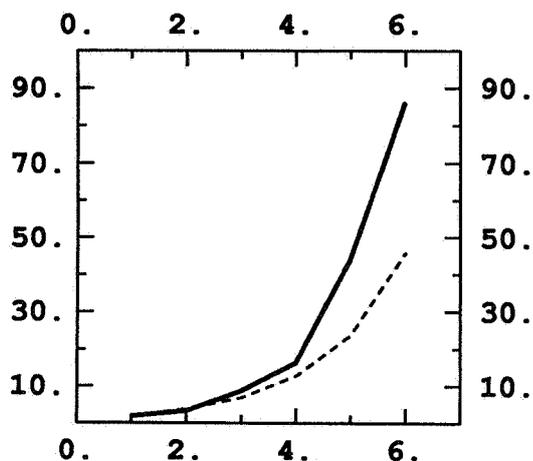


Figure 3.4 : Temps moyen de simulation d'un vecteur gaussien tronqué par la méthode d'acceptation et rejets ; décomposition de Cholesky

$$(\Sigma = [\rho^{|i-j|}]_{1 \leq i, j \leq n} \text{ avec } \rho = 0.7 \text{ et } x_c = 0.0)$$

Loi majorante: — loi gaussienne *a priori*
 ---- loi gaussienne tronquée

Les algorithmes présentés — simulation par inversion ou simulation par acceptation et rejets — permettent de simuler de manière exacte de petits vecteurs gaussiens tronqués ($n \leq 10$). Ils seront utilisés dans les algorithmes approximatifs présentés dans la partie suivante. Ces derniers sont utilisés pour simuler des vecteurs beaucoup plus grands.

3.2 Simulations approximatives

Les algorithmes de simulation étudiés dans cette partie fournissent, au lieu de la loi μ modélisée, une loi μ_ν dépendant d'un paramètre entier choisi par l'utilisateur, ν . On peut donc construire une suite $(\mu_\nu)_{\nu \in \mathbb{N}}$. L'algorithme proposé sera une approximation si la suite $(\mu_\nu)_{\nu \in \mathbb{N}}$ converge vers la loi à simuler μ . Généralement, la quantité de calculs nécessaires pour produire la loi μ_ν est une fonction croissante de ν : les moyens de calcul disponibles déterminent la précision de la simulation retenue.

Tout d'abord, deux méthodes sont présentées: la méthode incrémentale et la méthode itérative. Ne disposant par de critère théorique pour déterminer la précision de la simulation, celle-ci sera étudiée de manière empirique sur quelques exemples, dont on tentera de dégager certaines règles pratiques.

3.2.1 Méthode incrémentale

Le principe de la simulation d'un vecteur par la méthode incrémentale consiste à simuler chacune des composantes l'une après l'autre. Si le vecteur noté $X = (X_1, \dots, X_n)$ admet la densité $h(x_1, \dots, x_n)$, il faut simuler la première composante selon la loi marginale,

$$h_1(x) = \int h(x, u_2, \dots, u_n) du_2 \dots du_n$$

et la i -ème composante ($1 < i \leq n$) est simulée conditionnellement à toutes les composantes préalablement simulées, c'est-à-dire selon la densité

$$h_i(x | (x_j)_{1 \leq j \leq i}) = \frac{\int h(x_1, \dots, x_{i-1}, x, u_{i+1}, \dots, u_n) du_{i+1} \dots du_n}{\int h(x_1, \dots, x_{i-1}, u_i, u_{i+1}, \dots, u_n) du_i du_{i+1} \dots du_n}$$

La densité du vecteur ainsi simulé est bien h .

Dans le cas extrêmement favorable d'un vecteur multigaussien centré et sans contrainte, la mise en œuvre de la méthode incrémentale est particulièrement simple. En effet,

$$\begin{aligned} X_i &= X_i^{KS}(X_1, \dots, X_{i-1}) + \sigma_{KS}(i) U_i \\ &= \sum_{1 \leq j \leq i} l_j^i U_j \end{aligned}$$

où $X_i^{KS}(X_1, \dots, X_{i-1})$ est le krigeage simple de X_i à l'aide des $(X_j)_{1 \leq j \leq i}$, $\sigma_{KS}^2(i)$ la variance de krigeage et $(U_i)_{1 \leq i \leq n}$ des gaussiennes sans corrélation. On montre aisément que cette méthode est équivalente à la méthode de simulation matricielle utilisant la décomposition de Cholesky de la matrice de covariance. Mais dès que l'on impose une contrainte, on perd pour une grande part la simplicité des variables gaussiennes.

Les $i - 1$ premières composantes ayant été simulées, il reste à simuler un vecteur gaussien de dimension $n - i + 1$ avec la condition :

$$\bigcap_{j < i} \{X_j = x_j\} \cap \bigcap_{i \leq k \leq n} \{a_k \leq X_k \leq b_k\}$$

On peut toujours récrire les composantes du vecteur sous la forme,

$$\forall j \geq i \quad X_j = X_j^{KS} + \sigma_{KS}(j)U_j$$

où X_j^{KS} est le krigeage simple de X_j à partir des X_1, \dots, X_{i-1} et $(U_j)_{i \leq j \leq n}$ est un vecteur centré normé dont la matrice de covariance est

$$C = [c_{kl}]_{i \leq k, l \leq n} \text{ avec } c_{kl} = \text{Cov}(U_k, U_l) = \frac{\sigma_{kl} - \sum_{\alpha, \beta} \sigma_{k\alpha} R^{\alpha\beta} \sigma_{\beta l}}{\sigma_{KS}(k)\sigma_{KS}(l)}$$

L'inverse de la matrice $[\sigma_{\alpha\beta}]_{1 \leq \alpha, \beta \leq i}$ est noté ici $R = [R^{\alpha\beta}]_{1 \leq \alpha, \beta \leq i}$. Le problème revient à simuler un vecteur gaussien tronqué des résidus de dimension $n - i + 1$, U , sous la contrainte,

$$i \leq j \leq n \quad \frac{a_j - x_j^{KS}}{\sigma_{KS}(j)} \leq U_j \leq \frac{b_j - x_j^{KS}}{\sigma_{KS}(j)}$$

La dimension du vecteur des résidus, auquel on est ramené, reste généralement importante, aussi a-t-on défini un voisinage de la composante i : les inégalités sur les points n'appartenant pas au voisinage modifient peu la loi de la composante i . Le vecteur aléatoire considéré étant un échantillon d'une fonction aléatoire, le voisinage de la composante i est composé du point à simuler et de ses plus proches voisins. C'est une hypothèse similaire qui justifie l'usage du voisinage pour le krigeage.

Si le voisinage est suffisamment petit, le vecteur des résidus des points appartenant à ce voisinage peut être simulé par la méthode d'acceptation et rejets (voir page 50). Seule la composante i est simulée avec le voisinage idoine, aussi du vecteur simulé seule la valeur de la composante i est conservée. La densité de la composante ainsi simulée est donc :

$$\frac{\mathbb{1}_{B_i}(x) \int_{B_{i+1} \times \dots \times B_{i+d}} g(x, u_{i+1}, \dots, u_{i+d} | x_1, \dots, x_{i-1}) du_{i+1} \dots du_{i+d}}{\int_{B_i \times B_{i+1} \times \dots \times B_{i+d}} g(u_i, u_{i+1}, \dots, u_{i+d} | x_1, \dots, x_{i-1}) du_i du_{i+1} \dots du_{i+d}}$$

(d est la taille du voisinage, c'est-à-dire le nombre de points qu'il comporte) alors que la densité du modèle est :

$$\frac{\mathbb{1}_{B_i}(x) \int_{B_{i+1} \times \dots \times B_n} g(x, u_{i+1}, \dots, u_n | x_1, \dots, x_{i-1}) du_{i+1} \dots du_n}{\int_{B_i \times B_{i+1} \times \dots \times B_n} g(u_i, u_{i+1}, \dots, u_n | x_1, \dots, x_{i-1}) du_i du_{i+1} \dots du_n}$$

Lorsque la taille du voisinage grandit, la convergence de la loi du vecteur ainsi simulé vers la loi du modèle est assurée car, la dimension du vecteur aléatoire étant finie, un voisinage glissant de taille supérieure ou égale à la dimension du vecteur est équivalent à un voisinage unique. Cependant cette configuration est théorique, elle n'est pas celle de la pratique où la taille du voisinage (praticable) et celle du champ diffèrent d'un ordre de grandeur. Il sera donc nécessaire d'étudier empiriquement la convergence de la loi en fonction de la taille du voisinage des inégalités.

3.2.2 La méthode itérative

Dans la méthode itérative, on renonce à simuler directement la loi du modèle : une simulation initiale, non satisfaisante, est modifiée jusqu'à ce qu'une solution satisfaisante soit obtenue. Si les modifications successives ne dépendent que de l'état courant, la suite des simulations constitue une chaîne de Markov. De plus les modifications seront générées de sorte que la chaîne soit stationnaire. La convergence de la loi des simulations successives vers celle d'un vecteur gaussien tronqué est assurée si la probabilité de transition est ergodique. Ces techniques de simulation introduites par Metropolis et al. [41] en mécanique statistique sont également utilisées pour simuler des images modélisées par des champs markoviens [20].

On rappelle tout d'abord les chaînes de Markov à temps discret et à états continus dans le cas général. Puis on précise les chaînes de Markov utilisées : deux probabilités de transition sont étudiées ; leur mise en œuvre est aisée dans le cas des gaussiennes tronquées.

Rappels sur les chaînes de Markov à temps discret et états continus

Dans ce paragraphe, B désigne un sous-ensemble de \mathbb{R}^n de mesure non nulle et \mathcal{B}_B la tribu des Boréliens de B .

Une chaîne de Markov à valeurs dans B est un processus $(X^{(\nu)})_{\nu \in \mathbb{N}}$ tel que :

$$\forall \mathcal{A} \in \mathcal{B}_B, P\{X^{(\nu)} \in \mathcal{A} | X^{(0)}, \dots, X^{(\nu-1)}\} = P\{X^{(\nu)} \in \mathcal{A} | X^{(\nu-1)}\}.$$

La connaissance de l'état du phénomène à un instant donné apporte autant d'information sur le futur que la connaissance de tout le passé.

Si de plus, la probabilité $P\{X^{(\nu)} \in \mathcal{A} | X^{(\nu-1)}\}$ ne dépend pas de ν ($\nu > 0$), la chaîne est homogène et est définie par :

- une loi initiale p_0 , c'est-à-dire une probabilité sur (B, \mathcal{B}_B) , et
- une probabilité de transition, P .

Une probabilité de transition P est une application de $B \times \mathcal{B}_B$ dans \mathbb{R}^+ telle que :

- à x fixé, $x \in B$, $P(x, \cdot)$ est une probabilité sur (B, \mathcal{B}_B) ,
- à \mathcal{A} fixé, $\mathcal{A} \in \mathcal{B}_B$, $P(\cdot, \mathcal{A})$ est une fonction mesurable de (B, \mathcal{B}_B) .

$P(x, \mathcal{A})$ est la probabilité que $X^{(\nu)} \in \mathcal{A}$ conditionnelle à $X^{(\nu-1)} = x$:

$$P(x, \mathcal{A}) = P\{X^{(\nu)} \in \mathcal{A} | X^{(\nu-1)} = x\}.$$

La loi d'une chaîne homogène s'exprime par : $\mathcal{A}_0 \in \mathcal{B}_B, \dots, \mathcal{A}_\nu \in \mathcal{B}_B$,

$$P\{X^{(0)} \in \mathcal{A}_0, \dots, X^{(\nu)} \in \mathcal{A}_\nu\} = \int_{\mathcal{A}_0} dp_0(x^0) \int_{\mathcal{A}_1} dP(x^0, x^1) \int_{\mathcal{A}_2} \dots \int_{\mathcal{A}_\nu} dP(x^{\nu-1}, x^\nu)$$

En particulier, la loi de la ν -ème itération est donnée par :

$$P\{X^{(\nu)} \in \mathcal{A}\} = \int dp_0(u)P^{(\nu)}(u, \mathcal{A})$$

où P^ν désigne la probabilité de transition d'ordre ν , $P^{(\nu)}(x, \mathcal{A}) = P\{X^{(k+\nu)} \in \mathcal{A} | X^{(k)} = x\}$, qui est définie par :

$$\begin{aligned} P^{(\nu)}(x, \mathcal{A}) &= P(x, \mathcal{A}) && \text{si } \nu = 1 \\ &= \int dP(x, u)P^{(\nu-1)}(u, \mathcal{A}) && \text{si } \nu > 1 \end{aligned}$$

Si la loi de la ν -ème itération est indépendante de ν alors p — la probabilité initiale — est une probabilité stationnaire de P et vérifie alors :

$$\forall \mathcal{A} \in \mathcal{B}_B \quad p(\mathcal{A}) = \int dp(u)P(u, \mathcal{A}).$$

Une probabilité de transition est ergodique si quelle que soit la loi initiale p_0 , la suite formée par les lois des itérations successives converge vers une loi p strictement positive (c'est-à-dire pour tout ouvert de B , $p(\mathcal{A}) > 0$). Cette loi p est alors l'unique loi stationnaire de P .

a - Une chaîne de Markov

On voudrait construire une suite de vecteurs aléatoires $(X^{(\nu)})_{\nu \in \mathbb{N}}$ convergeant en loi vers X , vecteur gaussien tronqué sur B :

$$\lim_{\nu \rightarrow +\infty} X^{(\nu)} \stackrel{\mathcal{L}}{=} X$$

L'ensemble des états de ce processus est constitué des points de \mathbb{R}^n satisfaisant à la contrainte, c'est-à-dire B . Parmi tous les processus possibles, seules les chaînes de Markov ont été étudiées. En effet, ces processus présentent deux avantages :

- leur mise en œuvre ne nécessite pas le stockage de toutes les itérations, mais seulement de l'itération courante,
- ils sont définis par une loi de probabilité initiale, p , et une loi de probabilité de transition, P .

On a recherché une probabilité de transition ergodique P dont l'unique loi stationnaire soit la loi gaussienne tronquée sur B . Deux probabilités de transition sont étudiées — la probabilité de transition de Metropolis[41] et celle de Geman[20]. Toutes deux admettent pour loi stationnaire la loi gaussienne tronquée mais l'ergodicité n'est démontrée que dans le cas où la contrainte est compacte.

Si on dispose d'une probabilité de transition ergodique P , une simulation approximative de la loi gaussienne tronquée est obtenue en calculant les premiers termes de la chaîne de Markov. On arrête la chaîne dès que l'itération courante est jugée satisfaisante, c'est-à-dire dès que la loi n'est pas trop différente de la loi du modèle.

b - Une probabilité de transition

Une façon simple de modifier un vecteur est de modifier une et une seule de ses composantes. Là encore, la transition ne dépend que de l'état courant et on note P_i la probabilité de transition ne modifiant que la seule composante i . On modifie séquentiellement toutes les composantes du vecteur. Une itération est définie par l'ensemble de ces modifications. Sa probabilité de transition s'exprime à l'aide des probabilités de transition élémentaires : $\forall (x, \mathcal{A}) \in B \times \mathcal{B}_B$,

$$P(x, \mathcal{A}) = \int dP_1(x, u^{(1)}) \int dP_2(u^{(1)}, u^{(2)}) \int \dots \int dP_n(u^{(n-1)}, u^{(n)}) \mathbb{1}_{\mathcal{A}}(u^{(n)})$$

Si les probabilités de transition élémentaires P_i vérifient les équations⁸

$$1 \leq i \leq n \quad \forall \mathcal{A} \in \mathcal{B}_B \quad \int dx f(x) P_i(x, \mathcal{A}) = \int_{\mathcal{A}} f(x) dx \quad (3.1)$$

où f est la densité de la loi gaussienne tronquée sur B , la probabilité de transition P admet pour loi stationnaire la loi gaussienne tronquée.

Démonstration: Les équations 3.1 impliquent que pour tout fonction mesurable h ,

$$\int f(x) dx dP_i(x, y) h(y) = \int h(x) f(x) dx$$

Aussi pour la probabilité de transition de l'itération on a :

$$\begin{aligned} \int dx f(x) P(x, \mathcal{A}) &= \int f(x) dx dP_1(x, u^{(1)}) \int dP_2(u^{(1)}, u^{(2)}) \dots \int dP_n(u^{(n-1)}, u^{(n)}) \mathbb{1}_{\mathcal{A}}(u^{(n)}) \\ &= \int f(x) dx dP_2(x, u^{(2)}) \dots \int dP_n(u^{(n-1)}, u^{(n)}) \mathbb{1}_{\mathcal{A}}(u^{(n)}) \\ &= \int_{\mathcal{A}} f(x) dx \end{aligned}$$

La probabilité de densité f est donc une probabilité stationnaire de P . ■

Avant de présenter les deux probabilités de transition, on donne quelques notations et relations utilisées par la suite :

x est un élément de \mathbb{R}^n . Sa i -ème composante est notée x_i et le vecteur privé de sa i -ème composante $\bar{x}_i = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$. On écrira donc $x = (x_i, \bar{x}_i)$. Ces notations sont transposées aux vecteurs aléatoires : $X = (X_i, \bar{X}_i)$.

Si f est la densité de la gaussienne tronquée sur B , on notera $f(x_i)$ la densité de la i -ème composante, $f(\bar{x}_i)$ la densité du vecteur privé de sa i -ème composante et $f(x_i | \bar{x}_i)$ la densité conditionnelle de la i -ème composante \bar{x}_i étant fixé. D'après les formules 1.3 et 1.4 (page 13),

⁸La loi $f dx$ est une loi stationnaire des probabilités de transition élémentaires, mais ces probabilités de transition ne sont pas ergodiques.

on a donc :

$$\begin{aligned} f(x_i) &= \int f(x_i, \bar{u}_i) d\bar{u}_i \\ f(\bar{x}_i) &= \int f(u_i, \bar{x}_i) du_i \\ f(x_i|\bar{x}_i) &= \frac{f(x_i, \bar{x}_i)}{f(\bar{x}_i)} \end{aligned}$$

La dernière formule se simplifie grandement. En effet, la densité d'un vecteur multigaussien se factorise sous la forme :

$$g(x) = g(\bar{x}_i)g(x_i|\bar{x}_i) = g(\bar{x}_i) \frac{1}{\sigma_{KS}(i)} g\left(\frac{x_i - x_i^{KS}(\bar{x}_i)}{\sigma_{KS}(i)}\right) \quad (3.2)$$

où $x_i^{KS}(\bar{x}_i)$ est la valeur du krigeage simple de X_i à l'aide de \bar{X}_i et $\sigma_{KS}^2(i)$ la variance de krigeage.

Démonstration : i étant fixé, $X_i = X_i^{KS}(\bar{X}_i) + \sigma_{KS}(i) U_i$. U_i est une variable aléatoire gaussienne centrée normée. Les variables \bar{X}_i et U_i sont orthogonales donc indépendantes car le couple (\bar{X}_i, U_i) est multigaussien. La densité du couple est donc $g(\bar{x}_i)g(u_i)$. On effectue le changement de variable $(\bar{x}_i, u_i) \mapsto (x_i^{KS}(\bar{x}_i) + \sigma_{KS}(i) u_i)$ qui donne donc la formule 3.2. ■

La densité conditionnelle $f(x_i|\bar{x}_i)$ se récrit donc :

$$f(x_i|\bar{x}_i) = \frac{f(x_i, \bar{x}_i)}{f(\bar{x}_i)} = \frac{g(\bar{x}_i)g(x_i|\bar{x}_i) \mathbb{1}_B(x_i, \bar{x}_i)}{\int g(\bar{x}_i)g(u_i|\bar{x}_i) \mathbb{1}_B(u_i, \bar{x}_i) du_i} = \frac{g\left(\frac{x_i - x_i^{KS}(\bar{x}_i)}{\sigma_{KS}(i)}\right) \mathbb{1}_{B_i}(x_i) \frac{dx_i}{\sigma_{KS}(i)}}{\int_{B_i} g\left(\frac{u_i - x_i^{KS}(\bar{x}_i)}{\sigma_{KS}(i)}\right) \frac{du_i}{\sigma_{KS}(i)}} \quad (3.3)$$

1 - Probabilité de transition de Metropolis

L'idée de base de la probabilité de transition de Metropolis[41] consiste à décomposer la transition en une modification *arbitraire* suivie d'une validation — le nouvel état est accepté avec une certaine probabilité, sinon l'état initial est conservé — qui assure que la loi de densité f soit une probabilité stationnaire. Pour cela, on se donne une probabilité de transition dont la densité $\phi(y|x)$ est symétrique (c'est-à-dire $\phi(x|y) = \phi(y|x)$). La probabilité de transition de Metropolis est alors :

$$dP(x, y) = \min\left(1, \frac{f(y)}{f(x)}\right) \phi(y|x) dy + \left(1 - \int \min\left(1, \frac{f(u)}{f(x)}\right) \phi(u|x) du\right) \delta_x(y)$$

L'intérêt de cette méthode est qu'il n'est pas nécessaire de connaître une valeur de f , mais seulement le rapport $f(y)/f(x)$. De plus, malgré son aspect, cette probabilité de transition se simule aisément :

(i) On simule tout d'abord un vecteur y selon la densité conditionnelle $\phi(y|x)$.

(ii) Puis cette transition est validée avec la probabilité $\min(1, \frac{f(y)}{f(x)})$, sinon l'état initial est conservé.

Sur ce principe, on construit une probabilité de transition, P_i , ne modifiant que la composante i :

$$dP_i(x, y) = \delta_{\bar{x}_i}(\bar{y}_i) \left\{ \min(1, \frac{f(y)}{f(x)}) \phi_x^{(i)}(y_i) dy_i + (1 - \int \min(1, \frac{f(u_i, \bar{x}_i)}{f(x)}) \phi_x^{(i)}(u_i) du_i) \delta_{x_i}(y_i) \right\}$$

où $\phi^{(i)}$ est la densité d'une probabilité de transition telle que $\phi_{(x_i, \bar{x}_i)}^{(i)}(y_i) = \phi_{(y_i, \bar{x}_i)}^{(i)}(x_i)$ (donc symétrique) et que l'on sait simuler. La simulation de cette transition élémentaire se décompose en deux étapes:

(i) On simule y_i selon la densité $\phi_{(x_i, \bar{x}_i)}(\cdot)$.

(ii) Le vecteur $y = (y_i, \bar{x}_i)$ est validé avec la probabilité $\min(1, \frac{f(y)}{f(x)})$, sinon la valeur initiale $x = (x_i, \bar{x}_i)$ est conservée.

• f est la densité d'une probabilité stationnaire de P_i .

Démonstration: $\forall \mathcal{A} \in \mathcal{B}_B$,

$$\begin{aligned} \int dx f(x) P_i(x, \mathcal{A}) &= \int_{\mathcal{A}} dy f(y) + \\ &\int d\bar{x}_i \left[\int dx_i \min(f(x), f(y_i, \bar{x}_i)) \phi_x^{(i)}(y_i) \mathbb{1}_{\mathcal{A}}(y_i, \bar{x}_i) dy_i \right. \\ &\quad \left. - \int dx_i \min(f(x), f(y_i, \bar{x}_i)) \phi_x^{(i)}(y_i) \mathbb{1}_{\mathcal{A}}(x_i, \bar{x}_i) dy_i \right] \end{aligned}$$

La deuxième intégrale est nulle car $\phi_{(x_i, \bar{x}_i)}^{(i)}(y_i) = \phi_{(y_i, \bar{x}_i)}^{(i)}(x_i)$. ■

• Cette probabilité de transition est bien adaptée à la simulation d'un vecteur multigaussien tronqué, car le rapport des densités permet d'éliminer la constante de normalisation difficile, sinon impossible, à calculer. De plus comme seule la composante i est modifiée (formule 3.2) le rapport des densités devient extrêmement simple.

$$\frac{f(y)}{f(x)} = \mathbb{1}_{B_i} \frac{g(y_i | \bar{x}_i)}{g(x_i | \bar{x}_i)} = \mathbb{1}_{B_i}(y_i) \exp \left\{ -\frac{1}{2\sigma_{KS}^2(i)} [y_i^2 - x_i^2 - 2(y_i - x_i)x_i^{KS}(\bar{x}_i)] \right\}$$

• La densité symétrique utilisée est $\phi_x^{(i)}(\cdot) = h(x_i - y_i)$ où h est une densité symétrique que l'on sait simuler; le candidat à tester est dans ce cas

$$Y_i = X_i + \epsilon$$

où ϵ est une variable aléatoire de densité h . Les seules conditions sur h sont que h soit symétrique et qu'elle soit aisément simulable. Par la suite, cette perturbation centrée sera ou

bien une variable aléatoire uniforme ou bien une gaussienne ; la densité h sera donc spécifiée par sa variance.

A la ν -ème itération, la modification de la i -ème composante s'effectue de la manière suivante :

Algorithme 4

$$X_i^{(\nu+1)} = \begin{cases} X_i^{(\nu)} + \epsilon & \text{si } U \leq \min \left\{ 1, \mathbb{1}_{B_i}(X_i^{(\nu)} + \epsilon) \exp\left[-\frac{1}{2\sigma_{KS(i)}^2}(\epsilon^2 + 2\epsilon(X_i^{(\nu)} - X_i^{KS}))\right] \right\} \\ X_i^{(\nu)} & \text{sinon} \end{cases}$$

où

- X_i^{KS} est le krigeage simple de X_i par $(X_1^{(\nu+1)}, \dots, X_{i-1}^{(\nu+1)}, X_{i+1}^{(\nu)}, \dots, X_n^{(\nu)})$,
- ϵ et U sont des variables aléatoires de densité respective h et $\mathbb{1}_{]0,1[}$.

2 - Probabilité de transition de Geman S. et Geman D.

Le principe de cette probabilité de transition[20] est d'oublier la valeur initiale de la composante à modifier et de la remplacer par une valeur simulée selon sa loi conditionnelle. Cette probabilité de transition se construit donc à l'aide de la densité conditionnelle de la i -ème composante à \bar{x}_i fixé, $f(\cdot|\bar{x}_i)$:

$$dP_i(x, y) = \delta_{\bar{x}_i}(\bar{y}_i) f(y_i|\bar{x}_i) dy_i$$

- f est la densité d'une probabilité stationnaire de P_i .

Démonstration : $\forall \mathcal{A} \in \mathcal{B}_B$,

$$\begin{aligned} \int dx f(x) P_i(x, \mathcal{A}) &= \int dx f(x) \delta_{\bar{x}_i}(\bar{y}_i) dy_i f(y_i|\bar{x}_i) \mathbb{1}_{\mathcal{A}}(y) \\ &= \int \mathbb{1}_{\mathcal{A}}(y) \frac{f(x)}{f(\bar{x}_i)} f(y_i, \bar{x}_i) dy_i dx_i d\bar{x}_i \delta_{\bar{x}_i}(\bar{y}_i) \\ &= \int \mathbb{1}_{\mathcal{A}}(y) f(y_i, \bar{x}_i) f(x|\bar{x}_i) dy_i dx_i d\bar{y}_i \delta_{\bar{y}_i}(\bar{x}_i) \\ &= \int_{\mathcal{A}} dy \int f(y_i, \bar{x}_i) \delta_{\bar{y}_i}(\bar{x}_i) \int f(x|\bar{x}_i) dx_i \\ &= \int_{\mathcal{A}} dy \int f(y_i, \bar{x}_i) \delta_{\bar{y}_i}(\bar{x}_i) \\ &= \int_{\mathcal{A}} dy f(y) \end{aligned}$$

- Dans le cas d'un vecteur multigaussien tronqué, on a d'après la formule 3.3 : ■

$$dP_i(x, y) = \delta_{\bar{x}_i}(\bar{y}_i) \frac{g\left(\frac{y_i - x_i^{KS}(\bar{x}_i)}{\sigma_{KS(i)}}\right) \mathbb{1}_{B_i}(y_i) \frac{dy_i}{\sigma_{KS(i)}}}{\int_{B_i} g\left(\frac{u_i - x_i^{KS}(\bar{x}_i)}{\sigma_{KS(i)}}\right) \frac{du_i}{\sigma_{KS(i)}}$$

On peut récrire ceci en termes de variables aléatoires, ce qui conduit à l'algorithme de simulation suivant :

Algorithme 5

$$X_i^{(\nu+1)} = X_i^{KS} + \sigma_{KS}^2(i) U$$

où X_i^{KS} désigne le krigeage simple de X_i par $(X_1^{(\nu+1)}, \dots, X_{i-1}^{(\nu+1)}, X_{i+1}^{(\nu)}, \dots, X_n^{(\nu)})$ et U est une gaussienne scalaire tronquée sur $B_i^* = \{u \in \mathbb{R}, X_i^{KS} + \sigma_{KS}(i)u \in B_i\}$.

3 - Convergence vers la loi gaussienne

Les deux probabilités de transition élémentaires présentées ci-dessus admettent pour loi stationnaire la loi de la gaussienne tronquée. Les probabilités de transition d'une itération admettent donc elles aussi la loi gaussienne tronquée pour loi stationnaire. De ces deux probabilités de transition, la seconde est plus simple à mettre en œuvre dès que l'on dispose d'une méthode efficace pour simuler une gaussienne tronquée scalaire. De plus, on verra que la convergence empirique vers l'état stationnaire de la chaîne de Markov construite en utilisant cette probabilité élémentaire est plus rapide.

Si B est compact, la probabilité de transition P est ergodique et sa loi stationnaire est donc la loi gaussienne tronquée sur B .

Démonstration: Selon le théorème de la décomposition de Lebesgue [18], la probabilité de transition s'écrit comme la somme de deux mesures, l'une absolument continue et l'autre singulière — une mesure est singulière par rapport à la mesure de Lebesgue si elle est concentrée sur un ensemble de mesure nulle — par rapport à la mesure de Lebesgue :

$$P(x, \mathcal{A}) = \int_{\mathcal{A}} \psi(x, y) dy + \Delta_x(\mathcal{A})$$

Dans le cas de la probabilité de Metropolis, on peut choisir la loi des modifications — par exemple si la densité h ne s'annule pas sur \mathbb{R} — telle que la densité ψ soit continue et strictement positive; la partie singulière, due aux transitions élémentaires qui n'ont pas été validées, est une mesure ne comportant que des atomes. Quant à la probabilité de Geman et Geman, elle est absolument continue et sa densité ne s'annule pas. Si B est compact,

$$\inf_{(x,y) \in B \times B} \psi(x, y) = \min_{(x,y) \in B \times B} \psi(x, y) = \delta > 0$$

Doob [16] montre que dans ce cas la probabilité stationnaire est unique et telle que :

$$\forall (x, \mathcal{A}) \in B \times \mathcal{B}_B \quad |P^{(\nu)}(x, \mathcal{A}) - \int_{\mathcal{A}} f(u) du| < (1 - \delta|B|)^{\nu-1} .$$

■

Cependant, la valeur de δ est inconnue. D'autre part, si B n'est pas compact, je n'ai pu trouver une démonstration de l'unicité de la probabilité stationnaire. En pratique aucun cas pathologique n'a pu être mis en évidence. La convergence de ces méthodes de simulation a donc été étudiée empiriquement.

d - La simulation initiale

Pour simuler le vecteur initial, on peut négliger les corrélations entre les coordonnées qui peuvent alors être simulées indépendamment les unes des autres. En effet, si les corrélations entre les composantes sont négligées la densité gaussienne se factorise : $g(x) = \prod_{1 \leq i \leq n} g(x_i)$ et le numérateur comme le dénominateur de la densité de la loi tronquée se factorisent :

$$\frac{\mathbb{1}_B(x) g(x)}{\int_B g(u) du} = \prod_{1 \leq i \leq n} \frac{\mathbb{1}_{B_i}(x_i) g(x_i)}{\int_{B_i} g(u_i) du_i}$$

Une seconde méthode pour initialiser la chaîne est d'utiliser la méthode incrémentale avec un voisinage d'inégalités réduit au minimum. En procédant de la sorte, on ne tient pas compte des inégalités que doivent vérifier les composantes pas encore simulées du vecteur. Aussi le vecteur initial ainsi simulé n'est-il pas un vecteur gaussien tronqué.

3.3 Tests des algorithmes approximatifs

Les algorithmes présentés dans les parties précédentes permettent théoriquement de simuler un vecteur gaussien tronqué. Cependant la vitesse de convergence de ces algorithmes et donc les temps de calcul nécessaires n'ont pas été déterminés. L'étude empirique de ces éléments importants pour la mise en œuvre pratique est maintenant abordée. Il faudra tout d'abord choisir les exemples traités ainsi que la manière de les tester. Puis on détermine pour la méthode incrémentale et la méthode itérative respectivement la taille de voisinage et le nombre d'itérations nécessaires pour obtenir des simulations acceptables. Ceci permet d'évaluer le temps de calcul propre à chaque algorithme.

3.3.1 Tests des algorithmes

Afin de mesurer l'adéquation d'un modèle à un jeu de données, expérimental ou simulé, on calcule quelques statistiques — par exemple, moyenne, variance, histogramme et variogrammes. Dans le modèle, le jeu de données est interprété comme une réalisation d'un vecteur aléatoire, il en est de même pour les statistiques. Dans le cadre du modèle testé, on peut théoriquement calculer la loi des statistiques, ou au moins leurs deux premiers moments. La variance des statistiques calculées sur une simulation est généralement petite, le modèle est donc accepté si les valeurs des statistiques ne sont pas trop différentes de leur espérance.

Cette démarche va être utilisée pour tester les algorithmes de simulation des gaussiennes tronquées. Cependant la loi d'un vecteur gaussien tronqué ne permet pas un calcul explicite des deux premiers moments des statistiques élémentaires. On ne peut donc pas exploiter les statistiques calculées sur la simulation d'un vecteur gaussien tronqué pour juger des méthodes de simulation. Par contre si on *randomise* la contrainte, le vecteur tronqué redevient un banal vecteur multigaussien et un test est alors possible. On précise tout d'abord ce que l'on entend par randomisation de la contrainte.

On considère un échantillon de taille n d'une fonction aléatoire gaussienne, $X = (X_k)_{1 \leq k \leq n}$ et une famille d'intervalles de \mathbb{R} , $(B_k)_{1 \leq k \leq n}$ — par la suite, on se fixe un seuil x_c et les intervalles sont $B_k = [x_c, +\infty[$. Le vecteur gaussien tronqué dont la contrainte est randomisée se simule en deux étapes :

- (i) On simule tout d'abord un vecteur T de même loi que X . Ceci définit pour chaque composante une contrainte, $T_k \in B_k$ ou $T_k \in B_k^c$.
- (ii) On simule un vecteur gaussien tronqué vérifiant les mêmes contraintes.

Le vecteur ainsi simulé est un vecteur multigaussien. La fonction de \mathbb{R}^n dans $\{0, 1\}^n$, $I(u) = (\mathbb{1}_{B_\alpha}(u_\alpha))_{\alpha \in A}$, définit pour toute valeur i de $\{0, 1\}^n$ un sous-ensemble de $\mathbb{R}^n : I^{-1}(\{i\})$. On note la famille $(X_{(i)})_{i \in \{0, 1\}^n}$ des vecteurs gaussiens tronqués sur $I^{-1}(\{i\})$ dont la loi est donc

$$d\mu_i(x) = \frac{\mathbb{1}_{I^{-1}(\{i\})}(x)g(x)dx}{\int_{I^{-1}(\{i\})}g(u)du}$$

On définit d'autre part, le vecteur aléatoire à valeur dans $\{0, 1\}^n$ dont la loi est

$$\forall i \in \{0, 1\}^n \quad P\{I = i\} = \int_{I^{-1}(\{i\})}g(u)du$$

La loi de la randomisation de $(X_{(i)})_{i \in \{0, 1\}^n}$ par I est donc

$$\forall \mathcal{A} \in \mathcal{B}_{\mathbb{R}^n} \quad \sum_{i \in \{0, 1\}^n} P\{I = i\} \mu_i(\mathcal{A}) = \int_{\mathcal{A}} g(u)du$$

Le vecteur randomisé, X_I , est donc bien un vecteur multigaussien. Supposant correcte la simulation du vecteur T , donc la simulation de I , le test du modèle gaussien de X_I va permettre d'évaluer les algorithmes de simulation d'un vecteur gaussien tronqué.

3.3.2 Le cas d'un processus gaussien markovien

Tout d'abord, l'étude est réalisée sur un processus gaussien de covariance exponentielle : on simule le processus conditionné par la contrainte définie par une indicatrice préalablement simulée ; les trois méthodes présentées aux paragraphes précédents — la méthode incrémentale et les méthodes itératives avec pour probabilité de transition celle de Metropolis ou celle de Geman — sont utilisées. Quelques statistiques — moyenne, variance et variogrammes — sont calculées sur plusieurs simulations indépendantes, elles sont comparées avec les valeurs théoriques, celles du modèle gaussien *a priori* puisque l'indicatrice est simulée au préalable.

a - La méthode incrémentale

Le voisinage des inégalités considéré pour la simulation d'un point est constitué des plus proches voisins non encore simulés. On cherche donc à évaluer la taille du voisinage nécessaire pour reproduire les statistiques du modèle. Cette taille dépend de la covariance entre deux points voisins ρ et la valeur de la coupure x_c . Expérimentalement (figure 3.5), on constate

que pour une taille de voisinage réduite à un — seule l'inégalité portant sur le point simulé est prise en compte — la moyenne est satisfaisante tant que la valeur absolue de la coupure n'est pas trop importante. Par contre, la variance expérimentale est inférieure à la variance théorique — ceci correspond à un regroupement des valeurs autour du seuil visible sur l'histogramme calculé sur les simulations (figure 3.6) — et le variogramme (figure 3.7) n'est pas correctement rendu : au voisinage de l'origine on observe une rapide destructuration et le palier, correspondant à la variance expérimentale, est inférieur au palier théorique. Ces effets indésirables sont d'autant plus marqués que les corrélations sont fortes et que la valeur absolue de la coupure est petite (figure 3.5). Dans un cas défavorable ($\rho = 0.8$ et $x_c = 0.5$), la variance et les variogrammes sont reportés aux figures 3.8 et 3.7 pour différentes tailles du voisinage. On constate que la taille du voisinage nécessaire pour que les simulations deviennent acceptables croît avec la corrélation : un voisinage de deux points suffit lorsque la corrélation entre deux points adjacents vaut 0.6 mais cinq points sont nécessaires pour une corrélation de 0.8.

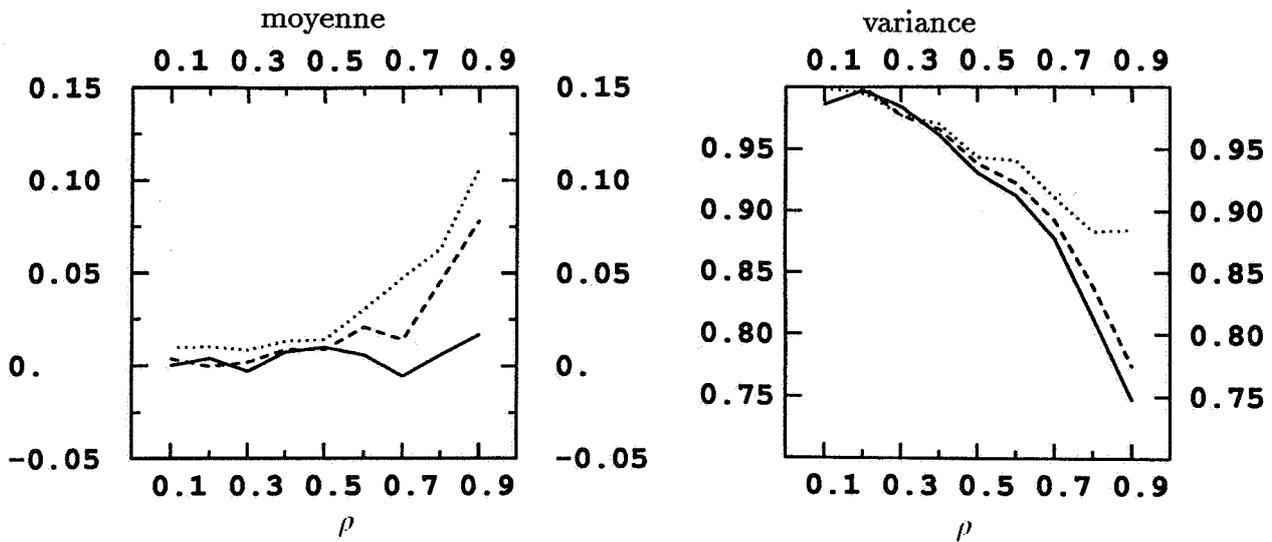


Figure 3.5: Moyenne et variance expérimentales de simulations produites par la méthode incrémentale avec une taille de voisinage égale à un (600 simulations de 100 points)

$x_c = 0.0$ —
 $x_c = 0.5$ ---
 $x_c = 1.0$

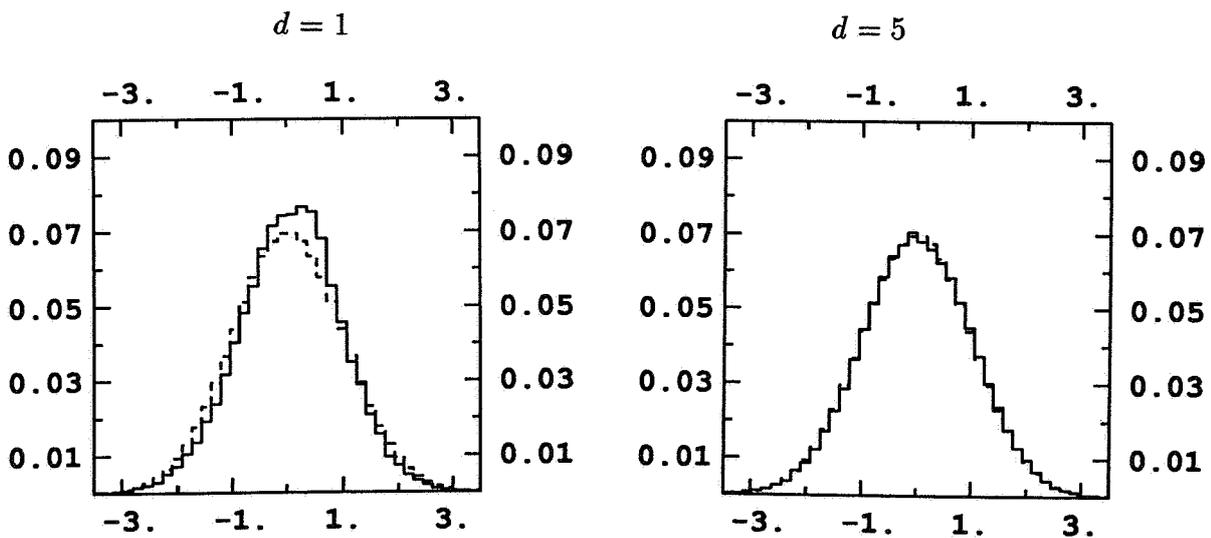


Figure 3.6: Histogramme de simulations produites par la méthode incrémentale pour deux tailles de voisinage (600 simulations de 100 points; $\rho = 0.8$ et $x_c = 0.5$)

--- histogramme théorique
 — histogramme expérimental

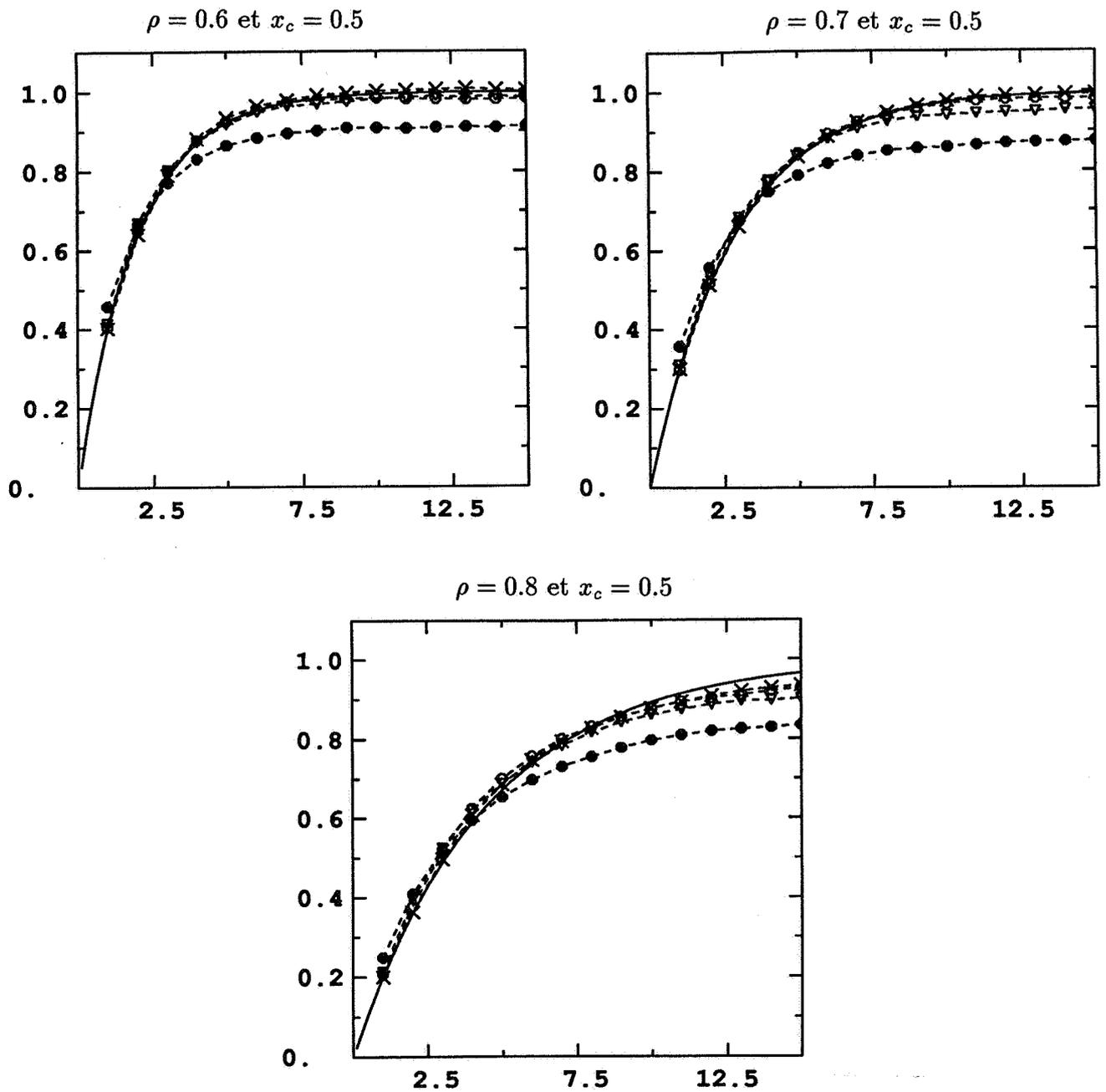


Figure 3.7: Variogrammes des simulations par la méthode incrémentale
 symbole taille du voisinage des inégalités

●	1
△	2
○	3
×	4

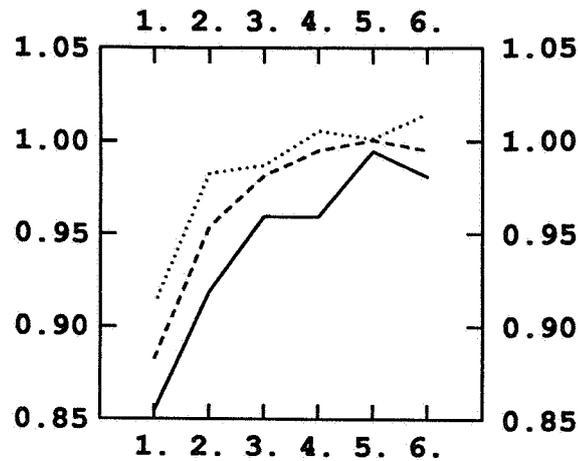


Figure 3.8: Variance des simulations produites par la méthode incrémentale pour différentes tailles de voisinage ($x_c = 0.5$)

$\rho = 0.6$
 0.8 ---
 0.9 ———

b - Les méthodes itératives

La méthode de Geman

La probabilité de transition de Geman est totalement spécifiée par la loi du vecteur à simuler. Il faut donc étudier l'influence de la simulation initiale sur la convergence de la suite des simulations. Pour cela on présente tout d'abord l'évolution de la variance expérimentale en fonction du rang de l'itération : la convergence de la variance expérimentale vers la variance du modèle est une condition nécessaire. Sur la figure 3.9, on constate que la simulation initiale n'influe guère sur le nombre des itérations nécessaires pour que la variance se stabilise. Par contre, la valeur des corrélations conditionne fortement la convergence : il faut cinq itérations lorsque ρ vaut 0.6 mais au moins 25 itérations sont nécessaires lorsque ρ vaut 0.8. Les variogrammes expérimentaux calculés sur les premières itérations (figure 3.10), confirment ces observations : les premiers points du variogramme sont correctement rendus dès les premières itérations, alors que le palier expérimental ne rejoint le palier théorique que très progressivement, en particulier lorsque les points simulés sont fortement corrélés. Or ce palier expérimental est lié à la variance expérimentale qui apparaît donc comme un bon indicateur de la convergence expérimentale.

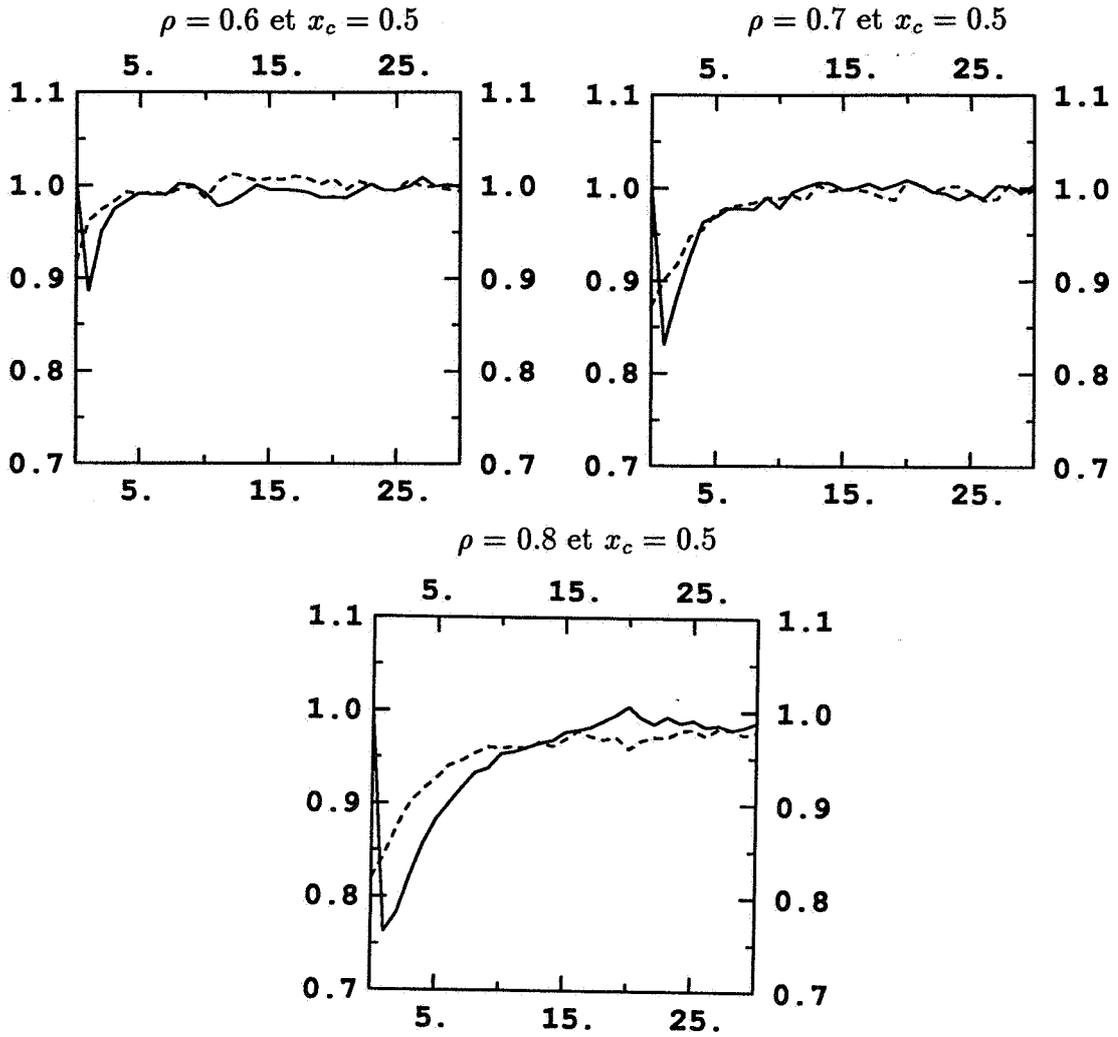


Figure 3.9: Variance expérimentale en fonction du rang de l'itération — Méthode de Geman simulation initiale

$\left\{ \begin{array}{l} \rho = 0.0 \text{ —} \\ \text{méthode incrémentale } d = 1 \text{ ---} \end{array} \right.$

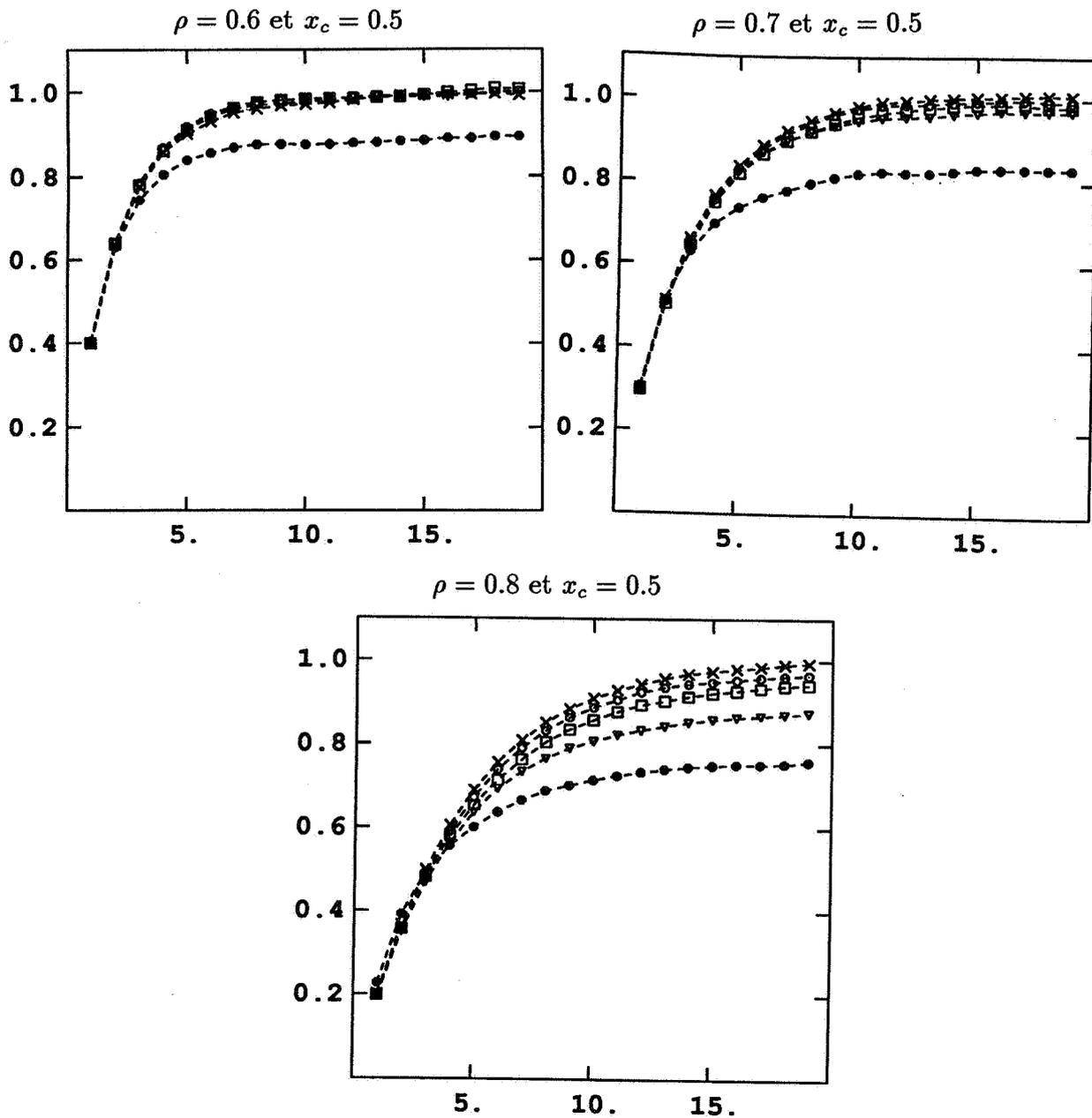


Figure 3.10: Variogrammes expérimentaux — Méthode de Geman
rang de l'itération symbole

1	•
5	▽
10	□
15	○
20	×

La méthode de Metropolis

La probabilité de transition de Metropolis utilisée dépend d'une loi de perturbation arbitraire — on impose simplement que sa densité conditionnelle soit symétrique. Deux types de perturbations ont été utilisés : un bruit additif indépendant gaussien ou uniforme ; tout deux sont définis à un facteur d'échelle près, égal à l'écart type.

De même que pour la méthode de Geman, la rapidité de la convergence des statistiques est très fortement liée aux corrélations entre les valeurs simulées. De plus, la variance des perturbations est prépondérante (figure 3.11) : il semble qu'il y ait une valeur optimale de cette variance. Si les perturbations sont trop faibles la convergence est lente, mais il en est de même lorsque les perturbations sont trop importantes car alors peu d'entre elles sont validées. Sur l'exemple présenté, l'optimum est obtenu pour une variance des perturbations égale à 1.

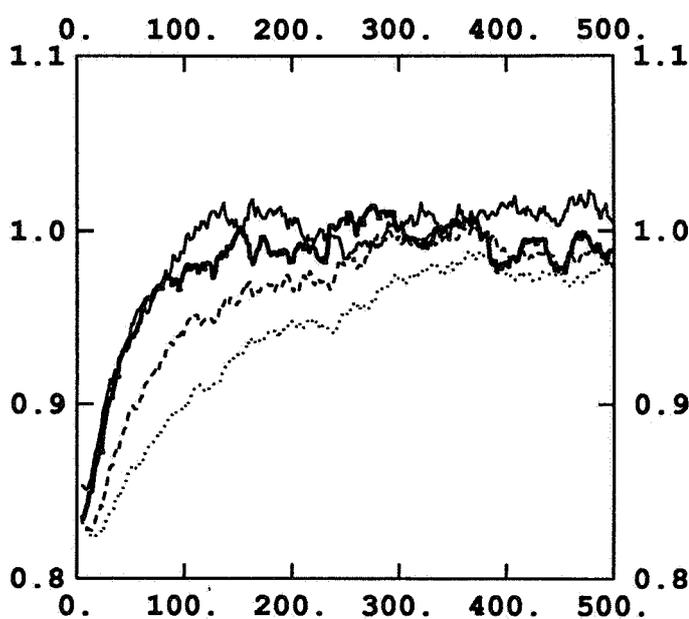


Figure 3.11 : Variance en fonction du temps de calcul — Méthode de Metropolis
($\rho = 0.8, x_c = 0.5$, perturbation uniforme)

courbe	variance des perturbations
.....	0.25
----	0.50
—	1.00
—	1.50

A la figure 3.12, sont reportées les variances en fonction du temps de calcul, pour les deux types d'initialisation et les deux types de perturbations, uniformes ou gaussiennes, lorsque

la variance des perturbations est égale à un. Les perturbations gaussiennes conduisent à une convergence légèrement plus lente en termes de temps de calcul — en termes du rang d'itération, les convergences sont très similaires — car la simulation d'une valeur gaussienne nécessite plus de calcul que la simulation d'une variable uniforme. Si la simulation ne tient pas compte des corrélations entre les valeurs à simuler, on observe tout d'abord une décroissance de la variance expérimentale ; puis la courbe croît pour atteindre la valeur théorique. Il semble donc que la chaîne de Markov de Metropolis la plus performante soit obtenue avec des perturbations uniformes et une simulation initiale calculée par la méthode incrémentale — le voisinage d'inégalité est réduit au seul point à simuler. Là encore, les premiers points du variogramme expérimental sont satisfaisants dès les premières itérations ; le palier expérimental suit les variations de la variance expérimentale.

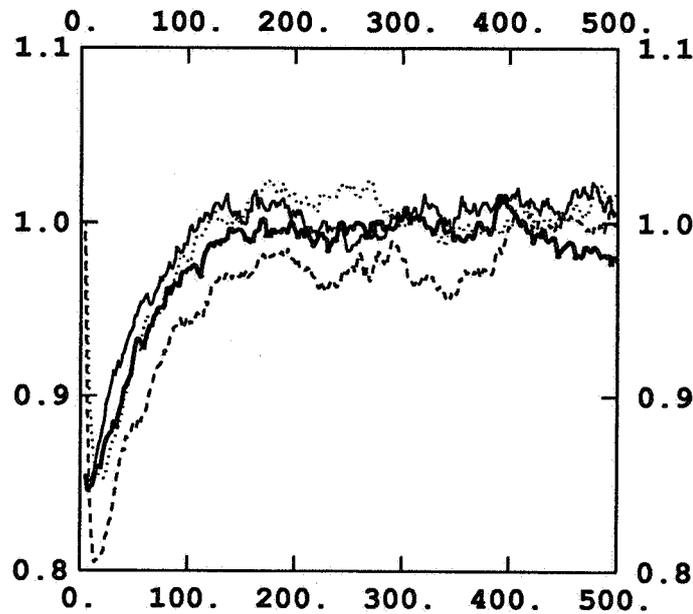


Figure 3.12: Variance en fonction du temps calcul — Méthode de Metropolis
($\rho = 0.8, x_c = 0.5$)

symbole	type des perturbations	simulation initiale
—	gaussien ($Var(\epsilon) = 1.0$)	méthode incrémentale
- - -	uniforme ($Var(\epsilon) = 1.0$)	méthode incrémentale
- - -	gaussien ($Var(\epsilon) = 1.0$)	$\rho = 0.0$
.....	uniforme ($Var(\epsilon) = 1.0$)	$\rho = 0.0$

Conclusion

Pour les trois méthodes, les courbes de la variance expérimentale en fonction du temps calcul sont reportées à la figure 3.13. Sur l'exemple présenté, la vitesse de convergence expérimentale par les trois méthodes est équivalente.

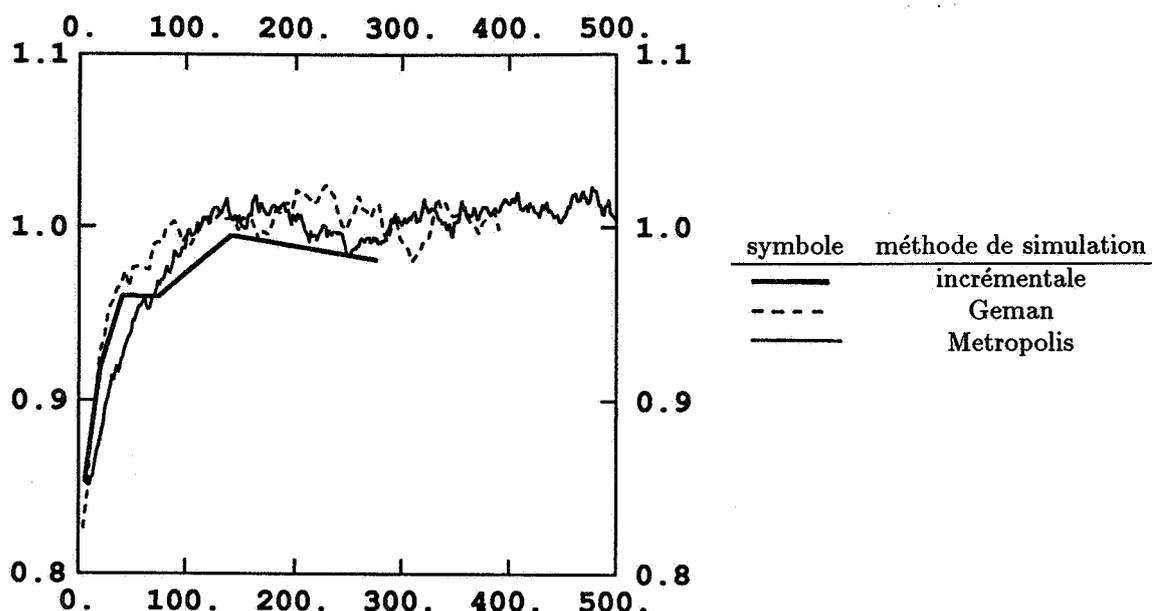


Figure 3.13: Variance en fonction du temps calcul ($\rho = 0.8$, $x_c = 0.5$)

Cependant, la méthode incrémentale est légèrement plus lente; de plus l'extension de la méthode incrémentale à \mathbb{R}^2 et \mathbb{R}^3 n'est guère possible. En effet, la taille du voisinage augmente très rapidement avec la dimension de l'espace. Or le temps de calcul nécessaire — ou le nombre moyen d'essais — pour simuler par la méthode d'acceptation et rejets croît de manière exponentielle en fonction du nombre de composantes de celui-ci. Par contre, les méthodes itératives permettent de simuler dans \mathbb{R}^2 ou \mathbb{R}^3 : toutes deux nécessitent le calcul en voisinage unique du krigeage et de sa variance en tout point à simuler. Ceci est possible tant que la taille du vecteur à simuler reste raisonnable. Dans ce cas, la relation matricielle suivante permet un calcul aisé de tous les éléments du krigeage simple nécessaires à la simulation: $\Sigma = [\sigma_{\alpha\beta}]_{(\alpha,\beta) \in A^2}$ est la matrice de covariance du vecteur à simuler. Pour une composante fixée, α_0 , la matrice Σ se réécrit classiquement sous la forme

$$\Sigma = \begin{pmatrix} \tilde{\Sigma}_{\alpha_0} & u_{(\alpha_0)} \\ u'_{(\alpha_0)} & \sigma_{\alpha_0\alpha_0} \end{pmatrix} \text{ où } \begin{cases} \tilde{\Sigma}_{\alpha_0} = [\sigma_{\alpha\beta}]_{\alpha,\beta \neq \alpha_0} \\ u_{(\alpha_0)} = [\sigma_{\alpha_0\beta}]_{\beta \neq \alpha_0} \end{cases}$$

L'inverse de Σ s'écrit alors [38]:

$$\Sigma^{-1} = \begin{pmatrix} \tilde{\Sigma}_{\alpha_0}^{-1} + \frac{\Lambda'_{(\alpha_0)} \Lambda_{(\alpha_0)}}{\sigma_{KS}^2(\alpha_0)} & -\frac{\Lambda_{(\alpha_0)}}{\sigma_{KS}^2(\alpha_0)} \\ -\frac{\Lambda'_{(\alpha_0)}}{\sigma_{KS}^2(\alpha_0)} & \frac{1}{\sigma_{KS}^2(\alpha_0)} \end{pmatrix} \text{ où } \begin{cases} \Lambda_{(\alpha_0)} = \tilde{\Sigma}_{\alpha_0}^{-1} u_{(\alpha_0)} \\ \sigma_{KS}^2(\alpha_0) = \sigma_{\alpha_0\alpha_0} - u'_{(\alpha_0)} \tilde{\Sigma}_{\alpha_0}^{-1} u_{(\alpha_0)} \end{cases}$$

Λ et $\sigma_{KS}^2(\alpha_0)$ sont donc les coefficients et la variance du krigeage simple de la composante α_0 à partir des autres composantes. Cette décomposition est valable pour tout α_0 . Aussi pour calculer tous les coefficients de krigeage, on inverse tout d'abord la matrice de covariance, $\Sigma^{-1} = B = [b^{\alpha\beta}]_{(\alpha,\beta) \in A^2}$, puis on calcule

$$\alpha \in A \quad \begin{cases} \sigma_{KS}^2(\alpha) = \frac{1}{b^{\alpha\alpha}} \\ \beta \neq \alpha \quad \lambda_\alpha^\beta = -\frac{1}{b^{\alpha\beta}} \end{cases}$$

Parfois, le voisinage unique se réduit de lui-même à un voisinage glissant restreint. Ceci est le cas de l'exemple présenté ci-dessus : le processus est markovien, le voisinage de krigeage est réduit aux deux plus proches voisins, de part et d'autre, du point à simuler. A deux ou trois dimensions, on a des propriétés similaires lorsque la covariance des gaussiennes est une exponentielle factorisée :

$$\rho(h) = \exp\left\{-\left(\frac{|h_x|}{a_x} + \frac{|h_y|}{a_y} + \frac{|h_z|}{a_z}\right)\right\}$$

où h_x , h_y et h_z sont les composantes du vecteur h selon les trois axes. Le voisinage de krigeage est alors constitué des plus proches voisins. Les propriétés de cette covariance sont utilisées pour simuler des réservoirs pétroliers[40]. Dans le cas général, on peut limiter la taille du voisinage de krigeage. En effet, si la covariance des gaussiennes tend vers zéro pour de grandes distances, les pondérateurs du krigeage simple tendent également vers zéro lorsque la distance au point à kriger augmente. Procéder de la sorte revient à substituer à la matrice de covariance du modèle une approximation de celle-ci. $(\lambda_\alpha^\beta)_{\beta \neq \alpha}$ sont les coefficients du krigeage simple de α à partir des points du voisinage V_α (si $\beta \notin V_\alpha$, $\lambda_\alpha^\beta = 0$) et σ_α^2 la variance de krigeage. Le vecteur effectivement simulé est le modèle auto-normal [2] défini par ses densités conditionnelles :

$$\alpha \in A \quad g(x_\alpha | (x_\beta)_{\beta \neq \alpha}) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_\alpha} \exp\left\{-\frac{1}{2} \left(\frac{x_\alpha - \sum_\alpha \lambda_\alpha^\beta x_\beta}{\sigma_\alpha}\right)^2\right\}$$

Si $\lambda_\alpha^\beta / \sigma_\alpha^2 = \lambda_\beta^\alpha / \sigma_\beta^2$, la densité du vecteur est

$$g(x) = \frac{1}{(2\pi)^{\frac{n}{2}}} |B_*|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} x' B_* x\right\}$$

où la matrice B_* est définie par

$$b_*^{\alpha\beta} = \begin{cases} \frac{1}{\sigma_\alpha^2} & \text{si } \alpha = \beta \\ -\frac{\lambda_{\alpha\beta}}{\sigma_\alpha^2} & \text{sinon} \end{cases}$$

En voisinage unique, on retrouve bien $B_*^{-1} = \Sigma$. Par contre le voisinage glissant n'assure ni la symétrie $\lambda_\alpha^\beta / \sigma_\alpha^2 = \lambda_\beta^\alpha / \sigma_\beta^2$ — lorsque la simulation s'effectue sur une grille régulière, on peut choisir des coefficients de krigeage constants, mais ceci fait apparaître des effets de bords — ni la condition de positivité de la matrice B . Cependant en pratique, l'utilisation du voisinage glissant permet de reproduire correctement les covariances modélisées.

Sur l'exemple présenté plus haut, la méthode de Metropolis nécessite plus d'itérations que la méthode de Geman pour atteindre l'état stationnaire. Par contre, le calcul d'une itération par la première méthode est plus rapide que par la seconde. En effet la modification d'une composante se décompose en deux termes : le calcul du krigeage et la simulation du résidu. Or sur l'exemple, l'étape importante — en temps calcul — est la simulation du résidu : la simulation et le test d'une perturbation uniforme sont beaucoup plus rapides que la simulation d'une gaussienne tronquée. Par contre si le voisinage de krigeage est grand — ce qui est le cas lorsque l'on travaille à deux ou trois dimensions avec une covariance quelconque — l'étape importante est le calcul du krigeage. La méthode de Metropolis est alors plus lente que la méthode de Geman.

Un deuxième argument joue en défaveur de la méthode de Metropolis : la variance des perturbations est un paramètre important de la probabilité de transition, pour lequel on ne dispose pas de détermination théorique. Il faut donc procéder à des essais préalables pour la déterminer expérimentalement.

Par suite, on utilisera donc la méthode de Geman pour simuler des exemples de gaussiennes tronquées. La courbe *rang de l'itération - variance expérimentale* permet de contrôler la convergence vers la loi conditionnelle à simuler.

Pour les deux méthodes itératives, l'initialisation peut s'effectuer ou bien en négligeant les corrélations entre les composantes ou bien par la méthode incrémentale, le voisinage des inégalités étant réduit au point à simuler. La seconde méthode est plus compliquée, il faut calculer de multiples krigeages, et ne permet pas d'accélérer la convergence expérimentale. C'est pourquoi par la suite les chaînes seront initialisées en négligeant les corrélations.

3.3.3 Trois exemples de simulation 2D par la méthode de Geman

La méthode de Geman est utilisée pour simuler des gaussiennes tronquées sur une grille régulière. Trois exemples sont présentés : une covariance exponentielle factorisée, une covariance exponentielle isotrope — ces deux covariances exponentielles coïncident sur les axes mais diffèrent ailleurs ; la corrélation entre deux pixels adjacents est 0.8 — et une covariance sphérique isotrope de portée 15. Une image 200×200 est simulée sous la contrainte $\mathbb{1}_{X \geq 0.67}$ (la proportion théorique des points de la grille appartenant à l'ensemble simulé est donc 0.25) par la méthode de Geman ; l'image initiale est calculée en négligeant les corrélations. La chaîne est arrêtée lorsque la variance expérimentale se stabilise (figure 3.14) : 200 itérations sont nécessaires dans le premier cas, 100 dans le second et 300 dans le dernier. Pour le premier modèle, le voisinage se réduit au huit plus proches voisins. Pour les deux derniers modèles, un voisinage glissant 9×9 est utilisé. Dans tous les cas, le variogramme expérimental correspond au modèle à simuler, ceci en dépit de l'approximation introduite par le voisinage glissant. Les images et variogrammes sont reproduits à la figure figure 3.15.

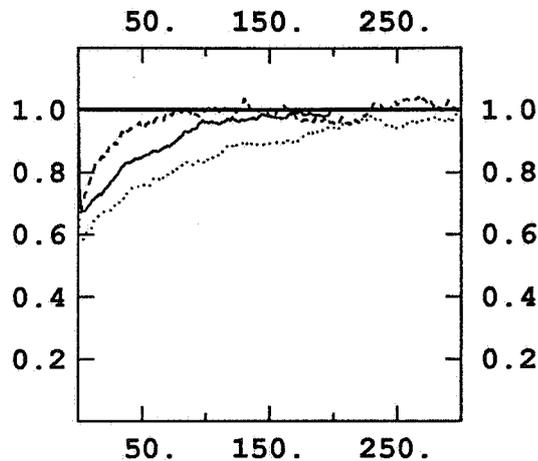
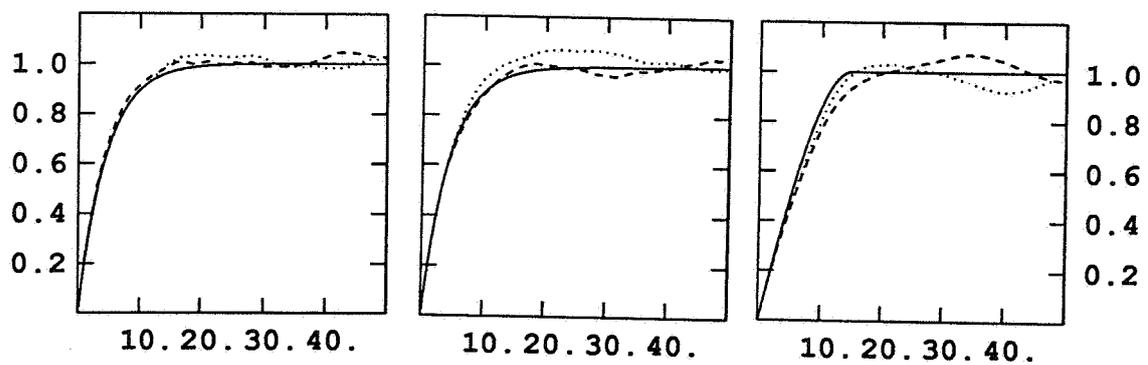
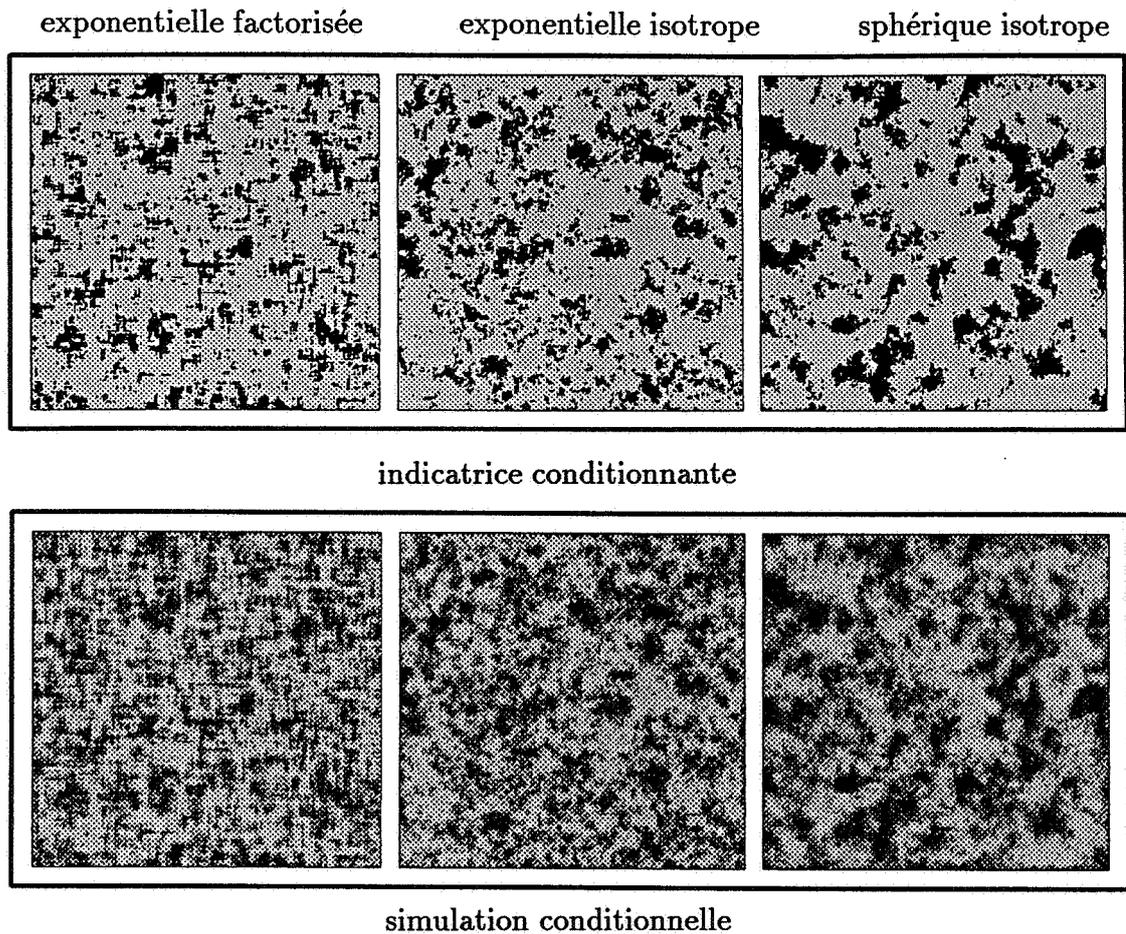


Figure 3.14: Variance en fonction du rang de l'itération

- : covariance exponentielle factorisée
- - - : covariance exponentielle isotrope
- : covariance sphérique isotrope

Figure 3.15: Simulations conditionnées par l'indicateur $\mathbb{1}_{X \geq 0.67}$



variogrammes expérimentaux des simulations conditionnelles

- : modèle
- : variogramme expérimental horizontal
- : variogramme expérimental vertical

3.4 Simulation et estimation

Disposant de plusieurs simulations conditionnelles, il est possible de calculer une approximation de l'espérance conditionnelle. Les méthodes itératives sont bien adaptées à la résolution de tels problèmes. En effet, bien que deux simulations successives ne soient pas indépendantes, on a la convergence presque sûre :

$$\lim_{\nu \rightarrow +\infty} \frac{1}{\nu} \sum_{i=1}^{\nu} \tilde{X}^{(i)} = \mathbf{E}\{X|X \in B\}$$

Cette utilisation des simulations conditionnelles pour calculer une espérance conditionnelle empirique est illustrée sur un exemple très simple ; elle sera plus amplement exposée au chapitre suivant : l'indicatrice $\mathbb{1}_{X \geq 0.67}$ d'un processus markovien ($\rho = 0.8$) est observée. A partir de ces données expérimentales, on voudrait estimer le processus lui-même ou une fonction de celui-ci. Disposant de la covariance simple de l'indicatrice et la covariance croisée entre l'indicatrice et la variable à estimer, on peut estimer par un cokrigeage simple — dans la mesure où le krigeage disjonctif est un cokrigeage simple à partir de toutes les indicatrices, la procédure utilisée ici peut être interprétée comme un *sous-krigeage* disjonctif car seule ici une indicatrice est disponible — ou bien par l'espérance conditionnelle. Trois fonctions du processus ont été estimées : une transformation linéaire, $y = x$, une transformation cubique, $y = x^3 + 3x^2 + 3x + 57$, et une autre indicatrice, $\mathbb{1}_{X \geq 0.0}$.

A la figure 3.16 sont reportées la variable à estimer et les deux estimations :

- a) L'estimateur de la gaussienne par co-krigeage et celui par espérance conditionnelle empirique donnent des résultats très similaires.
- b) Pour l'estimation de la fonction cubique, on observe un léger décalage entre les deux estimateurs pour les faibles valeurs. L'espérance conditionnelle empirique tient mieux compte de la dissymétrie de l'histogramme de la variable à estimer.
- c) L'estimateur de l'indicatrice par espérance conditionnelle empirique est comprise entre 0 et 1. Exceptés les points où le co-krigeage simple ne respecte pas ces inégalités, on observe peu de différences entre les deux estimateurs.

Dans tous les cas, les différences entre le co-krigeage simple et l'espérance conditionnelle empirique sont minimales. En modélisant correctement les covariances simples et croisées de la variable observée et la variable à estimer, un estimateur linéaire fournit d'excellent résultats.

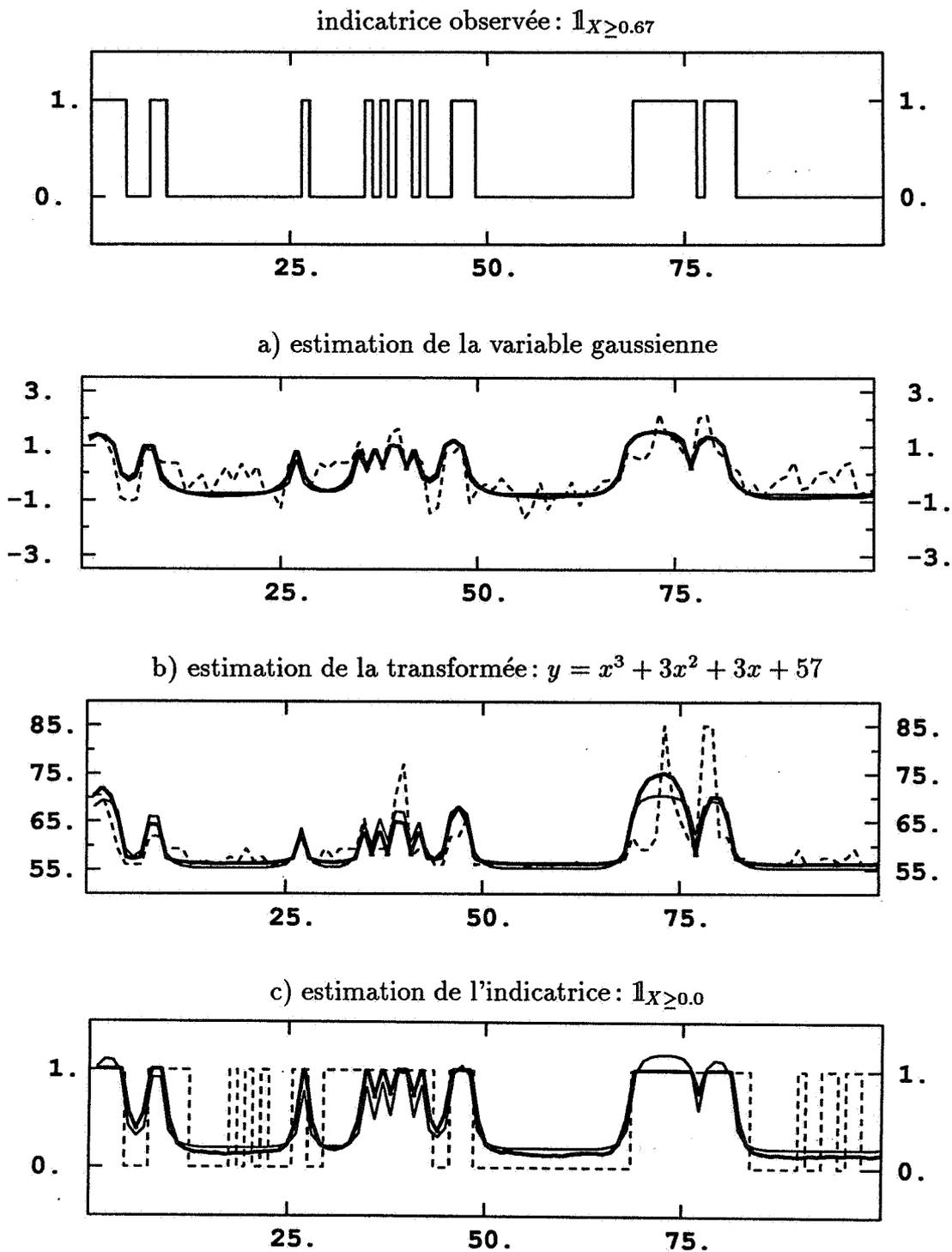


Figure 3.16 : Estimations à partir de l'indicateur $\mathbb{1}_{X \geq 0.5}$

- : variable à estimer
- : estimation par co-krigeage simple
- : estimation par espérance conditionnelle empirique



Chapitre 4

Simulations et filtrage

Dans les chapitres qui précèdent, nous nous sommes efforcés de construire des simulations conditionnelles contraintes par des inégalités ou des égalités, sans beaucoup nous préoccuper de l'utilisation qui en est faite. L'objet de ce chapitre est double : on présente une méthode pour construire des simulations gaussiennes conditionnées par des variables poissonisées, puis ces simulations sont utilisées pour calculer une approximation de l'espérance conditionnelle dite espérance conditionnelle empirique. Les résultats sont comparés avec deux estimateurs classiques de la géostatistique, le krigeage simple et le krigeage disjonctif.

Comme on l'a rappelé au chapitre 1, de nombreux bruits peuvent être modélisés comme une randomisation de la variable régionalisée étudiée. On se restreint ici au cas d'une poissonisation. La variable mesurée est alors modélisée par une variable aléatoire de Poisson dont le paramètre est la variable d'intérêt. Ce cas est illustré à l'aide d'images microsondes : un point d'un échantillon métallique est bombardé par un faisceau d'électrons ; pour une fréquence caractéristique d'un élément, le nombre de photons émis est modélisé par une variable aléatoire de Poisson dont le paramètre est la concentration de l'élément étudié. C. Daly [10] a montré que l'estimation de la variable sous-jacente peut être résolue de façon très satisfaisante par des techniques linéaires, telles que le krigeage simple ou le krigeage ordinaire mais que l'estimation d'une indicatrice, $\mathbb{1}_{y(u) \geq y_c}$, ou de sa régularisée, l'*histogramme local*,

$$\frac{1}{|V_x|} \int_{V_x} \mathbb{1}_{y(u) \geq y_c} du$$

nécessite une modélisation de la loi bivariable. Le krigeage disjonctif fournit alors un estimateur non biaisé de l'histogramme local. Pour ce type de bruit, une approche bayésienne est possible : la loi *a priori* de l'image sous-jacente et la modélisation du bruit par une randomisation détermine la loi conditionnelle de l'image. Sur cette loi conditionnelle sont construits des estimateurs maximisant la vraisemblance [20, 3] ou minimisant la variance d'estimation [4]. Cette dernière approche sera utilisée ici : on calculera l'espérance conditionnelle empirique de la variable sous-jacente et de son histogramme local.

Tout d'abord, il est nécessaire de présenter le modèle d'image qui a été utilisé. Puis la technique de simulation conditionnelle étudiée et les approximations effectuées sont précisées.

Enfin, la méthode sera validée sur des données simulées, pour être appliquée à un jeu de données réelles.

4.1 Un modèle d'image

Le modèle d'image utilisé par la suite est constitué, d'un part, par la modélisation du bruit et d'autre part, par la modélisation de la variable sous-jacente à estimer. Les mesures sont effectuées sur une grille régulière, elles sont indexées sur l'ensemble A . On considère les deux vecteurs aléatoires Z et Y indexés tous deux sur A : $Z = (Z_\alpha)_{\alpha \in A}$ correspond aux mesures bruitées et $Y = (Y_\alpha)_{\alpha \in A}$ correspond à la variable sous-jacente.

4.1.1 Bruit poissonien

L'image observée est modélisée comme la poissonisée — c'est-à-dire la randomisée ¹ par une variable de Poisson — de l'image sous-jacente. On suppose donc que pour tout pixel α , y_α étant fixé, Z_α est indépendante des autres mesures et des autres variables sous-jacentes, $(Z_\beta, Y_\beta)_{\beta \neq \alpha}$, et que la loi conditionnelle de l'observée, Z_α , est une loi de Poisson ayant comme paramètre la valeur, y_α , de la variable sous-jacente. La loi conditionnelle du pixel α est donc :

$$P\{Z_\alpha = z_\alpha | y_\alpha\} = f(z_\alpha | y_\alpha) = e^{-y_\alpha} \frac{(y_\alpha)^{z_\alpha}}{z_\alpha!}$$

L'indépendance conditionnelle des observations connaissant l'image sous-jacente entraîne la factorisation de la loi conditionnelle de Z :

$$P\{Z = z | y\} = f(z | y) = \prod_{\alpha \in A} f(z_\alpha | y_\alpha)$$

4.1.2 Modèle gaussien anamorphosé de l'image sous-jacente

L'image sous-jacente est modélisée comme la transformée d'une image multigaussienne par une fonction d'anamorphose. On suppose donc qu'il existe une fonction aléatoire stationnaire, X , multigaussienne, centrée, normée de covariance ρ , et une fonction d'anamorphose ϕ , élément de $L^2(\mathbb{R}, g)$ telles que :

$$\forall \alpha \in A \quad Y_\alpha = \phi(X_\alpha)$$

Les paramètres de ce modèle sont donc d'une part la loi marginale de la concentration sous-jacente, dont on déduit la fonction d'anamorphose, et d'autre part la covariance des gaussiennes. On rappelle tout d'abord comment ces paramètres peuvent être inférés à partir d'une réalisation de l'image bruitée. Puis, on cherche à construire une fonction des données expérimentales minimisant la variance d'estimation, c'est-à-dire l'espérance conditionnelle. L'expression de cet estimateur ne peut être utilisée pour un calcul direct. Une approximation est obtenue en calculant la moyenne de plusieurs simulations conditionnelles.

¹Une définition de la randomisée est donnée page 16 ou dans Feller [18, pages 52–53]

4.1.3 Inférence des paramètres du modèle

Les paramètres du modèle sont la loi marginale de Y , ou la fonction de transformation ϕ ainsi que la covariance de la fonction aléatoire gaussienne ρ . C. Daly[10] a montré qu'il est possible d'inférer ces paramètres sur une réalisation :

- L'estimation par moindres carrés de la loi sous-jacente à partir de la loi expérimentale est un problème mal posé. Cependant une solution acceptable est obtenue par régularisation. On obtient un estimateur de la loi de Y sous la forme,

$$d\hat{F}(y) = \sum_i \hat{p}_i \delta_{y_i}(y)$$

où \hat{p}_i est la fréquence estimée associée à la valeur y_i . La fonction d'anamorphose empirique [29] s'en déduit par

$$\phi(x) = \sum_i y_i \mathbb{1}_{[a_i, a_{i+1}[}(x)$$

avec $a_i = G^{-1}(\sum_{j < i} \hat{p}_j)$. Par la suite, la fonction ϕ sera connue pour les données simulées ; pour l'image micro-sonde, on a retenu la fonction d'anamorphose obtenue par C. Daly [10].

- Sur l'image bruitée, on calcule le variogramme expérimental,

$$\hat{\gamma}(h) = \frac{1}{N(h)} \sum_{\alpha-\beta=h} \frac{1}{2} (Z_\alpha - Z_\beta)^2$$

dont l'espérance est :

$$h \neq 0 \quad \mathbb{E}\{\hat{\gamma}(h)\} = m + C(0) - C(h)$$

où m est l'espérance de Y et C sa covariance. Z étant la poissonisée de Y , on a $\text{Cov}(Z_\alpha, Z_\beta) = m \delta_{\alpha\beta} + \text{Cov}(Y_\alpha, Y_\beta)$. C étant la covariance de la transformée par la fonction ϕ d'une fonction aléatoire de lois bivariées gaussiennes, ne peut être ajustée par une fonction de covariance quelconque [39]. En effet, ϕ , élément de $L^2(\mathbb{R}, g)$, s'écrit sous la forme,

$$\phi = \sum_0^{+\infty} \phi_n \eta_n$$

où les η_n sont les polynômes d'Hermite normés. On a donc les relations suivantes :

$$m = \phi_0 \quad C(h) = \sum_1^{+\infty} \phi_n^2 \rho^n(h)$$

Connaissant les $(\phi_n)_{n>0}$, on choisit la covariance des gaussiennes, ρ , telle que le variogramme expérimental $\hat{\gamma}$ soit ajusté au mieux par :

$$h \neq 0 \quad \phi_0 + \sum_1^{+\infty} \phi_n^2 (1 - \rho^n(h))$$

Cette procédure pour ajuster les paramètres du modèle permet de spécifier un modèle cohérent. On vérifie simplement une relation entre les covariances ; on ne teste pas de manière exhaustive l'hypothèse gaussienne. Cependant la physique du phénomène étudié peut suggérer la loi multivariable, par exemple un modèle diffusif. D'autre part, on constatera sur un exemple de modèle non adapté que le conditionnement, ici défini par le modèle de bruit, l'emporte sur le modèle *a priori* de l'image.

4.2 Simulation et espérance conditionnelle empirique

Le modèle présenté ci-dessus permet de calculer la loi conditionnelle du vecteur des gaussiennes. Sa densité est donnée par la formule de Bayes (formule 13 page 13) :

$$g(x|z) = \frac{\prod_{\alpha \in A} f(z_\alpha | \phi(x_\alpha)) g(x)}{\int \prod_{\alpha \in A} f(z_\alpha | \phi(u_\alpha)) g(u) du}$$

où g est la densité de Gauss de la loi *a priori* du vecteur X et x (resp. u) désigne le vecteur $(x_\alpha)_{\alpha \in A}$ (resp. $(u_\alpha)_{\alpha \in A}$). L'espérance conditionnelle d'une fonction des gaussiennes sous-jacente, ψ est donc

$$\mathbf{E}\{\psi(X)|z\} = \frac{\int \psi(u) \prod_{\alpha \in A} f(u_\alpha | \phi(x_\alpha)) g(u) du}{\int \prod_{\alpha \in A} f(z_\alpha | \phi(u_\alpha)) g(u) du}$$

En particulier, toute fonction des concentrations sous-jacentes peut s'exprimer comme une fonction des gaussiennes, $\psi(Y) = \psi((\phi(X_\alpha)_{\alpha \in A}))$, et peut donc être estimée à l'aide de la formule précédente. Bien souvent la fonction ψ est une fonction vectorielle et à chacune de ses composantes est associé un pixel de l'image. On a par exemple, $\forall \alpha \in B \subset A$,

$$\psi_\alpha(\tilde{X}^{(i)}) = \begin{cases} \phi(\tilde{X}_\alpha^{(i)}) \\ \text{ou} \\ \frac{1}{|V_\alpha|} \sum_{\beta \in V_\alpha} \mathbb{1}_{\tilde{X}_\beta^{(i)} \geq x_c} \end{cases}$$

Ces deux cas correspondent à l'estimation de la variable sous-jacente ou de l'histogramme local. Dans la pratique le calcul direct de l'espérance conditionnelle n'est pas possible. Ici une approximation sera calculée par la méthode de Monte Carlo [23]. A l'aide de simulations conditionnelles de l'image des gaussiennes, $(\tilde{X}^{(i)})_{i>0}$ — c'est-à-dire des réalisations de la loi conditionnelle de densité $g(x|z)$ — on calcule la valeur moyenne² de ψ :

$$\psi^*(n) = \frac{1}{n} \sum_{i=1}^n \psi(\tilde{X}^{(i)})$$

$\psi^*(n)$ est une approximation de l'espérance conditionnelle de $\psi(X)$ si

$$\lim_{n \rightarrow +\infty} \psi^*(n) = \mathbf{E}\{\psi(X)|z\} \quad \text{p.s.}$$

² $(\tilde{X}^{(i)})_{i>0}$ constitue un échantillon de la loi conditionnelle. Au lieu de la valeur moyenne, il est donc possible de calculer d'autres statistiques telles que la médiane ou des quantiles.

Pour assurer la convergence, on peut prendre des simulations indépendantes — ceci est la méthode de Monte Carlo la plus simple — ou bien les états successifs d'une chaîne de Markov ergodique — ceci est la méthode proposée par Metropolis et al. [41]. Par la suite, ces deux méthodes sont appelées respectivement la méthode de Monte-Carlo et la méthode de Metropolis. Dans ces méthodes, l'approximation de l'espérance conditionnelle est une variable aléatoire. Sa précision sera donnée par sa variance.

On présente tout d'abord, la méthode de simulation. Ensuite, on détermine la méthode d'estimation — Monte Carlo ou Metropolis — qui nécessite une moindre quantité de calculs pour obtenir la précision désirée.

4.2.1 Simulation conditionnelle

Pour simuler la loi conditionnelle des gaussiennes, $g(x|z)$, on a utilisé l'algorithme proposé par Geman et Geman[20]. Cette méthode a été utilisée au chapitre 3 pour simuler des gaussiennes tronquées (page 58) : La loi à simuler est l'unique probabilité stationnaire d'une probabilité de transition ergodique. A l'aide de cette probabilité et d'une probabilité initiale quelconque, on construit une chaîne de Markov, $(\tilde{X}^{(\nu)})_{\nu \in \mathbb{N}}$. La loi des itérations converge vers la loi stationnaire.

a - Simulation initiale

Dans cette application, pour générer l'image initiale, on choisit de ne pas tenir compte des corrélations entre les gaussiennes — la fonction aléatoire gaussienne est alors supposée pépétique. La densité *a priori* des gaussiennes se factorise $g(x) = \prod_{\alpha \in A} g(x_\alpha)$ et la densité de la loi initiale est donnée par la formule 1.5 :

$$\frac{\prod_{\alpha \in A} f(z_\alpha | \phi(x_\alpha)) g(x)}{\int \prod_{\alpha \in A} f(z_\alpha | \phi(u_\alpha)) g(u) du} = \frac{\prod_{\alpha \in A} f(z_\alpha | \phi(x_\alpha)) g(x_\alpha)}{\int \prod_{\alpha \in A} f(z_\alpha | \phi(u_\alpha)) g(u_\alpha) du_\alpha} = \prod_{\alpha \in A} \frac{f(z_\alpha | \phi(x_\alpha)) g(x_\alpha)}{\int f(z_\alpha | \phi(u_\alpha)) g(u_\alpha) du_\alpha}$$

La densité se factorisant, chacune des composantes se simule indépendamment l'une de l'autre. On doit donc simuler la densité,

$$\alpha \in A \quad \frac{f(z_\alpha | \phi(x_\alpha)) g(x_\alpha)}{\int f(z_\alpha | \phi(u_\alpha)) g(u_\alpha) du_\alpha}$$

b - Probabilité de transition

La probabilité de transition est définie de la sorte :

- Une itération est constituée par une modification séquentielle de toute les composantes de l'image.
- Pour modifier la composante α du vecteur des gaussiennes, les valeurs des autres gaussiennes, $(X_\beta)_{\beta \neq \alpha}$, étant fixées, on remplace la valeur de X_α par une valeur simulée selon

sa loi conditionnelle. On note $x = (x_\alpha, \bar{x}_\alpha)$ où $\bar{x}_\alpha = (x_\beta)_{\beta \neq \alpha}$. La densité à simuler est donc :

$$\frac{g(x_\alpha, \bar{x}_\alpha | z)}{\int g(u_\alpha, \bar{x}_\alpha | z) du_\alpha} = \frac{f(z_\alpha | \phi(x_\alpha)) g(x_\alpha | \bar{x}_\alpha)}{\int f(z_\alpha | \phi(u_\alpha)) g(u_\alpha | \bar{x}_\alpha) du_\alpha} = \frac{f(z_\alpha | \phi(x_\alpha)) g\left(\frac{x_\alpha - x_\alpha^{KS}}{\sigma_{KS}(\alpha)}\right) \frac{1}{\sigma_{KS}(\alpha)}}{\int f(z_\alpha | \phi(u_\alpha)) g\left(\frac{u_\alpha - x_\alpha^{KS}}{\sigma_{KS}(\alpha)}\right) \frac{du_\alpha}{\sigma_{KS}(\alpha)}}$$

où x_α^{KS} est la valeur du krigeage simple de X_α par les $(X_\beta)_{\beta \neq \alpha}$ et $\sigma_{KS}^2(\alpha)$ sa variance. On peut récrire ceci en termes de variables aléatoires :

$$X_\alpha = X_\alpha^{KS}(\bar{x}_\alpha) + \sigma_{KS}(\alpha) U_\alpha$$

où la densité du résidu U_α est proportionnelle à $f(z_\alpha | \phi(x_\alpha^{KS} + \sigma_{KS}(\alpha) u_\alpha)) g(u_\alpha)$.

On vérifie (voir page 63) que la loi de densité $g(x|z)$ est une loi stationnaire de la chaîne de Markov ainsi construite. L'ergodicité de la probabilité de transition n'est pas démontrée. Cependant, dans aucun des exemples traités, il n'a été rencontré de cas pathologique.

c - Simulation du résidu

Pour simuler l'image initiale ³ ou les itérations successives, il faut simuler une densité proportionnelle à

$$f(z_\alpha | \phi(x_\alpha^{KS} + \sigma_{KS}(\alpha) u_\alpha)) g(u_\alpha)$$

Pour cela, on fait l'hypothèse ⁴ que la fonction d'anamorphose, ϕ , est une fonction constante par morceau :

$$\phi = \sum_{i \in I} y_i \mathbb{1}_{B_i} \text{ p.p.}$$

où $(B_i)_{i \in I}$ est une famille finie d'intervalles disjoints, non réduits à un point, et $(y_i)_{i \in I}$ des réels. La densité à simuler est alors proportionnelle à

$$\sum_{i \in I} \mathbb{1}_{B'_i}(u) f(z_\alpha | y_i) g(u)$$

avec $B'_i = \{u, x_\alpha^{KS} + \sigma_{KS}(\alpha) u \in B_i\}$. Cette densité est simulée par la méthode d'acceptation et rejets. La méthode de simulation utilisée ici est très similaire à celle utilisée pour simuler une gaussienne tronquée scalaire (voir page 48) ; la seule différence est que les intervalles B'_i sont pondérés par les probabilités conditionnelles $f(z_\alpha | y_i)$. En effet, sur chacun des $(B'_i)_i$, la densité de Gauss est majorée par une fonction sommable, $u \in B'_i, g(u) \leq f_i(u)$. La densité

³Pour la simulation initiale, les corrélations entre les composantes gaussiennes ayant été sciemment négligées, on a $x_\alpha^{KS} = 0$ et $\sigma_{KS}^2(\alpha) = 1$.

⁴Le calcul de la fonction d'anamorphose empirique à partir d'un estimateur sous-jacent (voir page 85) fournit bien une fonction de cette forme. Si cette hypothèse n'est pas vérifiée, on approximera la fonction d'anamorphose par une fonction monotone et constante par morceau.

du résidu est alors majorée sur \mathbb{R} par

$$\frac{\sum_i f(z_\alpha|y_i) \mathbb{1}_{B'_i} g(u)}{\sum_i f(z_\alpha|y_i) \int_{B'_i} g(v) dv} \leq C \sum_i \varpi_i \frac{f_i(u) \mathbb{1}_{B'_i}(u)}{\int_{B'_i} f_i(v) dv} \quad \text{avec} \quad \begin{cases} C = \frac{\sum_j f(z_\alpha|y_j) \int_{B'_j} f_j(v) dv}{\sum_j f(z_\alpha|y_j) \int_{B'_j} g(w) dw} \\ \varpi_i = \frac{f(z_\alpha|y_i) \int_{B'_i} f_i(v) dv}{\sum_j f(z_\alpha|y_j) \int_{B'_j} g(w) dw} \end{cases}$$

L'algorithme de simulation de U par la méthode d'acceptation et rejets est donc :

(i) simulation de l'intervalle \underline{i} selon la loi $(\varpi)_i$,

(ii) simulation des variables aléatoires U et V de densité respective

$$\frac{f_{\underline{i}}(u) \mathbb{1}_{B'_{\underline{i}}}(u)}{\int_{B'_{\underline{i}}} f_{\underline{i}}(v) dv} \quad \text{et} \quad \mathbb{1}_{]0,1[}(v),$$

(iii) si $V f_{\underline{i}}(U) \leq g(U)$ le candidat U est retenu. Sinon retour en (i).

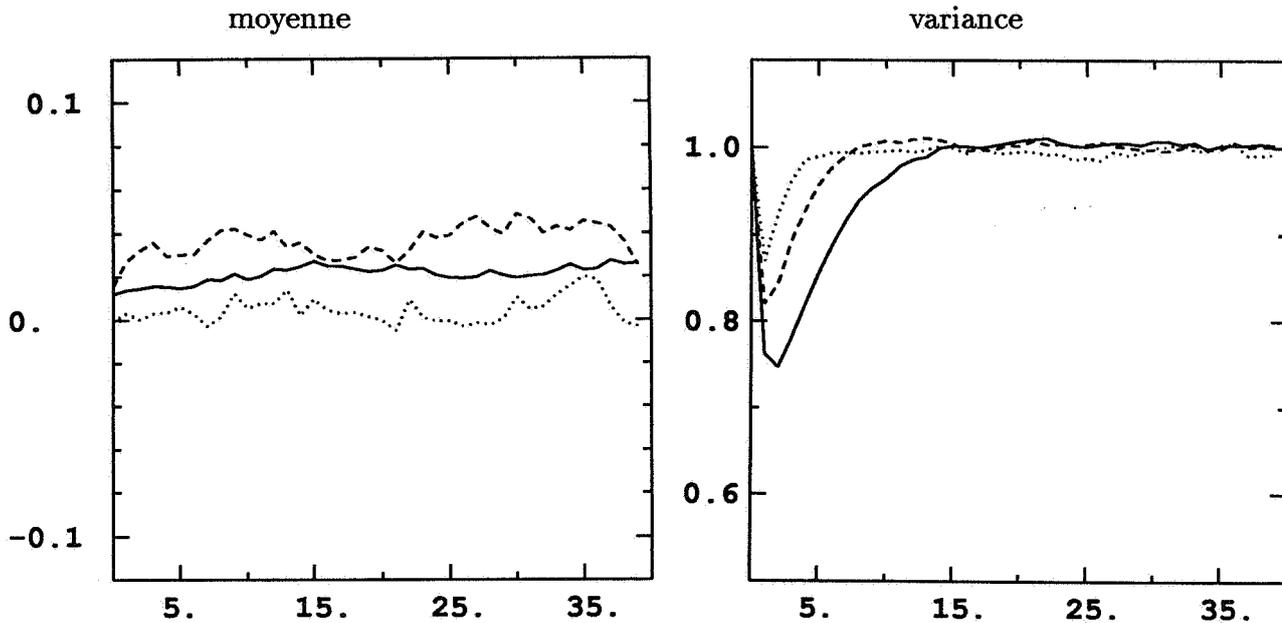


Figure 4.1 : Statistiques calculées sur les itérations successives

image 100×100 , $\phi(x) = 100 + 10x$, $\rho(h) = \rho^{|h_x|+|h_y|}$

..... : $\rho = 0.5$
 ---- : $\rho = 0.6$
 — : $\rho = 0.7$

Pour simuler la densité $g(\cdot|z)$, on simule tout d'abord une image initiale $\tilde{X}^{(0)}$ puis les itérations successives $(\tilde{X}^{(n)})_{n>0}$. On arrête la chaîne lorsque l'on estime avoir atteint l'état stationnaire. Cependant n'ayant pas trouvé de critère théorique pour déterminer le nombre d'itérations nécessaires pour atteindre cet état, nous avons étudié expérimentalement les deux premiers moments des statistiques calculées sur les itérations successives. L'état stationnaire est estimé atteint lorsque ces statistiques sont égales — aux fluctuations statistiques près — aux valeurs du modèle.

Sur les figures 4.1 et 4.3, est présentée l'étude réalisée sur un cas particulier : les statistiques élémentaires — moyenne et variance — et le variogramme sont calculés sur les itérations successives pour 10 simulations indépendantes. On constate que dès les premières itérations, la moyenne est satisfaisante. Par contre, la variance ne croît que lentement. Quant aux variogrammes expérimentaux, les premiers points sont très rapidement acceptables, mais la valeur du palier suit l'évolution de la valeur de la variance. La variance expérimentale semble un bon indicateur de l'état stationnaire. La vitesse de convergence expérimentale dépend de la corrélation entre les pixels voisins mais ne dépend pas de la taille du champ à simuler (figure 4.2).

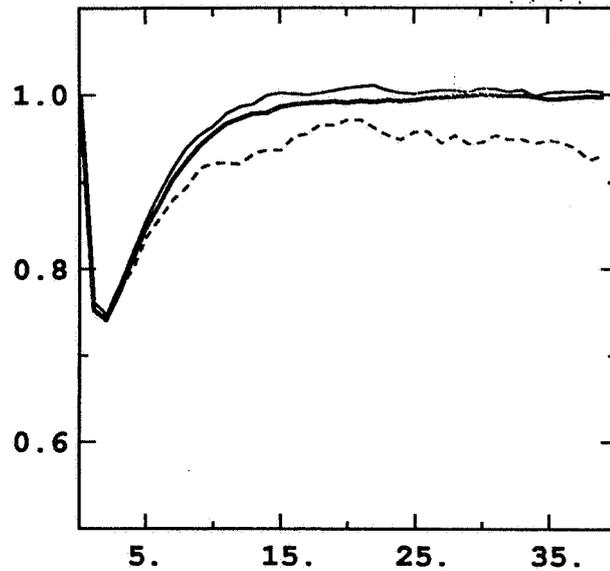


Figure 4.2: Variances calculées sur les itérations successives

$$(\rho(h) = 0.7^{|h_x|+|h_y|}, \phi(x) = 100 + 10x)$$

- - - : grille 50 × 50
- : grille 100 × 100
- : grille 200 × 200

4.2.2 Calcul de l'espérance conditionnelle empirique

La méthode de simulation présentée aux paragraphes précédents utilise une chaîne de Markov : la loi conditionnelle à simuler est une loi stationnaire de la probabilité de transition que l'on a supposée ergodique. Cette même chaîne peut donc être utilisée pour calculer l'espérance conditionnelle par la méthode de Metropolis. Ainsi les quantités de calculs effectués dans les deux méthodes — moyenne de simulations indépendantes ou moyenne des itérations successives de la chaîne de Markov — peuvent être comparées par le nombre total d'itérations calculées. On choisit la méthode fournissant pour un même effort de calcul la plus grande précision, c'est-à-dire la plus petite variance de ψ^* . Calculons tout d'abord la précision obtenue pour chacune des deux méthodes par le calcul de n_{tot} itérations. n_i désigne le nombre d'itérations pour atteindre l'état stationnaire. Dans la méthode de Monte Carlo, pour chaque simulation il faut calculer n_i itérations d'initialisation. Le nombre de simulations indépendantes obtenues par le calcul de n_{tot} itérations est donc $\frac{n_{tot}}{n_i}$. La variance de l'approximation de l'espérance conditionnelle est alors $\text{Var}(\psi^*) = \frac{n_i}{n_{tot}} \text{Var}(\psi(X)|z)$. Dans la méthode de Metropolis, il suffit d'initialiser une fois la chaîne. Pour n_{tot} itérations calculées, l'approximation est donc

$$\psi^* = \frac{1}{n_{tot} - n_i} \sum_{k=1}^{n_{tot} - n_i} \psi(\tilde{X}^{(n_i+k)})$$

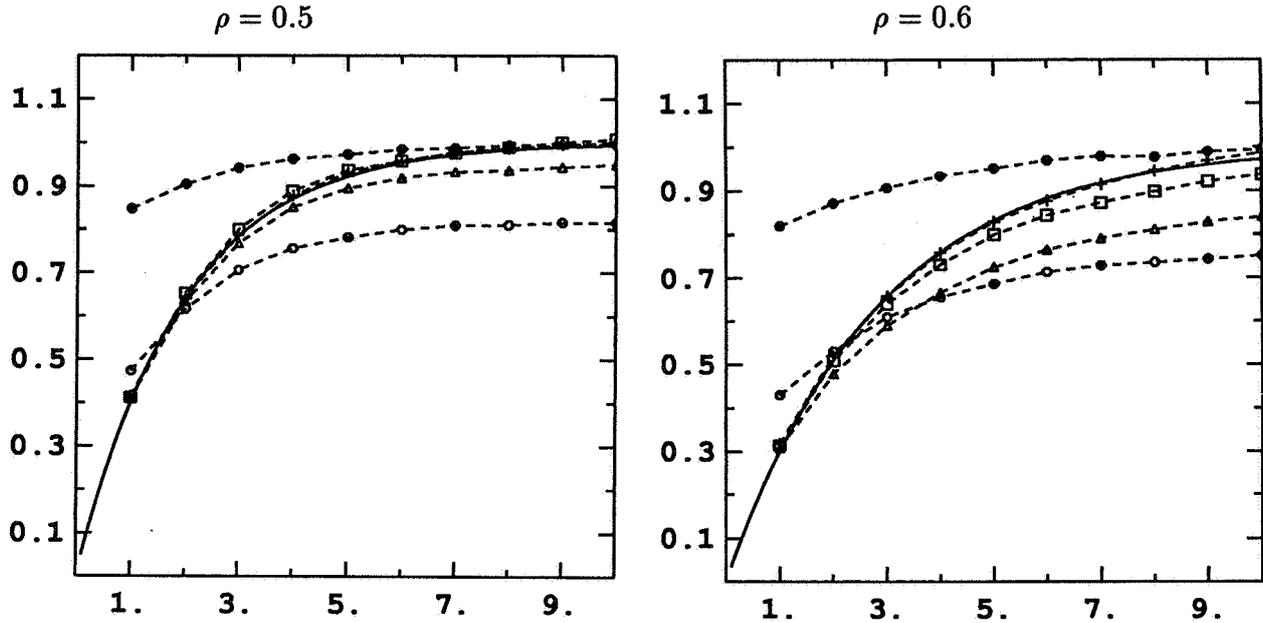


Figure 4.3: Variogrammes spatiaux calculés sur les itérations successives
(image 100×100 , $\rho(h) = \rho^{|h_x|+|h_y|}$, $\phi(x) = 100 + 10x$)

	: modèle
•	: 0
○	: 1
△	: 5
□	: 10
+	: 20

et sa variance est

$$\frac{1}{(n_{tot} - n_i)^2} \sum_{i,j=1}^{n_{tot}-n_i} \text{Cov}(\psi(\tilde{X}^{(n_i+i)}), \psi(\tilde{X}^{(n_i+j)}))$$

Mais au delà de n_i , la chaîne est à peu près stationnaire, il en est de même pour la covariance temporelle, $\text{Cov}(\psi(\tilde{X}^{(n_i+i)}), \psi(\tilde{X}^{(n_i+j)}))$, qui ne dépend donc plus que de la différence $i - j$. Elle sera notée par la suite $\mathcal{C}(i - j)$. On a de plus $\text{Var}(\psi(X|z)) = \mathcal{C}(0)$. Cependant là encore, l'expression théorique de cette covariance est restée inconnue. On a donc adopté une approche expérimentale: le variogramme moyen a été calculé sur les itérations successives dans la direction du temps.

$$\hat{\gamma}(k) = \frac{1}{n - k} \sum_{i=1}^{n-k} \frac{1}{2} [\psi(\tilde{X}^{(n_i+i)}) - \psi(\tilde{X}^{(n_i+i+k)})]^2$$

Le palier de ce variogramme donne la variance d'estimation à z fixé:

$$\mathbf{E}\{[\psi(X) - \mathbf{E}\{\psi(X)|z\}]^2|z\}$$

Cependant il faut disposer d'un très grand nombre d'itérations pour obtenir un variogramme expérimental structuré. De plus, ψ est bien souvent une fonction vectorielle dont les composantes sont associées à des pixels de l'image — par exemple, la variable estimée est la

variable sous-jacente ou l'histogramme local. On peut alors calculer le variogramme moyen dans la direction temporelle sur un sous-ensemble du champ, $B \subset A$:

$$\hat{\gamma}(k) = \frac{1}{|B|} \sum_{\alpha \in B} \frac{1}{n-k} \sum_{l=1}^{n-k} \frac{1}{2} [\psi_{\alpha}(\tilde{X}^{(n_i+l)}) - \psi_{\alpha}(\tilde{X}^{(n_i+l+k)})]^2$$

Travaillant sur un modèle stationnaire et ergodique, calculer une valeur moyenne dans l'espace donne une approximation de l'espérance. Le variogramme temporel moyen est alors un estimateur de

$$k > n_i \quad \frac{1}{2} \mathbf{E}\{[\psi(\tilde{X}^{(k)}) - \psi(\tilde{X}^{(k+n)})]^2\}$$

Ce variogramme expérimental (figure 4.4) peut être ajusté à l'aide d'un modèle stationnaire⁵.

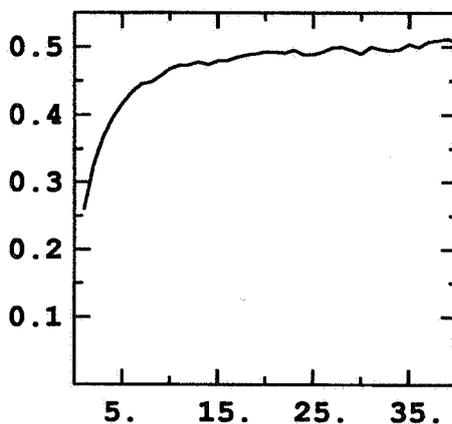


Figure 4.4: Variogramme temporel
(image 100×100 , $\rho(h) = 0.7^{|h|}$, $\phi(x) = 5 + x$)

En procédant de la sorte, on détermine une précision moyenne de l'image, qui est plus opératoire qu'une précision associée à chacun des pixels estimés. De plus, le palier du variogramme temporel donne la variance d'estimation associée à l'espérance conditionnelle empirique. Dans tous les cas traités, la méthode de Metropolis donne une variance de l'approximation plus faible pour un même effort de calcul.

Un deuxième argument joue en faveur de l'algorithme de Metropolis : si on a sous-estimé le nombre d'itérations d'initialisation, dans la méthode de Monte Carlo simple la loi des simulations conditionnelles indépendantes n'est pas exacte alors que le biais s'estompe avec la méthode de Metropolis, puisqu'on ne relance pas la chaîne.

Les résultats présentés par la suite utilisent donc la méthode de Métropolis.

⁵Sur cet exemple, la variance d'estimation vaut 0.5, c'est à dire la moitié de la variance du modèle. Cependant le rapport signal sur bruit avant et après filtrage passe de 0.2 à 2.

4.2.3 Conclusion

La méthode présentée permet de calculer une espérance conditionnelle empirique: théoriquement optimale — cette estimation minimise la variance d'estimation — la méthode présente cependant des inconvénients qu'il faut maintenant récapituler.

- La méthode s'appuie sur un modèle de bruit et un modèle gaussien de la variable sous-jacente. La modélisation du bruit s'appuie sur la physique de la mesure; le modèle gaussien est une hypothèse qu'il est difficile de tester sur les données expérimentales.
- On ne connaît pas les valeurs théoriques des paramètres de calcul, par exemple la variance d'estimation et le nombre d'itérations d'initialisation. Cependant ces valeurs peuvent être estimées expérimentalement: les premiers termes de la chaîne de Markov sont simulés puis les statistiques sont calculées.
- Cette méthode est beaucoup plus lente que les méthodes classiques, krigeage simple, krigeage ordinaire ou krigeage disjonctif gaussien.

Dans les paragraphes qui suivent, on se propose d'utiliser cette méthode sur des données simulées: la valeur des concentrations bruitées et celle des concentrations sous-jacentes sont connues; il est donc possible d'estimer les performances de l'espérance conditionnelle empirique — EC — et de les comparer à celles des autres méthodes (krigeage simple — KS — et krigeage disjonctif gaussien — KD). On testera aussi le cas où le modèle d'image n'est pas adapté en utilisant des modèles d'images mosaïques. Enfin, la méthode sera utilisée sur des données réelles.

4.3 Comparaison de trois méthodes d'estimation

A partir d'observations de la variable poissonisée, on estime la variable sous-jacente $Y_\alpha = \phi(X_\alpha)$ ainsi qu'une indicatrice $\mathbb{1}_{X_\alpha \geq x_c}$ et sa régularisée sur le voisinage V_α — l'*histogramme local*. Les estimateurs seront calculés sur des images simulées. On connaît la valeur de la variable estimée. Il est donc possible de calculer l'erreur d'estimation expérimentale. Deux statistiques ont été retenues, la moyenne et l'écart quadratique moyen:

$$* \in \{KS, KD, EC\} \quad \begin{cases} E_*(\psi) &= \frac{1}{|A|} \sum_{\alpha \in A} (\psi(X_\alpha) - \psi^*(X_\alpha)) \\ D_*^2(\psi) &= \frac{1}{|A|} \sum_{\alpha \in A} (\psi(X_\alpha) - \psi^*(X_\alpha))^2 \end{cases}$$

La première permet de vérifier que l'estimateur est sans biais, la seconde doit être comparée à la variance d'estimation théorique. De plus, dans le cas du filtrage — c'est-à-dire l'estimation de la variable sous-jacente — on définit [10] le rapport *signal sur bruit* avant et après filtrage par:

$$(S/B)_I = \frac{\text{Var}(Y)}{\text{Var}(Y - Z)} = \frac{\text{Var}(Y)}{\mathbf{E}\{Y\}} \quad (S/B)_F = \frac{\text{Var}(Y)}{\text{Var}(Y - Y^*)}$$

Nous avons étudié l'influence de la fonction de transformation ainsi que l'influence du modèle. Mais tout d'abord, on rappelle les estimateurs classiques auxquels l'espérance conditionnelle empirique sera comparée.

Le krigeage simple

Le krigeage simple [33] de la variable Y_α à partir des observations $Z = (Z_\alpha)_{\alpha \in A}$ est l'estimateur linéaire,

$$\alpha \in A \quad Y_\alpha^{KS} = \mathbf{E}\{Y\} + \sum_{\beta \in A} \lambda_{(\alpha)}^\beta (Z_\beta - \mathbf{E}\{Z\})$$

minimisant la variance d'estimation :

$$\mathbf{Var}(Y_\alpha - Y_\alpha^{KS})$$

Les $(\lambda_{(\alpha)}^\beta)_{\beta \in A}$ vérifient donc le système,

$$\forall \beta \in A \quad \sum_{\gamma \in A} \mathbf{Cov}(Z_\beta, Z_\gamma) \lambda_{(\alpha)}^\gamma = \mathbf{Cov}(Z_\beta, Y_\alpha)$$

La variance d'estimation est $\sigma_{KS}^2(\alpha) = \mathbf{Var}(Y_\alpha) - \sum_{\beta} \lambda_{(\alpha)}^\beta \mathbf{Cov}(Z_\beta, Y_\alpha)$.

Le krigeage disjonctif gaussien

On désire estimer une fonction de Y , $f(Y_\alpha)$ à l'aide d'un estimateur de la forme $f^* = \sum_{\beta} f_\beta(Z_\beta)$ minimisant la variance d'estimation. Pour mettre en œuvre le krigeage disjonctif gaussien [34, 10], on modélise la corégionalisation $(Z_\alpha, Y_\alpha)_{\alpha \in A}$ comme la transformée d'une corégionalisation de facteurs gaussiens :

$$\begin{cases} Z_\alpha = \tilde{\phi}(\tilde{X}_\alpha) \\ Y_\alpha = \phi(X_\alpha) \end{cases}$$

où $\tilde{X}_\alpha = rX_\alpha + \sqrt{1-r^2}U_\alpha$. X est la gaussienne de la variable sous-jacente et U un facteur gaussien pépitique. $f \circ \phi$ et les $(f_\beta \circ \tilde{\phi})_\beta$ appartenant à $L^2(\mathbb{R}, g)$ dont les polynômes d'Hermite constituent une base orthonormée complète, la variance d'estimation par krigeage disjonctif s'exprime comme la somme des variances d'estimation par krigeage simple des facteurs :

$$\mathbf{Var}(f \circ \phi(X_\alpha) - \sum_{\beta} f_\beta(\tilde{X}_\beta)) = \sum_{n>0} \mathbf{Var}(\eta_n(X_\alpha) - \sum_{\beta} \lambda_n^\beta(\alpha) \eta_n(\tilde{X}_\beta)) f_n^2$$

où les $(f_n)_{n \geq 0}$ sont les coordonnées de $f \circ \phi$ par rapport aux $(\eta_n)_{n \geq 0}$ et les $(\lambda_n^\beta(\alpha))_{\beta \in A, n > 0}$ sont solutions des systèmes de krigeage simple suivants,

$$n \in \mathbb{N}^* , \beta \in A \quad \sum_{\gamma \in A} [(1-r^2)\delta_{\beta\gamma} + r^2\rho(\beta-\gamma)] \lambda_n^\beta(\alpha) = [r\rho(\beta-\alpha)]^n$$

L'estimateur et sa variance s'écrivent alors :

$$\begin{cases} f^{KD} = f_0 + \sum_{n>0} f_n [\sum_{\beta} \lambda_n^\beta(\alpha) \eta_n(\tilde{X}_\beta)] \\ \sigma_{KD}^2 = \sum_{n>0} [1 - \sum_{\beta} \lambda_n^\beta(\alpha) (r\rho(\beta-\alpha))^n] f_n^2 \end{cases}$$

Estimateurs et modèles

Pour calculer le krigeage simple, il faut disposer des covariances entre les données expérimentales et les covariances croisées entre les mesures et la variable à estimer. Dans le cas de l'estimation de la variable sous-jacente, le modèle de bruit poissonien permet d'exprimer ces covariances en fonction de l'espérance de la variable sous-jacente et de sa covariance :

$$\begin{aligned} \text{Cov}(Z_\alpha, Z_\beta) &= \mathbf{E}\{Y\}\delta_{\alpha\beta} + \text{Cov}(Y_\alpha, Y_\beta) \\ \text{Cov}(Y_\alpha, Z_\beta) &= \text{Cov}(Y_\alpha, Y_\beta) \end{aligned}$$

On doit modéliser cette covariance à l'aide d'un modèle qui soit cohérent [39] avec l'histogramme estimé par ailleurs. Ne disposant pas de critères généraux pour assurer cette compatibilité, on doit expliciter un modèle de fonction aléatoire présentant les caractéristiques requises. On a retenu la covariance du modèle utilisé pour le calcul de l'espérance conditionnelle empirique, car, outre la cohérence entre la covariance et l'histogramme, ce choix permet de calculer le krigeage simple et l'espérance conditionnelle avec la même covariance.

Pour calculer le krigeage simple de l'indicatrice, $\mathbb{1}_{Y \geq y_c}$, il faut connaître la covariance croisée entre la variable bruitée et l'indicatrice. Là encore, le modèle de bruit permet de ramener cette inférence à celle de la covariance entre la variable sous-jacente et une de ses indicatrices. En effet, $\text{Cov}(Z_\beta, \mathbb{1}_{Y_\alpha \geq y_c}) = \text{Cov}(Y_\beta, \mathbb{1}_{Y_\alpha \geq y_c})$. Pour assurer la cohérence du modèle, on fait l'hypothèse que Y est l'anamorphosée d'une fonction aléatoire dont les lois bivariées sont gaussiennes. On a $Y_\alpha = \phi(X_\alpha)$ et $\mathbb{1}_{Y_\alpha \geq y_c} = \mathbb{1}_{X_\alpha \geq x_c}$ où $x_c = \phi^{-1}(y_c)$ ⁶. Les covariances simples et croisées vérifient alors les relations :

$$\begin{cases} C^{YY}(h) = \sum_{n \geq 1} \phi_n^2 \rho^n(h) \\ C^{Y\mathbb{1}}(h) = -g(x_c) \sum_{n \geq 1} \phi_n \frac{\eta_{n-1}(x_c)}{\sqrt{n}} \rho^n(h) \end{cases}$$

Ces deux hypothèses diffèrent grandement : dans le premier cas — filtrage — elle conduit à limiter la classe des fonctions de covariance dans laquelle sera choisie une fonction ajustant au mieux le variogramme expérimental, dans l'autre cas — estimation d'une indicatrice — on s'impose un type de covariance croisée qui n'est pas contrôlé sur des données expérimentales.

Pour mettre en œuvre le krigeage disjonctif, on suppose que la variable sous-jacente comme la variable bruitée sont les anamorphosées de fonctions aléatoires de lois bivariées gaussiennes et que ces variables sont liées par une certaine relation qui sous l'hypothèse d'un bruit poissonien ne peut être vérifiée. Ceci revient donc à substituer au modèle de bruit poissonien un modèle approximatif pour lequel la résolution du krigeage disjonctif soit possible. Sur simulations, on constate que dans certains cas la relation est une bonne approximation. Ceci limite l'usage qui peut être fait du krigeage disjonctif gaussien pour résoudre ce type de problème. En particulier, il faut associer une gaussienne à la variable bruitée entière. La fonction d'anamorphose, associant des entiers à des réels, n'est pas bijective. On convient d'un pseudo-inverse pour le calcul des valeurs de la gaussienne de la variable bruitée. Mais l'effet de cet artefact doit être contrôlé : en particulier la loi bivariable des gaussiennes associées à la variable bruitée doit être bigaussienne.

⁶Si ϕ n'est pas bijective, $x_c = \inf \phi^{-1}(\{y_c\})$.

L'espérance conditionnelle empirique s'appuie quant à elle sur l'hypothèse selon laquelle la variable sous-jacente est l'anamorphosée d'une fonction aléatoire multigaussienne.

Ainsi le modèle de bruit est poissonien pour le krigeage simple et l'espérance conditionnelle empirique, le modèle de bruit utilisé pour le krigeage disjonctif est une approximation. Par contre, si on adopte la même covariance pour le calcul des trois estimateurs, les modèles de l'image sous-jacente sont compatibles : le modèle du krigeage disjonctif est déduit du modèle du krigeage simple en précisant que la loi bivariable est gaussienne ; celui de l'espérance conditionnelle exige de plus que la loi de l'image soit multigaussienne.

L'hypothèse multigaussienne ne peut être testée sur une réalisation. Dans la pratique, on se limite généralement au test de la loi bivariable, ce qui conduit à confondre sciemment les fonctions aléatoires de lois bivariées gaussiennes et les fonctions aléatoires multigaussiennes⁷. Mais, dans le cas d'une poissonisation, le test de cette loi bivariable pose problème puisque les données expérimentales ne sont pas liées de manière simple avec la variable gaussienne dont les valeurs ne peuvent être calculées. C'est pourquoi nous avons comparé ces estimateurs sur simulations.

Influence de la fonction d'anamorphose

On simule tout d'abord une image multigaussienne de taille 100×100 dont la covariance est $\rho(h) = \exp(-|h|/2.8)$. Cette image est transformée à l'aide d'une fonction d'anamorphose (tableau 4.1), puis la poissonisée de cette image est calculée. A partir de l'image (très) bruitée, figures 4.7 et 4.8, la variable sous-jacente, $\phi(X)$, une indicatrice, $\mathbb{1}_{X \geq 1.64}$, et sa régularisée sur un carré de 5 pixels de côté ont été estimées à l'aide des trois méthodes présentées ci-dessus, figures 4.8, 4.11 et 4.12. La variance d'estimation et les statistiques de l'erreur d'estimation expérimentales⁸ sont reportées aux tableaux 4.2 et 4.3. Les cas envisagés se divisent en deux groupes : d'une part le cas où la fonction d'anamorphose est linéaire et d'autre part les autres fonctions de transformation.

a) anamorphoses linéaires

Pour l'estimation de la variable sous-jacente, les trois méthodes donnent des résultats équivalents. On peut expliquer cette équivalence expérimentale à l'aide des approximations suivantes. Si m est suffisamment grand, on substitue au bruit poissonien le modèle de bruit⁹ : $Z_\alpha = Y_\alpha + \sqrt{Y_\alpha}U_\alpha$ où U_α est un facteur gaussien pépétique. Si de plus on peut négliger¹⁰

⁷Il n'est pas facile d'exhiber des contre-exemples de fonctions aléatoires non multigaussiennes dont toutes les lois bivariées sont gaussiennes. Feller [18, page 99] donne un exemple de vecteur aléatoire de dimension n non multigaussien mais dont toutes les lois de rang inférieur à n sont gaussiennes.

⁸Ces trois estimateurs sont non biaisés.

⁹D'après le théorème de la limite centrale, $\frac{Pe-\theta}{\sqrt{\theta}}$ converge en loi vers une variable aléatoire gaussienne centrée normée lorsque θ tend vers $+\infty$

¹⁰Les valeurs de σ et m sont choisies de sorte que $m + \sigma X$ soit positif avec une probabilité très voisine de 1, ce qui entraîne que l'hypothèse $\sigma X \ll m$ est vérifiée.

Tableau 4.1 : Fonctions d'anamorphose

	moyenne	variance	rapport signal/bruit
anamorphose linéaire			
$\phi(x) = 5 + x$	5.0	1.0	0.2
$\phi(x) = 20 + 3.16x$	20.0	10.0	0.5
$\phi(x) = 100 + 10x$	100.0	100.0	1.0
anamorphose lognormale			
$\phi(x) = me^{sy-s^2/2} \quad s = 1.$	1.0	1.7	1.7
anamorphose cubique			
$\phi(x) = x^3 + 3x^2 + 3x + 57$	60.0	60.0	1.0
transformation discrète			
figure 4.5	1.8	3.8	2.1

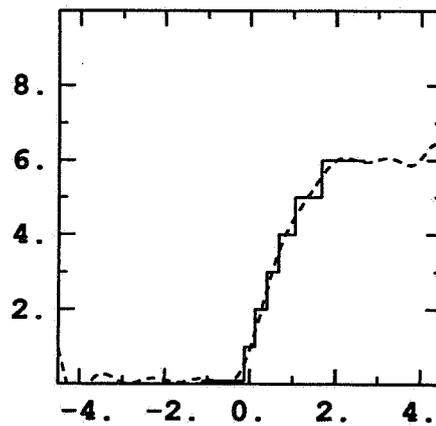


Figure 4.5 : Anamorphose discrète

- : fonction d'anamorphose
 --- : fonction modélisée par les 30 premiers polynômes d'Hermite

σX_α devant m , on a l'équivalence suivante,

$$Z_\alpha = m + \sigma X_\alpha + \sqrt{m + \sigma X_\alpha} U_\alpha \simeq m + \sqrt{m + \sigma^2} \left(\frac{\sigma}{\sqrt{m + \sigma^2}} X_\alpha + \sqrt{\frac{m}{m + \sigma^2}} U_\alpha \right)$$

On retrouve donc l'hypothèse du krigeage disjonctif gaussien et l'anamorphose associée à la variable bruitée est elle aussi linéaire. Le nuage de corrélation entre la gaussienne de la variable sous-jacente et celle de la variable bruitée (figure 4.6) confirme cette interprétation. On estime dans le cadre multigaussien une combinaison linéaire : les trois estimateurs utilisés sont équivalents. Dans ce cas, la méthode la plus simple, le krigeage simple, doit être retenue. On peut noter que l'espérance conditionnelle fournit des résultats très légèrement inférieurs, imputables aux approximations nécessaires pour mener à bien les calculs.

Pour l'estimation de l'indicatrice ou de l'histogramme local, l'espérance conditionnelle, utilisant toutes les propriétés du modèle, fournit l'estimateur le plus précis. On a reporté à la figure 4.8 les estimateurs de l'histogramme local lorsque la fonction d'anamorphose est $\phi(x) = 100 + 10x$: l'image la plus réaliste est celle obtenue par l'espérance conditionnelle empirique. Le krigeage simple lisse beaucoup trop le phénomène à estimer ; le krigeage disjonctif fournit un résultat intermédiaire entre le krigeage simple et l'espérance conditionnelle. De plus seule l'espérance conditionnelle empirique assure que l'estimateur soit compris entre 0 et 1.

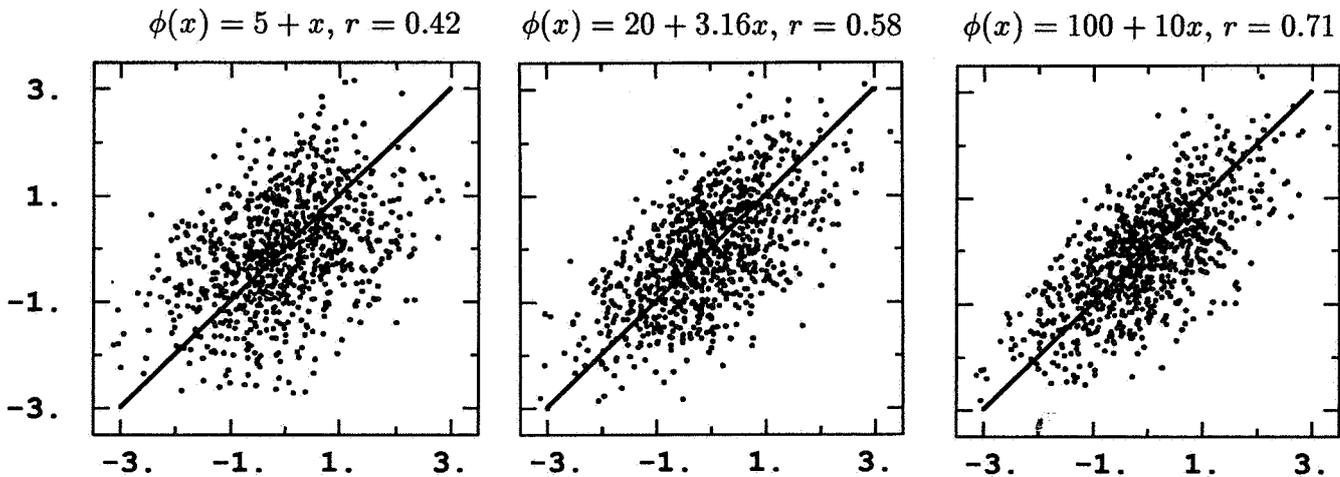


Figure 4.6 : Nuages de corrélation entre la gaussienne de la variable sous-jacente et la gaussienne de la variable bruitée

Tableau 4.2: Erreur d'estimation

*(image multigaussienne, anamorphose linéaire)*1- estimation de la variable sous-jacente

anamorphose	KS			KD			EC		
	E	σ_{KS}^2	D^2	E	σ_{KD}^2	D^2	E	σ_{EC}^2	D^2
$5 + x$	-0.0038	0.50	0.50	-0.0053	0.50	0.50	0.0144	0.50	0.52
$20 + 3.16x$	-0.0300	3.70	3.70	-0.0100	3.70	3.68	0.0300	3.70	3.86
$100 + 10x$	0.1200	28.32	27.92	-0.0053	28.12	27.90	0.1700	27.50	28.77

2- estimation de l'indicatrice $\mathbb{1}_{X>1.64}$

anamorphose	KS			KD			EC		
	E	σ_{KS}^2	D^2	E	σ_{KD}^2	D^2	E	σ_{EC}^2	D^2
$5 + x$	0.0011	0.0421	0.0421	0.0012	0.0413	0.0409	0.0026	0.0320	0.0374
$20 + 3.16x$	0.0025	0.0408	0.0408	0.0010	0.0372	0.0377	-0.0022	0.0310	0.0352
$100 + 10x$	0.0028	0.0398	0.0402	0.0012	0.0331	0.0331	0.0016	0.0280	0.0307

3- estimation de l'histogramme local ($V = 5 \times 5$)

anamorphose	KS		KD		EC		
	E	D^2	E	D^2	E	σ_{EC}^2	D^2
$5 + x$	0.0000	0.0070	-0.0005	0.0063	-0.0010	0.0039	0.0040
$20 + 3.16x$	0.0019	0.0066	0.0002	0.0045	-0.0045	0.0029	0.0031
$100 + 10x$	0.0024	0.0065	0.0004	0.0030	0.0006	0.0020	0.0020

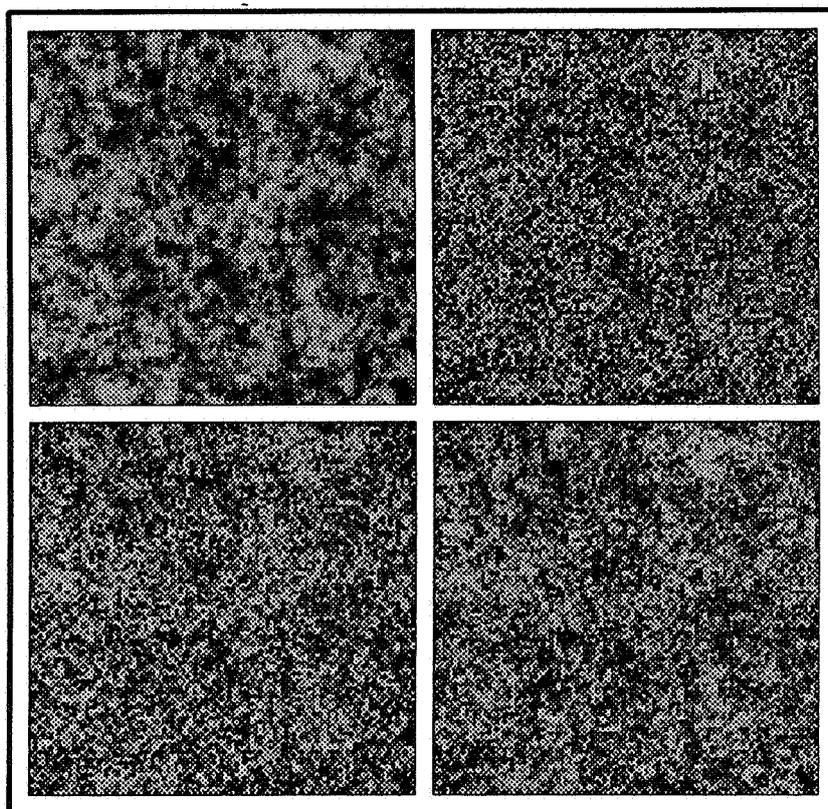


Figure 4.7: Images bruitées simulées

- en haut à gauche* : variable gaussienne
en haut à droite : variable bruitée — $\phi(x) = 5 + x$
en bas à gauche : variable bruitée — $\phi(x) = 20 + 3.16x$
en bas à droite : variable bruitée — $\phi(x) = 100 + 10x$

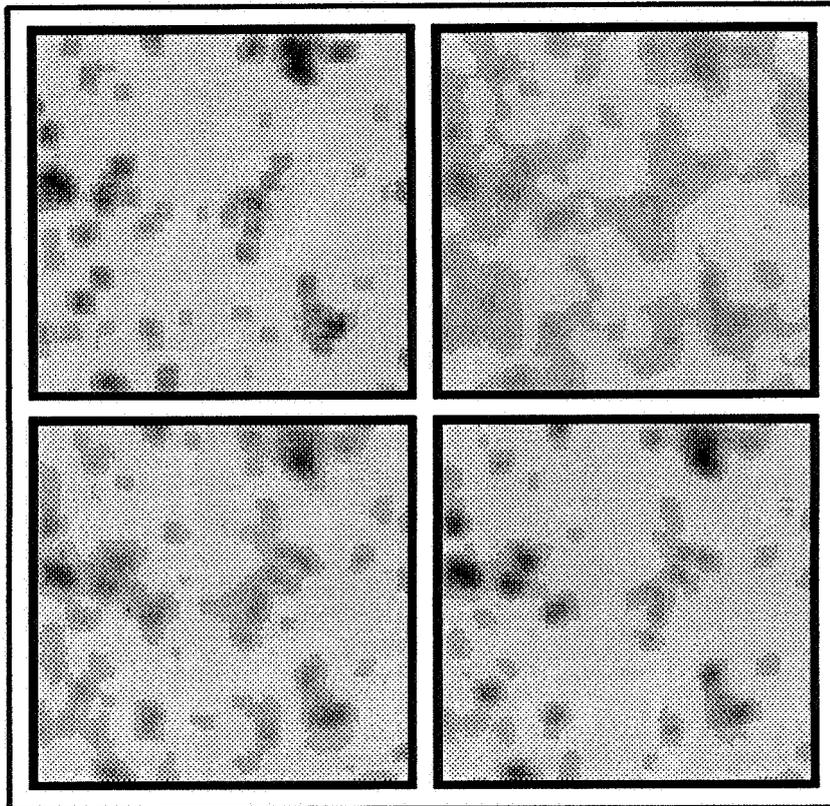


Figure 4.8 : Estimation de l'histogramme local ($\phi(x) = 100 + 10x$)

- en haut à gauche* : histogramme local calculé sur l'image simulée ($V = 5 \times 5$)
- en haut à droite* : histogramme local estimé par krigeage simple
- en bas à gauche* : histogramme local estimé par krigeage disjointif
- en bas à droite* : histogramme local estimé par espérance conditionnelle empirique

b) anamorphoses non linéaires

Lorsque la fonction d'anamorphose n'est pas linéaire, l'hypothèse sur laquelle le krigeage disjonctif s'appuie n'est pas vérifiée (figure 4.9). Expérimentalement le krigeage disjonctif est cependant plus précis que le krigeage simple mais il y a des divergences entre la variance d'estimation calculée dans le modèle et la variance expérimentale de l'erreur d'estimation. Pour toutes les variables estimées, l'espérance conditionnelle empirique fournit l'estimateur le plus précis. Les images estimées par krigeage simple et espérance conditionnelle sont reportées aux figures 4.11 et 4.12.

a - anamorphose lognormale b - anamorphose cubique c - anamorphose discrète (p. 98)

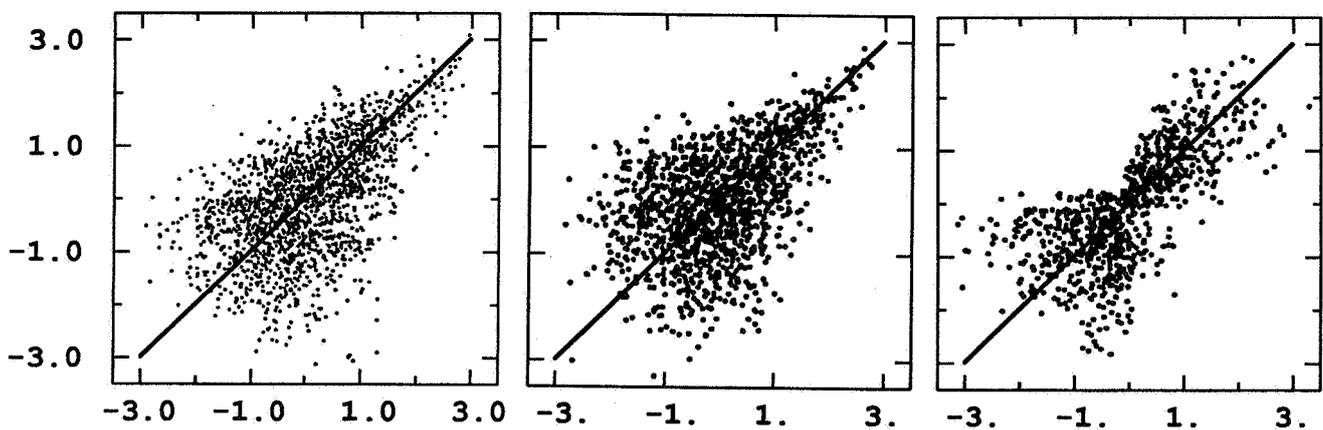


Figure 4.9 : Nuages de corrélation entre la gaussienne de la variable sous-jacente et la gaussienne de la variable bruitée

On doit noter cependant que la variance d'estimation de l'espérance conditionnelle empirique, dont la valeur est déterminée comme le palier du variogramme temporel, sous-estime parfois l'erreur d'estimation expérimentale.

Tableau 4.3: Erreur d'estimation

*(image multigaussienne, anamorphose non linéaire)*1- estimation de la variable sous-jacente

	KS			KD			EC		
	E	σ_{KS}^2	D^2	E	σ_{KD}^2	D^2	E	σ_{EC}^2	D^2
anamorphose									
lognormale	-0.0152	0.45	0.46	0.0080	0.49	0.48	0.0600	0.35	0.40
cubique	0.0800	19.96	20.58	0.0200	19.32	15.86	0.1100	15.50	13.14
discrète	-0.0200	0.86	0.85	-0.0100	0.79	0.78	-0.0009	0.79	0.71

2- estimation de l'indicatrice $\mathbb{1}_{X>1.64}$

	KS			KD			EC		
	E	σ_{KS}^2	D^2	E	σ_{KD}^2	D^2	E	σ_{EC}^2	D^2
anamorphose									
lognormale	-0.0010	0.0289	0.0282	0.0010	0.0269	0.0250	0.0008	0.0210	0.0226
cubique	0.0022	0.0297	0.0295	0.0010	0.0286	0.0246	0.0010	0.0240	0.0210
discrète	0.0008	0.0379	0.0377	0.0010	0.0242	0.0378	-0.0009	0.0290	0.0316

3- estimation de l'histogramme local ($V = 5 \times 5$)

	KS		KD		EC		
	E	D^2	E	D^2	E	σ_{EC}^2	D^2
anamorphose							
lognormale	-0.0013	0.0021	0.0003	0.0016	0.0004	0.0010	0.0011
cubique	0.0020	0.0023	0.0005	0.0014	0.0004	0.0008	0.0010
discrète	0.0005	0.0054	0.0008	0.0042	-0.0006	0.0020	0.0022

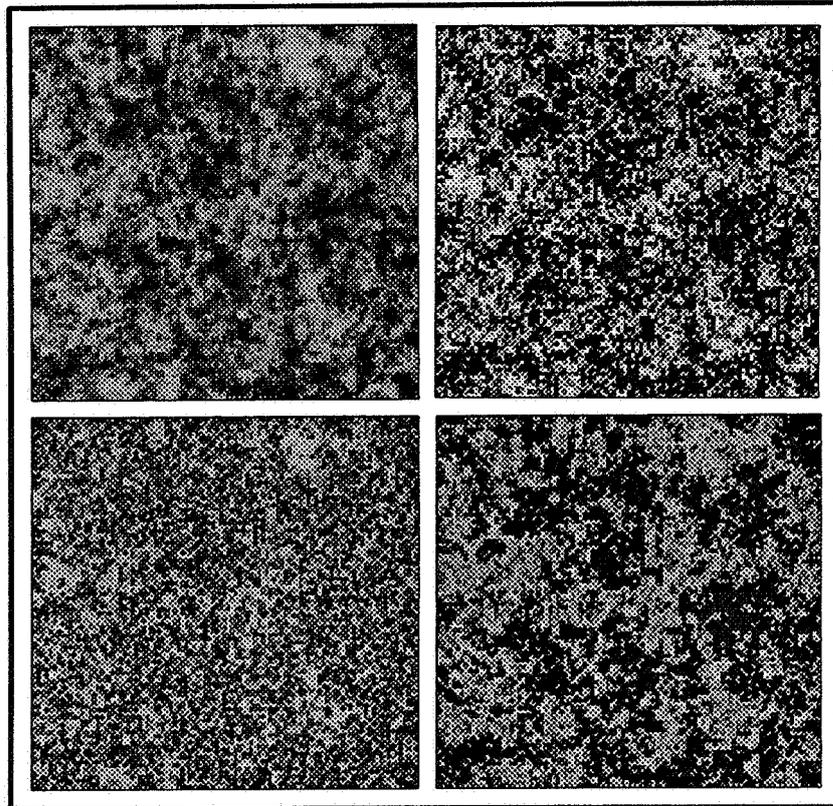
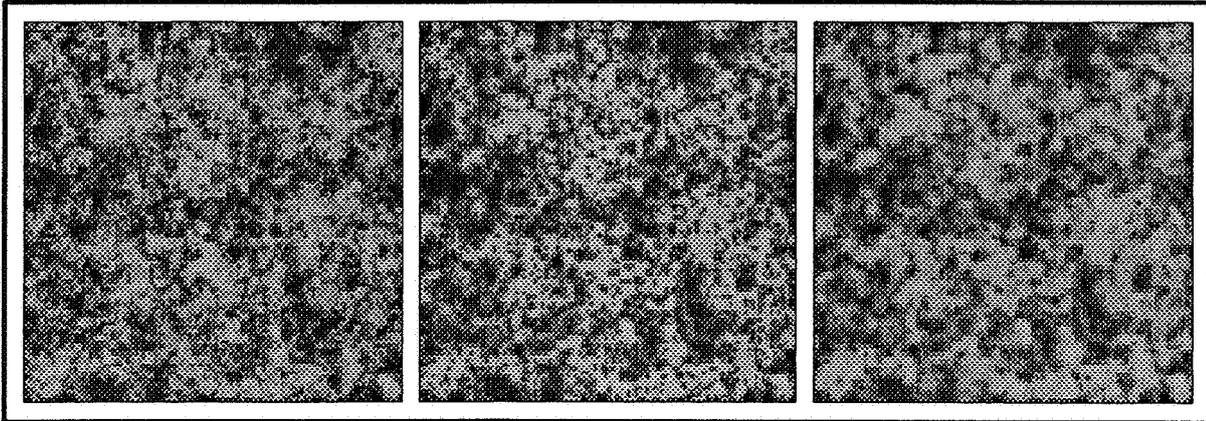


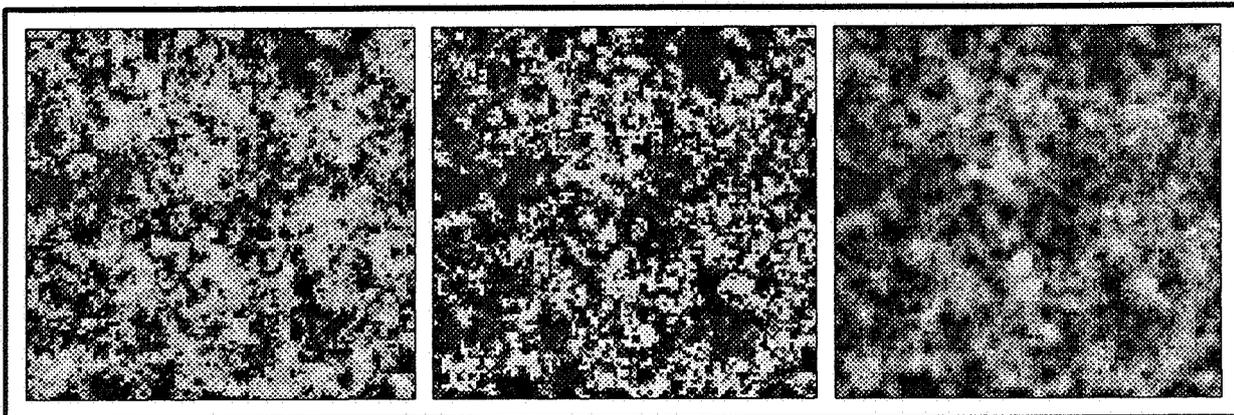
Figure 4.10: Images bruitées simulées

- en haut à gauche* : variable gaussienne
a - en haut à droite : variable bruitée — anamorphose lognormale
b - en bas à gauche : variable bruitée — anamorphose cubique
c - en bas à droite : variable bruitée — anamorphose discrète

a - anamorphose lognormale



b - anamorphose cubique



c - anamorphose discrète

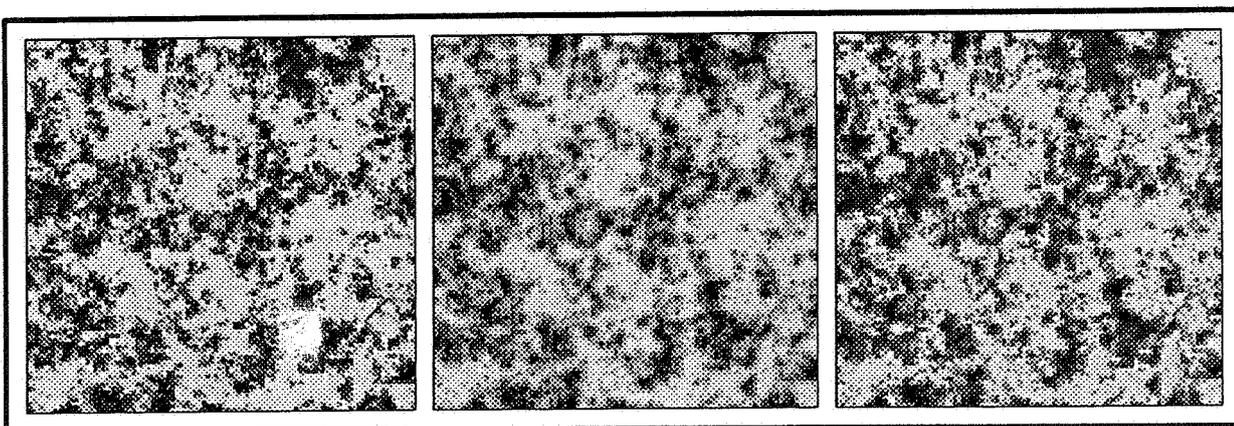
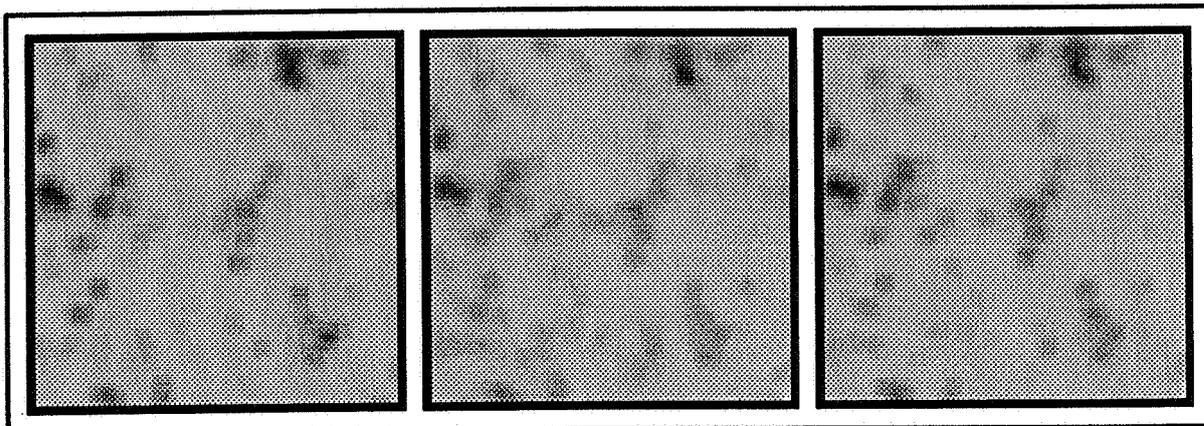
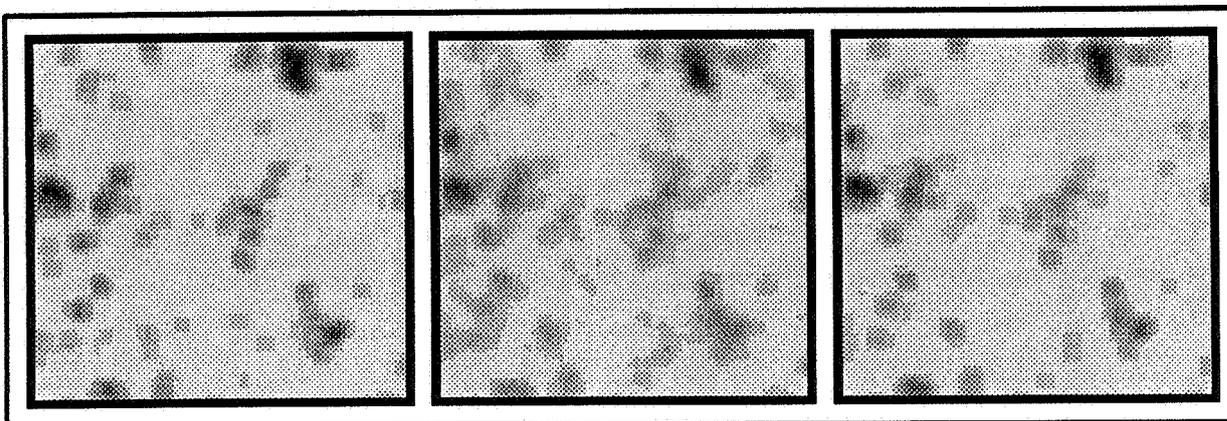
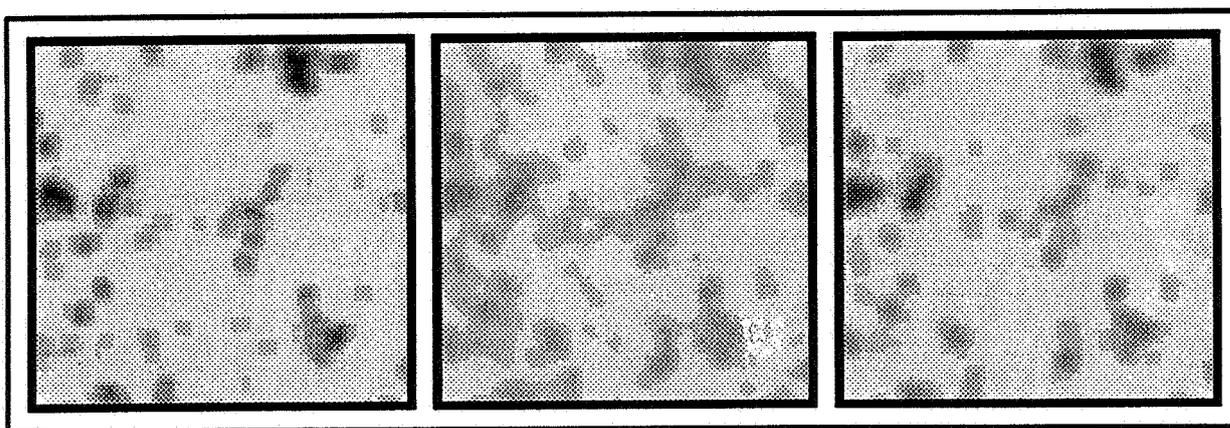


Figure 4.11 : Estimation de la variable sous-jacente

- à gauche* : variable à estimer
- au centre* : krigeage simple
- à droite* : espérance conditionnelle empirique

a - anamorphose lognormaleb - anamorphose cubiquec - anamorphose discrèteFigure 4.12: Estimation de l'histogramme local ($V = 5 \times 5$)

à gauche : variable à estimer
au centre : krigeage simple
à droite : espérance conditionnelle empirique

Influence du modèle

Le calcul de l'espérance conditionnelle empirique s'appuie sur le modèle multigaussien de l'image sous-jacente. Les exemples suivants illustrent comment réagissent les estimateurs sur des images non multigaussiennes. Les images mosaïques considérées sont construites sur la partition générée par des droites poissonniennes. Les polygones poissonniens sont valués par des variables aléatoires gaussiennes mutuellement indépendantes. La covariance de cette fonction aléatoire stationnaire est exponentielle: $Cov(X_x, X_{x+h}) = \exp(-2\theta|h|)$ où θ est la densité de Poisson des droites. A l'aide d'images mosaïques indépendantes, $(M^{(i)})_{i \geq 1}$, on construit les images $\tilde{X}^{(n)} = \frac{1}{\sqrt{n}} \sum_{i=1}^n M^{(i)}$. La suite des $(\tilde{X}^{(n)})_{n > 0}$ converge en loi vers l'image multigaussienne de même covariance exponentielle.

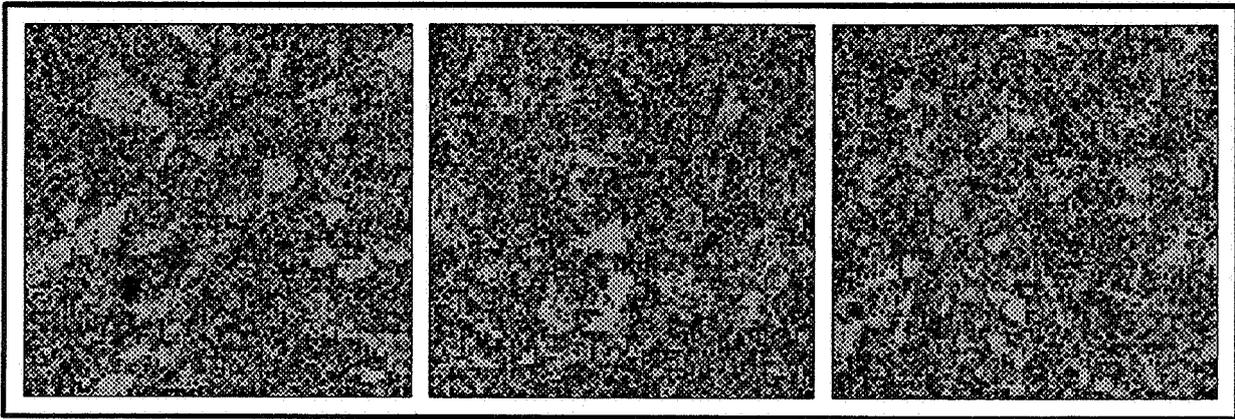


Figure 4.13 : Images mosaïques bruitées
à gauche: $n=1$, au centre: $n=2$ et à droite: $n=4$

La variable sous-jacente obtenue par anamorphose cubique de l'image mosaïque — $n = 1, 2$ ou 3 — est estimée à partir de la variable bruitée (figure 4.13. Pour les deux estimateurs, krigeage simple et krigeage disjonctif, on calcule l'erreur d'estimation expérimentale (tableau 4.4). Les images obtenues sont reportées à la figure 4.14. On constate que, bien que le modèle sous-jacent ne soit pas multigaussien — ni bigaussien —, l'espérance conditionnelle empirique fournit là encore l'estimation la plus précise. L'estimateur proposé, le calcul de l'espérance conditionnelle par la méthode de Metropolis, semble aussi robuste que les estimateurs classiques par rapport au modèle d'image. Ici la modélisation du bruit, c'est-à-dire le conditionnement des simulations, semble l'emporter sur le modèle de l'image.

Tableau 4.4 : Erreur d'estimation de la variable sous-jacente (*image mosaïque, anamorphose cubique*)

n	KS			EC		
	E	σ_{KS}^2	D^2	E	σ_{EC}^2	D^2
1	-0.0074	14.53	15.13	-0.0900	9.50	11.38
2	-0.0059	19.64	19.61	-0.0064	15.50	12.72
4	-0.0700	22.01	22.82	-0.1100	17.50	15.50

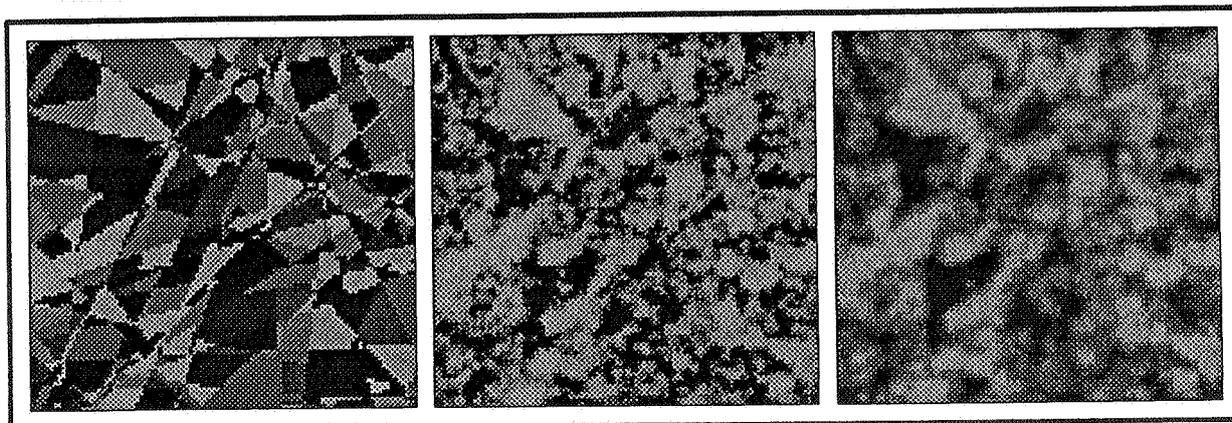
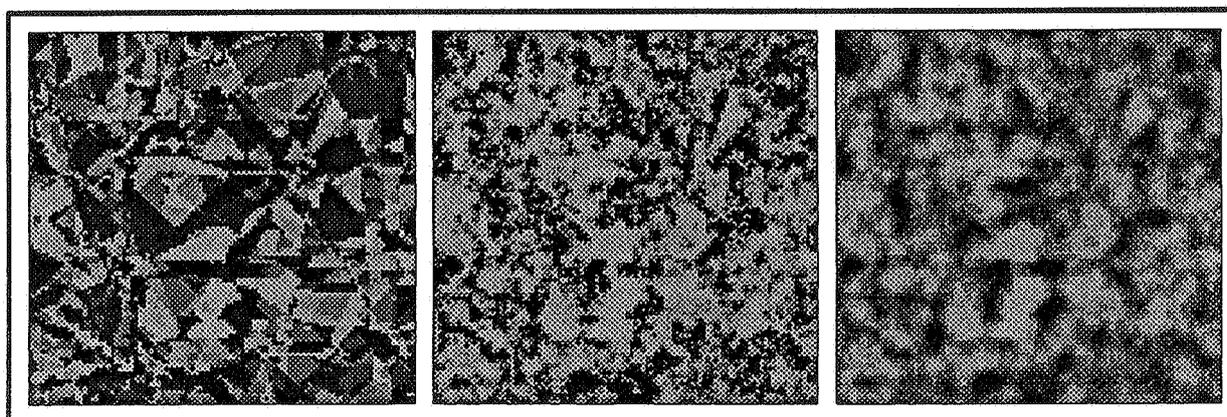
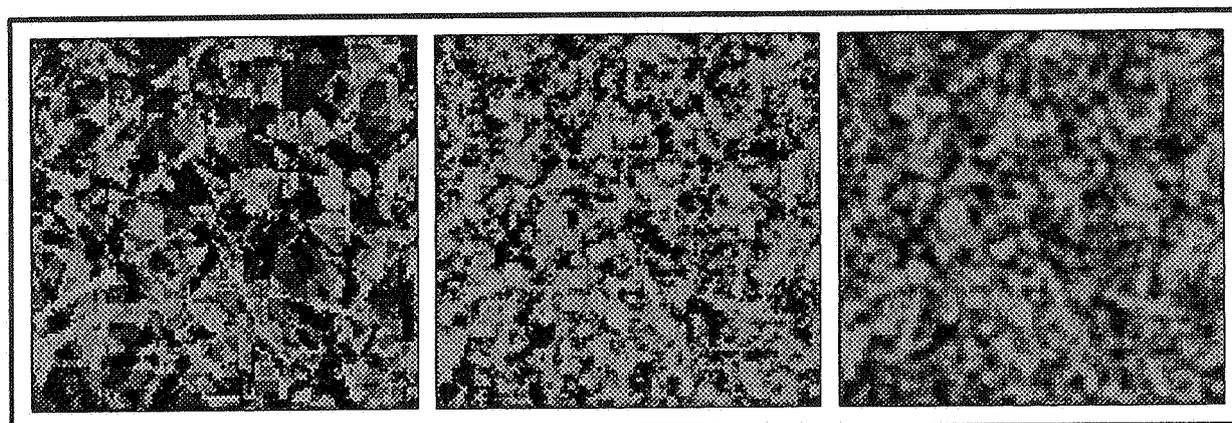
a - $n=1$ b - $n=2$ c - $n=4$ 

Figure 4.14 : Estimation de la variable sous-jacente sur des images mosaïques

à gauche : variable à estimer

au centre : krigeage simple

à droite : espérance conditionnelle empirique

4.4 Une image *micro-sonde*

L'espérance conditionnelle est calculée maintenant sur des données réelles : l'image, fournie par l'I.R.S.I.D., est obtenue par l'étude d'un échantillon métallique à la micro-sonde (figure 4.17). Les statistiques calculées sur cette image de taille 256×256 sont données au tableau 4.5 et les variogrammes expérimentaux à la figure 4.6. La variable mesurée est la poissonisée d'une concentration.

Tableau 4.5: Image *micro-sonde*, Statistiques

	nbr.	Minimum	Maximum	Moyenne	Variance
image bruitée	65377	801.	1815.	1053.	7623.
krigeage simple	65536	911.	1554.	1053.	6051.
krigeage disjonctif	65536	960.	1476.	1051.	5782.
espérance conditionnelle	65536	969.	1420.	1053.	6112.

(Pour 159 pixels, la variable bruitée n'est pas définie.)

La loi de cette variable sous-jacente, estimée par moindres carrés régularisés, à été fournie par C. Daly [10]. On a :

$$\begin{cases} m &= 1051 \\ \sigma_Y^2 &= 5721 \end{cases}$$

L'image à estimer est modélisée comme l'anamorphosée d'une fonction aléatoire stationnaire gaussienne : $\alpha \in A$, $Y_\alpha = \phi(X_\alpha)$. La fonction d'anamorphose (figure 4.16) est déduite de la loi de Y . Les variogrammes expérimentaux sont ajustés à l'aide du modèle stationnaire suivant :

$$h \neq 0 \quad m + \sum_{n \geq 1} \psi_n^2 [1 - \rho^n(h)]$$

où la covariance¹¹ de la gaussienne est :

$$\rho = 0.02sph(4) + 0.36sph(6) + 0.62exp(20)$$

(La distance entre deux pixels adjacents est égale à 1.)

¹¹ $sph(a)$ et $exp(a)$ désignent respectivement les fonctions de covariance :

$$\frac{[1 - 3/2|h|/a + 1/2(|h|/a)^3] \mathbf{1}_{|h| < a}}{\exp[-|h|/a]}$$

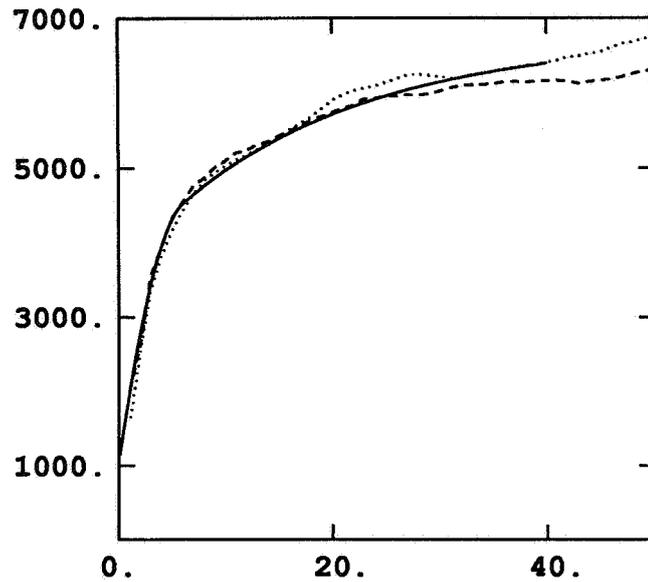


Figure 4.15: Variogrammes de l'image micro-sonde

- : variogramme calculé dans la direction horizontale
- : variogramme calculé dans la direction verticale
- : modèle

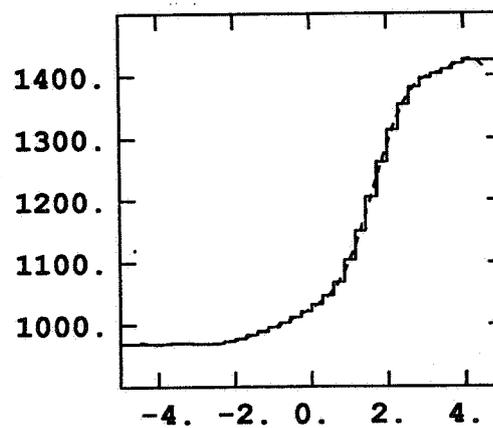


Figure 4.16: Fonction d'anamorphose de la concentration sous-jacente

- : fonction d'anamorphose
- : fonction modélisée par les 30 premiers polynômes d'Hermite

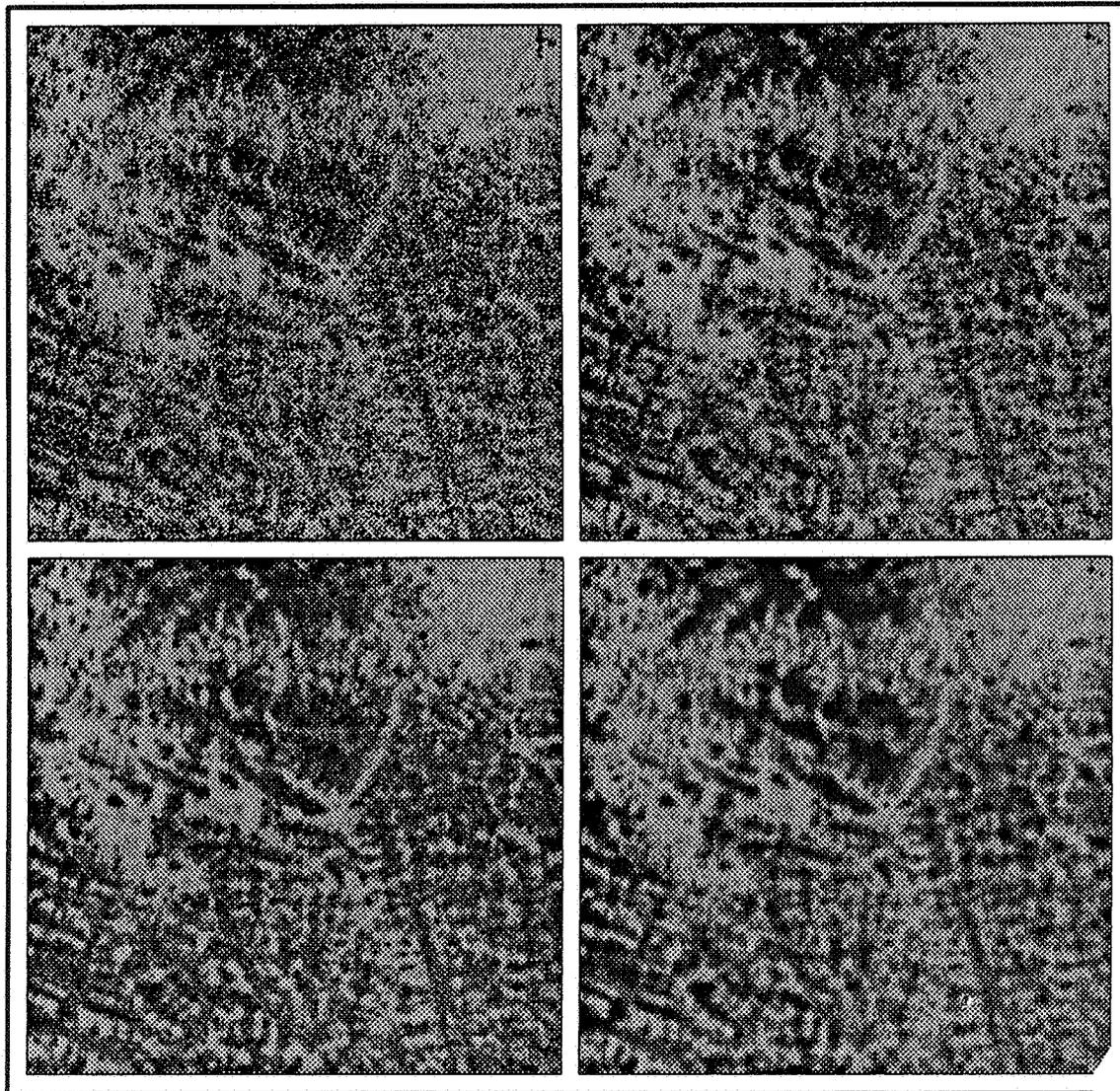


Figure 4.17: Image microsonde et trois estimateurs de la concentration sous-jacente

- en haut à gauche : image bruitée
- en haut à droite : krigeage simple
- en bas à gauche : krigeage disjonctif
- en bas à droite : espérance conditionnelle empirique

Cette image a été filtrée à l'aide des trois méthodes comparées aux paragraphes précédents (figure 4.17). La variance d'estimation et le rapport *signal sur bruit* sont reportés au tableau 4.6.

Tableau 4.6: Image *micro-sonde*, performances des estimateurs

estimateur	variance d'estimation	rapport signal/bruit
image bruitée	1051.	5.4
krigeage simple	500.	11.4
krigeage disjonctif	650.	8.8
espérance conditionnelle	390.	14.6

- Le krigeage simple est calculé avec le voisinage 7×7 .
- Le krigeage disjonctif gaussien est calculé avec un coefficient de corrélation entre les gaussiennes égal à 0.91 et un voisinage 9×9 . Cependant la loi bivariable des gaussiennes associées à la variable bruitée n'est pas bigaussienne (figure 4.18). La variance d'estimation, plus forte que celle obtenue par krigeage simple, n'est pas vraisemblable car le résultat du filtrage est très satisfaisant.

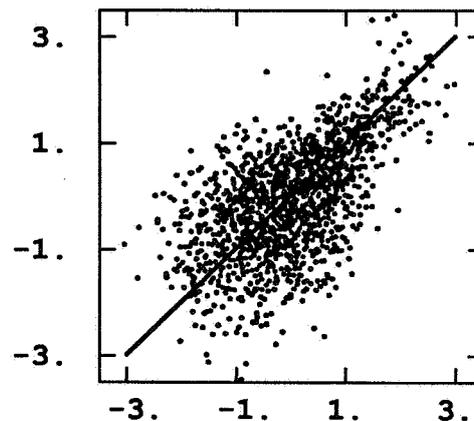


Figure 4.18: Nuage de corrélation différée des gaussiennes associées à la variable bruitée ($h = 1 \text{ pixel}$)

- L'espérance conditionnelle empirique est calculée par la méthode de Metropolis. 550 itérations dont 50 d'initialisation ont été calculées. Le variogramme temporel moyen (figure 4.19) est pépitique, son palier vaut 390. La précision de l'approximation est donc $\frac{390}{500}$.

L'espérance conditionnelle empirique permet donc un filtrage plus précis de l'image. Cependant, le krigeage simple, infiniment plus simple, fournit déjà un très bon résultat.

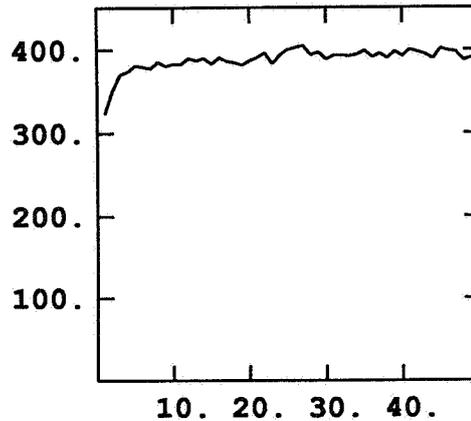


Figure 4.19: Variogramme temporel des simulations conditionnelles

4.5 Conclusion

La méthode proposée permet de produire des simulations conditionnées par l'image poissonisée. Dans ce cas la relation entre la variable observée et la variable à simuler est de nature probabiliste. Ces simulations ont été utilisées pour calculer l'espérance conditionnelle de la variable sous-jacente ou un histogramme local. Cette méthode est lourde — les temps de calcul nécessaire varient de 1 à 10 — mais les résultats obtenus sont plus précis. De plus, le modèle sur lequel s'appuie l'estimation est cohérent et peut être appliqué à une classe plus vaste de problèmes. En particulier, la poissonisée d'une variable dont les valeurs sont petites ne peut être étudiée comme une variable continue: le calcul du krigeage disjonctif gaussien est dans ce cas impossible.

Enfin, les estimateurs par espérance conditionnelle vérifient les mêmes inégalités que celles vérifiées par la variable à estimer. Ainsi, l'estimateur de l'histogramme local est compris entre 0 et 1, celui d'une variable positive est positif.

Conclusion

Lorsqu'un problème d'estimation ne peut être résolu directement, on a très souvent recours à des techniques de Monte-Carlo: l'estimateur est construit par moyenne de simulations de la loi conditionnée par les observations. Ceci implique deux hypothèses: on doit spécifier complètement le modèle probabiliste décrivant le phénomène étudié et le lien entre les observations et la variable.

Le modèle utilisé au cours de ce travail est le modèle gaussien anamorphosé; deux types de relation ont été étudiées pour expliciter le lien entre les observations et la variable simulée:

relation fonctionnelle: la variable observée est la transformée par une fonction déterministe¹² du vecteur à simuler.

randomisation: la variable observée est une variable aléatoire dont le paramètre est la variable à simuler.

La simulation de la loi conditionnelle peut alors se ramener à la simulation d'un vecteur multigaussien dont chaque composante doit vérifier une ou plusieurs inégalités. La simulation d'un tel vecteur peut s'effectuer à l'aide d'une chaîne de Markov ergodique dont la loi stationnaire est la loi conditionnelle du vecteur gaussien. La probabilité de transition de Geman permet de construire aisément une telle chaîne.

Ces techniques de simulation ont été utilisées sur deux exemples:

- la simulation conditionnelle d'une corégionalisation entre la teneur et un faciès lithologique du gisement de Laporte.
- la simulation d'une concentration conditionnée par sa mesure entachée d'un bruit poissonien.

Cependant, une réponse satisfaisante n'a pu être donnée aux points suivants:

- Pour les deux applications, la convergence de ces méthodes itératives n'a été abordée que d'un point de vue pratique. Il serait souhaitable d'étudier théoriquement la vitesse de convergence, ou bien pour conforter le critère utilisé — la stabilisation des deux premiers moments expérimentaux — ou bien pour proposer un autre critère.

¹²Dans le cas d'une fonction aléatoire de substitution, cette fonction est un processus indépendant du vecteur à simuler.

- Pour le filtrage d'un bruit poissonien, un test de l'adéquation du modèle gaussien anamorphosé n'a pas été étudié.
- Actuellement la simulation conditionnelle par des techniques itératives est lente. Ces temps de calcul devraient être réduits par une détermination de critère plus strict de convergence ou par une optimisation des algorithmes.

Annexes : simulations non conditionnelles

En modèle gaussien, une simulation conditionnée par la connaissance de la valeur de la fonction aléatoire en des points expérimentaux peut être construite par un krigeage conditionnant (voir page 9). Il faut pour cela disposer d'une simulation non conditionnelle sur la réunion du domaine à simuler et de l'ensemble des points de mesure. De très nombreuses méthodes de simulation sont disponibles ; Boulanger[6] en donne une présentation générale. Dans les annexes suivantes, sont présentées les techniques utilisées au cours du travail. Les phénomènes étudiés se déployant dans \mathbb{R}^2 ou \mathbb{R}^3 , la méthode des bandes tournantes a été utilisée pour produire des simulations de fonctions aléatoires gaussiennes, elle permet de ramener la simulation dans \mathbb{R}^n ($n = 2, 3$) à de multiples simulations dans \mathbb{R}^1 , mais en contrepartie, il faut discrétiser la demi-sphère de \mathbb{R}^n . L'étude de la discrétisation de la demi-sphère de \mathbb{R}^3 a été l'objet d'une présentation aux journées de Juin 1991 ; l'article[19] écrit en collaboration avec Ch. de Fouquet est reproduit. Les processus sur les droites ont été simulés à l'aide des techniques proposées par Boulanger[6] : la droite réelle est discrétisée et la restriction du processus sur cette grille régulière est modélisée par des processus auto-regressifs ou des moyennes mobiles discrètes. Utilisant les liens entre ces deux méthodes, il est possible de calculer les paramètres des deux modélisations en un seul calcul. Enfin, la mise en œuvre des moyennes mobiles par des méthodes spectrales permet de réduire très sensiblement le temps de calcul. Ces méthodes sont donc étudiées à une dimension pour simuler les droites des bandes tournantes mais aussi à deux dimensions.

Annexe A

Remarques sur la pratique des bandes tournantes à 3D

Remarques sur la pratique des bandes tournantes à 3D**X. FREULON et C. de FOUQUET**

*Ecole des Mines de Paris
Centre de Géostatistique
35 rue Saint-Honoré
77305 FONTAINEBLEAU
FRANCE.*

RÉSUMÉ

La méthode de simulation par bandes tournantes permet de construire des fonctions aléatoires isotropes sur R^n à partir de processus sur R . Dans le cas de R^3 , la relation entre la covariance de la fonction aléatoire, C_3 , et celle des processus, C_1 , est très simple, par contre, il n'est pas possible de répartir uniformément plus de 15 points sur la sphère. Aussi avons nous étudié plusieurs méthodes pour distribuer des points sur la sphère et proposons une pondération minimisant l'anisotropie introduite par une discrétisation non-uniforme de la sphère. Puis pour une simulation sur un champ plan, nous comparons en pratique le passage de R^1 à R^2 et celui de R^1 et R^3 . Une alternative à l'inversion de l'équation fonctionnelle reliant C_1 et C_2 est une discrétisation fine de la demi-sphère de R^3 .

ABSTRACT

In order to simulate an isotropic random function on R^n using the turning bands method, we have to simulate random processes on the line. In the three dimensional case, the equation which links the covariance of the random function, C_3 , to the covariance of the processes, C_1 , is very simple, but it is not possible to spread more than 15 directions uniformly on the sphere. In this paper, we study several series of points on the sphere. We propose an optimal weighing, which minimizes the anisotropy introduced by these non-uniform discretizations of the sphere. Then on the plane, we compare the direct simulation of a 2D random function and the simulation induced by a 3D random function, both with the same covariance: either we have to invert the equation between C_2 and C_1 or to perform a fine discretization of the sphere.

SOMMAIRE

0 – Introduction

1 – Suites de points sur la demi-sphère de \mathbb{R}^3

- 1.1. Icosaèdre et suites de points denses sur la demi-sphère
- 1.2. Points définis par un sous-groupe dénombrable de rotations
- 1.3. Points aléatoires sur la demi-sphère

2 – Calcul de la covariance

- 2.1. Erreur quadratique moyenne
- 2.2. Pondération optimale

3 – Simulations

- 3.1. Simulation des droites
- 3.2. Mise en œuvre des simulations

4 – Conclusion

5 – Bibliographie

0 - Introduction

G. Matheron a montré (1972) que pour toute covariance, C, stationnaire et isotrope dans \mathbb{R}^n , il existe une covariance sur \mathbb{R}^1 , C_1 , telle que :

$$C(h) = C_n(|h|) = 2 \int_{S_{n-1}^+} C_1(\langle h, s \rangle) d\varpi_{n-1}(s) \tag{1}$$

où S_{n-1}^+ est la demi-sphère unité de \mathbb{R}^n ($S_{n-1}^+ = \{x \in \mathbb{R}^n \mid \langle x, x \rangle = 1, \langle x, u \rangle \geq 0\}$, u étant un axe arbitraire) et ϖ_{n-1} est la loi uniforme sur la sphère S_{n-1} .

S'il existe une fonction $f \in L^2(\mathbb{R})$ telle que $C_1 = f \cdot \check{f}$, il est possible de construire une F.A. isotrope de \mathbb{R}^n , de covariance C, à l'aide d'une mesure aléatoire, M, orthogonale sur $S_{n-1}^+ \times \mathbb{R}$ et telle que, si A et B sont deux boréliens de $S_{n-1}^+ \times \mathbb{R}$, $Cov(M(A), M(B)) = |A \cap B|$:

$$Z_x = \sqrt{2} \int_{S_{n-1}^+ \times \mathbb{R}} f(\langle x, s \rangle - t) dM(s, t)$$

La méthode de simulation par bandes-tournantes est le calcul des intégrales sur la demi-sphère par quadrature. On approxime la mesure aléatoire M par la mesure : $M^{BT} = \frac{1}{\sqrt{2}} \sum_{i=1}^n (\omega_i)^{1/2} \delta_{\alpha_i} \otimes T^{(i)}$ où les $T^{(i)}$ sont des mesures aléatoires orthogonales mutuellement indépendantes sur \mathbb{R} (si $(F_i, F_j) \in (\mathfrak{B}_{\mathbb{R}})^2$ alors $E\{T^{(i)}(F_i) \cdot T^{(j)}(F_j)\} = \delta^{ij} |F_i \cap F_j|$) et $(\alpha_i)_{1 \leq i \leq n}$ n points de la demi-sphère.

La FA simulée par bandes-tournantes est $Z_x^{BT} = \sum_{i=1}^n (\omega_i)^{1/2} \int f(\langle \alpha_i, x \rangle - t) dT^{(i)}(t)$

La simulation se décompose donc en deux étapes :

- simulation de n processus indépendants de covariance C_1 : $1 \leq i \leq n \quad X_i^{(t)} = f * T^{(i)}(t)$
- calcul de la simulation sur \mathbb{R}^n , par "épannage" des processus simulés : $Z_x^{BT} = \sum_{i=1}^n (\omega_i)^{1/2} X_{\langle \alpha_i, x \rangle}^{(i)}$

La covariance de la F.A. ainsi construite est :

$$C^{BT}(x-y) = Cov(Z_x^{BT}, Z_y^{BT}) = \sum_{i=1}^n \omega_i C_1(\langle \alpha_i, x-y \rangle)$$

qui est une approximation de $2 \int_{S_{n-1}^+} C_1(\langle \alpha, x-y \rangle) d\varpi_{n-1}(\alpha)$.

De plus, la relation (1) entre les covariances C_1 et C_n reste valable dans le cas de covariances généralisées (Matheron 1973).

Pour mettre en œuvre cette méthode, il faut donc résoudre deux problèmes :

- l'inversion de l'équation (1), c'est-à-dire le calcul de la covariance, C_1 , à simuler sur les droites,
- le choix de la quadrature $(\alpha_i, \omega_i)_{1 \leq i \leq n}$.

Dans cet article, nous étudions tout d'abord différentes quadratures pour la simulation d'une F.A. isotrope dans \mathbb{R}^3 . Puis, nous regardons en pratique la simulation par la méthode des bandes tournantes d'une FA dans le plan : nous comparons la simulation par passage de \mathbb{R}^1 à \mathbb{R}^2 et la simulation par passage de \mathbb{R}^1 à \mathbb{R}^3 . La FA à simuler est alors considérée comme la restriction au plan d'une FA à 3D.

1 – Suites de points sur la demi-sphère de \mathbb{R}^3

Nous avons étudié trois types de suites de points :

- les points obtenus par recouvrement des faces de l'icosaèdre
- les points de l'orbite du sous-groupe dénombrable engendré par trois rotations d'angle $\text{Arc cos}(-3/5)$ autour des trois axes orthogonaux (Ox, Oy, Oz)
- les points dérivant des méthodes de simulations de points aléatoires uniformément répartis sur la sphère.

Ces suites sont denses sur la sphère, on obtient donc, moyennant une pondération adéquate, une approximation de l'intégrale recherchée. De plus, le troisième type fournit des suites équi-réparties c'est à dire :

$$\forall f \in \mathcal{C}(S_2^+) \quad \lim_{n \rightarrow \infty} \left(\frac{1}{n} \sum_{i=1}^n f(x_i) \right) = \int_{S_2^+} f(x) \, d\omega(x)$$

Dans ce cas, la pondération des directions est simple.

1.1. Icosaèdre et suites de points denses sur la demi-sphère

Dans \mathbb{R}^2 , une discrétisation de la sphère est obtenue en considérant les sommets des polygones réguliers. Cependant dans \mathbb{R}^3 , le nombre des polyèdres réguliers est limité à cinq. L'icosaèdre, comportant 20 sommets et 20 faces, qui sont des triangles équilatéraux, définit la plus fine discrétisation régulière de la sphère : les sommets de l'icosaèdre et les milieux des arêtes fournissent respectivement 10 et 15 directions régulièrement réparties (Guibal, 1972).

Afin d'aller au-delà de cette limite des 15 droites, tout en conservant les propriétés de symétrie de l'icosaèdre, plusieurs solutions ont été envisagées :

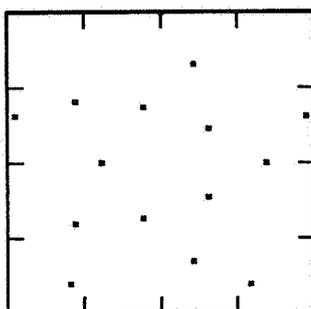
- implanter plusieurs icosaèdres selon des directions uniformément réparties sur la sphère (un icosaèdre est un solide, on tire une direction aléatoire, puis une rotation aléatoire autour de cet axe) ;
- recouvrir les 10 triangles sphériques définis par les faces de l'icosaèdre par des points répartis de manière la plus uniforme possible. Deux méthodes sont présentées ici :
 - la triangulation des faces de l'icosaèdre permet de définir des directions qui, denses sur la sphère, ne sont pas équi-réparties
 - des points aléatoires uniformément répartis sur les polygones sphériques correspondant aux faces du polyèdre sont simulés en utilisant la méthode d'acceptation et rejets. Ceci correspond donc à un échantillonnage aléatoire-stratifié de la demi-sphère.

La projection des points obtenus par ces méthodes est reportée sur la figure 1. La projection utilisée conserve les surfaces :

$$(u, v, w) \in S_2^+ \mapsto \left(\sqrt{1-w} \frac{u}{\sqrt{u^2+v^2}}, \sqrt{1-w} \frac{v}{\sqrt{u^2+v^2}} \right)$$

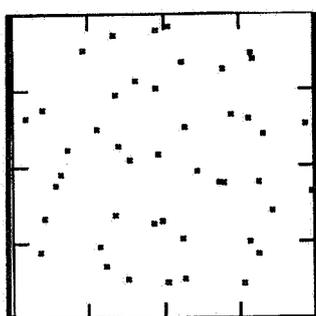
figure 1 : icosaèdre et points sur la demi-sphère de R^3

- milieu des arêtes de l'icosaèdre (I.0)

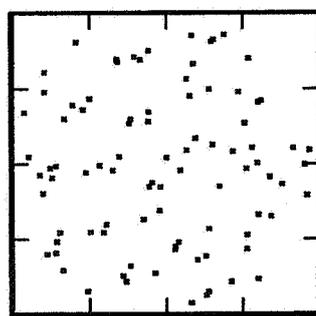


15 points

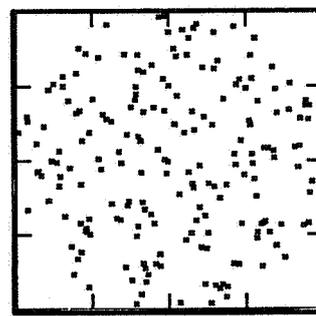
- milieu des arêtes de plusieurs icosaèdres aléatoires (I.1)



45 points

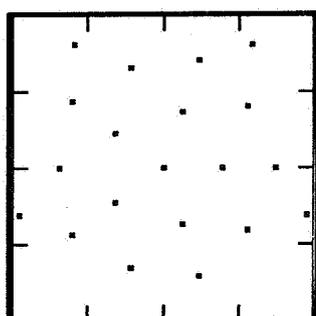


90 points

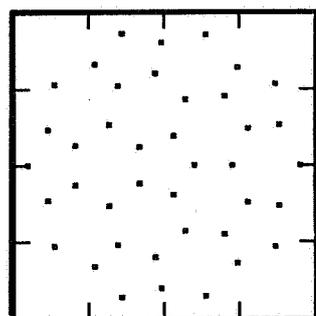


195 points

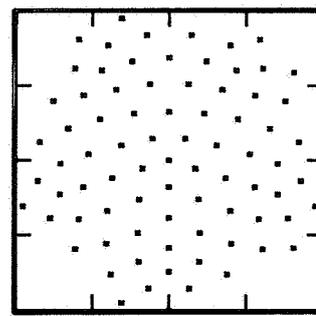
- recouvrement des faces par triangulation (I.2)



40 points

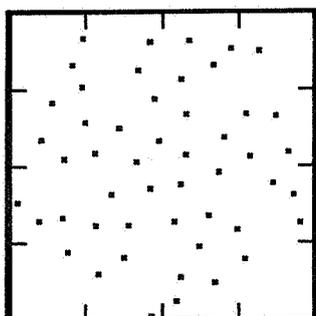


81 points

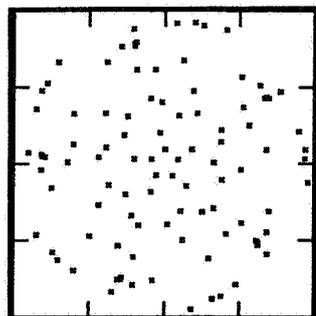


160 points

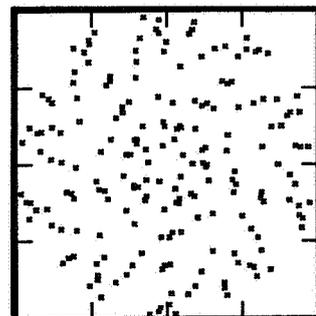
- recouvrement des faces par des points pseudo-aléatoires (I.3)



50 points



100 points



200 points

1.2. Points définis par un sous-groupe dénombrable de rotations (Lubotsky et al.,1986)

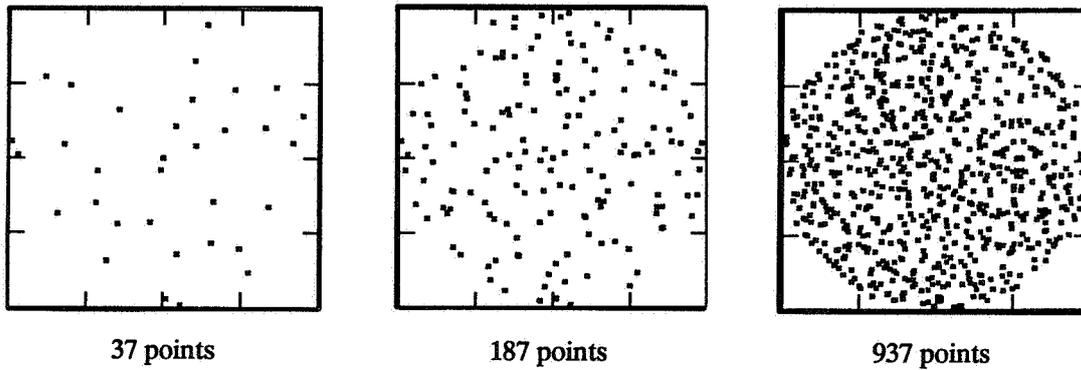
Soit Γ le sous-groupe dénombrable engendré par les trois rotations d'angle $\text{Arc cos } (-3/5)$ autour des axes O_x, O_y et O_z , notées A, B et C. Pour tout point x de la sphère, on construit le sous-ensemble de la sphère, $\{ \gamma x \mid \gamma \in \Gamma \}$, des images de x par les éléments de Γ . Il est possible d'ordonner cet ensemble pour construire une suite dense sur la sphère.

On considère les éléments γ de Γ , différents de l'identité et pouvant s'écrire comme le produit (non commutatif) de s éléments de $\{ A, B, C, A^{-1}, B^{-1}, C^{-1} \}$, deux termes consécutifs ne pouvant être l'inverse l'un de l'autre. $s(\gamma)$ est la longueur du mot représentant γ (l'identité est de longueur nulle). On considère la suite :

$$(\{ \gamma x \mid s(\gamma) = n \})_{n \geq 0}$$

Cependant la croissance du nombre de directions en fonction de s, $1 + (3/2)(5^s - 1)$, est très rapide : un grand nombre de points est nécessaire avant que les directions semblent uniformément répartie sur la sphère (figure 2).

figure 2 : orbite de l'action d'un groupe discret de rotations (II)



1.3. Points aléatoires sur la demi-sphère

Il existe des fonctions, f , continues sur $]0,1[^d$ sur S_2 telles que l'image de la mesure de Lebesgue sur $]0,1[^d$ soit la mesure de Lebesgue sur S_2^+ . Ces fonctions fournissent des méthodes pour simuler des points uniformément répartis sur la sphère. Nous avons utilisé deux méthodes de simulation de points aléatoires sur la sphère et trois suites équi-réparties sur $]0,1[^d$:

Ces deux simulations de points aléatoires sur la sphère sont les suivantes :

- La direction \underline{a} du vecteur aléatoire dont les coordonnées sont trois gaussiennes centrées, normées et indépendantes est uniformément répartie sur la sphère (Lantuejoul, 1982). Ces trois gaussiennes peuvent être simulées par la méthode d'inversion. On peut donc simuler \underline{a} à l'aide de $u = (U_1, U_2, U_3)$ uniformément réparti sur $]0,1[^3$: $i = 1, 2$ ou 3 $X_i = G^{-1}(U_i)$ et

$$\underline{a} = f_1(u) = \left(\frac{X_1}{\sqrt{X_1^2 + X_2^2 + X_3^2}}, \frac{X_2}{\sqrt{X_1^2 + X_2^2 + X_3^2}}, \frac{X_3}{\sqrt{X_1^2 + X_2^2 + X_3^2}} \right)$$

- La relation entre les coordonnées cartésiennes et les coordonnées sphériques fournit une autre méthode pour simuler un point uniformément sur la demi-sphère :

$$\underline{a} = f_2(u) = (\cos(2\pi U_1) \sqrt{1 - U_2^2}, \sin(2\pi U_1) \sqrt{1 - U_2^2}, U_2)$$

où $u = (U_1, U_2)$ est uniformément réparti sur $]0,1[^2$.

Trois suites de points équi-réparties sur $]0,1[^d$ (c'est à dire que les mesures discrètes $\frac{1}{n} \sum_1^n \delta_{x_i}$ convergent étroitement vers la mesure de Lebesgue sur $]0,1[^d$) ont été utilisées (Bouleau, 1987) :

- tirage d'une suite de points aléatoires indépendants uniformément répartis sur $]0,1[^d$;
- partage régulier de $]0,1[^d$:

$$\left(\left(\frac{i_1}{n+1}, \dots, \frac{i_d}{n+1} \right)_{1 \leq i_1, \dots, i_d \leq n} \right)_{n > 0}$$

- la suite de Van der Corput :

Soient les d premiers nombres premiers, $2, 3, \dots, n \in \mathbb{N}$ s'écrit en base $p_i (1 \leq i \leq d)$:

$$n = a_0 + a_1 p_1 + \dots + a_k p_1^k + \dots$$

On pose :

$$\varphi_{p_i}(n) = \frac{a_0}{p_i} + \dots + \frac{a_k}{p_i^{k+1}} + \dots$$

La suite de Van der Corput d -dimensionnelle est alors la suite : $(\varphi_{p_1}(n), \dots, \varphi_{p_d}(n))_{n > 0}$.

Si $(x_n)_n$ est une suite équi-répartie sur $]0,1[^d$, alors $(f(x_n))_n$ est une suite équi-répartie sur S_2^+ . En combinant les deux méthodes de simulation aux trois suites équi-réparties sur $]0,1[^d$, on dispose donc de six suites équi-réparties sur la demi-sphère (figure 3).

Le tableau 1 présente un récapitulatif des ces différentes suites.

Tableau 1 : Suites de points sur la sphère

a) Point de l'icosaèdre

milieu des arêtes :

- d'un icosaèdre I.0
- de plusieurs icosaèdres aléatoires I.1

recouvrement des faces :

- par triangulation I.2
- par des points pseudo-aléatoires I.3

b) Orbite de l'action d'un groupe discret de rotation II

c) Points pseudo-aléatoires

Suites équi-réparties sur $]0,1[^d$	Méthodes de simulation	
	des "trois gaussiennes"	des "coordonnées sphériques"
points aléatoire	III.1.1.	III.2.1.
partage régulier	III.1.2.	III.2.2.
suite de Van der Corput	III.1.3.	III.2.3.

figure 3 : points pseudo-aléatoires

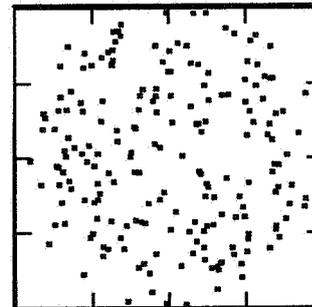
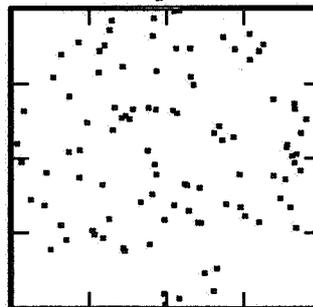
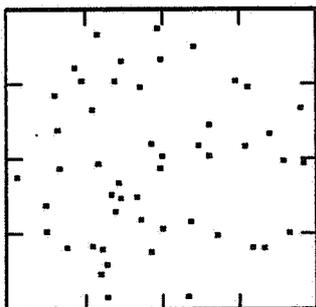
a) méthode des trois gaussiennes

50 points

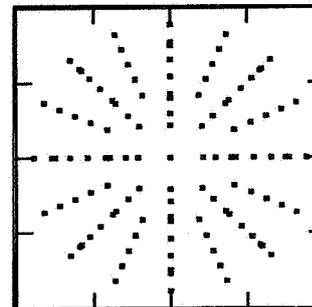
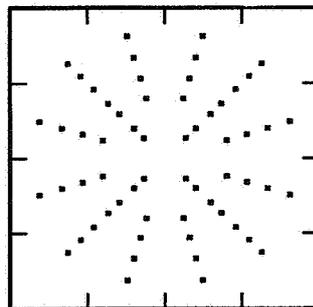
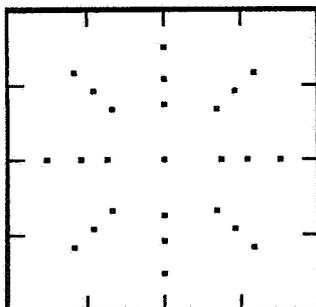
100 points

200 points

points aléatoires III.1.1



partage régulier III.1.2



suite de Van der Corput III.1.3

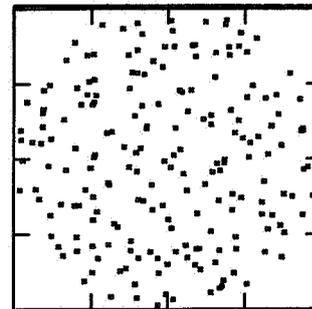
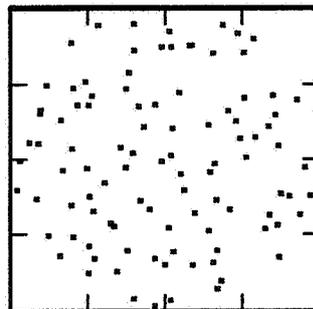
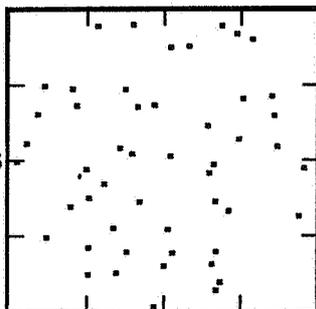
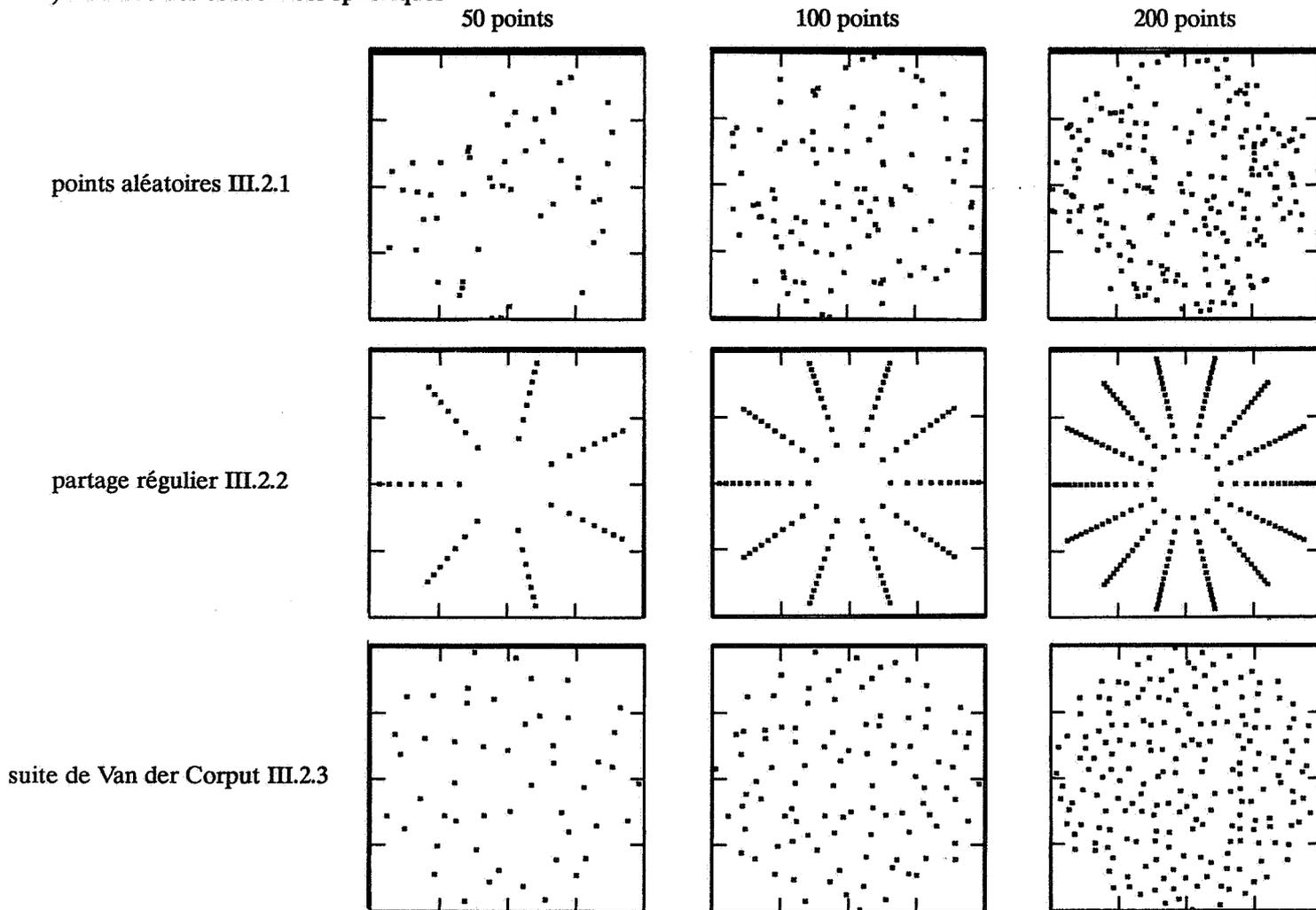


fig. 3 : points pseudo-aléatoires

b) méthode des coordonnées sphériques



2 - Calcul de la covariance

Afin de comparer l'efficacité pour le calcul de l'intégrale par quadrature de chacune des suites de points présentées au paragraphe précédent, nous avons utilisé l'erreur quadratique moyenne sur le champ à simuler entre le modèle de covariance et la modélisation de la covariance par la méthode des bandes tournantes.

2.1. Erreur quadratique moyenne

Soit $(a_i)_{1 \leq i \leq n}$ n directions de \mathbf{R}^3 . Soit D le champ, compact, sur lequel la F.A. est simulée. On définit le domaine de la covariance : $E = \{x-y \mid (x, y) \in D^2\}$. Soit P la restriction, normée, de la mesure de Lebesgue à E . L'erreur quadratique moyenne de la modélisation par bandes tournantes est :

$$\|C - C^{BT}\|_{L^2(\mathbf{R}^3, P)}^2 = \int_{\mathbf{R}^3} (C(h) - C^{BT}(h))^2 dP(h)$$

Nous n'avons pu calculer l'expression exacte de l'erreur quadratique moyenne que dans le cas d'un variogramme linéaire simulé sur domaine sphérique (le diamètre du champ intervenant par une constante multiplicative qui a été prise égale à un) :

$$\|C - C^{BT}\|_{L^2(\mathbf{R}^3, P)}^2 = 1 - 2 \sum_{i=1}^n \omega_i + \sum_{i,j=1}^n \omega_i \mu_{a_i a_j} \omega_j$$

$$\forall (s, t) \in S_2 \quad \mu_{st} = (C_1(\langle s, \cdot \rangle), C_1(\langle t, \cdot \rangle))_{L^2(\mathbf{R}^3, P)} = \frac{2}{3\pi} \left(\sin \alpha_{st} + \left(\frac{\pi}{2} - \alpha_{st}\right) \cos \alpha_{st} \right)$$

avec $\alpha_{st} = \arccos(\langle s, t \rangle_{\mathbf{R}^3})$.

Au lieu d'une pondération sur les distances, on aurait pu considérer les pondérations sur les points d'appui :

$$\|C - C^{BT}\|_g = \int (C(x-y) - C^{BT}(x-y))^2 g(x)g(y) \, dx dy = \int (C(h) - C^{BT}(h))^2 K(h) \, dh$$

où g est une densité et $K = g * g$. Dans le cas du variogramme linéaire simulé sur un champ sphérique, on montre que, si g est isotrope, toutes ces normes ne diffèrent que d'une constante multiplicative fonction de g .

Ceci nous permet donc de comparer l'efficacité des différentes suites dans le cas du variogramme linéaire.

2.2. Pondération optimale

Les directions ne sont pas strictement uniformément réparties dans \mathbf{R}^3 . En particulier, si les points sont aléatoires, il apparaît des agrégats fortuits qu'il conviendrait d'éliminer. Une solution consiste à ne pas conserver une pondération uniforme des droites. Les pondérateurs optimaux sont ceux minimisant l'erreur quadratique moyenne. On impose que ces pondérateurs soient positifs, ceci est une condition suffisante pour que C^{BT} soit une covariance, et que leur somme soit égale à un.

On peut interpréter cette minimisation en termes de Krigeage. Comme P est une probabilité sur \mathbf{R}^3 , $c_\alpha = C_1(\langle \alpha, \cdot \rangle) \in L^2(\mathbf{R}^3, P)$ et $\{c_\alpha, \alpha \in S_2\}$ est donc une FA d'ordre 2 (stationnaire dans le cas d'un champ isotrope) dont la covariance non centrée est : $E\{c_\alpha c_\beta\} = \mu_{\alpha\beta}$.

On a de plus :

$$C(h) = \frac{1}{|S_2|} \int c_\alpha(h) \, ds$$

La covariance optimale est donc le krigeage ordinaire de la valeur moyenne sur le champ de la FA $\{c_\alpha, \alpha \in S_2\}$ avec la contrainte de positivité sur les pondérateurs.

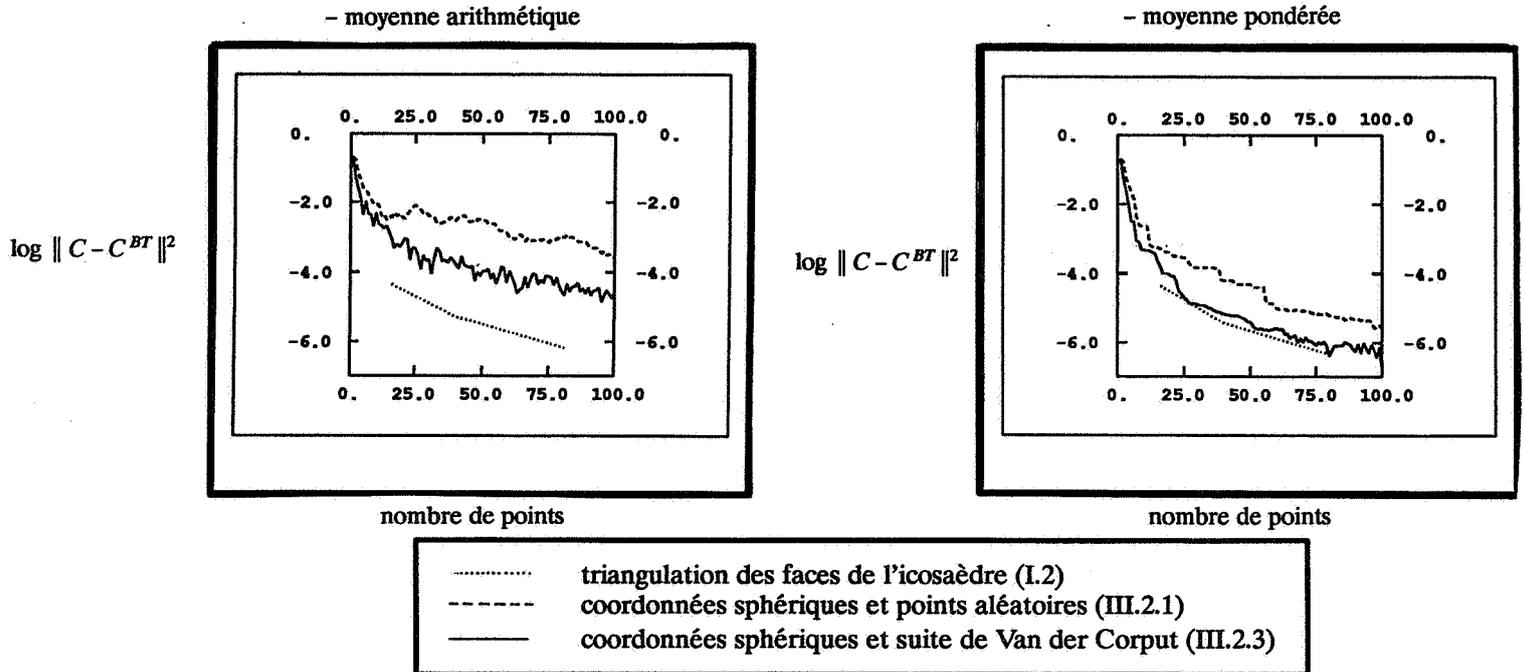
Il faut donc résoudre le problème de programmation quadratique suivant (Barnes et Johnson, 1984):

$$\begin{aligned} \text{Min} \quad & \left(\sum_{\alpha, \beta} \omega_\alpha \mu_{\alpha\beta} \omega_\beta \right) \\ \omega_\alpha \geq 0 \quad & \\ \sum_{\alpha} \omega_\alpha = 1 \quad & \end{aligned}$$

Sur la figure 4, on a reporté l'erreur quadratique moyenne, pour une moyenne arithmétique et une moyenne pondérée, en fonction du nombre de points, pour trois des suites présentées au paragraphe précédent.

Expérimentalement, à nombre de directions fixé, l'erreur quadratique moyenne minimale est obtenue avec les suites I.2 (triangulation des faces de l'icosaèdre) et III.2.3. (coordonnées sphériques, Van der Corput). La pondération accroît très sensiblement l'efficacité des suites de points pseudo-aléatoires en gommant l'anisotropie introduite par des agrégats fortuits.

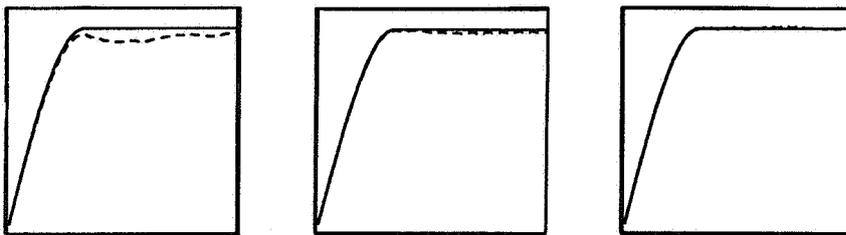
figure 4 : erreur quadratique moyenne pour le variogramme linéaire



Pour ces deux suites, on a regardé dans une direction particulière [111] le variogramme sphérique modélisé (figure 5). Dès 40 directions, les résultats sont très satisfaisants et dans ce cas la précision apportée par la pondération n'est guère appréciable.

figure 5 : variogramme sphérique modélisé dans la direction [111]

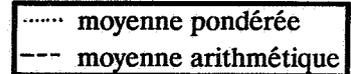
Suite I.2 : triangulation des face de l'icosaèdre



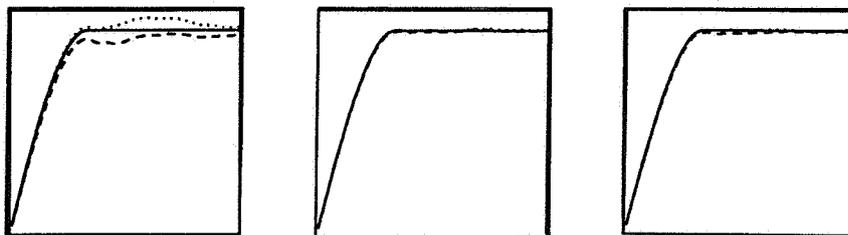
15 points

40 points

81 points



Suite III.2.3 : coordonnées sphériques et suite de Van der Corput



15 points

40 points

81 points

Les suites présentées permettent une discrétisation fine de la demi-sphère de \mathbb{R}^3 . Deux suites nous sont apparues particulièrement efficaces : I.2 (triangulation des faces de l'icosaèdre) et III.2.3. (coordonnées sphériques et suites de Van der Corput). Avec moins de 100 directions, il est possible de modéliser avec précision la covariance à simuler.

3 – Simulations

Dans cette partie, on compare deux types de simulations par bandes tournantes sur un champ plan. Dans le cas où les modèles à simuler sont autorisés dans R^2 et R^3 , on peut simuler dans le plan par passage de R à R^2 ou par passage de R à R^3 , le plan étant alors considéré comme un sous-ensemble de R^3 .

Dans le premier cas, la simulation 2D, le principal problème est la modélisation de la covariance à simuler sur les droites. La relation entre les covariances C_1 et C_2 (tableau 2) ne permet pas toujours un calcul explicite de C_1 . Cependant pour simuler les processus sur une grille régulière, il est nécessaire de connaître la valeur de la covariance, C_1 , en un nombre fini de points. Ces valeurs peuvent être calculées numériquement : connaissant la dérivée de C_2 , on calcule l'intégrale, et donc C_1 , par quadrature de Gauss (Les polynômes orthogonaux utilisés sont les polynômes de Tchebichef de 1^{ère} espèce). Par contre, il est très facile de répartir régulièrement des points sur le demi-cercle.

Dans le second cas, la simulation 3D, la relation entre C_1 et C_3 permet un calcul explicite de C_1 . Les points de la demi-sphère de R^3 et la pondération associée, présentés dans les parties précédentes, sont utilisés pour calculer l'intégrale sur la demi-sphère.

Tableau 2 : relation entre les covariances pour la méthode des bandes tournantes dans R^2 et R^3

Relation entre les covariances		
R^2	$C_2(\rho) = \frac{2}{\pi} \int_0^\rho \frac{C_1(u)}{\sqrt{\rho^2 - u^2}} du$	$C_1(\rho) = C_2(0) + \rho \int_0^\rho \frac{1}{\sqrt{\rho^2 - u^2}} \frac{d}{du} (C_2(u)) du$
R^3	$C_3(\rho) = \frac{1}{\rho} \int_0^\rho C_1(u) du$	$C_1(\rho) = \frac{d}{d\rho} (\rho C_3(\rho))$

Ces simulations se décomposent donc en deux parties : simulation sur les droites, puis "épandage" sur le plan.

3.1 Simulation sur les droites

Les droites sont simulées en utilisant des processus discrets, mutuellement indépendants, $(\tilde{X}_n^{(i)})_{n \in Z}$, dont la covariance est $Cov(\tilde{X}_n^{(i)}, \tilde{X}_m^{(i)}) = C_1(\Delta t |n - m|)$. Ceci est équivalent à considérer les processus continus, constants sur des intervalles de longueur Δt :

$$X_x^{(i)} = \sum_{n \in Z} \tilde{X}_n^{(i)} 1_{[n\Delta t, (n+1)\Delta t]}(x)$$

L'épandage se fait donc sur des bandes de largeur Δt , selon la formule :

$$Z_x^{BT} = \sum_{i=1}^n (\omega_i)^{\frac{1}{2}} \tilde{X}_{\lfloor \frac{\langle a_i, x \rangle}{\Delta t} \rfloor}$$

($\lfloor \cdot \rfloor$ note ici la partie entière.). La covariance entre deux points du plan est donc :

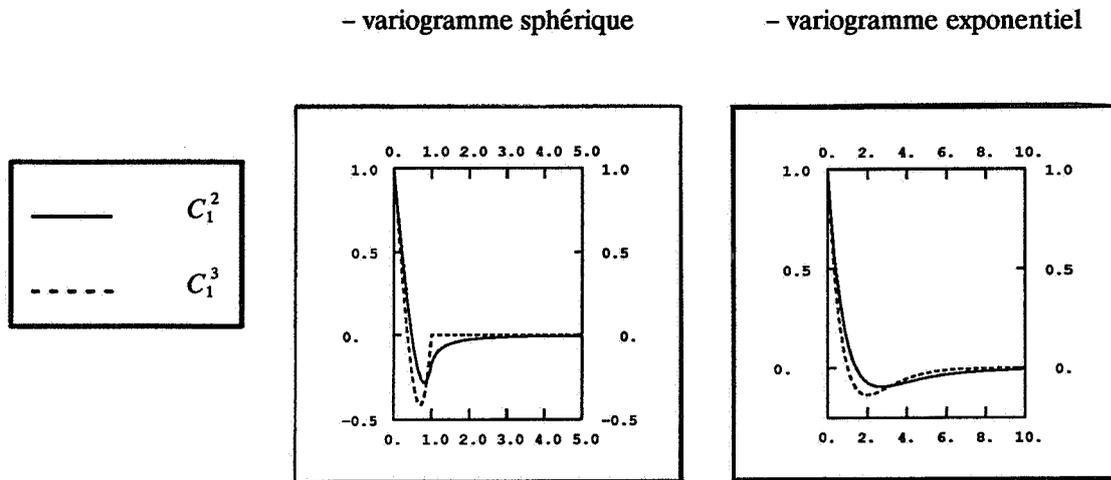
$$Cov(Z_x^{BT}, Z_y^{BT}) = \sum_{i=1}^n \omega_i C_1(\Delta t | \lfloor \frac{\langle a_i, x \rangle}{\Delta t} \rfloor - \lfloor \frac{\langle a_i, y \rangle}{\Delta t} \rfloor |)$$

Cette covariance n'est pas stationnaire et il est bien difficile de préciser le biais introduit par la discrétisation sur les droites. Cependant ce biais tend vers zéro lorsque Δt tend vers zéro. Nous avons donc choisi un pas

de discrétisation proportionnel à la portée du schéma à simuler sur le plan. Ceci est équivalent à la simulation sur les droites d'un schéma de portée unitaire : la modélisation des droites ne dépend pas de la portée à simuler sur le plan. Pour simuler des schéma de portée différente sur le plan, on peut donc utiliser la même modélisation des droites à un facteur d'échelle près. Dans le cas d'un schéma de transition, on choisit le nombre de points sur lequel cette transition est modélisée. Dans le cas d'une covariance tendant asymptotiquement vers zéro, la portée pratique retenue est la plus petite distance au delà de laquelle la valeur absolue de la covariance est inférieure à 10^{-2} . Dans cette étude, la portée pratique correspond à 20 points de discrétisation.

Pour simuler ces processus à temps discret, nous avons utilisé dans le cas du variogramme linéaire, des processus de Wiener-Levy et dans le cas d'une covariance stationnaire, les méthodes proposées par Boulanger (1990). Dans ce dernier cas, les covariances sur R ont des propriétés bien distinctes selon que l'on passe de R à R^2 ou de R à R^3 : figure 6. Les covariances pour le passage de R à R^2 tendent beaucoup moins rapidement vers zéro, elles sont beaucoup plus difficile à simuler (Par exemple, les modélisations par processus auto-régressif ou par moyenne mobile sont en général d'ordre plus élevé.)

figure 6 : covariance sur R^1 pour la simulation par bandes tournantes sur R^2 ou R^3



3.2 Mise en œuvre des simulations

Nous avons étudié en pratique le nombre de droites nécessaire.

La grille simulée est carrée de taille 256-256. Deux modèles sont utilisés :

- un variogramme sphérique de portée 10,
- un variogramme linéaire de pente 1.

Les points utilisés pour calculer l'intégrale sur la demi-sphère sont à 2D les sommets des polygones réguliers et à 3D les points des suites I.2 et III.2.3. Le nombre considéré de points varie de 15 à 81.

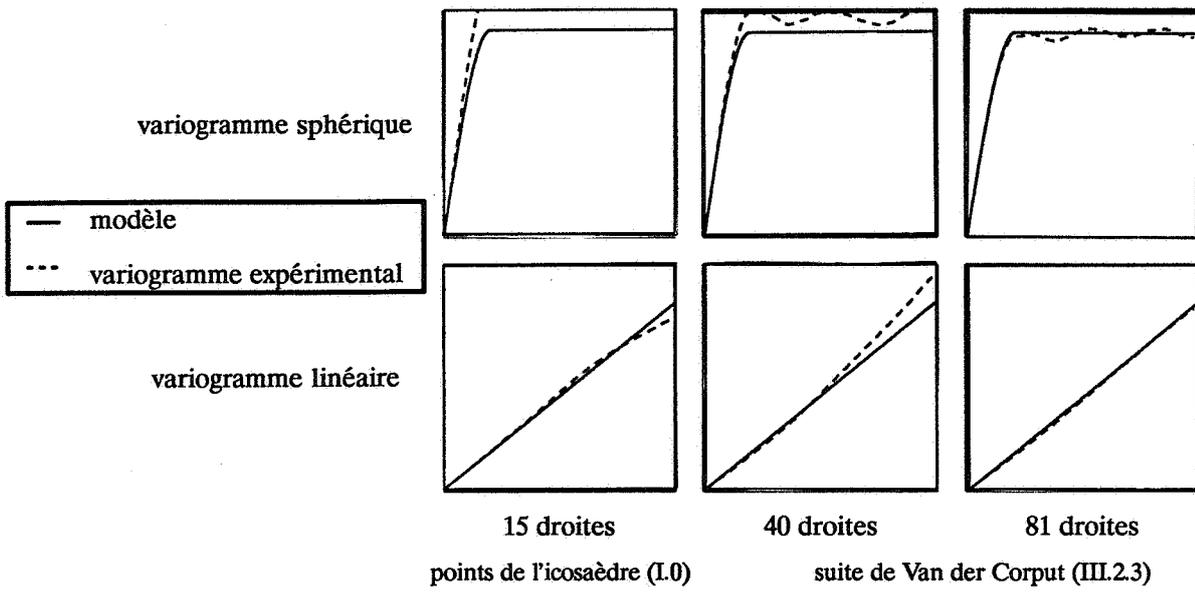
Avec 15 directions, on observe des fluctuations fortes sur les statistiques et variogrammes expérimentaux (figure 7) et les linéaments induits par les bandes sont très visibles (figures 8 et 9).

Dès 40 droites, les fluctuations statistiques, compatibles avec le modèle, sont plus importantes que le biais introduit par la discrétisation (figure 7). Ainsi, avec 40 droites, les linéaments s'estompent dans le cas d'un simulation 2D. Ils restent très visibles avec les directions déduites de l'icosaèdre et perceptibles avec les directions de la suite III.2.3. Avec 81 droites, des linéaments ne sont plus distinguables pour une simulation 2D ou une simulation 3D avec la suite III.2.3. Par contre les droites restent visibles sur la simulation 3D avec les points I.2 (figures 8 et 9).

Avec moins de 100 droites, la méthode de simulation par bandes tournantes permet une simulation satisfaisante des schéma considérés. Ceci rejoint les recommandations de Tompson et al. (1989) qui discrétisaient la sphère de R^3 à l'aide de 100 points aléatoires. Cependant, la suite II.2.3 (coordonnées

sphériques et suite de Van der Corput) donne à 3D les meilleurs résultats : les projections des $2^{n+1} - 1$ premières droites sur le plan horizontal sont réparties régulièrement avec les angles $(\pi k/2^n)_{1 \leq k \leq 2^{n+1}-1}$. Dans ce cas, l'épandage se fait comme un passage de \mathbf{R} à \mathbf{R}^2 et la portée sur les droites projetée est inversement proportionnelle au cosinus de l'azimuth.

figure 7 : quelques variogrammes expérimentaux calculés dans la direction [100]
simulation par passage de \mathbf{R}^1 à \mathbf{R}^3



4 - Conclusion

Il existe des suites de points permettant une discrétisation fine de la demi-sphère de \mathbf{R}^3 . Ceci permet une modélisation par "bandes tournantes" de toute covariance isotrope de \mathbf{R}^3 . Une simulation dans le plan peut être obtenue comme coupe d'une simulation dans l'espace. Cette démarche n'était pas conseillée dans la mesure où les droites utilisées pour simuler à 3D étaient les 15 droites de l'icosaèdre (Chilès, 1977). Elle peut s'avérer particulièrement intéressante dans le cas d'une covariance à portée finie car le passage de C_3 à C_1 conserve cette portée, ce qui n'est pas le cas du passage de C_2 à C_1 .

figure 8 : simulations par la méthode des bandes tournantes

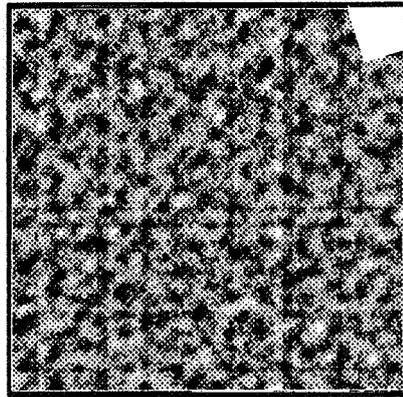
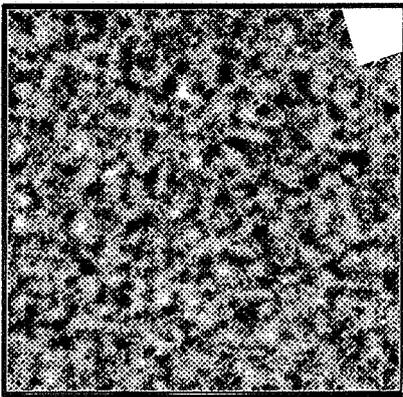
$$\gamma(h) = sph(|h|/10)$$
$$n_x = n_y = 256$$

R^2

R^3

points de l'icosaèdre (I.0)

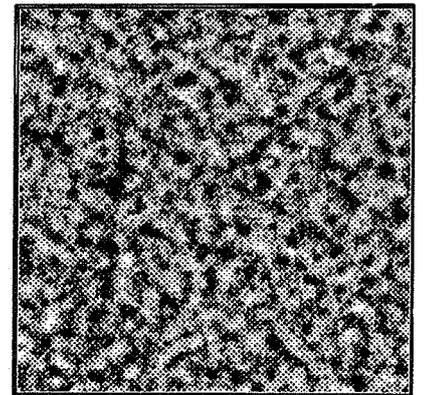
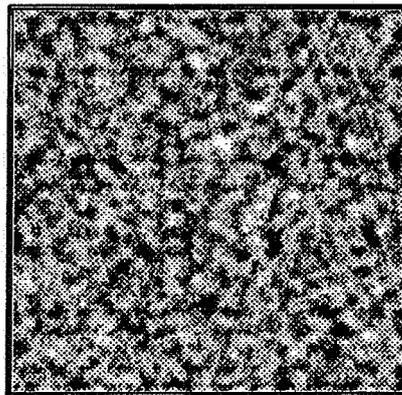
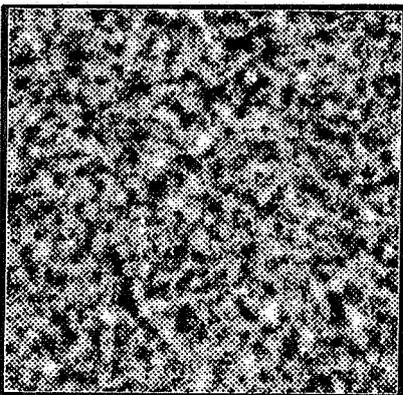
15 droites



triangulation (I.2)

Suite de Van der Corput (III.2.3)

40 droites



81 droites

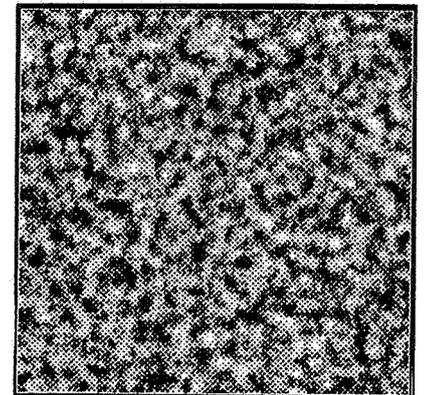
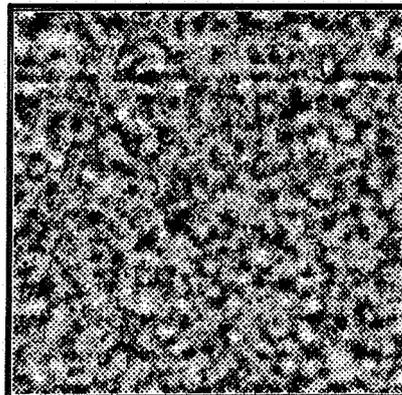
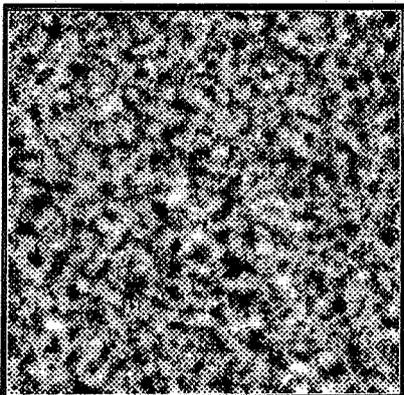


figure 9 : simulations par la méthode des bandes tournantes

$$\gamma(h) = |h|$$

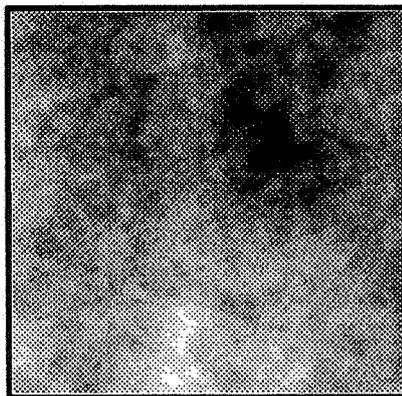
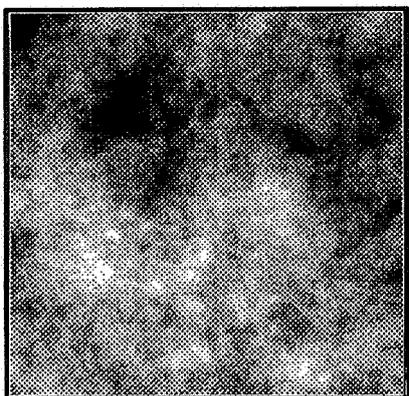
$$n_x = n_y = 256$$

R^2

R^3

points de l'icosaèdre (I.0)

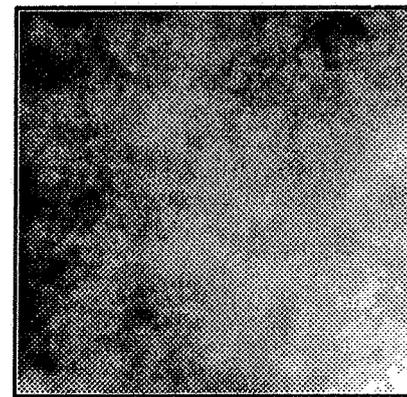
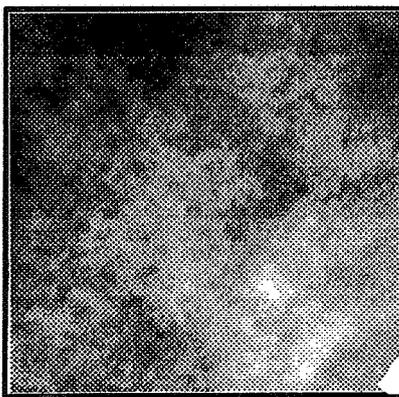
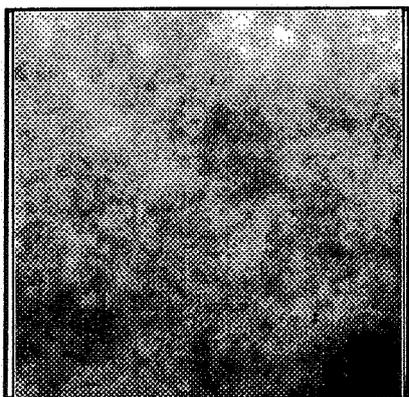
15 droites



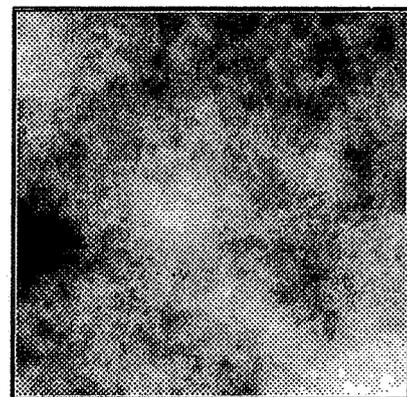
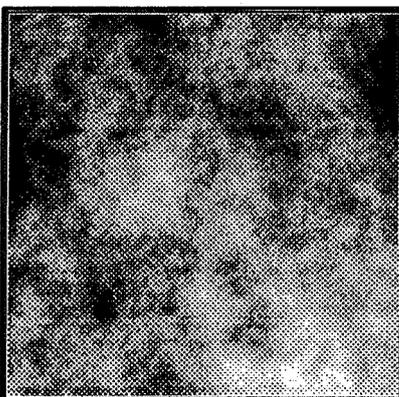
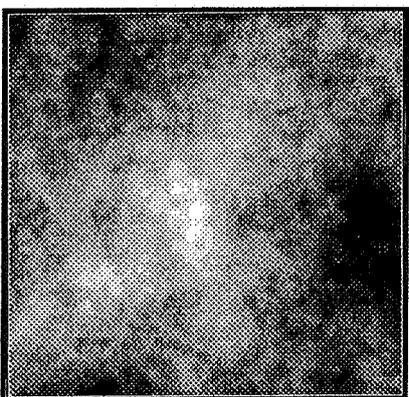
triangulation (I.2)

Suite de Van der Corput (III.2.3)

40 droites



81 droites



5 – Bibliographie

- Barnes R. J. and Johnson T. B. (1984), *Positive kriging*, Proceeding of NATO–ASI Geostatistics for natural resources characterization, South Lake Tahoe, California, 1983, D. Reidel Publishing Company, Dordrecht, Holland, 1984, p.231–244.
- Boulanger, F. (1990), *Modélisation et simulation de variables régionalisées par des fonctions aléatoires stables*, Thèse de Docteur en Géostatistique, Ecole des Mines de Paris, Centre de Géostatistique, Fontainebleau, 1990.
- Bouleau, N. (1986), *Probabilités de l'ingénieur, variables aléatoires et simulation*, Hermann, 1986.
- Chilès, J. P. (1977), *Géostatistique des phénomènes non stationnaires*, Thèse de Docteur–ingénieur, Nancy, 1986.
- Guibal, D. (1972), *Simulations de schémas intrinsèques*, N-291, Ecole des Mines de Paris, Centre de Géostatistique, Fontainebleau, 1972.
- Lubotzky A. , Phillips R. , Sarnak P. (1986), *Hecke operators and distributing points on sphere I*, CPAM, 1986, Vol. 39, p. 149–186.
- Matheron, G. (1972), *Quelques aspects de la montée*, N-271, Ecole des Mines de Paris, Centre de Géostatistique, Fontainebleau, 1972.
- Matheron, G. (1973), *The intrinsic random functions and their applications*, Adv. Appl. Prob., 1973, Vol. 5, p. 435–468.
- Tompson A. F. B. , Abadou R. and Gelhar L. W. (1989), *Implementation of a three–dimensional Turning Bands random field generator*, Water Ressources. Res., 1989, Vol. 25, no 10, p. 2227–2243.



Annexe B

Modèles auto-regressifs et moyennes mobiles

Dans cette partie, on étudie la simulation d'un processus stationnaire admettant une covariance fixée. La droite réelle est discrétisée à pas régulier : on se limitera donc à présenter la simulation de processus sur \mathbb{Z} . Le domaine de simulation est toujours borné, on peut théoriquement se ramener à la simulation d'un vecteur multigaussien dont la matrice de covariance se déduit de la covariance stationnaire. Une technique de simulation bien adaptée est alors la méthode matricielle. Cependant dans la pratique, la taille des vecteurs à simuler limite souvent l'utilisation d'une telle méthode. Aussi modélise-t-on la covariance à simuler par des processus auto-regressifs ou des moyennes mobiles. Pour une covariance fixée, les paramètres de ces deux modélisations peuvent être calculés simultanément, ce qui permet de choisir pour une même précision la modélisation la plus économique. L'utilisation de cet algorithme proposé par Boulanger[6] est présentée sur quelques exemples.

B.1 Méthode matricielle de simulation

Le processus gaussien stationnaire $(X_{(t)})_{t \in \mathbb{R}}$, centré et de covariance ρ , est échantillonné aux points $(\alpha_i)_{1 \leq i \leq n}$: on considère donc le vecteur multigaussien $X = (X_{(\alpha_i)})_{1 \leq i \leq n}$ — on notera $X_{(\alpha_i)} = X_i$ — dont la matrice des covariances est

$$\Sigma = [\sigma_{ij}]_{1 \leq i, j \leq n} \quad \text{où} \quad \sigma_{ij} = \mathbf{E}\{X_i X_j\} = \rho(\alpha_i - \alpha_j).$$

La méthode matricielle de simulation du vecteur multigaussien, X , s'appuie sur la factorisation de la matrice des covariances : $\Sigma = LL'$. Le vecteur à simuler se réécrit donc sous la forme $X = LU$ où U est un échantillon gaussien, centré, normé et sans corrélation. Le vecteur construit par combinaisons linéaires de variables aléatoires gaussiennes indépendantes est multigaussien et sa covariance est bien :

$$\mathbf{E}\{X X'\} = L \mathbf{E}\{U U'\} L' = \Sigma.$$

Deux factorisations classiques d'une matrice définie positive sont la décomposition de Cholesky et la décomposition par éléments propres (voir page 51). Toutes deux s'interprètent

en termes de variables aléatoires : si Σ est strictement définie positive¹, les $(X_i)_{1 \leq i \leq n}$ engendrent une sous-espace vectoriel de dimension n , dont les $(U_i)_{1 \leq i \leq n}$ constituent une base orthonormée. La matrice L n'est rien d'autre que la matrice de changement de base. Les deux factorisations correspondent donc à deux procédés d'orthogonalisation.

B.1.1 Décomposition par éléments propres

On procède à une analyse en composantes principales[45] des vecteurs $(X_i)_{1 \leq i \leq n}$, c'est-à-dire on recherche une base orthonormée $(U_i)_{1 \leq i \leq n}$ du sous-espace engendré maximisant l'inertie des $(X_i)_{1 \leq i \leq n}$ sur chacun des vecteurs $(U_i)_{1 \leq i \leq n}$: $J(k) = \sum_{i=1}^n (\mathbf{E}\{X_i U_k\})^2$. Posant $U_k = \sum_{i=1}^n a_k^i X_i$, on doit maximiser la forme quadratique

$$J(k) = \sum_{1 \leq j, j' \leq n} a_k^j a_k^{j'} \sum_{1 \leq i \leq n} \sigma_{ij} \sigma_{ij'}$$

sous les contraintes : $k' \leq k$, $\sum_{1 \leq j, j' \leq n} a_k^j \sigma_{jj'} a_k^{j'} = \delta_{kk'}$. Les facteurs sont donc les vecteurs propres de la matrice Σ , normalisés par rapport à Σ ; l'inertie associée à un vecteur propre est le carré de sa valeur propre. Si Σ est diagonalisée sous la forme $\Sigma = P \Lambda P'$ avec $P'P = I$ et Λ est la matrice diagonale des valeurs propres, on a $U_k = 1/\sqrt{\lambda_k} \sum_{i=1}^n p_{ik} X_i$. Sous forme matricielle, le vecteur des X_i se réécrit donc

$$X = P \Lambda^{\frac{1}{2}} U .$$

Décomposition de Cholesky

Les vecteurs $(U_i)_{1 \leq i \leq n}$ sont construits de la manière suivante (*orthogonalisation de Schmidt*) :

$$\begin{aligned} U_1 &= X_1 / \sigma_{11}^{\frac{1}{2}} \\ 1 < i \leq n, \quad U_i &= (X_i - \sum_{j < i} \mathbf{E}\{X_i U_j\} U_j) / \text{Var}(X_i - \sum_{j < i} \mathbf{E}\{X_i U_j\} U_j) \end{aligned}$$

Le terme $X_i - \sum_{j < i} \mathbf{E}\{X_i U_j\} U_j$ est la projection de X_i sur le sous-espace vectoriel engendré par les $(U_j)_{j < i}$. Aussi cette construction peut s'interpréter en termes de krigeage car le sous-espace vectoriel engendré par les $(U_j)_{j < i}$ coïncide avec celui engendré par les $(X_j)_{j < i}$. Le krigeage simple de X_i par les $(X_j)_{j < i}$ et sa variance sont donc :

$$\begin{aligned} X_i^{KS}(X_1, \dots, X_{i-1}) &= X_i - \sum_{j < i} \mathbf{E}\{X_i U_j\} U_j \\ \sigma_{KS}^2(i) &= \text{Var}(X_i - \sum_{j < i} \mathbf{E}\{X_i U_j\} U_j) \end{aligned}$$

¹Si Σ n'est pas strictement définie positive, le sous-espace engendré est de dimension $r < n$ et la matrice $L = [\mathbf{E}\{X_i U_k\}]_{1 \leq i \leq n, 1 \leq k \leq r}$ est rectangle.

U_i est l'erreur de krigeage normalisée. La i -ème composante s'exprime de deux manières :

$$\begin{aligned} X_i &= X_i^{KS}(X_1, \dots, X_{i-1}) + \sigma_{KS}(i) U_i \\ &= \sum_{j \leq i} \mathbf{E}\{X_i U_j\} U_j \end{aligned}$$

On en déduit deux manières, strictement équivalente pour simuler le vecteur X :

méthode séquentielle : Pour chaque composante, on calcule son krigeage par les composantes préalablement simulées et on simule l'erreur de krigeage².

méthode matricielle : On simule tout d'abord l'échantillon gaussien puis calcule le produit matriciel $X = LU$.

Sous forme matricielle, ces deux méthodes s'expriment :

$$\begin{aligned} X &= \Lambda X + S U \\ &= L U \end{aligned}$$

où Λ est la matrice des pondérateurs de krigeage — la ligne i correspond aux poids du krigeage de la composante i à l'aide des seules composantes $(X_j)_{j < i}$ —, S est la matrice diagonale dont les termes non nuls sont les racines carrées des variances des krigeage successifs et L la matrice obtenue par décomposition de Cholesky de la matrice des covariances. De l'équivalence des méthodes, on déduit les relations matricielles :

$$\Lambda = I - S L^{-1} \quad \text{et} \quad L = (I - \Lambda)^{-1} S .$$

B.2 Simulation d'un processus multigaussien sur \mathbb{Z}

B.2.1 Processus auto-régressifs

Le processus $(X_n)_{n \in \mathbb{Z}}$ est un processus auto-régressif d'ordre p s'il peut se mettre sous la forme,

$$n \in \mathbb{Z}, \quad X_n = \sum_{i=0}^p \lambda^i X_{n-i} + \sigma U_n$$

où les $(\lambda^i)_{1 \leq i \leq p}$ et σ sont les paramètres du processus et $(U_n)_{n \in \mathbb{Z}}$ est une suite de variables aléatoires centrées normées et indépendantes. Ce processus est stationnaire si et seulement si les p racines $(\rho_i)_{1 \leq i \leq p}$ de l'équation $\rho^p - \sum_{i=1}^p \lambda^i \rho^{p-i} = 0$ sont de module inférieur à un. La covariance est alors de la forme

$$\mathbf{E}\{X_n X_{n+k}\} = \rho(k) = \sum_{i=0}^p \alpha_i \rho_i^{|k|}$$

²Si le vecteur aléatoire à simuler est constitué des valeurs sur une grille régulière d'un processus stationnaire et que pour chaque point simulé les seuls pondérateurs de krigeage non nuls sont ceux associés aux p points précédents, le vecteur simulé est un échantillon de taille n d'un processus auto-régressif d'ordre p .

et elle vérifie les équations $k \in \mathbb{Z}$, $\rho(k) = \sum_{i=0}^p \lambda^i \rho(k-i) + \sigma^2 \delta_0(k)$. Inversement, connaissant la covariance ρ et son ordre p , on peut déterminer les paramètres du processus par les équations de Yule et Walker :

$$1 \leq i \leq p, \quad \sum_{j=1}^p \rho(i-j) \lambda^j = \rho(i)$$

$$\sigma^2 = \rho(0) - \sum_{j=1}^p \lambda^j \rho(j)$$

On reconnaît là les équations du krigeage simple de X_n par les p points précédents.

Une covariance quelconque, ρ , n'est pas toujours la covariance d'un processus auto-régressif. Cependant, si cette covariance tend asymptotiquement vers zéro, une modélisation par un processus auto-régressif d'ordre p permet de simuler une approximation, $\rho^{ar(p)}$, de celle-ci. En effet résolvant les équations de Yule et Walker, on détermine un processus auto-régressif d'ordre p : $n \in \mathbb{Z}$, $X_n^{ar(p)} = \sum_{i=0}^p \lambda^i X_{n-i}^{ar(p)} + \sigma U_n$ dont la covariance vérifie: $0 \leq k \leq p$, $\rho^{ar(p)}(k) = \rho(k)$. On a de plus [6] :

$$\lim_{p \rightarrow +\infty} \|\rho - \rho^{ar(p)}\|_{\infty} = 0.$$

Si les variables aléatoires indépendantes $(U_n)_{n \in \mathbb{Z}}$ sont des gaussiennes, le processus est multigaussien.

B.2.2 Les moyennes mobiles

Un processus sur \mathbb{Z} modélisé par moyennes mobiles s'écrit sous la forme

$$X_n = \sum_{i=0}^p f_i U_{n-i}$$

où $(f_i)_{0 \leq i \leq p}$ sont les paramètres du processus et $(U_n)_{n \in \mathbb{Z}}$ est une suite de variables aléatoires centrées normées et indépendantes. Sa covariance est :

$$\mathbf{E}\{X_n X_{n+k}\} = \rho(k) = \sum_{i=0}^{p-k} f_i f_{i+k}.$$

Soit ρ une covariance sur \mathbb{Z} que l'on voudrait simuler par moyennes mobiles. Il est possible, si $\sum_{k \leq 0} \rho(k) < \infty$, de construire une suite de covariances $\rho^{ma(p)}(k) = \sum_{i=0}^{p-k} f_i f_{i+k}$ telle que

$$\lim_{p \rightarrow +\infty} \|\rho - \rho^{ma(p)}\|_{\infty} = 0.$$

La fonction sur \mathbb{N} , f , est obtenue par décomposition de Cholesky de la matrice de covariance $\Sigma_n = [\rho(i-j)]_{1 \leq i, j \leq n}$. En effet, Σ_n étant une matrice définie positive, elle peut se factoriser sous la forme $\Sigma_n = L_n L_n'$ ($L_n = [l_{i,j}]_{1 \leq i, j \leq n}$). On définit les fonctions sur \mathbb{N} , $(f^{(n)})_{n > 0}$, par

$$i \in \mathbb{N}, \quad f^{(n)}(i) = \begin{cases} l_{n, n-k} & \text{si } 0 \leq i \leq n \\ 0 & \text{sinon} \end{cases}$$

La fonction f est la limite ponctuelle des $(f^{(n)})_{n > 0}$, c'est à dire, $\lim_{n \rightarrow +\infty} f^{(n)}(i) = f(i)$.

B.3 Détermination des paramètres des modèles

Si la covariance à modéliser est sommable, il est possible de la simuler par les deux méthodes présentées aux paragraphes précédents. Or pour ces deux méthodes, une simulation nécessite approximativement — généralement l'ordre du processus, p , étant très inférieur à la taille du vecteur à simuler, n , on néglige la phase d'initialisation — la simulation de n gaussiennes indépendantes puis le calcul de n combinaisons linéaires de p termes. On choisit donc la modélisation qui pour un ordre minimal assure que la différence entre la covariance à modéliser et celle effectivement modélisée n'excède pas une valeur fixée. Or les paramètres des deux modélisations se calculent par un même algorithme itératif. Il est donc possible de déterminer, sans multiplier les calculs, la modélisation optimale.

Le calcul des paramètres de la moyenne mobile nécessite le calcul de la décomposition de Cholesky de matrices des covariances de plus en plus grande. On note

$$n \in \mathbb{N}^*, \quad \Sigma_n = [\sigma_{i,j}]_{1 \leq i,j \leq n} = \begin{bmatrix} \Sigma_{n-1} & c_n \\ c_n' & \sigma_{n,n} \end{bmatrix}$$

La décomposition de Cholesky de la matrice Σ_n , L_n , s'exprime en fonction de celle de la matrice Σ_{n-1} , L_{n-1} :

$$L_n = \begin{bmatrix} L_{n-1} & 0 \\ u_n' & v_n \end{bmatrix}$$

Aussi, pour calculer L_n connaissant L_{n-1} , seul le vecteur u_n et le scalaire v_n sont à calculer :

$$\begin{aligned} u_n &= L_{n-1}^{-1} c_n \\ v_n &= (\sigma_{n,n} - u_n' u_n)^{\frac{1}{2}} \end{aligned}$$

Or, on a de même une expression de L_n^{-1} en fonction de L_{n-1}^{-1} :

$$L_n^{-1} = \begin{bmatrix} L_{n-1}^{-1} & 0 \\ -\frac{1}{v_n} u_n' L_{n-1}^{-1} & \frac{1}{v_n} \end{bmatrix}$$

Le calcul de L_n^{-1} nécessite donc le calcul de $u_n' L_{n-1}^{-1}$. Les composantes de ce vecteur sont les coefficients du krigeage d'un point par les $n-1$ précédents ; ils constituent avec la variance de krigeage — égal au terme diagonal v_n — les paramètres du processus auto-régressif d'ordre $n-1$.

Les paramètres des deux modélisations sont déterminés en calculant successivement la décomposition de Cholesky des matrices Σ_n : à l'itération n , les paramètres de la moyenne mobile d'ordre $n-1$ sont $(f^{(n-1)}(k))_{0 \leq k \leq n-1}$ et ceux du processus auto-régressif d'ordre $n-1$, $(\lambda_{n-1}^i)_{1 \leq i \leq n-1}$ et $\sigma_{KS}(n-1)$. On évalue les erreurs de modélisation

$$\begin{aligned} \epsilon_{ma(n-1)} &= \max_k |\rho(k) - \rho^{ma(n-1)}(k)| \\ \epsilon_{ar(n-1)} &= \max_k |\rho(k) - \rho^{ar(n-1)}(k)| \end{aligned}$$

Ces itérations sont poursuivies jusqu'à ce qu'une des erreurs de modélisation soit inférieure à une valeur fixée. La modélisation correspondante est alors retenue. Le calcul des covariances effectivement modélisées à l'itération n s'effectue à l'aide des équations suivantes :

$$\begin{aligned}\rho^{ma(n-1)}(k) &= \sum_{i=0}^{n-1-k} f_i^{(n-1)} f_{i+k}^{(n-1)} \\ \rho^{ar(n-1)}(k) &= \sum_{i=1}^{n-1} \rho^{ar(n-1)}(k-i) \lambda_{n-1}^i + \sigma_{KS}^2(n-1) \delta_0(k)\end{aligned}$$

L'algorithme présenté ci-dessus est utilisé sur deux exemples pour modéliser les covariances sur \mathbb{R}^1 utilisées pour simuler par la méthode des bandes tournantes les covariances sur \mathbb{R}^3 isotropes exponentielle et sphérique :

- Covariance exponentielle, $C_1(h) = (1 - |h|)e^{-|h|}$,
- Covariance sphérique, $C_1(h) = (1 - 3|h| + 2|h|^3) \mathbb{1}_{|h| < 1}$

Sur la figure B.1, on a reporté le nombre d'itérations nécessaires en fonction du pas de discrétisation des droites pour trois erreurs de modélisation $\epsilon = 0.1, 0.01$ et 0.001 . Ces exemples montrent qu'une modélisation par moyenne mobile est préférable, excepté le cas d'une modélisation grossière ($\epsilon = 0.1$) de la covariance *exponentielle* sur \mathbb{R} . Les variogrammes modélisés sont reportés à la figure B.2: la modélisation par processus auto-régressif induit des effets de trous qui ne sont atténués que par le recours à de grands voisinages de krigeage.

Figure B.1 : Modélisation par processus auto-régressif ou moyenne mobile en fonction du pas de discrétisation

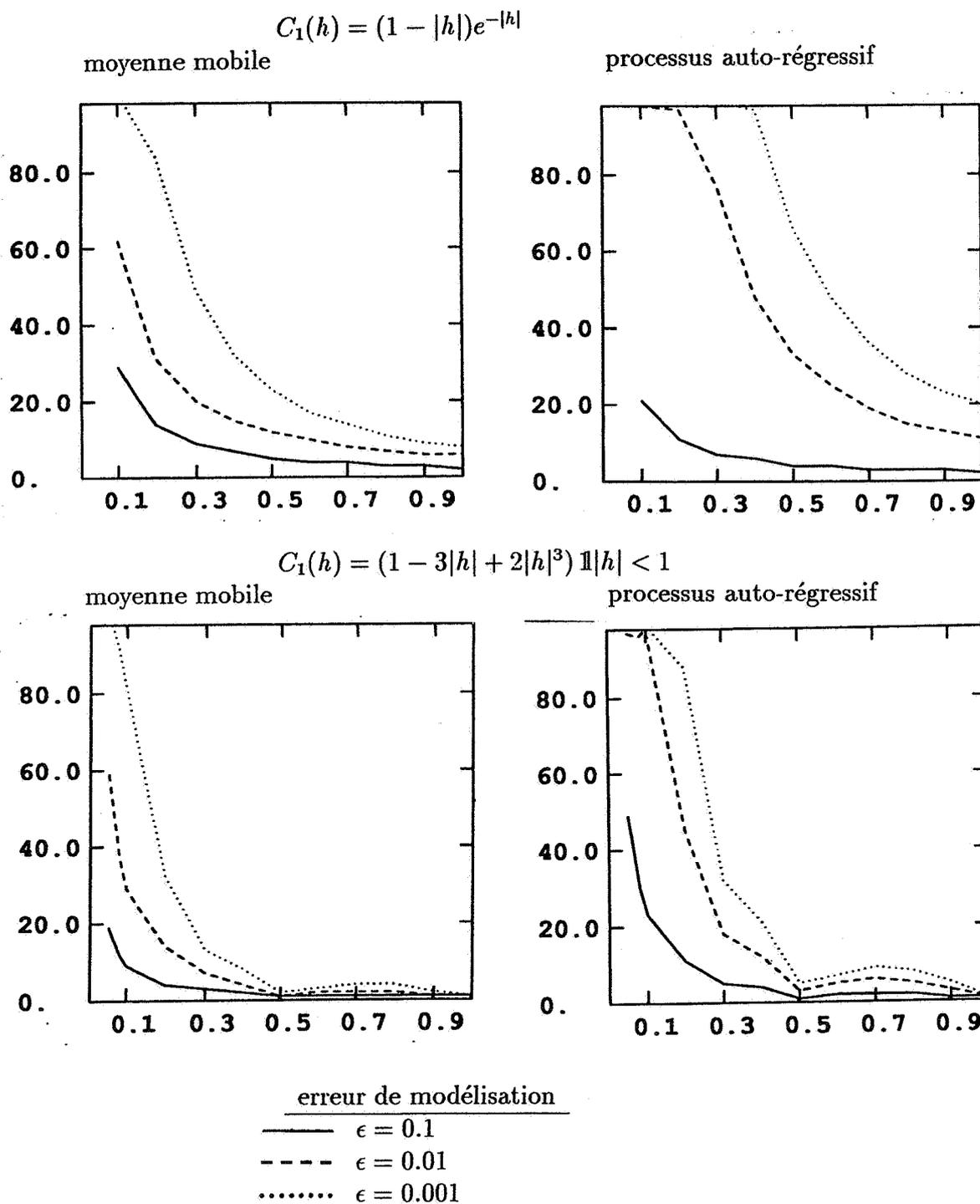
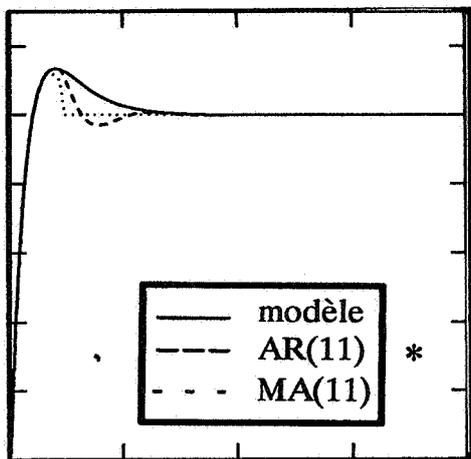
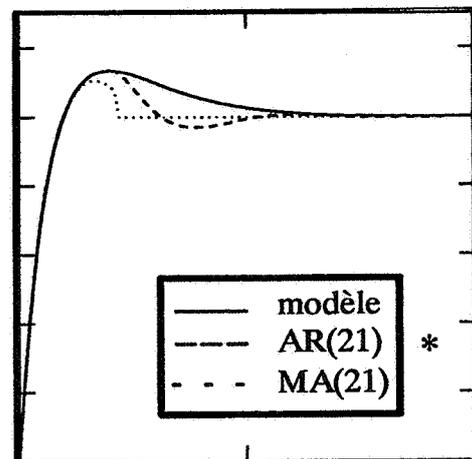


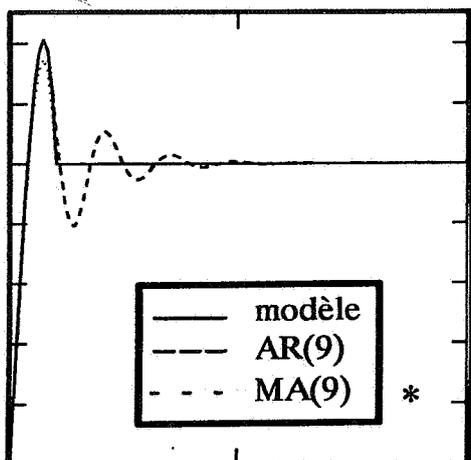
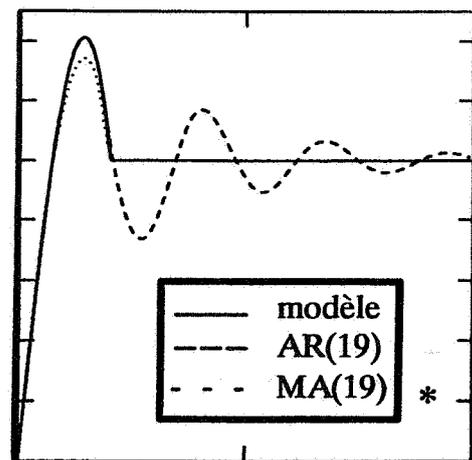
Figure B.2: Variogrammes modélisés par processus auto-régressif ou moyenne mobile

1 - Erreur de modélisation $\epsilon = 0.1$

$$C_1(h) = (1 - |h|)e^{-|h|}$$

 $dt = 0.2$  $dt = 0.1$

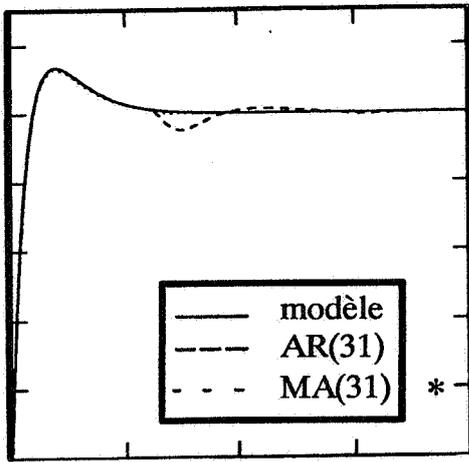
$$C_1(h) = (1 - 3|h| + 2|h|^3) \mathbb{1}_{|h| < 1}$$

 $dt = 0.1$  $dt = 0.05$

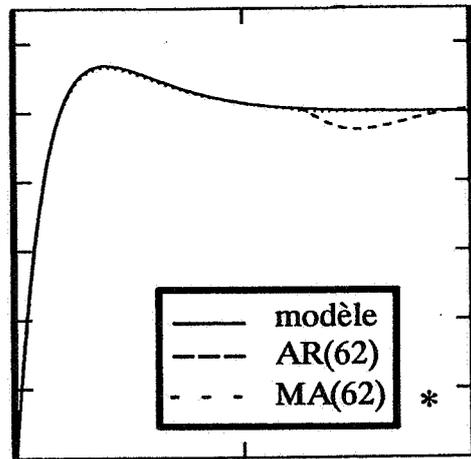
* modélisation retenue

2 - Erreur de modélisation $\epsilon = 0.01$

$$C_1(h) = (1 - |h|)e^{-|h|}$$

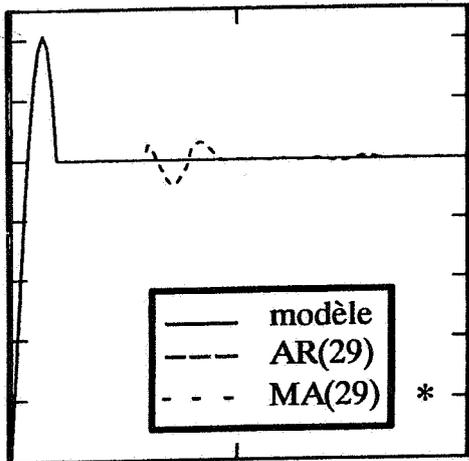


$dt = 0.2$

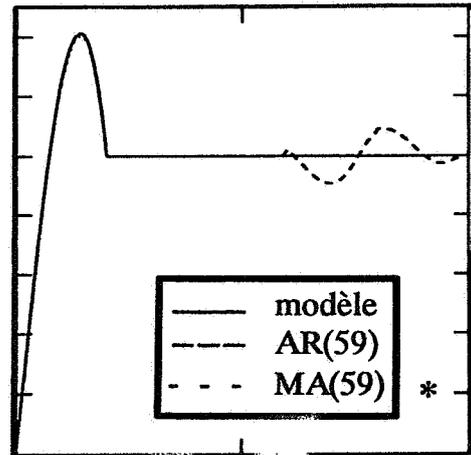


$dt = 0.1$

$$C_1(h) = (1 - 3|h| + 2|h|^3) \mathbb{1}_{|h| < 1}$$



$dt = 0.1$



$dt = 0.05$

* modélisation retenue

Annexe C

Simulation par méthodes spectrales

C.1 Transformée de Fourier discrète

L'ensemble $\mathbb{Z}/n\mathbb{Z} = \{0, \dots, n-1\}$ muni de l'addition, $(x, y) \in (\mathbb{Z}/n\mathbb{Z})^2 \mapsto x + y[n]$ — pour tout élément x de \mathbb{Z} , $x[n]$ désigne la reste de la division euclidienne de x par n — est un groupe fini commutatif. Sur cet ensemble, on peut donc définir le produit de convolution de deux fonctions,

$$k \in \mathbb{Z}/n\mathbb{Z}, \quad f \star g(k) = \sum_{i=0}^{n-1} f(i)g(k-i),$$

ainsi que la transformée de Fourier (TFD) et son inverse (TFD^{-1}),

$$TFD(f)(k) = \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} e^{-2i\pi \frac{jk}{n}} f(j)$$
$$TFD^{-1}(f)(k) = \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} e^{+2i\pi \frac{jk}{n}} f(j)$$

La transformée de Fourier sur $\mathbb{Z}/n\mathbb{Z}$ transforme le produit de convolution en un simple produit de deux fonctions :

$$\widehat{f \star g} = \sqrt{n} \hat{f} \hat{g}$$

Disposant d'un algorithme rapide pour calculer une transformée de Fourier discrète (FFT[9]), la simulation de processus stationnaires par des méthodes spectrales permet de réduire très sensiblement les temps de calcul.

C.2 Simulation d'un processus sur \mathbb{Z}

Dans cette partie, on s'intéresse à la simulation d'un processus stationnaire sur \mathbb{Z} de moyenne nulle et dont la covariance est l'auto-convoluée d'une fonction f de support $\{0, \dots, p\}$:

$$k \in \mathbb{Z}, \quad \rho(k) = \sum_i f(i)f(i+k)$$

Par la suite, on conviendra que $\rho(0) = 1$. Le processus à simuler est la convolution d'un bruit blanc¹ par la fonction f :

$$k \in \mathbb{Z}, \quad X_{(k)} = f * M_{(k)}$$

Cependant, on ne simule jamais un processus sur \mathbb{Z} tout entier, mais sur un nombre fini, l , de points consécutifs — on a $n = l + p$. La simulation de $(X_{(i)})_{1 \leq i \leq l}$ se réduit donc à la simulation du vecteur $(U_i)_{1 \leq i \leq n}$ puis on calcule la régularisée :

$$1 \leq k \leq l, \quad X_{(k)} = \sum_{i=0}^p f_i U_{k+p-i}.$$

Pour utiliser les méthodes spectrales, il est nécessaire d'abandonner la covariance initiale ρ pour conserver la covariance périodique σ construite de la manière suivante : on considère la restriction à $\{0, \dots, n-1\}$ de la fonction de pondération f comme une fonction de $\mathbb{Z}/n\mathbb{Z}$; la covariance périodique est $\sigma = f \star \check{f}$ (\star est ici le produit de convolution de $\mathbb{Z}/n\mathbb{Z}$). Si $p \leq n/2$, la valeur de $\sigma(k)$ ne diffère pas de la valeur de $\rho(k)$ pour $k \in \{0, \dots, n-p-1\}$ (figure C.1). Pour simuler un échantillon du processus de covariance ρ sur \mathbb{Z} , on simule sur $\mathbb{Z}/n\mathbb{Z}$ un processus de covariance σ , dont seul $n-p$ consécutifs sont retenus. En effet, les distances dans \mathbb{Z} entre les points retenus n'excèdent pas $n-p-1$ aussi les valeurs modélisées de covariance sont bien celles du modèle dans \mathbb{Z} .

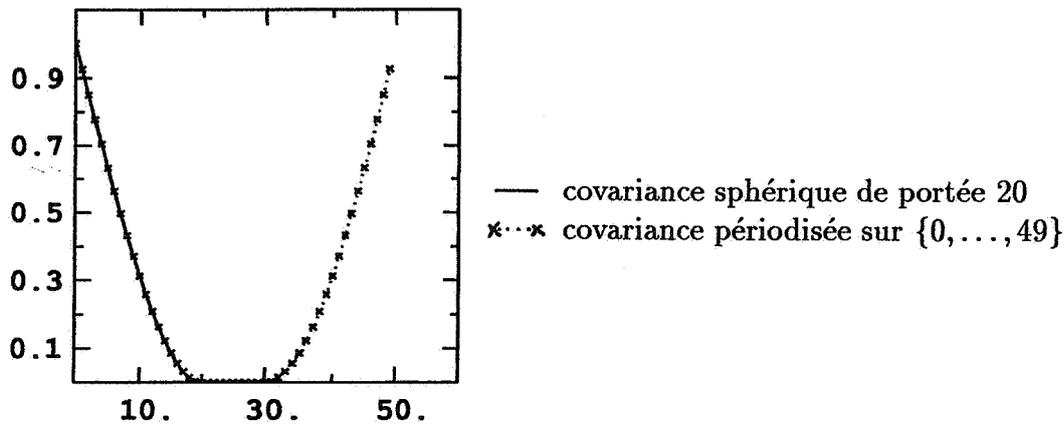


Figure C.1: Covariance sphérique sur \mathbb{R} et covariance périodisée sur $\mathbb{Z}/50\mathbb{Z}$

Sur $\mathbb{Z}/n\mathbb{Z}$, une fonction aléatoire stationnaire d'ordre 2, X , est définie par sa moyenne et sa covariance :

$$\begin{aligned} \mathbf{E}\{X_i\} &= m \\ \mathbf{E}\{(X_i - m)(X_j - m)\} &= \sigma(i - j) \end{aligned}$$

Comme précédemment, on supposera que $m = 0$ et $\sigma(0) = 1$. Mais $\mathbb{Z}/n\mathbb{Z}$ étant un groupe commutatif fini, les covariances stationnaires sont caractérisées par le théorème de

¹Sur \mathbb{Z} , un bruit blanc n'est autre chose qu'une suite infinie, $M = (U_i)_{i \in \mathbb{Z}}$, de variables aléatoires indépendantes et identiquement distribuées telles que $\mathbf{E}\{U_i\} = 0$ et $\mathbf{Var}(U_i) = 1$.

Bochner[22]: σ est une covariance stationnaire de $\mathbb{Z}/n\mathbb{Z}$ si et seulement si il existe une fonction positive $\hat{\sigma}$ telle que $\sigma = TFD(\hat{\sigma})$. Dans ce cas, la simulation par moyenne mobile et la simulation par méthode spectrale sont équivalentes. En simulant des processus sur $\mathbb{Z}/n\mathbb{Z}$, la transformée de Fourier discrète de la covariance utilisée est exacte alors que la discrétisation de la transformée de Fourier de la covariance définie sur \mathbb{R} [5, 31] entraîne une troncature des hautes fréquences; le comportement de la covariance au voisinage de l'origine peut alors être mal reproduit.

Moyenne Mobile

Si on connaît la fonction f — par exemple, celle fournie par la décomposition de Cholesky de la matrice des covariances sur \mathbb{Z} (page 140) —, le processus sur $\mathbb{Z}/n\mathbb{Z}$ s'écrit $X = f \star M$: on simule tout d'abord le vecteur M , puis on calcule les transformées de Fourier inverses de f et M , $\hat{f} = TFD^{-1}(f)$ et $\hat{M} = TFD^{-1}(M)$, enfin on calcule la transformée directe, $X = TFD(n^{\frac{1}{2}} \hat{f} \hat{M})$. Cette méthode "fréquentielle" ne diffère d'une moyenne mobile discrète que par la façon de calculer le produit de convolution.

Méthode spectrale

Une méthode spectrale de simulation consisterait à simuler directement un vecteur complexe \hat{X} sur $\mathbb{Z}/n\mathbb{Z}$ tel que:

$$0 \leq i, j \leq n, \quad \mathbf{E}\{\hat{X}_i \overline{\hat{X}_j}\} = \delta_{ij} \sqrt{n} \hat{\sigma}_i$$

puis de calculer la transformée de Fourier: $X = TFD(\hat{X})$. Cependant procéder de la sorte assure que la covariance de X est σ mais non que X est réel. Cette dernière condition complique la manière de simuler \hat{X} . Aussi, on recherche simplement un vecteur \hat{X} tel que $X = \Re(TFD(\hat{X}))$ admette σ comme covariance. Un vecteur particulièrement simple est un vecteur de la forme

$$\hat{X}_k = (\sqrt{n} \hat{\sigma}_k)^{\frac{1}{2}} (M_1(k) + i M_2(k))$$

où les vecteurs réels, M_1 et M_2 , vérifient

$$\begin{cases} \mathbf{E}\{M_i(k)\} & = 0 \\ \mathbf{E}\{M_i(k) M_j'(k')\} & = \delta_{ij} \delta_{kk'} \end{cases}$$

Il est très facile de construire de tels vecteurs à l'aide de variables aléatoires uniformes:

- $M_1(k) + i M_2(k) = \sqrt{2} e^{2i\pi U_k}$ où les $(U_k)_k$ sont des variables aléatoires indépendantes uniformément réparties sur $]0, 1[$.
- $M_1(k) + i M_2(k)$ où les $(M_1(k))_k$ et les $(M_2(k))_k$ sont des variables aléatoires indépendantes uniformément réparties sur $] -\sqrt{3}, \sqrt{3}[$

Si on pose $X = \Re(TFD(\hat{X}))$ et $Y = \Im(TFD(\hat{X}))$, on vérifie que

$$1 \leq i, j \leq n \quad \begin{cases} \text{Cov}(X_i, X_j) & = \sigma(i-j) \\ \text{Cov}(Y_i, Y_j) & = \sigma(i-j) \\ \text{Cov}(X_i, Y_j) & = 0 \end{cases}$$

Cette méthode présente un double avantage :

- Toute variable simulée est une combinaison linéaire de n variables indépendantes — n étant grand — aussi la loi marginale des processus est asymptotiquement gaussienne. Expérimentalement, on vérifie qu'il en est de même pour la loi bivariable.
- Deux processus orthogonaux sont simulés à l'aide d'une seule transformée de Fourier.

Il faut cependant calculer $(\hat{\sigma}(k))_{0 \leq k \leq n-1}$:

- Si on connaît la fonction f , on calcule alors $\hat{f} = TFD^{-1}(f)$ et $\hat{\sigma} = \sqrt{n} \hat{f} \overline{\hat{f}}$.
- Si on connaît la fonction de covariance périodisée σ , il suffit de calculer la transformée de Fourier $\hat{\sigma}$. Ce dernier calcul est possible si le modèle dans \mathbb{Z} , ρ , admet une portée finie p et que $p < n/2$: la covariance périodisée se calcule alors directement par la formule $\sigma(k) = \rho(k) + \rho(n - k)$.

C.2.1 Comparaison des méthodes de simulation

La méthode des moyennes mobiles et les deux méthodes de simulation présentées ci-dessus — moyennes mobiles par méthode spectrale et méthode spectrale — sont utilisées pour simuler k échantillons indépendants et de taille l d'un processus stationnaire sur \mathbb{Z} dont la covariance a une portée p . Ceci correspond au problème de la simulation des droites pour une simulation par bandes tournantes dans \mathbb{R}^3 d'une fonction aléatoire dont la covariance est isotrope et de portée finie². La covariance isotrope dans \mathbb{R}^3 utilisée ici est la covariance sphérique, la covariance à simuler sur les droites est donc $C_1(h) = (1 - 3|h| + 2|h|^3) \mathbb{1}_{|h| < 1}$; la discrétisation des droites est prise égale à $1/(p+1)$, aussi faut-il simuler sur \mathbb{Z} la covariance $\rho(k) = (1 - 3|h|/p + 2(|h|/p)^3) \mathbb{1}_{|k| \leq p}$.

Au tableau C.1, sont reportés les temps de calcul pour simuler 200 droites avec différentes valeurs de l et p ; on a distingué le temps de modélisation du temps de simulation :

- Le temps de modélisation pour les moyennes mobiles — le calcul de la fonction $(f(i))_{0 \leq i \leq p}$ par décomposition unilatérale de la covariance — est indépendant du domaine de simulation. Par contre, le temps de modélisation pour la méthode spectrale — il faut calculer la transformée de Fourier discrète de la covariance périodisée sur le champ — dépend de la taille du champ mais reste très inférieur au temps de modélisation par moyennes mobiles.

- La simulation par moyenne mobile nécessite pour chaque droite le tirage de $l + p$ gaussiennes indépendantes puis le calcul de l combinaisons linéaires de p termes. Le temps de simulation est approximativement $t_{MM} \approx k[(l + p)t_I + lpt_{II}]$. Expérimentalement, on trouve $t_I = 1.6 \cdot 10^{-5}$ et $t_{II} = 3.1 \cdot 10^{-7}$; t_I correspond au temps nécessaire pour simuler une gaussienne et t_{II} à celui nécessaire au calcul d'un terme de la combinaison linéaire.

Le calcul du produit de convolution par méthode spectrale nécessite le produit de $l + p$ nombres complexes et le calcul de deux transformées de Fourier discrètes. Le calcul de la

²La relation liant la covariance dans \mathbb{R}^1 et celle dans \mathbb{R}^3 , $C_1(\rho) = dC_3(\rho)/d\rho$, implique que C_1 admet la même portée que C_3 .

transformée de Fourier discrète de la fonction de pondération se fait dans la phase d'initialisation. Le temps de simulation est donc : $t_{MM*} \approx k[(l+p)t_I + (l+p)(t'_I + 2t_{FFT} \log_2(l+p))]$ où $t_I = 1.6 \cdot 10^{-5}$, $t'_I = 2.9 \cdot 10^{-5}$ et $t_{FFT} = 2.2 \cdot 10^{-7}$; t_I correspond au temps de simulation d'une variable gaussienne et t'_I au temps pour multiplier deux complexes et t_{FFT} est un coefficient associé à l'algorithme de FFT. Le calcul du produit de convolution par méthode fréquentielle ne se justifie que si $lp t_I > (l+p)(t'_I + 2t_{FFT} \log_2(l+p))$, c'est-à-dire si l'ordre de la moyenne mobile est très grande.

Pour la méthode spectrale, les droites sont simulées par paire. Il faut tout d'abord simuler $2(l+p)$ variables uniformes puis calculer une transformée de Fourier discrète.

On a donc $t_{MS} \approx k/2[2(l+p)t_I^* + t_{FFT}(l+p) \log_2(l+p)]$ où $t_I^* = 1.3 \cdot 10^{-5}$ et $t_{II}^* = 2.2 \cdot 10^{-7}$; t_I^* est le temps de simulation d'une variable uniforme.

Dans les exemples présentés, on a toujours $t_{MS} < t_{MM}$; l'économie de temps calcul est obtenue en substituant la simulation de gaussiennes indépendantes par la simulation de variables uniformes. La méthode des moyennes mobiles avec le calcul du produit de convolution par des méthodes spectrales nécessite toujours plus de temps calcul que les deux méthodes précédentes.

Le calcul de la transformée de Fourier discrète d'un vecteur de dimension l par FFT nécessite un temps proportionnel à $l \log_2(l)$, si l est une puissance de 2. Dans la méthode des bandes tournantes, cette condition n'est pas restrictive car on peut choisir la discrétisation sur les droites de sorte que cette condition soit satisfaite: si t_{max} désigne la longueur des droites à simuler, la relation liant le pas de discrétisation des droites et le nombre total de points simulés, $n = 2^\nu$, est $dt = (1 + t_{max})2^{-\nu}$; l'unité de longueur est la portée de la covariance et t_{max} est le diamètre du champ à simuler.

Pour un cas particulier ($k = 200$, $l = 482$ et $p = 30$), les statistiques calculées sur les droites simulées par méthode spectrale sont reportées à la figure C.2: pour le variogramme et l'histogramme, on ne distingue pas la courbe expérimentale de la courbe théorique; le nuage des corrélations différées est bien d'aspect bigaussien. Ceci justifie donc de simuler la mesure spectrale en utilisant des variables aléatoires uniformes.

C.2.2 Simulation 2D par méthode spectrale

Les résultats sur $\mathbb{Z}/n\mathbb{Z}$ utilisés dans la section précédente se généralisent à 2 dimensions: l'ensemble sur lequel une fonction aléatoire stationnaire est simulée est $\mathbb{Z}/n_1\mathbb{Z} \times \mathbb{Z}/n_2\mathbb{Z}$. La transformée de Fourier discrète ainsi que son inverse sont définies par:

$$TFD(f)(k_1, k_2) = \frac{1}{\sqrt{n_1 n_2}} \sum_{j_1=0}^{n_1-1} \sum_{j_2=0}^{n_2-1} e^{-2i\pi \frac{j_1 k_1}{n_1} + \frac{j_2 k_2}{n_2}} f(j_1, j_2)$$

$$TFD^{-1}(f)(k_1, k_2) = \frac{1}{\sqrt{n_1 n_2}} \sum_{j_1=0}^{n_1-1} \sum_{j_2=0}^{n_2-1} e^{+2i\pi \frac{j_1 k_1}{n_1} + \frac{j_2 k_2}{n_2}} f(j_1, j_2)$$

Tableau C.1 : Temps de simulation de 200 processus indépendants de covariance $\mathbb{1}_{|k|<p} (1 - 2(k/p) + 2(k/p)^3)$

1 - la portée de la covariance est 20

l	Moyenne Mobile		Moyenne Mobile *		Méthode Spectrale	
	modélisation	simulation	modélisation	simulation	modélisation	simulation
108	0.25	0.52	0.26	1.10	<0.01	0.25
236	0.25	1.06	0.26	2.24	0.02	0.69
492	0.25	2.15	0.26	4.85	0.03	1.50
1004	0.25	4.42	0.26	10.18	0.04	3.10

2 - $l + p = 1024$

p	l	Moyenne Mobile		Moyenne Mobile *		Méthode Spectrale	
		modélisation	simulation	modélisation	simulation	modélisation	simulation
10	1014	0.25	3.77	0.26	10.20	0.04	3.10
20	1004	0.25	4.42	0.26	10.20	0.04	3.10
30	994	0.25	5.05	0.26	10.20	0.04	3.10
40	984	0.25	5.68	0.26	10.20	0.04	3.10
50	974	0.25	6.28	0.26	10.20	0.04	3.10

* Le calcul du produit de convolution s'effectue par méthode spectrale.

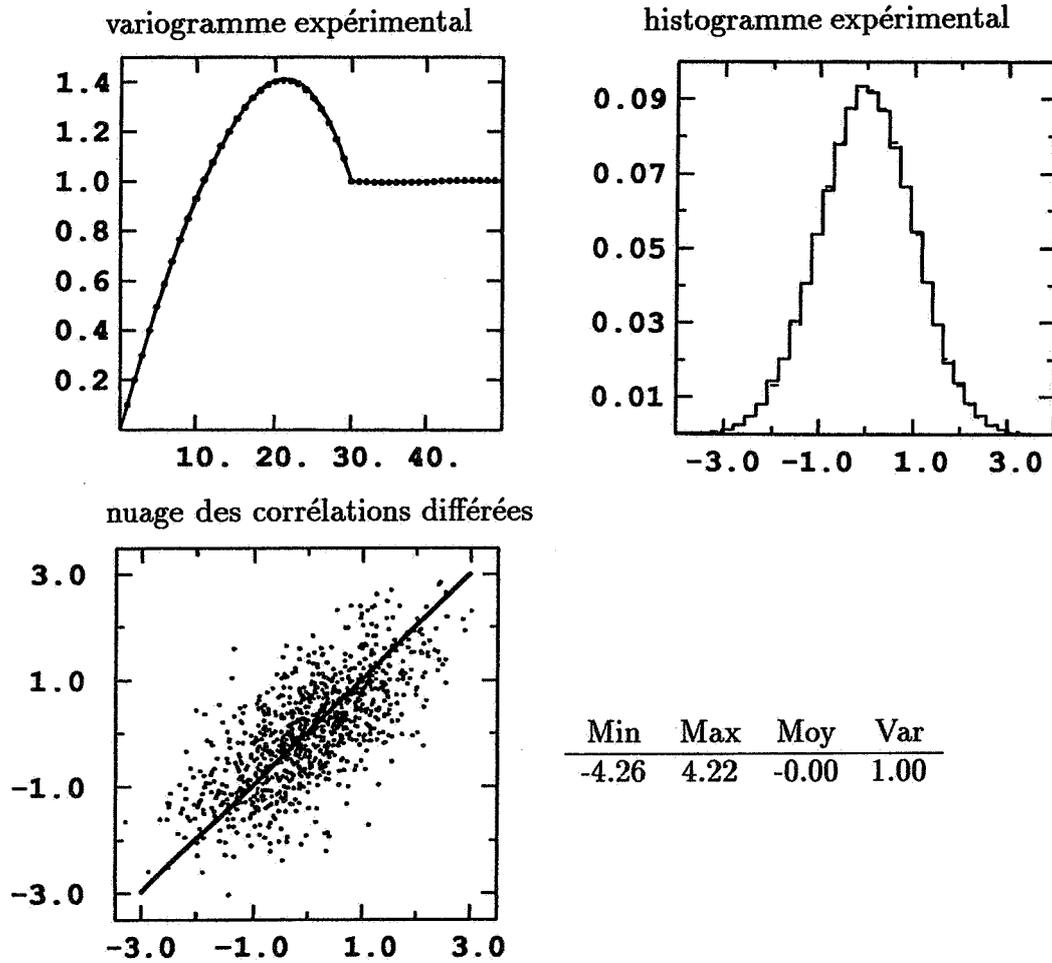


Figure C.2: Simulation par méthode spectrale de $(1 - 3(k/30) + 2(k/30)^3) \mathbb{1}_{|k| < 30}$
 (Statistiques calculées sur 200 échantillons indépendants de 482 points)

Là encore, la transformée de Fourier transforme le produit de convolution en simple produit,

$$\widehat{f \star g} = \sqrt{n_1 n_2} \hat{f} \hat{g}$$

Pour simuler la covariance ρ de portée finie, on la périodise par la formule

$$\sigma(k_1, k_2) = \rho(k_1, k_2) + \rho(n_1 - k_1, k_2) + \rho(k_1, n_2 - k_2) + \rho(n_1 - k_1, n_2 - k_2),$$

puis, après avoir calculé sa transformée de Fourier inverse $\hat{\sigma} = TFD^{-1}(\sigma)$, on simule la mesure orthogonale

$$\hat{X}(k_1, k_2) = (\sqrt{n_1 n_2} \hat{\sigma}(k_1, k_2))^{\frac{1}{2}} (M_1(k_1, k_2) + i M_2(k_1, k_2)).$$

Enfin, on calcule la transformée de Fourier directe. La variable simulée est donc $X = \Re(TFD(\hat{X}))$ ou $Y = \Im(TFD(\hat{X}))$. Ces deux champs sont orthogonaux (voire indépendants car ils sont asymptotiquement gaussiens).

La méthode spectrale à deux dimensions est utilisée pour simuler un schéma sphérique isotrope: à la figure C.3 sont reportés une image 200×200 et les variogrammes expérimentaux, la portée de la covariance est égale à 20.

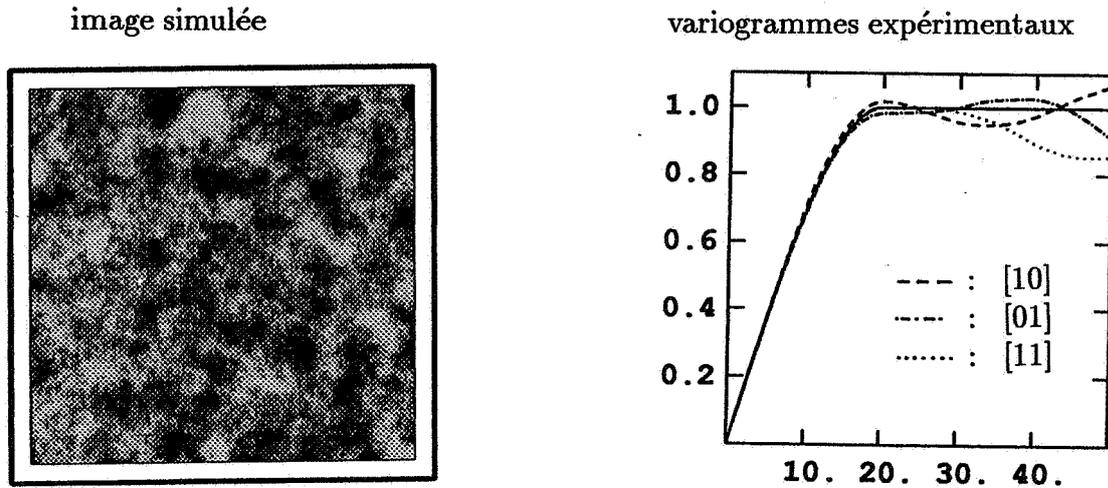


Figure C.3: Simulation par méthode spectrale d'une image 200×200 de covariance $sph(20)$

Un tel schéma peut être simulé également par d'autres méthodes: moyennes mobiles et bandes tournantes. L'image simulée est une grille carrée $n \times n$; la portée de la covariance est p . Pour ces trois méthodes, les temps de simulation sont décomposés de la manière suivante:

- La simulation par moyennes mobiles nécessite la simulation de $(n+p)^2$ gaussiennes puis pour chaque pixel de l'image d'un calcul d'une combinaison linéaire de p^2 termes dans le cas où le domaine de récursion est le quadrant supérieur gauche. Le temps de simulation est $t_{MM} \approx (n+p)^2 t_I + n^2 p^2 t_{II}$ où $t_I = 2.8 \cdot 10^{-5}$ et $t_{II} = 5.7 \cdot 10^{-7}$. Le temps de calcul nécessaire pour identifier les coefficients de la moyenne mobile n'a pas été évalué ici.

• Dans la méthode des bandes tournantes, les droites sont simulées par la méthode spectrale; on a donc considéré le passage de \mathbb{R}^1 à \mathbb{R}^3 car la covariance à simuler sur les droites est elle aussi de portée finie. Le temps de simulation des droites ne dépend que du nombre total des points simulés N (page 151). Pour chacun des n^2 points de l'image, on calcule une combinaison linéaire de k termes où k est le nombre de droites. On a donc: $t_{BT} \approx k/2[2Nt_I^* + t_{II}^*N \log_2(N)] + n^2 k t_{III}$. t_I^* et t_{II}^* sont les constantes de la méthode spectrale à une dimension et $t_{III} = 3.3 \cdot 10^{-6}$ s. Pour la méthode des bandes tournantes, 100 droites³ discrétisées sur 512 points ont été utilisées.

• Par la méthode spectrale, on simule une image $N \times N$, où $N = \min\{2^\nu; \nu \in \mathbb{N} \text{ et } n + p \leq 2^\nu\}$, dont on ne retient qu'une image $n \times n$. La modélisation nécessite le calcul d'une transformée de Fourier discrète et de N^2 racines carrées. Quant à la simulation, elle nécessite la simulation de $2N^2$ variables uniformes indépendantes puis le calcul d'une transformée de Fourier discrète. Le temps de simulation est donc $t_{MS} \approx 2N^2 t_I^* + t_{II}^* N^2 \log_2(N^2)$ où $t_I^* = 2.3 \cdot 10^{-3}$ et $t_{II}^* = 2.2 \cdot 10^{-7}$.

Les temps de calcul effectifs sont reportés au tableau C.2 et les courbes des temps de simulation en fonction de la taille du champ à la figure C.4.

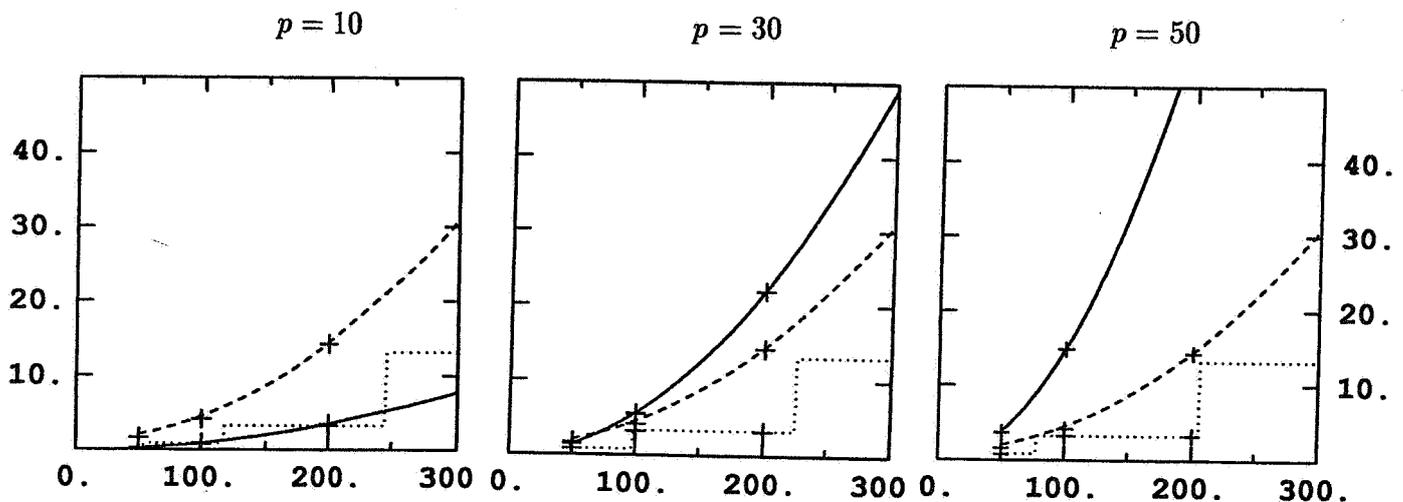


Figure C.4 : Temps de simulation d'une image $n \times n$ en fonction de n

moyennes mobiles : —
 bandes tournantes : ---
 méthode spectrale :
 points expérimentaux : +

³Dans l'annexe A, on a vu qu'une centaine de droites était nécessaire pour que les linéaments induits par les droites disparaissent

Tableau C.2: Temps de simulation d'un schéma sphérique de portée p sur une grille carrée1 - grille 50×50

p	Moyenne Mobile	Bandes Tournantes	Méthode Spectrale	
	simulation	simulation	modélisation	simulation
10	0.24	1.64	0.16	0.19
20	0.70	1.64	0.64	0.78
30	1.43	1.64	0.64	0.78
40	2.41	1.64	0.64	0.78
50	3.66	1.64	0.64	0.78

2 - grille 100×100

p	Moyenne Mobile	Bandes Tournantes	Méthode Spectrale	
	simulation	simulation	modélisation	simulation
10	0.90	4.14	0.64	0.78
20	2.65	4.14	0.64	0.78
30	5.52	4.14	2.67	3.21
40	9.51	4.14	2.67	3.21
50	14.82	4.14	2.67	3.21

3 - grille 200×200

p	Moyenne Mobile	Bandes Tournantes	Méthode Spectrale	
	simulation	simulation	modélisation	simulation
10	3.51	14.22	2.67	3.21
20	10.53	14.22	2.67	3.21
30	21.92	14.22	2.67	3.21
40	37.80	14.22	2.67	3.21
50	58.54	14.22	2.67	3.21

Les moyennes mobiles ne sont intéressantes que pour de très petites portées, c'est alors la méthode la plus performante. Pour des portées plus importantes, la méthode des bandes tournantes est légèrement moins rapide que la méthode spectrale.

Cette dernière méthode devient très performante lorsqu'il faut calculer plusieurs simulations : la densité spectrale n'est calculée qu'une fois et les simulations sont produites par paires, la partie réelle et la partie imaginaire de l'image complexe simulée par transformée de Fourier discrète.

Cependant si le domaine à simuler n'est pas une grille régulière, la seule méthode applicable est la méthode des bandes tournantes ; les droites peuvent toujours être simulées par méthode spectrale.

Bibliographie

- [1] M. Armstrong, editor. *Geostatistics*, Dordrecht Holland, 1989. Kluwer Academic Press.
- [2] J. Besag. Spatial interaction and statistical analysis of lattice systems. *J. R. Statist. Soc. B*, 36:192–236, 1974.
- [3] J. Besag. On statistical analysis of dirty pictures. *J. R. Statist. Soc. B*, 48:259–302, 1986.
- [4] J. Besag, J. York, and A. Mollié. Bayesian image restoration, with two applications in spatial statistics. *Ann. Inst. Statist. Math.*, 43(1):1–59, 1991.
- [5] L. Borgman, M. Tahari, and R. Hagan. Three dimensional, frequency domain simulations of geological variables. In Verly et al [17], pages 517–541.
- [6] F. Boulanger. *Modélisation et simulation de variables régionalisées par des fonctions aléatoires stables*. thèse de docteur en géostatistique, ENSMP, Fontainebleau, France, 1990.
- [7] J. P. Chilès. *Géostatistique des phénomènes non stationnaires*. thèse de docteur-ingénieur, Université de Nancy-I, Nancy, France, 1977.
- [8] P. G. Ciarlet. *Introduction à l'analyse matricielle et à l'optimisation*. Masson, Paris, 1982.
- [9] J. W. Cooley and J. W. Tukey. An algorithm for the machine computation of complex fourier series. *Mathematics of Computation*, 19:297–301, 1965.
- [10] C. Daly. *Applications de la géostatistique à quelques problèmes de filtrage*. thèse de docteur en géostatistique, ENSMP, Fontainebleau, France, 1991.
- [11] M. W. Davis. Generation of large stochastic simulations—the matrix polynomial approximation method. *Math. Geol.*, 19:99–107, 1987.
- [12] M. W. Davis. Production of conditional simulation via LU decomposition of the covariance matrix. *Math. Geol.*, 19:91–98, 1987.
- [13] C. de Fouquet. *L'estimation des réserves récupérées sur modèle géostatistique de gisements non homogènes*. thèse de docteur-ingénieur en Sciences et Techniques Minières, Ecole des Mines, Paris, France, 1985.

- [14] C. de Fouquet et al. Conditional simulation of random sets — application to argillaceous sandstone reservoir. In Armstrong [1], pages 517–530.
- [15] L. Devroye. *Non Uniform Random Variable Generation*. Springer Verlag, Berlin, 1986.
- [16] J. L. Doob. *Stochastic processes*. John Wiley & Sons, Inc., New York, 1953.
- [17] G. Verly et al, editor. *Geostatistics for Natural Resources Characterisation*, Dordrecht Holland, 1984. NATO A.S.I., D. Reidel.
- [18] W. Feller. *An introduction to probability theory and its applications*, volume II. John Wiley & Sons, Inc., New York, 1966.
- [19] X. Freulon and Ch. de Fouquet. Remarques sur la pratique des bandes tournantes à 3d. *Cahiers de Géostatistique*, 1:101–117, 1991. Compte rendu des Journées de Géostatistique, 6–7 juin 1991, Fontainebleau.
- [20] S. Geman and D. Geman. Stochastic relaxation, Gibbs distribution and the bayesian restoration of images. *I.E.E.E. transactions : Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.
- [21] M. Guarascio, editor. *Advanced Geostatistics in the Mining industry*, Dordrecht Holland, 1976. NATO A.S.I., D. Reidel.
- [22] A. Guichardet. *Analyse harmonique commutative*. Dunod, Paris, 1968.
- [23] J. M. Hammersley and D. C. Handscomb. *Monte Carlo Methods*. Methuen, London, 1964.
- [24] E. Isaaks. Indicator simulation: Application to the simulation of a high grade uranium mineralization. In Verly et al [17], pages 1057–1069.
- [25] A. G. Journal. *Simulations conditionnelles de gisements miniers*. thèse de docteur-ingénieur, Université de Nancy-I, Nancy, France, 1977.
- [26] V. Langlais. *Estimation sous contraintes d'inégalités*. thèse de docteur en géostatistique, ENSMP, Fontainebleau, France, 1990.
- [27] Ch. Lantuéjoul. Simulation de lois monovariabiles. Cours du Centre de Géostatistique, C-125, Ecole des Mines de Paris, Fontainebleau, 1987.
- [28] Ch. Lantuéjoul. Fonctions aléatoires de substitution. à paraître (Troia 92), 1992.
- [29] Ch. Lantuéjoul and J. Rivoirard. Une méthode de détermination d'anamorphose. Note du Centre de Géostatistique, N-916, Ecole des Mines de Paris, Fontainebleau, 1984.
- [30] D. H. Lehmer. Mathematical methods in large-scale computing units. *Ann. Comp. Lab. Harvard Univ.*, 26:141–146, 1951.

- [31] A. Mantoglou. Digital simulation of multivariate two and three dimensional stochastic processes with a spectral turning bands method. *Math. Geol.*, 19:129–149, 1987.
- [32] G. Matheron. *Les variables régionalisées et leur estimation*. Masson, Paris, 1965.
- [33] G. Matheron. *La théorie des variables régionalisées et ses applications*, volume 5 des *Cahiers du Centre de Morphologie Mathématique*. E.N.S.M.P., Fontainebleau, 1970.
- [34] G. Matheron. Le krigeage disjonctif. Note du Centre de Géostatistique, N-360, Ecole des Mines de Paris, Fontainebleau, 1973.
- [35] G. Matheron. Forecasting block grade distributions : the transfer functions. In Guarascio [21], pages 237–257.
- [36] G. Matheron. A simple substitute for conditional expectation: the disjunctive kriging. In Guarascio [21], pages 221–236.
- [37] G. Matheron. *Estimer et choisir*, volume 7 des *Cahiers du Centre de Morphologie Mathématique*. E.N.S.M.P., Fontainebleau, 1978.
- [38] G. Matheron. Remarques sur le krigeage et son dual. Note du Centre de Géostatistique, N-695, Ecole des Mines de Paris, Fontainebleau, 1981.
- [39] G. Matheron. The internal consistency of models in geostatistics. In Armstrong [1], pages 21–38.
- [40] G. Matheron et al. Simulation conditionnelle à trois faciès dans une falaise de la formation du brent. *Sciences de la Terre*, 28:213–249, 1988. Etudes Géostatistiques V - Séminaire C.F.S.G. sur la géostatistique 15-16 juin 1987, Fontainebleau.
- [41] N. Metropolis, A. W. Rosenbluth, A. H. Teller, and E. Teller. Equations of state calculations by fast computing machines. *J. Chem. Phys.*, 21:1087–1091, 1953.
- [42] J. Neveu. *Bases mathématiques du calcul des probabilités*. Masson, Paris, 1964.
- [43] J. Rivoirard. Modèles à résidus d'indicatrices autokrigeables. *Sciences de la Terre*, 28:303–326, 1988. Etudes Géostatistiques V - Séminaire C.F.S.G. sur la géostatistique 15-16 juin 1987, Fontainebleau.
- [44] J. Rivoirard. Models with orthogonal indicator residuals. In Armstrong [1], pages 91–107.
- [45] G. Saporta. *Probabilités, analyse des données et statistiques*. Technip, Paris, 1990.
- [46] J. von Neumann. Various techniques used in connection with random digits. *National Bureau of Standards Applied Mathematics Series*, 12:36–38, 1951.
- [47] H. Wackernagel. *L'inférence d'un modèle linéaire en géostatistique multivariable*. thèse de docteur de 3^e cycle en Sciences et Techniques Minières, Ecole des Mines, Paris, France, 1985.



Table des matières

1	Des exemples de conditionnement	5
1.1	Le modèle gaussien anamorphosé	6
1.1.1	Fonction aléatoire multigaussienne	6
1.1.2	Anamorphose	6
1.2	Conditionnement	7
1.3	Simulations conditionnelles <i>classiques</i>	9
1.3.1	Rappels : le krigeage conditionnant	9
1.3.2	Accélération des calculs par krigeage dual	10
1.4	La loi conditionnelle	11
1.4.1	Formules de probabilités totales et composées	11
1.4.2	Relation fonctionnelle	13
1.4.3	Randomisation	16
1.5	Des observations	17
1.5.1	Corégionalisation linéaire de facteurs gaussiens	17
1.5.2	Changement de support	18
1.5.3	Erreurs de mesures	21
1.5.4	Contraintes d'inégalités	22
1.6	Conclusion	23
2	Le gisement d'uranium de Laporte	25
2.1	Le gisement de Laporte	26
2.2	Covariances et géométrie	28
2.2.1	Les covariances d'indicatrice, des outils structuraux	28
2.2.2	Etude de Laporte	34
2.3	Modélisation de Laporte	36
2.3.1	Modèle de corégionalisation teneur-géologie	36
2.3.2	Analyse structurale	37
2.3.3	Simulations	41
2.4	Conclusion	42
3	Simulation de gaussiennes tronquées	45
3.1	Simulations exactes	46
3.1.1	Simulation par inversion	46

3.12	Variance en fonction du temps calcul — Méthode de Metropolis	74
3.13	Variance en fonction du temps calcul ($\rho = 0.8, x_c = 0.5$)	75
3.14	Variance en fonction du rang de l'itération	78
3.15	Simulations conditionnées par l'indicatrice $\mathbb{1}_{X \geq 0.67}$	79
3.16	Estimations à partir de l'indicatrice $\mathbb{1}_{X \geq 0.5}$	81
4.1	Statistiques calculées sur les itérations successives	90
4.2	Variances calculées sur les itérations successives	91
4.3	Variogrammes spatiaux calculés sur les itérations successives	92
4.4	Variogramme temporel	93
4.5	Anamorphose discrète	98
4.6	Nuages de corrélation entre la gaussienne de la variable sous-jacente et la gaussienne de la variable bruitée	99
4.7	Images bruitées simulées	101
4.8	Estimation de l'histogramme local ($\phi(x) = 100 + 10x$)	102
4.9	Nuages de corrélation entre la gaussienne de la variable sous-jacente et la gaussienne de la variable bruitée	103
4.10	Images bruitées simulées	105
4.11	Estimation de la variable sous-jacente	106
4.12	Estimation de l'histogramme local ($V = 5 \times 5$)	107
4.13	Images mosaïques bruitées	108
4.14	Estimation de la variable sous-jacente sur des images mosaïques	109
4.15	Variogrammes de l'image micro-sonde	111
4.16	Fonction d'anamorphose de la concentration sous-jacente	111
4.17	Image microsonde et trois estimateurs de la concentration sous-jacente	112
4.18	Nuage de corrélation différée des gaussiennes associées à la variable bruitée	113
4.19	Variogramme temporel des simulations conditionnelles	114
B.1	Modélisation par processus auto-regressif ou moyenne mobile en fonction du pas de discrétisation	143
B.2	Variogrammes modélisés par processus auto-regressif ou moyenne mobile	144
C.1	Covariance sphérique sur \mathbb{R} et covariance périodisée sur $\mathbb{Z}/50\mathbb{Z}$	148
C.2	Simulation par méthode spectrale de $(1 - 3(k/30) + 2(k/30)^3) \mathbb{1}_{ k < 30}$	153
C.3	Simulation par méthode spectrale d'une image 200×200 de covariance $sph(20)$	154
C.4	Temps de simulation d'une image $n \times n$ en fonction de n	155