

SIMULATION DE FONCTIONS ALEATOIRES ADMETTANT UN VARIOGRAMME CONCAVE DONNE

G. MATHERON



TABLE DES MATIERES

<i>RESUME</i>	196	<i>ABSTRACT</i>	196
A - INTRODUCTION	197	E - FORME GENERALE D'UN VARIOGRAMME CONCAVE	203
B - MOSAIQUE DE RENOUVELLEMENT.....	197	F - PREMIER PROCEDE	205
C - SIMULATION D'UN VARIOGRAMME CONCAVE LINEAIRE EN $h = 0$	199	G - REALISATION PRATIQUE	208
D - REMARQUE IMPORTANTE POUR LES APPLICATIONS	202	H - EXEMPLE DU VARIOGRAMME $ h ^\alpha$	209
		I - SECOND PROCEDE : SIMULATION DE FAI-1 DE TYPE SPLINE	210
		REFERENCES.....	212

*Centre de Géostatistique, Ecole Nationale Supérieure des Mines de Paris, 35 rue Saint-Honoré
77305 FONTAINEBLEAU CEDEX - FRANCE

Etudes Géostatistiques V - Séminaire C.F.S.G. sur la Géostatistique 15-16 Juin 1987, Fontainebleau. Sci. de la Terre, Sér. Inf., Nancy, 1988, 28, pp. 195 à 212.

RESUME

La technique connue des mosaïques de renouvellement permet de construire, à une dimension, des simulations de F.A. stationnaires admettant une covariance convexe et linéaire en $h = 0$, à cela près quelconque. Je propose ci-dessous un procédé un peu plus général, permettant la simulation facile de FAI-0 à variogramme concave quelconque, linéaire ou non en $h = 0$. En particulier, on peut ainsi simuler des FAI-0 à variogramme en $|h|^\alpha$ ($0 < \alpha \leq 1$), et aussi des FAI-1 de type spline, c'est-à-dire admettant la covariance généralisée $h^2 \log |h|$.

ABSTRACT

*Simulation of random functions having a given
concave variogram*

The well-known renewal mosaics technique allows constructing 1-D simulations of stationary random functions having any covariance convex and linear in $h = 0$.

Hereunder, I am proposing a more general approach, allowing an easy simulation of I.R.F.-0 with any concave variogram, linear or not in $h = 0$. Especially, we can thus simulate I.R.F.-0 with a variogram in $|h|^\alpha$ ($0 < \alpha \leq 1$) and also spline-type IRF-1, that is, having a generalized covariance $h^2 \log |h|$.

KEYWORDS : Simulations, fractal, self-similar, splines, renewal process, infinitely divisible distributions, stable distributions.

A - INTRODUCTION

La littérature est avare de renseignements au sujet des techniques permettant de simuler les FAI-0 dites, improprement, fractales, qui sont en fait autohomothétiques et caractérisées par un variogramme de la forme $\gamma(h) = A|h|^\alpha$ ($0 < \alpha \leq 2$). Il suffit, en fait, de savoir résoudre ce problème dans l'espace à une seule dimension, car la méthode des bandes tournantes conserve la forme en $|h|^\alpha$, à un coefficient près, de sorte que l'on obtiendra sans peine des simulations à 2 et 3 dimensions si on sait en construire à une dimension seulement (G. Matheron, (1973)). Je propose ci-dessous un procédé très simple et très économique pour construire de telles simulations à une dimension, mais seulement pour $0 < \alpha \leq 1$. Plus généralement, comme on le verra, ce procédé permet de simuler n'importe quelle fonction aléatoire intrinsèque d'ordre 0 (FAI-0) à variogramme concave donné (on verra, d'ailleurs, que n'importe quelle fonction concave sur \mathbb{R}^+ , positive et nulle en $h = 0$ est un variogramme sur l'espace à une dimension). Une extension immédiate nous conduira aussi à un procédé permettant la simulation de FAI-1 à covariance généralisée en $h^2 \log |h|$.

Notre point de départ sera la technique connue qui consiste à construire une mosaïque sur un processus de renouvellement stationnaire : on obtient ainsi toutes les fonctions de covariance stationnaires convexes et linéaires en $h = 0$ (paragraphe B).

En modifiant de manière adéquate la loi d'entrée du processus de renouvellement, on s'affranchit de la clause de stationnarité (ou, ce qui revient au même, de la condition $m < \infty$, où m est la longueur moyenne des intervalles séparant les points du processus). On peut ainsi simuler n'importe quel variogramme concave pourvu qu'il soit linéaire en $h = 0$ (paragraphe C).

Cette dernière restriction est levée dans les paragraphes qui suivent, et la théorie des processus à accroissements indépendants et stationnaires nous conduit à deux procédés permettant de simuler n'importe quel variogramme concave, linéaire ou non à l'origine. Les applications aux cas $|h|^\alpha$ et $|h|^2 \log |h|$ s'en déduisent de manière très simple.

B - MOSAÏQUE DE RENOUVELLEMENT

Commençons par le procédé le plus simple, et d'ailleurs classique, qui consiste à construire un modèle mosaïque sur un processus de renouvellement stationnaire. A l'état stationnaire, un processus de renouvellement délimite une partition aléatoire de la droite réelle en segments de droite dont les longueurs X_i sont des variables aléatoires mutuellement indépendantes et admettant la même distribution $F(dx)$. Cette loi $F(dx)$, concentrée sur la demi-droite ouverte $\{x > 0\}$, admet une moyenne m finie, mais, à cela près, peut être absolument quelconque. Le processus mosaïque $Z(x)$ est alors une fonction aléatoire prenant, sur chacun des seg-

ments délimités par le processus de renouvellement, des valeurs constantes ε_i , indépendantes et identiquement distribuées. La loi de ces valeurs ε_i n'interviendra dans ce qui suit que par sa moyenne, que nous supposerons nulle sans nuire à la généralité, et par sa variance σ^2 que nous supposerons finie. En pratique, on pourra prendre des ε_i gaussiennes, ou uniformes, ou encore à deux valeurs seulement 0 et 1 : dans ce dernier cas, notre mosaïque de renouvellement peut être considérée aussi comme un ensemble aléatoire.

Calculons la covariance (centrée) $C(h)$ de ce processus mosaïque $Z(x)$. On posera :

$$(1) \quad T(x) = P(X > x) = \int_{x+}^{\infty} F(dz) \quad ; \quad B(x) = \int_x^{\infty} T(\xi) d\xi$$

Considérons deux points x_0 et $x_0 + h$ de la droite, avec, par exemple, $h > 0$. On sait qu'en l'état stationnaire, le premier point du processus immédiatement à droite de x_0 tombe en $x_0 + Y$, avec une variable $Y \geq 0$ admettant la densité $T(y)/m$. Voir W. Feller (1967), Chapitre XI, pour une justification rigoureuse. Ce résultat se retrouve d'ailleurs intuitivement en remarquant deux choses : tout d'abord, un point x_0 donné ayant plus de chances de tomber dans un grand segment que dans un petit, la loi de la longueur du segment auquel il appartient n'est plus $F(dx)$, mais $(x/m) F(dx)$, la fréquence de la longueur x devant être pondérée par cette longueur x elle-même. En second lieu, cette longueur x étant fixée, le point x_0 peut occuper n'importe quelle position entre les deux extrémités, i.e. apparait comme uniformément distribuée sur le segment de longueur x . On aura donc à x fixé pour la variable Y la densité $(1/x) 1_{0 \leq y \leq x}$, et donc finalement, en déconditionnalisant, on trouve pour Y la densité

$$(2) \quad \theta(y) = \int_0^{\infty} 1_{0 \leq y \leq x} \frac{F(dx)}{m} = \frac{T(y)}{m}$$

Par suite, la probabilité pour que x_0 et $x_0 + h$ appartiennent au même segment sera

$$\rho(h) = \int_h^{\infty} \theta(y) dy = B(h)/m$$

D'où la covariance cherchée :

$$(3) \quad C(h) = \sigma^2 B(h)/m$$

La fonction $B(h)$ possède les propriétés suivantes : elle est convexe, positive, décroissante, nulle à l'infini et admet en $h = 0$ une dérivée à droite finie $B'(0) = -T(0) = -1$. Inversement, toute fonction $B(h)$ possédant ces propriétés est biunivoquement associée à une loi $F(dx)$ par les relations (1).

Compte tenu du fait que la variance σ^2 peut être choisie arbitrairement, on voit que pour toute fonction $C(h)$ convexe sur \mathbb{R}^+ , positive, nulle à l'infini (et donc décroissante) et linéaire au voisinage de $h = 0$, le procédé précédent permet de simuler une F.A. stationnaire admettant $C(h)$ comme covariance (le fait que $C(h)$ soit effectivement une covariance découle

immédiatement de ce qu'une telle fonction est limite de combinaisons linéaires positives de fonctions triangles). La loi $F(dx)$ définissant le processus de renouvellement, et sa moyenne m sont, en effet, bien définies par :

$$T(x) = \frac{C'(x)}{C'(0)} ; m = \frac{-C(0)}{C'(0)} = \frac{-\sigma^2}{C'(0)}$$

Deux limitations apparaissent, liées au fait que m doit rester fini pour assurer la stationnarité du processus de renouvellement : $C(0)$ doit être fini, ce qui exclut les variogrammes indéfiniment croissants ; et $C'(0)$ doit exister, ce qui exclut les comportements non linéaires en $h = 0$, par exemple les comportements en $|h|^\alpha$ ($0 < \alpha < 1$). Levons d'abord la première limitation.

C – SIMULATION D'UN VARIOGRAMME CONCAVE LINEAIRE EN $h = 0$

Notre objectif réel est de construire une simulation sur un segment $(0, L)$ de longueur L , peut-être très grande, mais, obligatoirement, finie. Il faut pour cela se donner la loi d'entrée $F_0(dy)$ du premier point du processus de renouvellement à droite de 0, et la loi $F(dy)$ des longueurs des segments suivants. Dans le cas stationnaire, on a vu que cette loi d'entrée F_0 admettait la densité $\theta(y) = T(y)/m$, relation (2) ci-dessus. Toutefois, avec la probabilité

$$\int_L^\infty \theta(y) dy = B(L)/m$$

cette loi d'entrée admet l'éventualité d'un premier point Y tombant au-delà de l'extrémité L de l'intervalle qui nous intéresse. Dans ce cas, $Z(x)$ reste constant de 0 à L , et le variogramme empirique est identiquement nul. L'éventualité $Y > L$ apportant une contribution nulle au variogramme doit pouvoir être écartée sans inconvénient. Cela revient à remplacer la loi d'entrée par la même loi conditionnée par $Y \leq L$.

Choisissons donc comme loi d'entrée la loi de densité :

$$(4) \quad \theta_L(y) = \frac{T(y)}{\int_0^L T(x) dx} \cdot 1_{0 \leq y \leq L}$$

On remarque que cette loi existe même si m est infinie. Les points suivants du processus de renouvellement sont alors $Y + X_1$, $Y + X_1 + X_2$..., les X_i étant indépendantes et admettant toujours la loi $F(dx)$. Choisissons deux points x_0 et $x_0 + h$ ($h > 0$) appartenant au segment $(0, L)$, et calculons la probabilité $\rho(x_0; h)$ pour que ces deux points ne soient pas séparés par les points du processus de renouvellement ainsi défini. Cette non séparation peut se produire de deux manières différentes :

- Ou bien le point d'entrée Y tombe déjà à droite de $x_0 + h$, ce qui a lieu avec la probabilité

$$(5) \quad P(Y > x_0 + h) = \int_{x_0+h}^L \theta_L(y) dy$$

- Ou bien le point d'entrée Y tombe à gauche de x_0 , en un point $y < x_0$. Dans ce cas, il y a à gauche de x_0 un dernier point ξ de renouvellement (éventuellement égal à y lui-même) et le segment suivant a une longueur supérieure à $x_0 + h - \xi$.

La probabilité correspondante est donc :

$$(6) \quad p(x_0; h) = \int_0^{x_0} \theta_L(y) \int_0^{x_0-y} U(dx) T(h + x_0 - y)$$

On reconnaît, sur cette expression, le produit de convolution $\theta_L * U * T_h$, où T_h désigne la fonction translaturée de T : $T_h(x) = T(x+h)$. La mesure $U(dx)$ qui figure dans ce produit de convolution est le potentiel du processus de renouvellement :

$$U(dx) = \delta(dx) + \sum_{n=1}^{\infty} F^{(n)}(dx)$$

(δ est la mesure de Dirac, $F^{(n)}$ la $n^{\text{ième}}$ autoconvoluée de la loi F , i.e. la loi de la somme $X_1 + \dots + X_n$ de n variables indépendantes de même loi F). Si l'on désigne par $\hat{U}(\lambda)$ et $\hat{\phi}(\lambda)$ les transformées de Laplace du potentiel et de la loi F , on a simplement

$$(7) \quad \hat{U}(\lambda) = \frac{1}{1 - \hat{\phi}(\lambda)}$$

Pour un borélien B donné de \mathbb{R}^+ , $U(B)$ représente l'espérance du nombre de points de renouvellement tombant dans B sachant que le point $x = 0$ est déjà lui-même un point de renouvellement (i.e. lorsque la loi d'entrée est $\delta(dx)$), et $U(dx)$ représente donc la probabilité pour qu'il y ait un point de renouvellement dans le segment très petit $(x, x+dx)$. D'où la relation (6).

Maintenant, la fonction $T(x)$ admet elle-même la transformée de Laplace

$$\hat{T}(\lambda) = \frac{1 - \hat{\phi}(\lambda)}{\lambda}$$

D'après (7), on a donc :

$$\hat{U}(\lambda) \hat{T}(\lambda) = 1/\lambda$$

c'est-à-dire $U * T = 1$, ou, explicitement :

$$(7') \quad \int_0^x U(d\xi) T(x-\xi) = 1$$

Pour $y \leq L$, la densité $\theta_L(y)$, définie en (4), ne diffère de $T(y)$ que par un facteur constant. Ainsi, la relation :

$$(8) \quad \int_0^x U(d\xi) \theta_L(x-\xi) = C_L = 1 / \int_0^L T(y) dy$$

est vraie pour $x < L$ (bien que l'on ne puisse pas écrire $U * \theta_L = C_L$, puisque la relation (8) devient fausse pour $x > L$). Or, dans la relation (6), nous avons $x_0 < L$. Pour évaluer la valeur en x_0 du produit de convolution $\theta_L * U * T_h$, nous pouvons donc remplacer la fonction $\theta_L * U$ par la constante C_L . Il vient ainsi :

$$(9) \quad p(x_0;h) = \frac{\int_0^{x_0} T(h+x) dx}{\int_0^L T(y) dy} = \int_h^{x_0+h} \theta_L(y) dy$$

Finalement, la probabilité cherchée $\rho(x_0;h)$ pour que nos deux points x_0 et $x_0 + h$ (compris entre 0 et L) ne soient pas séparés, qui est la somme des deux probabilités évaluées en (5) et en (9), est donc donnée par :

$$(10) \quad \rho(x_0;h) = \int_h^L \theta_L(y) dy$$

Elle ne dépend pas du choix du point x_0 (sous réserve seulement d'avoir $0 < x_0 < x_0 + h < L$), mais seulement de h . Par suite, le processus mosaïque $Z(x)$, construit sur $(0,L)$ par ce procédé, admettra sur l'intervalle $(0,L)$ la covariance stationnaire

$$(11) \quad C(h) = \sigma^2 \rho(x_0;h) = \sigma^2 \frac{\int_0^L T(x) dx}{\int_0^L T(y) dy}$$

Or, dans cette construction, nous n'avons à aucun moment supposé que la loi F admettait une espérance finie m . Même si m est infinie, nous obtenons une F.A. (localement) stationnaire admettant la covariance (11), ou, si l'on préfère, le variogramme

$$\gamma(h) = \frac{\sigma^2}{C_L} \int_0^h T(x) dx$$

Comme σ^2 peut être choisie à volonté, la constante C_L ne joue pas un rôle essentiel, et la procédé nous donne un variogramme proportionnel à l'intégrale de $T(x)$:

$$(12) \quad \gamma(h) = C \int_0^h T(x) dx$$

Plus précisément : le processus mosaïque $Z(x)$ ainsi construit constitue une représentation localement stationnaire (sur l'intervalle $(0,L)$) d'une FAI-0 admettant le variogramme $\gamma(h)$.

Ce variogramme $\gamma(h)$ est nécessairement concave, donc non décroissant et linéaire au voisinage de $h = 0$. Inversement, si $\gamma(h)$ est une fonction concave, positive et nulle en $h = 0$, donc non décroissante, et admettant en $h = 0$ une dérivée à droite $\gamma'(0) < \infty$, alors $\gamma(h)$ est un variogramme et le procédé ci-dessus permet de simuler une réalisation d'une FAI-0 admettant ce variogramme. Il suffit, en effet, de prendre pour la loi d'entrée la fonction de répartition :

$$(13) \quad F_L(y) = \frac{\gamma(y)}{\gamma(L)} \quad (\text{densité : } \theta_L(y) = \gamma'(y) / \gamma(L))$$

et pour la loi F du processus de renouvellement, la fonction de répartition

$$(14) \quad F(y) = 1 - \gamma'(y) / \gamma'(0)$$

D - REMARQUE IMPORTANTE POUR LES APPLICATIONS

Nous avons ainsi obtenu un procédé permettant, théoriquement, de simuler n'importe quel variogramme concave linéaire à l'origine, et nous allons dans un instant lever cette dernière limitation. Toutefois, si l'on simule un seul processus mosaïque sur l'intervalle $(0,L)$, le variogramme empirique sera très sensible au résultat y du tirage au sort de la variable Y d'entrée (on l'obtient en prenant pour y la solution de l'équation $F_L(y) = u$ où u a été tirée au sort selon la loi uniforme sur $(0,1)$). Il faut donc, en pratique, faire la somme d'un certain nombre de processus mosaïques de ce type simulés indépendamment.

Pour contrôler au mieux le variogramme résultant, il est même préférable de ne pas tirer au sort les variables d'entrée, mais d'utiliser les quantiles de la loi d'entrée : si l'on a choisi de superposer N simulations, on choisira, pour chacune d'elles, les variables d'entrée y_1, y_2, \dots, y_N définies par

$$\frac{\gamma(y_k)}{\gamma(L)} = \frac{k - 1/2}{N}$$

Cette remarque s'applique évidemment aussi aux procédés que nous allons maintenant décrire, et qui vont nous permettre de simuler des variogrammes concaves quelconques, même non linéaires à l'origine.

E - FORME GENERALE D'UN VARIOGRAMME CONCAVE

Le procédé ci-dessus ne convient qu'à un variogramme concave linéaire à l'origine, car, d'après (12), on trouve $\gamma'(0) = C T(0)$, et $T(0) = 1$, puisqu'il s'agit d'une fonction de répartition. Pour avoir $\gamma'(0) = \infty$, il faudrait $T(0) = \infty$. En termes intuitifs, cela signifie que notre processus ponctuel (qui ne serait plus réellement un processus de renouvellement) devrait présenter des points d'accumulation, ou, si l'on préfère, il devrait y avoir une probabilité non nulle pour qu'un intervalle donné, de longueur finie, contienne une infinité de points du processus. Or, il existe un procédé classique pour construire des processus ponctuels possédant des points d'accumulation. Ce procédé consiste à considérer les sauts d'un processus à accroissements indépendants et stationnaires. Il existe, en effet, des processus X_t de ce type qui accomplissent, en un temps fini, un nombre infiniment grand de sauts infiniment petits. En projection sur l'axe des x , les valeurs de X_t immédiatement avant et après ces sauts dessinent alors un processus ponctuel à points d'accumulation. On sait, depuis les travaux de P. Lévy entre autres, que les processus ont des lois indéfiniment divisibles. Partant de $X_0 = 0$ en $T = 0$, la valeur X_t du processus en $t > 0$ admet la loi définie par

$$E [e^{-\lambda X_t}] = e^{-t \psi(\lambda)}$$

avec une fonction $\psi(\lambda)$ de la forme :

$$\psi(\lambda) = \int_0^{\infty} \frac{1 - e^{-\lambda y}}{y} H(dy)$$

où $H(dy)$ est une mesure positive sur \mathbb{R}^+ , non nécessairement sommable, mais telle que :

$$\int_0^{\infty} \frac{H(dy)}{a+y} < \infty \quad \text{ou} \quad \int_a^{\infty} \frac{H(dy)}{y} < \infty \quad \text{pour } a > 0$$

Dans un langage intuitif, $H(dy)/y$ représente le nombre moyen par unité de temps de sauts d'amplitude comprise entre y et $y + dy$ (loi en nombre), et $H(dy)$ leur amplitude cumulée (loi en mesure). Les conditions imposées :

$$\int_0^a H(dy) < \infty \quad ; \quad \int_a^{\infty} \frac{H(dy)}{y} < \infty$$

signifient : la première, que l'amplitude cumulée des petits sauts reste finie, même si leur nombre est infini ; la seconde, que le nombre des grands sauts reste fini, même si leur amplitude cumulée prend une espérance infinie. Cette dernière circonstance (de temps à autre, des

sauts de très grande amplitude, conduisant à une espérance infinie) correspond au cas déjà examiné d'un processus de renouvellement avec $m = \infty$. La première (un nombre infiniment grand de sauts infiniment petits) va nous permettre de lever la restriction $\gamma'(0) < \infty$.

Pour $y > 0$ donné, la fonction définie, pour $h \geq 0$, par :

$$y \wedge h = \begin{cases} y & \text{si } y \leq h \\ h & \text{si } y \geq h \end{cases}$$

(et complétée par symétrie pour $h < 0$) est évidemment un variogramme. D'après les conditions imposées à la mesure $H(dy)$ de P. Lévy, l'intégrale :

$$(15) \quad \gamma(h) = \int_0^{\infty} (y \wedge h) \frac{H(dy)}{y}$$

est convergente pour tout $h \geq 0$, et définit donc un variogramme. Explicitement, en effet, on a :

$$\gamma(h) = \int_0^h H(dy) + h \int_h^{\infty} \frac{H(dy)}{y}$$

et chacune des deux intégrales est convergente.

Ce variogramme est concave et croissant (puisqu'il appartient au cône convexe engendré par les fonctions $y \wedge h$), mais, en général, il n'est pas linéaire en $h = 0$, sauf si l'intégrale

$$\int_0^{\infty} \frac{H(dy)}{y} \text{ est finie.}$$

Inversement, il est facile de voir que toute fonction positive, concave sur \mathbb{R}^+ , nulle en $h = 0$, est un variogramme. Elle est, en effet, nécessairement de la forme :

$$\gamma(h) = c h + \int_0^{\infty} (y \wedge h) \frac{H(dy)}{y}$$

avec une mesure $H(dy)$ du type de P. Lévy, et, éventuellement, un terme linéaire $\bar{c} h$ (avec $c \geq 0$). Nous supposerons $c = 0$ dans ce qui suit. Cela ne nuira pas à la généralité, puisqu'il est très facile de simuler des FAI-0 à variogramme linéaire. (On peut aussi, comme on le verra, retrouver le variogramme linéaire comme cas limite). Nous allons montrer que le processus à accroissements indépendants et stationnaires associé à cette mesure $H(dy)$ permet de construire des simulations de FAI-0 admettant ce variogramme. Naturellement, ces simulations ne seront effectivement construites que sur une maille discrète.

F - PREMIER PROCEDE

Considérons donc le processus X_t à accroissements indépendants et stationnaires, défini par la condition $X_0 = 0$ et

$$\left\{ \begin{array}{l} E[e^{-\lambda X_t}] = e^{-t\psi(\lambda)} \\ \psi(\lambda) = \int_0^{\infty} \frac{1 - e^{-\lambda y}}{y} H(dy) \end{array} \right.$$

A ce processus, on associe classiquement son potentiel $U(dx)$, qui est la mesure associant à tout borélien B de \mathbb{R}^+ l'espérance $U(B)$ du temps passé par le processus X_t dans le borélien B . Si $G_t(dx)$ est la loi de X_t , on a :

$$U(B) = \int_0^{\infty} E[1_B(X_t)] dt = \int_0^{\infty} G_t(B) dt$$

c'est-à-dire

$$U(.) = \int_0^{\infty} G_t(.) dt$$

Comme la transformée de Laplace de la loi G_t est $\exp(-t\psi(\lambda))$, celle du potentiel U est donc :

$$(16) \quad \tilde{U}(\lambda) = \int_0^{\infty} e^{-\lambda x} U(dx) = \frac{1}{\psi(\lambda)}$$

Nous considérerons aussi la fonction :

$$(17) \quad \theta(x) = \int_x^{\infty} \frac{H(dy)}{y}$$

qui jouera dans ce qui suit le rôle de la fonction $T(x) = 1 - F(x)$ des paragraphes précédents, à ceci près que nous pouvons maintenant avoir $\theta(0) = \infty$. Calculons sa transformée de Laplace $\vartheta(\lambda)$. Il vient :

$$\vartheta(\lambda) = \int_0^{\infty} e^{-\lambda x} dx \int_0^{\infty} 1_{x \leq y} \frac{H(dy)}{y}$$

et, par application du théorème de Fubini

$$\vartheta(\lambda) = \int_0^{\infty} \frac{H(dy)}{y} \frac{(1 - e^{-\lambda y})}{\lambda}$$

et donc :

$$(18) \quad \tilde{\theta}(\lambda) = \frac{\psi(\lambda)}{\lambda}$$

Comparant (16) et (18), nous trouvons :

$$\tilde{U}(\lambda) \tilde{\theta}(\lambda) = \frac{1}{\lambda}$$

c'est-à-dire

$$(19) \quad \int_0^x U(d\xi) \theta(x-\xi) = 1$$

Cette relation généralise (7'), et jouera un rôle essentiel dans ce qui suit.

Le point suivant consiste à calculer la probabilité $Q(x_0;h)$ pour que les points x_0 et $x_0 + h$, sur l'axe des x , ne soient séparés par aucun des points du processus ponctuel obtenu en projetant sur cet axe des x les valeurs prises par X_t . Il s'agit donc de trouver la probabilité pour que, à un certain instant $t \in (0, \infty)$, le processus X_t ait sauté d'une valeur inférieure à x_0 à une valeur supérieure à $x_0 + h$. Toujours avec la condition initiale $X_0 = 0$, on a donc

$$Q(x_0;h) = \int_0^\infty dt \int_0^{x_0} G_t(dx) \int_{x_0+h-x}^\infty \frac{H(dy)}{y}$$

c'est-à-dire :

$$(20) \quad Q(x_0;h) = \int_0^{x_0} U(dx) \theta(h + x_0 - x)$$

Il s'agit donc de la valeur en x_0 du produit de convolution $U * \theta_h$ où θ_h est défini par $\theta_h(x) = \theta(h+x)$. En particulier, la relation (19) nous garantit que l'on a bien $Q(x_0;0) = 1$ en $h = 0$.

Ce résultat concerne le processus soumis à la condition initiale $x_0 = 0$. Nous allons maintenant prendre pour X_0 la loi d'entrée de densité :

$$\theta_L(y) = \frac{\theta(y)}{\int_0^L \theta(x) dx}$$

Désignons par $Q_L(x_0;h)$ la probabilité pour que les points x_0 et $x_0 + h$ de l'axe des x ne soient séparés par aucun point du processus ponctuel, i.e. pour que le processus saute d'une valeur inférieure à x_0 à une valeur supérieure à $x_0 + h$, lorsque X_0 admet la loi d'entrée θ_L . On a évidemment :

$$Q_L(x_0;h) = \int_{x+h}^L \theta_L(y) dy + \int_0^{x_0} \theta_L(y) dy Q(x_0-y;h)$$

La deuxième intégrale est la valeur en x_0 du produit de convolution

$$\theta_L * Q(\cdot;h) = \theta_L * U * \theta_h$$

Mais on se limite au cas $x_0 \leq L$, et, pour $x \leq L$, la valeur de $\theta_L * U$ est

$$\frac{1}{\int_0^L \theta(y) dy} \int_0^x U(d\xi) \theta(x-\xi) = \frac{1}{\int_0^L \theta(y) dy}$$

à cause de (19). Il vient alors :

$$Q_L(x_0;h) = \frac{1}{\int_0^L \theta(y) dy} \left[\int_{x_0+h}^L \theta(x) dx + \int_0^{x_0} \theta(h+y) dy \right]$$

et finalement :

$$(21) \quad Q_L(x_0;h) = \frac{\int_0^L \theta(x) dx}{\int_0^L \theta(y) dy}$$

Ainsi, avec la loi d'entrée θ_L , cette probabilité $Q(x_0;h)$ ne dépend que de h et pas de x_0 (du moins pour x_0 et $x_0 + h \leq L$), et le processus mosaïque correspondant $Z(x)$ admettra sur $(0,L)$ la covariance stationnaire

$$C(h) = \frac{\int_0^L \theta(x) dx}{\int_0^L \theta(y) dy} \times \sigma^2$$

ou, si l'on préfère, le variogramme :

$$(22) \quad \gamma(h) = \sigma^2 \frac{\int_0^h \theta(x) dx}{\int_0^L \theta(y) dy}$$

Or, compte tenu de (17), il vient :

$$\int_0^h \theta(x) dx = \int_0^\infty (y \wedge x) \frac{H(dy)}{y}$$

Ainsi, à un facteur près égal à :

$$c = \frac{\sigma^2}{\int_0^L \theta(y) dy}$$

ce variogramme coïncide bien avec celui que l'on se proposait de simuler.

G - REALISATION PRATIQUE

En pratique donc, ayant choisi un entier L , on aura pour la loi d'entrée F_L la fonction de répartition

$$F_L(y) = \frac{\gamma(y)}{\gamma(L)}$$

où $\gamma(y)$ est le variogramme à simuler. La variable d'entrée $Y = y$ est donc simulée en prenant $F_L(y) = u$, où u est un tirage de la loi uniforme sur $(0,1)$, ou, mieux, un quantile choisi d'avance (voir paragraphe 3). De $i = 0$ jusqu'à $i_0 = \text{Int}(y)$ ($\text{Int} =$ partie entière), on pose donc $Z(x) = \varepsilon_0$, où ε_0 est tirée au sort selon une loi de variance $\sigma^2 = \gamma(L)$.

Pour informer les points suivants $i_0 + 1/\dots$, nous sommes amenés à considérer le processus partant de y au temps $t = 0$. Ainsi, le premier point du processus ponctuel postérieur à $i_0 + 1$ tombe en un point $i_0 + 1 + Y_1$, la loi de Y_1 étant donnée par :

$$P(Y_1 > h) = Q(i_0 + 1 - y; h)$$

avec la fonction $Q(x_0; h)$ définie en (20).

De $i_0 + 1$ à $i_1 = \text{Int}(i_0 + 1 + y_1)$, on prendra $Z(x) = \varepsilon_1$, où ε_1 est un nouveau tirage de la même variable ε de variance $\sigma^2 = \gamma(L)$. A partir de cette nouvelle origine $x_1 = i_0 + 1 + Y_1$, on recommence la même construction : on tire au sort Y_2 selon la loi

$$P(Y_2 > h) = Q(i_1 + 1 - x_1; h)$$

et on pose $Z(x) = \varepsilon_2$ de $i_1 + 1$ à $i_2 = \text{Int}(i_1 + 1, y_2)$, et ainsi de suite.

De la sorte, un nombre fini d'opérations permet d'informer la maille discrète $i = 0, 1, \dots, L$, et de construire la restriction à cette maille discrète d'une réalisation d'une FAI-0 admettant le variogramme souhaité $\gamma(h)$: il s'agit même, ici encore, d'une réalisation d'une représentation localement stationnaire sur $(0, L)$.

H - EXEMPLE DU VARIOGRAMME $|h|^\alpha$

Considérons le cas particulier où la fonction $\psi(\lambda)$ est

$$\psi(\lambda) = \lambda^{-\beta} \quad (0 < \beta < 1)$$

Le processus X_t correspondant admet donc la loi stable dont la transformée de Laplace est $\exp(-t \lambda^\beta)$. Son potentiel admet la transformée

$$\hat{U}(\lambda) = 1/\lambda^\beta$$

Donc le potentiel $U(dx)$ admet la densité :

$$u(x) = \frac{x^{\beta-1}}{\Gamma(\beta)}$$

De même, la fonction $\theta(x)$ admet la transformée

$$\hat{\theta}(\lambda) = \lambda^{\beta-1}$$

et par suite :

$$\theta(x) = \frac{x^{-\beta}}{\Gamma(1-\beta)}$$

D'après (22), le variogramme que l'on obtiendra par la technique de simulation exposée ci-dessus sera donc :

$$\gamma(h) = \sigma^2 \left(\frac{h}{L} \right)^{1-\beta}$$

Pour simuler $\gamma(h) = |h|^\alpha$, il convient donc de prendre :

$$\beta = 1 - \alpha \quad (0 < \alpha < 1)$$

La loi d'entrée admet alors la fonction de répartition très simple

$$(23) \quad F_L(y) = \left(\frac{y}{L} \right)^\alpha \quad (0 \leq y \leq L)$$

Il faut ensuite trouver la forme explicite de $Q(x_0; h)$. D'après (20) et les résultats déjà obtenus pour les fonctions $\theta(x)$ et $u(x)$, il vient :

$$(24) \quad Q(x_0; h) = \frac{1}{\Gamma(\alpha) \Gamma(1-\alpha)} \int_0^{x_0} x^{-\alpha} (x_0 + h - x)^{\alpha-1} dx$$

Cette expression évoque la fonction de répartition d'une variable beta de paramètre $1-\alpha, \alpha$. Soit, en effet, Π une variable admettant la densité

$$g(\omega) = \frac{\omega^{-\alpha} (1-\omega)^{\alpha-1}}{\Gamma(\alpha) \Gamma(1-\alpha)}$$

Il vient immédiatement :

$$Q(x_0; h) = \int_0^{x_0/(x_0+h)} g(\omega) d\omega$$

Ainsi, lorsque la condition initiale est $Y_0 = 0$, le premier point $x_0 + Y$ du processus ponctuel tombant à droite de x_0 admet la loi de probabilité définie par :

$$P(Y > h) = Q(x_0; h) = P\left(\Pi < \frac{x_0}{x_0+h}\right)$$

D'où l'équivalence en loi :

$$(25) \quad Y \equiv x_0 \left(\frac{1}{\Pi} - 1\right)$$

où Π est $\text{beta}(1-\alpha, \alpha)$.

Ceci permet la construction effective de la simulation :

- ~ On choisit la variable d'entrée selon les quantiles de $F_i(y) = (y/L)^\alpha$
- ~ De $i = 0$ à $i_0 = \text{Int}(y)$, on met $Z(i) = \varepsilon_0$
- ~ On prend $x_0 = i_0 + 1 - y$, on tire au sort ω selon la loi beta $(1-\alpha, \alpha)$ et on prend $y_1 = x_0(1/\omega - 1)$. De $i = i_0 + 1$ à $i_1 = \text{Int}(i_0 + 1 + y_1)$, on met $Z_i = \varepsilon_1$
- ~ et ainsi de suite, jusqu'à atteinte du point terminal $i = L$.

REMARQUE - Dans le cas limite $\alpha = 1$, c'est-à-dire pour un variogramme linéaire $\gamma(h) = |h|$, la loi d'entrée (23) se réduit à la loi uniforme sur $(0, L)$. Pour $\alpha = 1$, la variable Π est presque sûrement égale à 0, de sorte que, d'après (25), Y est toujours presque sûrement égal à l'infini : il y a donc un seul point de séparation y_0 , uniformément distribué sur $(0, L)$, et $Z(x) = \varepsilon_0$ pour $x \leq y_0$, $Z(x) = \varepsilon_1$ pour $x > y_0$: on retrouve ainsi un procédé connu de simulation du variogramme linéaire.

I - SECOND PROCÉDE : SIMULATION DE FAI-1 DE TYPE SPLINE

Voici maintenant un second procédé, plus facile à réaliser lorsque l'on ne connaît pas l'expression explicite du potentiel $U(dx)$. Nous prenons cette fois tout à fait au sérieux la

discrétisation : comme nous ne construirons, de toute façon, la réalisation de $Z(x)$ que sur les points d'abscisses entières $x = 0, 1, 2, \dots, L$, nous pouvons remplacer $\gamma(h)$ par n'importe quel variogramme prenant, aux points d'abscisses entières $h = n$, les valeurs correctes γ_n . Par exemple, nous pouvons prendre :

$$(26) \quad \gamma(h) = \gamma_n + (h-n) (\gamma_{n+1} - \gamma_n) \quad (n = \text{Int}(h))$$

Ce variogramme est concave et, de plus, linéaire au voisinage de $h = 0$: la simulation correspondante est donc toujours possible par les techniques exposées au paragraphe 2.

La variable d'entrée admet la fonction de répartition $F_L(y) = \gamma(y) / \gamma(L)$: la variable correspondante Y est la somme d'une variable U uniforme sur $(0, 1)$ et d'une variable N à valeurs entières telle que :

$$P(N = n) = (\gamma_{n+1} - \gamma_n) / \gamma_L \quad (n = 0, 1, \dots, L-1)$$

En pratique, la valeur exacte prise par U ne joue aucun rôle, et on peut prendre pour la variable d'entrée :

$$Y = N + 1/2$$

Pour un quantile d donné, Y_d est $n_d + 1/2$, où n_d est le plus petit entier n tel que $d < \gamma_{n+1} / \gamma_L$.

La loi F des intervalles séparant les points suivants du processus de renouvellement est concentrée sur les entiers strictement positifs. Elle est définie par

$$P(Y = n) = \frac{2 \gamma_n - \gamma_{n-1} - \gamma_{n+1}}{\gamma_1} \quad (n = 1, 2, \dots)$$

et la fonction de répartition correspondante est :

$$T(y) = 1 - F(y) = (\gamma_{n+1} - \gamma_n) / \gamma_1 \quad (n = \text{Int}(y))$$

D'où la technique de simulation : on tire au sort une variable u uniforme sur $(0, 1)$, et on prend pour y le plus petit entier n tel que $1 - u$ ait

$$u \geq (\gamma_{n+1} - \gamma_n) / \gamma_1$$

EXEMPLE - On peut, par ce procédé, simuler un variogramme en $|h|^\alpha$ avec $0 < \alpha < 1$. Un exemple plus intéressant concerne la simulation de FAI-1 dé type spline, c'est-à-dire admettant une covariance généralisée de la forme

$$K(h) = h^2 \log |h|$$

En réalité, nous souhaitons seulement simuler une réalisation d'une représentation $Z(x)$ de cette FAI-1 sur les points d'abscisse entière $x = 0, 1, \dots, L$, c'est-à-dire, en fait, une FAI-1 discrète Z_n admettant la covariance généralisée $n^2 \log |n|$. Choisissons la représentation nulle en $n = 0$ et $n = 1$, et posons

$$Y_n = Z_{n+1} - Z_n$$

Y_n est la représentation, nulle en $n = 0$, de la FAI-0 discrète admettant le variogramme :

$$\gamma_n = (n+1)^2 \log (n+1) - 2 n^2 \log n + (n-1)^2 \log (n-1)$$

Il se trouve qu'avec ces valeurs de γ_n , le variogramme $\gamma(h)$ linéaire par morceaux, défini en (26) est une fonction concave de h (pour $h \geq 0$). Il est donc possible, et même très facile, de simuler cette FAI-0 discrète Y_n par le procédé indiqué ci-dessus. Prenant ensuite $Z_0 = 0$, et, par récurrence

$$Z_{n+1} = Z_n + Y_n$$

on en déduit une simulation de la FAI-1 Z_n admettant la covariance généralisée $n^2 \log n$.

REFERENCES

- FELLER, W.- Introduction to Probability Theory and its Applications. Wiley and Sons, New York, 1966.
- MATHERON, G.- The intrinsic random functions, and their applications. Advances in Applied Probability, 5 (1973), pp. 439-468.