



ECOLE DES MINES
DE PARIS
Collège doctoral

T H E S E

pour obtenir le grade de
Docteur de l'Ecole des Mines de Paris
Spécialité «Géostatistique»

présentée et soutenue publiquement

par
Xavier EMERY

le 6 février 2004

SIMULATION CONDITIONNELLE DE MODÈLES ISOFACTORIELS

Directeur de thèse : Christian LANTUÉJOUL

Jury composé de :

M.	BACRO Jean-Noël	Président
M.	ALFARO Marco	Rapporteur
M.	HU Lin-Ying	Rapporteur
M.	CHILÈS Jean-Paul	Examineur
M.	JEULIN Dominique	Examineur

Remerciements

Je tiens à exprimer ma profonde gratitude à Christian LANTUÉJOUL, directeur de thèse et ami, pour son soutien et l'encadrement apporté au long de ces trois années. Ses commentaires et sa rigueur intellectuelle ont permis d'orienter et d'enrichir ce travail. Qu'il se rassure : la limite des 200 pages a été respectée, au prix d'un grand effort de rédaction et de concision.

Je suis très reconnaissant à Marco ALFARO et Lin-Ying HU d'avoir accepté d'être rapporteurs et d'avoir corrigé ce mémoire. Mes remerciements vont aussi à Jean-Paul CHILÈS pour ses observations et l'aide apportée non seulement dans le cadre de cette thèse, à Dominique JEULIN pour son examen minutieux du travail réalisé, et à Jean-Noël BACRO pour ses remarques et suggestions.

Je sais gré à tous les membres du Centre de Géostatistique, chercheurs, thésards, et personnel administratif, pour l'accueil chaleureux et la collaboration dont ils ont fait preuve au cours de mes séjours à Fontainebleau.

Une grande partie de cette thèse a été préparée à l'Université du Chili et a bénéficié de l'apport et du soutien de nombreuses personnes que je tiens à remercier : mes collègues de la chaire d'Evaluation des Gisements, Eduardo MAGRI et Julián ORTIZ, la direction du Département d'Ingénierie minière, ses professeurs et son personnel administratif, sans oublier mes amis et collègues du monde minier, en particulier Pedro CARRASCO et Alex CARPANETTI.

La réalisation de ce travail a été possible grâce au programme de formation de jeunes professeurs mis en place par la Faculté des Sciences Physiques et Mathématiques de l'Université du Chili, au patronage de Codelco-Chile destiné au Département d'Ingénierie minière, ainsi qu'à Meridian Gold et aux professionnels de la mine El Peñón qui m'ont généreusement fourni la base de données utilisée dans ce mémoire.

Enfin et surtout, je souhaite rendre hommage à ma famille et à Natalia, qui m'ont accompagné et constamment encouragé et à qui je dédie ce travail.

Table des matières

Introduction	1
I. L'alternative estimation - simulation.....	1
II. Limitations du modèle multigaussien	1
III. Principaux algorithmes de simulation non gaussienne	3
III.1. Méthodes “exactes” (simulation stochastique).....	3
III.1.1. Modèle mosaïque	3
III.1.2. Modèles construits par implantation d’objets.....	3
III.1.3. Modèles de substitution	4
III.1.4. Simulations gaussiennes tronquées et plurigaussiennes	4
III.2. Méthodes “passe-partout” (imagerie stochastique)	5
III.2.1. Techniques d’optimisation	5
III.2.2. Simulation séquentielle d’indicatrices.....	5
III.3. Une posture intermédiaire : la simulation séquentielle isofactorielle	7
IV. Notations	8
Chapitre 1 : comparaison entre krigeage d’indicatrices, krigeage disjonctif et espérance conditionnelle	9
I. Objectifs.....	9
II. Modèle multigaussien	10
II.1. Une donnée conditionnante	10
II.2. Deux données conditionnantes	12
III. Modèle mosaïque	20
III.1. Une donnée conditionnante.....	21
III.2. Deux données conditionnantes.....	21
III.2.1. Mosaïque de Poisson	22
III.2.2. Mosaïque des feuilles mortes	26
III.2.3. Conclusions	30
IV. Modèle bigamma	31
IV.1. Une donnée conditionnante	31
IV.2. Deux données conditionnantes.....	33
V. Fonctions aléatoires en tout ou rien.....	37
V.1. Fonctions aléatoires en damier.....	37
V.1.1. Deux données conditionnantes	39
V.1.2. Trois données conditionnantes	41
V.2. Processus autoduaux	43
V.3. Gaussienne tronquée.....	44
VI. Discussion	49

Chapitre 2 : modèles de fonctions aléatoires associés à l'algorithme séquentiel	51
I. Objectif	51
II. Modèle mosaïque à valuations continues	51
II.1. Condition pour reproduire les lois marginale et bivariable	52
II.2. Loi de probabilité multivariable des valeurs simulées	53
II.3. Conditionnement à un ensemble de données	54
II.4. Cohérence des lois multivariables	55
II.5. Exemple : déplacement aléatoire du point médian	58
II.6. Extension à plusieurs dimensions	60
III. Modèle d'ensemble aléatoire	61
III.1. Conditions sur les pondérateurs de krigeage.....	61
III.2. Cas d'une indicatrice de moyenne $1/2$	62
III.2.1. Point de vue matriciel.....	62
III.2.2. Point de vue du krigeage dual	62
III.2.3. Conjecture.....	63
III.2.4. Autodualité.....	63
III.2.5. Exemple : processus autodual à covariance exponentielle sur la droite réelle	65
III.3. Ensembles aléatoires non autoduaux	68
IV. Modèles à indicatrices multiples.....	69
IV.1. Condition pour reproduire les moyennes et covariances d'indicatrices.....	69
IV.2. Exemple : indicatrices à covariance exponentielle	70
IV.3. Covariances croisées d'indicatrices.....	71
IV.4. Simulation hiérarchique	73
IV.5. Variables catégorielles	75
V. Modèles isofactoriels	75
V.1. Condition pour reproduire les lois marginale et bivariable	75
V.2. Sur la positivité des pseudo-densités du krigeage disjonctif	76
V.3. Processus markoviens à une dimension	80
VI. Extensions des modèles précédents.....	83
VI.1. Composition de modèles	83
VI.2. Modèles interrompus	83
 Chapitre 3 : modèles isofactoriels à plusieurs dimensions	 85
I. Objectif	85
II. Fonctions aléatoires de substitution.....	85
II.1. Définition et application à la simulation de modèles isofactoriels	85
II.2. Modèle hermitien.....	87
II.2.1. Quelques exemples.....	87
II.2.2. Inférence	91
II.2.3. Extensions pour obtenir des modèles à portées intégrales finies	92
III. Modèles construits à partir d'ensembles aléatoires	95
III.1. Quelques exemples d'ensembles aléatoires	95
III.1.1. Modèle de substitution avec un processus de codage binaire	95
III.1.2. Troncature de processus isofactoriels hermitiens	96
III.1.3. Schéma booléen stationnaire	97
III.1.4. Ensemble aléatoire des feuilles mortes.....	98
III.2. Modèle à résidus d'indicatrices orthogonaux	99
III.2.1. Modèle discret sans effets de bord en descendant	99

III.2.2. Modèle discret sans effets de bord en montant.....	100
III.2.2. Modèle continu.....	101
III.3. Modèle binomial.....	104
IV. Modèles à lois indéfiniment divisibles	107
IV.1. Simulation par bandes tournantes.....	107
IV.1.1. Principe des bandes tournantes.....	107
IV.1.2. Applicabilité aux modèles isofactoriels ?.....	108
IV.1.3. Exemple de la loi de Laguerre.....	109
IV.1.4. Conditionnement à des données.....	112
IV.1.5. Extension aux modèles gigognes et anisotropes.....	112
IV.1.6. Généralisation aux modèles à lois indéfiniment divisibles.....	112
IV.1.7. Difficultés pratiques et limitations	114
IV.2. Dilution poissonienne	116
IV.3. Addition de mosaïques indépendantes.....	119
IV.3.1. Modèle de Laguerre.....	119
IV.3.2. Exemple de la mosaïque des feuilles mortes.....	120
IV.3.3. Prise en compte d'un effet de pépite	121
IV.3.4. L'inférence des paramètres du modèle	122
IV.3.5. Autres modèles à loi marginale indéfiniment divisible	124
V. Le modèle bigamma	125
V.1. Carré d'une fonction aléatoire bigaussienne	125
V.2. Modèle bigamma de paramètre demi-entier	128
Chapitre 4 : l'inférence statistique	131
I. Détermination de la loi marginale	131
I.1. Considérations générales.....	131
I.2. Commentaires sur la procédure d'anamorphose.....	132
II. Modélisation des lois bivariées	134
II.1. Méthodes d'indicatrices.....	134
II.1.1. Analyse variographique de chaque indicatrice	134
II.1.2. Analyse variographique conjointe	135
II.2. Modèles isofactoriels : généralités.....	135
II.3. L'exemple du modèle de Laguerre	137
II.3.1. Retour sur la propriété de déstructuration des teneurs extrêmes	137
II.3.2. Stratégie à adopter pour la modélisation des lois bivariées.....	143
II.4. Les variogrammes d'ordres inférieurs à deux	144
II.4.1. Modèles bigaussien et bigamma	144
II.4.2. Modèles hermitien et laguerrien	145
III. Lois multivariées	146
Chapitre 5 : application à des données minières.....	149
I. Présentation des données	149
II. Etude exploratoire et choix des données à analyser.....	151
III. Choix du modèle	154
III.1. Anamorphose	155
III.2. Un modèle bivariable	155
III.3. Mise en oeuvre pratique	160
IV. Simulation conditionnelle	164
V. Réflexions finales.....	167

Conclusions	169
Annexe A : rappels sur le krigeage disjonctif à support ponctuel	171
I. Définition générale	171
II. Les familles génératrices.....	172
III. Seconde définition du krigeage disjonctif	172
IV. Krigeage disjonctif et krigeage d'indicatrices	173
V. Cas de simplification du krigeage disjonctif	173
V.1. Modèle mosaïque	174
V.2. Modèles isofactoriels	174
Annexe B : modèles isofactoriels à support ponctuel	175
I. Les lois isofactorielles symétriques	175
I.1. Définition	175
I.2. Propriétés	176
I.3. Le krigeage disjonctif isofactoriel.....	177
II. Exemples	177
II.1. Modèles bigaussien et hermitien.....	177
II.2. Modèles bigamma et de Laguerre.....	178
II.3. Autres modèles.....	178
II.4. Modèles de type mosaïque	180
III. Deux modes de construction	180
IV. Aspects pratiques	180
IV.1. Anamorphose	180
IV.2. Choix des paramètres de la loi bivariable	181
V. Un critère pour identifier un modèle isofactoriel.....	182
Annexe C : simulations par chaînes de Markov	185
I. Principes de la simulation par itérations markoviennes	185
II. Conditionnement des simulations	188
II.1. Restriction du noyau de transition	188
II.2. Recuit simulé	188
Bibliographie	191

Introduction

I. L'alternative estimation - simulation

Les techniques de krigeage permettent d'estimer localement les valeurs prises par une variable régionalisée à partir d'un échantillonnage de cette dernière. Cependant, la propriété de lissage induite par le krigeage empêche d'apprécier la variabilité des valeurs inconnues (non mesurées) : ainsi, l'histogramme des valeurs estimées est moins dispersé que celui des valeurs réelles ; le variogramme des valeurs estimées n'est pas non plus représentatif de la réalité.

Les techniques de simulation conditionnelle visent à construire des images réalistes de la variable régionalisée, respectant ses propriétés statistiques (histogramme, variogramme, etc.) et restituant les mesures effectuées aux sites d'échantillonnage. Leur mise en œuvre fait appel à l'interprétation de la variable régionalisée en tant que réalisation d'une fonction aléatoire, et consiste à construire d'*autres* réalisations de cette même fonction aléatoire. De telles réalisations constituent des scénarii plausibles et équiprobables et permettent de mesurer l'incertitude portant sur les valeurs inconnues de la variable régionalisée.

Les simulations conditionnelles sont utiles pour estimer des grandeurs qui ne sont pas directement accessibles par krigeage. Ainsi, en géostatistique minière, la probabilité que la teneur en métal d'une unité de sélection (de support plus volumineux que celui des échantillons) soit supérieure à une teneur de coupure est évaluée par la fréquence avec laquelle les teneurs simulées de cette unité dépassent la teneur de coupure. Un tel outil est une aide précieuse pour la prise de décision et la planification du projet minier.

II. Limitations du modèle multigaussien

La construction de simulations est particulièrement simple dans le cadre des fonctions aléatoires de loi spatiale multigaussienne, c'est-à-dire telles que toute combinaison linéaire de valeurs suit une distribution gaussienne. De nombreux algorithmes ont été développés pour obtenir des simulations non conditionnelles : décomposition matricielle, moyennes mobiles discrètes, méthodes spectrales, bandes tournantes, méthode séquentielle, pour n'en citer que quelques-uns (Alabert, 1987 ; Boulanger, 1990 ; Chilès, 1995 ; Davis, 1987 ; Lantuéjoul, 1994, 2002 ; Matheron, 1972b, 1973a) ; le conditionnement à un ensemble de données se résout en général à l'aide d'un krigeage simple (Journel, 1974).

La variable régionalisée doit être préalablement transformée en une variable dont la distribution marginale est gaussienne, laquelle est supposée posséder une loi spatiale multigaussienne. Une telle transformation est connue sous le nom "d'anamorphose" gaussienne (Matheron, 1974 ; Journel and Huijbregts, 1978 ; Rivoirard, 1994) ; il s'agit par exemple d'un passage au logarithme si la variable suit une distribution lognormale. Une fois la procédure de simulation achevée dans le cadre gaussien, l'anamorphose inverse est appliquée pour revenir à la variable initiale.

La distribution multigaussienne est extrêmement commode à utiliser car elle ne dépend que des moments d'ordre 1 et 2 (espérance et fonction de covariance ou variogramme), ce qui facilite la tâche d'inférence statistique. Toutefois, elle présente plusieurs caractéristiques limitatives :

- difficulté de réaliser l'anamorphose lorsque l'histogramme de la variable régionalisée est fort dissymétrique ou présente une proportion importante de valeurs nulles (Matheron, 1980a ; Rivoirard, 1994) ;
- symétrie des distributions de probabilité par rapport à la valeur médiane : les caractéristiques structurales des valeurs élevées (anisotropie, continuité...) se produisent également pour les valeurs faibles ;
- “désordre” dans l'agencement spatial des valeurs, ce qui traduit un phénomène de type diffusif (Matheron, 1989b). Parmi les lois spatiales admettant une même fonction de covariance, la loi multigaussienne maximise l'entropie, qui peut s'interpréter comme une mesure du désordre spatial (Jones, 1979 ; Christakos, 1990 ; Chilès and Delfiner, 1999) ;
- propriété de “déstructuration des hautes teneurs” (Matheron, 1982 ; Journel, 1993 ; Goovaerts, 1997) : l'occurrence des valeurs extrêmes devient purement aléatoire, ce qui empêche que les valeurs très fortes ou très faibles soient spatialement connectées.

Ces caractéristiques motivent la recherche d'algorithmes de simulation non gaussienne. Or, quatre problèmes se posent lorsque l'on cherche à réaliser des simulations en dehors du cadre multigaussien :

- 1) spécification d'un modèle de loi spatiale, cohérent mathématiquement
- 2) inférence des paramètres du modèle à partir des données expérimentales
- 3) construction de réalisations de ce modèle
- 4) conditionnement de la simulation aux données.

Souvent, les points 1) et 3) sont interdépendants : la loi spatiale est fréquemment définie à partir d'un procédé de construction de fonctions aléatoires. Citons comme exemples les modèles booléen (Matheron, 1975a), des feuilles mortes (Matheron, 1968a ; Jeulin, 1997), de substitution (Lantuéjoul, 1993, 2002) et la technique des plurigaussiennes (Armstrong *et al.*, 2003). Cependant, le dernier point est souvent difficile à réaliser.

Certains algorithmes dérogent à la règle précédente, en ce sens que leur mise en œuvre ne requiert pas de spécifier un modèle de loi spatiale, mais seulement un nombre limité de paramètres (souvent, la loi marginale et une partie de la loi bivariable). De tels algorithmes se veulent “passe-partout” et ont été élaborés en vue de s'adapter à tous les types de variable régionalisée, ce qui explique leur côté séduisant pour l'utilisateur.

Les questions que soulève cette dernière approche sont de deux ordres :

- 1) qu'un algorithme produise une image conforme à l'idée que l'on se fait de la réalité, n'est pas une garantie de sa qualité et son utilisation peut conduire à des résultats erronés. Il est donc important de connaître le degré d'approximation des algorithmes “passe-partout”. Le vocabulaire utilisé par les promoteurs de tels algorithmes (“imagerie” stochastique) trahit la différence conceptuelle avec la “simulation” stochastique, qui se réfère à un modèle complet de loi spatiale.
- 2) peut-on se passer d'un modèle de fonction aléatoire et des limitations inhérentes à ce modèle ? Comme écrivait G. Matheron (1982), une méthode passe-partout, qui serait utilisable quelle que soit la variable étudiée, a peu de chance de tomber juste.

III. Principaux algorithmes de simulation non gaussienne

III.1. Méthodes “exactes” (simulation stochastique)

Nous énumérons ci-après plusieurs méthodes fournissant des réalisations exactes de fonctions aléatoires non gaussiennes. Le choix de telle ou telle méthode est dicté par les propriétés structurales observées sur la variable régionalisée.

III.1.1. Modèle mosaïque

L'espace est partitionné en cellules aléatoires stationnaires, dont les caractéristiques ne dépendent pas de leur position dans l'espace. On assigne à tous les sites d'une même cellule une valuation, tirée selon une loi de probabilité de moyenne et variance finies. Les valeurs tirées dans les cellules sont indépendantes les unes des autres. On montre que la fonction aléatoire générée (dite “mosaïque”) est stationnaire d'ordre deux et sa covariance est proportionnelle au covariogramme géométrique des cellules, lequel mesure la probabilité que deux sites appartiennent à la même cellule en fonction de leur séparation (Matheron, 1982 ; Rivoirard, 1994). Il existe de nombreux modèles de partition ou “tessellation”, par exemple :

- **mosaïque de Poisson** (Miles, 1961, 1974 ; Matheron, 1975a ; Alfaro, 1979) : la partition est définie par un réseau d'hyperplans poissoniens. La covariance simulée est un schéma exponentiel dont la portée est inversement proportionnelle à la densité des hyperplans poissoniens ;
- **mosaïque de Voronoï** (Gilbert, 1962 ; Miles, 1974) : la tessellation est constituée par les frontières des polyèdres d'influence de points poissoniens dans l'espace ;
- **mosaïque de Johnson-Mehl** (Lantuéjoul, 1977) : la construction est “dynamique” : elle consiste à faire apparaître, au cours du temps, des points poissoniens en chacun desquels une cellule se développe de manière isotrope. La croissance d'une cellule dans une direction s'arrête lorsqu'elle rencontre une autre cellule ;
- **mosaïque des feuilles mortes** (Matheron, 1968a ; Jeulin, 1997), qui est aussi définie de manière dynamique : des points poissoniens apparaissent successivement, en chacun desquels est implanté un ensemble aléatoire ou *grain primaire* ; les grains nouveaux recouvrent les grains anciens. Au bout d'un temps très long, les frontières apparentes des grains constituent une partition aléatoire.

Le conditionnement des simulations à un ensemble de données est un problème qui ne possède pas encore de solution générale. Certains cas sont analysés par Lantuéjoul (2002).

III.1.2. Modèles construits par implantation d'objets

Dans ces modèles, la construction de la simulation est réalisée de manière “dynamique”. Des points de position aléatoire apparaissent au cours du temps ; en général, ils sont issus d'un processus ponctuel de Poisson dont la densité peut être uniforme ou variable dans l'espace ou dans le temps. Chaque point est le centre de gravité d'un “objet”, par exemple une forme géométrique prédéterminée, de taille et orientation aléatoires et de valuation uniforme ; la valuation peut aussi être une fonction de la distance au centre de gravité ou dépendre de l'instant où apparaît l'objet. Les objets sont appelés *grains primaires* ou *fonctions primaires*, selon que leur valuation est constante ou non.

Lorsqu'un objet nouveau rencontre un objet ancien, une convention est adoptée pour définir la valeur des sites où les deux objets se chevauchent, par exemple :

- valeur du dernier objet ou du premier objet : *modèle des feuilles mortes* (Jeulin, 1989, 1997)
- somme des valeurs des objets : *modèle de dilution* (Serra, 1968 ; Alfaro, 1979)
- produit des valeurs des objets : *modèle des jetons multiplicatifs* (Guiblin *et al.*, 1995)
- valeur maximale : *modèle booléen* (Jeulin and Jeulin, 1981 ; Serra, 1988).

La simulation s'arrête au bout d'un temps suffisamment long. Cette méthode de simulation par lancement d'objets fournit de nombreux modèles de fonctions aléatoires non gaussiennes. La mise en œuvre pratique pose certaines difficultés :

- 1) la détermination de l'instant où arrêter la simulation, car la convergence du processus simulé vers le processus asymptotique peut être lente ;
- 2) l'inférence des paramètres du modèle (densité du processus poissonien, taille et forme des objets, type de valuation, convention pour le recouvrement des objets) ;
- 3) le conditionnement de la simulation à des données : souvent, il est nécessaire de recourir à des procédures itératives d'acceptation et rejet (Lantuéjoul, 2002).

III.1.3. Modèles de substitution

Une fonction aléatoire de substitution s'écrit sous la forme d'une composition de deux processus : une *fonction directrice* et un *processus de codage* (Matheron, 1989b ; Lantuéjoul, 1993, 2002). Des hypothèses doivent être vérifiées par ces deux processus pour que la fonction aléatoire résultante ait de bonnes propriétés statistiques telle la stationnarité d'ordre deux. La procédure de substitution est très flexible pour imposer aux réalisations certaines propriétés morphologiques (anisotropie, stratification, relations topologiques entre les valeurs, etc.). Cette technique sera revue en détail au chapitre 3.

III.1.4. Simulations gaussiennes tronquées et plurigaussiennes

La technique des gaussiennes tronquées simule un ensemble aléatoire, que l'on représente par une variable binaire ou "indicatrice", par seuillage d'une fonction aléatoire de loi spatiale multigaussienne (Journel and Isaaks, 1984 ; Matheron *et al.*, 1987 ; de Fouquet *et al.*, 1989). On peut étendre cette approche et simuler une partition de l'espace en plusieurs ensembles aléatoires en travaillant sur plusieurs fonctions aléatoires gaussiennes. Cette généralisation, connue sous le nom de méthode des plurigaussiennes (Le Loc'h *et al.*, 1994 ; Armstrong *et al.*, 2003), fournit une vaste classe de modèles capables de reproduire de nombreuses caractéristiques morphologiques des ensembles aléatoires à simuler, notamment les relations de contiguïté et d'emboîtement.

III.2. Méthodes “passe-partout” (imagerie stochastique)

III.2.1. Techniques d'optimisation

L'idée est de perturber progressivement une image initiale, afin d'améliorer la reproduction de certaines caractéristiques (histogramme, variogramme, variogrammes d'indicatrices...). La nouveauté est que le problème de simulation est entièrement formulé en termes d'optimisation, sans se référer à un modèle de fonction aléatoire : il s'agit de construire une image qui minimise une *fonction objectif* donnée, mesurant l'écart entre les caractéristiques structurales expérimentales de l'image et celles souhaitées (Deutsch, 1992 ; Deutsch and Cockerham, 1994 ; Deutsch and Journel, 1998).

En pratique, l'image initiale est soit une simulation jugée insatisfaisante, soit une image respectant déjà l'histogramme à simuler. On lui applique une perturbation élémentaire, en vue de diminuer la valeur de la fonction objectif. Typiquement, on échange les valeurs prises en deux sites quelconques. L'échange est accepté s'il est favorable, c'est-à-dire s'il diminue la valeur de la fonction objectif, rejeté sinon. En fait, on autorise aussi certaines perturbations non favorables, avec une probabilité dépendant d'un paramètre appelé *température*. Plus la température est élevée, plus une perturbation défavorable est susceptible d'être acceptée ; au contraire, si la température devient nulle, seules les perturbations favorables sont acceptées.

Usuellement, on part d'une température donnée que l'on abaisse peu à peu (ni trop lentement, ni trop rapidement) : plus la température baisse, plus l'image a tendance à se figer. Ce procédé présente des analogies avec celui utilisé en métallurgie pour le recuit d'un métal, d'où le nom de *recuit simulé* donné à la méthode. Cette terminologie est malheureuse, car elle prête à confusion avec une technique proposée par Kirkpatrick *et al.* (1983) (*cf.* annexe C) ; c'est pourquoi on préférera ici la terminologie “technique d'optimisation” à celle de “recuit simulé”.

L'algorithme s'arrête lorsque la fonction objectif atteint une valeur prédéfinie, signifiant un bon accord de l'image avec les caractéristiques souhaitées. Toutefois, outre sa lourdeur algorithmique, la méthode ne garantit en rien la reproduction des caractéristiques structurales qui n'entrent pas dans la fonction objectif, comme par exemple le variogramme le long des directions ou pour les classes de distances non spécifiées. Il n'est d'ailleurs pas rare de voir apparaître des structures “parasites” sur les réalisations obtenues (anisotropies non souhaitées, décrochage du variogramme au-delà de la dernière distance spécifiée, fluctuation anormale du variogramme entre deux distances consécutives, etc.).

III.2.2. Simulation séquentielle d'indicatrices

Il s'agit d'une technique itérative : la valeur en un site est simulée selon sa fonction de répartition conditionnée par les données initiales et les valeurs précédemment simulées (Alabert, 1987 ; Deutsch and Journel, 1998 ; Goovaerts, 1997 ; Journel, 1989). Cependant, hormis quelques rares cas particuliers comme le cas multigaussien, une telle fonction de répartition est inaccessible et l'on se contente de l'estimer par krigeage d'indicatrices ; cette fonction estimée, imparfaite et incomplète, fait l'objet de corrections (figure 1).

L'avantage de cette méthode est sa flexibilité car elle s'affranchit de l'hypothèse multigaussienne. En utilisant un krigeage d'indicatrices, les traits structuraux de la variable régionalisée sont modélisés par les covariances des indicatrices associées à plusieurs seuils de coupure : on peut ainsi rendre compte d'une éventuelle structuration des valeurs extrêmes. Par ailleurs, l'emploi des indicatrices permet de coder aisément les incertitudes de mesures et les informations secondaires, qualitatives ou non (Zhu and Journel, 1993). Enfin, les réalisations obtenues sont directement conditionnées aux données initiales car celles-ci sont utilisées dans chaque krigeage d'indicatrices.

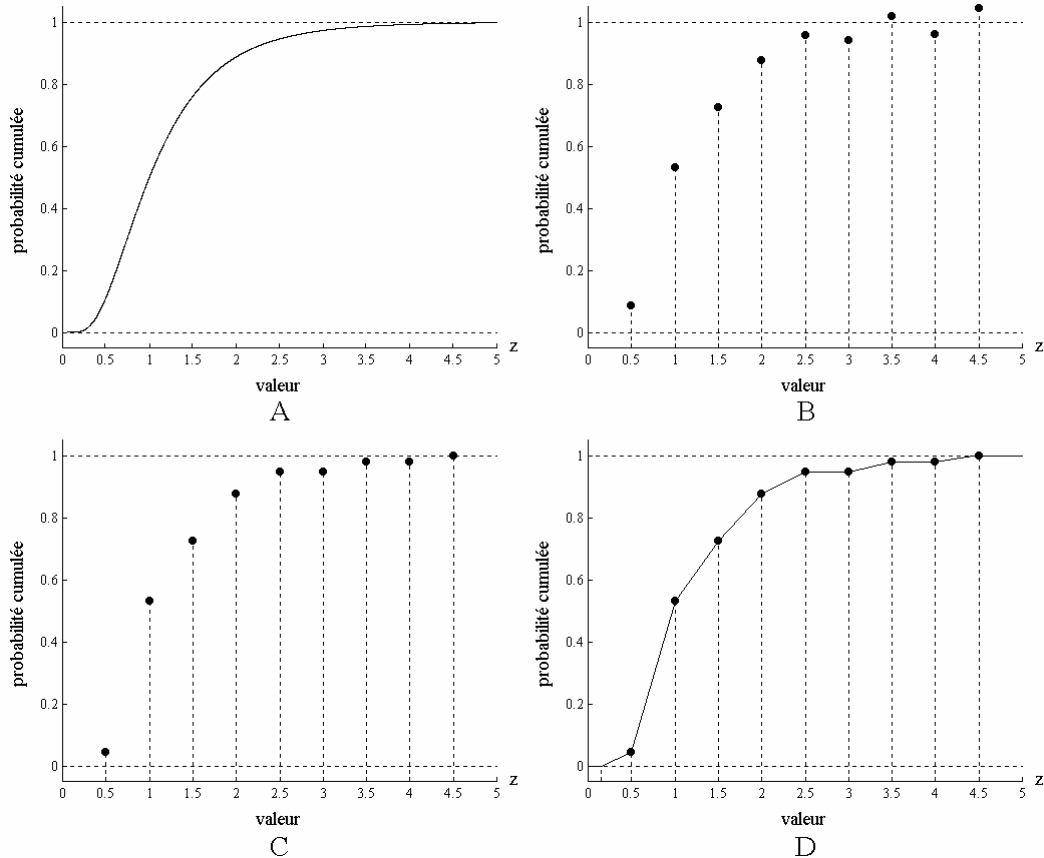


Figure 1. Fonctions de répartition conditionnelles en un site : A, théorique, B, estimée par krigeage d'indicatrices, C, corrigée et D, interpolée et extrapolée

Bien que sa simplicité conceptuelle et sa flexibilité puissent séduire l'utilisateur, l'algorithme séquentiel d'indicatrices souffre de plusieurs défauts dont les effets ne se compensent pas mais tendent au contraire à s'accumuler :

- 1) analyse variographique : n'importe quel variogramme n'est pas autorisé pour décrire une variable binaire (Matheron, 1987, 1989a, 1993) ; de surcroît, les variogrammes des indicatrices associées à des seuils différents doivent satisfaire certaines contraintes de compatibilité car ces indicatrices ne sont jamais indépendantes (Rivoirard, 1993) ;
- 2) caractère non optimal des estimations d'indicatrices (typiquement, un krigeage séparé de chaque indicatrice, au lieu d'une espérance conditionnelle, d'où une perte importante de précision) ;
- 3) problèmes de "relations d'ordre" entre les estimations d'indicatrices, qui requièrent une correction *a posteriori* (Deutsch and Journel, 1998 ; Goovaerts, 1997 ; Sullivan, 1984) ;
- 4) discrétisation : les fonctions de répartition obtenues par krigeage d'indicatrices sont incomplètes et doivent faire l'objet d'une interpolation entre les seuils de coupure choisis et d'une extrapolation au-delà des seuils extrêmes (Deutsch and Journel, 1998 ; Goovaerts, 1997) ;
- 5) changement de support : les fonctions de répartition estimées se réfèrent au même support que celui des échantillons, d'où l'impossibilité de simuler directement des valeurs sur un support différent, à moins d'avoir recours à une approximation grossière, par exemple soumettre les lois locales à une correction affine (Chilès and Liao, 1993 ; Journel, 1985, 1987) ou recourir à la technique des *variables utiles* consistant à utiliser des pseudo-données se référant au support à estimer (Bouchind'homme, 1980 ; Sullivan, 1984) ;

- 6) conditionnement limité aux valeurs voisines du site à simuler, afin de réduire les temps de calcul qui deviendraient vite prohibitifs si toutes les valeurs étaient gardées (Deutsch and Journel, 1998 ; Goovaerts, 1997) ;
- 7) instabilité numérique lorsque l'on simule des sites très proches les uns des autres¹, surtout si les variogrammes d'indicatrices sont continus à l'origine (apparition de poids de krigeage négatifs qui entraînent l'obtention de densités de probabilité estimées négatives). Ces incohérences soulèvent une autre difficulté conceptuelle, celle de l'*existence* même de la fonction aléatoire (définie continûment en tout site de l'espace) que l'on prétend simuler ; nous aborderons ce point qui peut sembler surprenant dans le deuxième chapitre de ce mémoire.

La section suivante présente une technique qui, bien qu'elle reste imparfaite, résout les points 1, 4 et 5 ci-dessus et améliore le deuxième point.

III.3. Une posture intermédiaire : la simulation séquentielle isofactorielle

L'idée directrice de cette méthode consiste à remplacer, dans l'algorithme séquentiel, le krigeage d'indicatrices par un krigeage disjonctif (Emery, 2002). Ce dernier tire parti de la loi bivariable complète de la fonction aléatoire (Matheron, 1973b, 1976), tandis que le krigeage d'indicatrices n'utilise qu'une partie de cette loi, à savoir les fonctions de covariance de quelques indicatrices. On démontre que le krigeage disjonctif équivaut à un cokrigeage simple de toutes les indicatrices associées à la fonction aléatoire (Rivoirard, 1994) car la loi bivariable est caractérisée par la donnée de toutes les covariances simples et croisées d'indicatrices. Les équations du krigeage disjonctif peuvent être résolues analytiquement lorsque la loi bivariable est *isofactorielle*, c'est-à-dire lorsqu'il existe une famille complète de fonctions sans corrélations spatiales croisées appelées *facteurs* (cf. annexes A et B).

Le krigeage disjonctif améliore l'estimation des fonctions de répartition conditionnelles, élimine les problèmes de discrétisation inhérents aux méthodes d'indicatrices et réduit les contraintes d'analyse variographique². Mais à l'instar du krigeage d'indicatrices, il ne résout pas les problèmes de relations d'ordre : en général, les densités de probabilité conditionnelles estimées peuvent prendre des valeurs négatives, bien que leur intégrale soit égale à 1, ce qui requiert une correction *a posteriori*.

Par rapport aux techniques précédentes, la simulation séquentielle isofactorielle apparaît comme une méthodologie "hybride", au sens où elle repose sur la spécification d'un modèle (un même modèle ne permettra pas de décrire n'importe quelle régionalisation), mais cette spécification reste incomplète (lois bivariées au lieu de multivariées). C'est un compromis entre deux exigences souvent difficiles à concilier : la cohérence mathématique interne du modèle et sa flexibilité d'usage, qualifiée de "cohérence externe" par Journel (1993). Cet argument rejoint une réflexion de Matheron (1978b) selon laquelle un modèle entièrement spécifié de loi spatiale dépasse ce que cet auteur dénomme *seuil de réalisme*.

¹ Ce problème se pose notamment lorsque l'on cherche à simuler des valeurs non ponctuelles par la moyenne de simulations ponctuelles réalisées sur une grille de maille très fine. Il serait donc évité si le changement de support pouvait être pris en compte directement au moment de la simulation (problème n°5).

² Dans les modèles isofactoriels usuels, les covariances des facteurs se déduisent du corrélogramme de la fonction aléatoire, qui est le seul outil variographique requis. Le fait d'utiliser un unique corrélogramme comme dans le modèle multigaussien n'est guère gênant, car certaines propriétés structurales de la variable régionalisée (par exemple, une connectivité des valeurs élevées ou une dissymétrie entre la structuration des valeurs faibles et celle des valeurs fortes) peuvent être prises en compte, selon le modèle choisi, à travers de paramètres scalaires.

Toutefois, il se pose encore un problème de cohérence du modèle : en effet, étant donné un modèle de loi isofactorielle, quels variogrammes sont compatibles avec cette loi ? et existe-t-il une fonction aléatoire admettant de telles lois bivariées isofactorielles ? (Matheron, 1973b, 1987, 1989a). A l'heure actuelle il n'existe pas de critères généraux pour répondre à cette question.

Les algorithmes de simulation séquentielle d'indicatrices et isofactorielle seront le point de départ de cette thèse qui s'intéresse aux simulations de modèles isofactoriels et à leur conditionnement à des données expérimentales. Le choix des lois isofactorielles n'est pas fortuit, car elles constituent une vaste classe de lois bivariées et permettent de s'adapter à la description de nombreux phénomènes. Jusqu'à présent, on sait mener l'estimation locale par krigeage disjonctif, mais il n'existe guère de méthodes de simulation : c'est cette lacune que l'on cherchera à combler.

IV. Notations

Nous introduisons ci-dessous quelques notations usuelles en géostatistique :

- d indique la dimension de l'espace de travail ; en général, $d = 2$ ou 3 , c'est-à-dire que l'on se place dans l'espace géographique \mathbb{R}^2 ou \mathbb{R}^3
- $|D|$ est la mesure (longueur, aire, volume...) du domaine D inclus dans \mathbb{R}^d
- \mathbf{x} est un vecteur de localisation dans l'espace, qui varie continûment dans \mathbb{R}^d
- (x_1, \dots, x_d) sont les coordonnées du vecteur \mathbf{x}
- $\{z(\mathbf{x}), \mathbf{x} \in \mathbb{R}^d\}$ constitue l'ensemble des valeurs prises par la variable régionalisée
- n représente le nombre de sites de données
- $\{\mathbf{x}_\alpha, \alpha = 1 \dots n\}$ sont les sites de données dans \mathbb{R}^d
- $\{Z(\mathbf{x}), \mathbf{x} \in \mathbb{R}^d\}$ est la fonction aléatoire représentant la variable régionalisée
- $\{Y(\mathbf{x}), \mathbf{x} \in \mathbb{R}^d\}$ est la fonction aléatoire "anamorphosée", dont la loi marginale coïncide avec celle du modèle isofactoriel considéré
- ϕ est la fonction d'anamorphose reliant les valeurs initiales et anamorphosées : $Z(\mathbf{x}) = \phi[Y(\mathbf{x})]$
- $I(Y(\mathbf{x}); y)$ est l'indicatrice cumulée associée au seuil de coupure y :

$$I(Y(\mathbf{x}); y) = \begin{cases} 1 & \text{si } Y(\mathbf{x}) < y \\ 0 & \text{sinon} \end{cases}$$

En général, les indices grecs α, β sont réservés pour désigner les sites de données. Les matrices et vecteurs sont notés en caractères gras. Les lettres minuscules désignent les quantités déterministes et les lettres majuscules les quantités aléatoires. Les estimateurs utilisés sont généralement repérés par un indice ou une astérisque placée en exposant ; par exemple, $Y^*(\mathbf{x})$ représente un estimateur de $Y(\mathbf{x})$. Par abus de notation, on écrira parfois $Y(\mathbf{x})$ au lieu de $\{Y(\mathbf{x}), \mathbf{x} \in \mathbb{R}^d\}$ pour désigner la fonction aléatoire.

Terminons cette introduction par une remarque d'ordre typographique : pour faciliter la lecture des chapitres suivants, les démonstrations et points secondaires seront fréquemment écrits en caractères plus petits et mis légèrement en retrait du texte principal.

Chapitre 1 :

Comparaison entre krigeage d'indicatrices, krigeage disjonctif et espérance conditionnelle

I. Objectifs

Dans le chapitre introductif, nous avons présenté des algorithmes de simulation approchée basés sur une estimation des lois de probabilité conditionnelles par un krigeage d'indicatrices ou un krigeage disjonctif (simulation séquentielle d'indicatrices et séquentielle isofactorielle), alors que les véritables lois de probabilité sont obtenues par l'espérance conditionnelle des indicatrices.

Ces trois estimateurs sont hiérarchisés : l'espérance conditionnelle repose sur la loi spatiale de la fonction aléatoire, le krigeage disjonctif sur la loi bivariante complète et le krigeage d'indicatrices sur une partie de la loi bivariante (covariances de plusieurs indicatrices). Que perd-on en passant de l'un à l'autre de ces estimateurs ? Pour répondre à cette question, nous allons nous intéresser à quelques cas simples de fonctions aléatoires et un nombre limité de sites de données, situations pour lesquelles on peut expliciter des lois multivariées et mener à bien les calculs.

L'objectif de cette étude est double : d'une part, fournir un ordre de grandeur des approximations commises par les algorithmes de simulation séquentielle dans des modèles particuliers ; d'autre part, mettre en évidence des situations dans lesquelles le krigeage disjonctif ou d'indicatrice coïncide avec l'espérance conditionnelle, ce qui permettra de trouver les conditions sous lesquelles les algorithmes séquentiels sont exacts.

Les modèles de fonctions aléatoires qui seront analysés sont :

- le *modèle multigaussien* (Anderson, 1958 ; Maréchal, 1978 ; Verly, 1983)
- le *modèle mosaïque* (Matheron, 1982 ; Rivoirard, 1994)
- le *modèle de diffusion bigamma* (Matheron, 1973b ; Hu, 1988)
- *modèles en tout ou rien* : modèles en damier, processus autoduaux, gaussiennes tronquées.

II. Modèle multigaussien

Le premier cas étudié est celui d'une fonction aléatoire $\{Y(\mathbf{x}), \mathbf{x} \in \mathbb{R}^d\}$ de loi spatiale gaussienne stationnaire, centrée et réduite : la densité de probabilité conjointe *a priori* d'un ensemble de valeurs $\{Y(\mathbf{x}_1), \dots, Y(\mathbf{x}_n)\}$ est entièrement caractérisée par la fonction de corrélation $\rho(\mathbf{h})$:

$$g(y_1, \dots, y_n) = \frac{1}{(\sqrt{2\pi})^n \sqrt{\det(\mathbf{R})}} e^{-\frac{1}{2} \mathbf{y}^t \mathbf{R}^{-1} \mathbf{y}}$$

où $\mathbf{y}^t = (y_1, \dots, y_n)$ et \mathbf{R} est la matrice de corrélation de terme générique $\rho(\mathbf{x}_i - \mathbf{x}_j)$ en ligne i et colonne j .

Dans ce modèle, la loi de probabilité conditionnelle d'une valeur est encore une gaussienne, de moyenne égale au krigeage simple de cette valeur par les données conditionnantes et de variance égale à la variance de krigeage associée (Journel and Huijbregts, 1978 ; Rivoirard, 1994). Cette extrême simplicité des lois conditionnelles explique la popularité du modèle multigaussien. En contrepartie, les réalisations de processus multigaussiens exhibent des caractéristiques spécifiques, notamment une déstructuration des valeurs extrêmes et un désordre dans l'agencement spatial des valeurs (*cf.* chapitre introductif).

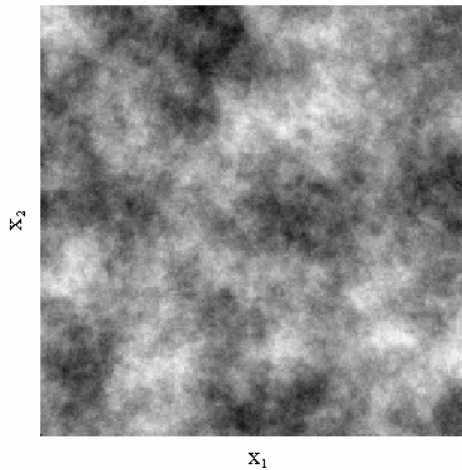


Figure 1.1. Réalisation d'un processus multigaussien dans le plan, de corrélogramme exponentiel (représentation en niveaux de gris)

II.1. Une donnée conditionnante

Examinons tout d'abord ce qui se passe en présence d'une seule donnée conditionnante, notée Y_1 . On cherche à déterminer la probabilité qu'une valeur Y_0 soit inférieure à un seuil y . On désigne par ρ la corrélation entre Y_0 et Y_1 et par r celle entre les indicatrices $I(Y_0; y)$ et $I(Y_1; y)$. Ces corrélations sont reliées par la formule (Matheron, 1975b) :

$$r = \frac{g(y)^2}{G(y)[1 - G(y)]} \sum_{p \geq 1} \frac{H_{p-1}(y)^2}{p} \rho^p \quad (1.1)$$

où g et G sont les densités de probabilité standard et cumulée de la loi normale réduite $\{H_p, p \in \mathbb{N}\}$ sont les polynômes d'Hermite normés (Rivoirard, 1994).

Les différents estimateurs s'écrivent :

- krigeage d'indicateur : $I(Y_0 ; y)^{KS} = rI(y_1 ; y) + (1-r)G(y)$
- krigeage disjonctif : $I(Y_0 ; y)^{KD} = G(y) + \sum_{p \geq 1} \frac{1}{\sqrt{p}} H_{p-1}(y) g(y) \rho^p H_p(y_1)$
- espérance conditionnelle : $I(Y_0 ; y)^{EC} = \frac{1}{g(y_1)} \int_{-\infty}^y g(y_0, y_1) dy_0 = G\left(\frac{y - \rho y_1}{\sqrt{1 - \rho^2}}\right)$

En développant la densité conjointe $g(y_0, y_1)$ en polynômes d'Hermite normés, on s'aperçoit que les expressions de l'espérance conditionnelle et du krigeage disjonctif sont identiques. Cette identité n'est pas surprenante : les calculs ne font intervenir que les lois marginales et bivariable du couple $\{Y_0, Y_1\}$, qui sont entièrement modélisées et prises en compte dans le krigeage disjonctif et l'espérance conditionnelle. En revanche, le krigeage d'indicateur fournit une estimation différente, à cause de la codification de la donnée en indicatrice (variable binaire), ce qui implique une perte d'information.

La figure 1.2 visualise l'écart absolu entre krigeage d'indicateur et espérance conditionnelle, pour trois valeurs du coefficient de corrélation ρ . On observe que l'estimation par krigeage d'indicateur est éloignée de l'espérance conditionnelle lorsque la donnée est très proche du seuil de coupure ou lorsque la donnée est extrême et le seuil médian. Une explication intuitive est que la codification en indicatrice perd l'information de la distance entre la donnée et le seuil de coupure et suppose implicitement que la donnée est "moyennement" éloignée de ce seuil.

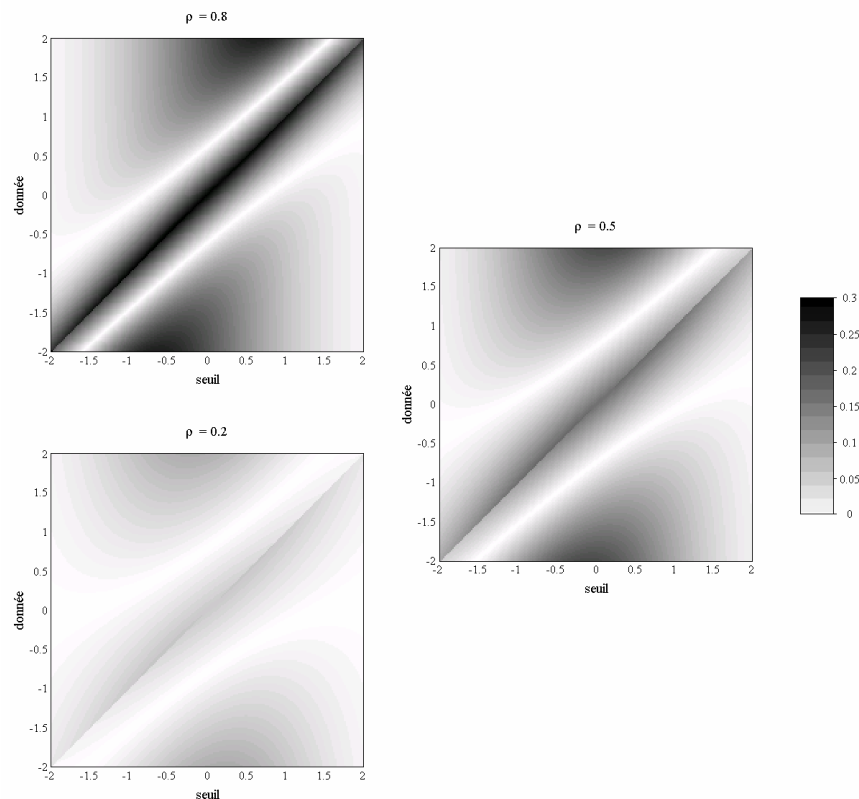


Figure 1.2. Ecart absolu entre krigeage d'indicateur et espérance conditionnelle en fonction de la valeur du seuil de coupure et de la donnée conditionnante

L'amplitude de l'erreur commise est d'autant plus élevée que la donnée est corrélée avec le site à estimer. Par exemple, au point ($y = 0, y_1 = 0$), l'écart vaut $r / 2$, soit 0.295 si $\rho = 0.8$. Cette remarque montre que, si l'effet de pépite est peu important, l'estimation des lois conditionnelles par krigeage d'indicatrices est fortement erronée lorsque le site que l'on cherche à estimer est très proche d'une donnée. En conséquence, on peut mettre en doute la possibilité de réaliser un changement de support à partir de la simulation de valeurs ponctuelles par la méthode séquentielle d'indicatrices sur une grille très dense, comme le recommandent certains auteurs (Deutsch and Journel, 1998 ; Glacken, 1997 ; Journel, 1993).

II.2. Deux données conditionnantes

Dans ce qui suit, la lettre grecque ρ désigne la corrélation entre valeurs gaussiennes, tandis que la lettre latine r indique la corrélation entre indicatrices. Les données sont indexées par 1 et 2, la valeur à simuler par 0. On obtient :

- krigeage d'indicatrice

$$I(Y_0 ; y)^{KS} = \frac{1}{1 - r_{12}^2} [(r_{10} - r_{12} r_{20}) I(y_1 ; y) + (r_{20} - r_{12} r_{10}) I(y_2 ; y)] + (1 - \frac{r_{10} + r_{20}}{1 + r_{12}}) G(y)$$

$$\text{avec } \forall i, j \in \{0,1,2\}, r_{ij} = \frac{g(y)^2}{G(y)[1 - G(y)]} \sum_{p \geq 1} \frac{H_{p-1}(y)^2}{p} \rho_{ij}^p \quad [\text{Eq. (1.1)}]$$

- krigeage disjonctif

$$I(Y_0 ; y)^{KD} = G(y) + \sum_{p \geq 1} \frac{1}{\sqrt{p}} H_{p-1}(y) g(y) [H_p(Y_0)]^{KS}$$

$$\text{où } [H_p(Y_0)]^{KS} = \frac{1}{1 - \rho_{12}^{2p}} [(\rho_{10}^p - \rho_{12}^p \rho_{20}^p) H_p(y_1) + (\rho_{20}^p - \rho_{12}^p \rho_{10}^p) H_p(y_2)]$$

- espérance conditionnelle

$$I(Y_0 ; y)^{EC} = G\left(\frac{y - y_0^{KS}}{\sigma_0^{KS}}\right) = G\left(\frac{(1 - \rho_{12}^2) y - (\rho_{10} - \rho_{12} \rho_{20}) y_1 - (\rho_{20} - \rho_{12} \rho_{10}) y_2}{\sqrt{1 - \rho_{12}^2} \sqrt{1 - \rho_{12}^2 - \rho_{10}^2 - \rho_{20}^2 + 2 \rho_{12} \rho_{20} \rho_{10}}}\right)$$

Si la corrélation ρ_{12} entre données tend vers 1, on obtient les convergences suivantes :

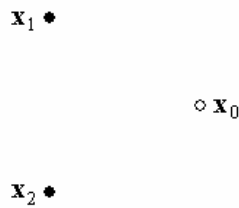
$$y_1 \rightarrow y_2, I(Y_1 ; y) \rightarrow I(Y_2 ; y), \rho_{10} \rightarrow \rho_{20}, r_{12} \rightarrow 1, r_{10} \rightarrow r_{20}.$$

Si $(\rho_{20} - \rho_{10}) / (1 - \rho_{12})$ reste borné¹, on retrouve la situation précédente à une donnée égale à une moyenne pondérée de Y_1 et Y_2 (les termes en $1 - \rho_{12}$ ou $1 - r_{12}$ se factorisent et se simplifient dans l'expression des estimateurs) : les deux données Y_1 et Y_2 deviennent totalement redondantes et ne comptent plus que pour une donnée unique.

¹ La démonstration est relativement aisée si l'on suppose que $(\rho_{20} - \rho_{10}) / (1 - \rho_{12})$ tend vers une constante finie ; en réalité, le résultat reste valide si ce quotient reste borné sans présenter nécessairement de limite (il suffit de manipuler des inégalités, au lieu d'équivalents). En revanche, lorsque le quotient tend vers l'infini, ce qui se produit avec des covariances très régulières à l'origine, la convergence peut être mise en défaut (considérer par exemple un site au milieu des données avec une covariance gaussienne), bien qu'elle soit toujours vraie en espérance.

L'expression des estimateurs dépend de six paramètres : les corrélations entre valeurs gaussiennes $\rho_{12}, \rho_{10}, \rho_{20}$, les valeurs conditionnantes y_1 et y_2 et le seuil y . Pour faciliter leur comparaison, nous allons étudier quelques configurations élémentaires des sites $\{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2\}$ associés à $\{Y_0, Y_1, Y_2\}$: répartition aux sommets d'un triangle équilatéral ou sur une même droite. L'estimation porte sur deux indicatrices cumulées, celles associées aux seuils $y = 0$ (seuil médian) et $y = 1$ (seuil correspondant au quantile de probabilité cumulée 84%). On suppose que la fonction aléatoire étudiée est isotrope, de covariance exponentielle.

A) Triangle équilatéral



Les résultats sont présentés ci-dessous, sur trois graphiques correspondant à des niveaux de corrélation entre les valeurs de plus en plus faibles. A cause de l'hypothèse d'isotropie, les corrélations entre deux valeurs sont égales : $\rho_{12} = \rho_{10} = \rho_{20}$. Le krigeage d'indicatrice est constant dans chacun des quadrants d'origine (y,y) , tandis que l'espérance conditionnelle prend des valeurs égales sur les droites d'équation $y_1 + y_2 = \text{constante}$; ceci explique les motifs qui se dessinent sur les graphiques.

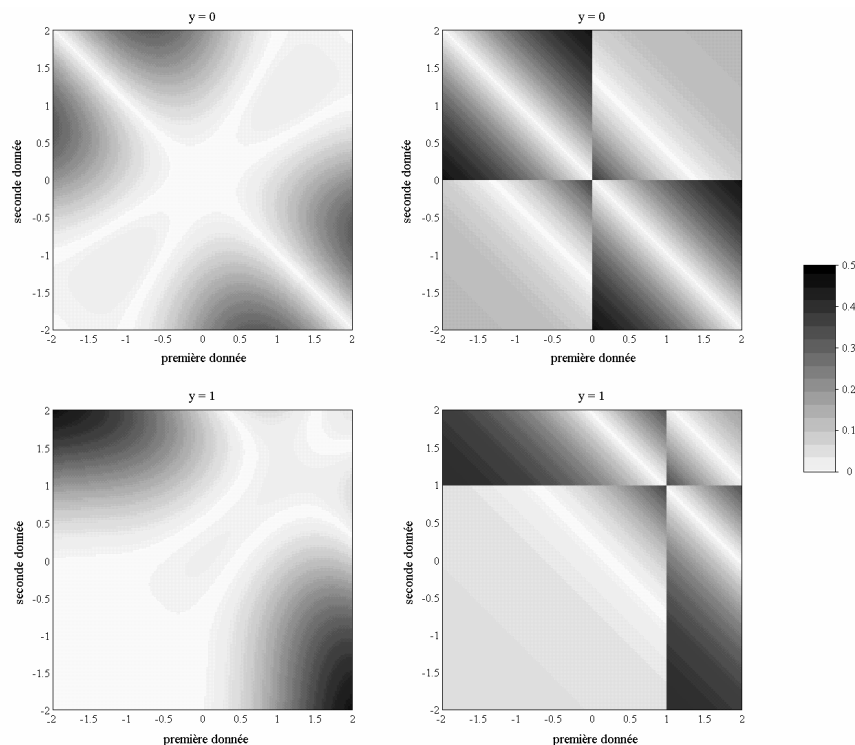


Figure 1.3. *Ecart absolu entre espérance conditionnelle et krigeage disjonctif (gauche) et entre espérance conditionnelle et krigeage d'indicatrice (droite), pour $\mathbf{r}_{12} = \mathbf{r}_{10} = \mathbf{r}_{20} = 0.8$ et deux seuils : $y = 0$ (haut) et $y = 1$ (bas)*

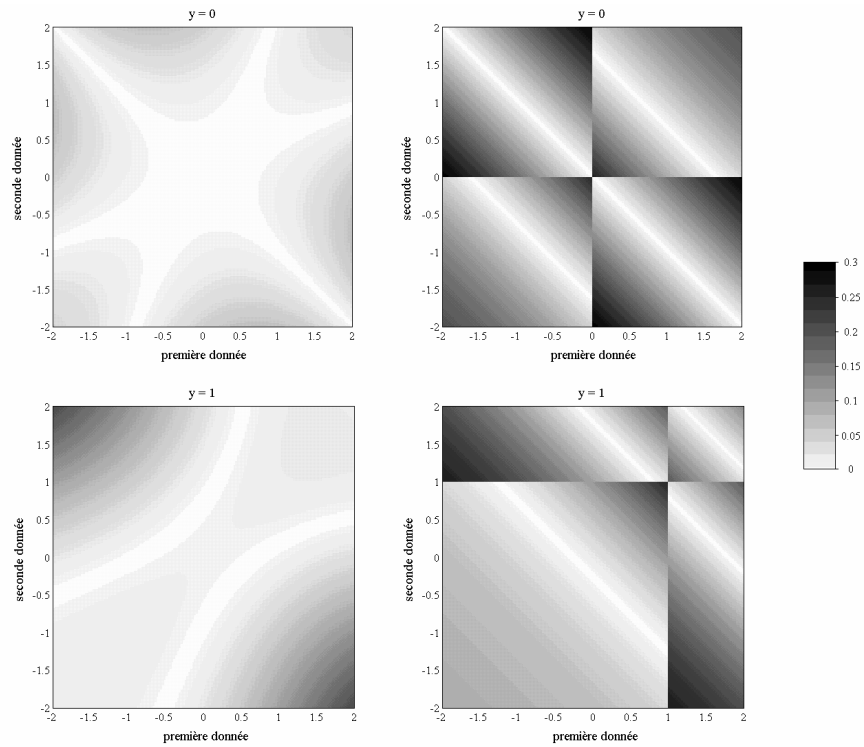


Figure 1.4. *Ecart absolu entre espérance conditionnelle et krigeage disjonctif (gauche) et entre espérance conditionnelle et krigeage d'indicateur (droite), pour $\mathbf{r}_{12} = \mathbf{r}_{10} = \mathbf{r}_{20} = 0.5$ et deux seuils : $y = 0$ (haut) et $y = 1$ (bas)*

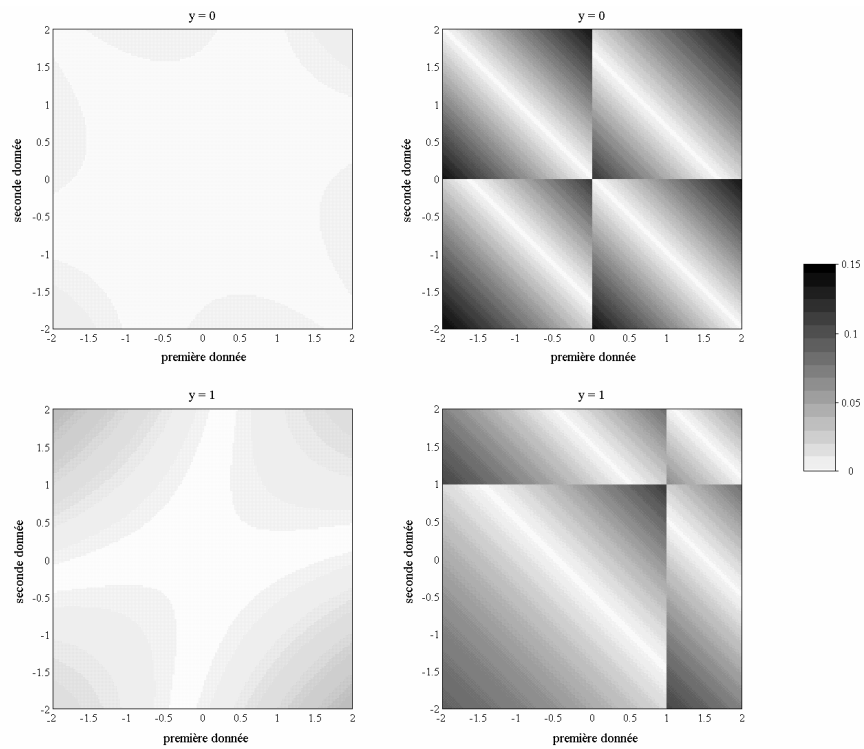


Figure 1.5. *Ecart absolu entre espérance conditionnelle et krigeage disjonctif (gauche) et entre espérance conditionnelle et krigeage d'indicateur (droite), pour $\mathbf{r}_{12} = \mathbf{r}_{10} = \mathbf{r}_{20} = 0.2$ et deux seuils : $y = 0$ (haut) et $y = 1$ (bas)*

A titre indicatif, les résultats sont résumés par le calcul approché des quantités suivantes (tableau 1.1), où l'astérisque représente l'un des deux estimateurs testés (krigeage disjonctif ou d'indicatrice) :

- écart absolu moyen : $\int |I(Y_0; y)^* - I(Y_0; y)^{EC}| g(y_1, y_2) dy_1 dy_2$
- racine de l'écart quadratique moyen : $\sqrt{\int [I(Y_0; y)^* - I(Y_0; y)^{EC}]^2 g(y_1, y_2) dy_1 dy_2}$

En moyenne, l'erreur absolue commise par le krigeage d'indicatrice est très supérieure à celle du krigeage disjonctif. La différence relative entre les performances des deux estimateurs s'accroît lorsque les corrélations diminuent : le krigeage disjonctif fournit une erreur absolue moyenne trois à quatre fois plus faible lorsque le niveau de corrélation est élevé, et jusqu'à quarante fois plus faible lorsque le niveau de corrélation est bas.

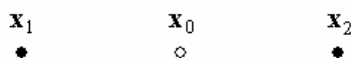
Par ailleurs, les erreurs dues au krigeage d'indicatrice sont élevées non seulement pour des valeurs extrêmes des données, mais aussi lorsque les données sont proches du seuil à estimer (figures 1.3 à 1.5 selon le quadrant dans lequel on se situe), une circonstance qui risque de se présenter fréquemment en pratique. Au contraire, les erreurs les plus élevées du krigeage disjonctif ont lieu lorsque les données conditionnantes prennent des valeurs extrêmes, ce qui se produit plus rarement.

		krigeage disjonctif		krigeage d'indicatrice	
		y = 0	y = 1	y = 0	y = 1
$\rho_{12} = 0.8$	écart absolu moyen	0.026	0.019	0.108	0.070
	racine de l'écart quadratique moyen	0.044	0.036	0.134	0.097
$\rho_{12} = 0.5$	écart absolu moyen	0.008	0.014	0.085	0.061
	racine de l'écart quadratique moyen	0.013	0.021	0.103	0.076
$\rho_{12} = 0.2$	écart absolu moyen	0.001	0.004	0.040	0.031
	racine de l'écart quadratique moyen	0.001	0.005	0.049	0.038

Tableau 1.1. Statistiques approchées des écarts par rapport à l'espérance conditionnelle (sites en triangle équilatéral)

B) Sites alignés : x_0 milieu de $[x_1, x_2]$

Une deuxième configuration à laquelle nous allons porter intérêt est celle où le site à estimer se trouve au milieu des deux données.



Les performances des deux estimateurs sont comparables à celles obtenues dans le cas du triangle équilatéral. La supériorité relative du krigeage disjonctif sur le krigeage d'indicatrice s'accroît quand les corrélations sont faibles (cf. figures 1.6 à 1.8 et tableau 1.2).

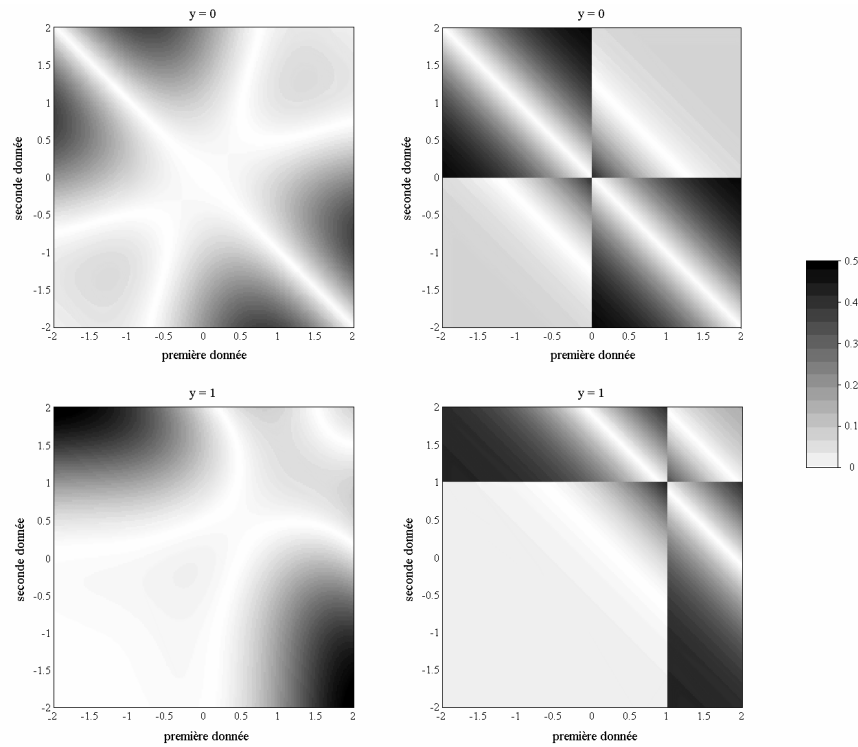


Figure 1.6. *Ecart absolu entre espérance conditionnelle et krigeage disjonctif (gauche) et entre espérance conditionnelle et krigeage d'indicateur (droite), pour $\mathbf{r}_{12} = 0.64$, $\mathbf{r}_{10} = \mathbf{r}_{20} = 0.8$ et deux seuils : $y = 0$ (haut) et $y = 1$ (bas)*

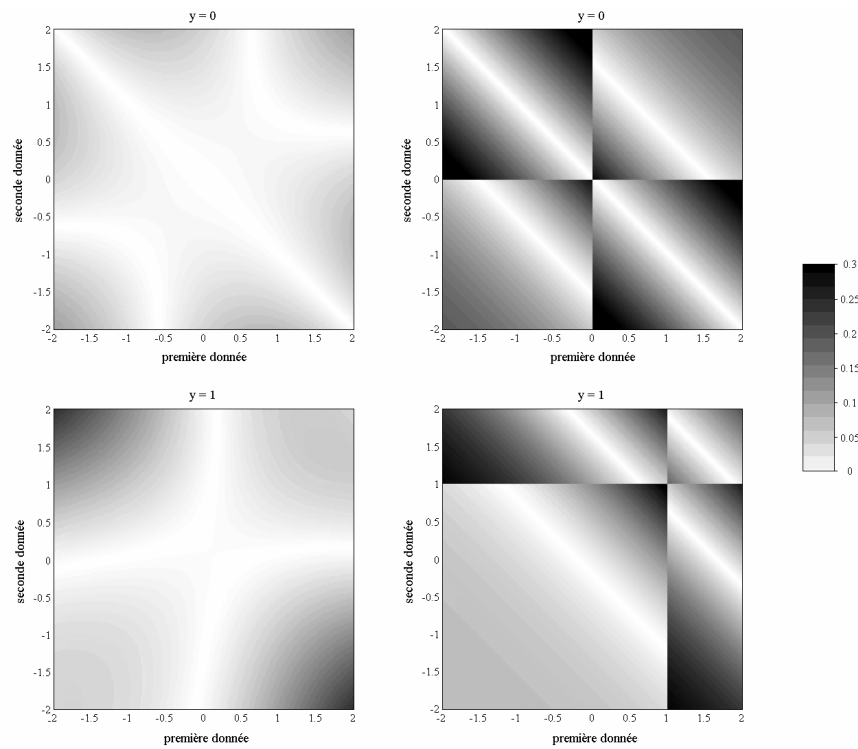


Figure 1.7. *Ecart absolu entre espérance conditionnelle et krigeage disjonctif (gauche) et entre espérance conditionnelle et krigeage d'indicateur (droite), pour $\mathbf{r}_{12} = 0.25$, $\mathbf{r}_{10} = \mathbf{r}_{20} = 0.5$ et deux seuils : $y = 0$ (haut) et $y = 1$ (bas)*

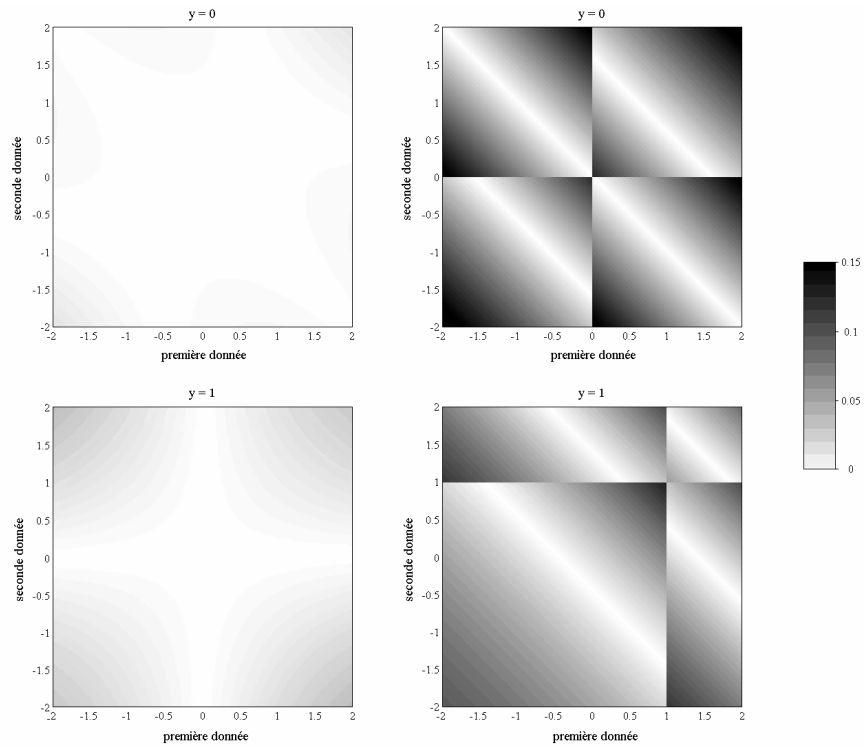


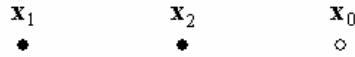
Figure 1.8. *Ecart absolu entre espérance conditionnelle et krigeage disjonctif (gauche) et entre espérance conditionnelle et krigeage d'indicateur (droite), pour $\mathbf{r}_{12} = 0.04$, $\mathbf{r}_{10} = \mathbf{r}_{20} = 0.2$ et deux seuils : $y = 0$ (haut) et $y = 1$ (bas)*

		krigeage disjonctif		krigeage d'indicateur	
		y = 0	y = 1	y = 0	y = 1
$\rho_{10} = 0.8$	écart absolu moyen	0.051	0.037	0.117	0.074
	racine de l'écart quadratique moyen	0.073	0.061	0.156	0.115
$\rho_{10} = 0.5$	écart absolu moyen	0.017	0.025	0.099	0.070
	racine de l'écart quadratique moyen	0.023	0.035	0.122	0.091
$\rho_{10} = 0.2$	écart absolu moyen	0.001	0.005	0.045	0.035
	racine de l'écart quadratique moyen	0.002	0.007	0.056	0.043

Tableau 1.2. *Statistiques approchées des écarts par rapport à l'espérance conditionnelle (sites alignés)*

C) Sites alignés : x_2 milieu de $[x_0, x_1]$

La troisième et dernière situation qui sera analysée est celle où le site à estimer est aligné avec les données selon la configuration suivante :



Les résultats présentés ci-après montrent que le krigeage disjonctif ne souffre aucune erreur. En effet, la propriété d'écran de la covariance exponentielle (Matheron, 1970 ; Rivoirard, 1984), implique que l'espérance conditionnelle ne fait pas intervenir la valeur de Y_1 : il y a indépendance entre Y_0 et Y_1 conditionnellement à Y_2 . Quant au krigeage disjonctif, les facteurs des lois bivariées (polynômes d'Hermite normés) possèdent eux-aussi une covariance à décroissance exponentielle, de sorte que la valeur Y_1 est totalement écrantée dans l'estimation. Dans ces conditions, il n'est pas étonnant que le krigeage disjonctif coïncide avec l'espérance conditionnelle puisqu'en définitive seule entre en compte la loi bivariable de $\{Y_0, Y_2\}$.

Toutefois cette circonstance ne se produit pas pour le krigeage d'indicateur, car le corrélogramme de l'indicateur n'est plus une exponentielle décroissante, ce qui lève l'effet d'écran. L'estimation dépend donc de la première donnée, ou plus exactement de la position de sa valeur par rapport au seuil de coupure.

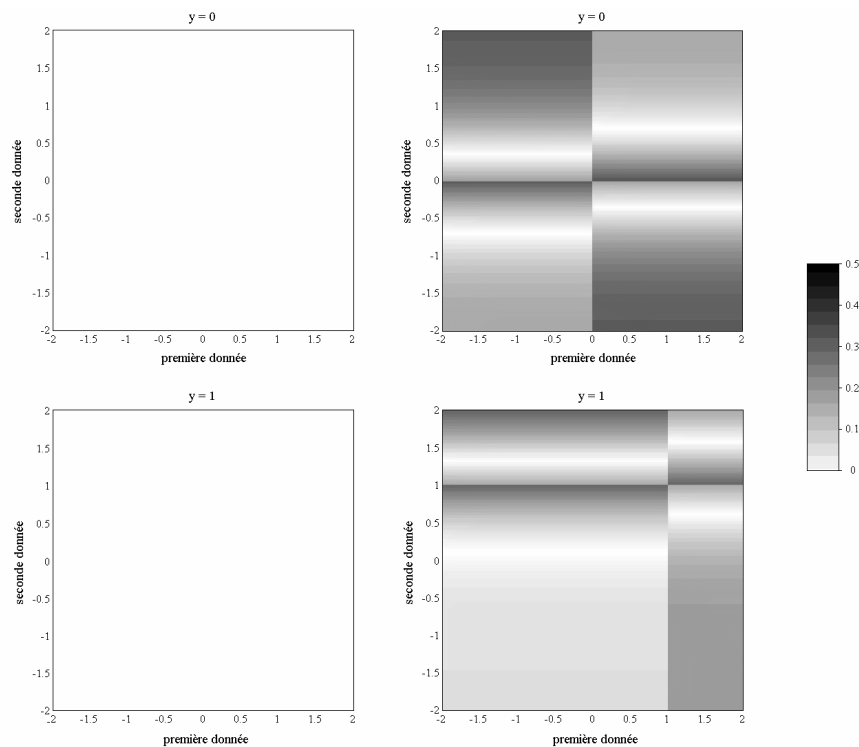


Figure 1.9. *Ecart absolu entre espérance conditionnelle et krigeage disjonctif (gauche) et entre espérance conditionnelle et krigeage d'indicateur (droite), pour $r_{10} = 0.64$, $r_{12} = r_{20} = 0.8$ et deux seuils : $y = 0$ (haut) et $y = 1$ (bas)*

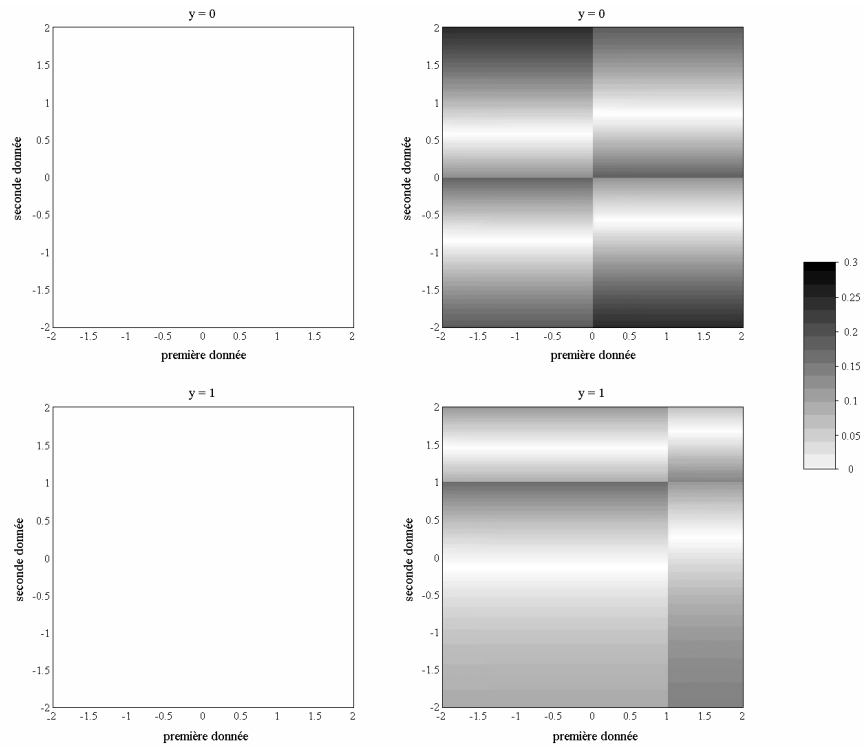


Figure 1.10. *Ecart absolu entre espérance conditionnelle et krigeage disjonctif (gauche) et entre espérance conditionnelle et krigeage d'indicateur (droite), pour $\mathbf{r}_{10} = 0.25$, $\mathbf{r}_{12} = \mathbf{r}_{20} = 0.5$ et deux seuils : $y = 0$ (haut) et $y = 1$ (bas)*

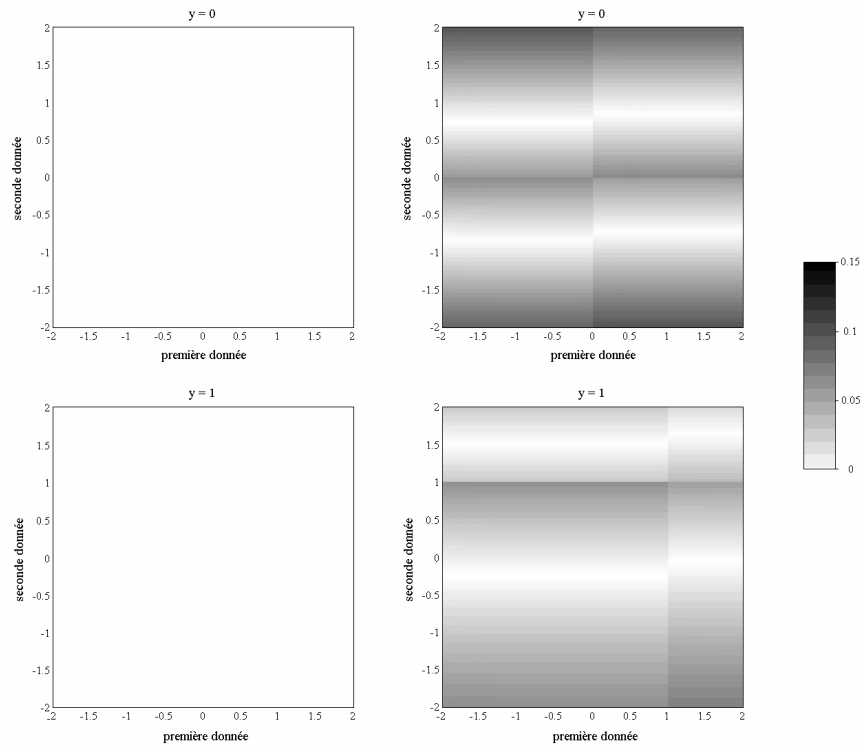


Figure 1.11. *Ecart absolu entre espérance conditionnelle et krigeage disjonctif (gauche) et entre espérance conditionnelle et krigeage d'indicateur (droite), pour $\mathbf{r}_{10} = 0.04$, $\mathbf{r}_{12} = \mathbf{r}_{20} = 0.2$ et deux seuils : $y = 0$ (haut) et $y = 1$ (bas)*

		krigeage disjonctif		krigeage d'indicatrice	
		y = 0	y = 1	y = 0	y = 1
$\rho_{12} = 0.8$	écart absolu moyen	0.000	0.000	0.123	0.079
	racine de l'écart quadratique moyen	0.000	0.000	0.144	0.101
$\rho_{12} = 0.5$	écart absolu moyen	0.000	0.000	0.085	0.059
	racine de l'écart quadratique moyen	0.000	0.000	0.100	0.070
$\rho_{12} = 0.2$	écart absolu moyen	0.000	0.000	0.034	0.026
	racine de l'écart quadratique moyen	0.000	0.000	0.041	0.031

Tableau 1.3. Statistiques approchées des écarts par rapport à l'espérance conditionnelle (sites alignés)

III. Modèle mosaïque

Le deuxième modèle que nous allons examiner est celui d'une fonction aléatoire de loi mosaïque (Matheron, 1982 ; Rivoirard, 1994) : l'espace est partitionné en cellules aléatoires de loi stationnaire ; dans chacune de ces cellules, la fonction aléatoire prend une valeur constante indépendante des valeurs des autres cellules.

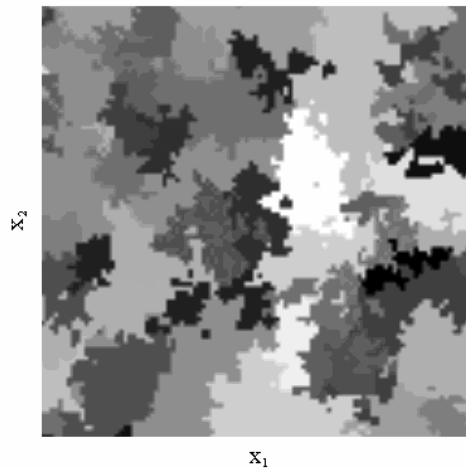


Figure 1.12. Réalisation d'un processus mosaïque dans le plan

Une particularité de ce modèle est que la covariance de toute fonction de $Y(\mathbf{x})$ est proportionnelle à un unique corrélogramme $\rho(\mathbf{h})$, qui mesure la probabilité que deux sites séparés de \mathbf{h} soient dans la même cellule. Par suite, les indicatrices sont autokrigeables et le krigeage d'indicatrices coïncide avec le krigeage disjonctif, puisque ce dernier équivaut à un *co*krigeage complet de toutes les indicatrices (Rivoirard, 1994). Le krigeage d'indicatrices apparaît comme la technique d'estimation implicitement associée au modèle mosaïque.

III.1. Une donnée conditionnante

Soient ρ le coefficient de corrélation entre la donnée Y_1 et la valeur à estimer Y_0 et $F(\cdot)$ la fonction de répartition de la loi marginale de $Y(\mathbf{x})$. Les différents estimateurs (krigeage d'indicatrice, krigeage disjonctif et espérance conditionnelle) sont identiques :

$$I(Y_0 ; y)^{KS} = I(Y_0 ; y)^{KD} = I(Y_0 ; y)^{EC} = \rho I(y_1 ; y) + (1 - \rho) F(y).$$

La connaissance du seul paramètre ρ équivaut à spécifier la loi bivariable entière, ce qui suffit à calculer l'estimateur optimal.

III.2. Deux données conditionnantes

Soient \mathbf{x}_1 et \mathbf{x}_2 les localisations des données Y_1 et Y_2 et \mathbf{x}_0 celle de la valeur Y_0 à estimer. Les estimateurs du krigeage d'indicatrice et krigeage disjonctif sont identiques :

$$\begin{aligned} I(Y_0 ; y)^{KS} &= I(Y_0 ; y)^{KD} \\ &= \frac{1}{1 - \rho_{12}^2} [(\rho_{10} - \rho_{12} \rho_{20}) I(y_1 ; y) + (\rho_{20} - \rho_{12} \rho_{10}) I(y_2 ; y)] + (1 - \frac{\rho_{10} + \rho_{20}}{1 + \rho_{12}}) F(y) \end{aligned} \quad (1.2)$$

L'expression de l'espérance conditionnelle fait intervenir la loi trivariable entre les deux valeurs conditionnantes et la valeur à estimer. Pour expliciter les paramètres de cette loi, nous adopterons la convention suivante, où la virgule indique un changement de cellule :

$p_{1,2,0}$: probabilité que \mathbf{x}_1 et \mathbf{x}_2 soient dans la même cellule et \mathbf{x}_0 dans une cellule différente
 $p_{1,2,0}$: probabilité que $\mathbf{x}_1, \mathbf{x}_2$ et \mathbf{x}_0 soient dans trois cellules différentes

Pour des raisons de compatibilité mathématique, on doit avoir :

$$\begin{cases} p_{1,2,0} + p_{12,0} + p_{10,2} + p_{20,1} + p_{120} = 1 \\ \rho_{12} = p_{12,0} + p_{120} \\ \rho_{10} = p_{10,2} + p_{120} \\ \rho_{20} = p_{20,1} + p_{120} \end{cases}$$

Par rapport à la loi bivariable, il suffit de spécifier un paramètre supplémentaire, par exemple p_{120} (fonction des localisations de $\mathbf{x}_1, \mathbf{x}_2$ et \mathbf{x}_0), pour connaître la loi trivariable. Pour calculer l'espérance conditionnelle, on distingue les différents cas :

- si \mathbf{x}_1 et \mathbf{x}_2 sont dans la même cellule, ce qui se produit avec la probabilité ρ_{12}

$$I(Y_0 ; y)^{EC} = \frac{1}{\rho_{12}} [p_{120} I(y_1 ; y) + p_{12,0} F(y)] = \frac{1}{\rho_{12}} [p_{120} I(y_2 ; y) + p_{12,0} F(y)] \quad (1.3)$$

- si \mathbf{x}_1 et \mathbf{x}_2 sont dans des cellules distinctes, ce qui se produit avec la probabilité $1 - \rho_{12}$

$$I(Y_0 ; y)^{EC} = \frac{1}{1 - \rho_{12}} [p_{10,2} I(y_1 ; y) + p_{20,1} I(y_2 ; y) + p_{1,2,0} F(y)] \quad (1.4)$$

Considérant en outre que $Y(\mathbf{x})$ est une variable à loi marginale absolument continue (gaussienne, par exemple), l'appartenance de deux sites à une même cellule est synonyme d'égalité des valeurs associées : la probabilité que deux cellules distinctes aient la même valeur est nulle.

III.2.1. Mosaïque de Poisson

Dans cette section, nous allons considérer la *mosaïque de Poisson* (Miles, 1961, 1974 ; Matheron, 1975a) évaluée selon une distribution gaussienne réduite (moyenne 0, variance 1) dans l'espace à deux dimensions. Le corrélogramme, qui représente la probabilité que deux sites soient dans la même cellule de la mosaïque en fonction de la séparation entre ces sites, est une fonction exponentielle :

$$\rho(\mathbf{h}) = \exp(-2\lambda |\mathbf{h}|)$$

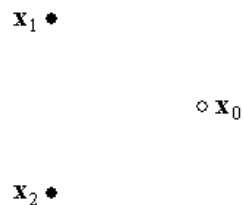
où λ est la densité du réseau de droites poissonniennes définissant la mosaïque.

Plus généralement, la probabilité qu'un domaine borné convexe X soit inclus dans une cellule est liée à son périmètre $P(X)$ par la formule :

$$\text{Prob}(X \subseteq 1 \text{ cellule}) = \exp(-\lambda P(X)).$$

Pour faciliter la comparaison avec le cas multigaussien, les résultats sont présentés selon la même échelle de gris. Les valeurs inscrites sur les diagonales des graphiques correspondent au cas où les deux données sont égales (qui se produit avec une probabilité non nulle).

A) Triangle équilatéral



La probabilité que les trois sites se trouvent dans la même cellule est $p_{120} = \exp(-3\lambda a)$, où a est la longueur du côté du triangle équilatéral. Cette relation permet de déterminer toutes les probabilités trivariées :

$$\begin{cases} p_{120} = \exp(-3\lambda a) = (\rho_{12})^{3/2} \\ p_{12,0} = \exp(-2\lambda a) - \exp(-3\lambda a) \\ p_{10,2} = \exp(-2\lambda a) - \exp(-3\lambda a) \\ p_{20,1} = \exp(-2\lambda a) - \exp(-3\lambda a) \\ p_{1,2,0} = 1 - 3\exp(-2\lambda a) + 2\exp(-3\lambda a) \end{cases}$$

Les résultats (figure 1.13) seront commentés ultérieurement.

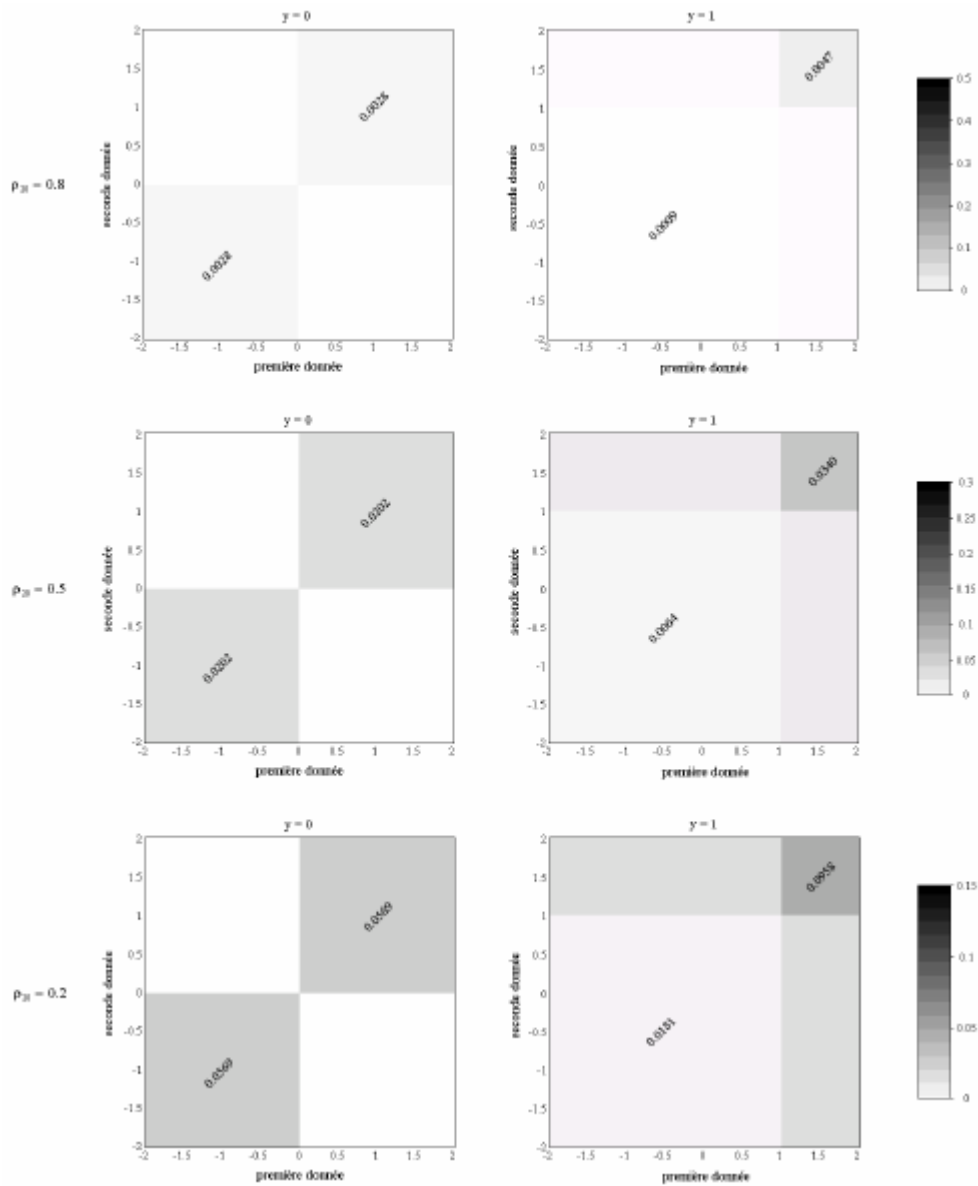
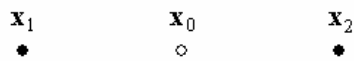


Figure 1.13. *Ecart absolu entre krigeage d'indatrice et espérance conditionnelle*

B) Sites alignés : x_0 milieu de $[x_1, x_2]$



Les deux données jouent des rôles symétriques. La loi trivariante est déterminée par la donnée de $p_{12,0} = 0$ (les cellules de la mosaïque poissonnienne sont convexes) :

$$\begin{cases} p_{1,2,0} = 1 + \rho_{12} - \rho_{10} - \rho_{20} = (1 - \rho_{20})^2 \\ p_{120} = \rho_{12} = (\rho_{20})^2 \\ p_{10,2} = \rho_{10} - \rho_{12} = \rho_{20}(1 - \rho_{20}) \\ p_{20,1} = \rho_{20} - \rho_{12} = \rho_{20}(1 - \rho_{20}) \end{cases}$$

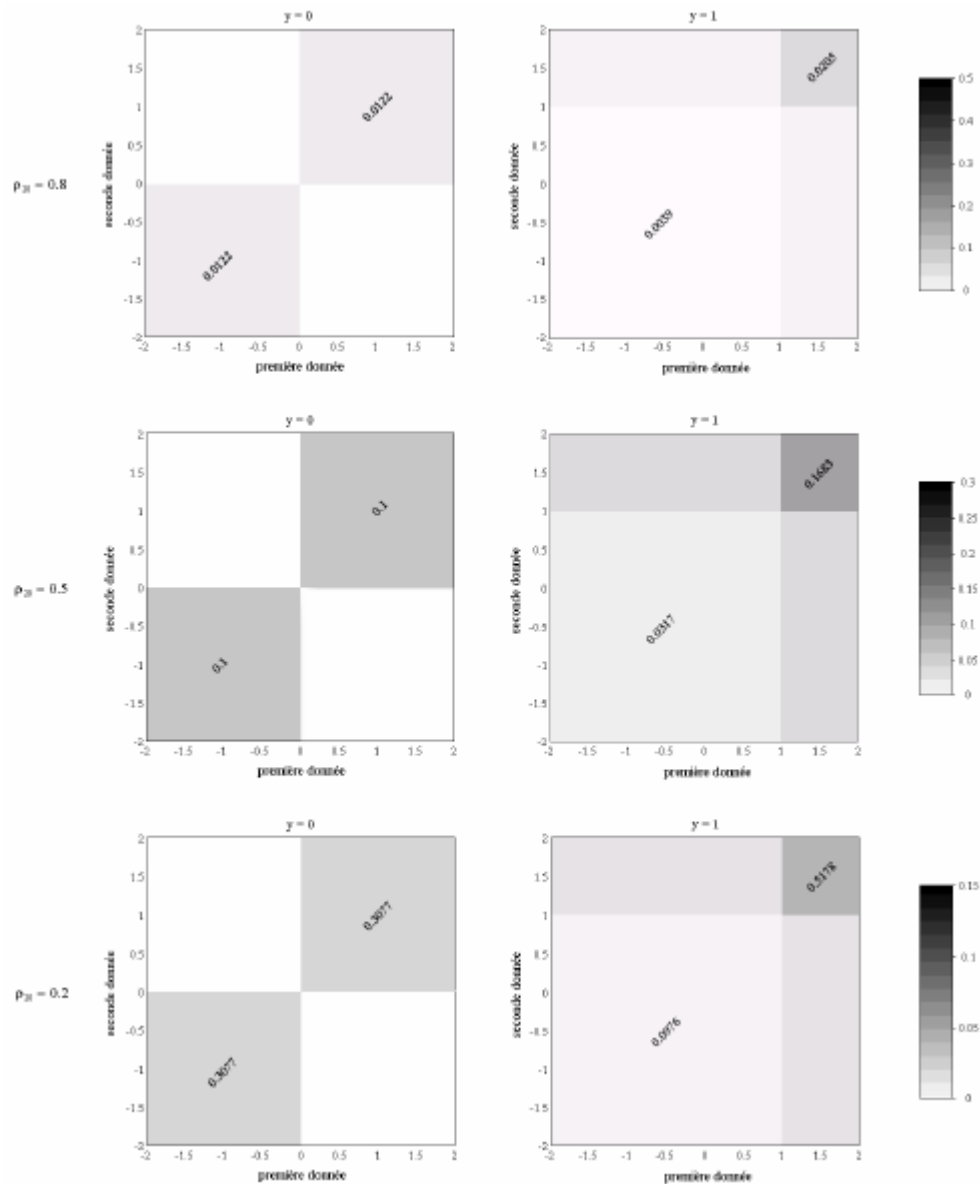


Figure 1.14. *Ecart absolu entre krigeage d'indicatrice et espérance conditionnelle*

Commentaires sur les situations A) et B)

- 1) Seuls deux critères interviennent dans les graphiques précédents :
 - (i) appartenance ou non des données conditionnantes à la même cellule
 - (ii) position des valeurs des données par rapport au seuil à estimer.

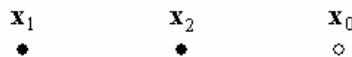
A cause du caractère mosaïque du processus, on a vu que l'espérance conditionnelle ne prend pas en compte les valeurs des données elles-mêmes, mais seulement leur éventuelle égalité (i) et les indicatrices associées à ces données (ii). Pour sa part, par construction, le krigeage d'indicatrice ne considère que la propriété (ii). En particulier, aucun des estimateurs ne dépend de la proximité des valeurs conditionnantes par rapport au seuil considéré.

- 2) Dans le cas B), le krigeage d'indicatrice ne prend pas en compte la convexité des cellules de la mosaïque et s'écarte de l'espérance conditionnelle lorsque les données sont égales ; l'écart est d'autant plus important que les corrélations sont faibles : le krigeage d'indicatrice se rapproche de l'espérance *a priori*, tandis que la propriété de convexité fournit la valeur exacte du site \mathbf{x}_0 . Une situation analogue se produit dans le cas A).
- 3) Lorsque le seuil de coupure est médian ($y = 0$) et que les données sont situées de part et d'autre de ce seuil, l'écart entre le krigeage d'indicatrice et l'espérance conditionnelle est nul : ces deux estimateurs fournissent en effet une estimation égale à 0.5 (cette situation correspond aux deux quadrants de l'anti-diagonale des graphiques associés au seuil nul).
- 4) Bien que la loi marginale soit la même que dans le cadre multigaussien, les écarts entre espérance conditionnelle et krigeage d'indicatrice sont moindres lorsque les corrélations sont élevées (cf. tableau 1.4). Cependant, si les corrélations sont faibles, le krigeage d'indicatrice est à peine meilleur que dans le cas multigaussien et ses résultats sont moins bons que ceux du krigeage disjonctif bigaussien.

		sites en triangle équilatéral		sites alignés	
		y = 0	y = 1	y = 0	y = 1
$\rho_{10} = 0.8$	écart absolu moyen	0.005	0.003	0.016	0.011
	racine de l'écart quadratique moyen	0.007	0.005	0.021	0.015
$\rho_{10} = 0.5$	écart absolu moyen	0.020	0.015	0.050	0.036
	racine de l'écart quadratique moyen	0.025	0.018	0.065	0.047
$\rho_{10} = 0.2$	écart absolu moyen	0.023	0.016	0.025	0.018
	racine de l'écart quadratique moyen	0.031	0.023	0.064	0.047

Tableau 1.4. Statistiques des écarts entre le krigeage d'indicatrice et l'espérance conditionnelle

C) Sites alignés : x_2 milieu de $[x_0, x_1]$



Dans ce cas, la probabilité que \mathbf{x}_1 et \mathbf{x}_0 soient dans la même cellule et \mathbf{x}_2 soit en dehors est nulle (convexité des cellules) : $p_{10,2} = 0$. On en déduit :

$$\begin{cases} p_{120} = \rho_{10} \\ p_{12,0} = \rho_{12} - \rho_{10} \\ p_{20,1} = \rho_{20} - \rho_{10} \\ p_{1,2,0} = 1 + \rho_{10} - \rho_{12} - \rho_{20} \end{cases}$$

En utilisant la relation $\rho_{10} = \rho_{12} \rho_{20}$, on trouve que les trois estimateurs coïncident :

$$I(Y_0; y)^{KS} = I(Y_0; y)^{KD} = I(Y_0; y)^{EC} = \rho_{20} I(y_2; y) + (1 - \rho_{20}) F(y).$$

La covariance exponentielle implique un effet d'écran de \mathbf{x}_2 sur \mathbf{x}_1 . En conséquence, seule la loi bivariable entre \mathbf{x}_0 et \mathbf{x}_2 entre en jeu dans l'espérance conditionnelle, ce qui explique l'identité avec le krigeage disjonctif, donc aussi le krigeage d'indicatrice.

III.2.2. Mosaïque des feuilles mortes

Le second modèle que nous allons étudier est la mosaïque des feuilles mortes (Matheron, 1968a ; Jeulin, 1997). Dans ce modèle, des points poissonniens ou "germes" apparaissent au cours du temps, en chacun desquels est implanté une forme géométrique A de valuation constante (*grain primaire*). Les grains nouveaux sont cachés par les grains plus anciens. Le procédé de construction s'arrête dès que le domaine est entièrement recouvert ; on obtient alors une mosaïque dans laquelle la probabilité que deux sites \mathbf{x} et $\mathbf{x} + \mathbf{h}$ soient dans la même cellule est :

$$\rho(\mathbf{h}) = \frac{K(\mathbf{h})}{2K(\mathbf{0}) - K(\mathbf{h})}$$

où $K(\mathbf{h})$ est le covariogramme géométrique du grain primaire A , c'est-à-dire le volume moyen entre le grain primaire et le même grain translaté du vecteur \mathbf{h} (Matheron, 1965 ; Serra, 1982).

Plus généralement, la probabilité qu'un compact K soit inclus dans une cellule est donnée par la formule (Matheron, 1968a) :

$$\text{Prob}(K \subseteq 1 \text{ cellule}) = \frac{E(|A \ominus \tilde{K}|)}{E(|A \oplus \tilde{K}|)}$$

où \tilde{K} est le transposé du compact K (image de K par une symétrie de centre l'origine)

\ominus et \oplus désignent les opérations d'érosion et de dilatation respectivement

$|D|$ représente la mesure du domaine D .

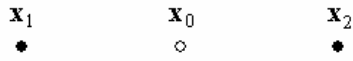
Pour éviter une grande complexité mathématique, nous n'allons considérer que le cas où les trois sites $\{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2\}$ sont alignés et où les grains primaires sont tels que le corrélogramme de la mosaïque soit une exponentielle décroissante, soit $\rho(\mathbf{h}) = \exp(-|\mathbf{h}|)$. D'après Matheron (1986), cette circonstance se produit dans l'espace à une dimension si les feuilles sont des segments aléatoires de covariogramme géométrique :

$$K(\mathbf{h}) = \frac{4 \exp(-|\mathbf{h}|)}{1 + \exp(-|\mathbf{h}|)}.$$

La valuation des grains primaires est supposée gaussienne, ce qui assure que la loi marginale du processus est également gaussienne. Le covariogramme géométrique est lié à la fonction de répartition de la loi en longueur des segments par :

$$\forall \ell \geq 0, F(\ell) = 1 + K'(\ell) = \left[\frac{1 - \exp(-\ell)}{1 + \exp(-\ell)} \right]^2.$$

A) Sites alignés : x_0 milieu de $[x_1, x_2]$



Notons $2L$ la longueur séparant x_1 et x_2 . Les paramètres des lois bivariable et trivariable sont liés à cette longueur par les relations suivantes :

$$\rho_{10} = \rho_{20} = \frac{K(L)}{2K(0) - K(L)}, \quad \rho_{12} = \frac{K(2L)}{2K(0) - K(2L)} = \rho_{10}^2 \quad \text{et} \quad \rho_{120} = \frac{2e^{-2L}(1+e^{-L})}{(1+e^{-2L})(3-e^{-L})}$$

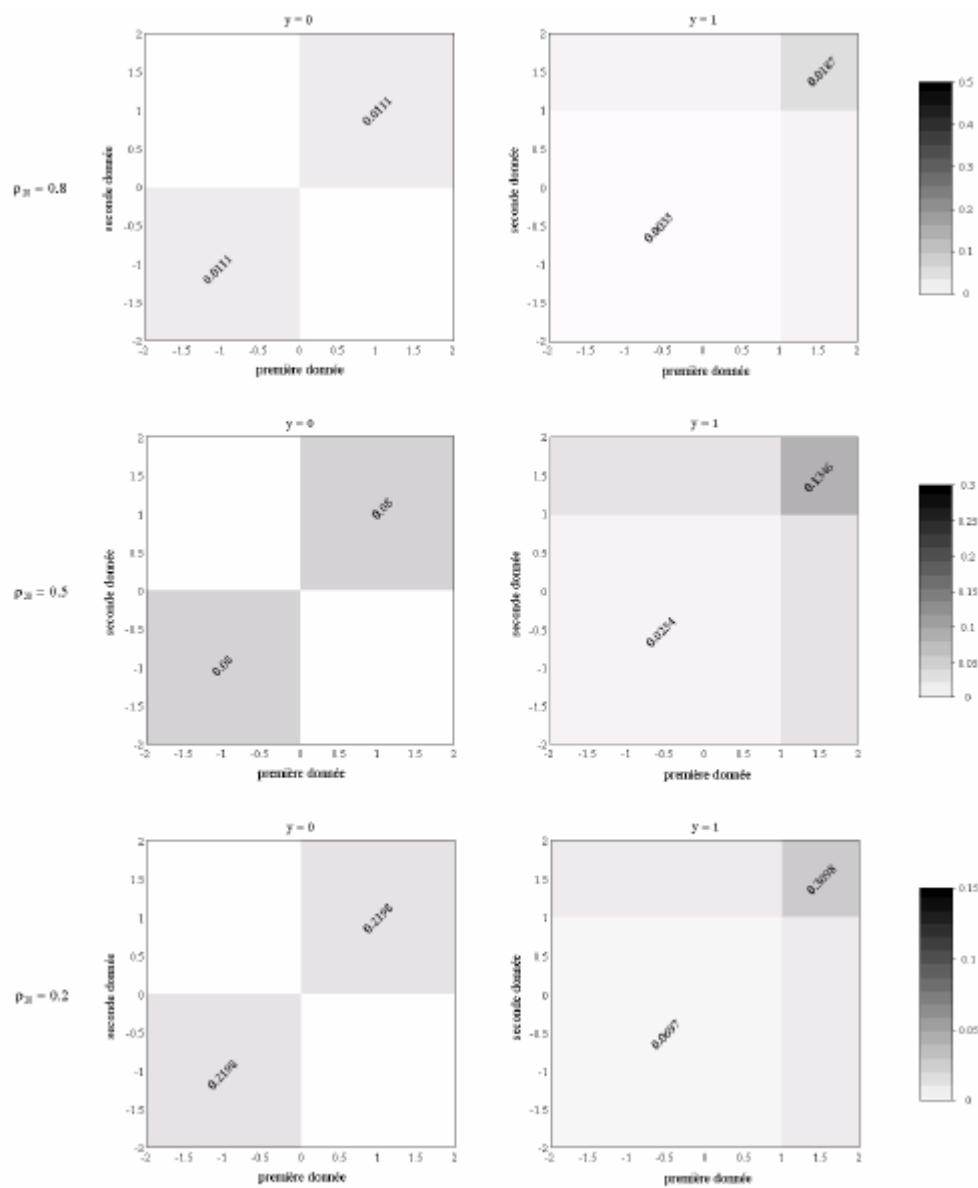


Figure 1.15. Ecart absolu entre krigeage d'indicatrice et espérance conditionnelle

Démonstration de l'expression de la loi trivariable

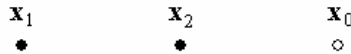
La probabilité que les trois sites soient dans la même cellule de la mosaïque est le rapport entre les deux quantités suivantes :

- $E(A \ominus \tilde{K}) = \int_{2L}^{+\infty} (\ell - 2L) dF(\ell) = - \int_{2L}^{+\infty} [F(\ell) - 1] d\ell = K(2L)$
- $E(A \oplus \tilde{K}) = \int_0^L 3\ell dF(\ell) + \int_L^{+\infty} (2L + \ell) dF(\ell)$
 $= \{3LF(L) - 3 \int_0^L F(\ell) d\ell\} - \{3L[F(L) - 1] + \int_L^{+\infty} [F(\ell) - 1] d\ell\}$
 $= 3LF(L) - 3[L + K(L) - K(0)] - 3L[F(L) - 1] + K(L)$
 $= 3K(0) - 2K(L)$

On trouve finalement l'expression annoncée :

$$P_{120} = \frac{K(2L)}{3K(0) - 2K(L)} = \frac{2e^{-2L}(1 + e^{-L})}{(1 + e^{-2L})(3 - e^{-L})}$$

B) Sites alignés : x_2 milieu de $[x_0, x_1]$



Soit $2L$ la longueur entre x_0 et x_1 .

Par rapport à la situation précédente, la probabilité que les trois sites soient contenus dans la même feuille est inchangée, tandis que les corrélations entre valeurs sont :

$$\rho_{12} = \rho_{20} = \frac{K(L)}{2K(0) - K(L)} \text{ et } \rho_{10} = \frac{K(2L)}{2K(0) - K(2L)} = \rho_{12}^2$$

Contrairement à la mosaïque de Poisson, l'espérance conditionnelle ne bénéficie pas de la propriété markovienne et dépend de la donnée Y_1 . Le krigeage d'indicatrice reste indépendant de cette donnée. Par conséquent, il existera un écart non nul entre les deux estimateurs.

Pour chaque configuration étudiée, l'écart le plus important se produit lorsque la corrélation entre valeurs est intermédiaire ($\rho_{20} = 0.5$). Lorsque x_0 est le milieu de $[x_1, x_2]$, on note une légère amélioration par rapport au cas de la mosaïque de Poisson, bien que les ordres de grandeur des écarts soient similaires (tableau 1.5). Cette amélioration est toutefois compensée par la détérioration qui se produit lorsque x_2 est le milieu de $[x_0, x_1]$, situation pour laquelle l'écart n'est plus nul.

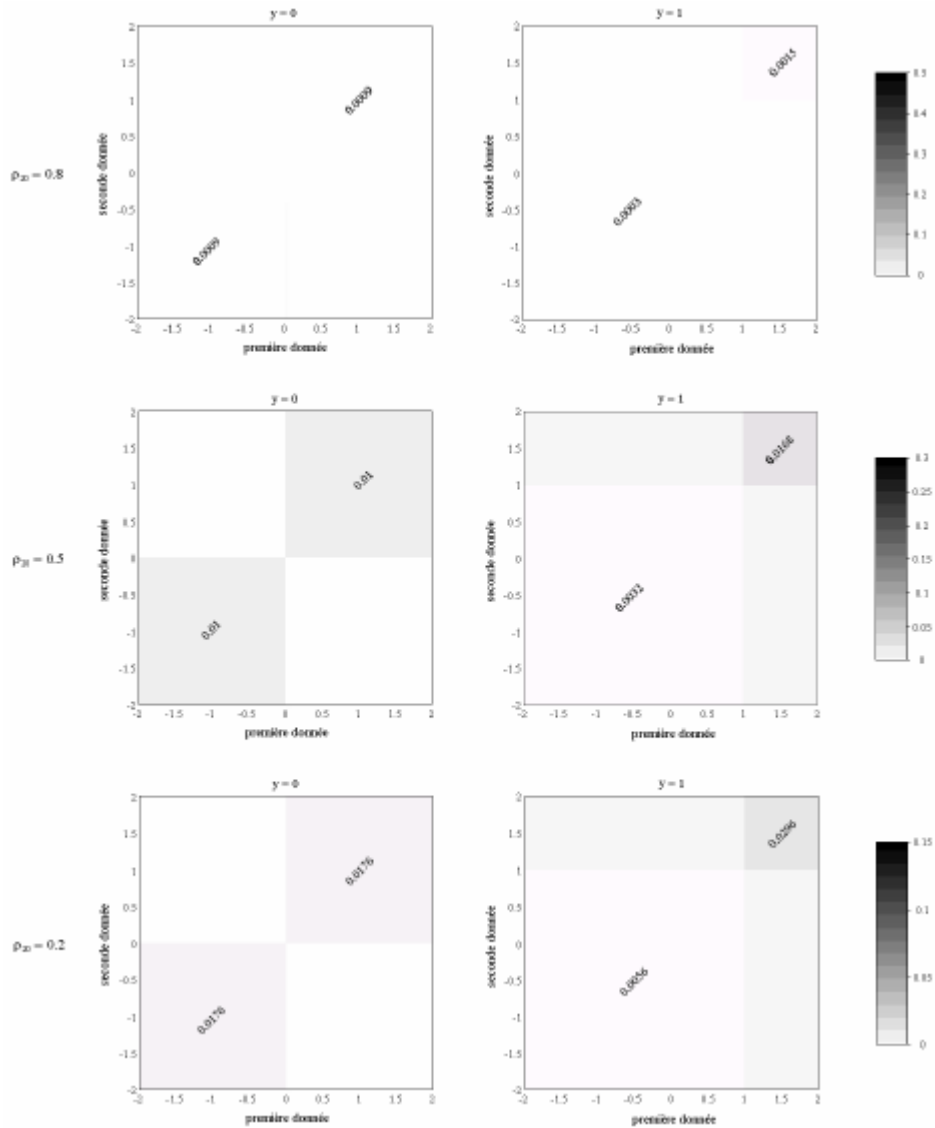


Figure 1.16. Ecart absolu entre krigeage d'indicatrice et espérance conditionnelle

		x_0 milieu de $[x_1, x_2]$		x_2 milieu de $[x_0, x_1]$	
		y = 0	y = 1	y = 0	y = 1
$\rho_{20} = 0.8$	écart absolu moyen	0.014	0.010	0.001	0.001
	racine de l'écart quadratique moyen	0.019	0.014	0.002	0.002
$\rho_{20} = 0.5$	écart absolu moyen	0.040	0.029	0.010	0.007
	racine de l'écart quadratique moyen	0.052	0.038	0.012	0.009
$\rho_{20} = 0.2$	écart absolu moyen	0.018	0.013	0.007	0.005
	racine de l'écart quadratique moyen	0.046	0.033	0.010	0.007

Tableau 1.5. Statistiques des écarts entre le krigeage d'indicatrice et l'espérance conditionnelle

III.2.3. Conclusions

De la comparaison des résultats précédents avec ceux obtenus dans le cadre multigaussien, on tire la conclusion qualitative suivante : le krigeage d'indicatrice gagne en précision lorsque l'on passe du modèle de diffusion au modèle mosaïque (dans ce dernier cas, la covariance d'indicatrice réussit à modéliser la loi bivariable entière) ; au contraire, le krigeage disjonctif gagne en précision lorsque l'on passe du modèle mosaïque au modèle de diffusion.

On est alors enclin à penser que la simulation séquentielle d'indicatrices est un algorithme adapté au modèle mosaïque, tout comme le krigeage d'indicatrices est implicitement associé à ce modèle². Cette réflexion soulève la question suivante : à quelle condition est-il possible d'avoir coïncidence entre le krigeage d'une indicatrice et son espérance conditionnelle ? Une condition nécessaire est que les résultats du krigeage de l'indicatrice soient toujours compris entre 0 et 1, ce qui constitue d'ores et déjà une circonstance rarissime car elle interdit la possibilité d'obtenir des pondérateurs de krigeage négatifs.

Une telle condition n'est réalisable que dans l'espace unidimensionnel et pour certaines fonctions de covariance³ telle la covariance exponentielle (Matheron, 1986). Or, même dans ce cas de figure, il n'est pas possible d'obtenir n'importe quel modèle de mosaïque. En témoigne la mosaïque des feuilles mortes unidimensionnelle à covariance exponentielle : comme on a pu le vérifier avec l'exemple de trois sites alignés, elle ne correspond à aucune séquence de simulation. Une exception notable est fournie par la mosaïque de Poisson, laquelle correspond à une simulation séquentielle d'indicatrices en ordre régulier, indépendamment du nombre de données conditionnantes, à cause de l'effet d'écran exercé par la covariance exponentielle.

Lorsque la covariance de la mosaïque est une exponentielle décroissante et que le site \mathbf{x}_0 à simuler est au milieu de deux sites de données $\{\mathbf{x}_1, \mathbf{x}_2\}$, un calcul basé sur les équations (1.2) à (1.4) montre que le krigeage d'indicatrice coïncide avec l'espérance conditionnelle à ces deux données si et seulement si la loi trivariable est définie par :

$$p_{120} = \frac{2\rho_{12}^{3/2}}{1 + \rho_{12}}.$$

Cette loi diffère de celles associées aux mosaïques de Poisson et des feuilles mortes ; nous en donnerons une interprétation par un modèle de recouvrement au chapitre suivant. Signalons que les compartiments d'une telle mosaïque ne sont pas convexes, puisque l'on a :

$$p_{12,0} = \rho_{12} - \frac{2\rho_{12}^{3/2}}{1 + \rho_{12}} = \rho_{12} \frac{(1 - \rho_{12}^{1/2})^2}{1 + \rho_{12}} > 0 \text{ dès que } 0 < \rho_{12} < 1.$$

² Hors du cadre mosaïque, la simulation séquentielle d'indicatrices n'a aucune chance de fournir une image exacte du modèle de fonction aléatoire considéré, puisque le krigeage d'indicatrices qu'elle met en jeu ne coïncide jamais avec le krigeage disjonctif (Matheron, 1982), ni *a fortiori* avec l'espérance conditionnelle.

³ Une conséquence est que dans l'espace multidimensionnel la simulation séquentielle d'indicatrices ne correspond à aucun modèle de fonction aléatoire à loi marginale absolument continue, quelle que soit la covariance choisie (hormis, bien sûr, l'effet de pépite pur).

IV. Modèle bigamma

Les fonctions aléatoires précédemment étudiées sont de loi marginale gaussienne, ce qui implique notamment une symétrie de la structuration spatiale de leurs indicatrices par rapport au seuil médian. Il existe des processus qui s'appuient sur des lois marginales non gaussiennes et pour lesquels il est possible de former le krigeage disjonctif ; nous prendrons l'illustration des processus isofactoriels à loi marginale gamma (Matheron, 1973b ; Hu, 1988).

La loi gamma dépend d'un paramètre positif, noté α , et présente une dissymétrie d'autant plus importante que ce paramètre est proche de zéro. La loi isofactorielle bigamma est une extension de la loi bigaussienne, qui apparaît comme un cas limite lorsque α tend vers l'infini. Les facteurs de la loi bigamma sont les polynômes de Laguerre normés d'ordre α définis par la formule de Rodrigues (Chilès and Delfiner, 1999) :

$$\forall p \in \mathbf{N}, L_p^\alpha(y) = \sqrt{\frac{\Gamma(\alpha)}{\Gamma(p+1)\Gamma(\alpha+p)}} \frac{1}{e^{-y} y^{\alpha-1}} \frac{d^p}{dy^p} (e^{-y} y^{p+\alpha-1}).$$

Dans cette section, nous allons considérer un processus bigamma particulier, à savoir une fonction aléatoire *de diffusion* (Lantuéjoul, 1984 ; Matheron, 1983, 1984b ; Hu, 1988) le long de l'axe temporel. Un tel processus est markovien : connaissant la valeur présente du processus, les valeurs futures ne dépendent pas des valeurs passées. Les fonctions de covariance des facteurs sont toutes des fonctions à décroissance exponentielle. Pour s'éloigner du cas multigaussien, on fixe le paramètre α de la loi gamma égal à 0.5.

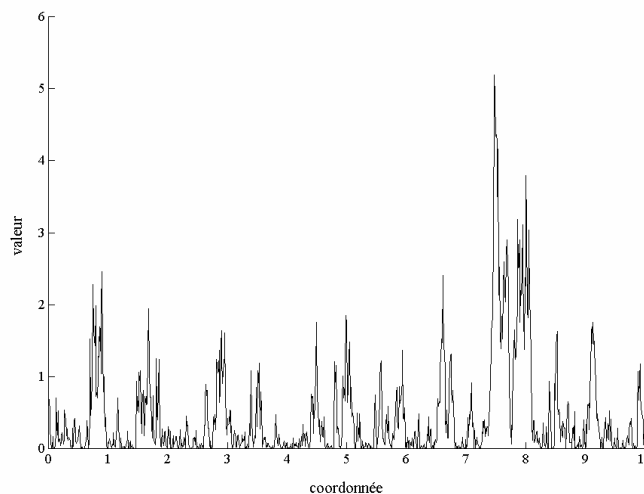


Figure 1.17. Réalisation d'un processus de diffusion gamma ($\alpha = 0.5$)

IV.1. Une donnée conditionnante

Le krigeage d'indicatrice s'écrit classiquement :

$$I(Y_0; y)^{KS} = r I(y_1; y) + (1 - r) F_\alpha(y)$$

avec $r = \frac{\alpha f_{\alpha+1}^2(y)}{F_\alpha(y)[1-F_\alpha(y)]} \sum_{p \geq 1} \frac{L_{p-1}^{\alpha+1}(y)^2}{p} \rho^p$, cf. équation (1.1)

$f_\alpha(y) = \frac{1}{\Gamma(\alpha)} \exp(-y) y^{\alpha-1}$: densité de la loi gamma de paramètre α

$F_\alpha(y)$: fonction de répartition associée.

Le krigeage disjonctif et l'espérance conditionnelle sont identiques :

$$I(Y_0; y)^{EC} = I(Y_0; y)^{KD} = \frac{1}{f_\alpha(y_1)} \int_0^y f_\alpha(y_0, y_1) dy_0 \quad (1.5)$$

Dans cette expression, $f_\alpha(y_0, y_1)$ est la densité bivariable du couple $\{Y_0, Y_1\}$; sous forme analytique (Hu, 1988) :

$$f_\alpha(y_0, y_1) = \frac{1}{\Gamma(\alpha)(1-\rho)} \exp\left(-\frac{y_0 + y_1}{1-\rho}\right) \left(\frac{y_0 y_1}{\rho}\right)^{(\alpha-1)/2} I_{\alpha-1}\left(2\sqrt{\frac{y_0 y_1 \rho}{1-\rho}}\right)$$

où ρ est le coefficient de corrélation entre Y_0 et Y_1 et où $I_{\alpha-1}$ est la fonction de Bessel modifiée de première espèce, définie par (Abramowitz and Stegun, 1972) :

$$I_{\alpha-1}(z) = \left(\frac{z}{2}\right)^{\alpha-1} \sum_{p=0}^{+\infty} \frac{1}{p! \Gamma(\alpha + p)} \left(\frac{z}{2}\right)^{2p}$$

Les estimateurs (1.5) s'obtiennent par intégration numérique. Il est aussi possible de développer la densité bivariable en polynômes de Laguerre normés d'ordre α $\{L_p^\alpha, p \in \mathbf{N}\}$:

$$f_\alpha(y_0, y_1) = f_\alpha(y_0) f_\alpha(y_1) \left[1 + \sum_{p \geq 1} \rho^p L_p^\alpha(y_0) L_p^\alpha(y_1)\right]$$

et d'intégrer terme à terme, ce qui donne, après quelques calculs :

$$I(Y_0; y)^{EC} = I(Y_0; y)^{KD} = F_\alpha(y) + \sum_{p \geq 1} \rho^p \sqrt{\frac{\alpha}{p}} f_{\alpha+1}(y) L_{p-1}^{\alpha+1}(y) L_p^\alpha(y_1)$$

Les écarts les plus importants entre le krigeage d'indicatrice et l'espérance conditionnelle ont lieu lorsque la valeur de la donnée est proche du seuil de l'indicatrice à estimer ; ils s'accroissent aussi avec le niveau de corrélation entre la donnée et la valeur inconnue (figure 1.18). L'amplitude des écarts est du même ordre de grandeur que dans le cas multigaussien (figure 1.2). La commodité de manipulation des indicatrices, avancée par certains auteurs comme un avantage décisif du krigeage d'indicatrices et autres méthodes "non paramétriques" (Isaaks, 1984 ; Journel, 1983, 1984), ne résout pas la difficulté due à la dissymétrie de la distribution marginale, en ce sens qu'elle n'améliore pas la précision des estimations.

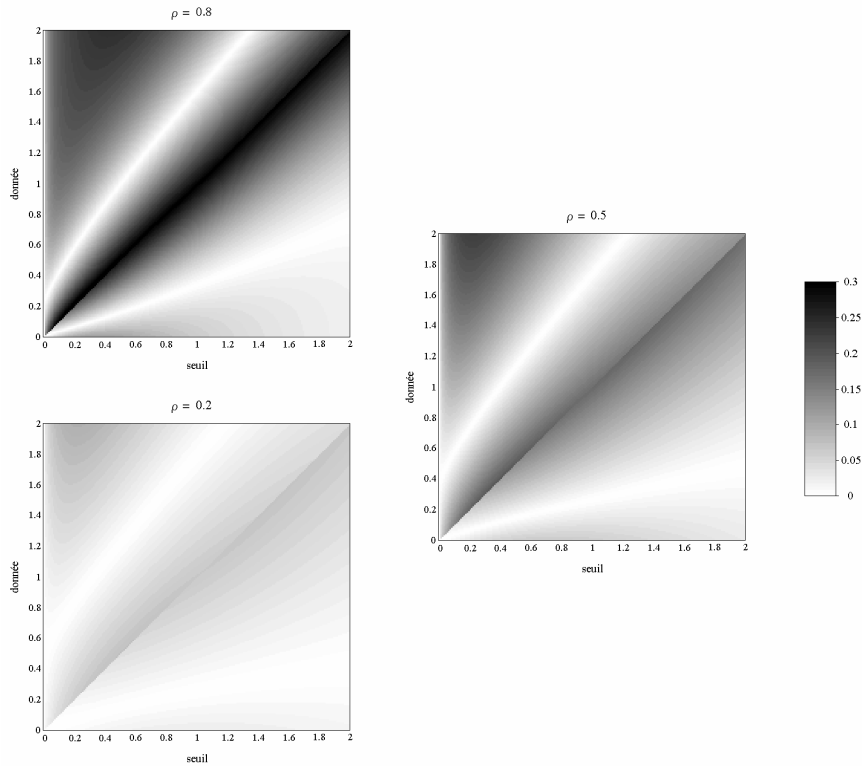
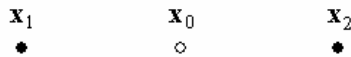


Figure 1.18. *Ecart absolu entre krigeage d'indicateur et espérance conditionnelle en fonction de la valeur du seuil de coupure et de la donnée conditionnante*

IV.2. Deux données conditionnantes

On décide de n'étudier qu'un seuil de coupure, égal à 0.5 ; cette valeur coïncide avec la moyenne de la loi marginale.

A) Sites alignés : x_0 milieu de $[x_1, x_2]$



Compte tenu de la propriété markovienne, la loi de Y_1 conditionnée par Y_0 et Y_2 est égale à la loi de Y_1 conditionnée par Y_0 , soit :

$$f_{\alpha}(y_1 | y_0, y_2) = \frac{f_{\alpha}(y_1, y_0, y_2)}{f_{\alpha}(y_0, y_2)} = f_{\alpha}(y_1 | y_0) = \frac{f_{\alpha}(y_0, y_1)}{f_{\alpha}(y_0)}$$

ce qui donne l'expression de la densité trivariante :

$$f_{\alpha}(y_1, y_0, y_2) = \frac{f_{\alpha}(y_0, y_1) f_{\alpha}(y_0, y_2)}{f_{\alpha}(y_0)} \tag{1.6}$$

L'espérance conditionnelle de l'indicatrice associée à la valeur Y_0 est donc :

$$I(Y_0; y)^{EC} = \frac{1}{f_\alpha(y_1, y_2)} \int_0^y f_\alpha(y_1, y_0, y_2) dy_0$$

Le calcul numérique de cette expression est délicat. Un moyen efficace consiste à développer la densité bivariable sous la forme d'un mélange poissonien de lois gamma (Matheron, 1983 ; Lantuéjoul, 1984) :

$$f_\alpha(y_0, y_1) = f_\alpha(y_1) e^{-\theta_{10} y_1} \sum_{p \geq 0} \frac{(\theta_{10} y_1)^p}{p!} \eta_{10} f_{p+\alpha}(\eta_{10} y_0) \text{ avec } \eta_{10} = \frac{1}{1-\rho_{10}} \text{ et } \theta_{10} = \eta_{10} \rho_{10} .$$

$$f_\alpha(y_0, y_2) = f_\alpha(y_2) e^{-\theta_{20} y_2} \sum_{q \geq 0} \frac{(\theta_{20} y_2)^q}{q!} \eta_{20} f_{q+\alpha}(\eta_{20} y_0) \text{ avec } \eta_{20} = \frac{1}{1-\rho_{20}} \text{ et } \theta_{20} = \eta_{20} \rho_{20} .$$

On obtient alors l'expression de l'espérance conditionnelle sous la forme d'une série dont le calcul est plus précis qu'une intégration numérique de la densité trivariable :

$$\begin{aligned} I(y_0, y)^{EC} &= \frac{\Gamma(\alpha) f_\alpha(y_1) f_\alpha(y_2)}{f_\alpha(y_1, y_2)} e^{-\theta_{10} y_1 - \theta_{20} y_2} \sum_{p, q \geq 0} \frac{(\theta_{10} y_1)^p}{p! \Gamma(p+\alpha)} \frac{(\theta_{20} y_2)^q}{q! \Gamma(q+\alpha)} \int_0^y \frac{\eta_{10}^{p+\alpha} \eta_{20}^{q+\alpha} \Gamma(p+q+\alpha) f_{p+q+\alpha}(\eta y_0) dy_0}{\eta^{p+q+\alpha-1}} \\ &= \frac{\Gamma(\alpha) f_\alpha(y_1) f_\alpha(y_2)}{f_\alpha(y_1, y_2)} e^{-\theta_{10} y_1 - \theta_{20} y_2} \sum_{p, q \geq 0} \frac{(\theta_{10} y_1)^p}{p! \Gamma(p+\alpha)} \frac{(\theta_{20} y_2)^q}{q! \Gamma(q+\alpha)} \frac{\eta_{10}^{p+\alpha} \eta_{20}^{q+\alpha} \Gamma(p+q+\alpha) F_{p+q+\alpha}(\eta y)}{\eta^{p+q+\alpha}} \\ &= \frac{(\rho_{12} y_1 y_2)^{(\alpha-1)/2} e^{(\eta_{12}-\eta_{10})y_1 + (\eta_{12}-\eta_{20})y_2}}{\eta_{12} I_{\alpha-1}(2\eta_{12} \sqrt{\rho_{12} y_1 y_2})} \sum_{p, q \geq 0} \frac{(\theta_{10} y_1)^p}{p! \Gamma(p+\alpha)} \frac{(\theta_{20} y_2)^q}{q! \Gamma(q+\alpha)} \frac{\eta_{10}^{p+\alpha} \eta_{20}^{q+\alpha} \Gamma(p+q+\alpha) F_{p+q+\alpha}(\eta y)}{\eta^{p+q+\alpha}} \end{aligned}$$

où l'on a posé $\eta_{12} = \frac{1}{1-\rho_{12}}$ et $\eta = \eta_{10} + \eta_{20} - 1$.

Bien qu'elle comporte une double sommation, la série précédente converge plus rapidement que celle obtenue par développement de la densité trivariable en polynômes de Laguerre et intégration terme à terme. L'expression se simplifie encore si l'on écrit :

$$\eta = \frac{1-\rho_{10} \rho_{20}}{(1-\rho_{10})(1-\rho_{20})} = \frac{\eta_{10} \eta_{20}}{\eta_{12}} \text{ et } a_{12} = \eta_{12} \sqrt{\rho_{12} y_1 y_2}$$

de sorte que

$$I(y_0, y)^{EC} = \frac{a_{12}^{\alpha-1} e^{(\eta_{12}-\eta_{10})y_1 + (\eta_{12}-\eta_{20})y_2}}{I_{\alpha-1}(2a_{12})} \sum_{p, q \geq 0} \frac{(\eta_{10} \theta_{10} y_1)^p}{p! \Gamma(p+\alpha)} \frac{(\eta_{20} \theta_{20} y_2)^q}{q! \Gamma(q+\alpha)} \frac{\Gamma(p+q+\alpha) F_{p+q+\alpha}(\eta y)}{\eta^{p+q}}$$

Le krigeage disjonctif met en jeu un développement en polynômes de Laguerre d'ordre α normés $\{L_p^\alpha, p \in \mathbb{N}\}$ et le krigeage simple de chacun de ces polynômes :

$$I(Y_0; y)^{KD} = F_\alpha(y) + \sum_{p \geq 1} \sqrt{\frac{\alpha}{p}} f_{\alpha+1}(y) L_{p-1}^{\alpha+1}(y) [L_p^\alpha(Y_0)]^{KS}$$

$$\text{où } [L_p^\alpha(Y_0)]^{KS} = \frac{1}{1-\rho_{12}^{2p}} [(\rho_{10}^p - \rho_{12}^p \rho_{20}^p) L_p^\alpha(y_1) + (\rho_{20}^p - \rho_{12}^p \rho_{10}^p) L_p^\alpha(y_2)] .$$

Enfin, le krigeage d'indicatrice s'écrit sous la forme :

$$I(Y_0; y)^{KS} = \frac{1}{1-r_{12}^2} [(r_{10} - r_{12} r_{20}) I(y_1; y) + (r_{20} - r_{12} r_{10}) I(y_2; y)] + \left(1 - \frac{r_{10} + r_{20}}{1 + r_{12}}\right) F_\alpha(y) \quad (1.7)$$

Sur la figure 1.19, on observe que le krigeage d'indicatrice et le krigeage disjonctif s'écartent davantage de l'espérance conditionnelle quand le niveau de corrélation est élevé. Cependant, le krigeage disjonctif constitue une bien meilleure approximation de l'espérance conditionnelle que le krigeage d'indicatrice, dont les écarts sont importants dès lors que les valeurs des deux données sont situées du même côté du seuil à estimer et proches de ce seuil. La prise en compte du comportement conjoint des indicatrices par une modélisation de la loi bivariable entière améliore considérablement les estimations ; chercher un algorithme non paramétrique qui ignore une grande partie de cette loi bivariable conduit à des résultats grossiers.

Il est certain qu'une modélisation au-delà de la loi bivariable améliorerait encore la performance de l'estimateur. Une comparaison des résultats relatifs aux krigeage disjonctif et d'indicatrice montre cependant que le gain principal s'opère au niveau de la loi bivariable, plus qu'il ne serait possible d'obtenir à l'aide de lois d'ordre supérieur.

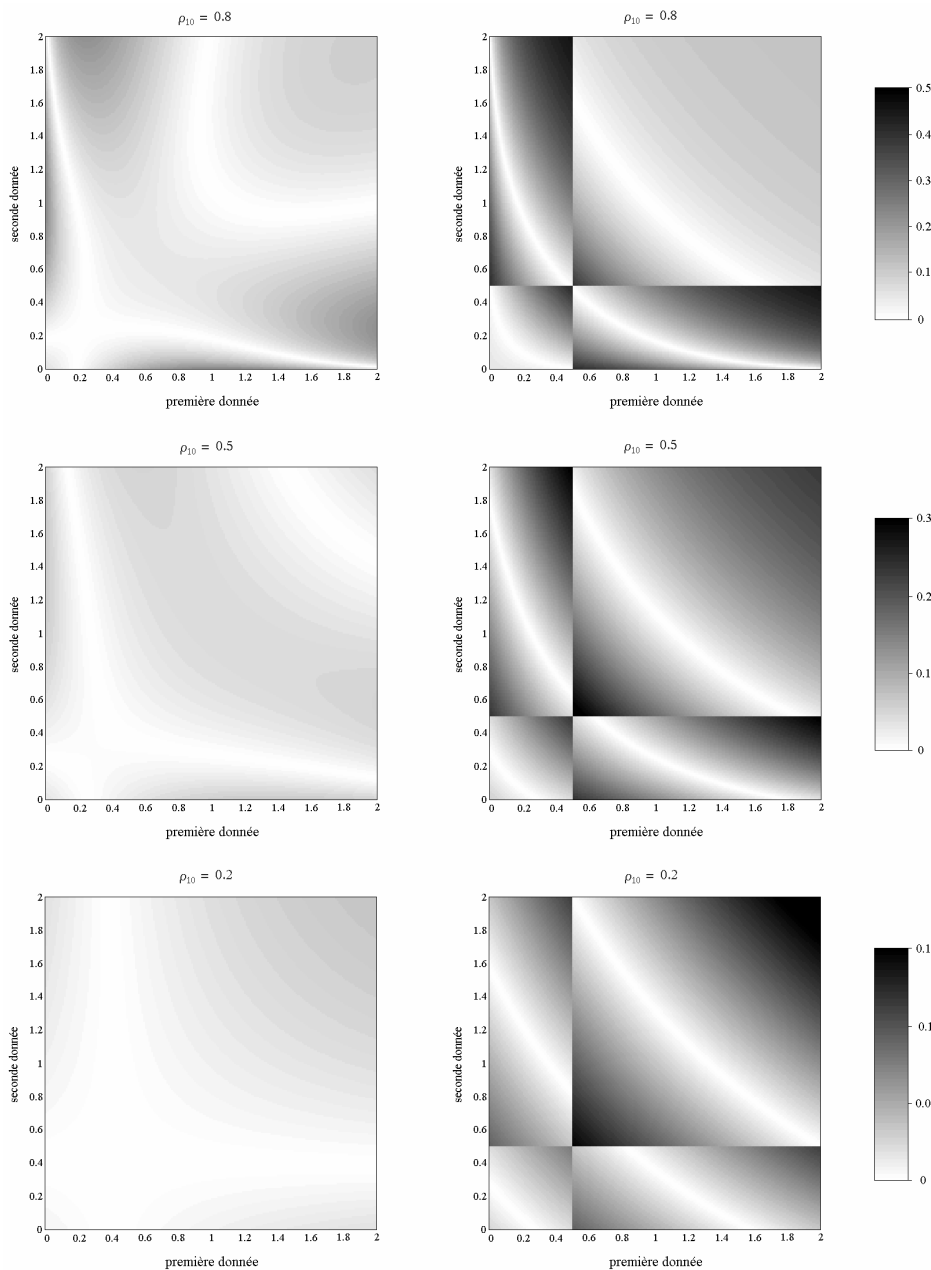


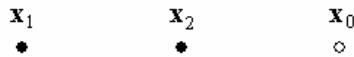
Figure 1.19. *Ecart absolu entre krigeage disjonctif et espérance conditionnelle (gauche) et entre krigeage d'indicatrice et espérance conditionnelle (droite), pour trois valeurs du coefficient de corrélation \mathbf{r}_{10}*

Le tableau 1.6 fournit des statistiques approchées par calcul numérique des écarts par rapport à l'espérance conditionnelle. Les valeurs sur les axes de coordonnées (une donnée nulle, correspondant à une densité de probabilité infinie) ont été négligées dans le calcul, ce qui, au vu des graphiques précédents, favorise le krigeage d'indicatrice.

		krigeage disjonctif	krigeage d'indicatrice
$\rho_{10} = 0.8$	écart absolu moyen	0.027	0.072
	racine de l'écart quadratique moyen	0.040	0.113
$\rho_{10} = 0.5$	écart absolu moyen	0.014	0.056
	racine de l'écart quadratique moyen	0.020	0.085
$\rho_{10} = 0.2$	écart absolu moyen	0.003	0.024
	racine de l'écart quadratique moyen	0.005	0.037

Tableau 1.6. Statistiques approchées des écarts par rapport à l'espérance conditionnelle

B) Sites alignés : x_2 milieu de $[x_0, x_1]$



Du fait de la propriété markovienne, l'espérance conditionnelle en x_0 ne dépend pas de la valeur au site x_1 et coïncide avec le krigeage disjonctif :

$$I(Y_0; y)^{EC} = I(Y_0; y)^{KD} = \frac{1}{f_\alpha(y_1, y_2)} \int_0^y f_\alpha(y_1, y_2, y_0) dy_0 = \frac{1}{f_\alpha(y_2)} \int_0^y f_\alpha(y_2, y_0) dy_0 .$$

L'expression du krigeage d'indicatrice est quant à elle inchangée [Eq. (1.7)].

Bien que l'une des données soit écrantée, le krigeage d'indicatrice fournit une précision du même ordre de grandeur que pour la configuration précédente, comme l'indiquent la figure et le tableau ci-après.

	$\rho_{12} = 0.8$	$\rho_{12} = 0.5$	$\rho_{12} = 0.2$
écart absolu moyen	0.079	0.047	0.017
racine de l'écart quadratique moyen	0.112	0.068	0.026

Tableau 1.7. Statistiques approchées des écarts entre krigeage d'indicatrice et espérance conditionnelle

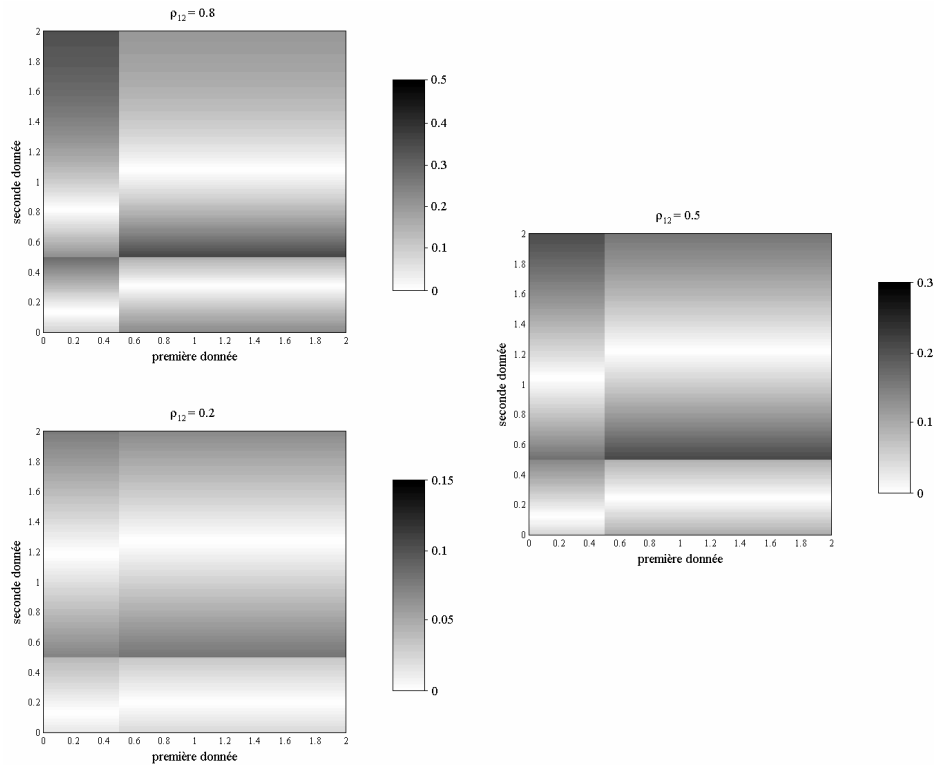


Figure 1.20. Ecart absolu entre krigeage d'indicatrice et espérance conditionnelle, pour trois valeurs du coefficient de corrélation \mathbf{r}_{12}

V. Fonctions aléatoires en tout ou rien

On s'intéresse à présent à des fonctions aléatoires binaires (à valeurs 0 ou 1), qui peuvent être représentées par une fonction indicatrice. Le krigeage de l'indicatrice coïncide avec son krigeage disjonctif puisque la loi bivariable est entièrement caractérisée par la fonction de covariance de l'unique indicatrice. Le résultat est égal à l'espérance conditionnelle dans le cas d'une seule donnée conditionnante, car tous les estimateurs ne mettent en jeu que la loi bivariable entre la donnée (Y_1) et la valeur à estimer (Y_0) :

$$Y_0^{EC} = Y_0^{KD} = Y_0^{KI} = \rho_{10} Y_1 + (1 - \rho_{10}) E(Y).$$

Désormais, seules seront étudiées les situations avec deux ou trois données conditionnantes.

V.1. Fonctions aléatoires en damier

Considérons deux modèles particuliers de fonctions aléatoires binaires (d'après Ch. Lantuéjoul) :

- (a) Dans le plan, on construit deux réseaux de droites (l'un vertical, l'autre horizontal) séparées par des distances exponentielles indépendantes de moyenne a . On obtient une partition du plan en rectangles, que l'on value indépendamment à 0 ou 1 avec la même probabilité 1/2. On peut voir ce modèle comme un cas particulier de mosaïque de Poisson, associée à un processus poissonien de densité anisotrope (concentrée le long des axes de coordonnées).

- (b) Même situation sauf que les droites sont séparées par des distances exponentielles de moyenne $2a$, tandis que les valuations des rectangles sont alternativement mises à 0 ou 1.

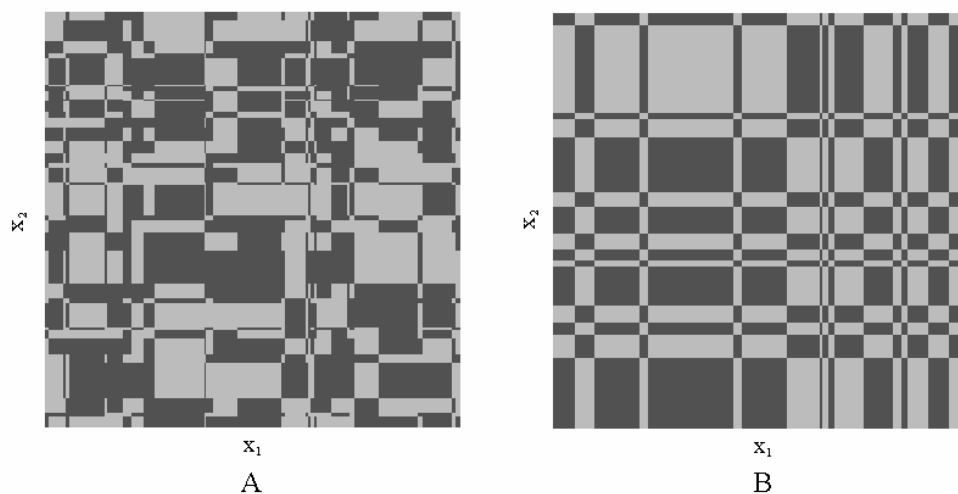


Figure 1.21. Réalisations de modèles en “damier”

Ces deux fonctions aléatoires ont en commun la loi marginale (loi à deux atomes, 0 et 1, avec la même probabilité) ainsi que les lois bivariable et trivariable. Leur différence se note seulement à partir de la loi quadrivariable. En effet, s’agissant de fonctions aléatoires binaires, la loi bivariable est caractérisée par la fonction de covariance ; en outre, la propriété d’autodualité⁴ implique que la loi trivariable (et plus généralement, la loi d’ordre $2p+1$) se déduit de la loi bivariable (loi d’ordre $2p$). Nous allons démontrer l’identité des fonctions de covariance, ce qui établira le résultat annoncé.

Processus (a)

En utilisant la propriété d’absence de mémoire de la loi exponentielle, on peut exprimer la probabilité que deux sites séparés de $\mathbf{h} = (h_1, h_2)$ se trouvent dans le même rectangle :

$$\rho(\mathbf{h}) = \int_{|h_1|/a}^{+\infty} e^{-u} du \int_{|h_2|/a}^{+\infty} e^{-v} dv = \exp\left(-\frac{|h_1| + |h_2|}{a}\right).$$

⁴ Un ensemble fermé aléatoire (représenté par une fonction indicatrice) est autodual s’il admet la même loi spatiale que la fermeture de son complémentaire, c’est-à-dire si l’on peut échanger les valeurs 0 et 1 sans modifier la loi spatiale. Cette propriété implique en particulier que les proportions de 0 et de 1 sont égales. La formule de Poincaré s’écrit sous la forme :

$$\text{Prob}\left[\bigcup_{i=1}^{2p+1} A_i\right] = \sum_{i=1}^{2p+1} \text{Prob}[A_i] - \sum_{\substack{i,j=1 \\ i \neq j}}^{2p+1} \text{Prob}[A_i \cap A_j] + \dots + (-1)^{2p} \text{Prob}\left[\bigcap_{i=1}^{2p+1} A_i\right]$$

où A_i est un événement qui se réfère au site \mathbf{x}_i (appartenance ou non à l’ensemble aléatoire). Or, l’hypothèse d’autodualité implique, en notant A_i^c l’événement complémentaire de A_i :

$$\text{Prob}\left[\bigcap_{i=1}^{2p+1} A_i\right] = \text{Prob}\left[\bigcap_{i=1}^{2p+1} A_i^c\right] = 1 - \text{Prob}\left[\bigcup_{i=1}^{2p+1} A_i\right]$$

d’où

$$2\text{Prob}\left[\bigcup_{i=1}^{2p+1} A_i\right] = \sum_{i=1}^{2p+1} \text{Prob}[A_i] - \sum_{\substack{i,j=1 \\ i \neq j}}^{2p+1} \text{Prob}[A_i \cap A_j] + \dots + 1$$

ce qui permet d’exprimer les caractéristiques de la loi à $(2p+1)$ variables à l’aide seulement de caractéristiques de la loi à $(2p)$ variables.

Cette fonction $\rho(\mathbf{h})$ est le corrélogramme de la fonction aléatoire construite, puisque sa covariance est :

$$\text{cov}[Y(\mathbf{x}+\mathbf{h}), Y(\mathbf{x})] = \rho(\mathbf{h}) \times C(\mathbf{0}) + [1 - \rho(\mathbf{h})] \times 0 = \frac{1}{4} \exp\left(-\frac{|\mathbf{h}_1| + |\mathbf{h}_2|}{a}\right).$$

Processus (b)

Deux sites distants de $\mathbf{h} = (h_1, h_2)$ ont la même valeur si les nombres de droites horizontales et verticales qui les séparent ont même parité, ce qui se produit avec la probabilité

$$\begin{aligned} P(\mathbf{h}) &= \sum_{p,q=0}^{+\infty} e^{-\frac{|h_1|}{2a}} \frac{(|h_1|/2a)^{2p}}{(2p)!} e^{-\frac{|h_2|}{2a}} \frac{(|h_2|/2a)^{2q}}{(2q)!} + \sum_{p,q=0}^{+\infty} e^{-\frac{|h_1|}{2a}} \frac{(|h_1|/2a)^{2p+1}}{(2p+1)!} e^{-\frac{|h_2|}{2a}} \frac{(|h_2|/2a)^{2q+1}}{(2q+1)!} \\ &= e^{-\frac{|h_1|}{2a}} \text{ch}(|h_1|/2a) e^{-\frac{|h_2|}{2a}} \text{ch}(|h_2|/2a) + e^{-\frac{|h_1|}{2a}} \text{sh}(|h_1|/2a) e^{-\frac{|h_2|}{2a}} \text{sh}(|h_2|/2a) \\ &= \frac{1}{4} (1 + e^{-\frac{|h_1|}{a}}) (1 + e^{-\frac{|h_2|}{a}}) + \frac{1}{4} (1 - e^{-\frac{|h_1|}{a}}) (1 - e^{-\frac{|h_2|}{a}}) \\ &= \frac{1}{2} + \frac{1}{2} \exp\left(-\frac{|\mathbf{h}_1| + |\mathbf{h}_2|}{a}\right) \end{aligned}$$

On a utilisé le fait que les nombres de droites horizontales et verticales entre les deux sites suivent des lois de Poisson de paramètres respectifs $|h_1|/2a$ et $|h_2|/2a$. Par suite, on a :

$$\begin{aligned} E[Y(\mathbf{x}+\mathbf{h})Y(\mathbf{x})] &= \text{Prob}[Y(\mathbf{x})=1] \times P(\mathbf{h}) = \frac{1}{2} P(\mathbf{h}) \\ \text{cov}[Y(\mathbf{x}+\mathbf{h}), Y(\mathbf{x})] &= \frac{1}{2} P(\mathbf{h}) - \frac{1}{4} = \frac{1}{4} \exp\left(-\frac{|\mathbf{h}_1| + |\mathbf{h}_2|}{a}\right) \\ \rho(\mathbf{h}) &= \text{corr}[Y(\mathbf{x}+\mathbf{h}), Y(\mathbf{x})] = 2P(\mathbf{h}) - 1 = \exp\left(-\frac{|\mathbf{h}_1| + |\mathbf{h}_2|}{a}\right). \end{aligned}$$

Examinons à présent l'écart entre l'espérance conditionnelle et le krigeage d'indicatrice. Nous étudierons le cas de deux et trois données conditionnantes, de façon à mettre en jeu les lois tri et quadrivariées. Ce n'est que dans la seconde situation que l'on observera une différence entre les estimateurs associés aux modèles (a) et (b) ; pour cette raison, les démonstrations relatives au cas de deux données conditionnantes se réfèrent à la fonction aléatoire de type (b).

V.1.1. Deux données conditionnantes

On reprend les notations précédemment introduites, où ρ désigne le coefficient de corrélation :

$$\begin{aligned} P_{12} &= (1 + \rho_{12})/2 : \text{probabilité que } Y_1 \text{ et } Y_2 \text{ soient égales} \\ P_{10} &= (1 + \rho_{10})/2 : \text{probabilité que } Y_1 \text{ et } Y_0 \text{ soient égales} \\ P_{20} &= (1 + \rho_{20})/2 : \text{probabilité que } Y_2 \text{ et } Y_0 \text{ soient égales.} \end{aligned}$$

A cause de la propriété d'autodualité, ces probabilités se divisent par moitié selon que les valeurs sont égales à 1 ou au contraire égales à 0. De même, la probabilité $1 - P_{12}$ que Y_1 et Y_2 soient différentes se divise par moitié selon que $(Y_1, Y_2) = (1, 0)$ ou $(0, 1)$. On note P_{120} la probabilité que Y_1, Y_2 et Y_0 soient toutes trois égales. On a :

$$P_{120} + (P_{12} - P_{120}) + (P_{10} - P_{120}) + (P_{20} - P_{120}) = 1, \text{ soit } P_{120} = (1 + \rho_{12} + \rho_{10} + \rho_{20})/4.$$

De même que précédemment, la moitié de cette probabilité correspond au cas où les trois valeurs sont égales à 1, l'autre moitié au cas où elles sont nulles.

On est à présent en mesure d'exprimer l'espérance conditionnelle, en distinguant tous les cas possibles pour les valeurs des données conditionnantes :

- si $Y_1 = 1$ et $Y_2 = 0$, ce qui se produit avec la probabilité $(1 - P_{12}) / 2$:

$$Y_0^{EC} = \text{Prob}(Y_0 = Y_1) = \frac{(P_{10} - P_{120})/2}{(1 - P_{12})/2} = \frac{1}{2} \left(1 + \frac{\rho_{10} - \rho_{20}}{1 - \rho_{12}}\right)$$

- si $Y_1 = 0$ et $Y_2 = 1$, ce qui se produit avec la probabilité $(1 - P_{12}) / 2$:

$$Y_0^{EC} = \text{Prob}(Y_0 = Y_2) = \frac{(P_{20} - P_{120})/2}{(1 - P_{12})/2} = \frac{1}{2} \left(1 + \frac{\rho_{20} - \rho_{10}}{1 - \rho_{12}}\right)$$

- si $Y_1 = 1$ et $Y_2 = 1$, ce qui se produit avec la probabilité $P_{12} / 2$:

$$Y_0^{EC} = \text{Prob}(Y_0 = Y_1) = \frac{P_{120}/2}{P_{12}/2} = \frac{1}{2} \left(1 + \frac{\rho_{10} + \rho_{20}}{1 + \rho_{12}}\right)$$

- si $Y_1 = 0$ et $Y_2 = 0$, ce qui se produit avec la probabilité $P_{12} / 2$:

$$Y_0^{EC} = \text{Prob}(Y_0 \neq Y_1) = 1 - \frac{P_{120}/2}{P_{12}/2} = \frac{1}{2} \left(1 - \frac{\rho_{10} + \rho_{20}}{1 + \rho_{12}}\right).$$

Or, il se trouve que ces formules coïncident avec le krigeage d'indicatrice

$$Y_0^{KS} = \frac{1}{1 - \rho_{12}^2} [(\rho_{10} - \rho_{12} \rho_{20}) Y_1 + (\rho_{20} - \rho_{12} \rho_{10}) Y_2] + \frac{1}{2} \left(1 - \frac{\rho_{10} + \rho_{20}}{1 + \rho_{12}}\right).$$

Dans les deux modèles autoduaux proposés, le krigeage d'indicatrice est identique à l'espérance conditionnelle lorsqu'il y a une ou deux données conditionnantes, quelle que soit leur configuration géométrique par rapport au site à estimer.

L'interprétation que l'on tire de ce résultat est la suivante : lorsque les proportions de 0 et 1 sont égales, le krigeage d'indicatrice suppose que la loi du processus est autoduale, c'est-à-dire que l'on peut échanger les valeurs 0 et 1 sans modifier la loi spatiale, et déduit la loi trivariante à partir de la loi bivariante (covariance de l'indicatrice).

V.1.2. Trois données conditionnantes

On s'intéresse au cas où les données et la valeur à estimer constituent les sommets d'un rectangle de côtés horizontaux et verticaux.

x_1 * $\diamond x_0$

x_2 * * x_3

Le krigeage d'indicatrice donne comme estimation en \mathbf{x}_0

$$Y_0^{KS} = \lambda_1 Y_1 + \lambda_2 Y_2 + \lambda_3 Y_3 + (1 - \lambda_1 - \lambda_2 - \lambda_3) \frac{1}{2}$$

où les pondérateurs de krigeage sont donnés par le système matriciel:

$$\begin{pmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \begin{pmatrix} \rho_{10} \\ \rho_{20} \\ \rho_{30} \end{pmatrix}$$

avec $\rho_{23} = \rho_{10}$, $\rho_{12} = \rho_{30}$, $\rho_{13} = \rho_{12} \rho_{23}$ et $\rho_{20} = \rho_{12} \rho_{23}$.

Le site \mathbf{x}_2 est affecté d'un poids non nul, bien que cela ne serait pas le cas en l'absence du site \mathbf{x}_3 à cause de la propriété d'écran d'une covariance factorisable entre deux directions orthogonales dans l'espace à deux dimensions (Rivoirard, 1984). L'expression des poids de krigeage est relativement compliquée et leur explicitation n'a guère d'intérêt ici.

Pour calculer l'espérance conditionnelle, il convient de distinguer le type de processus.

Processus (a)

Sachant que le coefficient de corrélation entre deux sites représente la probabilité qu'ils se trouvent dans le même rectangle, on a :

- si $Y_1 = Y_3 \neq Y_2$, alors $Y_0^{EC} = \frac{1}{2}$
- si $Y_2 = Y_3 \neq Y_1$, alors $Y_0^{EC} = \frac{2}{1 + \rho_{23}} (\rho_{23} Y_1 + \frac{1 - \rho_{23}}{2} \frac{1}{2}) = \frac{1}{2} + \frac{\rho_{23}}{1 + \rho_{23}} (2 Y_1 - 1)$
- si $Y_1 = Y_2 \neq Y_3$, alors $Y_0^{EC} = \frac{2}{1 + \rho_{12}} (\rho_{12} Y_3 + \frac{1 - \rho_{12}}{2} \frac{1}{2}) = \frac{1}{2} + \frac{\rho_{12}}{1 + \rho_{12}} (2 Y_3 - 1)$
- si $Y_1 = Y_2 = Y_3$, alors

$$Y_0^{EC} = \frac{4(\rho_{12} \rho_{23} Y_2 + \rho_{12} \frac{1 - \rho_{23}}{2} Y_3 + \rho_{23} \frac{1 - \rho_{12}}{2} Y_1 + \frac{1 - \rho_{12}}{2} \frac{1 - \rho_{23}}{2} \frac{1}{2})}{1 + \rho_{12} + \rho_{23} + \rho_{12} \rho_{23}}$$

$$= \frac{1}{2} + \frac{(\rho_{12} + \rho_{23})}{(1 + \rho_{12})(1 + \rho_{23})} (2 Y_1 - 1)$$

Processus (b)

Compte-tenu de la structure en "damier" du processus, trois cas se dessinent :

- si $Y_1 = Y_3$, alors $Y_0^{EC} = Y_2$
- si $Y_1 = Y_2 \neq Y_3$, alors $Y_0^{EC} = Y_3$
- si $Y_2 = Y_3 \neq Y_1$, alors $Y_0^{EC} = Y_1$

Contrairement au processus de type (a) l'erreur d'estimation est nulle, c'est-à-dire que l'estimation est parfaite.

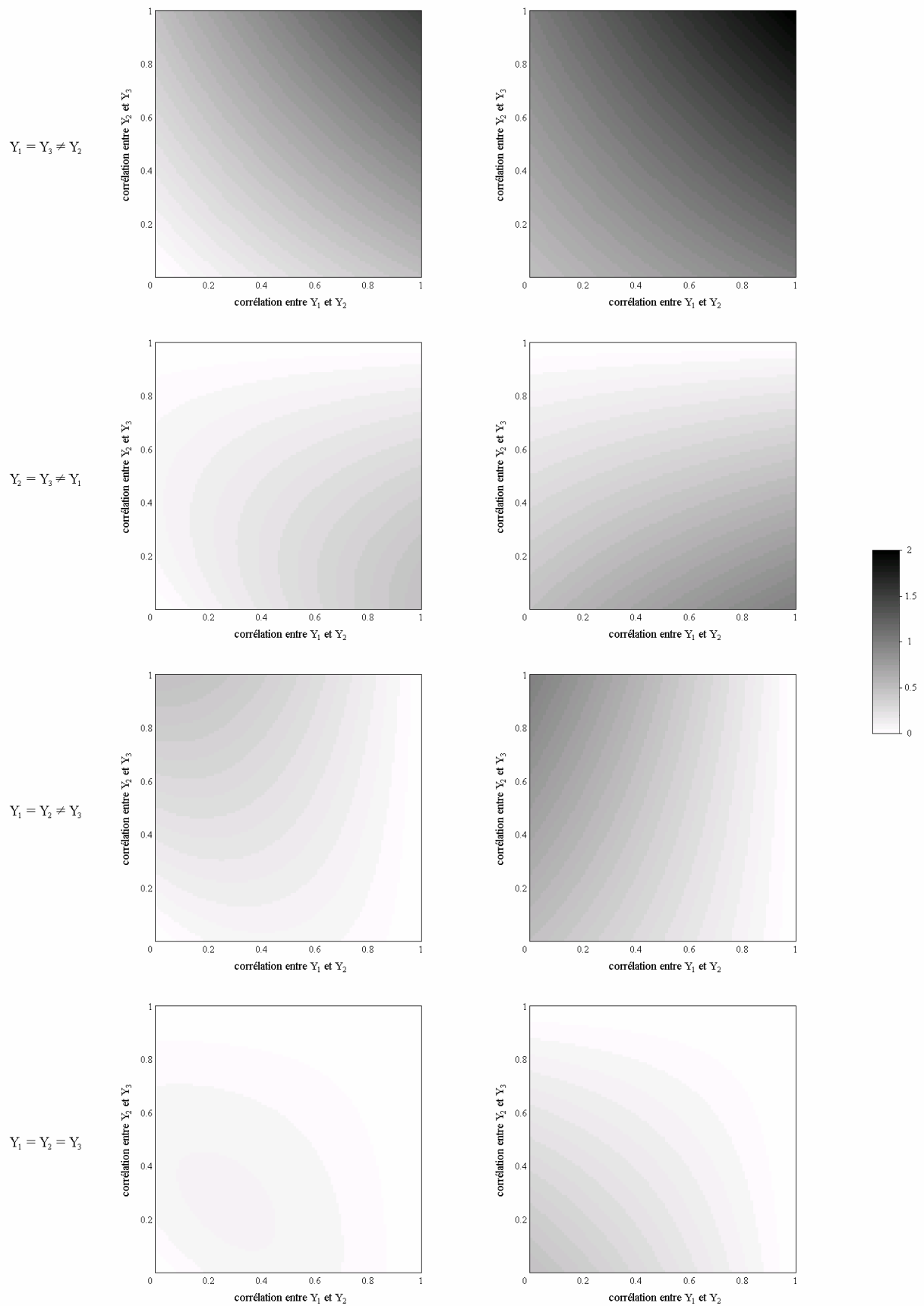


Figure 1.22. *Ecart absolu entre krigeage d'indicatrice et espérance conditionnelle en fonction des coefficients de corrélation entre données conditionnantes gauche : processus (a), droite : processus (b)*

Les fortes erreurs correspondent à des situations peu fréquentes, par exemple lorsque ρ_{12} et ρ_{23} tendent vers 1 alors que Y_2 diffère de Y_1 et Y_3 : dans une telle circonstance, l'estimation par krigeage d'indicatrice sort nettement de l'intervalle $[0,1]$ (quasi-singularité de la matrice de krigeage, ce qui entraîne l'apparition de pondérateurs négatifs), alors que l'espérance conditionnelle se situe toujours dans cet intervalle.

En comparant les graphiques de la colonne de gauche avec ceux de la colonne de droite, on constate que l'écart entre krigeage d'indicatrice et espérance conditionnelle est moindre dans le cas du processus (a). Ce dernier présente une organisation spatiale des valeurs moins "prévisible", en ce sens que l'espérance conditionnelle est affectée d'une erreur d'estimation, tandis que le processus (b) conduit à une espérance conditionnelle sans erreur. Le krigeage d'indicatrice favorise le processus dans lequel la valeur non mesurée est la plus incertaine ; il ne voit pas les liaisons présentes dans les lois à plus de deux variables, qui sont particulièrement marquées dans l'exemple du processus (b) pour la configuration étudiée.

V.2. Processus autoduaux

Nous avons signalé que, pour les fonctions aléatoires binaires, le krigeage d'indicatrice coïncide avec l'espérance conditionnelle en présence d'une seule donnée conditionnante. Ce résultat s'étend au cas de deux données conditionnantes lorsque l'on a affaire à un processus autodual, tels ceux présentés au paragraphe précédent. Un processus binaire est dit autodual lorsque l'on peut échanger les valeurs 0 et 1 sans modifier la loi spatiale.

Dans le cas d'un processus binaire, l'espérance d'une valeur Y_0 conditionnée par deux données Y_1 et Y_2 s'écrit sous la forme :

$$Y_0^{EC} = a + b Y_1 + c Y_2 + d Y_1 Y_2$$

puisque $\{1, Y_1, Y_2, Y_1 Y_2\}$ forment une base sur l'ensemble des fonctions de Y_1 et Y_2 .

Par linéarité de l'espérance conditionnelle, il vient :

$$\begin{aligned} (1 - Y_0)^{EC} &= 1 - a - b Y_1 - c Y_2 - d Y_1 Y_2 \\ &= (1 - a - b - c - d) + (b + d)(1 - Y_1) + (c + d)(1 - Y_2) - d(1 - Y_1)(1 - Y_2) \end{aligned}$$

Sous l'hypothèse d'autodualité, les coefficients sont inchangés entre l'estimation de Y_0 à partir de Y_1 et Y_2 et celle de $1 - Y_0$ à partir de $1 - Y_1$ et $1 - Y_2$, d'où :

$$\begin{cases} 1 - a - b - c - d = a \\ b + d = b \\ c + d = c \\ d = -d \end{cases} \quad \text{soit} \quad \begin{cases} 2a + b + c = 1 \\ d = 0 \end{cases}$$

La nullité du coefficient d implique que l'espérance conditionnelle s'écrit comme une combinaison linéaire des deux données conditionnantes ; on en déduit qu'elle est identique au krigeage, en tant que combinaison linéaire des données qui minimise la variance d'estimation. Ce résultat se généralise à un nombre pair de données conditionnantes : dans l'expression de l'espérance conditionnelle, le produit de plus haut degré est nécessairement d'ordre impair. L'espérance conditionnelle coïncidera avec le krigeage d'indicatrice si, à cause des effets d'écran ou de l'indépendance entre certaines valeurs, son expression ne fait intervenir aucun produit d'ordre supérieur ou égal à 3.

V.3. Gaussienne tronquée

Soient y un réel et $\{Y(\mathbf{x}), \mathbf{x} \in \mathbb{R}^d\}$ une fonction aléatoire gaussienne de moyenne nulle, variance unité et corrélogramme $\rho(\mathbf{h})$. On obtient une fonction aléatoire binaire en considérant l'indicatrice cumulée associée au seuil y , soit $I(Y(\mathbf{x}); y)$. La valeur du seuil détermine les proportions de 0 et de 1 ; si $y = 0$, ces proportions sont égales et le processus est autodual.

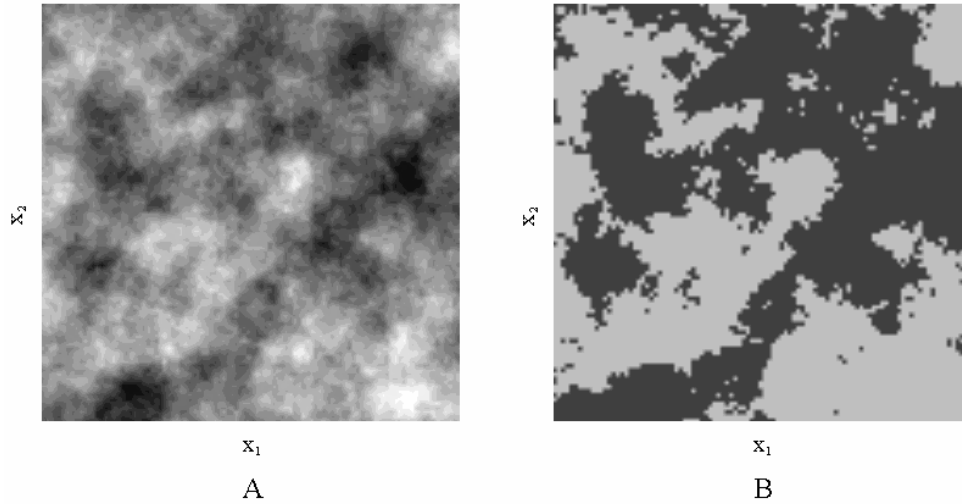


Figure 1.23. A, réalisation d'un processus gaussien isotrope et B, image obtenue par coupure au seuil médian

Le corrélogramme $r(\mathbf{h})$ de la fonction indicatrice se déduit de celui de la gaussienne (Matheron, 1975b ; Xiao, 1985) :

$$r(\mathbf{h}) = \frac{1}{2\pi G(y)[1-G(y)]} \int_0^{\arcsin[\rho(\mathbf{h})]} \exp\left[-\frac{y^2}{1+\sin \theta}\right] d\theta .$$

Pour les calculs pratiques, il est commode d'introduire le développement de la fonction indicatrice en polynômes d'Hermite normés $\{H_p, p \in \mathbb{N}\}$; on obtient [Eq. (1.1)] :

$$r(\mathbf{h}) = \frac{g(y)^2}{G(y)[1-G(y)]} \sum_{p \geq 1} \frac{H_{p-1}(y)^2}{p} \rho(\mathbf{h})^p$$

où g et G sont les densités standard et cumulée de la loi normale réduite.

L'expression du krigeage d'indicatrice est identique à celle vue dans le cadre multigaussien. En revanche, l'espérance conditionnelle est différente, puisque les données conditionnantes sont ici des valeurs binaires au lieu de valeurs continues. Dans la suite de cette section, nous analyserons le cas de deux données conditionnantes. Pour cela, commençons par déterminer les probabilités mono, bi et tri-variables susceptibles d'intervenir dans les calculs :

- $E[I(Y_1; y)] = E[I(Y_2; y)] = E[I(Y_0; y)] = G(y)$
- $E[I(Y_1; y)I(Y_2; y)] = G(y)[1-G(y)]r_{12} + G(y)^2 .$

Les autres probabilités bivariées se déduisent par permutation des indices.

La probabilité trivariante $E[I(Y_1; y) I(Y_2; y) I(Y_0; y)]$ peut s'exprimer à l'aide des coefficients du développement des indicatrices en polynômes d'Hermite normés ; toutefois, son calcul pose des problèmes de convergence numérique, de sorte que l'on préférera l'évaluer par intégration numérique, selon la situation considérée. On est à présent en mesure de spécifier l'espérance conditionnelle, selon les valeurs des données conditionnantes :

- si $I(Y_1; y) = I(Y_2; y) = 1$:

$$I(Y_0; y)^{EC} = \frac{\text{Prob}[Y_1 < y, Y_2 < y, Y_0 < y]}{\text{Prob}[Y_1 < y, Y_2 < y]} = \frac{E[I(Y_1; y) I(Y_2; y) I(Y_0; y)]}{E[I(Y_1; y) I(Y_2; y)]}$$

- si $I(Y_1; y) = 1$ et $I(Y_2; y) = 0$:

$$I(Y_0; y)^{EC} = \frac{\text{Prob}[Y_1 < y, Y_2 > y, Y_0 < y]}{\text{Prob}[Y_1 < y, Y_2 > y]} = \frac{E\{I(Y_1; y)[1 - I(Y_2; y)]I(Y_0; y)\}}{E\{I(Y_1; y)[1 - I(Y_2; y)]\}}$$

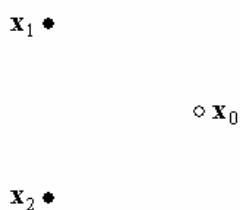
- si $I(Y_1; y) = 0$ et $I(Y_2; y) = 1$:

$$I(Y_0; y)^{EC} = \frac{\text{Prob}[Y_1 > y, Y_2 < y, Y_0 < y]}{\text{Prob}[Y_1 > y, Y_2 < y]} = \frac{E\{[1 - I(Y_1; y)]I(Y_2; y)I(Y_0; y)\}}{E\{[1 - I(Y_1; y)]I(Y_2; y)\}}$$

- si $I(Y_1; y) = I(Y_2; y) = 0$:

$$I(Y_0; y)^{EC} = \frac{\text{Prob}[Y_1 > y, Y_2 > y, Y_0 < y]}{\text{Prob}[Y_1 > y, Y_2 > y]} = \frac{E\{[1 - I(Y_1; y)][1 - I(Y_2; y)]I(Y_0; y)\}}{E\{[1 - I(Y_1; y)][1 - I(Y_2; y)]\}}$$

A) Triangle équilatéral



Le calcul de la probabilité que les trois gaussiennes soient simultanément inférieures au seuil y est facilité en considérant la situation suivante. Si T, U, V, W sont quatre gaussiennes réduites et indépendantes entre elles et ρ un réel compris entre 0 et 1, alors les variables aléatoires définies par

$$\begin{cases} Y_0 = \sqrt{\rho} T + \sqrt{1-\rho} U \\ Y_1 = \sqrt{\rho} T + \sqrt{1-\rho} V \\ Y_2 = \sqrt{\rho} T + \sqrt{1-\rho} W \end{cases}$$

sont trigaussiennes, de coefficients de corrélation $\rho_{12} = \rho_{10} = \rho_{20} = \rho$.

En conditionnant par la valeur de T, on obtient la probabilité cherchée :

$$\begin{aligned}
 E[I(Y_1; y)I(Y_2; y)I(Y_0; y)] &= \text{Prob}[Y_1 < y, Y_2 < y, Y_0 < y] \\
 &= \int_{-\infty}^{+\infty} g(t) \text{Prob}\left[U < \frac{y - \sqrt{\rho} t}{\sqrt{1-\rho}}, V < \frac{y - \sqrt{\rho} t}{\sqrt{1-\rho}}, W < \frac{y - \sqrt{\rho} t}{\sqrt{1-\rho}}\right] dt \\
 &= \int_{-\infty}^{+\infty} g(t) \left[G\left(\frac{y - \sqrt{\rho} t}{\sqrt{1-\rho}}\right)\right]^3 dt
 \end{aligned}$$

Une intégration numérique fournit la valeur de l'intégrale avec un degré d'approximation très satisfaisant.

Pour être les plus complets possibles, les résultats seront visualisés en fonction de la corrélation ρ_{12} entre les données gaussiennes et du seuil de troncature y . On observe que le krigeage d'indicatrice coïncide avec l'espérance conditionnelle lorsque le seuil de coupure est nul (cas d'auto-dualité). L'écart absolu est faible lorsque la corrélation tend vers 0 (valeurs indépendantes) ou vers 1 (valeurs presque sûrement égales) et augmente dans les situations les moins fréquentes : données inférieures à un seuil très bas ou supérieures à un seuil très élevé, données fortement corrélées mais situées de part et d'autre du seuil de troncature. On note enfin que l'écart entre les deux estimateurs dépend davantage du seuil de troncature que du niveau de corrélation entre données (du moins, tant que ce dernier ne tend pas vers 0 ou 1).

Par rapport au cas multigaussien, les écarts entre krigeage d'indicatrice et espérance conditionnelle sont ici beaucoup plus faibles (au plus, de 1.5%, au lieu de 50% dans le cas multigaussien) : s'agissant d'un processus binaire, le codage en indicatrice n'entraîne pas de perte d'information et se révèle "presque" optimal.

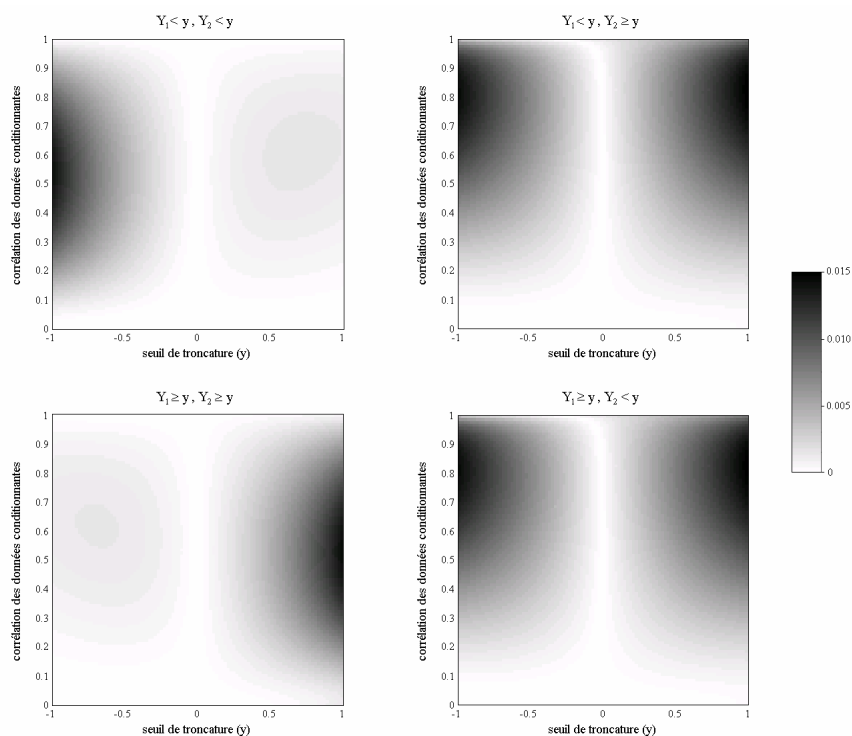
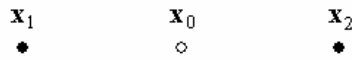


Figure 1.24. Ecart absolu entre krigeage d'indicatrice et espérance conditionnelle en fonction du coefficient de corrélation entre données conditionnantes et du seuil de troncature

B) Sites alignés : x_0 milieu de $[x_1, x_2]$



On suppose que la covariance de la fonction aléatoire gaussienne décroît de façon exponentielle, ce qui implique $\rho_{12} = \rho_{10}^2 = \rho_{20}^2$. Cette fois, les valeurs trigaussiennes peuvent s'écrire sous la forme

$$\begin{cases} Y_0 = T \\ Y_1 = \sqrt{\rho} T + \sqrt{1-\rho} V \\ Y_2 = \sqrt{\rho} T + \sqrt{1-\rho} W \end{cases}$$

avec T, V, W gaussiennes réduites et indépendantes entre elles et $\rho = \rho_{12} \in [0,1]$.

La probabilité trivariante s'exprime comme suit :

$$\begin{aligned} E[I(Y_1; y)I(Y_2; y)I(Y_0; y)] &= \int_{-\infty}^y g(t) \text{Prob}\left[V < \frac{y - \sqrt{\rho} t}{\sqrt{1-\rho}}, W < \frac{y - \sqrt{\rho} t}{\sqrt{1-\rho}}\right] dt \\ &= \int_{-\infty}^y g(t) \left[G\left(\frac{y - \sqrt{\rho} t}{\sqrt{1-\rho}}\right)\right]^2 dt \end{aligned}$$

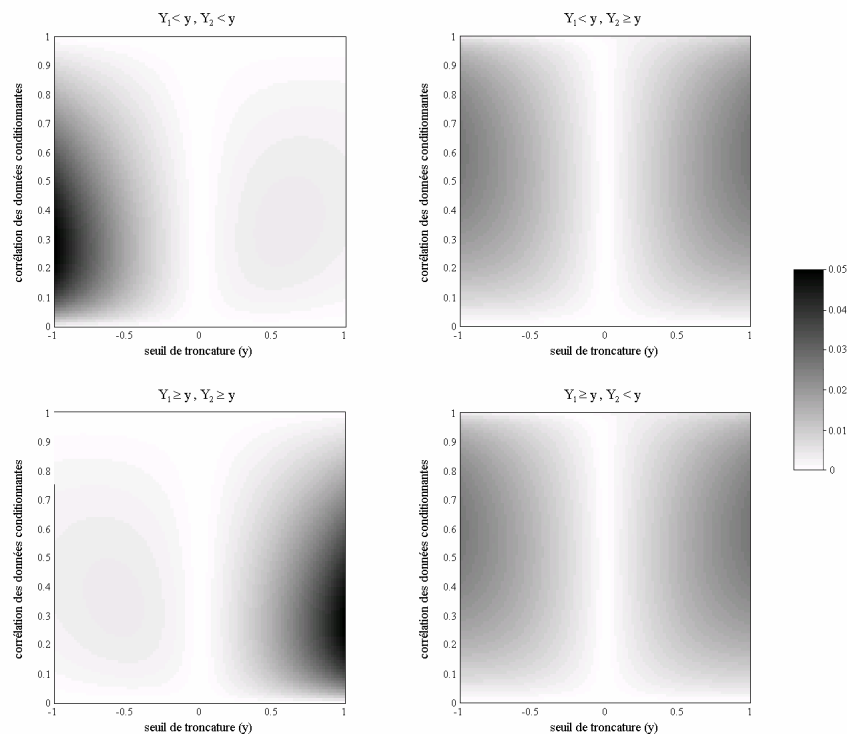
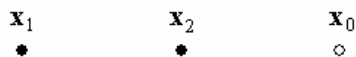


Figure 1.25. *Ecart absolu entre krigeage d'indicatrice et espérance conditionnelle en fonction du coefficient de corrélation entre données conditionnantes et du seuil de troncature*

Les conclusions sont qualitativement identiques à celles de la configuration en triangle équilatéral. Toutefois, l'amplitude de l'écart absolu entre krigeage d'indicatrice et espérance conditionnelle est plus importante (jusqu'à 5% pour les niveaux de corrélation et les seuils de troncature considérés).

C) Sites alignés : x_2 milieu de $[x_0, x_1]$



Si $\rho = \rho_{12}$ et si la covariance décroît de manière exponentielle, la probabilité trivariée admet une expression analogue au cas précédent :

$$E [I(Y_1 ; y)I(Y_2 ; y)I(Y_0 ; y)] = \int_{-\infty}^y g(t) \left[G\left(\frac{y - \rho t}{\sqrt{1 - \rho^2}}\right) \right]^2 dt .$$

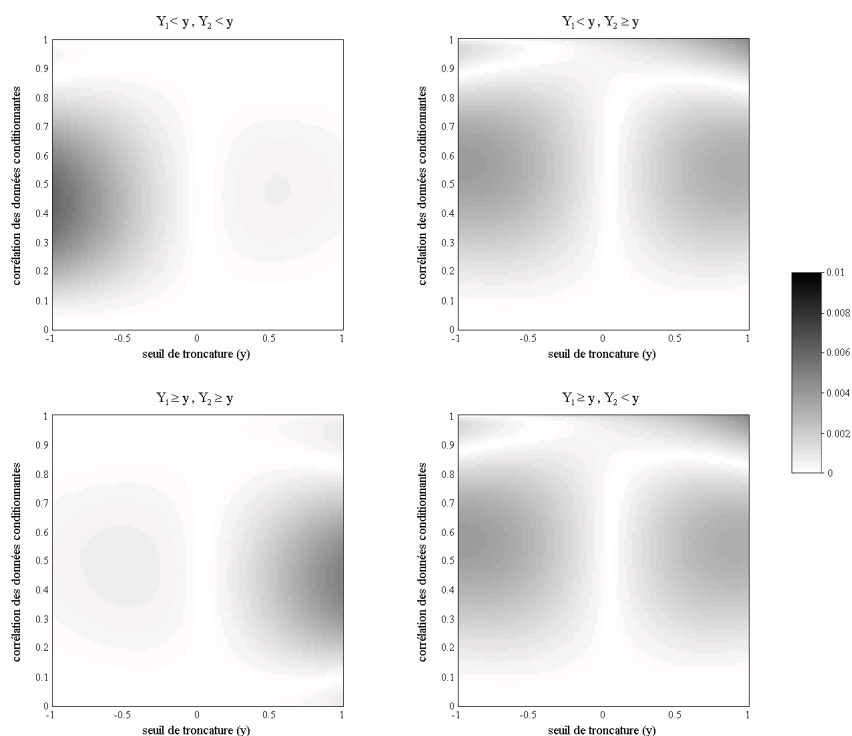


Figure 1.26. *Ecart absolu entre krigeage d'indicateur et espérance conditionnelle en fonction du coefficient de corrélation entre données conditionnantes et du seuil de troncature*

Ici, les écarts entre estimateurs sont bien moindres que dans le cas précédent : la cause est l'effet d'écran exercé par la donnée Y_2 sur la donnée Y_1 , laquelle joue peu de rôle dans les estimations.

VI. Discussion

Les modèles et configurations analysés établissent que le krigeage d'une indicatrice ne coïncide presque jamais avec son espérance conditionnelle. En effet, dans l'espace à plusieurs dimensions, il peut fournir des estimations négatives car il est toujours possible d'obtenir des pondérateurs de krigeage négatifs. La seule exception qui ait été vue est celle de la mosaïque de Poisson avec des sites alignés. Quant au krigeage disjonctif, une autre situation est apparue où l'estimation coïncide exactement avec une espérance conditionnelle : celle des modèles isofactoriels markoviens (bigaussien ou bigamma) dans l'espace à une dimension. Ainsi, ce tour d'horizon laisse entrevoir que ce n'est que dans l'espace à une dimension et pour ces familles de modèles que l'on peut espérer obtenir une simulation exacte. Nous reviendrons sur ce sujet au chapitre suivant.

Suivant Matheron (1989b), si l'on différencie les modèles isofactoriels en modèles de *diffusion* et modèles de type *mosaïque*, on peut avancer que le krigeage disjonctif s'éloigne moins de l'espérance conditionnelle dans les modèles de diffusion que dans le modèle mosaïque, tandis que le krigeage d'indicatrices est au contraire plus proche de l'espérance conditionnelle dans le cas mosaïque que dans les modèles de diffusion. Dans tous les cas, le krigeage disjonctif est toujours meilleur que le krigeage d'indicatrices ; l'amélioration est spectaculaire dans les modèles de diffusion. En termes d'imagerie stochastique, cela signifie que l'algorithme séquentiel isofactoriel est bien plus précis que l'algorithme séquentiel d'indicatrices, surtout pour simuler des processus de diffusion.

Les facteurs expliquant l'écart entre krigeage d'indicatrice et espérance conditionnelle sont :

- 1) La proximité des valeurs des données par rapport au seuil à estimer : le krigeage d'indicatrice suppose implicitement que la valeur de chaque donnée est "moyennement" éloignée du seuil à estimer (d'un côté ou de l'autre), ce qui conduit à des écarts importants par rapport à l'espérance conditionnelle lorsque cette valeur est très proche de ce seuil ; la codification de la donnée en indicatrice est alors trompeuse.
- 2) L'effet d'écran (partiel ou total) qui se produit lorsque les données sont alignées. Cet effet n'est pas systématique, même lorsque la fonction de covariance est exponentielle (*cf.* le cas de la mosaïque des feuilles mortes) : l'argument selon lequel les données proches écrantent les données éloignées, invoqué pour limiter la taille des voisinages de krigeage, n'est donc pas toujours justifié.
- 3) Les corrélations entre valeurs : dans les modèles de diffusion, plus elles sont élevées, plus le krigeage d'indicatrice s'écarte de l'espérance conditionnelle. La modélisation des lois de diffusion par une covariance d'indicatrice apparaît comme très incomplète et les résultats obtenus ne sont acceptables que lorsque l'on se rapproche de l'indépendance bivariable.
- 4) Les liens multivariés entre les valeurs (par exemple, les relations de connexité ou de contiguïté), invisibles dans la covariance de l'indicatrice : plus ces relations réduisent l'incertitude portant sur la valeur non mesurée, plus l'estimation optimale par espérance conditionnelle s'écarte de celle donnée par le krigeage d'indicatrice ; des exemples sont fournis par le processus en damier ou par la mosaïque poissonnienne où l'on cherche à simuler la valeur située au milieu de deux données égales.

Le krigeage disjonctif coïncide avec le krigeage d'indicatrice dans le cadre mosaïque mais apporte de nettes améliorations dans le cadre de processus de diffusion (gaussien ou gamma). En un sens, il dépend moins du caractère "mosaïque" ou "diffusif" de la régionalisation. En outre, la précision de ses résultats est peu sensible à la proximité des valeurs des données par rapport au seuil à estimer : l'explication vient du fait que le krigeage disjonctif évite l'écueil de codifier chaque donnée en une valeur binaire, ce qui réduit considérablement la quantité d'information disponible. En revanche, les liens multivariés entre valeurs restent ignorés par le krigeage disjonctif tout comme le krigeage d'indicatrice.

Chapitre 2 :

Modèles de fonctions aléatoires associés à l'algorithme séquentiel

I. Objectif

Dans le chapitre précédent, nous avons comparé les performances du krigeage disjonctif et du krigeage d'indicateur, qui interviennent dans les algorithmes de simulation séquentielle (isofactorielle et d'indicateurs) via l'estimation des lois de probabilité conditionnelles. Ce chapitre s'intéresse à plusieurs modèles de fonctions aléatoires et cherche les conditions sous lesquelles les algorithmes séquentiels restituent exactement les caractéristiques du modèle qui ont été spécifiées. On explicitera alors les lois bivariées et multivariées obtenues et dans quelle mesure la simulation conditionnelle de telles lois est possible. Les modèles étudiés sont les suivants :

- *modèle mosaïque*, dans lequel les covariances simples et croisées de toutes les indicateurs sont proportionnelles entre elles ;
- *modèle d'ensemble aléatoire*, qui peut être vu comme un cas particulier du modèle mosaïque, pour lequel les valeurs sont binaires ;
- *modèles "multi-indicateurs"*, où l'on spécifie les espérances et covariances simples de plusieurs indicateurs ;
- *modèles isofactoriels*, où l'on spécifie entièrement les lois marginales et bivariées du processus (espérances et covariances simples et croisées de toutes les indicateurs).

II. Modèle mosaïque à valeurs continues

Dans cette section, nous allons examiner en détail la simulation d'un modèle mosaïque par la méthode séquentielle d'indicateurs. Dans ce modèle, les indicateurs sont en *corrélation intrinsèque*, c'est-à-dire que leurs covariances simples et croisées sont proportionnelles à un même corrélogramme $\rho(\mathbf{h})$, qui caractérise la loi bivariée de la fonction aléatoire mosaïque. Les valeurs en deux sites $\mathbf{x} + \mathbf{h}$ et \mathbf{x} sont égales avec la probabilité $\rho(\mathbf{h})$ et indépendantes avec la probabilité complémentaire. L'espace est partitionné en compartiments au sein desquels le processus est constant et sa valeur indépendante de celles des autres compartiments (Matheron, 1982 ; Rivoirard, 1994).

Dans la suite de cette section, on supposera que la loi marginale est absolument continue. Cette hypothèse simplifie grandement l'étude du modèle, puisque l'appartenance de plusieurs sites à un même compartiment de la mosaïque est synonyme d'égalité de leurs valeurs (en effet, deux valeurs indépendantes prennent presque sûrement des valeurs différentes).

II.1. Condition pour reproduire les lois marginale et bivariable

Dans l'espace unidimensionnel, avec une covariance à décroissance exponentielle, une simulation séquentielle dans un ordre régulier fournit une mosaïque de Poisson, mais ce n'est plus le cas si la simulation s'effectue dans un ordre irrégulier : on obtient alors une mosaïque à cellules non convexes (cf. chapitre 1). En somme, **la loi spatiale dépend de l'ordre selon lequel est réalisée la simulation**. Pourtant, un calcul simple montre que la covariance entre deux valeurs simulées quelconques est bien une exponentielle décroissante.

Cette remarque n'est du reste qu'un cas particulier d'un résultat plus général qui s'énonce comme suit : à condition que les poids du krigeage d'indicatrice soient toujours positifs et de somme inférieure à l'unité, les valeurs simulées reproduisent correctement les lois marginale et bivariable du modèle.

La démonstration s'effectue par récurrence sur le nombre total de sites simulés. Supposons que l'on dispose de valeurs $\{Y_1, \dots, Y_n\}$ aux sites $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ telles que

$$\begin{aligned} \forall i \in \{1, \dots, n\}, \text{Prob}(Y_i < y) &= F(y) \\ \forall i, j \in \{1, \dots, n\}, \text{cov}(Y_i, Y_j) &= \rho(\mathbf{x}_i - \mathbf{x}_j) \end{aligned}$$

Construisons la fonction de répartition conditionnelle d'un nouveau site \mathbf{x}_{n+1} par krigeage d'indicatrice. A la valeur prise par la donnée Y_i , cette fonction présente un saut d'amplitude égale au poids de krigeage λ_i de cette donnée (supposé positif) ; le poids λ_0 de la moyenne, complémentaire au poids cumulé des données, est attribué à la fonction de répartition *a priori* $F(y)$. Cela signifie que la valeur simulée Y_{n+1} est égale à Y_i avec la probabilité λ_i ou indépendante de $\{Y_1, \dots, Y_n\}$ avec la probabilité λ_0 : le caractère mosaïque des lois bivariées est donc bien respecté. Il reste à vérifier que la loi marginale et la covariance sont celles attendues :

$$\begin{aligned} \text{Prob}(Y_{n+1} < y) &= \sum_{i=1}^n \lambda_i \text{Prob}(Y_i < y) + \lambda_0 F(y) = F(y) \\ \forall j \in \{1, \dots, n\}, \text{cov}(Y_{n+1}, Y_j) &= \sum_{i=1}^n \lambda_i \text{cov}(Y_i, Y_j) = \sum_{i=1}^n \lambda_i \rho(\mathbf{x}_i - \mathbf{x}_j) = \rho(\mathbf{x}_{n+1} - \mathbf{x}_j) \end{aligned}$$

La dernière égalité provient des équations de krigeage simple. Ceci établit le résultat annoncé.

Peu de covariances stationnaires garantissent que les poids de krigeage simple soient toujours positifs et de somme inférieure à l'unité ; hormis l'effet de pépité pur, c'est seulement le cas à une dimension des **covariances "complètement monotones"**¹ sur \mathbf{R}_+ (exponentielles gigognes). Matheron (1986) a démontré que de telles covariances sont du type "feuilles mortes" et correspondent à des modèles autorisés pour décrire des indicatrices :

$$\rho(\mathbf{h}) = \frac{K(\mathbf{h})}{2K(\mathbf{0}) - K(\mathbf{h})}$$

où $K(\mathbf{h})$ est le covariogramme géométrique d'un ensemble aléatoire.

¹ Une fonction $f(x)$ est complètement monotone si le signe de ses dérivées successives alterne : $\forall p \in \mathbf{N}, (-1)^p f^{(p)}(x) \geq 0$. Un théorème dû à Bernstein (Feller, 1966) établit que les fonctions complètement monotones sur \mathbf{R}_+ sont proportionnelles à la transformée de Laplace d'une loi de probabilité et peuvent donc s'écrire sous la forme d'un mélange de covariances exponentielles.

La simulation séquentielle d'indicatrices fournit donc un procédé pour construire des réalisations exactes de mosaïques unidimensionnelles à covariances complètement monotones. Inversement, les modèles variographiques théoriquement interdits pour décrire des indicatrices ne peuvent faire partie des covariances donnant toujours des poids de krigeage positifs.

Remarque

Utiliser un krigeage ordinaire d'indicatrice n'assure plus la reproduction de la covariance ni même le caractère mosaïque de la loi bivariable : l'algorithme ne peut plus générer des valeurs indépendantes de celles préexistantes, car le poids de la moyenne est nul. En outre, la covariance simulée ne sera pas stationnaire et s'écartera de plus en plus du modèle théorique à mesure que se déroulera la simulation. C'est pourquoi tous les krigeages considérés ici seront des krigeages simples, à moyenne connue.

II.2. Loi de probabilité multivariable des valeurs simulées

Soient $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ les sites à simuler. L'ordre selon lequel ces sites sont traités est représenté par une permutation $\sigma = \{i_1, \dots, i_n\}$ de l'ensemble $\{1, \dots, n\}$. Par application récursive de la formule de Bayes, la densité de probabilité multivariable des valeurs simulées $\{Y_1, \dots, Y_n\}$ se factorise comme suit :

$$f_{\sigma, n}(y_1, \dots, y_n) = \prod_{k=1}^n \left[\sum_{\ell < k} \lambda_{\ell, k}^{(\sigma)} \delta_{y_{i_k} = y_{i_\ell}} + (1 - \sum_{\ell < k} \lambda_{\ell, k}^{(\sigma)}) f(y_{i_k}) \right] \quad (2.1)$$

où $f(y)$ est la densité marginale des valeurs simulées

$\{\lambda_{\ell, k}^{(\sigma)}, \ell < k\}$ sont les poids de krigeage du site \mathbf{x}_{i_k} à partir des sites $\{\mathbf{x}_{i_\ell}, \ell < k\}$.

L'intégrale de $f_{\sigma, n}(y_1, \dots, y_n)$ sur \mathbb{R}^n vaut 1 (on le voit en intégrant successivement par rapport à y_{i_n}, \dots, y_{i_1}). Si $f(y)$ est absolument continue, la positivité de $f_{\sigma, n}(y_1, \dots, y_n)$ est assurée *si et seulement si* tous les poids de krigeage sont positifs et de somme inférieure à 1 : dans ce cas, il s'agit bien d'une densité de probabilité multivariable.

Les densités de probabilité à m variables ($m < n$) sont définies par intégration par rapport aux $n - m$ indices restants ; en général, elles ne sont pas invariantes par translation (car elles dépendent de configurations de krigeage qui ne se répètent pas dans l'espace), sauf les lois bivariées qui sont de type mosaïque et associées à la covariance stationnaire $\rho(\mathbf{h})$:

$$\begin{aligned} \forall i, j \in \{1, \dots, n\}, f(y_i, y_j) &= \int f_{\sigma, n}(y_1, \dots, y_n) dy_1 \dots dy_{i-1} dy_{i+1} \dots dy_{j-1} dy_{j+1} \dots dy_n \\ &= \rho(\mathbf{x}_i - \mathbf{x}_j) \delta_{y_i = y_j} f(y_i) + [1 - \rho(\mathbf{x}_i - \mathbf{x}_j)] f(y_i) f(y_j) \end{aligned}$$

Contrairement aux densités multivariées, ces densités bivariées sont indépendantes de n et σ (nombre et ordre des sites simulés). Par suite, il est possible de traiter les sites avec un ordre différent d'une simulation à l'autre, sans modifier les lois bivariées simulées. On préfère alors considérer que l'ordre de traitement des sites est aléatoire (Isaaks, 1990), choisi au hasard "pur" parmi toutes les permutations de $\{1, \dots, n\}$. La densité multivariable devient :

$$f_n(y_1 \dots y_n) = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} f_{\sigma, n}(y_1 \dots y_n) \quad (2.2)$$

où \mathfrak{S}_n désigne l'ensemble des permutations de $\{1, \dots, n\}$.

II.3. Conditionnement à un ensemble de données

Pour conditionner les simulations à un ensemble de données, l'approche la plus fréquente consiste à considérer les sites de données comme des sites *déjà simulés* (Journel, 1989 ; Goovaerts, 1997). Bien que d'apparence simple, cette opération revient à imposer le début de la séquence de simulation, ce qui empêche d'atteindre certaines lois multivariées (seul un sous-ensemble de \mathfrak{S}_n est utilisé).

Par ailleurs, les lois multivariées *a priori* qui restent accessibles dépendent de la localisation des échantillons, ce qui laisse croire que les propriétés *a priori* de la fonction aléatoire sont altérées parce que celle-ci a été échantillonnée ! Une telle approche est fort discutable.

Par exemple, soient trois valeurs $\{Y_1, Y_2, Y_3\}$ placées en des sites $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ alignés et régulièrement espacés. Supposons que le modèle soit une mosaïque à covariance exponentielle et cherchons à quantifier la probabilité que \mathbf{x}_2 interrompe une cellule contenant \mathbf{x}_1 et \mathbf{x}_3 , c'est-à-dire que Y_1 et Y_3 soient égales et différentes de Y_2 . Les calculs se simplifient grâce à l'effet d'écran induit par la covariance exponentielle (Rivoirard, 1984).

Si aucun échantillon n'est effectué, la probabilité cherchée est strictement positive (tableau 2.1) :

$$\text{Prob}[Y_1 = Y_3 \neq Y_2] = \frac{\rho^2 (1 - \rho)^2}{3(1 + \rho^2)} \text{ où } \rho \text{ est le coefficient de corrélation entre deux sites consécutifs.}$$

séquence de simulation	Prob [$Y_1 = Y_3 \neq Y_2$]	probabilité moyenne
$\mathbf{x}_1 - \mathbf{x}_2 - \mathbf{x}_3$ $\mathbf{x}_2 - \mathbf{x}_1 - \mathbf{x}_3$ $\mathbf{x}_2 - \mathbf{x}_3 - \mathbf{x}_1$ $\mathbf{x}_3 - \mathbf{x}_2 - \mathbf{x}_1$	0	$\frac{\rho^2 (1 - \rho)^2}{3(1 + \rho^2)}$
$\mathbf{x}_1 - \mathbf{x}_3 - \mathbf{x}_2$ $\mathbf{x}_3 - \mathbf{x}_1 - \mathbf{x}_2$	$\rho^2 [1 - \frac{2\rho}{1 + \rho^2}]$	

Tableau 2.1. Distribution *a priori* de $\{Y_1, Y_2, Y_3\}$

Si à présent \mathbf{x}_2 est échantillonné, les seules séquences considérées seront celles commençant par ce site, de sorte que la probabilité devient nulle, quelle que soit la valeur de l'échantillon. Autrement dit, le simple fait de savoir où se trouve un échantillon (sans connaître sa valeur) modifie les statistiques *a priori* de la simulation.

Pour contourner ces difficultés, une solution consiste à séparer le problème de la simulation d'une mosaïque en deux étapes : d'abord la détermination des compartiments, puis leur valuation selon la loi marginale souhaitée, indépendamment d'un compartiment à l'autre. Seule la première étape pose réellement problème pour la simulation conditionnelle, les données introduisant des contraintes sur la géométrie des compartiments.

Pour simplifier la description, on symbolise l'appartenance de deux sites à un même compartiment par une connexion entre ces sites : un compartiment de la mosaïque est alors constitué par l'ensemble des sites connectés entre eux, qui forment une classe d'équivalence.

La séquence de simulation, que l'on fixe une fois pour toutes, induit un ordre chronologique entre les sites. D'après le paragraphe II.1, la valeur simulée en un site est égale à celle d'un site plus ancien avec une probabilité égale au poids de krigeage associé, ou indépendante avec une probabilité égale au poids de la moyenne. Ceci suggère l'algorithme d'acceptation et de rejet suivant pour mener à bien la simulation conditionnelle.

1) Initialisation

Trouver un ensemble de connexions respectant les conditions sur les données (i.e. tel que deux sites de même valeur appartiennent à la même classe d'équivalence et que deux sites de valeurs différentes ne soient pas connectés entre eux).

2) Phase itérative

On choisit un site \mathbf{x} au hasard parmi ceux à simuler. Soient $\{\mathbf{x}_i, i = 1 \dots v\}$ les sites plus anciens que \mathbf{x} , $\{\lambda_i, i = 1 \dots v\}$ leurs poids de krigeage et λ_0 le poids de la moyenne. On simule une variable U uniforme sur $[0,1]$. Si $U \leq \lambda_0$, on propose d'effacer toute connexion entre \mathbf{x} et les sites plus anciens. Sinon, on propose de connecter \mathbf{x} au site \mathbf{x}_i tel que

$$\sum_{j=0}^{i-1} \lambda_j \leq U \leq \sum_{j=0}^i \lambda_j .$$

La proposition est acceptée si elle ne contredit pas les conditions sur les données.

Exemple : simulation unidimensionnelle en ordre régulier avec un schéma exponentiel

Parmi les sites plus anciens que le site à simuler, seul compte le site immédiatement antérieur, à cause de l'effet d'écran de la covariance exponentielle. Par suite, un changement de compartiment entre ces sites est une alternative de Bernoulli, dont la probabilité de succès est $\lambda_0 = 1 - \exp(-a\delta)$ où δ est la distance entre deux sites consécutifs et a un paramètre positif lié à la portée de la covariance. L'application de l'algorithme d'acceptation et de rejet est donc ici très simple.

Cette description permet d'ailleurs de retrouver la définition de la mosaïque poissonnienne à partir d'un processus ponctuel de Poisson. En effet, soit un intervalle de longueur $l = m\delta$ avec m entier. Le nombre N de changements de compartiments sur cet intervalle est la somme de m alternatives de Bernoulli indépendantes entre elles ; c'est une loi binomiale de paramètres m et $\eta = 1 - \exp(-a\delta)$. A présent, faisons tendre vers 0 la distance entre deux sites consécutifs tout en maintenant constante la longueur l : m tend vers l'infini, η tend vers 0 mais $m\eta$ tend vers $\theta = l a$. Par suite, la loi de N tend vers une loi de Poisson de paramètre θ , directement proportionnel à la longueur de l'intervalle. On retrouve une définition de la mosaïque poissonnienne en continu, où les compartiments sont délimités par un processus ponctuel de Poisson (Alfaro, 1979 ; Lantuéjoul, 2002).

La mosaïque poissonnienne à une dimension constitue en fait un cas particulier, car elle bénéficie d'une propriété markovienne qui permet de mener directement la simulation conditionnelle, à partir de la connaissance des lois de probabilité trivariées, sans passer par une procédure itérative (cf. section ultérieure sur les modèles isofactoriels).

II.4. Cohérence des lois multivariées

On se pose à présent la question de savoir si la loi multivariée [Eq. (2.2)] correspond à un modèle de fonction aléatoire. Le problème peut se formuler de la façon suivante : si l'on décide de simuler $n+1$ valeurs au lieu de n , obtiendra-t-on la même loi de probabilité conjointe pour les n valeurs communes, c'est-à-dire la loi des valeurs simulées est-elle indépendante du niveau de "résolution" et de la taille du domaine de simulation choisis par l'utilisateur ?

La réponse à cette question est négative. En général², on aura :

$$f_{n+1}(y_1, \dots, y_n) = \int f_{n+1}(y_1, \dots, y_n, y_{n+1}) dy_{n+1} \neq f_n(y_1, \dots, y_n).$$

Un exemple illustrant cette inégalité est donné ci-après.

Considérons quatre sites alignés et régulièrement répartis $\{x_1, x_2, x_3, x_4\}$, dont on simule les valeurs à partir d'une covariance exponentielle. On cherche à comparer la loi de $\{Y_1, Y_2, Y_3\}$, selon que l'on simule ou non la valeur Y_4 . Cette loi trivariée est caractérisée par la probabilité que Y_1 et Y_3 soient égales et Y_2 différente (probabilité que x_2 coupe un compartiment contenant x_1 et x_3). On note ρ le coefficient de corrélation entre deux sites consécutifs.

séquence de simulation	Prob[$Y_1 = Y_3 \neq Y_2$]	probabilité moyenne
$x_1 - x_2 - x_3 - x_4$ $x_1 - x_2 - x_4 - x_3$ $x_2 - x_1 - x_3 - x_4$ $x_2 - x_1 - x_4 - x_3$ $x_2 - x_3 - x_1 - x_4$ $x_2 - x_3 - x_4 - x_1$ $x_2 - x_4 - x_1 - x_3$ $x_2 - x_4 - x_3 - x_1$ $x_3 - x_2 - x_1 - x_4$ $x_3 - x_2 - x_4 - x_1$ $x_3 - x_4 - x_2 - x_1$ $x_4 - x_2 - x_1 - x_3$ $x_4 - x_2 - x_3 - x_1$ $x_4 - x_3 - x_2 - x_1$	0	$\frac{\rho^2 (1-\rho)^2 (2+\rho^2 + 2\rho^3)}{6(1+\rho^2)(1+\rho^3)}$
$x_1 - x_3 - x_2 - x_4$ $x_1 - x_3 - x_4 - x_2$ $x_3 - x_1 - x_2 - x_4$ $x_3 - x_1 - x_4 - x_2$ $x_3 - x_4 - x_1 - x_2$ $x_4 - x_3 - x_1 - x_2$	$\frac{\rho^2 (1-\rho)^2}{1+\rho^2}$	
$x_1 - x_4 - x_2 - x_3$ $x_4 - x_1 - x_2 - x_3$	$\frac{\rho^4 (1-\rho)^2 (1+\rho)}{(1+\rho^2)(1+\rho^3)}$	
$x_1 - x_4 - x_3 - x_2$ $x_4 - x_1 - x_3 - x_2$	$\frac{\rho^2 (1-\rho)^2}{1+\rho^3}$	

Tableau 2.2. Loi de $\{Y_1, Y_2, Y_3\}$ lorsque l'on simule Y_4

² Il existe quelques cas où la simulation d'une valeur supplémentaire ne modifie pas la loi des valeurs initiales :

- cas où l'on ne considère que deux valeurs :

$$\forall n \geq 2, \forall i, j \leq n, f_n(y_i, y_j) = f_2(y_i, y_j)$$

- cas où l'on tient compte de l'ordre des sites simulés et où le site supplémentaire est ajouté en dernier ou en avant-dernier :

$$\forall \sigma = \{i_1, \dots, i_n, i_{n+1}\} \in \mathfrak{S}_{n+1}, f_{\sigma, n+1}(y_{i_1}, \dots, y_{i_n}, y_{i_{n+1}}) = f_{\sigma', n}(y_{i_1}, \dots, y_{i_{n-1}}, y_{i_n}) \text{ où } \sigma' = \{i_1, \dots, i_{n-1}, i_n\}$$

$$\forall \sigma = \{i_1, \dots, i_n, i_{n+1}\} \in \mathfrak{S}_{n+1}, f_{\sigma, n+1}(y_{i_1}, \dots, y_{i_{n-1}}, y_{i_{n+1}}) = f_{\sigma', n}(y_{i_1}, \dots, y_{i_{n-1}}, y_{i_{n+1}}) \text{ où } \sigma' = \{i_1, \dots, i_{n-1}, i_{n+1}\}$$

La dernière égalité découle des relations entre pondérateurs de krigeage (Emery, 2003) :

$$\forall \ell < k < n, \lambda_{\ell, k}^{(\sigma)} = \lambda_{\ell, k}^{(\sigma')} \text{ et } \forall \ell < n, \lambda_{\ell, n+1}^{(\sigma)} = \lambda_{\ell, n+1}^{(\sigma')} + \lambda_{\ell, n}^{(\sigma)} \lambda_{n, n+1}^{(\sigma')}.$$

Ce tableau est à comparer avec le tableau 2.1, qui donne la même probabilité lorsque x_4 n'est pas considéré dans la séquence de simulation. En conclusion, la probabilité cherchée change dès que l'on simule une valeur supplémentaire. La différence n'est pas négligeable, surtout lorsque le coefficient de corrélation ρ est compris entre 0.5 et 0.75 (figure 2.1).

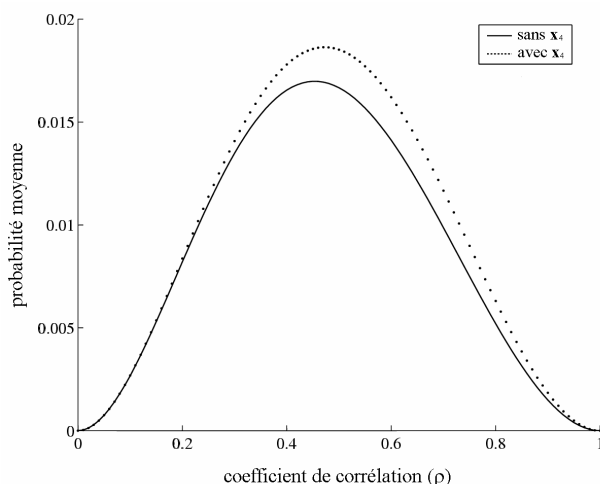


Figure 2.1. Comparaison des probabilités trivariées selon que l'on simule ou non la valeur Y_4 , en fonction du coefficient de corrélation ρ

Le fait que les lois de probabilité dépendent du nombre total de sites simulés signifie que le vecteur aléatoire simulé $\{Y_1, \dots, Y_n\}$ ne correspond pas à un modèle sous-jacent de fonction aléatoire $\{Y(\mathbf{x}), \mathbf{x} \in \mathbb{R}^d\}$. Plus exactement, on voit sur l'exemple précédent que plus le nombre de sites simulés est important, plus les séquences possibles sont nombreuses et plus on enrichit l'ensemble des lois multivariées accessibles par l'algorithme séquentiel. En simulant de plus en plus densément pour chercher à s'approcher d'une loi spatiale, on multiplierait à l'infini les lois multivariées accessibles, ce qui suggère qu'il existe une infinité de fonctions aléatoires dont la restriction aux sites simulés coïncide avec le vecteur $\{Y_1, \dots, Y_n\}$.

Conséquences

Hormis les lois bivariées, les lois multivariées dépendent du nombre total de valeurs simulées, donc sont laissées à l'arbitraire de l'utilisateur. Toutefois, deux situations évitent cet inconvénient.

- L'espace ne compte qu'un nombre fini de localisations et on les simule toutes. C'est par exemple le cas lorsque l'on cherche à simuler des blocs de support non ponctuel sur un domaine fixé. Cette situation n'est pas pleinement satisfaisante, car on ne peut augmenter la taille du domaine sans à nouveau modifier le nombre total de localisations.
- On peut ordonner préalablement tous les sites de l'espace, indépendamment des données et des sites à simuler choisis par l'utilisateur. En suivant cet ordre, la poursuite à l'infini de l'algorithme séquentiel permet de simuler l'espace entier et de converger vers un modèle unique de loi spatiale. Cette approche n'est possible qu'à une dimension, par exemple avec une séquence dichotomique telle le "déplacement aléatoire du point médian"³ (Voss, 1985 ; Chilès, 1995). L'ordre régulier est aussi envisageable sous certaines conditions, par exemple dans le cas de la mosaïque de Poisson, laquelle possède la propriété markovienne. Sans cette dernière propriété, il n'est plus sûr que les lois résultant d'une simulation en ordre régulier soient indépendantes du nombre de sites simulés (maille de discrétisation choisie).

³ On simule d'abord les extrémités du domaine à simuler, puis leur milieu, et ainsi de suite : à chaque étape, on simule les milieux des segments définis par les sites simulés aux étapes précédentes. Tout site du domaine à simuler est atteint au bout d'un nombre fini ou dénombrable d'itérations. Cette technique permet d'affiner une simulation entre deux sites déjà simulés.

A l'exception de ces quelques cas, l'usage des simulations comme modèle d'incertitude ne sera satisfaisant que si l'on fait appel à des statistiques relatives à une ou deux valeurs à la fois (distribution des valeurs en un site, d'où l'on déduit la moyenne ou la probabilité de dépasser un seuil donné...). Cette situation était prévisible, puisque l'algorithme séquentiel n'exige de spécifier que la loi bivariable des valeurs ; il faut donc considérer que **les lois d'ordre supérieur sont indéterminées**. En revanche, il n'est pas recommandé de mettre en jeu des statistiques multivariées, en particulier réaliser un **changement de support** sur les réalisations des simulations, puisque la loi des valeurs régularisées dépendra de la maille de discrétisation (arbitraire) choisie par l'utilisateur.

II.5. Exemple : déplacement aléatoire du point médian

Dans ce paragraphe, nous nous intéressons au processus mosaïque unidimensionnel de covariance exponentielle obtenu par déplacement aléatoire du point médian. Contrairement à la simulation en ordre régulier, cette procédure fournit des mosaïques à compartiments non convexes.

Supposons que l'on cherche à simuler (2^N+1) valeurs en des sites répartis à maille régulière sur l'intervalle $[0,1]$, de coordonnées $\{k/2^N, k = 0 \dots 2^N\}$. Hormis 0, chaque coordonnée peut s'écrire sous forme de fraction irréductible : $k/2^N = k'/2^{N'}$ avec k' entier impair. On conviendra d'appeler *ordre du site* la fraction $1/2^{N'}$ et l'on notera ρ le coefficient de corrélation entre les deux sites extrêmes $\{0,1\}$.

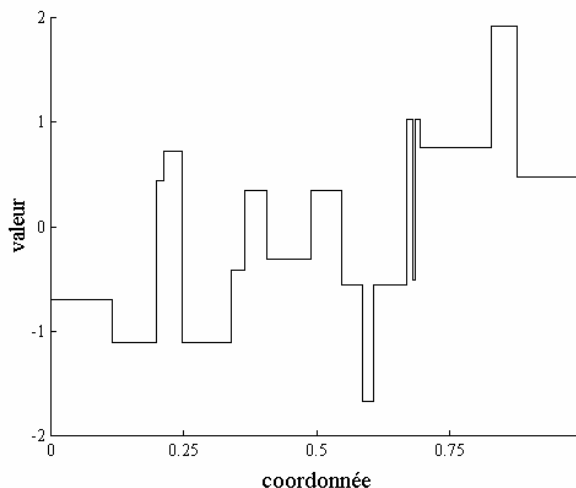


Figure 2.2. Simulation d'une mosaïque à covariance exponentielle et loi marginale gaussienne par déplacement aléatoire du point médian

Nous allons donner une interprétation de ce processus par une construction voisine de celle du modèle des feuilles mortes colorées (Jeulin, 1980, 1997 ; Lantuéjoul, 2002).

- **Initialisation** : simulation des sites extrêmes

On recouvre le segment $[0,1]$ par une "feuille" valuée selon la loi marginale souhaitée. Le site de coordonnée 0 est attaché à cette feuille et prend sa valeur ; il devient une barrière empêchant toute feuille future de la traverser (à cause de l'effet d'écran induit par la covariance exponentielle).

Le site de coordonnée 1 est alors simulé. Avec la probabilité ρ , la feuille attachée en 0 est attachée en 1 ; avec la probabilité complémentaire, on attache une feuille nouvelle valuée indépendamment de la feuille précédente et recouvrant l'intervalle $]0,1[$. Le site prend la valeur de la feuille à laquelle il est attaché et devient une barrière pour les futures feuilles.

- Phase récursive : simulation des sites d'ordre $i = 1/2^N$

Pour tout site x d'ordre i , on effectue l'opération suivante. Avec la probabilité $2p^i/(1+p^{2i})$, on attache le site à une feuille choisie au hasard parmi celles déjà présentes en x . Avec la probabilité complémentaire, on attache en x une feuille vannée indépendamment des autres feuilles ; dans ce dernier cas, la feuille nouvelle recouvre l'intervalle situé entre les deux sites adjacents $]-i, x+i[$. Le site x prend la valeur de la feuille à laquelle il est attaché et devient une barrière qui coupe toutes les feuilles passant par ce site. Les bouts de feuilles non attachés tombent et disparaissent.

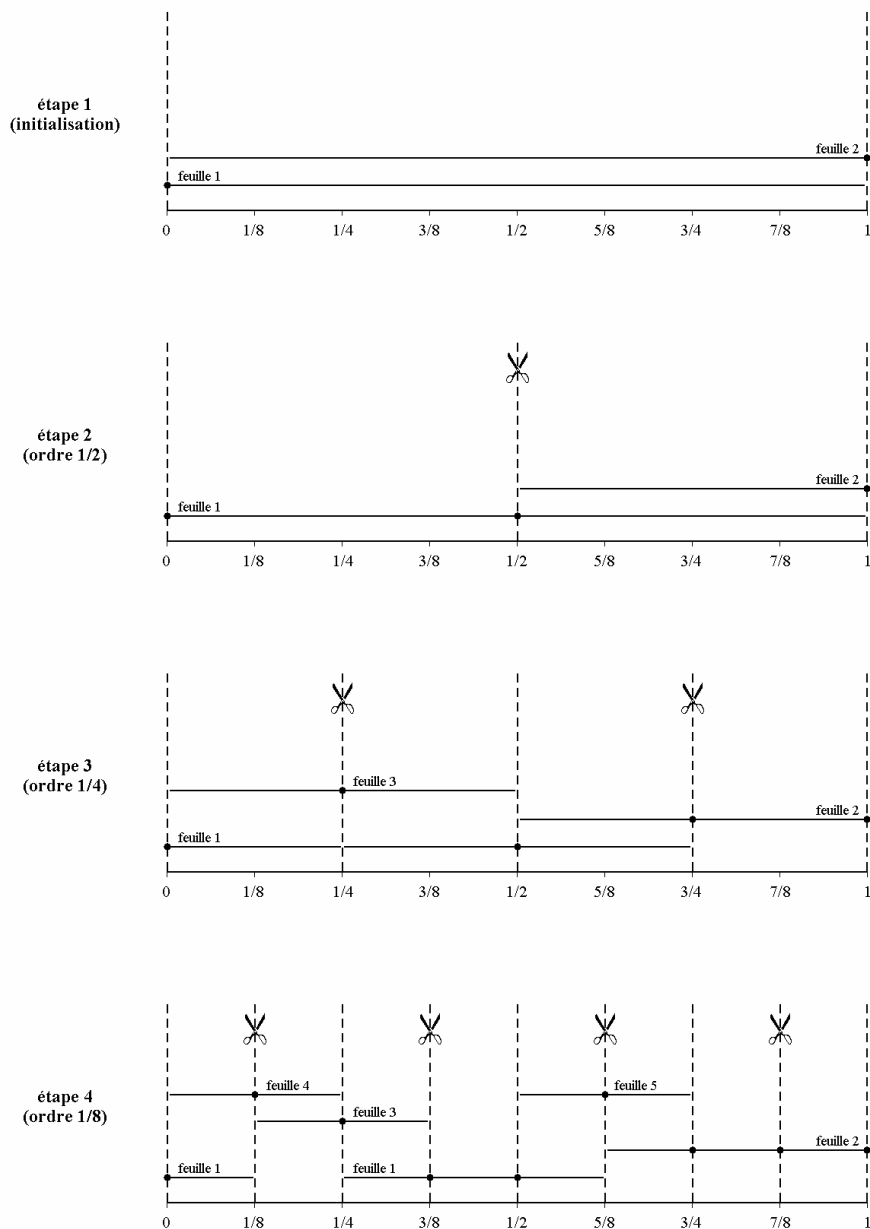


Figure 2.3. *Modèle de recouvrement*

Remarques

- 1) *Convergence de la construction* : les feuilles apparaissant sur les sites d'ordre i couvrent en moyenne une longueur $l(i) = 1 - 2p^i/(1+p^{2i})$, qui converge vers 0 lorsque i diminue ; la vitesse de convergence est proportionnelle à i^2 :

$$l(i) \underset{i \rightarrow 0}{\sim} i^2 \ln^2(\rho) / 2.$$

Il existe un réel i_0 tel que $\forall i \leq i_0, l(i) \leq i$. Le cumul des longueurs pour les ordres inférieurs à i_0 est majoré par :

$$l^+(i_0) = \sum_{i \leq i_0} l(i) \leq \sum_{i \leq i_0} i = 2i_0$$

car les ordres sont en progression géométrique de raison 1/2. En conclusion, étant arrivé à un ordre i , la longueur susceptible d'être recouverte par de futures feuilles est une variable aléatoire positive dont l'espérance tend vers 0 à mesure que i diminue.

- 2) On peut étendre la construction à un ordre irrégulier quelconque de la séquence simulée : la probabilité de choisir une feuille déjà présente est donnée par les poids de krigeage des sites adjacents à \mathbf{x} ; en particulier, les feuilles attachées à ces sites ne sont plus équiprobables. On peut également généraliser la construction à une covariance non exponentielle (pourvu qu'elle soit complètement monotone, ce qui garantit des poids de krigeage toujours positifs) : dans ce cas, les sites simulés ne font plus écran et les feuilles ne sont plus sectionnées ; le choix d'une feuille présente est donné par les poids de krigeage de tous les sites déjà simulés, d'une feuille nouvelle par la probabilité complémentaire.
- 3) En inversant l'échelle de temps, c'est-à-dire en commençant par les sites d'ordres les plus faibles et en remontant vers ceux d'ordres plus élevés, on obtient un modèle équivalent mais dont la construction présente deux avantages. D'une part, on peut simuler l'espace entier (les sites sont traités par ordres croissants et l'ordre peut être aussi élevé qu'on le souhaite) et pas seulement le segment $[0,1]$; on doit alors définir un ordre maximal (ordre à partir duquel le domaine à simuler est entièrement recouvert), ainsi qu'un ordre minimal en dessous duquel on considère que la probabilité d'apparition d'une feuille est nulle. D'autre part, en randomisant la position de l'un des sites d'ordre minimal (pris comme origine), on évite de donner une situation privilégiée aux bords du champ à simuler.

II.6. Extension à plusieurs dimensions

Une fonction aléatoire à loi mosaïque dans \mathbb{R}^d peut être obtenue via le "produit" de d mosaïques unidimensionnelles $\{Y_1, \dots, Y_d\}$: deux sites \mathbf{x} et \mathbf{x}' de \mathbb{R}^d appartiennent au même compartiment si, pour tout $i \in \{1, \dots, d\}$, les sites de \mathbb{R} correspondant aux i èmes coordonnées de \mathbf{x} et \mathbf{x}' appartiennent au même compartiment de Y_i . Une fois définis les compartiments dans l'espace \mathbb{R}^d , leur valuation se fait en fonction de la loi marginale souhaitée, indépendamment d'un compartiment à l'autre.

Le corrélogramme de la mosaïque multidimensionnelle est le produit des corrélogrammes des mosaïques unidimensionnelles, par exemple une exponentielle factorisée si l'on part de mosaïques à corrélogrammes exponentiels.

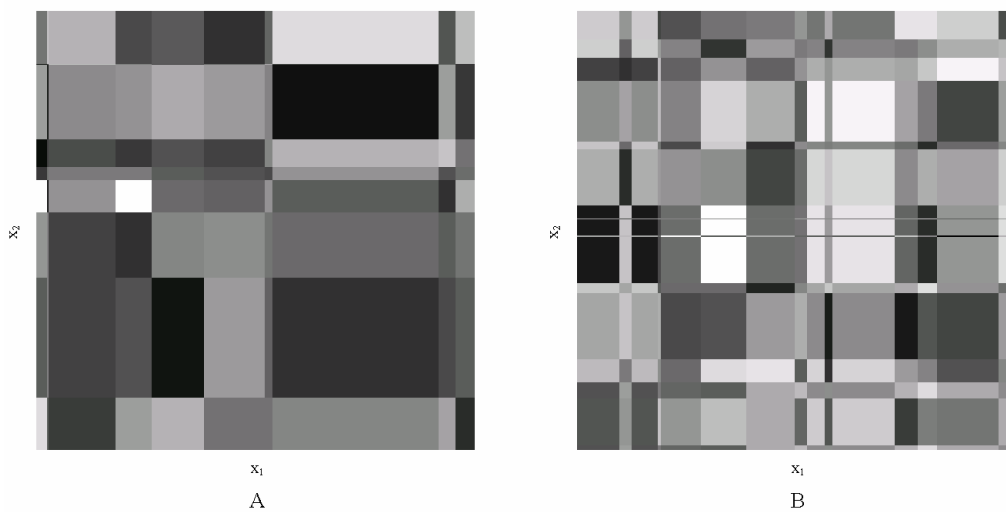


Figure 2.4. Réalisations obtenues par le produit de deux mosaïques unidimensionnelles à covariance exponentielle : A, de Poisson et B, déplacement aléatoire du point médian

III. Modèle d'ensemble aléatoire

Lorsque la valuation de la mosaïque est binaire, on simule une fonction indicatrice qui représente un ensemble aléatoire. On a le résultat suivant, qui particularise celui obtenu dans le cadre mosaïque : si le krigeage de l'indicatrice reste toujours compris entre 0 et 1, l'algorithme séquentiel d'indicatrice reproduit correctement la moyenne et la covariance de la fonction indicatrice.

III.1. Conditions sur les pondérateurs de krigeage

Le krigeage simple d'une indicatrice $I(\mathbf{x})$ de moyenne p à partir d'un ensemble de valeurs situées en $\{\mathbf{x}_\alpha, \alpha = 1 \dots n\}$, s'écrit sous la forme :

$$I^*(\mathbf{x}) = \sum_{\alpha=1}^n \lambda_\alpha I(\mathbf{x}_\alpha) + (1 - \sum_{\alpha=1}^n \lambda_\alpha) p.$$

Soient Σ^+ et Σ^- les sommes des poids positifs et négatifs respectivement. Les valeurs extrémales de l'estimateur sont :

$$I_{\max}^*(\mathbf{x}) = \Sigma^+ + (1 - \Sigma^+ - \Sigma^-) p$$

$$I_{\min}^*(\mathbf{x}) = \Sigma^- + (1 - \Sigma^+ - \Sigma^-) p$$

L'estimateur est toujours compris entre 0 et 1 si et seulement si $I_{\max}^*(\mathbf{x}) \leq 1$ et $I_{\min}^*(\mathbf{x}) \geq 0$, soit :

$$0 \leq \Sigma^+ \leq 1 + \frac{p}{1-p} \Sigma^- \text{ et } (\Sigma^+ - 1) \frac{p}{1-p} \leq \Sigma^- \leq 0.$$

Dans le plan paramétré par Σ^+ et Σ^- , la zone autorisée est un triangle délimité par les axes de coordonnées et la droite d'équation

$$\Sigma^+ = 1 + \max\left(\frac{p}{1-p}, \frac{1-p}{p}\right) \Sigma^-.$$

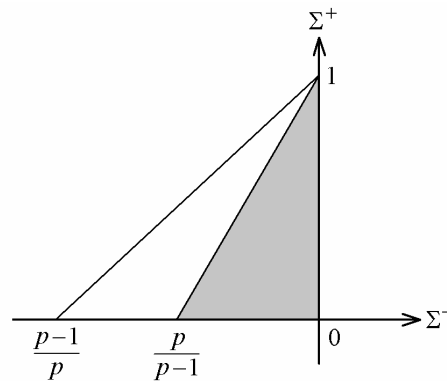


Figure 2.5. Condition sur les poids de krigeage

III.2. Cas d'une indicatrice de moyenne 1/2

La condition la moins sévère se produit lorsque $p = 1/2$ (l'indicatrice compte autant de 1 que de 0) ; on obtient alors : $\Sigma^+ - \Sigma^- \leq 1$. Cette inégalité traduit synthétiquement que la somme des *valeurs absolues* des poids de krigeage doit être inférieure à 1 :

$$\sum_{\alpha=1}^n |\lambda_{\alpha}| \leq 1 \quad (2.3)$$

Comme dans le modèle mosaïque à valuations continues, une condition suffisante est que les poids de krigeage soient tous positifs et de somme inférieure à 1 (cas d'un effet de pépite pur ou d'une covariance complètement monotone dans l'espace unidimensionnel), mais à première vue elle n'est pas nécessaire.

III.2.1. Point de vue matriciel

La condition antérieure peut s'exprimer sous forme matricielle. Soient $\{\mathbf{x}_{\alpha}, \alpha = 1 \dots n\}$ un ensemble de sites, $\mathbf{C} = [C(\mathbf{x}_{\alpha} - \mathbf{x}_{\beta})]_{\alpha, \beta = 1 \dots n}$ leur matrice de variance-covariance et \mathbf{B} la matrice inverse⁴ de \mathbf{C} . On démontre que les pondérateurs de krigeage de \mathbf{x}_{α} par les $\{\mathbf{x}_{\beta}, \beta \neq \alpha\}$ valent $\{-b_{\alpha\beta}/b_{\alpha\alpha}, \beta \neq \alpha\}$, où $b_{\alpha\beta}$ désigne le terme générique de \mathbf{B} (Dubrule, 1983). Ainsi, l'équation (2.3) équivaut à écrire :

$$\forall \alpha = 1 \dots n, b_{\alpha\alpha} \geq \sum_{\substack{\beta=1 \\ \beta \neq \alpha}}^n |b_{\alpha\beta}|$$

c'est-à-dire que la matrice \mathbf{B} est à *diagonale dominante*. Cette propriété doit être vérifiée pour tout choix des sites $\{\mathbf{x}_{\alpha}, \alpha = 1 \dots n\}$.

III.2.2. Point de vue du krigeage dual

Le krigeage de l'indicatrice à partir des données situées en $\{\mathbf{x}_{\alpha}, \alpha = 1 \dots n\}$ peut s'exprimer sous forme duale (Matheron, 1980b, 1981b ; Chauvet, 1999) :

$$\varepsilon(\mathbf{x}) = 2\mathbf{I}^*(\mathbf{x}) - 1 = \sum_{\alpha=1}^n b_{\alpha} C(\mathbf{x}_{\alpha} - \mathbf{x})$$

où les coefficients $\{b_{\alpha}, \alpha = 1 \dots n\}$ vérifient le système d'équations linéaires qui exprime les conditions d'interpolation exacte aux sites de données :

$$\forall \alpha = 1 \dots n, \sum_{\beta=1}^n b_{\beta} C(\mathbf{x}_{\alpha} - \mathbf{x}_{\beta}) = \varepsilon(\mathbf{x}_{\alpha}) = \pm 1.$$

La fonction de covariance doit être telle que la fonction interpolatrice $\varepsilon(\mathbf{x})$ soit comprise entre -1 et +1, pour tout choix des sites $\{\mathbf{x}_{\alpha}, \alpha = 1 \dots n\}$ et des valeurs $\{\varepsilon(\mathbf{x}_{\alpha}), \alpha = 1 \dots n\}$ dans $\{-1, +1\}$.

⁴ Tout comme \mathbf{C} , \mathbf{B} est une matrice symétrique définie positive. Ses termes diagonaux sont toujours positifs ; ce sont en fait les inverses des variances de krigeage (Dubrule, 1983).

En notant ici encore \mathbf{B} la matrice inverse de \mathbf{C} , les coefficients $\{b_\alpha, \alpha = 1 \dots n\}$ s'expriment de la façon suivante :

$$\forall \alpha = 1 \dots n, b_\alpha = \sum_{\alpha=1}^n b_{\alpha\beta} \varepsilon(\mathbf{x}_\beta) = b_{\alpha\alpha} \varepsilon(\mathbf{x}_\alpha) + \sum_{\substack{\alpha=1 \\ \alpha \neq \beta}}^n b_{\alpha\beta} \varepsilon(\mathbf{x}_\beta).$$

Dire que \mathbf{B} est à diagonale dominante équivaut à dire que b_α est du même signe que $\varepsilon(\mathbf{x}_\alpha)$ pour tout $\alpha \in \{1 \dots n\}$ et tout choix des $\{\varepsilon(\mathbf{x}_\alpha), \alpha = 1 \dots n\}$ dans $\{-1, +1\}$. Comme la fonction de covariance présente un maximum absolu à l'origine, cette condition permet que la fonction d'interpolation

$$\varepsilon(\mathbf{x}) = \sum_{\alpha=1}^n b_\alpha C(\mathbf{x}_\alpha - \mathbf{x})$$

présente un extremum local en chaque site de données (maximum ou minimum, selon le signe de la donnée). Ceci suggère que les "points névralgiques" de $\varepsilon(\mathbf{x})$ à étudier sont les sites de données.

III.2.3. Conjecture

On conjecture que seules les covariances complètement monotones dans l'espace unidimensionnel vérifient l'équation (2.3). Il est facile d'exhiber des configurations mettant en défaut cette condition pour les autres modèles usuels de covariance (sphérique, triangulaire, quadratique, gaussienne...), même lorsqu'on leur ajoute une forte composante pépitique. Par exemple, un corrélogramme composé d'un effet de pépite d'amplitude 0.95 et d'un schéma sphérique de palier 0.05 et portée 1 conduit à une somme des valeurs absolues des poids égale à 1.0298 lorsque l'on estime le site origine à partir de 1000 données régulièrement réparties sur le segment $[-1, 1]$.

III.2.4. Autodualité

Si l'on simule une indicatrice de moyenne 1/2, sous réserve que le krigeage d'indicatrice soit toujours compris entre 0 et 1, l'ensemble aléatoire simulé est *autodual*, c'est-à-dire que l'on peut échanger les 0 et les 1 sans modifier la loi spatiale. Une première conséquence est que l'algorithme séquentiel est incapable de simuler des ensembles aléatoires non autoduaux tels la plupart des schémas booléens.

Si l'on représente l'ordre des sites simulés par une permutation $\sigma = \{i_1, \dots, i_n\}$ de $\{1, \dots, n\}$, la loi multivariable des valeurs simulées $\{I(\mathbf{x}_1), \dots, I(\mathbf{x}_n)\}$ s'exprime de la manière suivante [Eq. (2.1)] :

$$f_{\sigma, n}(y_1, \dots, y_n) = \prod_{k=1}^n \left[\sum_{\ell < k} \lambda_{\ell, k}^{(\sigma)} \delta_{y_{i_k} = y_{i_\ell}} + (1 - \sum_{\ell < k} \lambda_{\ell, k}^{(\sigma)}) / 2 \right] = \frac{1}{2^n} \prod_{k=1}^n \left[1 + \sum_{\ell < k} \lambda_{\ell, k}^{(\sigma)} (2\delta_{y_{i_k} = y_{i_\ell}} - 1) \right]$$

où $\{y_1, \dots, y_n\}$ sont des valeurs de $\{0, 1\}$

$\{\lambda_{\ell, k}^{(\sigma)}, \ell < k\}$ sont les poids de krigeage du site \mathbf{x}_{i_k} à partir des sites $\{\mathbf{x}_{i_\ell}, \ell < k\}$.

L'autodualité se manifeste par l'invariance de la loi multivariable lorsque l'on change $\{y_1, \dots, y_n\}$ en $\{1 - y_1, \dots, 1 - y_n\}$. On peut également décrire les lois multivariées des valeurs simulées par les fonctionnelles caractéristiques de l'ensemble fermé aléatoire \mathbf{E} associé à l'indicatrice $I(\mathbf{x})$ (Matheron, 1975a) :

$$Q(K) = \text{Prob}[K \cap \mathbf{E} = \emptyset] \text{ pour tout compact } K \text{ de } \mathbf{R}^d$$

$$T(K) = \text{Prob}[K \cap \mathbf{E} \neq \emptyset] = 1 - Q(K)$$

Un ensemble fermé aléatoire est autodual s'il possède les mêmes fonctionnelles caractéristiques que la fermeture de son complémentaire.

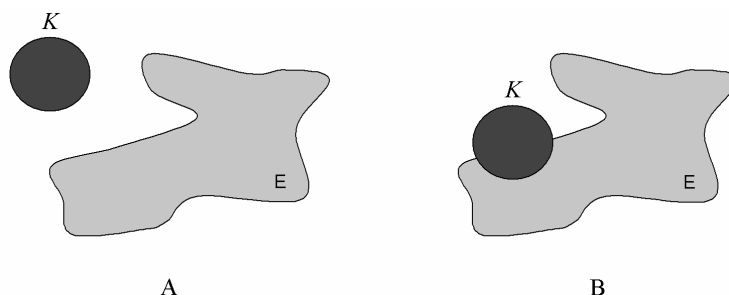


Figure 2.6. Le compact K évite ou, au contraire, rencontre l'ensemble E

Remarque : loi spatiale et fonctionnelles caractéristiques

La **loi spatiale** d'une indicatrice ne caractérise pas complètement l'ensemble aléatoire associé. Ainsi, un ensemble constitué d'un nombre fini ou dénombrable de points situés au hasard dans l'espace a une loi spatiale nulle (égale à la loi spatiale de l'ensemble vide). Plusieurs propriétés de l'ensemble aléatoire ne sont pas décrites par sa loi spatiale, en particulier les propriétés de connexité et convexité dont la définition met en jeu une quantité infinie et non dénombrable de sites.

Les **fonctionnelles caractéristiques** sont des outils plus complets pour caractériser un ensemble aléatoire *fermé*. L'idée de base est de remplacer l'ensemble fini de sites $\{x_1, \dots, x_n\}$ par des ensembles plus volumineux, en l'occurrence des volumes fermés et bornés (compacts). La fonctionnelle T satisfait les propriétés suivantes :

- 1) $0 \leq T \leq 1$, avec $T(\emptyset) = 0$.
- 2) Continuité monotone séquentielle : pour toute suite décroissante $\{K_p, p \in \mathbb{N}\}$ de compacts qui converge vers un compact limite K , la suite $\{T(K_p), p \in \mathbb{N}\}$ décroît et tend vers $T(K)$.
- 3) Pour tout compact K et tout ensemble de compacts $\{K_i, i \in I < \infty\}$, on a l'inégalité :

$$\sum_{j \in I} (-1)^{|j|} T\left(\bigcup_{j \in I} K_j \cup K\right) \leq 0 \text{ où } |j| \text{ indique le cardinal de l'ensemble } j.$$

Cette propriété exprime la positivité de la probabilité que l'ensemble aléatoire rencontre chaque ensemble K_i sans toucher K . Elle implique que T est croissante : si $K \subseteq K'$, alors $T(K) \leq T(K')$.

On qualifie de **capacité alternée d'ordre infini** toute fonction qui vérifie les deux dernières propriétés et de **capacité de Choquet** toute fonction qui satisfait les trois propriétés. Un résultat fondamental est le suivant (Choquet, 1954 ; Kendall, 1974 ; Matheron, 1969a, 1972c, 1975a) : si T est une capacité de Choquet définie sur les compacts de \mathbb{R}^d , alors T caractérise un unique fermé aléatoire. La fonctionnelle caractéristique est donc l'équivalent pour un fermé aléatoire d'une fonction de répartition pour une variable aléatoire.

En revanche, étant donnée une loi spatiale d'indicatrice, il existe un nombre infini de fermés aléatoires associés à cette loi spatiale. Toutefois, l'un d'entre eux (noté E) est contenu dans tous les autres et se caractérise par :

$$E = \overline{E \cap D}$$

pour un ensemble dénombrable D dense dans \mathbb{R}^d (la barre indique la fermeture topologique). On dit que E est séparable par D . Cette propriété signifie que, grâce à un échantillonnage dénombrable et dense (représenté par D), il est possible de reconstituer complètement le fermé aléatoire, bien que disposer d'un échantillonnage dénombrable ne permet que d'inférer la loi spatiale de l'indicatrice. Cependant, en général, l'ensemble D dépend de E . Pour que ceci n'ait pas lieu, on peut restreindre l'étude aux fermés aléatoires dont la frontière a une mesure nulle (Schmitt, 2001). Sous cette contrainte, l'ensemble minimal E est **régulier**, c'est-à-dire qu'il coïncide avec la fermeture de son intérieur, et l'on a :

$$E = \overline{E \cap D}$$

pour n'importe quel ensemble dénombrable D dense dans \mathbb{R}^d .

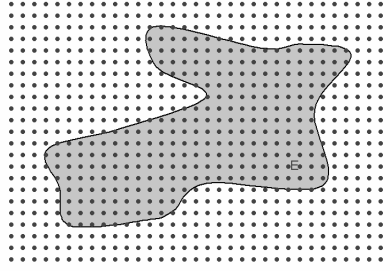


Figure 2.7. Lorsque la maille d'échantillonnage tend vers zéro, on obtient une image complète de l'ensemble aléatoire si celui-ci est régulier

L'algorithme de simulation séquentielle d'indicatrices se place dans l'optique de la loi spatiale, car il simule les sites de l'espace les uns après les autres, donc ne peut traiter qu'un nombre fini de sites. Pour définir sans ambiguïté l'ensemble aléatoire simulé, on doit se restreindre à la famille des fermés aléatoires dont la frontière a une mesure nulle.

III.2.5. Exemple : processus autodial à covariance exponentielle sur la droite réelle

Fonctionnelle caractéristique

Soient $C(\mathbf{h}) = \exp(-a|\mathbf{h}|)$ la fonction de covariance et E l'ensemble aléatoire simulé (ensemble des sites pour lesquels l'indicatrice est égale à 1). Un résultat remarquable est que l'ensemble aléatoire E est indépendant de l'ordre de simulation et que sa fonctionnelle caractéristique est :

$$Q(K_I) = \text{Prob}[K_I \cap E = \emptyset] = \frac{1}{2^I} \prod_{i=1}^{I-1} \{1 + \exp[-a(\mathbf{x}_{i+1} - \mathbf{x}_i)]\}$$

pour toute union finie de sites $K_I = \bigcup_{i=1}^I \{\mathbf{x}_i\}$.

La démonstration s'effectue par récurrence sur le nombre de sites simulés. La propriété est vérifiée lorsqu'il n'y a qu'un ou deux sites. Supposons-la valide jusqu'à l'ordre $I-1$ et simulons un site supplémentaire \mathbf{x}_I . Notons \mathbf{x}_α et \mathbf{x}_β les sites adjacents à \mathbf{x}_I avec $\alpha, \beta \in \{1, \dots, I-1\}$. A cause de l'effet d'écran induit par la covariance exponentielle, la valeur Y_I est égale à Y_α ou Y_β ou est indépendante de ces valeurs, avec des probabilités données par les poids de krigeage et le poids de la moyenne respectivement. En particulier, si Y_α et Y_β sont toutes deux nulles, Y_I vaut 1 avec une probabilité égale à la moitié du poids de la moyenne, 0 avec la probabilité complémentaire.

En notant K_{I-1} l'union de tous les sites $\{\mathbf{x}_1, \dots, \mathbf{x}_{I-1}\}$, ordonnés par abscisse croissante, on a :

$$\begin{aligned} Q(K_I) &= \text{Prob}[K_I \cap E = \emptyset] = \text{Prob}[K_{I-1} \cap E = \emptyset] \times \text{Prob}[Y_I = 0 | K_{I-1} \cap E = \emptyset] \\ &= \frac{1}{2^{I-1}} \prod_{i=1}^{I-2} \{1 + \exp[-a(\mathbf{x}_{i+1} - \mathbf{x}_i)]\} \times \left\{1 - \frac{\exp[-a(\mathbf{x}_\beta - \mathbf{x}_I)] + \exp[-a(\mathbf{x}_I - \mathbf{x}_\alpha)]}{1 + \exp[-a(\mathbf{x}_\beta - \mathbf{x}_\alpha)]}\right\} \frac{1}{2} \\ &= \frac{1}{2^I} \prod_{i=1}^{I-2} \{1 + \exp[-a(\mathbf{x}_{i+1} - \mathbf{x}_i)]\} \times \frac{1 + \exp[-a(\mathbf{x}_\beta - \mathbf{x}_\alpha)] + \exp[-a(\mathbf{x}_\beta - \mathbf{x}_I)] + \exp[-a(\mathbf{x}_I - \mathbf{x}_\alpha)]}{1 + \exp[-a(\mathbf{x}_\beta - \mathbf{x}_\alpha)]} \\ &= \frac{1}{2^I} \prod_{i=1}^{I-2} \{1 + \exp[-a(\mathbf{x}_{i+1} - \mathbf{x}_i)]\} \times \frac{\{1 + \exp[-a(\mathbf{x}_\beta - \mathbf{x}_I)]\} \{1 + \exp[-a(\mathbf{x}_I - \mathbf{x}_\alpha)]\}}{1 + \exp[-a(\mathbf{x}_\beta - \mathbf{x}_\alpha)]} \\ &= \frac{1}{2^I} \prod_{i=1}^{I-1} \{1 + \exp[-a(\mathbf{x}_{i+1} - \mathbf{x}_i)]\} \end{aligned}$$

On vérifie sans peine que la formule de récurrence est vérifiée même si \mathbf{x}_I n'a pas de voisin d'un côté.

La connaissance de la fonctionnelle caractéristique sur toute union finie de sites équivaut à la donnée de la loi spatiale de l'ensemble aléatoire. D'après ce qui précède, l'ensemble aléatoire est alors entièrement caractérisé s'il est choisi dans la classe des fermés aléatoires dont la frontière est de mesure nulle.

Ce résultat d'unicité facilite la simulation conditionnelle d'un tel ensemble aléatoire, puisque l'on peut considérer les sites de données comme des sites déjà simulés (la contrainte que cela impose sur la séquence des sites simulés est sans incidence sur la loi spatiale du processus simulé).

Coïncidence avec d'autres modèles

L'ensemble aléatoire obtenu coïncide avec plusieurs autres modèles :

- a) mosaïque de Poisson à valuations binaires et équiprobables, dont les compartiments sont définis par un processus de Poisson de paramètre a
- b) processus binaire dont les valeurs alternent aux points d'un processus de Poisson de paramètre $a/2$
- c) schéma booléen stationnaire de covariance exponentielle et moyenne $1/2$.

Pour établir le premier cas, il suffit de considérer la simulation séquentielle en ordre régulier (cf. section sur le modèle mosaïque). Le deuxième cas se démontre en remarquant que les changements de valeurs dans la mosaïque décrite en a) dessinent un processus induit par un processus de Poisson dans lequel chaque point peut être éliminé avec probabilité $1/2$; il s'agit encore d'un processus de Poisson, mais de paramètre $a/2$. La démonstration du troisième cas est plus fastidieuse et requiert de vérifier l'identité des fonctionnelles caractéristiques des deux ensembles aléatoires.

Dans le modèle booléen, la covariance est liée au covariogramme géométrique $K(\mathbf{h})$ et à la densité θ du processus de Poisson associé à la chute des grains primaires (Matheron, 1987) :

$$C(\mathbf{h}) = \frac{1}{4} \exp[\theta K(\mathbf{h})] - \frac{1}{4} \quad \text{d'où} \quad K(\mathbf{h}) = \frac{\ln[1 + \exp(-a|\mathbf{h}|)]}{\theta}$$

Ce covariogramme géométrique est celui d'un segment de longueur aléatoire distribuée selon la loi

$$\forall \ell \geq 0, F_b(\ell) = 1 + K'(\ell) = 1 - \frac{a \exp(-a\ell)}{\theta[1 + \exp(-a\ell)]} \quad (\text{on doit avoir } \theta = a/2 \text{ pour assurer } F_b(0) = 0).$$

A présent, la fonctionnelle caractéristique du schéma booléen stationnaire est (Lantuéjoul, 2002) :

$$Q(K) = \exp[-\theta E(|A \oplus \tilde{K}|)] \quad \text{pour tout compact } K \text{ de } \mathbb{R}$$

où A est le grain primaire (segment de longueur distribuée selon F_b)

$$A \oplus \tilde{K} = \bigcup_{\mathbf{x} \in K} A_{-\mathbf{x}} \quad \text{est le dilaté de } A \text{ par le compact } K \quad (A_{-\mathbf{x}} \text{ est le translaté de } A \text{ par le vecteur } -\mathbf{x}).$$

Soit K_I une union finie ou dénombrable de sites : $K_I = \bigcup_{i=1}^I \{\mathbf{x}_i\}$ avec $\forall i \in \{1, \dots, I-1\}, \mathbf{x}_{i+1} > \mathbf{x}_i$.

Comme A est un segment, la mesure moyenne du dilaté de A par K_I vaut (figure 2.8) :

$$E(|A \oplus \tilde{K}_I|) = I \times K(\mathbf{0}) - \sum_{i=1}^{I-1} K(\mathbf{x}_{i+1} - \mathbf{x}_i)$$

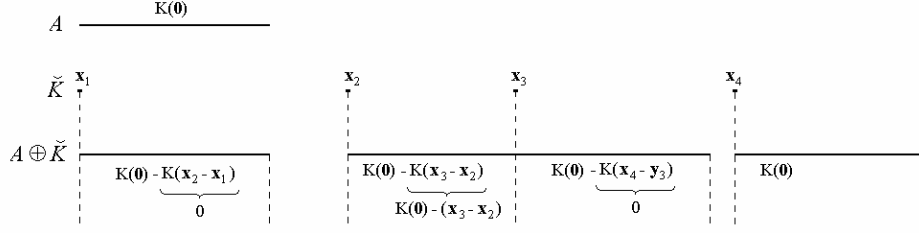


Figure 2.8. Dilatation d'un segment A par le compact K

La fonctionnelle caractéristique du schéma booléen s'écrit finalement :

$$Q(K_1) = \frac{1}{2^I} \prod_{i=1}^{I-1} \{1 + \exp[-a(\mathbf{x}_{i+1} - \mathbf{x}_i)]\}, \text{ C.Q.F.D.}$$

Remarque

Une conséquence est que le schéma booléen de moyenne $1/2$ et de covariance exponentielle à une dimension est autodual. La propriété d'autodualité se produit seulement avec la covariance exponentielle. En effet, le corrélogramme $\rho(\mathbf{h})$ d'un schéma booléen est associé au grain primaire de covariogramme géométrique

$$K(\mathbf{h}) = \ln [1 + \rho(\mathbf{h})] / \theta.$$

La fonctionnelle caractéristique du compact $K_1 = \bigcup_{i=1}^I \{\mathbf{x}_i\}$ devient : $Q(K_1) = \frac{1}{2^I} \prod_{i=1}^{I-1} \{1 + \rho(\mathbf{x}_{i+1} - \mathbf{x}_i)\}$.

Considérons un ensemble de trois sites $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ alignés dans cet ordre. La probabilité que les trois valeurs associées $\{Y_1, Y_2, Y_3\}$ valent 1 s'exprime comme suit :

$$\begin{aligned} \text{Prob}[Y_1 = Y_2 = Y_3 = 1] &= Q(\{\mathbf{x}_1, \mathbf{x}_2\}) + Q(\{\mathbf{x}_1, \mathbf{x}_3\}) + Q(\{\mathbf{x}_2, \mathbf{x}_3\}) - Q(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}) - \frac{1}{2} \\ &= \frac{1}{8} \{1 + \rho(\mathbf{x}_2 - \mathbf{x}_1) + \rho(\mathbf{x}_3 - \mathbf{x}_2) + 2\rho(\mathbf{x}_3 - \mathbf{x}_1) - \rho(\mathbf{x}_2 - \mathbf{x}_1)\rho(\mathbf{x}_3 - \mathbf{x}_2)\} \end{aligned}$$

Cette probabilité est égale à $Q(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\})$ si et seulement si $\rho(\mathbf{x}_2 - \mathbf{x}_1)\rho(\mathbf{x}_3 - \mathbf{x}_2) = \rho(\mathbf{x}_3 - \mathbf{x}_1)$, i.e. si et seulement si $\rho(\mathbf{h})$ est une exponentielle décroissante sur \mathbf{R}_+ . On en déduit que l'ensemble aléatoire de moyenne $1/2$ obtenu par simulation séquentielle d'indicatrice ne coïncide avec un schéma booléen que dans le cas d'une covariance exponentielle à une dimension. Dans les autres cas, le premier ensemble est autodual tandis que le second ne l'est pas.

Unicité du processus ?

Les exemples précédents correspondant curieusement au même ensemble aléatoire, il est légitime de chercher à savoir s'il existe d'autres processus binaires autoduaux de covariance exponentielle dans l'espace à une dimension. De tels processus existent effectivement, comme l'illustre l'exemple suivant.

Soit un processus poissonien qui divise la droite réelle en intervalles indépendants de longueur aléatoire régie selon une loi exponentielle de moyenne $2a$. Sur la première moitié de chaque intervalle on pose $Y^{(1)}(\mathbf{x}) = 1$ alors que sur la seconde moitié $Y^{(1)}(\mathbf{x}) = -1$. Cette technique dite de "migration" conduit à un processus $Y^{(1)}(\mathbf{x})$ admettant comme fonction de covariance (Haas *et al.*, 1967 ; Matheron, 1968b)

$$C^{(1)}(\mathbf{h}) = \left(1 - \frac{|\mathbf{h}|}{a}\right) \exp\left(-\frac{|\mathbf{h}|}{a}\right)$$

On considère alors la fonction aléatoire binaire définie par

$$\forall \mathbf{x} \in \mathbb{R}, Y(\mathbf{x}) = Y^{(1)}(\mathbf{x} U)$$

où U est une variable aléatoire uniforme sur $[-1,1]$. $Y(\mathbf{x})$ n'est autre que la restriction à un axe de coordonnée du processus tridimensionnel obtenu à partir de $Y^{(1)}(\mathbf{x})$ par bandes tournantes (Matheron, 1972b, 1973a). La covariance de $Y(\mathbf{x})$ décroît de manière exponentielle (Lantuéjoul, 1994, 2002) :

$$C(\mathbf{h}) = \exp\left(-\frac{|\mathbf{h}|}{a}\right)$$

La loi de U étant symétrique par rapport à 0, le processus $Y(\mathbf{x})$ représente un ensemble aléatoire autodual, mais diffère du schéma booléen vu précédemment car les alternances de valeurs n'obéissent pas à un processus de Poisson (deux compartiments consécutifs ont la même longueur, ce qui implique que la propriété markovienne n'est pas satisfaite). Pour rendre le processus ergodique, il suffit de le composer avec une mosaïque de Poisson : dans chaque compartiment de la mosaïque, on définit un ensemble aléatoire tel que ci-dessus, indépendamment d'un compartiment à l'autre.

III.3. Ensembles aléatoires non autoduaux

L'identité entre le schéma booléen de covariance exponentielle et l'ensemble aléatoire obtenu par l'algorithme séquentiel d'indicatrice en **ordre régulier** reste vérifiée quand la proportion p de l'espace occupée par l'ensemble aléatoire (moyenne de l'indicatrice) diffère de $1/2$.

En effet, le schéma booléen admet la covariance exponentielle $C(\mathbf{h}) = p(1-p)\exp(-a|\mathbf{h}|)$ si le covariogramme géométrique du grain primaire est

$$K(\mathbf{h}) = \frac{1}{\theta} \ln\left[1 + \frac{p}{1-p} \exp(-a|\mathbf{h}|)\right]$$

qui correspond à un segment de longueur distribuée selon la fonction de répartition

$$\forall \ell \geq 0, F_b(\ell) = 1 + K'(\ell) = 1 - \frac{\exp(-a\ell)}{1-p + p\exp(-a\ell)} \quad (\text{on a posé } \theta = ap).$$

Sous ces conditions, on trouve comme fonctionnelle caractéristique :

$$Q_{\text{booléen}}(K_I) = (1-p)^I \prod_{i=1}^{I-1} \left\{ 1 + \frac{p}{1-p} \exp[-a(\mathbf{x}_{i+1} - \mathbf{x}_i)] \right\}$$

pour tout compact $K_I = \bigcup_{i=1}^I \{\mathbf{x}_i\}$ tel que $\forall i \in \{1, \dots, I-1\}, \mathbf{x}_{i+1} > \mathbf{x}_i$.

Quant à l'algorithme séquentiel en ordre régulier, on obtient une mosaïque de Poisson dont les compartiments sont valués par 1 avec la probabilité p et par 0 avec la probabilité complémentaire. Le nombre de compartiments contenant deux sites \mathbf{x}_i et \mathbf{x}_{i+1} est égal à 1 si l'intervalle $[\mathbf{x}_i, \mathbf{x}_{i+1}]$ ne contient aucun point poissonien, ce qui se produit avec la probabilité $\pi_i = \exp(-a|\mathbf{x}_{i+1} - \mathbf{x}_i|)$, ou à 2 avec la probabilité complémentaire $1 - \pi_i$. Un raisonnement analogue pour un compact K_I constitué de I sites consécutifs $\{\mathbf{x}_i, i = 1 \dots I\}$ donne la fonctionnelle caractéristique cherchée :

$$Q_{\text{séquentiel}}(K_I) = E\left[(1-p)^{1 + \sum_{i=1}^{I-1} B(\pi_i)} \right]$$

où $B(\pi_i)$ est une alternative de Bernoulli de probabilité de succès $1 - \pi_i$.

D'après la propriété d'absence de mémoire du processus de Poisson, les variables $\{B(\pi_i), i = 1 \dots I - 1\}$ sont indépendantes entre elles. Par suite :

$$\begin{aligned}
 Q_{\text{séquentiel}}(K_I) &= (1-p) \prod_{i=1}^{I-1} E[(1-p)^{B(\pi_i)}] \\
 &= (1-p) \prod_{i=1}^{I-1} [e^{-a(x_{i+1}-x_i)} + (1-e^{-a(x_{i+1}-x_i)}) (1-p)] \\
 &= (1-p) \prod_{i=1}^{I-1} [p e^{-a(x_{i+1}-x_i)} + (1-p)] \\
 &= (1-p)^I \prod_{i=1}^{I-1} [1 + \frac{p}{1-p} e^{-a(x_{i+1}-x_i)}]
 \end{aligned}$$

Si la simulation séquentielle s'effectue en **ordre irrégulier**, la loi spatiale des valeurs simulées change. Pour le voir, considérons trois sites $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$, tels que \mathbf{x}_2 soit le milieu de $[\mathbf{x}_1, \mathbf{x}_3]$, que l'on simule selon la séquence $\{\mathbf{x}_1, \mathbf{x}_3, \mathbf{x}_2\}$. En notant ρ le coefficient de corrélation entre les valeurs en \mathbf{x}_1 et \mathbf{x}_2 , la fonctionnelle caractéristique est :

$$Q_{\text{séquentiel}}(\{\mathbf{x}_1, \mathbf{x}_3, \mathbf{x}_2\}) = (1-p) \{1 - p(1-\rho^2)\} \{1 - p(1 - \frac{2\rho}{1+\rho^2})\}$$

qui diffère de celle du schéma booléen.

IV. Modèles à indicatrices multiples

IV.1. Condition pour reproduire les moyennes et covariances d'indicatrices

Les résultats établis dans le cadre mosaïque peuvent se généraliser lorsque les indicatrices ne sont pas en corrélation intrinsèque. Il est alors nécessaire de discrétiser l'ensemble des valeurs de la fonction aléatoire à l'aide de plusieurs seuils de coupure et de spécifier un modèle variographique pour chaque indicatrice de seuil.

Si toutes les fonctions de répartition conditionnelles construites par krigeage d'indicatrices sont croissantes, la simulation reproduira correctement les moyennes et les covariances d'indicatrices. Ce résultat a déjà été mentionné par Journel (1989), sans que cet auteur mette l'accent sur la contrainte de croissance des fonctions de répartition obtenues par krigeage d'indicatrices. Il peut être démontré par récurrence sur le nombre de sites simulés, à l'instar de ce qui a été fait dans le cadre mosaïque.

La contrainte de croissance des fonctions de répartition estimées est très exigeante, car à présent les poids de krigeage évoluent avec le seuil de l'indicatrice considérée. Elle est impossible à satisfaire dès lors que l'on utilise un modèle de covariance interdit pour décrire une indicatrice (Matheron, 1987, 1989a, 1993) ou dès que les modèles associés à plusieurs indicatrices ne sont pas compatibles entre eux : les indicatrices ne sont jamais indépendantes, de sorte que des contraintes lient leurs covariances simples et croisées (Rivoirard, 1993 ; Chilès and Delfiner, 1999).

De plus, les difficultés mentionnées dans le cadre mosaïque et liées à l'utilisation d'un algorithme séquentiel subsistent ; en particulier, les lois multivariées dépendent du nombre de sites et de l'ordre dans lequel ils sont traités. Le mieux est de recourir à un ordre préétabli, par exemple l'ordre régulier ou le déplacement aléatoire du point médian dans l'espace à une dimension.

IV.2. Exemple : indicatrices à covariance exponentielle

Cet exemple a pour but de trouver des modèles qui garantissent la croissance des fonctions de répartition conditionnelles estimées par krigeage d'indicatrices. Dans l'espace à une dimension, on considère un processus dont les indicatrices ont des corrélogrammes exponentiels de portée dépendant du seuil :

$$\forall y \in \mathbf{R}, r_y(\mathbf{h}) = \text{corr}[I(Y(\mathbf{x}+\mathbf{h});y), I(Y(\mathbf{x});y)] = e^{-a(y)|\mathbf{h}|} \text{ avec } a(y) > 0.$$

Si la simulation s'effectue en **ordre régulier**, l'estimation de la fonction de répartition au site à simuler \mathbf{x}_{n+1} ne fait intervenir que le site immédiatement antérieur \mathbf{x}_n à cause de l'effet d'écran de la covariance exponentielle. Cette fonction s'écrit donc :

$$I^*(Y_{n+1}; y) = e^{-a(y)\delta} I(Y_n; y) + [1 - e^{-a(y)\delta}] F(y)$$

où δ est la distance séparant \mathbf{x}_{n+1} de \mathbf{x}_n et $F(y)$ la fonction de répartition *a priori*.

La valeur de la donnée Y_n étant quelconque, la croissance de cette fonction de répartition n'est assurée que si chacune des deux fonctions suivantes est croissante :

$$\phi_1(y) = [1 - e^{-a(y)\delta}] F(y) \text{ et } \phi_2(y) = e^{-a(y)\delta} + [1 - e^{-a(y)\delta}] F(y) = [1 - e^{-a(y)\delta}] [F(y) - 1] + 1$$

ce qui revient à garantir la croissance de $\phi_1(y) = [1 - e^{-a(y)\delta}] F(y)$ et $\phi_2(y) = [1 - e^{-a(y)\delta}] [F(y) - 1]$.

Loi marginale exponentielle

Si $F(y)$ est la fonction de répartition d'une loi exponentielle : $\forall y \in \mathbf{R}_+, F(y) = 1 - e^{-y}$, plusieurs solutions sont envisageables, par exemple :

$$\forall y \in \mathbf{R}_+, a(y) = a e^{by} \text{ avec } 0 \leq a \text{ et } 0 \leq b \leq 1$$

$$\forall y \in \mathbf{R}_+, a(y) = a + b y \text{ avec } 0 \leq b \leq a.$$

Loi exponentielle inversée

Si à présent $\forall y \in \mathbf{R}_-, F(y) = e^y$, on peut proposer les exemples suivants pour la fonction $a(y)$:

$$\forall y \in \mathbf{R}_-, a(y) = a e^{-by} \text{ avec } 0 \leq a \text{ et } 0 \leq b \leq 1$$

$$\forall y \in \mathbf{R}_-, a(y) = a - b y \text{ avec } 0 \leq b \leq a.$$

L'intérêt de cet exemple est que la portée – inversement proportionnelle à la fonction $a(y)$ – augmente avec le seuil, tandis que dans le cas de la loi exponentielle la portée diminue avec le seuil.

Loi quelconque

On se ramène à l'un des cas précédents par anamorphose : on associe à tout seuil la portée du seuil de la loi exponentielle de même fréquence cumulée. On peut aussi composer les cas précédents, par exemple associer les portées de la loi exponentielle inversée jusqu'à un certain seuil, puis celles de la loi exponentielle, ou encore changer les paramètres (a et b) qui régissent l'évolution des portées à partir d'un certain seuil. Ceci confère une très grande souplesse d'utilisation.

Cette procédure fonctionne pour des lois marginales de toute nature ; elle permet d'associer à un histogramme donné des covariances d'indicatrices compatibles avec un tel histogramme, au sens où l'ensemble de ces paramètres pourront être reproduits exactement par l'algorithme séquentiel à une dimension.

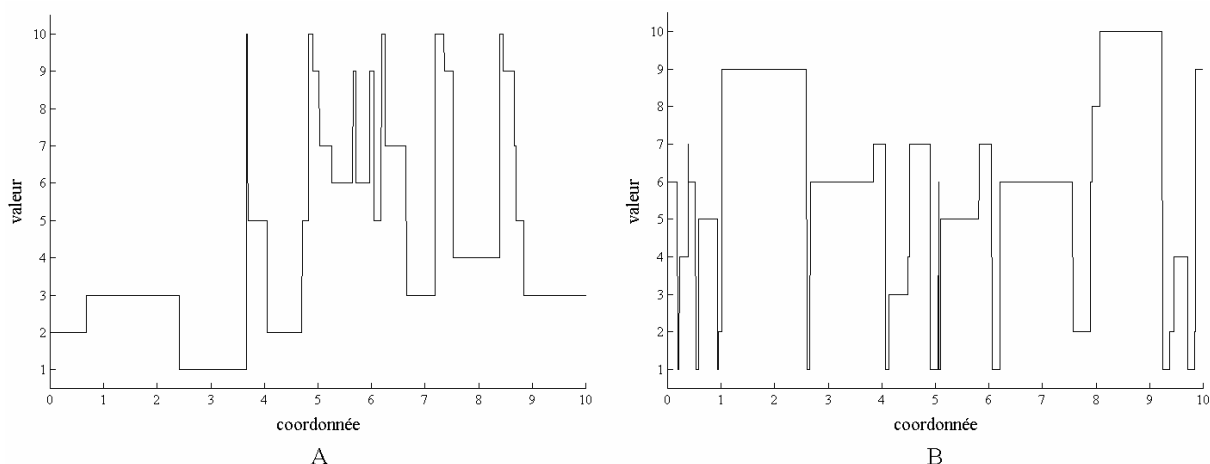


Figure 2.9. Réalisations de processus discrets de loi uniforme dans $\{1, \dots, 10\}$ pour lequel la portée des indicatrices diminue (A) ou augmente (B) avec le seuil

Remarque

Les réalisations obtenues ont un aspect de mosaïque : elles présentent une valeur constante dans un compartiment de l'espace, puis une transition brusque vers une autre valeur. La raison est que la probabilité qu'un site ait la même valeur que le site précédent n'est pas nulle :

$$\text{Prob}[Y_{n+1} = y \mid Y_n = y] = e^{-a(y)\delta}.$$

Contrairement au cas mosaïque, cette probabilité de transition dépend de la valeur prise par le site Y_n , ce qui signifie que la longueur des compartiments dépend de leur valuation.

Si l'on souhaite impérativement obtenir des réalisations dont les trajectoires ne sont pas divisées en compartiments, il faut améliorer la procédure séquentielle, par exemple en tenant compte des covariances croisées des indicatrices et en remplaçant le krigeage propre des indicatrices par un cokrigeage (simulation séquentielle isofactorielle, cf. section suivante).

IV.3. Covariances croisées d'indicatrices

En général les covariances croisées des indicatrices simulées ne sont pas stationnaires, c'est-à-dire qu'elles ne dépendent pas seulement de la distance de séparation mais aussi de la position des points d'appui. En outre, elles dépendent de la séquence selon laquelle les sites sont simulés.

En effet, considérons un ensemble de données $\{Y_1, \dots, Y_n\}$ situées en $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ et telles que :

$$\begin{aligned} \forall y \in \mathbb{R}, \forall i \in \{1, \dots, n\}, E\{I(Y_i; y)\} &= \text{Prob}(Y_i < y) = F(y) \\ \forall y, y' \in \mathbb{R}, \forall i, j \in \{1, \dots, n\}, \text{cov}\{I(Y_i; y), I(Y_j; y')\} &= \alpha(y, y') r_{yy'}(\mathbf{x}_i - \mathbf{x}_j) \end{aligned}$$

où $\alpha(y, y') = F(y) [1 - F(y')]$ si $y \leq y'$, $F(y') [1 - F(y)]$ sinon
 $r_{yy'}(\mathbf{x}_i, \mathbf{x}_j)$ est une fonction égale à 1 lorsque $\mathbf{x}_i = \mathbf{x}_j$ et telle que $r_{yy'}(\mathbf{x}_i, \mathbf{x}_j) = r_{y'y}(\mathbf{x}_j, \mathbf{x}_i)$.

La fonction de répartition conditionnelle en \mathbf{x}_{n+1} est estimée par krigeage d'indicatrices :

$$E[I(Y_{n+1}; y) | \text{données}] = \sum_{i=1}^n \lambda_i(y) I(Y_i; y) + \lambda_0(y) F(y) \quad \text{où} \quad \lambda_0(y) = 1 - \sum_{i=1}^n \lambda_i(y)$$

Par rapport au cas mosaïque, les corrélogrammes d'indicatrices et les poids de krigeage dépendent du seuil y . La formule ci-dessus implique :

$$\begin{aligned} E[I(Y_{n+1}; y) I(Y_j; y')] &= E\{I(Y_j; y') E[I(Y_{n+1}; y) | \text{données}]\} \\ &= E\{I(Y_j; y') [\sum_{i=1}^n \lambda_i(y) I(Y_i; y) + \lambda_0(y) F(y)]\} \\ &= \sum_{i=1}^n \lambda_i(y) \alpha(y, y') r_{yy'}(\mathbf{x}_i - \mathbf{x}_j) + F(y) F(y') \end{aligned}$$

d'où

$$\text{cov}\{I(Y_{n+1}; y), I(Y_j; y')\} = \alpha(y, y') \sum_{i=1}^n \lambda_i(y) r_{yy'}(\mathbf{x}_i - \mathbf{x}_j)$$

Cette expression dépend des configurations de krigeage, donc de la séquence de simulation, et n'a pas de raison de se simplifier (en particulier, elle ne dépend pas seulement de la distance séparant \mathbf{x}_{n+1} et \mathbf{x}_j) :

$$\text{cov}\{I(Y_{n+1}; y), I(Y_j; y')\} \neq \alpha(y, y') r_{yy'}(\mathbf{x}_{n+1} - \mathbf{x}_j) .$$

Même si l'on randomise la séquence de simulation, il n'est pas certain que les covariances croisées deviennent stationnaires (fonctions de $\mathbf{x}_{n+1} - \mathbf{x}_j$ seulement). En effet, les configurations de krigeage ne sont pas invariantes par translation car il y aura inévitablement des effets de bords en périphérie du domaine simulé. Par ailleurs, cette approche pose le problème de la dépendance des lois multivariées vis-à-vis du nombre total de sites considérés dans la simulation (*cf.* section sur le modèle mosaïque).

En conséquence, certaines caractéristiques de la loi bivariable vont dépendre de la séquence de simulation et ne seront pas stationnaires dans l'espace : c'est le cas en particulier de la covariance, qui s'exprime comme la somme de toutes les covariances croisées d'indicatrices lorsque les seuils (y, y') balayent \mathbb{R}^2 (Matheron, 1982). Avoir une covariance ou un variogramme dépendant de la séquence de simulation signifie que, d'une réalisation à l'autre, le variogramme régional risque de présenter des différences dont l'amplitude dépasse celle d'une simple "fluctuation" statistique. C'est là une situation fort inconfortable.

En ce qui concerne la non-stationnarité, le seul moyen de l'éviter est de trouver une séquence de simulation qui laisse les covariances croisées d'indicatrices invariantes par translation. Par exemple, dans l'espace à une dimension, la simulation en **ordre régulier** avec des corrélogrammes exponentiels donnera, à cause des effets d'écran :

$$\forall j \in \{1, \dots, n\}, r_{yy'}(\mathbf{x}_{n+1}, \mathbf{x}_j) = \lambda_n(y) r_{yy'}(\mathbf{x}_n, \mathbf{x}_j) = \prod_{i=j}^n \lambda_i(y) \times 1 = \prod_{i=j}^n r_y(\mathbf{x}_{i+1} - \mathbf{x}_i) = r_y(\mathbf{x}_{n+1} - \mathbf{x}_j) .$$

Ainsi, la covariance croisée entre les indicatrices associées à deux coupures distinctes ne dépend que de la distance de séparation entre valeurs. Pour les distances positives, elle est proportionnelle au corrélogramme de la coupure relative à la dernière valeur simulée ; pour les distances négatives, elle est proportionnelle au corrélogramme de la coupure relative à la première valeur simulée :

$$\forall \mathbf{h} \geq 0, r_{yy'}(\mathbf{h}) = r_y(\mathbf{h}) \quad \text{et} \quad r_{yy'}(-\mathbf{h}) = r_{y'y}(\mathbf{h}) = r_{y'}(\mathbf{h}) .$$

Bien que stationnaire, la loi bivariable représentée par ces covariances d'indicatrices ne peut être isofactorielle, car elle n'est pas symétrique (invariante lorsque l'on change \mathbf{h} en $-\mathbf{h}$). En revanche, le processus bénéficie de la propriété de Markov, ce qui facilite sa simulation conditionnelle (*cf.* section suivante sur les modèles isofactoriels).

L'ordre régulier est sans doute la seule séquence fournissant des covariances croisées stationnaires. Par exemple, considérons un ordre irrégulier tel le déplacement aléatoire du point médian. Soient \mathbf{x}_1 et \mathbf{x}_2 les deux sites extrêmes, \mathbf{x}_3 , \mathbf{x}_4 et \mathbf{x}_5 les milieux respectifs de $[\mathbf{x}_1, \mathbf{x}_2]$, $[\mathbf{x}_1, \mathbf{x}_3]$ et $[\mathbf{x}_3, \mathbf{x}_2]$. Supposant à nouveau que les covariances d'indicatrices sont des schémas exponentiels, un calcul montre que les covariances croisées entre \mathbf{x}_1 et \mathbf{x}_3 d'une part et entre \mathbf{x}_4 et \mathbf{x}_5 d'autre part, ne sont pas égales :

$$r_{yy'}(\mathbf{x}_1, \mathbf{x}_3) \neq r_{yy'}(\mathbf{x}_4, \mathbf{x}_5)$$

d'où l'absence de stationnarité du processus simulé. Ce défaut de stationnarité se présente dès qu'une même distance sépare des paires de sites simulés à partir de configurations différentes.

IV.4. Simulation hiérarchique

Considérons une variable à valeurs dans $\{0, \dots, m\}$ et définissons des ensembles aléatoires par

$$E_0 = \mathbb{R}^d \text{ et } \forall i \in \{1, \dots, m\}, E_i = \{\mathbf{x} \in \mathbb{R}^d / Y(\mathbf{x}) \geq i\}$$

Une procédure connue sous le nom de *simulation hiérarchique* (Dimitrakopoulos and Dagbert, 1993) consiste à simuler chacun de ces ensembles aléatoires de façon indépendante par l'algorithme séquentiel d'indicatrices, puis à poser :

$$\forall \mathbf{x} \in \mathbb{R}^d, Y(\mathbf{x}) = \max\{i / \mathbf{x} \in E_i\} \quad (2.4)$$

Cette procédure permet de diminuer les temps de calcul par rapport à la simulation d'indicatrices classique, puisque l'ensemble E_i n'a besoin d'être simulé que dans le complémentaire des ensembles d'ordre supérieur ; c'est pourquoi il est avantageux de simuler E_m , puis E_{m-1} , et ainsi de suite jusqu'à E_1 (d'où le qualificatif de "hiérarchique" donné à la méthode).

Loi bivariable

Le processus simulé ne présente pas d'*effets de bord* en descendant : si en un site \mathbf{x} on sait que la valeur est plus faible qu'un seuil j (c'est-à-dire si \mathbf{x} n'appartient à aucun des ensembles E_j, \dots, E_m), alors la valeur simulée en \mathbf{x} est indépendante de l'éventuelle appartenance d'un site voisin $\mathbf{x} + \mathbf{h}$ à E_j, \dots, E_m . La simulation hiérarchique est donc implicitement associée au modèle isofactoriel à *résidus d'indicatrices orthogonaux* (Rivoirard, 1988, 1989). Il s'agit d'une extension du modèle mosaïque, où les effets de bord ne se manifestent que lorsque l'on passe d'une valeur faible à une valeur forte. On démontre que la covariance croisée entre les indicatrices associées aux seuils i et j est proportionnelle au corrélogramme de l'indicatrice du seuil le plus élevé. Ce modèle reste donc relativement limité ; les modèles de diffusion du type multigaussien sont inaccessibles.

Lois multivariées

Comme le modèle mosaïque est un cas particulier de ce modèle sans effets de bord en descendant, on sait par avance que les lois à plus de deux variables dépendront de la séquence de simulation et du nombre total de sites simulés. C'est pourquoi il est conseillé de recourir à un ordre prédéterminé, le plus simple étant de choisir le même ordre pour toutes les simulations d'indicatrices (par exemple dans l'espace à une dimension, l'ordre régulier avec des covariances exponentielles ou le déplacement aléatoire du point médian).

Plus précisément, la loi multivariable de l'ensemble des valeurs simulées $\{Y(\mathbf{x}_1), \dots, Y(\mathbf{x}_n)\}$ peut se factoriser sous la forme suivante :

$$\text{Prob}[Y(\mathbf{X}_1) = 1, \dots, Y(\mathbf{X}_m) = m] = \prod_{i=1}^m \text{Prob}[\mathbf{X}_i \subset E_i \text{ et } (\bigcup_{j=1}^{i-1} \mathbf{X}_j) \cap E_i = \emptyset]$$

où $\{\mathbf{X}_i, i = 1 \dots m\}$ forme une partition de $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$; \mathbf{X}_i représente l'ensemble des sites de $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ prenant la valeur i . La loi multivariable exprime que l'ensemble E_i doit couvrir tous les sites valant i et aucun des sites prenant une valeur inférieure à i . Autrement dit, l'indicatrice simulée $I(Y(\mathbf{x}) ; i)$ doit valoir 0 sur \mathbf{X}_i et 1 sur tous les $\{\mathbf{X}_j, j < i\}$; la probabilité d'occurrence de ces événements dépend de l'ordre selon lequel les sites $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ sont simulés.

Exemple : covariances exponentielles à une dimension

La figure suivante donne une illustration de simulation hiérarchique à une dimension. Les portées des covariances (des modèles exponentiels) dépendent du seuil considéré, tandis que les proportions des ensembles aléatoires sont fixées de façon que le processus simulé suive une distribution uniforme sur $\{0, \dots, 4\}$. Ces derniers sont simulés sur une maille très fine en ordre régulier.

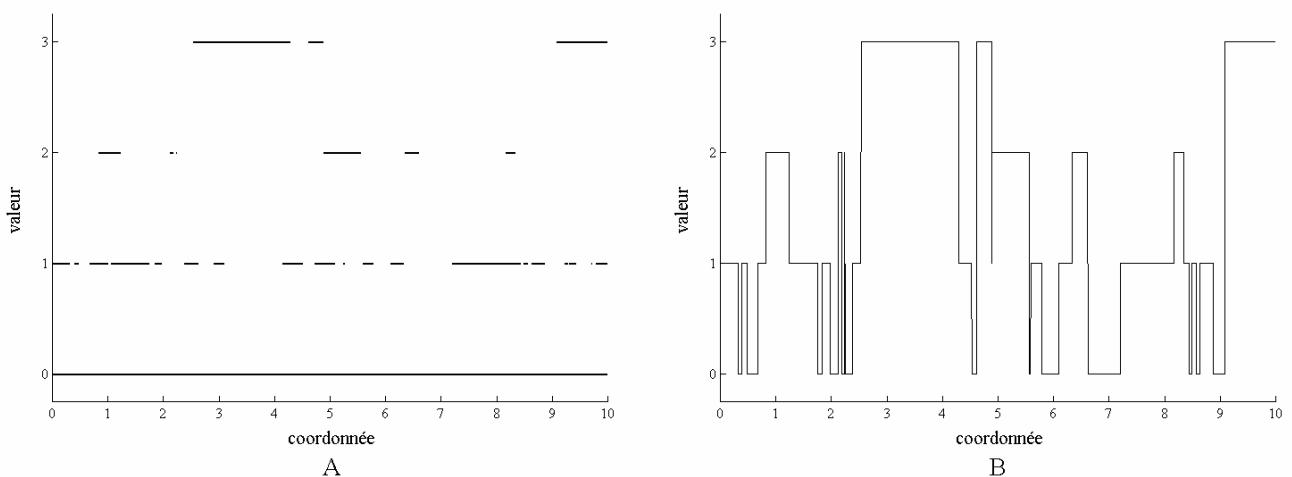


Figure 2.10. A, simulation de trois ensembles aléatoires de covariance exponentielle et B, modèle à résidus d'indicatrices associé, sans effets de bord en descendant

Conditionnement à des données

L'équation (2.4) montre que la simulation conditionnelle du processus $\{Y(\mathbf{x}), \mathbf{x} \in \mathbb{R}^d\}$ se ramène à celle des ensembles aléatoires $\{E_i, i = 0 \dots m\}$, qui sont mutuellement indépendants. Une donnée $Y(\mathbf{x}) = y$ crée une condition sur les ensembles $\{E_i, i = y \dots m\}$ et est inactive sur les autres ensembles $\{E_i, i = 0 \dots y-1\}$:

$$\mathbf{x} \in E_y \text{ et } \forall i > y, \mathbf{x} \notin E_i .$$

IV.5. Variables catégorielles

De telles variables codifient l'appartenance d'un site à une catégorie (*faciès* ou *unité géologique*) sous forme d'un nombre entier : $Y(\mathbf{x}) = i$ si \mathbf{x} appartient au i -ème faciès ; on suppose que l'ensemble des faciès, au nombre de m , forment une partition de l'espace. La valeur de $Y(\mathbf{x})$ étant le produit d'une codification, l'ordre induit par la numérotation des faciès est arbitraire et ne correspond en général à aucune réalité physique.

Dans ce cas, il est tentant d'abandonner le concept d'indicatrice *cumulée* $I(Y(\mathbf{x}); y)$ (fonction valant 1 si $Y(\mathbf{x}) < y$, 0 sinon) et de recourir à celui d'indicatrice *standard* $i(Y(\mathbf{x}); y)$ (fonction valant 1 si $Y(\mathbf{x}) = y$, 0 sinon) (Deutsch and Journel, 1998). Cependant, une telle approche conduit à une impasse dans l'algorithme séquentiel d'indicatrice car la somme des indicatrices standards estimées ne sera pas toujours égale à 1, ce qui signifie que les probabilités d'appartenance aux différents faciès ne totalisent pas 1, à moins de se placer dans le cadre du modèle mosaïque.

Pour le voir, il suffit de considérer l'estimation en un site \mathbf{x} à partir d'un seul site de donnée $\mathbf{x} + \mathbf{h}$:

$$\forall i \in \{1, \dots, m\}, i(Y(\mathbf{x}); i)^* = \rho_i(\mathbf{h})i(Y(\mathbf{x} + \mathbf{h}); i) + [1 - \rho_i(\mathbf{h})]m_i$$

où m_i et $\rho_i(\mathbf{h})$ désignent la moyenne et le corrélogramme de l'indicatrice $i(Y(\mathbf{x}); i)$. En notant i_0 l'indice du faciès présent au site de donnée $\mathbf{x} + \mathbf{h}$, la sommation sur tous les indices donne :

$$\sum_{i=1}^m i(Y(\mathbf{x}); i)^* = \rho_{i_0}(\mathbf{h}) + \sum_{i=1}^m [1 - \rho_i(\mathbf{h})]m_i = \rho_{i_0}(\mathbf{h}) + 1 - \sum_{i=1}^m \rho_i(\mathbf{h})m_i .$$

Cette somme est égale à 1 en toutes circonstances seulement si elle ne dépend pas de i_0 , c'est-à-dire si tous les corrélogrammes sont identiques (modèle mosaïque).

La simulation de variables catégorielles est donc relativement limitée. On devra recourir aux indicatrices cumulées (avec un modèle mosaïque ou à résidus d'indicatrices orthogonaux) au risque de devoir hiérarchiser arbitrairement les faciès. Une autre possibilité revient à employer un cokrigage des indicatrices au lieu d'un krigeage séparé de chacune d'elles : la somme des valeurs estimées sera alors toujours égale à 1. Or, le cokrigage d'indicatrices équivaut à utiliser un krigeage disjonctif ; sa mise en œuvre n'est commode que dans le cadre des modèles isofactoriels qui sont l'objet de la section suivante.

V. Modèles isofactoriels

V.1. Condition pour reproduire les lois marginale et bivariable

Dans le cadre de la simulation séquentielle isofactorielle, on a le résultat suivant : si, au cours de la simulation, toutes les pseudo-densités du krigeage disjonctif sont positives, les bis marginale et bivariable du processus obtenu coïncident exactement avec le modèle spécifié.

La démonstration s'effectue par récurrence. Supposons qu'aux sites $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$, des données $\{Y_1, \dots, Y_n\}$ suivent une loi bivariable isofactorielle caractérisée par les facteurs $\{\chi_p, p \in \mathbf{N}\}$ et leurs corrélogrammes $\{T_p(\mathbf{h}), p \in \mathbf{N}\}$ (cf. annexe B). On simule en \mathbf{x}_{n+1} une valeur Y_{n+1} à partir de sa pseudo-densité $f_{KD}(y | \text{données})$, supposée positive. Pour toute fonction φ développable en série de facteurs, on peut écrire :

$$\forall j \in \{1, \dots, n\}, E[\varphi(Y_{n+1})\varphi(Y_j)] = E\{E[\varphi(Y_{n+1})\varphi(Y_j) | \text{données}]\} .$$

Or, sous réserve de positivité, la pseudo-densité joue le rôle de densité conditionnelle :

$$E[\varphi(Y_{n+1})\varphi(Y_j) | \text{données}] = \left[\int_{-\infty}^{+\infty} \varphi(y) f_{\text{KD}}(y | \text{données}) dy \right] \varphi(y_j) = \varphi(y_{n+1})^{\text{KD}} \varphi(y_j)$$

Par suite, en restituant aux valeurs Y_{n+1} et Y_j leur caractère aléatoire et en notant $\{\varphi_p, p \in \mathbf{N}\}$ les coefficients du développement de φ en facteurs :

$$\forall j \in \{1, \dots, n\}, E[\varphi(Y_{n+1})\varphi(Y_j)] = E[\varphi(Y_{n+1})^{\text{KD}} \varphi(Y_j)] = \varphi_0^2 + E\left[\sum_{p \geq 1} \varphi_p^2 \chi_p(Y_{n+1})^{\text{KS}} \chi_p(Y_j) \right].$$

Désignant par $\{\lambda_i^p, i = 1 \dots n\}$ les poids de krigeage de $\chi_p(Y_{n+1})$ et utilisant les équations du krigeage simple, il vient :

$$\forall j \in \{1, \dots, n\}, E[\varphi(Y_{n+1})\varphi(Y_j)] = \varphi_0^2 + \sum_{p \geq 1} \varphi_p^2 \sum_{i=1}^n \lambda_i^p T_p(\mathbf{x}_i - \mathbf{x}_j) = \varphi_0^2 + \sum_{p \geq 1} \varphi_p^2 T_p(\mathbf{x}_{n+1} - \mathbf{x}_j).$$

La covariance entre $\varphi(Y_{n+1})$ et $\varphi(Y_j)$ admet alors l'expression théorique souhaitée

$$\forall j \in \{1, \dots, n\}, \text{cov}[\varphi(Y_{n+1}), \varphi(Y_j)] = \sum_{p \geq 1} \varphi_p^2 T_p(\mathbf{x}_{n+1} - \mathbf{x}_j)$$

ce qui démontre que la loi bivariable de l'ensemble $\{Y_1, \dots, Y_n, Y_{n+1}\}$ est bien isofactorielle.

En ce qui concerne la loi marginale, sa correcte reproduction découle de la propriété de non-biais du krigeage disjonctif :

$$\forall y \in \mathbf{R}, E[I(Y_{n+1}; y)] = E[I(Y_{n+1}; y)^{\text{KD}}].$$

Là encore, la loi spatiale du processus simulé dépend de l'ordre dans lequel les sites sont traités. Ainsi, dans le cas unidimensionnel avec une séquence de simulation régulière et un corrélogramme exponentiel, la simulation d'une loi bigaussienne est multigaussienne, tandis que la simulation d'une loi bigamma donne un processus de diffusion gamma. Par contre, dans un ordre irrégulier, les lois spatiales diffèrent de ces lois de diffusion (*cf.* chapitre 1).

V.2. Sur la positivité des pseudo-densités du krigeage disjonctif

La question est à présent de savoir quelle condition sur les corrélogrammes $\{T_p(\mathbf{h}), p \in \mathbf{N}\}$ assure la positivité des pseudo-densités du krigeage disjonctif. Pour cela, nous allons examiner en détail le cas bigaussien (loi marginale gaussienne). La pseudo-fonction de répartition et sa dérivée, la pseudo-densité, s'expriment sous la forme (Matheron, 1976 ; Rivoirard, 1994) :

$$F_{\text{KD}}(y | \text{données}) = G(y) + \sum_{p \geq 1} \frac{1}{\sqrt{p}} H_{p-1}(y) g(y) H_p(Y_{n+1})^{\text{KS}}$$

$$f_{\text{KD}}(y | \text{données}) = g(y) + \sum_{p \geq 1} H_p(y) g(y) H_p(Y_{n+1})^{\text{KS}}$$

où $\{H_p, p \in \mathbf{N}\}$ représentent les polynômes d'Hermite normés.

On s'intéresse au cas où l'on dispose de deux données conditionnantes $\{Y_1, Y_2\}$ situées en $\{\mathbf{x}_1, \mathbf{x}_2\}$ pour simuler la valeur Y_0 en un site \mathbf{x}_0 . La pseudo-densité du krigeage disjonctif s'écrit :

$$f_{\text{KD}}(y | y_1, y_2) = g(y) + g(y) \sum_{p \geq 1} \left\{ \frac{\rho_{10}^p - \rho_{12}^p \rho_{20}^p}{1 - \rho_{12}^{2p}} H_p(y_1) + \frac{\rho_{20}^p - \rho_{12}^p \rho_{10}^p}{1 - \rho_{12}^{2p}} H_p(y_2) \right\} H_p(y).$$

Imaginons que les deux données $\{Y_1, Y_2\}$ soient égales à une même valeur y' . On obtient ainsi une condition nécessaire pour que la pseudo-densité soit positive : il faut qu'il en soit de même de la fonction bivariable

$$f(y, y') = g(y)g(y') \sum_{p \geq 0} a_p H_p(y) H_p(y') \text{ où } a_0 = 1 \text{ et } \forall p \geq 1, a_p = \frac{\rho_{10}^p + \rho_{20}^p}{1 + \rho_{12}^p}.$$

L'intégrale de cette fonction vaut 1, car les termes d'ordre $p \geq 1$ ont une intégrale nulle. D'après Matheron (1975b), la condition cherchée est que les coefficients $\{a_p, p \in \mathbf{N}\}$ constituent les moments successifs d'une variable aléatoire à valeurs dans $[-1, 1]$. Or, un théorème (Hausdorff, 1921 ; Widder, 1941) établit qu'une suite de réels $\{a_p, p \in \mathbf{N}\}$ coïncide avec les moments d'une variable aléatoire sur \mathbf{R} si et seulement si pour toute famille finie $\{\alpha_0, \dots, \alpha_p\}$ de scalaires, on a :

$$S(\alpha_0, \dots, \alpha_p) = \sum_{i,j=0}^p \alpha_i \alpha_j a_{i+j} \geq 0.$$

Particularisons au cas où $p = 1$: $S(\alpha_0, \alpha_1) = \alpha_0^2 a_0 + 2\alpha_0 \alpha_1 a_1 + \alpha_1^2 a_2$.

Il s'agit d'un polynôme du second degré en α_0 et α_1 ; un tel polynôme sera toujours positif si son discriminant réduit est négatif ou nul. Or, ce discriminant est du signe de :

$$\Delta = a_1^2 - a_0 a_2 = \frac{2(\rho_{10} - \rho_{12} \rho_{20})(\rho_{20} - \rho_{10} \rho_{12})}{(1 + \rho_{12})^2 (1 + \rho_{12}^2)}$$

Examinons le cas où les trois sites $\{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2\}$ forment un triangle isocèle et supposons que le corrélogramme soit isotrope. Deux situations sont possibles :

- si \mathbf{x}_0 est équidistant de \mathbf{x}_1 et \mathbf{x}_2 (figure 2.11A), on a :

$$\Delta = \frac{2\rho_{10}^2(1 - \rho_{12})^2}{(1 + \rho_{12})^2 (1 + \rho_{12}^2)}$$

La condition cherchée $\Delta \leq 0$ n'est jamais satisfaite, à moins que le corrélogramme soit purement pépitique ($\rho_{10} = 0$ en toutes circonstances). Hormis ce cas, la pseudo-densité du krigeage disjonctif prendra inévitablement des valeurs négatives.

- si \mathbf{x}_2 est équidistant de \mathbf{x}_0 et \mathbf{x}_1 (figure 2.11B), alors on doit avoir :

$$\Delta = \frac{2\rho_{12}(\rho_{10} - \rho_{12}^2)(1 - \rho_{10})}{(1 + \rho_{12})^2 (1 + \rho_{12}^2)} \leq 0.$$

Or, un raisonnement identique au précédent lorsque les deux données $\{Y_1, Y_2\}$ sont égales à y' et $-y'$ respectivement conduit à une autre condition nécessaire :

$$\Delta' = -\frac{2\rho_{12}(\rho_{10} - \rho_{12}^2)(1 - \rho_{10})}{(1 - \rho_{12})^2 (1 + \rho_{12}^2)} \leq 0.$$

Ainsi, les seuls modèles variographiques possibles sont l'effet de pépite pur ($\rho_{12} = 0$) et le schéma exponentiel à une dimension, pour lequel $\rho(2\mathbf{h}) = \rho(\mathbf{h})^2$, soit $\rho_{10} = \rho_{12}^2$.

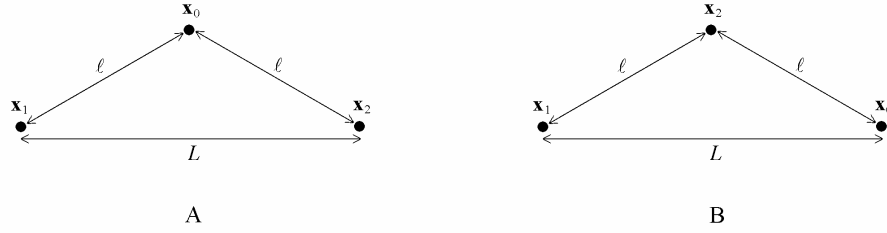


Figure 2.11. Les données et le site à estimer forment un triangle isocèle

Réciproquement, les conditions ci-dessus suffisent à assurer la positivité des pseudo-densités :

- l'effet de pépite conduit à un modèle de “bruit blanc”, où toutes les valeurs sont mutuellement indépendantes ;
- le corrélogramme exponentiel à une dimension fournit un processus de diffusion gaussien lorsque la simulation est effectuée en ordre régulier (cf. chapitre 1) ; il s'agit d'un modèle markovien, ce qui explique pourquoi la connaissance de la seule loi bivariable suffit à reproduire le modèle.

Toutefois, la simulation en ordre régulier n'a d'intérêt que si le résultat ne dépend pas de la maille choisie (niveau de résolution de la simulation). On s'en assure en vérifiant l'identité des densités obtenues en considérant ou non la donnée Y_2 :

$$f_{KD}(y | y_1) = E\{ f_{KD}(y | y_1, y_2) | y_1 \}.$$

Démonstration

$$\begin{aligned} E\{ f_{KD}(y | y_1, y_2) | y_1 \} &= \int f_{KD}(y | y_1, y_2) g(y_2 | y_1) dy_2 \\ &= g(y) + g(y) \sum_{p \geq 1} \frac{\rho_{10}^p - \rho_{12}^p \rho_{20}^p}{1 - \rho_{12}^{2p}} H_p(y_1) H_p(y) + g(y) \sum_{p \geq 1} \frac{\rho_{20}^p - \rho_{12}^p \rho_{10}^p}{1 - \rho_{12}^{2p}} H_p(y) \int H_p(y_2) g(y_2 | y_1) dy_2 \\ &= g(y) + g(y) \sum_{p \geq 1} \frac{\rho_{10}^p - \rho_{12}^p \rho_{20}^p}{1 - \rho_{12}^{2p}} H_p(y_1) H_p(y) + g(y) \sum_{p \geq 1} \frac{\rho_{20}^p - \rho_{12}^p \rho_{10}^p}{1 - \rho_{12}^{2p}} H_p(y) \rho_{12}^p H_p(y_1) \\ &= g(y) + g(y) \sum_{p \geq 1} \rho_{10}^p H_p(y_1) H_p(y) \\ &= f_{KD}(y | y_1) \end{aligned}$$

En bref, hormis le cas du modèle purement pépitique, **seule la simulation en ordre régulier avec une covariance à décroissance exponentielle conduit à une pseudo-densité du krigeage disjonctif qui soit toujours positive**. Ce résultat s'étend aisément aux cadres bigamma, poissonien et binomial négatif : il implique que la simulation séquentielle d'un processus de diffusion est inexacte, quel que soit le modèle variographique retenu, à l'exception des processus markovien et pépitique pur.

Revenant au cas bigaussien, il est intéressant de quantifier jusqu'à quel point la pseudo-densité du krigeage disjonctif peut être négative. Pour qu'une telle étude ait un sens, il est nécessaire de pondérer les valeurs de la pseudo-densité par la probabilité d'occurrence des données $\{Y_1, Y_2\}$. On est ainsi amené à définir une distribution de probabilité de la pseudo-densité. La figure 2.12 montre plusieurs quantiles de cette distribution (quartiles, déciles et centiles extrêmes) lorsque l'on simule une valeur située au milieu de deux données et le modèle variographique est un schéma exponentiel (figure 2.12 gauche) et un schéma exponentiel de palier 0.85 plus un effet de pépite d'amplitude 0.15 (figure 2.12 droite).

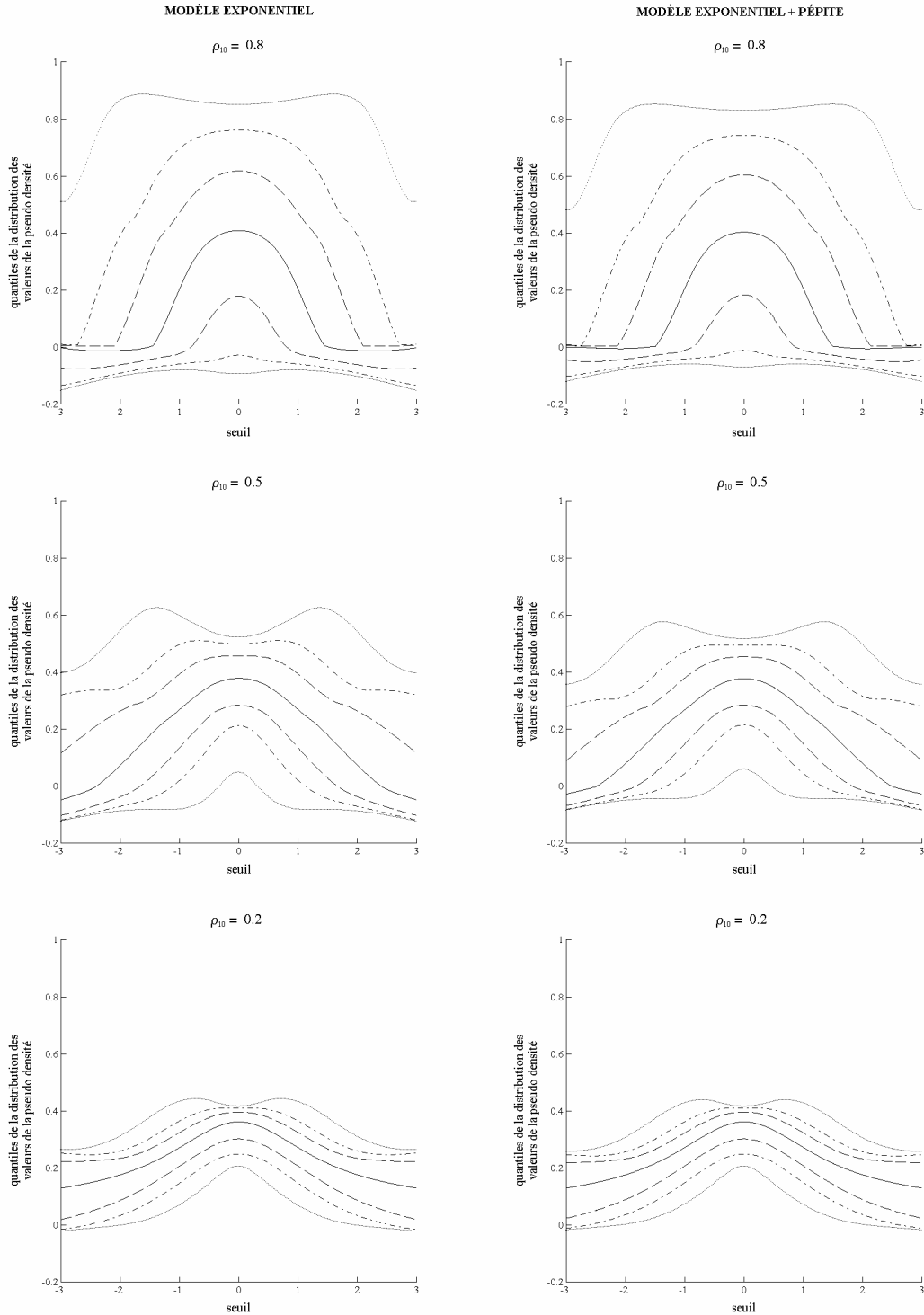


Figure 2.12. Centiles extrêmes, déciles extrêmes et quartiles de la distribution des valeurs de la pseudo-densité lorsque les données varient

La conclusion que l'on tire de ces représentations est que la pseudo-densité du krigeage disjonctif prend souvent des valeurs négatives lorsque la corrélation est importante (jusqu'à -0.15 dans les circonstances les plus défavorables) et lorsque les seuils sont éloignés de 0. L'ajout d'un effet de pénite modéré n'améliore la situation que très légèrement. En conclusion, la méthode séquentielle isofactorielle n'est guère recommandable, sauf pour travailler en ordre régulier dans l'espace à une dimension avec des processus à covariances exponentielles.

Toutefois, si l'on cherche à conditionner des simulations à des données préexistantes, l'ordre de simulation ne peut être régulier. Ceci montre que, si l'on souhaite construire des simulations exactes, il est nécessaire d'explicitier les lois de probabilité au-delà des lois marginales et bivariées. Ici encore, le cas le plus prometteur est celui de processus de covariances exponentielles à une dimension, car la propriété markovienne de cette covariance laisse espérer qu'une loi trivariée suffit.

V.3. Processus markoviens à une dimension

Un processus unidimensionnel, défini le long de l'axe temporel, est *markovien* si, connaissant la valeur présente du processus, les valeurs futures ne dépendent pas des valeurs passées.

Signalons tout d'abord qu'un processus à covariance exponentielle dans l'espace à une dimension n'est pas forcément markovien. Un contre-exemple est fourni par la mosaïque des feuilles mortes avec des segments de longueur aléatoire (cf. chapitre 1). D'autres mosaïques à covariance exponentielle peuvent être construites à l'aide de l'algorithme séquentiel, sans pour autant être markoviennes : les compartiments ne sont pas convexes dès que l'ordre de simulation n'est pas régulier.

En revanche, il est facile d'imaginer la construction de processus markoviens admettant des lois bivariées isofactorielles. On connaît déjà les processus de diffusion (bigaussien ou bigamma), qui sont de covariance exponentielle, mais nous verrons que de tels processus peuvent être généralisés si l'on randomise la portée de leur covariance.

Nous allons commencer par définir des processus markoviens à temps discret ; pour simplifier la présentation, les temps seront pris entiers. La procédure de construction est séquentielle : connaissant les valeurs $\{Y_1, \dots, Y_n\}$, on simule Y_{n+1} selon une loi isofactorielle conditionnée uniquement par la valeur immédiatement antérieure :

$$f(y | y_n) = f(y) \sum_{p \geq 0} T_p(n, n+1) \chi_p(y) \chi_p(y_n).$$

où $\{\chi_p, p \in \mathbb{N}\}$ sont les facteurs des lois bivariées et $\{T_p, p \in \mathbb{N}\}$ leurs covariances.

Le processus ainsi construit est markovien, puisque la valeur Y_{n+1} ne dépend que de Y_n . Grâce à la formule de Bayes, on exprime la densité de probabilité entre trois valeurs consécutives, par exemple $\{Y_1, Y_2, Y_3\}$, en fonction des lois bivariées (voir la formule (1.6) du chapitre 1) :

$$f(y_1, y_2, y_3) = \frac{f(y_1, y_2) f(y_2, y_3)}{f(y_2)} \quad (2.5)$$

L'intégration de cette densité par rapport à y_1 ou y_3 redonne les densités bivariées des couples $\{Y_2, Y_3\}$ et $\{Y_1, Y_2\}$, lesquelles sont isofactorielles. La vérification que la densité de $\{Y_1, Y_3\}$ est isofactorielle requiert un petit calcul :

$$\begin{aligned} f(y_1, y_3) &= \int f(y_1, y_2, y_3) dy_2 \\ &= \sum_{p \geq 0} \sum_{q \geq 0} f(y_1) f(y_3) T_p(1,2) T_q(2,3) \chi_p(y_1) \chi_q(y_3) \int \chi_p(y_2) \chi_q(y_2) f(y_2) dy_2 \\ &= \sum_{p \geq 0} f(y_1) f(y_3) T_p(1,3) \chi_p(y_1) \chi_p(y_3) \end{aligned}$$

avec $\forall p \in \mathbb{N}, T_p(1,3) = T_p(1,2) T_p(2,3)$, qui est bien un moment d'ordre p en tant que produit de deux moments d'ordre p .

Plus généralement, on obtient toujours une loi isofactorielle entre deux sites non consécutifs :

$$\forall p \in \mathbf{N}, T_p(1, n) = T_p(1, 2) \dots T_p(n-1, n) = [T_p(1, 2)]^{n-1} \text{ sous réserve de stationnarité.}$$

Cette équation montre que les covariances des facteurs sont des exponentielles décroissantes :

$$\forall p \in \mathbf{N}, T_p(1, n) = \exp[(n-1) \ln(T_p(1, 2))].$$

En conclusion, si à chaque étape l'on utilise une loi conditionnelle isofactorielle dans le procédé de construction, on obtient un processus markovien dont toutes les lois bivariées sont isofactorielles et dont les covariances des facteurs sont toutes exponentielles. Par rapport au paragraphe précédent, on dispose d'une formule explicite et cohérente de toute loi trivariée. La simulation séquentielle de tels processus dans un ordre irrégulier sera donc possible de manière exacte, puisque la valeur en un site ne dépendra que des valeurs connues adjacentes à ce site ; on pourra donc déterminer la loi de probabilité conditionnelle au site à simuler, puis tirer une valeur aléatoire selon cette loi de probabilité.

Pour généraliser la classe des processus construits, une idée consiste à randomiser le coefficient de corrélation ρ_n entre Y_{n+1} et Y_n selon une distribution de probabilité $\varpi(d\rho)$ sur $[-1, 1]$. Supposant que ce coefficient est indépendant des précédents $\{\rho_1, \dots, \rho_{n-1}\}$ et que, ρ_n étant fixé, la loi entre Y_{n+1} et Y_n est de diffusion, alors la covariance du p -ième facteur pour une distance unité vaut :

$$T_p(1, 2) = \int \rho^p \varpi(d\rho).$$

Donnons quelques exemples simples de lois de probabilité pour ρ_n :

- *Dirac en ρ* : $\varpi(d\rho) = \delta_\rho$

$$T_p(1, 2) = \rho^p : \text{modèle de diffusion}$$

- *Loi à deux atomes en 0 et 1* : $\varpi(d\rho) = (1-\rho)\delta_0 + \rho\delta_1$

$$T_p(1, 2) = \rho : \text{modèle mosaïque}$$

- *Loi beta de paramètres $\{\mathbf{br}, \mathbf{b}(1-\mathbf{r})\}$ sur $[0, 1]$* : $\varpi(d\rho) = \frac{\Gamma(\beta)\rho^{\beta\rho-1}(1-\rho)^{\beta(1-\rho)-1}d\rho}{\Gamma(\beta\rho)\Gamma(\beta(1-\rho))}$ avec $\beta > 0$

$$T_p(1, 2) = \frac{\Gamma(\beta)\Gamma(\beta\rho+p)}{\Gamma(\beta\rho)\Gamma(\beta+p)}.$$

Lorsque la corrélation est toujours positive, une autre possibilité consiste à l'écrire sous la forme $\rho_n = \exp(-a_n)$ où a_n suit une loi de probabilité $\varpi(da)$ sur \mathbf{R}_+ , par exemple :

- *Loi de Poisson de paramètre \mathbf{q}* : $\varpi(da) = e^{-\theta} \sum_{i \in \mathbf{N}} \frac{\theta^i}{i!} \delta_i$

$$T_p(1, 2) = \exp(-\theta(1 - e^{-p})).$$

- *Loi gamma de paramètre \mathbf{a} et facteur d'échelle \mathbf{b}* : $\varpi(da) = \frac{1}{\beta^\alpha \Gamma(\alpha)} \exp(-a/\beta) a^{\alpha-1} da$

$$T_p(1, 2) = \exp(-\alpha \ln(1 + \beta p)).$$

Lorsque β tend vers 0, on se rapproche du processus de diffusion.

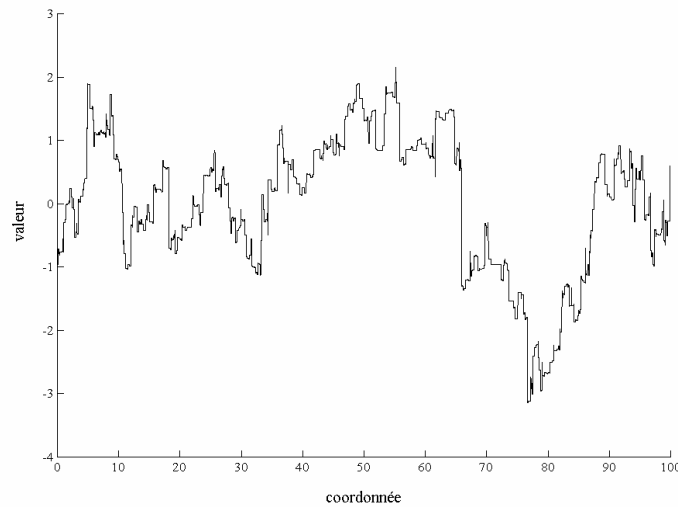


Figure 2.13. Réalisation d'un modèle hermitien (le paramètre de la covariance suit une loi gamma de paramètre $\mathbf{a} = 0.01$)

On obtient ainsi un procédé de construction de processus isofactoriels markoviens, qui ont la particularité de généraliser la classe des processus de diffusion. Pour faciliter l'inférence du modèle et de ses paramètres à partir de données expérimentales, il est commode d'examiner le variogramme d'ordre 1 ou madogramme, noté γ_1 . Dans le cas d'un processus de diffusion (coefficient de corrélation déterministe et constant), le madogramme est proportionnel à la racine carrée du variogramme. Dans le cas général, on a la formule :

$$\forall n \geq 1, \frac{\gamma_1(1,n)}{\gamma_1(\infty)} = E[\sqrt{1-\rho_{1,n}}] = -\frac{1}{2\Gamma(1/2)} \sum_{p \geq 0} \frac{\Gamma(p-1/2)}{p!} T_p(1,n).$$

Nous reviendrons sur ce point au chapitre 4.

Remarques

- 1) La loi bigaussienne autorise l'existence de corrélations négatives, qui ne peuvent donc se mettre sous la forme $\exp(-a)$. Par exemple, écrivons la corrélation entre deux sites consécutifs sous la forme $\cos(a)$ et randomisons a selon une loi uniforme entre 0 et π :

$$T_p(1,2) = \frac{1}{\pi} (1 + (-1)^p) \int_0^{\pi/2} \cos^p(a) da = (1 + (-1)^p) \frac{2^{p-1}}{\pi} \frac{\Gamma[(p+1)/2]^2}{p!}.$$

Dans ce modèle, les covariances des facteurs d'ordre impair sont purement pépétiques, tandis que les facteurs d'ordre pair (fonctions du carré des valeurs) manifestent une structure spatiale : tout se passe comme si le signe des valeurs générées était aléatoire.

- 2) L'hypothèse markovienne a permis d'éviter l'occurrence de densités de probabilité négatives lors de la simulation séquentielle. Ce résultat n'est pas le fruit du hasard, puisque cette seule hypothèse revient en fait à déterminer un type de loi spatiale à partir des seules lois bivariées : le modèle de fonction aléatoire est alors entièrement spécifié.

VI. Extensions des modèles précédents

Pour terminer ce tour d'horizon sur l'algorithme séquentiel, les paragraphes suivants proposent deux extensions des modèles précédents.

VI.1. Composition de modèles

Donnons un exemple simple : il consiste à simuler d'abord une mosaïque unidimensionnelle puis, indépendamment dans chaque compartiment de cette mosaïque, un processus isofactoriel markovien.

Notons $\rho(\mathbf{h})$ la covariance de la mosaïque unidimensionnelle, $f(y)$ et $f(y_1, y_2)$ les densités marginale et bivariable du processus isofactoriel markovien, $\{\chi_p, p \in \mathbf{N}\}$ ses facteurs et $\{T_p(\mathbf{h}), p \in \mathbf{N}\}$ leurs corrélogrammes. La composition des deux modèles donne un processus $\{Y(\mathbf{x}), \mathbf{x} \in \mathbf{R}^d\}$ tel que la densité du couple $\{Y(\mathbf{x} + \mathbf{h}), Y(\mathbf{x})\}$ s'écrit sous la forme :

$$\begin{aligned} f'(y_1, y_2) &= \rho(\mathbf{h})f(y_1, y_2) + [1 - \rho(\mathbf{h})]f(y_1)f(y_2) \\ &= \rho(\mathbf{h})[f(y_1)f(y_2) \sum_{p \geq 0} T_p(\mathbf{h})\chi_p(y_1)\chi_p(y_2)] + [1 - \rho(\mathbf{h})]f(y_1)f(y_2) \\ &= f(y_1)f(y_2) \sum_{p \geq 0} T'_p(\mathbf{h})\chi_p(y_1)\chi_p(y_2) \end{aligned}$$

avec $T'_0(\mathbf{h}) = 1$ et $\forall p \geq 1, T'_p(\mathbf{h}) = \rho(\mathbf{h})T_p(\mathbf{h})$.

Le processus résultant est donc bien isofactoriel de même loi marginale, mais en général il ne sera plus markovien, sauf si la mosaïque unidimensionnelle est elle-même markovienne (mosaïque de Poisson). C'est ce dernier cas qui retiendra notre attention, car il facilite grandement la simulation conditionnelle. En effet, à cause de la propriété markovienne, seule la loi trivariante du processus est requise, laquelle s'exprime en fonction des lois marginale et bivariable [Eq. (2.5)].

VI.2. Modèles interrompus

Partant d'un processus markovien unidimensionnel $\{Y(\mathbf{x}), \mathbf{x} \in \mathbf{R}^d\}$, on élimine à intervalles réguliers un segment de longueur constante l . Soit L la longueur restante entre deux coupures ; l'origine \mathbf{x}_0 des coupures est choisie au hasard dans le segment de coordonnées $[0, L+l]$ (figure 2.14).

Soient deux sites $\{\mathbf{x} + \mathbf{h}, \mathbf{x}\}$. On pose

- $E_h =$ partie entière de $|\mathbf{h}| / L$
- $F_h =$ partie fractionnaire de $|\mathbf{h}| / L$
- $f_h(y_1, y_2) = f(y_1)f(y_2) \sum_{p \geq 0} T_p(\mathbf{h})\chi_p(y_1)\chi_p(y_2)$: loi bivariable de $\{Y(\mathbf{x} + \mathbf{h}), Y(\mathbf{x})\}$

Les sites $\{\mathbf{x} + \mathbf{h}, \mathbf{x}\}$ occupaient initialement d'autres positions qui ont été séparées par un nombre N de coupures : $N = E_h$ avec la probabilité $1 - F_h$, $E_h + 1$ avec la probabilité complémentaire F_h .

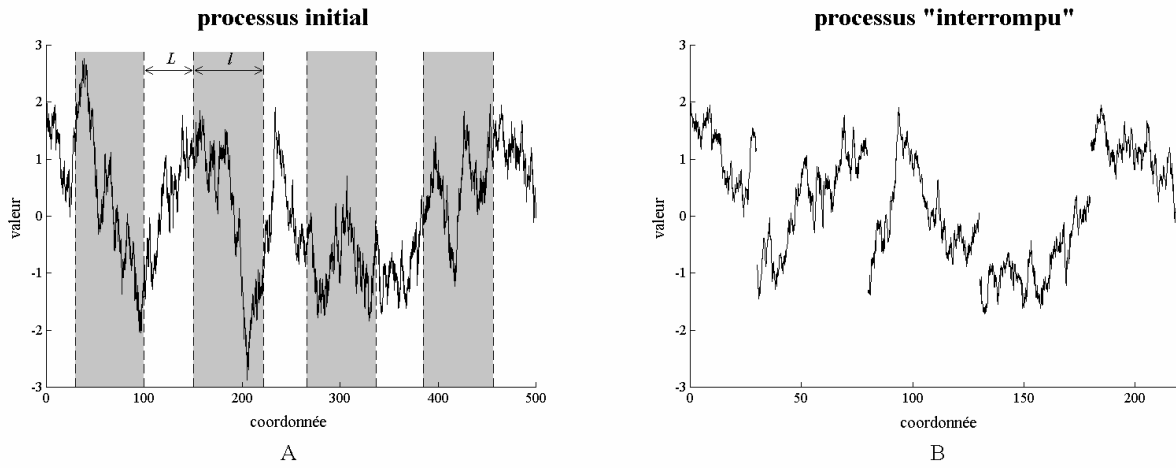


Figure 2.14. A, processus initial markovien et B, processus interrompu

La loi bivariable du processus “interrompu” $Y(\mathbf{x})$ s’exprime comme suit :

$$\begin{aligned}
 f'_h(y_1, y_2) &= (1 - F_h) [f(y_1) f(y_2) \sum_{p \geq 0} T_p(\mathbf{h} + l E_h) \chi_p(y_1) \chi_p(y_2)] \\
 &\quad + F_h [f(y_1) f(y_2) \sum_{p \geq 0} T_p(\mathbf{h} + l E_h + l) \chi_p(y_1) \chi_p(y_2)] \\
 &= f(y_1) f(y_2) \sum_{p \geq 0} T'_p(\mathbf{h}) \chi_p(y_1) \chi_p(y_2)
 \end{aligned}$$

avec $T'_0(\mathbf{h}) = 1$ et $\forall p \geq 1, T'_p(\mathbf{h}) = (1 - F_h) T_p(\mathbf{h} + l E_h) + F_h T_p(\mathbf{h} + l E_h + l)$.

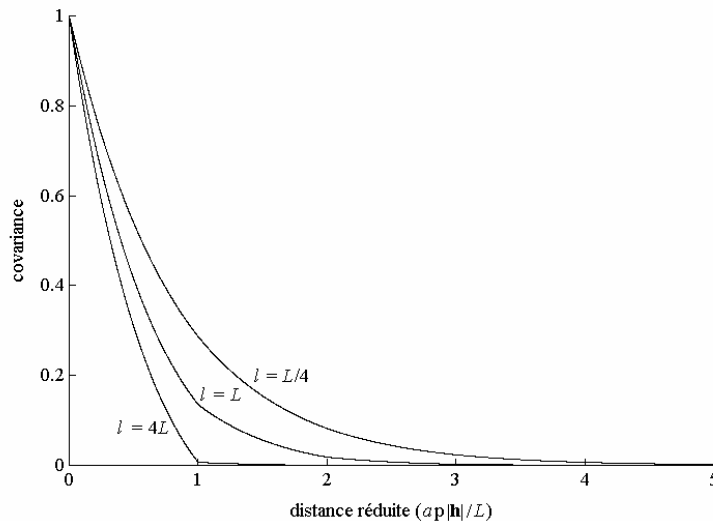


Figure 2.15. Covariance du facteur d’ordre p en fonction du rapport l/L

Ainsi, le processus résultant est encore isofactoriel. Cependant, il perd la propriété markovienne car, connaissant les valeurs jusqu’à un site choisi comme origine, la valeur en un site “futur” dépend non seulement de la valeur présente, mais aussi de l’instant de la dernière coupure, lequel dépend lui-même de toutes les valeurs passées.

Chapitre 3 :

Modèles isofactoriels à plusieurs dimensions

I. Objectif

Le chapitre précédent a permis d'identifier les conditions sous lesquelles l'algorithme séquentiel reproduit exactement les paramètres du modèle théorique. En général, ces modèles doivent être définis dans l'espace à une dimension et posséder la propriété markovienne, ce qui constitue un cadre très restrictif pour les applications.

Ce chapitre vise à étendre la simulation de modèles isofactoriels à l'espace multidimensionnel. A ce titre, deux pistes seront explorées. La première voie est la méthode dite de *substitution* (Matheron, 1989b ; Lantuéjoul, 1993, 2002), qui permet de construire des fonctions aléatoires isofactorielles à plusieurs dimensions à partir de processus isofactoriels unidimensionnels ; sous certaines conditions, ceux-ci pourront être simulés par l'algorithme séquentiel. Cet algorithme sera ensuite laissé de côté et l'on introduira des modèles nouveaux et des méthodes de simulation adaptées. On s'intéressera en particulier aux modèles isofactoriels construits à partir d'ensembles aléatoires, de modèles mosaïques ou de fonctions aléatoires multigaussiennes.

II. Fonctions aléatoires de substitution

II.1. Définition et application à la simulation de modèles isofactoriels

Les méthodes de substitution s'intéressent aux fonctions aléatoires qui s'écrivent sous la forme :

$$Y(\mathbf{x}) = X[T(\mathbf{x})]$$

où $\{T(\mathbf{x}), \mathbf{x} \in \mathbb{R}^d\}$ est une fonction aléatoire appelée *fonction directrice*
 $\{X(t), t \in \mathbb{R}\}$ est un processus unidimensionnel appelé *processus de codage*.

L'explicitation des lois marginales et bivariées de la fonction aléatoire de substitution est possible sous l'hypothèse que $X(t)$ est stationnaire, $T(\mathbf{x})$ intrinsèque stricte et indépendante de $X(t)$ (Lantuéjoul, 2002). Dans ce cas, $Y(\mathbf{x})$ est stationnaire, de même loi marginale que $X(t)$ et de covariance

$$C_Y(\mathbf{h}) = E\{C_X[T(\mathbf{h}) - T(\mathbf{0})]\}.$$

Une propriété importante est la conservation du caractère **isofactoriel** : si $X(t)$ est un processus de loi isofactorielle, il en est de même de $Y(\mathbf{x})$. Cette propriété permet de définir des processus de loi isofactorielle dans des espaces à plusieurs dimensions à partir de processus unidimensionnels. Les facteurs sont inchangés, mais la loi isofactorielle de $Y(\mathbf{x})$ peut être plus générale que celle de $X(t)$ (par exemple, hermitienne au lieu de bigaussienne). Les covariances des facteurs d'ordre p de $X(t)$ et $Y(\mathbf{x})$ sont reliées par la formule (Matheron, 1989b) :

$$C_p^{(Y)}(\mathbf{h}) = E\{C_p^{(X)}[T(\mathbf{h}) - T(\mathbf{0})]\} \quad (3.1)$$

Le conditionnement du modèle à un ensemble de données expérimentales $\{Y(\mathbf{x}_\alpha), \alpha = 1 \dots n\}$ peut être réalisé en plusieurs étapes (Lantuéjoul, 2002) :

- a) simulation de $\{T(\mathbf{x}_\alpha), \alpha = 1 \dots n\}$ conditionnellement à $\{Y(\mathbf{x}_\alpha) = y_\alpha, \alpha = 1 \dots n\}$;
- b) simulation de $T(\mathbf{x})$ conditionnellement à $\{T(\mathbf{x}_\alpha) = t_\alpha, \alpha = 1 \dots n\}$;
- c) simulation de $X(t)$ conditionnellement à $\{X(t_\alpha) = y_\alpha, \alpha = 1 \dots n\}$;
- d) calcul de $X[T(\mathbf{x})]$.

Les étapes b) et c) requièrent de disposer d'algorithmes de simulation conditionnelle, à la fois de la fonction directrice et du processus de codage. Pour leur mise en œuvre, on pose les deux conditions simplificatrices suivantes :

- le **processus de codage** est **markovien**
- la **fonction directrice** est un processus à **accroissements multigaussiens** et son variogramme est la somme d'un schéma linéaire et d'un variogramme à palier. On sait simuler exactement de tels variogrammes dans l'espace multidimensionnel par la méthode des bandes tournantes (Orfeuill, 1972 ; Chilès, 1995) ; le conditionnement s'effectue par krigeage ordinaire.

L'étape a) peut être menée à bien par un procédé itératif basé sur l'échantillonneur de Gibbs (cf. annexe C) ; l'idée consiste à partir d'une simulation non conditionnelle et modifier progressivement les valeurs simulées de façon à se rapprocher de la distribution conditionnelle. Plus précisément, on propose l'algorithme suivant (Lantuéjoul, 2002) :

- a1) simulation non conditionnelle de $\{T(\mathbf{x}_\alpha), \alpha = 1 \dots n\}$; obtention de valeurs $\{t_\alpha, \alpha = 1 \dots n\}$
- a2) sélection d'un indice α_0 au hasard dans $\{1, \dots, n\}$
- a3) simulation de $T(\mathbf{x}_{\alpha_0})$ conditionnellement à $\{T(\mathbf{x}_\alpha) = t_\alpha, \alpha \neq \alpha_0\}$: obtention d'une nouvelle valeur t'_{α_0} candidate pour remplacer t_{α_0}
- a4) calcul des probabilités des valeurs t_{α_0} et t'_{α_0} , sachant $\{X(t_\alpha), \alpha \neq \alpha_0\}$:

$$p_{\alpha_0} = \text{Prob}[X(t_{\alpha_0}) = y_{\alpha_0} \mid \{X(t_\alpha) = y_\alpha, \alpha \neq \alpha_0\}]$$

$$p'_{\alpha_0} = \text{Prob}[X(t'_{\alpha_0}) = y_{\alpha_0} \mid \{X(t_\alpha) = y_\alpha, \alpha \neq \alpha_0\}]$$

- a5) simulation d'une valeur u uniforme dans $[0,1]$

- a6) si $u \leq \frac{p'_{\alpha_0}}{p_{\alpha_0} + p'_{\alpha_0}}$, remplacement de la valeur ancienne t_{α_0} par la valeur nouvelle t'_{α_0}

a7) itération du procédé depuis a2).

Telle qu'écrite ci-dessus, l'étape a4) suppose que $X(t)$ suit une loi discrète. Si la loi de $X(t)$ est absolument continue, il suffit de remplacer la probabilité par la densité de probabilité. La propriété markovienne du processus de codage entraîne des simplifications importantes, car seules les valeurs adjacentes à t_{α_0} et t'_{α_0} interviennent dans le calcul.

A ce stade, on dispose de peu de degrés de liberté, car le variogramme du processus de codage est un schéma exponentiel. Nous allons voir qu'une plus grande souplesse d'utilisation est possible dans le cadre du modèle hermitien.

II.2. Modèle hermitien

Supposons que $T(\mathbf{x})$ et $X(t)$ soient des processus multigaussiens, le premier intrinsèque strict de variogramme γ_T et le second stationnaire de corrélogramme C_X . La méthode de substitution fournit alors un processus $Y(\mathbf{x})$ isofactoriel hermitien. L'intérêt de ce modèle est que l'on peut s'affranchir du caractère markovien du processus de codage sans compliquer la tâche de simulation. Les manipulations restent relativement simples dans le cadre multigaussien : ainsi, les lois conditionnelles requises au cours de l'étape a4) sont gaussiennes, dont les moyennes et variances sont données par un système de krigeage. Par rapport au cas général, on a levé une restriction, puisque le variogramme du processus de codage peut maintenant être quelconque parmi les variogrammes à palier.

II.2.1. Quelques exemples

Pour se faire une idée des modèles hermitiens qu'il est possible d'obtenir, nous donnons ci-après quatre exemples illustratifs où l'on teste plusieurs covariances pour le processus de codage.

1) Covariance exponentielle + pépité (figure 3.1A) : $C_X(\Delta t) = (1 - C_0) \delta_0 + C_0 \exp(-a |\Delta t|)$ où $C_0 \leq 1$.

Le facteur d'ordre p du processus de substitution a pour covariance :

$$\forall \mathbf{h} \neq \mathbf{0}, C_p^{(Y)}(\mathbf{h}) = 2 (C_0)^p \exp\left[\frac{p^2 a^2 \gamma_T(\mathbf{h})}{2}\right] G[-p a \sqrt{\gamma_T(\mathbf{h})}]$$

où G est la fonction de répartition de la gaussienne réduite.

En effet, la formule générale donnant la covariance des facteurs en fonction de celle du processus de codage [Eq. (3.1)] se développe comme suit :

$$\begin{aligned} \forall \mathbf{h} \neq \mathbf{0}, C_p^{(Y)}(\mathbf{h}) &= E\{C_p^{(X)}[T(\mathbf{h}) - T(\mathbf{0})]\} \\ &= (C_0)^p E\{\exp[-p a |T(\mathbf{h}) - T(\mathbf{0})|]\} \\ &= (C_0)^p E\{\exp[-p a \sqrt{\gamma_T(\mathbf{h})} |U|]\} \quad (U : \text{variable gaussienne réduite}) \\ &= 2 (C_0)^p \int_0^{+\infty} \exp[-p a \sqrt{\gamma_T(\mathbf{h})} y] g(y) dy \\ &= 2 (C_0)^p \exp\left[\frac{p^2 a^2 \gamma_T(\mathbf{h})}{2}\right] \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} \exp\left\{-\frac{1}{2}[p a \sqrt{\gamma_T(\mathbf{h})} + y]^2\right\} dy \\ &= 2 (C_0)^p \exp\left[\frac{p^2 a^2 \gamma_T(\mathbf{h})}{2}\right] G[-p a \sqrt{\gamma_T(\mathbf{h})}] \end{aligned}$$

Le développement de la fonction de répartition gaussienne en fraction continue (Abramowitz and Stegun, 1972) :

$$\forall u \geq 0, G(u) = 1 - \frac{1}{\sqrt{2\pi}} \exp(-u^2/2) \frac{1}{u + \frac{1}{u + \frac{2}{u + \frac{3}{u + \dots}}}}$$

fournit les convergences suivantes :

- comportement asymptotique : comme $\gamma_T(\mathbf{h})$ est un variogramme sans palier, on obtient

$$C_p^{(Y)}(\mathbf{h}) \underset{|\mathbf{h}| \rightarrow +\infty}{\sim} \sqrt{\frac{2}{\pi}} \frac{(C_0)^p}{p a} \frac{1}{\sqrt{\gamma_T(\mathbf{h})}}$$

Cette covariance tend vers 0 mais n'est pas intégrable, car $\gamma_T(\mathbf{h})$ a une croissance au plus parabolique.

- vitesse de convergence lorsque $|\mathbf{h}|$ est fixé et p tend vers l'infini

$$\forall \mathbf{h} \neq \mathbf{0}, C_p^{(Y)}(\mathbf{h}) \underset{p \rightarrow +\infty}{\sim} \frac{1}{a} \sqrt{\frac{2}{\pi \gamma_T(\mathbf{h})}} \frac{(C_0)^p}{p}$$

La déstructuration des covariances des facteurs de $Y(\mathbf{x})$ est plus lente que celle des facteurs du processus de codage $X(t)$, lesquels se déstructurent comme $[C_X(\mathbf{h})]^p$ avec $C_X(\mathbf{h}) < C_0$.

2) Covariance gaussienne + pépite (figure 3.1B) : $C_X(\Delta t) = (1 - C_0) \delta_0 + C_0 \exp(-a |\Delta t|^2)$.

On obtient :

$$\forall \mathbf{h} \neq \mathbf{0}, C_p^{(Y)}(\mathbf{h}) = \frac{(C_0)^p}{\sqrt{1 + 2 p a \gamma_T(\mathbf{h})}}$$

Ainsi, si $\gamma_T(\mathbf{h})$ est un schéma linéaire, les facteurs ont pour covariances des modèles gamma de paramètre 1/2.

De manière explicite, le calcul donne successivement :

$$\begin{aligned} \forall \mathbf{h} \neq \mathbf{0}, C_p^{(Y)}(\mathbf{h}) &= (C_0)^p E\{\exp[-p a \gamma_T(\mathbf{h}) U^2]\} \quad (U : \text{variable gaussienne réduite}) \\ &= (C_0)^p \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp[-p a \gamma_T(\mathbf{h}) y^2] \exp(-y^2/2) dy \\ &= (C_0)^p \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\{-\frac{1}{2}[1 + 2 p a \gamma_T(\mathbf{h})] y^2\} dy \\ &= (C_0)^p \frac{1}{\sqrt{1 + 2 p a \gamma_T(\mathbf{h})}} \end{aligned}$$

Là encore, ces covariances tendent vers 0 lorsque $|\mathbf{h}|$ devient infini mais ne sont pas intégrables ; le comportement asymptotique de $C_p^{(Y)}(\mathbf{h})$ est encore en $[\gamma_T(\mathbf{h})]^{-1/2}$. Toutefois, la déstructuration des facteurs est plus lente, puisqu'elle est en $(C_0^p / p^{1/2})$.

3) **Covariance cosinus + pépité** (figure 3.1C) : $C_X(\Delta t) = (1 - C_0) \delta_0 + C_0 \cos(\omega \Delta t)$.

Les covariances des facteurs du processus de substitution ont pour expression :

$$\forall \mathbf{h} \neq \mathbf{0}, C_p^{(Y)}(\mathbf{h}) = \left(\frac{C_0}{2} \right)^p \sum_{k=0}^p C_p^k \exp\left[-\frac{1}{2}(2k-p)^2 \omega^2 \gamma_T(\mathbf{h})\right].$$

Cette formule sera justifiée dans un cadre plus général au paragraphe suivant.

Les covariances des facteurs d'ordre impair sont intégrables, car leur comportement asymptotique est en $\exp[-\omega^2 \gamma_T(\mathbf{h})/2]$. En revanche, il n'en est pas de même de celles des facteurs d'ordre pair, car le terme correspondant à $k = p/2$ est constant. Lorsque \mathbf{h} est fixé et p tend vers l'infini, le terme dominant est celui où $k = p/2$ ou $k = (p \pm 1) / 2$ selon la parité de p : C_p^k est alors équivalent à $(2^p / p^{1/2})$, donc la convergence des covariances des facteurs est dominée par $(C_0^p / p^{1/2})$.

Lorsque la composante pépitique est nulle, le processus de codage est une sinusoïde. La fonction aléatoire de substitution s'écrit alors sous la forme :

$$Y(\mathbf{x}) = A \cos[\omega T(\mathbf{x})] + B \sin[\omega T(\mathbf{x})]$$

où $T(\mathbf{x})$ est une fonction aléatoire intrinsèque à accroissements multigaussiens et A et B sont deux variables gaussiennes réduites indépendantes.

4) **Covariance cosinus amorti + pépité** (figure 3.1D) : $C_X(\Delta t) = (1 - C_0) \delta_0 + C_0 \exp(-a |\Delta t|^2) \cos(\omega \Delta t)$

Les covariances des facteurs du processus de substitution ont pour expression :

$$\forall \mathbf{h} \neq \mathbf{0}, C_p^{(Y)}(\mathbf{h}) = \left(\frac{C_0}{2} \right)^p \frac{1}{\sqrt{2 a p \gamma_T(\mathbf{h}) + 1}} \sum_{k=0}^p C_p^k \exp\left[-\frac{1}{2} \frac{(2k-p)^2 \omega^2 \gamma_T(\mathbf{h})}{2 a p \gamma_T(\mathbf{h}) + 1}\right].$$

En effet, en utilisant la formule de Moivre et le développement en série du cosinus :

$$\begin{aligned} \forall \mathbf{h} \neq \mathbf{0}, C_p^{(Y)}(\mathbf{h}) &= (C_0)^p E\{\cos^p[\omega \sqrt{\gamma_T(\mathbf{h})} U] \exp[-a p \gamma_T(\mathbf{h}) U^2]\} \\ &= (C_0)^p \frac{1}{2^p} \sum_{k=0}^p C_p^k E\{\cos[(2k-p) \omega \sqrt{\gamma_T(\mathbf{h})} U] \exp[-a p \gamma_T(\mathbf{h}) U^2]\} \\ &= (C_0)^p \frac{1}{2^p} \sum_{k=0}^p C_p^k \sum_{n \geq 0} \frac{(-1)^n}{(2n)!} [(2k-p) \omega \sqrt{\gamma_T(\mathbf{h})}]^{2n} E\{U^{2n} \exp[-a p \gamma_T(\mathbf{h}) U^2]\} \end{aligned}$$

où U désigne une variable gaussienne réduite. En posant $\theta = a p \gamma_T(\mathbf{h})$, une intégration par parties donne :

$$I_{2n} = E(U^{2n} e^{-\theta U^2}) = \sqrt{\frac{2}{\pi}} \int_0^{+\infty} u^{2n} e^{-u^2(\theta + \frac{1}{2})} du = \frac{2\theta + 1}{2n + 1} \sqrt{\frac{2}{\pi}} \int_0^{+\infty} u^{2(n+1)} e^{-u^2(\theta + \frac{1}{2})} du = \frac{2\theta + 1}{2n + 1} I_{2(n+1)}$$

$$d'où : I_{2n} = \frac{2n-1}{2\theta+1} I_{2(n-1)} = \dots = \frac{2n-1}{2\theta+1} \times \frac{2n-3}{2\theta+1} \times \dots \times \frac{3}{2\theta+1} \times \frac{1}{2\theta+1} \times I_0 = \frac{(2n)!}{2^n n! (2\theta+1)^{n+1/2}}.$$

On en déduit finalement :

$$\begin{aligned} \forall \mathbf{h} \neq \mathbf{0}, C_p^{(Y)}(\mathbf{h}) &= (C_0)^p \frac{1}{2^p} \sum_{k=0}^p C_p^k \sum_{n \geq 0} \frac{(-1)^n}{2^n n! (2\theta+1)^{n+1/2}} [(2k-p) \omega \sqrt{\gamma_T(\mathbf{h})}]^{2n} \\ &= (C_0/2)^p \frac{1}{\sqrt{2 a p \gamma_T(\mathbf{h}) + 1}} \sum_{k=0}^p C_p^k \exp\left[-\frac{1}{2} \frac{(2k-p)^2 \omega^2}{2 a p \gamma_T(\mathbf{h}) + 1} \gamma_T(\mathbf{h})\right] \end{aligned}$$

$\gamma_T(\mathbf{h})$ n'ayant pas de palier, le comportement asymptotique de ces covariances est en $[\gamma_T(\mathbf{h})]^{-1/2}$, qu'elles soient d'ordre pair ou impair, dès que le paramètre d'amortissement a n'est pas nul. On retrouve la situation où les covariances ne sont pas intégrables. A \mathbf{h} fixé, la convergence des covariances des facteurs est dominé par (C_0^p / p) quand p tend vers l'infini.

Lorsque la composante pépitique est nulle, la fonction aléatoire de substitution s'écrit sous la forme :

$$Y(\mathbf{x}) = A[T(\mathbf{x})]\cos[\omega T(\mathbf{x})] + B[T(\mathbf{x})]\sin[\omega T(\mathbf{x})]$$

où $T(\mathbf{x})$ est une fonction aléatoire intrinsèque à accroissements multigaussiens et $A(\mathbf{x})$ et $B(\mathbf{x})$ deux fonctions aléatoires gaussiennes indépendantes de même corrélogramme (en l'occurrence, un schéma gaussien de paramètre a). Un modèle similaire a été introduit par Matheron (1982) qui a proposé comme exemple de modèle hermitien :

$$Y(\mathbf{x}) = A(\mathbf{x})\cos[\omega T(\mathbf{x})] + B(\mathbf{x})\sin[\omega T(\mathbf{x})].$$

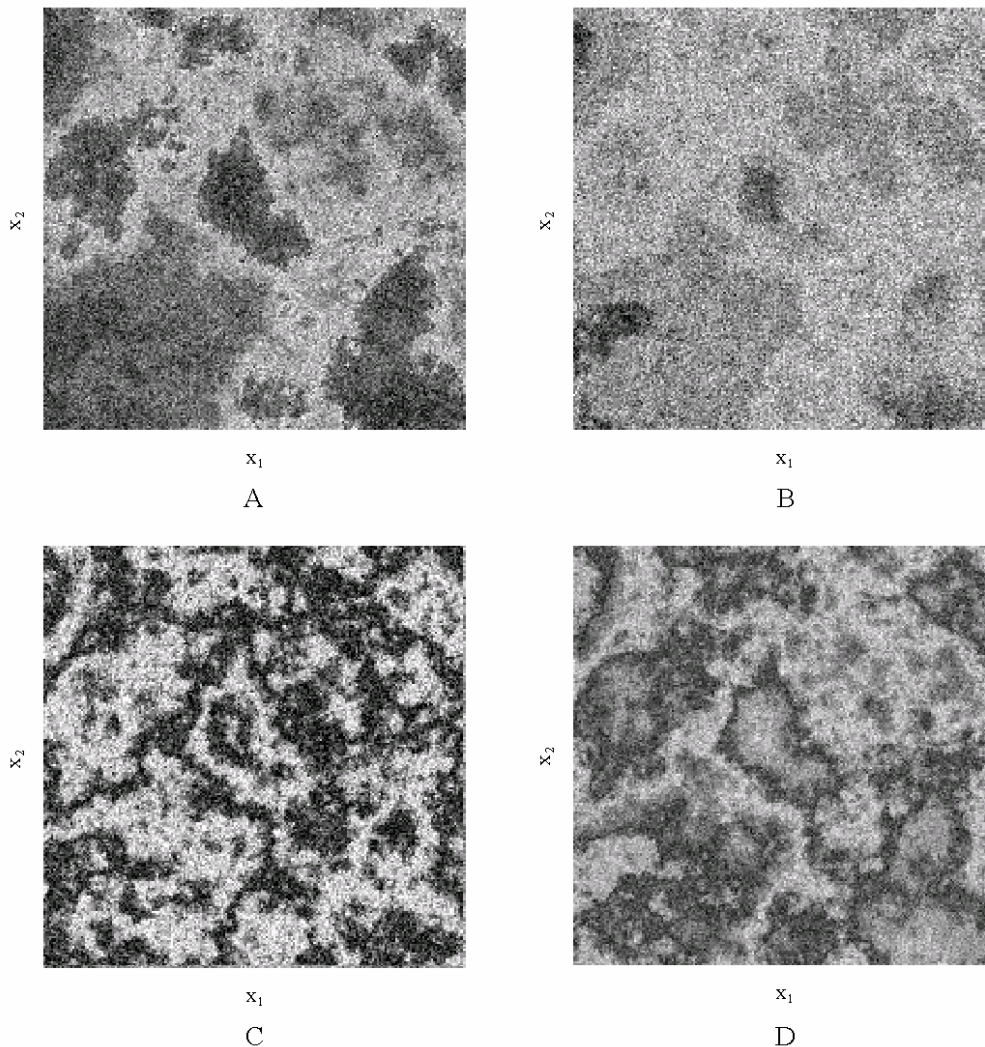


Figure 3.1. Quatre réalisations de modèles hermitiens. Dans tous les cas, la fonction directrice a un variogramme linéaire ; seul change le processus de codage

Conclusion provisoire

Comme l'illustrent les quatre exemples précédents, une difficulté des procédures de substitution est l'obtention de modèles dont les facteurs ont des *portées intégrales* finies ; la portée intégrale d'une fonction aléatoire est définie comme l'intégrale de sa covariance sur tout l'espace (Matheron, 1978b ; Yaglom, 1987). Que celle-ci soit infinie restreint l'ensemble des modèles variographiques accessibles, mais cela a aussi des conséquences sur les caractéristiques du modèle de fonction aléatoire.

La portée intégrale est en particulier liée à la notion d'ergodicité (Lantuéjoul, 1991). L'ergodicité est une propriété selon laquelle une espérance mathématique peut être approchée par une moyenne spatiale. Elle implique que les paramètres de la loi spatiale (histogramme, variogramme, etc.) peuvent être inférés à partir d'une unique réalisation du processus. Or, lorsque la portée intégrale est infinie, il faut disposer d'un domaine très étendu pour espérer avoir une estimation précise de ces paramètres (cela reste possible, mais les conditions sont beaucoup moins favorables).

II.2.2. Inférence

La loi marginale de la fonction aléatoire étudiée peut être transformée par anamorphose, de sorte que le problème de l'inférence concerne surtout les lois bivariées. On introduit le variogramme d'ordre 1 ou *madogramme* de $Y(\mathbf{x})$, défini par :

$$\gamma_1(\mathbf{h}) = E | Y(\mathbf{x} + \mathbf{h}) - Y(\mathbf{x}) | / 2 .$$

Le madogramme est la somme de tous les variogrammes d'indicatrices lorsque le seuil de coupure balaie la droite réelle (Matheron, 1982) et constitue à ce titre un paramètre qui "résume" la distribution du couple $\{Y(\mathbf{x} + \mathbf{h}), Y(\mathbf{x})\}$. Toujours d'après Matheron (1982), le modèle hermitien vérifie :

$$\sqrt{\pi} \gamma_1(\mathbf{h}) = E \{ \sqrt{1 - C_X[T(\mathbf{h}) - T(\mathbf{0})]} \} .$$

La racine carrée peut se développer en série, ce qui laisse apparaître les covariances des facteurs :

$$\sqrt{\pi} \gamma_1(\mathbf{h}) = 1 + \sum_{p \geq 1} \frac{\Gamma(p - 1/2)}{p! \Gamma(-1/2)} C_p^{(Y)}(\mathbf{h}) \quad (3.2)$$

D'autre part, le variogramme est défini par : $\gamma(\mathbf{h}) = E\{1 - C_X[T(\mathbf{h}) - T(\mathbf{0})]\}$.

L'examen du madogramme en fonction du variogramme indique la vitesse de déstructuration des facteurs du processus isofactoriel.

On a toujours : $\sqrt{\pi} \gamma_1(\mathbf{h}) \leq \sqrt{\gamma(\mathbf{h})}$ et, si le corrélogramme C_X est positif, $\sqrt{\pi} \gamma_1(\mathbf{h}) \geq \gamma(\mathbf{h})$.

Ces inégalités s'obtiennent en écrivant :

- $\text{var}\{\sqrt{1 - C_X[T(\mathbf{h}) - T(\mathbf{0})]}\} = E\{1 - C_X[T(\mathbf{h}) - T(\mathbf{0})]\} - E\{\sqrt{1 - C_X[T(\mathbf{h}) - T(\mathbf{0})]}\}^2 \geq 0$
- si C_X est une fonction positive, $\sqrt{1 - C_X[T(\mathbf{h}) - T(\mathbf{0})]} \geq 1 - C_X[T(\mathbf{h}) - T(\mathbf{0})]$.

Les cas d'égalité correspondent aux modèles bigaussien (le madogramme est proportionnel à la racine carrée du variogramme) et mosaïque (madogramme et variogramme sont proportionnels). Les deux premiers modèles présentés précédemment conduisent à des situations intermédiaires (figure 3.2).

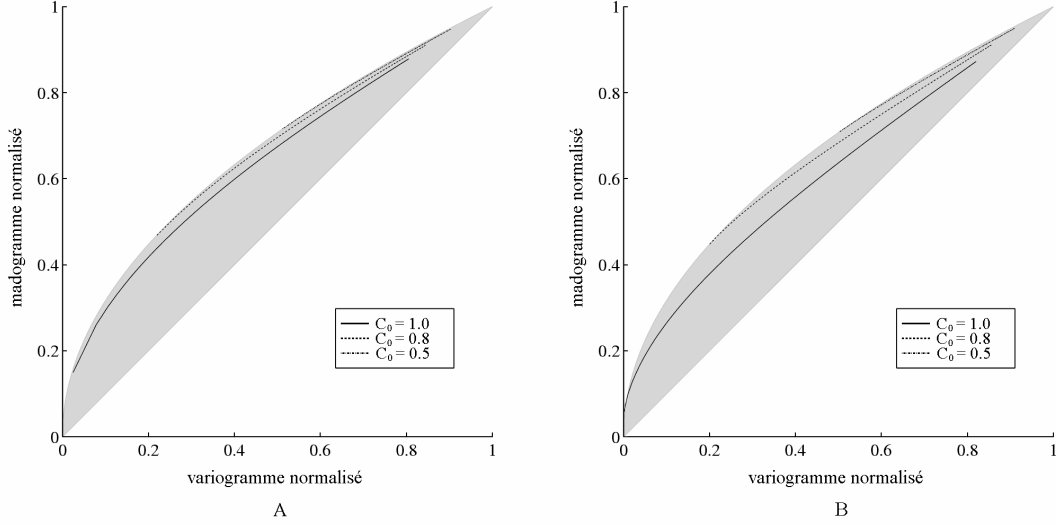


Figure 3.2. Comparaison entre le variogramme et le madogramme pour un processus de codage de covariance (A) exponentielle + pépité et (B) gaussienne + pépité. La zone autorisée est grisée.

II.2.3. Extensions pour obtenir des modèles à portées intégrales finies

Fonction directrice multivariable

Nous proposons ci-dessous une extension de la procédure de substitution qui fournit des portées intégrales finies. Les ingrédients sont :

- une fonction directrice multivariable $\mathbf{T}(\mathbf{x}) = \{T_1(\mathbf{x}), \dots, T_N(\mathbf{x})\}$ à composantes mutuellement indépendantes, chacune de même loi et même variogramme ;
- un processus de codage multigaussien à N dimensions $X(\mathbf{t})$ à covariance factorisée :

$$C_X(\Delta\mathbf{t}) = C_X(\Delta t_1, \dots, \Delta t_N) = \prod_{i=1}^N \rho(\Delta t_i).$$

La loi bivariable du processus de codage étant bigaussienne, la fonction aléatoire de substitution $Y(\mathbf{x}) = X[\mathbf{T}(\mathbf{x})] = X[T_1(\mathbf{x}), \dots, T_N(\mathbf{x})]$ est hermitienne. Son p-ième facteur a pour covariance [Eq. (3.1)] :

$$\begin{aligned} C_p^{(Y)}(\mathbf{h}) &= E\left\{ \left[\prod_{i=1}^N \rho(T_i(\mathbf{h}) - T_i(\mathbf{0})) \right]^p \right\} \quad (\text{caractère bigaussien de } X(\mathbf{t})) \\ &= \prod_{i=1}^N E\left\{ [\rho(T_i(\mathbf{h}) - T_i(\mathbf{0}))]^p \right\} \quad (\text{indépendance de } T_1(\mathbf{x}), \dots, T_N(\mathbf{x})) \\ &= E\left\{ \rho^p [T_1(\mathbf{h}) - T_1(\mathbf{0})] \right\}^N \end{aligned}$$

Par rapport à la section précédente, les covariances des facteurs obtenus sont élevées à la puissance N. Lorsque ce nombre est élevé, la portée intégrale devient finie. On a cette fois trois degrés de liberté qui facilitent la modélisation :

- le variogramme de la fonction directrice, lequel peut notamment inclure des structures gigognes et des anisotropies ;
- la covariance $\rho(\Delta\mathbf{t})$ des processus de codage élémentaires ;
- le nombre N de composantes de la fonction directrice.

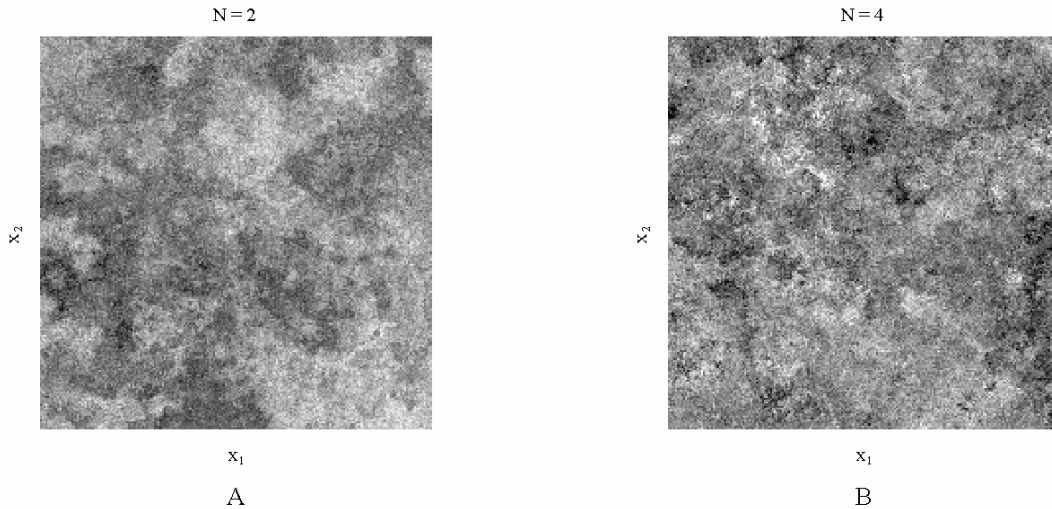


Figure 3.3. Réalisations de fonctions aléatoires hermitiennes utilisant une fonction directrice (A) bivariable et (B) quadrivariable et un processus de codage à covariance exponentielle factorisée

L'**inférence des paramètres** est toujours possible grâce à une comparaison des variogramme et madogramme du processus de substitution [Eq. (3.2)].

Le **conditionnement** des simulations à un ensemble de données $\{Y(\mathbf{x}_\alpha), \alpha = 1 \dots n\}$ est mis en œuvre selon l'algorithme décrit au paragraphe II.1. L'étape c) se divise en deux parties : une simulation non conditionnelle d'un processus factorisé (directement ou par le théorème de la limite centrale) et un conditionnement classique par krigeage. L'étape a) est inchangée, à ceci près que l'on travaille avec des composantes vectorielles $\{\mathbf{t}_\alpha, \alpha = 1 \dots n\}$ au lieu de scalaires. Pour accélérer la convergence de l'algorithme, on peut n'actualiser qu'une seule composante de la fonction directrice à la fois : ainsi, au cours de l'étape a3), on choisit un indice i dans $\{1, \dots, N\}$ et on obtient une valeur candidate pour remplacer la i -ème composante de $\mathbf{T}(\mathbf{x}_{\alpha_0})$.

Introduction de termes de dérivés dans la fonction directrice

Soit d la dimension de l'espace où l'on cherche à simuler la fonction de substitution (i.e. le nombre de composantes du vecteur \mathbf{x}). La seconde extension proposée considère les ingrédients suivants :

- une fonction directrice multivariable $\mathbf{T}(\mathbf{x}) = \{T_1(\mathbf{x}), \dots, T_{d+N}(\mathbf{x})\}$ dont les N dernières composantes sont des fonctions aléatoires à accroissements multigaussiens mutuellement indépendantes, de variogramme $a^2 \gamma_1(\mathbf{h})$, et les d premières composantes sont des termes de "dérive" de la forme

$$\forall i \in \{1, \dots, d\}, T_i(\mathbf{x}) = \varepsilon_i b x_i$$

avec $\varepsilon_i = 1$ ou -1 avec la même probabilité $1/2$

b : coefficient scalaire positif

x_i : i -ème coordonnée du vecteur \mathbf{x} .

- un processus de codage multigaussien $\mathbf{X}(\mathbf{t})$ à $d+N$ dimensions qui s'écrit sous la forme

$$\mathbf{X}(\mathbf{t}) = X(t_1, \dots, t_d, \dots, t_{d+N}) = X_1(t_1, \dots, t_d) \times X_{d+1}(t_{d+1}) \times \dots \times X_{d+N}(t_{d+N})$$

où X_1 est un processus d -dimensionnel isotrope de covariance $\rho_1(\Delta t)$, tandis que X_{d+1}, \dots, X_{d+N} sont des processus unidimensionnels indépendants de même covariance $\rho(\Delta t)$.

La covariance de $\mathbf{X}(\mathbf{t})$ se factorise de la façon suivante :

$$C_{\mathbf{X}}(\Delta\mathbf{t}) = C_{\mathbf{X}}(\Delta t_1, \dots, \Delta t_{d+N}) = \rho_1(\sqrt{\Delta t_1^2 + \dots + \Delta t_d^2}) \prod_{i=d+1}^{d+N} \rho(\Delta t_i).$$

La loi bivariable du processus de codage étant bigaussienne, la fonction aléatoire de substitution $\mathbf{Y}(\mathbf{x}) = \mathbf{X}[\mathbf{T}(\mathbf{x})]$ est hermitienne. Son p -ième facteur a pour covariance [Eq. (3.1)] :

$$\begin{aligned} C_p^{(Y)}(\mathbf{h}) &= E\{C_p^{(X)}[\mathbf{T}(\mathbf{h}) - \mathbf{T}(\mathbf{0})]\} \\ &= \rho_1^p(b|\mathbf{h}) \times E\{\rho^p[\mathbf{T}_{d+1}(\mathbf{h}) - \mathbf{T}_{d+1}(\mathbf{0})]\}^N \\ &= \rho_1^p(b|\mathbf{h}) \times E\{\rho^p[a\sqrt{\gamma_T(\mathbf{h})} | U]\}^N \end{aligned}$$

où U est une variable gaussienne réduite.

Par rapport à l'extension précédente, la fonction de covariance du p -ième facteur est multipliée par la covariance isotrope ρ_1 élevée à la puissance p . Il est alors très facile d'obtenir des covariances de portées intégrales finies ; il suffit par exemple que ρ_1 ait une portée finie. La simulation conditionnelle de tels modèles est identique à celle présentée dans le cadre des fonctions directrices multivariées.

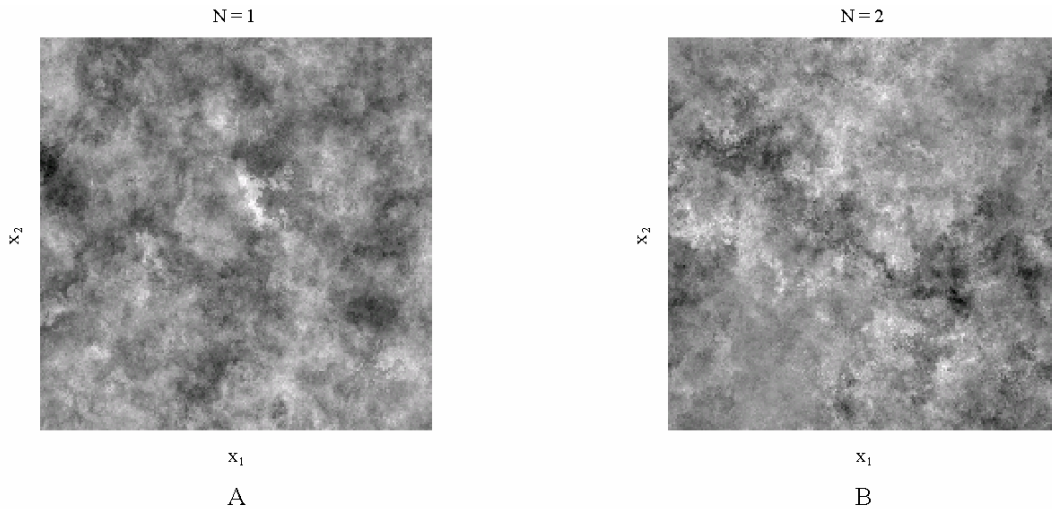


Figure 3.4. Réalisations de fonctions aléatoires hermitiennes utilisant une fonction directrice contenant des termes de dérive

Cette procédure est satisfaisante car elle conduit aux situations extrêmes suivantes :

- si a est infiniment petit, les termes de dérive sont prépondérants et l'on obtient un processus multigaussien ;
- si b est infiniment petit, les termes de dérive n'ont guère d'impact dans les covariances des facteurs et l'on retrouve le processus de substitution proposé antérieurement.

III. Modèles construits à partir d'ensembles aléatoires

Un ensemble aléatoire peut être représenté par son indicatrice, qui est une fonction aléatoire à valeurs binaires. Nous détaillons ci-dessous quelques modèles élémentaires permettant de simuler des ensembles aléatoires. Ceux-ci serviront ensuite d'outils pour construire deux modèles isofactoriels : le modèle à résidus d'indicatrices orthogonaux et le modèle binomial.

III.1. Quelques exemples d'ensembles aléatoires

III.1.1. Modèle de substitution avec un processus de codage binaire

L'obtention d'un ensemble aléatoire par substitution requiert de choisir un processus de codage ne pouvant prendre que deux valeurs. On pose donc :

$$Y(\mathbf{x}) = X[\mathbf{T}(\mathbf{x})]$$

où la fonction directrice $\mathbf{T}(\mathbf{x})$ est intrinsèque à accroissements multigaussiens
le processus de codage $X(\mathbf{t})$ est un processus stochastique binaire.

La loi bivariable de $Y(\mathbf{x})$ est caractérisée par sa covariance :

$$C_Y(\mathbf{h}) = E\{C_X[\mathbf{T}(\mathbf{h}) - \mathbf{T}(\mathbf{0})]\}.$$

Ainsi, une fonction directrice monovariante de variogramme $\gamma_T(\mathbf{h})$ et un processus de codage de covariance exponentielle (de paramètre a) donnent :

$$C_Y(\mathbf{h}) = C_Y(\mathbf{0})\{2 \exp[a^2 \gamma_T(\mathbf{h}) / 2] G[-a \sqrt{\gamma_T(\mathbf{h})}]\}.$$

Si la fonction directrice comprend N composantes indépendantes et si le processus de codage a une covariance exponentielle factorisée, on obtient la même covariance élevée à la puissance N :

$$C_Y(\mathbf{h}) = C_Y(\mathbf{0})\{2 \exp[a^2 \gamma_T(\mathbf{h}) / 2] G[-a \sqrt{\gamma_T(\mathbf{h})}]\}^N.$$

Dans le premier chapitre, on a étudié deux exemples de processus binaires à covariance factorisée (processus "en damier"). Le plus flexible est celui où l'espace est partitionné en compartiments valués indépendamment les uns des autres, car ce modèle permet de choisir les proportions de 0 et 1 (figure 3.5).

Cependant, le conditionnement de ces modèles à des données expérimentales n'est guère réalisable en pratique : les procédures classiques comme l'échantillonneur de Gibbs (*cf.* annexe C) conduisent à une combinatoire très élevée (l'expression des lois de probabilité multivariées est compliquée dès que la propriété markovienne disparaît). Nous allons chercher d'autres modèles permettant d'obtenir des ensembles aléatoires.

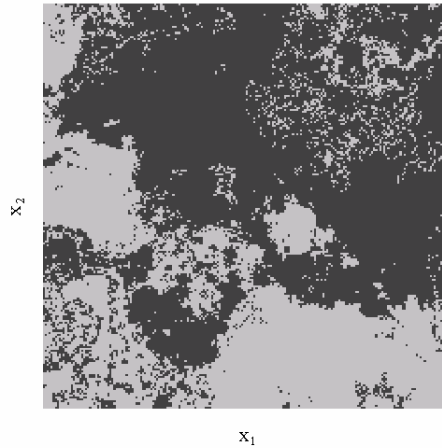


Figure 3.5. Réalisation d'un ensemble aléatoire (fonction directrice bivariable ; processus de codage autodual à covariance exponentielle factorisée)

III.1.2. Troncature de processus isofactoriels hermitiens

Si le processus de codage est une fonction en échelon déterministe et si la fonction directrice est multigaussienne, alors la fonction aléatoire de substitution est une gaussienne tronquée. Une extension multivariable de ce modèle est la technique des plurigaussiennes (Armstrong *et al.*, 2003) qui considère une fonction directrice multivariable et un processus de codage déterministe. La section précédente utilise des processus de codage stochastiques.

De fait, on peut adapter la technique des gaussiennes tronquées au cas des processus isofactoriels. En particulier, si $Y(\mathbf{x})$ est un processus hermitien obtenu par substitution, la covariance de l'ensemble aléatoire coïncidant avec les sites de l'espace où $Y(\mathbf{x})$ est supérieur à un seuil y vaut :

$$C(\mathbf{h}) = g(y)^2 \sum_{p=1}^{\infty} \frac{1}{P} H_{p-1}^2(y) C_p^{(Y)}(\mathbf{h})$$

où $\{H_p, p \in \mathbb{N}\}$ désignent les polynômes d'Hermite normés
 $\{C_p^{(Y)}, p \in \mathbb{N}\}$ sont les covariances des facteurs successifs de $Y(\mathbf{x})$.

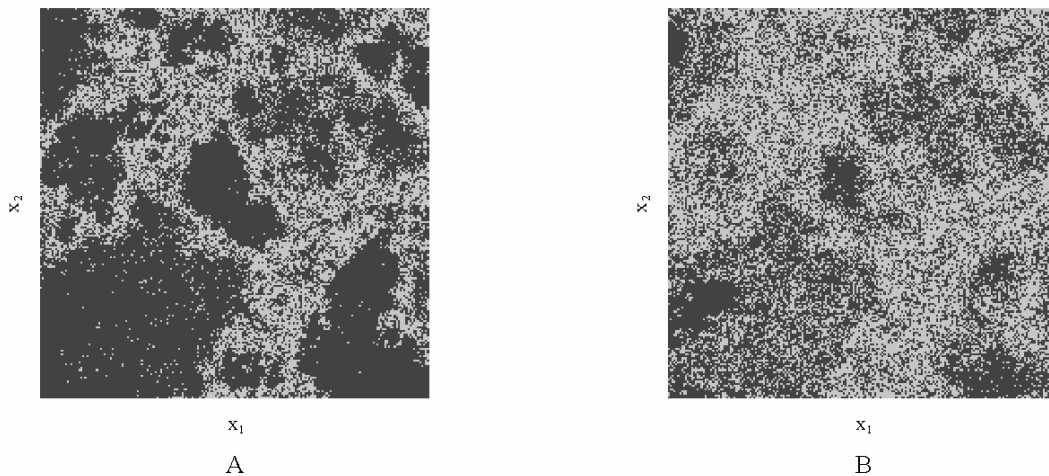


Figure 3.6. Réalisation d'ensembles aléatoires obtenus par troncature de processus hermitiens : A, processus de la figure 3.1A et B, processus de la figure 3.1B

En principe, on obtient une classe plus vaste de covariances d'indicatrices que celle associée aux processus gaussiens tronqués, pour lesquels la covariance du p -ième facteur est égale à la covariance du processus gaussien élevée à la puissance p .

La simulation conditionnelle de ces ensembles aléatoires est identique à celle des gaussiennes tronquées (Freulon, 1992 ; Freulon and de Fouquet, 1993 ; Lantuéjoul, 2002). On génère d'abord un ensemble de valeurs gaussiennes aux sites de données, puis l'on simule le processus hermitien conditionnellement à ces valeurs, enfin on effectue la troncature. La génération des valeurs gaussiennes est basée sur la technique de l'échantillonneur de Gibbs (*cf.* annexe C).

Lorsque le seuil de troncature est le seuil médian ($y = 0$), l'ensemble aléatoire obtenu est autodual. Il est donc clair que la classe des ensembles aléatoires qu'il est possible d'obtenir est loin d'être la plus générale. On peut imaginer étendre la procédure à la troncature de plusieurs processus hermitiens indépendants (généralisation de la méthode des plurigaussiennes au cadre hermitien). Toutefois, l'expression des fonctions de covariance des ensembles aléatoires, déjà compliquée dans le cadre multigaussien, devient inextricable (à plus forte raison, celle de la loi spatiale).

III.1.3. Schéma booléen stationnaire

Cet ensemble aléatoire est constitué par la réunion d'objets ou "grains primaires" implantés aux points d'un processus de Poisson (Matheron, 1967, 1975a ; Lantuéjoul, 2002 ; Stoyan *et al.*, 1996). Sous l'hypothèse de stationnarité, la densité poissonnienne (notée θ) est constante et tous les grains primaires ont des caractéristiques invariantes par translation. Soient G un grain de référence centré à l'origine, $K(\mathbf{h})$ son covariogramme géométrique, B le schéma booléen et $Y(\mathbf{x})$ sa fonction indicatrice.

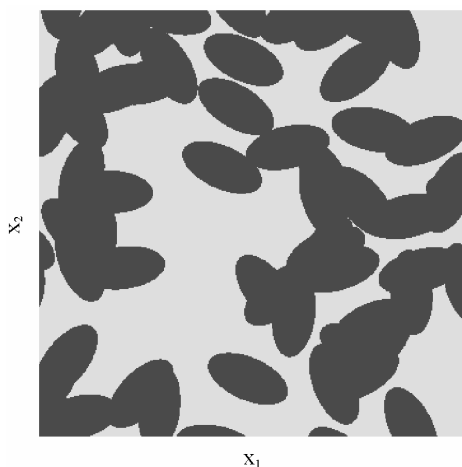


Figure 3.7. Réalisation d'un schéma booléen dont les grains primaires sont des ellipses de taille fixe et orientation aléatoire

Le schéma booléen possède les propriétés suivantes :

- **loi marginale** : $\text{Prob}[Y(\mathbf{x}) = 0] = \text{Prob}[\{\mathbf{x}\} \cap B = \emptyset] = \exp[-\theta K(\mathbf{0})]$ (porosité du schéma booléen)
- **loi bivariable** : $\text{Prob}[Y(\mathbf{x} + \mathbf{h}) = 0, Y(\mathbf{x}) = 0] = \text{Prob}[\{\mathbf{x} + \mathbf{h}, \mathbf{x}\} \cap B = \emptyset] = \exp\{-\theta[2K(\mathbf{0}) - K(\mathbf{h})]\}$
- **fonctionnelle caractéristique** : $\forall K \in \mathcal{K}, Q(K) = \text{Prob}[K \cap B = \emptyset] = \exp\{-\theta E|G \oplus \tilde{K}| \}$

où \mathcal{K} est l'ensemble des compacts de l'espace \mathbb{R}^d

$G \oplus \tilde{K} = \{\mathbf{x} \in \mathbb{R}^d \text{ tels que } K \text{ translaté du vecteur } \mathbf{x} \text{ rencontre } G\}$ est le dilaté de G par K .

Pour éviter des effets de bord, la simulation de B dans un domaine de l'espace requiert de dilater ce domaine du diamètre maximal des grains primaires. C'est dans ce domaine dilaté, noté D , que sont implantés les points poissoniens. Pour conditionner la simulation à un ensemble de données binaires, on utilise une procédure itérative (Lantuéjoul, 2002) :

- *Initialisation* : placer des grains primaires afin de satisfaire les contraintes sur les données ; ces grains "primitifs" devront disparaître ou être recouverts au cours de la simulation, car ils ne respectent pas les caractéristiques structurales ou géométriques de l'ensemble aléatoire à simuler.
- *Phase itérative* : on propose le noyau de transition suivant entre deux états :

a) simuler une variable aléatoire U à valeurs dans $\{-1,0,1\}$ avec le système de probabilité :

$$\begin{cases} \text{Prob}(U = 1) = \frac{\Theta}{\Theta + v + 1} \\ \text{Prob}(U = -1) = \frac{v}{\Theta + v} \\ \text{Prob}(U = 0) = \frac{\Theta}{(\Theta + v + 1)(\Theta + v)} \end{cases}$$

où v est le nombre de grains primaires présents

Θ est le nombre moyen de grains primaires dans le domaine D (produit de la densité θ par la mesure de D).

- b) Si $U = 0$, on ne modifie pas le processus en cours de simulation. Si $U = 1$, on ajoute un grain primaire, tandis que si $U = -1$, on retire un grain choisi au hasard parmi ceux déjà simulés.

D'après la dynamique de Barker (*cf.* annexe C), ce noyau de transition admet pour limite ergodique une loi de Poisson de moyenne Θ : il permet donc la simulation non conditionnelle d'un schéma booléen, vu comme une population de grains primaires dont l'effectif est poissonien (Lantuéjoul, 2002). A présent, pour assurer le conditionnement aux données, la transition est restreinte aux seuls états compatibles avec les données, c'est-à-dire que l'ajout ou le retrait d'un grain primaire n'est accepté que si l'opération n'interrompt pas le conditionnement aux données expérimentales.

On arrête la simulation au bout d'un nombre suffisamment élevé d'itérations.

III.1.4. Ensemble aléatoire des feuilles mortes

Cet ensemble aléatoire n'est autre qu'une mosaïque des feuilles mortes stationnaire avec des grains primaires de valuations binaires (Jeulin, 1997). Il peut être généralisé par le modèle des feuilles mortes colorées à deux phases : à chaque valeur (0 ou 1) sont associés un processus de Poisson et une forme de grain primaire différents. Sauf mention contraire, on utilisera le premier modèle, qui revient en fait à attribuer à chaque phase les mêmes caractéristiques (mêmes densités et mêmes grains primaires).

La figure 3.8 illustre les deux modèles avec des grains primaires elliptiques d'orientation aléatoire et taille fixe ; les proportions de 0 et 1 sont égales. Dans le cas des feuilles mortes colorées biphasées, les grains associés à la valeur 1 ont des rayons deux fois plus élevés que ceux associés à la valeur nulle mais leur proportion est quatre fois moindre afin de conserver les mêmes proportions de 0 et de 1.

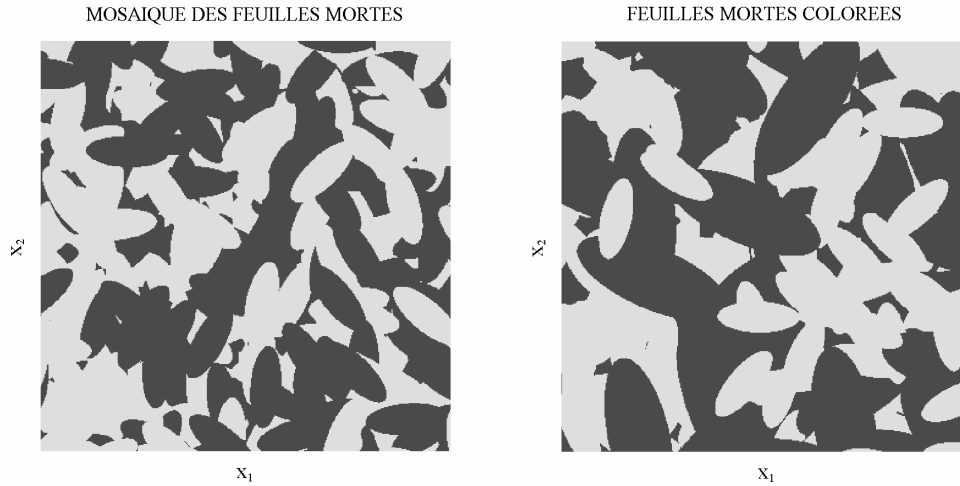


Figure 3.8. Réalisation d'ensembles aléatoires des feuilles mortes dont les grains primaires sont des ellipses de taille fixe et orientation aléatoire

III.2. Modèle à résidus d'indicatrices orthogonaux

III.2.1. Modèle discret sans effets de bord en descendant

Ce modèle peut être obtenu par simulation hiérarchique (cf. chapitre 2) en posant :

$$\forall \mathbf{x} \in \mathbb{R}^d, Y(\mathbf{x}) = \max \{i / \mathbf{x} \in E_i\} \quad (3.3)$$

où $E_0 = \mathbb{R}^d$ et $\{E_i, i = 1 \dots m\}$ sont des ensembles aléatoires mutuellement indépendants. La figure 3.9 illustre le mode de construction dans les cas où les ensembles aléatoires sont des schémas booléens.

Le conditionnement d'un tel modèle à des données expérimentales requiert seulement de savoir mener des simulations conditionnelles des ensembles aléatoires. Une donnée conditionnante $Y(\mathbf{x}) = y$ est active sur les ensembles $\{E_i, i = y \dots m\}$ et inactive sur les autres ensembles $\{E_i, i = 1 \dots y-1\}$:

$$\mathbf{x} \in E_y \text{ et } \forall i > y, \mathbf{x} \notin E_i .$$

Ce modèle discret a été introduit par J. Rivoirard (1988, 1989), qui a démontré l'isofactorialité des lois bivariées et exhibé les facteurs ("résidus d'indicatrices") :

$$\chi_0[Y(\mathbf{x})] = 1 \text{ et } \forall p \in \{1, \dots, m\}, \chi_p[Y(\mathbf{x})] = \frac{I(Y(\mathbf{x}); p)}{E[I(Y(\mathbf{x}); p)]} - \frac{I(Y(\mathbf{x}); p+1)}{E[I(Y(\mathbf{x}); p+1)]} .$$

Il s'agit d'une généralisation du modèle mosaïque : l'espace est divisé en compartiments au sein desquels la fonction aléatoire prend une valeur constante, mais cette valeur dépend de la taille du compartiment. Une propriété intéressante est l'absence d'*effets de bord en descendant* : si, pour tout i dans $\{1, \dots, m\}$, on définit les ensembles A_i des sites à valeurs inférieures ou égales à i , la probabilité de trouver une valeur inférieure à $j < i$ est uniforme dans A_i . Autrement dit, les sites proches de la frontière de A_i ont autant de chance de prendre des valeurs inférieures à i que les sites centraux de A_i . On peut identifier l'absence d'effets de bord en descendant en comparant les covariances simples et croisées d'indicatrices : la covariance croisée entre deux indicatrices est proportionnelle à la covariance simple de l'indicatrice associée au seuil le plus élevé.

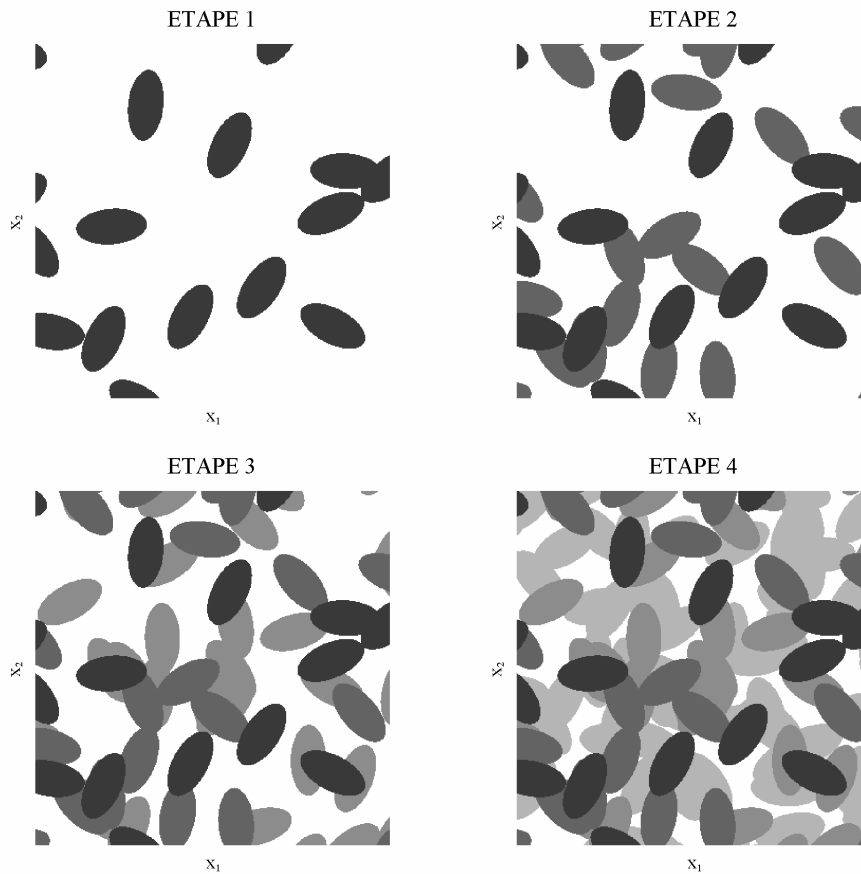


Figure 3.9. Construction du modèle à résidus d'indicateurs orthogonaux par hiérarchisation d'ensembles aléatoires booléens indépendants

III.2.2. Modèle discret sans effets de bord en montant

En inversant la construction précédente et en posant

$$\forall \mathbf{x} \in \mathbb{R}^d, Y(\mathbf{x}) = \min \{i / \mathbf{x} \in E_i\} \text{ avec } E_m = \mathbb{R}^d \quad (3.4)$$

on définit un modèle sans **effets de bord en montant** : sur l'ensemble des sites à valeurs supérieures ou égales à i , il existe une probabilité uniforme de trouver une valeur supérieure à $j > i$. Cette fois, la covariance croisée entre deux indicateurs est proportionnelle à la covariance simple de l'indicateur associée au seuil le plus faible.

En géostatistique minière, le modèle sans effets de bord en montant peut servir à décrire des veines minéralisées à l'intérieur desquelles les teneurs sont totalement pépétiques. L'appartenance à une veine minéralisée (teneur positive), que le site soit proche ou non de la frontière de la veine, ne renseigne pas sur la teneur effective, qui peut prendre n'importe quelle valeur positive. La probabilité de trouver une teneur élevée est identique en tout site de la veine.

Seul le cas où toutes les covariances simples et croisées d'indicateurs sont proportionnelles réunit l'absence d'effets de bord en montant et en descendant : les indicateurs sont en corrélation intrinsèque, ce qui correspond au *modèle mosaïque*. Toutefois, les constructions proposées [Eq. (3.3) et (3.4)] ne permettent pas d'obtenir un tel modèle où les effets de bord sont absents dans les deux sens. Elles coïncident avec des fonctions aléatoires booléennes à grains primaires cylindriques (Serra, 1988), ou encore avec des modèles de feuilles mortes colorées dont les grains primaires dépendent du temps.

Inférence des paramètres (modèle sans effets de bord en montant)

On détermine les caractéristiques des ensembles aléatoires de départ via leur impact sur les lois marginale et bivariable de $Y(\mathbf{x})$. Dans la suite, on notera $\{p_i, i = 1 \dots m\}$ les moyennes des ensembles aléatoires $\{E_i, i = 1 \dots m\}$ et $\{\rho_i(\mathbf{h}), i = 1 \dots m\}$ leurs corrélogrammes.

- Loi marginale

$$\forall i \in \{0, \dots, m\}, T_i = \text{Prob}[Y(\mathbf{x}) > i] = \prod_{k=0}^i \text{Prob}[\mathbf{x} \notin E_k] = \prod_{k=0}^i (1 - p_k).$$

- Loi bivariable : si \wedge et \vee désignent les opérateurs minimum et maximum respectivement, alors

$$\forall i, j \in \{0, \dots, m\}, T_{i,j}(\mathbf{h}) = \text{Prob}[Y(\mathbf{x} + \mathbf{h}) > i, Y(\mathbf{x}) > j] = \prod_{k=0}^{i \wedge j} \{1 - p_k [1 - \rho_k(\mathbf{h})]\} \prod_{k=0}^{i \vee j} (1 - p_k)$$

d'où l'on tire la relation

$$\forall i, j \in \{0, \dots, m\}, T_{i,j}(\mathbf{h}) = \frac{T_{i \vee j}}{T_{i \wedge j}} T_{i \wedge j, i \wedge j}(\mathbf{h})$$

En effet, la probabilité cherchée s'exprime sous la forme :

$$\text{Prob}[Y(\mathbf{x} + \mathbf{h}) > i, Y(\mathbf{x}) > j] = \prod_{k=0}^{i \wedge j} \text{Prob}[\mathbf{x} + \mathbf{h} \notin E_k, \mathbf{x} \notin E_k] \prod_{k=i \vee j}^{i \vee j} \text{Prob}[\mathbf{x} \notin E_k]$$

Or, pour l'ensemble aléatoire E_k , de variogramme $\gamma_k(\mathbf{h}) = p_k(1 - p_k)[1 - \rho_k(\mathbf{h})]$, la probabilité que deux sites séparés de \mathbf{h} soit hors de cet ensemble est :

$$\text{Prob}[\mathbf{x} + \mathbf{h} \notin E_k, \mathbf{x} \notin E_k] = 1 - \text{Prob}[\mathbf{x} \in E_k] - \text{Prob}[\mathbf{x} + \mathbf{h} \in E_k, \mathbf{x} \notin E_k] = 1 - p_k - \gamma_k(\mathbf{h}) = (1 - p_k)\{1 - p_k[1 - \rho_k(\mathbf{h})]\}$$

On en déduit la formule annoncée ci-dessus.

III.2.2. Modèle continu

Dans le modèle décrit précédemment, si $Y(\mathbf{x})$ prend un nombre élevé de valeurs, il est possible de les regrouper en un petit nombre de classes adjacentes ; un tel regroupement préserve le caractère isofactoriel. En effet, si l'on décide de supprimer la valeur i , il suffit d'éliminer le résidu χ_i associé et de remplacer le résidu suivant χ_{i+1} par $\chi_i + \chi_{i+1}$, les autres résidus restant inchangés. Cette opération est réversible : on peut ajouter des valeurs au lieu d'en supprimer, c'est-à-dire que l'on peut augmenter le niveau de discrétisation (Rivoirard, 1989).

En augmentant à l'extrême la discrétisation, on obtient un modèle où les ensembles aléatoires sont indexés par un paramètre positif qui varie continûment et sera assimilé à un temps. La description du modèle est la suivante : entre les instants t et $t + dt$, on laisse tomber un ensemble aléatoire de moyenne $p(dt)$ infinitésimale¹ et corrélogramme $\rho_t(\mathbf{h})$; on retient pour $Y(\mathbf{x})$ la valeur de t à laquelle le site \mathbf{x} a été recouvert pour la première fois :

$$\forall \mathbf{x} \in \mathbb{R}^d, Y(\mathbf{x}) = \inf\{t / \mathbf{x} \in E_t\} \quad (3.5)$$

¹ En pratique, cet ensemble aléatoire sera un *schéma booléen infinitésimal*, c'est-à-dire un schéma booléen dont les points poissonniens ont une densité $\theta(dt)$ infiniment petite.

Pour obtenir les équations du modèle continu, on remplace l'opérateur produit par l'exponentielle de l'intégrale du logarithme et l'on utilise le développement de la fonction logarithme au premier ordre : si ε est infiniment petit, $\ln(1 - \varepsilon) \approx -\varepsilon$.

- Loi marginale

$$\forall t \in \mathbf{R}_+, 1 - F(t) = \text{Prob}[Y(\mathbf{x}) > t] = \exp\left\{\int_0^t \ln[1 - p(du)]\right\} = \exp\left\{-\int_0^t p(du)\right\}.$$

L'inversion de cette relation permet de choisir les moyennes des ensembles aléatoires de manière à reproduire une distribution donnée :

$$\forall t \in \mathbf{R}_+, p(dt) = \frac{F(dt)}{1 - F(t)}.$$

Au passage, on s'assure que tout point \mathbf{x} est recouvert en un temps presque sûrement fini.

- Loi bivariable

$$\begin{aligned} \forall t, t' \in \mathbf{R}_+, \text{Prob}[Y(\mathbf{x} + \mathbf{h}) > t, Y(\mathbf{x}) > t'] &= \exp\left\{\int_0^{t \wedge t'} -[1 - \rho_u(\mathbf{h})] p(du)\right\} \exp\left\{-\int_0^{t \vee t'} p(du)\right\} \\ &= [1 - F(t \wedge t')][1 - F(t \vee t')] \exp\left\{\int_0^{t \wedge t'} \rho_u(\mathbf{h}) p(du)\right\} \\ &= [1 - F(t)][1 - F(t')] \exp\left\{\int_0^{t \wedge t'} \rho_u(\mathbf{h}) p(du)\right\} \end{aligned}$$

En particulier, si le corrélogramme est indépendant du temps, il vient :

$$\forall t, t' \in \mathbf{R}_+, \text{Prob}[Y(\mathbf{x} + \mathbf{h}) > t, Y(\mathbf{x}) > t'] = [1 - F(t)][1 - F(t')][1 - F(t \wedge t')]^{-\rho(\mathbf{h})}.$$

Une telle loi est associée à la "densité"

$$\forall t, t' \in \mathbf{R}_+, F_h(dt, dt') = \rho(\mathbf{h})[1 - F(t)]^{1-\rho(\mathbf{h})} F(dt) \delta_{tt'} + [1 - \rho(\mathbf{h})][1 - F(t \wedge t')]^{-\rho(\mathbf{h})} F(dt) F(dt').$$

En effet, supposons que $t \geq t'$ et montrons que l'on a : $\text{Prob}[Y(\mathbf{x} + \mathbf{h}) > t, Y(\mathbf{x}) > t'] = \int_t^{+\infty} \int_{t'}^{+\infty} F_h(du, dv)$

D'une part

$$\int_t^{+\infty} \int_{t'}^{+\infty} \rho(\mathbf{h})[1 - F(u)]^{1-\rho(\mathbf{h})} F(du) \delta_{uv} = \int_t^{+\infty} \rho(\mathbf{h})[1 - F(u)]^{1-\rho(\mathbf{h})} F(du) = \frac{\rho(\mathbf{h})}{2 - \rho(\mathbf{h})} [1 - F(t)]^{2-\rho(\mathbf{h})}$$

D'autre part,

$$\begin{aligned} &\int_t^{+\infty} \int_{t'}^{+\infty} [1 - \rho(\mathbf{h})][1 - F(u \wedge v)]^{-\rho(\mathbf{h})} F(du) F(dv) \\ &= \int_t^{+\infty} \int_{t'}^u [1 - \rho(\mathbf{h})][1 - F(v)]^{-\rho(\mathbf{h})} F(du) F(dv) + \int_t^{+\infty} \int_u^{+\infty} [1 - \rho(\mathbf{h})][1 - F(u)]^{-\rho(\mathbf{h})} F(du) F(dv) \\ &= \int_t^{+\infty} \{[1 - F(t')]^{1-\rho(\mathbf{h})} - [1 - F(u)]^{1-\rho(\mathbf{h})}\} F(du) + \int_t^{+\infty} [1 - \rho(\mathbf{h})][1 - F(u)]^{-\rho(\mathbf{h})} [1 - F(u)] F(du) \\ &= [1 - F(t')]^{1-\rho(\mathbf{h})} [1 - F(t)] - \int_t^{+\infty} \rho(\mathbf{h})[1 - F(u)]^{1-\rho(\mathbf{h})} F(du) \\ &= [1 - F(t')]^{1-\rho(\mathbf{h})} [1 - F(t)] - \frac{\rho(\mathbf{h})}{2 - \rho(\mathbf{h})} [1 - F(t)]^{2-\rho(\mathbf{h})} \end{aligned}$$

On notera l'analogie entre ces formules et celles du modèle mosaïque (Matheron, 1982, 1984c). Elles peuvent d'ailleurs être retrouvées grâce à la théorie des fonctions aléatoires booléennes à grains primaires cylindriques (Serra, 1988), avec lesquelles la construction proposée [Eq. (3.5)] coïncide.

Isofactorialité des lois bivariées

Les lois bivariées sont isofactorielles dès que l'on considère un regroupement des valeurs en un nombre fini ou dénombrable de classes, puisque l'on retrouve alors le modèle discret (*cf.* paragraphe III.2.1). Toutefois, le modèle continu n'est pas isofactoriel, car la propriété de commutation évoquée en annexe B est mise en défaut :

$$\forall t, v \in \mathbb{R}_+, \forall \mathbf{h}, \mathbf{h}' \in \mathbb{R}^d, \int_{u=0}^{+\infty} \frac{F_{\mathbf{h}}(dt, du)}{\sqrt{F(dt)F(du)}} \frac{F_{\mathbf{h}'}(du, dv)}{\sqrt{F(du)F(dv)}} \neq \int_{u=0}^{+\infty} \frac{F_{\mathbf{h}'}(dt, du)}{\sqrt{F(dt)F(du)}} \frac{F_{\mathbf{h}}(du, dv)}{\sqrt{F(du)F(dv)}}$$

Cette inégalité peut être établie en considérant un corrélogramme indépendant du temps et une loi marginale exponentielle (ce qui revient à supposer que la densité des schémas booléens infinitésimaux est constante dans le temps, situation à laquelle on peut toujours se ramener par anamorphose).

Simulation conditionnelle

La simulation du modèle continu, construit à partir de schémas booléens infinitésimaux, revient à simuler un processus ponctuel "marqué", c'est-à-dire un ensemble de germes poissoniens auxquels sont associés des "objets" (les grains primaires valués). Ces points sont situés dans l'espace-temps $\mathbb{R}^d \times \mathbb{R}_+$; leurs coordonnées spatiales sont en fait restreintes à un domaine fini, noté D ci-dessous, correspondant au domaine à simuler dilaté du rayon maximal des grains primaires. La simulation peut se décomposer en trois étapes :

- 1) initialisation : recouvrement des données conditionnantes $\{y(\mathbf{x}_\alpha), \alpha = 1 \dots n\}$;
- 2) simulation des valeurs comprises entre zéro et la plus forte donnée conditionnante, soit y_{\max} ;
- 3) simulation des valeurs supérieures à y_{\max} .

Pour mener à bien l'étape 1), on ordonne les données par valeurs croissantes. Chaque groupe de valeurs égales doit être recouvert par un grain primaire² germé dans D qui ne rencontre aucune des données de plus faibles valeurs ; une procédure d'acceptation et de rejet peut être utilisée à cet effet. Chaque grain primaire est ensuite valué selon la valeur commune des données qu'il recouvre, ce qui définit sa "date" d'apparition le long de l'axe temporel.

L'étape 2) revient à simuler une population poissonienne de germes sur $D \times [0, y_{\max}]$ qui respecte les contraintes induites par les données. Elle peut être réalisée par itérations markoviennes. A chaque itération, un germe poissonien associé à un grain primaire peut être ajouté ou retiré ; l'ajout ou le retrait est accepté s'il n'interrompt pas les conditions sur les données (Lantuéjoul, 2002 ; voir aussi l'annexe C). Les grains obtenus lors de l'étape 1) sont inamovibles, car il est en principe impossible de les retirer sans interrompre le conditionnement et également impossible d'ajouter un grain primaire dont la valuation coïncide exactement avec une donnée.

L'étape 3) consiste simplement à poursuivre la simulation sans contrainte de conditionnement sur $D \times]y_{\max}, \infty[$, jusqu'à ce que le domaine à simuler soit totalement recouvert.

² En théorie, on devrait être capable de recouvrir toutes les données égales avec un seul grain primaire car, dans le modèle, deux grains primaires distincts ont des valuations presque sûrement différentes.

III.3. Modèle binomial

D'après Matheron (1980a), un modèle isofactoriel de diffusion à loi marginale binomiale de taille N et paramètre p peut être obtenu de façon exacte en additionnant N ensembles aléatoires stationnaires indépendants de même moyenne p et même corrélogramme $\rho(\mathbf{h})$. Dans ce modèle, la covariance du p -ième facteur (polynôme de Krawtchouk normé) est égale à $\rho(\mathbf{h})$ élevé à la puissance p . Si N tend vers l'infini, le modèle binomial converge vers un modèle de diffusion gaussien (Matheron, 1984a), ce qui est cohérent avec la construction proposée en vertu du théorème de la limite centrale.

Parmi les modèles d'ensembles aléatoires, deux retiendront notre attention (le schéma booléen et les feuilles mortes), car ils facilitent considérablement le conditionnement de la simulation. Dans les deux cas, les ensembles aléatoires sont construits par lancement de grains primaires, dont les germes sont situés dans un domaine, noté D , égal au domaine à simuler dilaté du diamètre maximal des grains primaires. Bien que ces modèles ne soient pas les plus généraux³, ils peuvent s'adapter à la plupart des situations pratiques (la classe des corrélogrammes accessibles est déjà très vaste).

Le conditionnement est basé sur la technique du recuit simulé (Kirkpatrick *et al.*, 1983 ; Hegstad *et al.*, 1994) dont le principe est détaillé en annexe C. Le point de départ peut être une simulation non conditionnelle ; la mise en œuvre du recuit simulé repose sur les noyaux de transitions ci-dessous.

Somme de schémas booléens stationnaires (figure 3.10)

On utilise le noyau de transition défini comme suit :

- a) sélectionner un indice i au hasard dans $\{1, \dots, N\}$;
- b) simuler une variable aléatoire U pouvant prendre les valeurs $\{-1, 0, 1\}$ avec les probabilités

$$\text{Prob}(U = 1) = \frac{\Theta}{\Theta + v_i + 1}, \quad \text{Prob}(U = -1) = \frac{v_i}{\Theta + v_i} \quad \text{et} \quad \text{Prob}(U = 0) = \frac{\Theta}{(\Theta + v_i + 1)(\Theta + v_i)}$$

où v_i est le nombre de grains primaires présents dans le i -ème schéma booléen
 Θ est le nombre moyen de grains primaires souhaité dans D pour chaque schéma booléen.

- c) Si $U = 1$, on ajoute un grain primaire au i -ème schéma booléen, tandis que si $U = -1$, on retire un grain choisi au hasard parmi ceux déjà simulés dans le i -ème schéma booléen.

Somme d'ensembles aléatoires des feuilles mortes (figure 3.11)

Un noyau de transition pour simuler une mosaïque des feuilles mortes est proposé en annexe C. Dans le cas qui nous intéresse, il est précédé de la sélection d'un indice i au hasard dans $\{1, \dots, N\}$ et appliqué au i -ème ensemble aléatoire des feuilles mortes.

³ Le corrélogramme obtenu $\rho(\mathbf{h})$ est toujours positif, alors qu'en théorie le coefficient de corrélation d'une loi de diffusion binomiale peut prendre des valeurs négatives. Par ailleurs, $\rho(\mathbf{h})$ est nécessairement le corrélogramme d'une variable binaire, ce qui interdit les modèles incompatibles avec cette distribution (schéma gaussien par exemple) ; toutefois, on ignore quelle classe de modèles variographiques est compatible avec la loi de diffusion binomiale.

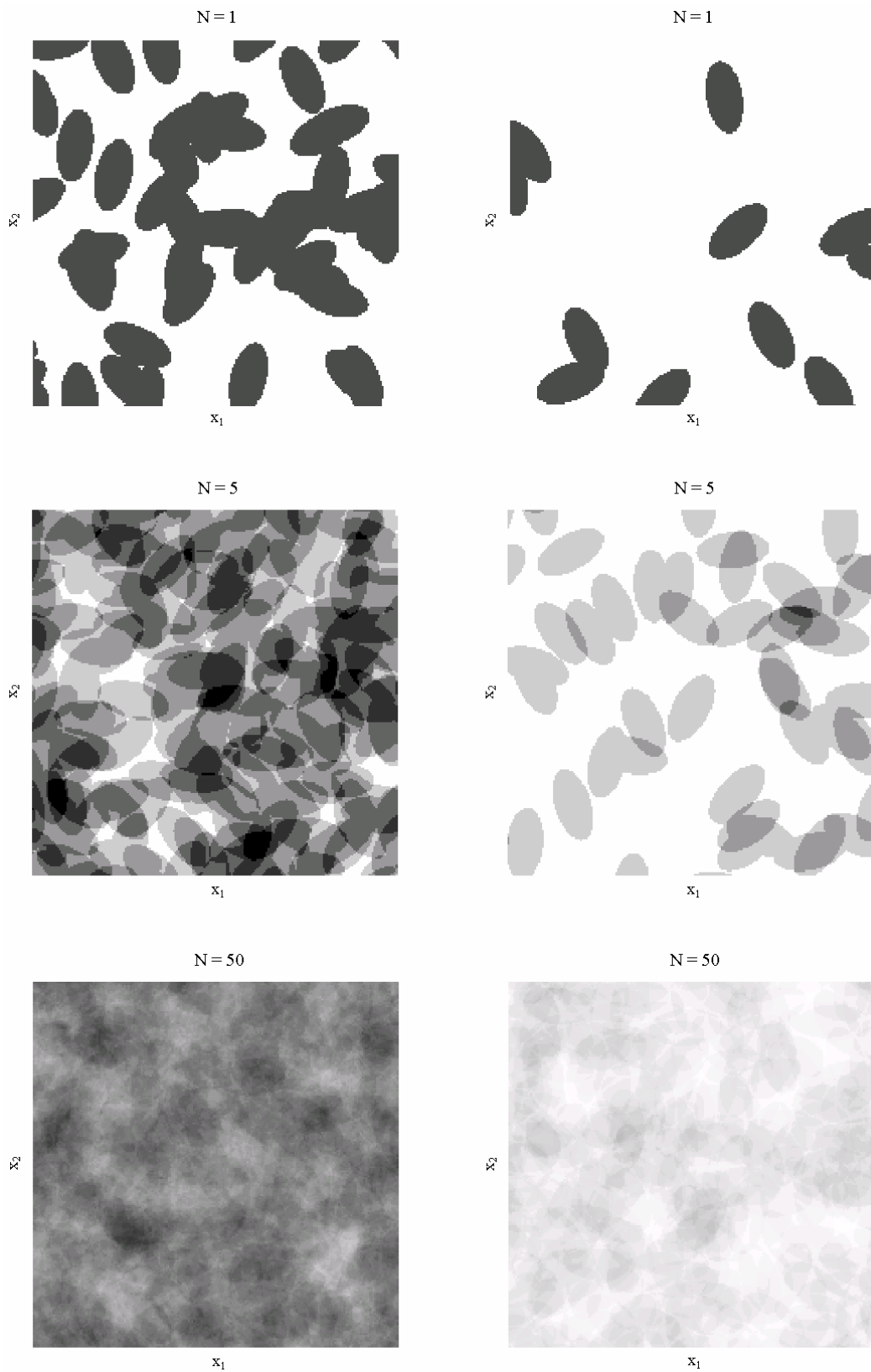


Figure 3.10. Réalisations de processus binomiaux : $p = 0.5$ (gauche) et $p = 0.1$ (droite) par addition de schémas booléens dont les grains primaires sont des ellipses de taille fixe et orientation aléatoire

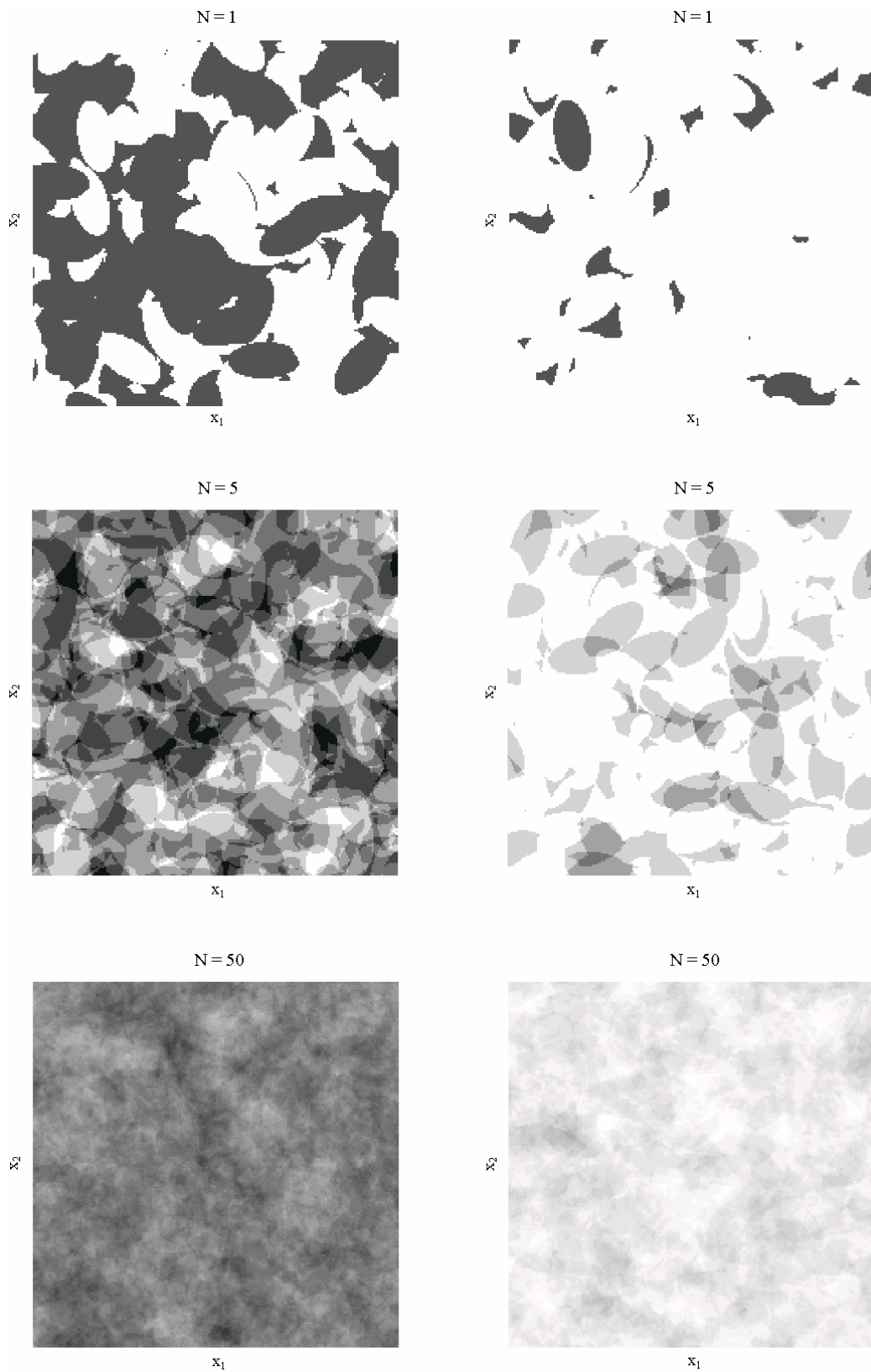


Figure 3.11. Réalisations de processus binomiaux : $p = 0.5$ (gauche) et $p = 0.1$ (droite) par addition d'ensembles aléatoires de feuilles mortes dont les grains primaires sont des ellipses de taille fixe et orientation aléatoire

IV. Modèles à lois indéfiniment divisibles

Une variable aléatoire est indéfiniment divisible si, pour n'importe quel entier N , elle peut s'écrire comme la somme de N variables aléatoires indépendantes et identiquement distribuées. Cette propriété, qui concerne la loi marginale, peut être étendue aux lois bivariées. Elle va permettre de simuler des modèles isofactoriels indéfiniment divisibles par sommation de plusieurs modèles élémentaires. Dans la suite, nous analyserons trois procédés pour mener à bien cette sommation : le premier est celui des bandes tournantes, le second est le modèle de dilution poissonnienne, tandis que le troisième consiste à additionner plusieurs modèles mosaïques à loi marginale indéfiniment divisible.

IV.1. Simulation par bandes tournantes

Dans le cas multigaussien, on connaît une technique générale permettant de passer d'un processus unidimensionnel à un processus multidimensionnel : la méthode des **bandes tournantes** (Matheron, 1972b, 1973a ; Freulon et de Fouquet, 1991 ; Lantuéjoul, 1994). Dans cette section, nous proposons une extension de cette méthode à d'autres modèles isofactoriels.

IV.1.1. Principe des bandes tournantes

Les ingrédients de la méthode sont les suivants :

- une fonction aléatoire unidimensionnelle $\{X(t), t \in \mathbb{R}\}$ de covariance continue, notée $C_X(\Delta t)$;
- une droite Δ de \mathbb{R}^d et un vecteur directeur unitaire \mathbf{u} de Δ .

Dans un premier temps, on pose :

$$\forall \mathbf{x} \in \mathbb{R}^d, Y_{\mathbf{u}}(\mathbf{x}) = X(\underbrace{\langle \mathbf{x} | \mathbf{u} \rangle}_{\substack{\text{projection} \\ \text{de } \mathbf{x} \text{ sur } \Delta}}) \text{ (formule d'épandage).}$$

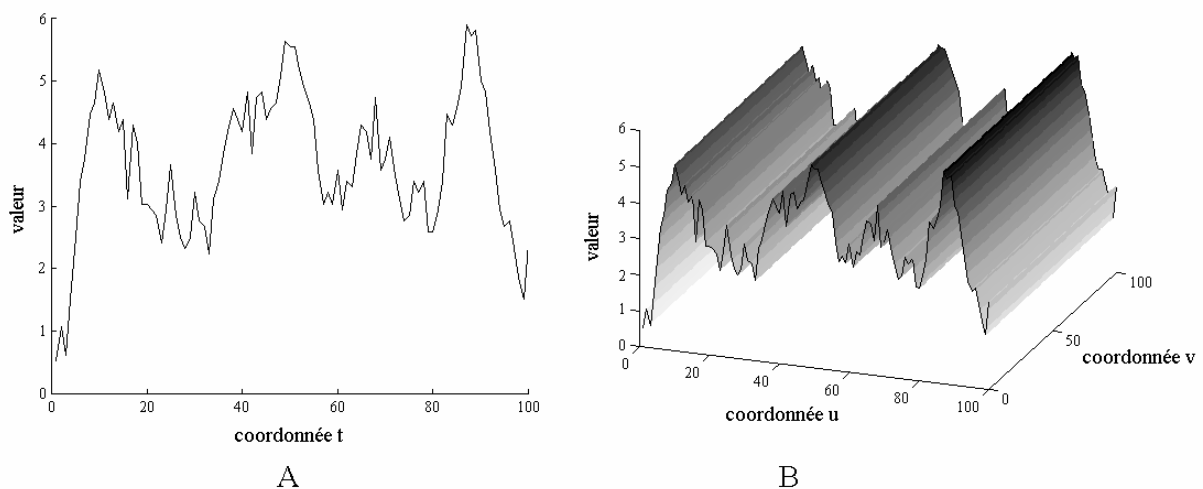


Figure 3.12. A, réalisation d'un processus unidimensionnel et B, processus bidimensionnel obtenu par épandage

La fonction aléatoire $Y_{\mathbf{u}}$ n'est pas isotrope car ses réalisations sont constantes sur les hyperplans orthogonaux à \mathbf{u} . On la rend isotrope en randomisant l'orientation de la droite Δ : on remplace \mathbf{u} par une variable aléatoire \mathbf{U} uniforme sur la sphère unité de \mathbb{R}^d , d'où la définition finale de $Y(\mathbf{x})$:

$$\forall \mathbf{x} \in \mathbb{R}^d, Y(\mathbf{x}) = X(\langle \mathbf{x} | \mathbf{U} \rangle).$$

La covariance de $Y(\mathbf{x})$ s'exprime de la manière suivante :

$$C_Y(\mathbf{h}) = E \{ C_X(\langle \mathbf{h} | \mathbf{U} \rangle) \}.$$

La technique des bandes tournantes peut être vue comme un cas particulier de substitution, pour laquelle la fonction directrice $T(\mathbf{x})$ est une projection orthogonale sur une droite d'orientation \mathbf{U} aléatoire et uniforme :

$$\forall \mathbf{x} \in \mathbb{R}^d, T(\mathbf{x}) = \langle \mathbf{x} | \mathbf{U} \rangle.$$

Ce processus correspond à une fonction aléatoire intrinsèque de variogramme parabolique.

IV.1.2. Applicabilité aux modèles isofactoriels ?

La méthode des bandes tournantes, utilisable dans le cadre multigaussien, peut-elle être étendue à d'autres modèles isofactoriels ? *A priori*, la situation semble favorable : si l'on utilise un processus de codage $\{X(t), t \in \mathbb{R}\}$ à loi isofactorielle, on obtiendra une fonction aléatoire également isofactorielle et définie dans l'espace \mathbb{R}^d car on sait que la substitution conserve le caractère isofactoriel.

Or, dans la mise en œuvre des bandes tournantes, la procédure d'épandage est répétée pour de nombreuses directions $\{\mathbf{U}_i, i = 1 \dots N\}$, tirées au hasard ou quasi-régulièrement espacées⁴ sur la sphère unité de \mathbb{R}^d . La simulation finale est la somme des différentes simulations élémentaires :

$$Y(\mathbf{x}) = \sum_{i=1}^N Y_i(\mathbf{x}) = \sum_{i=1}^N X_i(\langle \mathbf{x} | \mathbf{U}_i \rangle) \quad (3.6)$$

où $\{X_i, i = 1 \dots N\}$ sont des processus de moyenne nulle et variance $1/N$.

Cette pratique répond à une double préoccupation :

- d'une part, assurer la convergence du processus simulé vers un processus multigaussien d'après le théorème de la limite centrale (Feller, 1966) ;
- d'autre part, garantir l'ergodicité de ce processus : les réalisations de chaque $Y_i(\mathbf{x})$ sont constantes le long d'hyperplans de l'espace, donc présentent une anisotropie non contenue dans la covariance théorique ; cette situation disparaît lorsque l'on additionne un grand nombre de composantes élémentaires [Eq. (3.6)].

⁴ Dans l'espace à trois dimensions, le plus grand nombre de droites régulièrement réparties qu'il est possible de construire est égal à 15 : il s'agit des droites joignant les milieux des arêtes opposées d'un icosaèdre régulier (Guibal, 1972 ; Journel, 1974). Pour augmenter ce nombre, une alternative consiste à utiliser une séquence de directions *équiréparties*, c'est-à-dire une suite de points sur la sphère unité telle que la proportion de points tombant dans un angle solide déterminé parmi les N premiers points converge vers la proportion de l'espace occupée par cet angle solide lorsque N devient infini. A titre d'exemple, citons les séquences de Van der Corput (Freulon et de Fouquet, 1991 ; Freulon, 1992 ; Lantuéjoul, 1994) qui sont basées sur les décompositions binaires et ternaires des entiers $i = 1 \dots N$, N étant le nombre total de droites. Dans l'espace bidimensionnel, il n'existe pas de telles restrictions et l'on peut toujours construire une séquence contenant un nombre quelconque de droites régulièrement réparties.

Bien que l'on cherche à s'affranchir du caractère multigaussien de la simulation, la contrainte d'ergodicité rend inévitable la sommation de plusieurs composantes. Pour s'en sortir, il faut utiliser une famille de lois isofactorielles qui soit stable par addition. Parmi les lois de diffusion absolument continues, c'est le cas de la loi bigaussienne et de la loi bigamma⁵. Malheureusement, cette dernière ne peut être obtenue par la procédure des bandes tournantes, car l'épandage d'un processus bigamma unidimensionnel conduit à un processus de loi de Laguerre.

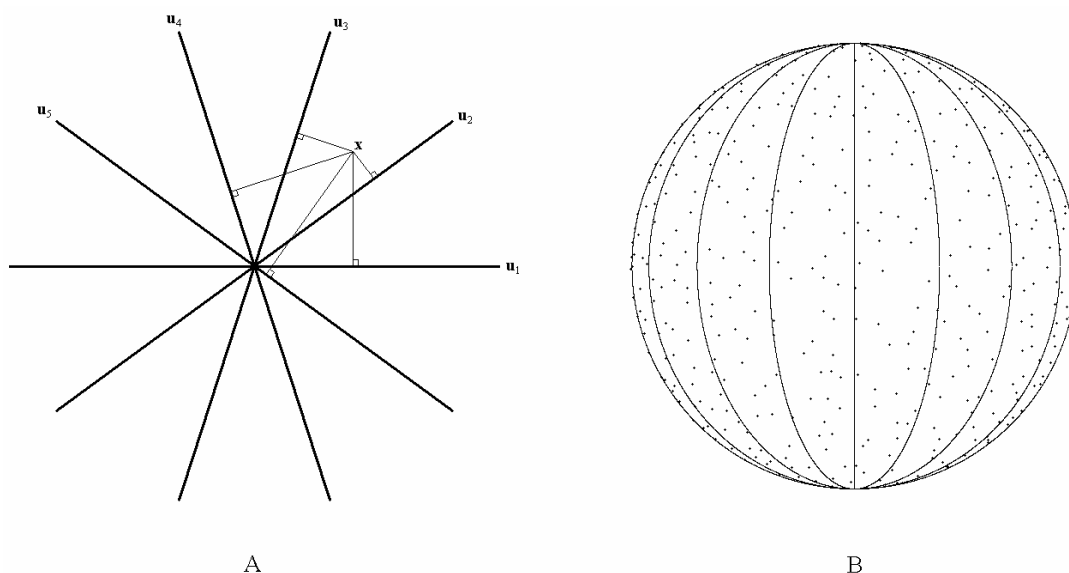


Figure 3.13. Directions A, régulières dans le plan et B, équiréparties dans l'espace tridimensionnel

IV.1.3. Exemple de la loi de Laguerre

En général, la propriété de stabilité pour l'addition est mise en défaut pour la loi de Laguerre. Une exception notable a lieu lorsque les covariances des facteurs successifs sont les moments d'une loi beta de paramètres $\{C_Y(\mathbf{h}), C_Y(\mathbf{0}) - C_Y(\mathbf{h})\}$. Matheron (1973b) a en effet montré qu'un couple $\{Y(\mathbf{x} + \mathbf{h}), Y(\mathbf{x})\}$ admettant cette loi laguerrienne s'écrit sous la forme

$$\begin{cases} Y(\mathbf{x} + \mathbf{h}) = U + V \\ Y(\mathbf{x}) = U + W \end{cases} \quad (3.7)$$

où U, V, W sont trois variables indépendantes à loi gamma, de paramètres $C_Y(\mathbf{h})$ pour la première et $C_Y(\mathbf{0}) - C_Y(\mathbf{h})$ pour les deux autres ; $C_Y(\mathbf{h})$ est la covariance entre $Y(\mathbf{x} + \mathbf{h})$ et $Y(\mathbf{x})$.

⁵ Considérons la transformée de Laplace d'un couple de valeurs $\{Y_1, Y_2\}$ ayant une loi bigamma de paramètre α et corrélation ρ (Matheron, 1973b ; Hu, 1988) :

$$\phi(\lambda, \mu) = E[e^{-\lambda Y_1 - \mu Y_2}] = [1 + \lambda + \mu + \lambda\mu(1 - \rho)]^{-\alpha}.$$

La somme de deux processus indépendants de loi bigamma de paramètres α et α' respectivement et de même corrélogramme $\rho(\mathbf{h})$ aura pour transformée de Laplace le produit des transformées de Laplace de chaque processus. Celle-ci s'identifie à la transformée de Laplace d'un processus bigamma de corrélogramme $\rho(\mathbf{h})$ et paramètre $\alpha + \alpha'$, ce qui établit le résultat annoncé. Cette propriété de stabilité pour l'addition ne contredit pas le théorème de la limite centrale, car la loi bigamma converge vers la loi bigaussienne lorsque son paramètre tend vers l'infini.

Sous cette forme, on voit que la somme de deux processus indépendants admettant cette loi de Laguerre, de paramètres α_1 et α_2 et covariances $C_1(\mathbf{h})$ et $C_2(\mathbf{h})$ respectivement, fournit un processus de même loi, mais de paramètre $\alpha_1 + \alpha_2$ et covariance $C_1(\mathbf{h}) + C_2(\mathbf{h})$.

D'après Matheron (1973b), un tel processus laguerrien peut être obtenu par régularisation d'une mesure aléatoire orthogonale à loi gamma. Lantuéjoul (1984) explicite un mode de construction sur une maille régulière dans l'espace unidimensionnel. Le processus de départ est une "marche aléatoire" à loi gamma $\{W(t), t \in \mathbb{R}\}$, laquelle se caractérise par les propriétés suivantes :

- 1) pour tout $\Delta t \geq 0$, l'accroissement entre deux valeurs $W(t + \Delta t) - W(t)$ a une loi gamma de paramètre $\alpha \Delta t / a$ où a est un paramètre positif fixé une fois pour toutes ; en conséquence, le variogramme de $W(t)$ est un schéma linéaire de pente $\alpha / 2a$;
- 2) les accroissements $W(t_1 + \Delta t_1) - W(t_1)$ et $W(t_2 + \Delta t_2) - W(t_2)$ sont indépendants dès lors que les intervalles $[t_1, t_1 + \Delta t_1]$ et $[t_2, t_2 + \Delta t_2]$ ont au plus un point commun.

On pose finalement : $\forall t \geq 0, X(t) = W(t + a) - W(t)$. La génération des valeurs est aisée si la maille de simulation est discrète, même lorsqu'elle est irrégulière.

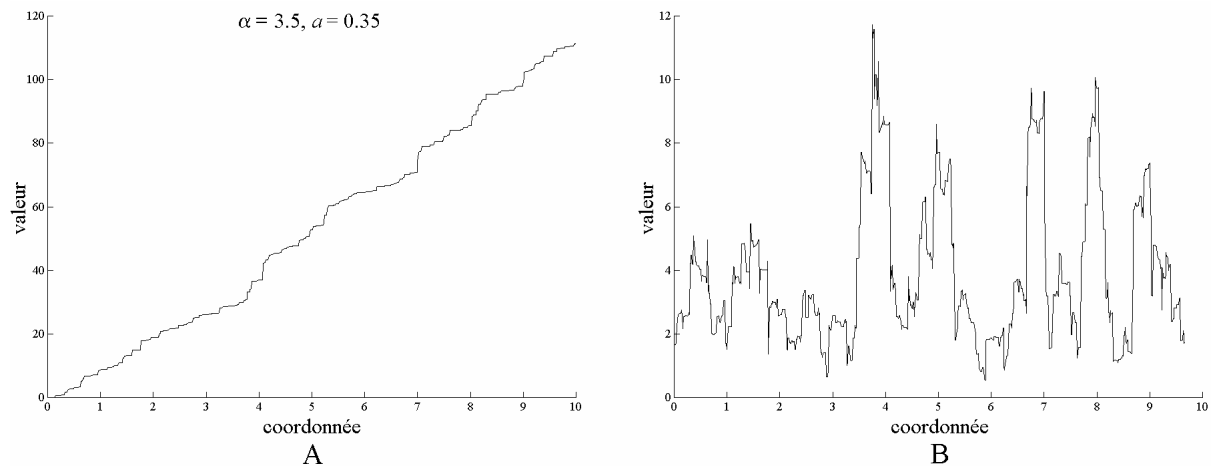


Figure 3.14. A, réalisation d'une marche aléatoire gamma et B, processus obtenu par accroissement sur un intervalle fixé

$X(t)$ a une loi marginale gamma de paramètre α et une loi bivariable isofactorielle de Laguerre [Eq. (3.7)]. Contrairement à $W(t)$, ce processus n'est pas markovien. Sa covariance est proportionnelle au covariogramme géométrique du segment de longueur a , c'est-à-dire un schéma triangulaire :

$$\forall \Delta t \geq 0, C_X(\Delta t) = \begin{cases} \alpha(1 - \Delta t / a) & \text{si } \Delta t \leq a \\ 0 & \text{sinon} \end{cases}$$

L'épandage de $X(t)$ perpendiculairement à une direction \mathbf{u} de l'espace conduit à un processus isofactoriel laguerrien tel que les covariances des facteurs successifs sont les moments d'une variable aléatoire beta de paramètres $\{C_X(\langle \mathbf{h} | \mathbf{u} \rangle), C_X(0) - C_X(\langle \mathbf{h} | \mathbf{u} \rangle)\}$ Pour le voir, il suffit de remarquer qu'une écriture sous la forme de l'équation (3.7) est conservée par épandage.

Soit à présent un ensemble de directions $\{\mathbf{u}_i, i = 1 \dots N\}$ équiréparties sur la sphère unité de \mathbb{R}^d et $\{X_i(t), i = 1 \dots N\}$ des processus laguerriens indépendants obtenus par régularisation d'une mesure orthogonale à loi gamma de même paramètre α/N . L'application de la technique des bandes tournantes conduit à poser

$$Y(\mathbf{x}) = \sum_{i=1}^N Y_i(\mathbf{x}) = \sum_{i=1}^N X_i(\langle \mathbf{x} | \mathbf{u}_i \rangle).$$

A cause de la propriété de stabilité pour l'addition de la loi laguerrienne, $Y(\mathbf{x})$ admet une loi de Laguerre de paramètre α et de covariance égale à la somme des covariances des processus épanchés :

$$C_Y(\mathbf{h}) = \sum_{i=1}^N C_{X_i}(\langle \mathbf{h} | \mathbf{u}_i \rangle) = \frac{1}{N} \sum_{i=1}^N C_X(\langle \mathbf{h} | \mathbf{u}_i \rangle).$$

En toute rigueur, cette covariance n'est pas isotrope. Toutefois, si la suite $\{\mathbf{u}_i, i = 1 \dots N\}$ constitue une séquence équirépartie et si le nombre N est élevé, l'anisotropie induite par la discrétisation de la sphère unité à partir d'un nombre fini de directions devient négligeable (Chilès, 1977 ; Freulon et de Fouquet, 1991). On peut alors écrire⁶ :

$$C_Y(\mathbf{h}) \approx E\{C_X(\langle \mathbf{h} | \mathbf{U} \rangle)\}$$

où \mathbf{U} est une variable aléatoire d'orientation uniforme sur la sphère unité de \mathbb{R}^d . Cette écriture signifie que C_Y est la covariance associée à C_X (schéma triangulaire) par l'opérateur des bandes tournantes (Matheron, 1973a). On en déduit :

- à deux dimensions

$$C_Y(\mathbf{h}) = \begin{cases} \alpha \left(1 - \frac{2|\mathbf{h}|}{\pi a}\right) & \text{si } |\mathbf{h}| \leq a \\ \alpha - \frac{2\alpha}{\pi} \left\{ \frac{|\mathbf{h}|}{a} - \sqrt{\frac{|\mathbf{h}|^2}{a^2} - 1} + \arccos\left(\frac{a}{|\mathbf{h}|}\right) \right\} & \text{sinon} \end{cases}$$

- à trois dimensions

$$C_Y(\mathbf{h}) = \begin{cases} \alpha \left(1 - \frac{|\mathbf{h}|}{2a}\right) & \text{si } |\mathbf{h}| \leq a \\ \frac{\alpha a}{2|\mathbf{h}|} & \text{sinon} \end{cases}$$

Dans les deux cas, $C_Y(\mathbf{h})$ décroît à l'infini comme une hyperbole.

On peut d'ailleurs rendre le processus simulé exactement isotrope en définissant les directions $\{\mathbf{u}_i, i = 1 \dots N\}$ à une rotation aléatoire près, ce qui n'altère pas le caractère isofactoriel des lois bivariées. Cependant, cette procédure est purement théorique, puisqu'elle ne rend pas davantage isotrope la covariance expérimentale obtenue sur les réalisations.

⁶ En particulier, la loi d'un couple de valeurs $\{Y(\mathbf{x}_1), Y(\mathbf{x}_2)\}$ ne dépend plus de N lorsque ce nombre devient infini. En ce qui concerne les lois multivariées, tout ensemble de valeurs $\{Y(\mathbf{x}_1), \dots, Y(\mathbf{x}_n)\}$ peut s'écrire sous la forme d'une décomposition analogue à l'équation (3.7) ; du fait que la suite de directions $\{\mathbf{u}_i, i = 1 \dots N\}$ est équirépartie, les paramètres des variables aléatoires intervenant dans cette décomposition sont asymptotiquement indépendants de N : la construction converge donc vers un unique modèle de loi spatiale lorsque le nombre de directions devient infini.

IV.1.4. Conditionnement à des données

Le conditionnement du modèle à des données expérimentales peut être réalisé par la technique de **recuit simulé** (cf. annexe C) avec le noyau de transition défini comme suit :

- a) sélectionner un nombre i au hasard parmi $\{1, \dots, N\}$;
- b) sélectionner une valeur t_0 uniformément sur $[0, L_i]$ où L_i est la longueur maximale de la projection du domaine à simuler sur la droite de direction \mathbf{u}_i ;
- c) le long de cette droite, modifier la marche aléatoire $W_i(t)$ sur un intervalle $[t_0, t_0 + \Delta t_0]$ où Δt_0 est une longueur fixée à l'avance ; raccorder en $t_0 + \Delta t_0$ les valeurs anciennes de $W_i(t)$;
- d) calculer les nouvelles valeurs de $X_i(t)$ sur l'intervalle $[t_0 - a, t_0 + \Delta t_0]$ et les valeurs correspondantes de $Y(\mathbf{x})$; les autres valeurs ne sont pas altérées.

La vitesse de convergence de l'algorithme peut être améliorée en choisissant un noyau de transition non stationnaire (Gidas, 1985). En l'occurrence, il s'agit de partir d'une valeur Δt_0 relativement élevée et de la faire diminuer à mesure que se déroule la simulation et que le paramètre de température décroît. Le but est de pouvoir n'accepter de grandes transitions que pour les températures fortes, car ces transitions sont presque systématiquement rejetées lorsque la température baisse.

IV.1.5. Extension aux modèles gigognes et anisotropes

Pour simuler un modèle isofactoriel de **covariance gigogne**, le plus simple consiste à simuler séparément chaque composante de la covariance : leur superposition conduira à la loi voulue en vertu de la propriété de stabilité du modèle laguerrien pour l'addition. Le conditionnement par recuit simulé utilise un noyau de transition similaire à celui présenté ci-dessus, à ceci près qu'à chaque étape on ne modifie qu'une seule composante, choisie au hasard parmi celles contenues dans le modèle gigogne.

Quant à la simulation de **modèles anisotropes**, il est facile de tenir compte des anisotropies géométriques et zonales puisqu'il suffit d'effectuer un changement de référentiel (rotation suivie d'une homothétie) pour se ramener au cadre isotrope. Les anisotropies plus complexes peuvent souvent être représentées par un modèle gigogne dans lequel les diverses composantes présentent des anisotropies géométriques ou zonales d'orientations et rapports différents.

IV.1.6. Généralisation aux modèles à lois indéfiniment divisibles

Le modèle présenté repose sur la stabilité pour l'addition de la loi laguerrienne, laquelle découle de la décomposition d'un couple de valeurs $\{Y(\mathbf{x} + \mathbf{h}), Y(\mathbf{x})\}$ selon [Eq. (3.7)] :

$$\begin{cases} Y(\mathbf{x} + \mathbf{h}) = U + V \\ Y(\mathbf{x}) = U + W \end{cases}$$

où U, V, W sont trois variables gamma indépendantes.

Or, ce type d'identité n'est pas spécifique à la **loi de Laguerre** et est associé à plusieurs autres lois isofactorielles, notamment les modèles **bigaussien**, de **Poisson** (Matheron, 1980a) et de **Meixner**⁷ (Matheron, 1973b). Les variables U, V et W suivent alors des distributions gaussiennes, poissonniennes et binomiales négatives respectivement. Dans les deux premiers cas, le modèle donne une loi de diffusion, pour laquelle la covariance du p-ième facteur est égale à celle du premier facteur élevée à la puissance p. Dans le dernier cas, à l'instar du modèle de Laguerre, les covariances des facteurs successifs sont les moments d'une variable aléatoire beta, ce qui n'est pas étonnant car le modèle de Meixner est la version discrétisée du modèle de Laguerre (Matheron, 1983, 1984a).

Dans les modèles de Laguerre et Meixner, les covariances des facteurs sont les moments d'une variable beta $R(\mathbf{h})$ de paramètres $\{\alpha \rho(\mathbf{h}), (1 - \alpha) \rho(\mathbf{h})\}$ où α est un scalaire positif et $\rho(\mathbf{h})$ est le corrélogramme de la fonction aléatoire $Y(\mathbf{x})$; $R(\mathbf{h})$ tend vers son espérance $\rho(\mathbf{h})$ lorsque α devient infini, c'est-à-dire qu'elle devient déterministe. On en déduit les convergences de ces modèles vers les modèles de diffusion :

- la loi de Laguerre de paramètre α tend vers la loi bigaussienne lorsque α tend vers l'infini ;
- la loi de Meixner de paramètres (η, α) tend vers la loi de Poisson de paramètre θ quand $\eta \rightarrow 0$, $\alpha \rightarrow \infty$ et $\alpha\eta \rightarrow \theta$.

Ces quatre modèles ont en commun la propriété d'**infinie divisibilité**, aussi bien des distributions marginales que bivariées : quel que soit l'entier N, un processus suivant l'un de ces modèles peut s'écrire comme la somme de N processus indépendants de même loi isofactorielle mais de paramètres différents, ce qui induit la stabilité du modèle pour l'addition. Dans la procédure des bandes tournantes cette propriété joue un double rôle :

- 1) elle permet de conserver le caractère isofactoriel des distributions lorsque l'on ajoute les processus élémentaires épanchés ;
- 2) elle rend possible la construction de chaque processus unidimensionnel à épancher à partir d'une marche aléatoire, ce qui rejoint la procédure proposée par Matheron pour définir un processus isofactoriel à loi indéfiniment divisible par régularisation d'une mesure aléatoire orthogonale stationnaire (Armstrong and Matheron, 1986a ; Matheron, 1973b).

D'autres modèles isofactoriels bénéficient de la propriété d'infinie divisibilité, mais ils sont d'un emploi moins courant car leurs facteurs ne sont pas polynomiaux. Citons celui proposé par Hu (1988) : il s'agit d'un modèle à **loi de Gauss inverse**, appelée parfois *loi de Sichel continue* par référence aux travaux de Sichel (1971), dont la construction est basée sur un processus de marche aléatoire et est identique à celle présentée dans le cadre laguerrien. La loi de Gauss inverse dépend de deux paramètres $\{\alpha, \beta\}$ ⁸ et peut décrire des distributions très dissymétriques, d'où son intérêt ; une méthode pour simuler des variables aléatoires suivant cette loi est donnée par Matheron (1985b). La loi de Gauss inverse fait à son tour partie d'une famille plus générale de lois, les lois de Gauss inverses généralisées, dont la loi gamma est un cas limite. Bien que la construction de modèles isofactoriels à lois de Gauss inverse généralisée soit envisageable, l'inférence des paramètres est plus délicate.

⁷ Ce modèle a été introduit sous le nom de modèle binomial négatif, mais cette terminologie est à présent réservée au modèle de diffusion (Chilès and Delfiner, 1999).

⁸ Les densité de probabilité et fonction de répartition sont définies sur \mathbb{R}_+ et s'expriment comme suit :

$$\forall y > 0, \begin{cases} f_{\alpha, \beta}(y) = \frac{\alpha}{\sqrt{2\pi}} y^{-3/2} \exp\left(\alpha\beta - \frac{\alpha^2}{2y} - \frac{\beta^2 y}{2}\right) \\ F_{\alpha, \beta}(y) = G\left(\beta\sqrt{y} - \frac{\alpha}{\sqrt{y}}\right) + \exp(2\alpha\beta) \left[1 - G\left(\beta\sqrt{y} + \frac{\alpha}{\sqrt{y}}\right)\right] \end{cases}$$

Dans un cadre similaire aux processus à lois indéfiniment divisibles, Boulanger (1990) a proposé d'utiliser la méthode des bandes tournantes pour simuler des processus à *lois stables*, qui généralisent les processus gaussiens mais possèdent une variance infinie.

IV.1.7. Difficultés pratiques et limitations

La simulation par bandes tournantes présente plusieurs difficultés. Citons en particulier, dans le cas du modèle de Laguerre, les points suivants :

- simulation de variables aléatoires gamma de paramètre proche de zéro : pour construire un modèle à loi gamma de paramètre α à partir de N directions $\{\mathbf{u}_i, i = 1 \dots N\}$, il faut simuler le long de chacune de ces directions un processus gamma de paramètre α/N . Pour ce faire, il est conseillé de recourir à un algorithme d'acceptation et rejet (von Neumann, 1951 ; Lantuéjoul, 1984) plutôt que la méthode classique de simulation de variables aléatoires par inversion.
- l'épandage d'une valeur forte conduit à des artefacts dans la simulation finale, surtout lorsque le paramètre α est faible (aspect de "toile d'araignée", cf. figure 3.15).
- lenteur de la convergence, ce qui oblige à prendre un nombre très important de droites.
- limitation sur la classe des fonctions de covariance qu'il est possible de reproduire : cette classe est l'image par l'opérateur des bandes tournantes de la covariance triangulaire ou, dans la version plus générale où l'on simule un modèle gigogne, d'une somme de schémas triangulaires, c'est-à-dire d'une covariance convexe. Dans tous les cas, la covariance multidimensionnelle décroît de façon hyperbolique, donc sa portée intégrale est infinie.

Par exemple, dans le cas isotrope à trois dimensions, la décroissance de $C_Y(\mathbf{h})$ est au moins aussi lente que $1/|\mathbf{h}|$, car :

$$C_X(r) \geq 0 \text{ (fonction convexe tendant vers 0 à l'infini) et } C_Y(\mathbf{h}) = \frac{1}{|\mathbf{h}|} \int_0^{|\mathbf{h}|} C_X(r) dr \underset{|\mathbf{h}| \rightarrow +\infty}{\sim} \frac{A}{|\mathbf{h}|}$$

où A est la portée intégrale de la covariance C_X .

Dans le cadre multigaussien, il est possible de simuler des processus multidimensionnels à portée intégrale finie à partir de processus unidimensionnels de portée intégrale nulle, dont la covariance présente un fort effet de trou ; des exemples connus sont les modèles sphérique et exponentiel (Matheron, 1972a ; Lantuéjoul, 1994). Ceci n'est plus possible pour la loi de Laguerre, car son coefficient de corrélation doit nécessairement être positif.

En fin de compte, la méthode des bandes tournantes est décevante pour la simulation de processus à loi marginale fortement dissymétrique, tandis qu'elle devient acceptable lorsque la loi gamma se rapproche d'une loi gaussienne (paramètre α élevé). Néanmoins, bien que la méthodologie proposée ait peu d'intérêt pratique (le but est de simuler des processus les plus éloignés possibles de la loi de diffusion gaussienne), l'idée d'ajouter un grand nombre de processus élémentaires inspirera les méthodes de simulation de modèles isofactoriels présentées dans les sections suivantes.

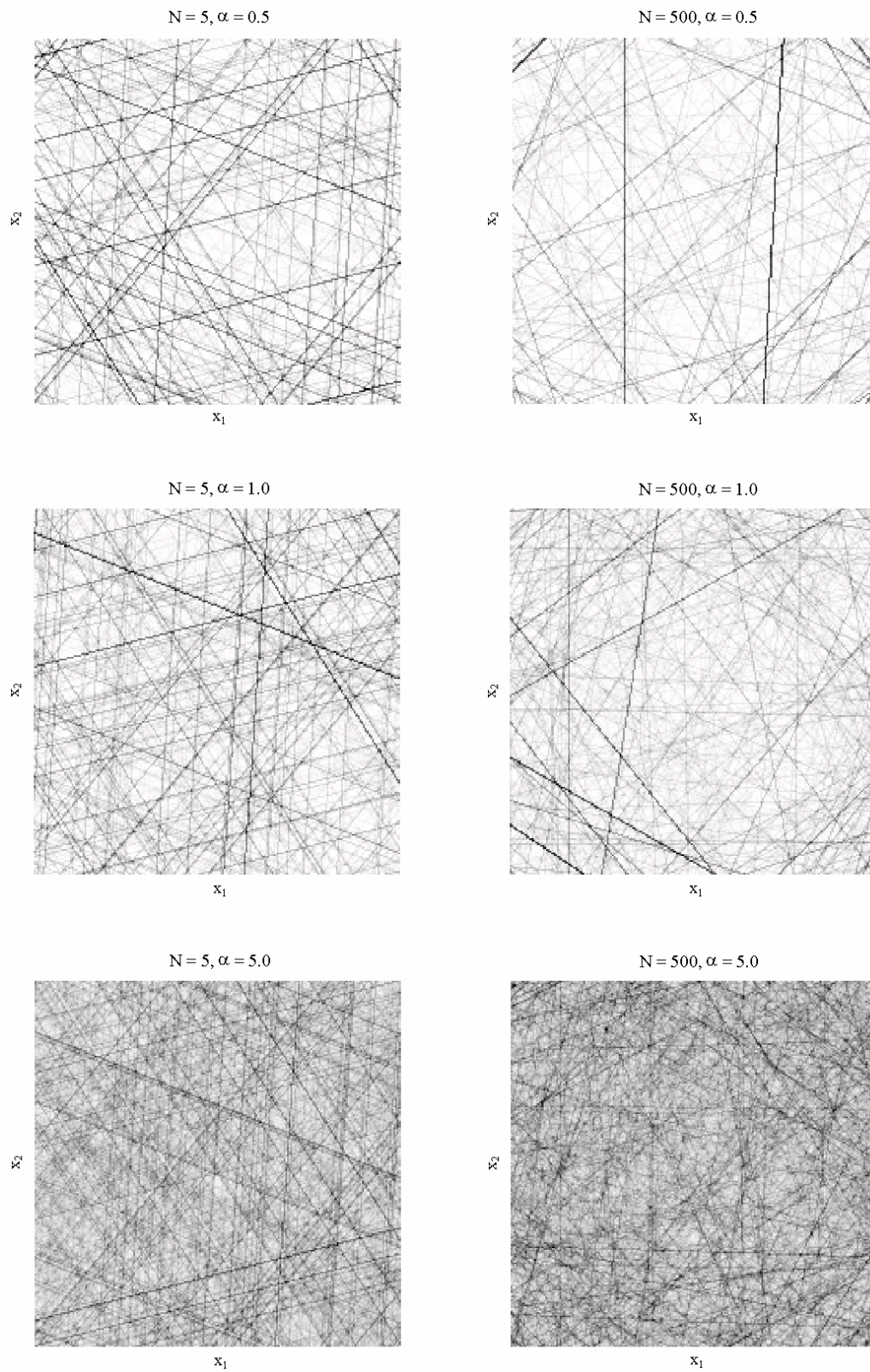


Figure 3.15. Réalisations par bandes tournantes de processus laguerriens

IV.2. Dilution poissonnienne

Le modèle de dilution considère un ensemble de points poissonniens dans l'espace. Chaque point devient le centre de gravité d'un grain primaire valué ; on calcule ensuite la somme des valuations de tous les grains primaires (Serra, 1968 ; Alfaro, 1979 ; Lantuéjoul, 2002) :

$$\forall \mathbf{x} \in \mathbb{R}^d, Y(\mathbf{x}) = \sum_{\mathbf{x} - \mathbf{X}_i \in G_i} A_i$$

où $\{\mathbf{X}_i, i \in I\}$ est un processus de Poisson ponctuel de densité θ constante dans l'espace

$\{G_i, i \in I\}$ est une famille de grains primaires déterministes ou aléatoires centrés à l'origine ; leur forme géométrique ne dépend pas de $\{\mathbf{X}_i, i \in I\}$

$\{A_i, i \in I\}$ sont des variables aléatoires de même loi, d'espérance μ et variance σ^2 , indépendantes entre elles et indépendantes de $\{\mathbf{X}_i, i \in I\}$ et de $\{G_i, i \in I\}$.

La covariance du processus simulé est proportionnelle au covariogramme géométrique moyen des grains primaires $K(\mathbf{h})$, par exemple un schéma sphérique de portée a si les grains primaires sont des sphères de diamètre a :

$$C(\mathbf{h}) = \theta(\sigma^2 + \mu^2)K(\mathbf{h})$$

Calculons la transformée de Laplace des lois marginale et bivariable.

- **Loi marginale** : pour un site \mathbf{x} fixé, soient N le nombre aléatoire de points poissonniens $\{\mathbf{X}_i, i \in I\}$ tels que $\mathbf{x} - \mathbf{X}_i \in G_i$ et ϕ la transformée de Laplace de la loi des valuations $\{A_i, i \in I\}$, alors :

$$\phi(\lambda) = E[\exp(-\lambda Y(\mathbf{x}))] = E\{E[\exp(-\lambda \sum_{i=1}^n A_i) \mid N = n]\} = E\{\phi(\lambda)^N\}$$

Or, N suit une distribution de Poisson de paramètre $\theta K(\mathbf{0})$, d'où :

$$\phi(\lambda) = \exp\{\theta K(\mathbf{0})[\phi(\lambda) - 1]\}$$

Cette loi marginale est indéfiniment divisible en tant que somme de variables aléatoires en nombre poissonnien (Feller, 1966 ; Matheron, 1969b).

Dans le cas de valuations discrètes, la même relation lie les fonctions caractéristiques $g(s)$ et $G(s)$ des valuations et du processus de dilution :

$$G(s) = \exp\{\theta K(\mathbf{0})[g(s) - 1]\}$$

- **Loi bivariable** : les grains primaires recouvrant les sites \mathbf{x} et $\mathbf{x} + \mathbf{h}$ peuvent se décomposer en trois processus indépendants : ceux qui recouvrent $\mathbf{x} + \mathbf{h}$ et pas \mathbf{x} , en nombre N_1 , ceux qui recouvrent \mathbf{x} et pas $\mathbf{x} + \mathbf{h}$, en nombre N_2 , et ceux qui recouvrent à la fois \mathbf{x} et $\mathbf{x} + \mathbf{h}$, en nombre N_3 . Les nombres N_1 , N_2 et N_3 sont des variables poissonniennes indépendantes, de paramètres $\theta [K(\mathbf{0}) - K(\mathbf{h})]$ pour les deux premières et $\theta K(\mathbf{h})$ pour la dernière. On en déduit :

$$\begin{aligned} \phi(\lambda, \mu) &= E\{E[\exp(-\lambda Y(\mathbf{x} + \mathbf{h}) - \mu Y(\mathbf{x})) \mid N_1 = n_1, N_2 = n_2, N_3 = n_3]\} \\ &= E\{\phi(\lambda)^{N_1} \phi(\mu)^{N_2} \phi(\lambda + \mu)^{N_3}\} \\ &= \exp\{\theta [K(\mathbf{0}) - K(\mathbf{h})][\phi(\lambda) + \phi(\mu) - 2] + \theta K(\mathbf{h})[\phi(\lambda + \mu) - 1]\} \end{aligned}$$

Cette transformée de Laplace vérifie l'équation caractéristique des modèles isofactoriels à lois indéfiniment divisibles (Matheron, 1973b) :

$$\phi(\lambda, \mu) = \phi(\lambda)^{1-\rho(\mathbf{h})} \phi(\mu)^{1-\rho(\mathbf{h})} \phi(\lambda + \mu)^{\rho(\mathbf{h})}$$

où $\rho(\mathbf{h}) = K(\mathbf{h}) / K(\mathbf{0})$ est le coefficient de corrélation entre les valeurs prises aux sites \mathbf{x} et $\mathbf{x} + \mathbf{h}$.

On retrouve donc un mode de construction général des modèles isofactoriels à lois indéfiniment divisibles. Parmi ces lois, toutes celles qui sont positives sont accessibles (Matheron, 1969b), mais le modèle bigaussien par exemple ne l'est pas (la loi gaussienne ne peut s'écrire comme un mélange poissonien de variables aléatoires). La simulation conditionnelle des modèles de dilution est détaillée par Lantuéjoul (2002) ; elle consiste en une procédure itérative au cours de laquelle les conditions sur les données sont respectées à chaque itération.

Exemples

- *Processus de diffusion de Poisson*

Ce processus (Matheron, 1973b) correspond au cas où la valuation des grains primaires est presque sûrement égale à 1 (figure 3.16). Ce modèle porte le nom de *jetons aléatoires* (Alfaro, 1979). La valeur simulée en un site est une variable entière égale au nombre de grains primaires ("jetons") qui recouvrent ce site ; elle suit une loi de Poisson.

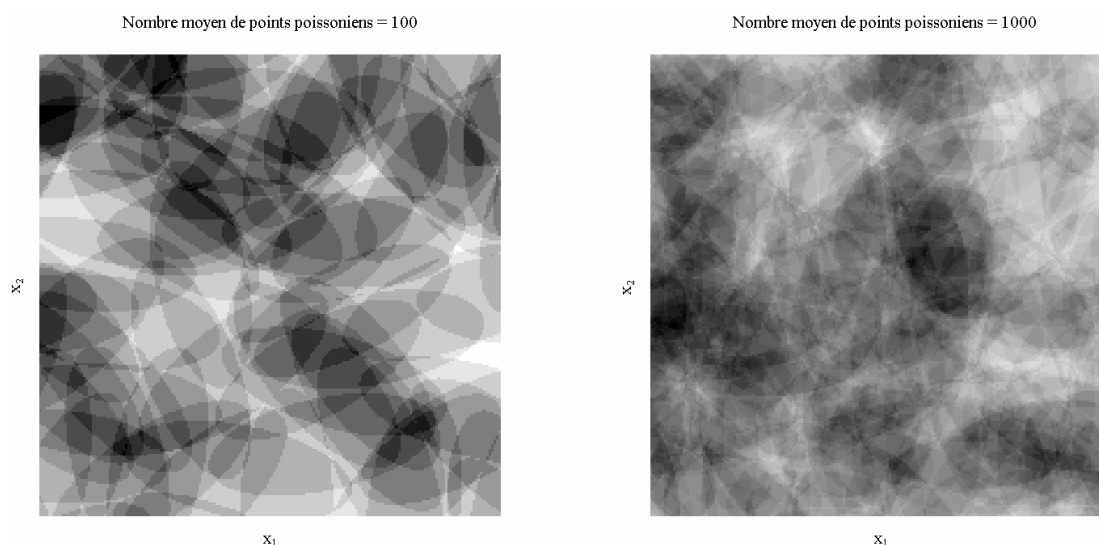


Figure 3.16. Réalisations de modèles de jetons aléatoires de densités poissonniennes différentes. Les grains primaires sont des ellipses de taille fixe et orientation aléatoire

- *Modèle stable*

Etant donné un scalaire α dans $]0,1[$, on considère la somme suivante :

$$Y = \sum_{n=1}^N T_n^{-1/\alpha}$$

où N est une variable aléatoire de Poisson de paramètre $t / \Gamma(1 - \alpha)$
 $\{T_n, n \in \mathbb{N}\}$ sont des variables aléatoires uniformes et indépendantes sur $[0,t]$

Alors Y converge vers une variable aléatoire de loi stable de paramètre α quand t tend vers l'infini (Lantuéjoul, 1987), caractérisée par sa transformée de Laplace $\phi(\lambda) = \exp(-\lambda^\alpha)$.

Pour obtenir un modèle isofactoriel à loi marginale stable de paramètre α , il suffit donc de prendre un modèle de dilution avec une densité poissonnienne θ très élevée et des grains primaires valués selon une loi uniforme sur $[0, \theta \Gamma(1 - \alpha)]$ élevée à la puissance $-1/\alpha$.

- **Modèle de Meixner**

La loi binomiale négative de paramètres η et α est un mélange poissonnien (de densité $\theta K(\mathbf{0})$) de variables entières définies par le système de probabilité suivant :

$$p(0) = 1 + \frac{\alpha}{\theta K(\mathbf{0})} \ln(1 - \eta) \text{ et } \forall n \geq 1, p(n) = \frac{\alpha}{\theta K(\mathbf{0})} \frac{\eta^n}{n}$$

sous la contrainte que $p(0)$ soit positif. On vérifie en effet que, si N suit le système de probabilité précédent, sa fonction caractéristique est donnée par :

$$g(s) = E(s^N) = p(0) + \sum_{n=1}^{+\infty} s^n p(n) = p(0) - \frac{\alpha}{\theta K(\mathbf{0})} \ln(1 - s\eta) = 1 + \frac{\alpha}{\theta K(\mathbf{0})} \ln\left(\frac{1 - \eta}{1 - s\eta}\right)$$

de sorte que le mélange poissonnien admet la fonction caractéristique

$$G(s) = \exp \{ \theta K(\mathbf{0}) [g(s) - 1] \} = \left(\frac{1 - \eta}{1 - s\eta} \right)^\alpha$$

qui coïncide avec celle de la loi binomiale négative de paramètres η et α .

On peut simuler la variable N par itérations markoviennes (procédure de Métropolis-Hastings, cf. annexe C) ou par acceptation et rejet, car sa loi de probabilité aux grandes valeurs est majorée par une loi géométrique (Lantuéjoul, 2002).

Les modèles de dilution manquent cependant de flexibilité, car ils ne permettent pas de contrôler le caractère plus ou moins mosaïque du processus simulé : ainsi, le modèle isofactoriel de Charlier, à loi marginale poissonnienne, est inaccessible ; seul le modèle de diffusion poissonnien peut être obtenu.

Le modèle présenté ci-après permet au contraire de contrôler le caractère diffusif ou mosaïque du processus obtenu. Il ne possède plus la propriété d'infinie divisibilité des lois bivariées, bien que celle des lois marginales reste conservée. Pour faciliter la lecture, nous commencerons l'exposé par la loi gamma.

IV.3. Addition de mosaïques indépendantes

IV.3.1. Modèle de Laguerre

Soit une fonction aléatoire $\{Y(\mathbf{x}), \mathbf{x} \in \mathbb{R}^d\}$ mosaïque, de loi marginale gamma de paramètre α et corrélogramme $\rho(\mathbf{h})$ à valeurs positives. Avec la probabilité $\rho(\mathbf{h})$, deux valeurs $Y(\mathbf{x} + \mathbf{h})$ et $Y(\mathbf{x})$ sont égales ; avec la probabilité complémentaire elles sont indépendantes (Matheron, 1982). Par suite, l'espace est divisé en compartiments aléatoires à l'intérieur de chacun desquels la valeur est constante et indépendante des autres compartiments.

Additionnons à présent plusieurs fonctions aléatoires mosaïques $\{Y_i, i = 1 \dots N\}$ indépendantes de même loi marginale gamma de paramètre α/N et corrélogramme $\rho(\mathbf{h})$:

$$\forall \mathbf{x} \in \mathbb{R}^d, Y(\mathbf{x}) = \sum_{i=1}^N Y_i(\mathbf{x}).$$

Soit K le nombre (aléatoire) de fois où les sites \mathbf{x} et $\mathbf{x} + \mathbf{h}$ appartiennent au même compartiment de mosaïque ; K est une variable binomiale de taille N et paramètre $\rho(\mathbf{h})$. Conditionnellement à $K = k$, le couple de valeurs $\{Y(\mathbf{x} + \mathbf{h}), Y(\mathbf{x})\}$ se décompose de la manière suivante :

$$\begin{cases} Y(\mathbf{x} + \mathbf{h}) = U + V \\ Y(\mathbf{x}) = U + W \end{cases} \quad (3.8)$$

où U , V et W sont trois variables aléatoires gamma indépendantes⁹, la première de paramètre $\alpha k / N$, les deux dernières de paramètre $\alpha (1 - k / N)$.

Le coefficient de corrélation conditionnel à $K = k$ est égal à k / N ; il est toujours positif.

D'après la section précédente [Eq. (3.7)], le couple $\{Y(\mathbf{x} + \mathbf{h}), Y(\mathbf{x})\}$ suit une loi isofactorielle de Laguerre telle que les covariances des facteurs successifs sont les moments d'une variable beta de paramètres $\{\alpha k / N, \alpha (1 - k / N)\}$.

En déconditionnant par rapport à K , on voit que la loi *a priori* du couple $\{Y(\mathbf{x} + \mathbf{h}), Y(\mathbf{x})\}$ est une loi de Laguerre où la corrélation est caractérisée par une variable aléatoire $R(\mathbf{h})$ à loi beta randomisée par une binomiale :

$$\begin{aligned} \forall \mathbf{h} \neq \mathbf{0}, C_p^{(Y)}(\mathbf{h}) &= E\{R(\mathbf{h})^p\} = E\left\{\frac{\Gamma(\alpha)\Gamma(\alpha K / N + p)}{\Gamma(\alpha K / N)\Gamma(\alpha + p)}\right\} \\ &= \frac{\Gamma(\alpha)[1 - \rho(\mathbf{h})]^N}{\Gamma(\alpha + p)} \sum_{k=1}^N C_N^k \left(\frac{\rho(\mathbf{h})}{1 - \rho(\mathbf{h})}\right)^k \frac{\Gamma(\alpha k / N + p)}{\Gamma(\alpha k / N)} \end{aligned} \quad (3.9)$$

Examinons les deux situations extrêmes :

- si $N = 1$, on retrouve le modèle mosaïque à loi marginale gamma :

$$C_p^{(Y)}(\mathbf{h}) = \rho(\mathbf{h}) \text{ indépendant de } p.$$

⁹ On admet qu'une variable gamma de paramètre nul est presque sûrement égale à zéro ; cette convention permet d'inclure les cas où $k = 0$ et $k = N$.

- si N tend vers l'infini, d'après la loi forte des grands nombres, la variable K / N converge presque sûrement vers son espérance, c'est-à-dire $\rho(\mathbf{h})$. On retrouve donc le modèle de Laguerre présenté dans la section précédente où les covariances des facteurs sont les moments d'une variable beta de paramètres $\{\alpha \rho(\mathbf{h}), \alpha [1 - \rho(\mathbf{h})]\}$:

$$C_p^{(Y)}(\mathbf{h}) = \frac{\Gamma(\alpha)\Gamma(\alpha\rho(\mathbf{h})+p)}{\Gamma(\alpha\rho(\mathbf{h}))\Gamma(\alpha+p)}.$$

Le choix du nombre N permet d'évoluer entre ces deux extrêmes. On dispose donc d'un paramètre supplémentaire pour réguler certaines caractéristiques de la simulation comme la propriété plus ou moins marquée de déstructuration des hautes teneurs.

IV.3.2. Exemple de la mosaïque des feuilles mortes

Contrairement à la méthode des bandes tournantes, il est aisé de construire des modèles dont les facteurs ont des covariances de portée intégrale finie. Il suffit en effet de choisir une mosaïque dont la covariance a une portée finie. Parmi les modèles mosaïques connus, l'un des plus intéressants est celui des **feuilles mortes** (Matheron, 1968a ; Jeulin, 1997) car il correspond à de nombreux modèles de covariance et permet de réaliser le conditionnement du modèle isofactoriel à un ensemble de données.

Détaillons brièvement le modèle des feuilles mortes, qui se construit dynamiquement. Des points poissonniens apparaissent successivement, en chacun desquels est implanté un ensemble aléatoire valué ou *grain primaire*. Les grains nouveaux recouvrent les grains plus anciens (*modèle de recouvrement*). Au bout d'un temps infini, on obtient un modèle mosaïque dont la loi marginale coïncide avec celle de la valuation des grains primaires et dont la loi bivariable est caractérisée par le corrélogramme

$$\rho(\mathbf{h}) = \frac{K(\mathbf{h})}{2K(\mathbf{0}) - K(\mathbf{h})}$$

où $K(\mathbf{h})$ est le covariogramme géométrique moyen du grain primaire. Cette formule suppose que la densité du processus de Poisson et les caractéristiques du grain primaire ne dépendent pas du temps. On obtiendra un corrélogramme de portée finie s'il en est de même du covariogramme géométrique $K(\mathbf{h})$ (grain primaire de taille presque sûrement finie et bornée).

Une variante de ce modèle, qui fournit la même loi spatiale, consiste à considérer que les grains nouveaux sont "cachés" par les grains anciens (*modèle de germination*) ; cela revient à réordonner chronologiquement les grains primaires en renversant l'échelle de temps. Vis-à-vis du modèle initial, cette variante présente l'avantage de donner une image exacte de la mosaïque dès que le domaine à simuler est recouvert par les grains primaires (ajouter d'autres grains n'altère plus l'image, puisque les grains nouveaux sont cachés par les grains déjà tombés) ; ceci permet d'obtenir de façon exacte une simulation non conditionnelle du modèle, en tant que somme de N mosaïques indépendantes (figure 3.17).

Le conditionnement du modèle à un ensemble de données peut être mené à bien par **recuit simulé** à partir d'une simulation non conditionnelle, à l'aide du noyau de transition proposé en annexe C pour la mosaïque des feuilles mortes.

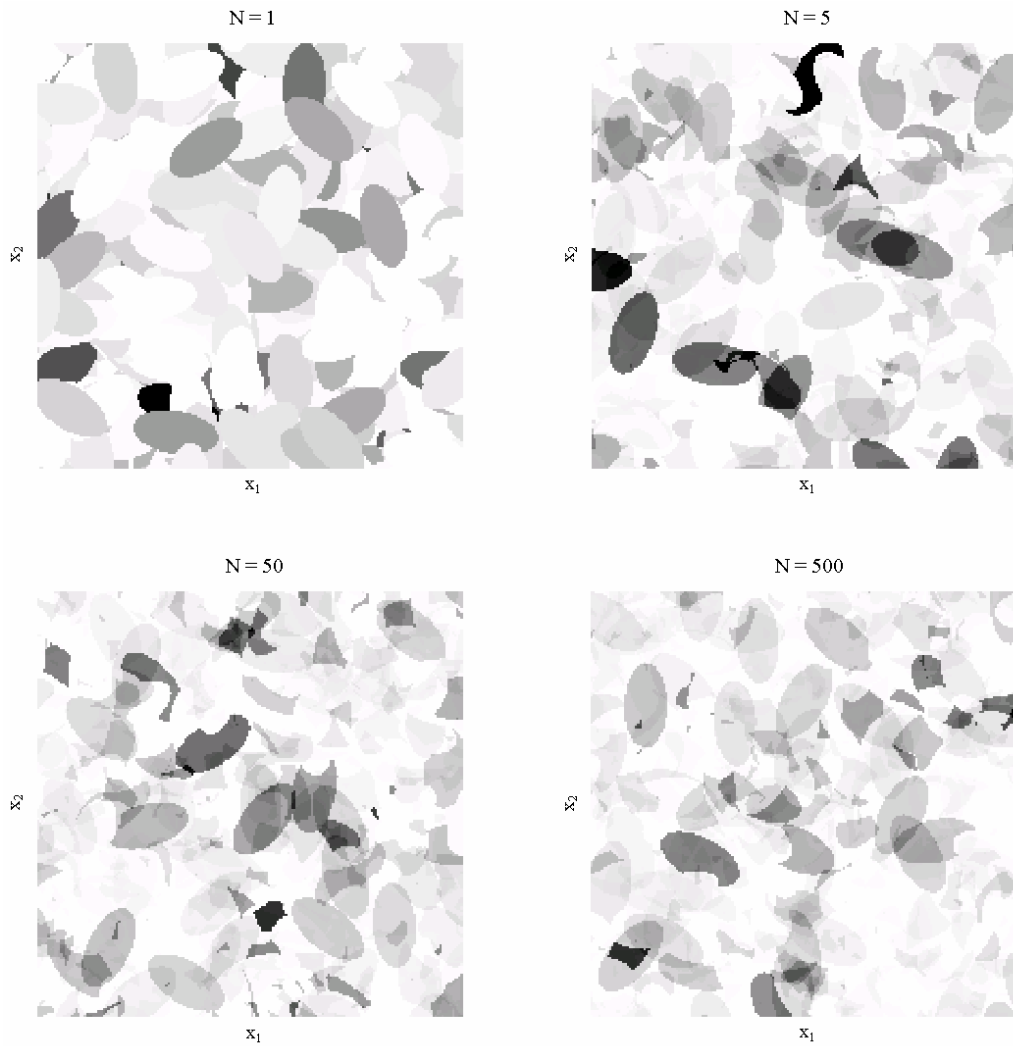


Figure 3.17. *Quatre réalisations de processus laguerriens de paramètre $a = 0.5$
Les grains primaires sont des ellipses de taille fixe et orientation aléatoire*

IV.3.3. Prise en compte d'un effet de pépite

La présence d'un effet de pépite – vu comme une microstructure – se révèle gênante, car elle correspond à un grain primaire de dimensions très petites mais de densité poissonnienne très élevée. Un modèle de feuilles mortes contenant une telle microstructure risque donc de provoquer une quantité considérable de calculs, surtout lors de la procédure de recuit simulé.

Une solution consiste à intégrer l'effet de pépite en tant que composante supplémentaire qui vient s'ajouter aux mosaïques. On choisit donc d'ajouter N mosaïques de même loi marginale gamma de paramètre $(\alpha - \alpha_0)/N$ et corrélogramme $\rho(\mathbf{h})$ et un processus pépitique de loi marginale gamma de paramètre α_0 . L'équation (3.8) devient :

$$\begin{cases} Y(\mathbf{x} + \mathbf{h}) = U + V \\ Y(\mathbf{x}) = U + W \end{cases}$$

où U , V et W sont trois variables gamma indépendantes, la première de paramètre $(\alpha - \alpha_0) k / N$, les deux dernières de paramètre $\alpha - (\alpha - \alpha_0) k / N$.

L'équation (3.9) donnant les covariances des facteurs successifs devient :

$$\begin{aligned} \forall \mathbf{h} \neq \mathbf{0}, C_p^{(Y)}(\mathbf{h}) &= E\{R(\mathbf{h})^p\} = E\left\{\frac{\Gamma(\alpha)\Gamma[(\alpha - \alpha_0)K/N + p]}{\Gamma[(\alpha - \alpha_0)K/N]\Gamma(\alpha + p)}\right\} \\ &= \frac{\Gamma(\alpha)[1 - \rho(\mathbf{h})]^N}{\Gamma(\alpha + p)} \sum_{k=1}^N C_N^k \left(\frac{\rho(\mathbf{h})}{1 - \rho(\mathbf{h})}\right)^k \frac{\Gamma[(\alpha - \alpha_0)k/N + p]}{\Gamma[(\alpha - \alpha_0)k/N]} \end{aligned} \quad (3.10)$$

Le conditionnement par recuit simulé peut s'effectuer à partir d'une simulation non conditionnelle à l'aide du même noyau de transition que précédemment. Il suffit de considérer que la valuation des grains primaires est la somme d'une valeur gamma de paramètre $(\alpha - \alpha_0)/N$ et d'un effet de pépite de loi gamma de paramètre α_0/N .

IV.3.4. L'inférence des paramètres du modèle

Outre le corrélogramme $\rho(\mathbf{h})$, on dispose de paramètres scalaires pour caractériser le modèle : α , α_0 et N . Le premier est un paramètre de dissymétrie dans la distribution marginale des valeurs : plus il est proche de zéro, plus la loi marginale est concentrée sur les valeurs faibles et présente une queue allongée. En pratique, bien qu'une anamorphose est toujours possible, ce paramètre est choisi de façon à ne pas trop déformer l'histogramme des valeurs mesurées (Demange *et al.*, 1987 ; Hu, 1988). Le paramètre α_0 est destiné à modéliser l'effet de pépite.

Le nombre N peut quant à lui s'interpréter comme un paramètre de déstructuration des hautes teneurs (propriété de la loi bivariable qui traduit le "désordre" observé dans l'organisation spatiale des valeurs). Un critère synthétique permettant de l'inférer à partir des données disponibles consiste à comparer leurs variogramme et madogramme. D'après Matheron (1982), dans le modèle laguerrien, le madogramme $\gamma_1(\mathbf{h})$ vérifie :

$$\gamma_1(\mathbf{h}) / \gamma_1(\infty) = E\{\sqrt{1 - R(\mathbf{h})}\}.$$

Après quelques calculs, on obtient l'expression suivante :

$$\frac{\gamma_1(\mathbf{h})}{\gamma_1(\infty)} = \rho(\mathbf{h})^N \frac{\Gamma(\alpha)}{\Gamma(\alpha + 1/2)} \sum_{k=0}^N C_N^k \left(\frac{1 - \rho(\mathbf{h})}{\rho(\mathbf{h})}\right)^k \frac{\Gamma[\alpha_0 + (\alpha - \alpha_0)k/N + 1/2]}{\Gamma[\alpha_0 + (\alpha - \alpha_0)k/N]} \quad (3.11)$$

En effet, on a le développement :

$$\begin{aligned} \forall \mathbf{h} \neq \mathbf{0}, \frac{\gamma_1(\mathbf{h})}{\gamma_1(\infty)} &= E\{\sqrt{1 - R(\mathbf{h})}\} = 1 + \sum_{p \geq 1} \frac{\Gamma(p - 1/2)}{p! \Gamma(-1/2)} C_p^{(Y)}(\mathbf{h}) \\ &= 1 + [1 - \rho(\mathbf{h})]^N \sum_{k=0}^N C_N^k \left(\frac{\rho(\mathbf{h})}{1 - \rho(\mathbf{h})}\right)^k \frac{\Gamma(\alpha)}{\Gamma[(\alpha - \alpha_0)k/N] \Gamma(-1/2)} \sum_{p \geq 1} \frac{\Gamma(p - 1/2) \Gamma[(\alpha - \alpha_0)k/N + p]}{\Gamma(\alpha + p) p!} \\ &= 1 + [1 - \rho(\mathbf{h})]^N \sum_{k=0}^N C_N^k \left(\frac{\rho(\mathbf{h})}{1 - \rho(\mathbf{h})}\right)^k [{}_2F_1(-1/2, (\alpha - \alpha_0)k/N, \alpha, 1) - 1] \\ &= \rho(\mathbf{h})^N \sum_{k=0}^N C_N^k \left(\frac{1 - \rho(\mathbf{h})}{\rho(\mathbf{h})}\right)^k {}_2F_1(-1/2, (\alpha - \alpha_0)(1 - k/N), \alpha, 1) \quad (\text{changement de variable } k \rightarrow N - k) \end{aligned}$$

où ${}_2F_1$ est la fonction hypergéométrique de Gauss (Abramowitz and Stegun, 1972) définie par :

$${}_2F_1(a, b, c, x) = 1 + \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)} \sum_{p \geq 1} \frac{\Gamma(a + p)\Gamma(b + p)}{\Gamma(c + p)} \frac{x^p}{p!}.$$

La série converge pour $-1 < x < 1$ et en $x = \pm 1$ si $c > a + b$. L'expression annoncée découle de l'identité de Gauss

$${}_2F_1(a, b, c, 1) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}$$

En absence d'effet de pépité ($\alpha_0 = 0$), les cas extrêmes sont les suivants :

- $N = 1$ (modèle mosaïque)

$$\frac{\gamma_1(\mathbf{h})}{\gamma_1(\infty)} = 1 - \rho(\mathbf{h}) = \frac{\gamma(\mathbf{h})}{\gamma(\infty)} : \text{variogramme et madogramme sont proportionnels}$$

- $N \rightarrow \infty$ (modèle beta, cf. Chilès and Delfiner, 1999)

$$\forall \mathbf{h} \neq \mathbf{0}, \frac{\gamma_1(\mathbf{h})}{\gamma_1(\infty)} = \frac{\Gamma(\alpha)}{\Gamma(\alpha + 1/2)} \frac{\Gamma[\gamma(\mathbf{h}) + 1/2]}{\Gamma[\gamma(\mathbf{h})]}$$

Lorsque le paramètre α augmente, ce rapport tend vers la racine carrée du variogramme divisé par son palier (cas bigaussien).

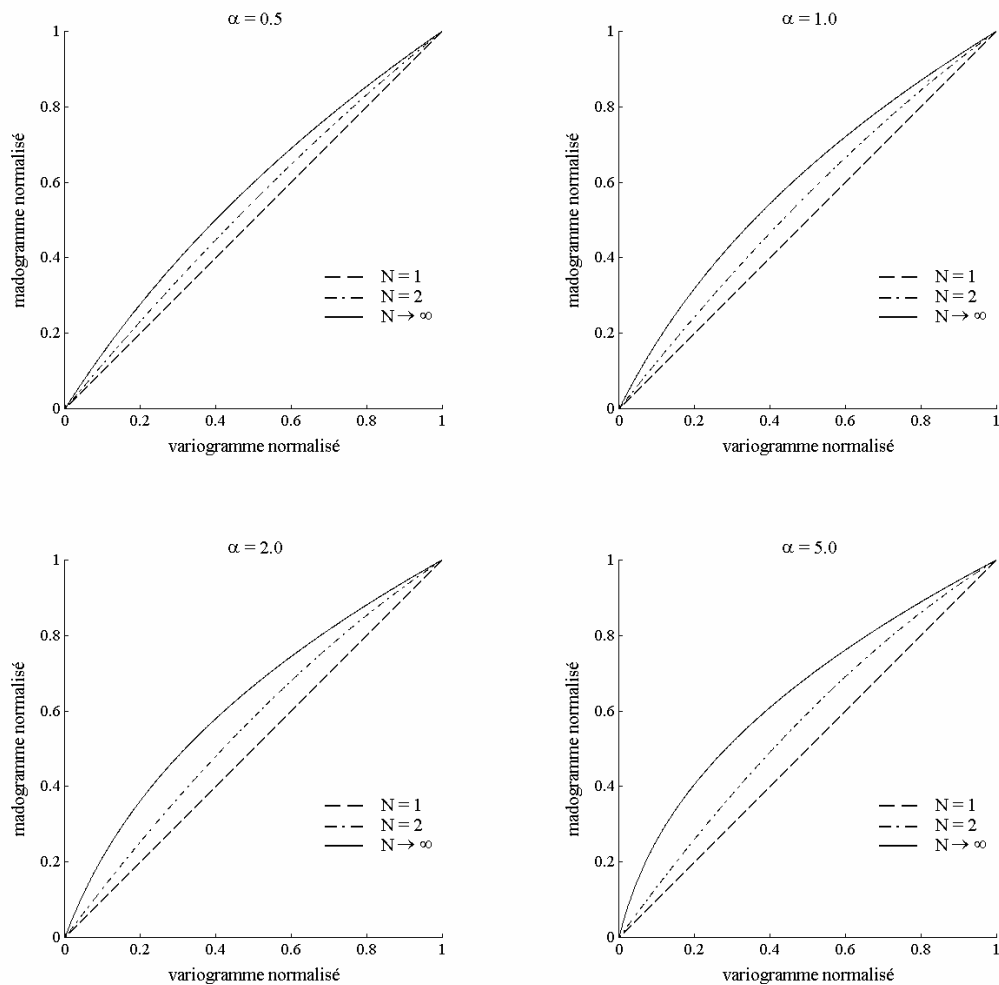


Figure 3.18. Comparaison du variogramme et du madogramme, en fonction du paramètre α et du nombre N de mosaïques ($\alpha_0 = 0$)

IV.3.5. Autres modèles à loi marginale indéfiniment divisible

La construction s'étend aux autres modèles isofactoriels à lois marginales indéfiniment divisibles (gaussienne, poissonnienne, binomiale négative, gaussienne inverse). On obtient les modèles suivants :

- **modèle hermitien** : la loi marginale est une gaussienne réduite et les covariances des facteurs sont les moments de la variable $R(\mathbf{h}) = K(1 - \alpha_0) / N$, où α_0 représente la variance de l'effet de pépité :

$$\forall \mathbf{h} \neq \mathbf{0}, C_p^{(Y)}(\mathbf{h}) = \frac{(1 - \alpha_0)^p}{N^p} E(K^p) = \frac{(1 - \alpha_0)^p [1 - \rho(\mathbf{h})]^N}{N^p} \sum_{k=0}^N C_N^k k^p \left(\frac{\rho(\mathbf{h})}{1 - \rho(\mathbf{h})} \right)^k$$

- **modèle de Charlier**¹⁰ : la loi marginale est poissonnienne de paramètre α et les covariances des facteurs sont les moments de la variable $R(\mathbf{h}) = K(\alpha - \alpha_0) / N \alpha$;
- **modèle de Meixner** : la loi marginale est binomiale négative ; les covariances des facteurs sont les moments successifs d'une variable $R(\mathbf{h})$ dont la loi est celle d'une variable beta de paramètres $\{(\alpha - \alpha_0) K / N, \alpha - (\alpha - \alpha_0) K / N\}$ randomisés par la binomiale K ;
- **modèle inverse gaussien** : la loi marginale est une loi de Gauss inverse de paramètres $\{\alpha, \beta\}$.

Dans le modèle de Meixner, le madogramme est lié au corrélogramme $\rho(\mathbf{h})$ par la même formule que celle donnée dans le cadre laguerrien [Eq. (3.11)]. En revanche, dans les modèles hermitien et de Charlier, cette formule devient :

$$\frac{\gamma_1(\mathbf{h})}{\gamma_1(\infty)} = \rho(\mathbf{h})^N \sum_{k=1}^N C_N^k \sqrt{\frac{k}{N} \left(1 - \frac{\alpha_0}{\alpha}\right) + \frac{\alpha_0}{\alpha}} \left(\frac{1 - \rho(\mathbf{h})}{\rho(\mathbf{h})} \right)^k$$

où α représente la variance *a priori* (1 pour le modèle hermitien, paramètre de la loi poissonnienne pour le modèle de Charlier) et α_0 l'amplitude de l'effet de pépité.

En effet, le calcul donne successivement :

$$\begin{aligned} \forall \mathbf{h} \neq \mathbf{0}, \frac{\gamma_1(\mathbf{h})}{\gamma_1(\infty)} &= E\{\sqrt{1 - R(\mathbf{h})}\} = 1 + \sum_{p \geq 1} \frac{\Gamma(p - 1/2)}{p! \Gamma(-1/2)} C_p^{(Y)}(\mathbf{h}) \\ &= 1 + [1 - \rho(\mathbf{h})]^N \sum_{k=0}^N C_N^k \left(\frac{\rho(\mathbf{h})}{1 - \rho(\mathbf{h})} \right)^k \sum_{p \geq 1} \frac{\Gamma(p - 1/2)}{p! \Gamma(-1/2)} \left(\frac{k(\alpha - \alpha_0)}{N\alpha} \right)^p \\ &= 1 + [1 - \rho(\mathbf{h})]^N \sum_{k=0}^N C_N^k \left(\frac{\rho(\mathbf{h})}{1 - \rho(\mathbf{h})} \right)^k \left(\sqrt{1 - \frac{k(\alpha - \alpha_0)}{N\alpha}} - 1 \right) \\ &= [1 - \rho(\mathbf{h})]^N \sum_{k=0}^N C_N^k \left(\frac{\rho(\mathbf{h})}{1 - \rho(\mathbf{h})} \right)^k \sqrt{1 - \frac{k(\alpha - \alpha_0)}{N\alpha}} \end{aligned}$$

La formule annoncée s'obtient par le changement de variable $k \rightarrow N - k$.

En ce qui concerne le processus à loi inverse gaussienne, bien que les facteurs de la loi bivariable n'aient pas encore été explicités¹¹, on connaît la relation liant le madogramme du processus isofactoriel et son corrélogramme (Hu, 1988) :

¹⁰ Cette terminologie se réfère aux modèles isofactoriels à loi marginale poissonnienne mais dont la loi bivariable n'est pas de diffusion (Chilès and Delfiner, 1999).

$$\gamma_1(\mathbf{h}) = E\{S(\alpha - (\alpha - \alpha_0)K/N, \beta)\} = [1 - \rho(\mathbf{h})]^N \sum_{k=0}^N C_N^k \left(\frac{\rho(\mathbf{h})}{1 - \rho(\mathbf{h})} \right)^k S(\alpha - (\alpha - \alpha_0)k/N, \beta)$$

où $S(a,b)$ est la sélectivité ou indicateur de dispersion de la loi de Gauss inverse de paramètres $\{a,b\}$, calculable par intégration numérique à partir de la formule :

$$S(a,b) = \int_0^{+\infty} F_{a,b}(y)[1 - F_{a,b}(y)]dy.$$

La section suivante s'intéresse à un autre modèle isofactoriel qui, bien qu'indéfiniment divisible, ne peut être obtenu par dilution poissonienne ou addition de mosaïques indépendantes : le modèle de diffusion bigamma.

V. Le modèle bigamma

V.1. Carré d'une fonction aléatoire bigaussienne

Si U est une fonction aléatoire dont les lois bivariées sont bigaussiennes réduites, alors $Y = U^2/2$ suit une loi bigamma de paramètre $\alpha = 0.5$.

Démonstration

La densité bigamma de paramètre $\alpha = 0.5$ et de coefficient de corrélation ρ s'exprime sous la forme (Matheron, 1973b ; Hu, 1988) :

$$f_\rho(y, y') = \frac{1}{\sqrt{\pi(1-\rho)}} \exp\left(-\frac{y+y'}{1-\rho}\right) \left(\frac{yy'}{\rho}\right)^{-1/4} I_{-1/2}\left(2\frac{\sqrt{yy'\rho}}{1-\rho}\right)$$

où $I_{1/2}$ est la fonction de Bessel modifiée de première espèce (Abramowitz and Stegun, 1972) : $I_{-1/2}(z) = \sqrt{\frac{2}{\pi z}} \text{ch}(z)$

En reportant dans l'expression de la densité bivariable, il vient :

$$\begin{aligned} f_\rho(y, y') &= \frac{1}{2\pi\sqrt{yy'(1-\rho)}} \exp\left(-\frac{y+y'}{1-\rho}\right) \left\{ \exp\left(2\frac{\sqrt{yy'\rho}}{1-\rho}\right) + \exp\left(-2\frac{\sqrt{yy'\rho}}{1-\rho}\right) \right\} \\ &= \frac{1}{2\pi\sqrt{yy'(1-\rho)}} \left\{ \exp\left(-\frac{1}{2(1-\rho)}(2y+2y'-4\sqrt{yy'\rho})\right) + \exp\left(-\frac{1}{2(1-\rho)}(2y+2y'+4\sqrt{yy'\rho})\right) \right\} \\ &= \frac{1}{2\sqrt{yy'}} \{ g_{\varepsilon\sqrt{\rho}}(\sqrt{2y}, \sqrt{2y'}) + g_{\varepsilon\sqrt{\rho}}(\sqrt{2y}, -\sqrt{2y'}) + g_{\varepsilon\sqrt{\rho}}(-\sqrt{2y}, \sqrt{2y'}) + g_{\varepsilon\sqrt{\rho}}(-\sqrt{2y}, -\sqrt{2y'}) \} \end{aligned}$$

où $\varepsilon \in \{+1, -1\}$ et $g_r(u, u')$ est la densité bigaussienne réduite de coefficient de corrélation r :

$$g_r(u, u') = \frac{1}{2\pi\sqrt{1-r^2}} \exp\left(-\frac{1}{2(1-r^2)}(u^2 - 2ruu' + u'^2)\right) = \frac{\partial^2 G_r(u, u')}{\partial u \partial u'}$$

¹¹ D'après Matheron (1973b), ce ne sont pas des polynômes ; Hu (1988) propose deux familles de fonctions orthonormées pour la loi marginale gaussienne inverse, mais ces familles ne restent pas orthogonales pour la loi bivariable.

Soit $\{Y, Y'\}$ un couple bigamma de paramètre $\alpha = 0.5$ et coefficient de corrélation ρ . D'après ce qui précède, on a :

$$\begin{aligned} \text{Prob}[Y < y, Y' < y'] &= \int_0^y \int_0^{y'} f_\rho(t, t') dt dt' \\ &= G_{\varepsilon\sqrt{\rho}}(\sqrt{2y}, \sqrt{2y'}) - G_{\varepsilon\sqrt{\rho}}(\sqrt{2y}, -\sqrt{2y'}) - G_{\varepsilon\sqrt{\rho}}(-\sqrt{2y}, \sqrt{2y'}) + G_{\varepsilon\sqrt{\rho}}(-\sqrt{2y}, -\sqrt{2y'}) \end{aligned}$$

On vérifie que cette fonction est nulle si $y = 0$ ou $y' = 0$. Soit à présent $\{U, U'\}$ un couple bigaussien réduit de corrélation $\varepsilon\rho^{1/2}$:

$$\begin{aligned} \text{Prob}[Y < y, Y' < y'] &= \text{Prob}[U < \sqrt{2y}, U' < \sqrt{2y'}] - \text{Prob}[U < \sqrt{2y}, U' < -\sqrt{2y'}] \\ &\quad - \text{Prob}[U < -\sqrt{2y}, U' < \sqrt{2y'}] + \text{Prob}[U < -\sqrt{2y}, U' < -\sqrt{2y'}] \\ &= \text{Prob}[U < \sqrt{2y}, -\sqrt{2y'} < U' < \sqrt{2y'}] - \text{Prob}[U < -\sqrt{2y}, -\sqrt{2y'} < U' < \sqrt{2y'}] \\ &= \text{Prob}[-\sqrt{2y} < U < \sqrt{2y}, -\sqrt{2y'} < U' < \sqrt{2y'}] \\ &= \text{Prob}[U^2/2 < y, U'^2/2 < y'] \end{aligned}$$

ce qui établit que la loi de $\{Y, Y'\}$ coïncide avec celle de $\{U^2/2, U'^2/2\}$.

Ce résultat montre que, pour construire une fonction aléatoire à loi bigamma de paramètre $\alpha = 0.5$, il suffit d'élever au carré une fonction aléatoire multigaussienne. On obtient des images à textures très variées selon le choix du corrélogramme de la fonction gaussienne (figure 3.19).

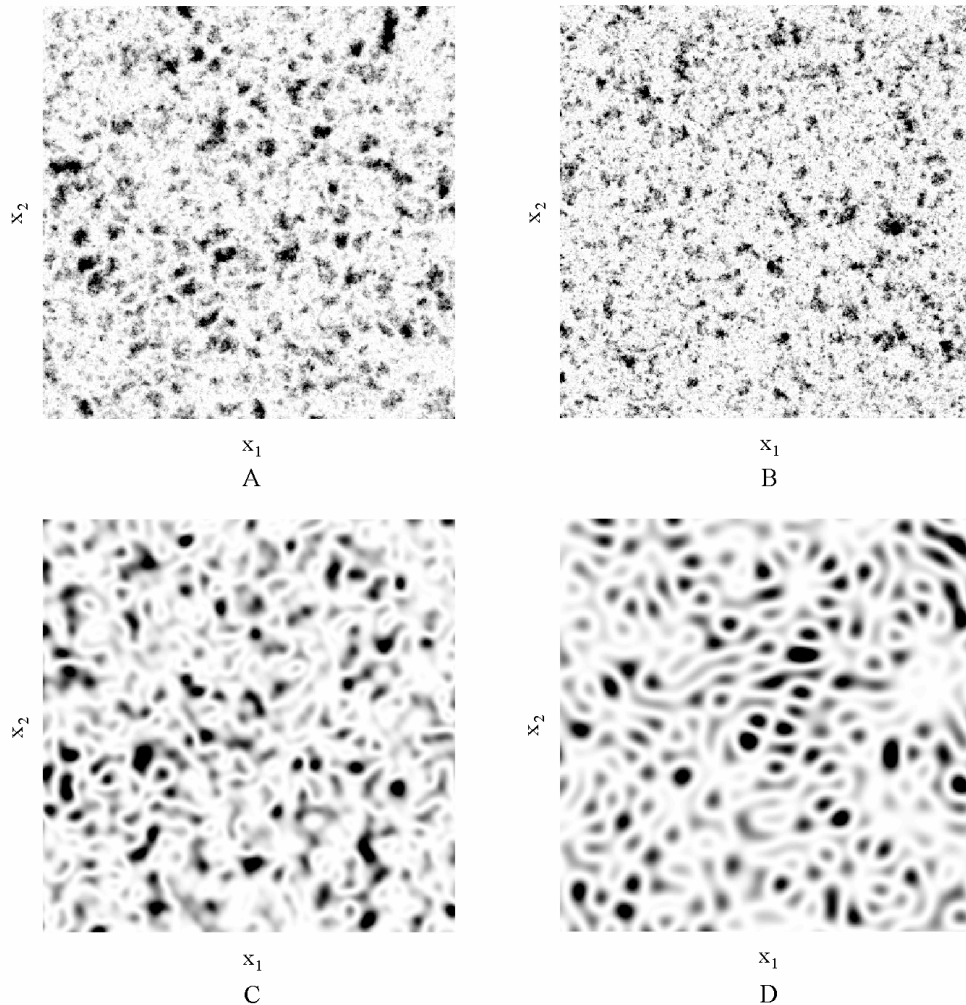


Figure 3.19. Réalisations de processus bigamma construits à partir de fonctions aléatoires gaussiennes de corrélogramme A, sphérique, B, exponentiel, C, gaussien et D, sinus cardinal

Plusieurs conséquences découlent de cette construction.

1) Relation entre polynômes d'Hermite et de Laguerre d'ordre 1/2

Si U est une fonction aléatoire multigaussienne, alors $Y = U^2/2$ est un processus isofactoriel dont les facteurs s'expriment en fonction des polynômes d'Hermite normés (Rivoirard, 1994) d'ordre pair :

$$\forall p \in \mathbf{N}, \chi_p(Y) = H_{2p}(\sqrt{2Y}).$$

En effet, H_{2p} est un polynôme pair de degré $2p$, de sorte que χ_p est un polynôme de degré p . Ainsi, la suite $\{\chi_p, p \in \mathbf{N}\}$ forme un système polynomial complet et, de plus, orthonormé :

$$\forall p, q \in \mathbf{N}, \text{cov}[\chi_p(Y), \chi_q(Y')] = 0 \text{ si } p \neq q.$$

Or, d'après ce qui précède, Y est une fonction aléatoire à loi bigamma, donc ses facteurs sont les polynômes de Laguerre normés d'ordre $\alpha = 0.5$ (Hu, 1988) :

$$\forall p \in \mathbf{N}, \chi_p(Y) = (-1)^p L_p^{1/2}(Y).$$

Le coefficient $(-1)^p$ a été introduit afin que le signe du terme dominant soit positif, à l'instar de celui des polynômes d'Hermite d'ordre pair. On en déduit une identité entre polynômes d'Hermite et de Laguerre :

$$\forall p \in \mathbf{N}, \forall y \geq 0, L_p^{1/2}(y) = (-1)^p H_{2p}(\sqrt{2y}).$$

2) Cohérence interne du modèle

Le corrélogramme $\rho(\mathbf{h})$ du processus bigamma est le carré de celui du processus multigaussien. On retrouve la condition nécessaire $\rho(\mathbf{h}) \geq 0$ pour que la densité bigamma soit positive (Matheron, 1973b). Mais on obtient également une condition suffisante pour que $\rho(\mathbf{h})$ soit un modèle autorisé : il suffit qu'il s'exprime comme le carré d'un corrélogramme. Parmi les modèles les plus connus, c'est le cas des schémas exponentiel, stable, gamma et de Cauchy ; en revanche, les modèles triangulaire et sphérique ne sont pas autorisés, même dans l'espace à une dimension.

3) Proportionnalité entre le madogramme et la racine carrée du variogramme

L'écart entre deux valeurs gamma de coefficient de corrélation ρ peut s'écrire sous la forme

$$Y - Y' = \sqrt{1-\rho} AB \text{ où l'on a posé } A = \frac{U - U'}{\sqrt{2 - 2\sqrt{\rho}}} \text{ et } B = \frac{U + U'}{\sqrt{2 + 2\sqrt{\rho}}}.$$

A et B sont deux variables gaussiennes indépendantes, de moyennes nulles et variances unitaires. D'après Matheron (1982), on a :

$$E|A| = E|B| = \sqrt{2/\pi}$$

de sorte que le madogramme de Y et Y' est proportionnel à la racine carrée de leur variogramme :

$$\frac{1}{2} E|Y - Y'| = \frac{1}{\pi} \sqrt{1-\rho}.$$

V.2. Modèle bigamma de paramètre demi-entier

Une fonction aléatoire Y bigamma de paramètre $\alpha = m/2$ et corrélogramme $\rho(\mathbf{h})$ peut être obtenue en additionnant m fonctions aléatoires mutuellement indépendantes¹², à loi bigamma de paramètre $1/2$ et corrélogramme $\rho(\mathbf{h})$:

$$\forall \mathbf{x} \in \mathbb{R}^d, Y(\mathbf{x}) = \sum_{i=1}^m Y_i(\mathbf{x}) = \frac{1}{2} \sum_{i=1}^m U_i^2(\mathbf{x})$$

où $\{U_i, i = 1 \dots m\}$ sont des fonctions aléatoires multigaussiennes mutuellement indépendantes. Dans la suite, on notera $\mathbf{U}(\mathbf{x})$ le vecteur aléatoire dont les composantes sont $\{U_i(\mathbf{x}), i = 1 \dots m\}$:

$$\forall \mathbf{x} \in \mathbb{R}^d, 2Y(\mathbf{x}) = |\mathbf{U}(\mathbf{x})|^2$$

Ce modèle pourrait être baptisé de “chi carré” (car sa loi marginale est un chi carré à m degrés de liberté). Dans le cas où $m = 2$, Alfaro (1979) a qualifié de “processus de Rayleigh” la racine carrée de $Y(\mathbf{x})$.

La simulation d’un tel modèle revient à simuler m fonctions aléatoires gaussiennes indépendantes, soumises à des conditions sur la norme du vecteur \mathbf{U} aux sites d’échantillonnage $\{\mathbf{x}_\alpha, \alpha = 1 \dots n\}$. Pour inclure de telles contraintes non linéaires, il est avantageux de recourir à un échantillonneur de Gibbs (Freulon, 1992, 1994 ; voir aussi l’annexe C).

Initialisation

En chaque site \mathbf{x}_α ($\alpha = 1 \dots n$), on peut par exemple poser

$$\forall i \in \{1 \dots m\}, U_i(\mathbf{x}_\alpha) = \varepsilon_{i,\alpha} \sqrt{\frac{2y(\mathbf{x}_\alpha)}{m}} \text{ où } \varepsilon_{i,\alpha} = 1 \text{ ou } -1 \text{ avec équiprobabilité}$$

Phase itérative

- 1) choisir un site \mathbf{x}_{α_0} au hasard parmi $\{\mathbf{x}_\alpha, \alpha = 1 \dots n\}$;
- 2) remplacer $\mathbf{U}(\mathbf{x}_{\alpha_0})$ par un vecteur simulé selon la distribution conditionnelle à $\{\mathbf{U}(\mathbf{x}_\alpha), \alpha \neq \alpha_0\}$ et à $|\mathbf{U}(\mathbf{x}_{\alpha_0})|^2 = 2y(\mathbf{x}_{\alpha_0})$.

Conditionnellement aux valeurs $\{\mathbf{U}(\mathbf{x}_\alpha), \alpha \neq \alpha_0\}$, $\mathbf{U}(\mathbf{x}_{\alpha_0})$ est un vecteur gaussien dont la moyenne est le vecteur \mathbf{u}^* des krigeages de chaque composante et dont la matrice de variance-covariance est proportionnelle à la matrice identité¹³. Ainsi, dans l’espace \mathbb{R}^m , la densité de probabilité du vecteur $\mathbf{U}(\mathbf{x}_{\alpha_0})$ conditionnée par $\{\mathbf{U}(\mathbf{x}_\alpha), \alpha \neq \alpha_0\}$ est

$$f(\mathbf{u}) = \frac{1}{(2\pi)^{m/2} \sigma^{*m}} \exp \left\{ -\frac{1}{2\sigma^{*2}} |\mathbf{u} - \mathbf{u}^*|^2 \right\}$$

¹² Cette propriété découle de l’indéfinie divisibilité de la transformée de Laplace d’une loi bigamma de paramètre α :

$$E[e^{-\lambda Y - \mu Y'}] = [1 + \lambda + \mu + (1 - \rho)\lambda\mu]^{-\alpha} .$$

¹³ Les composantes de $\mathbf{U}(\mathbf{x})$ sont indépendantes et ont le même corrélogramme, ce qui explique que les termes de covariance soient nuls et que les termes de variance égaux à la variance de krigeage de l’une quelconque des composantes, notée σ^{*2} .

Les courbes d'isodensité sont des sphères centrées en \mathbf{u}^* . Il s'agit à présent d'échantillonner cette densité multivariable sous la contrainte que la norme $\|\mathbf{U}(\mathbf{x}_{\alpha_0})\|$ soit égale à $R = [2y(\mathbf{x}_{\alpha_0})]^{1/2}$, ce qui revient à choisir un point sur la sphère $S(\mathbf{0}, R)$ centrée en $\mathbf{0}$ et de rayon R (figure 3.20). Ce choix peut être réalisé grâce à l'algorithme de Metropolis (*cf.* annexe C).

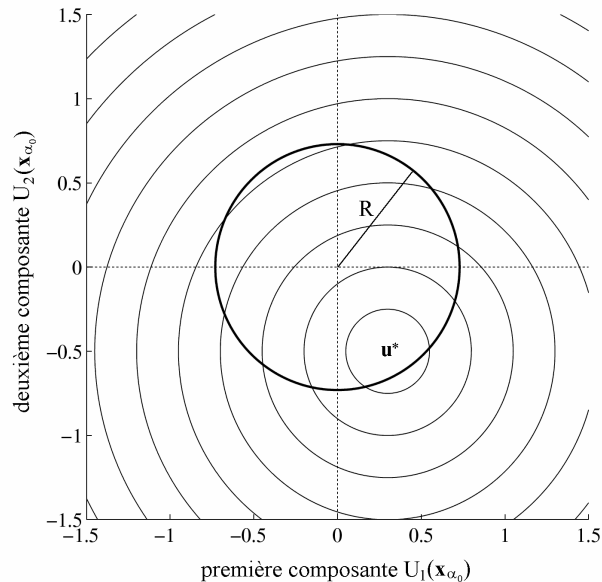


Figure 3.20. Courbes d'isodensité du vecteur $\mathbf{U}(\mathbf{x}_{\alpha_0})$ et contrainte sur la norme de ce vecteur

L'étape 2) peut alors être décomposée en :

- 2a) poser $\mathbf{u} = (R, 0, \dots, 0)$;
- 2b) choisir un état candidat \mathbf{u}' uniforme sur la sphère $S(\mathbf{0}, R)$;
- 2c) choisir V uniforme sur $[0, 1]$;
- 2d) si $f(\mathbf{u}) V < f(\mathbf{u}')$, remplacer \mathbf{u} par \mathbf{u}' ;
- 2e) retourner en 2b).

A l'issue de la phase itérative, le vecteur \mathbf{U} est simulé en chaque site de donnée. On obtient une simulation du processus bigamma dans l'espace \mathbb{R}^d entier en simulant chacune des composantes $U_i(\mathbf{x})$ conditionnellement à $\{U_i(\mathbf{x}_\alpha), \alpha = 1 \dots n\}$.

A titre d'illustration, une simulation non conditionnelle d'une fonction aléatoire à loi exponentielle (paramètre $\alpha = 1$) est menée sur une grille de taille 400×400 (figure 3.21A). Le corrélogramme des composantes gaussiennes comprend deux modèles sphériques gigognes de palier 0.5 et portées 20 et 100 respectivement. Une centaine d'échantillons sont ensuite extraits de la réalisation (figure 3.21B) et sont utilisés comme données conditionnantes pour obtenir deux simulations conditionnelles (figures 3.21C et 3.21D).

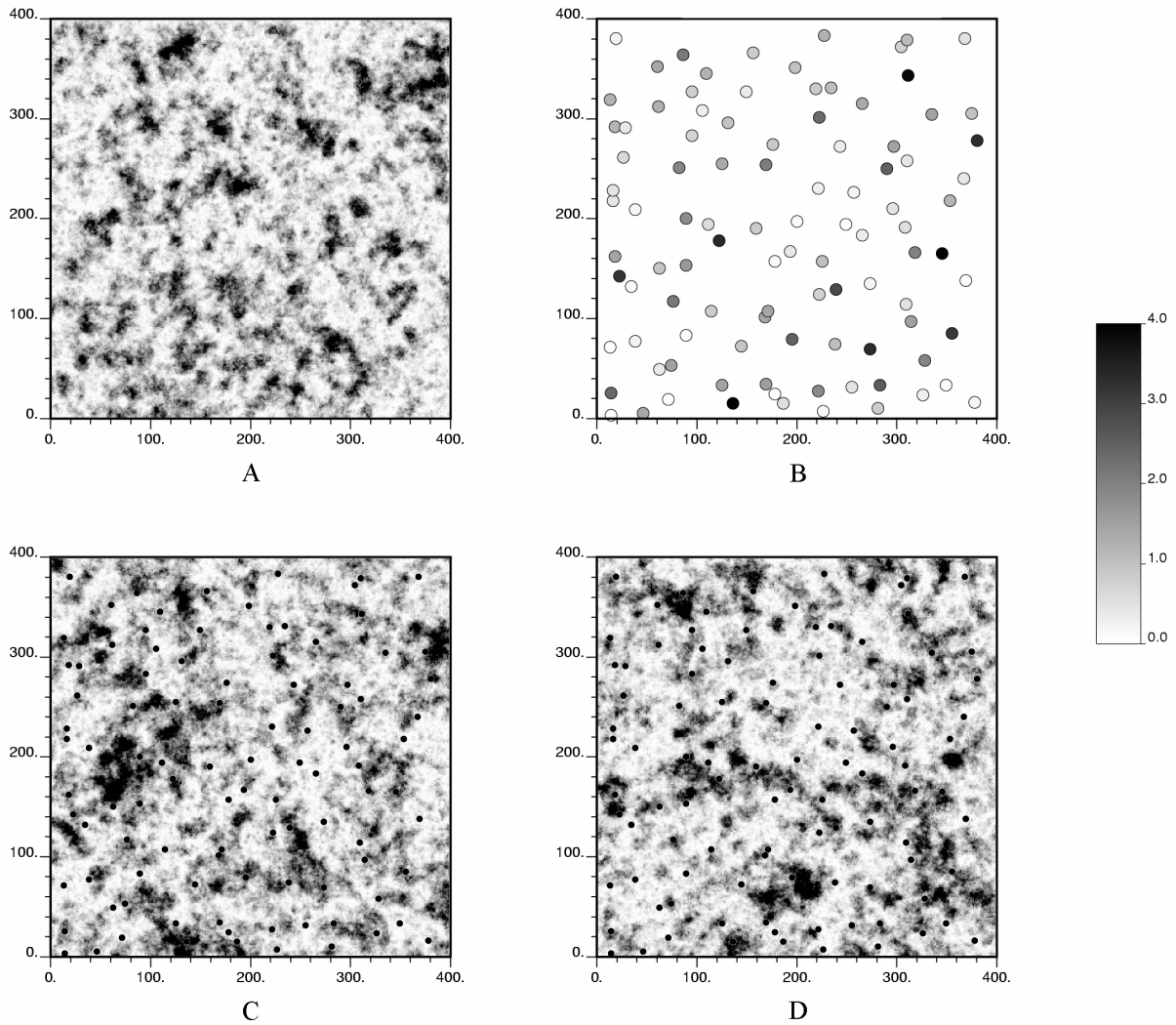


Figure 3.21. A, simulation non conditionnelle ; B, données conditionnantes ; C et D, deux simulations conditionnelles (les données sont superposées)

Remarques

- 1) La classe des corrélogrammes que l'on peut simuler par cette procédure est très vaste. Il est aussi possible d'envisager la simulation de modèles laguerriens par la procédure de substitution, avec les mêmes extensions que celles proposées dans le cas hermitien (fonction directrice multivariable et processus de codage à covariance factorisée).
- 2) La construction proposée du modèle bigamma est une extension au cadre continu de la technique des plurigaussiennes (Armstrong *et al.*, 2003) : dans l'espace \mathbb{R}^m dont les axes sont les gaussiennes $\{U_i(\mathbf{x}), i = 1 \dots m\}$, la valeur du processus bigamma est constante sur chacune des sphères centrées à l'origine et est égale à la moitié du carré du rayon de cette sphère.

Chapitre 4 :

L'inférence statistique

Ce chapitre vise à définir les conditions sous lesquelles les paramètres des lois (histogramme, variogramme, etc.) sont compatibles avec le modèle choisi et comment déterminer ces paramètres.

I. Détermination de la loi marginale

I.1. Considérations générales

La modélisation d'une loi marginale présente deux difficultés majeures :

- 1) la représentativité de l'histogramme empirique des données lorsque la répartition spatiale de ces dernières est très irrégulière voire préférentielle. Pour minimiser le risque d'erreur radicale, il est recommandé de pondérer les données dans le calcul de l'histogramme expérimental. Plusieurs techniques ont été proposées à cet effet, reposant sur des critères de nature géométrique (Journal, 1983 ; Goovaerts, 1997 ; Deutsch and Journel, 1998) ou structurale (Matheron, 1977 ; Bourgault, 1997). Par exemple, citons la *méthode des cellules*, qui consiste à paver le domaine échantillonné par des cellules rectangulaires identiques, puis à assigner à chaque donnée un poids inversement proportionnel au nombre d'échantillons présents dans la cellule à laquelle appartient cette donnée.
- 2) la modélisation des queues de distribution, particulièrement incommode lorsque l'histogramme des données est fort dissymétrique. Il est nécessaire de choisir un comportement pour les valeurs extrêmes, c'est-à-dire celles qui dépassent la plus grande valeur échantillonnée. Un tel choix doit être guidé par des considérations pratiques (connaissance du phénomène, voire intuition) et peut-être aussi par des raisons de cohérence avec les autres paramètres du modèle, en particulier le variogramme¹.

La distribution marginale des valeurs est requise dans l'algorithme de simulation séquentielle d'indicateurs pour spécifier les moyennes de ces dernières, ainsi que dans la simulation des modèles isofactoriels pour transformer la variable initiale en une variable suivant la loi marginale du modèle (procédure d'anamorphose). Une hypothèse de stationnarité est nécessaire pour pouvoir inférer correctement cette distribution (Matheron, 1974).

¹ Par exemple, un modèle variographique aussi usuel que le schéma sphérique n'est pas toujours compatible avec une fonction aléatoire bilognormale (Matheron, 1989a). Il convient toutefois de signaler que, au cours de la modélisation des processus isofactoriels, l'analyse variographique repose non sur la variable lognormale mais sur son logarithme, pour lequel le schéma sphérique est autorisé. Aussi, la modélisation de la loi marginale n'est-elle pas directement liée à l'analyse variographique, car entre les deux intervient une transformation ou anamorphose des valeurs.

I.2. Commentaires sur la procédure d'anamorphose

La pratique actuelle (Lantuéjoul et Rivoirard, 1984 ; Hu et Lantuéjoul, 1988 ; Chilès and Delfiner, 1999) consiste à calculer une *anamorphose empirique* et à lui ajuster un polynôme, dont le degré est généralement inférieur à 100. Dans le cadre bigaussien ou hermitien, l'anamorphose s'écrit comme un développement en série de polynômes d'Hermite, tronqué à un certain ordre ; dans le cadre bigamma ou laguerrien, il est avantageux d'utiliser des polynômes de Laguerre.

L'avantage de cette procédure est que les coefficients du développement polynomial permettent de mener à bien le calcul des "réserves récupérables" par krigeage disjonctif. Elle présente toutefois plusieurs inconvénients :

- problème pour la modélisation des queues de la distribution, car le développement polynomial (de degré fini) diverge à l'infini (figures 4.1C et 4.1D) ; une correction est inévitable (Lajaunie, 1993) ;
- manque de robustesse de l'ajustement polynomial : l'anamorphose empirique est une fonction en escalier, de sorte qu'un ajustement global par un polynôme est sensible à l'existence de marches de grande amplitude (données extrêmes ou atypiques, cf. figure 4.1D). Ceci détériore la qualité de l'ajustement des marches de faible amplitude qui renferment souvent beaucoup de détails (figure 4.1C) et peut conduire à des valeurs anamorphosées de qualité médiocre (figure 4.1B).

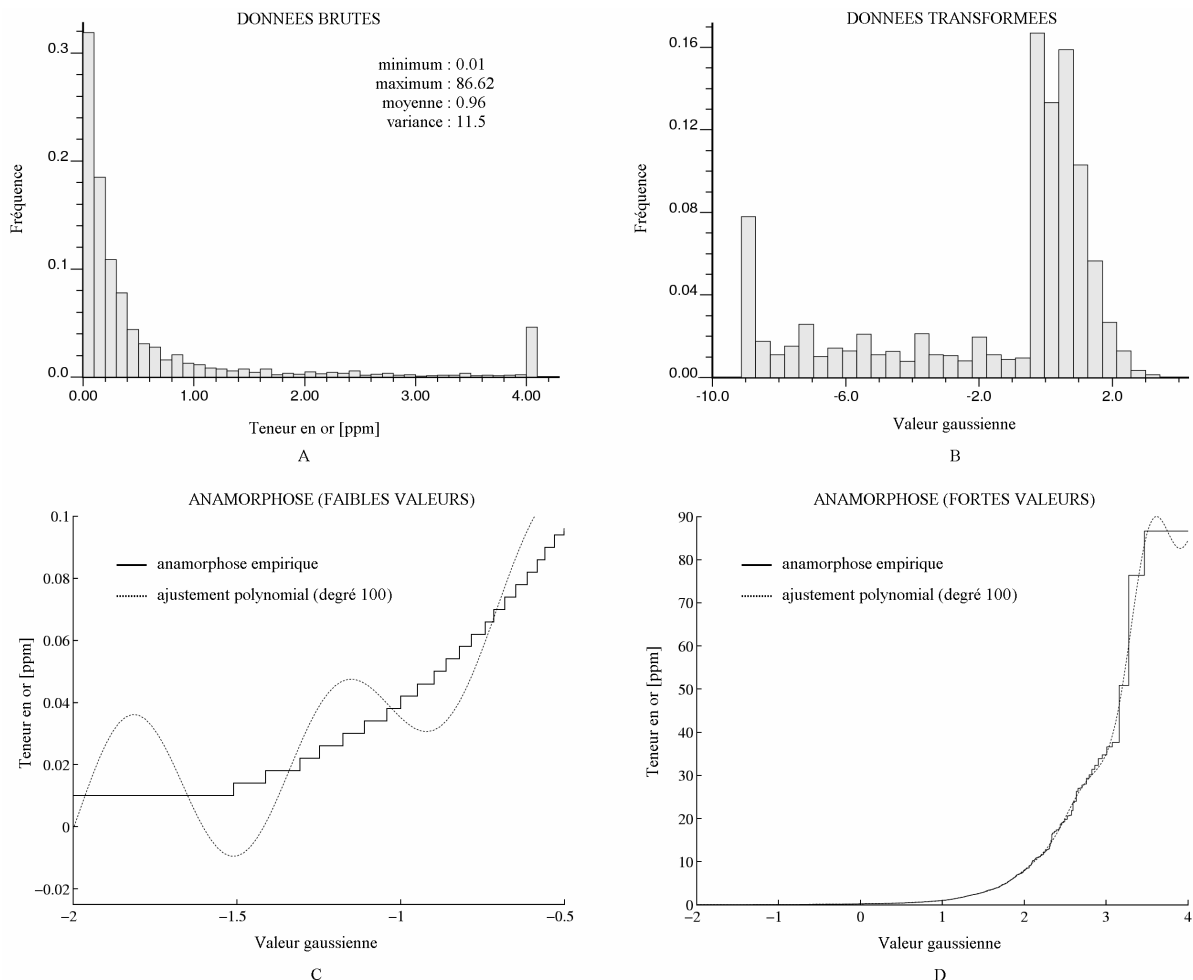


Figure 4.1. Anamorphose gaussienne sur des données à distribution dissymétrique

Une procédure alternative consiste à effectuer un ajustement *local* de l'anamorphose empirique, vue comme un nuage de points et non plus comme une fonction en escalier (figure 4.2C) : par exemple, on peut joindre les points expérimentaux par des segments de droite (figure 4.2D). La principale difficulté est la définition des extrema de la variable initiale, en particulier la valeur maximale possible, ainsi que le mode d'extrapolation au-delà du dernier point expérimental, par exemple par une fonction exponentielle. Sur les données de la figure 4.1, l'anamorphose conduit à des résultats plus satisfaisants (figure 4.2B), bien qu'imparfaits à cause de la présence d'un effet zéro (6.5% des données initiales sont égales à la valeur minimale de 0.01 ppm et conduisent à un atome dans l'histogramme gaussien).

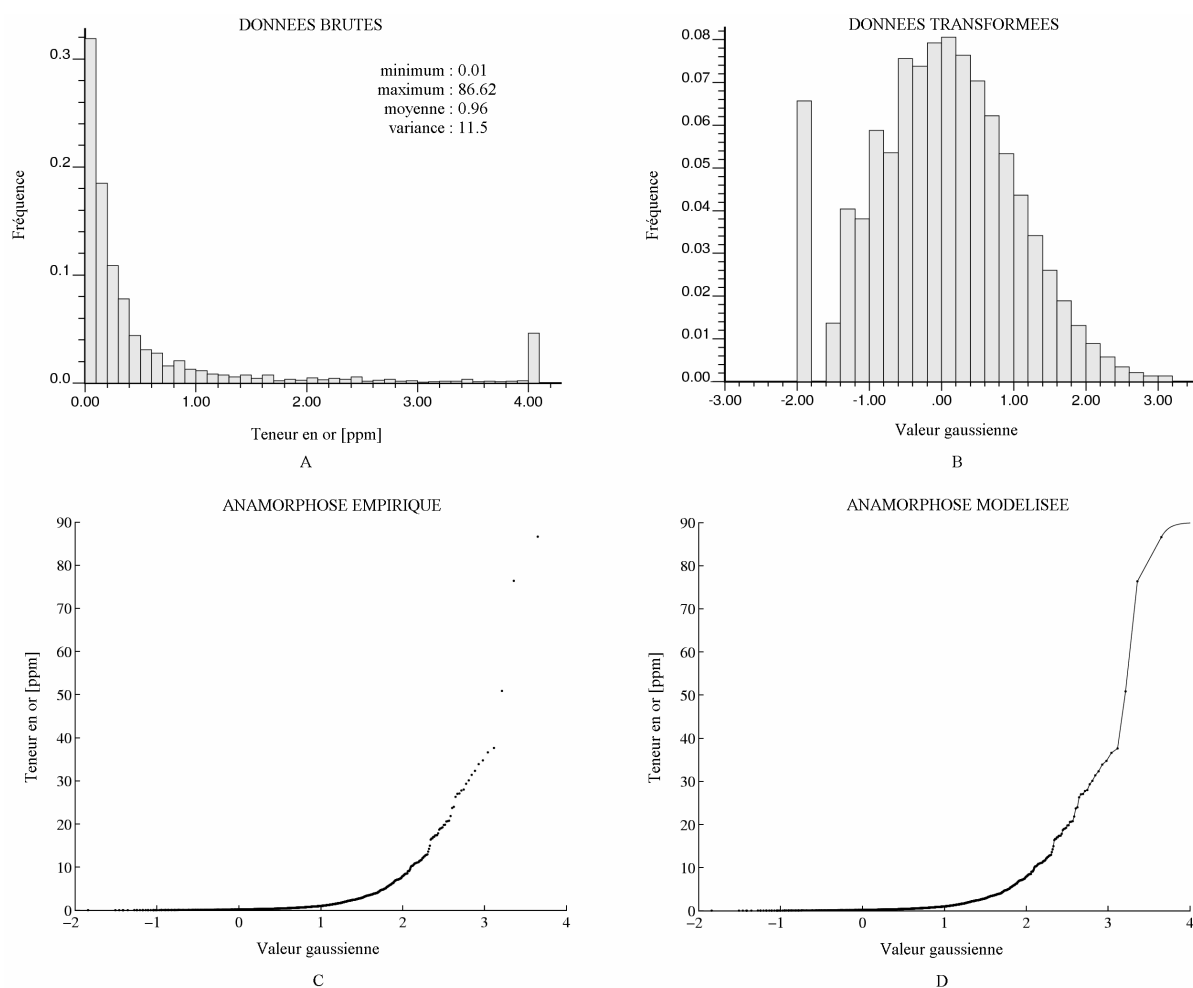


Figure 4.2. Anamorphose gaussienne sur des données à distribution dissymétrique

Bien entendu, sur l'exemple précédent, le choix d'une anamorphose gaussienne est discutable et le recours à une loi gamma de faible paramètre serait sans doute plus adéquat. La détermination exacte de ce paramètre reste à définir, principalement en fonction des lois bivariées (Hu, 1988), mais aussi de la loi marginale : il est en effet souhaitable que l'anamorphose ne déforme pas trop l'histogramme des données initiales. Ce souhait répond à des considérations d'ordre pratique plutôt que statistique : une transformation violente risque d'induire des interprétations sujettes à caution, par exemple si elle amplifie démesurément les écarts entre des valeurs initialement très voisines et ne différant en fait que par un bruit de mesure. On évitera ainsi de transformer un histogramme extrêmement dissymétrique en un histogramme gaussien, ce qui est d'ailleurs déconseillé par plusieurs auteurs (Hu, 1988 ; Rivoirard, 1994).

II. Modélisation des lois bivariabes

II.1. Méthodes d'indicatrices

Une limitation des méthodes qui reposent sur le krigeage d'indicatrices est le choix d'un modèle variographique cohérent. Cette section a pour but de mettre en relief les principales difficultés de la modélisation.

II.1.1. Analyse variographique de chaque indicatrice

Les modèles autorisés pour décrire des variables indicatrices sont très restrictifs. En effet, outre la propriété de négativité conditionnelle classique, le variogramme $\gamma_{I,y}$ de l'indicatrice associée au seuil de coupure y doit être borné (maximum = 1/2, palier maximal = 1/4) et vérifier une autre propriété de négativité conditionnelle (Matheron, 1993) :

$$\forall k > 2, \forall \mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^d, \forall \varepsilon_1, \dots, \varepsilon_k \in \{-1, 1\} \text{ tels que } \sum_{i=1}^k \varepsilon_i = 1, \sum_{i=1}^k \sum_{j=1}^k \varepsilon_i \varepsilon_j \gamma_{I,y}(\mathbf{x}_i - \mathbf{x}_j) \leq 0.$$

En particulier, en choisissant trois sites $\{\mathbf{x}, \mathbf{x} + \mathbf{h}_1, \mathbf{x} - \mathbf{h}_2\}$ affectés des poids $\{-1, 1, 1\}$, on obtient une inégalité triangulaire qui interdit que le comportement à l'origine soit convexe (Matheron, 1987, 1989a) :

$$\forall \mathbf{h}_1, \mathbf{h}_2 \in \mathbb{R}^d, \gamma_{I,y}(\mathbf{h}_1 + \mathbf{h}_2) \leq \gamma_{I,y}(\mathbf{h}_1) + \gamma_{I,y}(\mathbf{h}_2).$$

Cette inégalité empêche que le comportement à l'origine soit plus régulier qu'un schéma linéaire :

$$\text{si } \gamma_{I,y}(\mathbf{h}) \underset{|\mathbf{h}| \rightarrow 0}{\sim} |\mathbf{h}|^\theta, \text{ alors } \theta \leq 1.$$

On ignore si les conditions précédentes suffisent pour qu'un variogramme soit un variogramme d'indicatrice. Parmi les modèles autorisés dans l'espace multidimensionnel, citons l'effet de pépite, le schéma exponentiel et les mélanges de ces derniers (covariances complètement monotones). Dans l'espace à une dimension, tout corrélogramme convexe sur \mathbb{R}_+ est compatible avec une indicatrice, par exemple les modèles triangulaire et sphérique.

Ce dernier résultat vient du fait qu'il est toujours possible de construire une mosaïque unidimensionnelle admettant un corrélogramme convexe donné. En effet, une fonction convexe $K(h)$ peut s'interpréter comme une courbe de bénéfice conventionnel associée à une fonction de répartition F (Matheron, 1981a ; Lantuéjoul, 1990) et s'écrire sous la forme :

$$K(h) = \int_h^{+\infty} (\ell - h) F(d\ell).$$

On reconnaît l'expression du covariogramme géométrique moyen d'un segment dont la longueur ℓ est aléatoire et régie par la loi F . On obtient la mosaïque souhaitée en juxtaposant des segments indépendants ayant ce covariogramme géométrique. Cette construction a d'ailleurs été proposée par Matheron (1988) pour simuler des fonctions aléatoires à variogramme concave sur \mathbb{R}_+ .

Ajoutons que toute fonction paire convexe sur \mathbb{R}_+ et s'annulant à l'infini est une fonction de covariance dans l'espace à une dimension, d'après un théorème dû à Pólya (1918). Les fonctions complètement monotones ne constituent qu'une classe restreinte de fonctions convexes.

II.1.2. Analyse variographique conjointe

La modélisation conjointe d'un ensemble de variogrammes d'indicatrices est très délicate. D'une part, il existe des contraintes entre les variogrammes simples et croisés, car les indicatrices ne sont jamais indépendantes (Rivoirard, 1993). D'autre part, même le modèle linéaire de corégionalisation – d'usage courant en géostatistique multivariable (Wackernagel, 1998) – peut se révéler inadapté, car il impose que tous ces variogrammes aient la même portée, ce qui est souvent incompatible avec la propriété de déstructuration des hautes teneurs observée sur la plupart des régionalisations.

II.2. Modèles isofactoriels : généralités

La plupart des lois bivariées isofactorielles stationnaires associées à une distance vectorielle \mathbf{h} s'expriment sous la forme (cf. annexe B)

$$F_{\mathbf{h}}(dy, dy') = F(dy) F(dy') \sum_{p \geq 0} E[R(\mathbf{h})^p] \chi_p(y) \chi_p(y')$$

où $F(dy)$ est la densité marginale

$\{\chi_p, p \in \mathbb{N}\}$ sont les facteurs de la loi isofactorielle

$R(\mathbf{h})$ est une variable aléatoire, dont les moments constituent la suite des covariances des facteurs.

Le type de modèle isofactoriel étant défini, seule la variable $R(\mathbf{h})$ doit être déterminée. Il convient donc de chercher les conditions sur cette variable assurant la cohérence du modèle. Matheron (1973b) énonce une condition nécessaire pour que la famille de lois bivariées $\{F_{\mathbf{h}}(dy, dy'), \mathbf{h} \in \mathbb{R}^d\}$ puisse être associée à une fonction aléatoire : pour tout choix de l'entier k , on doit avoir

$$\forall \mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^d, \forall \varphi_1, \dots, \varphi_k \in L^2(\mathbb{R}), E\left[\left(\sum_{i=1}^k \varphi_i(Y(\mathbf{x}_i))\right)^2\right] = \sum_{i,j=1}^k \int \varphi_i(y) \varphi_j(y') F_{\mathbf{x}_i - \mathbf{x}_j}(dy, dy') \geq 0.$$

Mais une telle condition ne suffit sans doute pas à assurer l'existence d'un processus isofactoriel ayant ces lois bivariées.

En général, les corrélogrammes des facteurs – donc la variable $R(\mathbf{h})$ – se déduisent de celui de la variable anamorphosée, noté $\rho(\mathbf{h})$ ci-après, par une formule simple ne mettant en jeu que des paramètres scalaires. La question reste de savoir quels corrélogrammes sont compatibles avec une loi isofactorielle donnée. Si le modèle bigaussien, restriction du modèle multigaussien, ne pose aucune limitation du point de vue variographique², il n'en est pas de même des autres modèles isofactoriels. Ainsi, dans le cadre mosaïque, le corrélogramme du processus est égal au corrélogramme de ses indicatrices, donc doit vérifier les contraintes propres aux fonctions aléatoires en tout ou rien (Matheron, 1993). D'ores et déjà, par rapport aux méthodes d'indicatrices, le recours à un modèle isofactoriel facilite grandement la recherche de conditions de cohérence interne, puisque l'on ne manipule désormais qu'un seul corrélogramme (celui de la variable anamorphosée) : le problème de l'analyse variographique multivariable est résolu.

Pour trouver des conditions nécessaires de cohérence, un moyen élémentaire consiste à examiner le madogramme $\gamma_1(\mathbf{h})$ du processus isofactoriel. D'après Matheron (1982), ce dernier est la somme des variogrammes d'indicatrices : il doit satisfaire l'inégalité triangulaire et ne peut être plus régulier qu'un schéma linéaire à l'origine. On pourra examiner la contrainte que cela implique sur le modèle de corrélogramme $\rho(\mathbf{h})$.

² Pour toute fonction de type positif, un théorème dû à Kolmogorov assure l'existence d'un processus gaussien admettant cette fonction comme covariance.

Par exemple, si la variable $R(\mathbf{h})$ suit une loi beta de paramètres $\{\alpha\rho(\mathbf{h}), \alpha[1-\rho(\mathbf{h})]\}$ (cf. chapitre 3 pour l'exemple des modèles de Laguerre et de Meixner), le madogramme normé a pour expression (Chilès and Delfiner, 1999) :

$$\frac{\gamma_1(\mathbf{h})}{\gamma_1(\infty)} = \frac{\Gamma[\alpha(1-\rho(\mathbf{h})) + 1/2] \Gamma(\alpha)}{\Gamma[\alpha(1-\rho(\mathbf{h}))] \Gamma(\alpha + 1/2)}$$

Par suite, si $1 - \rho(\mathbf{h}) \underset{|\mathbf{h}| \rightarrow 0}{\sim} \omega |\mathbf{h}|^\theta$, alors

$$\begin{cases} \gamma_1(\mathbf{h}) \underset{|\mathbf{h}| \rightarrow 0}{\sim} \omega' |\mathbf{h}|^{\theta/2} & \text{si } \alpha = \infty \text{ (modèle de diffusion)} \\ \gamma_1(\mathbf{h}) \underset{|\mathbf{h}| \rightarrow 0}{\sim} \omega' |\mathbf{h}|^\theta & \text{sinon} \end{cases}$$

En conclusion, excepté le modèle de diffusion, on ne peut utiliser au plus qu'un corrélogramme linéaire à l'origine ($\theta \leq 1$).

Quelques résultats

- 1) Les corrélogrammes compatibles avec un modèle mosaïque sont également compatibles avec les modèles isofactoriels à lois indéfiniment divisibles ; ceci résulte de la construction de ces derniers par addition de modèles mosaïques (cf. chapitre 3). De même, les corrélogrammes d'indicatrices sont compatibles avec le modèle binomial, celui-ci étant une somme d'ensembles aléatoires.
- 2) La covariance exponentielle est compatible avec tous les types de loi isofactorielle, du moins dans l'espace à une dimension : pour cela, il suffit de construire un modèle markovien à temps discret (cf. chapitre 2 ; se référer aussi à Matheron, 1975b et 1989b). Plus généralement, les covariances complètement monotones sont compatibles avec tous les types de loi isofactorielle dans l'espace unidimensionnel.

En effet, un corrélogramme complètement monotone s'écrit comme la superposition de schémas exponentiels dont la portée a est régie par une loi de probabilité $\varpi(da)$:

$$\rho(\mathbf{h}) = \int_0^{+\infty} e^{-3\mathbf{h}/a} \varpi(da) .$$

Il suffit donc de considérer le processus égal, avec la probabilité $\varpi(da)$, à un processus isofactoriel markovien dont la covariance est une exponentielle de portée a . Un tel processus n'est toutefois pas ergodique.

Il est possible de définir un processus ergodique si la mesure de probabilité ϖ est à support borné. Dans ce cas, $\rho(\mathbf{h})$ est le produit d'un schéma exponentiel de portée finie et d'un corrélogramme complètement monotone : il s'agit du corrélogramme du processus composé obtenu en définissant un processus identique au précédent dans chaque compartiment d'une mosaïque de Poisson, indépendamment d'un compartiment à l'autre (cf. chapitre 2). Sur un intervalle d'espace très grand, on aura une succession de nombreux processus indépendants, ce qui assure l'ergodicité du processus composé.

A défaut de critères généraux de cohérence interne, on se contentera de choisir la loi marginale et le corrélogramme de la variable anamorphosée parmi les modèles dont on a pu expliciter un mode de construction, par exemple les modèles à résidus d'indicatrices, à loi marginale indéfiniment divisible ou à loi bigamma (cf. chapitre 3). Au contraire, on évitera de recourir à des algorithmes "aveugles", comme l'algorithme séquentiel qui ne comporte aucun garde-fou sur la cohérence interne du modèle et dont les résultats sont très imparfaits (cf. chapitre 2).

II.3. L'exemple du modèle de Laguerre

Le modèle de Laguerre (Matheron, 1973b, 1975b ; Hu, 1988 ; Chilès and Liao, 1993) est l'un des modèles isofactoriels les plus connus après les modèles bigaussien et hermitien. Son intérêt réside dans la possibilité de modéliser des distributions très dissymétriques. Pour réaliser l'inférence statistique d'un tel modèle, on se penchera sur la propriété de déstructuration des teneurs extrêmes, selon laquelle les variables indicatrices deviennent pépitiqes lorsque le seuil de coupure est très faible ou très élevé.

II.3.1. Retour sur la propriété de déstructuration des teneurs extrêmes

Considérons une fonction aléatoire $Y(\mathbf{x})$ dont les lois bivariées sont laguerriennes et notons α le paramètre de la loi marginale gamma. La densité d'un couple $\{Y(\mathbf{x} + \mathbf{h}), Y(\mathbf{x})\}$ s'écrit sous la forme

$$f_{\alpha, \mathbf{h}}(y, y') = f_{\alpha}(y) f_{\alpha}(y') \sum_{p \geq 0} C_p^{(Y)}(\mathbf{h}) L_p^{\alpha}(y) L_p^{\alpha}(y')$$

où $f_{\alpha}(y)$ est la densité de la loi gamma de paramètre α
 $\{L_p^{\alpha}, p \in \mathbf{N}\}$ sont les polynômes de Laguerre normés
 $C_p^{(Y)}(\mathbf{h})$ sont les corrélogrammes de ces polynômes.

La déstructuration des teneurs extrêmes peut être étudiée en examinant le comportement des corrélogrammes des fonctions indicatrices $I(Y(\mathbf{x}); y)$ en fonction du seuil de coupure y . Comme toutes les fonctions de variance finie, les indicatrices se développent en polynômes de Laguerre normés d'ordre α $\{L_p^{\alpha}, p \in \mathbf{N}\}$:

$$I(Y(\mathbf{x}); y) = F_{\alpha}(y) + \sum_{p=1}^{+\infty} \sqrt{\frac{\alpha}{p}} f_{\alpha+1}(y) L_{p-1}^{\alpha+1}(y) L_p^{\alpha}[Y(\mathbf{x})]$$

où $F_{\alpha}(y)$ désigne la fonction de répartition de la loi gamma de paramètre α .

En effet, dans le développement en facteurs, le coefficient d'ordre p est l'espérance de l'indicatrice multipliée par le polynôme de Laguerre de degré p (cf. annexe B) :

$$I(Y(\mathbf{x}); y) = \sum_{p=0}^{+\infty} \varphi_p(\alpha, y) L_p^{\alpha}[Y(\mathbf{x})] \text{ avec } \forall p \in \mathbf{N}, \varphi_p(\alpha, y) = \int_0^{+\infty} I(t; y) L_p^{\alpha}(t) f_{\alpha}(t) dt = \int_0^y L_p^{\alpha}(t) f_{\alpha}(t) dt$$

Pour le terme d'ordre 0, il vient :

$$\varphi_0(\alpha, y) = \int_0^y f_{\alpha}(t) dt = F_{\alpha}(y).$$

Par ailleurs, les polynômes de Laguerre normés sont définis par la formule de Rodrigues (Hu, 1988) :

$$\forall p \in \mathbf{N}, L_p^{\alpha}(y) = \sqrt{\frac{\Gamma(\alpha+p)}{\Gamma(\alpha)\Gamma(p+1)}} \frac{f_{\alpha+p}^{(p)}(y)}{f_{\alpha}(y)} \quad (4.1)$$

d'où

$$\forall p \geq 1, \varphi_p(\alpha, y) = \int_0^y L_p^{\alpha}(t) f_{\alpha}(t) dt = \sqrt{\frac{\Gamma(\alpha+p)}{\Gamma(\alpha)\Gamma(p+1)}} f_{\alpha+p}^{(p-1)}(y) = \sqrt{\frac{\alpha}{p}} f_{\alpha+1}(y) L_{p-1}^{\alpha+1}(y)$$

Par suite, le corrélogramme de l'indicatrice s'exprime en fonction des corrélogrammes $\{C_p^{(y)}(\mathbf{h}), p \in \mathbb{N}^*\}$ des polynômes de Laguerre :

$$r_y(\mathbf{h}) = \sum_{p=1}^{+\infty} \bar{\omega}_p(\alpha, y) C_p^{(y)}(\mathbf{h}) \text{ avec } \bar{\omega}_p(\alpha, y) = \frac{1}{F_\alpha(y)[1-F_\alpha(y)]} \frac{\alpha}{p} [f_{\alpha+1}(y) L_{p-1}^{\alpha+1}(y)]^2 \quad (4.2)$$

Hormis le cas mosaïque, le modèle de Laguerre est caractérisé par une déstructuration des facteurs d'ordre élevé, dont les corrélogrammes se rapprochent d'un effet de pépite. Le coefficient $\bar{\omega}_p(\alpha, y)$ avec $p \in \mathbb{N}^*$ indique la contribution du p-ième corrélogramme dans celui de l'indicatrice associée au seuil y. A titre d'exemple, prenons la valeur $\alpha = 0.25$ et examinons l'évolution des dix premiers coefficients en fonction du seuil (tableau 4.1). Il apparaît que les coefficients d'ordre faible ont une contribution plus importante dans le corrélogramme des indicatrices de seuil élevé que dans ceux des indicatrices de seuil faible ; on doit donc s'attendre à ce que **la déstructuration des teneurs faibles soit plus prononcée que celle des hautes teneurs**.

y	p = 1	p = 2	p = 3	p = 4	p = 5	p = 6	p = 7	p = 8	p = 9	p = 10
10^{-5}	0.0165	0.0103	0.0078	0.0063	0.0054	0.0047	0.0042	0.0038	0.0035	0.0032
10^{-4}	0.0310	0.0194	0.0145	0.0118	0.0100	0.0880	0.0078	0.0071	0.0065	0.0060
10^{-3}	0.0609	0.0380	0.0285	0.0231	0.0196	0.0171	0.0153	0.0138	0.0126	0.0117
10^{-2}	0.1314	0.0808	0.0597	0.0477	0.0399	0.0343	0.0302	0.0269	0.0242	0.0221
10^{-1}	0.3307	0.1749	0.1103	0.0748	0.0526	0.0377	0.0273	0.0198	0.0143	0.0102
1	0.6505	0.0163	0.0182	0.0486	0.0482	0.0317	0.0145	0.0038	0.0001	0.0012

Tableau 4.1. Contributions des corrélogrammes des premiers facteurs dans celui des indicatrices (modèle de Laguerre de paramètre $\alpha = 0.25$)

Pour corroborer cet argument, trois cas sont examinés : le modèle "betifié" (modèle bigamma dans lequel le coefficient de corrélation est randomisé selon une loi beta de paramètres α, ρ et $\alpha(1-\rho)$) ; le modèle obtenu par addition de plusieurs mosaïques indépendantes (cf. chapitre 3) ; le modèle de diffusion bigamma.

Modèle betifié

Les corrélogrammes des facteurs s'expriment en fonction du corrélogramme $\rho(\mathbf{h})$ de la variable gamma (Chilès and Delfiner, 1999) :

$$\forall p \geq 1, C_p^{(y)}(\mathbf{h}) = \frac{\Gamma(\alpha)\Gamma(\alpha\rho(\mathbf{h})+p)}{\Gamma(\alpha\rho(\mathbf{h}))\Gamma(\alpha+p)} \quad (4.3)$$

Connaissant $\rho(\mathbf{h})$, il est donc possible de calculer les corrélogrammes d'indicatrices pour divers seuils, par exemple les quartiles de la distribution marginale. On observe que, lorsque le paramètre α est faible, les teneurs sont d'autant mieux structurées qu'elles sont élevées (figure 4.3A). Ce n'est que lorsque le paramètre α devient très grand que la situation tend à se symétriser : les teneurs médianes sont alors mieux structurées que les teneurs extrêmes (figure 4.3D) ; ceci était prévisible, car le modèle de Laguerre betifié converge vers le modèle bigaussien lorsque α devient infini, donc on retrouve la propriété connue de "déstructuration des hautes teneurs" (Matheron, 1982). Toutefois, les teneurs fortes sont toujours mieux structurées que les faibles.

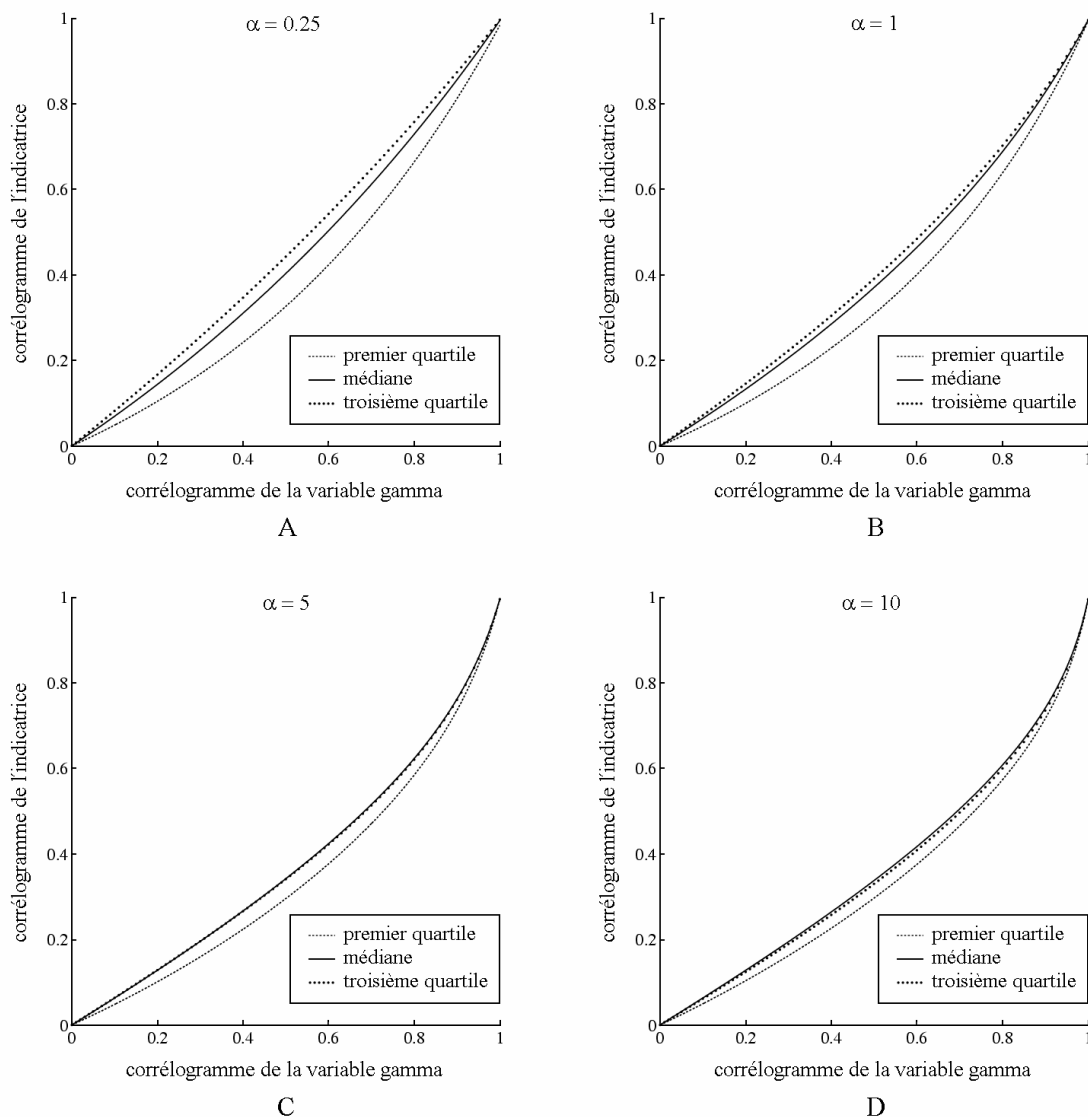


Figure 4.3. La déstructuration des indicatrices dans le modèle de Laguerre betifié, pour quatre valeurs du paramètre α

Addition de mosaïques indépendantes

Dans le cas de l'addition de plusieurs mosaïques, nous comparerons graphiquement les ensembles aléatoires obtenus par seuillage de deux réalisations de la figure 3.17 du chapitre 3 (celles associées à la somme de cinq et cinquante mosaïques de paramètres 0.1 et 0.01 respectivement). On observe que les ensembles associés aux teneurs inférieures au premier quartile sont plus morcelés que ceux associés aux teneurs supérieures au troisième quartile (figure 4.4). Une explication intuitive est que, à cause de la forte dissymétrie de l'histogramme, un site de l'espace appartient à l'ensemble des hautes teneurs dès lors qu'il a une teneur élevée dans l'une des mosaïques, indépendamment des autres mosaïques dont les contributions ont toutes les chances d'être peu importantes. En revanche, les compartiments de faibles teneurs de chaque mosaïque sont fragmentés par les compartiments des autres mosaïques : même si ces derniers n'ont qu'une faible contribution, elle est susceptible de provoquer un changement important de quantile à cause de la forte dissymétrie de la loi gamma.

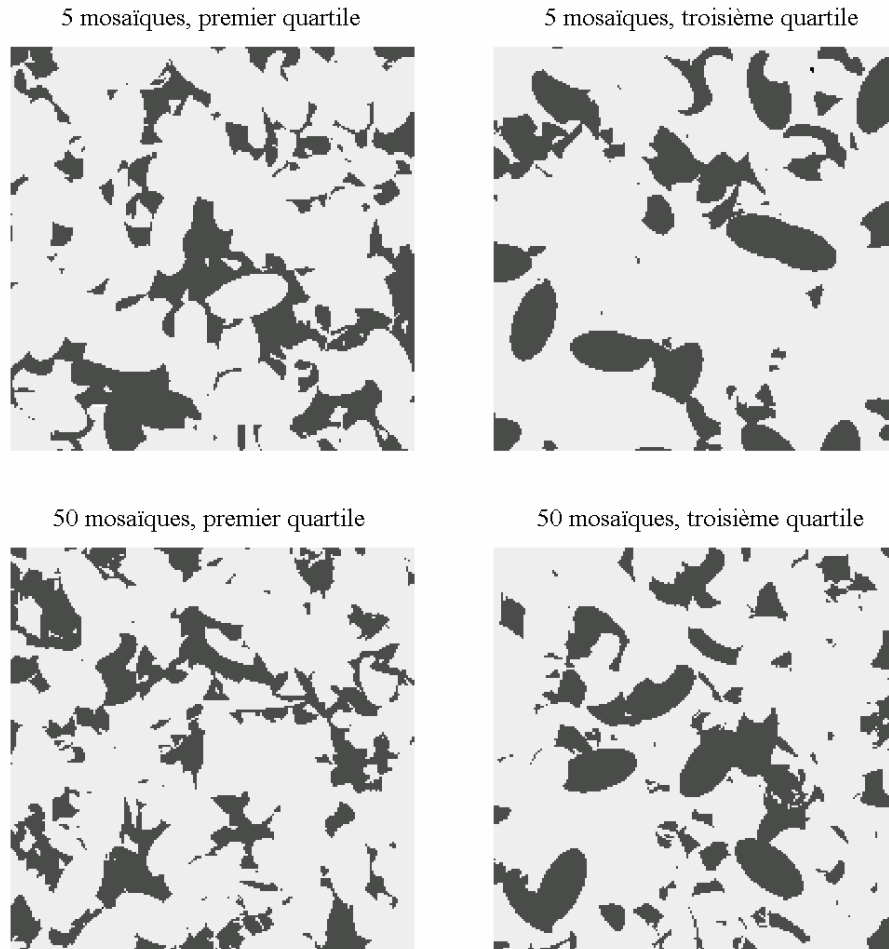


Figure 4.4. Seuillage des réalisations de modèles obtenus par addition de mosaïques

Modèle de diffusion bigamma

Dans le modèle de diffusion, la covariance du p -ième facteur est égale au corrélogramme de $Y(\mathbf{x})$ élevé à la puissance p ; la densité bivariable de coefficient de corrélation ρ s'écrit :

$$f_{\alpha,p}(y, y') = f_{\alpha}(y) f_{\alpha}(y') \sum_{p \geq 0} \rho^p L_p^{\alpha}(y) L_p^{\alpha}(y') \quad (4.4)$$

On établit que le variogramme de l'indicatrice associée au seuil de coupure y s'exprime en fonction du corrélogramme de la variable gamma :

$$\gamma_{1,y}(\mathbf{h}) = \alpha \int_{\rho(\mathbf{h})}^1 f_{\alpha+1,r}(y, y) dr \quad (4.5)$$

et que, en l'absence d'effet de pépite, ce variogramme est proportionnel au madogramme $\gamma_1(\mathbf{h})$ du processus bigamma au voisinage de l'origine :

$$\gamma_{1,y}(\mathbf{h}) \underset{|\mathbf{h}| \rightarrow 0}{\sim} f_{\alpha+1/2}(y) \gamma_1(\mathbf{h}) \quad (4.6)$$

où $f_{\alpha+1/2}$ désigne la densité de probabilité de la loi gamma de paramètre $\alpha + 1/2$.

Démonstration

Dans le cadre bigamma, la formule (4.2) donnant le corrélogramme d'indicatrice s'écrit :

$$r_y(\alpha, \rho) = \sum_{p=1}^{+\infty} \rho^p \bar{\omega}_p(\alpha, y) \quad \text{avec} \quad \bar{\omega}_p(\alpha, y) = \frac{1}{F_\alpha(y)[1-F_\alpha(y)]} \frac{\alpha}{p} [f_{\alpha+1}(y) L_{p-1}^{\alpha+1}(y)]^2$$

En dérivant cette équation par rapport à ρ , il vient :

$$\frac{\partial r_y(\alpha, \rho)}{\partial \rho} = \sum_{p=1}^{+\infty} p \bar{\omega}_p(\alpha, y) \rho^{p-1} = \frac{\alpha}{F_\alpha(y)[1-F_\alpha(y)]} \sum_{p=0}^{+\infty} \rho^p [f_{\alpha+1}(y) L_p^{\alpha+1}(y)]^2.$$

On reconnaît le développement isofactoriel de la densité bigamma de paramètre $\alpha + 1$ [Eq. (4.4)], d'où :

$$\frac{\partial r_y(\alpha, \rho)}{\partial \rho} = \frac{\alpha}{F_\alpha(y)[1-F_\alpha(y)]} f_{\alpha+1,p}(y, y), \quad \text{soit} \quad r_y(\alpha, \rho) = \frac{\alpha}{F_\alpha(y)[1-F_\alpha(y)]} \int_0^\rho f_{\alpha+1,r}(y, y) dr \quad (\text{nul lorsque } \rho = 0).$$

En termes de variogramme, cette égalité s'écrit :

$$\gamma_{1,y}(\mathbf{h}) = \gamma_{1,y}(\infty) [r_y(\alpha, 0) - r_y(\alpha, \rho(\mathbf{h}))] = \alpha \int_{\rho(\mathbf{h})}^1 f_{\alpha+1,r}(y, y) dr$$

A présent, la densité bigamma admet l'expression suivante (Matheron, 1973b ; Hu, 1988) :

$$f_{\alpha+1,p}(y, y') = \frac{1}{\Gamma(\alpha+1)(1-\rho)} \exp\left(-\frac{y+y'}{1-\rho}\right) \left(\frac{yy'}{\rho}\right)^{\alpha/2} I_\alpha\left(2\sqrt{\frac{yy'\rho}{1-\rho}}\right)$$

où I_α est la fonction de Bessel modifiée de première espèce (Matheron, 1965 ; Abramowitz and Stegun, 1972) :

$$I_\alpha(z) = \left(\frac{z}{2}\right)^\alpha \sum_{p=0}^{+\infty} \frac{1}{p! \Gamma(\alpha+p+1)} \left(\frac{z}{2}\right)^{2p} \underset{z \rightarrow +\infty}{\sim} \frac{e^z}{\sqrt{2\pi z}}.$$

Après quelques calculs, on trouve :

$$f_{\alpha+1,p}(y, y) \underset{\rho \rightarrow 1}{\sim} \frac{\exp(-y) y^{\alpha-1/2}}{2\sqrt{\pi} \Gamma(\alpha+1)} \frac{1}{\sqrt{1-\rho}} = \frac{f_{\alpha+1/2}(y) \Gamma(\alpha+1/2)}{2\alpha\sqrt{\pi} \Gamma(\alpha)} \frac{1}{\sqrt{1-\rho}}$$

de sorte que le variogramme de l'indicatrice près de l'origine ($\rho \rightarrow 1$, en supposant qu'il n'y ait pas d'effet de pépité) vérifie

$$\begin{aligned} \gamma_{1,y}(\mathbf{h}) &\underset{|\mathbf{h}| \rightarrow 0}{\sim} \frac{f_{\alpha+1/2}(y) \Gamma(\alpha+1/2)}{\sqrt{\pi} \Gamma(\alpha)} \int_{\rho(\mathbf{h})}^1 \frac{dr}{2\sqrt{1-r}} \\ &= f_{\alpha+1/2}(y) \frac{\Gamma(\alpha+1/2)}{\sqrt{\pi} \Gamma(\alpha)} \sqrt{1-\rho(\mathbf{h})} \end{aligned}$$

Aux courtes distances, le variogramme d'indicatrice est donc proportionnel à la racine carrée du variogramme du processus bigamma, tout comme l'est d'ailleurs son madogramme (Matheron, 1982). Il est intéressant d'exprimer le variogramme d'indicatrice en fonction de ce dernier, ce qui conduit à l'équation (4.6).

Pour étudier la vitesse de déstructuration des indicatrices, il faut considérer leurs variogrammes "standardisés" (normés autour d'un palier unité) :

$$\frac{\gamma_{1,y}(\mathbf{h})}{\gamma_{1,y}(\infty)} \underset{|\mathbf{h}| \rightarrow 0}{\sim} \frac{f_{\alpha+1/2}[F_\alpha^{-1}(u)]}{u(1-u)} \gamma_1(\mathbf{h}) = \xi(\alpha, u) \gamma_1(\mathbf{h})$$

où $u = F_\alpha(y)$ est le quantile (entre 0 et 1) associé au seuil y .

La fonction $\xi(\alpha, u)$ est infinie lorsque u tend vers 0 ou 1, ce qui signifie que la déstructuration des quantiles extrêmes est très prononcée. Toutefois, on observe que la vitesse de déstructuration est plus lente pour les quantiles élevés que pour les quantiles plus faibles (figure 4.5).

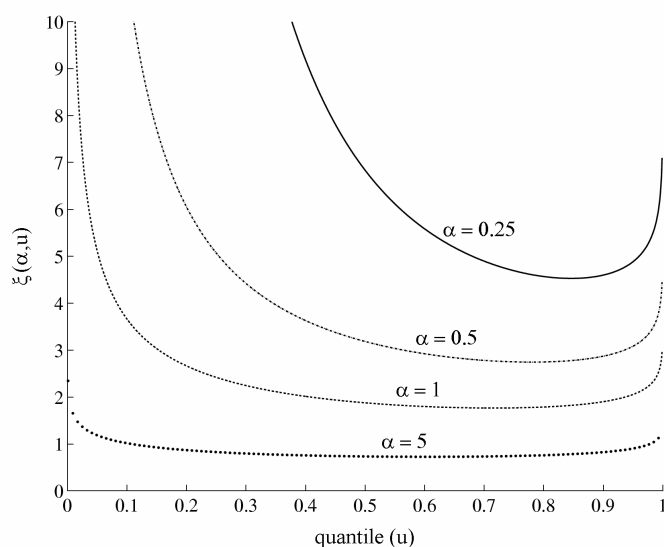


Figure 4.5. Vitesse de déstructuration des indicatrices en fonction du quantile

paramètre de la loi gamma	0.25	0.5	1	5
quantile le mieux structuré	84.6%	78.0%	71.5%	60.6%

Tableau 4.2. Quantiles les mieux structurés en fonction du paramètre de la distribution gamma

Remarques

- 1) Le modèle bigaussien est régi par une formule analogue à l'équation (4.6) (Matheron, 1975b ; Chilès and Delfiner, 1999), ce qui n'est pas étonnant car il s'agit d'un cas limite de modèle bigamma :

$$\gamma_{1,y}(\mathbf{h}) = \int_{\rho(\mathbf{h})}^1 g_r(y, y) dr \underset{|\mathbf{h}| \rightarrow 0}{\sim} \frac{e^{-y^2/2}}{\sqrt{2} \pi} \sqrt{1 - \rho(\mathbf{h})}$$

soit, compte tenu de la proportionnalité entre le madogramme et la racine carrée du variogramme :

$$\gamma_{1,y}(\mathbf{h}) \underset{|\mathbf{h}| \rightarrow 0}{\sim} g(y) \gamma_1(\mathbf{h}).$$

- 2) L'équation (4.6) reste valable dans le modèle de Laguerre dès lors que la variable aléatoire $R(\mathbf{h})$ qui remplace le corrélogramme $\rho(\mathbf{h})$ est presque sûrement égale à 1 lorsque $|\mathbf{h}|$ tend vers 0. Ceci exclut le modèle mosaïque ou le modèle betifié.
- 3) Pour contrôler la dissymétrie de structuration entre les faibles et fortes teneurs, on peut jouer sur le paramètre α (plus il est élevé, plus le modèle se rapproche du cas hermitien) ou sur le caractère diffusif ou mosaïque de la loi bivariable : à la limite, dans le modèle mosaïque, les indicatrices ont toutes le même corrélogramme et l'on a une parfaite symétrie entre les faibles et fortes teneurs.

II.3.2. Stratégie à adopter pour la modélisation des lois bivariées

En pratique, l'ajustement d'un modèle de Laguerre à un ensemble de données se réalise souvent selon les étapes suivantes (Hu, 1988 ; Chilès and Liao, 1993 ; Chilès and Delfiner, 1999) :

- 1) transformation des données initiales en données à loi gamma, dont le paramètre est choisi de sorte à ne pas trop déformer l'histogramme initial;
- 2) analyse variographique des données anamorphosées ;
- 3) spécification du modèle par ajustement du madogramme comme une fonction paramétrique du variogramme.

Le troisième point est discutable. En effet, le madogramme est la somme de tous les variogrammes d'indicatrices (Matheron, 1982) donc ce qui est réellement ajusté est en quelque sorte le comportement "moyen" des indicatrices. En particulier, la dissymétrie de structuration entre valeurs fortes et valeurs faibles échappe à la modélisation. La procédure ne permet même pas de vérifier que les valeurs fortes sont mieux structurées que les valeurs faibles. Dans de nombreuses situations (teneurs en or dans les gisements aurifères, par exemple), c'est plutôt la circonstance contraire qui se produit : les teneurs fortes sont très pépitiées, tandis que les teneurs faibles sont bien structurées dans l'espace.

Le modèle de Laguerre est intéressant pour décrire des variables régionalisées à histogrammes très dissymétriques et dont les valeurs fortes présentent une meilleure structuration spatiale que les valeurs faibles. En revanche, il n'est guère adapté si ces deux conditions ne sont pas réunies et un autre modèle doit alors être recherché (par exemple, celui à résidus d'indicatrices orthogonaux ou un modèle qui ne bénéficie pas de la propriété isofactorielle). Il n'est en effet guère recommandable de construire une anamorphose qui distord fortement l'histogramme des données initiales, car elle risque d'amplifier démesurément les écarts entre des valeurs initialement très voisines et de rendre la modélisation très sensible aux bruits de mesure.

Le paramètre de la loi gamma doit être choisi en fonction de deux critères : d'une part, la parenté de la loi gamma avec la distribution des données initiales ; d'autre part, la capacité du modèle à différencier les propriétés des valeurs fortes et des valeurs faibles et à accorder aux premières une meilleure structuration spatiale qu'aux secondes.

Ainsi, la stratégie proposée pour ajuster un modèle de Laguerre à un ensemble de données est la suivante :

- 1) calcul de quelques variogrammes d'indicatrices, par exemple ceux associés aux quartiles de la distribution globale ;
- 2) choix du paramètre de la loi gamma en fonction de l'histogramme initial et du degré de dissymétrie observé entre la structuration des valeurs fortes et celle des valeurs faibles ;
- 3) anamorphose des données ;
- 4) spécification des paramètres du modèle (analyse variographique, ajustement du madogramme) ;
- 5) validation du modèle proposé (examen des variogrammes d'indicatrices induits par le modèle ou examen des variogrammes d'ordres inférieurs à 2, *cf.* section suivante) ; s'il n'est pas satisfaisant, retour en 2).

II.4. Les variogrammes d'ordres inférieurs à deux

Pour tout scalaire ω compris entre 0 et 2, le variogramme d'ordre ω est défini par (Matheron, 1987) :

$$\gamma_\omega(\mathbf{h}) = \frac{1}{2} E\{|Y(\mathbf{x} + \mathbf{h}) - Y(\mathbf{x})|^\omega\}$$

Le variogramme usuel correspond à $\omega = 2$, le madogramme à $\omega = 1$. Lorsque ω décrit l'intervalle $]0,2]$, l'ensemble des variogrammes d'ordre ω déterminent la loi de l'accroissement $Y(\mathbf{x} + \mathbf{h}) - Y(\mathbf{x})$. Il s'agit donc d'un outil plus complet que le seul madogramme, à défaut de caractériser entièrement la loi du couple $\{Y(\mathbf{x} + \mathbf{h}), Y(\mathbf{x})\}$. Dans les sections suivantes, nous allons exprimer ces variogrammes en fonction du variogramme usuel dans les modèles les plus courants : bigaussien, bigamma, hermitien et laguerrien.

II.4.1. Modèles bigaussien et bigamma

Dans le **modèle bigaussien**, l'accroissement $Y(\mathbf{x} + \mathbf{h}) - Y(\mathbf{x})$ a une loi normale de moyenne nulle et variance égale à deux fois le variogramme :

$$Y(\mathbf{x} + \mathbf{h}) - Y(\mathbf{x}) \stackrel{d}{=} [2\gamma(\mathbf{h})]^{1/2} \mathcal{N}(0,1) \quad (\text{égalité en distribution})$$

où $\mathcal{N}(0,1)$ est une variable gaussienne réduite. Par suite, le variogramme d'ordre ω vaut :

$$\gamma_\omega(\mathbf{h}) = \frac{1}{2} [2\gamma(\mathbf{h})]^{\omega/2} E\{|\mathcal{N}(0,1)|^\omega\} = \frac{1}{2} [2\gamma(\mathbf{h})]^{\omega/2} \frac{1}{\sqrt{\pi}} 2^{\omega/2} \int_0^{+\infty} u^{(\omega-1)/2} e^{-u} du$$

soit :

$$\gamma_\omega(\mathbf{h}) = [\gamma(\mathbf{h})]^{\omega/2} \frac{2^{\omega-1}}{\sqrt{\pi}} \Gamma\left(\frac{\omega+1}{2}\right) \quad (4.7)$$

En particulier, pour tout ω , on doit avoir:

$$\forall \mathbf{h} \in \mathbb{R}^d, \frac{\gamma_\omega(\mathbf{h})}{\gamma_\omega(\mathbf{h}_0)} = \left[\frac{\gamma(\mathbf{h})}{\gamma(\mathbf{h}_0)} \right]^{\omega/2} \quad (4.8)$$

où \mathbf{h}_0 est une distance de référence (en pratique, il s'agit souvent de la portée des variogrammes empiriques). En coordonnées logarithmiques, les points associés aux variogrammes "standardisés" doivent être alignés avec une pente $\omega/2$. Cette propriété peut servir à valider le caractère bigaussien d'un ensemble de données dont l'histogramme est déjà gaussien.

Dans le **modèle bigamma** de paramètre α et corrélogramme $\rho(\mathbf{h})$, l'accroissement d'un couple a pour loi

$$Y(\mathbf{x} + \mathbf{h}) - Y(\mathbf{x}) \stackrel{d}{=} \sqrt{2[1-\rho(\mathbf{h})]\mathcal{G}(\alpha)} \mathcal{N}(0,1)$$

où $\mathcal{N}(0,1)$ est une variable gaussienne réduite

$\mathcal{G}(\alpha)$ est une variable gamma de paramètre α indépendante de $\mathcal{N}(0,1)$.

Démonstration

La transformée de Laplace d'un couple bigamma $\{Y, Y'\}$ de paramètre α et corrélation ρ est (Hu, 1988) :

$$E\{\exp(-\lambda Y - \mu Y')\} = [1 + \lambda + \mu + (1 - \rho)\lambda\mu]^{-\alpha}$$

En remplaçant λ et μ par $-iu$ et iu respectivement, on obtient la transformée de Fourier de l'accroissement $Y - Y'$:

$$E\{\exp[iu(Y - Y')]\} = [1 + (1 - \rho)u^2]^{-\alpha}$$

qui peut être développée comme suit :

$$\begin{aligned} E\{\exp[iu(Y - Y')]\} &= \frac{1}{\Gamma(\alpha)} \int_0^{+\infty} t^{\alpha-1} \exp\{-t[1 + (1 - \rho)u^2]\} dt \\ &= \frac{1}{\Gamma(\alpha)} \int_0^{+\infty} t^{\alpha-1} e^{-t} \exp[-2(1 - \rho)t \frac{u^2}{2}] dt \end{aligned}$$

On reconnaît la transformée de Fourier de $\sqrt{2(1 - \rho)}\mathcal{G}(\alpha) \mathcal{N}(0,1)$, C.Q.F.D.

En conséquence, le variogramme d'ordre ω s'exprime comme suit :

$$\begin{aligned} \gamma_\omega(\mathbf{h}) &= \frac{1}{2} 2^{\omega/2} [1 - \rho(\mathbf{h})]^{\omega/2} E\{\mathcal{G}(\alpha)^{\omega/2}\} E\{|\mathcal{N}(0,1)|^\omega\} \\ &= \frac{1}{2} 2^{\omega/2} [1 - \rho(\mathbf{h})]^{\omega/2} \left\{ \frac{1}{\Gamma(\alpha)} \int_0^{+\infty} t^{\alpha-1+\omega/2} e^{-t} dt \right\} \left\{ \sqrt{\frac{2}{\pi}} \int_0^{+\infty} y^\omega e^{-y^2/2} dy \right\} \\ &= \frac{1}{2} 2^{\omega/2} [1 - \rho(\mathbf{h})]^{\omega/2} \frac{\Gamma(\alpha + \omega/2)}{\Gamma(\alpha)} \frac{1}{\sqrt{\pi}} 2^{\omega/2} \int_0^{+\infty} u^{(\omega-1)/2} e^{-u} du \end{aligned}$$

soit, en tenant compte de la relation $\gamma(\mathbf{h}) = \alpha [1 - \rho(\mathbf{h})]$:

$$\gamma_\omega(\mathbf{h}) = [\gamma(\mathbf{h})]^{\omega/2} \frac{2^{\omega-1}}{\alpha^{\omega/2} \sqrt{\pi}} \Gamma\left(\frac{\omega+1}{2}\right) \frac{\Gamma(\alpha + \omega/2)}{\Gamma(\alpha)} \quad (4.9)$$

Matheron (1982) a obtenu cette même formule dans le cas où $\gamma(\mathbf{h}) = \alpha$ (valeurs non corrélées). La formule (4.8) reste valable et constitue une manière de tester le modèle.

II.4.2. Modèles hermitien et laguerrien

Ces modèles s'obtiennent à partir des modèles bigaussien et bigamma en remplaçant le coefficient de corrélation $\rho(\mathbf{h})$ par une variable aléatoire $R(\mathbf{h})$, à valeurs dans $[-1,1]$ pour le modèle hermitien et $[0,1]$ pour le modèle laguerrien. Les formules (4.7) et (4.9) deviennent :

- modèle hermitien : $\gamma_\omega(\mathbf{h}) = E\{[1 - R(\mathbf{h})]^{\omega/2}\} \frac{2^{\omega-1}}{\sqrt{\pi}} \Gamma\left(\frac{\omega+1}{2}\right)$ (4.10)

- modèle laguerrien : $\gamma_\omega(\mathbf{h}) = E\{[1 - R(\mathbf{h})]^{\omega/2}\} \frac{2^{\omega-1}}{\sqrt{\pi}} \Gamma\left(\frac{\omega+1}{2}\right) \frac{\Gamma(\alpha + \omega/2)}{\Gamma(\alpha)}$ (4.11)

Ces expressions peuvent se développer en fonction des covariances $\{C_p^{(Y)}(\mathbf{h}), p \in \mathbb{N}^*\}$ des facteurs successifs :

$$\frac{\gamma_\omega(\mathbf{h})}{\gamma_\omega(\infty)} = 1 + \sum_{p \geq 1} \frac{\Gamma(p - \omega/2)}{p! \Gamma(-\omega/2)} C_p^{(Y)}(\mathbf{h})$$

Pour pouvoir aller plus loin et expliciter cette formule, il faut adopter un modèle de distribution pour la variable $R(\mathbf{h})$. Un exemple intéressant est celui où cette variable suit une loi beta de paramètres $\{\beta \rho(\mathbf{h}), \beta [1 - \rho(\mathbf{h})]\}$ ($\beta > 0$), de façon que l'espérance de $R(\mathbf{h})$ soit égale au corrélogramme $\rho(\mathbf{h})$. Alors :

$$E\{[1 - R(\mathbf{h})]^{\omega/2}\} = \frac{\Gamma(\beta) \Gamma[\beta - \beta\rho(\mathbf{h}) + \omega/2]}{\Gamma(\beta + \omega/2) \Gamma[\beta - \beta\rho(\mathbf{h})]} \quad (4.12)$$

Les formules (4.10) à (4.12) sont utiles non seulement pour valider le caractère isofactoriel mais aussi pour déterminer la valeur du paramètre β . Les modèles bigaussien et bigamma correspondent au cas où β est infini. Au contraire, le modèle mosaïque s'obtient en faisant tendre β vers 0 ; la variable $R(\mathbf{h})$ est alors égale à 0 ou 1 et tous les variogrammes d'ordre ω sont proportionnels entre eux :

$$E\{[1 - R(\mathbf{h})]^{\omega/2}\} = E\{[1 - R(\mathbf{h})]\} = 1 - \rho(\mathbf{h})$$

Les variogrammes d'ordres faibles (pour fixer les idées, $\omega \leq 1$) sont des outils plus robustes que le variogramme classique, surtout lorsque l'histogramme des données est très dissymétrique³. Ils sont donc une aide précieuse pour faciliter l'inférence statistique des paramètres du modèle et l'analyse variographique : sous réserve d'avoir choisi un type de modèle, ils peuvent servir de substitut au variogramme classique, via les équations (4.7) à (4.12).

III. Lois multivariées

Dans le chapitre précédent où sont proposées plusieurs méthodes de simulation, les lois à plus de deux variables n'ont volontairement pas été explicitées. En fait, l'inférence statistique des modèles isofactoriels est pratiquement toujours limitée aux lois marginales et bivariées ; même dans le cadre multigaussien, la spécification des lois d'ordre supérieur est laissée de côté et l'on se borne à vérifier partiellement l'adéquation des lois bivariées aux données. La raison est qu'une inférence raisonnable des lois multivariées requiert un nombre très élevé de données dans des configurations géométriques se répétant dans l'espace, ce qui est souvent bien trop exigeant. Comme le souligne Matheron (1978b), « nous n'aurons, en général, aucun espoir de tirer de nos données des renseignements crédibles sur les lois à plus de deux variables ».

En limitant l'inférence aux lois marginales et bivariées et en *imposant* les lois d'ordre supérieur (par exemple, en adoptant un modèle de diffusion), certaines caractéristiques du processus aléatoire échappent à la modélisation, comme une éventuelle connexité existant entre les valeurs extrêmes : la définition d'une telle relation met en jeu un nombre infini de sites de l'espace.

³ A ce titre, le variogramme d'ordre $\frac{1}{2}$ ou "rodogramme" a été proposé comme outil par Cressie et Hawkins (1980) pour réaliser l'analyse variographique, en évitant le recours au variogramme expérimental traditionnel. D'autres auteurs suggèrent d'examiner les variogrammes d'ordre 1 et $\frac{1}{2}$ pour compléter le variogramme usuel et détecter des caractéristiques structurales aux grandes distances, telles que la portée ou les anisotropies (Journel, 1987 ; Goovaerts, 1997).

Toutefois, cette limitation aux lois bivariées n'est que partielle, car ces lois contiennent déjà beaucoup d'information. A titre d'exemple, il est relativement difficile d'exhiber des processus à lois bivariées *bigaussiennes* autres que les processus *multigaussiens*, bien qu'il soit facile de construire des contre-exemples parmi les vecteurs aléatoires⁴. Chilès et Delfiner (1999) donnent des exemples de processus binaires à corrélogrammes purement pépitiques mais différents d'un bruit blanc (les valeurs sont indépendantes deux à deux mais non mutuellement indépendantes) ; à partir d'un tel processus à valeurs dans $\{+1,-1\}$, noté $\{\varepsilon(\mathbf{x}), \mathbf{x} \in \mathbb{R}^d\}$, on peut définir une fonction aléatoire bigaussienne à effet de pépite en posant :

$$\forall \mathbf{x} \in \mathbb{R}^d, Y(\mathbf{x}) = \varepsilon(\mathbf{x}) | A(\mathbf{x}) |$$

où $A(\mathbf{x})$ est une fonction aléatoire multigaussienne pépitique (bruit blanc). Bien que les lois marginales et bivariées de $Y(\mathbf{x})$ soient gaussiennes, il n'en est pas de même des lois d'ordre supérieur puisque les valeurs de $\varepsilon(\mathbf{x})$ ne sont pas mutuellement indépendantes.

L'ajustement d'un modèle géostatistique à partir des lois marginales et bivariées réalise un juste compromis entre deux exigences difficiles à concilier :

- la possibilité de réaliser l'inférence statistique et de définir un modèle rigoureux (cohérence "interne") ;
- la capacité du modèle à décrire des caractéristiques complexes de la variable régionalisée (cohérence "externe").

Contrairement à ce qu'exprime Journel (1993), il n'est jamais conseillé de sacrifier la première exigence à la seconde : le modèle est un garde-fou qui évite de faire "n'importe quoi" et dont il faut savoir accepter les limitations. *A contrario*, une technique de simulation qui prétend échapper à tout modèle prédéterminé de fonctions aléatoires, tel l'algorithme séquentiel d'indicateurs, est indéfendable lorsque l'on sait que, sauf de rares exceptions, les propriétés des images obtenues dépendent du niveau de résolution choisi par l'utilisateur ou de la taille du champ (*cf.* chapitre 2).

⁴ Soit $g(y_1, \dots, y_n)$ la densité de probabilité d'un vecteur multigaussien de taille n . Alors la fonction définie par

$$f(y_1, \dots, y_n) = g(y_1, \dots, y_n) [1 + \sin(y_1) \dots \sin(y_n)]$$

est la densité de probabilité d'un vecteur aléatoire dont toutes les lois bivariées sont bigaussiennes : en effet, à cause de leur imparité, les termes en sinus disparaissent dès lors qu'ils sont intégrés.

Chapitre 5 :

Application à des données minières

I. Présentation des données

Le gisement aurifère *El Peñón* est situé dans la deuxième région du Chili (désert d'Atacama), à 160km au sud-est du port d'Antofagasta et à une altitude de 1800m environ. Deux veines d'orientation nord-sud sont en cours d'exploitation : la première, *Quebrada Colorada*, a une faible épaisseur (2 à 5 mètres) et des frontières marquées, perceptibles à l'œil nu depuis une galerie ; la seconde, *Quebrada Orito*, est plus épaisse (5 à 15 mètres environ) mais ses teneurs sont plus faibles et ses frontières avec la matière stérile plus diffuses.

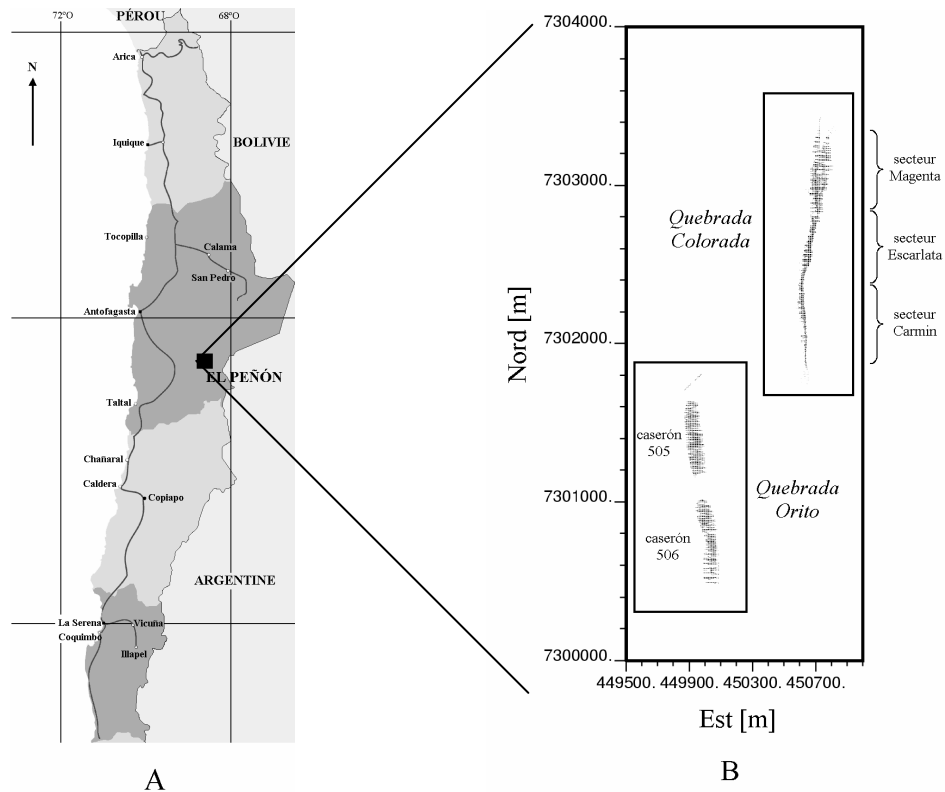


Figure 5.1. A, localisation du gisement et B, implantation des échantillons (coordonnées UTM)

Du point de vue géologique, la minéralisation est un système de veines épithermales de “faible sulfuration” (en anglais, *low sulfidation*) associé à un important système de failles. Le modèle utilisé pour la planification minière considère des unités de sélection de $2\text{m} \times 5\text{m} \times 2\text{m}$, divisées en sous-blocs de 0.5m de côté pour *Quebrada Colorada* et de 1m de côté pour *Quebrada Orito*.

Dans la suite de ce chapitre, nous allons nous intéresser au secteur *Escarlata* de la première veine (*Quebrada Colorada*), qui s’étend approximativement sur 350 mètres le long de l’axe nord-sud et sur 200 mètres le long de la verticale. Deux types de données sont disponibles pour ce secteur :

- 16 069 données d’exploration disposées le long de sondages rectilignes d’orientation variable. Il s’agit d’échantillons composites représentatifs d’une portion de sondage de 0.5 mètre de longueur. Les composites de support inférieur à 0.4 mètre présents dans la base de données initiale ont été éliminés ; ils ne représentaient qu’une faible fraction du total (207 échantillons).
- 5 094 données extraites dans la galerie au moment de l’exploitation (*rainurages*). Une série d’échantillons est prise tous les trois mètres environ sur toute la largeur de la veine, à raison de 0.5 mètre au maximum par échantillon.

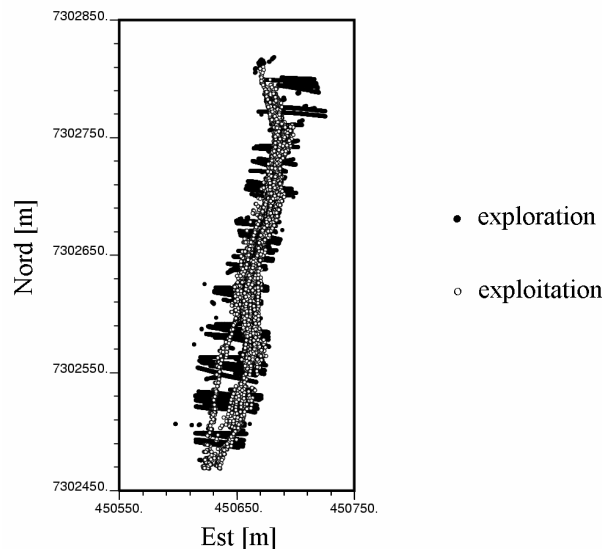


Figure 5.2. Secteur Escarlata (veine *Quebrada Colorada*)

Dans chaque cas, deux variables numériques sont disponibles, à savoir les teneurs des échantillons en or et en argent, chacune exprimée en parties par million (ppm ou g/t). Une troisième variable, codifiant l’appartenance ou non des échantillons à la veine, était originellement contenue dans la base de sondages d’exploration, mais elle ne sera pas prise en compte dans la suite de l’étude pour deux raisons :

- l’appartenance d’un échantillon à la veine n’est pas définie objectivement (il semble qu’une teneur de coupure ait été utilisée comme critère) ;
- on ne dispose d’aucun modèle géologique permettant de décider si un point non échantillonné se trouve à l’intérieur ou à l’extérieur de la veine.

II. Etude exploratoire et choix des données à analyser

Les échantillons étant disposés irrégulièrement dans l'espace, il convient de les pondérer dans le calcul des histogrammes et des statistiques élémentaires. Plusieurs méthodes ont été développées pour réaliser cette pondération ; en général, elles ne prennent en compte que la configuration géométrique des échantillons et laissent de côté leur structuration spatiale. Leur principal intérêt est d'éviter de commettre une erreur radicale au moment de déterminer la distribution marginale des échantillons, ce qui pourrait se produire lorsque l'échantillonnage est préférentiel.

En l'occurrence, les poids des échantillons ont été calculés par la *méthode des cellules* (Deutsch and Journel, 1998) : l'espace est partitionné en cellules parallélépipédiques ; chaque cellule reçoit le même poids, lequel est ensuite réparti équitablement entre les échantillons qu'elle contient. La taille des cellules a été fixée à $1\text{m} \times 3\text{m} \times 5\text{m}$ (ce choix est arbitraire et correspond à une taille "raisonnable" au vu du mode d'échantillonnage).

Les histogrammes obtenus (figure 5.3) montrent une très forte dissymétrie des distributions de teneurs : le maximum est égal à plusieurs dizaines ou centaines de fois la moyenne et le coefficient de variation est très supérieur à 1. On note également un important *effet zéro*, c'est-à-dire une proportion non négligeable de teneurs nulles ou sub-nulles. Un tel effet compromet d'ores et déjà le recours à une anamorphose gaussienne.

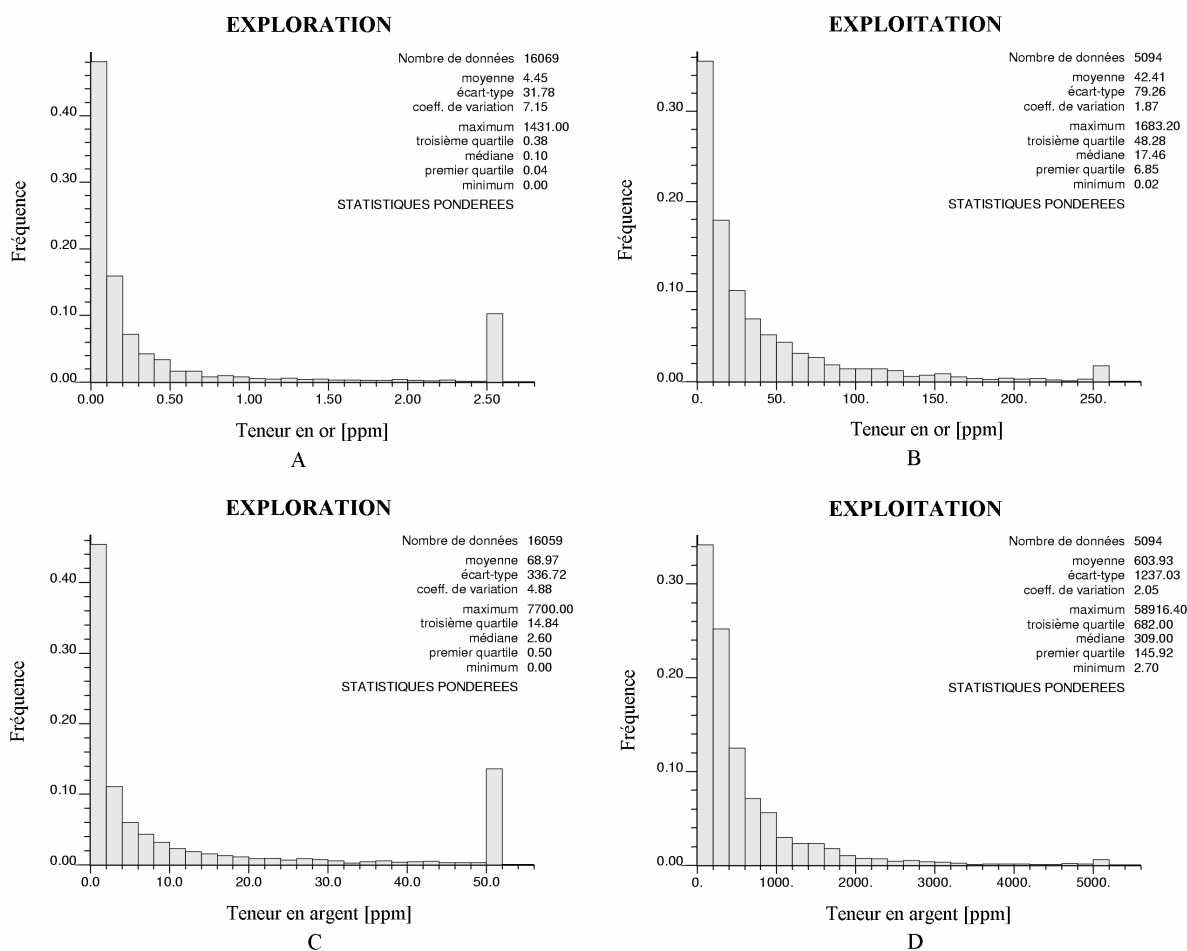


Figure 5.3. Histogrammes empiriques des données (A et B, teneurs en or ; C et D, teneurs en argent)
La dernière barre représente les échantillons dont la teneur dépasse la dernière valeur en abscisse

Une difficulté se pose pour l'analyse géostatistique : au regard des teneurs moyennes et quantiles des histogrammes, les deux séries de données ne représentent manifestement pas la même réalité. En fait, la plupart des échantillons issus des sondages d'exploration correspondent à du matériel stérile situé hors de la veine, sans intérêt économique et aux caractéristiques géologiques fort différentes de la zone minéralisée. Il est donc prudent de séparer leur étude de celles des données d'exploitation. Les données d'exploration posent d'ailleurs un autre problème, lié à la qualité de leur positionnement dans l'espace : les coordonnées fournies sont entachées d'une erreur pouvant atteindre quelques mètres ; s'il s'agit d'un niveau d'erreur acceptable vis-à-vis de la longueur totale des sondages (plusieurs centaines de mètres), cela l'est beaucoup moins en comparaison d'une veine de deux à cinq mètres d'épaisseur...

Pour ces raisons, l'étude géostatistique est désormais restreinte aux seules données d'exploitation. Toutefois, de telles données ne se prêtent guère à une modélisation en trois dimensions : comment calculer un variogramme sur une longueur maximale de quelques mètres (une dizaine d'échantillons au plus) ? Pour contourner cette difficulté, une solution consiste à **passer en accumulations**, c'est-à-dire sommer les teneurs le long de chaque rainurage en les pondérant par leur longueur. L'avantage de cette procédure est triple :

- on se ramène à une étude en deux dimensions ;
- on élimine un problème de stationnarité au sein de la veine car, même si ses frontières sont nettes, les teneurs les plus hautes sont très souvent au centre et les teneurs les plus faibles en bordure ;
- le passage en accumulation a pour effet de régulariser les valeurs et facilitent leur modélisation, surtout du point de vue variographique.

La figure 5.4 visualise les cartes et histogrammes des données d'accumulations. Les cartes sont des projections des centres de gravité des rainurages sur le plan nord / cote ; les valeurs sont repérées par un niveau de gris.

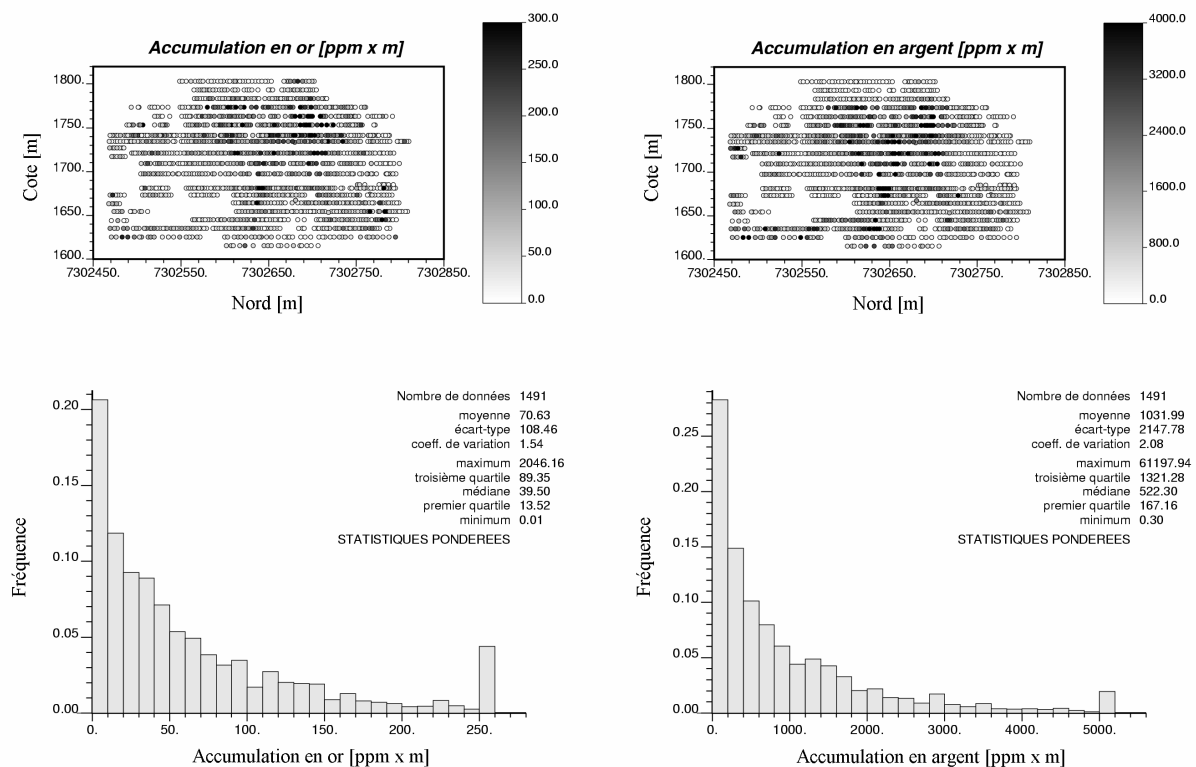


Figure 5.4. Cartes et histogrammes des données d'accumulations

Les accumulations sont mieux adaptées que les teneurs pour décrire les gisements en veine, car leur définition ne dépend pas (ou peu) de la quantité de stérile qui a été échantillonnée en dehors de la veine, correspondant à des teneurs nulles ou sub-nulles. De plus, dans le cas qui nous occupe, la faible épaisseur de la veine (2 à 5 mètres) ne permet pas une exploitation sélective, de sorte que le détail des teneurs le long de cette épaisseur n'a guère d'importance. On aurait pu compléter l'étude en ajoutant la variable puissance, qui mesure l'épaisseur de la veine ; cette dernière a été laissée de côté pour ne pas compliquer le modèle qui comporte déjà deux variables corrélées entre elles (figure 5.5).

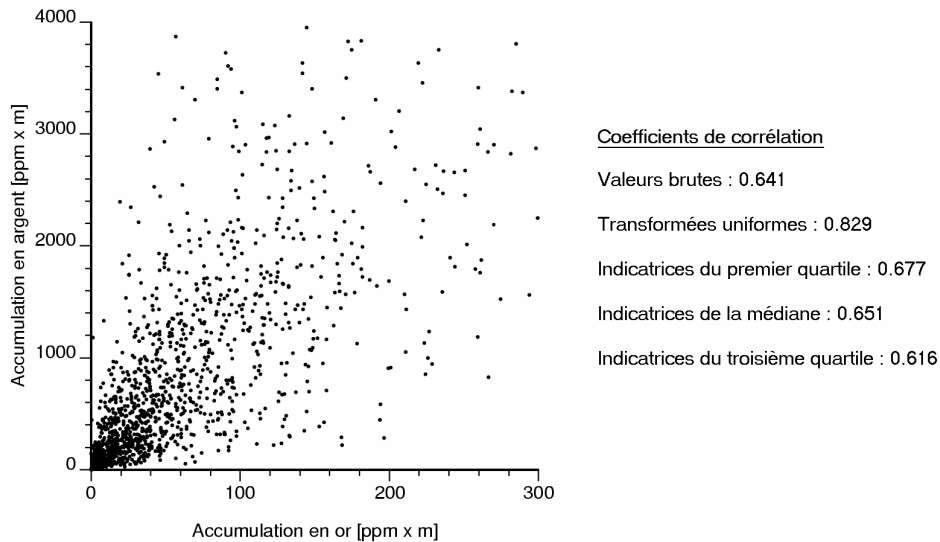


Figure 5.5. Nuage de corrélation entre les accumulations en or et en argent

Bien que l'on se soit formellement ramené à un problème bidimensionnel, la modélisation et la simulation des accumulations s'effectueront dans l'espace tridimensionnel, car la veine est une surface non plane (figure 5.6). Pour leur visualisation, les résultats seront projetés sur le plan nord/cote comme sur la figure 5.4 ou, mieux, sur la surface de la veine une fois "dépliée" (ce qui induit de légères distorsions dans la représentation plane).

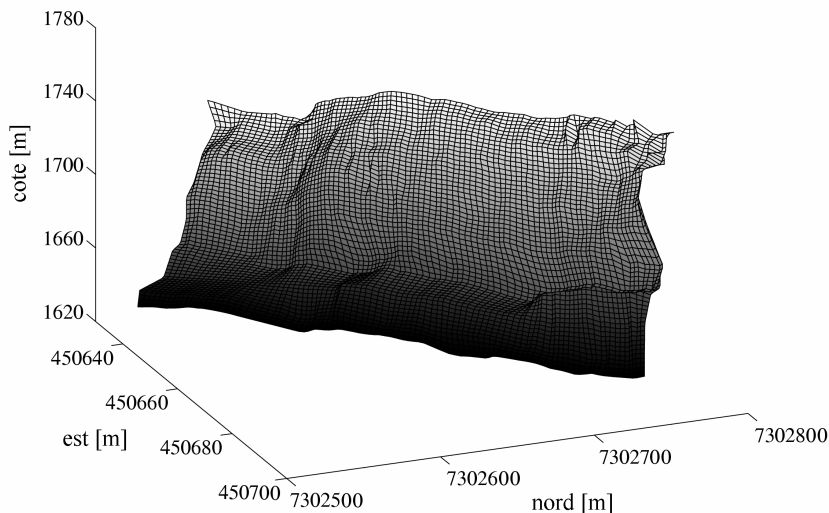


Figure 5.6. Modélisation de la géométrie d'une partie de la veine à partir des centres de gravité des rainurages

III. Choix du modèle

Se pose à présent le choix d'un modèle isofactoriel destiné à décrire les accumulations en or et en argent. Celui-ci doit être recherché parmi les modèles à lois continues.

Un premier candidat est le *modèle à résidus d'indicateurs orthogonaux*. Pour tester l'adéquation de ce modèle aux données, on compare les variogrammes simples et croisés de plusieurs indicateurs, en l'occurrence celles associées aux quartiles de la distribution marginale. Le variogramme croisé entre deux indicateurs est proportionnel au variogramme simple de l'indicateur associé au seuil le plus élevé s'il n'y a pas d'effets de bord en descendant, ou à celui de l'indicateur associé au seuil le plus faible s'il n'y a pas d'effets de bord en montant (Rivoirard, 1988, 1989). Dans le cas qui nous occupe, le test appliqué aux variogrammes omnidirectionnels se révèle infructueux, aussi bien pour les accumulations en or (figures 5.7A et 5.7B) que pour celles en argent (figures 5.7C et 5.7D) : aussi le modèle à résidus d'indicateurs sera-t-il écarté.

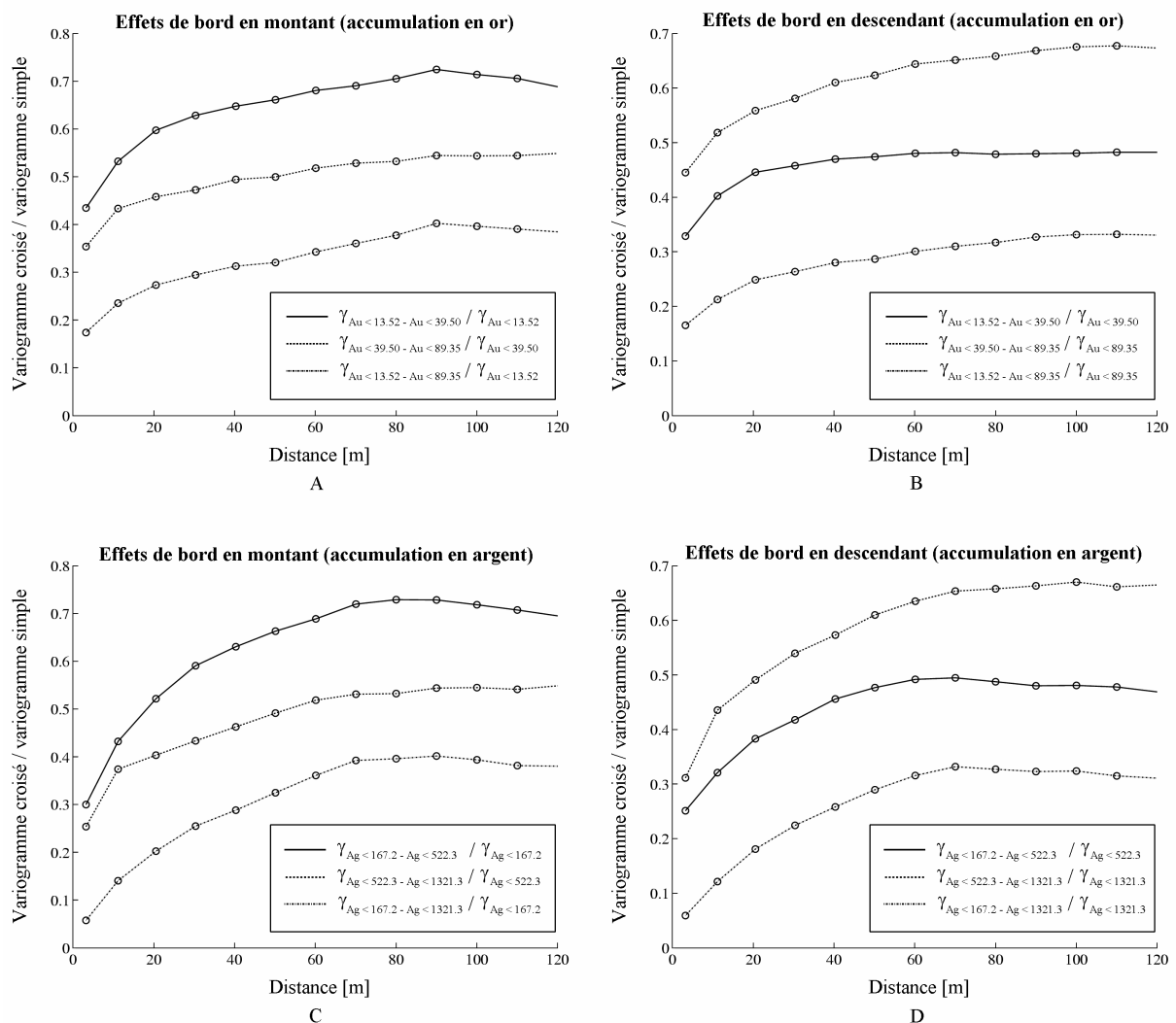


Figure 5.7. Test du modèle à résidus d'indicateurs orthogonaux : en l'absence d'effets de bord, le rapport entre les variogrammes simples et croisés d'indicateurs doit être constant

On s'oriente vers les modèles isofactoriels à loi marginale indéfiniment divisible obtenus par addition de mosaïques indépendantes, car ces modèles permettent de contrôler l'aspect plus ou moins diffusif de la régionalisation. Parmi les lois continues susceptibles de décrire les accumulations, le choix se porte sur la loi gamma (*modèle de Laguerre*).

III.1. Anamorphose

Il est tout d'abord nécessaire de transformer les valeurs initiales en données dont la distribution marginale est une loi gamma. Or, cette loi dépend d'un paramètre scalaire α qui contrôle son degré de dissymétrie. Plusieurs critères ont été proposés pour déterminer ce paramètre (Demange *et al.*, 1987 ; Hu, 1988), l'idée étant généralement de ne pas trop "déformer" l'histogramme des valeurs initiales.

Le point-clé est toutefois la modélisation des lois bivariées (par exemple, le paramètre choisi influe sur le degré de dissymétrie entre la structuration des faibles valeurs et celle des fortes valeurs), de sorte qu'il est difficile d'avancer avec certitude la meilleure valeur de α . On choisit ici un paramètre égal à 0.5, aussi bien pour les accumulations en or que pour celles en argent ; au besoin, ce choix sera revu ultérieurement.

III.2. Un modèle bivarié

Nous allons étendre les modèles présentés au chapitre 3 afin de décrire le comportement conjoint des accumulations en or et en argent. Pour cela, on crée une dimension supplémentaire en affectant à chaque variable un indice u (par exemple, 1 pour l'or et 2 pour l'argent), ce qui ramène le problème à une modélisation sur $\mathbb{R}^d \times \{1,2\}$. La démarche proposée ci-dessous peut aisément être généralisée à plus de deux variables.

Construction d'une tessellation sur $\mathbb{R}^d \times \{1,2\}$

Une première idée consiste à construire une mosaïque des feuilles mortes dont les grains primaires sont des cylindres de hauteur L (figure 5.8). Sur chacun des hyperplans $\mathbb{R}^d \times \{1\}$ et $\mathbb{R}^d \times \{2\}$, ce modèle induit une mosaïque des feuilles mortes dont le covariogramme géométrique est celui de la base des grains primaires cylindriques. Les deux mosaïques sont corrélées, puisque certains grains primaires sont communs aux deux hyperplans tandis que d'autres n'en rencontrent qu'un seul. Cependant, elles possèdent le même corrélogramme, ce qui est assez restrictif.

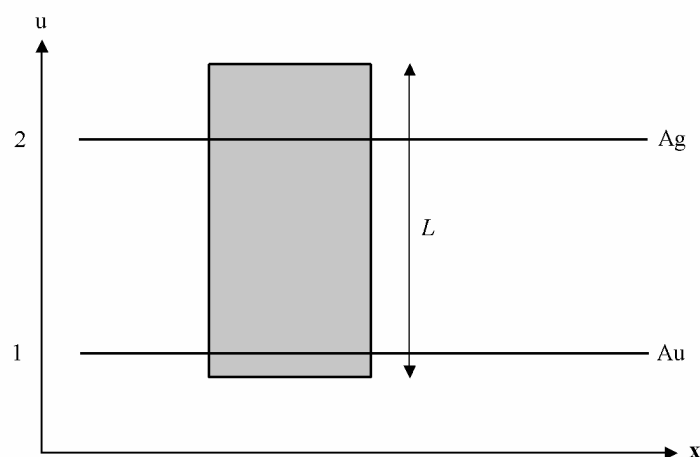


Figure 5.8. Grain primaire cylindrique sur $\mathbb{R}^d \times \mathbb{R}$

Pour pallier cette difficulté, une construction plus générale est envisagée, sur le principe du modèle des feuilles mortes colorées. On considère une famille de grains primaires cylindriques $\{A_i, i = 1 \dots I\}$ tels que :

- leur hauteur est $L = 1$;
- leurs bases ont pour covariogrammes géométriques $\{K_i(\mathbf{h}), i = 1 \dots I\}$ sur \mathbb{R}^d ;
- leurs germes ont une densité régionalisée $\{\theta_i(\mathbf{x}, u), i = 1 \dots I\}$ nulle hors des hyperplans $\mathbb{R}^d \times \{1\}$, $\mathbb{R}^d \times \{3/2\}$ et $\mathbb{R}^d \times \{2\}$, constante sur ces hyperplans mais différente d'un hyperplan à l'autre. Dans la suite, on écrira $\theta_i(u)$ au lieu de $\theta_i(\mathbf{x}, u)$ et l'on posera $\theta_i = \theta_i(1) + \theta_i(3/2) + \theta_i(2)$.

Entre les instants t et $t + dt$, on implante I schémas booléens indépendants, de grains primaires $\{A_i, i = 1 \dots I\}$ et de densités $\{\theta_i(u) dt, i = 1 \dots I\}$. Les grains primaires nouveaux recouvrent les grains déjà présents à l'instant t . Au bout d'un temps infiniment long, les frontières des grains primaires apparents dessinent une tessellation sur $\mathbb{R}^d \times \{1, 2\}$.

Les probabilités d'appartenance de deux sites distincts à une même cellule de la mosaïque (que l'on peut symboliser par la relation d'équivalence \equiv) valent :

$$\text{Prob}[(\mathbf{x} + \mathbf{h}, 1) \equiv (\mathbf{x}, 1)] = \rho_1(\mathbf{h}) = \frac{\sum_{i=1}^I [\theta_i(1) + \theta_i(3/2)] K_i(\mathbf{h})}{\sum_{i=1}^I [\theta_i(1) + \theta_i(3/2)] [2 K_i(\mathbf{0}) - K_i(\mathbf{h})]}$$

$$\text{Prob}[(\mathbf{x} + \mathbf{h}, 2) \equiv (\mathbf{x}, 2)] = \rho_2(\mathbf{h}) = \frac{\sum_{i=1}^I [\theta_i(3/2) + \theta_i(2)] K_i(\mathbf{h})}{\sum_{i=1}^I [\theta_i(3/2) + \theta_i(2)] [2 K_i(\mathbf{0}) - K_i(\mathbf{h})]}$$

$$\text{Prob}[(\mathbf{x} + \mathbf{h}, 1) \equiv (\mathbf{x}, 2)] = \rho_{12}(\mathbf{h}) = \frac{\sum_{i=1}^I \theta_i(3/2) K_i(\mathbf{h})}{\sum_{i=1}^I \theta_i [2 K_i(\mathbf{0}) - K_i(\mathbf{h})]}$$

Les deux premières équations expriment que, sur les hyperplans $\mathbb{R}^d \times \{1\}$ et $\mathbb{R}^d \times \{2\}$, le modèle proposé induit des mosaïques des feuilles mortes stationnaires de covariogrammes géométriques

$$K_{Au}(\mathbf{h}) = \sum_{i=1}^I a_i K_i(\mathbf{h}) \quad (5.1)$$

$$K_{Ag}(\mathbf{h}) = \sum_{i=1}^I b_i K_i(\mathbf{h}) \quad (5.2)$$

$$\text{avec } a_i = \frac{\theta_i(1) + \theta_i(3/2)}{\sum_{j=1}^I [\theta_j(1) + \theta_j(3/2)]}$$

$$b_i = \frac{\theta_i(3/2) + \theta_i(2)}{\sum_{j=1}^I [\theta_j(3/2) + \theta_j(2)]}.$$

Démonstration des formules

Pour qu'à un instant T les deux sites $(\mathbf{x} + \mathbf{h}, 1)$ et $(\mathbf{x}, 2)$ se trouvent dans une même cellule de la mosaïque, il faut qu'ils aient été recouverts à un instant t par un grain germé sur l'hyperplan $\mathbf{R}^d \times \{3/2\}$ et non recouverts entre t et T . Ceci se produit avec la probabilité :

$$\begin{aligned} \text{Prob}[(\mathbf{x} + \mathbf{h}, 1) \equiv (\mathbf{x}, 2)] &= \sum_{i=1}^I \int_0^T \underbrace{K_i(\mathbf{h}) \theta_i(3/2) dt}_{\substack{(\mathbf{x} + \mathbf{h}, 1) \text{ et } (\mathbf{x}, 2) \\ \text{recouverts à l'instant} \\ t \text{ par un grain } A_i}} \prod_{j=1}^I \underbrace{\exp \left\{ - \int_t^T \theta_j [2K_j(\mathbf{0}) - K_j(\mathbf{h})] du \right\}}_{\substack{(\mathbf{x} + \mathbf{h}, 1) \text{ et } (\mathbf{x}, 2) \text{ non recouverts} \\ \text{entre } t \text{ et } T \text{ par un grain } A_j}} \\ &= \sum_{i=1}^I \frac{K_i(\mathbf{h}) \theta_i(3/2)}{\sum_{j=1}^I \theta_j [2K_j(\mathbf{0}) - K_j(\mathbf{h})]} \left\{ 1 - \exp \left\{ - T \sum_{j=1}^I \theta_j [2K_j(\mathbf{0}) - K_j(\mathbf{h})] \right\} \right\} \end{aligned}$$

De même, les sites $(\mathbf{x} + \mathbf{h}, 1)$ et $(\mathbf{x}, 1)$ se trouvent dans une même cellule de la mosaïque à l'instant T s'ils ont été couverts à un instant t par un grain germé sur l'un des hyperplans $\mathbf{R}^d \times \{3/2\}$ ou $\mathbf{R}^d \times \{1\}$ et non recouverts entre t et T :

$$\begin{aligned} \text{Prob}[(\mathbf{x} + \mathbf{h}, 1) \equiv (\mathbf{x}, 1)] &= \sum_{i=1}^I \int_0^T K_i(\mathbf{h}) [\theta_i(1) + \theta_i(3/2)] dt \prod_{j=1}^I \exp \left\{ - \int_t^T [\theta_j(1) + \theta_j(3/2)] [2K_j(\mathbf{0}) - K_j(\mathbf{h})] du \right\} \\ &= \frac{\sum_{i=1}^I K_i(\mathbf{h}) [\theta_i(1) + \theta_i(3/2)]}{\sum_{j=1}^I [\theta_j(1) + \theta_j(3/2)] [2K_j(\mathbf{0}) - K_j(\mathbf{h})]} \left\{ 1 - \exp \left\{ - T \sum_{j=1}^I [\theta_j(1) + \theta_j(3/2)] [2K_j(\mathbf{0}) - K_j(\mathbf{h})] \right\} \right\} \end{aligned}$$

En changeant l'indice u de 1 à 2, il vient :

$$\text{Prob}[(\mathbf{x} + \mathbf{h}, 2) \equiv (\mathbf{x}, 2)] = \frac{\sum_{i=1}^I K_i(\mathbf{h}) [\theta_i(3/2) + \theta_i(2)]}{\sum_{j=1}^I [\theta_j(3/2) + \theta_j(2)] [2K_j(\mathbf{0}) - K_j(\mathbf{h})]} \left\{ 1 - \exp \left\{ - T \sum_{j=1}^I [\theta_j(3/2) + \theta_j(2)] [2K_j(\mathbf{0}) - K_j(\mathbf{h})] \right\} \right\}$$

Les probabilités annoncées s'obtiennent en faisant tendre T vers l'infini.

On obtient une image approchée de la tessellation en arrêtant la simulation à un instant T suffisamment grand. Une image exacte est obtenue en changeant de point de vue dans la construction du modèle des feuilles mortes. En effet, si l'on pouvait commencer la simulation à partir de l'instant $t = -\infty$, on obtiendrait à $t = 0$ un processus dont les propriétés statistiques ont atteint l'équilibre (cela revient à translater l'échelle de temps conventionnelle, qui part de $t = 0$ et tend vers $t = \infty$). Maintenant, si l'on inverse cette échelle de temps, ce modèle équivaut à partir de $t = 0$ et remonter dans le temps, en changeant l'opérateur "dernier grain observé" en "premier grain observé". La différence fondamentale est que le processus n'est plus altéré dès lors que le domaine à simuler est entièrement recouvert ; à partir de cet instant, on obtient donc une image parfaite du processus asymptotique.

Valuation des cellules

Une fois construite la tessellation sur $\mathbf{R}^d \times \{1, 2\}$, deux valuations sont attribuées à chaque cellule, l'une concernant l'accumulation en or, l'autre l'accumulation en argent. La raison pour laquelle on ne considère pas la même valuation pour les deux variables est que leur corrélation n'est que partielle (et accessoirement, leurs lois marginales pourraient être différentes). On pose donc, pour tout site \mathbf{x} de \mathbf{R}^d :

$$\begin{cases} Y(\mathbf{x}, 1) = U + V \\ Y(\mathbf{x}, 2) = U + W \end{cases}$$

où U , V et W sont trois variables aléatoires gamma indépendantes, de paramètres respectifs α , $\alpha_1 - \alpha$ et $\alpha_2 - \alpha$, constantes sur chaque cellule et indépendantes d'une cellule à l'autre.

Les covariances simples et croisées des deux mosaïques sont finalement :

$$\begin{aligned} \text{cov}[Y(\mathbf{x} + \mathbf{h}, 1), Y(\mathbf{x}, 1)] &= \alpha_1 \rho_1(\mathbf{h}) \\ \text{cov}[Y(\mathbf{x} + \mathbf{h}, 2), Y(\mathbf{x}, 2)] &= \alpha_2 \rho_2(\mathbf{h}) \\ \text{cov}[Y(\mathbf{x} + \mathbf{h}, 1), Y(\mathbf{x}, 2)] &= \text{cov}[Y(\mathbf{x} + \mathbf{h}, 2), Y(\mathbf{x}, 1)] = \alpha \rho_{12}(\mathbf{h}) \end{aligned} \quad (5.3)$$

Addition de plusieurs mosaïques

Additionnons N mosaïques indépendantes telles que définies ci-dessus :

$$\begin{cases} Y_{Au}(\mathbf{x}) = \sum_{n=1}^N Y_n(\mathbf{x}, 1) \\ Y_{Ag}(\mathbf{x}) = \sum_{n=1}^N Y_n(\mathbf{x}, 2) \end{cases}$$

Pour chacune des variables (accumulations en or et en argent), on obtient un modèle de Laguerre défini par une somme de mosaïques à valuations gamma. Les lois marginales sont des lois gamma de paramètres respectifs $N \alpha_1$ et $N \alpha_2$ et les corrélogrammes sont $\rho_1(\mathbf{h})$ et $\rho_2(\mathbf{h})$. Les covariances des facteurs sont (cf. chapitre 3, équation (3.9)) :

$$\forall \mathbf{h} \neq \mathbf{0}, \forall u \in \{1, 2\}, C_p^{(u)}(\mathbf{h}) = \frac{\Gamma(N\alpha_u)[1-\rho_u(\mathbf{h})]^N}{\Gamma(N\alpha_u + p)} \sum_{k=1}^N C_N^k \left(\frac{\rho_u(\mathbf{h})}{1-\rho_u(\mathbf{h})} \right)^k \frac{\Gamma(\alpha_u k + p)}{\Gamma(\alpha_u k)}$$

Voyons à présent la loi conjointe d'un couple de valeurs $\{Y_{Au}(\mathbf{x} + \mathbf{h}), Y_{Ag}(\mathbf{x})\}$. Soit K le nombre de fois où les sites $(\mathbf{x} + \mathbf{h}, 1)$ et $(\mathbf{x}, 2)$ appartiennent à la même cellule de mosaïque ; il s'agit d'une variable binomiale de taille N et paramètre $\rho_{12}(\mathbf{h})$ en tant que somme de N alternatives de Bernoulli indépendantes de probabilité de succès $\rho_{12}(\mathbf{h})$. Conditionnellement à $K = k$, le couple $\{Y_{Au}(\mathbf{x} + \mathbf{h}), Y_{Ag}(\mathbf{x})\}$ s'écrit sous la forme :

$$\begin{cases} Y_{Au}(\mathbf{x} + \mathbf{h}) = U + V \\ Y_{Ag}(\mathbf{x}) = U + W \end{cases} \quad (5.4)$$

où U , V et W sont trois variables aléatoires gamma indépendantes, de paramètres αk , $\alpha_1 N - \alpha k$ et $\alpha_2 N - \alpha k$ respectivement.

La relation (5.4) établit que le couple $\{Y_{Au}(\mathbf{x} + \mathbf{h}), Y_{Ag}(\mathbf{x})\}$ suit, conditionnellement à $K = k$, une loi de Laguerre dissymétrique (Hu, 1988) telle que les covariances croisées des facteurs d'ordre p sont

$$\forall \mathbf{h} \neq \mathbf{0}, (C_p^{(12)}(\mathbf{h}) | K = k) = \sqrt{\frac{\Gamma(N\alpha_1 + p)\Gamma(N\alpha_2)}{\Gamma(N\alpha_2 + p)\Gamma(N\alpha_1)}} \frac{\Gamma(N\alpha_1)\Gamma(\alpha k + p)}{\Gamma(\alpha k)\Gamma(N\alpha_1 + p)}$$

En déconditionnant par rapport à K , on obtient les covariances croisées *a priori* des facteurs :

$$\begin{aligned} \forall \mathbf{h} \neq \mathbf{0}, C_p^{(12)}(\mathbf{h}) &= \sqrt{\frac{\Gamma(N\alpha_1 + p)\Gamma(N\alpha_2)}{\Gamma(N\alpha_2 + p)\Gamma(N\alpha_1)}} \mathbb{E} \left\{ \frac{\Gamma(N\alpha_1)\Gamma(\alpha K + p)}{\Gamma(\alpha K)\Gamma(N\alpha_1 + p)} \right\} \\ &= \sqrt{\frac{\Gamma(N\alpha_2)\Gamma(N\alpha_1)}{\Gamma(N\alpha_2 + p)\Gamma(N\alpha_1 + p)}} [1-\rho_{12}(\mathbf{h})]^N \sum_{k=1}^N C_N^k \left(\frac{\rho_{12}(\mathbf{h})}{1-\rho_{12}(\mathbf{h})} \right)^k \frac{\Gamma(\alpha k + p)}{\Gamma(\alpha k)} \end{aligned} \quad (5.5)$$

Modèle final

On peut compléter le modèle en ajoutant, pour l'accumulation en or, N_1 mosaïques monovariabiles de paramètre α_1 et covariogramme $K_{Au}(\mathbf{h})$ [Eq. (5.1)], et pour l'accumulation en argent, N_2 mosaïques monovariabiles de paramètre α_2 et covariogramme $K_{Ag}(\mathbf{h})$ [Eq. (5.2)]. Il est aussi possible d'ajouter des composantes purement pépitiques $\{Y_0(\mathbf{x}), Y_0'(\mathbf{x}), Y_0''(\mathbf{x})\}$ à lois gamma de paramètres respectifs α_0, α_0' et α_0'' :

$$\begin{cases} Y_{Au}(\mathbf{x}) = \sum_{n=1}^N Y_n(\mathbf{x},1) + \sum_{n=1}^{N_1} Y_n'(\mathbf{x}) + Y_0(\mathbf{x}) + Y_0'(\mathbf{x}) \\ Y_{Ag}(\mathbf{x}) = \sum_{n=1}^N Y_n(\mathbf{x},2) + \sum_{n=1}^{N_2} Y_n''(\mathbf{x}) + Y_0(\mathbf{x}) + Y_0''(\mathbf{x}) \end{cases}$$

Lois directes

Les accumulations en or et en argent sont distribuées selon des lois gamma de paramètres

$$\begin{aligned} \alpha_{Au} &= (N + N_1) \alpha_1 + \alpha_0 + \alpha_0' \\ \alpha_{Ag} &= (N + N_2) \alpha_2 + \alpha_0 + \alpha_0'' \end{aligned}$$

Leurs covariances simples sont :

$$\begin{aligned} C_{Au}(\mathbf{h}) &= (N + N_1) \alpha_1 \rho_1(\mathbf{h}) + (\alpha_0 + \alpha_0') \delta(\mathbf{h}) \\ C_{Ag}(\mathbf{h}) &= (N + N_2) \alpha_2 \rho_2(\mathbf{h}) + (\alpha_0 + \alpha_0'') \delta(\mathbf{h}) \end{aligned} \quad (5.6)$$

Quant aux covariances des facteurs, elles s'expriment comme suit (cf. chapitre 3, équation (3.10)) :

$$\begin{aligned} \forall \mathbf{h} \neq \mathbf{0}, C_p^{(Au)}(\mathbf{h}) &= \frac{\Gamma(\alpha_{Au}) [1 - \rho_1(\mathbf{h})]^N}{\Gamma(\alpha_{Au} + p)} \sum_{k=1}^N C_N^k \left(\frac{\rho_1(\mathbf{h})}{1 - \rho_1(\mathbf{h})} \right)^k \frac{\Gamma[(\alpha_{Au} - \alpha_0 - \alpha_0') k / N + p]}{\Gamma[(\alpha_{Au} - \alpha_0 - \alpha_0') k / N]} \\ \forall \mathbf{h} \neq \mathbf{0}, C_p^{(Ag)}(\mathbf{h}) &= \frac{\Gamma(\alpha_{Ag}) [1 - \rho_2(\mathbf{h})]^N}{\Gamma(\alpha_{Ag} + p)} \sum_{k=1}^N C_N^k \left(\frac{\rho_2(\mathbf{h})}{1 - \rho_2(\mathbf{h})} \right)^k \frac{\Gamma[(\alpha_{Ag} - \alpha_0 - \alpha_0'') k / N + p]}{\Gamma[(\alpha_{Ag} - \alpha_0 - \alpha_0'') k / N]} \end{aligned} \quad (5.7)$$

Loi croisée

Dans l'équation (5.4), si $\mathbf{h} \neq \mathbf{0}$, le paramètre de la variable U est inchangé, par contre les variables V et W ont cette fois pour paramètres $\alpha_{Au} - \alpha k$ et $\alpha_{Ag} - \alpha k$. Les covariances croisées des facteurs [Eq. (5.5)] deviennent :

$$\forall \mathbf{h} \neq \mathbf{0}, C_p^{(12)}(\mathbf{h}) = \sqrt{\frac{\Gamma(\alpha_{Ag}) \Gamma(\alpha_{Au})}{\Gamma(\alpha_{Ag} + p) \Gamma(\alpha_{Au} + p)}} [1 - \rho_{12}(\mathbf{h})]^N \sum_{k=1}^N C_N^k \left(\frac{\rho_{12}(\mathbf{h})}{1 - \rho_{12}(\mathbf{h})} \right)^k \frac{\Gamma(\alpha k + p)}{\Gamma(\alpha k)} \quad (5.8)$$

Pour $\mathbf{h} = \mathbf{0}$, le paramètre de la variable U est $\alpha k + \alpha_0$, tandis que ceux des variables V et W sont $\alpha_{Au} - \alpha k - \alpha_0$ et $\alpha_{Ag} - \alpha k - \alpha_0$. Les covariances croisées des facteurs pour cette distance sont donc :

$$C_p^{(12)}(\mathbf{0}) = \sqrt{\frac{\Gamma(\alpha_{Ag}) \Gamma(\alpha_{Au})}{\Gamma(\alpha_{Ag} + p) \Gamma(\alpha_{Au} + p)}} [1 - \rho_{12}(\mathbf{0})]^N \sum_{k=1}^N C_N^k \left(\frac{\rho_{12}(\mathbf{0})}{1 - \rho_{12}(\mathbf{0})} \right)^k \frac{\Gamma(\alpha k + \alpha_0 + p)}{\Gamma(\alpha k + \alpha_0)}$$

La covariance croisée entre les accumulations en or et en argent s'écrit :

$$C_{\text{Au-Ag}}(\mathbf{h}) = N \alpha \rho_{12}(\mathbf{h}) + \alpha_0 \delta(\mathbf{h}) \quad (5.9)$$

Bilan

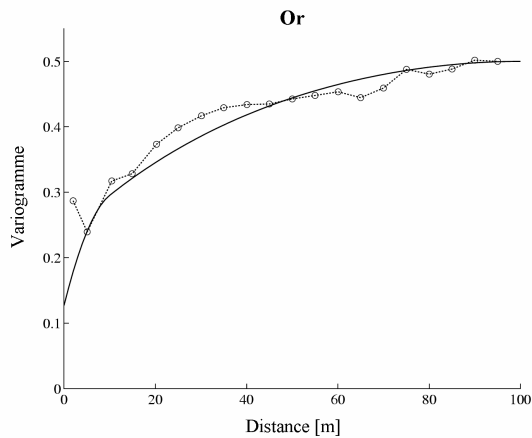
Nous disposons en fin de compte d'un modèle très riche pour décrire le comportement conjoint des accumulations en or et en argent. Résumons le rôle de chacun des paramètres :

- les scalaires α_0 , α_0' et α_0'' servent à modéliser l'effet de pépite des covariances simples et croisée des accumulations [Eq. (5.6) et (5.9)] ;
- les covariogrammes géométriques $\{K_i(\mathbf{h}), i = 1 \dots I\}$ et les densités régionalisées $\{\theta_i(u), i = 1 \dots I\}$ servent à modéliser la composante continue de ces covariances simples et croisée. Ces paramètres jouent en quelque sorte le même rôle que les schémas gigognes et leurs paliers dans le modèle linéaire de corégionalisation (Wackernagel, 1998) ;
- les nombres N_1 , N_2 et N servent à contrôler le caractère plus ou moins diffusif des lois simples et croisées ; ils permettent une modélisation de ces lois au-delà des covariances, par exemple via les covariances d'indicatrices qui se déduisent des équations (5.7) et (5.8) (revoir à ce titre la formule (4.2) du chapitre 4) ;
- les scalaires α_1 et α_2 régulent les variances *a priori* des accumulations en or et en argent. Le scalaire α influe sur leur coefficient de corrélation ponctuel [Eq. (5.9)].

III.3. Mise en oeuvre pratique

L'inférence des paramètres et la validation du modèle va se baser sur les variogrammes simples et croisés des accumulations anamorphosées ainsi que sur ceux des indicatrices associées aux quartiles de la distribution globale des accumulations. La raison pour laquelle les autres outils structuraux (comme le madogramme ou les variogrammes d'ordre inférieur à deux, *cf.* chapitre 4) ne sont pas considérés est qu'ils n'ont pas d'extension au cadre bivariable (une telle extension constituerait un intéressant sujet de recherche) ; au contraire, les variogrammes croisés d'indicatrices permettent une description de la loi conjointe des accumulations en or et en argent.

On commence par ajuster les variogrammes simples et croisés des accumulations anamorphosées (figure 5.9). De la sorte, tous les paramètres sont spécifiés, à l'exception des nombres de mosaïques N , N_1 et N_2 . Les calculs sont effectués de manière omnidirectionnelle, car les variables ne manifestent aucune anisotropie évidente. Les modèles élémentaires considèrent des grains primaires sphériques, de diamètre 10m et 100m.



PARAMETRES DE MODELISATION

Structure 1

portée : $a = 10m$

densités : $\theta_1(1) = 57, \theta_1(3/2) = 143, \theta_1(2) = 0$

Structure 2

portée : $a = 100m$

densités : $\theta_2(1) = 0.1, \theta_2(3/2) = 0.9, \theta_2(2) = 0.2$

Effet de pépite

paliers : $\alpha_0 = 0.076, \alpha'_0 = 0.051, \alpha''_0 = 0$

Nombre de mosaïques et paramètres des lois gamma

$(N+N_1) \alpha_1 = 0.373, (N+N_2) \alpha_2 = 0.424, N\alpha = 0.373$

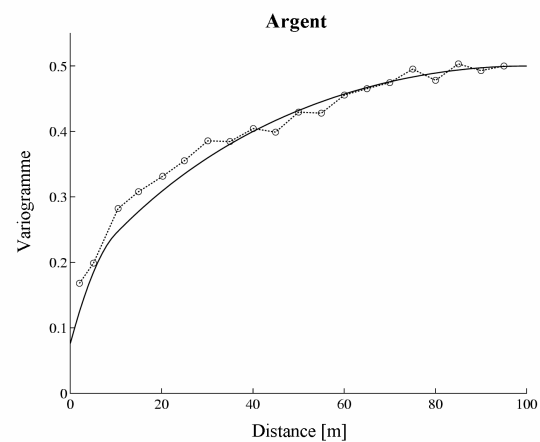
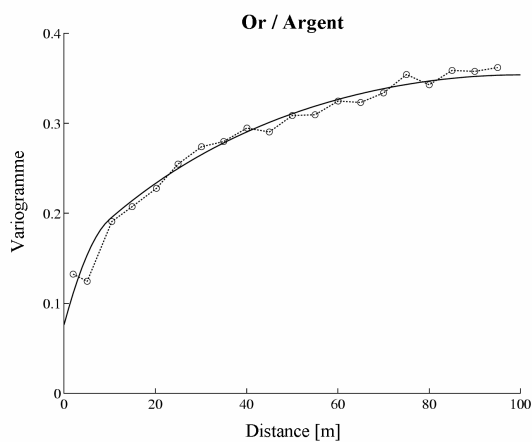


Figure 5.9. Variogrammes simples et croisés expérimentaux et modélisés

Remarque générale sur le calcul des variogrammes expérimentaux

Aux grandes distances, les variogrammes expérimentaux se stabilisent généralement au-dessus de leurs paliers théoriques (0.5 pour les variables gamma ; 0.1875 et 0.25 pour les indicatrices associées aux quartiles de la distribution). Ce problème n'est pas spécifique aux modèles isofactoriels étudiés : il se pose également dans le cadre multigaussien lorsque le variogramme expérimental de la transformée gaussienne a un palier supérieur à 1. Plusieurs causes peuvent être identifiées :

- données peu "compatibles" avec l'hypothèse de stationnarité ;
- échantillonnage irrégulier, qui détériore la qualité des variogrammes expérimentaux ;
- domaine trop petit par rapport à l'échelle du phénomène étudié, ce qui provoque un écart non négligeable entre la variance expérimentale des données et la variance *a priori* du modèle¹.

Une solution approximative consiste à normaliser les variogrammes simples et croisés autour des variances et covariances expérimentales. C'est ce parti qui a été adopté, en convenant que le palier est la valeur obtenue pour une distance de 100m.

¹ Après la procédure d'anamorphose, l'histogramme des données coïncide exactement avec une loi gamma. En particulier, la variance expérimentale est égale à la variance *a priori* du modèle, alors que la théorie des fluctuations ergodiques (Matheron, 1965, 1978b) indique qu'il ne devrait pas en être ainsi. La racine du problème se situe donc déjà au niveau de l'anamorphose des données et de la modélisation de la loi marginale.

Pour déterminer les paramètres manquants (nombres de mosaïques) on examine les variogrammes des indicatrices associés aux quartiles de la distribution marginale, ainsi que les variogrammes croisés entre indicatrices de mêmes quartiles mais de variables différentes². Ces outils sont jugés suffisants pour fournir une description adéquate des lois bivariées simples et croisées. Inclure davantage de variogrammes d'indicateurs rendrait la tâche d'inférence trop laborieuse et n'enrichirait sans doute pas la modélisation.

L'ajustement est satisfaisant en prenant $N = 1$ et $N_1 = N_2 = 0$ (figures 5.10 à 5.12) : les variables anamorphosées sont donc ajustées individuellement par un modèle mosaïque plus un effet de pépite.

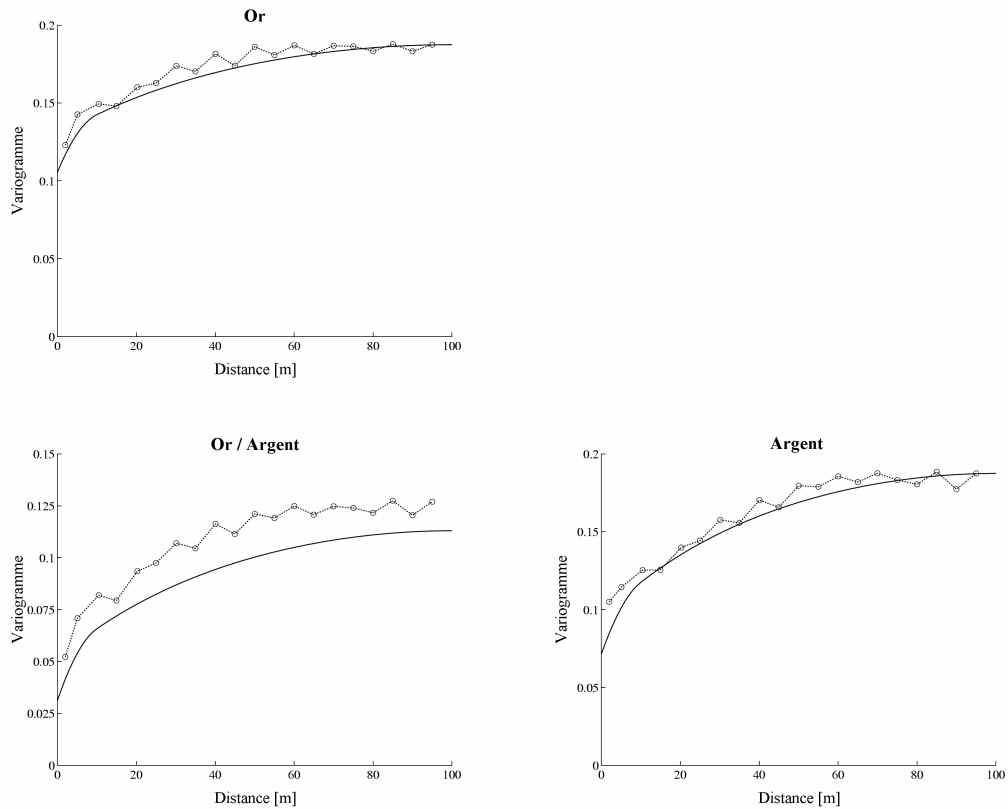


Figure 5.10. Variogrammes simples et croisés des indicatrices associées au premier quartile

² L'expression de ces variogrammes se déduit des coefficients du développement des indicatrices en polynômes de Laguerre normés et des covariances de ces derniers [Eq. (5.7) et (5.8)], cf. chapitre 3.

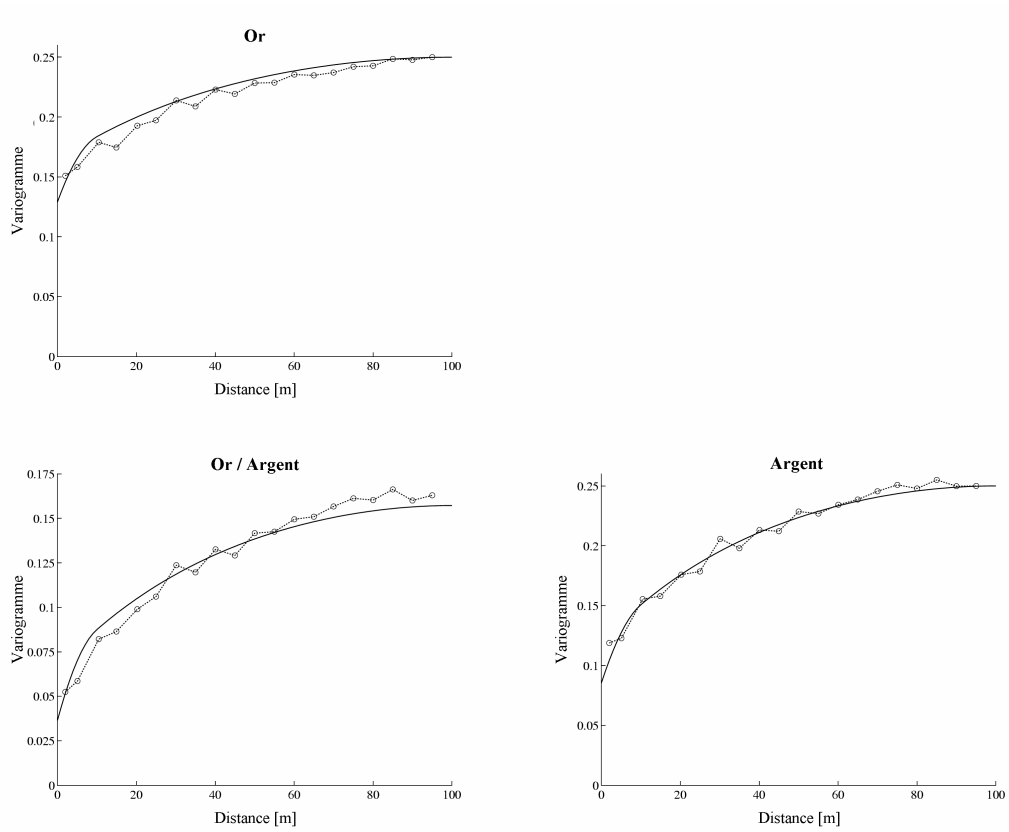


Figure 5.11. Variogrammes simples et croisés des indicatrices associées à la médiane

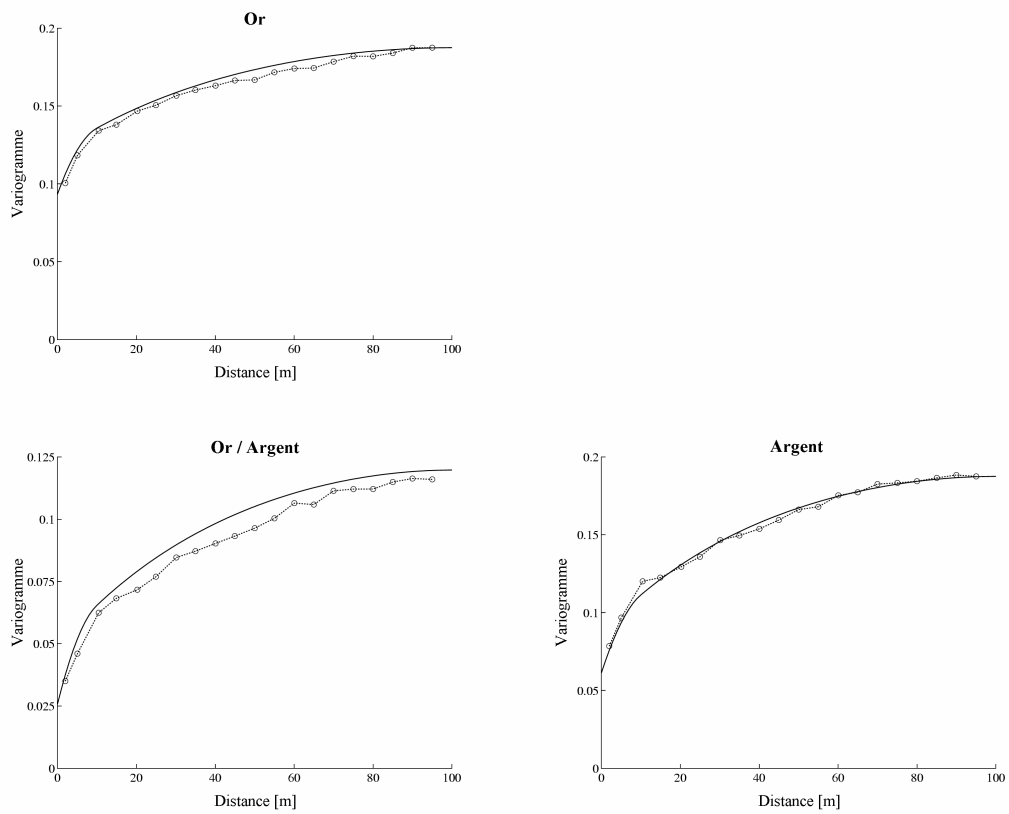


Figure 5.12. Variogrammes simples et croisés des indicatrices associées au troisième quartile

IV. Simulation conditionnelle

Rappelons le principe de l’algorithme de simulation proposé au chapitre 3. L’idée est de regarder le modèle bivariable précédent comme un processus ponctuel marqué. A chaque germe poissonien est affecté un “objet” constitué de plusieurs éléments :

- un diamètre (10m ou 100m) ;
- un indice indiquant si l’objet concerne exclusivement l’une des variables (or ou argent) ou s’il est commun aux deux variables ; le choix de cet indice dépend des densités des germes poissoniens $\theta_i(u)$;
- une ou plusieurs (selon qu’une seule variable est concernée ou les deux) valuations à loi gamma uniformes sur l’objet ;
- un ou plusieurs (idem, selon qu’une seule variable est concernée ou les deux) effets de pépite à loi gamma^3 venant s’ajouter aux valuations précédentes ;
- le numéro de la mosaïque à laquelle appartient l’objet (ce paramètre est inutile dans le cas présent, puisque le modèle ne compte qu’une seule mosaïque).

Le point de départ est une simulation non conditionnelle des deux variables (accumulations en or et en argent), effectuée sur $D \times \mathbb{R}_+$, D étant le domaine à simuler dilaté du rayon maximal des objets. Le conditionnement repose sur le recuit simulé, qui consiste à construire une chaîne markovienne d’états dont la loi converge vers la loi conditionnelle cherchée (cf. annexe C). A chaque itération, un état candidat est généré en renouvelant tous les objets germés dans un sous-domaine de $D \times \mathbb{R}_+$ dont l’emplacement est choisi au hasard. L’état candidat est ensuite accepté ou refusé selon la dynamique de Metropolis, dépendant à la fois d’une fonction objectif et d’un paramètre de température contrôlant la probabilité d’accepter des transitions non favorables.

Les paramètres choisis pour la mise en œuvre pratique du recuit simulé sont détaillés ci-après. La fonction objectif est la somme des écarts absolus entre les données et les valeurs simulées aux sites d’échantillonnage, après application de l’anamorphose gamma. A la k -ième itération, la température est prise inversement proportionnelle à $\ln(k+1)$. Le sous-domaine de renouvellement est une cellule parallélépipédique de $D \times \mathbb{R}_+$ dont les côtés sont $D_x / 5$, $D_y / 16$, $D_z / 12$ et $\tau / 15$, avec :

- D_x, D_y, D_z : longueurs du domaine D le long de chaque axe de coordonnées, soit 140m, 450m et 300m environ ;
- τ : date d’apparition du dernier grain recouvrant D lors de la simulation non conditionnelle.

Les figures 5.13 et 5.14 montrent une partie des données d’accumulations, projetées sur la surface de la veine (figure 5.6), les images initiales de la procédure de recuit (simulations non conditionnelles) et celles obtenues après 50 000 et 500 000 itérations. La coordonnée en abscisse mesure la position horizontale le long de la veine ; elle correspond à la direction nord légèrement inclinée vers l’est. Dans l’état initial, la fonction objectif vaut 1898 pour les 1491 données d’or et d’argent ; sa valeur n’est plus que de 834 dans l’état final, indiquant un écart absolu moyen de 0.28 en chaque site d’échantillonnage, à mettre en comparaison avec la moyenne (0.5) des données gamma.

³ En pratique, on garde en mémoire les valeurs pépitiqes grâce à un seul nombre (le numéro de “germe” utilisé dans la génération des nombres aléatoires), d’où une importante simplification algorithmique.

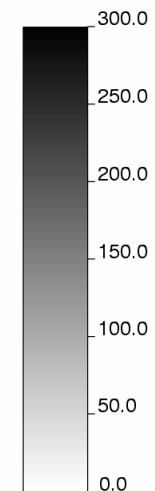
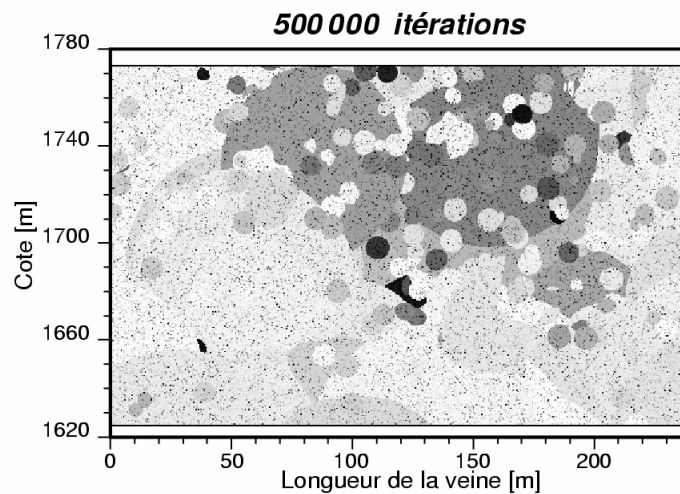
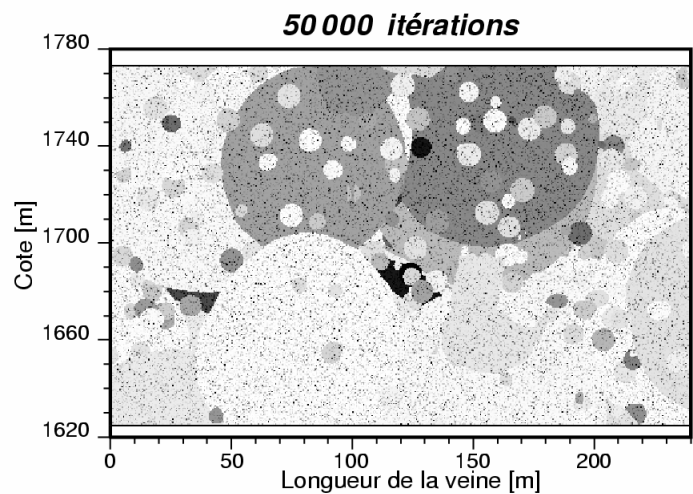
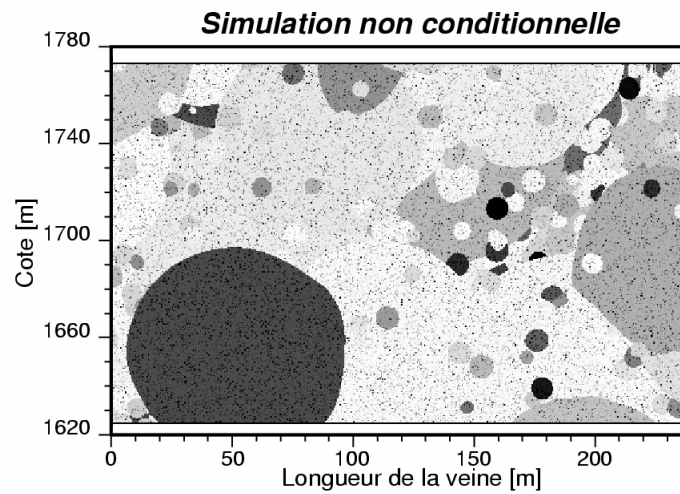
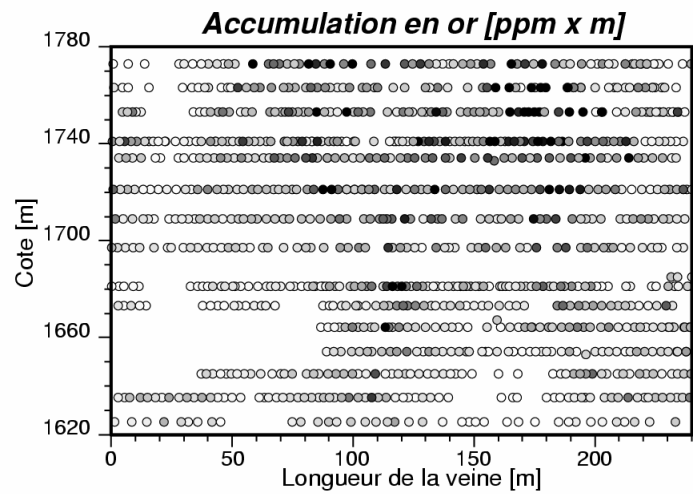


Figure 5.13. Données d'accumulation en or, simulation non conditionnelle et images obtenues par itérations markoviennes

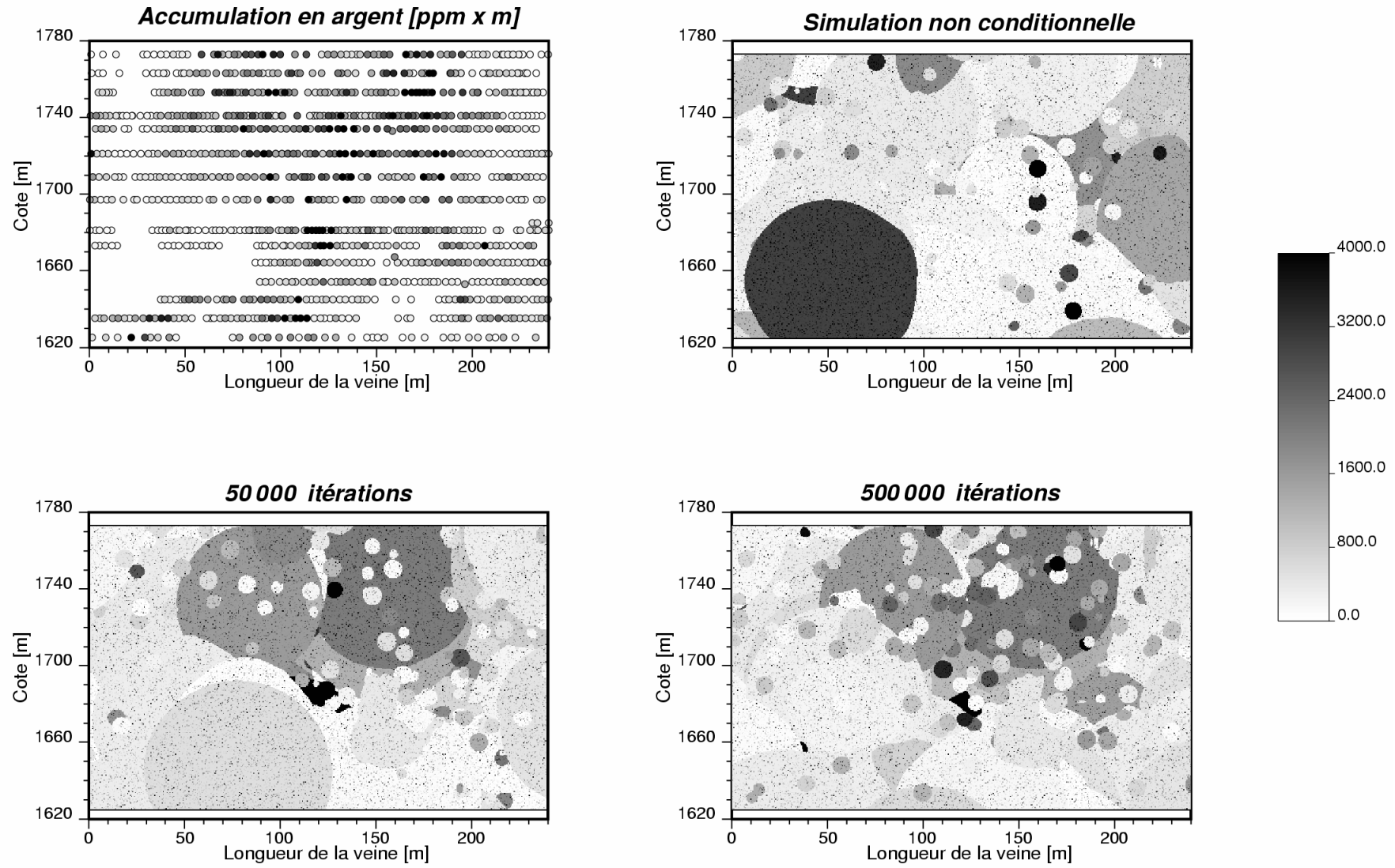


Figure 5.14. Données d'accumulation en argent, simulation non conditionnelle et images obtenues par itérations markoviennes

A titre de remarque générale, la fonction objectif décroît de plus en plus lentement à mesure que se déroule la simulation. Ainsi, elle perd environ un tiers de sa valeur initiale après 50 000 itérations, mais à partir de cet instant sa valeur stagne pratiquement : plusieurs centaines de milliers d'itérations sont alors nécessaires pour la réduire significativement. Cette lenteur de la convergence est sans doute le principal inconvénient du recuit simulé lorsque l'on a affaire à des variables à lois continues.

Pour résoudre cette difficulté, des recherches complémentaires sont nécessaires ; on pourrait par exemple

- tester plusieurs états candidats à chaque transition ;
- identifier les grains primaires ayant la plus forte contribution à la fonction objectif et effectuer les transitions préférentiellement (mais non exclusivement) dans les sous-domaines susceptibles de recouvrir ou remplacer ces grains primaires ;
- diminuer la taille du sous-domaine de renouvellement à mesure que se déroule la simulation, de manière à permettre des transitions de plus en plus fines (au bout d'un certain nombre d'itérations, les états candidats sont presque toujours rejetés). Ceci revient à définir un noyau de transition non stationnaire ; une telle piste a été évoquée par Gidas (1985) et Boulanger (1990).

V. Réflexions finales

Au-delà des difficultés d'implémentation et de la lenteur de la convergence de l'algorithme de simulation conditionnelle, les paragraphes précédents ont montré la capacité du modèle de Laguerre de décrire des variables à distributions très dissymétriques et à caractère non diffusif. Cependant, d'après les résultats du chapitre 4, un tel modèle ne permet pas de représenter des variables pour lesquelles les hautes teneurs sont moins bien structurées que les faibles teneurs. Les autres modèles isofactoriels présentés (par exemple, le modèle binomial ou celui à résidus d'indicatrices orthogonaux) sont eux-aussi limités et ne sont pas toujours adaptés aux applications pratiques.

A titre d'exemple, examinons brièvement la structuration spatiale des teneurs en or dans un autre secteur du gisement, à savoir le secteur "Caserón 506" de la veine *Quebrada Orito* (cf. figure 5.1B). Les données disponibles, composites de 1m de longueur, sont issues de sondages d'exploration et sont au nombre de 23 701 (figure 5.15).

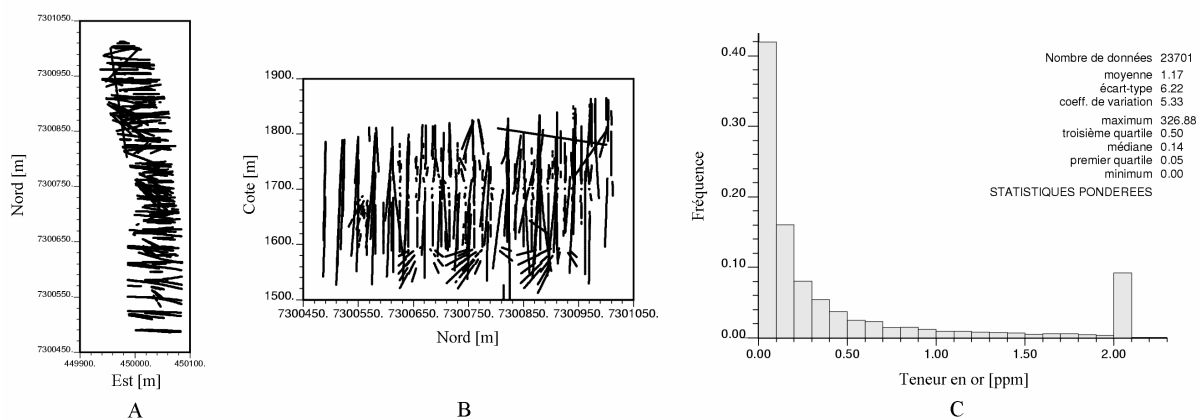


Figure 5.15. A et B, localisation des échantillons (coordonnées UTM) et C, histogramme des teneurs en or

La figure 5.16 montre les variogrammes simples et croisés des indicatrices associées aux quartiles de la distribution marginale des teneurs en or le long des directions principales d'anisotropie (à savoir, le plan horizontal et la direction verticale). Sur cette figure, on observe que l'indicateur du troisième quartile présente une portée nettement plus faible que celle du premier quartile, indiquant que les hautes teneurs sont moins bien structurées que les faibles teneurs, ce qui exclut les modèles hermitien et laguerrien. En outre, la proportion relative d'effet de pépite et l'anisotropie entre les directions horizontale et verticale diffèrent d'un variogramme à l'autre et ne permettent pas l'utilisation d'un modèle à résidus d'indicateurs orthogonaux. Ainsi, aucun des modèles isofactoriels à loi continue présentés dans cette thèse n'est satisfaisant pour décrire un tel ensemble de données.

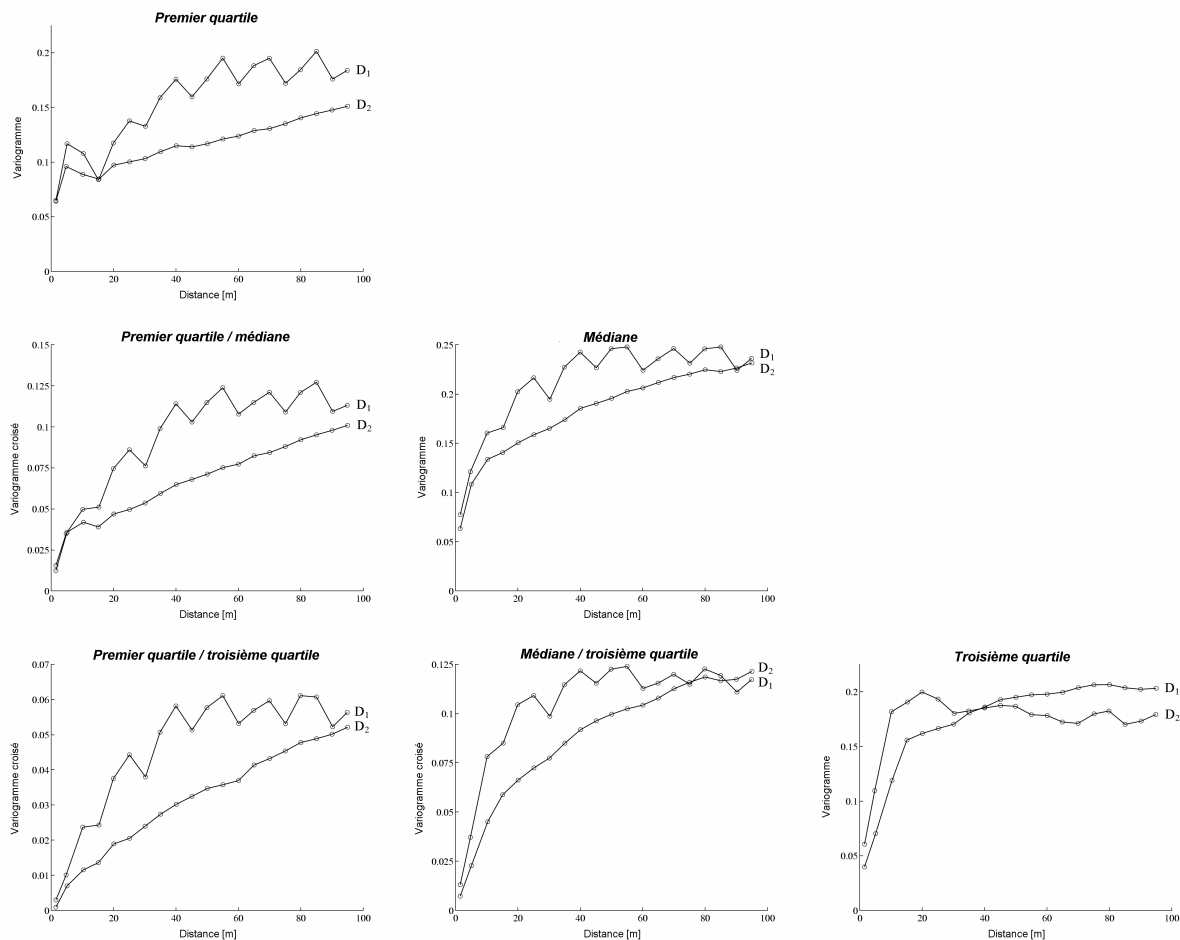


Figure 5.16. Variogrammes simples et croisés expérimentaux des indicatrices associées aux quartiles de la distribution des teneurs en or ; D_1 : plan horizontal, D_2 : direction verticale

Ces remarques montrent que l'un des points les plus délicats concerne la rigidité des modèles et la difficulté à les ajuster à une distribution empirique, même lorsque l'inférence statistique se réduit aux lois marginale et bivariées. C'est précisément cette limitation que les méthodes non paramétriques (tel l'algorithme séquentiel d'indicateurs) cherchent à contourner, au prix de fortes approximations et d'un manque de cohérence interne du modèle.

Conclusions

La simulation de fonctions aléatoires présente une importance grandissante dans les applications géostatistiques car elle permet de quantifier l'incertitude portant sur les valeurs non mesurées d'une variable régionalisée. Ce travail s'est intéressé aux fonctions aléatoires à lois bivariées isofactorielles, qui sont notamment utilisées en géostatistique minière pour estimer les réserves récupérables des gisements grâce à la technique du krigeage disjonctif. L'intérêt des modèles isofactoriels réside dans leur flexibilité, car ils peuvent décrire des variables régionalisées très diverses : discrètes ou continues, de type diffusif ou au contraire de type mosaïque.

Pour traiter le problème de la simulation, une première approche consiste à recourir à une méthode dite "non paramétrique", dont le prototype est l'algorithme séquentiel d'indicatrices. Le principe est de simuler les sites de l'espace les uns après les autres, dans un ordre aléatoire, en estimant en chaque site la distribution de probabilité conditionnelle de la valeur inconnue par un krigeage d'indicatrices. Cette approche évite de définir complètement le modèle de fonction aléatoire et se contente de spécifier quelques paramètres de sa loi bivariable (les covariances de plusieurs indicatrices) ; elle est censée être applicable en toutes circonstances, ce qui explique son succès auprès des utilisateurs. Toutefois, elle présente plusieurs graves défauts, que l'on peut résumer en trois points :

- 1) En général, le krigeage d'indicatrices est peu précis pour estimer des distributions de probabilité conditionnelles : les réalisations obtenues par l'algorithme séquentiel sont donc approximatives. Ce défaut s'explique en partie par la codification des données conditionnantes en valeurs binaires, qui provoque une perte d'information et affecte la qualité de reproduction des paramètres du modèle. Remplacer le krigeage d'indicatrices par un krigeage disjonctif améliore substantiellement la situation, surtout pour simuler des modèles de diffusion, mais reste néanmoins imparfait.
- 2) Rien ne permet d'assurer la cohérence interne du modèle, c'est-à-dire l'existence d'un processus aléatoire qui puisse être associé aux covariances d'indicatrices spécifiées. En laissant une totale liberté de choix à l'utilisateur, l'algorithme séquentiel a toutes les chances d'accepter un modèle contradictoire.
- 3) Les propriétés *a priori* des réalisations dépendent de facteurs "externes" tels que le nombre total de sites simulés, le nombre et la position des données conditionnantes. En particulier, un changement de support sur les réalisations n'a pas de sens puisqu'il dépend du niveau de résolution de l'image simulée. Il s'agit là d'une limitation très forte, car elle indique que les *images* produites par l'algorithme séquentiel ne correspondent pas à un modèle sous-jacent de fonction aléatoire ; elles ne peuvent donc être qualifiées de *simulations* stricto sensu.

Les situations où ces défauts peuvent être évités sont extrêmement rares et sont réduites à l'espace unidimensionnel ou au modèle de bruit blanc (effet de pépite pur). En dehors de ces quelques cas, les méthodes non paramétriques doivent être abandonnées : le fait qu'elles soient "passe-partout" se révèle finalement un avantage illusoire.

La seconde approche pour résoudre le problème de la simulation est plus laborieuse et se base sur la définition complète d'un modèle de fonction aléatoire. Bien que ce dernier comporte inévitablement des limitations (les degrés de liberté laissés à l'utilisateur sont restreints, d'où une certaine rigidité dans l'ajustement du modèle), il permet de garantir la cohérence des résultats.

Plusieurs modèles isofactoriels ont été examinés au cours de ce travail de thèse : modèles obtenus à partir de processus isofactoriels unidimensionnels par les procédures de substitution ou de bandes tournantes ; modèles définis par hiérarchisation ou par addition d'ensembles aléatoires indépendants ; modèles de dilution poissonnienne ; modèles construits par l'addition de mosaïques indépendantes à loi marginale indéfiniment divisible ; modèle de diffusion bigamma obtenu par sommation des carrés de fonctions aléatoires gaussiennes indépendantes. Le conditionnement des simulations à un ensemble de données est souvent réalisable par itérations markoviennes, en utilisant le procédé de recuit simulé ou l'échantillonneur de Gibbs.

L'inférence des paramètres du modèle se réduit aux lois marginales (détermination d'une fonction d'anamorphose) et bivariées. A ce titre, un outil intéressant pour synthétiser l'information bivariable est le madogramme ou variogramme d'ordre 1, mais il n'est pas toujours suffisant car il ne renseigne pas sur la dissymétrie structurale entre les valeurs fortes et faibles. Il peut être complété par l'examen des variogrammes d'ordre inférieur à 2 ou par celui des variogrammes de quelques indicatrices. Tous ces outils sont utiles pour diagnostiquer le caractère isofactoriel, ajuster et valider les paramètres du modèle.

L'application pratique à des données minières provenant d'un gisement aurifère a mis en évidence certaines difficultés qui restent à résoudre :

- 1) les temps de calcul des procédures itératives comme le recuit simulé sont très importants lorsque l'on cherche à simuler des variables à lois continues. Des recherches complémentaires sont sans doute souhaitables pour élaborer des procédures de conditionnement plus opératoires, même si les contraintes en termes de calcul ont de plus en plus tendance à disparaître ;
- 2) l'ajustement d'un modèle isofactoriel à une loi bivariable empirique n'est jamais une opération triviale à cause de la rigidité du modèle et du manque de robustesse des statistiques expérimentales lorsque la distribution des données est fort dissymétrique ;
- 3) les modèles isofactoriels classiques, tel le modèle de Laguerre, ne peuvent décrire des variables régionalisées pour lesquelles les faibles teneurs ont une meilleure structuration spatiale que les fortes teneurs. La recherche d'autres modèles (non forcément isofactoriels) est nécessaire pour pouvoir rendre compte de telles variables.

Bien qu'un gain de flexibilité considérable ait été apporté par rapport au modèle multigaussien, on ne dispose toujours pas de modèles "universels", capables de s'adapter à n'importe quelle variable régionalisée. Une telle perspective semble hors de portée.

Annexe A :

Rappels sur le krigeage disjonctif à support ponctuel

I. Définition générale

On cherche à estimer une grandeur à support ponctuel qui s'exprime comme une fonction de la variable étudiée, soit $\varphi[Y(\mathbf{x})]$, à partir d'un ensemble de données $\{Y(\mathbf{x}_\alpha), \alpha = 1 \dots n\}$. Le krigeage disjonctif de $\varphi[Y(\mathbf{x})]$, qui sera désormais indexé par KD , est une combinaison linéaire de fonctions *monovariabiles* des données (Matheron, 1973b, 1976 ; Armstrong and Matheron, 1986a, 1986b) :

$$\varphi^{KD}[Y(\mathbf{x})] = \sum_{\alpha=1}^n \psi_\alpha[Y(\mathbf{x}_\alpha)]$$

telle que les fonctions $\{\psi_\alpha, \alpha = 1 \dots n\}$ minimisent l'erreur quadratique moyenne

$$E \{ (\varphi[Y(\mathbf{x})] - \varphi^{KD}[Y(\mathbf{x})])^2 \}.$$

En géostatistique linéaire, le krigeage (simple, ordinaire, intrinsèque) ne permet d'estimer qu'une fonction φ *linéaire* et recherche l'estimateur parmi les combinaisons *linéaires* des données. Le krigeage disjonctif est à la fois plus général (φ peut être quelconque) et plus précis, l'estimateur étant pris parmi une classe plus vaste de fonctions. En contrepartie, les prérequis sont plus importants pour le krigeage disjonctif que pour le krigeage classique.

Les fonctions $\{\psi_\alpha, \alpha = 1 \dots n\}$ sont définies de manière univoque de sorte que l'erreur d'estimation soit non corrélée à toute fonction monovariabiles construite sur les données :

$$\forall \alpha = 1 \dots n, \forall \psi \text{ telle que } E\{\psi[Y(\mathbf{x}_\alpha)]^2\} < +\infty, \text{cov}\{\varphi[Y(\mathbf{x})] - \varphi^{KD}[Y(\mathbf{x})], \psi[Y(\mathbf{x}_\alpha)]\} = 0 \quad (\text{A.1})$$

En général, ces équations ne peuvent être résolues que de manière approchée, sauf dans certains cas particuliers comme les modèles *isofactoriels* que nous allons examiner ultérieurement.

Auparavant, nous allons donner une définition plus opératoire du krigeage disjonctif, à l'aide d'un développement de $\varphi[Y(\mathbf{x})]$ selon une base de fonctions.

II. Les familles génératrices

On dit qu'une famille de fonctions $\{\chi_p, p \in \mathbf{N}\}$ est **génératrice** ou **complète** pour la loi marginale de $Y(\mathbf{x})$, notée $f(y)$ ci-après, lorsqu'elle permet d'exprimer toute fonction de $Y(\mathbf{x})$ de variance finie sous la forme d'une combinaison linéaire :

$$\forall \varphi \text{ telle que } \int \varphi^2(y) f(y) dy < +\infty, \quad \varphi(y) = \sum_{p=0}^{+\infty} \varphi_p \chi_p(y).$$

On peut donner comme exemple les bases de polynômes, comme les monômes $\{y^p, p \in \mathbf{N}\}$, qui constituent une famille génératrice pour les densités marginales définies dans des intervalles bornés ou décroissant suffisamment rapidement à l'infini. Un autre exemple important est fourni par les fonctions **indicatrices**, qui sont complètes pour les lois marginales discrètes. En effet, si $Y(\mathbf{x})$ est une variable à valeurs dans un espace discret ou dénombrable, par exemple \mathbf{N} , toute fonction $\varphi[Y(\mathbf{x})]$ peut s'écrire :

$$\varphi[Y(\mathbf{x})] = \sum_{i \geq 1} [\varphi(i-1) - \varphi(i)] I(Y(\mathbf{x}); i).$$

III. Seconde définition du krigeage disjonctif

Soit $\{\chi_p, p \in \mathbf{N}\}$ une famille génératrice pour la loi marginale de $Y(\mathbf{x})$. Le krigeage disjonctif est obtenu par cokrigeage simple (indexé ci-après par *CKS*) des fonctions $\{\chi_p[Y(\mathbf{x})], p \in \mathbf{N}\}$:

$$\varphi^{KD}[Y(\mathbf{x})] = \sum_{p=0}^{+\infty} \varphi_p \{\chi_p[Y(\mathbf{x})]\}^{CKS}.$$

Le krigeage disjonctif apparaît comme le cokrigeage d'un codage disjonctif de $\varphi[Y(\mathbf{x})]$, d'où son nom. En particulier, pour une variable à valeurs discrètes ou entières, le krigeage disjonctif coïncide avec le **cokrigeage simple d'indicatrices** (Matheron, 1982 ; Rivoirard, 1994) :

$$\varphi^{KD}[Y(\mathbf{x})] = \sum_{i \geq 1} [\varphi(i-1) - \varphi(i)] \{I(Y(\mathbf{x}); i)\}^{CKS}.$$

Cette seconde définition du krigeage disjonctif sous forme d'un cokrigeage montre que le krigeage disjonctif est additif. Ainsi, le krigeage disjonctif de la somme de deux fonctions est égal à la somme de leur krigeage disjonctif.

Remarques

- 1) Pour montrer l'équivalence des deux définitions du krigeage disjonctif, la démonstration la plus rapide consiste à s'assurer que les équations du krigeage disjonctif [Eq. (A.1)] sont vérifiées par le cokrigeage simple des fonctions génératrices $\{\chi_p[Y(\mathbf{x})], p \in \mathbf{N}\}$:
 - pour une fonction ψ prise parmi $\{\chi_p, p \in \mathbf{N}\}$, à cause de la propriété d'orthogonalité de l'erreur de cokrigeage simple avec les données ;
 - pour toute fonction ψ (qui s'exprime comme une combinaison linéaire des $\{\chi_p, p \in \mathbf{N}\}$), à cause de la propriété de bilinéarité de l'opérateur de covariance.

- 2) On peut utiliser n'importe quelle famille génératrice (base polynomiale, indicatrices cumulées...) : d'après la propriété d'additivité du cokrigeage, le résultat est indépendant de la famille considérée, toute famille génératrice se déduisant de toute autre par combinaisons linéaires.
- 3) Le calcul du krigage disjonctif requiert de spécifier les lois marginales et bivariable de la fonction aléatoire. Pour le voir, il suffit de choisir les fonctions indicatrices comme famille génératrice. Leur cokrigeage simple nécessite de connaître :

- leurs espérances, qui constituent la *loi marginale* de $Y(\mathbf{x})$:

$$E[I(Y(\mathbf{x});i)] = \text{Prob}[Y(\mathbf{x}) < i].$$

- leurs covariances simples et croisées, donnant les *lois bivariées* des couples $\{Y(\mathbf{x} + \mathbf{h}), Y(\mathbf{x})\}$ pour toute séparation \mathbf{h} . En effet, la covariance non centrée entre les indicatrices $I(Y(\mathbf{x}); i)$ et $I(Y(\mathbf{x} + \mathbf{h}); j)$ s'écrit :

$$C_{ij}(\mathbf{h}) = E\{I(Y(\mathbf{x});i) I(Y(\mathbf{x} + \mathbf{h});j)\} = \text{Prob}[Y(\mathbf{x}) < i, Y(\mathbf{x} + \mathbf{h}) < j].$$

IV. Krigage disjonctif et krigage d'indicatrices

Une simplification courante consiste à réaliser un krigage séparé des indicatrices, au lieu de leur cokrigeage. Cette méthode, connue sous le nom de **krigeage d'indicatrices** (Journel, 1983, 1984 ; Matheron, 1982), doit être vue comme une solution approximative : en général, le cokrigeage ne se réduit pas au krigage de l'indicatrice concernée. En effet, les indicatrices ne sont pas indépendantes puisque :

$$\forall i \leq j, I(Y(\mathbf{x});i) \leq I(Y(\mathbf{x});j)$$

Cette hiérarchie entraîne l'existence de corrélations spatiales entre les diverses indicatrices (leurs covariances croisées ne sont jamais nulles). Par suite, le krigage d'indicatrices conduit à une perte de l'information contenue dans les données, puisqu'il néglige ces covariances croisées.

Un autre inconvénient du krigage d'indicatrices concerne l'analyse variographique : l'ajustement séparé des variogrammes d'indicatrices conduit souvent à un modèle incohérent, car il ne prend pas en compte les relations entre indicatrices ni même le caractère binaire de ces variables, qui restreint considérablement la classe des modèles possibles (cf. chapitre 4).

V. Cas de simplification du krigage disjonctif

Le krigage disjonctif correspond à un cokrigeage de fonctions génératrices. Or, il existe deux cas remarquables où le cokrigeage d'un ensemble de variables se simplifie, c'est-à-dire où il coïncide avec un krigage séparé de chaque variable (Wackernagel, 1998) : celui de la corrélation intrinsèque en situation d'homotopie (toutes les variables sont connues en tous les sites de mesure) et celui où les variables sont spatialement non corrélées. Ces cas conduisent à deux classes de modèles particuliers : le modèle mosaïque et les modèles isofactoriels respectivement.

V.1. Modèle mosaïque

Considérons une famille génératrice $\{\chi_p[Y(\mathbf{x})], p \in \mathbf{N}\}$ en corrélation intrinsèque, c'est-à-dire dont les covariances simples et croisées sont proportionnelles à un même corrélogramme $\rho(\mathbf{h})$. Si toutes ces fonctions sont connues aux sites de mesure, leur cokrigeage se réduit à un krigeage séparé, de sorte que le krigeage disjonctif s'identifie à un **krigeage simple** :

$$\varphi^{KD}[Y(\mathbf{x})] = \sum_{p=0}^{+\infty} \varphi_p \{\chi_p[Y(\mathbf{x})]\}^{KS} = \left\{ \sum_{p=0}^{+\infty} \varphi_p \chi_p[Y(\mathbf{x})] \right\}^{KS} = \varphi^{KS}[Y(\mathbf{x})].$$

$\varphi[Y(\mathbf{x})]$ étant une combinaison linéaire des $\{\chi_p[Y(\mathbf{x})], p \in \mathbf{N}\}$, sa covariance est une combinaison linéaire de leurs covariances simples et croisées, donc est proportionnelle à $\rho(\mathbf{h})$. Par conséquent, il suffit de résoudre un seul système de krigeage avec le corrélogramme $\rho(\mathbf{h})$: les pondérateurs obtenus sont valables pour n'importe quelle fonction de $Y(\mathbf{x})$. Cette situation correspond au **modèle mosaïque** (Matheron, 1982 ; Rivoirard, 1994).

Si l'on prend les fonctions indicatrices comme famille génératrice, on s'aperçoit que le krigeage disjonctif est identique au krigeage d'indicatrices. Dans ce cas particulier, le krigeage d'indicatrices ne conduit pas à une simplification du krigeage disjonctif.

Il convient toutefois de souligner le caractère exceptionnel de ce modèle. En général, on observe une déstructuration des indicatrices relatives aux seuils élevés, qui deviennent pépitiqes. Par suite, associer le même corrélogramme à toutes les indicatrices tend à exagérer la structure des seuils élevés et surestimer leur portée, ce qui introduit un biais dans les estimations.

V.2. Modèles isofactoriels

Le second cas de simplification du krigeage disjonctif est celui où il existe une famille génératrice de fonctions sans corrélations spatiales croisées, appelées *facteurs*. Les modèles correspondants sont qualifiés d'**isofactoriels** (Matheron, 1973b, 1975b, 1976, 1978a, 1980a, 1984a). De tels modèles sont plus intéressants que le modèle mosaïque, car ils permettent d'explicitier le krigeage disjonctif, tout en restant adaptés à la description d'une vaste classe de phénomènes. Du reste, le modèle mosaïque est un cas particulier de nombreuses catégories de modèles isofactoriels.

Annexe B :

Modèles isofactoriels à support ponctuel

I. Les lois isofactorielles symétriques

I.1. Définition

On dit qu'une fonction aléatoire $\{Y(\mathbf{x}), \mathbf{x} \in \mathbb{R}^d\}$ est **isofactorielle** si la densité $f(y_1, y_2)$ de tout couple de valeurs $\{Y(\mathbf{x}_1), Y(\mathbf{x}_2)\}$ admet un développement de la forme :

$$f(y_1, y_2) = f(y_1)f(y_2) \sum_{p=0}^{+\infty} T_p(\mathbf{x}_1, \mathbf{x}_2) \chi_p(y_1) \chi_p(y_2)$$

où $f(y)$ est la densité marginale de la fonction aléatoire

$\{T_p(\mathbf{x}_1, \mathbf{x}_2), p \in \mathbb{N}\}$ est une suite de coefficients dépendant de \mathbf{x}_1 et \mathbf{x}_2 , ou plus simplement de leur séparation $\mathbf{x}_1 - \mathbf{x}_2$ (cas stationnaire)

$\{\chi_p, p \in \mathbb{N}\}$ est une famille de fonctions réelles, appelées **facteurs**, formant une base orthonormée des fonctions de carré intégrable pour la loi marginale $f(y)$:

(a) la famille $\{\chi_p, p \in \mathbb{N}\}$ est génératrice pour la loi $f(y)$:

$$\forall \varphi / \int \varphi^2(y) f(y) dy < +\infty, \varphi(y) = \sum_{p=0}^{+\infty} \varphi_p \chi_p(y).$$

(b) la famille $\{\chi_p, p \in \mathbb{N}\}$ est orthonormée pour cette loi marginale :

$$\forall p, q \in \mathbb{N}, \int \chi_p(y) \chi_q(y) f(y) dy = E\{\chi_p[Y(\mathbf{x})] \chi_q[Y(\mathbf{x})]\} = \begin{cases} 0 & \text{si } p \neq q \\ 1 & \text{sinon} \end{cases}$$

Le préfixe "iso" vient du fait que les facteurs sont les mêmes pour tous les couples de valeurs $\{Y(\mathbf{x}_1), Y(\mathbf{x}_2)\}$, indépendamment de leur position et de leur séparation (Matheron, 1989b).

Remarques

- 1) Usuellement, on impose $\chi_0(y) \equiv 1$.
- 2) Il existe d'autres types de lois isofactorielles (les lois *dissymétriques*), qui font intervenir deux familles de facteurs $\{\chi_p^1, p \in \mathbb{N}\}$ et $\{\chi_p^2, p \in \mathbb{N}\}$. Ces lois sont utilisées dans les problèmes multivariés ou dans les problèmes de changement de support, pour modéliser la distribution bivariable entre les valeurs ponctuelles et de bloc.

I.2. Propriétés

De la définition des lois isofactorielles découlent les propriétés suivantes :

- (1) les coefficients du développement de φ en série s'obtiennent par les relations :

$$\forall p \in \mathbb{N}, \varphi_p = \int \varphi(y) \chi_p(y) f(y) dy$$

En particulier, φ_0 n'est autre que l'espérance de $\varphi[Y(\mathbf{x})]$.

- (2) Hormis χ_0 , les facteurs appliqués à une valeur $Y(\mathbf{x})$ de loi $f(y)$ sont d'espérance nulle, de variance unité et non corrélés entre eux :

$$\forall p, q > 0, E\{\chi_p[Y(\mathbf{x})]\} = 0, \text{var}\{\chi_p[Y(\mathbf{x})]\} = 1 \text{ et } \text{cov}\{\chi_p[Y(\mathbf{x})], \chi_q[Y(\mathbf{x})]\} = 0 \text{ si } p \neq q$$

- (3) La variance de $\varphi[Y(\mathbf{x})]$ se développe au moyen des coefficients $\{\varphi_p, p \in \mathbb{N}\}$:

$$\text{var}\{\varphi[Y(\mathbf{x})]\} = \sum_{p=1}^{+\infty} \varphi_p^2$$

- (4) Plus généralement, pour deux fonctions φ et φ' , $\text{cov}\{\varphi[Y(\mathbf{x})], \varphi'[Y(\mathbf{x})]\} = \sum_{p=1}^{+\infty} \varphi_p \varphi'_p$.

- (5) L'orthogonalité des facteurs s'étend à la loi bivariable :

$$\forall p, q \in \mathbb{N}, E\{\chi_p[Y(\mathbf{x}_1)] \chi_q[Y(\mathbf{x}_2)]\} = \begin{cases} 0 & \text{si } p \neq q \\ T_p(\mathbf{x}_1, \mathbf{x}_2) & \text{sinon} \end{cases}$$

- (6) Si le couple $\{Y(\mathbf{x}_1), Y(\mathbf{x}_2)\}$ suit la loi isofactorielle de densité $f(y_1, y_2)$, alors

$$\forall p \in \mathbb{N}, E\{\chi_p[Y(\mathbf{x}_1)] | Y(\mathbf{x}_2)\} = T_p(\mathbf{x}_1, \mathbf{x}_2) \chi_p[Y(\mathbf{x}_2)].$$

Cette propriété est connue sous le nom de *relation d'échange des facteurs*.

- (7) Les propriétés précédentes entraînent : $\text{cov}\{\varphi[Y(\mathbf{x}_1)], \varphi'[Y(\mathbf{x}_2)]\} = \sum_{p=1}^{+\infty} T_p(\mathbf{x}_1, \mathbf{x}_2) \varphi_p \varphi'_p$.

En particulier, pour $p > 0$, $\text{cov}\{\chi_p[Y(\mathbf{x}_1)], \chi_p[Y(\mathbf{x}_2)]\} = T_p(\mathbf{x}_1, \mathbf{x}_2)$.

D'après l'inégalité de Schwarz, la valeur absolue des coefficients $\{T_p(\mathbf{x}_1, \mathbf{x}_2), p \in \mathbb{N}\}$ est inférieure ou égale à 1 : $\forall p > 0, |T_p(\mathbf{x}_1, \mathbf{x}_2)| \leq 1$.

I.3. Le krigeage disjonctif isofactoriel

On suppose que la fonction aléatoire $Y(\mathbf{x})$ suit, après anamorphose, une loi bivariable isofactorielle caractérisée par la famille de facteurs $\{\chi_p, p \in \mathbf{N}\}$ et leurs corrélogrammes $\{T_p(\mathbf{h}), p \in \mathbf{N}\}$ [propriété (7)]. On cherche à estimer une grandeur $\varphi[Y(\mathbf{x})]$ de variance finie :

$$\varphi[Y(\mathbf{x})] = \sum_{p=0}^{+\infty} \varphi_p \chi_p[Y(\mathbf{x})].$$

D'après l'annexe A, le krigeage disjonctif équivaut au cokrigeage simple des facteurs. Or, à cause de la propriété (5), un tel cokrigeage s'identifie au krigeage simple (KS) de chaque facteur séparément, d'où finalement :

$$\varphi^{KD}[Y(\mathbf{x})] = \sum_{p=0}^{+\infty} \varphi_p \chi_p^{KS}[Y(\mathbf{x})].$$

Ainsi, s'il existe une famille à la fois génératrice et orthonormée, le krigeage disjonctif s'obtient en décomposant la fonction à estimer sur cette famille et en krigeant séparément chaque terme. La variance d'estimation s'écrit :

$$\text{var}\{\varphi[Y(\mathbf{x})] - \varphi^{KD}[Y(\mathbf{x})]\} = \sum_{p=0}^{+\infty} \varphi_p^2 \text{var}\{\chi_p[Y(\mathbf{x})] - \chi_p^{KS}[Y(\mathbf{x})]\}.$$

Comme la variance de krigeage simple, la variance de krigeage disjonctif prend en compte la structure spatiale de la régionalisation et la configuration géométrique formée par le site à estimer et les sites d'observation, mais est indépendante des valeurs prises par les données $\{Y(\mathbf{x}_\alpha), \alpha = 1 \dots n\}$.

II. Exemples

II.1. Modèles bigaussien et hermitien

La loi marginale est une gaussienne centrée et réduite (moyenne 0 et variance 1), tandis que les facteurs sont les polynômes d'Hermite normés, définis par la formule de Rodrigues :

$$\forall p \in \mathbf{N}, H_p(y) = \frac{1}{\sqrt{p!}} \frac{d^p g(y)}{dy^p}.$$

Dans le **modèle bigaussien** (Matheron, 1973b, 1974, 1975b ; Rivoirard, 1994), les covariances des facteurs sont les puissances successives du corrélogramme $\rho(\mathbf{h})$ de la gaussienne :

$$\forall p > 0, T_p(\mathbf{h}) = [\rho(\mathbf{h})]^p$$

Ces covariances tendent vers un effet de pépite lorsque p devient infini, de sorte que le krigeage disjonctif ne nécessite que le krigeage simple d'un nombre limité de facteurs (les termes d'ordre élevé étant pratiquement nuls). Il s'agit d'un *modèle de diffusion*.

Dans le **modèle hermitien** (Matheron, 1974, 1975b, 1976), les covariances des facteurs sont les moments d'une variable aléatoire $R(\mathbf{h})$ à valeurs dans $[-1,1]$:

$$\forall p > 0, T_p(\mathbf{h}) = E\{[R(\mathbf{h})]^p\}$$

On retrouve le modèle bigaussien en prenant $R(\mathbf{h})$ presque sûrement constante et égale à $\rho(\mathbf{h})$. Un autre cas remarquable est celui où $R(\mathbf{h})$ ne prend que les valeurs 0 ou 1 : les variables $Y(\mathbf{x} + \mathbf{h})$ et $Y(\mathbf{x})$ sont égales ou indépendantes ; il s'agit, à une anamorphose près, du *modèle mosaïque*. Dans le cas général, la loi hermitienne est un mélange de lois bigaussiennes dont le coefficient de corrélation est randomisé entre -1 et 1 .

Le krigeage disjonctif est difficile à mettre en œuvre lorsque la convergence de $\{T_p(\mathbf{h}), p \in \mathbf{N}\}$ vers un effet de pépite est lente, puisqu'alors les termes d'ordres élevés ne s'annulent pas dans les développements en polynômes d'Hermite. Une solution approchée est cependant possible (Emery, 2002).

II.2. Modèles bigamma et de Laguerre

Dans ces modèles, la loi marginale est une distribution gamma qui dépend d'un paramètre positif α contrôlant son degré de dissymétrie (Matheron, 1973b, 1975b, 1983 ; Hu, 1988). Les facteurs sont les polynômes de Laguerre normés d'ordre α :

$$\forall p \in \mathbf{N}, L_p^\alpha(y) = \sqrt{\frac{\Gamma(\alpha)}{\Gamma(p+1)\Gamma(\alpha+p)}} \frac{1}{e^{-y} y^{\alpha-1}} \frac{d^p}{dy^p} (e^{-y} y^{p+\alpha-1}).$$

Dans le **modèle bigamma**, la covariance du facteur d'ordre p est égale au corrélogramme $\rho(\mathbf{h})$ de la variable gamma (ce dernier ne pouvant jamais prendre de valeurs négatives) élevé à la puissance p :

$$\forall p > 0, T_p(\mathbf{h}) = [\rho(\mathbf{h})]^p$$

alors que dans le **modèle de Laguerre**, il s'agit du moment d'ordre p d'une variable aléatoire $R(\mathbf{h})$ à valeurs dans $[0,1]$:

$$\forall p > 0, T_p(\mathbf{h}) = E\{[R(\mathbf{h})]^p\}$$

Si α devient infini, le modèle bigamma se rapproche du modèle bigaussien. L'intérêt des modèles bigamma et laguerrien est leur capacité à modéliser des variables régionalisées dont l'histogramme présente une forte dissymétrie ou une proportion importante de valeurs nulles ou quasi-nulles.

II.3. Autres modèles

D'autres modèles isofactoriels ont été développés, notamment :

- le **modèle beta** (Matheron, 1975b, 1984a) : la loi marginale est une loi beta, laquelle dépend de deux paramètres positifs $\{\alpha, \beta\}$. Les facteurs sont les polynômes de Jacobi ; leurs covariances sont liées au corrélogramme $\rho(\mathbf{h})$ de la variable beta par :

$$\forall p > 0, T_p(\mathbf{h}) = [\rho(\mathbf{h})]^{p(\alpha+\beta+p-1)/(\alpha+\beta)}$$

- le **modèle de Walsh** (Matheron, 1975c) : la loi marginale est uniforme tandis que les facteurs sont les fonctions de Walsh. Dans le cas le plus général, leurs covariances sont les moments successifs d'une variable aléatoire $R(\mathbf{h})$ à valeur dans $[-1,1]$;
- le **modèle de Gauss inverse** (Hu, 1988) : la distribution marginale est une loi de Gauss inverse ou loi de Sichel ; les facteurs de la loi bivariable n'ont cependant toujours pas été explicités ;
- les **modèles binomial négatif et de Meixner** : la loi marginale est une loi binomiale négative qui dépend de deux paramètres η et α avec $\eta \in [0,1[$ et $\alpha > 0$. Les facteurs de la loi bivariable sont les polynômes de Meixner normés. Leurs covariances sont les puissances du corrélogramme de $Y(\mathbf{x})$ pour le modèle binomial négatif (Matheron, 1975c, 1980a) ou les moments d'une variable aléatoire $R(\mathbf{h})$ à valeurs dans $[0,1]$ pour le modèle de Meixner (Matheron, 1973b, 1983). Si η tend vers 1, la loi de $(1-\eta)Y(\mathbf{x})$ converge vers une loi laguerrienne de paramètre α ;
- les **modèles binomial et de Krawtchouk** : la distribution marginale est binomiale ; elle est définie sur un intervalle borné $\{0,1\dots m\}$ et dépend d'un paramètre $\eta \in]0,1[$. Les polynômes orthonormés pour cette loi marginale sont appelés polynômes de Krawtchouk. Dans le modèle diffusif binomial (Matheron, 1975c, 1980a), les covariances des facteurs sont les puissances du corrélogramme de la variable binomiale. Plus généralement, dans le modèle de Krawtchouk, ce sont les moments d'une variable aléatoire $R(\mathbf{h})$ à valeurs dans $[-1,1]$. Si m tend vers l'infini, la loi de la variable $Y(\mathbf{x})$, une fois centrée et réduite, converge vers une loi hermitienne ;
- les **modèles de Poisson et de Charlier** : la loi de Poisson dépend d'un paramètre θ positif et ses facteurs sont les polynômes de Charlier. A l'instar des modèles précédents, il est possible de définir un modèle de diffusion et un modèle composé – les modèles de Poisson et de Charlier (Matheron, 1973b, 1975c, 1980a, 1984a) – selon que les covariances des facteurs sont les puissances du corrélogramme de $Y(\mathbf{x})$ ou les moments d'une variable aléatoire $R(\mathbf{h})$ à valeurs dans $[0,1]$. Le modèle de Poisson est un cas limite des modèles précédents. En effet, dans le modèle binomial négatif, si $\eta \rightarrow 0$ et $\alpha \rightarrow +\infty$ tel que $\alpha\eta \rightarrow \theta$, on retrouve le modèle de Poisson de paramètre θ . De même, le modèle binomial redonne le modèle de Poisson lorsque $\eta \rightarrow 0$ et $m \rightarrow +\infty$ avec $m\eta \rightarrow \theta$;
- le **modèle de Jacobi** (Matheron, 1984a, 1989b) : la loi de Jacobi est une loi binomiale de taille m dont le paramètre η est randomisé par une loi beta de paramètres $\{\alpha,\beta\}$ positifs. Les facteurs de la loi bivariable sont les polynômes de Jacobi discrets, qui dépendent de trois paramètres $\{\alpha,\beta,m\}$. Ce modèle permet de retrouver asymptotiquement la plupart des exemples précédents (binomial, binomial négatif, Poisson et beta), ce qui lui confère une grande souplesse d'application. Les corrélogrammes des facteurs sont définis à partir de celui de la variable $Y(\mathbf{x})$:

$$\forall p > 0, T_p(\mathbf{h}) = [\rho(\mathbf{h})]^{p(\alpha+\beta+p-1)/(\alpha+\beta)}$$

- le **modèle anti-Jacobi** (Matheron, 1984a, 1989b) : la loi marginale dépend de deux réels positifs $\{\alpha,\beta\}$ et d'un entier m qui est la borne maximale atteinte par le processus discret. Les facteurs sont définis à partir des polynômes de Jacobi discrets. Ce modèle présente moins d'intérêt pratique que le modèle de Jacobi discret, à cause de la forme en cloche des lois marginales (peu fréquente dans les applications) et de la difficulté à lui associer un modèle de changement de support. Les corrélogrammes des facteurs sont liés à celui de la variable $Y(\mathbf{x})$ par :

$$\forall p > 0, T_p(\mathbf{h}) = [\rho(\mathbf{h})]^{p(2m+\alpha+\beta-p-1)/(2m+\alpha+\beta-2)}$$

- le **modèle hypergéométrique** : ce modèle est mentionné par Matheron (1975c) mais ses facteurs n'ont pas encore été explicités ;

- les **modèles discrets construits à partir des lois empiriques** (Matheron, 1984a ; Lajaunie and Lantuéjoul, 1989) : ces modèles font appel à la théorie des processus de naissance et de mort dans l'espace unidimensionnel (l'axe temporel). L'aspect multidimensionnel de $Y(\mathbf{x})$ et son éventuelle anisotropie sont pris en compte par un variogramme sans palier, qui relie l'espace "géographique" à l'espace "temporel" du processus de naissance et de mort.

II.4. Modèles de type mosaïque

Le **modèle mosaïque** est un cas extrême de nombreux modèles précédents. L'espace est divisé en cellules de géométrie aléatoire en chacune desquelles la fonction aléatoire prend une valeur constante, indépendante d'une cellule à l'autre. Toute fonction de la variable admet le même corrélogramme que celle-ci, en particulier ses indicatrices. Il n'existe en conséquence aucune déstructuration des valeurs élevées (Matheron, 1982).

Un modèle un peu plus général est le **modèle à résidus d'indicatrices autokrigeables** (Rivoirard, 1989, 1994 ; Matheron, 1989b), dans lequel la fonction aléatoire étudiée n'admet pas d'effets de bord en montant ou en descendant. Dans le premier cas, le variogramme croisé entre deux indicatrices est proportionnel au variogramme de l'indicatrice associée au seuil le plus bas ; dans le second cas, il est proportionnel au variogramme de l'indicatrice associée au seuil le plus élevé.

III. Deux modes de construction

Deux méthodologies ont été élaborées pour définir des lois bivariées isofactorielles. La première consiste à régulariser une mesure aléatoire orthogonale (Matheron, 1973b ; Armstrong and Matheron, 1986a). Cette approche met en évidence les modèles dont les lois marginales et bivariées sont indéfiniment divisibles : bigaussien, laguerrien, de Poisson, de Meixner et à loi de Gauss inverse.

La seconde approche repose sur la théorie des processus markoviens de diffusion dans le cadre continu ou des processus de naissance et de mort dans le cadre discret (Armstrong and Matheron, 1986b ; Matheron, 1975b, 1975c, 1984a, 1985a, 1989b ; Lajaunie and Lantuéjoul, 1989). Ce procédé est plus général que la régularisation de mesures orthogonales et permet d'obtenir tous les modèles mentionnés précédemment. Dans un premier temps, il fournit des processus isofactoriels à une dimension dont les facteurs ont des covariances exponentielles. L'extension à plusieurs dimensions est possible par substitution (Matheron, 1989b ; Lantuéjoul, 1993, 2002).

IV. Aspects pratiques

IV.1. Anamorphose

Le choix de la distribution marginale après anamorphose (caractère discret ou continu, type de distribution et paramètres) est guidé par le degré de dissymétrie que présente la variable initiale ou par l'éventuelle présence d'une proportion non négligeable de valeurs nulles (Matheron, 1980a). Dans le cadre gamma, des critères pratiques sont proposés par Demange *et al.* (1987), Hu (1988), Hu et Lantuéjoul (1988).

IV.2. Choix des paramètres de la loi bivariable

La loi bivariable de $Y(\mathbf{x})$ est caractérisée par les corrélogrammes $\{T_p(\mathbf{h}), p \in \mathbf{N}\}$ ou, de façon équivalente, par la variable aléatoire $R(\mathbf{h})$, à valeurs dans $[0,1]$ ou $[-1,1]$ selon le modèle. Pour faciliter l'inférence statistique, il est commode d'introduire un modèle de distribution pour $R(\mathbf{h})$, dépendant de peu de paramètres. Un modèle intéressant est celui où $R(\mathbf{h})$ suit une **loi beta** ; cette loi est comprise entre 0 et 1 et dépend de deux réels positifs $a(\mathbf{h})$ et $b(\mathbf{h})$; la densité de probabilité de $R(\mathbf{h})$ s'écrit :

$$\varpi_{\mathbf{h}}(d\rho) = \frac{\Gamma(a(\mathbf{h}) + b(\mathbf{h}))}{\Gamma(a(\mathbf{h}))\Gamma(b(\mathbf{h}))} \rho^{a(\mathbf{h})-1} (1-\rho)^{b(\mathbf{h})-1} d\rho.$$

Si $a(\mathbf{h})$ et $b(\mathbf{h})$ tendent vers l'infini, on obtient un modèle de diffusion ; au contraire, s'ils tendent vers 0, on retrouve le modèle mosaïque. Entre ces deux extrêmes, la combinaison des paramètres $a(\mathbf{h})$ et $b(\mathbf{h})$ permet de décrire une vaste classe de modèles, où la structuration des valeurs extrêmes est plus ou moins accentuée. Dans les modèles où les corrélogrammes $\{T_p(\mathbf{h}), p \in \mathbf{N}\}$ sont les moments de $R(\mathbf{h})$, l'écriture générale de $T_p(\mathbf{h})$ est la suivante :

$$T_p(\mathbf{h}) = \frac{\Gamma(a(\mathbf{h}) + b(\mathbf{h}))\Gamma(a(\mathbf{h}) + p)}{\Gamma(a(\mathbf{h}))\Gamma(a(\mathbf{h}) + b(\mathbf{h}) + p)}.$$

En pratique, il est plus simple d'employer comme paramètres $\rho(\mathbf{h})$ et $\beta = a(\mathbf{h}) + b(\mathbf{h})$, que l'on suppose constant :

$$T_p(\mathbf{h}) = \frac{\Gamma(\beta)\Gamma(\beta\rho(\mathbf{h}) + p)}{\Gamma(\beta\rho(\mathbf{h}))\Gamma(\beta + p)} \text{ avec } \beta > 0.$$

Il suffit de connaître le corrélogramme $\rho(\mathbf{h})$ de la variable anamorphosée et le paramètre β pour déterminer le modèle. Ce dernier peut être choisi à partir de la relation entre le madogramme $\gamma_1(\mathbf{h})$ de $Y(\mathbf{x})$ et son variogramme $\gamma(\mathbf{h})$ (Chilès and Delfiner, 1999) :

$$\frac{\gamma_1(\mathbf{h})}{\gamma_1(\infty)} = \frac{\Gamma(\beta)}{\Gamma(\beta + 1/2)} \frac{\Gamma(\beta\gamma(\mathbf{h})/\gamma(\infty) + 1/2)}{\Gamma(\beta\gamma(\mathbf{h})/\gamma(\infty))}.$$

Le modèle de diffusion correspond à $\beta \rightarrow \infty$: le variogramme est alors proportionnel au carré du madogramme. Au contraire, le modèle mosaïque correspond à $\beta \rightarrow 0$: dans ce cas, le variogramme est proportionnel au madogramme.

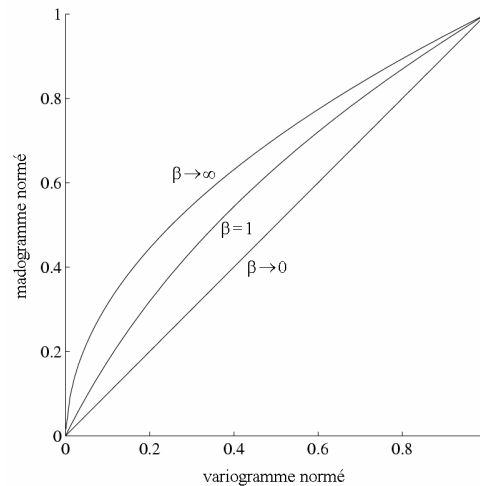


Figure B.1. Madogramme en fonction du variogramme dans le modèle betifié

V. Un critère pour identifier un modèle isofactoriel

Comment savoir si une fonction aléatoire donnée est isofactorielle ? Nous mentionnons ci-dessous un critère très simple permettant de tester la propriété isofactorielle sans expliciter les facteurs ; ce critère a récemment été proposé par Subramanyam et Pandalai (2001). On commencera par examiner le cas d'une fonction aléatoire dont les valeurs sont en nombre fini, par exemple les entiers compris entre 0 et m. Les densités bivariées isofactorielles s'expriment sous la forme

$$P_{\mathbf{h}}(i, j) = \text{Prob}[Y(\mathbf{x} + \mathbf{h}) = i, Y(\mathbf{x}) = j] = P_i P_j \sum_{p=0}^m T_p(\mathbf{h}) \chi_p(i) \chi_p(j) \text{ avec } \sum_{i=0}^m \chi_p(i) \chi_q(i) P_i = \delta_{pq}$$

Cette équation peut s'écrire sous forme matricielle :

$$\mathbf{P}(\mathbf{h}) = {}^t \mathbf{X} \mathbf{T}(\mathbf{h}) \mathbf{X} \text{ avec } {}^t \mathbf{X} \mathbf{X} = \mathbf{I}$$

avec

- $\mathbf{P}(\mathbf{h})$: matrice symétrique à termes positifs de terme générique (ligne i, colonne j) $P_{\mathbf{h}}(i, j) / \sqrt{P_i P_j}$
- \mathbf{X} : matrice carrée dont le terme générique (ligne p, colonne i) est $\sqrt{P_i} \chi_p(i)$
- $\mathbf{T}(\mathbf{h})$: matrice diagonale dont le terme générique (ligne p, colonne p) est $T_p(\mathbf{h})$

On a donc exprimé la propriété isofactorielle sous la forme d'un problème de diagonalisation de la matrice $\mathbf{P}(\mathbf{h})$ dans une base orthonormée. Comme $\mathbf{P}(\mathbf{h})$ est une matrice symétrique réelle, une première conséquence est qu'une telle diagonalisation est toujours possible. Autrement dit, toute loi bivariée discrète est "factorielle".

La question est de savoir à quelle condition l'ensemble des matrices $\mathbf{P}(\mathbf{h})$ induites par un modèle de loi spatiale sont associées aux mêmes facteurs $\{\chi_p, p = 0 \dots m\}$ (caractère isofactoriel). Cela revient à chercher la condition sous laquelle toutes ces matrices sont co-diagonalisables, i.e. diagonalisables dans la même base orthonormée. Une condition nécessaire et suffisante est que pour toute paire de distances $(\mathbf{h}, \mathbf{h}')$, les matrices $\mathbf{P}(\mathbf{h})$ et $\mathbf{P}(\mathbf{h}')$ commutent.

La condition est suffisante : si $\mathbf{P}(\mathbf{h}) = {}^t \mathbf{X} \mathbf{T}(\mathbf{h}) \mathbf{X}$ et $\mathbf{P}(\mathbf{h}') = {}^t \mathbf{X} \mathbf{T}(\mathbf{h}') \mathbf{X}$, alors

$$\mathbf{P}(\mathbf{h}) \mathbf{P}(\mathbf{h}') = {}^t \mathbf{X} \mathbf{T}(\mathbf{h}) \mathbf{X} {}^t \mathbf{X} \mathbf{T}(\mathbf{h}') \mathbf{X} = {}^t \mathbf{X} \mathbf{T}(\mathbf{h}) \mathbf{T}(\mathbf{h}') \mathbf{X}$$

$$\mathbf{P}(\mathbf{h}') \mathbf{P}(\mathbf{h}) = {}^t \mathbf{X} \mathbf{T}(\mathbf{h}') \mathbf{X} {}^t \mathbf{X} \mathbf{T}(\mathbf{h}) \mathbf{X} = {}^t \mathbf{X} \mathbf{T}(\mathbf{h}') \mathbf{T}(\mathbf{h}) \mathbf{X}$$

et le critère est bien satisfait puisque deux matrices diagonales commutent.

Inversement, soit un ensemble de matrices symétriques réelles $\{\mathbf{P}(\mathbf{h}_i), i \in I\}$ commutant deux à deux. Introduisons les endomorphismes $\{u_i, i \in I\}$ canoniquement associés à ces matrices ; ces endomorphismes sont diagonalisables en base orthonormée. Pour montrer qu'ils sont co-diagonalisables, un lemme préliminaire est utile : si u_i et u_j commutent, alors u_j laisse stables les sous-espaces propres de u_i . En effet, soit x un vecteur propre de u_i associé à la valeur propre λ :

$$u_i(x) = \lambda x \text{ d'où } u_i \circ u_j(x) = u_j \circ u_i(x) = \lambda u_j(x)$$

donc $u_j(x)$ est un vecteur propre de u_i associé à la valeur propre λ et appartient au sous-espace propre correspondant.

On établit la co-diagonalisation des endomorphismes $\{u_i, i \in I\}$ par récurrence sur leur rang: si ce rang vaut 1, le résultat est immédiat ; supposons-le acquis jusqu'au rang n-1. Deux cas se présentent : soit tous les endomorphismes sont des homothéties, auquel cas ils sont diagonaux dans toute base de \mathbb{R}^n , soit l'un d'entre eux au moins (par exemple u_1) n'est pas une homothétie. Les sous-espaces propres de u_1 (de rang < n) forment une décomposition de \mathbb{R}^n ; les autres endomorphismes commutent avec u_1 donc, d'après le lemme, laissent stables ses sous-espaces propres : ils y induisent par restriction des endomorphismes qui sont co-diagonalisables (hypothèse de récurrence). En recollant les sous-bases diagonalisantes, on obtient une base de \mathbb{R}^n qui co-diagonalise tous les endomorphismes, C.Q.F.D.

Exemples

- Modèle mosaïque de corrélogramme $\rho(\mathbf{h})$:

$$\mathbf{P}(\mathbf{h}) = \rho(\mathbf{h}) \mathbf{I} + [1 - \rho(\mathbf{h})] \mathbf{A}$$

où $\mathbf{A} = \mathbf{P}(8)$ est la matrice de terme générique (ligne i , colonne j) $\sqrt{P_i P_j}$.

Il est aisé de voir que le critère est vérifié, car \mathbf{A} commute avec la matrice identité.

- Modèle des feuilles mortes colorées (Jeulin, 1997 ; Lantuéjoul, 2002)

Ce modèle est une généralisation de la mosaïque des feuilles mortes. On considère un ensemble de valuations $\{0, \dots, m\}$ auxquelles on associe des densités de points poissonniens $\{\theta_i, i = 0 \dots m\}$ et des covariogrammes géométriques $\{K_i(\mathbf{h}), i = 0 \dots m\}$. Entre les instants t et $t + dt$, m processus infinitésimaux (chacun associé à une valuation i , une forme de grain primaire $K_i(\mathbf{h})$ et une densité $\theta_i dt$) apparaissent et recouvrent partiellement le processus déjà construit ; la probabilité que deux processus infinitésimaux correspondant à des valuations différentes se chevauchent est nulle.

Si l'on inverse l'échelle de temps, ce modèle équivaut à partir de $t = 0$ et remonter dans le temps, en changeant l'opérateur "dernier grain observé" en "premier grain observé". La différence est que le processus n'est plus altéré dès lors que le domaine à simuler est entièrement recouvert ; à partir de cet instant, on obtient donc une image exacte du processus asymptotique.

En posant

$$\begin{cases} \forall i \in \{0, \dots, m\}, a_i = \theta_i K_i(\mathbf{0}) & a = \sum_{i=0}^m a_i \\ \forall i \in \{0, \dots, m\}, b_i(\mathbf{h}) = \theta_i K_i(\mathbf{h}) & b(\mathbf{h}) = \sum_{i=0}^m b_i(\mathbf{h}) \end{cases}$$

les lois marginales et bivariable sont définies par (Lantuéjoul, 2002) :

$$\begin{cases} \forall i \in \{0, \dots, m\}, P_i = \frac{a_i}{a} \\ \forall i, j \in \{0, \dots, m\}, P_{\mathbf{h}}(i, j) = \frac{a_i [a_j - b_j(\mathbf{h})] + a_j [a_i - b_i(\mathbf{h})]}{a [2a - b(\mathbf{h})]} + \delta_{ij} \frac{b_i(\mathbf{h})}{2a - b(\mathbf{h})} \end{cases}$$

Des essais numériques montrent que, en général, la propriété de commutation n'est pas vérifiée, sauf si tous les covariogrammes géométriques sont identiques (mosaïque des feuilles mortes).

Cas continu

Les résultats précédents peuvent être étendus au cadre continu en faisant appel à la théorie des espaces de Hilbert. La condition de commutation des couples d'opérateurs $\{\mathbf{P}(\mathbf{h}), \mathbf{P}(\mathbf{h}')\}$ est nécessaire mais n'est plus suffisante pour assurer la propriété isofactorielle, car leurs spectres peuvent ne pas être discrets (Matheron, 1989b). Il faut donc également s'assurer que les valeurs propres sont en nombre au plus dénombrable, ce qui peut se faire en exhibant une écriture factorielle (mais non *isofactorielle*) des lois bivariées.

Annexe C :

Simulations par chaînes de Markov

I. Principes de la simulation par itérations markoviennes

On considère un espace probabilisé défini par un triplet (Ω, \mathcal{A}, p) :

- Ω désigne l'ensemble des états possibles ou *univers*.
- \mathcal{A} est une tribu (σ -algèbre) sur Ω ; elle décrit l'ensemble des propositions ou *événements* que l'on peut énoncer sur les éléments de Ω .
- p est une distribution de probabilité (i.e. une mesure positive de poids total 1) définie sur \mathcal{A} .

Dans la suite de cette annexe, on supposera que Ω est un espace fini ou dénombrable, bien que la théorie markovienne peut être étendue à des espaces plus complexes (Tierney, 1994 ; Guyon, 1999), et que p ne s'annule pas sur Ω (au besoin, on élimine les états n'ayant aucune probabilité d'être atteints).

On se pose le problème de simuler la loi p sur Ω . Pour cela, on cherche à construire une suite $\{X_n, n \in \mathbb{N}\}$ d'états aléatoires dont la loi converge vers p lorsque n tend vers l'infini. Deux simplifications sont adoptées :

- 1) la construction d'un état X_n ne dépend que de l'état antérieur X_{n-1} (propriété *markovienne*) ;
- 2) la loi de transition d'un état à l'état suivant est stationnaire (propriété d'*homogénéité*).

Sous ces conditions, la suite $\{X_n, n \in \mathbb{N}\}$ est une chaîne de Markov stationnaire. Sa construction ne requiert que la donnée d'un état initial X_0 et d'un noyau de transition P :

$$P(x, y) = \text{Prob}[X_{n+1} = y \mid X_n = x] \quad (\text{C.1})$$

Le noyau de transition élevé à la puissance n représente la loi de transition d'ordre n :

$$P^n(x, y) = \text{Prob}[X_n = y \mid X_0 = x]$$

de sorte que la convergence de la chaîne vers la loi p équivaut à écrire :

$$\forall x, y \in \Omega, \lim_{n \rightarrow +\infty} P^n(x, y) = p(y) \quad (\text{C.2})$$

On dit que la loi p est la **limite ergodique** du noyau de transition P . La relation (C.2) implique certaines propriétés de la chaîne de Markov :

- a) **irréductibilité** : comme $p(y)$ est strictement positif, il existe un rang (dépendant de x et y) à partir duquel on a

$$P^n(x, y) > 0 \quad (\text{C.3})$$

Ceci signifie qu'il existe une probabilité non nulle de passer d'un état à tout autre, ou encore que tous les états communiquent. On dit que la chaîne de Markov est **irréductible**.

- b) **apériodicité** : la *période* d'une chaîne de Markov se définit comme le plus grand commun diviseur des éléments de $\{n \in \mathbb{N} \text{ tels que } P^n(x, x) > 0\}$. Cette définition est indépendante du choix de x dans Ω . Or, d'après l'équation (C.3), il existe un entier n tel que $P^n(x, x) > 0$ et $P^{n+1}(x, x) > 0$. La chaîne a donc une période égale à 1 ; on dit aussi qu'elle est **apériodique**.
- c) **invariance de p par le noyau de transition P** : à cause de la convergence de la chaîne, si X_n suit la loi p , alors X_{n+1} suit également la loi p (on dit que l'on est en régime stationnaire). En utilisant la relation (C.1), il vient :

$$\sum_{x \in \Omega} p(x) P(x, y) = p(y) \quad (\text{C.4})$$

Réciproquement, on montre que, si une chaîne de Markov homogène, irréductible et apériodique admet une loi p invariante par son noyau de transition, alors p est la limite ergodique du noyau de transition [Eq. (C.2)]. Par conséquent, la simulation de la loi p revient à trouver un noyau de transition vérifiant les trois propriétés ci-dessus. Avant de donner des algorithmes permettant de résoudre cette question, mentionnons une dernière propriété qui sera utile : la **réversibilité**. Une chaîne de Markov est p -réversible si

$$\forall x, y \in \Omega, p(x) P(x, y) = p(y) P(y, x) \quad (\text{C.5})$$

Il est aisé de voir que cette propriété entraîne l'invariance de p par le noyau de transition P [Eq. (C.4)].

Premier exemple : l'algorithme de Metropolis

L'algorithme de Metropolis permet de construire un noyau de transition P assurant l'ergodicité de la chaîne [Eq. (C.2)]. Il se décompose en deux étapes :

- 1) la chaîne se trouvant dans un état x , proposition d'un état candidat y selon une transition définie par un noyau irréductible $Q(x, y)$;
- 2) acceptation de l'état candidat avec une probabilité $a(x, y)$.

Cette procédure fournit comme noyau de transition final :

$$\forall x, y \in \Omega, P(x, y) = a(x, y) Q(x, y) \text{ si } x \neq y \text{ et } \forall x \in \Omega, P(x, x) = 1 - \sum_{z \in \Omega} P(x, z)$$

Plusieurs possibilités sont envisageables pour choisir le noyau $Q(x,y)$ (proposition de changement) et la fonction d'acceptation de changement $a(x,y)$:

- **dynamique de Metropolis** (Metropolis *et al.*, 1953) : soit Q un noyau symétrique, c'est-à-dire tel que $Q(x,y) = Q(y,x)$ pour tous les états x et y ; on pose :

$$\forall x, y \in \Omega, a(x, y) = \min \left(1, \frac{p(y)}{p(x)} \right)$$

- **dynamique de Hastings** (Hastings, 1970) :

$$\forall x, y \in \Omega, a(x, y) = \min \left(1, \frac{p(y) Q(y, x)}{p(x) Q(x, y)} \right)$$

- **dynamique de Barker** (Barker, 1965) :

$$\forall x, y \in \Omega, a(x, y) = \frac{p(y) Q(y, x)}{p(x) Q(x, y) + p(y) Q(y, x)}$$

Si Q est symétrique, cette fonction se réduit à : $\forall x, y \in \Omega, a(x, y) = \frac{p(y)}{p(x) + p(y)}$.

Dans tous les cas, la probabilité p n'a besoin d'être connue qu'à un facteur multiplicatif près.

Second exemple : l'échantillonneur de Gibbs

Supposons que l'on cherche à simuler un vecteur aléatoire \mathbf{X} dont les composantes, au nombre de n , prennent leur valeur sur un espace E et sont distribuées selon une loi conjointe p . L'échantillonneur de Gibbs consiste à balayer l'ensemble des composantes et à remplacer, au i -ème pas de ce balayage, la i -ème composante x_i par une valeur issue de sa distribution conditionnelle aux autres composantes (on dit que l'on "relaxe" la valeur x_i).

Une transition est définie par un balayage complet, soit :

$$\forall x, y \in \Omega, P(x, y) = \prod_{i=1}^n p(x_i | y_i, \dots, y_{i-1}, x_{i+1}, \dots, x_n)$$

Dans le conditionnement, les $i - 1$ premières composantes ont déjà été relaxées, contrairement aux $n - i$ dernières composantes. Là encore, la probabilité conjointe p n'a besoin d'être connue qu'à une constante multiplicative près.

On démontre que la chaîne ainsi construite converge vers la loi p à simuler (Geman and Geman, 1984). Plusieurs variantes ont été proposées, par exemple effectuer un balayage aléatoire à chaque itération au lieu d'un balayage systématique ; d'autres modes de balayage sont possibles, à condition que chaque composante soit relaxée un nombre infini de fois.

En géostatistique, cette algorithmes sert à simuler des vecteurs gaussiens tronqués (Freulon, 1992 ; Freulon and de Fouquet, 1993), des fonctions aléatoires de substitution ou un point uniforme sur un sous-ensemble de \mathbf{R}^d (Lantuéjoul, 2002), pour ne citer que quelques exemples.

Vitesse de convergence

Dans les applications, la chaîne de Markov doit être stoppée après un nombre fini d'itérations, d'où la question : comment s'assurer que la chaîne est entrée dans son régime stationnaire ? Bien que de nombreux travaux y aient été consacrés (Meyn and Tweedie, 1993 ; Tierney, 1994 ; Guyon, 1999 ; Galli and Gao, 2001), les résultats théoriques concernant la vitesse de convergence des chaînes de Markov sont souvent très spécifiques ou peu pratiques.

Pour contourner ces difficultés, d'autres approches permettent d'atteindre un état distribué selon p en nombre fini d'itérations : simulation "exacte" par composition inverse (Propp and Wilson, 1996) ou par composition directe (Diaconis and Freedman, 1999). Ces techniques récentes dépassent le cadre de cette présentation.

II. Conditionnement des simulations

La probabilité à simuler peut être altérée par une information partielle sur le phénomène étudié (prise d'échantillons, contraintes physiques, informations qualitatives, etc.). Dans ce cas, tous les états de Ω ne sont pas autorisés, mais seulement un sous-ensemble Ω_c , et l'on cherche à simuler la loi de probabilité conditionnelle p_c définie par

$$\forall A \in \mathcal{A}_c, p_c(A) = \frac{p(A)}{p(\Omega_c)} \text{ sous réserve que } p(\Omega_c) > 0$$

où \mathcal{A}_c est la tribu induite par A sur Ω_c .

Deux techniques de conditionnement sont envisageables : la restriction du noyau de transition et le recuit simulé.

II.1. Restriction du noyau de transition

Une première idée consiste à ne retenir que les transitions qui, à chaque étape, restent en accord avec les données conditionnantes, c'est-à-dire que l'on exclut toute excursion de la chaîne en dehors de Ω_c . On obtient une chaîne de Markov sur Ω_c de noyau de transition P_c défini par :

$$\forall x, y \in \Omega_c, P_c(x, y) = P(x, y) \text{ si } x \neq y \text{ et } \forall x \in \Omega_c, P_c(x, x) = P(x, x) + \sum_{z \in \Omega - \Omega_c} P(x, z)$$

Pour garantir que p_c est la limite ergodique de P_c [Eq. (C.2)] il suffit que P soit réversible [Eq. (C.5)] et que P_c soit p_c -irréductible. L'état initial doit appartenir à Ω_c . A titre d'exemple, cette méthode est utilisée pour construire des simulations conditionnelles du schéma booléen (Lantuéjoul, 2002).

II.2. Recuit simulé

On définit une *fonction objectif* (O) sur l'ensemble des états de Ω , s'annulant quand les conditions sur les données sont satisfaites et restant strictement positive dans le cas contraire :

$$\forall x \in \Omega_c, O(x) = 0 \text{ et } \forall x \in \Omega - \Omega_c, O(x) > 0.$$

En pratique, il est naturel de prendre une fonction qui décroît lorsque la réalisation se rapproche de l'ensemble des états autorisés, par exemple la somme des valeurs absolues des écarts entre les valeurs simulées aux sites de données et les valeurs réellement mesurées en ces sites.

On construit ensuite une chaîne de Markov dont la loi de transition s'effectue selon une étape de proposition suivie d'une étape d'acceptation ou rejet. L'état proposé (noté y) est généré selon un noyau de transition P réversible admettant la loi *a priori* à simuler (p) comme limite ergodique, puis accepté avec la probabilité

$$\min \left(1, \frac{\exp(-O(y)/t)}{\exp(-O(x)/t)} \right)$$

où x désigne l'état courant de la chaîne de Markov et t est un paramètre scalaire appelé *température*. Ce critère d'acceptation et de rejet correspond à la dynamique de Metropolis vue antérieurement.

Ainsi, toutes les transitions favorables (celles diminuant la fonction objectif) sont acceptées. Les transitions non favorables n'ont qu'une certaine probabilité d'être acceptées, dépendant de la valeur de la température ; plus celle-ci est faible, plus les transitions non favorables ont tendance à être rejetées. On montre que l'on obtient asymptotiquement un état de Ω_c suivant la loi p_c si on laisse évoluer cette chaîne de Markov tout en faisant tendre simultanément la température vers zéro (Kirkpatrick *et al.*, 1983 ; Boulanger, 1990 ; Hegstad *et al.*, 1994 ; Lantuéjoul, 2002).

Pour éviter que la fonction objectif soit piégée dans un minimum local, il est conseillé de partir d'une température élevée et de l'abaisser lentement, par exemple à la k -ième itération prendre une température inversement proportionnelle à $\ln(k+1)$ (Geman and Geman, 1984 ; Lantuéjoul, 2002).

L'état initial de la chaîne de Markov peut être un état quelconque de Ω . En pratique toutefois, il peut être avantageux de partir d'une simulation non conditionnelle.

Le recuit simulé présente plusieurs avantages sur la restriction du noyau de transition :

- il peut être généralisé au cas où $p(\Omega_c) = 0$, c'est-à-dire lorsque l'ensemble des états autorisés est de mesure nulle ;
- il peut incorporer des contraintes complexes, par l'intermédiaire de la fonction objectif ;
- il peut être étendu aux chaînes de Markov inhomogènes, pour lesquelles le noyau de transition dépend de l'itération (Gidas, 1985 ; Guyon, 1999). Le principe est que les transitions potentielles soient de plus en plus modérées à mesure que la température diminue (aux faibles températures, les grandes transitions perdent leur intérêt car elles sont presque toujours rejetées).

Exemple : mosaïque des feuilles mortes

Ce modèle (Matheron, 1968a ; Jeulin, 1997) possède deux descriptions équivalentes :

- (a) **modèle de recouvrement** : des points poissoniens ou "germes" apparaissent au cours du temps, en chacun desquels est implanté un objet valué ou *grain primaire* ; les grains nouveaux recouvrent les grains anciens. Au bout d'un temps suffisamment long, les grains apparents constituent une mosaïque des feuilles mortes ;
- (b) **modèle de germination** : on renverse l'échelle de temps, ce qui revient à considérer que les grains nouveaux sont cette fois cachés par les grains plus anciens.

En termes figurés, dans le premier cas, on regarde le tapis de feuilles mortes de dessus, tandis que dans le second cas, ce tapis est vu de dessous. Il est aisé d'obtenir des images exactes de la mosaïque des feuilles mortes en adoptant la seconde description (modèle de germination) : il suffit d'attendre que les grains primaires aient recouvert l'intégralité du domaine D à simuler¹, puisqu'alors l'ajout de nouveaux grains n'altère plus l'image obtenue. C'est cette description qui sera désormais retenue.

On se pose à présent le problème du conditionnement des simulations à un ensemble de données expérimentales. Pour simplifier, on suppose que les valuations du modèle sont en nombre fini, de sorte que $p(\Omega_c) > 0$, et l'on prend pour état initial une simulation non conditionnelle.

La fonction objectif et le schéma de refroidissement (évolution de la température en fonction des itérations) étant choisis, il reste à déterminer le noyau de transition réversible P . Pour cela, on regarde la mosaïque des feuilles mortes comme une population d'objets (les grains primaires valués), repérés par leurs germes qui constituent un ensemble de points poissonniens dans $D \times \mathbb{R}_+$. On partitionne D en sous-domaines $\{D_i, i = 1 \dots I\}$ et l'axe temporel en intervalles de la forme $[j \tau, (j+1) \tau[, j \in \mathbb{N}$ où τ est une date de référence. Le noyau de transition proposé est le suivant :

- 1) choisir un indice i au hasard dans $\{1, \dots, I\}$;
- 2) choisir un indice j dans \mathbb{N} , selon une loi géométrique ou de Poisson ;
- 3) remplacer tous les objets germés dans $D_i \times [j \tau, (j+1) \tau[$ par un ensemble d'objets générés selon la loi non conditionnelle.

On notera la parenté de cet algorithme avec celui proposé par Lantuéjoul (2002) pour simuler un modèle de dilution poissonnienne. Les valuations des feuilles mortes étant par hypothèse en nombre fini, il est également possible de partir d'un état respectant les données conditionnantes et de restreindre le noyau de transition précédent (i.e. exclure toutes les transitions qui interrompent le conditionnement des données), mais l'obtention de l'état initial s'avère délicate lorsque le nombre de valuations est très élevé.

En pratique, à la k -ième itération, seuls comptent les objets germés jusqu'à une date $T(k)$:

- $T(0)$ est la date d'apparition du dernier grain primaire lors de la simulation non conditionnelle qui constitue l'état initial.
- Si la transition proposée ne laisse aucune donnée conditionnante à découvert, $T(k) = T(k-1)$. Dans le cas contraire, il est nécessaire de compléter l'état courant par une simulation non conditionnelle jusqu'à ce que toutes les données soient à nouveau recouvertes ; on pose alors comme valeur de $T(k)$ la date d'apparition du dernier grain primaire.
- Il est inutile de tester la transition lorsque la valeur de l'indice j choisi est supérieure à $T(k-1) / \tau$. Formellement, il faut cependant considérer l'existence d'objets germés sur $]T(k-1), +\infty[$ suivant la loi non conditionnelle du modèle. Ces grains n'ont pas besoin d'être explicités tant qu'ils sont recouverts.

¹ En fait, le domaine D doit être le domaine à simuler dilaté du rayon maximal des grains primaires, car un grain germé en dehors du domaine à simuler peut recouvrir partiellement ce domaine.

Bibliographie

- [1] Abramowitz M. and Stegun I., 1972, *Handbook of Mathematical Functions*, Ninth printing, New York: Dover Publications Inc., 1046 p.
- [2] Alabert F., 1987, *Stochastic imaging of spatial distributions using hard and soft information*, Master's thesis, Stanford: Department of Applied Earth Sciences, Stanford University, 198 p.
- [3] Alfaro Sironvalle M., 1979, *Etude de la robustesse des simulations de fonctions aléatoires*, Thèse de docteur-ingénieur en sciences et techniques minières, Fontainebleau: Centre de Géostatistique, Ecole des Mines de Paris, 161 p.
- [4] Anderson T., 1958, *An introduction to multivariate statistical analysis*, New York: Wiley, 374 p.
- [5] Armstrong M., Galli A., Le Loc'h G., Geffroy F. and Eschard R., 2003, *Plurigaussian simulations in geosciences*, Berlin: Springer-Verlag, 160 p.
- [6] Armstrong M. and Matheron G., 1986a, Disjunctive kriging revisited, Part I, *Mathematical Geology*, **18** (8), pp. 711-728.
- [7] Armstrong M. and Matheron G., 1986b, Disjunctive kriging revisited, Part II, *Mathematical Geology*, **18** (8), pp. 729-742.
- [8] Barker A.A., 1965, Monte-Carlo calculations of radial distribution functions for a proton-electron plasma, *Aust. J. Phys.*, **18**, pp. 119-133.
- [9] Bouchind'homme J.-F., 1980, *Estimation de l'uranium récupérable sur les gisements sédimentaires, stratiformes, exploitables à ciel ouvert*, Thèse de docteur-ingénieur, Nancy: Institut National polytechnique de Lorraine, 185 p.
- [10] Boulanger F., 1990, *Modélisation et simulation de variables régionalisées, par des fonctions aléatoires stables*, Thèse de docteur en géostatistique, Fontainebleau: Centre de Géostatistique, Ecole des Mines de Paris, 385 p.
- [11] Bourgault G., 1997, Statistical declustering and convex estimation using determinant of redundant matrix, in: E.Y. Baafi and N.A. Schofield (Eds.), *Geostatistics Wollongong'96 (Quantitative geology and geostatistics)*, Dordrecht: Kluwer Academic Publishers, **1**, pp. 103-114.
- [12] Chauvet P., 1999, *Aide-mémoire de géostatistique linéaire*, Paris: Presses de l'Ecole des Mines de Paris, 370 p.
- [13] Chilès J.P., 1977, *Géostatistique des phénomènes non stationnaires (dans le plan)*, Thèse de docteur-ingénieur, Nancy: Université de Nancy I, 152 p.
- [14] Chilès J.P., 1995, Quelques méthodes de simulation de fonctions aléatoires intrinsèques, in: *Compte-rendu des journées de géostatistique 15-16 juin 1995, Cahiers de géostatistique, Fascicule 5*, Fontainebleau: Centre de Géostatistique, Ecole des Mines de Paris, pp. 97-112.

- [15] Chilès J.P. and Delfiner P., 1999, *Geostatistics: Modeling Spatial Uncertainty*, New York: Wiley, 696 p.
- [16] Chilès J.P. and Liao H.-T., 1993, Estimating the recoverable reserves of gold deposits: comparison between disjunctive kriging and indicator kriging, in: A. Soares (Ed.), *Geostatistics Tróia'92*, Dordrecht: Kluwer Academic Publishers, **2**, pp. 1053-1064.
- [17] Choquet G., 1954, Theory of capacities, *Annales de l'Institut Fourier*, Grenoble, **V**, pp. 131-295.
- [18] Christakos G., 1990, A Bayesian / maximum-entropy view of the spatial estimation problem, *Mathematical Geology*, **22** (7), pp. 763-777.
- [19] Cressie N. and Hawkins D.M., 1980, Robust estimation of the variogram, *Mathematical Geology*, **12** (2), pp. 115-125.
- [20] Davis M.W., 1987, Production of conditional simulations via the LU triangular decomposition of the covariance matrix, *Mathematical Geology*, **19** (2), pp. 91-98.
- [21] Demange C., Lajaunie C., Lantuéjoul C. and Rivoirard J., 1987, Global recoverable reserves: testing various changes of support models on uranium data, in: M. Armstrong and G. Matheron (Eds.), *Geostatistical Case Study*, Dordrecht, Reidel & Co, pp. 187-208.
- [22] Deutsch C.V., 1992, *Annealing techniques applied to reservoir modeling and the integration of geological and engineering (well test) data*, Ph. D. Thesis, Stanford: Stanford University.
- [23] Deutsch C.V. and Cockerham P., 1994, Practical considerations in the application of simulated annealing to stochastic simulation, *Mathematical Geology*, **26** (1), pp. 67-82.
- [24] Deutsch C.V. and Journel A.G., 1998, *GSLIB: Geostatistical Software Library and User's Guide*, Second edition, New York: Oxford University Press, 369 p.
- [25] Diaconis P. and Freedman D., 1999, Iterated random functions, *SIAM Review*, **41** (1), pp. 45-76.
- [26] Dimitrakopoulos R. and Dagbert M., 1993, Sequential modelling of relative indicator variables: dealing with multiple lithology types, in: A. Soares (Ed.), *Geostatistics Tróia'92*, Dordrecht: Kluwer Academic Publishers, **1**, pp. 413-424.
- [27] Dubrule O., 1983, Cross-validation of kriging in a unique neighbourhood, *Mathematical Geology*, **15** (6), pp. 687-699.
- [28] Emery X., 2002, Conditional simulation of non-gaussian random functions, *Mathematical Geology*, **34** (1), pp. 79-100.
- [29] Emery X., 2003, Disjunctive kriging with hard and imprecise data, *Mathematical Geology*, **35** (6), pp. 699-718.
- [30] Feller W., 1966, *An introduction to probability theory and its applications*, Second edition, Volume 2, New York: Wiley, 626 p.
- [31] de Fouquet C., Beucher H., Galli A. and Ravenne C., 1989, Conditional simulation of random sets. Application to an argileous sandstone reservoir, in: M. Armstrong (Ed.), *Geostatistics, Proceedings of the 3rd International Geostatistics Congress, September 5-9, 1988, Avignon, France*, Dordrecht: Kluwer Academic Publishers, pp. 517-530.

- [32] Freulon X. et de Fouquet C., 1991, Remarques sur la pratique des bandes tournantes à trois dimensions, in: *Cahiers de géostatistique, Fascicule 1*, Fontainebleau: Centre de Géostatistique, Ecole des Mines de Paris, pp. 101-117.
- [33] Freulon X., 1992, Conditionnement du modèle gaussien par des inégalités ou des randomisées, Thèse de docteur en géostatistique, Fontainebleau: Centre de Géostatistique, Ecole des Mines de Paris, 168 p.
- [34] Freulon X., 1994, Conditional simulation of a Gaussian random vector with nonlinear and/or noisy observations, in: M. Armstrong and P.A. Dowd (Eds.), *Geostatistical simulations*, Dordrecht: Kluwer Academic Publishers, pp. 57-71.
- [35] Freulon X. and de Fouquet C., 1993, Conditioning a gaussian model with inequalities, in: A. Soares (Ed.), *Geostatistics Tróia '92*, Dordrecht: Kluwer Academic Publishers, **1**, pp. 201-212.
- [36] Galli A. and Gao H., 2001, Rate of convergence of the Gibbs sampler in the Gaussian case, *Mathematical Geology*, **33** (6), pp. 653-677.
- [37] Geman S. and Geman D., 1984, Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images, *I.E.E.E. Transactions on Pattern Analysis and Machine Intelligence*, **6**, pp. 721-741.
- [38] Gidas B., 1985, Non-stationary Markov chains and convergence of the annealing algorithm, *Journal of Statistical Physics*, **39**, pp. 73-131.
- [39] Gilbert E.N., 1962, Random subdivisions of space into crystals, *Annals of mathematical statistics*, **33** (958).
- [40] Glaken I.M., 1997, Change of support and use of economic parameters for block selection, in: E.Y. Baafi and N.A. Schofield (Eds.), *Geostatistics Wollongong'96 (Quantitative geology and geostatistics)*, Dordrecht: Kluwer Academic Publishers, **2**, pp. 811-821.
- [41] Goovaerts P., 1997, *Geostatistics for natural resources evaluation*, New York: Oxford University Press, 480 p.
- [42] Guibal D., 1972, *Simulations de schémas intrinsèques*, rapport N-291, Fontainebleau: Centre de Géostatistique, Ecole des Mines de Paris, 38 p.
- [43] Guiblin P., Rivoirard J. et Simmonds E.J., 1995, Analyse structurale de données à distribution dissymétrique, in: *Compte-rendu des journées de géostatistique 15-16 juin 1995, Cahiers de géostatistique, Fascicule 5*, Fontainebleau: Centre de Géostatistique, Ecole des Mines de Paris, pp. 137-159.
- [44] Guyon X., 1999, *Méthodes numériques par chaînes de Markov*, Ecole d'été de Mathématiques, Mérida, Venezuela: Université de Paris I et Universidad de los Andes, 94 p.
- [45] Haas A., Matheron G. et Serra J., 1967, Morphologie mathématique et granulométrie en place, *Annale des Mines*, **11**, pp. 736-753 et **12**, pp. 767-782.
- [46] Hastings W.K., 1970, Monte-Carlo sampling methods using Markov chains and their applications, *Biometrika*, **57**, pp. 97-109.
- [47] Hausdorff F., 1921, Summationsmethoden und Momentfolgen, *Mathematische Zeitschrift*, **9**, pp. 74-109 & pp. 280-299.

- [48] Hegstad B.K., Omre H., Tjelmeland H. and Tyler K., 1994, Stochastic simulation and conditioning by annealing in reservoir description, in: M. Armstrong and P.A. Dowd (Eds.), *Geostatistical simulations*, Dordrecht: Kluwer Academic Publishers, pp. 43-55.
- [49] Hu L.-Y., 1988, *Mise en œuvre du modèle gamma pour l'estimation des distributions spatiales*, Thèse de docteur en géostatistique, Fontainebleau: Centre de Géostatistique, Ecole des Mines de Paris, 142 p.
- [50] Hu L.-Y. et Lantuéjoul C., 1988, Recherche d'une fonction d'anamorphose pour la mise en œuvre du krigeage disjonctif isofactoriel gamma, *Sciences de la Terre, série informatique géologique*, **28**, Nancy, pp. 145-173.
- [51] Isaaks E.H., 1984, *Risk qualifed mappings for hazardous waste sites: a case study in distribution free geostatistics*, Master's thesis, Stanford: Department of Applied Earth Sciences, Stanford University, 85 p.
- [52] Isaaks E.H., 1990, *The application of Monte-Carlo methods to the analysis of spatially correlated data*, Ph.D. Thesis, Stanford: Stanford University, 213 p.
- [53] Jeulin D., 1980, Multi-component random models for the description of complex microstructures, *Mikroskopie*, **37**, pp. 130-137.
- [54] Jeulin D., 1989, Sequential random function models, in: M. Armstrong (Ed.), *Geostatistics, Proceedings of the 3rd International Geostatistics Congress, September 5-9, 1988, Avignon, France*, Dordrecht: Kluwer Academic Publishers, **1**, pp. 189-200.
- [55] Jeulin D., 1997, Dead leaves models: from space tessellation to random functions, in: D. Jeulin (Ed.), *Advances in theory and applications of random sets, Proceedings of the International Symposium, October 9-11, 1996, Fontainebleau, France*, Singapore: World Scientific Publishing Company, pp. 137-156.
- [56] Jeulin D. and Jeulin P., 1981, Synthesis of rough surfaces of random morphological functions, *Stereologica Jugoslavia*, **3** (suppl. 1), pp. 239-246.
- [57] Jones D., 1979, *Elementary information theory*, Oxford: Clarendon Press, 182 p.
- [58] Journel A.G., 1974, *Simulation conditionnelle – théorie et pratique*, Thèse de docteur-ingénieur, Fontainebleau: Centre de Géostatistique, Ecole des Mines de Paris, 110 p.
- [59] Journel A.G., 1983, Non parametric estimation of spatial distributions, *Mathematical Geology*, **15** (3), pp. 445-468.
- [60] Journel A.G., 1984, The place of non-parametric geostatistics, in: G. Verly, M. David, A.G. Journel and A. Maréchal (Eds.), *Geostatistics for natural resources characterization*, Dordrecht: NATO ASI Series C-122, Part 1, Reidel & Co., pp. 307-335.
- [61] Journel A.G., 1985, Recoverable reserves estimation – the geostatistical approach, *Mining Engineering*, **37** (6), pp. 563-568.
- [62] Journel A.G., 1987, *Geostatistics for the environmental sciences*, Stanford: Department of Applied Earth Sciences, Stanford University, 135 p.
- [63] Journel A.G., 1989, *Fundamentals of geostatistics in five lessons*, Volume 8, Short course in geology, Washington D.C.: American geophysical union, 90 p.

- [64] Journel A.G., 1993, Geostatistics: roadblocks and challenges, in: A. Soares (Ed.), *Geostatistics Tróia'92*, Dordrecht: Kluwer Academic Publishers, **1**, pp. 213-224.
- [65] Journel A.G. and Huijbregts C. J., 1978, *Mining geostatistics*, London: Academic Press, 600 p.
- [66] Journel A.G. and Isaaks E.H., 1984, Conditional indicator simulation: application to a Saskatchewan uranium deposit, *Mathematical Geology*, **16** (7), pp. 685-718.
- [67] Kendall D.G., 1974, Foundations of a theory of random sets, in: E.F. Harding and D.G. Kendall (Eds.), *Stochastic Geometry*, London: Wiley, pp. 322-376.
- [68] Kirkpatrick S., Gelatt C.D. and Vecchi M.P., 1983, Optimization by simulated annealing, *Science*, **220** (4598), pp. 671-680.
- [69] Lajaunie C., 1993, *L'estimation géostatistique non linéaire*, cours C-152, Fontainebleau: Centre de Géostatistique, Ecole des Mines de Paris, 93 p.
- [70] Lajaunie C. and Lantuéjoul C., 1989, Setting up the general methodology for discrete isofactorial models, in: M. Armstrong (Ed.), *Geostatistics, Proceedings of the 3rd International Geostatistics Congress, September 5-9, 1988, Avignon, France*, Dordrecht: Kluwer Academic Publishers, **1**, pp. 323-334.
- [71] Lantuéjoul C., 1977, *Sur le modèle de Johnson-Mehl généralisé*, rapport N-514, Fontainebleau: Centre de Géostatistique et de Morphologie Mathématique, Ecole des Mines de Paris, 15 p.
- [72] Lantuéjoul C., 1984, Trois modèles de changement de support à partir d'une distribution ponctuelle gamma, *Sciences de la Terre, série informatique géologique*, **21**, Nancy, pp. 139-173.
- [73] Lantuéjoul C., 1987, *Simulation de lois monovariées*, cours C-125, Fontainebleau: Centre de Géostatistique et de Morphologie Mathématique, Ecole des Mines de Paris, 32 p.
- [74] Lantuéjoul C., 1990, *Cours de sélectivité*, cours C-140, Fontainebleau: Centre de Géostatistique, Ecole des Mines de Paris, 72 p.
- [75] Lantuéjoul C., 1991, Ergodicity and integral range, *Journal of Microscopy*, **161** (3), pp. 387-404.
- [76] Lantuéjoul C., 1993, Substitution random functions, in: A. Soares (Ed.), *Geostatistics Tróia'92*, Dordrecht: Kluwer Academic Publishers, **1**, pp. 37-48.
- [77] Lantuéjoul C., 1994, Non conditional simulation of stationary isotropic multigaussian random functions, in: M. Armstrong and P.A. Dowd (Eds.), *Geostatistical simulations*, Dordrecht: Kluwer Academic Publishers, pp. 147-177.
- [78] Lantuéjoul C., 2002, *Geostatistical simulation, models and algorithms*, Berlin: Springer-Verlag, 256 p.
- [79] Lantuéjoul C. et Rivoirard J., 1984, *Une méthode de détermination d'anamorphose*, rapport N-916, Fontainebleau: Centre de Géostatistique et de Morphologie Mathématique, Ecole des Mines de Paris, 10 p.
- [80] Le Loc'h G., Beucher H., Galli A., Doligez B., and Heresim Group, 1994, Improvement in the truncated Gaussian method: combining several gaussian functions, in: *Proc. ECMOR IV, Fourth European Conference on the Mathematics of Oil Recovery*, Røros, Norway, 7-10 June 1994, Topic B: Heterogeneity Description and Assessment of Uncertainty, 13 p.

- [81] Maréchal A., 1978, *Gaussian anamorphosis models*, cours C-72, Fontainebleau: Centre de Morphologie Mathématique, Ecole des Mines de Paris, 22 p.
- [82] Matheron G., 1965, *Les variables régionalisées et leur estimation, une application de la théorie des fonctions aléatoires aux sciences de la nature*, Paris: Masson, 305 p.
- [83] Matheron G., 1967, *Eléments pour une théorie des milieux poreux*, Paris: Masson, 166 p.
- [84] Matheron G., 1968a, *Schéma booléen séquentiel de partition aléatoire*, rapport N-83, Fontainebleau: Centre de Morphologie Mathématique, Ecole des Mines de Paris, 16 p.
- [85] Matheron G., 1968b, *Processus de renouvellement et pseudo-périodicités*, rapport N-71, Fontainebleau: Centre de Morphologie Mathématique, Ecole des Mines de Paris, 4 p.
- [86] Matheron G., 1969a, *Théorie des ensembles aléatoires*, Les cahiers du Centre de Morphologie Mathématique de Fontainebleau, Fascicule 4, Fontainebleau: Centre de Morphologie Mathématique, Ecole des Mines de Paris, 54 p.
- [87] Matheron G., 1969b, *Cours de processus stochastiques*, cours C-37, Fontainebleau: Centre de Géostatistique et de Morphologie Mathématique, Ecole des Mines de Paris, 139 p.
- [88] Matheron G., 1970, *La théorie des variables régionalisées et ses applications*, Les cahiers du Centre de Morphologie Mathématique de Fontainebleau, Fascicule 5, Fontainebleau: Ecole des Mines de Paris, 212 p.
- [89] Matheron G., 1972a, *Quelques aspects de la montée*, rapport N-271, Fontainebleau: Centre de Géostatistique et de Morphologie Mathématique, Ecole des Mines de Paris, 36 p.
- [90] Matheron G., 1972b, *The turning bands: a method for simulating random functions in R^n* , rapport N-303, Fontainebleau: Centre de Géostatistique et de Morphologie Mathématique, Ecole des Mines de Paris.
- [91] Matheron G., 1972c, *Ensembles fermés aléatoires, ensembles semi-markoviens et polyèdres poissoniens*, *Advances in Applied Probability*, **4** (3), pp. 508-541.
- [92] Matheron G., 1973a, *The intrinsic random functions and their applications*, *Advances in Applied Probability*, **5**, pp. 439-468.
- [93] Matheron G., 1973b, *Le krigeage disjonctif*, rapport N-360, Fontainebleau: Centre de Géostatistique, Ecole des Mines de Paris, 40 p.
- [94] Matheron G., 1974, *Les fonctions de transfert des petits panneaux*, note n°127, rapport N-395, Fontainebleau: Centre de Morphologie Mathématique, Ecole des Mines de Paris, 73 p.
- [95] Matheron G., 1975a, *Random sets and integral geometry*, New York: Wiley, 261 p.
- [96] Matheron G., 1975b, *Compléments sur les modèles isofactoriels*, rapport N-432, Fontainebleau: Centre de Géostatistique, Ecole des Mines de Paris, 20 p.
- [97] Matheron G., 1975c, *Modèles isofactoriels discrets et modèle de Walsh*, note n°133, rapport N-449, Fontainebleau: Centre de Géostatistique, Ecole des Mines de Paris, 24 p.
- [98] Matheron G., 1976, *A simple substitute for conditional expectation: the disjunctive kriging*, in: M. Guarascio, M. David and Ch. J. Huijbregts (Eds.), *Advanced geostatistics in the mining industry*, Dordrecht: NATO ASI Series C-24, Reidel & Co., pp. 221-236.

- [99] Matheron G., 1977, *Peut-on imposer des conditions d'universalité au krigeage disjonctif ?*, note n°135, rapport N-539, Fontainebleau: Centre de Géostatistique, Ecole des Mines de Paris, 24 p.
- [100] Matheron G., 1978a, *Le krigeage disjonctif et le paramétrage local des réserves*, cours C-76, Fontainebleau: Centre de Géostatistique, Ecole des Mines de Paris, 22 p.
- [101] Matheron G., 1978b, *Estimer et choisir: un essai sur les probabilités appliquées*, Les cahiers du Centre de Morphologie Mathématique de Fontainebleau, Fascicule 7, Fontainebleau: Centre de Morphologie Mathématique, Ecole des Mines de Paris, 175 p.
- [102] Matheron G., 1980a, *Modèles isofactoriels pour l'effet zéro*, rapport N-659, Fontainebleau: Centre de Géostatistique et de Morphologie Mathématique, Ecole des Mines de Paris, 33 p.
- [103] Matheron G., 1980b, *Splines et krigeage: leur équivalence formelle*, rapport N-667, Fontainebleau: Centre de Géostatistique et de Morphologie Mathématique, Ecole des Mines de Paris, 26 p.
- [104] Matheron G., 1981a, *La sélectivité des distributions*, rapport N-686, Fontainebleau: Centre de Géostatistique, Ecole des Mines de Paris, 44 p.
- [105] Matheron G., 1981b, *Remarques sur le krigeage et son dual*, rapport N-695, Fontainebleau: Centre de Géostatistique et de Morphologie Mathématique, Ecole des Mines de Paris, 36 p.
- [106] Matheron G., 1982, *La déstructuration des hautes teneurs et le krigeage des indicatrices*, rapport N-761, Fontainebleau: Centre de Géostatistique et de Morphologie Mathématique, Ecole des Mines de Paris, 33 p.
- [107] Matheron G., 1983, *Modèles isofactoriels et changement de support*, rapport N-821, Fontainebleau: Centre de Géostatistique, Ecole des Mines de Paris, 46 p.
- [108] Matheron G., 1984a, Une méthodologie générale pour les modèles isofactoriels discrets, *Sciences de la Terre, série informatique géologique*, **21**, Nancy, pp. 1-64.
- [109] Matheron G., 1984b, Isofactorial models and change of support, in: G. Verly, M. David, A.G. Journel and A. Maréchal (Eds.), *Geostatistics for natural resources characterization*, Dordrecht: NATO ASI Series C-122, Part 1, Reidel & Co., pp. 449-467.
- [110] Matheron G., 1984c, *Changement de support en modèle mosaïque*, rapport N-880, Fontainebleau: Centre de Géostatistique, Ecole des Mines de Paris, 21 p.
- [111] Matheron G., 1985a, Change of support for diffusion-type random functions, *Mathematical Geology*, 17 (2), pp. 137-166.
- [112] Matheron G., 1985b, Comparaison de quelques distributions du point de vue de la sélectivité, *Sciences de la Terre, série informatique géologique*, **24**, Nancy, pp. 1-21.
- [113] Matheron G., 1986, *Sur la positivité des poids de krigeage*, rapport N-30/86/G, Fontainebleau: Centre de Géostatistique et de Morphologie Mathématique, Ecole des Mines de Paris, 44 p.
- [114] Matheron G., 1987, Suffit-il, pour une covariance, d'être de type positif ?, *Sciences de la Terre, série informatique géologique*, **26**, Nancy, pp. 51-66.
- [115] Matheron G., 1988, Simulation de fonctions aléatoires admettant un variogramme concave donné, *Sciences de la Terre, série informatique géologique*, **28**, Nancy, pp. 195-212.

- [116] Matheron G., 1989a, The internal consistency of models in geostatistics, in: M. Armstrong (Ed.), *Geostatistics, Proceedings of the 3^d International Geostatistics Congress, September 5-9, 1988, Avignon, France*, Dordrecht: Kluwer Academic Publishers, **1**, pp. 21-38.
- [117] Matheron G., 1989b, Two classes of isofactorial models, in: M. Armstrong (Ed.), *Geostatistics, Proceedings of the 3^d International Geostatistics Congress, September 5-9, 1988, Avignon, France*, Dordrecht: Kluwer Academic Publishers, **1**, pp. 309-322.
- [118] Matheron G., 1993, Une conjecture sur la covariance d'un ensemble aléatoire, in: *Cahiers de géostatistique, Fascicule 3*, Fontainebleau: Centre de Géostatistique, Ecole des Mines de Paris, pp. 107-113.
- [119] Matheron G., Beucher H., de Fouquet C., Galli A., Guérillot D. and Ravenne C., 1987, Conditional simulation of a fluvio-deltaic reservoir, in: *62nd Annual Technical Conference and Exhibition of SPE*, SPE paper 16753, Dallas, pp. 591-599.
- [120] Metropolis N., Rosenbluth A.W., Teller A.H. and Teller E., 1953, Equation of state calculation by fast computing machines, *Journal of Chemical Physics*, **21** (6), pp. 1087-1092.
- [121] Meyn S.P. and Tweedie R.L., 1993, *Markov chains and stochastic stability*, London: Springer-Verlag.
- [122] Miles R.E., 1961, *Random polytopes: the generalization to n dimensions of the intervals of a Poisson point process*, Ph.D. Thesis, Cambridge: University of Cambridge.
- [123] Miles R.E., 1974, A synopsis of Poisson flats in Euclidean spaces, in: E.F. Harding & D.G. Kendall (Eds.), *Stochastic Geometry*, London: Wiley, pp. 202-227.
- [124] Orfeuil J.P., 1972, *Simulation du Wiener-Lévy et de ses intégrales*, rapport N-290, Fontainebleau: Centre de Géostatistique et de Morphologie Mathématique, Ecole des Mines de Paris, 57 p.
- [125] Pólya G., 1918, Über die Nullstellen gewisser ganzer Funktionen, *Mathematische Zeitschrift*, **2**, pp. 352-383.
- [126] Propp J.P. and Wilson D.B., 1996, Exact sampling with coupled Markov chains and applications to statistical mechanics, *Random structures and algorithms*, **9**, pp. 223-252.
- [127] Rivoirard J., 1984, Le comportement des poids de krigeage, Thèse de docteur-ingénieur en sciences et techniques minières, Fontainebleau: Centre de Géostatistique, Ecole des Mines de Paris, 72 p.
- [128] Rivoirard J., 1988, Modèles à résidus d'indicatrices autokrigeables, *Etudes Géostatistiques V – séminaire C.F.S.G. sur la géostatistique, 15-16 juin 1987, Fontainebleau, Sciences de la Terre, série informatique géologique*, **28**, Nancy, pp. 303-326.
- [129] Rivoirard J., 1989, Models with orthogonal indicator residuals, in: M. Armstrong (Ed.), *Geostatistics, Proceedings of the 3^d International Geostatistics Congress, September 5-9, 1988, Avignon, France*, Dordrecht: Kluwer Academic Publishers, **1**, pp. 91-107.
- [130] Rivoirard J., 1993, Relations between the indicators related to a regionalized variable, in: A. Soares (Ed.), *Geostatistics Tróia'92*, Dordrecht: Kluwer Academic Publishers, **1**, pp. 273-284.
- [131] Rivoirard J., 1994, *Introduction to disjunctive kriging and nonlinear geostatistics*, Oxford: Clarendon Press, 181 p.

- [132] Schmitt M., 2001, Random sets, sampling and connectivity, in: W.J. Kleingeld and D.G. Krige (Eds.), *Geostatistics 2000, Proceedings of the 6th International Geostatistics Congress*, Cape Town, South Africa, pp. 205-213.
- [133] Serra J., 1968, *Fonctions aléatoires de dilution*, rapport C-12, Fontainebleau: Centre de Morphologie Mathématique, Ecole des Mines de Paris.
- [134] Serra J., 1982, *Image analysis and mathematical morphology*, London: Academic Press.
- [135] Serra J., 1988, Boolean random functions, in: J. Serra (Ed.), *Image analysis and mathematical morphology, Theoretical advances*, London: Academic Press, pp. 317-342.
- [136] Sichel H.S., 1971, On a family of discrete distributions particularly suited to represent long-tailed frequency data, in: *Proceedings of the Third Symposium on Mathematical Statistics*, Pretoria: S.A. C.S.I.R., pp. 1-97.
- [137] Stoyan D., Kendall W.S. and Mecke J., 1996, *Stochastic geometry and its applications*, 2nd Edition, New York: Wiley, 456 p.
- [138] Subramanyam A. and Pandalai H.S., 2001, A characterization of symmetric isofactorial models, *Mathematical Geology*, **33** (1), pp. 103-114.
- [139] Sullivan J.A., 1984, *Non-parametric estimation of spatial distributions*, Ph.D. Thesis, Stanford: Department of Applied Earth Sciences, Stanford University, 367 p.
- [140] Tierney L., 1994, Markov chains for exploring posterior distributions, *The Annals of statistics*, **22** (4), pp. 1701-1762.
- [141] Verly G., 1983, The Multigaussian approach and its applications to the estimation of local reserves, *Mathematical Geology*, **15** (2), pp. 259-286.
- [142] von Neumann J., 1951, Various techniques used in connection with random digits, *United States National Bureau of Standards, Applied Mathematical Series*, **12**, pp. 36-38.
- [143] Voss R.F., 1985, Random fractal forgeries, in: R.A. Earnshaw (Ed.), *Fundamental algorithms for computer graphics*, Berlin Heidelberg: Springer-Verlag, pp. 805-835.
- [144] Wackernagel H., 1998, *Multivariate geostatistics: an introduction with applications*, Berlin: Springer-Verlag, 291 p.
- [145] Widder D.V., 1941, *The Laplace transform*, Princetown: Princetown University Press.
- [146] Xiao H., 1985, *A description of the behavior of indicator variograms for a bivariate normal distribution*, Master's Thesis, Stanford: Stanford University, 51 p.
- [147] Yaglom A.M., 1987, *Correlation theory of stationary and related random functions*, New York: Springer Series in Statistics.
- [148] Zhu H. and Journel A.G., 1993, Formatting and integrating soft data: stochastic imaging via the Markov-Bayes algorithm, in: A. Soares (Ed.), *Geostatistics Tróia'92*, Dordrecht: Kluwer Academic Publishers, **1**, pp. 1-12.

Résumé

Cette thèse vise à développer des algorithmes de simulation conditionnelle de fonctions aléatoires à lois bivariées isofactorielles, c'est-à-dire telles qu'il existe une base de fonctions (*facteurs*) sans corrélations spatiales croisées. Les lois isofactorielles réalisent un compromis entre deux exigences souvent difficiles à concilier : d'une part, elles forment une vaste classe de modèles et permettent de s'adapter à la description de nombreux phénomènes ; d'autre part, l'inférence statistique du modèle repose sur un faible nombre de paramètres.

La première partie de la thèse met en relief les limitations et les approximations commises par une technique réputée "passe-partout" : l'algorithme *séquentiel*, qui consiste à simuler les sites de l'espace les uns après les autres dans un ordre aléatoire, en estimant en chaque site la distribution de probabilité conditionnelle de la valeur inconnue par un krigeage d'indicatrices ou par un krigeage disjonctif.

La seconde partie propose des modèles nouveaux et des algorithmes adaptés à leur simulation et au conditionnement à des données expérimentales. Plusieurs outils structuraux sont introduits et étudiés pour mener l'inférence des lois bivariées, notamment le madogramme, les variogrammes d'ordre inférieur à deux et les variogrammes d'indicatrices. Les concepts et méthodes sont finalement appliqués à un ensemble de données minières (accumulations en or et argent dans un gisement chilien en veine) caractérisées par une très forte dissymétrie des histogrammes.

Abstract

This thesis aims at developing conditional simulation algorithms for random fields with bivariate isofactorial distributions, i.e. such that there exists a basis of functions (*factors*) with no spatial cross-correlation. Isofactorial distributions fulfill two constraints that are often difficult to conciliate: on the one hand, they constitute a wide class of models and adapt to the description of numerous phenomena; on the other hand, their statistical inference relies only on a few parameters.

The first part of the thesis highlights the limits and approximations of an "all-purpose" technique: the *sequential* algorithm. It consists in simulating the field locations one after the other, according to a random sequence, by estimating at each location the conditional distribution of the unknown value by either an indicator or a disjunctive kriging.

In the second part, new models and algorithms are proposed in order to achieve the simulation and the conditioning to experimental values. Several structural tools are introduced and studied to enable the inference of bivariate distributions, in particular the madogram, variograms of order lower than two and indicator variograms. The concepts and methods are finally applied to a mining dataset (gold and silver accumulations in a Chilean vein-type deposit) characterized by highly skewed histograms.