

8. Simulations

8.1 Introduction

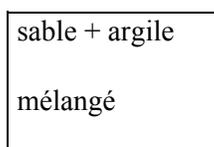
En géostatistique, le domaine de recherche le plus actif au cours des 10 dernières années fut sans contredit celui des simulations géostatistiques. Les simulations sont nécessaires pour tout problème impliquant des transformations non-linéaires des variables mesurées. On reconnaît 3 grands types d'application:

- Les problèmes non-linéaire de changement d'échelle. La conductivité hydraulique ou la transmissivité en sont des exemples courants. Ici l'échelle varie en fonction du test utilisé et en fonction de décisions arbitraires reliées à l'utilisation de simulateurs d'écoulements (taille des éléments, taille des cellules, conditions frontières).
- Les problèmes impliquant des relations non-linéaires entre différentes variables. La charge hydraulique et la transmissivité en sont un exemple; le champ gravimétrique et les densités en sont un autre exemple.
- Les problèmes impliquant des fonctions de transfert non-linéaires ou linéaires mais dont les équations analytiques sont inextricables. Ex. design de pile d'homogénéisation; variabilité des teneurs au concentrateur pour un schéma X d'exploitation; temps de transport d'un contaminant, etc.

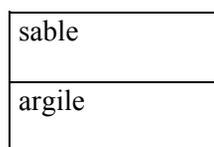
Pour certains problèmes, il n'est pas suffisant de connaître la concentration des teneurs dans un domaine donné pour pouvoir prédire le comportement de ce domaine. La distribution spatiale des valeurs ponctuelles à l'intérieur du domaine donné est tout aussi importante que la valeur moyenne.

Ex.

- Pour connaître le profit généré par un bloc, il est suffisant de connaître la teneur moyenne du bloc et les coûts pour exploiter ce bloc. Il s'agit là d'un problème où la distribution exacte du minéral économique dans le bloc n'a pas d'importance.
- Pour déterminer, la conductivité hydraulique d'un bloc constitué de deux phases (argile et sable), il n'est pas suffisant de connaître la proportion de sable. Il faut connaître comment ce sable se distribue à l'intérieur du bloc.



bloc 1



bloc 2



bloc3

Supposons que la conductivité hydraulique du sable est de 1×10^{-3} cm/s et celle de l'argile de 1×10^{-7} cm/s. On obtiendrait alors les résultats suivants:

$$\text{bloc 1: } k_{\text{hor}} = k_{\text{vert}} = 1 \times 10^{-5} \text{ cm/s (moyenne géométrique: } m = \left(\prod_{i=1}^n Z_i \right)^{1/n} \text{)}$$

$$\text{bloc 2: } k_{\text{hor}} = 0.5 \times 10^{-3} \text{ cm/s (moyenne arithmétique)}$$

$$k_{\text{vert}} = 2 \times 10^{-7} \text{ cm/s (moyenne harmonique } m = \left(\frac{1}{n} \sum_{i=1}^n \frac{1}{Z_i} \right)^{-1} \text{)}$$

bloc 3: $k_{\text{hor}} = k_{\text{vert}} = 2 \times 10^{-7}$ cm/s (si on suppose que l'écoulement est strictement horizontal ou strictement vertical; en réalité, une connexion s'effectuera probablement par les sommets et le résultat dépendra de l'importance relative de cette connexion).

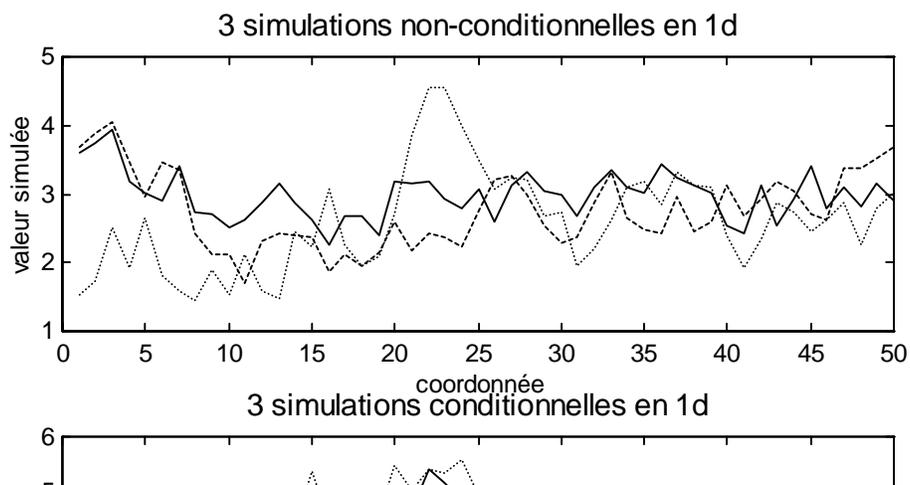
Note: Pour les 3 blocs ci-haut, on a exactement la même proportion de sable et d'argile, pourtant les valeurs de conductivité varient considérablement et même des anisotropies du tenseur de conductivité apparaissent. Toutefois, les 3 blocs ci-haut ne présenteraient pas le même variogramme. Celui de gauche serait le moins structuré, celui du centre serait anisotrope et celui de droite montrerait une structure spatiale isotrope.

8.1 Objectifs des simulations et type de simulation

Une méthode de simulation *non-conditionnelle* vise à produire des champs montrant la même structure spatiale (variogramme) et le même histogramme que ceux inférés à partir des données observées. Lorsqu'on applique l'algorithme de simulation correspondant à une méthode donnée, on obtient une *réalisation*. On peut appliquer l'algorithme autant de fois que l'on veut et obtenir ainsi plusieurs réalisations. Les réalisations sont indépendantes les unes des autres.

Une méthode de simulation *conditionnelle* vise les mêmes objectifs que la simulation non-conditionnelle mais, en plus, elle doit assurer que chaque réalisation soit compatible avec les valeurs observées aux points échantillons. Ainsi, si on simule une valeur en un point coïncidant avec une observation, il faut que la valeur simulée soit égale à la valeur réellement observée.

Exemples de simulations en 1D



Notes:

- Pour répondre à des problèmes concrets, les simulations conditionnelles sont habituellement plus appropriées puisqu'elles utilisent l'information connue. Les simulations non-conditionnelles sont surtout utilisées pour tester des méthodes, des modèles, des algorithmes.
- Les méthodes de simulation permettent de reproduire les statistiques d'ordre 1 (histogramme) et 2 (variogramme). Les statistiques d'ordre supérieures ne peuvent généralement pas être spécifiées et dépendent de la méthode de simulation choisie et, dans le cas des simulations conditionnelles, des données conditionnantes. Une exception à cette règle est la méthode du *recuit simulé* qui permet d'inclure des statistiques d'ordre supérieur.
- Les données conditionnantes ont une forte influence sur les caractéristiques statistiques des réalisations même pour les deux premiers ordres. On ne peut pas, par exemple, simuler un modèle de variogramme qui serait totalement incompatible avec la structure spatiale des données conditionnantes. De même, les données conditionnantes confèrent une certaine robustesse des résultats face à la méthode choisie pour effectuer les simulations.
- Nous ne considérons ici que la simulation d'une variable continue dans un domaine défini. Parfois, le problème est d'abord de simuler des objets (ex. lentilles, chenaux, etc.) dans lesquels, ensuite, on simule des variables quantitatives. Également, on peut vouloir simuler des variables discrètes (ex. type de roches). Nous n'aborderons pas ces méthodes de simulation.

8.2 Méthodes de simulation

Il existe une panoplie de méthodes de simulation géostatistique. Si on cherche à regrouper ces méthodes, une subdivision possible est la suivante:

- les méthodes gaussiennes
 - ⇒ méthodes matricielles basées sur une décomposition de la matrice de covariance
 - ⇒ méthode gaussienne séquentielle
 - ⇒ moyennes mobiles
 - ⇒ bandes tournantes
 - ⇒ méthodes autorégressives
 - ⇒ méthodes fréquentielles
 - ⇒ etc.
- les méthodes non-gaussiennes
 - ⇒ recuit simulé
 - ⇒ simulation séquentielle d'indicatrices
 - ⇒ méthodes utilisant des champs de probabilité ("p-fields")
 - ⇒ etc.

Dans ce qui suit, nous examinons quelques unes de ces méthodes.

8.3 Méthode de simulation matricielle (LU, décomposition de Cholesky)

Cette méthode est la plus simple à programmer et sans doute la plus efficace lorsqu'on doit simuler de petits champs. Cette méthode permet de réaliser presque aussi aisément des simulations conditionnelles que non-conditionnelles.

8.3.1 Simulations non-conditionnelles

Soit n points à simuler montrant une covariance $C(h)$. On construit la matrice de covariance $K_{n \times n}$ (c'est la même matrice que celle utilisée pour un krigeage simple. On décompose cette matrice par $K=LU$, avec $L=U'$. On tire n valeurs, aléatoirement, d'une $N(0,1)$: $y_i, i=1 \dots n$. On pose $z=Ly$; z est alors une réalisation de Z ayant le bon covariogramme.

En effet,

$$\text{Cov}[LY,LY]=E[LYY'L']=LE[YY'L']=LL'=LU=K$$

8.3.2 Simulations conditionnelles

La démarche est très similaire au cas précédent.

Soit N points conditionnants et n points à simuler. On place dans un vecteur les N points conditionnants suivi des n points à simuler. La matrice $K_{(N+n)(N+n)}$ est décomposée comme auparavant en LU . On partitionne L en 4 blocs distincts:

$$L = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix}$$

Le bloc L_{11} correspond aux points conditionnants, le bloc L_{22} aux points à simuler, le bloc L_{21} mets en relation les points à simuler et les points conditionnants.

On forme un vecteur $y_{(N+n)}$ de valeurs $N(0,1)$. On a alors:

$$\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$$

d'où on tire:

$$Z_1 = L_{11} Y_1$$

$$Z_2 = L_{21} Y_1 + L_{22} Y_2$$

On a observé $Z_1=z_1$. Il faut donc imposer $y_1=L_{11}^{-1} z_1$. Connaissant y_1 , il ne reste qu'à tirer aléatoirement " n " valeurs d'une $N(0,1)$ pour former y_2 et ensuite calculer z_2 en utilisant l'expression ci-haut (notez que $L_{21}y_1$ est constant et n'a pas besoin d'être recalculé pour chaque nouvelle réalisation)..

Notes:

- Cette méthode de simulation est extrêmement efficace et rapide car une seule inversion de matrice est requise. Des réalisations additionnelles peuvent être obtenues aisément à très faible coût. La principale limitation en est une de mémoire. En effet, $(N+n)$ ne peut excéder quelques milliers car alors la matrice K devient trop large et ne peut être stockée en mémoire.
- Il faut que K soit non-singulière pour pouvoir effectuer la décomposition de Cholesky. Ceci implique que l'on ne doit pas faire coïncider un point à simuler avec un point observé.

- Il existe d'autres décompositions possibles de K . Ainsi, on peut utiliser une décomposition en vecteurs propres, valeurs propres de K .

8.4 Méthode de simulation séquentielle gaussienne (SGS)

Les fondements théoriques de cette méthode remontent à Paul Levy (1937). Nous la présentons pour le cas conditionnel car le cas non-conditionnel n'est qu'un cas particulier de celui-ci.

On a N valeurs conditionnantes normales de covariance $C(h)$ (si $N=0$, c'est le cas non-conditionnel) et n points à simuler. On génère d'abord un chemin aléatoire permettant de visiter à tour de rôle chacun des n points à simuler. Pour chaque point, à tour de rôle, on effectue le krigeage simple de ce point en utilisant l'ensemble des N points conditionnants et tous les points ayant déjà été simulés. Puisqu'on est dans un contexte gaussien, la valeur krigée et la variance de krigeage correspondent respectivement à l'espérance conditionnelle et à la variance conditionnelle de la v.a. au point visité. On tire aléatoirement une valeur d'une $N(Z_0^*, \sigma_k^2)$ et on ajoute le nouveau point simulé à l'ensemble des points connus (incluant points conditionnants et points déjà simulés). On visite le point suivant et ainsi de suite jusqu'à ce qu'on ait visité tous les points à simuler.

Notes:

- Théoriquement, cette méthode ne permet de simuler que des champs de faible taille car, en fin d'algorithme, les systèmes de krigeage simple sont de la taille du nombre total de points. En pratique, on peut limiter la taille des systèmes de krigeage simple, comme on le fait habituellement, en raison de l'effet d'écran du krigeage. (Plus la covariance est "markovienne", plus la reproduction sera fidèle).
- Si on a un effet de pépite dans le modèle, pour le cas non-conditionnel, il est préférable de simuler l'effet de pépite à posteriori, en ajoutant une erreur aléatoire de variance égale à C_0 . En effet, l'effet de pépite diminue l'effet d'écran du krigeage. Si on a des données conditionnantes, ce n'est pas possible car les valeurs sont observées avec l'influence de l'effet de pépite.
- Certaines structures sont difficiles à reproduire parfaitement par cet algorithme (ex. gaussien, sphérique). Dans le cas du gaussien, ceci peut être relié au faible effet d'écran de cette covariance. Dans le cas sphérique, c'est sans doute à cause de la portée finie, mais on ne comprend pas parfaitement le phénomène.
- La stratégie de visite des points à simuler peut influencer sur le résultat. Il faut éviter tout déplacement systématique selon une direction. Certains auteurs recommandent d'adopter une visite en 2 temps: une première visite aléatoire sur une maille très large pour assurer la reproduction des structures de grande échelle suivi d'une visite aléatoire sur une maille plus fine pour compléter la grille. On peut évidemment adopter une stratégie du même type mais en plusieurs temps.

8.4.1 Démonstration que la méthode SGS fournit les bons variogrammes (ou covariogrammes)

On invoque un argument inductif. Nous supposons que n points ont été simulés par SGS et qu'ils ont la bonne structure. Nous démontrons que ceci implique qu'à l'étape $n+1$, les $n+1$ points simulés auront également la bonne structure.

Supposons,

$$Cov(Z_n^S, Z_n^{S'}) = Cov(Z_n, Z_n') = K_{nn}$$

où Z_n^s est un vecteur $n \times 1$ contenant les n v.a. correspondant aux points déjà simulés.

à l'étape $n+1$, la valeur krigée au point $n+1$ s'écrit:

$$Z_{n+1}^{s*} = Z_n^s \lambda$$

où

$$\begin{aligned} \lambda &= K_{nn}^{-1} k \\ k &= Cov(Z_n, Z_{n+1}) \end{aligned}$$

L'algorithme SGS permet de représenter la prochaine valeur simulée comme:

$$Z_{n+1}^s = Z_{n+1}^{s*} + e$$

où e est une erreur de moyenne 0 et de variance σ_{ks}^2

On calcule alors:

$$Cov(Z_n^s, Z_{n+1}^s) = Cov(Z_n^s, Z_{n+1}^{s*}) = Cov(Z_n^s, Z_n^s) K_{nn}^{-1} k = k$$

ce qui montre que la covariance entre le $n+1$ ^{ème} point simulé et les n précédents est bien donnée par le covariogramme que l'on veut simuler. Pour la variance de Z_{n+1}^s on a:

$$Var(Z_{n+1}^s) = Var(Z_{n+1}^{s*}) + \sigma_k^2 = \sigma^2$$

ce dernier résultat découlant des équations du krigeage simple (relation de lissage, voir chap. 5).

Notes:

- L'algorithme est basé au départ sur les distributions conditionnelles. Il faudrait donc que Z soit gaussien sinon les résultats du KS ne correspondent pas aux paramètres d'une loi conditionnelle. Toutefois, la preuve précédente n'exige aucune distribution particulière pour les valeurs Z . En effet, l'argument inductif employé demeure valide peu importe la distribution puisque seules les équations de krigeage simple sont utilisées dans la preuve et le fait que le KS coïncide ou non avec les paramètres conditionnels n'est pas utilisé. Bref, on pourrait simuler un résidu " e " de moyenne 0 et de variance σ_{ks}^2 sans qu'il soit normal et on reproduirait encore le bon variogramme. Toutefois, on ne serait pas assuré de reproduire l'histogramme des données (l'histogramme des valeurs simulées tendrait davantage vers une loi normale que les valeurs originales). Ceci est dû au fait que l'on utilise la valeur krigée (par KS) dans la simulation et que celle-ci est une combinaison linéaire des valeurs observées (conséquence que l'on peut déduire, par analogie, du théorème de la limite centrale pour des données indépendantes).
- On peut très facilement combiner cet algorithme à l'algorithme LU précédent, soit pour simuler simultanément une variable en plusieurs points, soit pour simuler simultanément plusieurs variables en un même point ou plusieurs variables en plusieurs points.

8.5 Méthode de recuit simulé

8.5 Méthode de recuit simulé

Cette méthode de simulation est, de fait, basée sur une technique d'optimisation: le recuit simulé ("simulated annealing"). Cette méthode est d'une grande flexibilité. Brièvement, il s'agit de définir une fonction objectif à optimiser. Cette fonction objectif peut être la différence entre le variogramme expérimental des données simulées et le variogramme théorique que l'on désire reproduire. On peut aussi inclure plusieurs autres termes dans la fonction objectif à minimiser pour incorporer diverses contraintes géologiques, le respect d'une courbe tonnage-teneur connue, des relations déterministes, etc. Il existe plusieurs variantes du recuit simulé.

On amorce la simulation en tirant au hasard, pour chaque point à simuler, une valeur et en s'assurant que l'histogramme est bien reproduit par ce tirage. On place ensuite les valeurs conditionnantes dans le champ et on calcule le variogramme expérimental avec tous ces points et on définit la fonction objectif combine une mesure de la distance entre le variogramme expérimental et le modèle souhaité et entre l'histogramme simulé et celui désiré.

Ensuite on applique l'algorithme suivant:

- i) On tire au hasard un point dans le champ. Si le point est un point conditionnant, on ne fait rien. Sinon, on tire une valeur d'une distribution globale et on évalue la valeur de la fonction objectif.
- ii) Si la valeur de la fonction objectif a diminué, on conserve la valeur tirée. Si la fonction objectif a augmenté, on conserve la valeur tirée avec probabilité :

$$p_a = e^{(O_i - O_{i+1})/t} \text{ (et donc on replace les points à leur position originale avec probabilité } 1-p_a\text{).}$$

Dans cette équation, $O_i - O_{i+1}$ est nécessairement négatif et t est une "température" que l'on fait décroître lentement tout au long de l'application de l'algorithme. Le but est de permettre à la fonction objectif, surtout au début de l'algorithme, de croître afin de s'échapper d'optima très locaux et d'atteindre un optimum plus global.

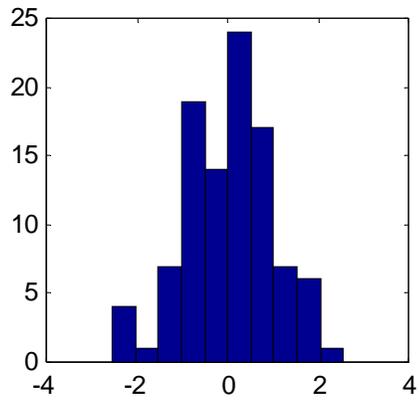
- iii) On détermine si la nouvelle valeur de la fonction objectif rencontre un certain critère de convergence. Si oui, on arrête, sinon, on retourne à l'étape i).

On peut aussi amorcer l'algorithme avec une réalisation obtenue d'une autre méthode de simulation plutôt que par tirage aléatoire. Évidemment on peut inclure d'autres contraintes dans la fonction objectif.

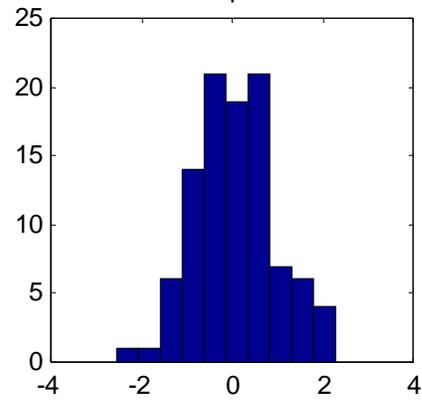
8.5.1 Exemples de recuit simulé

- i. Simuler une matrice de taille 10x10 en imposant des sommes égales pour chaque ligne et chaque colonne.

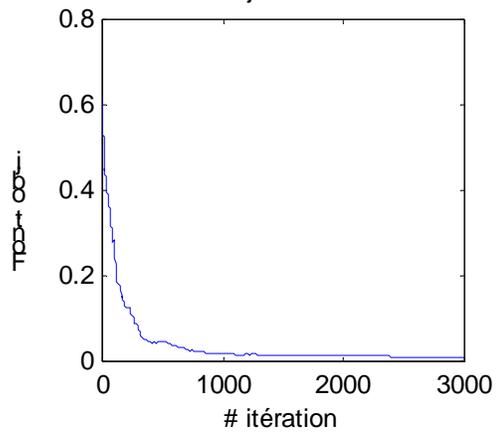
Recuit simulé données normales, 10x10, moyennes lignes et colonnes égales
Données initiales



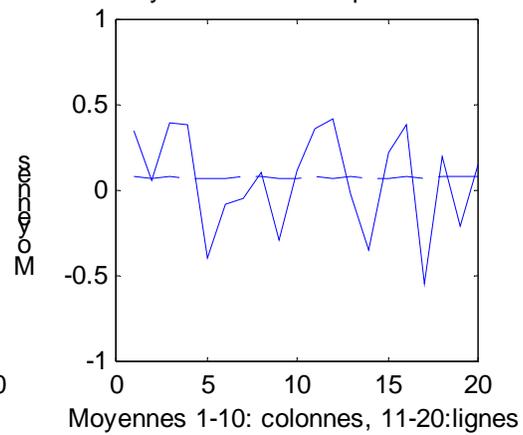
Données après recuit



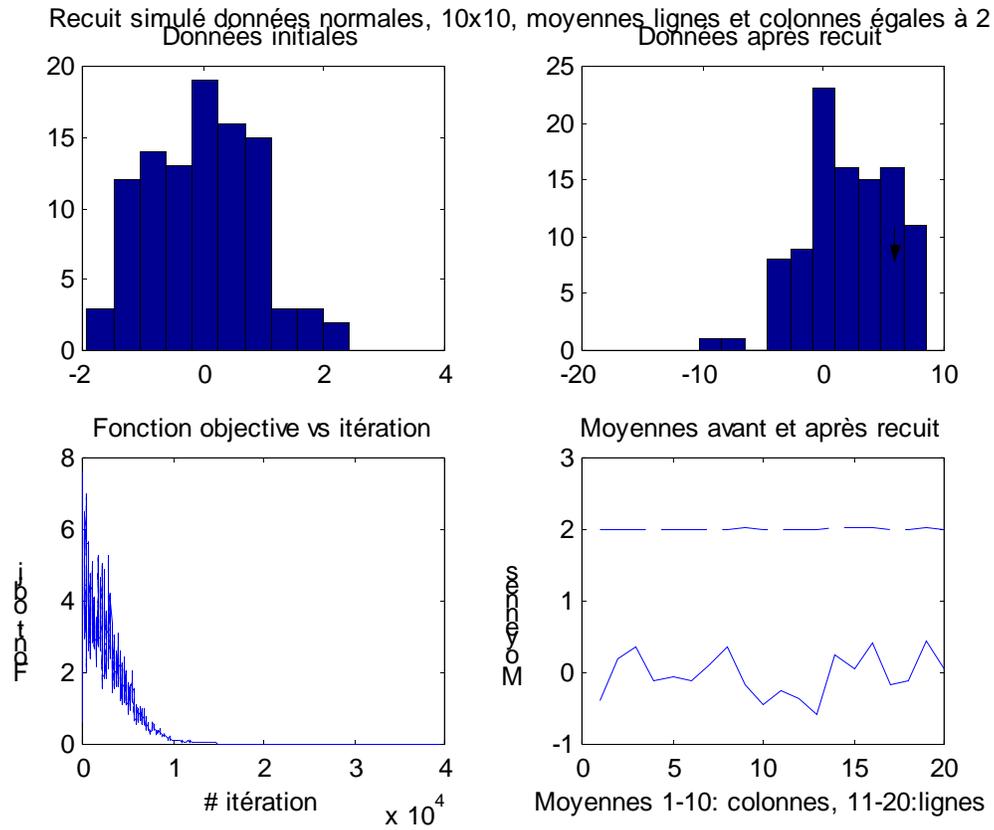
Fonction objective vs itération



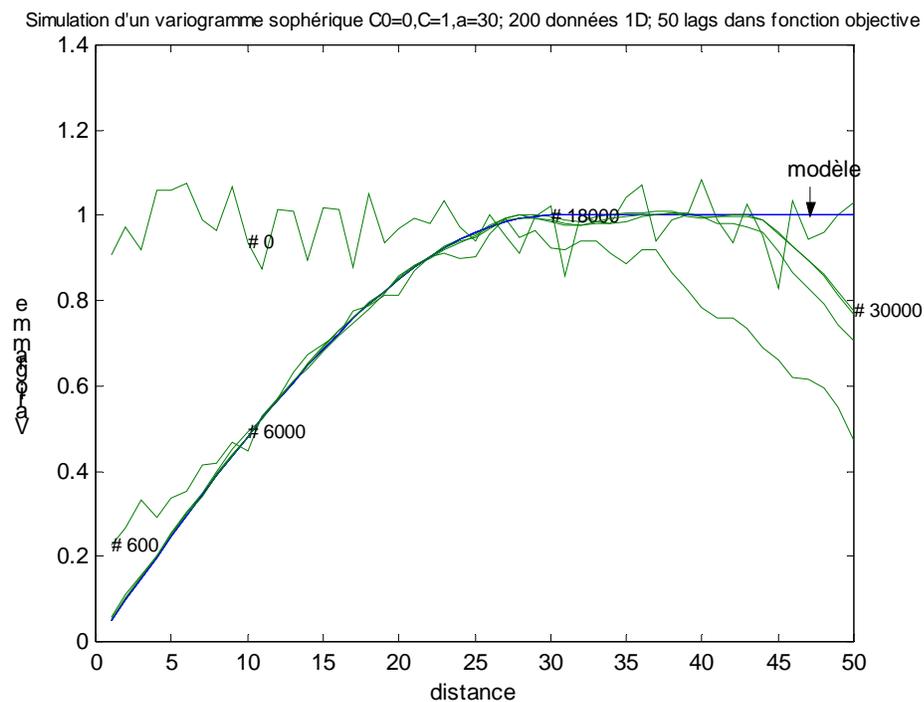
Moyennes avant et après recuit



ii. Même exemple, cette fois imposant une moyenne de 2 pour chaque ligne et chaque colonne.



iii. Simulation d'un variogramme sphérique avec $a=30$, $C_0=0$, $C=1$, 200 données en 1D, fonction objectif calculée sur les 50 premiers lags.



8.6 Exemples d'application des simulations conditionnelles.

Hydrogéologie

- Domingue (1994, mémoire de maîtrise) a évalué l'efficacité du piège hydraulique de Ville-Mercier pour confiner les contaminants présents dans le roc fracturé. Pour ce faire, il a utilisé les mesures de capacité spécifique (tirées de l'annuaire des puits et forages du Québec) pour estimer les transmissivités. Ces transmissivités ont servi de points conditionnants pour des simulations de champs de transmissivités. Ces champs ont été utilisés, conjointement avec les charges connues et les conditions aux limites, pour simuler le réseau d'écoulement. Domingue a pu ainsi définir la zone de confinement associé au piège hydraulique pour 3 réalisations différentes du champ de transmissivité. Pour au moins une des réalisations, la zone contaminée n'était pas incluse totalement dans la zone de captage. Conséquemment, le piège hydraulique pourrait être inefficace.
- Gutjahr et al. (1995), ont utilisé les simulations conditionnelles afin de comparer le pouvoir conditionnant des charges et de la transmissivité pour déterminer le trajet d'un contaminant et le temps de déplacement avant d'arriver à la frontière du champ étudié. Il a montré, entre autres, que les charges possédaient un fort pouvoir conditionnant pour le trajet suivi par le contaminant mais un très faible pouvoir conditionnant pour les temps de parcours (et donc les vitesses). Au contraire, les transmissivités ont un faible pouvoir conditionnant pour les trajets, mais un fort pouvoir pour les temps de parcours. Toutefois, c'est la combinaison des deux informations (charge+transmissivité) qui conditionne le mieux les temps de parcours.
- Gomez-Hernandez (1993) a utilisé les simulations conditionnelles pour effectuer les changements d'échelle des transmissivités à partir de la mesure quasi-ponctuelle vers des blocs de la taille de ceux utilisés dans la simulation d'écoulement. Il a défini un champ de la taille du bloc désiré et a simulé les transmissivités sur des supports comparables aux données observées. Il a imposé des frontières imperméables d'abord sur les côtés nord et sud des blocs pour déterminer la transmissivité est-ouest, puis sur les côtés est et ouest pour déterminer les transmissivités nord-sud. Il a ensuite déduit les modèles de covariance croisée liant les transmissivités ponctuelles aux transmissivités de blocs (nord-sud et est-ouest) et les covariances simples et croisées des transmissivités de blocs. Il a utilisé ces modèles de covariance dans une simulation conditionnelle pour produire des champs de transmissivité de blocs partant de valeurs ponctuelles de transmissivité.
- Gomez-Hernandez (1996) a utilisé les simulations conditionnelles comme champs initiaux pour solutionner le problème inverse en hydrogéologie. Il obtient plusieurs champs de transmissivités différents, tous satisfaisant les charges observées. Ces champs pourraient être utilisés pour modéliser le rendement d'un aquifère, l'efficacité d'un piège hydraulique, la sensibilité à la contamination, etc.

Mines

- Naraghi et Marcotte (1996) ont utilisé les simulations conditionnelles pour déterminer le nombre (et l'épaisseur) de couches à utiliser dans une pile d'homogénéisation pour obtenir le degré d'homogénéité désiré.
- Marcotte et al. (1996) a utilisé les simulations conditionnelles pour décrire l'impact du biais conditionnel des estimateurs sur les profits d'une mine type. Ils ont étudié l'influence de divers facteurs sur ces profits, en particulier de la quantité d'observations disponibles lors de l'estimation finale menant à la sélection des blocs.
- Froideveaux (1984) a utilisé les simulations conditionnelles pour modéliser les précisions sur les estimations des réserves récupérables (i.e. après application d'une teneur de coupure).

Pétrole

Les exemples d'application en pétrole sont les plus nombreux mais aussi les plus complexes. La complexité provient en premier lieu de la complexité géologique des réservoirs (failles, paléochenaux, types de roche, etc.), du peu de données directes (forer un puits coûte extrêmement cher) et de l'abondance des données indirectes (relevés de sismique réflexion, diagraphies, observations de champs pétroliers comparables, données de production (débits, pressions, etc.), expertise géologique, etc. La plupart des études combinent plusieurs méthodes de simulation (méthodes booléennes, points marqués, simulations géostatistiques).

- Haas et Dubrule(1994) ont utilisé les simulations conditionnelles séquentielles pour effectuer l'inversion 3D des impédances acoustiques. Leur approche consiste à générer un vecteur (vertical) d'impédance, de convertir la coordonnée z en coordonnée t, de calculer l'amplitude sismique théorique correspondante et de comparer la trace synthétique ainsi obtenue à la trace mesurée. Si les 2 traces sont assez semblables, le vecteur d'impédance est accepté et on passe à un autre vecteur vertical. Sinon, on resimule au même point un autre vecteur vertical et ainsi de suite.

Autres applications possibles

On pourrait utiliser les simulations conditionnelles pour:

- calculer la précision pour l'estimation de volumes contaminés au delà d'une certaine norme.
- modéliser la variabilité des teneurs observées au concentrateur et étudier l'impact de modifications au plan de minage sur cette variabilité.
- étudier différents scénarios pour le design d'une fosse à ciel ouvert et déterminer les emplacements optimaux pour les équipements. On pourrait envisager la même chose pour les exploitations souterraines mais le problème est beaucoup plus complexe.
- pour une cimenterie, fournir une base de données de "carrières" possibles sur lesquelles on peut tester l'effet d'homogénéisation de différents procédés (sautage, récupération et chargement, broyage, etc.) en vue d'optimiser les modes d'exploitation.
- en environnement, pour l'emmagasinage de déchets nucléaires, fournir des bases de données permettant d'étudier les probabilités d'occurrence de scénarios pessimistes.
- en environnement ou en hydrogéologie, identifier les emplacements névralgiques où des échantillons supplémentaires devraient être prélevés.
- etc.

8. 6 Une remarque à propos de la normalité des observations

Comme mentionné précédemment, les méthodes gaussiennes nécessitent que les données conditionnantes suivant une distribution multinormale. On ne peut jamais être certain que cette hypothèse est réaliste. On peut toutefois s'assurer que la distribution des valeurs soit normale par une transformation appropriée. On peut aussi tester en partie si la distribution est binormale, mais guère plus.

En pratique, il faut donc généralement :

- 1- Transformer les données vers une distribution normale. Ceci peut être fait par une transformation graphique ou autre.
- 2- Calculer et modéliser le variogramme des valeurs transformées.
- 3- Effectuer la simulation dans le domaine normal.
- 4- Effectuer la transformation inverse pour revenir au domaine original.

On notera que cette procédure assure la reproduction du variogramme de la variable transformée, non celui de la variable originale. La qualité de la reproduction du variogramme original dépendra largement du caractère réaliste, face au comportement des données, des hypothèses de stationnarité et de multinormalité des méthodes de simulation. Toutefois, comme mentionné précédemment, le conditionnement par des données libère en quelque sorte le modèle de ces contraintes puisque les caractéristiques implicitement contenues dans les données seront préservées et propagées aux autres données dans une certaine mesure.

Remarque: Il faut aussi garder à l'esprit que l'histogramme et le variogramme ne sont que les deux premiers moments du processus. Il est possible de générer plusieurs processus ayant même histogramme et même variogramme et pourtant présentant une apparence très différente.

8.7 Extension au cas multivariable

La simulation de données multivariées ne pose pas de problèmes particuliers et est immédiate dans le cas de la méthode LU. Pour la méthode SGS, il suffit de substituer au krigeage simple un cokrigeage simple et de déterminer la distribution multivariable conditionnelle à partir des propriétés de la loi multinormale (les valeurs impliquées s'obtiennent du cokrigeage).

8.8 Post-conditionnement

Certaines méthodes ne permettent pas de réaliser directement des simulations conditionnelles. Dans ce cas, on peut effectuer un post-conditionnement de la simulation. Le principe est le suivant (on suppose $Z(x)$ gaussien et de moyenne 0):

- i. Réaliser une simulation non-conditionnelle aux points désirés pour la simulation et aux points échantillons. Soit Z_g^s et Z_i^s les valeurs simulées aux points de simulation et aux points échantillons respectivement.
- ii. Effectuer un krigeage (simple ou ordinaire) aux points que l'on veut simuler en utilisant les valeurs observées pour obtenir Z_g^* . Faire de même avec les valeurs simulées aux points échantillons pour obtenir Z_g^{s*} .
- iii. Former: $Z_g^{sc} = Z_g^* + (Z_g^s - Z_g^{s*})$ pour obtenir la simulation conditionnelle.

Notes :

- i. Si un point à simuler coïncide avec un point échantillon, on aura $Z_g^{s*} = Z_g^s = Z_i^s$, $Z_g^* = Z_g = Z_i$ et donc $Z_g^{sc} = Z_g = Z_i$, i.e. on reproduit les données observées. Très loin des données, on aura $Z_g^{s*} = Z_g^* = 0$ et donc $Z_g^{sc} = Z_g^s$, i.e. on retrouve une simulation non-conditionnelle.
- ii. On peut montrer que $\text{Var}(Z_g^{sc} - Z_g) = 2\sigma_k^2$, i.e. une simulation conditionnelle est 2 fois moins précise qu'une estimation par krigeage simple (ou ordinaire). Il faut donc éviter d'utiliser une réalisation unique d'une simulation conditionnelle comme estimateur.
- iii. Si l'on effectue un grand nombre de réalisations ainsi conditionnées et que l'on en fait la moyenne, on retrouvera l'estimateur de krigeage simple (ordinaire). La variance d'estimation obtenue en prenant cette moyenne comme estimation est donc la variance de krigeage simple (ordinaire). La variance des différentes réalisations autour de cette moyenne est aussi la variance de krigeage simple (ordinaire).