

# GÉOSTATISTIQUE



La pratique de la Géostatistique est, d'abord, une occasion de rencontres: rencontres entre des champs d'application variés et parfois inattendus, entre des problématiques sans cesse renouvelées, et également une confrontation entre des objectifs et contraintes purement techniques d'une part et les exigences sociales, économiques, environnementales d'un monde complexe d'autre part. Autrement dit, tout en étant fiers de ce que le néologisme "géostatistique" jadis forgé à l'École des mines ait trouvé droit de cité dans le Petit Larousse, il est satisfaisant d'observer, au fil des ans, que l'immuable définition qu'en donne le dictionnaire s'éloigne de plus en plus de la réalité, et que notre discipline trouve à s'exprimer bien au-delà

de la simple estimation des gisements miniers. De fait, dans tout domaine où des jeux de données numériques présentent une organisation spatiale ou temporelle, la Géostatistique a les outils pour apporter un éclairage original, à la fois constructif et sans concession. Il semble que cet aspect transversal et non-conformiste de la Géostatistique constitue désormais son caractère dominant au regard des optionnaires, et nous ne manquerons pas dans le futur de justifier cette appréciation. Ainsi, chaque année, la diversité des vœux des étudiants constitue une chance exceptionnelle de tester des méthodes nouvelles et de parcourir des domaines nouveaux, et la garantie d'insuffler un surplus de dynamisme à l'équipe encadrante. Mais la mise en œuvre d'une Géostatistique de qualité exige en permanence d'assurer un équilibre, parfois délicat, entre des exigences souvent contradictoires: garantir une rigueur théorique indispensable à la fiabilité des résultats tout en conservant un point de vue pragmatique et réaliste afin que les conclusions abstraites trouvent à s'appliquer sur le terrain. Sans oublier une indispensable déontologie, dans des domaines où souvent les contraintes économiques ou environnementales soumettent le géostatisticien à des pressions qui ne relèvent plus de la science ou de la technique...

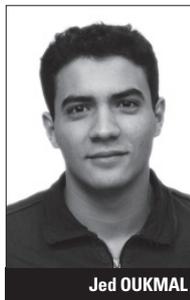
Ouverture et équilibre: c'est dans cet esprit que nous avons continué à proposer à la promotion 2010 un voyage de deux semaines en Guyane où, dans le contexte inhabituel et parfois tourmenté d'un DOM, les visites à des laboratoires, à des industriels et à des organismes institutionnels ont permis tout à la fois d'élargir l'horizon des optionnaires et de susciter un échange ouvert et fructueux avec nos interlocuteurs. Partie intégrante de la scolarité, la mission en Guyane constitue pour les optionnaires la phase d'initiation à la réalité du terrain.

Enfin, le souci d'ouverture s'est à nouveau exprimé au niveau des sujets de l'option. Outre la variété des champs d'application, une multiplicité de méthodes statistiques est employée. Nous avons ici l'illustration que notre démarche méthodologique peut trouver à s'appliquer dans de multiples domaines et cela souligne le caractère généraliste de l'option Géostatistique et Probabilités Appliquées, tant en ce qui concerne les champs applications abordés que les méthodes mathématiques mises en jeu.

**Hans Wackernagel**

# GÉOSTATISTIQUE

LUNDI 1<sup>ER</sup> JUILLET 2013 - AMPHI L108



Jed OUKMAL

**14h-15h**

## SIMULATION GÉOSTATISTIQUE D'UN CORPS GÉOLOGIQUE SOUS CONTRAINTE DE CONNECTIVITÉ

La géomodélisation des réservoirs pétroliers consiste à construire des modèles numériques maillés 3D. Ce sont des représentations de l'architecture, de la géologie et des propriétés des gisements pétroliers permettant d'évaluer leur potentiel et leur performance.

Le choix probablement le plus problématique est celui des modèles géologiques à utiliser pour reproduire le comportement dynamique des réservoirs, d'estimer l'incertitude sur les réserves et sur les profils de production. L'approche actuellement privilégiée est basée sur la génération de nombreux modèles alternatifs à un cas de base et sur un choix a posteriori de modèles d'après un certain nombre de critères. Les critères principaux utilisés sont le contenu statique des modèles et leur connectivité car ces propriétés guident en grande partie l'écoulement des fluides dans les réservoirs.

Le travail présenté ici se propose d'éliminer cette part d'a posteriori dans les modèles géologiques. Ceci est rendu possible par la construction de simulations stochastiques « sur mesure » intégrant le volume cible, les connexions et les déconnexions entre puits ainsi que le volume connecté observé pour chaque puits. Il s'agit là d'un nouveau paradigme pour la géomodélisation. Au cours du stage, nous avons exploré deux pistes. La première proposition introduit un nouveau concept de simulations stochastiques utilisant une fonction objectif. La deuxième est directement inspirée des travaux de Denis Allard au Centre de Géostatistique de Mines ParisTech sur « la simulation de lithofaciès sous contrainte de connectivité » utilisant un modèle de gaussienne tronquée.

**Total (La Défense)**



Clément WALTER

**15h-16h**

## CONSTRUCTION ET UTILISATION DE MÉTAMODÈLES POUR L'ESTIMATION DE PROBABILITÉS D'ÉVÈNEMENTS RARES ET DE QUANTILES FAIBLES

L'estimation de probabilités d'événements rares et de quantiles faibles représente un enjeu majeur dans la problématique de la quantification d'incertitudes et de la maîtrise des risques. Ces problèmes se présentent quasiment toujours sous la forme « estimer la probabilité  $P(g(X) > q)$  pour un  $q$  donné grand » (problème d'estimation de la probabilité d'un événement rare) ou « estimer le quantile  $q$  tel que  $P(g(X) > q) = p$  pour un niveau

de probabilité  $p$  petit » (problème d'estimation d'un quantile faible). Ici  $X$  est un vecteur aléatoire de loi donnée et  $g$  est une fonction coûteuse en temps de calculs. On souhaite proposer des solutions à ces problèmes d'estimation en minimisant le nombre d'appels à la fonction.

Deux directions sont privilégiées actuellement :  
- d'une part des techniques d'échantillonnage séquentiel qui permettent la mise au point d'al-

gorithmes de type Monte Carlo accéléré. Elles n'exploitent cependant pas les résultats intermédiaires pour essayer de reconstruire tout ou partie de la fonction, ce qui peut conduire à l'appeler et à la rappeler en des points inintéressants.

- d'autre part des techniques basées sur la construction (séquentielle ou non) d'un modèle réduit qui approche au mieux la fonction dans la zone de l'espace des paramètres d'entrée la plus importante pour l'estimation. Le principal défaut

de ces méthodes est que le premier plan d'expériences doit être déjà assez bon pour avoir une première idée de la zone de défaillance.

L'objectif du stage est donc de voir comment combiner ces méthodes (en particulier via le subset sampling et le krigeage) pour l'estimation directe de la probabilité ou dans une technique d'échantillonnage préférentiel adaptative.

**CEA (Bruyère-le-Chatel)**

**16h-17h**

## **DU GÉNOTYPE AU PHÉNOTYPE POUR LES LEVURES**

Pour améliorer la qualité des levures utilisées dans différents secteurs industriels, nous cherchons à étudier en collaboration avec l'équipe de « Recombinaison et Instabilité Génétique » de l'Institut Curie le lien entre le génome et les phénotypes, à savoir les propriétés macroscopiques, de différentes levures. Ces levures sont obtenues par le croisement de deux souches pures distinctes. A cause du phénomène de recombinaison génétique, le génome des cellules filles est une combinaison aléatoire de celui des deux parents. Les propriétés statistiques de cette combinaison déterminent très largement la palette des phénotypes observés. Afin d'évaluer le nombre de cellules nécessaires pour nos expériences et éventuellement

d'améliorer notre protocole expérimental, nous cherchons, dans un premier temps, à caractériser ces propriétés statistiques en les modélisant par des processus stochastiques de renouvellement. Ces propriétés sont étudiées à partir de données en libre accès et comparées à celles de nos premières expériences. Dans un second temps, nous adapterons ensuite le nombre de nos expériences et tenterons d'appliquer des méthodes d'apprentissage automatisé - aussi appelé « machine learning » - pour relier les phénotypes observés aux zones du génome qui en sont responsables.

**Centre for Computational Biology,  
MINES ParisTech (Paris)**



**Carl Johann  
SIMON-GABRIEL**