



HAL
open science

La modélisation géostatistique des milieux anthropisés

Chantal de Fouquet

► **To cite this version:**

Chantal de Fouquet. La modélisation géostatistique des milieux anthropisés. Mathématiques [math].
Université Pierre et Marie Curie - Paris VI, 2006. tel-00536237

HAL Id: tel-00536237

<https://theses.hal.science/tel-00536237v1>

Submitted on 15 Nov 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ACADEMIE DE PARIS
UNIVERSITE PIERRE ET MARIE CURIE

Mémoire des Sciences de la Terre
N° 2006-13

La modélisation géostatistique
des milieux anthropisés

Chantal de Fouquet

Habilitation à Diriger des Recherches
7 juillet 2006

Jury

Roberto BRUNO	Université de Bologne	rapporteur
Jean-Paul CHILES	Ecole des Mines de Paris	examineur
Denis MARCOTTE	Ecole Polytechnique Montréal	rapporteur
Ghislain de MARSILY	Université Pierre et Marie Curie	rapporteur
Alain TABBAGH	Université Pierre et Marie Curie	président

UFR 928, Sciences de la Terre
4, place Jussieu – 75252 Paris cedex 05

La modélisation géostatistique des milieux anthropisés

*Die Wanderjahre sind nun angetreten
und jeder Schritt des Wandrers ist bedenklich.*

Goethe (Wilhelm Meisters Wanderjahre)

*Traduction possible :
Les années de voyage sont maintenant arrivées,
et chaque pas du marcheur est sujet à réflexion.*

RESUME

Que sait-on du niveau des pollutions dans les différents milieux, comme l'air des agglomérations, les sols de friches industrielles, les cours d'eau ou les nappes ? Le mémoire présente l'application de la géostatistique à la quantification de ces pollutions et à l'évaluation de l'incertitude associée.

L'analyse exploratoire des données vise à caractériser la variabilité spatiale ou temporelle des concentrations, ainsi que leurs relations avec le milieu. Obtenus pour des sites particuliers, mais généralisables à des contextes analogues, les résultats empiriques aident à mieux comprendre la pollution et guident la modélisation. Les modèles classiques sont alors adaptés, ou de nouveaux sont développés, par exemple pour décrire les concentrations le long d'un réseau hydrographique. Pour des problèmes complexes, le recours aux simulations devient nécessaire.

Les applications sont nombreuses : prise en compte d'informations semi-quantitatives pour améliorer la précision de l'estimation, cartographie d'un risque de dépassement de seuil, adaptation de l'échantillonnage, améliorations conceptuelles pour la réglementation environnementale.

Divisé en quatre parties, le mémoire débute par une synthèse bibliographique, avec une revue des principales méthodes utilisées en « géostatistique environnementale ». La deuxième partie présente des exemples de modélisation de pollutions à l'échelle locale ou régionale, dans les sols, les cours d'eau, et l'air. La troisième partie traite des simulations : le modèle gaussien seuillé pour représenter les hétérogénéités des formations, ainsi que les variables liées par des équations aux dérivées partielles, avec l'application à un écoulement simplifié. Des perspectives de recherche sont proposées dans la dernière partie.

REMERCIEMENTS

Nombreux sont celles et ceux que je suis heureuse de remercier car ce parcours leur doit beaucoup.

Je tiens à exprimer toute ma gratitude à Jean-Paul Chilès, pour sa cordialité, la confiance accordée, ainsi que pour les précieux et discrets conseils avisés.

Je remercie Ghislain de Marsily, venu à Fontainebleau nous inciter à passer notre HDR. Qu'il reconnaisse ici l'un des résultats de cette initiative. Son intérêt bienveillant, son soutien lors de diverses publications, et sa disponibilité m'ont grandement encouragée.

De la mine à l'environnement, le parallélisme des parcours a créé une connivence avec Roberto Bruno. Les voyages de plusieurs mois à Rome, en 1985 puis en 1990, ont constitué une excellente ouverture scientifique, évidemment associée à une expérience culturelle majeure. Depuis une dizaine d'années, cette fois-ci avec l'université de Bologne, une fructueuse coopération s'est développée avec l'envoi à Fontainebleau de ses meilleurs étudiants, avec qui j'ai pu librement approfondir quelques sujets novateurs. Je l'en remercie bien vivement.

Même parallélisme avec Denis Marcotte, avec qui les discussions se poursuivent d'un congrès à l'autre, dans un sympathique mélange d'exigence, d'attention et d'humour. Je lui dois bien des remerciements pour le rapportage simultané sur la thèse de Caroline Bernard-Michel. J'espère que les échanges pourront s'intensifier.

Merci à Alain Tabbagh de son intérêt pour la géostatistique, dont il soutient activement l'enseignement à Jussieu.

J'ai plaisir à travailler avec mes collègues, parmi lesquels Hélène Beucher, François Geffroy, Christian Lajaunie, Jacques Rivoirard, Didier Renard, sans oublier Nathalie Dietrich, Françoise Poirier, Isabelle Schmitt... Comme pour les meilleurs crus, certaines relations se sont bonifiées au cours du temps, et de professionnels, les rapports sont devenus amicaux. Je suis très reconnaissante de toutes les aides reçues à l'occasion de cette habilitation.

J'ai toujours en mémoire la chaleureuse solidarité manifestée par Pierre Chauvet, Anne Dong, et Valérie Langlais.

La contribution des optionnaires, étudiants et thésards est fondamentale, par leurs attentes et leur motivation toujours renouvelées. Je garde un excellent souvenir, entre autres, du travail avec Chris Roth, guidé par Georges Matheron et Jean-Paul Chilès ; des trop brefs stages de DEA d'Alexander Prechtel et de Maarten Cuypers ; de la gentillesse de Pamela Mariotti et de l'entrain de Luca Pellicciotta ; de l'efficacité des passages à Fontainebleau de Nathalie Courtois puis d'Hélène Demougeot-Renard. La qualité du travail de Xavier Freulon et de Nicolas Jeannée durant leur thèse s'est poursuivie par celle de leur accueil, à l'occasion des visites professionnelles. Enfin, le tandem avec Caroline Bernard-Michel a fonctionné dans une excellente ambiance, avec au final les deux soutenances le même jour.

Les sujets sur l'environnement m'ont permis de retrouver avec grand plaisir d'anciens collègues, en particulier à Géovariations Jacques Deraisme et Régine Dumay-Broggi, qui m'avait si bien accueillie à mon arrivée, Henri Fraisse à Total, ainsi que Hu Lin-Ying à l'IFP.

Le soutien des collègues d'autres centres, en particulier Emmanuel Ledoux, m'a été très précieux, et je leur en suis reconnaissante. Isabelle Cojan est venue opportunément me rappeler le côté plaisant de l'exercice d'HDR, m'en facilitant l'achèvement.

Les collègues « du CIG » sont venus nombreux à ma soutenance, ce qui m'a beaucoup touchée, et j'espère que le travail en commun va se développer. Je remercie aussi l'équipe aux Mines de Saint-Etienne, qui évoluant avec le temps, m'a toujours réservé un accueil très sympathique.

L'accueil par les collègues extérieurs est essentiel, car il montre aussi qu'avec des applications effectives, l'intérêt de la recherche déborde largement du cadre académique. Parmi d'autres, j'ai beaucoup apprécié le travail avec Jean-Jacques Fry à EDF à Chambéry, Alain Puech à Fugro-France, le SHOM à Brest, l'ASPA à Strasbourg, ou avec l'équipe de l'Ineris.

Je suis très reconnaissante à Louis-Charles Oudin de l'Agence de Bassin Loire-Bretagne, grâce à qui le travail sur les cours d'eau a pu se réaliser. Sa contribution a été matérielle (avec d'efficaces conseils pour le financement) et scientifique, avec sa totale ouverture d'esprit, sa grande connaissance du contexte réglementaire, et son évaluation critique des résultats.

Enfin je dédie très volontiers ce travail à mes parents, en remerciement de leur constant soutien.

Je ne voudrais pas finir ces remerciements sans évoquer trois figures importantes dans mon parcours.

Georges Matheron, à qui l'on doit l'essentiel de la géostatistique, et qui faisait grand cas des applications. Une quarantaine d'années plus tard, les notions comme l'effet de support et l'effet d'information, la composition des variances élémentaires, le krigeage et ses multiples variantes, ainsi que la formulation de la géostatistique non linéaire et des simulations se transposent directement de la mine à l'environnement. Sans parler des FAI-k pour décrire certains phénomènes. Cette fécondité intellectuelle s'accompagnait d'une grande attention aux chercheurs du centre de géostatistique. Je n'ai pas oublié la confiance accordée, ni son intervention pour m'obtenir un meilleur poste. Il reste un grand regret sur les conditions de son départ.

Philippe Formery m'avait incitée à candidater pour une thèse auprès du centre de géostatistique. J'y ai effectivement trouvé le mélange de théorie et d'applications auquel j'aspirais. Il était le professeur sachant faire (brièvement) rire les étudiants découvrant le sujet de TD ou d'examen, mais également celui dont on continue à consulter les cours, car sous une présentation parfois bonhomme, « tout y est ».

Enfin, Jean Teissié était le mineur au jugement sûr et modeste. C'est avec une grande sympathie qu'il avait accompagné et soutenu les développements en géostatistique minière, durant ma thèse et après.

SOMMAIRE

<i>Introduction</i>	1
<i>I. L'essor d'une géostatistique environnementale</i>	3
<i>I.1. De l'exploitation minière à l'environnement</i>	3
<i>I.2. Retour sur les variables et les modèles</i>	5
<i>I.2.1. Additivité</i>	5
<i>I.2.2. Modèles multivariables</i>	6
<i>I.2.3. Éléments de modélisation spatio-temporelle</i>	10
<i>I.3. Dépassement de seuil</i>	13
<i>II. Quantification de pollutions</i>	15
<i>II.1. Sols de friches industrielles</i>	15
<i>II.1.1. Emergence de la problématique « sites et sols pollués »</i>	15
<i>II.1.2. Les sites étudiés</i>	16
<i>II.1.3. Teneurs dans les sols de friches industrielles</i>	18
<i>II.1.4. Recommandations pour l'échantillonnage</i>	24
<i>II.1.5. Améliorations et développements méthodologiques</i>	24
<i>II.2. Pollution des eaux</i>	27
<i>II.2.1. Cartographie de nappes</i>	27
<i>II.2.2. Vers un réseau représentatif de la qualité des cours d'eau</i>	28
<i>II.2.3. Moyenne et quantiles par station</i>	31
<i>II.2.4. Modèles spatiaux de débits et de concentrations le long des cours d'eau</i>	33
<i>II.3. Pollution atmosphérique</i>	36
<i>II.3.1. Éléments bibliographiques</i>	36
<i>II.3.2. Vallée de la Thur : influence de la distance à la route</i>	38
<i>II.3.3. Relation entre concentrations saisonnières et milieu à Mulhouse</i>	39
<i>II.3.4. Représentativité des stations pour le NO₂ à Rouen</i>	44
<i>II.3.5. Conséquences de la réglementation</i>	44
<i>III. Simulations</i>	45
<i>III.1. Le modèle gaussien seuillé et ses applications</i>	45
<i>III.1.1. Le modèle gaussien seuillé</i>	45
<i>III.1.2. Modélisation de réservoirs hétérogènes</i>	49
<i>III.1.3. Corégionalisation entre variable catégorielle et teneur</i>	50
<i>III.2. Fonctions Aléatoires liées par des EDP linéaires</i>	51
<i>III.2.1. Retour sur la dérivation stochastique</i>	51
<i>III.2.2. Simulation conjointe par tranches tournantes</i>	53
<i>III.2.3. Simulation d'un écoulement simplifié</i>	54
<i>IV. Perspectives</i>	57
<i>Références bibliographiques</i>	63
<i>Curriculum vitae</i>	71
<i>Liste des publications et rapports</i>	79

Publications annexées

- de Fouquet C., Beucher H., Galli A. et Ravenne C. 1989. Conditional simulation of random sets : Application to an argillaceous sandstone reservoir. In : M. Armstrong (ed.), *Geostatistics*, Dordrecht : Kluwer, vol. 2, 517-530
- Roth C., Chilès J.-P., de Fouquet C. 1996. Adapting geostatistical transmissivity simulations to finite difference flow simulators. *Water resources research*, 32 (10) 3237-3242.
- de Fouquet C. 2000. Geostatistical simulation of transmissivity and flow : simplified case of the linearized diffusion equation in steady state. In : A. Dassargues (ed.), *Tracers and modelling in hydrogeology*. IAHS publication n°. 262. 103-108
- Courtois N., Gerbaux-François O., Grenier C., Maugis P., Mouche E., de Fouquet C., Goblet P., Ledoux E. 2000. Characterization of dispersion in an alluvial aquifer by tracing techniques and stochastic modelling. In : F. Stauffer et al. (eds.), *Calibration and reliability in groundwater modelling*, 84-89.
- Jeannée N., de Fouquet C. 2003. Apport d'informations qualitatives pour l'estimation des teneurs en milieux hétérogènes : cas d'une pollution de sols par des hydrocarbures aromatiques polycycliques (HAP). *Comptes-rendus Géosciences*, 335 (5) 441-449.
- de Fouquet C. Des statistiques contre la pollution. 2003. *La recherche*, n° 368, 92-93
- Demougeot-Renard H., de Fouquet C. 2004. Geostatistical approach for assessing soil volumes requiring remediation: validation using lead-polluted soils underlying a former smelting works. *Environmental science & technology*, 38 (19) 5120-5126.
- de Fouquet C., Bernard-Michel C. 2006. Modèles géostatistiques de concentrations ou de débits le long des cours d'eau. *Comptes-rendus Géosciences*, 338 (5) 307-318

INTRODUCTION

Ce mémoire présente une sélection de travaux traitant des applications environnementales de la géostatistique et de simulations. « L'environnement » désigne ici principalement des milieux anthropisés (air, cours d'eau ou nappes, sols) dont on cherche à quantifier certaines pollutions. Développées dans des contextes différents (l'estimation minière ou pétrolière, l'hydrogéologie) les simulations peuvent également trouver des applications en environnement. Cette thématique est donc retenue comme fil directeur.

Les contributions plus ponctuelles, notamment en hydrographie, en géotechnique ou en hydrogéologie stochastique, sont omises. Des publications traitant de ce dernier sujet sont jointes en annexe.

Le mémoire commence par une bibliographie raisonnée des problèmes abordés et des principaux modèles utilisés en géostatistique environnementale.

La deuxième partie présente la quantification des pollutions dans les différents milieux, à l'échelle locale ou régionale, en insistant sur l'étude exploratoire des données. Les connaissances acquises par « l'auscultation » des données dans des cas particuliers ont souvent une portée beaucoup plus générale. Les études sur les sols de friches industrielles sont synthétisées sous forme de « retour d'expérience ». Pour les nappes, les cours d'eau, ou l'air, quelques exemples instructifs sont détaillés. Les développements méthodologiques rendus nécessaires pour ces études sont résumés dans leur contexte.

Dans la troisième partie, deux méthodes de simulation spécifiques sont développées, « les gaussiennes seuillées » pour la modélisation de formations hétérogènes, et la simulation conjointe de variables liées par des équations aux dérivées partielles, évitant certaines approximations numériques par discrétisation.

Des pistes pour des développements complémentaires sont mentionnées au long de ces deux chapitres. Quelques perspectives majeures sont abordées dans la dernière partie.

A propos de la terminologie : un « polluant » désigne toute substance considérée pour ses effets nocifs. Epanchus comme nutriments, les nitrates en ruissellement vers les cours d'eau deviennent ainsi des polluants.

Un « seuil » désigne une valeur à laquelle une concentration est comparée, indépendamment de la terminologie réglementaire (objectif de qualité, seuil de recommandation, valeur limite de protection de la santé humaine par exemple).

L'ESSOR D'UNE GEOSTATISTIQUE ENVIRONNEMENTALE

Reflétant l'évolution des sociétés occidentales, une « géostatistique environnementale » s'est développée depuis plus d'une décennie. Ce chapitre tente d'en dégager certaines caractéristiques : problèmes traités, espaces définissant la modélisation, propriétés des variables. Une revue des principaux modèles est présentée. Enfin, des rappels sont effectués sur les estimateurs non linéaires ; légèrement modifiés pour intégrer des informations auxiliaires, ces estimateurs classiques permettent en effet le calcul de risques de dépassement de seuil.

Ce chapitre reprend des éléments de la contribution sur l'environnement d'un ouvrage en hommage à G. Matheron [Bruno & de Fouquet, 2005].

I.1. De l'exploitation minière à l'environnement

Le développement des applications « environnementales » de la géostatistique apparaît à travers les rendez-vous quadriennaux de cette discipline : absent d'Avignon en 1988, "l'environnement" figure parmi les "autres applications" en 1992 à Troia (Portugal) ; devenue autonome en 1996 à Wollongong (Australie), cette rubrique regroupe le quart des contributions en 2000 au Cap (Afrique du Sud). Le congrès de Banff (Canada) en 2004 a pâti de la programmation concurrente d'un congrès "Géoenv" ; ces réunions biennales sur la « géostatistique pour les applications environnementales » drainent depuis 1996 des présentations méthodologiques et de nombreux cas d'études.

Une grande diversité d'applications

Au sens large, la « géostatistique environnementale » recoupe des disciplines variées, ayant pour objet de caractériser ou de quantifier les ressources renouvelables, les milieux naturels ou anthropisés (sols, eaux superficielles ou nappes, air), les phénomènes climatiques. Les travaux précurseurs remontent à Delhomme [1970] en hydrogéologie ou Guibal [1973] et Marbeau [1976], en agronomie forestière. La géostatistique est désormais couramment appliquée en agronomie [par exemple Webster et al., 1989 ; Almeida et al. 2001], entre autres pour le recensement de certaines espèces végétales ou animales [Pereira et al., 1999] ou en halieutique [Petitgas, 1991] ; en pédologie [Grzebyk et al. 1994], pour cartographier la profondeur de différents horizons, classifier des sols sur une parcelle ou quantifier des paramètres physico-chimiques pour l'agriculture de précision ; en climatologie, pour cartographier des champs de températures [Gohin, 1989] ou de pluies [Biau et al. 1999 ; Raspa et al., 1997]. Sont également étudiés l'exposition aux poussières ou au bruit, la stabilité ou l'évolution temporelle de reliefs [Bourgine et al., 2001], l'activité volcanique [Jaquet, 2005], etc.

Par sa généralité, le « problème inverse » pour les écoulements souterrains déborde largement du champ environnemental. La prévision du développement d'un panache de pollution devrait désormais associer modélisation hydrogéochimique et modélisation géostatistique, afin de mieux contraindre les paramètres comme la perméabilité par les observations. En contrepartie d'une modélisation plus complexe, la géostatistique évite certains biais d'estimation et rend plus objective l'évaluation des incertitudes entachant les prévisions.

Au sens restreint, la « géostatistique environnementale » a pour objet la caractérisation des milieux anthropisés en vue de leur contrôle ou de leur remise en état.

Etudiant la pollution des huîtres du bassin du Morbihan, Marbeau [1974] recherche ainsi des corrélations avec de possibles variables explicatives. Orfeuil [1977] associe analyse factorielle des correspondances et krigeage disjonctif pour construire un modèle de prédiction temporelle à court terme de la pollution atmosphérique. Aujourd'hui encore, de nombreux problèmes conduisent à rechercher des relations entre les concentrations mesurées et des variables mieux informées décrivant les différents milieux, afin d'améliorer les prévisions en incorporant ces informations dans l'estimation. La pollution de l'air, celle des nappes ou des cours d'eau sont notamment reliées à la densité de population ou à l'occupation des sols, informations désormais facilement valorisables grâce aux systèmes d'information géographique.

Pour un même milieu, les mêmes questions sont abordées à des échelles très différentes. La pollution atmosphérique (composés azotés ou sulfurés, particules, dépôts salins) est caractérisée aussi bien à l'échelle continentale ou nationale, qu'à proximité immédiate d'une cheminée d'usine. Dans les nappes phréatiques, les concentrations en nutriments ou en micropolluants sont étudiées à l'échelle régionale, ou localement à proximité de la "source" de pollution.

Des acquis méthodologiques nombreux

Quelles caractéristiques dégager du foisonnement des modèles proposés dans la littérature, comme en témoignent par exemple les actes des congrès Géoenvironnement ? Sous une présentation différente, beaucoup d'études de pollution se ramènent à la formulation développée pour l'estimation minière il y a plus d'une trentaine d'années [Matheron, 1965, 1973, 1976]. Les problèmes abordés concernent en effet :

- la caractérisation et la quantification de la variabilité spatiale d'une ou de plusieurs variables conjointement ;
- l'estimation linéaire et l'évaluation de la précision associée, pour la cartographie des concentrations ou celle de limites de couches ;
- l'adaptation des mailles de reconnaissance, pour atteindre une précision fixée ou rechercher un compromis favorable entre le coût d'échantillonnage et la précision de l'estimation ;
- l'utilisation d'informations auxiliaires, moins précises mais généralement plus abondantes, pour améliorer l'estimation ;
- la comparaison à un seuil, avec le calcul de la probabilité de dépassement associée.

L'analogie minière est particulièrement évidente en pollution de sols. S'y retrouve l'effet de support avec la quantification du volume de sols à excaver pour extraire la pollution, en fonction du seuil, des unités de sélection, ainsi que des éventuelles contraintes techniques d'accessibilité. L'effet d'information commence à être pris en compte. La délimitation des zones polluées est toujours approximative, puisque les teneurs réelles restent inconnues, seules des estimations étant disponibles. En dépollution de site, ces erreurs conduisent à laisser en place des teneurs fortes estimées à tort faibles, et réciproquement à traiter parfois inutilement des sols non pollués. Ne pas détecter des zones de teneur forte peut se révéler critique, si un contrôle en fin de dépollution prouve que celle-ci est incomplète. En pratique, des marges de sécurité parfois assez importantes sont appliquées empiriquement. Pour les chantiers de grande ampleur, il serait sans doute plus économique et moins risqué d'adapter l'échantillonnage, ou de calculer sur site un estimateur éventuellement simplifié mais mieux adapté.

En géostatistique minière, la "sélectivité" d'une distribution caractérise la possibilité de récupérer le plus de "métal" en extrayant le moins de "minerai" possible. Le métal représente ici la masse de polluant (métaux, hydrocarbures, etc.), et le "minerai" les sols. A même quantité de "métal" présent sur le site, une répartition plus sélective nécessite le traitement de volumes

de sols plus réduits. La sélectivité est caractérisée par l'une ou l'autre des fonctions de récupération suivantes [Matheron, 1984]:

- la courbe $Q(T)$ exprimant le "méta" récupéré en fonction du "minera" sélectionné ;
- la courbe $B(s)$ exprimant le bénéfice conventionnel, défini comme la masse de polluant récupéré au-delà du seuil s .

Ces courbes synthétiques mesurent la dégradation de la sélectivité lorsque le support augmente ou à cause de l'effet d'information. Peu utilisées en environnement, elles constitueraient une aide utile au choix du seuil de dépollution et du volume des unités de sélection.

En pollution atmosphérique, un indicateur comme l'AOT40 (accumulated exposure over a threshold of 40ppb) qui mesure la dose d'Ozone en excès au-dessus d'un seuil de 40 ppb, correspond au bénéfice conventionnel à cette coupure. L'usage de tels indicateurs mérite d'être généralisé, pour faciliter la comparaison de l'évolution interannuelle des dépassements de seuil [rapport d'étude PREDIT, 2005]

Problématiques émergentes

Aux débuts de la « géostatistique environnementale » beaucoup des problèmes abordés avaient déjà reçu une solution théorique. Superficiellement, « l'environnement » pouvait apparaître comme un vaste champ d'application de méthodes bien rodées, que le praticien se contenterait d'adapter au cas par cas.

Les problèmes presque systématiquement multivariés ont ainsi été l'occasion de formuler diverses variantes du krigeage multivarié [par exemple Monestiez et al., 1999]. La méthode de dérive externe, qui s'accommode d'une relation localement linéaire qu'il n'est pas nécessaire de complètement spécifier entre variables auxiliaires et variable principale, a été largement utilisée pour introduire une certaine connaissance du milieu dans l'estimation [par exemple Bobbia et al., 2000]. Les phénomènes à composante temporelle ont été souvent traités de façon satisfaisante en élargissant la classe des modèles usuels de covariance [Gneiting 2002]. Néanmoins l'estimation des concentrations en polluants dans les cours d'eau, l'étude du transfert sol-nappe d'une pollution ou les prévisions du niveau de la pollution atmosphérique, rendent nécessaires de nouveaux développements : changement de topologie, traitement de grandeurs non additives, couplage d'un modèle déterministe complexe et d'un modèle probabiliste par exemple. Enfin l'espace géographique est parfois déformé pour modéliser des phénomènes fortement contrastés localement comme la pluviométrie.

Des problématiques nouvelles apparaissent. La quantification des effets épidémiologiques des pollutions implique de rechercher une régionalisation du risque à travers la régionalisation d'un paramètre, à partir de données fortement agrégées. L'évaluation d'une contribution spécifique aux émissions globales, par exemple pour le CO₂ rejoint le problème de la « conciliation » entre modèle phénoménologique et mesures, qui nécessite de corrélérer des variables définies sur des supports qui diffèrent parfois de plusieurs ordres de grandeur.

I.2. Retour sur les variables et les modèles

Des méthodes « bien rodées » comme l'estimation linéaire ne sont pas toujours appropriées, même pour des problèmes apparemment simples. Un retour sur la pertinence physique de la modélisation est toujours nécessaire.

I.2.1. Additivité

En physique, les variables intrinsèquement additives représentent des grandeurs « extensives », comme les dimensions ou la quantité de matière d'un système. La masse du mélange de plusieurs échantillons est ainsi la somme de leurs masses respectives. Le traitement des grandeurs additives doit tenir compte des éventuelles différences de support entre données, ou entre les données et le volume à estimer.

Soit une couche d'épaisseur variable, dont la teneur est échantillonnée par des sondages comportant un seul tronçon homogénéisé et analysé sur toute la couche. L'estimation s'effectue alors par l'intermédiaire de « l'accumulation », produit de la teneur par l'épaisseur du niveau. En pédologie ou en pollution de sols, les taux (de matière organique, d'argiles...) sont donnés pour certaines classes granulométriques, principalement les particules fines. Pour l'estimation, il faudrait tenir compte de la masse de matière organique ou d'argile contenues dans l'échantillon, et non plus seulement de leur taux. L'information nécessaire manque parfois, en particulier pour l'échantillonnage des sols pollués.

Les variables « intensives », décrivant les propriétés associées aux points de l'espace, notamment les grandeurs définies par des lois physiques (Darcy, Ohm ...) ne sont généralement pas additives : perméabilité, résistivité électrique, impédance acoustique... La plupart des « paramètres » caractérisant les sols sont ainsi intrinsèquement non additifs : au signe près, le pH est défini comme le logarithme de la concentration en ions H⁺. La « capacité d'échange cationique », mesure de la quantité totale de cations échangeables que le sol peut adsorber, s'exprime en milliéquivalents par 100g de sol ou d'argile et dépend du pH... L'estimation par krigeage reste pertinente, mais uniquement au même support que celui, supposé identique, de tous les échantillons.

La surface terrière (surface relative occupée par les troncs « coupés » à 1,50 m) ou la densité de tiges à l'hectare sont additives, mais pas le nombre des espèces. En l'absence de toute autre information, nous savons seulement que le nombre d'espèces présentes dans la réunion de deux placettes est compris entre le maximum et la somme du nombre d'espèces présentes dans chacune d'elles, selon que toutes les espèces de la placette la moins diversifiée sont ou non communes aux deux placettes. Estimer par krigeage « de bloc » le nombre d'espèces présentes dans une parcelle provoque nécessairement une sous-estimation radicale de la valeur recherchée. On constaterait d'ailleurs que le nombre d'espèces ainsi estimé sur une parcelle *S* serait assez fréquemment inférieur au nombre recensé dans les placettes de superficie *s* les plus diversifiées échantillonnées à l'intérieur de *S*. Le krigeage « de bloc » fournit en fait une estimation du nombre moyen d'espèces présentes dans les placettes *s* dont la réunion forme la parcelle *S*. La biodiversité dans un champ, quantifiée par le nombre d'espèces qui y sont présentes, reste donc inaccessible à partir du seul relevé du nombre d'espèces présentes sur des placettes d'échantillonnage.

Une application aveugle du krigeage, négligeant les propriétés des variables étudiées, peut évidemment fournir des résultats aberrants. D'où la nécessité, en abordant un nouveau domaine d'application de la géostatistique, d'un examen détaillé du problème et des propriétés des variables. Dans l'exemple de la biodiversité, une information plus précise, à savoir les espèces recensées par parcelle, est indispensable pour évaluer le nombre d'espèces dans tout un champ.

I.2.2. Modèles multivariés

Le « contexte environnemental » se caractérise par le grand nombre de « paramètres » intervenant dans les phénomènes, et par suite par l'aspect multivariable des problèmes.

A) Typologie des liaisons entre variables

Laissant de côté les questions de classification [Wackernagel et al., 1988], on distingue de façon non exhaustive, différents types de relations.

Variables homologues

Il s'agit par exemple des concentrations en différents polluants dans un milieu donné.

Sur les anciens sites industriels, les activités passées ont généré simultanément ou successivement différents cortèges de polluants. La pétrochimie génère ainsi des pollutions

métalliques (plomb et autres adjuvants) et organiques (entre autres les Hydrocarbures aromatiques polycycliques). Les corrélations dépendent de l'origine et des propriétés physico-chimiques de ces produits : les corrélations entre HAP de compositions chimiques voisines sont plus élevées, à l'exception cependant des plus volatils ; les concentrations entre composés métalliques se différencient notamment suivant leur solubilité.

Les polluants issus d'activités différentes (un traitement métallurgique et une usine de production d'acide, par exemple [Thiry et al., 2002]) sont a priori sans corrélation, mais peuvent être communs à plusieurs de ces activités. Bien que certaines activités aient engendré des cortèges caractéristiques, il n'est donc pas toujours possible de connaître a priori les liaisons entre les différentes substances présentes,

En pollution atmosphérique, les réactions photochimiques et la température provoquent des réactions complexes entre l'ozone et les oxydes d'azotes, dont il est nécessaire de tenir compte pour une modélisation à l'échelle journalière ou à petite distance.

Causalité entre « variables explicatives » et concentrations

Certaines variables caractérisant le milieu, comme le type et la densité d'occupation des sols, l'indication d'une activité (implantation ou caractéristiques d'infrastructures industrielles ou routières, intensité du trafic) fournissent une information sur la présence et l'intensité possibles d'une source de pollution. On peut chercher à quantifier le degré « prédictif » de cette information souvent indirecte par rapport aux concentrations (pollution atmosphérique locale, pollution des sols, pollution des nappes par les engrais). Mettre en évidence, puis quantifier ces relations permet d'améliorer l'estimation des concentrations, sous réserve que le milieu soit connu avec suffisamment de détail à l'échelle concernée ; ce peut être le cas pour la pollution atmosphérique sur une agglomération ou à l'échelle régionale, ou pour la pollution d'origine agricole dans les cours d'eau ou les nappes.

Concentrations et « marqueurs »

En pollution de sol, des observations « organoleptiques » comme la présence de remblais ou de scories, les traces ou les odeurs de « goudrons », n'expliquent pas la pollution mais, comme conséquences de celle-ci, peuvent en guider la recherche. Parmi divers critères statistiques, le rapport de corrélation, défini comme la variance des moyennes des concentrations par classes de la variable qualitative, rapportée à la variance des concentrations permet de mesurer le degré de liaison entre concentrations et observations, même qualitatives [Saporta, 1990]. En présence d'une liaison marquée, ces observations qualitatives, imprécises mais peu coûteuses à acquérir, et donc potentiellement abondantes sont parfois utilisables pour améliorer la précision de l'estimation. Un exemple en pollution de sols par des hydrocarbures est présenté au chapitre suivant.

Concentrations et variables synthétiques

Lorsque les variables « explicatives » ou les « marqueurs » sont nombreux mais chacun faiblement lié aux concentrations, l'introduction de variables synthétiques peut améliorer ces liaisons. Une fonction empirique des variables auxiliaires est fréquemment utilisée par les AASQUA (associations agréées de surveillance de la qualité de l'air) pour approcher la pollution atmosphérique sur une agglomération. Les facteurs de l'analyse en composantes principales des variables auxiliaires sont également utilisables.

Relations phénoménologiques

En hydrogéologie, l'équation de diffusivité relie charge et transmissivité. Pour des conditions d'écoulement simplifiées (gradient de charge moyen uniforme, régime permanent...) la linéarisation de cette équation au premier ordre conduit à une équation différentielle entre la perturbation de la transmissivité et celle de la charge. Dans le cadre de la théorie des FAI-k [Matheron, 1973], A. Dong [1990] déduit de ces équations le degré de stationnarité des différentes variables, ainsi que les relations entre leurs covariances généralisées ; l'écriture de

ces relations sous forme intégrale permet d'éviter la dérivation de la transmissivité, physiquement peu réaliste.

Asli et al. [2000] exploitent de même la linéarité des relations pour inverser par cokrigage le contraste de densité à partir de mesures d'anomalie gravimétrique.

La puissance de calcul toujours croissante facilite les approches numériques qui permettent de s'affranchir de certaines simplifications restrictives. Plusieurs auteurs calculent numériquement les covariances dans des cas plus réalistes, à partir d'un grand nombre de tirages d'une simulation géostatistique de la transmissivité, la charge associée en régime permanent étant obtenue par un code d'écoulement.

Ces méthodes ont été étendues au régime transitoire. Le transfert réactif élargit considérablement le nombre et le type de relations à prendre en compte.

Systèmes dynamiques complexes

Les calculs d'évolution en climatologie, ou les prévisions météorologiques à court terme, et les prévisions associées de qualité de l'air sur les grandes agglomérations, imposent de tenir compte de la dynamique du système. Dans ces domaines, le nombre de paramètres rend le système très complexe, et il paraît illusoire d'en déduire directement un système spatio-temporel cohérent de covariances simples et croisées entre les différentes variables. De plus, seules certaines de ces variables s'avèrent « utiles » pour le problème posé : ainsi en écologie, on s'intéresse à la pluviométrie, non au champ de pressions pourtant nécessaire aux prévisions de hauteur de pluies. Pour ces phénomènes, la modélisation géostatistique des variables d'intérêt se concentre plutôt sur l'écart entre les mesures et les prévisions d'un modèle généralement déterministe. Ces techniques « d'assimilation de données » sont évoquées dans la suite.

B) Modèles associés

Ces diverses relations sont décrites à l'aide d'une palette de modèles. En particulier, la régression non linéaire $z=f(\mathbf{y})$ entre une concentration z et un vecteur \mathbf{y} de « variables explicatives » offre de nombreuses variantes, de la relation déterministe à la liaison probabiliste.

Relations à composante déterministe comme les modèles à résidus

Ces modèles, du type $Z(x)=f(y(x))+R(x)$ ou $Z(x)=f(y(x))+g(y(x))R(x)$ sont très utilisés. Le biais pour l'inférence du variogramme des résidus, lorsque les données ont d'abord servi au calage de la « dérive », est rarement examiné.

Relations déterministes non complètement spécifiées comme la dérive externe

La relation $Z(x)=a_0+\sum_i a_i f_i(y_i(x))+R(x)$ est supposée admissible localement sur chaque voisinage de krigeage, les coefficients a_i , inaccessibles en pratique, étant filtrés par les conditions de non biais. Ainsi, le logarithme translaté $f(y)=\log(I+y/m)$ est souvent commode pour linéariser les relations entre concentration et variables décrivant le milieu (occupation du sol, densité de population par exemple).

Relations probabilistes de type linéaire, comme le modèle linéaire de corégionalisation

Dans ce modèle très courant, les covariances spatiales ou spatio-temporelles simples et croisées s'expriment comme combinaisons linéaires de structures élémentaires : $\gamma_{ij}(h)=\sum_{\ell} c_{ij}^{\ell} \gamma_{\ell}(h)$. Les variables sont des combinaisons linéaires de composantes spatiales correspondant aux différentes échelles de variabilité, décomposées en facteurs communs. L'analyse krigeante consiste à estimer les composantes à partir des variables.

Relations différées

Ces modèles prennent en charge les corrélations différées, par exemple de la forme $Z(x) = Y(x - d) + R(x)$. Peuvent être jointes à cette classe les corrélations entre dérivées partielles, la corrélation entre Y et ses dérivées $\partial Y / \partial x_i$ au même point étant nulle.

De nombreux autres modèles sont possibles en combinant convolution, dérivation, etc.

Relations non linéaires

Il s'agit notamment des relations entre des variables et leurs transformées, comme les indicatrices, les variables seuillées à une « limite de quantification », le logarithme translaté, les anamorphosées, etc. Leur modélisation nécessite de préciser les lois bivariées ou la loi spatiale complète.

Factorisation spatiale

Orthogonaux point à point, les facteurs d'une analyse en composantes principales (aussi notés EOF pour empirical orthogonal factors) sont généralement spatialement corrélés. Obled et Braud [1989] effectuent une interpolation en conservant les fonctions propres associées aux plus fortes valeurs propres. L'interpolateur ainsi obtenu n'est plus « exact » mais ne nécessite pas l'hypothèse d'un champ homogène. Cette méthode a été notamment utilisée pour interpoler des champs pluviométriques fortement contrastés dans les Cévennes [Obled et Creutin, 1986].

Bourgault et al. [1993] comparent ACP et factorisation spatiale d'abord sur simulation puis sur des données de prospection géochimique. La factorisation spatiale est moins sensible à des données bruitées ; l'interprétation géologique des facteurs spatiaux est plus claire.

Pour introduire la corrélation spatiale dans la décomposition factorielle, Switzer et Green [1984] introduisent les facteurs d'autocorrélation min/max ou MAF (min/max autocorrelation factor) ; ils décomposent un vecteur de fonctions aléatoires spatialement corrélées en autant de facteurs orthogonaux en $h=0$ ainsi qu'en un pas $h = \mathbf{D}$ fixé, ordonnés par structure spatiale croissante. Ces auteurs filtrent ainsi le bruit dans des images multicanaux en télédétection. Desbarats et Dimitrakopoulos [2000] utilisent cette décomposition en facteurs spatialement peu corrélés pour ramener la simulation conjointe de 15 classes granulométriques à la simulation monovariée des six facteurs les mieux structurés spatialement. Le choix de la distance \mathbf{D} semble pour partie arbitraire, ou imposé par le nombre de couples disponibles par pas du variogramme expérimental.

Apparue en géophysique depuis une vingtaine d'années, la transformation en ondelettes fournit une décomposition multi-échelle et localisée. Utilisée notamment en traitement du signal et en traitement d'images, elle reste peu appliquée en environnement.

Transformation ou changement d'espace

Pour des phénomènes fortement contrastés comme la pluviométrie, les caractéristiques du milieu expliquent parfois mieux les valeurs observées que la proximité géographique. Des modèles de la forme $Z(f(x))$ sont alors recherchés, la fonction f étant une variable explicative (altitude, facteur synthétique) ou résultant d'une optimisation numérique. Ces transformations changent l'espace de travail ou le déforment.

Sampson et al. [2001] présentent une revue de différentes méthodes de déformation de l'espace. Une famille paramétrique de variogrammes étant choisie, la fonction f est calculée de façon à ce que le variogramme expérimental non stationnaire $\mathbf{g}_z(x, x')$ soit approché au mieux par un variogramme ne dépendant que de l'incrément dans l'espace transformé $\mathbf{g}_z(x, x') = \mathbf{g}(f(x) - f(x'))$. La méthode d'optimisation varie selon les auteurs : moindres carrés pondérés, spline, ajustement par maximum de vraisemblance, approche bayésienne, algorithmes itératifs de type recuit simulé ... L'une des difficultés consiste à obtenir une

fonction f suffisamment régulière et évitant « les plis ». Des questions importantes restent en suspens : les critères définissant la déformation restent pour partie arbitraires, notamment le choix de la famille de variogrammes dans l'espace déformé ; les conditions de cohérence sont délicates à assurer, comme la bijection de la transformation ; la variance de krigeage devrait être complétée par une incertitude sur la modélisation. Pour lever certaines difficultés dans la détermination de la fonction f , Perrin et al. [1999] utilisent une famille paramétrique de fonctions radiales, fondée sur l'exponentielle de Gauss.

Avant le développement de l'informatique, les déformations de l'espace, comme les transformations conformes conservant les angles, étaient couramment utilisées pour se ramener à un cas simple dont la solution analytique était connue. C. Roth [1995] calcule ainsi numériquement une transformation conforme ramenant une portion de nappe à un rectangle. Pour un milieu de perméabilité constante dont la géométrie et les segments aux limites sont ceux de la nappe réelle, le code d'écoulement fournit en effet un réseau de lignes orthogonales (lignes de courant et équipotentielles), compatible avec les conditions aux limites imposées. Mais cette méthode n'est adaptée qu'à des conditions aux limites de type perméamètre (quatre segments en limites, deux de type imperméables et deux à charge imposée constante).

La transformation f peut également être déduite de variables auxiliaires ou de facteurs synthétiques. Traitant d'un problème bivariable, la relation entre hauteur de pluie et pression atmosphérique au sud-ouest de l'Europe, Biau [1999] ramène ainsi un krigeage temporel de la hauteur d'eau par station à un krigeage spatial dans l'espace des coordonnées définies par les deux premiers facteurs d'une ACP (EOF). Les conditions météorologiques fixant les « coordonnées » pour un jour donné, les points expérimentaux ont dans ce nouvel espace la même localisation pour toutes les stations, alors la pluviométrie en ces « points » et son variogramme changent selon le site. Ce modèle particulier ne permet pas l'interpolation de la hauteur de pluie entre stations.

Dans cet exemple, un phénomène spatio-temporel est ramené à une modélisation spatiale (par changement d'espace) et multivariable (la pluviométrie aux différents sites).

1.2.3. Eléments de modélisation spatio-temporelle

La composante temporelle constitue une troisième caractéristique importante des phénomènes environnementaux. Le temps intervient de diverses façons dans la modélisation.

Les phénomènes discrétisés, évacuant le temps de la modélisation probabiliste

La discrétisation temporelle permet de ramener certains problèmes spatio-temporels à une modélisation multivariable, notamment en cas de mesures synchrones à des dates peu nombreuses. Le traitement multivariable offre une grande souplesse de modélisation, facilitant une description assez fine de la non stationnarité temporelle, lorsque la variabilité évolue au cours du temps, sous l'effet saisonnier ou sous l'influence d'une « dérive ».

Dans certaines décompositions en une prévision déterministe et une « innovation », la dynamique du système est prise en charge par le modèle déterministe, le modèle probabiliste portant sur les « résidus ». Le modèle multivariable décrit alors les relations entre les écarts successifs entre les « observations » et le modèle dynamique. Ainsi, en pollution de l'air, les phénomènes sont complexes du fait de l'interaction d'un grand nombre de paramètres environnementaux et météorologiques. La concentration Z à la date t_i est alors décomposée en deux termes, la prévision issue d'un modèle déterministe f , et un résidu aléatoire, spatialement corrélé $Z(x, t_i) = f(x, t_i) + R_i(x)$. Intégrant la dynamique du phénomène, la composante déterministe retrouve les non stationnarités spatiales majeures ; les hypothèses de stationnarité sont plus strictes sur les résidus R_i que sur la concentration Z , du moins si les sorties du modèle ne présentent pas de décalage spatial ou d'écart systématique sur les niveaux.

Dans sa version avancée, « l'assimilation de données » consiste à faire évoluer une matrice de « covariance d'erreurs » entre la prévision phénoménologique et les mesures ou « observations » en quelques stations. Des prototypes simplifiés (modélisation écologique de l'évolution d'une population), ou opérationnels (prévisions météorologiques, prévision du niveau de pollution atmosphérique à différentes échelles), permettent la mise à jour régulière des prévisions à partir des mesures [Bertino, 2001 ; Sénégas et al., 2001 ; Blond, 2002].

Pour bon nombre de problèmes inverses, la dynamique est trop complexe pour une modélisation stochastique explicite. Contraindre les perméabilités par les essais de puits est cependant une pratique courante en modélisation pétrolière. Le conditionnement repose sur une procédure de minimisation des écarts entre l'historique observé et l'historique simulé, par exemple à l'aide de la méthode de déformation graduelle [Hu, 2000 ; Hu et al., 2004]. Effectués sous la direction d'E. Mouche, des tests préliminaires sur des essais de traçage ont donné des résultats satisfaisants [Rolin, 2005]. La démarche associant modélisation géostatistique et modélisation phénoménologique via les codes de calcul gagnerait à être étendue.

Mentionnons enfin une approche récente du problème inverse, plus globale, consistant à estimer les paramètres à partir des observations, en considérant la « surface de réponses » dans un espace de grandes dimensions [par exemple Jones et al., 1998 ; Huang et al., 2006].

Les phénomènes temporels sans modélisation dynamique explicite

Les premiers modèles géostatistiques spatio-temporels consistaient souvent en une simple extension des modèles spatiaux usuels, le temps étant traité comme une coordonnée supplémentaire. Les covariances $C(h, \mathbf{t})$, h désignant l'incrément spatial et \mathbf{t} l'incrément temporel, utilisaient une anisotropie géométrique ou zonale, respectivement de la forme $C\left(\sqrt{\frac{h^2}{a^2} + \frac{\mathbf{t}^2}{b^2}}\right)$ ou $C_s(h) + C_t(\mathbf{t})$. Les périodicités, annuelles, hebdomadaires ou journalières selon les phénomènes, se traduisent par une composante périodique non amortie sur le variogramme temporel, modélisable par un terme en cosinus. Ce schéma n'étant admissible qu'en dimension un nécessite un modèle adapté d'anisotropie spatio-temporelle.

La factorisation de la covariance $C(h, \mathbf{t}) = C_s(h) \cdot C_t(\mathbf{t})$, fréquemment utilisée pour séparer les composantes spatiales et temporelles, correspond à la propriété markovienne, ou en langage géostatistique, à « l'effet d'écran » suivant : conditionnellement à $Z(x, t)$ les variables $Z(x', t)$ à la même date en un autre point et $Z(x, t')$ au même point à une autre date sont indépendantes. Ce modèle, auparavant appliqué à la modélisation spatiale de milieux sédimentaires, a des propriétés très commodes pour l'estimation ou la simulation : lorsque les mesures sont synchrones aux différentes stations, l'estimation de la concentration Z en un point quelconque à une date de mesure t_i ne dépend que des mesures à cette seule date. Parmi ces modèles, la covariance « gaussienne » $C(h, \mathbf{t}) = \exp\left(-\left(\frac{h^2}{a^2} + \frac{\mathbf{t}^2}{b^2}\right)\right)$ présente la particularité d'être simultanément factorisée et géométriquement anisotrope (ou même isotrope).

Ces anisotropies ne permettent pas de décrire des phénomènes « propagatifs » ; la covariance d'un phénomène se propageant à vitesse constante v admet une composante $C(h - v\mathbf{t})$ non séparable. Gneiting [2002, 2005] examine différents modèles. La symétrie est dite complète lorsque la covariance ou le variogramme sont des fonctions symétriques de l'incrément spatial et temporel, ce qui s'écrit : $Var(Z(s, t) - Z(s', t')) = Var(Z(s, t') - Z(s', t))$.

Cartographiant la bathymétrie à proximité d'une côte, à partir de mesures effectuées durant presque trois décennies, Bourguin et al. [2001] factorisent ainsi la covariance en trois

termes : une covariance exponentielle pour la composante temporelle, une autre pour la composante spatiale parallèlement à la côte, et le sinus cardinal pour la composante spatiale perpendiculairement à la côte. Cette composante périodique atténuée décrit l'implantation des différentes barres de sables rencontrées à distance croissante de la côte (il s'agit en fait des résidus normés par rapport à la profondeur moyenne sur environ trois décennies). Le modèle temporel n'est pertinent que discrétisé au pas d'un an, car il ne prend pas en charge le mouvement du sable, en particulier la migration des barres aux petits intervalles de temps ; leur modélisation nécessiterait le calage d'une composante de type propagatif aux petites distances spatiales et temporelles.

Pour représenter des phénomènes de pollution, stationnaires temporellement et non stationnaires spatialement, Matheron [1974] a proposé un modèle multivariable et non stationnaire, généralisant le modèle gaussien anamorphosé. La stationnarité temporelle des concentrations permet de définir en chaque station et pour chaque variable i une fonction d'anamorphose $Z(i, x_a, t) = \mathbf{j}_{ix} (Y_{i, x_a, t})$, dont les coefficients du développement hermitien sont interpolés par krigeage entre les stations. Dans ce modèle, une covariance tridimensionnelle « classique » $\mathbf{r}_{ij}(h, \mathbf{t})$ des transformées gaussiennes est alors admissible, le temps n'intervenant plus que comme troisième coordonnée.

Comment valider ce modèle ? Les variogrammes temporels aux points expérimentaux x_a permettent de vérifier l'hypothèse de stationnarité temporelle. Si l'anamorphose « glissante » décrit correctement la non-stationnarité spatiale, les variogrammes spatiaux des transformées gaussiennes $Y_{i, x_a, t}$ (déterminées séparément par station) sont stationnaires. Les variogrammes spatiaux des coefficients du développement hermitien permettent d'en contrôler la régionalisation. On peut vérifier la positivité de l'anamorphose en-dehors des points expérimentaux en traçant la fonction correspondante, ou par la carte des quantiles des concentrations. Enfin, plusieurs critères classiques permettent de vérifier le caractère bi-gaussien des $Y_{i, x_a, t}$.

Ce modèle est généralisable à des phénomènes stationnaires spatialement mais non temporellement, par exemple du fait de la saisonnalité. On infère alors une anamorphose variable temporellement, interpolée à partir des anamorphoses calées aux différentes dates.

Les phénomènes dynamiques

La modélisation spatio-temporelle est indispensable pour prévoir à court terme un niveau de pollution atmosphérique, pour étudier le transfert d'un polluant du sol vers la nappe, ou évaluer le temps d'arrivée d'une pollution à un captage. Une description réaliste s'appuie sur un modèle phénoménologique, via les techniques d'assimilation de données pour des prévisions opérationnelles, ou une résolution du « problème inverse » pour contraindre les paramètres des modèles numériques.

En régime transitoire, Tartakovsky et al. [1998], et Hughson et al. [1998] construisent ainsi un grand nombre de simulations géostatistiques de la transmissivité utilisées en entrée d'un code d'écoulement. Les covariances simples et croisées entre charge et transmissivité, obtenues numériquement, sont alors automatiquement compatibles avec les équations de l'écoulement et les conditions aux limites. Ces covariances bivariées spatio-temporelles sont utilisées pour conditionner la transmissivité aux mesures de charge. Assez commode, cette approche numérique comporte des limitations, comme la fixation de « conditions aux limites » parfois mal connues, mais surtout la difficulté de conditionner directement par les essais de traçage (plus généralement par le transport).

Enfin, des modèles de fonctions aléatoires intrinsèques spatio-temporelles, construits par Matheron (1992) comme les solutions de certaines équations physiques n'ont malheureusement jamais été publiés.

I.3. Dépassement de seuil

La comparaison à un seuil de qualité (norme, valeur guide) est un problème récurrent en environnement. La bibliographie reflète un certain engouement pour les techniques de simulation, y compris lorsqu'un calcul direct de la grandeur recherchée est possible, avec un résultat formellement plus précis. Il semble donc utile de revenir sur les propriétés du modèle gaussien anamorphosé. Dans quels cas le recours aux calculs sur simulations est-il nécessaire ?

Estimation non linéaire ou simulations ?

Dans le modèle gaussien anamorphosé, la concentration Z est la transformée par une fonction d'anamorphose j d'une fonction aléatoire Y de loi spatiale gaussienne réduite : $Z(x) = j(Y(x))$. Lorsque l'anamorphose est strictement croissante sur son domaine de définition, ce que nous supposons désormais dans cette partie, la connaissance des valeurs expérimentales $Z_\alpha = Z(x_\alpha)$ est équivalente à celle des transformées gaussiennes $Y_\alpha = \varphi^{-1}(Z_\alpha)$. La loi conditionnelle de Z se déduit alors directement de celle de Y , ce qui s'écrit $Z(x) = \varphi(Y^*(x) + \sigma_K(x)W(x))$, $Y^*(x)$ désignant le krigeage à moyenne nulle de $Y(x)$ par les Y_α , $\sigma_K(x)$ l'écart-type de l'erreur associé et W une fonction aléatoire de loi spatiale gaussienne réduite (et de covariance non stationnaire), spatialement indépendante de Y^* .

Toute grandeur ne faisant intervenir qu'un seul point est accessible directement à partir de l'histogramme de W , sans recourir à une approximation par un calcul empirique sur simulations. C'est par exemple le cas pour tracer des cartes de quantiles ou de probabilité de dépassement de seuil, ou pour calculer un intervalle de probabilité.

Dans d'autres cas, des approximations réalistes peuvent fournir des bornes. Les problèmes plus complexes de nature géométrique comme la connexité, l'évaluation de fonctions définies en plusieurs points simultanément, ou le calcul des statistiques locales ou globales du rapport de deux variables imposent cependant le recours aux simulations.

Le modèle gaussien discrétisé permet de prendre en compte le changement de support.

A deux dimensions, soit par exemple T la superficie sur laquelle la concentration dépasse le seuil s . T correspond au « minerai » à la coupure s sur le champ D , qui peut être estimé par son espérance conditionnelle $E[T|Z_\alpha] = \int_D P(Z(x) > s | Z_\alpha) dx$. La discrétisation aux nœuds d'une grille fournit une valeur approchée de cette intégrale. Un calcul empirique sur simulations est généralement nécessaire pour évaluer des statistiques globales de T , comme les quantiles ou un intervalle de probabilité.

Un calcul simplifié d'exposition

Considérons le cas d'une brève pollution accidentelle, ce qui permet de négliger les déplacements de la population. La population b_i étant supposée connue aux nœuds d'une grille régulière, l'effectif soumis au dépassement du seuil s est $B_s \approx \sum_i b_i I_{Z(x_i) \geq s}$. Au risque statistique d sur le dépassement local de seuil, la population exposée est estimée par $B_d^*(s) \approx \sum_i b_i I_{P[Z(x_i) \geq s | Z_\alpha] \geq d}$, directement calculable sur le produit de la carte de la population par la carte des probabilités conditionnelles seuillée à d .

Les quantiles de la population exposée $B(s)$ faisant intervenir la loi spatiale de la concentration sont calculés par simulations [Deraisme et Bobbia, 2003]. On note que la moyenne de cette grandeur, directement accessible comme $E[B_s / Z_a] = \sum_i b_i P[Z(x_i) \geq s / Z_a]$, ne constitue pas un estimateur pertinent de la population exposée.

Stationnarité

Loin des points expérimentaux, la loi conditionnelle tend vers la loi a priori. Pour éviter ce calage, l'application du krigeage à moyenne inconnue nécessite d'adapter les calculs pour tenir compte de la corrélation entre le krigeage et l'erreur associée. Dès 1977 P. Chauvet conditionnait des simulations par krigeage ordinaire, avec une correction multiplicative de variance sur le résidu, de la forme $Z_{sc} = Z^{KO} + p[Z_s - Z_s^{KO}]$, avec $p^2 = 1 - \frac{2m}{s_{ko}}$, s désignant une simulation non conditionnelle, sc la simulation conditionnelle, et m le pondérateur de Lagrange. En pratique, on peut vérifier si la covariance ainsi obtenue approche correctement le modèle initial. Rivoirard pour le krigeage lognormal [1991], et Emery pour le modèle gaussien anamorphosé [2006] approchent de même la loi conditionnelle de Z par $\mathbf{j} \left(Y^{KO} + \sqrt{s_{ko}^2 - 2m} W \right)$.

Utilisation d'informations sur le milieu

En environnement, des « variables auxiliaires » connues parfois exhaustivement, comme la densité de population ou l'occupation des sols, expliquent partiellement les concentrations. Celles-ci sont alors décrites par des modèles à résidus, du type $Z(x) = f(x) + t(x)\mathbf{j}(Y(x))$, où $f(x)$ et $t(x)$ sont des fonctions déterministes. Le modèle stationnaire correspond au cas particulier où f et t sont constantes. Les calculs s'adaptent sans difficulté. Par exemple, la probabilité que

la concentration $Z(x)$ dépasse le seuil s s'écrit $P(Z(x) \geq s | Z_a) = 1 - G \left(\frac{\mathbf{j}^{-1} \left(\frac{s-f(x)}{t(x)} \right) - Y^*(x)}{s_K(x)} \right)$.

QUANTIFICATION DE POLLUTIONS

Cette partie présente l'apport de l'analyse exploratoire de données à la quantification de la pollution dans les différents milieux : sols, eaux, air.

II.1. Sols de friches industrielles

II.1.1. Emergence de la problématique « sites et sols pollués »

L'ampleur de la problématique des « sites et sols pollués » s'est révélée ces quinze dernières années. En France la décision de constituer un inventaire des sites pollués ou potentiellement pollués remonte à 1993, la première version des guides pour « l'évaluation simplifiée » et « l'évaluation détaillée des risques » ayant été publiée en 2000. En janvier 2005, le « tableau de bord » du site Internet du ministère français en charge de l'environnement recensait 1470 sites avec une pollution avérée, les sites traités disparaissant de la liste. Les friches industrielles présentées dans cette partie y figurent ou en faisaient partie avant traitement.

Le contexte géologique et hydrogéologique des friches industrielles, leur historique et notamment les remaniements lors du démantèlement des installations, font de chaque site de quelque ampleur un cas particulier. L'évolution des techniques et par suite la diversité des substances rejetées, la présence de drains (canalisations) ou de barrières (fondations), les différents remblais déposés sur le site ont généralement créé de fortes hétérogénéités. Cette importante variabilité spatiale a comme conséquence de fortes incertitudes sur les substances présentes, leur localisation et sur le niveau des teneurs. L'extension verticale des pollutions et leur évolution (migration vers la nappe, niveau d'atténuation « naturelle ») restent généralement très mal caractérisées. Or une connaissance suffisamment précise du « terme source » ainsi que des milieux affectés est indispensable pour évaluer les risques d'exposition (dans les sols, dans l'eau ou dans l'air) ou pour choisir et dimensionner un traitement adapté. Mais malgré l'importance des risques éventuels et des enjeux économiques, en particulier les coûts de remise en état d'un site ou les surcoûts éventuels consécutifs à une dépollution insuffisante, la délimitation des zones polluées et l'évaluation des volumes à traiter sont souvent conduites de façon empirique.

Les premières applications de la géostatistique aux sites pollués portent sur la reconnaissance. Par référence aux méthodes géostatistiques usuelles, Flatman et al. [1984] recommandent un échantillonnage en deux étapes, afin d'adapter la maille aux caractéristiques du site, notamment à une éventuelle anisotropie. Procédant par simulations conditionnelles, Englund et al. [1993] introduisent les coûts d'échantillonnage et différencient les coûts de traitement des sols, pollués ou non pollués (ou supposés tels). Le nombre d'échantillons à prélever en une seule campagne est alors optimisé suivant un critère économique. Cette optimisation « en espérance » paraît peu réaliste, pouvant s'avérer inadaptée pour un site particulier. Le réseau Européen NICOLE aborde de façon élémentaire le dimensionnement d'un maillage par le risque de rater une tache de pollution de géométrie élémentaire et de superficie fixées. Comme d'autres auteurs traitant de pollution dans les différents milieux, Okx et al. [1993] calculent par « probability kriging » le risque de dépassement de seuil, en cokrigeant les indicatrices pour différents seuils par la transformée uniforme de la teneur. Dans l'exemple

donné, une pollution au plomb sur une zone d'environ 30km² échantillonnée à maille kilométrique, la portée change selon l'indicatrice, avec une déstructuration des valeurs élevées par rapport aux valeurs médianes. Un modèle de type diffusif comme le modèle gaussien anamorphosé, malheureusement non testé, pourrait alors améliorer la cohérence des estimations aux différents seuils considérés. Les variogrammes simples présentent un effet de pépite élevé.

Un fort effet de pépite, associé à une structure spatiale d'échelle hectométrique, se retrouve pour les teneurs en PCB (polychlorobiphényles) des sédiments d'une zone portuaire échantillonnée à maille de 150m complétée localement [Reed et al., 2001]. Les auteurs établissent un lien entre teneur et paramètres environnementaux, comme la granulométrie et le Carbone Organique Total ; les estimations sont ensuite effectuées par krigeage.

Les travaux que je présente s'inscrivent dans cette démarche. L'analyse exploratoire apporte une certaine compréhension de l'organisation de la pollution sur le site, ce qui permet de sérier les problèmes et de choisir des méthodes d'estimation appropriées, compte tenu du problème posé (notamment du seuil) et des données disponibles. A l'issue de plusieurs études de sites, le retour d'expérience permet de dégager quelques conclusions générales.

L'échantillonnage effectué pour le diagnostic initial d'un site vise à identifier la pollution, non à la quantifier. Dans un contexte de forte variabilité spatiale, les prévisions effectuées empiriquement à partir de données disparates et partielles se sont révélées éloignées de la réalité des sites, faisant douter la profession de la faisabilité d'une estimation « raisonnée ». Les premières études que j'ai effectuées, financées par les programmes du ministère en charge de l'environnement, comportaient donc un caractère démonstratif : était-il effectivement possible de mettre en évidence une structure spatiale des teneurs, pour des pollutions métalliques (Mortagne, 1995) ou organiques (Homécourt et Thionville, 1998) ; le krigeage pouvait-il s'adapter à la variabilité des sols des friches industrielles ?

Les sites étudiés sont d'abord présentés succinctement. Les principaux résultats sont illustrés par un ou deux exemples. Les développements auxquels ces études ont donné lieu sont enfin précisés.

II.1.2. Les sites étudiés

La présentation suivante utilise entre autres des informations de la base de données BASOL sur les sites actuellement pollués (<http://basol.environnement.gouv.fr>).

Certains sites avaient été échantillonnés spécifiquement pour un travail de recherche.

- la friche de *Mortagne-du-Nord*, divisée en deux par un canal, couvre plus de 20 ha. Les activités de métallurgie des non-ferreux (plomb, zinc) et de chimie (acide sulfurique) ont cessé au cours des années 1960. Démantelé puis remblayé par les débris de démolition et localement par des boues de dragage, ce site présente une pollution en plomb, zinc et cadmium. Sa « pelouse métallicole » de quatre ha est classée réserve naturelle protégée.

Thiry et al. [1997] ont échantillonné le site à maille quasi-régulière de 50m, en vue d'une étude géochimique ; verticalement, l'échantillonnage varie suivant les niveaux rencontrés. Les 94 échantillons sont répartis entre les remblais et les alluvions, sur les deux parties du site. Les oxydes (CuO, Fe₂O₃, PbO, ZnO) et certains éléments traces (Ba, Ni) ont été mesurés.

- le *site X* couvre sur environ 6 ha l'emprise d'installations de traitement des gaz et d'une usine à sous-produits sur une friche de cokeries démantelées, située au Nord de la France. Etudiée par N. Jeannée dans sa thèse [2001] que j'ai dirigée, cette zone avait été échantillonnée à maille 20m sur deux niveaux, un rainurage vertical de fosse de 0.5m de profondeur, prolongé par un sondage (idéalement, un mètre échantillonné en fond de fosse).

Les concentrations des 16 HAP de la liste de l'agence américaine de l'environnement (US EPA, voir par exemple [Kneith et al., 1979]) ont été mesurées en 58 nœuds de grille. Des

mesures de différents types (géophysique, analyse de gaz) sont malheureusement inutilisables car effectuées sur des zones différentes ou suivant un maillage recouvrant insuffisamment le précédent.

- le site Y, d'environ un ha, est situé près de Douai (59). Zone d'épandage d'une cokerie, cette friche remaniée a été échantillonnée par N. Jeannée et P. Wavrer suivant la même méthode que pour le site X [Jeannée, 2001]. Les 52 points expérimentaux sont implantés à maille 10m localement resserrée à 5m.

Pour les autres sites, l'échantillonnage provenait du diagnostic ou de la conduite de la dépollution.

- L'ancien site *pétrochimique*, étudié par A. Prechtel dans son DEA [1998] que j'ai encadré, est situé en France. Il a été échantillonné sur environ 150 ha sur deux niveaux dont la profondeur n'est pas indiquée, suivant un maillage irrégulier d'environ cent cinquante mètres, comportant une soixantaine de sondages répartis sur tout le site. Sur la zone des anciennes unités de production, le maillage est resserré approximativement à 35m, avec environ 130 points. L'unité des teneurs en hydrocarbures totaux n'est pas précisée.

- La *cokerie de Thionville* était située dans un ancien méandre de la Moselle. Après le démantèlement des installations, le site, d'environ onze ha, a été remblayé par divers débris.

Thiry et al. [2000] ont prospecté le site par creusement de fosses à la pelle mécanique jusqu'à une profondeur maximale de six mètres, faisant apparaître une très grande hétérogénéité des remblais, verticalement et latéralement (Figure 1). Par endroit, des huiles et « goudrons » sont retenus au-dessus de morceaux de dalles en béton. Baignant dans la nappe, les alluvions sous-jacentes sont localement imprégnées par des « goudrons » sur environ un mètre d'épaisseur.

Les données, assez mal répertoriées, proviennent d'au moins trois campagnes, menées par différents intervenants. Les premières comportent 66 mesures d'une somme de HAP, et une trentaine pour les métaux (plomb, zinc). Une campagne ultérieure comporte 72 teneurs en HAP, dont la majorité est supposée « ponctuelle », et 25 autres échantillons analysés pour les métaux.

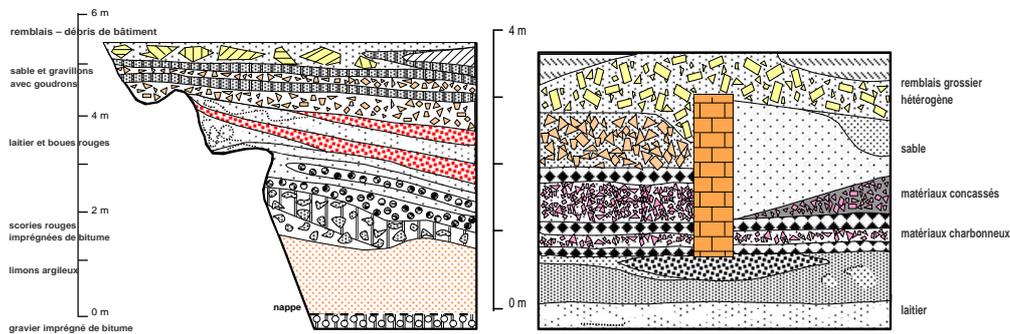


Figure 1. Friche de Thionville, coupe verticale des fosses 700 et 715. L'échelle verticale est adaptée à chaque schéma [Thiry et al., 2000].

- L'ancienne *cokerie d'Homécourt* (54), désormais site atelier pour l'étude de pollutions organiques, présente sur environ neuf ha une pollution avérée par des HAP, des métaux et des cyanures. Les sondages proviennent de deux campagnes mal répertoriées, la technique de prélèvement n'étant pas précisée. Les 56 sondages issus de la première campagne comportent environ quatre-vingt échantillons, et les 49 sondages de la deuxième campagne 55 échantillons. Les teneurs de différents métaux (Pb, Zn, As, Cr) ont été mesurées, ainsi qu'une somme de HAP et le B(a)P.

Enfin, deux sites ont été étudiés avec H. Demougeot-Renard dans le cadre de sa thèse [2004], dont j'ai encadré la partie géostatistique.

- Les sols de l'ancienne *usine à gaz de Gennevilliers* sont constitués de remblais d'épaisseur métrique déposés sur des alluvions. D'une quarantaine d'hectares, ce site a été échantillonné pour le diagnostic par une centaine de sondages verticaux de 4m (environ 70% des sondages) à un mètre de longueur, et à peu près autant de tranchées, mesurés en B(a)P. La dépollution a été conduite localement à partir de plus de 500 échantillons composites de quatre prélèvements chacun, prélevés suivant un maillage aléatoire stratifié mais lacunaire de 1.5mx1.5mx0.5m, principalement dans les deux premiers mètres du sol. A cette étape, la somme des teneurs en HAP était mesurée.

- Le site de 3 ha de l'ancienne *fonderie de plomb de Belin-Béliet* (33), fermée en 1983, avait été acquis par un particulier ; les concentrations y causeront la mort de trois chevaux entre juin 1991 et janvier 1992. Le site a fait l'objet d'une dépollution par l'ADEME, avec excavation partielle. Les données de diagnostic utilisables comportent 67 mesures de la teneur en Pb, sur 97 données appartenant à différents types de sondages ou prélevées en tranchées, irrégulièrement réparties. Les estimations du diagnostic ont pu être comparées aux mesures effectuées pour conduire la dépollution, consistant en 212 échantillons composites de cinq prélèvements chacun, acquis suivant un schéma centré sur des blocs de 10mx10mx0.30m.

Un constat s'impose : dans la majorité des cas, l'échantillonnage disponible pour le diagnostic des sites est disparate (différents types de prélèvements, différents intervenants) et mal répertorié ; en particulier, le support des mesures n'est pas toujours précisé.

II.1.3. Teneurs dans les sols de friches industrielles

Pour chaque site l'analyse exploratoire a été menée avec les outils usuels de la statistique inférentielle : résumés statistiques, transformation des données (logarithme, indicatrices), histogrammes, nuages de corrélation et régression empirique, analyse factorielle multivariable, complétés par différents « outils variographiques ». Les teneurs ont été estimées par krigeage ou cokrigeage. Dans certains cas, la probabilité de dépassement de seuil a été calculée dans le cadre du modèle gaussien anamorphosé, avec ou sans changement de support. Pour montrer les conséquences de l'effet de support ou de l'effet d'information, des simulations ont parfois été effectuées, amenant à revenir sur la détermination de l'anamorphose. Une modélisation bivariable entre teneur et relevés organoleptiques a montré qu'une information auxiliaire qualitative pouvait être utilisée pour améliorer la précision des estimations.

L'échantillonnage disponible pour les premières études était le plus souvent inadapté, insuffisant et mal répertorié, ce qui complique l'analyse exploratoire. Malgré la précision médiocre des estimations, des conclusions générales peuvent être dégagées. Elles sont illustrées, de façon non exhaustive, sur les différents sites.

Hétérogénéité des sites

A Mortagne, reflétant les activités passées, les teneurs en oxydes de zinc et de fer sont plus élevées dans la partie Nord, la partie sud étant plus « riche » en baryum ou en oxyde de cadmium. Les teneurs sont plus faibles dans les alluvions sous-jacentes, le contraste avec les remblais variant selon les substances et la partie du site.

Ces deux principales causes d'hétérogénéité, localisation des activités industrielles passées et formations pédologiques (naturelles ou anthropisées) se retrouvent presque systématiquement sur les friches. Lors de l'échantillonnage, il est indispensable de répertorier correctement les formations rencontrées.

L'historique de site, une information utile mais partielle

En accord avec l'historique de site, de fortes teneurs en plomb ont été mesurées à proximité de l'ancienne fonderie de Belin-Béliet ; sur le site Y, de fortes valeurs en HAP s'observent à proximité des anciennes mares à goudrons. Mais dans les deux cas, l'échantillonnage montre la présence de teneurs élevées en d'autres endroits. Pour le site Y, elles s'expliquent par le dépôt provisoire de remblais d'excavation des mares à goudrons. Le plan des anciennes installations ou l'historique des sites sont des informations utiles mais généralement incomplètes : ces informations indiquent des zones possibles de fortes teneurs mais n'en constituent pas un recensement exhaustif. En particulier, les déplacements de la pollution lors du réaménagement des sites sont souvent oubliés.

L'ancien site pétrochimique est exemplaire. L'échantillonnage est resserré sur la zone des anciennes unités de production, là où des teneurs plus fortes sont attendues. En « surface » les valeurs mesurées y sont effectivement plus élevées. Mais la couverture systématique par un maillage lâche a détecté de fortes teneurs ailleurs sur le site [A. Prechtel, 1998].

Parfois, les campagnes s'organisent d'abord autour des anciennes installations, dans les zones présumées a priori les plus polluées, la zone reconnue étant ensuite élargie. Ces campagnes successives peuvent refléter certaines réticences devant un échantillonnage coûteux et susceptible de dévoiler des problèmes : à court terme, il faut parfois chercher la pollution pour la trouver...

Hétérogénéité des sols

Après fermeture et démantèlement, les sites ont été remblayés par des déchets d'exploitation et des débris de démolition. Ces remblais peuvent présenter de très forts contrastes de teneurs. Quelle est la part de la variabilité aux très petites distances dans la variabilité spatiale ?

Pour le site X, le CNRSSP avait transmis les résultats de l'expérience suivante : un prélèvement à faible profondeur a été réparti en deux bocal, cinq échantillons étant analysés après homogénéisation par bocal. L'histogramme des dix teneurs en HAP issues du même prélèvement recouvre assez largement celui de la soixantaine de mesures réparties sur le site [Jeannée et al., 2000]. Le variogramme horizontal calculé à partir de la maille régulière à 20m confirme qu'un modèle pépitique pur semble dans ce cas adapté, du prélèvement à l'échelle hectométrique. Par ailleurs, la plus forte dispersion des teneurs du bocal le plus « riche » font présumer un effet proportionnel à cette échelle : localement, la variabilité croît avec la moyenne.

A Mortagne, les sondages ont été segmentés en tronçons d'épaisseur variable correspondant aux principales formations rencontrées, les prélèvements étant choisis pour être « représentatifs » de ces tronçons [Thiry et al., 1997]. Pour les prélèvements jugés hétérogènes dans les remblais, constitués entre autres de scories et de débris de creusets (fours, fonderies), deux à trois échantillons ont été analysés ; les mesures multiples dans un même prélèvement sont ainsi préférentielles et associées à une forte variabilité. Les teneurs en oxydes de plomb ou de zinc peuvent varier de plus d'un ordre de grandeur entre échantillons issus d'un même prélèvement. L'hétérogénéité de ces échantillons (environ 1/7^{ème} de l'ensemble des prélèvements) est-elle compatible avec une éventuelle structure spatiale ?

Calculés avec l'ensemble des données de remblais, les variogrammes horizontaux sont effectivement pépitiques. La variabilité est considérablement réduite lorsque l'on retient la teneur moyenne des prélèvements hétérogènes. Le premier point des variogrammes, jugé significatif, est inférieur aux suivants, qui oscillent autour du palier. Classiquement, la structure spatiale est plus apparente pour le zinc, plus soluble. Ces variogrammes sont compatibles avec la présence d'une structure de portée inférieure à la maille de reconnaissance. Reportée à l'origine du variogramme, la moyenne des écarts quadratiques pour les prélèvements

hétérogènes atteint (pour ZnO) ou dépasse (pour PbO) le palier expérimental (voir le gros point à l'origine sur les variogrammes de la figure 2).

Quelques prélèvements hétérogènes génèrent ainsi une variabilité localement très importante, masquant une possible structure horizontale d'échelle décimétrique. Cela correspond par exemple à une stratification subhorizontale de remblais fortement contrastés.

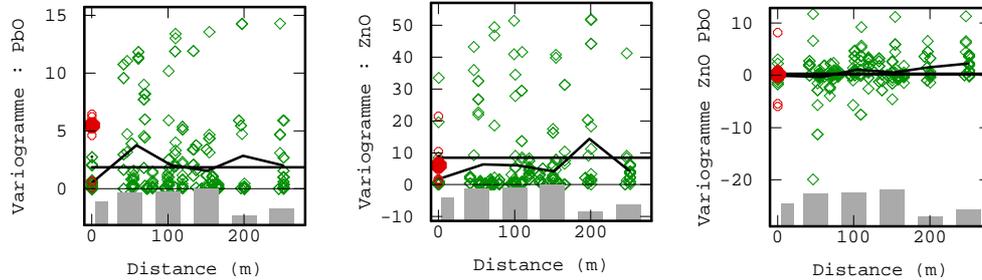


Figure 2. Teneurs en PbO et ZnO dans les remblais sud de la friche de Mortagne, exprimées en %. Comparaison de la variabilité interne aux prélèvements hétérogènes (cercles) à la variabilité spatiale. Nuées variographiques et variogrammes horizontaux simples (à gauche et au centre) et croisé (à droite), calculés sur la teneur moyenne des prélèvements (losanges) ; 41 données dont 5 prélèvements hétérogènes comportant deux ou trois échantillons. Les histogrammes reportés au bas des graphiques donnent une indication relative du nombre de couples par pas du variogramme.

Pour le diagnostic, un échantillonnage disparate

La forte variabilité des teneurs à l'échelle du prélèvement a comme conséquence un effet de support important : la composante pépitique prédominant à l'échelle décimétrique à métrique, le support des mesures influence très fortement leur variance. Or les données disponibles lors du diagnostic proviennent trop souvent de campagnes successives effectuées par différents intervenants, avec différentes techniques de prélèvement (comme à Gennevilliers, Belin-Béliet, Thionville, Homécourt ...). Constituant un échantillon hétérogène, ces différentes données ne doivent pas être mélangées pour l'étude exploratoire, effectuée en séparant les différentes familles de mesures, ou pour la population la plus nombreuse. Lorsque le support n'est pas systématiquement indiqué, il devient également difficile d'utiliser toutes les données pour l'estimation. Malgré le coût élevé de l'échantillonnage, une proportion importante de l'information devient alors inutilisable.

A Gennevilliers par exemple, l'échantillonnage se répartit pour moitié entre « tranchées » et « sondages ». La population statistique la mieux répertoriée, celle des sondages à 4m, représente moins de 40% des données initiales. La raison pour laquelle 30% des sondages sont plus courts, n'est pas précisée. Sur un domaine restreint d'environ 15 ha comportant approximativement une cinquantaine de « sondages » et de « tranchées », la moyenne des teneurs en B(a)P est environ 7 fois plus élevée pour les tranchées que pour les sondages. Il n'a pas été possible de faire la part entre un éventuel biais d'échantillonnage et le caractère possiblement préférentiel des rainurages en paroi des tranchées.

Le cokrigeage à moyennes inconnues permet d'utiliser l'ensemble des informations, les données des tranchées, moins nombreuses étant utilisées comme variable auxiliaire ; un effet de pépite additionnel est appliqué aux sondages « courts » pour tenir compte de la différence de supports. A même nombre de données, des mesures homogènes conduiraient à une estimation plus précise, et aussi plus simple.

Variation des teneurs avec la profondeur

Sur certains sites, les teneurs sont mesurées sur deux niveaux, suivant différentes techniques de prélèvement. Le support variant, on ne peut comparer sans précaution les histogrammes des teneurs pour les deux niveaux, ou les zones de dépassement d'un seuil réglementaire.

Pour le site Y, les statistiques et les nuages de corrélation indiquent des teneurs plus fortes en profondeur, avec un coefficient de corrélation élevé entre les deux niveaux. Pour le site X, la relation est plus complexe, les médianes et les quartiles étant plus forts en profondeur pour les HAP « légers » (deux ou trois cycles) et en surface pour les HAP « lourds » (4 cycles et plus). Pour ces friches de cokerie, on ne peut supposer a priori une variation systématique de la teneur avec la profondeur.

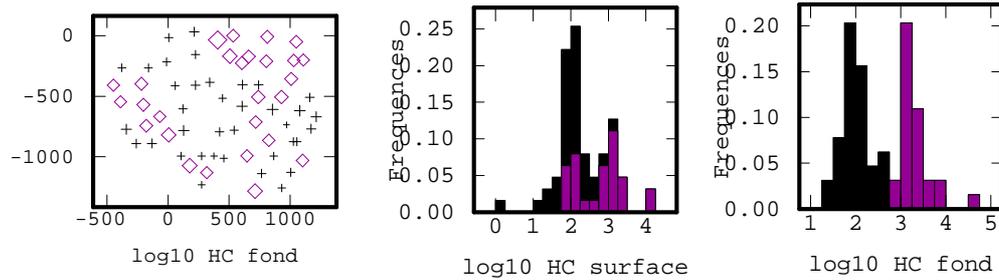


Figure 3. Ancien site pétrochimique, échantillonnage systématique à maille large. Carte d'implantation des prélèvements (à gauche) et histogrammes du logarithme décimal des teneurs en hydrocarbures totaux, en surface (au milieu) et en profondeur (à droite). Les deux modes des teneurs en « profondeur » sont visualisés sur l'histogramme de « surface » ainsi que sur la carte d'implantation (+ pour le mode faible, ? pour le mode fort).

Sur l'ancien site pétrochimique, le mode des fortes teneurs en surface, très atténué, ne permet de détecter qu'une partie des valeurs fortes du niveau plus profond (figure 3). Un échantillonnage de surface fournirait donc une image « optimiste » de la pollution.

La variation de la teneur avec la profondeur change suivant les endroits. Sur l'ensemble du site, les teneurs sont en moyenne deux fois plus fortes en profondeur alors que sur l'emprise des anciennes unités de production, les statistiques sont analogues pour les deux niveaux, indiquant leur possible mélange. Sur cette zone, la moyenne des deux niveaux, légèrement plus élevée que sur l'ensemble du site, traduit une pollution localement plus forte.

L'organisation des teneurs avec la profondeur varie donc suivant les substances et les sites, et parfois localement sur un site. Or fréquemment, seuls les deux à trois premiers mètres de remblais sont reconnus. Restreindre l'échantillonnage aux niveaux superficiels se justifie lorsque le risque provient de ces seuls niveaux, par exemple pour des poussières. Mais lorsque les niveaux plus profonds contribuent au risque, comme pour le transfert à la nappe, il faudrait pouvoir quantifier l'ensemble du « terme source », dans les remblais, les sédiments, etc. Avec un échantillonnage arbitrairement restreint aux tout premiers mètres, sans reconnaissance complémentaire du volume des remblais, l'ordre de grandeur du terme source est inaccessible. Pour de nombreux sites, les volumes pollués restent encore inconnus à l'issue du diagnostic.

Le cas de Thionville est symptomatique (figure 4). La profondeur reconnue par les « anciennes fosses » était limitée à deux mètres. La faible moyenne des teneurs aux profondeurs supérieures à un mètre, peu échantillonnées, résulte d'un effectif trop restreint : en présence d'un histogramme dissymétrique avec une forte proportion de valeur faibles, les teneurs élevées ont une faible probabilité d'apparition dans les niveaux peu échantillonnés. A Thionville, les régressions de la teneur sur la profondeur se raccordent effectivement pour les « anciennes » et les « nouvelles fosses », plus profondes.

A l'étape du diagnostic, la reconnaissance des niveaux plus profonds est indispensable, associée si nécessaire à un échantillonnage suivant un maillage lâche mais systématique.

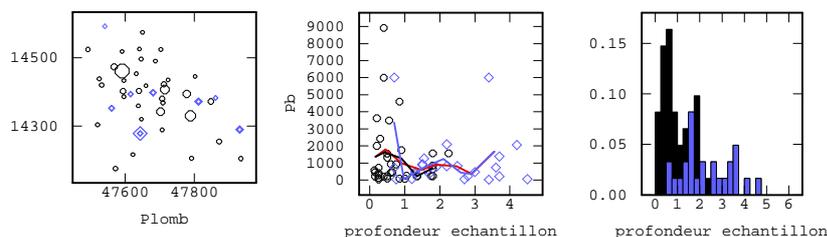


Figure 4. Thionville, teneur en plomb, exprimée en ppm. Implantation des données (à gauche ; les coordonnées sont exprimées en mètres et la taille des symboles est proportionnelle à la teneur). Cercle noirs, les « anciennes fosses », losanges bleus les « nouvelles ». Nuage de corrélation entre teneur et profondeur exprimée en mètre (au milieu). Les moyennes par classe sont reportées respectivement pour les « anciennes fosses », les « nouvelles », et en rouge pour l'ensemble. Histogramme des profondeurs d'échantillons (à droite), avec les « anciennes fosses » en noir et les « nouvelles » en clair.

Echantillonnage en conduite de dépollution

Lors de la dépollution, la sélection des sols s'effectue à partir d'un échantillonnage à maille systématique dans le plan. Pour certains chantiers, en chaque maille, l'échantillonnage (et par suite l'excavation) des niveaux plus profonds est interrompu verticalement dès le premier échantillon rencontré de teneur inférieure à la valeur réglementaire. Ce procédé suppose une décroissance systématique de la teneur avec la profondeur. A Belin-Béliet par exemple, cette décroissance est vérifiée en moyenne sur les échantillons disponibles, mais elle n'est pas systématique ; une dépollution conduite suivant ce protocole risque de s'avérer insuffisante.

Ce type de sélection est fréquent en pratique minière, la reconnaissance étant interrompue après *plusieurs* échantillons successifs de teneur inférieure à la coupure fixée, deux ou trois selon les cas. La dépollution des sites gagnerait à s'inspirer de cette pratique.

Corrélation entre substances

Dans différents contextes, des liaisons sont mises en évidence pour certaines familles de substances. Ainsi, les corrélations entre HAP sont souvent élevées avec des nuages linéaires, la liaison étant d'autant mieux marquée que les structures chimiques sont proches. Sur le site Y, tous les coefficients de corrélation entre HAP de même nombre de cycles sont ainsi supérieurs à 0.9 [Jeannée, 2001]. A Mortagne, de fortes corrélations non linéaires, avec des coefficients supérieurs à 0.8, sont observées entre certains oxydes (CuO-Ba ou CdO-PbO) pour les remblais de la partie Nord. Ces corrélations ne se retrouvent pas dans les remblais de la partie Sud.

Ces fortes liaisons sont compatibles avec l'importante variabilité à petite distance, qui traduit l'hétérogénéité du phénomène plutôt que des erreurs de mesures - ou alors celles-ci sont fortement corrélées, à cause du prélèvement et de la préparation des échantillons. Pour le site X, un modèle d'effets de pépite corrélés est admissible entre HAP.

Les associations entre substances varient avec l'origine de la pollution. Oosterbaan [2000] propose d'utiliser ces « signatures » pour préciser cette origine pour des sites mal répertoriés, ou pour alléger l'échantillonnage multivariable lorsque la typologie du site est connue.

Variabilité spatiale des teneurs

Pour la majorité des sites, un échantillonnage resserré, comme la reconnaissance à maille régulière effectuée pour la conduite de la dépollution, révèle une structure spatiale.

Pour l'ancien site pétrochimique, les données à maille d'environ 150m réparties sur tout le site présentent un variogramme irrégulier mais croissant, avec une composante pépitique importante mais non prépondérante. Sur une partie des anciennes unités de production, le variogramme des teneurs calculé au pas de 20m est bien structuré, avec un effet de pépite modéré et une portée d'environ 100m (figure 5).

L'étude variographique a révélé une zone adjacente d'environ 300m par 150m, dans laquelle quelques teneurs pratiquement nulles alternent avec des valeurs plus fortes ; le variogramme y est pépitique. Le remblai par des matériaux propres mélangés ultérieurement à ceux du site peut expliquer cette particularité, comme cela a été effectivement confirmé.

La forte hétérogénéité de ce site se traduit par des structures spatiales variables selon l'échelle considérée [de Fouquet et al., 2004].

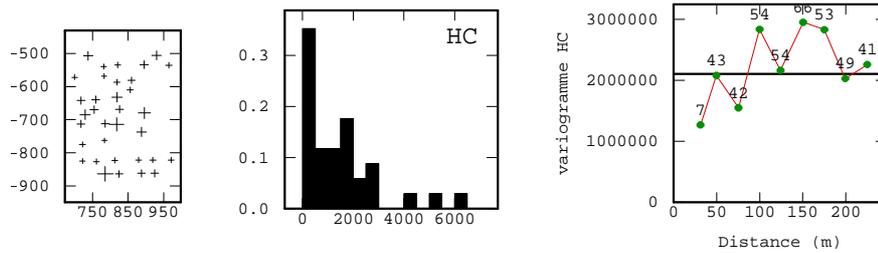


Figure 5. Ancien site pétrochimique, partie du périmètre des anciennes unités de production. Implantation des données (à gauche). La taille des symboles est proportionnelle à la teneur, dont l'unité n'est pas précisée. Histogramme des teneurs en profondeur (au milieu), et variogramme des teneurs (à droite), avec le report du nombre de couples.

La cartographie des teneurs est instructive. Au Nord du site, le krigeage de « blocs » de l'ordre de l'hectare fait ressortir une tache « en profondeur », détectée par un seul échantillon de la maille lâche, qui a été confirmée ultérieurement. Sur l'emprise des anciennes Unités de Production, le resserrement local de la reconnaissance permet l'estimation sur une grille à maille 25 ou 50m, faisant apparaître des contrastes de teneurs. Localement sur la petite étendue pépitique, aucune sélection n'est possible.

Lorsqu'une structure spatiale est détectée, le krigeage permet de différencier différents niveaux de concentrations, à condition de choisir la dimension des blocs en relation avec la portée et avec la maille de reconnaissance. Il est préférable de restreindre la visualisation des estimations aux endroits où l'écart-type de krigeage, absolu ou relatif, reste « acceptable ».

Visualisation des incertitudes

La probabilité de dépassement de seuil, qui peut être considérée comme un critère conventionnel lorsque les hypothèses de calcul ne sont pas tout à fait vérifiées, complète utilement les cartes du krigeage et de l'écart-type d'estimation. En référence à un seuil fixé, les valeurs intermédiaires de cette probabilité signalent une forte incertitude, pouvant guider un échantillonnage complémentaire. Souvent très dispersé, le nuage de corrélation entre teneur estimée et probabilité de dépassement de seuil, fournit une visualisation de l'incertitude entachant l'estimation de la teneur [de Fouquet, 1999].

Conciliation des prévisions à l'échantillonnage en dépollution

Dans sa thèse, H. Demougeot-Renard a tenté une « conciliation » entre l'estimation à partir des données du diagnostic et les mesures effectuées en conduite de dépollution [2004].

Pour l'un des sites, ce recalage s'avère impossible du fait de l'échantillonnage. En effet, les teneurs en B(a)P (mesurées pour le diagnostic) et en HAP (mesurées en dépollution) ne sont simultanément disponibles que pour une trentaine d'échantillons, ce qui est insuffisant pour

établir un modèle bivariable B(a)P-HAP. Mais ces données suffisent pour vérifier que la relation empirique utilisée sur le site pour convertir les mesures de HAP en teneurs en B(a)P, principal responsable en terme de risque sanitaire, surestime largement cette teneur.

A Belin-Béliet, sur les quatre niveaux de 0.30m d'épaisseur, les zones dépolluées se superposent de façon satisfaisante aux zones pour lesquelles la probabilité de dépassement de seuil, calculée à partir des données du diagnostic, est forte ou moyenne. L'estimation géostatistique indiquant une probabilité élevée sur d'autres secteurs, non dépollués, un contrôle sur le site aurait été utile.

II.1.4. Recommandations pour l'échantillonnage

A l'issue de ces études, des recommandations pour l'échantillonnage, assez « classiques », sont rappelées.

Après une reconnaissance préliminaire d'identification de la nature des pollutions, une étape d'échantillonnage systématique est indispensable. Une couverture systématique permet de détecter des zones de concentrations élevées non répertoriées par l'historique. La profondeur échantillonnée doit être suffisante et non arbitraire. Un échantillonnage en plusieurs étapes est préférable, pour adapter la maille au niveau et à la variabilité des teneurs dans les différentes formations. L'échantillonnage dépend également de l'objectif visé. Par exemple,

- pour évaluer le terme source des modèles hydro-géochimiques, une couverture régulière systématique, permettant de quantifier le volume total de remblais et leurs concentrations, est nécessaire ;
- pour cartographier les teneurs avec une précision fixée, en présence d'un effet proportionnel, on échantillonnera de façon plus dense les zones de fortes teneurs ;
- pour dimensionner un chantier de dépollution, en présence d'une structure spatiale suffisamment marquée, une première cartographie à partir d'une maille régulière sera utilement complétée par la carte de probabilité de dépassement du seuil réglementaire ; l'échantillonnage sera alors resserré dans les zones de transition autour du seuil, plutôt que dans celles de concentrations élevées par rapport au seuil.

Pour réduire les coûts d'analyse et améliorer la « représentativité » locale des échantillons, la dépollution est souvent conduite à partir d'échantillons composites : plusieurs prélèvements rapprochés sont mélangés, ce mélange étant échantillonné après homogénéisation. En dehors de cas « sensibles », comme l'implantation éventuelle d'une école ou d'une zone résidentielle, lors du diagnostic également, une reconnaissance par des échantillons composites prélevés suivant un « motif » fixé et convenablement homogénéisés, permettrait de mieux préciser la structure spatiale à l'échelle décimétrique à hectométrique.

Enfin, le resserrement local de la maille en quelques « croix de sondages » réparties sur le site est indispensable pour préciser la variabilité aux petites distances, et donc la faisabilité d'une sélection à l'échelle plurimétrique.

II.1.5. Améliorations et développements méthodologiques

Traitement de la limite de quantification

Les mesures de pollution font fréquemment référence à une « limite de quantification », en dessous de laquelle des traces de la substance peuvent être présentes à des niveaux très faibles. Suivant les milieux, les sites et les substances, ces données représentent seulement quelques échantillons ou constituent la majorité des données. Comment traiter cette information ?

Une méthode consiste à « reconstituer » les données inférieures à la limite de détection, par simulation sous contrainte d'intervalles, conditionnellement aux valeurs quantifiées. A une anamorphose près, le modèle gaussien seuillé, avec randomisation des valeurs entre 0 et la ou

les limites de quantification est applicable. Une telle modélisation peut être utile lorsque la limite de quantification est proche du seuil de qualité. Sur les exemples traités, aussi bien pour les sols que pour l'eau, ce raffinement apparaît inutile, les limites de quantification restant faibles relativement au seuil de qualité.

En pratique, on compare les statistiques pour deux hypothèses, l'une haute (en ramenant les valeurs à la limite de quantification), l'autre basse (en ramenant ces valeurs à 0 pour obtenir une borne inférieure, ou en les randomisant sans tenir compte de la structure spatiale). Pour les applications, les estimations sont plutôt effectuées sous la seule hypothèse « haute ».

Les pesticides dans les eaux posent un problème particulier : la réglementation fixe un seuil par substance, ainsi que sur la somme des substances mesurées, mais sans en préciser les termes ni même le nombre. Or plus d'une quarantaine de substances pouvant être recherchées, la somme des limites de quantification correspondantes peut dépasser le seuil de qualité. Dans ce cas, une modélisation de type gaussien seuillé, tenant compte des corrélations entre substances établies sur les mesures quantifiées est envisageable. Actuellement, les mesures sont trop peu systématiques pour tenter une modélisation.

Anamorphose

La probabilité de dépassement de seuil en modèle multigaussien se calcule directement sur les transformées gaussiennes des teneurs ; le seuil gaussien correspondant peut se déduire de façon cohérente de cette transformation, par une interpolation linéaire locale. Pour les simulations, l'expression analytique de l'anamorphose est nécessaire pour le retour des gaussiennes conditionnées en teneurs. Sous différentes variantes [Rivoirard, 1991 ; Chilès et Delfiner, 1999] le calcul se décompose en deux étapes :

- classement des teneurs z_a , et détermination des gaussiennes y_a de même quantile ;
- calcul de la fonction d'anamorphose j , par exemple via son développement hermitien.

L'écart $z_a - j(y_a)$ fournit un critère de qualité de l'ajustement. Lorsque l'anamorphose est déterminée à partir de données peu nombreuses ou fortement dispersées, cet écart n'est plus négligeable et peut provoquer des artefacts : les simulations ponctuelles présentent une valeur inférieure au seuil là où la donnée dépasse le seuil, ou réciproquement.

Cet artefact est facilement corrigé en transformant les teneurs avant anamorphose de façon que leur histogramme se rapproche de celui d'une variable normale. Effectuer l'anamorphose sur le logarithme translaté des teneurs permet de limiter le développement hermitien à quelques polynômes, la transformation composée étant effectivement monotone. Cette correction a été utilisée de façon satisfaisante pour simuler des perméabilités [rapport d'étude de Fouquet et Courtois, 1997] ainsi que des teneurs en polluants dans des sols [Demougeot-Renard, 2004]. Elle ne s'étend pas au cas d'un changement de support.

Inférence de la structure spatiale : le variogramme et ses variantes [N. Jeannée et al., 2000]

En pollution de sols, les échantillons statistiquement homogènes sont parfois peu nombreux. Les variogrammes expérimentaux présentent généralement un effet de pépite élevé, ainsi qu'une « légère structure spatiale ». Le variogramme expérimental permet-il d'inférer raisonnablement le variogramme régional dans le contexte de distributions fortement dissymétriques, ou l'interprétation tend-elle à exagérer la structuration spatiale ?

Afin d'atténuer l'influence des fortes valeurs sur l'écart quadratique moyen, l'analyse variographique est menée sur les teneurs ou sur leurs transformées : logarithme translaté, transformées gaussiennes, indicatrices pour différents quantiles, etc. En l'absence de structure spatiale des transformées, on suppose l'absence de structure spatiale des teneurs.

Confronté en halieutique à un échantillonnage anisotrope et à de forts contrastes de valeurs, Rivoirard [2000] a examiné diverses pondérations pour le calcul du variogramme, et introduit le calcul du variogramme expérimental à partir de la covariance non centrée $C(h)$:

$g(h) = C(0) - C(h)$. Contrairement au calcul usuel, ce variogramme expérimental n'est pas nécessairement positif.

Les propriétés des deux estimateurs du variogramme ont été examinées sur simulation, en étudiant notamment la dispersion de la nuée variographique et l'écart interquantile associé. Aux petites distances, la nuée du variogramme est moins dispersée que celle de la covariance (elle est nulle à l'origine), puis les résultats s'inversent à partir des «distances intermédiaires» relativement à la portée. Le calcul sur simulation de la variance d'estimation du variogramme expérimental a été validé par confrontation à des résultats antérieurs [Matheron, 1965 ; Alfaro 1979 ; Chilès et Delfiner, 1999].

Le variogramme « classique » reste préférable pour préciser le comportement à l'origine ainsi que le degré de stationnarité ; lorsque l'hypothèse de stationnarité d'ordre deux est admissible, le variogramme déduit de la covariance non centrée est utile pour préciser le palier.

Valorisation de l'information qualitative [N. Jeannée et al., 2003]

L'importante variabilité spatiale des teneurs rend les estimations peu précises. Cependant des observations qualitatives, potentiellement abondantes car peu coûteuses, peuvent renseigner sur le niveau de la pollution : parfois une pollution se voit ou se sent.

Lors de l'échantillonnage du site Y, des relevés organoleptiques avaient été effectués : présence de traces de goudrons, odeur (absente, légère ou forte), présence de débris de maçonnerie, présence de charbon dans les remblais ... Les statistiques usuelles entre variable catégorielle et variable quantitative (moyenne des rangs, moyenne et médiane par classe, etc. [Saporta, 1990]) montrent une liaison marquée entre la teneur et deux indices, la présence de goudron (ce qui est peu surprenant), et la présence de charbon dans les remblais. Le rapport de corrélation fournit un critère objectif de classement des indices selon leur caractère « explicatif ». Z désignant la variable quantitative et A une variable catégorielle, ce rapport est défini comme $h^2_{Z/A} = Var(E(Z/A)) / VarZ$ et compris entre 0 et 1. Des transformées des variables catégorielles, comme les facteurs d'une analyse de correspondances, présentent parfois un meilleur caractère explicatif.

Comment utiliser ces informations, notées A ? En l'absence de corrélation différée, l'espérance conditionnelle, autrement dit la moyenne par classe, est la fonction des observations la mieux corrélée point à point à la teneur Z . Cette régression $H(x) = E[Z(x)/A(x)]$ évite le codage arbitraire des variables catégorielles et permet de les convertir en une variable de même « dimension » que Z . En pratique, il est nécessaire de vérifier que la régression reste stationnaire sur le champ. On peut alors cokriquer la teneur Z par H .

Dans cette décomposition, la teneur Z est une dispersée de H . A même « minéral », la sélection d'après les indices organoleptiques (sur H) conduit à une moindre récupération de la pollution que la sélection sur la teneur Z .

Pour le site Y, l'expérimentation par validation croisée montre que le cokrigeage à partir des indices organoleptiques améliore l'estimation. Il serait intéressant de valider ces résultats sur un autre site pour lequel l'information qualitative serait plus abondante, et chercher comment adapter l'échantillonnage en conséquence.

Ce modèle simple présente plusieurs avantages : les hypothèses sont commodes à tester, même avec des données peu nombreuses, et sa souplesse le rend facilement adaptable, par exemple par des regroupements de classes.

Un modèle légèrement différent avait été utilisé pour cokriquer des variables quantitatives, densité d'arbres et surface terrière, par des observations plus nombreuses issues de l'interprétation de photographies aériennes, donnant l'âge et le type de peuplement forestier. Avec seulement un quart des mesures au sol, la précision du cokrigeage se détériorait faiblement [de Fouquet et Mandallaz, 1992].

Un modèle plus complexe, la corégionalisation d'une teneur et d'une indicatrice en modèle gaussien anamorphosé, est présenté au chapitre III.

La mise au point de méthodes de reconnaissance peu coûteuse pour localiser la pollution ou suivre le traitement d'un site constitue un enjeu important en pollution des sols.

II.2. Pollution des eaux

Des substances toujours plus nombreuses sont désormais recherchées dans les eaux superficielles et dans les nappes : nutriments (composés azotés comme les nitrates et les nitrites, composés phosphorés), produits phytosanitaires ou pesticides (insecticides, herbicides et fongicides), PCB (Polychlorobiphényles), certains médicaments, etc.

J'ai abordé la pollution en pesticides dans les nappes, et traité de façon plus détaillée celle en nutriments dans les cours d'eaux. La première section regroupe deux études de nappes, la cartographie du toit de la nappe d'Alsace (qui rejoint la problématique de l'anthropisation), et l'estimation des moyennes annuelles en pesticides dans la nappe de la craie.

II.2.1. Cartographie de nappes

Toit de la nappe d'Alsace [rapport d'étude de Fouquet & Lozzi, 2001]

La question posée portait sur l'amélioration de l'estimation *spatiale* du toit de la nappe, et sur l'évaluation de la précision associée, en vue de réduire si possible le nombre de stations de mesure : comment cartographier « au mieux » cette surface à partir des mesures à une date donnée ? La bonne synchronisation des mesures explique le cadre strictement spatial imposé.

L'étude variographique offre un bel exemple de la puissance des outils comme la nuée variographique pour détecter les zones présentant un comportement particulier. L'amplitude du battement de la nappe, la diversité de structure des variogrammes temporels de la cote piézométrique (palier, amplitude relative de la composante périodique annuelle) font apparaître un fonctionnement différencié de la nappe, en particulier vers certaines bordures.

L'écart-type de krigeage est fortement réduit lorsque l'altitude est utilisée en dérive externe, ou si l'on remplace l'estimation de la charge, fortement non stationnaire, par celle de la profondeur, dont le variogramme est effectivement stationnaire. Dans les deux cas, une régularisée du relief, malheureusement non disponible alors, aurait pu améliorer les résultats.

Moyenne annuelle en pesticides dans la nappe de la Craie [de Fouquet & D. Renard, 2005]

Malgré l'abondance des données collectées dans les nappes ces dernières années, l'évaluation des concentrations en pesticides et de leur évolution reste difficile. Les difficultés sont de différents ordres :

- phénoménologiques : les pesticides recouvrent une grande diversité de produits ainsi que leurs dérivés, épandus à des périodes et sur des durées différentes. Les propriétés physico-chimiques de ces différentes substances restent mal connues.

- analytiques : l'amplitude de l'erreur d'échantillonnage (du prélèvement au résultat d'analyse) évolue dans le temps, et varie selon les laboratoires. Les évaluations actuelles indiquent une incertitude analytique de 15 à 20% [Amalric et al., 2004].

- échantillonnage : le forage d'un puits répond d'abord à une demande d'alimentation, donc à la recherche de bas niveaux de pollution. Les chroniques par station sont réduites dans le temps. La fréquence des prélèvements varie selon les réseaux de mesure (réseau de « connaissance générale » ou patrimonial, ou réseau de surveillance de la qualité pour l'eau potable) ; elle est parfois adaptée en fonction des résultats des mesures. Les substances recherchées varient au cours du temps. Spatialement et temporellement, l'échantillonnage est pour partie préférentiel.

Avec quelle précision peut-on cartographier les concentrations ? Des évolutions temporelles ou des variations saisonnières sont-elles détectables ? Des relations apparaissent-elles entre le « milieu » et le niveau des différentes substances ?

Sa grande extension, le nombre important de mesures disponibles, ainsi que sa vulnérabilité aux pollutions à cause de la rapidité du transfert vers la nappe, ont conduit l'Institut Français de l'Environnement, en coordination avec le BRGM, à retenir la nappe de la Craie du Nord pour essayer de quantifier les concentrations en pesticides, et examiner les relations avec le milieu.

Interdite depuis plusieurs années, l'atrazine décroît en moyenne linéairement sur une période de cinq ans (1998-2002), mais avec de fortes fluctuations temporelles ainsi que par forage ; courant 2002, des niveaux supérieurs au seuil réglementaire de $0,1\text{mg/l}$ par substance sont encore mesurés. Sur les données du réseau patrimonial, une structure spatiale et temporelle a été mise en évidence, pour l'atrazine et son dérivé, l'atrazine déséthyl, corrélés positivement. Bien que les écarts restent proches des incertitudes de mesure (limite de quantification, incertitude analytique), il a été montré que pour l'atrazine, l'estimation à partir des données du « réseau patrimonial » vers le « réseau d'usage » était sans biais, alors qu'en sens inverse, les mesures du réseau d'usage tendent à sous-estimer les niveaux aux points du réseau patrimonial.

Aucune influence saisonnière n'a pu être établie sur les concentrations mesurées. Malgré la prise en compte de différents critères comme l>IDPR (Indice de Développement et de Persistance du sous-sol, élaboré par le BRGM) la seule liaison détectée entre concentration en atrazine et milieu est celle, assez ténue, avec la proportion de la superficie du bassin versant à l'amont du forage occupée par la culture de maïs. Cette faible liaison peut résulter des usages non agricoles de cette substance.

A cause du nombre de substances potentiellement considérées (plus d'une quarantaine), et de la limite de quantification (actuellement autour de $0,01$ à $0,02\text{mg/l}$), la cartographie de la probabilité que la somme des concentrations dépasse le seuil réglementaire ($0,5\text{mg/l}$) est un intéressant problème en suspens.

II.2.2. Vers un réseau représentatif de la qualité des cours d'eau

Contexte réglementaire et technique

Le travail sur les cours d'eau est motivé par l'évolution de la réglementation, suite à la directive cadre européenne sur l'eau (décembre 2000) qui impose un « bon état des masses d'eau » à l'horizon 2015. Pour le volet « chimie », les seuils réglementaires concernent la moyenne annuelle des concentrations pour les « micropolluants » comme les pesticides, et les quantiles d'ordre élevé pour les nutriments. Le « rapportage » actuellement préconisé est fondé sur des calculs statistiques usuels mais inadaptés. Dans ce contexte, les questions suivantes ont été examinées:

- est-il possible de construire un réseau de « stations » représentatif des cours d'eau à densité et précisions constantes, pour les nutriments ?

- avec quelle précision la moyenne et le quantile 90 annuels de la concentration en nutriments par station sont-ils estimés ?

Des modèles de fonctions aléatoires définies sur un support arborescent ont été développés pour le calcul des moyennes spatiales des concentrations le long des réseaux hydrographiques.

Relations entre concentrations et milieux [rapport d'étude : de Fouquet, 2000]

Afin d'assurer une homogénéité aux statistiques fournies par les différents états, l'Agence Européenne de l'Environnement préconise la construction d'un « réseau

représentatif » le long des cours d'eau, comportant en moyenne une station pour 1000 km². En France métropolitaine l'effectif de ce réseau « Eurowaternet » est largement inférieur aux 1200 stations du Réseau National de Bassins (RNB) ou aux 1660 stations de la Base Nationale de Données sur l'Eau. En 1999, une présélection de 550 stations avait été effectuée par l'IFEN en croisant différents critères de stratification, comme la classe des cours d'eau, caractérisée par la superficie des bassins versants, et le type de « pression » anthropique (principalement agricole ou urbaine) caractérisé par l'occupation des sols ou la densité de population. Mais Léonard et Crouzet [1999] posaient la question de la valorisation de l'ensemble des stations du RNB : pour une meilleure précision, une pondération adaptée utilisant l'ensemble des stations n'était-elle pas préférable à la sélection recommandée ?

Une étape d'analyse exploratoire a d'abord validé la présélection de l'IFEN, qui corrigeait l'implantation préférentielle des stations dans les bassins sous forte pression agricole ou urbaine. Peu échantillonnés, les bassins sous faible pression anthropique restent sous-représentés. Les graphes de quantiles ou « q-q plot » facilitent grandement la comparaison détaillée des différentes populations : tous bassins versants, toutes stations, stations présélectionnées.

L'étude des relations entre les concentrations et les paramètres de stratification, cumulés sur les bassins à l'amont, permet de quantifier des relations attendues :

- croissance de la concentration en nitrates avec la proportion d'occupation des sols par l'agriculture intensive;
- croissance de l'ammonium ou de l'orthophosphate avec la densité de population;
- décroissance de ces valeurs lorsque la proportion de couvert forestier, indicateur de faibles pressions, augmente.

La régression des concentrations sur les paramètres de stratification n'a mis en évidence aucun seuil particulier ; les limites définissant les strates de pression sont donc conventionnelles.

Définition de la moyenne annuelle

Pour les rendus statistiques, la moyenne annuelle par station est actuellement calculée comme la moyenne arithmétique des concentrations mesurées durant l'année, la variance d'estimation associée étant la variance statistique de ces valeurs rapportée à leur nombre. La moyenne par strate est la moyenne arithmétique des moyennes annuelles des stations de la strate. Implicitement, ces calculs reposent sur les hypothèses suivantes :

- par station, les concentrations sont des tirages indépendants d'une même variable aléatoire. L'espérance m et la variance sont donc supposées constantes ;
- la grandeur à estimer est ce paramètre m de la loi de la variable;
- les valeurs mesurées ne sont corrélées ni temporellement par station, ni spatialement entre stations.

Les données expérimentales montrent que les concentrations en nutriments présentent généralement des variations saisonnières marquées (figure 6). Pour les nitrates les concentrations « hivernales » sont généralement plus élevées que les valeurs « estivales », à cause du ralentissement de l'activité végétale coïncidant avec un ruissellement plus important. L'hypothèse d'une espérance constante dans le temps n'est donc pas réaliste. La plupart des variogrammes expérimentaux présentent d'ailleurs une composante temporelle structurée en plus de la composante périodique annuelle. L'hypothèse d'indépendance temporelle des concentrations par station est donc mise en défaut.

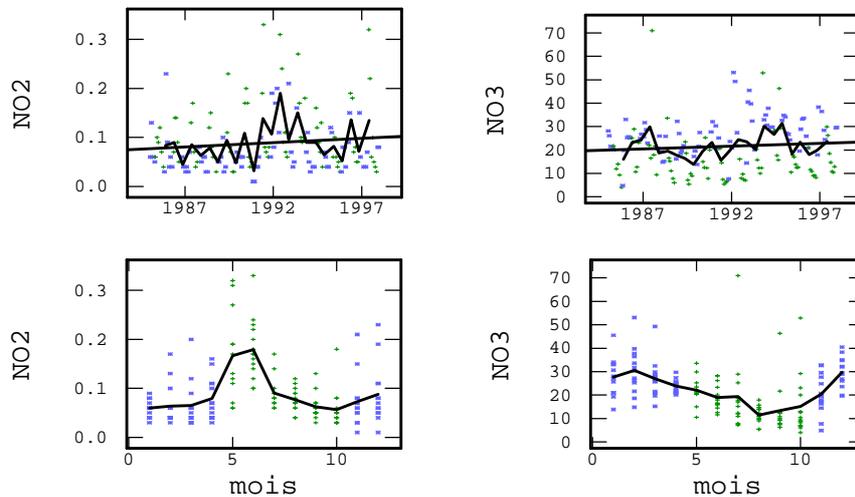
Les hypothèses précédentes ne sont cependant pas nécessaires. En effet, si $z(t)$ désigne la concentration en une station au cours du temps, la moyenne annuelle durant l'année

$[t_0, t_0 + T]$ est l'intégrale temporelle $z_T = \frac{1}{T} \int_{t_0}^{t_0+T} z(t) dt$. La version probabiliste de cette grandeur

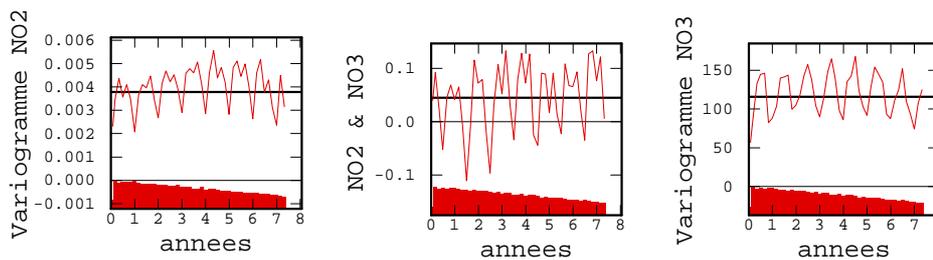
est l'intégrale stochastique $Z_T = \frac{1}{T} \int_{t_0}^{t_0+T} Z(t) dt$, qui reste définie en-dehors de toute hypothèse de stationnarité sur la fonction aléatoire Z . Lorsque Z admet une espérance m stationnaire, alors la relation entre ce paramètre et la moyenne annuelle recherchée est $E(Z_T) = m$. Appliqué sans recul, le calcul statistique usuel « se trompe » sur la variable à estimer.

Z_T^* désignant l'estimateur retenu, la variance d'estimation est $Var(Z_T - Z_T^*)$. Son calcul à l'aide du variogramme suppose nulle l'espérance des accroissements. Les hypothèses réellement nécessaires sont en fait plus faibles, car on vérifie que cette variance diffère peu du calcul approché en supposant l'indépendance des erreurs d'extension élémentaires. La stationnarité de la moyenne ne porte plus alors que sur le plus grand intervalle de temps entre deux mesures, soit environ trois mois dans le cas très défavorable d'une mesure par trimestre.

En de nombreuses stations, pour surveiller les fortes concentrations en nitrates, les mesures sont plus fréquentes en « hiver » (deux mesures par mois) qu'en « été » (une mesure par mois). L'échantillonnage étant préférentiel, la moyenne arithmétique fournit alors une estimation biaisée, facilement corrigée par un krigeage ou par une simple pondération par segment d'influence.



a) évolution temporelle des concentrations de 1985 à 1997 avec report des moyennes semestrielles (en haut) et moyennes mensuelles, montrant des valeurs supérieures en été pour les nitrites et en hiver pour les nitrates (en bas). *: valeurs « hivernales », de novembre à avril, + : valeurs « estivales », de mai à octobre.



b) variogrammes simples et variogramme croisé (au milieu), calculés au pas de deux mois. La composante périodique annuelle des concentrations apparaît clairement.

Figure 6. Concentrations en nitrites et en nitrates, exprimées en mg/l, mesurées à la station 2078 sur la Moselle. Le coefficient de corrélation est égal à 0.07

Variance d'estimation par strate [rapports d'étude : C. de Fouquet, 2000 ; C. de Fouquet et N. Bez, 2001]

La moyenne des valeurs annuelles estimées pour les stations de la strate n'est pas le seul estimateur possible de la moyenne annuelle par strate. Des pondérations sont envisageables, mais sur quel critère ? Le calcul géostatistique de la variance d'estimation impose de *définir la grandeur que l'on voudrait connaître parfaitement*.

Calculer la variance d'estimation de la strate comme la variance de la moyenne des erreurs temporelles commises sur les stations dans la strate revient à négliger la composante spatiale de la variance d'estimation. Si le plan représente l'espace géographique et l'axe vertical le temps, ce calcul ne prend en compte que le « terme de ligne » de la variance d'estimation, en oubliant le « terme de tranche ». Expérimentalement, les variogrammes spatiaux à l'échelle de quelques centaines de kilomètres, ou les variogrammes temporels sur une à quelques années sont du même ordre de grandeur. Le terme de tranche n'est donc pas négligeable a priori.

Assimiler une strate à la réunion des stations qu'elle contient revient à définir la strate à partir d'un mécanisme de calcul. D'autres définitions sont envisageables :

- la réunion des bassins versants correspondant aux critères de stratification du milieu. Le réseau hydrographique n'est pas pris spécifiquement en compte dans cette définition, qui ramène le calcul à une estimation dans le plan à partir d'un semis de points ;

- la réunion des tronçons de cours d'eau intersectant les bassins versants précédents. Cette dernière définition paraît physiquement plus pertinente. Elle nécessite de construire des modèles de fonctions aléatoires valides sur un réseau hydrographique.

Quelle que soit la définition des strates, il est impossible de construire un réseau à densité (linéaire ou surfacique) et à précision constantes. En effet, les variogrammes spatiaux (linéaires ou surfaciques) diffèrent fortement selon les « pressions », le palier croissant généralement avec la concentration. A précision constante, il faut alors adapter le réseau à l'occupation du sol. Mais les polluants concernés varient selon les pressions, les nitrates étant plus élevés dans les régions d'agriculture intensive, et l'orthophosphate sous influence urbaine. A précision constante, le réseau devrait donc être plus dense pour les nitrates dans les régions agricoles, et pour les phosphates, dans les régions urbaines. Or des réseaux différents selon les substances ne sont pas envisageables actuellement.

Une révision des objectifs assignés au réseau est donc nécessaire.

II.2.3. Moyenne et quantiles par station [C. Bernard-Michel et al., 2005]

Correction du biais pour l'estimation de la moyenne et des quantiles

Les estimations de la moyenne annuelle par krigeage et par pondération par segments d'influence sont souvent très proches. Le krigeage fournit la variance d'estimation associée, ainsi que celle sur l'écart interannuel lorsque celui-ci est estimé directement. La figure 7 montre l'influence de la pondération, l'écart à la moyenne statistique étant généralement bien marqué dans le cas d'un échantillonnage préférentiel, avec des mesures deux fois plus fréquentes en hiver (soit 18 mesures par an). La pondération diminue la moyenne pour les nitrates ; son influence est également sensible pour les autres nutriments. A cause de la composante périodique annuelle, prise en charge par le variogramme, la variance d'estimation géostatistique est largement inférieure à la variance prévue par le calcul classique. Dans certains cas, l'estimateur classique se situe en-dehors de l'intervalle de probabilité construit à partir du krigeage, en posant conventionnellement une hypothèse d'erreur gaussienne.

Pour qualifier les plus « mauvais états » observés durant l'année, la réglementation sur les nutriments préconise le quantile 90% comme indicateur. Or même dans le cas de variables indépendantes, les estimateurs de quantiles sont biaisés, avec un biais d'autant plus marqué que

Quelle distribution inférer ?

Durant l'année, la moyenne présente des variations saisonnières systématiques, l'amplitude des fluctuations étant généralement croissante avec la concentration. A la composante saisonnière, se superpose une corrélation temporelle. A quelle loi se rapporte le quantile recherché ?

Le modèle est le suivant. $z(t)$ désignant la variable régionalisée, désignons par \mathbf{q} une variable uniforme sur $[t_0, t_0 + T]$. La concentration en un point choisi aléatoirement et de façon équiprobable durant l'année, est une variable aléatoire $z(\mathbf{q})$, d'espérance $E[z(\mathbf{q})] = z_T$ et de variance $s^2(0/T) = \frac{1}{T} \int_{t_0}^{t_0+T} (z(t) - z_T)^2 dt$. Déconditionnons par rapport à la réalisation : $E[Z(\mathbf{q})] = Z_T$ est une variable aléatoire, régularisée de la concentration sur l'année considérée.

La variance de $Z(\mathbf{q})$, notée $S^2(0/T) = \frac{1}{T} \int_{t_0}^{t_0+T} (Z(t) - Z_T)^2 dt$ est également une variable aléatoire. La moyenne annuelle et la variance des concentrations « instantanées » ont comme espérance respective $E[Z_T] = \frac{1}{T} \int_{t_0}^{t_0+T} E[Z(t)] dt$ et $E[S^2(0/T)] = D^2(0/T)$. En moyenne sur les réalisations, la variance de la distribution annuelle est la variance de dispersion des concentrations « instantanées » durant une année [Matheron, 1965]. La définition de ces grandeurs ne nécessite aucune hypothèse de stationnarité ; leur calcul peut s'effectuer en modèle localement intrinsèque.

Retrouver en moyenne la variance de dispersion peut fournir un critère de calage de la fonction de répartition, par exemple pour choisir le maximum admissible de z durant l'année.

II.2.4. Modèles spatiaux de débits et de concentrations le long des cours d'eau [C. de Fouquet et C. Bernard-Michel, 2006]

Une fois estimés les moyennes annuelles ou les quantiles par station, comment étendre les résultats le long des cours d'eau ? Modéliser des concentrations ou des débits sur un réseau hydrographique nécessite de construire des covariances ou des variogrammes spécifiques. En effet, les modèles usuels ne sont généralement plus valides par changement de topologie, leur caractérisation spectrale faisant intervenir explicitement la distance euclidienne. Dans la suite, seule la composante spatiale est examinée, ce qui est suffisant pour décrire le débit annuel moyen ou la concentration annuelle moyenne le long d'un réseau hydrographique.

Des résultats théoriques ont été récemment établis pour des fonctions aléatoires définies sur les sommets d'un graphe [Kondor R. et al., 2004]. Le problème se pose différemment lorsque le support est continu.

R. Bruno et al. [2001] décomposent la concentration en vanadium des sédiments de rivières en deux composantes, un apport régional défini en deux dimensions, et un apport amont, défini le long du réseau hydrographique. Aux confluences, les composantes des différents affluents se combinent. Ce modèle, qui fait intervenir simultanément les deux topographies, le plan et le réseau, revient en fait à régulariser l'apport régional sur un support variable le long du réseau. Il est utilisé pour estimer par analyse krigéante un « indice de concentration locale » dans le plan.

Plus récemment, d'autres classes de fonctions aléatoires ont été proposées, construites des sources vers l'exutoire [Ver Hoef et al., 2006], ou en sens inverse, de l'exutoire vers les sources [Bailly et al., 2005; Monestiez et al., 2005].

Dans la suite, nous définissons à l'ordre deux des fonctions aléatoires sur un « arbre ». L'exutoire désigne la racine de l'arbre, les sources les sommets pendants autres que la racine, les confluences sont les sommets extrémité de plusieurs arêtes. Une « rivière » désigne un

chemin complet d'une source vers l'exutoire. Deux points dont l'un est à l'aval de l'autre sont dits reliés, sous-entendu « au fil de l'eau » ; ils sont non reliés dans le cas contraire. Tout point de l'arbre est repéré par son abscisse curviligne s comptée positivement depuis l'exutoire où $s = 0$, ainsi que par l'indice d'un cours d'eau J auquel il appartient ou l'indice i de son arête.

A) Combinaison de fonctions aléatoires monodimensionnelles

Considérons les affluents d'un réseau hydrographique comme des « filets d'eau » distincts, dont la réunion forme les « rivières » (figure 8). Sur chaque filet, de la source à l'exutoire, nous définissons un débit et une concentration, fonctions de l'abscisse curviligne comptée depuis l'exutoire. A l'aval des confluences successives, les débits des filets se cumulent pour former le débit de la rivière, et leurs concentrations se composent en proportion des débits relatifs. Ces combinaisons fournissent un procédé général de construction de FA sur un arbre.

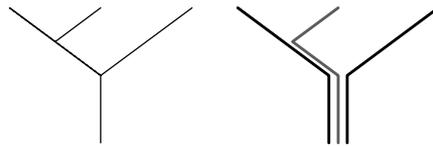


Figure 8. Décomposition de l'arbre en « filets »

Pour un arbre à N sources, soient N Fonctions Aléatoires ou « composantes » définies à une dimension sur chacun des filets. Sur l'arbre, la concentration Z est construite par combinaison des composantes à l'aval des confluences.

Si les composantes sont indépendantes, les concentrations sont sans corrélation entre points non reliés. Lorsque toutes les composantes ont la même covariance C_1 , la concentration résultante est stationnaire par arête et proportionnelle à C_1 entre points situés à l'aval l'un de l'autre, à un facteur près variable selon l'arête du point amont et celle du point aval. Aux confluences, la covariance est discontinue.

Ce modèle généralise celui proposé par Ver Hoef et al. [2006] dans le cas particulier où la covariance C_i est l'autoconvoluée d'un noyau f défini sur la demi droite. La construction par « filet » est plus simple et plus générale, car elle permet de combiner des FA monodimensionnelles de covariance quelconque. Elle s'étend à des composantes non stationnaires ou corrélées, ce qui offre de nombreuses possibilités.

Par exemple, intégrer des fonctions aléatoires positives le long des filets, en sommant à l'aval des confluences les fonctions aléatoires intrinsèques ainsi obtenues, permettrait de représenter des flux de substances croissants au fil de l'eau, en l'absence de bio-dégradation.

Dans un bassin versant, les concentrations des différentes substances sont fortement liées au milieu (nature des sols, type d'agriculture). Les modèles multivariés de type dérive externe ou résidu permettent d'introduire le contexte environnemental dans l'estimation, les modèles précédents s'appliquant alors aux résidus. Cependant, l'information auxiliaire n'étant pas toujours disponible ou n'expliquant pas nécessairement les liaisons observées, des modèles présentant des corrélations entre rivières à l'amont des confluences sont utiles.

B) Fonctions aléatoires stationnaires ou intrinsèques par cours d'eau

Pour modéliser la largeur d'un fluvisol ou celle des fossés d'un réseau de drainage, Bailly et Monestiez [2005] utilisent une décomposition du type dérive déterministe et résidu stationnaire et continu aux confluences. Ces modèles permettent aussi de décrire certaines caractéristiques végétales, ou les concentrations dans un delta. Le procédé est fondé sur une

hypothèse d'indépendance conditionnelle de Z entre cours d'eau à l'amont de leur confluence, connaissant exhaustivement Z entre la confluence et l'exutoire.

On construit une simulation monodimensionnelle sur un premier cours d'eau, de l'exutoire à la source. A partir du premier point de confluence, on poursuit la simulation sur le deuxième cours d'eau, indépendamment de celle précédemment construite sur le premier affluent, mais conditionnellement au tronçon commun depuis l'exutoire. On itère le processus à chaque nouvelle confluence rencontrée.

La covariance entre affluents se calcule simplement à l'aide de la loi conditionnelle dans le cas gaussien, qui se déduit du krigeage à moyenne connue. La construction se généralise au cas intrinsèque, avec le variogramme linéaire comme cas particulier.

Un calcul élémentaire montre que la covariance exponentielle et le variogramme linéaire sont les seuls modèles compatibles avec l'hypothèse d'indépendance conditionnelle et « stationnaires » sur un arbre, c'est-à-dire qui s'expriment en fonction de la distance curviligne seulement, indépendamment de la structure de l'arbre.

Lorsque le support est discrétisé, ces modèles se construisent par simulation séquentielle. Une construction sans discrétisation du support arborescent peut être obtenue à partir du processus de Poisson. D'après le théorème central limite, la sommation d'un grand nombre de Fonctions Aléatoires indépendantes fournit alors une Fonctions Aléatoire de loi spatiale gaussienne.

C) Vers les modèles multivariables

La version multivariable de ces modèles ne pose pas de difficulté. Ainsi, sur tout arbre, le modèle linéaire de corégionalisation s'obtient classiquement par combinaison de facteurs spatiaux mutuellement indépendants. Ce modèle peut décrire les concentrations en nutriments (nitrates, nitrites, phosphates ...) dans les cours d'eau. Il est aussi utilisable pour représenter la relation débit-flux pour différentes substances, dans un modèle à résidu analogue à celui utilisé pour la relation puissance-accumulation [J. Rivoirard, 1991].

Les débits étant fortement non stationnaires, la modélisation est plus commode, et surtout les estimations sont plus précises, lorsqu'une information auxiliaire est introduite en « dérive ». On utilisera par exemple une relation empirique entre le débit et la superficie du bassin drainé (par exemple [Sauquet et al., 2000]), ou les prévisions issues d'un modèle phénoménologique. Des modèles intégrés comme Pégase [Smitz et al., 1997], ou pour la Seine le couplage Sénèque-Prose [par exemple Flipo, 2005 ; Poulin et al., 2001] décrivent le fonctionnement hydrologique, hydrodynamique et écologique du bassin versant au tronçon de cours d'eau. Ces modèles fournissent donc également une « ébauche » pour les concentrations. Les conditions de cohérence aux confluences varient selon la variable considérée, l'information auxiliaire et le modèle géostatistique retenus (débit spécifique, c'est-à-dire rapporté à la surface drainée, ou modèle à résidu débit-concentration, par exemple).

Les premières applications pratiques sont en cours. Le couplage des modèles phénoménologiques à la modélisation géostatistique constitue un champ vaste de futurs développements certainement très instructifs.

D) Inférence des modèles

En une confluence, les concentrations dans les cours d'eau présentent une discontinuité : si les concentrations et les débits immédiatement à l'amont sont respectivement c_1, c_2, d_1, d_2 , juste après la confluence le débit est $d_1 + d_2$ et la concentration $\frac{c_1 d_1 + c_2 d_2}{d_1 + d_2}$. Un modèle « continu » convient par exemple pour les propriétés d'un végétal, ou les concentrations dans un delta, l'indépendance conditionnelle décrivant des concentrations évoluant indépendamment sur les différents bras depuis leur point de divergence. Un modèle continu est également admissible pour les résidus, lorsque la dérive prend en charge les discontinuités aux confluences.

En modèle continu, l'inférence du variogramme s'effectue le long des cours d'eau, en retenant dans le calcul les seuls couples de points reliés.

Pour le modèle discontinu aux confluences, les coefficients de la combinaison linéaire sont supposés connus ; pour des concentrations, ils sont par exemple déduits des surfaces drainées. Idéalement, l'inférence du modèle s'effectue en deux étapes : inférence du variogramme des composantes sur les « filets », et inférence de la corrélation entre les différentes composantes. Lorsque les coefficients sont constants par arête, le variogramme aux petites distances peut être inféré par arête. La comparaison des variogrammes sur les arêtes à l'aval immédiat des sources permet de vérifier si un même modèle est admissible pour les différentes composantes. Le calcul du variogramme pour les couples de points non reliés permet de décider de la corrélation entre composantes.

Dans sa thèse, C. Bernard-Michel [2006] observe que pour le bassin de la Loire et celui de la Moselle, chaque arête comporte généralement une seule station (exceptionnellement deux, ou aucune vers l'amont du réseau), ce qui complique l'étude variographique. On peut alors examiner la stationnarité d'accroissements du type $Z(s_i) - \mathbf{b}_{ij} Z(t_j)$, le coefficient \mathbf{b}_{ij} étant déduit des pondérations aux confluences. Si un même variogramme est admissible pour toutes les composantes, on peut ensuite tester la corrélation entre points non reliés.

L'étude de la qualité des cours d'eau nécessite d'inférer les modèles multivariés (débits, concentrations et flux, variables auxiliaires). Parmi la classe des modèles possibles, lesquels décrivent « au mieux » les données tout en assurant une certaine robustesse aux estimations ? Une première application est en cours, dans le cadre de la thèse de C. Bernard-Michel.

Actuellement, seule la composante spatiale a été examinée, à travers la moyenne annuelle. D'autres applications nécessiteront une modélisation spatio-temporelle.

II.3. Pollution atmosphérique

II.3.1. Eléments bibliographiques

Cartographie avec informations auxiliaires

On distingue usuellement trois échelles pour l'étude de la pollution atmosphérique: planétaire, à longue distance, et de proximité, cette dernière étant seule abordée ici. A cette échelle, les questions portent principalement sur l'estimation de moyennes temporelles en tirant parti d'informations auxiliaires, sur le dimensionnement de dispositifs d'échantillonnage et sur la « représentativité » des sites de mesures, et sur l'évaluation de l'exposition.

La réglementation impose de cartographier avec précision les concentrations à l'échelle de l'agglomération, ce que ne permettent pas les mesures aux quelques « analyseurs » fixes implantés dans les principales agglomérations. Ces mesures, temporellement très denses (des valeurs quart horaires durant plus d'une décennie sont parfois disponibles) sont complétées par des campagnes par « tubes à diffusion passive », effectuées sur une agglomération ou un département (parfois plusieurs). De la taille d'un tube à essai, ces « capteurs » sont placés de façon à ce que l'air y circule par convection naturelle, le polluant étant piégé sur un adsorbant spécifique. Quelques dizaines de sites soigneusement sélectionnés sont alors instrumentés, fournissant des valeurs hebdomadaires ou bimensuelles selon le polluant mesuré. Des variables auxiliaires décrivant plus ou moins directement le milieu et les sources de pollution sont généralement disponibles.

Depuis plus de quinze ans, les concentrations à différentes échelles sont cartographiées par krigeage [voir par exemple Lefohn et al., 1988 ; Casado et al., 1994; Liu et al., 1996]. La

littérature comporte de nombreux exemples de comparaison de méthodes utilisant la technique de validation croisée (par exemple Cressie et al. [1999]).

Le krigeage avec dérive externe permet d'incorporer des informations auxiliaires pour améliorer la précision de l'estimation. Cartographiant la concentration en sel marin dans des lichens, Figueira et al. [2001] comparent krigeage ordinaire, krigeage en dérive externe et « soft kriging », cette méthode décomposant l'estimation de la concentration en deux termes, le krigeage ordinaire des concentrations et un terme correctif, l'erreur de krigeage calculée sur la variable auxiliaire. La validation croisée indique un écart minime entre dérive externe et « soft kriging ». Un modèle spatio-temporel avec anisotropie géométrique est ensuite utilisé pour l'estimation à différentes dates.

Pour estimer les concentrations en benzène sur l'agglomération de Rouen, Bobbia et al. [2000] utilisent en dérivées externes la densité de population et le relief, ces variables améliorant grandement la qualité de l'estimation. Pour le dioxyde d'azote NO₂ à Paris et en proche banlieue, l'estimation de la médiane annuelle des concentrations journalières est effectuée par variable utile, avec les médianes aux stations comme données. Le cadastre des émissions n'étant pas disponible sur l'ensemble de la région, un cokrigeage est cette fois-ci effectué, faisant ressortir l'augmentation des concentrations à proximité des principaux axes routiers.

Parmi les autres questions traitées, le dépassement de seuil est souvent abordé par krigeage d'indicateurs ou par diverses techniques de simulation. A l'échelle plus vaste de la Norvège, Host et al. [1999] procèdent par simulation, pour évaluer l'AOT40 en ozone, cumul sur un an de l'écart positif entre les concentrations horaires et le seuil de $80 \text{ mg} / \text{m}^3$ (soit 40 ppb). L'AOT40 représente le « bénéfice conventionnel » à ce seuil. Les auteurs utilisent un modèle spatio-temporel de type dérive aléatoire et résidu ; la dérive est le produit d'un facteur spatial aléatoire et d'un facteur temporel déterministe, la covariance du résidu étant factorisée en une composante spatiale et une composante temporelle. Malheureusement, les simulations non conditionnelles de ces résidus sont construites comme le produit de deux fonctions aléatoires, de covariance respectivement égale à la composante spatiale ou à la composante temporelle du modèle factorisé. La validation est effectuée en trois stations non utilisées pour la modélisation. Si la précision effective est souvent meilleure que celle annoncée par le modèle, les résultats montrent généralement une sous-estimation de l'AOT ainsi calculée.

Les méthodes sont présentées sur des sites particuliers, la validation sur différents sites étant rarement abordée. La « sanction par la pratique » conduit cependant à retenir les méthodes jugées satisfaisantes sur différents cas. La dérive externe est ainsi largement utilisée. Certaines méthodes d'estimation non linéaire ne sont que rarement évaluées.

Concentration annuelle en dioxyde d'azote

Le dioxyde d'azote NO₂ est un polluant suivi avec attention à cause de ses effets et du niveau des concentrations. Il résulte principalement de la combustion incomplète des hydrocarbures fossiles (circulation routière, chauffage), et secondairement de certains procédés industriels. Le NO₂ aggrave certaines pathologies chez les sujets fragiles ou souffrant d'insuffisance respiratoire. Il contribue aux pluies acides, à l'altération de la couche d'ozone, et indirectement au processus d'effet de serre [voir par exemple ASPA, 2002]. Du fait de réactions photochimiques complexes, le NO₂ prédomine généralement en hiver et l'ozone en été. Selon les conditions météorologiques et chimiques locales, le NO₂ peut « consommer » l'ozone ou en générer. La réglementation européenne fixe à $40 \text{ mg} / \text{m}^3$ la moyenne annuelle en NO₂ à ne pas dépasser à partir de 2010. L'estimation de cette moyenne annuelle est l'objet de nombreuses études.

Les questions initialement posées portaient sur le « meilleur » choix des variables en dérive externe. L'analyse exploratoire détaillée pour l'agglomération de Mulhouse a permis de préciser les relations entre les concentrations saisonnières ainsi qu'avec la moyenne annuelle,

relations confirmées par la suite pour d'autres agglomérations [Gallois et al., 2005 ; de Fouquet et al., 2005]. Une adaptation des campagnes de mesure a été proposée, validée par l'INERIS pour d'autres agglomérations [Cardenas et al., 2004] et actuellement recommandée aux Associations Agréées chargées de la surveillance de la qualité de l'air (AASQUA). L'évaluation du terme temporel de la variance d'estimation a montré que ce terme était négligé à tort [D. Gallois, 2004]. Enfin le résultat de l'analyse krigeante des concentrations saisonnières à Mulhouse a été confronté à l'interprétation phénoménologique.

Comme illustration, deux études variographiques sont détaillées ci-dessous : la campagne de mesure par tubes à diffusion passive de la vallée de la Thur dans les Vosges et celle de l'agglomération de Mulhouse.

II.3.2. Vallée de la Thur : influence de la distance à la route

Préalablement au réaménagement de la RN66 dans la vallée de la Thur, une campagne de mesure du NO₂ a été effectuée en deux phases de 90 jours chacune, de juin à septembre 2001, et de novembre 2001 à février 2002. Les échantillonneurs passifs ont été disposés le long de la route et suivant des transects perpendiculaires, principalement dans les zones urbanisées. Quarante cinq sites comportent une série complète de mesures, parmi lesquels quatre ont été écartés comme « atypiques », car situés à proximité d'axes locaux à forte circulation, et par conséquent sous l'influence d'une importante source de pollution autre que la RN66.

L'étude de l'ASPA [2002] présente de nombreux résultats :

- les concentrations hivernales généralement plus élevées ;
- la diminution des concentrations vers le fond de la vallée, avec un gradient mieux marqué en été ;
- les fortes concentrations à proximité immédiate de la chaussée (1 ou 2m), et la forte décroissance des concentrations entre 2 et 50m, aucun site n'étant implanté à une distance intermédiaire. En plusieurs sites à proximité de carrefours à feux tricolores, la moyenne du NO₂ durant les 180 jours mesurés dépassait en 2001 l'objectif des 40 **mg / m³**.
- une légère dissymétrie des concentrations de part et d'autre de la route, explicable par la rose des vents, avec des valeurs plus élevées « sous le vent ».

Les mesures saisonnières sont fortement corrélées, avec un nuage approximativement linéaire par morceaux (figure 9). L'influence du trafic apparaît effectivement mieux marquée en été, avec une décroissance plus nette des concentrations en fonction de la distance à la route. Cette distance influence donc les relations entre concentrations saisonnières :

- à proximité immédiate de la chaussée, les mesures estivales diffèrent peu des mesures hivernales ou leur sont supérieures ;
- au-delà de 50m, les mesures hivernales sont systématiquement supérieures aux estivales.

En moyenne, le rapport de l'écart saisonnier à la valeur estivale est proche de 100%, l'écart saisonnier étant plus réduit lorsque la valeur estivale augmente.

Près d'une voie à fort trafic, la distance à la route est le facteur principal caractérisant le niveau des concentrations. A proximité immédiate de la route, les concentrations conservent un niveau analogue en été et en hiver, ces fortes valeurs estivales pouvant indiquer une « consommation » de l'ozone transformé entre autres en NO₂. Enfin, dans ce cas l'influence de la route est visible au moins jusqu'à 500m, valeur supérieure aux distances communément admises.

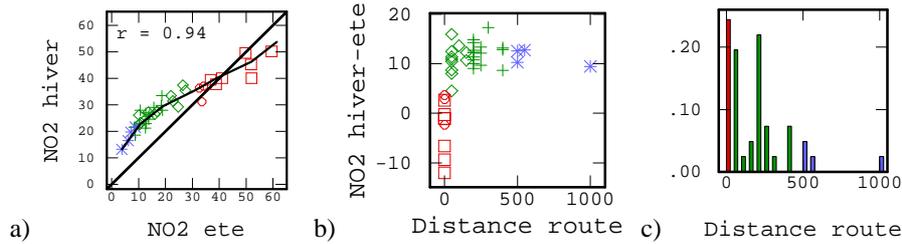


Figure 9. Vallée de la Thur. Relation entre les concentrations saisonnières en NO₂ (en µg/m³) et la distance à la route (en m). a) Nuage de corrélation des mesures saisonnières en fonction de la distance à la route (? 1m, ? 2m, ? de 50 à 150m, + de 200 à 400m, * 500m ou plus). La première bissectrice et la régression empirique sont reportées. b) nuage de corrélation de l'écart intersaisonnier en fonction de la distance à la route. c) histogramme de la distance à la route. Les couleurs se correspondent pour les trois graphiques.

II.3.3. Relation entre concentrations saisonnières et milieu à Mulhouse

Les mesures ont été effectuées en vue de cartographier la moyenne annuelle de la concentration en NO₂ sur l'agglomération. La liaison entre la moyenne annuelle en NO₂ et les variables décrivant le milieu (occupation des sols, cadastre des émissions, densité de population), disponibles sur l'ensemble de la zone d'étude, avait été examinée à l'ASPA, concluant à l'intérêt du krigeage avec dérive externe.

Relations entre valeurs saisonnières

Sur l'agglomération, 75 sites urbains ou périurbains ont été équipés durant trois quinzaines hivernales puis trois quinzaines estivales durant l'année 2001, soit un total de trois mois de mesures. Pour diverses raisons, une soixantaine de mesures sont disponibles pour chaque saison, fournissant seulement 50 mesures annuelles (figure 10). La carte confirme les concentrations annuelles généralement plus faibles en périphérie, et plus fortes au centre et vers le nord-est de l'agglomération, suivant un important axe économique en direction de Strasbourg.

A cause de la corrélation spatiale, les statistiques des concentrations sont indicatives, du fait de l'implantation préférentielle des mesures vers le centre de l'agglomération ; pour des estimations non linéaires, une pondération de l'histogramme est nécessaire. Les valeurs hivernales sont en moyenne supérieures de 75% aux valeurs estivales. Les écart-types indiquent une variabilité légèrement croissante avec la concentration ; mesurée par le coefficient de dispersion, la variabilité relative est supérieure en été avec 0.37, contre 0.25 en hiver.

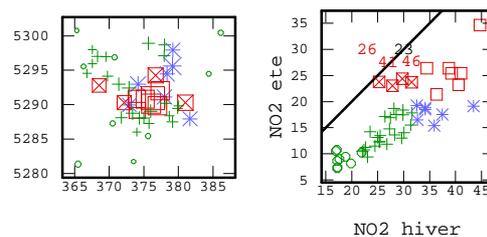


Figure 10. NO₂ à Mulhouse, campagne 2001. Implantation des sites « annuels » (à gauche ; la dimension des symboles est proportionnelle à la moyenne annuelle) et nuage de corrélation des mesures saisonnières. Les symboles se correspondent pour les deux graphiques. Les sites repérés par un X correspondent à de fortes valeurs estivales associées à des valeurs hivernales intermédiaires. Les concentrations sont exprimées en µg/m³, et les distances en km.

Le nuage de corrélation des mesures saisonnières indique des mesures hivernales toujours supérieures aux mesures estivales (figure 10). Les faibles concentrations coïncident en hiver et en été, l'implantation des fortes concentrations variant sensiblement entre les deux saisons, ce que confirment les cartes krigées. Ces différences reflètent l'origine de la pollution, avec l'influence supplémentaire du chauffage en hiver.

Le mode des fortes mesures estivales recouvre les plus fortes mesures hivernales et quelques valeurs hivernales intermédiaires, mesurées en des sites situés en proche périphérie de la ville (ces sites sont repérés sur la figure 10). Ces valeurs estivales élevées par rapport aux hivernales indiquent-elles comme pour la vallée de la Thur une forte influence routière ? Après vérification, il s'avère que trois de ces sites sont dans un environnement assez proche d'axes de circulation à fort trafic, le dernier étant localisé en zone urbaine dense (G. Perron, 2005).

Relations entre concentrations et variables auxiliaires [rapports d'étude, de Fouquet 2003]

L'étude des corrélations (nuages, analyse factorielle) confirme les liaisons attendues avec les variables explicatives, ou leur logarithme translaté (Figures 11 et 13) : la concentration estivale est corrélée aux émissions en oxydes d'azote (notés NO_x) incluant ici la circulation routière et les activités industrielles, tandis que la concentration hivernale est mieux corrélée au bâti dense, ce qui peut refléter l'influence du chauffage. A cause de leur liaison élevée avec les concentrations, l'introduction de ces variables en dérive externe améliore fortement l'estimation, en particulier pour les valeurs annuelles ou hivernales pour lesquelles les corrélations sont les mieux marquées. Différentes combinaisons des variables en dérive donnent des résultats analogues. En cohérence avec les variances expérimentales, la variance d'estimation est plus élevée pour les valeurs hivernales et la variance relative pour les valeurs estivales, avec des résultats intermédiaires pour les moyennes annuelles.

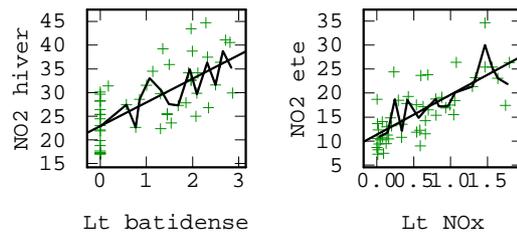


Figure 11. Mulhouse, campagnes 2001 pour le NO₂. Nuage de corrélation entre les mesures hivernales et le logarithme translaté du bâti dense, et entre les mesures estivales et le logarithme translaté du cadastre des émissions en oxydes d'azote. Les concentrations sont en µg/m³, la variable auxiliaire est rendue adimensionnelle.

Adaptation des campagnes saisonnières

Avec des coefficients supérieurs à 0.90, la corrélation est plus élevée entre les mesures « annuelles » et les mesures saisonnières qu'avec les variables auxiliaires. Cette forte liaison est utilisable pour améliorer l'estimation des valeurs annuelles ou saisonnières, la série de mesures étant incomplète pour environ 20% des sites. Le co-krigeage des mesures saisonnières assure de plus la cohérence des estimations sous la condition de « fermeture », l'estimation de la moyenne annuelle étant égale à la moyenne des estimations saisonnières.

Dans son stage d'option, D. Gallois [2004] a alors montré qu'une réduction de 30% du nombre de mesures ne détériorait que très faiblement l'estimation, à condition de mesurer chaque site durant au moins une saison. A coût d'échantillonnage constant, conserver suffisamment de sites communs aux deux saisons pour le calage du modèle bivarié, et alterner les autres sites, permet d'étendre largement la zone pouvant être cartographiée avec une précision suffisante. C'est effectivement la méthode qui avait été adoptée à Montpellier pour les campagnes 2001-2002, mais avec un nombre trop restreint de sites communs.

Composante temporelle de la variance d'estimation

Les données des campagnes par tubes ne permettant pas d'accéder à la variabilité temporelle des concentrations, la composante temporelle de la variance d'estimation est généralement négligée. Des analyseurs fixes, peu nombreux mais échantillonnant les concentrations en continu, sont utilisables pour approcher cette composante. Si l'on néglige la corrélation entre les erreurs d'estimation spatiales et temporelles, la composante temporelle de la variance d'estimation s'obtient par un simple calcul de variance d'extension, en dimension un.

Tenant compte d'un effet proportionnel ainsi que de la liaison entre la moyenne annuelle et les émissions en NO_x , D. Gallois [2004] a évalué pour Mulhouse l'écart type d'estimation temporel à environ 50% de l'écart type spatial. L'influence de la composante temporelle de l'erreur d'estimation n'est donc pas négligeable.

Les mesures quasi exhaustives aux analyseurs ont également permis de valider la durée et la période des campagnes. Expérimentalement, l'extension à quatre quinzaines n'améliore que très faiblement l'estimation de la moyenne annuelle pour Mulhouse ; le choix des périodes de mesure s'avère important. D. Gallois [2004] a mis en évidence une période « optimale », suivant un critère de précision mais aussi de « robustesse » relativement à un décalage temporel par rapport à cette période optimale.

Modélisation et interprétation

Pour préciser les relations spatiales entre mesures saisonnières, il est commode d'introduire l'écart D entre mesures hivernales et estivales. A Mulhouse, la valeur estivale apparaît non corrélée à l'écart au même point. L'ajustement en modèle linéaire de corégionalisation conduit à la décomposition $[NO_2]_{ETE} = \mathbf{a}_E P + \mathbf{b}_E S + \mathbf{b}'_E S' + \mathbf{d}_E L$, $D = \mathbf{a}_D P + \mathbf{b}'_D S' + \mathbf{d}_D L$, avec évidemment $[NO_2]_{HIVER} = [NO_2]_{ETE} + D$, P, S, S', L désignant des facteurs spatialement non corrélés, de variogramme respectivement pépitique, sphérique de portée 5km, et linéaire.

L'importance relative de chaque structure est donnée par les coefficients reportés au tableau 1. La composante pépitique et la composante linéaire correspondent chacune à une seule structure, d'amplitude variable selon la saison. La composante pépitique reflète vraisemblablement pour partie une structure spatiale de faible portée, qui tend à disparaître en hiver. La composante linéaire, et la première composante sphérique, peu présentes en été, se développent en hiver.

La deuxième composante sphérique conserve la même amplitude en été et en hiver. Cette composante peut présenter une corrélation différée de plus de 5km avec la première composante sphérique, qui apparaît sur la partie impaire de la covariance croisée des mesures saisonnières (figure 12). A. Cori [2005] a montré que la partie impaire de la covariance croisée est identiquement égale entre les deux saisons ou entre l'écart et chacune des saisons.

L'absence de corrélation entre mesure estivale et écart au même point semble finalement fortuite. Elle n'est d'ailleurs pas confirmée à Rouen [A. Cori, 2005].

composante		été	écart hiver-été	hiver	annuelle
pépitique		3.19	-2.47	0.72	1.95
sphériques	S	0.69	2.99	3.68	2.19
portée 5km	S'	4.81		4.81	4.81
linéaire		0.85	0.74	1.59	1.22

Tableau 1. Coefficients de la factorisation des concentrations saisonnières en NO_2 à Mulhouse.

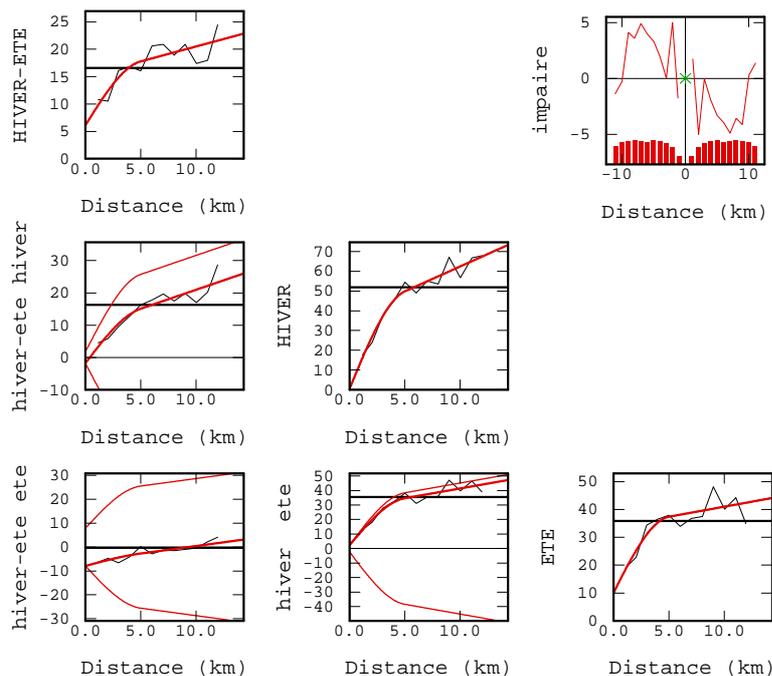


Figure 12. Mulhouse, campagne 2001 des mesures en NO₂. Variogrammes des concentrations saisonnières et de leur écart, et ajustement en modèle linéaire de corégionalisation. En haut à droite, partie impaire de la covariance croisée entre concentrations saisonnières ; l'histogramme indique le nombre de couples par pas, identique pour tous les variogrammes.

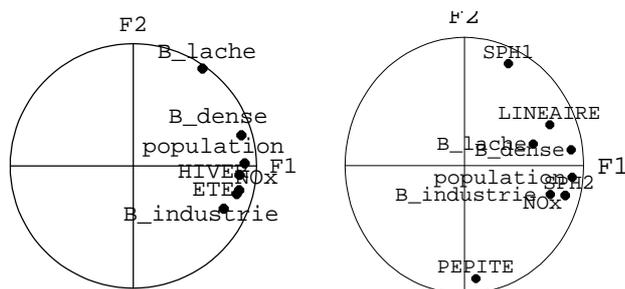


Figure 13. Mulhouse, campagne 2001 des mesures de NO₂. ACP des variables explicatives, des concentrations saisonnières et des facteurs spatiaux estimés par analyse krigéante. A gauche : variables explicatives et mesures saisonnières ; au milieu : concentrations saisonnières, facteurs spatiaux, et deux premiers facteurs d'ACP des cinq variables explicatives ; à droite : facteurs spatiaux et variables explicatives. SPH1 représente le facteur sphérique *S* commun à l'été et à l'écart, et SPH2 le facteur *S'* commun à l'été et à l'hiver seulement. Les variables sont rendues adimensionnelles

L'analyse krigéante aux points expérimentaux (effectuée à partir des seules mesures en NO₂) permet d'affiner l'interprétation.

La forte liaison entre variables explicatives et mesures saisonnières se retrouve avec certains facteurs seulement (figure 13). Les informations auxiliaires disponibles « expliquent » mal la composante pépitique. Le facteur *S*, qui augmente fortement entre l'été et l'hiver, semble « compenser » la disparition de la composante pépitique. Le facteur *S'*, de même

amplitude durant toute l'année, est fortement lié à l'anthropisation, en particulier aux différentes émissions (incluant le trafic automobile), ainsi qu'à la densité de bâti industriel et à la population. Enfin, prépondérant en hiver, le facteur linéaire apparaît principalement lié aux habitations (bâti dense ou lâche) ; il prend sans doute en compte le chauffage.

Selon G. Perron (ASPA, 2006) ces structures spatiales sont liées aux processus physico-chimiques de production, transformation et destruction du NO₂ :

- en été, l'ensoleillement et la température entraînent une rapide décroissance du NO₂ à courte distance des sources d'émission, ce qui explique les niveaux estivaux plus faibles avec des contrastes mieux marqués ;

- en hiver, l'accumulation des polluants « primaires » prédomine, les émissions provenant du chauffage domestique étant réparties sur toute l'agglomération. Le « dôme de pollution » couvrant l'agglomération favorise une certaine homogénéisation dans l'espace, expliquant la réduction de la composante pépitique et la forte contribution de la composante linéaire associée au développement de la composante sphérique ;

- l'absence de relation entre la composante pépitique et les variables « explicatives » proviendrait de l'implantation des tubes à l'écart des sources d'émission identifiées, pour la mesure de la pollution « de fond ».

Validation sur d'autres agglomérations

L'étude exploratoire des concentrations en NO₂ sur les agglomérations de Montpellier [Gallois et al., 2005] et de Rouen [A. Cori, 2005], ainsi qu'à Bourg-en-Bresse [Cardenas, 2004] montrent de nombreuses analogies avec les résultats obtenus à Mulhouse :

- forte corrélation entre concentrations saisonnières, ou pour la même saison à une année d'intervalle ;

- moyenne et variance supérieures en hiver, mais coefficient de variation supérieur en été ;

- concentrations analogues en été ou en hiver à proximité de fortes émissions routières ;

- liaison différente selon les saisons entre les concentrations et les émissions routières, ainsi qu'avec la densité de population.

- utilité du cokrigeage des concentrations saisonnières pour palier les séries de mesures incomplètes, ou pour valoriser une implantation des mesures variable suivant les saisons. Un allègement des campagnes de mesure ou, à coût constant, une extension des zones échantillonnées peut être proposé.

L'analyse exploratoire des données de Mulhouse a ainsi une portée beaucoup plus générale, et peut guider l'étude d'autres sites.

Pour mieux faire apparaître les relations entre concentrations saisonnières, il est préférable de conduire l'analyse exploratoire sur les séries complètes de mesures, l'ensemble des mesures étant ensuite utilisé pour l'estimation.

Le krigeage en dérive externe a été utilisé pour cartographier d'autres polluants, comme le benzène sur les départements du Rhône et de l'Isère. Dans son travail d'option, G. L'Hégaret [2002] montre la nécessité de mesures couvrant non seulement l'espace géographique mais échantillonnant également l'espace des variables auxiliaires. Extrapoler jusqu'à des altitudes voisines de 4000m un modèle multivariable calé sur des mesures effectuées à moins de 1000m, peut en effet fournir des résultats aberrants, comme des estimations fortement négatives.

II.3.4. Représentativité des stations pour le NO₂ à Rouen

La notion de représentativité est fréquemment mentionnée dans la réglementation sans être jamais définie. Dans son travail d'option [2005], A. Cori a examiné plusieurs définitions possibles, en les appliquant à la ville de Rouen.

La notion empirique est simple : la zone de représentativité d'une station de mesure est l'ensemble des points où les concentrations sont « proches » de celle de la station. La représentativité est donc définie pour un polluant et pour un critère fixés. Pour le NO₂, la moyenne annuelle est retenue du fait de la réglementation.

Comme les valeurs varient d'une année à l'autre, il paraît préférable d'introduire l'information sur le milieu à l'aide d'une « ébauche » f . L'écart est alors décomposé comme une somme $Z(x_0) - Z(x) = f(x_0) - f(x) + R(x_0) - R(x)$. Minorer l'écart quadratique de la concentration conduit à définir des distances maximales à partir du variogramme des concentrations ou de celui des résidus. Ces voisinages définis « en moyenne » sont peu satisfaisants : soit ils englobent des sites des mesures assez différents, soit ils deviennent trop réduits lorsque la contrainte est trop stricte.

Introduire la carte des concentrations estimées $Z^* = f + R^*$ assure une certaine cohérence entre les différents documents publiés. La décomposition précédente s'écrit alors $Z(x_0) - Z(x) = (f(x_0) - f(x)) + (R(x_0) - R^*(x)) + e(x)$, l'erreur de krigeage du résidu $e(x)$ d'espérance nulle et de variance connue, étant supposée gaussienne. Minorer chacun des termes (en valeur absolue) permet de « moduler » l'influence du milieu et d'introduire un « risque statistique » sur l'amplitude de l'incertitude liée à l'erreur d'estimation.

II.3.5. Conséquences de la réglementation

La réglementation sur les dépassements de seuils est parfois très complexe. Ainsi pour l'ozone, le maximum journalier des régularisées sur huit heures ne doit pas dépasser le seuil de $120 \text{ mg} / \text{m}^3$ durant plus de 25 jours par an ; mais le support spatial n'est pas précisé. On peut procéder par simulations pour établir la carte annuelle des dépassements de seuils, à partir des écarts entre les mesures aux analyseurs et les modèles phénoménologiques remis à jour. On peut également chercher à définir des « zones de représentativité » des stations à l'échelle régionale. Mais ce problème est complexe, le niveau d'ozone et sa structure spatiale variant suivant les conditions météorologiques.

La réglementation pourrait à terme concerner de nombreuses substances, ce qui nécessite le développement et la validation de méthodes d'estimation « praticables ». L'évaluation de l'exposition des individus constitue une autre question d'actualité.

SIMULATIONS

Ce chapitre présente deux méthodes de simulation spécifiques. Le modèle gaussien seuillé fournit une partition de l'espace, décrivant par exemple la géométrie des formations de gisements hétérogènes. La simulation conjointe de Fonctions Aléatoires liées par des équations aux dérivées partielles linéaires permet d'éviter les approximations d'un calcul par différences finies.

III.1. Le modèle gaussien seuillé et ses applications

Initialement développé pour représenter les différentes formations lithologiques de gisements pétroliers ou miniers hétérogènes, le modèle gaussien seuillé pourrait aussi décrire la répartition des unités pédologiques sur une parcelle, ou de différents remblais dans les sols de friches industrielles.

III.1.1. Le modèle gaussien seuillé

Un modèle approché

Au début des années 1980, une palette de modèles permettait de simuler des milieux homogènes ou hétérogènes : le modèle gaussien anamorphosé [Matheron, 1973], les modèles mosaïques [Matheron, 1984], le schéma booléen [Serra, 1982], des modèles génétiques de processus sédimentaires [Joathon, 1971]. Le conditionnement aux données « ponctuelles » était exact pour le modèle gaussien anamorphosé, ou approché par «essais et erreurs».

Le gisement d'uranium de Ben Lomond (Australie) étudié dans ma thèse, de type multifissuré, est constitué de lentilles minéralisées aplaties suivant un « plan de fracturation moyen ». Le long des sondages, les teneurs présentaient de forts contrastes aux intersections avec la minéralisation, avec une légère bimodalité de l'histogramme. Les limites des corps minéralisés étaient définies à partir des teneurs, les valeurs les plus fortes étant situées à l'intérieur des passées minéralisées plutôt qu'en bordure.

Pour un gisement de ce type, une simulation en modèle homogène avec changement de support est inadéquate, car elle ne reproduit pas un contraste suffisant des teneurs. L'idée vient naturellement de procéder en deux étapes : on construit d'abord une partition de l'espace en « minerais » et « stérile », chaque classe étant ensuite informée en teneurs.

Une simulation des lentilles minéralisées avait d'abord été effectuée suivant un schéma booléen, les teneurs étant construites indépendamment de la géométrie. Ce modèle « sans effet de bordure » a été jugé non conforme aux observations. Un modèle gaussien seuillé a alors été préféré : la minéralisation est construite par seuillage d'une fonction aléatoire de loi spatiale gaussienne, le seuil gaussien étant défini par la proportion minéralisée à la coupure donnée. La régularisée de l'indicatrice, qui représente la proportion minéralisée dans le bloc, était conditionnée par les transformées gaussiennes des teneurs, en modèle gaussien discrétisé. Les teneurs des blocs étaient ensuite simulées séparément pour le « minerai » et le « stérile », puis combinées selon la proportion minéralisée. Sur la simulation, l'histogramme des teneurs de blocs était effectivement bimodal.

Ce modèle approché fournit des simulations cohérentes de la géométrie et de la teneur. L'approximation réside dans l'ajustement séparé des variogrammes des trois variables gaussiennes : minéralisation, teneurs fortes et teneurs faibles. Comment généraliser ce modèle

lorsqu'aux points expérimentaux seules les classes de la partition sont connues, les gaussiennes associées étant inaccessibles ?

Le modèle gaussien seuillé [G. Matheron et al., 1987 ; C. de Fouquet et al., 1989]

Soit Y une fonction aléatoire stationnaire de loi spatiale gaussienne réduite, sa covariance $r(h)$ étant supposée continue. Considérant d'abord un seuil w constant, on caractérise par son indicatrice l'ensemble aléatoire A des points où Y dépasse le seuil w . La proportion p des points appartenant à A est l'espérance de son indicatrice : $p = E[I(x)] = 1 - G(w)$.

Ce modèle est entièrement spécifié par la covariance de Y et le seuil w , déduit de la proportion p . Les variogrammes de l'indicatrice de A ou de son complémentaire, identiques, correspondent à la probabilité que deux points distants de h soient l'un dans A , l'autre dans le complémentaire de A [J. Rivoirard, 1991] ; la covariance non centrée donne la probabilité que deux points appartiennent à A . Variogramme et covariance peuvent se calculer par intégration numérique, ou via le développement hermitien de l'indicatrice. Recaler à un même palier les variogrammes des indicatrices à différents seuils met en évidence la « déstructuration des hautes teneurs » : à même portée initiale de Y , la portée pratique de l'indicatrice diminue lorsque le seuil se rapproche des quantiles forts ou faibles (Figure 14.a).

Une famille de seuils $w_0 = -\infty \leq w_1 \leq \dots \leq w_{n-1} \leq w_N$ permet de partitionner l'espace en N classes ou « faciès » ordonnés. La loi (spatiale) gaussienne étant de type diffusif, tout segment joignant deux points appartenant à des faciès différents traverse tous les faciès intermédiaires.

Ce modèle présente une grande souplesse : outre le choix de la covariance, on peut regrouper des faciès (Figure 14.b.) ou faire varier les seuils. La figure 15 présente ainsi des coupes lithologiques à deux faciès, obtenues par seuillage d'une même réalisation gaussienne, pour différentes courbes de proportion verticale.

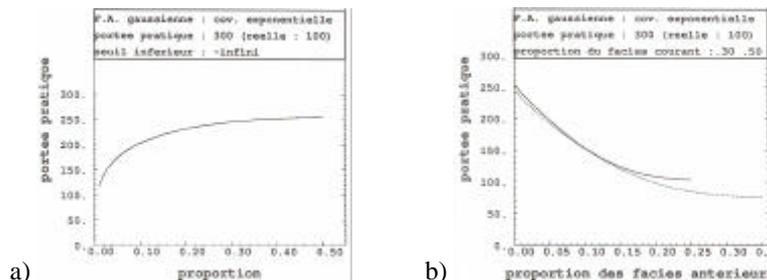


Figure 14. Portée pratique de l'indicatrice, à portée constante de la fonction aléatoire gaussienne. (a) en fonction de la proportion du faciès (b) en fonction de la proportion de la première classe, avec regroupement des deux classes extrêmes.

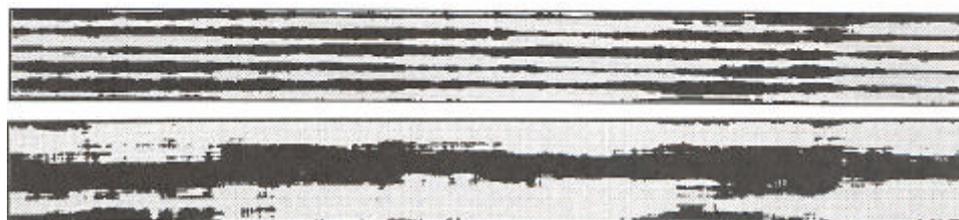


Figure 15. Coupes obtenues par seuillage d'une même réalisation gaussienne, pour deux courbes sinusoïdales de proportion verticale, de périodes différentes. La période représentée en haut $1/5^{\text{ème}}$ de la hauteur de la coupe, et en bas, elle lui est égale.

Conditionnement par un intervalle

Dans le cas d'une indicatrice, l'anamorphose n'étant pas bijective, les « gaussiennes conditionnantes » ne sont pas directement accessibles, le « faciès » en un point expérimental fournissant seulement un intervalle pour la gaussienne associée. Le conditionnement s'effectue alors en deux étapes : reconstruction des gaussiennes Y_a aux points expérimentaux, sous contraintes d'intervalles, puis conditionnement classique par les valeurs ainsi obtenues.

Dans une première phase, la reconstruction des gaussiennes utilisait comme données initiales les seuils gaussiens aux transitions entre faciès. Le vecteur gaussien Y est alors reconstruit par récurrence (« séquentiellement »), conditionnellement à ces « données exactes » et aux valeurs précédemment reconstruites: au $i^{\text{ème}}$ nœud, on pose $Y_i = Y_i^* + S_i R_i$, le résidu R_i étant choisi tel que Y_i appartienne à l'intervalle voulu.

Les tests ont montré un calage des « gaussiennes » ainsi reconstruites autour des seuils. Des méthodes itératives permettent de corriger ces anomalies. L'idée consiste à partir d'une initialisation imparfaite, modifiée peu à peu pour obtenir un vecteur approchant le modèle gaussien sous contraintes. Deux initialisations ont été envisagés :

- une fonction aléatoire pépitique, vérifiant les contraintes mais n'admettant pas la covariance voulue ;

- une fonction aléatoire admettant la covariance $r(h)$ voulue, mais ne vérifiant pas toutes les contraintes. A l'étape d'initialisation, seule est imposée la contrainte au point courant. Les contraintes aux points suivants n'étant pas prises en compte, la « gaussienne » présentait des discontinuités aux limites entre faciès.

Les écarts au modèle sont atténués en modifiant itérativement les simulations. A chaque étape, la valeur en tout point est successivement remplacée par une simulation conditionnée par tous les autres points. Le procédé est itéré jusqu'à ce que la covariance se rapproche suffisamment du modèle.

Ces méthodes de conditionnement ont été testées et pour certaines implantées dans la version initiale du logiciel HERESIM, dans le cadre d'un projet commun avec l'Institut Français du Pétrole.

Dans sa thèse que j'ai encadrée, X. Freulon [1992] a formalisé ces procédés empiriques en établissant des résultats de convergence, dans le cadre des chaînes de Markov à temps discret et états continus. Ces algorithmes empiriques se ramènent en effet à la méthode de Geman ou échantillonneur de Gibbs [Geman & Geman, 1984], consistant à remplacer itérativement les valeurs par leur simulation selon la loi conditionnelle. Après comparaison à un algorithme de Metropolis, consistant à valider ou à rejeter une modification arbitraire suivant une probabilité adaptée, X. Freulon préconise la méthode de Geman, avec comme simulation initiale un vecteur pépitique vérifiant les contraintes. Il étend ces méthodes au conditionnement de randomisées.

Dans son travail de fin d'études, L. Pellicciotta [2001] a comparé différentes variantes pour le conditionnement du modèle gaussien seuillé ; il recommande plutôt l'initialisation par une simulation non conditionnelle, admettant la covariance voulue mais ne vérifiant pas les contraintes, imposées ensuite itérativement par une méthode de Geman. Les meilleurs résultats sont obtenus en conservant les seules composantes vérifiant les contraintes dans la simulation non conditionnelle initialisant le processus.

Traitant de problèmes d'imagerie, J. Sénégas [2002] recense et développe de nombreux algorithmes applicables à un conditionnement quelconque. Dans la méthode de Geman, la modification pour des ensembles de points simultanément (relaxation par blocs) accélère fortement les calculs.

Factorisation de la covariance

La factorisation de la covariance $C(h_x, h_z) = C_1(h_x)C_2(h_z)$ suivant la verticale et le plan horizontal se justifie pour représenter un phénomène sédimentaire avec une forte anisotropie. Les effets d'écrans résultant de cette factorisation accélèrent fortement les temps de calcul pour une simulation par récurrence (« séquentielle »). Nous en rappelons le principe.

Soit à construire une section verticale, discrétisée suivant une maille régulière. Supposons les n premières colonnes Y_j simulées. Dans la décomposition $Y_j = Y_{ij}^* + \mathbf{s}_j R_j$, les poids du krigeage à moyenne nulle sont

- nuls pour les données situées sur les autres lignes
- identiques pour toutes les lignes, de même que la variance de krigeage.

Vectoriellement, la relation précédente s'écrit

$$Y_j = \sum_a I_j^a Y_{j-a}^* + \mathbf{s}_j R_j \quad (1).$$

Un calcul simple montre que la covariance du résidu R_j est la composante verticale de la covariance factorisée. Le procédé de simulation s'en déduit :

- simuler des vecteurs indépendants R_j , de covariance « verticale » C_2
- les combiner suivant la relation (1), le krigeage étant effectué avec la covariance « horizontale » C_1 .

Lorsque les données sont situées sur des sondages verticaux, le krigeage inclut les vecteurs conditionnants précédemment reconstruits.

Dans la première phase du projet HERESIM, la covariance tridimensionnelle était factorisée en trois exponentielles. Les différents krigeages ne faisaient intervenir qu'un ou deux points, correspondant à la « donnée » la plus proche située de part et d'autre du point courant.

La rapidité des calculs est importante. Elle facilite le tirage d'un grand nombre de simulations, pour visualiser la grande diversité des configurations compatibles avec les données ainsi que les études de sensibilité aux variations des paramètres du modèle.

La factorisation de la covariance dans le plan horizontal apparaît cependant peu réaliste, une anisotropie géométrique donnant des images plus « naturelles ». La simulation par transformation de Fourier discrète dans le plan, associée à une factorisation de la covariance entre le plan et la direction verticale permet de conserver des temps de calcul performants, la relation (1) consistant alors à combiner des plans horizontaux. Lorsque les puits conditionnants sont verticaux et complets (ou complétés), le conditionnement s'effectue alors à deux dimensions.

La simulation par « gaussiennes seuillées » s'étend aux autres modèles d'anisotropie, au détriment des temps de calcul pour la simulation et pour le conditionnement.

Variation verticale des proportions

Dans le cas de trois faciès ou plus, les transitions entre classes sont symétriques. Or les séquences sédimentaires sont orientées, des matériaux grossiers (grès) à la base vers les plus fins (argiles). La transition inverse est remplacée par des discontinuités, traduisant les ruptures de sédimentation. La continuité latérale des faciès permet de supposer leurs proportions stationnaires localement suivant l'horizontale. Les transitions de faciès suivant les séquences sédimentaires et les discontinuités entre séquences sont reproduites par une variation verticale des proportions, c'est-à-dire des seuils associés.

La non stationnarité des faciès a été par la suite étendue dans le plan, ce qui augmente la souplesse du modèle mais rend l'inférence plus problématique. Enfin, le modèle gaussien seuillé a été généralisé au seuillage de plusieurs Fonctions Aléatoires simultanément, permettant d'obtenir une grande diversité de structures [Le Loc'h et al. 1994, Beucher & Renard, 2002].

Inférence du modèle

La simulation d'un réservoir en modèle gaussien seuillé nécessite de préciser

- le nombre de lithofaciès, et leur séquençement ;
- les proportions, variables spatialement ;
- la covariance stationnaire de la Fonction aléatoire gaussienne sous-jacente, ou sa portée

lorsque le type est fixé a priori.

Certains paramètres sont accessibles expérimentalement. Le choix des « faciès » dépend des objectifs de la simulation, qui déterminent le degré de précision recherché. L'ordre des faciès est donné par des considérations géologiques. La stationnarité latérale étant supposée admissible localement, les courbes de proportions verticales sont calculées à partir des sondages disponibles.

Le variogramme de la gaussienne sous-jacente est ajusté indirectement, par comparaison des variogrammes simples et croisés expérimentaux à ceux déduits du modèle. Les conditions d'inférence sont généralement plus favorables pour la portée horizontale que pour la portée verticale, qui peut être « masquée » en présence de fortes variations des proportions verticales.

III.1.2. Modélisation de réservoirs hétérogènes

Etude méthodologique en milieu fluvio-deltaïque [Matheron et al., 1988, de Fouquet et al., 1989]

Au début des années 1980, la modélisation des réservoirs hétérogènes se heurtait au problème majeur de la méconnaissance des structures, à cause du nombre réduit des forages, de leur espacement et de la résolution sismique limitée. Des affleurements supposés similaires aux formations profondes servaient de modèle pour décrire les hétérogénéités, verticalement à l'échelle décimétrique à décamétrique, et horizontalement à l'échelle métrique à kilométrique.

Les affleurements du Jurassique Moyen situés en bord de mer dans le Yorkshire (G.-B.) présentant des alternances argilo-gréseuses similaires à certaines formations fluvio-deltaïques de Mer du Nord, la falaise de Long-Nab servait d'école de sédimentologie aux géologues de plusieurs compagnies pétrolières. Cet affleurement d'environ 30 mètres de hauteur et 650 de longueur a alors été complété pour l'étude tridimensionnelle par le forage d'une trentaine de puits sur une zone d'environ 1 km² à l'arrière de la falaise, suivant une maille d'environ 300m resserrée localement [Ravenne et al., 1987].

La formation a été décrite en trois ou en quatre faciès, évoluant du grès vers l'argile. La coupe géologique avait été discrétisée suivant un maillage d'environ 0.20 m verticalement et 10m horizontalement (Figure 16.a), permettant le calcul des courbes de proportion verticales avec une bonne précision. Ces courbes fournissent une description synthétique de la formation, avec un niveau continu gréseux à la base, surmonté des faciès intermédiaires puis de niveaux fins à nouveau continus, argileux puis gréseux. Les derniers mètres en haut de la falaise présentent des chenaux gréseux dans un ensemble plus argileux (Figure 16.b).

La courbe de proportion calée sur la falaise reste *grosso modo* valide pour les forages. Les variogrammes expérimentaux, calculés en deux dimensions puis en tenant compte des puits, ont été regroupés en quelques zones majeures, afin de contrôler assez précisément le modèle. L'ajustement apparaît globalement satisfaisant, avec des portées respectives de 10m verticalement, 175m le long de la falaise, et le double perpendiculairement, dans la direction présumée des apports. Les résultats auraient pu être améliorés en tenant compte des resserrlements locaux de la maille de sondages qui modifient la courbe de proportion associée aux différents pas.

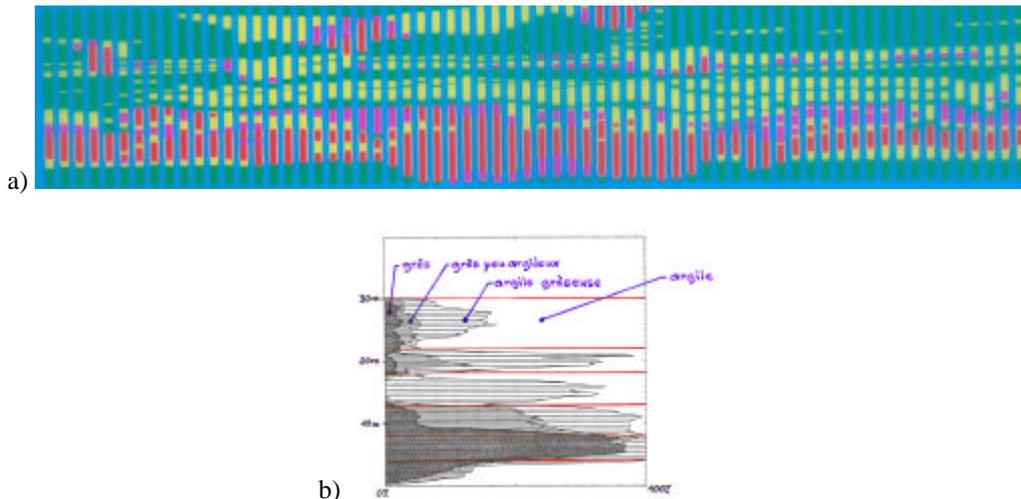


Figure 16. Falaise de Long Nab. a) discrétisation et interprétation en 4 faciès. b) courbe de proportion verticale.

Choix des lignes de corrélation stratigraphique [de Fouquet, Ikonomi et al., 1991]

Après leur dépôt, les unités lithostratigraphiques constituant les réservoirs ont été déformées, compactées ou érodées. Les paléosurfaces se sont modifiées au cours du temps. Différents types de contacts (concordances, discordances diverses) sont observables. Des transformations géométriques simples permettent de rendre compte de ces transformations : les translations, rotations et homothéties, ainsi que le masquage reproduisent des concordances, des compactations, des érosions, etc.

En l'absence de description géologique détaillée, l'étude variographique aide à préciser certaines caractéristiques, en comparant les variogrammes calculés en référence à différentes « horizontales », toit ou mur des principales unités. Aux incertitudes sur les paramètres (courbes de proportions, portées), s'ajoutent celles sur la morphologie de ces unités et sur leur agencement. Les simulations effectuées sous différentes hypothèses permettent de quantifier ces incertitudes relativement à un problème donné : volume de réservoir connecté à un puits, connexité des principales unités.

III.1.3. Corégionalisation entre variable catégorielle et teneur [Freulon et al., 1993]

Le modèle gaussien seuillé permet de représenter commodément la liaison entre formations lithologiques et teneur, en introduisant une corrélation entre variables gaussiennes associées.

Dans le gisement d'Uranium de Laporte (Cantal), les minéralisations se rencontrent à l'intersection d'un réseau de diaclases et de filons de lamprophyres subverticaux, de puissance métrique à plurimétrique, assez continus verticalement et suivant leur extension Nord-Sud. Le levé d'exploitation d'un niveau de la carrière montre que les valeurs fortes de la radioactivité se situent dans les lamprophyres, les valeurs plus faibles diffusant en dehors suivant un cheminement préférentiel le long des diaclases, avec quelques jonctions entre filons voisins. Les lamprophyres contiennent les valeurs les plus fortes, mais ne sont pas intégralement minéralisés.

Comment utiliser la continuité notamment verticale des filons pour construire un modèle « réaliste » des concentrations dans les niveaux inférieurs, à partir du lever du niveau d'exploitation et de sondages, qui mesurent principalement les teneurs ? L'implantation préférentielle mais non systématique des teneurs élevées est prise en compte par une corrélation.

Le modèle est le suivant :

- l'indicatrice des lamprophyres est modélisée par une gaussienne seuillée Y ;
- la teneur est classiquement la transformée par anamorphose d'une gaussienne T ;
- la liaison entre géométrie et teneur est décrite par la corrélation spatiale (Y, T) .

Négligeant un très léger décalage entre lamprophyres et teneurs maximales, un modèle linéaire de corégionalisation est ajusté indirectement sur les variogrammes simples et croisés de l'indicatrice et des transformées gaussiennes des teneurs. Le coefficient de corrélation des gaussiennes (de valeur 0.5) est déterminé par la corrélation expérimentale de l'indicatrice et de la teneur ou de sa transformée gaussienne, calculable lorsque la proportion p est connue. La continuité Nord-Sud des filons est prise en compte via une anisotropie zonale, et leur régularité par une covariance très régulière, le schéma cubique. La gaussienne T de la teneur s'écrit comme une combinaison de Y et d'un résidu de structure spatiale plus irrégulière. Les variogrammes simples et croisés sont ajustés correctement par ce modèle. La teneur simulée conditionnellement à la trace des lamprophyres fournit une image réaliste.

Le modèle gaussien seuillé est utilisable pour contraindre des simulations de teneur par des observations organoleptiques, pour des polluants notamment. Pour l'estimation avec des données en nombre réduit, un modèle plus simple de corrélation spatiale entre teneur et « moyenne par classe » a été jugé suffisant.

III.2. Fonctions Aléatoires liées par des EDP linéaires

La simulation de Fonctions Aléatoires liées par des équations aux dérivées partielles permet d'éviter les biais dus à la discrétisation des calculs par différences finies, et de conditionner de façon exacte la simulation d'une variable par la donnée de ses dérivées en certains points. Cette méthode fournit une solution exacte à des problèmes plus complexes comme l'équation de Poisson, prenant en compte les conditions aux limites là où elles sont connues, mais sans qu'il soit nécessaire de les fixer ailleurs. Une application à la simulation conjointe de la charge et de la transmissivité en écoulement simplifié (cas du perméamètre en régime permanent et en l'absence de recharge) est présentée.

III.2.1. Retour sur la dérivation stochastique

Les simulations conditionnelles de Fonctions Aléatoires gaussiennes ou de FAI-k dans \mathbb{R}^2 ou \mathbb{R}^3 sont connues depuis plus d'une trentaine d'années [Matheron, 1973 ; Journel, 1974 ; Chilès, 1977]. Par ailleurs, le cokrigage d'une fonction aléatoire et de ses dérivées, initié par A. Dong [1990] pour l'estimation d'un champ de transmissivité à partir de données de transmissivité et de charge, a été repris depuis [Roth, 1995 ; Grenier, 1996]. La méthode des tranches tournantes associée à la théorie des FAI-k fournit des solutions remarquablement simples pour la simulation de variables liées par des Equations aux Dérivées Partielles Linéaires.

Rappels sur la dérivation stochastique

Comme pour la continuité, la dérivation stochastique nécessite de distinguer entre les propriétés des Fonctions Aléatoires définies en moyenne quadratique, traduites par les covariances, et celles presque sûres sur les réalisations [Doob, 1953 ; Lévy, 1965]. Les méthodes de simulation présentées garantissent la dérivabilité en ces deux sens : toute réalisation est dérivable en (presque sûrement) tout point.

Rappelons qu'une fonction aléatoire d'ordre deux Z , définie sur \mathbb{R}^n , est différentiable en moyenne quadratique si, en tout point x , il existe un vecteur aléatoire $G_x = (G_1(x), \dots, G_n(x))$ et une fonction aléatoire d'ordre 2, $e_x(h)$, tels que

$$\forall h = (h_1, \dots, h_n), Z(x+h) - Z(x) = \sum_{i=1}^n h_i G_i(x) + e_x(h), \text{ avec } \lim_{|h| \rightarrow 0} E \left[\left(\frac{e_x(h)}{|h|} \right)^2 \right] = 0.$$

La Fonction aléatoire à n dimensions, G_x est le gradient de Z .

La théorie des FAI-k permet de décrire les relations de dérivation entre FA [Matheron 1971, 1973 ; Delfiner et Matheron, 1980]. Les résultats fondamentaux sont les suivants :

- toute primitive d'ordre $k+1$ d'une FASt est une FAI-k. Réciproquement, si une FA est différentiable et si toutes ses dérivées partielles d'ordre $k+1$ sont stationnaires d'ordre deux, alors Z est une FAI-k.

- Si Z est une FAI-k p fois différentiable (avec $k \geq p$), chacune des dérivées partielles d'ordre p est une FAI-(k-p). Si Z est une FAI-0 p fois différentiable, chacune des dérivées partielles d'ordre inférieur ou égal à p est une FASt.

La solution de l'équation de Poisson est utile pour de nombreuses applications :

- Soit Y une FASt. Il existe une unique FAI-1 dont toute représentation Z vérifie $DZ = Y$.

- Soit Y une FAI-k. Il existe une unique FAI-(k+2p) Z telle que $D^p Z = Y$.

Formellement les relations entre covariances (généralisées) se retrouvent en échangeant espérance et dérivation. K désignant la covariance (ou une covariance généralisée) de Z , on a par exemple $K_{ij}(h) = E[G_i(x).G_j(x+h)] = -\frac{\partial^2}{\partial h_i \partial h_j} K(h)$. Le degré des polynômes définissant une FAI-k comme une classe d'équivalence coïncide avec les constantes d'intégration successives. Il en est de même des indéterminations polynomiales des covariances généralisées.

Simulation d'une FA et de son gradient

Dans la suite, Z ou Z_0 désigne une fonction aléatoire, et $Z_i = \frac{\partial Z}{\partial x_i}$ ses dérivées partielles.

Dans le cas stationnaire, les relations différentielles constituent un modèle multivariable particulier, généralisable au cas intrinsèque en considérant la covariance non stationnaire. Les méthodes discrètes sont donc applicables.

La méthode spectrale continue se transpose de façon immédiate, la dérivation formelle étant justifiée. Lorsque la covariance est l'autoconvoluée d'une fonction f continûment dérivable, la simulation de Z et de ses dérivées s'obtient par dilution d'une même « mesure orthogonale » W , par exemple un processus de Poisson, ce qui s'écrit $Z = f * W$ et $Z_i = f_i * W$. Pour ces deux méthodes, les relations différentielles sont vérifiées en tout point sur les réalisations.

Classique en dimension un, la simulation par convolution d'une fonction aléatoire et de ses dérivées offre de nombreuses possibilités, en association avec la méthode des tranches tournantes. Ceci est développé dans la partie suivante.

Le conditionnement s'effectue classiquement par « repiquage » des résidus, effectué par cokrigeage. En effet, le poids pour le krigeage de la dérivée coïncide avec la dérivée du poids de krigeage. De cette propriété découle la commutativité du cokrigeage et de la dérivation. Si la méthode de simulation garantit que les relations différentielles sont vérifiées sur toutes les réalisations, alors il en est de même après conditionnement.

Le cokrigeage conditionnant nécessite toutefois l'expression analytique de toutes les covariances simples et croisées, relations qui ne sont pas nécessaires pour les techniques continues de simulation (spectrale dilution, tranches tournantes).

Dans le cas d'une anamorphose, les réalisations s'écrivent $z = \mathbf{j}(y)$ et $z_i = \mathbf{j}'(y).y_i$. Pour que Z soit différentiable, il suffit que sa transformée gaussienne le soit et que l'anamorphose \mathbf{j} soit dérivable. La simulation de Z et de ses dérivées en découle.

III.2.2. Simulation conjointe par tranches tournantes [de Fouquet, 2000]

Intimement liée à la décomposition spectrale, la méthode des tranches tournantes offre quelques simplifications remarquables notamment pour la résolution de l'équation de Poisson.

Simulation d'une FA et de son gradient

D'après le théorème de Bochner, la covariance K d'une FASt isotrope et différentiable s'écrit $K(h) = E[C(\langle h, S \rangle)]$, C étant une covariance monodimensionnelle et S un vecteur uniforme sur la demi-sphère unité de \mathbb{R}^n . Cette décomposition est le fondement de la méthode des tranches tournantes. Un calcul simple montre que la covariance des dérivées s'écrit $K_j(h) = E[-\langle s_j, S \rangle^2 C''(\langle h, S \rangle)]$. $-C''$ étant la covariance de la dérivée d'une FA de covariance C , les relations précédentes montrent que les FA Z et Z_i s'écrivent respectivement comme les randomisées d'une FA monodimensionnelle Y ou de sa dérivée Y' , à un facteur près pour les dérivées. Ce facteur est égal au produit scalaire entre la direction de dérivation s_i et la direction S uniforme : si $Z(x) = Y(\langle x, S \rangle)$ alors $Z_i(x) = \langle s_i, S \rangle Y'(\langle x, S \rangle)$.

Simuler Z et ses dérivées par tranches tournantes revient à épandre respectivement une FA Y ou sa dérivée. A partir d'un seul épandage de la dérivée Y' , les dérivées de Z dans les différentes directions s'obtiennent à un facteur correctif près, dépendant de la direction de dérivation. Si les relations différentielles au sens des fonctions sont vérifiées en dimension un sur les réalisations (y, y') , elles le restent après épandage entre z et z_i .

Dans son DEA, M. Cuypers [1997] a comparé la simulation conjointe des dérivées au calcul par différences finies, validant ces méthodes.

Ces résultats se généralisent aux FA non isotropes. La loi de S n'est plus nécessairement uniforme sur la demi sphère unité, et la loi du rayon vecteur dépend de la direction.

On construit ainsi par tranches tournantes des champs conservatifs, comme le rotationnel d'un champ vectoriel, ou irrotationnel, comme le gradient d'un champ scalaire.

Equation de Poisson

Z étant obtenue par épandage de Y , soit R la FA isotrope obtenue par épandage de la dérivée Y' . La covariance de R vérifie $G(h) = -E[C''(\langle h, S \rangle)]$. Un calcul simple sur les transformées de Fourier montre que cette covariance est, au signe près, le laplacien de la covariance K : $G = -DK$. Si Y' est stationnaire, alors Y et Z sont des FAI-0.

Matheron [1973] a montré que la covariance du laplacien est le laplacien itéré de la covariance : $D^{(2)}K_Z = K_{DZ}$. On construit des solutions de l'équation de Poisson en itérant le procédé.

FA		covariances	
sur \mathbb{R}	épandage dans \mathbb{R}^n	sur \mathbb{R}	épandage dans \mathbb{R}^n
Y	Z	C	K
Y'		-C''	-DK
Y''	DZ	C ⁽⁴⁾	DDK

Tableau 2. Relations différentielles par la méthode des tranches tournantes.

Soit X une FA isotrope construite par épandage de R , de covariances respectives K_X et C . En dimension trois (ou par plongement de \mathbb{R}^2 dans \mathbb{R}^3), la solution Z de l'équation $DZ = X$ est construite par épandage de T , intégrale seconde de R sur les droites : $T'' = R$. Si

X est stationnaire, alors R est également stationnaire, et leurs intégrales secondes T et Z sont des FAI-1. Le tableau 2 récapitule ces relations.

III.2.3. Simulation d'un écoulement simplifié

A deux dimensions, le modèle stochastique le plus simple est un écoulement macroscopiquement linéaire, en régime permanent et en l'absence de recharge, dans un champ de transmissivité isotrope.

Soit $T = T_0 e^Q$ une transmissivité lognormale, Q représentant la perturbation au premier ordre : $T \approx T_0(1+Q)$. Pour un gradient de charge macroscopique négatif $-J$ orienté suivant Ox , la charge en $x = (x_1, x_2)$ s'écrit $H(x_1, x_2) = J(x_0 - x_1) + F(x_1, x_2)$ (Figure 17). F représente donc la perturbation de la charge. La linéarisation de l'équation de la diffusivité $div(T.GradH) = 0$ conduit à la relation bien connue entre perturbations : $JQ = DZ$ et $F = \frac{\partial Z}{\partial x_1}$. Si la transmissivité est une FA stationnaire, alors Z est une FAI-1 et F une FAI-0 [A. Dong, 1990].

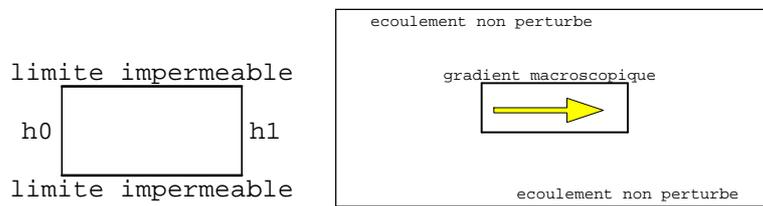


Figure 17. Conditions aux limites pour le calcul discrétisé (exemple du perméamètre) et pour le modèle géostatistique.

Perturbation		Simulation sur \mathbb{R}^n	Simulation sur \mathbb{R}
transmissivité		Q	U
			$V = J \int U$
gradient de charge	parallèle	$\partial F / \partial x_1$	$\langle S.s_1 \rangle^2 JU$
	perpendiculaire	$\partial F / \partial x_2$	$\langle S.s_1 \rangle \langle S.s_2 \rangle JU$
charge		F	$\langle S.s_1 \rangle V$

Tableau 3. Simulation d'un écoulement simplifié par tranches tournantes. Le gradient macroscopique est orienté suivant Ox . s_1 et s_2 désignent le vecteur directeur, respectivement parallèle ou perpendiculaire à ce gradient.

Les perturbations de la transmissivité et de la charge, et par suite ces grandeurs elles-mêmes, sont directement simulables par la méthode des tranches tournantes, à partir de la seule expression de la covariance des perturbations de la transmissivité. Le plus simple consiste à procéder par plongement de \mathbb{R}^2 dans \mathbb{R}^3 , en effectuant sur chaque droite une correction d'échelle pour la covariance monodimensionnelle.

Les FA à une dimension associées aux différentes perturbations (transmissivité, gradient de charge, charge) sont reportées au tableau 3. Le flux (ou sa perturbation) s'en déduisent en appliquant la loi de Darcy à la transmissivité et au gradient de charge simulés.

Que deviennent les conditions aux limites dans ce modèle ? Le long de chaque droite, U est nul en-dehors d'un segment central ; V est nulle à une extrémité, constante et d'espérance nulle à l'autre. Un raisonnement géométrique élémentaire montre qu'en un point extérieur au pavé central simulé, le nombre de droites contribuant à la variabilité décroît avec la distance au pavé. A l'infini, seule une droite a une contribution non nulle, mais négligeable. A l'infini, la simulation se raccorde ainsi à l'écoulement macroscopique non perturbé.

La comparaison de simulations non conditionnelles au calcul par différences finies montre pour la charge des écarts croissants avec la variance logarithmique de la transmissivité, le nuage restant globalement linéaire avec un coefficient de corrélation élevé [P. Mariotti et al., 2000]. Les limites à charge imposée du calcul par différences finies apparaissent clairement sur les perturbations, la simulation géostatistique autorisant la variabilité spatiale (Figure 18).

Le conditionnement aux données de charge ou de transmissivité s'effectue classiquement par cokrigage. M. Tijani (2004, communication personnelle) a proposé une méthode simplifiant le calcul analytique des covariances simples et croisées des perturbations. Là où elles sont connues, les conditions aux limites sont prises en compte par le conditionnement.

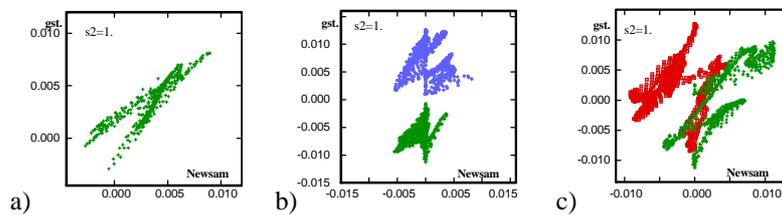


Figure 18. Nuage de corrélation des perturbations de la charge entre le calcul par différences finies (code Newsam, CIG - Ecole des Mines de Paris) et la simulation géostatistique, en ordonnée. a) zone centrale, 20mx20m. b) limites à charge imposée, 10mx100m. c) limites imperméables, 100mx10m. La variance logarithmique de la transmissivité est égale à 1.

Ces simulations conditionnelles sont utilisables dans le cadre du problème inverse, en complément aux calculs d'écoulement sur simulations géostatistiques de la transmissivité. Elles permettent de « relâcher » l'influence de conditions aux limites mal connues. L'application à un cas réel reste à effectuer.

PERSPECTIVES

Quelques développements ont été mentionnés aux chapitres précédents. Plusieurs perspectives sont proposées ici. Tout d'abord, « l'auscultation de données » mériterait d'être poursuivie. Après formulation dans le cadre des modèles usuels (principalement le modèle gaussien anamorphosé et ses variantes multivariées) des solutions exactes ou des approximations « praticables » seront recherchées aux problèmes posés par la réglementation environnementale. Enfin, des modélisations spécifiques sont à construire, comportant le plus souvent une forte composante de simulations numériques.

Quantifier l'évolution de la pollution dans les différents milieux

Les banques de données de la qualité de l'air ou de l'eau comportent une grande quantité de mesures quasiment inexplorées, avec des chroniques couvrant parfois plusieurs décennies. Malgré leurs imperfections, l'étude de ces données permettrait d'améliorer la connaissance du niveau des pollutions et d'en quantifier l'évolution, même de façon approchée. Ainsi, la comparaison de mesures de pollution atmosphérique par des analyseurs implantés dans des milieux périurbains supposés a priori « analogues », a montré une diversité « inattendue » des niveaux, mais aussi de l'évolution récente pour différents polluants [rapport d'étude PREDIT, de Fouquet, 2005].

Face aux données disponibles, il s'agit de sélectionner quelques questions (quelles substances, quels milieux, quels sites, quelles périodes, quels critères de comparaison), puis de mener l'analyse exploratoire appropriée. Les estimations spatiales sont rarement possibles, l'échantillonnage n'ayant pas été prévu à cette fin.

L'exploitation de ces données pourrait également devenir intéressante lorsque l'on cherche à reconstituer des expositions durant de longues périodes, par exemple pour des études épidémiologiques. Si une méthode de calcul d'exposition à un polluant est raisonnablement transposable aux substances pour lesquelles des mesures sont disponibles, que donne la confrontation des valeurs calculées aux données disponibles en différents sites ? Les « nouveaux polluants » (particules atmosphériques fines, pesticides dans l'air ou dans l'eau) sont cependant absents de ces bases, ou leurs chroniques trop récentes.

Réglementation et indicateurs environnementaux

La réglementation sur l'environnement s'étoffe régulièrement, concernant des substances toujours plus nombreuses, en particulier dans l'eau et dans l'air. L'application effective de cette réglementation pose d'intéressantes difficultés conceptuelles et pratiques.

Les problèmes liés aux *limites de quantification* ont été mentionnés précédemment, en particulier à propos des « micropolluants ». Dans tous les milieux se pose la question d'évaluer un risque de dépassement de seuil par substance, en présence de valeurs « non quantifiées » parfois nombreuses. Raisonner par famille de substances pose par ailleurs le problème de la comparaison à un seuil pour une somme pas toujours bien définie de constituants, de surcroît rarement mesurés simultanément. Des applications du modèle gaussien seuillé sont envisageables, ainsi que peut-être des modèles de variables poissonisées, c'est-à-dire la somme d'un nombre aléatoire de composantes.

Divers *indicateurs* sont préconisés, en vue de quantifier de façon synthétique l'influence de décisions réglementaires, comme la réduction ou l'interdiction de certaines substances (phosphates, pesticides, adjuvants de carburants ...), ou pour suivre l'évolution du niveau des

pollutions. Mais ces indicateurs ne sont pas toujours pertinents : les grandeurs à calculer sont parfois définies de façon inappropriée (confusion entre moyenne spatiale ou temporelle et espérance d'un processus, par exemple), les corrélations spatiales ou temporelles sont généralement négligées. En pollution atmosphérique, certains calculs sont compliqués, car ils font intervenir plusieurs supports temporels, variables selon les substances, ainsi que des opérateurs comme le seuillage ou parfois le maximum. Il devrait être possible de proposer des indicateurs adaptés aux objectifs visés mais plus commodes à calculer, ce qui permettrait en particulier de mieux quantifier les écarts interannuels.

Des approximations pour le calcul effectif de certains indicateurs actuellement préconisés pourront être recherchées. Préciser le domaine de validité de ces approximations nécessite d'en examiner la sensibilité à différentes hypothèses. Traiter des exemples pratiques permettra de vérifier la cohérence des calculs demandés par la réglementation et des mesures usuellement disponibles, et aussi d'évaluer l'ordre de grandeur des incertitudes. Des simulations seront ici commodes. Enfin, il sera utile de préciser des critères de comparaison pour caractériser les variations interannuelles, et en examiner le caractère significatif. La question se pose notamment pour les quantiles d'ordre élevé mais non extrêmes (portant sur un faible pourcentage de valeurs élevées), pour lesquels les méthodes examinées par C. Bernard-Michel [2006] sont applicables, ainsi que pour les d'indicateurs d'exposition en excès, équivalents au bénéfice conventionnel.

L'indicateur étant précisé, *comment le calculer en-dehors des sites de mesures ?* En pollution atmosphérique, plusieurs supports temporels interviennent simultanément (par exemple, huit et vingt-quatre heures pour la valeur cible en ozone pour la santé humaine (cf. II.5.5), une et trois heures pour le seuil d'alerte en NO₂). Ces questions se rattachent à l'estimation non linéaire et aux calculs de sélectivité, développés en géostatistique minière.

La solution, actuellement très lourde à implémenter, consistant à comparer sur simulation les concentrations aux seuils réglementaires pourrait être allégée, en réduisant par exemple les périodes simulées aux seuls épisodes « incertains ». Comment alors détecter ces épisodes et évaluer l'incertitude sur l'indicateur calculé ? Modélisation et validation par simulation sont ici nécessaires.

Toujours en pollution atmosphérique, le support et le champ temporel sont définis précisément, alors que le support spatial, sans doute supposé implicitement ponctuel, est omis. Examiner l'influence de ce support est important pour le couplage pratique entre les mesures aux stations et les prévisions des modèles phénoménologiques, effectuées sur des mailles souvent kilométriques ou hectométriques.

Un autre aspect de la réglementation concerne le *contrôle* sur le terrain, en dépollution des sols ou pour le suivi de la qualité des eaux d'anciens sites miniers par exemple. Supposons le support spatial et temporel précisé par la réglementation. Quelques mesures effectuées par intermittence en quelques stations ne peuvent garantir que la concentration dans toute une zone et durant une période données ne dépasse pas le seuil fixé, si ces données « négatives » restent proches du seuil. Réciproquement, une mesure ponctuelle « positive » ne signifie pas nécessairement qu'une moyenne spatiale ou temporelle dépasse le seuil. On pourrait construire des intervalles de probabilités conventionnels mais pertinents. Les modèles usuels permettent de tels calculs, tout en tenant compte du changement de support.

Informations auxiliaires et traitements intégrés

La prise en compte d'informations auxiliaires améliore la précision des estimations en les rendant plus réalistes. Ces informations auxiliaires peuvent aussi corriger les effets d'un échantillonnage préférentiel. Les développements techniques, en particulier la géophysique, fournissent de nouvelles « variables auxiliaires » issues de mesures indirectes, dont il faut préciser les caractéristiques et évaluer les liaisons avec les concentrations.

L'exigence d'une démarche plus intégrée se fait jour, par exemple en *pollution de sols*. Une couverture assez dense par des mesures de terrain peu coûteuses est complétée par des mesures en laboratoire, plus précises mais moins nombreuses. Quelles approximations mettre en œuvre pour automatiser les méthodes d'estimation linéaire et non-linéaire, cette fois-ci en contexte multivariable ? Il s'agit à terme de détecter les zones dans lesquelles un échantillonnage supplémentaire est nécessaire, et de fournir des cartes de concentration ou de risque de dépassement de seuil. Ceci nécessite d'évaluer les incertitudes, sachant que le traitement du site permettra ensuite un contrôle sur le terrain. Il faudra aussi élaborer des critères pour détecter les cas où le modèle géostatistique serait en désaccord avec les données expérimentales. Enfin, la « robustesse » des calculs devra être vérifiée (ajustement des variogrammes, choix du voisinage, détermination de l'anamorphose, calage du modèle bivariable des transformées gaussiennes, calcul du coefficient de changement de support). L'automatisation des calculs pose ainsi d'intéressantes questions de délimitation pratique de la validité des modèles.

Le programme LOQUAS (Localisation et QUantification d'une pollution organique de Sols, conduit par l'IFP et associant Vinci technologies, Arcadis ESG, l'Ecole des Mines de Paris -ARMINES et le BRGM), qui vient de démarrer pour une durée de trois ans, vise à une intégration géostatistique de données géophysiques et de mesures PollutEval sur site. Cet appareil permet une identification rapide d'hydrocarbures pétroliers par l'analyse de leur « pyrogramme ». Mais le volume des échantillons prélevés pour les analyses sur le terrain étant réduit d'un ordre de grandeur par rapport aux prélèvements classiques, il convient de commencer par préciser la « représentativité » de ces prélèvements, pour en déduire des recommandations pratiques pour l'échantillonnage sur site. Il est prévu des expériences de terrain pour examiner la variabilité spatiale à très petite distance, aux échelles inframétriques, et adapter en conséquence les schémas de reconnaissance. Une confrontation à des simulations numériques est prévue, pour valider les modèles.

Incertitudes dans les modèles hydrodynamiques

L'évaluation du terme source pour le transfert d'une pollution vers la nappe se révèle encore souvent impossible, l'échantillonnage restant limité aux niveaux superficiels à l'étape de l'évaluation des risques. Examinant, dans la thèse que je codirige, le cas d'une pollution très localisée, S. Mazuel cherche à évaluer les incertitudes sur le temps d'arrivée à la nappe d'un traceur de l'eau, donc en l'absence d'interaction chimique. La modélisation du transfert en zone non saturée nécessite le calage des paramètres des modèles hydrodynamiques, dont il examine par simulation l'influence respective.

Lorsque des chroniques de concentrations sont disponibles en quelques piézomètres, comment caler le modèle hydrodynamique ? Sur le résultat recherché, quelle est la part d'incertitude provenant respectivement des paramètres hydrodynamiques, du terme source, ou des fluctuations des simulations à modèle fixé ?

Ces questions rejoignent le vaste champ du problème inverse en hydrogéologie [de Marsily, 1999 ; 2000 ; 2004 ; Chilès, 2001 ; Pasquier et Marcotte, 2005]. En exploitation pétrolière, l'intégration de données de différentes origines, notamment les essais de puits et l'historique de production, a fortement progressé ces dernières années. Par exemple, la méthode de déformation graduelle permet de décrire une famille de modèles numériques de réservoirs avec un paramétrage réduit, la combinaison ajustant au mieux les « observations » étant alors sélectionnée [Hu, 2000 ; Hu et Le Ravalec-Dupin, 2004]. Transposer ces méthodes en hydrogéologie nécessite des adaptations à un contexte différent : cas des nappes libres et des écoulements monophasiques, tests de traçage remplaçant les essais de puits.

Sur un cas synthétique, il sera instructif de comparer ce conditionnement à la modélisation globale par surface de réponse (cf. I.2.3.). Quels sont en particulier les avantages

et les inconvénients de ces deux classes de méthodes, par exemple suivant le type et la densité d'information disponibles ?

Un cas particulier important concerne l'estimation ou le conditionnement de perméabilités tensorielles, par exemple par le croisement d'écoulements avec différentes conditions aux limites. Une fois caractérisée la structure spatiale, que devient le changement d'échelle pour ces milieux ?

La complexité des milieux et des problèmes implique de traiter ces questions à l'aide de simulations numériques. Mais préciser les modèles sous-jacents reste indispensable pour identifier certaines approximations et délimiter les performances de ces méthodes.

Géochimie

La modélisation géochimique en milieu hétérogène est une question émergente, motivée par l'évaluation du transfert à la nappe de polluants pour les études de risque, l'exploitation de gisements par lixiviation in situ, ou l'étude de la faisabilité du stockage de CO₂ dans des formations sédimentaires.

Les problèmes de transport réactif couplent de manière intime écoulement, transport en solution et géochimie, avec notamment la prise en compte des phénomènes de dissolution-précipitation. La variabilité spatiale intervient à travers les paramètres de l'écoulement, ainsi que par les concentrations initiales. Quelle en est l'influence par rapport aux prévisions données par un modèle homogène ? Cette question est abordée par Marco De Lucia dans sa thèse, suivie conjointement avec V. Lagneau (Ecole des Mines de Paris).

Plusieurs problèmes apparaissent. En raison des contraintes liées à la résolution du problème de transport réactif (temps de calcul, mais aussi forme particulière des maillages imposés pour résoudre des problèmes pratiques), il est nécessaire de traiter du changement d'échelle en géométrie complexe et en terme de transport. Différents calculs de perméabilité équivalente (interblocs ou intrablocs) ont d'abord été comparés, en retenant un critère de restitution du traceur.

La variabilité spatiale pouvant alors être prise en compte dans le calcul d'écoulement, reste à en quantifier l'influence. Le problème est complexe, les caractéristiques du milieu comme la porosité et la perméabilité évoluant avec la chimie du système, du fait de la modification du volume du solide lors des réactions de dissolution et de précipitation. Cette question est actuellement abordée dans le cadre simplifié du perméamètre, dans le cas d'un maillage régulier. Il s'agit actuellement de préciser les conditions d'apparition de phénomènes comme la formation de digitations ou le lissage des fronts.

La question du calage des « paramètres » du modèle (paramètres d'écoulement, mesures de concentrations initiales ou en régime transitoire) se posera pour les applications ; l'expérience acquise par l'étude des écoulements sera utile. Enfin, il sera intéressant de vérifier si des modèles simplifiés (changement d'échelle, simplification du « réacteur » chimique) permettent une approximation suffisante des prévisions données par la modélisation détaillée.

Ces développements impliquent de généraliser les résultats qui seront obtenus par simulations numériques pour des exemples synthétiques ou sur des cas particuliers.

La prévision des performances des méthodes de dépollution des sols élargit encore le problème. Plusieurs variantes de traitement in-situ consistent à faire circuler un fluide (eau, air) éventuellement chargé en nutriments, pour favoriser le développement de bactéries qui dégradent la pollution. Aux problèmes de modélisation du milieu (dans ce cas, le sol), des concentrations initiales et du transport (de l'eau ou de l'air), s'ajoute la modélisation des effets de l'activité bactérienne. Un état des lieux des connaissances dans ces divers domaines serait nécessaire, avant d'aborder le couplage de différents modèles pour les études par simulations.

Intégration des modèles phénoménologiques et géostatistiques

Associer géostatistique et modélisation phénoménologique permet une meilleure valorisation de l'information quantitative disponible. La géostatistique présente l'avantage de tenir compte du support des données, différant du maillage utilisé par les codes numériques. Lorsque les données sont suffisantes, l'analyse exploratoire permet d'autre part une « auscultation » parfois assez précise du phénomène.

Outre l'assimilation de données, développée pour le recalage des prévisions phénoménologiques aux mesures, associer ces deux modélisations est nécessaire dans différents contextes :

- la « validation de code », pour confronter les résultats d'un code numérique aux mesures. La question se pose par exemple pour l'évaluation de la part des émissions en CO₂ à attribuer à certaines activités économiques (transport, chauffage, etc.), préalablement à leur réduction réglementaire. Avant toute répartition par secteur, il convient de vérifier si les évaluations de l'état global actuel sont en accord avec les mesures. Le support spatial des mesures étant inférieur de plusieurs ordres de grandeur à celui mis en œuvre dans les codes phénoménologiques, il faut d'abord revenir sur les critères permettant de valider ou non les résultats numériques. Des études par simulation de corrélation entre résultat numérique et « mesures » peuvent par exemple être développées.

- le calage des modèles phénoménologiques aux mesures. Des méthodes simplifiées peuvent être utilisées en alternative à l'assimilation de données, par exemple pour modéliser le fonctionnement d'un bassin versant. Une comparaison de différentes méthodes de recalage serait intéressante, par exemple en réservant pour la validation une partie des mesures, lorsque celles-ci sont en nombre suffisant.

- ou réciproquement, le calage du modèle géostatistique (variographie, estimation) à la phénoménologie.

Mener l'étude exploratoire en relation étroite avec les connaissances phénoménologiques reste encore peu fréquent. La construction d'un modèle géostatistique des écarts entre les observations et le modèle phénoménologique implique le choix des variables à considérer, la prise en compte des différents supports, et l'intégration de différentes corrélations (par exemple pour un bassin versant : débits, flux des divers nutriments, etc...), voire leur réduction à quelques facteurs synthétiques principaux.

Le couplage de ces modèles serait également utile pour la gestion « durable » des ressources rares, par exemple de certaines nappes (en quantité et en qualité) en agriculture, ou pour la gestion « de précision » des eaux superficielles (en débits et en concentrations).

Modélisation spatio-temporelle sur des graphes

La modélisation géostatistique des débits et des concentrations le long de réseaux hydrographiques en est à ses débuts. La prochaine étape consistera à caler des modèles spatiaux débits-flux-concentrations sur quelques réseaux, pour plusieurs nutriments simultanément (nitrates et nitrites, autres produits azotés ou phosphorés), et en relation avec les milieux traversés. Ce calage s'effectue sur la base d'une analyse exploratoire, en référence à des modèles multivariés classiques (modèle linéaire de corégionalisation, modèle à résidu), mais avec des modèles variographiques spécifiques. La topologie du réseau nécessite l'adaptation des outils variographiques, pour si possible se ramener à des analogues d'incrément stationnaires sur l'arborescence. Les outils d'inférence à développer différeront, selon que les débits sont ou non supposés connus.

La modélisation spatio-temporelle est nécessaire pour tenir compte de mesures asynchrones, ou pour effectuer des estimations sur de courtes périodes. Le statut de la périodicité annuelle est à examiner : celle-ci est-elle mieux prise en compte à travers une dérive, ou par une composante périodique du variogramme ? La forme de la covariance spatio-

temporelle dépend de l'échelle de temps considérée, suivant qu'une composante propagative est ou non nécessaire.

Comme indiqué précédemment, la modélisation spatio-temporelle détaillée s'effectuera de préférence en relation avec un modèle phénoménologique de fonctionnement de bassins versants, capable par exemple de rendre compte de l'influence des pluies sur les débits. Le nécessaire recalage du modèle et des données associe là encore étude exploratoire et modélisation, pour tenir compte des différents supports, spatiaux et temporels.

Enfin, il sera intéressant d'étendre ces modèles à d'autres substances, comme les pesticides.

L'étude de la qualité de l'eau dans les réseaux de distribution urbains, ou de la pollution (sols, air) à proximité des axes routiers, nécessite enfin de généraliser la modélisation à d'autres types de graphes. On pourra chercher à combiner ou à composer des Fonctions Aléatoires monodimensionnelles définies sur les chemins présents sur le graphe. Pour préciser les modèles, il convient de revenir aux phénomènes étudiés, en tenant compte en particulier de l'orientation éventuelle du graphe. Comme dans le cas des réseaux hydrographiques, les méthodes d'inférence dépendront de l'implantation des mesures, selon que les arêtes comportent un ou plusieurs points expérimentaux.

« *En guise de conclusion* »

La géostatistique environnementale offre une grande diversité de sujets, de l'analyse exploratoire aux développements théoriques.

Lorsque l'information le permet, « l'auscultation des données » aboutit fréquemment à une description détaillée des cas examinés. Les observations issues de cas particuliers ont souvent une portée générale, précisant les connaissances phénoménologiques, ou les complétant. La synthèse des quelques exemples étudiés enrichit ainsi les connaissances et approfondit notre compréhension des phénomènes.

Par ailleurs, l'échelle de travail (un site pollué, une agglomération, une nappe) a plusieurs fois permis la validation expérimentale des hypothèses, pour expliquer les particularités observées, ou pour confirmer les résultats d'estimation. Pas si fréquent, ce retour d'expérience illustre utilement le bien-fondé de la démarche.

Réciproquement, « l'apport de l'environnement à la modélisation géostatistique » est prometteur. Résoudre de nouveaux problèmes pratiques mène très vite à aborder de nouveaux problèmes théoriques, comme la construction de fonctions aléatoires sur des graphes. Ainsi, la prévision du transfert à la nappe rejoint des sujets très actuels : le problème inverse pour les écoulements souterrains, offrant de nouveaux champs d'application aux développements les plus récents de la géostatistique pétrolière, et la modélisation hydrogéochimique.

Enfin, anticipant sur les questions futures concernant le niveau *actuel* des concentrations pour certains polluants, il serait utile d'examiner dès maintenant l'information qui pourra raisonnablement être reconstituée. Détecter les lacunes actuelles de l'échantillonnage aux différentes échelles pour répondre à de telles questions, et y remédier, améliorerait considérablement la précision des futures estimations.

REFERENCES BIBLIOGRAPHIQUES

- Alfaro M. 1979. Etude de la robustesse des simulations de fonctions aléatoires. Thèse de doctorat. Ecole des Mines de Paris.
- Almeida J., Santos C. 2001. Characterization of pine tree productivity index. *Proceedings of the Third European Conference on Geostatistics for Environmental Applications held in Avignon, France, November 22-24, 2000*. Kluwer Academic Publishers.
- Amalric L., Baran N. 2004. Caractérisation de la contamination des eaux souterraines par les pesticides – phase I. Rapport final. BRGM/RP-53341-FR
- Asli M., Marcotte D., Chouteau M. Direct inversion of Gravity data by cokriging. In : *Geostatistics 2000 Cape Town*, W. Kleingeld and D.G. Krige (eds.) 64-73
- ASPA. 2002. Observatoire de la qualité de l'air dans la vallée de la Thur. Rapport 02051701-ID. Schiltigheim.
- Bailly J.-S., Monestiez P., Lagacherie P. 2006. Exploring spatial variability along drainage networks with geostatistics. *Mathematical Geology*, 38(5).
- Bernard-Michel C., de Fouquet C. 2005. Geostatistical indicators of waterway quality for nutrients. In *Geostatistics Banff 2004*. Leuangthong O., Deutsch C.V. (eds) Springer.
- Bertino L. 2001. Assimilation de données pour la prédiction de paramètres hydrodynamiques et écologiques : cas de la lagune de l'Oder. Thèse de Doctorat de l'Ecole des Mines de Paris.
- Beucher H., Renard D. 2002. Reservoir characterization. Centre de Géostatistique, Ecole des Mines de Paris.
- Biau G. 1999. Downscaling of precipitation combining kriging and empirical orthogonal function analysis. In *geoENV II - Geostatistics for environmental applications*. Gomez-Hernandez, Soares and Froidevaux (eds) Kluwer.
- Biau G., Zorita E., von Storch H., Wackernagel H. 1999. Estimation of precipitation by kriging in the EOF Space of the Sea level pressure Field. *Journal of Climate*, 12, 1070-1085.
- Blond N. 2002. Assimilation de données photochimiques et prévision de la pollution atmosphérique. Thèse de doctorat de l'Ecole Polytechnique, Palaiseau, France.
- Bobbia M., Mietlicki F., Roth C. 2000. Surveillance de la qualité de l'air par cartographie : l'approche géostatistique. Prix du meilleur poster INRETS, 5-8 juin Avignon, France.
- Bourgault G., Marcotte D. 1993. Spatial filtering under the linear coregionalization model. *Geostatistics Troia '92*, A. Soares (ed). Kluwer 237-248.
- Bourgin B., Chilès J.-P., Watremez P. 2001. Space-time modelling of sand beach data. In *Geoenv III. Geostatistics for environmental applications*. P. Monestiez et al. (eds), 113-123. Kluwer Academic Publishers.
- Bruno R., Palumbo V., Bonduà S. 2001. Identification of regional variability component by geostatistical analysis of streams sediments. In *Geoenv III. Geostatistics for environmental applications*. P. Monestiez et al. (eds), 113-123. Kluwer Academic Publishers.
- Bruno R., de Fouquet C. 2005. The expansion of environmental geostatistics. In *Space, structure and randomness. Contribution in honour of Georges Matheron in the fields of geostatistics, random sets, and mathematical morphology*. Bilodeau M., Meyer F., Schmitt M. (eds) Springer 105-133
- Cardenas G., Malherbe L. 2004. Géostatistique et prise en compte de l'aspect temporel LCSQA, étude n°12. INERIS.
- Casado L.S., Rouhani S., Cardelino C.A., Ferrier A.J., 1994. Geostatistical analysis and visualization of hourly ozone data. *Atmospheric Environment* 28, n°12, 2105-2118.
- Chilès J.-P. 1977. Géostatistique des phénomènes non stationnaires (dans le plan). Thèse de docteur-

- ingénieur, université de Nancy I.
- Chilès J.-P., Delfiner P. 1999. Geostatistics : modelling spatial uncertainty. N.Y. Wiley.
- Chilès J.-P. 2001. On the contribution of hydrogeology to advances in geostatistics. *In Geoenv III. Geostatistics for environmental applications. P. Monestiez et al. (eds), 1-16. Kluwer Academic Publishers.*
- Cori A. 2005. Utilisation de données environnementales pour la cartographie du NO₂ sur la région Rouennaise. Rapport de stage d'option, 1^{ère} partie. Ecole des Mines de Paris & AirNormand/Remappa.
- Cori A. 2005. Représentativité spatiale des stations de mesure de la concentration moyenne annuelle en NO₂. Rapport de stage d'option, 2^{ème} partie. Ecole des Mines de Paris & AirNormand/Remappa.
- Cressie N., Kaiser M.S., Daniels M.J., Aldworth J., Lee J., Lahiri S. N., Cox L.H. 1999. Spatial analysis of particulate matter in an urban environment. *In geoENV II - Geostatistics for environmental applications. Gomez-Hernandez, Soares and Froidevaux (eds) Kluwer.*
- Cuyppers M. 1997. Simulation non conditionnelle d'une fonction aléatoire et de son gradient par dilution poissonnienne. Etude géostatistique d'un aquifère. Mémoire de DEA Méthodes quantitatives et modélisation des bassins sédimentaires.
- Delfiner P., Matheron G. 1980. Les fonctions aléatoires intrinsèques d'ordre k. Note du centre de géostatistique, Ecole des Mines de Paris.
- Delhomme J.-P. 1970. Application de la théorie des variables régionalisées à l'hydrogéologie. Construction automatique de cartes d'une grandeur mesurée ponctuellement. Ecole des Mines de Paris, CIG, Fontainebleau. LHM/R71/4
- Demougeot-Renard H. 2004. De la reconnaissance à la réhabilitation des sols industriels pollués : estimations géostatistiques pour une optimisation multicritère. Thèse de doctorat, ETH Zürich.
- Deraisme J., Bobbia M. 2003. L'apport de la géostatistique à l'étude des risques liés à la pollution atmosphérique. *Environnement, Risque & santé*. 2(3)
- Desbarats A.J., Dimitrakopoulos R. 2000. Geostatistical simulation of regionalized pore-size distributions using min/max autocorrelation factors. *Mathematical Geology* 32(8), 919-941.
- Dong A. 1990. Estimation géostatistique des phénomènes régis par des équations aux dérivées partielles. Thèse de docteur en géostatistique. Ecole des Mines de Paris.
- Doob J. 1953. Stochastic processes. Wiley.
- Emery X.. 2006. Ordinary multigaussien kriging for mapping conditional probabilities of soil properties. *Geoderma*, 132(1-2), 75-88.
- Englund E., Heravi N. 1993. Conditional simulation: practical application for sampling design optimization. *Geostatistics Troia '92. A. Soares (ed.)*. 613-624
- Flatman G., Yfantis A. 1984. Geostatistical strategy for soil sampling : the survey and the census. *Environmental monitoring and assessment*, 335-349.
- Figueira R., Sousa A.J., Durao F., Pacheco A.M.G., Catarino F. 2001. Spatio-temporel stochastic modelling of saline deposition measured by biomonitors. - *In Geostatistics 2000 Cape Tow: W. Kleingeld and D.G. Krige (eds) 489-506.*
- Flipo N. 2005. Modélisation intégrée des transferts d'azote dans les aquifères et les rivières. Application au bassin du Grand Morin. Thèse de doctorat, Ecole des Mines de Paris.
- de Fouquet C., Beucher H., Galli A. et Ravenne C. 1989. Conditional simulation of random sets : Application to an argillaceous sandstone reservoir. *In M. Arsmtrong (ed.), Geostatistics, Dordrecht : Kluwer*, 517-530.
- de Fouquet C., Ikonomi A., Beucher H., Groupe Heresim. 1991. Choix des lignes de corrélation stratigraphique pour les simulations conditionnelles de lithofaciès. *Sciences de la Terre, Sér. Inf.*, 30, 87-118.
- de Fouquet C., Mandallaz D. 1993. Using geostatistics for forest inventory with air cover : an example. *In Geostatistics Tróia '92, A. Soares (ed.)*, Dordrecht : Kluwer, 875-886.
- de Fouquet C., Courtois N. 1997. Suivi de l'étude géostatistique de la nappe alluviale au confluent du Drac et de l'Isère (Site M27, CEA-Grenoble). Rapport intermédiaire 1^{ère} partie, contrat CEA-ARMINES n° 7020.7B031360.

- de Fouquet C. 1999. Représentativité spatiale des mesures de pollution dans les sols : des échantillons aux estimations. *Les techniques de l'industrie minière*, no. 3, 3^e trim. 24-32, 145-155.
- de Fouquet C. 2000. Joint simulation of a random function and its derivatives. In *Geostats 2000 Cape Town*, W. Kleingeld and D.G. Krige (eds), vol. 1, 84-93
- de Fouquet C. 2000. Construction d'un réseau représentatif de qualité des cours d'eau. Phase I : analyse exploratoire des données. Contrat BETURE CEREC/ARMINES Construction d'un réseau représentatif EUROWATERNET de qualité des cours d'eau. Ecole des Mines de Paris, centre de Géostatistique.
- de Fouquet C., Bez N. 2001. Construction d'un réseau représentatif de qualité des cours d'eau. Rapport final. Contrat BETURE CEREC/ARMINES. Ecole des Mines de Paris, centre de Géostatistique.
- de Fouquet C., Lozzi G. 2001. Etude de la représentativité du réseau piézométrique de la nappe d'Alsace. Contrat APRONA/ARMINES. Ecole des Mines de Paris, centre de Géostatistique.
- de Fouquet C. 2003. Etude sur la réalisation de cartographies de la qualité de l'air dans les zones peu/pas couvertes par les réseaux de stations fixes à l'aide de méthodes géostatistiques. Contrat INERIS/ARMINES n° CNS 0032784. Rapport final. Ecole des Mines de Paris, centre de Géostatistique.
- de Fouquet C., Prechtel A., Setier J.-C. 2004. Estimation de la teneur en hydrocarbures totaux du sol d'un ancien site pétrochimiques : étude méthodologique. *Oil&Gas Science and Technology- Rev. IFP*, 59(3) 275-295.
- de Fouquet C., Renard D., 2005. Etude relative à la caractérisation de la contamination des eaux souterraines par les pesticides. Convention ARMINES/IFEN. Rapport final. Ecole des Mines de Paris, centre de Géostatistique.
- de Fouquet C., Gallois D., Malherbe L., Cardenas G. 2005. Mapping annual nitrogene dioxide concentration above Mulhouse urban area. In *geostatistics for environmental applications*. Renard P. et al (eds), CD, Springer
- de Fouquet C., Bernard-Michel C. 2006. Modèles géostatistiques de concentrations ou de débits le long des cours d'eau. *Comptes-rendus Géosciences*, 338(5) 307-318
- Freulon X. 1992. Conditionnement du modèle gaussien par des inégalités ou des randomisées. Thèse de doctorat en géostatistique. Ecole des Mines de Paris.
- Freulon X., de Fouquet C. 1993. Conditioning a gaussian model with inequalities. *Geostatistics Troia '92*. A. Soares (ed.) 201-212
- Gallois D. 2004. Optimisation des estimations spatio-temporelles par des méthodes géostatistiques. Rapport de stage ASPA. Ecole des mines de Paris.
- Gallois D., de Fouquet C., Le Loc'h G., Malherbe L., Cardenas G. 2005. Mapping annual nitrogene dioxide concentrations in urban areas. In *Geostatistics Banff 2004*. Leuangthong O., Deutsch C.V. (eds) Springer
- Gneiting T. 2002. Nonseparable stationary covariance functions for space-time data. *J. american statistical association*, 97 560-600.
- Gneiting T., Genton M. G., Guttorp P. 2005. Geostatistical space-time models, stationarity, separability and full symmetry. Technical Report n°475. Department of Statistics. University of Washington.
- Gohin F. 1989. Analyse structurale de la température de surface de la mer. In *Geostatistics*, Armstrong M. (ed.), Kluwer.
- Grenier C. 1996. Modélisation stochastique conditionnelle de l'écoulement et du transport particulaire dans un milieu hétérogène bidimensionnel. Thèse de Doctorat, Université P&M Curie, Paris VI.
- Grzebyk M., Dubroeuq D. 1994. Quantitative analysis of distribution of soil types : existence of an evolutionary sequence in Amazonia. *Geoderma*, 62, 285-298.
- Geman S., Geman D. 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transaction s on pattern Analysis and machine Intelligence*, 6(6) 721-741.
- Guibal D., 1973. L'estimation des okoumés du Gabon - problèmes méthodologiques. Note du Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau.

- Huang D., Allen T.T., Notz W.I., Zheng N. 2006. Global optimization of stochastic black-box systems via sequential kriging meta-models. *J. Global optimization*, 34 (3)
- Host G., Follestad T. 1999. Spatial interpolation of ozone exposure in Norway from space-time data. In *geoENV II - Geostatistics for environmental applications*. Gomez-Hernandez, Soares and Froidevaux (eds) Kluwer.
- Hu L.Y. 2000. Gradual deformation and iterative calibration of Gaussian related stochastic models. *Mathematical Geology* 32(1), 87-108
- Hu L.Y., Le Ravalec-Dupin M. 2004. Elements for an integrated geostatistical modeling of heterogeneous reservoirs. *Oil&Gas science and technology-rev. IFP*, 59(2), 141-155.
- Hughson D.L., Gutjahr A. 1998. Effect of conditioning randomly heterogeneous transmissivity on temporal hydraulic head measurements in transient two-dimensional aquifer flow. *Stochastic Hydrology and hydraulics*. 12 (3) 155-170
- Jaquet O., Carniel R., Namar R., Di Cecca M. 2005. Forecasting volcanic eruptions using geostatistical methods. In *Geostatistics for environmental applications*, P. Renard, H. Demougeot-Renard, R. Froidevaux (eds). Springer.
- Jeannée N. 2001. Caractérisation géostatistique de pollutions industrielles de sols. Cas des HAP sur d'anciens sites de cokeries. Thèse de Doctorat en Géostatistique, Ecole des Mines de Paris.
- Jeannée N., de Fouquet C. 2000. Characterization of soil pollutions from former coal processing sites.- In : *Geostatistics 2000 Cape Town*, Kleingeld and D.G. Krige (eds) 526-537.
- Jeannée N., de Fouquet C. 2001. Which experimental variogram for the structural inference : a methodological study. In *geoENV III - Geostatistics for environmental applications*. P. Monestiez et al (eds.), 357-368.
- Jeannée N., de Fouquet C. 2003. Apport d'informations qualitatives pour l'estimation des teneurs en milieux hétérogènes : cas d'une pollution de sols par des hydrocarbures aromatiques polycycliques (HAP). *Comptes rendus Geoscience*. 335 (5), 441-449.
- Joathon P. 1971. Représentation d'une série sédimentaire par un modèle probabiliste. Application aux processus d'Ambarzoumian et étude de la structure de Chémery. Thèse de Docteur-Ingénieur, Université de Nancy.
- Jones D.R., Schonlau M., Welch W. J. 1998. Efficient global optimization of expensive black-box functions. *J. Global optimization* 13(4) 455-492
- Journal A. 1974. Simulations conditionnelles : théorie et pratique. Thèse de Docteur-Ingénieur, Université de Nancy.
- Kneith L.H., Telliard W.A. 1979. Priority pollutants I – a perspective view. *Environmental Science and technology*, 13 (4) 416-423
- Kondor R., Vert J.-P. 2004. Diffusion kernels. In *kernel methods in computational biology*. (Schoelkopf B., Tsuda K., Vert J.-P., Eds.) MIT Press, 171-192.
- L'Hégaret G. 2002. Mise en place d'une méthodologie pour la cartographie du benzène et du dioxyde d'azote à l'échelle des départements du Rhône et de l'Isère. Rapport de stage ASCOPARG-COPARLY. Ecole des mines de Paris.
- Le Loc'h G., Beucher H., Galli A., Doligez B., Heresim Group. 1994. Improvement in the truncated gaussian method : combining several Gaussian functions. In *Proc. ECMOR IV, 4th European Conference on the Mathematics of Oil Recovery, Røros, Norway*.
- Lefohn A.S., Knudsen H.P., McEvoy L.R., 1988. The use of kriging to estimate monthly ozone exposure parameters for the Southern United States. *Environmental Pollution*, 53, 27-42.
- Léonard J., Crouzet P. 1999. Construction d'un réseau représentatif. Contribution au réseau « Eurowaternet/Qualité des cours d'eau » de l'Agence Européenne de l'environnement. IFEN, Note de méthode n. 13, Orléans.
- Liu L.-J. S., Rossini A.J., 1996. Use of kriging models to predict 12-hour mean ozone concentrations in metropolitan Toronto - A pilot study. *Environment International*, 6, 667-692.
- Lévy P. 1965. Processus stochastiques et mouvement brownien. Gauthiers-Villars. Paris.

- Marbeau J.-P., 1974. Pollution des huîtres du bassin du Morbihan. Note du Centre de Morphologie Mathématique, Ecole des Mines de Paris, Fontainebleau.
- Marbeau J.-P. 1976. Etat actuel et développements nouveaux pour l'aménagement en forêts tropicales. Thèse de Doctorat, Ecole des Mines de Paris.
- Mariotti P., de Fouquet C., Bruno R. 2000. Geostatistical characterization of an aquifer by joint simulation of a random function and its partial derivatives. In : *Geostatistics 2000 Cape Town*, W. Kleingeld and D.G. Krige (eds) 29-38.
- Marsily, G. de, J.P. Delhomme, F. Delay, et A. Buoro 1999. Regards sur 40 ans de problèmes inverses en hydrogéologie. *Comptes Rendus de l'Académie des Sciences de Paris, Sciences de la Terre et des Planètes*, t. 329, 73-87.
- Marsily, G. de, J.P. Delhomme, A. Coudrain-Ribstein, and A.M. LaVenue. 2000. Four decades of inverse problems in hydrogeology. In *Theory, Modeling, and Field Investigation in Hydrogeology: A Special Volume in Honor of Shlomo P. Neuman's 60th Birthday*, D. Zhang and C.L. Winter (eds.), Special Paper 348, Geological Society of America, Boulder, Colorado, 1-17.
- Marsily, G. de, F. Delay, J. Gonçalves, P. Renard, V. Teles, and S. Violette. 2005. Dealing with spatial heterogeneity. *Hydrogeology*, special issue "The Future of hydrogeology",
- Matheron G. 1965. Les variables régionalisées et leur estimation. Masson et C^{ie}, Paris.
- Matheron G. 1970. La théorie des variables régionalisées, et ses applications. *Les Cahiers du CMM, fasc. 5*, Ecole des Mines de Paris.
- Matheron G. 1971. La théorie des Fonctions Aléatoires intrinsèques généralisées. Note du Centre de Géostatistique, Ecole des Mines de Paris.
- Matheron G. 1973. The intrinsic random functions and their applications. *Advances in Applied probability*, 5, 439-468.
- Matheron G. 1973. Le krigeage disjonctif. Note du Centre de Géostatistique, Ecole des Mines de Paris.
- Matheron G. 1974 Les fonction de transfert des petits panneaux. Note du Centre de Géostatistique, Ecole des Mines de Paris.
- Matheron G. 1976. Forecasting block grade distribution : the transfert functions. In *Advanced geostatistics in the mining industry*. M. Guarascio et al. (eds) Reidel publishing company.
- Matheron G. 1978. Estimer et choisir. Les cahiers du centre de morphologie mathématique de Fontainebleau. Ecole des Mines de Paris.
- Matheron G. 1978. L'estimation globale des réserves récupérables. Note du Centre de Géostatistique, Ecole des Mines de Paris.
- Matheron G. 1978. Le krigeage disjonctif et le paramétrage local des réserves. Note du Centre de Géostatistique, Ecole des Mines de Paris.
- Matheron G. 1984. Modèle isofactoriel et changement de support. *Sciences de la terre*, n°18, Série Informatique géologique.
- Matheron G. 1984. The selectivity of the distributions and the "second principle of geostatistics". *Proc. 2nd NATO ASI, "geostatistics for natural resources characterization"*. G. Verly et al. eds. Dordrecht, Holland. Reidel Publ.
- Matheron G. 1985. Comparaison de quelques distributions du point de vue de la sélectivité. *Sciences de la terre, Sér. Inf.*, 24
- Matheron G., Beucher H., de Fouquet C., Galli A., Guérillot D. et Ravenne C. 1987. Conditional simulation of the geometry of fluvio-deltaic reservoirs. *SPE paper 16753, 62nd Annual Technical Conference and Exhibition, Dallas, Texas*.
- Matheron G., Beucher H., de Fouquet C., Galli A. et Ravenne C. 1988. Simulation conditionnelle à trois faciès dans une falaise de la formation du Brent. *Sciences de la Terre, Sér. Inf.*, 28, 213-249.
- Monestiez P., Allard D., Navarro Sanchez I., Courault D. 1999. Kriging with categorical external drift : use of thematic maps in spatial prediction and application to local climate interpolation for agriculture. In *geoENV II - Geostatistics for environmental applications*. Gomez-Hernandez H. et al. (eds) Kluwer.

- Monestiez P., Bailly J.-S., Lagacherie P., Voltz M. 2005. Geostatistical modelling of spatial processes on directed trees: Application to fluvisol extent. *Geoderma*.128(3-4) 179-191
- Obled Ch., Creutin J.-D. 1986. Some developments in the use of Empirical orthogonal Functions for mapping meteorological fields. *Journal of Applied meteorology*, 15(9), 1189-1204.
- Obled Ch., Braud I. 1989. Analogies entre géostatistique et analyse en composantes principales de processus ou analyse EOFs. In *geostatistics*, M. Armstrong (ed.). 177-188. Kluwer Academic Publishers.
- Okx J., Leenaers H., Krzanowski R.M. 1993. probability kriging as a decision support tool for local soil pollution problems. *Geostatistics Troia '92*. A. Soares (ed.). 673-684
- Oosterbaan J. 2000. Utilisation des méthodes d'analyse exploratoire de données environnementales pour l'établissement de descripteurs macroscopiques de la dynamique des termes sources polluants dans la géosphère.- Application à la caractérisation de la pollution par des hydrocarbures aromatiques polycycliques (HAP) de sites industriels anciens ayant utilisé la pyrolyse de la houille (usines à gaz et cokeries). Thèse de doctorat, Ecole des Mines de Paris.
- Orfeuill J.-P. 1977. Etude, mise en œuvre et test d'un modèle de prédiction à court terme de pollution atmosphérique. Note du Centre de Géostatistique, Ecole des Mines de Paris.
- Pasquier P., Marcotte D. 2005. Solving the groundwater inverse problem by successive flux estimation. *Geostatistics for Environmental Applications*. Renard, Demougeot-Renard and Froidevaux (eds) Springer.
- Pellicciotta L. 2001. Simulazione geostatistica di una funzione aleatoria definita da una distribuzione gaussiana troncata mediante tecniche iterative. Tesi di Laurea, Università degli Studi di Bologna, facoltà di Ingegneria.
- Pereira M., Almeida J., Rodrigues A., Moura D. 1999. Mapping the spatial distribution of waders on south Tagus riverside. In *geoENV II. Geostatistics for Environmental Applications*. Gomez-Hernandez, Soares and Froidevaux (eds) Kluwer.
- Perrin O., Monestiez 1999. P. Modelling of non-stationary spatial structure using parametric radial basis deformation. In *geoENV II. Geostatistics for Environmental Applications*. Gomez-Hernandez, Soares and Froidevaux (eds) Kluwer.
- Petitgas P. 1991. Contributions géostatistiques à la biologie des pêches maritimes. Thèse de Doctorat en Géostatistique, Ecole des Mines de Paris.
- Poulin M., Even S., Mouchel J.-M., Billen G., Garnier J., Levassor A., Leviandier T., de Marsily G. 2001. Models of a river system : the PIREN Seine program. In *Advances in environmental and ecological modelling*. Blasco F., Weill A. (eds) Elsevier 17-45
- Prechtel A. 1998. La pollution d'un site par des hydrocarbures. Aspects de la modélisation hydrogéologique et étude géostatistique. Mémoire de DEA Hydrologie, hydrogéologie, géostatistique et géochimie des eaux.
- Raspa G., Tucci M., Bruno R. 1997. How cokriging and kriging with external drift can be used to reconstruct rainfall maps by combining raingauges data with radar files. In *Geostatistics Wollongong '96*. Baafi-Schofield (Ed). Kluwer Academic Publishers.
- Ravenne C., Eschard R., Galli A., Mathieu Y., Montadert L., Rudkiewicz J.-L. 1987. Heterogeneities and geometry of sedimentary bodies in a fluvio-deltaic reservoir. *SPE paper 16752*, 62nd *Annual Technical Conference and Exhibition, Dallas, Texas*.
- Reed J., Chappel A., French J.R., Oliver M.A. 2001. Geostatistical analysis of PCB-contaminated sediment in a commercial dock, Swansea, UK. In *geoENV III – geostatistics for Environmental Applications*, P. Monestiez et al. (eds) 487-497. Kluwer Academic Publishers.
- Rivoirard J. 1991. Introduction au krigeage disjonctif et à la géostatistique non-linéaire. Document de cours. Centre de géostatistique, Ecole des Mines de Paris.
- Rivoirard J. 2000. Weighted variograms. In *Geostatistics 2000 Cape Town*. W. Kleingeld and D.G. Krige (eds) 145-155

- Rolin N. 2005. Estimation du champ de transmissivité de l'aquifère Culebra Dolomite du site de stockage américain du Wipp par la méthode de déformation graduelle. Rapport de stage CEA. Ecole des Mines de Paris.
- Roth Ch. 1995. Contribution de la géostatistique à la résolution du problème inverse en hydrogéologie. Thèse de Doctorat. Ecole des Mines de Paris. *Documents du BRGM 241*. Ed. BRGM.
- Sampson P.D., Damian D., Guttorp P. 2001. Advances in modelling and inference for environmental processes with nonstationary spatial covariance. In *GeoENVIII – geostatistics for Environmental Applications*, P. Monestiez et al. (eds), 17-32. Kluwer Academic Publishers.
- Saporta G. 1990. Probabilités, analyse des données et statistique. Editions Technip, Paris.
- Sauquet E., Gorrshalk L., Leblais E. 2000. Mapping average annual runoff : a hierarchical approach applying a stochastic interpolation scheme. *Hydrological Science Journal*, 45(6) 799-816.
- Sénégas J., Wackernagel H., Rosenthal W., Wolf T. 2001. Error covariance modelling in sequential data assimilation. – *Stochastic environmental research and risk assessment*. 1, 65-86.
- Sénégas J. 2002. Méthodes de Monte-Carlo en vision stéréoscopique. Application à l'étude de modèles numériques de terrain. Thèse de doctorat en géostatistique. Ecole des Mines de Paris.
- Serra J. 1982. Image analysis and mathematical morphology. Academic press.
- Smits J., Everbecq E., Delière J.-F., Descy J.-P., Wollast R., Vanderborght J.-P. 1997. Pégase, une méthodologie et un outil de simulation prévisionnelle pour la gestion de la qualité des eaux de surface. *Tribune de l'eau*, 558(4) 73-82
- Switzer P., Green A.A. 1984. Min/max autocorrelation factors for multivariate spatial imaging. Technical report n° 6., Department of Statistics, Stanford University, Etats-Unis.
- Tartakovsky D., Neuman S. 1998. Transient flow in bounded randomly heterogeneous domains. *Water resources research*, 34(1)1-12,13-20.
- Thiry M., Huet-Taillanter S., Maurel B., Raulo A., Forette N., van Oort F. 1997. Diagnostic de la pollution de la friche industrielle de Mortagne du Nord. I Travaux de reconnaissance préliminaires. ENSMP, Fontainebleau .
- Thiry M., N. Forette, S. Huet-Taillanter, A. Raulo, J.-L. Haag. 2000. Etude de la distribution spatiale d'une pollution multiple dans une friche industrielle de cokerie et son périmètre. I – prospection du site n°1. Note LHM/RC/2000/1. Ecole des Mines de Paris, Fontainebleau
- Thiry M., Huet-Taillanter S., Schmitt J.-M. 2002. La friche industrielle de Mortagne du Nord (59)-I-prospection du site, composition des scories, hydrochimie, hydrologie et estimation flux. *Bull. Soc. Géol. France*, 173/4, p369-381.
- Ver Hoef J. M., Peterson E., Theobald D. 2007. Spatial statistical models that use flow and stream distance. *Environmental and Ecological Statistics*, 14. Sous presse.
- Wackernagel H., Webster R., Oliver M. 1988. A geostatistical method for segmenting multivariate sequences of soil data. – In H.H. Bock, ed. Elsevier.
- Webster R., Oliver M.A. 1989. Disjunctive kriging in agriculture. In : M. Armstrong (ed.), *Geostatistics*, Dordrecht : Kluwer, vol. 1.

CURRICULUM VITAE

Chantal de FOUQUET

née le 6 Juin 1957 à Poitiers (Vienne)

Adresse : Ecole des Mines de Paris, Centre de Géosciences
35, rue Saint-Honoré - F 77305 Fontainebleau Cedex
Tél. 33 (0)1 64 69 47 61 - Télécopie 33 (0)1 64 69 47 05
Courriel : chantal.de_fouquet@ensmp.fr

FONCTIONS

- 1989 - Maître de Recherche en géostatistique à l'Ecole des Mines de Paris
- 1984 - Chargée de cours à l'Ecole des Mines de Saint-Etienne
- 2005 - Chargée de cours à l'Université Pierre-et-Marie Curie, master SDUEE

FORMATION SUPERIEURE

- 1985 Docteur-Ingénieur en Sciences et Techniques Minières - option Géostatistique
Sujet de thèse : L'estimation des réserves récupérées sur modèle géostatistique de gisements non homogènes.
- 1980 Ingénieur Civil des Mines - Ecole des Mines de Saint-Etienne.

PARCOURS PROFESSIONNEL

Chercheur en géostatistique à l'Ecole des Mines de Paris

- 1987-1989 Chargé de recherche
- 1982-1987 Attaché de recherche
- 1980-1982 Bourse de thèse

ACTIVITES DE RECHERCHE

Depuis 1992 Responsable de l'activité "environnement" au Centre de Géostatistique

Caractérisation et modélisation de pollutions dans les différents milieux (sols et sédiments, nappes et cours d'eau, air) : études et développements méthodologiques.

Modélisation géostatistique des écoulements.

Montage et conduite de projets, dans le cadre de programmes institutionnels ou d'études contractuelles.

Nombreux partenariats : Ministère en charge de l'environnement, Institut Français de l'Environnement (IFEN), INERIS, Agences de bassin, Associations Agréées de surveillance de la qualité de l'air, EDF, GDF, CEA, IFP, ADEME ...

- 1989 -2002 Applications de la géostatistique à des domaines variés : hydrographie, géotechnique, agronomie, contrôle de déchets, etc.
Montage de projets et études.
Partenariats : Service Hydrographique de la Marine, EDF, FUGRO France, CIRAD, ANDRA ...
- 1987-1989 Développement de la méthode des gaussiennes seuillées pour la modélisation des réservoirs hétérogènes.
Partenariat : Institut Français du Pétrole
- 1981-1986 Modélisation de gisements et développement d'algorithmes pour la simulation d'exploitation minière
Partenariats : Total Compagnie Minière, COGEMA, projets du ministère en charge de la recherche et de la technologie

PRINCIPAUX THEMES DE RECHERCHE

Modélisation de la variabilité spatiale et applications

Modélisation multivariable et non stationnaire, intégrant différents types de mesures (paramètres géotechniques) ou pour le calage de modèles numériques aux mesures (hydrographie). Exploitation des simulations géostatistiques pour évaluer les incertitudes.

Etude de la cohérence des mesures acquises par les sondeurs multifaisceaux (hydrographie), et mise au point de méthodes automatiques de détection de sondes aberrantes.

Etude exploratoire, modélisation et estimation dans des domaines d'application variés : agronomie, inventaire forestier, niveau piézométrique, etc...

Caractérisation de milieux anthropisés et quantification de pollutions

Analyse exploratoire de mesures de concentrations dans les différents milieux (sols, nappes et cours d'eau, air), et modélisation des relations avec l'environnement, caractérisé par des variables auxiliaires (occupation des sols, densité de bâti, etc.).

Utilisation d'informations auxiliaires, quantitatives ou qualitatives, et adaptation de l'échantillonnage, pour la cartographie du niveau des concentrations, en modèles multivariables. Estimation non linéaire et comparaison à des seuils de référence. Construction et étude d'indicateurs géostatistiques du niveau des pollutions.

Développements méthodologiques spécifiques : modèles de fonctions aléatoires définies sur des arbres, modélisation de l'histogramme en présence de corrélation spatiale, étude des propriétés du variogramme expérimental.

Méthodes de simulation

Etude, mise en œuvre et programmation de la méthode des gaussiennes seuillées pour la simulation de gisements pétroliers ou miniers hétérogènes. Application à la prise en compte d'informations catégorielles.

Simulation numérique et modélisation géostatistique de variables liées par des équations aux dérivées partielles. Application à la modélisation stochastique en hydrogéologie ainsi qu'à l'étude de l'influence de la variabilité spatiale sur le transfert réactif.

ANIMATION SCIENTIFIQUE

Direction ou tutorat de thèses

- 2003 Stéphanie Brégoïn. *Etude de la variabilité des argilites du callovo-oxfordien*.
Co-direction avec Bernard Beaudoin. Doctorat en « dynamique et ressources des bassins sédimentaires », Ecole des Mines de Paris.
Actuellement chargée d'étude en bureau d'étude.
- 2003 Anne Zanolin. *Irrigation de Précision en Petite-Beauce : mesures au champ et modélisation stochastique spatialisée du fonctionnement hydrique et agronomique d'une parcelle de maïs*.
Tutorat. Doctorat en Hydrologie, Hydrogéologie, Géostatistique et Géochimie des eaux, Université Pierre et Marie Curie, Paris VI.
Actuellement professeur de gestion et maîtrise de l'eau en BTS au lycée agricole de Vienne.
- 2002 Hélène Renard-Demougeot. *De la reconnaissance à la réhabilitation des sols pollués : estimations géostatistiques pour une optimisation multicritère*.
Tutorat. Doctorat en Sciences naturelles, Ecole Polytechnique Fédérale de Zürich.
Actuellement consultante senior auprès de FSS
- 2001 Nicolas Jeannée. *Caractérisation géostatistique de pollutions industrielles de sols. Cas des hydrocarbures aromatiques polycycliques sur d'anciens sites de cokeries*.
Direction. Doctorat en Géostatistique, Ecole des Mines de Paris.
Actuellement consultant-formateur à Géovariances.
- 1999 Nathalie Courtois. *Caractérisation de la dispersion en aquifère hétérogène par méthodes de traçages et modélisation stochastique*.
Tutorat. Doctorat en hydrologie et hydrogéologie quantitatives de l'Ecole nationale Supérieure des Mines de Paris.
Actuellement hydrogéologue au BRGM (Montpellier)
- 1995 Eric Antoinet. *Intégration de la variabilité spatiale dans l'estimation probabiliste du tassement et du dévers de grands ouvrages*.
Tutorat. Doctorat de l'Université de Bordeaux 1.
Actuellement ingénieur de projet à ANTEA.
- 1995 Etienne Castelier. *Estimation d'un champ de perméabilité à partir de mesures de charge hydraulique*.
Tutorat. Doctorat en Géostatistique, Ecole des Mines de Paris.
Actuellement Ingénieur au CEA
- 1995 Chris Roth. *Contribution de la géostatistique à la résolution du problème inverse en hydrogéologie*.
Co-direction avec Jean-Paul Chilès. Doctorat en Géostatistique, Ecole des Mines de Paris.
Actuellement responsable du pôle « études et développement » à l'Observatoire des Sciences et des techniques.
- 1992 Xavier Freulon. *Conditionnement du modèle gaussien par des inégalités ou des randomisées*.
Direction. Doctorat en Géostatistique, Ecole des Mines de Paris.
Actuellement ingénieur réservoir à Total.

Thèses en cours

- Caroline Bernard-Michel. *Construction d'indicateurs de la qualité des eaux superficielles*.
Direction. Doctorat en Géostatistique, Ecole des Mines de Paris. Soutenance en juillet 2006.
- Sylvain Mazuel. *Evaluation des incertitudes sur le développement d'une pollution en zone non saturée*.
Tutorat. Doctorat en Hydrologie, Hydrogéologie, Géostatistique et Géochimie des eaux, Université Pierre et Marie Curie, Paris VI. Soutenance prévue en 2006.
- Marco De Lucia. *Influence de la variabilité spatiale sur la modélisation hydrogéochimique*.
Cotutelle avec l'Université de Bologne. Soutenance prévue en 2007.

Jury de thèse

- 1997 Yann Gautier. *Estimation de paramètres géostatistiques à partir de tests de puits*.
Doctorat de l'Université Pierre et Marie Curie, Paris VI. Spécialité : Sciences de la Terre.
- 1997 Hui-Chao Zhu. *Geological controls to the radon distribution in the Ardennes Massif. Applications of statistical and geostatistical methods*. Doctorat, faculté polytechnique de Mons (Belgique).
- 1996 Christophe Grenier. *Modélisation stochastique conditionnelle de l'écoulement et du transport particulaire dans un milieu hétérogène bidimensionnel*. Doctorat en Hydrologie, Hydrogéologie, Géostatistique et Géochimie des eaux, Université Pierre et Marie Curie, Paris VI.
- 1991 Serge Séguret. *Géostatistique des phénomènes à tendance périodique (dans l'espace-temps)* Doctorat en Géostatistique, Ecole des Mines de Paris. Rapporteur.
- 1991 Claude Bacconnet. *Géostatistique et géotechnique. Application à la reconnaissance des sols*. Doctorat, Université Blaise Pascal, Clermont II. Spécialité Génie Civil.
- 1990 Chen Zhen Wu. *Simulation de l'exploitation souterraine de gisements complexes pour la planification à moyen terme*. Doctorat en techniques et économie de l'entreprise minière, Ecole des Mines de Paris.
- 1990 Jean-Luc Bordessoule. *Estimation géostatistique d'un gisement d'Uranium multicouche*.
Doctorat en Géostatistique, Ecole des Mines de Paris. Rapporteur.
- 1986 Rukmana Adhi. *Etude géostatistique des gisements d'étain alluvionnaires exploités par dragage*. Thèse de Docteur-Ingénieur en Sciences et Techniques Minières, option géostatistique. Ecole des Mines de Paris.

Organisation de séminaires

Journées de Géostatistique, séminaire biennal européen, à l'Ecole des Mines de Paris à Fontainebleau : 4 sessions de 1989 à 1995 avec publication, puis en 2003.

Relecture d'articles pour des revues ou des congrès

Mathematical Geology, Geoderma, Engineering Geology, Hydrological Sciences Journal

Comité scientifiques : congrès internationaux quadriennaux de géostatistique, congrès européens biennaux de géostatistique environnementale Geoenv, congrès de géotechnique ICASP 7.

Invitation à des conférences

- 2002 Problème inverse et changement d'échelle en hydrogéologie. Ecole CEA-INSTN *Estimation et simulation géostatistiques. Des perméabilités aux écoulements.*
- 2002 Géovariances, journées technique sur la pollution de l'air, Paris : *Quelques problèmes de géostatistique environnementale.*
- 2002 avec H. Demougeot-Renard. Première rencontre nationale de la recherche sur les sites et sols pollués : bilan et perspectives. Paris. *Estimation géostatistique des volumes de sols à dépolluer et quantification de la précision : application à un cas réel de pollution au plomb et validation des résultats.*
- 2001 Séminaire sur la cartographie en pollution de l'air. INERIS, ministère de l'Aménagement du territoire et de l'environnement. *La géostatistique au service de la représentation cartographique.*
- 2000 EGS, Nice. *Geostatistical approach of the inverse problem in hydrogeology :review of some recent developments.*
- 2000 Evaluation de la politique de préservation de la ressource en eau destinée à la consommation humaine : séminaire sur les systèmes d'information et de connaissances, Ecole de Santé Publique, Rennes : *Quelle est la représentativité des réseaux de suivi de la qualité des eaux de surface ?*
- 1999 Industrial land contamination by heavy metals, NICOLE Meeting, Bilbao. *Etude géostatistique de la friche industrielle de Mortagne du Nord.*
- 1999 Géovariances, Journées ISATIS, Fontainebleau. *Etude géostatistique d'une friche industrielle.*
- 1990 Politecnico, Turin. *Simulazione geostatistica di giacimenti eterogenei.*

Autres conférences, posters

Participation à de nombreux congrès et journées techniques (APCOM, SPE, AFS, EGS, AIH, Industrie minière...) avec ou sans publication : contribution orale ou poster.

Vulgarisation : "où va l'eau". 2003. Ecole des Mines de Paris, Fontainebleau. Poster.

ACTIVITES PEDAGOGIQUES

Enseignement de deuxième cycle

Enseignement de géostatistique aux écoles des Mines de Paris et de Saint-Etienne.

Contribution aux activités pédagogiques de l'Ecole des Mines de Paris : acte d'entreprendre, module d'intégration généraliste.

Suivi de stages d'option en troisième année de l'Ecole des Mines de Paris

- 2005 Anne Cori. *Utilisation de données environnementales pour la cartographie du NO₂ sur la région Rouennais ; représentativité spatiale des stations de mesure de la concentration moyenne en NO₂ & définition de zones homogènes vis-à-vis du dépassement de seuil pour la concentration en ozone.* Air Normand, Rouen.
- 2005 Nicolas Rolin. *Estimation du champ de transmissivité de l'aquifère Culebra Dolomite du site de stockage américain du Wipp par la méthode de déformation graduelle.* CEA

- 2004 David Gallois. *Optimisation des estimations spatio-temporelles par des méthodes géostatistiques*. ASPA, Strasbourg.
- 2002 Guillaume L'Hégaret. *Cartographie de l'ozone atmosphérique sur le département de l'Isère*. ASCOPARG, Grenoble.
- 2001 Laurent Bascouert. *Analyse géostatistique des données environnementales du site minier de l'Ecarpière*. COGEMA.
- 2000 Nicolas Ungari. *Etude géostatistique de paramètres géotechniques sur diagraphies*. ANDRA
- 1998 Xavier Emery. *Estimations de superficies agricoles par les méthodes géostatistiques (application à des données guadeloupéennes)*. CIRAD.
- 1998 Sébastien Favre. *Etude géostatistique d'une pollution organique et métallique d'une décharge*. Rhodia Eco Service - ATE
- 1990 A. Mallard. *Approche géostatistique de la stabilité des remblais argileux : l'exemple du barrage de Mirgenbach*. EDF.

Suivi de stages de deuxième année de l'École des Mines de Saint-Etienne

- 2004 Clément Chevalier. *Approche géostatistique du transfert de fluide en milieux poreux*.
- 2000 Jillali Nemmassi. *Comparaison de quelques méthodes de simulation de fonctions aléatoires gaussiennes*.

Suivi de travaux de fin d'études (Tési di Laurea) de l'Université de Bologne

- 2002 Alessandro Pozzi. *Ricostruzione di profile di resistenza di punta su cable-route sottomarine a partire da misure sismiche*. Etude des relations entre résistance de pointe et vitesse de sismique réfraction le long de routes de câbles sous-marins.
- 2001 Giovanni Lozzi. *Ottimizzazione della rete di monitoraggio della falda freatica dell'Alsazia : l'approccio geostatistico*. Etude géostatistique de la cote piézométrique de la nappe d'Alsace.
- 2001 Luca Pellicciotta. *Simulazione geostatistica di una funzione aleatoria definita da una distribuzione gaussiana troncata mediante tecniche iterative*. Comparaison de la convergence de différentes méthodes de simulation de gaussiennes seuillées.
- 1997 Pamela Mariotti. *Caratterizzazione di una georisorsa mediante co-simulazione geostatistica contestuale di una funzione aleatoria e delle sue derivate parziali : applicazione al corso di un acquifero*. Simulation géostatistique conjointe de la charge et de la transmissivité, et comparaison à un calcul par différences finies.

Autres formations

- DEA : enseignement de géostatistique linéaire et de simulations (MSOE, HHGG).
Suivi de stages :
- 2005 Anne Cori, HETRE. *Représentativité des stations de mesure de différents polluants atmosphériques en France. Application à l'étude du dioxyde d'azote à Rouen et à l'ozone en haute-Normandie*.
- 1998 Alexander Prechtel, HHGG. *La pollution d'un site par des hydrocarbures. Aspects de la modélisation hydrogéologique et étude géostatistique* (coencadrement, avec P. Goblet).

1997 Maarten Cuypers, Méthodes quantitatives et modélisation des bassins sédimentaires. *Simulation d'une Fonction Aléatoire et de son gradient.*

1997 Gregor LAAHA, HHGG. *Etude de la pollution des eaux souterraines. Application des méthodes de statistique et géostatistique multivariante* (coencadrement, avec P. Chauvet).

- 1983-2003 : Cours et suivi de stages (une douzaine) au CFSG, Cycle de Formation Spécialisée en Géostatistique.
- Interventions dans différentes formations post-diplôme : Cespromin, mastères spécialisés (ISIGE, Géotechnique et environnement)
- Depuis 2002 : Contribution aux “hivernales”, cours annuel de 3 semaines organisé par le Centre de Géostatistique. Géostatistique linéaire, multivariante, non linéaire, simulations.
- Contribution à de nombreuses formations, internes ou externes, en particulier les cours d'une semaine suivants :
 - 1983. ETS de Caminos à Barcelone
 - deux séries de cours européens COMETT de 1993 à 1995 sur les applications de la géostatistique aux études environnementales : Suisse (Institut Polytechnique de Lausanne), France (EMP), Allemagne (TU Freiberg) et Italie (Université La Sapienza, Rome). Cours donnés en allemand et en italien.
 - 2004 APAT (agenzia per la protezione del l'ambiente e per i servizi tecnici), Rome. *Geostatistica applicata alle problematiche ambientali.*

Organisation de cours

Ecoles d'été d'une semaine sur les applications de la géostatistique : pollution des sols (1999), écoulements souterrains (1996), géotechnique (1994), pollutions (1993).

INVITATIONS A L'ETRANGER

1998 Trois semaines comme professeur visiteur à l'Université Laval (Québec), département des sols et de génie agroalimentaire.

1990 Deux mois d'expertise à l'Université La Sapienza (Rome), département Arte Mineraria.

1985 Quatre mois d'expertise auprès de Mining Italiana (Rome), société de consultants miniers.

DIVERS

2003 Chevalier des palmes académiques

2002 - Correspondant national auprès de l'académie d'agriculture de France, section V : Physique et chimie des milieux et des êtres vivants

1981 - Représentante élue des personnels de recherche à l'Ecole des Mines de Paris : plusieurs mandats au comité de la recherche et au comité pédagogique.

mai 2006

LISTE DES PUBLICATIONS ET RAPPORTS

1. Revues ou actes de congrès avec comité de lecture sélectif

- Olives Baños J., Amouric M., **de Fouquet C.**, Baronnet A. 1983. Interlayering and interlayer slip in biotite as seen by HRTEM. *Am. Mineralogist*, 68, 754-758.
- Deraisme J. et **de Fouquet C.** 1984. Recent and future developments in «downstream» geostatistics. In: G. Verly et al. (eds.), *Geostatistics for natural resources characterization*, Dordrecht : Reidel, 979-999.
- Matheron G., Beucher H., **de Fouquet C.**, Galli A., Guérillot D. et Ravenne C. 1987. Conditional simulation of the geometry of fluvio-deltaic reservoirs. SPE 16753. In *Proc. SPE Annual Technical Conference and Exhibition, Dallas, Texas, Reservoir Engineering*, 591-599
- Matheron G., Beucher H., **de Fouquet C.**, Galli A. et Ravenne C. 1988. Simulation conditionnelle à trois faciès dans une falaise de la formation du Brent. *Sciences de la Terre, Sér. Inf.*, **28**, 213-249.
- de Fouquet C.**, Beucher H., Galli A. et Ravenne C. 1989. Conditional simulation of random sets : Application to an argillaceous sandstone reservoir. In M. Armstrong (ed.), *Geostatistics*, Dordrecht : Kluwer, 517-530.
- de Fouquet C.**, Ikonomi A., Beucher H., Groupe Heresim. 1991. Choix des lignes de corrélation stratigraphique pour les simulations conditionnelles de lithofaciès. *Sciences de la Terre, Sér. Inf.*, **30**, 87-118.
- de Fouquet C.** 1993. Géostatistique orientée vers le traitement des données territoriales : notions-clés et exemples. *Biométrie, Praximétrie*, 33, 113-146.
- Freulon X., **de Fouquet C.** 1993. Conditioning a Gaussian model with inequalities. In *Geostatistics Tróia '92*, Soares A. (ed.), Dordrecht : Kluwer, 201-212.
- de Fouquet C.**, Mandallaz D. 1993. Using geostatistics for forest inventory with air cover : an example. In *Geostatistics Tróia '92*, A. Soares (ed.), Dordrecht : Kluwer, 875-886.
- de Fouquet C.** 1994. Reminders on the conditioning kriging. In Armstrong M., Dowd P. (eds.), *Geostatistical Simulations*, Dordrecht : Kluwer, 131-145.
- Roth C., Chilès J. P., **de Fouquet C.** 1996. Adapting geostatistical transmissivity simulations to finite difference flow simulators. *Water resources research*, 32 (10), 3237-3242.
- Antoinet E., Fry J.-J., **de Fouquet C.**, Costaz J. L. 1996. Probabilistic estimation of settlement of nuclear islands using geostatistical simulations. In *Applications of statistics and probability, ICASP7 Paris*, Lemaire M. et al. (eds.), Balkema, 1205-1211.
- Castelier E., **de Fouquet C.**, Brunet C., Fry J. J. 1996. Estimation of the permeabilities of a dam. In *Applications of statistics and probability, ICASP7 Paris*, Lemaire M. et al. (eds.), Balkema 1303-1308.
- Roth C., **de Fouquet C.**, Chilès J. P., Matheron G. 1997. Geostatistics applied to hydrogeology's inverse problem: taking boundary conditions into account. In *Geostatistics Wollongong '96*, Baafi E.Y., Schofield N.A. (eds.), Dordrecht : Kluwer, 1085-1097.
- Bisquay H., Freulon X., **de Fouquet C.**, Lajaunie C. 1998. Multibeam data cleaning for hydrography using geostatistics. In *Oceans '98 Conference proceedings*, 1135-1143.
- Roth C., Chilès J.-P., **de Fouquet C.** 1998. Combining geostatistics and flow simulators to identify transmissivity. *Advances in water resources*, 21, 555-565.
- Decencièrre E., **de Fouquet C.**, Meyer F. 1998. Applications of kriging to image sequence coding. *Signal processing : image communication* 13, 227-249.
- * Primé "Eurasip Best-Paper 1998/99"

- de Fouquet C.** 2000. Geostatistical simulation of transmissivity and flow : simplified case of the linearized diffusion equation in steady state. In *Tracers and modelling in hydrogeology*. IAHS publication 262. Dassargues A. (ed.) 103-108
- Courtois N., Gerbaux-François O., Grenier C., Maugis P., Mouche E., **de Fouquet C.**, Goblet P., Ledoux E. 2000. Characterization of dispersion in an alluvial aquifer by tracing techniques and stochastic modelling. In F. Stauffer et al. (eds.), *Calibration and reliability in groundwater modelling*, 84-89.
- Jeannée N., **de Fouquet C.** 2000. Characterization of soil pollutions from former coal processing sites. In *Geostats 2000 Cape Town*, W. Kleingeld and D.G. Krige (eds), vol. 2, 526-537.
- Mariotti P., **de Fouquet C.**, Bruno R. 2000. Geostatistical characterization of an aquifer by joint simulation of a random function and its partial derivatives. In *Geostats 2000 Cape Town* W. Kleingeld and D.G. Krige (eds) vol. 1, 29-38.
- de Fouquet C.** 2000. Joint simulation of a random function and its derivatives. In *Geostats 2000 Cape Town*, W. Kleingeld and D.G. Krige (eds), vol. 1, 84-93
- Jeannée N., **de Fouquet C.** 2001. Which experimental variogram for the structural inference : a methodological study. In *geoENV III - Geostatistics for environmental applications*, P. Monestiez et al. (eds.), 357-368.
- Demougeot-Renard H., **de Fouquet C.** 2001. Managing heterogeneous sampling data for geostatistical estimations of Benzo(a)pyrene concentrations in a former gas works. In *geoENV III - Geostatistics for environmental applications*, P. Monestiez et al (eds.), 429-440.
- Arnaud M., Emery X., **de Fouquet C.**, Brouwers M., Fortier M. 2001. L'analyse krigeante pour le classement d'observations spatiales et multivariées. *Revue de statistique appliquée* XLIX (2), 45-67.
- Jeannée N., **de Fouquet C.** 2003. Apport d'informations qualitatives pour l'estimation des teneurs en milieux hétérogènes : cas d'une pollution de sols par des hydrocarbures aromatiques polycycliques (HAP). *Comptes rendus Geoscience*, 335 (5) 441-449.
- de Fouquet C.** 2003. Des statistiques contre la pollution. *La recherche*, 368, 92-93.
- Jeannée N., **de Fouquet C.** 2004. Estimating the grades of polluted industrial sites: use of categorical information and comparison with threshold values. In *GeoENV IV - Geostatistics for environmental applications*. Sanchez-Vila X. et al. (eds) Kluwer.
- Demougeot-Renard H., **de Fouquet C.**, Fritsch M. 2004. Optimizing sampling for acceptable accuracy levels on remediation volume and cost estimations. In *GeoENV IV - Geostatistics for environmental applications*. Sanchez-Vila X. et al. (eds) Kluwer.
- Demougeot-Renard H., **de Fouquet C.**, Renard P. 2004. Forecasting the number of soil samples required to reduce remediation cost uncertainty. *Journal of Environmental Quality*. 33 (3) 1694-1702.
- Demougeot-Renard H., **de Fouquet C.** 2004. Geostatistical approach for assessing soil volumes requiring remediation: validation using lead-polluted soils underlying a former smelting works. *Environmental science & technology*. 38 (19) 5120 – 5126.
- Bernard-Michel C., **de Fouquet C.** 2005. Estimating indicators of river quality by geostatistics. In *geostatistics for environmental applications*. Renard P. et al (eds), Springer 443-454.
- de Fouquet C.**, Gallois D., Malherbe L., Cardenas G. 2005. Mapping annual nitrogen dioxide concentration above Mulhouse urban area. In *geostatistics for environmental applications*. Renard P. et al (eds), CD, Springer
- Gallois D., **de Fouquet C.**, Le Loc'h G., Malherbe L., Cardenas G. 2005. Mapping annual nitrogen dioxide concentrations in urban areas. In *Geostatistics Banff 2004*. Leuangthong O., Deutsch C.V. (eds) Springer
- Bernard-Michel C., **de Fouquet C.** 2005. Geostatistical indicators of waterway quality for nutrients. - In *Geostatistics Banff 2004*. Leuangthong O., Deutsch C.V. (eds) Springer
- Bernard-Michel C., **de Fouquet C.** 2005. Geostatistical indicators of nutrients concentrations in streams. In *Proceedings IAMG2005 Toronto*. Cheng Q. & Bonham-Carter G. (eds), York University, Toronto Canada & GPMR, Wuhan, China. 716-721

de Fouquet C., Bernard-Michel C. 2006. Modèles géostatistiques de concentrations ou de débits le long des cours d'eau. *Comptes-rendus Géosciences*, 338 (5) 307-318

de Fouquet C., Goblet P., Jamet P., Ledoux E. 2006. Caractérisation de la pollution des sols de la friche de Thionville : typologie et estimation des termes sources. Soumis à *Hydrological science journal*

2. Autres articles

Deraisme J., **de Fouquet C.**, Fraisse H. 1983. Méthodologie de simulation d'exploitation souterraine dans les mines métalliques : application à un gisement d'uranium en Australie. *Revue de l'Industrie Minérale «Les Techniques»*, 513-523.

Deraisme J., **de Fouquet C.**, Fraisse H. 1984. Geostatistical orebody model for computer optimization of profits from different underground mining methods. In *Proc. 18th APCOM Symp.*, London. Institution of Mining and Metallurgy, 583-590.

Adhi R., **de Fouquet C.** 1987. Calcul des réserves d'un gisement d'étain alluvionnaire à bedrock karstique. *Revue de l'Industrie Minérale, Mines et Carrières «Les Techniques»*, 363-369.

Freulon X., **de Fouquet C.**, Rivoirard J. 1990. Simulation of the geometry and grades of a uranium deposit using a geological variable. In *Proc. XXII International Symposium APCOM, Berlin*, 649-659.

Freulon X., **de Fouquet C.** 1991. Remarques sur la pratique des bandes tournantes à 3D. In *Cahiers de géostatistique*, fasc. 1, ENSMP, 101-117.

Matheron G., Roth C., **de Fouquet C.** 1993. Modélisation et cokrigage de la charge et de la transmissivité avec conditions aux limites à distance finie. *Cahiers de Géostatistique*, fasc. 3, Ecole des Mines de Paris, 61-76.

Castelier E., **de Fouquet C.**, Brunet C. 1995. Estimation des perméabilités d'un barrage. In *Cahiers de Géostatistique*, fasc. 5, ENSMP, 15-22.

Deraisme J., **de Fouquet C.** 1996. The geostatistical approach for reserves estimation. *Mining magazine*, 309-313.

de Fouquet C. 1996. Some reminders about geostatistics. In *Applications of Statistics and Probability*, ICASP7 Paris, Lemaire M. et al. (eds.) Paris, vol. 3, Balkema, Rotterdam, 1537-1539.

de Fouquet C. 1996. Random processes, geostatistical methods : applications of geostatistics. (Synthesis of session 9c). In *Applications of statistics and probability*, ICASP 7, Paris, : Lemaire M. et al. (eds) vol. 3, Balkema, Rotterdam, 1477-1479.

de Fouquet C. 1997, Influence de la méthode d'estimation et de la maille de reconnaissance sur la quantification des pollutions : étude méthodologique à 2D. In *Echantillonnage et environnement*, Cebedoc, Liège, Nicolas J. (ed.), 39-63.

de Fouquet C. 1998. Simulation conjointe d'une fonction aléatoire et de ses dérivées partielles. Application à l'équation de diffusivité linéarisée en régime permanent. *Cahiers de Géostatistique*, fasc. 6, ENSMP, 45-59.

de Fouquet C. Lajaunie C. , Bisquay H., Freulon X. 1998. Méthodologie de détection des sondes aberrantes dans les données bathymétriques multifaisceaux. *Cahiers de Géostatistique*, fasc. 6, ENSMP, 123-136.

Bisquay H., Freulon X., **de Fouquet C.** 1999. Lajaunie C. Multibeam data cleaning using geostatistics. *Sea technology*, 43-51.

de Fouquet C. 1999, Représentativité spatiale des mesures de pollution dans les sols : des échantillons aux estimations. *Les techniques de l'industrie minière*, no. 3, 3^e trim. 24-32, 145-155.

Demougeot-Renard H., **de Fouquet C.** 2001. La géostatistique au service de l'environnement. Première partie. *Environnement et technique*, 208, 40-44

Demougeot-Renard H., **de Fouquet C.** 2001. La géostatistique au service de l'environnement. Deuxième partie. *Environnement et technique*, 209, 46-50.

- Demougeot-Renard H., de Fouquet C. 2002. Estimation géostatistique des volumes de sol à dépolluer et quantification de la précision : application à un cas réel de pollution au plomb et validation des résultats. Communication In *Première Rencontre de la Recherche sur les sites et sols pollués, Maison de la Chimie, Paris, 12-13 décembre 2002*. 6 p.
- de Fouquet C.**, Prechtel A., Setier J. C. 2004. Estimation de la teneur en hydrocarbures totaux du sol d'un ancien site pétrochimiques : étude méthodologique. *Oil&Gas Science and Technology- Rev. IFP*, 59(3) 275-295
- de Fouquet C.**, Renard D. 2005. Caractérisation de la contamination des eaux souterraines par les pesticides : l'exemple de la nappe de la Craie du Nord. In *L'eau souterraine et les produits phytosanitaires, Colloque CFAIH*. Orléans. 61-68. ISBN 2- 7159-0975-6
- Bernard-Michel C. **de Fouquet C.**. 2006. Vers des indicateurs géostatistiques pour les comptes de la qualité de l'eau. Soumis aux Journées de la Société Française de statistiques, Chatou
- Renard D., **de Fouquet C.** 2006. Characterization of the pesticide concentrations in the "Nappe de la craie du Nord". Poster. Gestion des grands aquifères. AIH, Dijon
- De Lucia M., Lagneau V., **de Fouquet C.** 2006. Effects of spatial heterogeneity and upscaling methods on hydrodynamic transport coupled with geochemical reactions. Groundwater hydraulics in complex environments. IAHR, Toulouse.
- Mazuel S., **de Fouquet C.**, Chilès J.-P., Goblet P., Krimissa M. 2006. geostatistical modelling for the quantification of uncertainties on the unsaturated zone and the groundwater transfer. Groundwater hydraulics in complex environments. IAHR, Toulouse.

3. Publications, contribution à des ouvrages

- de Fouquet C.** 1985. L'estimation des réserves récupérées sur modèle géostatistique de gisements non homogènes. Thèse de Docteur-Ingénieur en Sciences et Techniques Minières, Option Géostatistique, ENSMP. 150 p.
- de Chambure L., **de Fouquet C.**, Fraisse H. 1987. Calculating ore reserves subject to mining constraints, for a uranium deposit. In *geostatistical cases studies*. Matheron G., Armstrong M. (eds). Reidel, 208-246
- de Fouquet C.** (ed.). 1991. Compte-rendu des Journées de géostatistique, 6-7 Juin 1991, Fontainebleau. *Cahiers de géostatistique 1*. Paris : ENSMP, 261 p. ISSN 1168- 2574
- de Fouquet C.** (ed.). 1993. Compte-rendu des Journées de géostatistique, 25-26 Mai 1993, Fontainebleau. *Cahiers de géostatistique 3*. Paris : ENSMP, 205 p. ISSN 1168- 2574
- de Fouquet C.** (ed.). 1995. Compte rendu des journées de géostatistique, 15-16 Juin 1995, Fontainebleau. *Cahiers de Géostatistique 5*. - Paris : ENSMP, 248 p. ISSN 1168-2574
- de Fouquet C.** 1996. Mise en oeuvre de la géostatistique pour l'évaluation des réserves. In *Memento des mines et carrières : aide-mémoire de l'exploitant*, 38e éd., (2) Paris : Société de l'Industrie Minérale, 74-107.
- Chilès J.-P., **de Fouquet C.** 2005. Comment passer des analyses chimiques à la cartographie des teneurs en polluants. In *Fiches techniques Geosipol*.
- Bruno R., **de Fouquet C.** 2005. The expansion of environmental geostatistics. In *Space, structure and randomness. Contribution in honour of Georges Matheron in the fields of geostatistics, random sets, and mathematical morphology*. Bilodeau M., Meyer F., Schmitt M. (eds) Springer 105-133

4. Notes et rapports d'études non publiés

Les thèmes détaillés dans le mémoire sont indiqués de la façon suivante : (A) pollution atmosphérique, (G) méthode des gaussiennes seuillées, (H) hydrogéologie stochastique, (N) nappes, (R) pollution des cours d'eau, (S) pollution des sols et des sédiments.

Les autres thèmes sont notés comme : (L) agronomie, (M) simulation de gisements miniers ou d'exploitation minière, (O) hydrographie, (T) géotechnique, (Z) autres

- (M) Deraisme J., **de Fouquet C.** 1981. Etude géostatistique de la carrière de La Vernhe à Ranteil : estimation des réserves et simulation de l'exploitation. 54 p. + annexes. [Confidentiel](#)
- (Z) **de Fouquet C.**, Rivoirard J. 1981. Etude structurale géostatistique de nodules polymétalliques. Etude CNEOX. 100 p. [Confidentiel](#).
- (M) Deraisme J., Dumay R., **de Fouquet C.** 1982. Ben Lomond : construction du modèle numérique de gisement. Rapport d'avancement. 75 p. [Confidentiel](#).
- (M) Deraisme J., Fraisse H., **de Fouquet C.** Développement industriel d'une méthodologie de simulation d'exploitation : cas d'une mine d'uranium (gisement en stockwerk exploité en souterrain). Action concertée DDSTI no. 82 DO 116). Novembre 1982. 16 p. [Confidentiel](#).
- (M) Deraisme J., **de Fouquet C.** Développement industriel de la méthodologie d'étude d'un projet minier sur modèle géostatistique de gisement. Rapport de fin d'étude DDSTI. Janvier 1984. 100 p.
- (M) **de Fouquet C.** L'apport de la géostatistique aval au contrôle des réserves. Communication in Colloque *Valorisation des ressources du sous-sol*, DDSTI, Paris, 27 nov. 1984. Février 1985.
- (M) Tessié J., Deliac E., Cheimanoff N., **de Fouquet C.** Perspectives de l'utilisation de l'outil informatique pour le choix et la conduite des méthodes d'exploitation minière souterraine? Rapport de fin d'étude DDSTI, TCM-ARMINES no. 83 E 1078. Mars 1985. 76 p. + annexes.
- (M) **de Fouquet C.** 1985. Etude géostatistique de la carrière de Sultana (3^{ème} partie). 13 p. [Confidentiel](#)
- (M) **de Fouquet C.**, Coléou T. 1985. Etude géostatistique de la carrière de Beaucaire. 143 p. [Confidentiel](#)
- (M) **de Fouquet C.** 1986. Modélisation géostatistique du corps B de la mine de Bertholène. (Annexe 2). 51 p. [Confidentiel](#)
- (G) Beucher H., **de Fouquet C.** 1987. Discrétisation de l'image géologique de la falaise "Long Nab". 17 p. [Confidentiel](#)
- (M) Tessié, J., Deliac, E., **de Fouquet C.** Fraisse, H., Noubibou P., Cheimanoff N. 1988. Evaluation par simulation de la sélectivité des méthodes d'exploitation souterraine. Compte-rendu de fin d'étude DDSTI, action concertée VRSS, décision d'aide 85 E 0676. 36 p. + 2 annexes.
- (M) **de Fouquet C.**, Freulon X., Rivoirard J. 1988. Analyse structurale de Laporte (Nord) (uranium). Rapport d'avancement CEE, décision d'aide M-A1M-0022-C(CD). 29 p.
- (M) Cheimanoff N., Fraisse H., Teissié J., **de Fouquet C.** 1988. Sélectivité et infrastructures en mines souterraines. Présentation des actions VRSS 85 E 676 et 83 E 1078 au Colloque-bilan, Paris. 16 p.
- (G) Beucher H., **de Fouquet C.** 1988. Analyse structurale d'un réservoir hétérogène. Rapport d'avancement : étude de la falaise de Cloughton. 53 p. [Confidentiel](#)
- (G) Beucher H., **de Fouquet C.**, Galli A. 1988. Structural analysis of Sandstone Level N-2, Wara Unit of the Khafji field. (Annexe du rapport: *Analyse structurale du niveau gréseux N-2, Unité Wara du champ Khafji*). 109 p. [Confidentiel](#)
- (G) **de Fouquet C.** 1988. Description du programme SIMENS - Simulation conditionnelle d'ensembles aléatoires. 78 p. [Confidentiel](#)
- (G) **de Fouquet C.**, Beucher H. 1988. Calculs et ajustements de variogrammes d'indicatrices: modèles de gaussiennes seuillées. [Confidentiel](#).

- (M) Freulon X., **de Fouquet C.**, Rivoirard J. 1988. Analyse structurale du gisement d'uranium de Laporte (Sud). Rapport d'avancement. Décision d'aide M-A1M-0022-CD). 73 p. [Confidentiel](#).
- (G) **de Fouquet C.**, Beucher H., Galli A. et Mattieu, Y. 1989. Geostatistical modelling of macroscopic reservoir heterogeneities: the model parameter : a detailed structural analysis. 79 p. [Confidentiel](#).
- (G) Teissié, J., Rosenkranz, R., Freulon X., **de Fouquet C.**, Rivoirard J. 1990. Simulations de morphologies de gisements tectonisés. Rapport d'avancement n°4 du contrat CEE MA1M-0022-C. 54 p. [Confidentiel](#).
- (G) Teissié, J., Rosenkranz, R., Freulon X., **de Fouquet C.**, Rivoirard J. 1990. Simulations de morphologies de gisements tectonisés. Rapport final du contrat CEE MA1M-0022-C. 70 p. [Confidentiel](#).
- (G) Teissié, J., Rosenkranz, R., Freulon X., **de Fouquet C.**, Rivoirard J. 1990. Simulations de morphologies de gisements tectonisés. Rapport synthétique du contrat CEE MA1M-0022-C. 14 p.
- (N) **de Fouquet C.** 1990. Approche géostatistique de la teneur en métaux lourds d'une nappe fossile. 22 p. + 85 p. figures. [Confidentiel](#).
- (O) Chauvet P., **de Fouquet C.** 1991. Etude des coefficients de marée de la baie de Seine - Rapport d'avancement. 42 p. [Confidentiel](#).
- (O) **de Fouquet C.**, Chauvet P. 1991. Etude du coefficient de marée Q22 de la Baie de Seine - Krigeage avec dérive externe. 74 p. [Confidentiel](#).
- (T) **de Fouquet C.** 1992. Etude préliminaire des mesures pressiométriques du site de Nogent. 18 p. + Fig. [Confidentiel](#).
- (O) Castelier E., **de Fouquet C.** 1992. Etude des coefficients de concordance par espèce pour la baie de Seine. Rapport d'étude, contrat EPSHOM/CMO/ASF n° 2982, 03/12/91. 98 p. + annexes. [Confidentiel](#).
- (S) **de Fouquet C.** 1992. Remarques sur le rapport : Etude géostatistique du site de l'ancienne usine à gaz de Geislingen par la Société TAUW Infra Consult - B.V. 6 p. [Confidentiel](#).
- (G) **de Fouquet C.** 1992. Modélisation de réservoirs hétérogènes par seuillages de fonctions aléatoires multigaussiennes. Communication présentée à : *Statistics Earth and Space Sciences*, Leuven, Belgium, August 22-26, 1989. 26 p.
- (H) **de Fouquet C.** 1993. Modélisation des perturbations de la charge et de la transmissivité avec conditions aux limites pour un champ rectangulaire : cas d'une recharge. 1er rapport d'avancement. Contrat BRGM SGN/ISA. Suivi de thèse de Ch. Roth. 13 p. [Confidentiel](#).
- (H) **de Fouquet C.** 1993. Notions élémentaires sur le cokrigeage de la charge et de la transmissivité. 12 p.
- (T) **de Fouquet C.** 1993. Simulation non conditionnelle du BR1 de NOGENT. Modèle simplifié du module pressiométrique. 18 p. [Confidentiel](#).
- (H) **de Fouquet C.** 1993. Modélisation de la charge et de la transmissivité avec conditions aux limites à distance finie. 2ème rapport d'avancement. Contrat BRGM SGN/ISA. 9 p. [Confidentiel](#).
- (T) Guiblin P. et **de Fouquet C.** 1993. Etude variographique des mesures pressiométriques - champs du BR1 et du BR2. 56 p. + annexes. [Confidentiel](#).
- (Z) **de Fouquet C.** 1994. Contribution de la géostatistique au traitement de l'information environnementale et géographique. 7 p.
- (T) **de Fouquet C.** 1994. Modèles variographiques pour la simulation conditionnelle du BR1 et du BR2. 1er rapport d'avancement, Contrat EDF 2H 1082. [Confidentiel](#).
- (H) **de Fouquet C.** 1995. Applications de la géostatistique en hydrogéologie. Eléments bibliographiques. Rapport de contrat EDF/ARMINES ZH 1082. 24 p.
- (T) Antoinet E., Fry J. J., **de Fouquet C.** Costaz J.-L. 1995. Estimation probabiliste du tassement des îlots nucléaires par simulation géostatistique (avec traduction en anglais). 16 p.
- (H) Castelier E., **de Fouquet C.**, Brunet C., Fry J. J. 1995. Estimation des perméabilités d'un barrage (+ traduction en anglais). 12 p.
- (H) **de Fouquet C.** 1995. Modélisation de la charge et de la transmissivité avec conditions aux limites à distance finie. 3ème rapport d'avancement. Contrat BRGM SGN/ISA.

- (Z) **de Fouquet C.** 1995. Estimation globale du tonnage de boues du bassin de Resplandy (COMURHEX-Narbonne). 13 p. [Confidentiel](#).
- (L) Koudriavsteva T., **de Fouquet C.** 1995. Etude géostatistique multivariable de Carreau Blémur. 119 p.
- (O) Lajaunie C., **de Fouquet C.** 1995. Etude géostatistique des données du sondeur SIMRAD SM12 en vue de la détection automatique des sondes aberrantes. Rapport d'étude, contrat no. 3663 EPSHOM/EPASF. 136 p., annexe. [Confidentiel](#).
- (O) **de Fouquet C.** 1996. Construction de simulations géostatistiques pour l'étude de sensibilité à la réponse géotechnique de la maquette de Hualien. 21 p. [Confidentiel](#).
- (Z) **de Fouquet C.** 1996. Remarques sur l'étude variographique de Cs137 post Tchernobyl. Etude INTAS 94-2361, contribution au 1er rapport intermédiaire. 27 p.
- (O) **de Fouquet C.** 1996. Validation des mesures de sondeur SIMRAD EM12, modèle variographique de la bathymétrie mesurée. Première partie : méthodologie de détection des sondes aberrantes. Bon de commande no. 3861 EPSHOM/EP/ASF, premier rapport d'avancement. 42 p. [Confidentiel](#).
- (O) **de Fouquet C.** 1996. Validation des mesures du sondeur SIMRAD EM12, modèle variographique de la bathymétrie mesurée. Deuxième partie : étude variographique aux petites distances. Bon de commande no. 3861 EPSHOM/EP/ASF, premier rapport d'avancement (2ème partie). 42 p. [Confidentiel](#).
- (Z) **de Fouquet C.** 1995. Quelques rappels sur les méthodes géostatistiques. 5 p.
- (O) **de Fouquet C.** 1997. Méthodes géostatistiques automatiques de contrôle des données bathymétriques acquises par un sondeur multifaisceaux grands fonds : évaluation sur données réelles (Première partie). Bon de commande no. 3861 EPSHOM/EP/ASF, rapport final (1ère partie). 88 p. [Confidentiel](#).
- (O) **de Fouquet C.**, Lajaunie C. 1997. Méthodes géostatistiques automatiques de contrôle des données bathymétriques acquises par un sondeur multifaisceaux grands fonds : évaluation sur des données réelles. Bon de commande no. 3861 EPSHOM/EP/ASF, rapport de synthèse. 33 p. [Confidentiel](#).
- (H) **de Fouquet C.** 1997. Calcul de la covariance de la charge et de la transmissivité en régime permanent ; questions résolues et limitations. (Avec traduction en anglais : Determination of the covariance function of the head and the transmissivity ; resolved issues and limitations). Communication in *Stochastic hydrology seminar, Université Paris VI*. 21 p.
- (N) **de Fouquet C.** 1997. Etude préliminaire de mesures de la concentration en Atrazine. 55 p. [Confidentiel](#).
- (H) **de Fouquet C.** 1997. Suivi de l'étude géostatistique de la nappe alluviale au confluent du Drac et de l'Isère (Site M27). Rapport intermédiaire 1ère partie, contrat CEA-ARMINES n°7020.7B031360. 11 p. [Confidentiel](#).
- (L) Arnaud M., **de Fouquet C.**, Pichot J. P., Koudriavsteva T. 1997. Une méthodologie de classification à partir d'observations spatiales et multivariées. Rapport technique CIRAD/CG. 21 p.
- (H) **de Fouquet C.**, Courtois N. 1997. Suivi de l'étude géostatistique de la nappe alluviale au confluent du Drac et de l'Isère (Site M27, CEA-Grenoble). Rapport intermédiaire 1ère partie, contrat CEA-ARMINES n° 7020.7B031360. 81 p. [Confidentiel](#).
- (Z) **de Fouquet C.** 1998. Etude pour une méthodologie d'échantillonnage de déchets TFA. Première partie : analyse exploratoire de données expérimentales. Commande ANDRA 006018PFA. 33 p. [Confidentiel](#).
- (S) Emery X., **de Fouquet C.**, Jeannée N. 1998. Etude géostatistique de la friche industrielle de Mortagne-du-Nord. Première partie : analyses exploratoire et variographique des échantillons de sol. 1er rapport d'avancement, A.O. sur les friches industrielles du Ministère de l'Environnement, INRA/ME- DGAD/SRAE. 59 p.
- (S) Emery X., **de Fouquet C.** 1998. Etude géostatistique de la friche industrielle de Mortagne-du-Nord. Deuxième partie : étude des échantillons de sol, estimation des concentrations, probabilités de dépassement de seuils. 2ème rapport d'avancement, A.O. sur les friches industrielles du Ministère de l'Environnement, INRA/ME- DGAD/SRAE. 39 p.
- (O) Manuel L., **de Fouquet C.** 1998. Amélioration de la représentation des fonds marins à partir de mesures très denses acquises par sondeurs multifaisceaux. Rapport de synthèse. 38 p. [Confidentiel](#).

- (Z) **de Fouquet C.** 1998. Etude pour une méthodologie d'échantillonnage de déchets TFA. Rapport final. Commande ANDRA 006018PFA. 27 p. **Confidentiel.**
- (S) Favre S., **de Fouquet C.** 1998. Etude géostatistique de la friche industrielle de Mortagne-du-Nord. Troisième partie : étude des échantillons de sol, simulations conditionnelles. 3ème rapport d'avancement, A.O. sur les friches industrielles du Ministère de l'Environnement, INRA/ME-DGAD/SRAE. 24 p.
- (S) Favre S., **de Fouquet C.** 1999. Etude géostatistique de la friche industrielle de Mortagne-du-Nord. Quatrième partie : étude des échantillons d'eau, analyse exploratoire. 4ème rapport d'avancement, A.O. sur les friches industrielles du Ministère de l'Environnement, INRA/ME-DGAD/SRAE. 35 p.
- (S) **de Fouquet C.** 1999. Etude préliminaire des mesures 3D de la concentration en HAP d'un ancien site industriel. 14 p. **Confidentiel.**
- (S) **de Fouquet C.** 1999. Caractérisation spatiale des pollutions. 13 p. Contrat TAUW/ARMINES, no. 3630951.
- (S) **de Fouquet C.** 1999. Etude géostatistique de la friche industrielle de Mortagne-du-Nord : estimation des volumes et des concentrations des polluants d'une friche industrielle. Rapport final, A.O. sur les friches industrielles du Ministère de l'Environnement, INRA/ME-DGAD/SRAE. 89 p.
- (Z) **de Fouquet C.** 1999. Etude exploratoire de collectes sélectives d'emballages. Commande CREED no. F00184. **Confidentiel.**
- (S) Demougeot-Renard H., **de Fouquet C.** 1999. Etude géostatistique des concentrations en BAP et HAP des sols d'une ancienne usine à gaz. Rapport de synthèse, Convention no. 610 ARMINES/GDF *Développement de techniques d'estimation de tonnage de sol pollué et de quantification des incertitudes. Application au site de l'ancienne usine à gaz de Gennevilliers.* 129 p. **Confidentiel.**
- (S) Balabdaoui F., **de Fouquet C.**, Oosterbaan J. 1999. Outils d'analyse de données et de simulation des transferts pour une reconnaissance rationnelle d'un site industriel pollué. (2) Etude géostatistique du site T. Contrat Ministère de l'aménagement du territoire et de l'environnement (DGAD/SRAE)/ARMINES *Etude géostatistique et simulation de transferts en vue d'une quantification des risques induits.* 110 p. **Confidentiel.**
- (S) Okx J., Schlebaum W., Hanneman F., Oliver M.A., **de Fouquet C.** 2000. Good Survey practice. Balancing between effort and risk. R3630951/WSB/C01. Tauw b.v., Deventer (Pays-Bas)
- (T) Ungari N., **de Fouquet C.** 2000. Etude géostatistique des diagraphies du site de reconnaissance : prétraitement du sondage HTM102. Commande ANDRA n° 013059ANM. 28 p. **Confidentiel.**
- (R) **de Fouquet C.** 2000. Construction d'un réseau représentatif de qualité des cours d'eau. Phase I : analyse exploratoire des données. Contrat BETURE CEREC/ARMINES Construction d'un réseau représentatif EUROWATERNET de qualité des cours d'eau. 69 p.
- (T) Sadki N., **de Fouquet C.** 2000. Pré-traitement des données provenant du pénétromètre statique : troncature des artefacts du signal pénétrométrique. Contrat EDF/SEPTEN/ARMINES ND 2965-GC, Suivi technique d'une étude géostatistique de données géotechniques. Rapport intermédiaire. 52 p. **Confidentiel.**
- (T) Sadki N., **de Fouquet C.** 2000. Cartographie du toit des formations géotechniques. Contrat EDF/SEPTEN/ARMINES ND 2965-GC, Suivi technique d'une étude géostatistique de données géotechniques. Rapport final. 67 p. **Confidentiel.**
- (S) Demougeot-Renard H., **de Fouquet C.** 2000. Etude géostatistique des concentrations en plomb des sols d'une ancienne fonderie. Convention no. 610 ARMINES/GDF *Développement de techniques d'estimation de tonnage de sol pollué et de quantification des incertitudes. Application au site de l'ancienne usine à gaz de Gennevilliers.* 30 p. **Confidentiel.**
- (R) **de Fouquet C.**, Bez N. 2001. Construction d'un réseau représentatif de qualité des cours d'eau. Rapport final. Contrat BETURE CEREC/ARMINES. 153 p.
- (S) BALABDAOUI., F., Cardenas G., **de Fouquet C.**, GOBLET, P., JAMET, P., LEDOUX, E., OOSTERBAAN, J., VIENNOT, P. 2001. Outils d'analyse de données et de simulation des transferts pour une reconnaissance rationnelle d'un site industriel pollué : topologie des pollutions et des transferts. Rapport final. Contrat Ministère de l'Aménagement du Territoire et de l'Environnement, DGAD/SRAE-ARMINES, *Etude géostatistique et simulation des transferts en vue d'une quantification des risques induits.* 141 p. **Confidentiel.**

- (S) Cardenas G., **de Fouquet C.** 2001. Outils d'analyse de données et de simulation des transferts pour une reconnaissance rationnelle d'un site industriel pollué : étude géostatistique du site H. 4^{ème} rapport d'avancement. Contrat Ministère de l'aménagement du territoire et de l'environnement (DGAD/SRAE)/ ARMINES *Etude géostatistique et simulation de transferts en vue d'une quantification des risques induits*. 74 p. [Confidentiel](#).
- (O) El Goulli O., **de Fouquet C.** 2001. Cartographie géostatistique de la bathymétrie sur l'extension du port du Havre. Contrat FUGRO France-ARMINES, *Projet Espadon : corrélation, intégration et estimation spatiale des données de reconnaissance d'origine diverse*. Rapport d'avancement n°1. 27 p. [Confidentiel](#).
- (L) Cardenas G., **de Fouquet C.** 2001. Etude géostatistique des dégâts occasionnés par la tempête de décembre 1999 à la forêt d'Armainvilliers (Seine-et-Marne). 63 p.
- (S) Demougeot-Renard H., **de Fouquet C.** 2001. Etude géostatistique complémentaire des concentrations en plomb des sols d'une ancienne fonderie : calcul et validation des volumes. 36 p. [Confidentiel](#).
- (T) El Goulli O., **de Fouquet C.** 2001. Modélisation de la résistance de pointe. Contrat FUGRO France-ARMINES, *Projet Espadon : corrélation, intégration et estimation spatiale des données de reconnaissance d'origine diverse*. Rapport d'avancement no. 2. 19 p. [Confidentiel](#).
- (N) **de Fouquet C.**, Lozzi G. 2001. Etude de la représentativité du réseau piezométrique de la nappe d'Alsace. Contrat APRONA/ARMINES. 97 p.
- (T) Schmitt M., Cardenas G., **de Fouquet C.** 2001. Segmentation de données de sismique réfraction par traitement d'images. Projet Coïmbra FUGRO/ARMINES, commande N° 0064-SO-439 et 440.18 p.
- (T) Pozzi A., **de Fouquet C.** 2002. Recherche de corrélation entre vitesse sismique et résistance de pointe : étude sur de petits tronçons de routes. Contrat FUGRO France-ARMINES, *Projet Espadon : corrélation, intégration et estimation spatiale des données de reconnaissance d'origine diverse*. Rapport final. 79 p. [Confidentiel](#).
- (S) Cardenas G., Le Loc'h G., **de Fouquet C.** 2002. Etude géostatistique des volumes et teneurs en polluants de sédiments de retenues. Contrat EDF/ARMINES no. P76L69/F00311. 52 p. [Confidentiel](#).
- (R) Bernard-Michel C., **de Fouquet C.** 2002. Remarques sur les calculs statistiques pour l'évaluation de la qualité de l'eau. Convention Ministère de l'Ecologie et du Développement Durable (MEDD)/ARMINES no. CV02000187, 1er rapport d'avancement. 53 p.
- (A) **de Fouquet C.** 2003. Méthodologie de cartographie de la concentration annuelle en NO₂ sur l'agglomération de Mulhouse. Contrat INERIS/ARMINES no. CNS 0032784 *Etude sur la réalisation de cartographies de la qualité de l'air dans les zones peu/pas couvertes par les réseaux de stations fixes à l'aide de méthodes géostatistiques*. Rapport d'avancement n° 3. 41 p.
- (A) **de Fouquet C.** 2003. Etude sur la réalisation de cartographies de la qualité de l'air dans les zones peu/pas couvertes par les réseaux de stations fixes à l'aide de méthodes géostatistiques. Contrat INERIS/ARMINES no. CNS 0032784. Rapport final. 34 p.
- (R) Bernard-Michel C., **de Fouquet C.** 2003. Calculs statistiques et géostatistiques pour l'évaluation de la qualité de l'eau. Convention Ministère de l'Ecologie et du Développement Durable (MEDD)/ARMINES no. CV02000187, 2ème rapport d'avancement. 163 p.
- (N) **de Fouquet C.**, Renard D. 2003. Etude relative à la caractérisation de la contamination des eaux souterraines par les pesticides. Convention ARMINES/IFEN. 1^{ère} partie, 1^{er} rapport d'avancement. 87 p.
- (N) **de Fouquet C.**, Renard D. 2004. Etude relative à la caractérisation de la contamination des eaux souterraines par les pesticides. Convention ARMINES/IFEN. 1^{ère} partie, 2^{ème} rapport d'avancement. 98p.
- (N) **de Fouquet C.**, Renard D. 2004. Etude relative à la caractérisation de la contamination des eaux souterraines par les pesticides. Convention ARMINES/IFEN. 1^{ère} partie, rapport final. 57 p.
- (R) Bernard-Michel C., **de Fouquet C.** 2004. Calculs statistiques et géostatistiques pour l'évaluation de la qualité de l'eau : estimation du quantile 90. Convention Ministère de l'Ecologie et du Développement Durable (MEDD)/ARMINES no. CV02000187, 3ème rapport d'avancement. 138 p.

- (A) **de Fouquet C.** 2005. Construction d'indicateurs caractérisant l'évolution temporelle des concentrations. Travaux d'ARMINES-EMP dans le cadre du groupe de recherche « surveillance des pollutions atmosphériques en proximité des voies de circulation », groupe 7 PREDIT 3. Rapport intermédiaire. 36p.
- (N) **de Fouquet C., Renard D.,** 2005. Etude relative à la caractérisation de la contamination des eaux souterraines par les pesticides. Convention ARMINES/IFEN. 2^{ème} partie, 1^{er} rapport d'avancement
- (N) **Renard D., de Fouquet C.,** 2005. Etude relative à la caractérisation de la contamination des eaux souterraines par les pesticides. Convention ARMINES/IFEN. 2^{ème} partie, 2^{ème} rapport d'avancement
- (N) **de Fouquet C., Renard D.,** 2005. Etude relative à la caractérisation de la contamination des eaux souterraines par les pesticides. Convention ARMINES/IFEN. Rapport final.
- (N) **Renard D., de Fouquet C.,** 2005. Etude relative à la caractérisation de la contamination des eaux souterraines par les pesticides. Convention ARMINES/IFEN. Annexe, nappe d'Alsace.
- (R) **Bernard-Michel C., de Fouquet C., Oudin L.-C.** 2005. Calculs géostatistiques d'indicateurs des concentrations dans les cours d'eau. Convention Ministère de l'Ecologie et du Développement Durable (MEDD)/ARMINES no. CV02000187. Rapport final, 1^{ère} partie.

6. Documents de cours

- de Fouquet C.** 1993. Simulation conditionnelle de fonctions aléatoires : cas gaussien stationnaire et schéma linéaire. Document de cours. Ecole des Mines de Paris, Centre de géostatistique. 91 p.
- de Fouquet C.** 1994. Introduction à la géostatistique. Cours EDF-CNEH, Chambéry 13 Avril 1994. (Recueil de transparents).