



# LA GÉOSTATISTIQUE

## DESCRIPTION THÉORIQUE DU PRINCIPE DE LA TECHNIQUE

La géostatistique regroupe un ensemble de méthodes mathématiques qui permet, à partir de données géoréférencées (localisation X, Y et Z), de comprendre les phénomènes et de les représenter spatialement. L'analyse des données à partir de laquelle sont modélisés les phénomènes permet notamment leur interpolation, c'est-à-dire la prédiction d'une variable en un point non mesuré. Elle permet également de fournir un niveau d'incertitude de l'interpolation au niveau local.

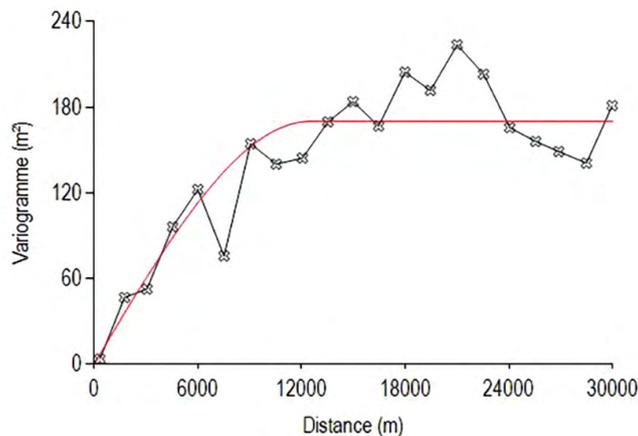
La géostatistique étudie des variables physiques régionalisées c'est-à-dire qui varient selon la position dans l'espace et dans le temps : c'est le cas d'une cote piézométrique, d'une concentration, d'un flux, ou de la profondeur et l'épaisseur d'une couche géologique. L'étude géostatistique permet de modéliser théoriquement les variations de la variable dans l'espace et éventuellement d'exploiter les relations entre cette variable et d'autres variables disponibles pour améliorer la qualité des prédictions.

Le variogramme, le krigeage et les simulations conditionnelles sont les outils et méthodes les plus courants. **Le variogramme permet de quantifier et d'analyser la continuité spatiale d'un phénomène, c'est à dire de**

**mesurer le degré de ressemblance d'une variable en fonction de la distance et de la direction.** L'ajustement du variogramme expérimental aboutit à un modèle théorique.

Le krigeage est un estimateur linéaire exact, non biaisé, basé sur le variogramme. Il tient compte de la distance entre les données et les points cibles (grille), d'éventuelles anisotropies et permet d'utiliser des variables auxiliaires corrélées avec le phénomène étudié. Cette technique minimise la variance d'estimation et assure la meilleure estimation globale d'une variable. Le krigeage est utilisé pour estimer une quantité moyenne locale d'un polluant dans un milieu.

Des **simulations conditionnelles** permettent d'obtenir des cartes équiprobables en conservant la même distribution statistique que le phénomène étudié ou de gérer des valeurs censurées pour prendre en compte les valeurs inférieures aux limites de détection par exemple. **Les simulations géostatistiques sont fréquemment utilisées pour calculer les risques de dépassement de seuil et pour les cartographier** (par exemple, 90 % des simulations obtenues dépassent le seuil choisi).



© B. Bourguin, BRGM.

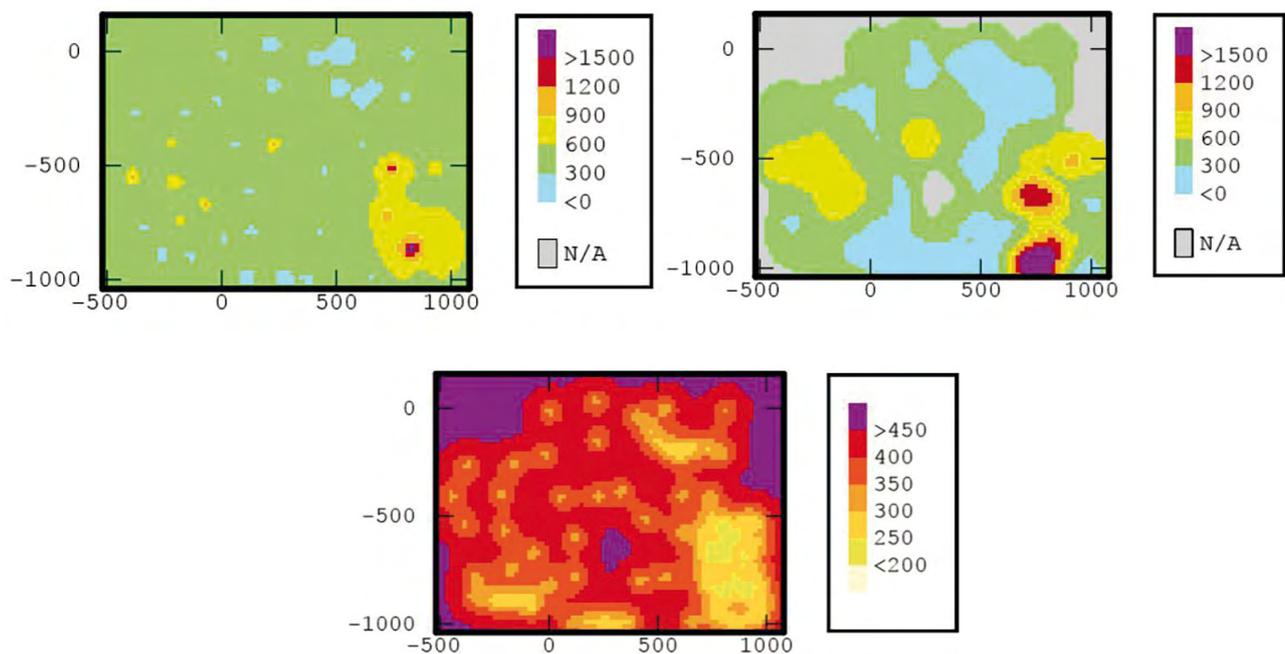
*Le variogramme théorique (en rouge) modélise la variation des teneurs mesurées (en noir) avec la distance*

Dans le cadre de l'étude des sites et sols pollués, un modèle géostatistique peut être utilisé dès la phase d'investigations et actualisé jusqu'à la phase de travaux. L'utilisation du modèle estimé est différente selon l'étape de gestion du site. La géostatistique peut ainsi permettre d'identifier les zones présentant des incertitudes (zones blanches ou lacunes de points) et nécessitant des investigations complémentaires. Elle permet l'optimisation du réseau de points de mesure pour délimiter une pollution sur la base d'éléments ou d'hypothèses factuelles. La géostatistique peut enfin permettre de mieux

estimer la quantité de polluant dans un milieu ou d'améliorer l'estimation du volume de terres susceptibles de dépasser une concentration seuil.

À l'issue d'une étude géostatistique, le niveau de connaissance du site est estimé à l'aide de la représentation graphique continue 2D ou 3D de la variable choisie et de l'indice de confiance associé à cette représentation graphique. La restitution visuelle peut correspondre au krigeage ou à des cartes de probabilité de dépassement de seuil. Le krigeage est toujours accompagné de sa carte d'incertitude.

# LA GÉOSTATISTIQUE



Cartographie des teneurs sans géostatistique, avec géostatistique et incertitude associée à la carte géostatistique – Source [1]

## CONTEXTE D'UTILISATION

La plus-value des géostatistiques réside dans les études de sites à enjeux et/ou les sites pour lesquels les jeux de données sont suffisamment importants. En revanche, elles ne sont pas nécessaires dans les situations d'actions simples et pas adaptées dans le cas de sites remblayés, profondément remaniés ou à géologie complexe. Il est possible d'utiliser la géostatistique uniquement pour l'analyse des données pour la définition de volumes de terres à traiter ou pour l'aménagement d'un site en lien avec une évaluation des risques sanitaires. L'ampleur de l'étude géostatistique doit dépendre de l'enjeu associé à l'étude et de la pollution.

Par ailleurs, le lien avec le métier des SSP est primordial afin de pouvoir comprendre les phénomènes modélisés et que la situation représentée corresponde le plus possible à la réalité du terrain. Lorsque les phénomènes modélisés sont issus d'une activité anthropique, l'analyse variographique peut être délicate.

**La réussite de cette méthode dépend de la qualité des données d'entrée. Les données doivent être en nombre suffisant au droit du site (avec un minimum de 30 valeurs) et de qualité homogène.** Des données sont de qualité homogène si par exemple les échantillons ont été collectés sur un même horizon, acquis avec la même méthodologie si les coupes lithologiques de sondages ont été relevées par un même opérateur, ou si les analyses proviennent d'un même laboratoire. Il est important d'examiner l'erreur de l'estimation qui est donnée par la variance de krigeage. La cartographie de cette erreur permet d'identifier les zones les moins informées par les données et pour lesquelles l'acquisition de données complémentaires peut être judicieuse. Attention cependant, la vraie incertitude a plusieurs sources : l'incertitude du krigeage liée à la modélisation théorique du variogramme et celles des erreurs issues de l'incertitude des données (représentativité de l'échantillon, erreur de mesures, erreur d'analyse).

**La géostatistique est un outil d'aide à la décision et de justification dans l'implantation de nouvelles investigations. Elle permet d'apporter des éléments techniques avec une représentation compréhensible par un public non-sachant.**

### À quelle étape ?

La géostatistique est essentiellement utilisée lors du **diagnostic**. Lors de cette étape, la géostatistique permet d'orienter la stratégie d'échantillonnage en produisant des cartes d'incertitudes. Elle permet une meilleure compréhension de la pollution et l'identification des sources ainsi que des voies de migration.

À la fin des phases d'investigations, les modélisations permettent une représentation graphique des variables d'études. Ces cartes donnent une information sur l'état des milieux et servent également de base pour l'élaboration de scénarios de gestion (bilan massique, volumes à traiter, masses de polluants mises en jeu...). L'estimation de l'incertitude associée aux masses permet le dimensionnement des travaux.

Lors des étapes de **dépollution et de suivi**, la géostatistique est utilisée à des fins de représentations graphiques d'une variable à un instant t.

Lors des étapes de réhabilitation des friches, la géostatistique, couplée à l'évaluation des risques sanitaires permet une optimisation des plans de réaménagement.

Enfin, à chaque étape de gestion, la géostatistique est un outil de communication avec les différents acteurs du projet (maître d'ouvrage, riverains, administration).

La géostatistique est également un des outils utilisés pour déterminer les valeurs de fonds pédogéochimiques anthropisés ou naturels à l'échelle d'un territoire.

## POLLUANTS CONCERNÉS

La géostatistique est applicable quel que soit le type de polluant. Néanmoins, la fiabilité des résultats dépend de la qualité des données d'entrée. L'analyse spatiale est

particulièrement adaptée à la modélisation d'un phénomène de diffusion ou d'interaction spatiale. Il est également possible d'utiliser les corrélations entre substances ou phénomènes.

## MATÉRIEL NÉCESSAIRE

Un ordinateur équipé d'un logiciel adapté aux méthodes géostatistiques est nécessaire pour la réalisation des traitements. Plusieurs logiciels sont disponibles et se

différencient selon les cas d'étude plus ou moins complexes. L'utilisation des logiciels de géostatistique nécessite néanmoins une expertise poussée.

## MÉTHODOLOGIE

Une étude géostatistique commence par l'analyse exploratoire des données qui consiste notamment à valider les données d'entrées, vérifier leur homogénéité, identifier des données aberrantes et rechercher les corrélations judicieuses pour utiliser le maximum de données représentatives. Des processus de traitement des données initiales peuvent être réalisés pour garantir leur homogénéité (régularisation pour palier des variations de taille du support d'analyse, declustering pour prendre en compte les effets de données groupées) ou pour satisfaire les hypothèses de calcul (transformation log-normale par exemple dans le cas d'une distribution non normale du jeu de données initiales).

À l'issue de la phase d'analyse exploratoire des données, la variographie permet de modéliser par un ou plusieurs variogrammes théoriques les variations spatiales des données.

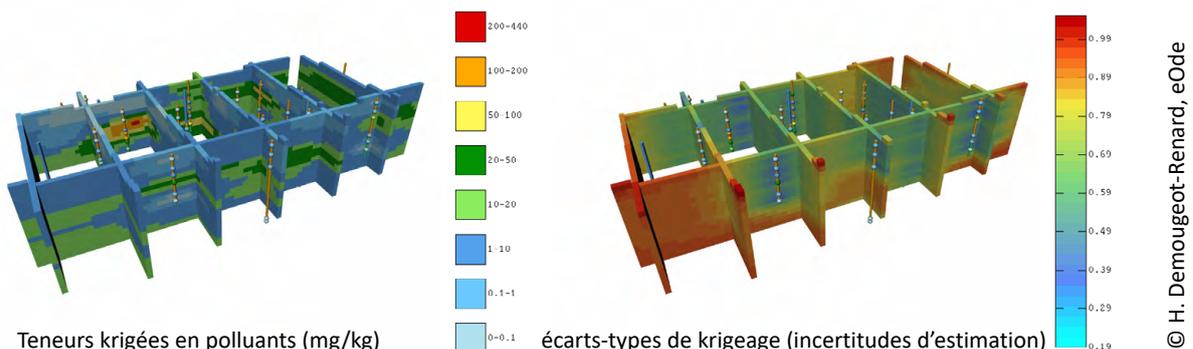
Le modèle de variogramme obtenu permet ensuite de réaliser une interpolation en chaque point de l'espace et de la représenter sur une cartographie. Selon les besoins, il est possible de réaliser un krigeage qui donne une estimation de la variable ou bien de générer un grand nombre de simulations respectant les données du phénomène étudié et sur lesquelles un traitement statistique permet de calculer des probabilités. Les probabilités de dépassement de seuil donnent également le risque qui est associé au choix d'une concentration seuil

(prise en compte d'une terre non polluée comme polluée ou inversement).

Les conclusions de l'analyse dépendent de l'objectif et du stade de gestion du site : elle peut conduire à une stratégie d'échantillonnage complémentaire, ou à un volume de terres impactées ou à traiter.

### Points de vigilance :

- Une étude géostatistique doit être réalisée en concertation avec des spécialistes en SSP.
- Les cartes diffusées à l'issue d'une étude géostatistique sont simples et efficaces pour la transmission d'un message mais il est essentiel que la méthode de calcul soit comprise par les acteurs manipulant ces cartes afin de les interpréter correctement. Les légendes utilisées (intervalles de valeurs, couleurs,...) doivent être réfléchies pour qu'elles n'orientent pas voire ne faussent pas la représentation des résultats.
- Les changements de support inhérents à la méthode d'échantillonnage (usage de plusieurs outils) ou au passage du support de point (échantillon) aux blocs de terre (unité de surface ou de volume de dépollution) sont des questions à première vue annexes mais délicates tant leur effet dans l'estimation de la quantité de terres polluées est significatif.



Résultat de traitement : krigeage 3D de teneurs en polluant et incertitude associée.



## AVANTAGES – INCONVÉNIENTS – MATURITÉ DE LA TECHNIQUE

### AVANTAGES

#### Polluants

- Utilisable avec l'ensemble des polluants,
- Exploitation de variables auxiliaires potentiellement corrélées.

#### Matériel

- Ne nécessite qu'un logiciel de géostatistique appliqué aux sites pollués,

#### Influence sur les résultats et incertitudes

- Quantification des incertitudes associées à la modélisation géostatistique,

#### Résultats d'interprétation

- Outil d'aide à la décision,
- Appui à la communication.

### INCONVÉNIENTS

#### Données

- Nécessite un nombre minimal de données géoréférencées (X, Y, Z),
- Importance de données de qualité et homogènes,
- Temps de qualification des données et de mise en forme potentiellement long.

#### Interprétation

- Nécessite une très bonne expertise du sujet malgré l'apparence de cartes simples.

### MATURITÉ DE LA TECHNIQUE



R&D aboutie, indicateurs développés, technique utilisée sur le terrain

## DÉLAIS DE MISE EN ŒUVRE

Les délais associés à une étude géostatistique dépendent de la quantité et de la qualité des données disponibles, de la connaissance du phénomène et du type de gestion possible ou attendue. La durée d'une étude géostatistique peut être d'une journée si les données sont de bonne qualité, informatisées, géoréférencées et s'il y a une bonne connaissance préalable

des phénomènes. Une étude plus poussée intégrant un travail préparatoire sur les données, une réflexion sur les phénomènes étudiés et une modélisation multicritères sur un site complexe peut durer plusieurs semaines voire plusieurs mois.

### PHASE

Délai associé

☒ : jour / ☒☒ : semaine / ☒☒☒ : mois

### TRAITEMENT



## ÉLÉMENTS DE COÛTS

Les coûts de traitement sont associés à la durée de la prestation et peuvent varier de plusieurs centaines d'euros pour une étude simple à plusieurs dizaines de milliers d'euros dans

le cas d'une étude très complexe. Le coût annuel de la licence des logiciels est de plusieurs milliers d'euro, il dépend des fonctionnalités du logiciel et du besoin en support technique.

### PHASE

Coût associé

€ < 100 € / €€ < 1000 € / €€€ > 1000 €

### TRAITEMENT



## POUR EN SAVOIR PLUS – RÉFÉRENCES BIBLIOGRAPHIQUES

[1] *Géostatistique appliquée aux sites et sols pollués- Manuel méthodologique et exemples d'applications – GeoSiPol – Novembre 2005*

[2] *Journée technique sur la géostatistique appliquée aux sites pollués – REX et perspectives ADEME - 23 janvier 2019*

[3] *Fiches techniques – GeoSiPol: Fiche n°1: Comment passer des analyses chimiques à la cartographie des teneurs en polluants?, Fiche n°2: Comment délimiter et quantifier les zones à dépolluer?, Fiche n°3: Comment intégrer toute l'information, de l'historique aux analyses de site?, Fiche n°4: Quel est l'apport de la Géostatistique à l'échantillonnage des sites pollués?*

[4] *Guide de détermination des valeurs de fonds dans les sols – Échelle d'un site – ADEME novembre 2018, pages 87 - 98*

[5] *La modélisation géostatistique des milieux anthropisés - HDR Chantal de Fouquet, Université Pierre et Marie Curie - Paris VI, 2006*

[6] *Caractérisation géostatistique de pollutions industrielles de sols: cas des hydrocarbures aromatiques polycycliques sur d'anciens sites de cokeries - Thèse Nicolas Jeannée, Mines de Paris – mai 2001*