# GML6402A : Géostatistique

## Cours 6 : Simulations géostatistiques

POLYTECHNIQUE
MONTRÉAL

UNIVERSITÉ
D'INGÉNIERIE

# Objectifs

- Identifier les problèmes types où les simulations s'appliquent;
- Expliquer les différences entre estimation et simulation;
- Expliquer les différences entre simulations non conditionnelles et conditionnelles;
- Être en mesure d'appliquer les méthodes de simulations de Choleski et SGS. Discuter de leurs avantages et inconvénients;
- Utiliser et interpréter les résultats d'une simulation pour l'estimation des ressources;
- Expliquer les principales propriétés des simulations.

# Plan du cours

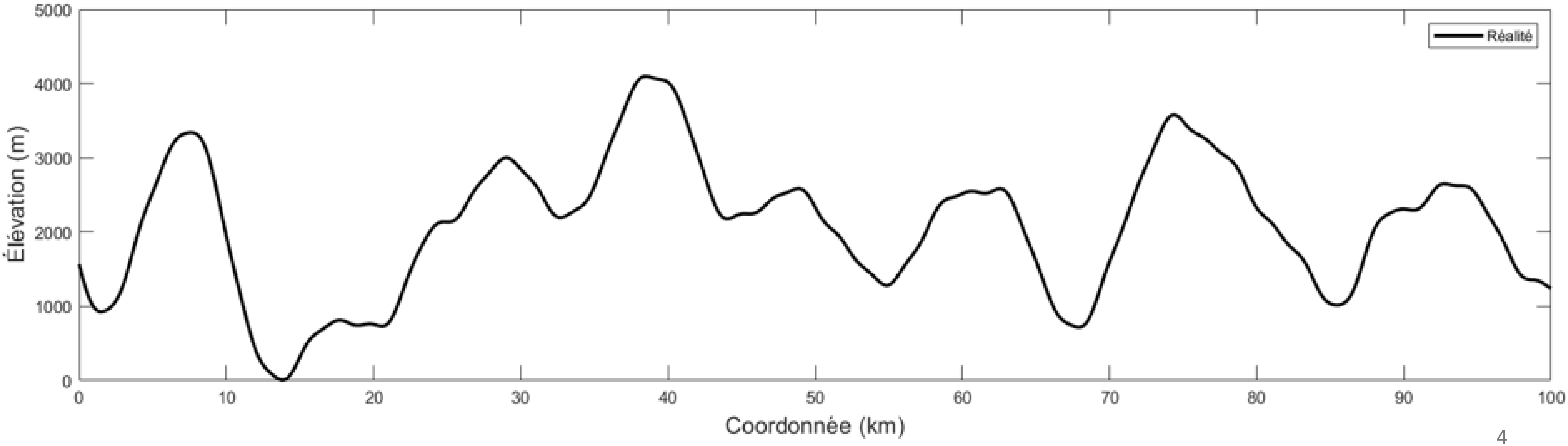
1. Introduction : contexte et problématique
   - Simulation vs krigeage
   - Fonction de transfert
2. Simulation conditionnelle et non-conditionnelle
3. Méthodes de simulations géostatistiques
   i. Transformation graphique (cas non gaussien)
   ii. Simulation Cholesky (LU)
   iii. Simulation séquentielle gaussienne
   iv. FFTMA
   v. Les bandes tournantes
4. Simulation de teneurs de blocs
5. Recuit simulé
6. Post-conditionnement par krigeage simple

# 1. Introduction : contexte et problématique

## Mise en contexte : profil d'un fond marin

Vous déposez un câble de fibre optique sur le fond marin.
Quelle longueur devrait avoir ce câble[1]?

Deux questions?
1) Quelle est la longueur totale du fond marin;
2) Quelle est la pente maximale du fond marin.

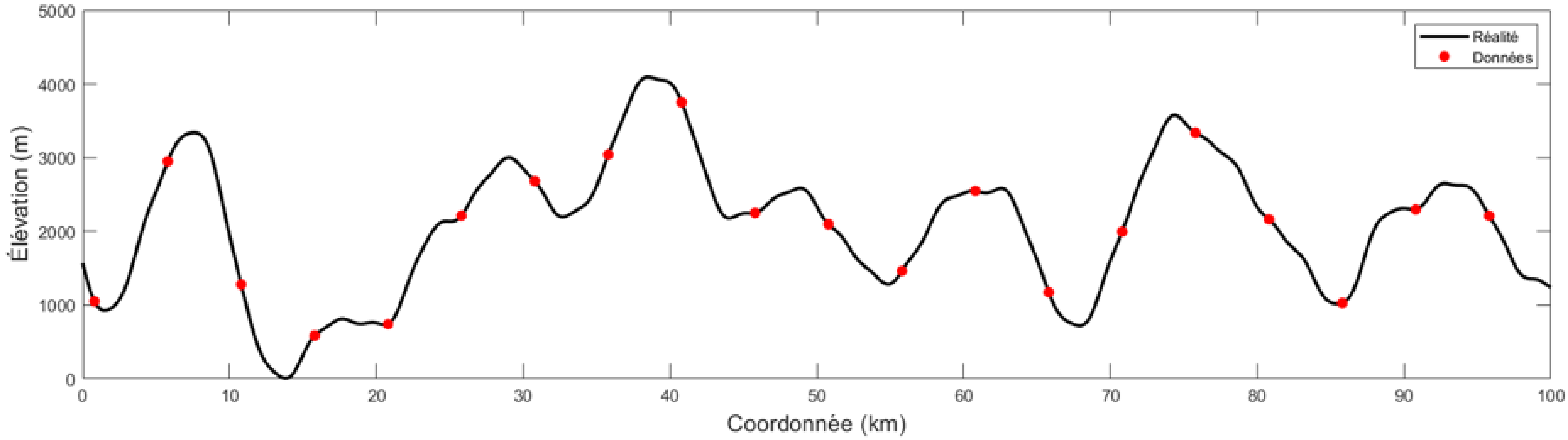



[1]Exemple inspiré de Chilès et Delfiner, 1999 et Alfaro, 1979

# 1. Introduction : contexte et problématique

## Mise en contexte : profil d'un fonds marin

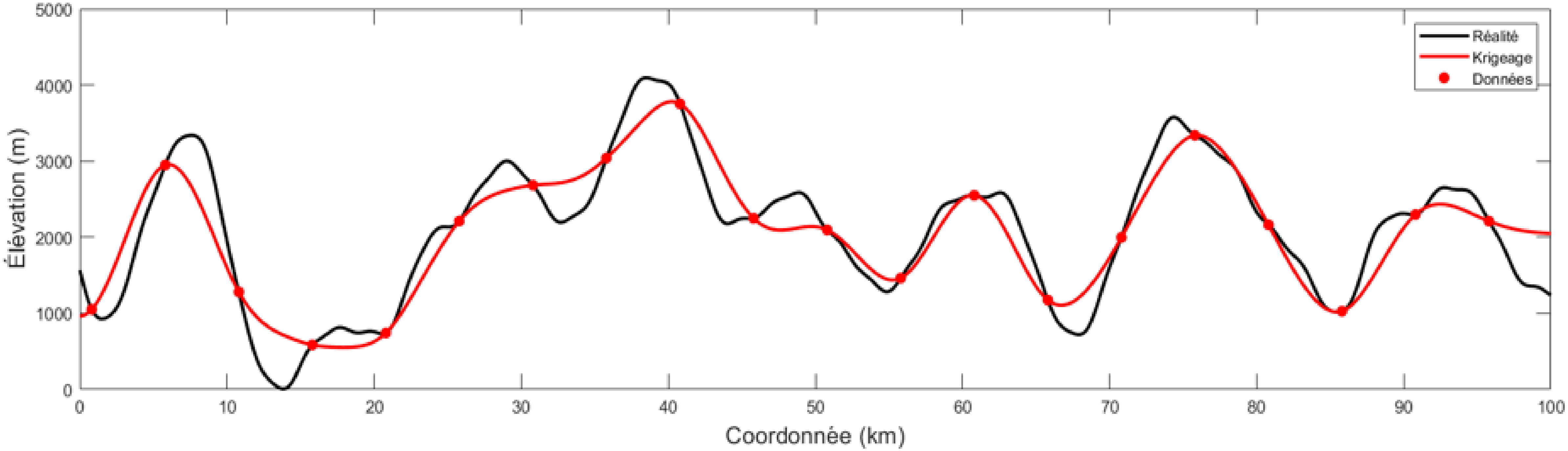
La profondeur exacte est connue uniquement aux points observations



# 1. Introduction : contexte et problématique

**Mise en contexte : profil d'un fonds marin**

On effectue un krigeage

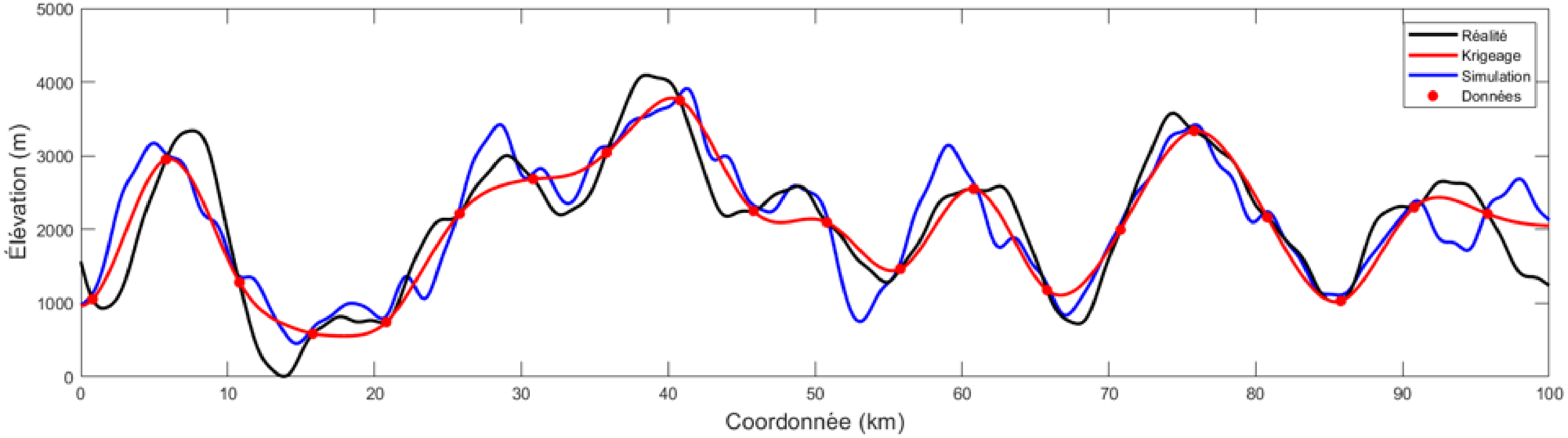


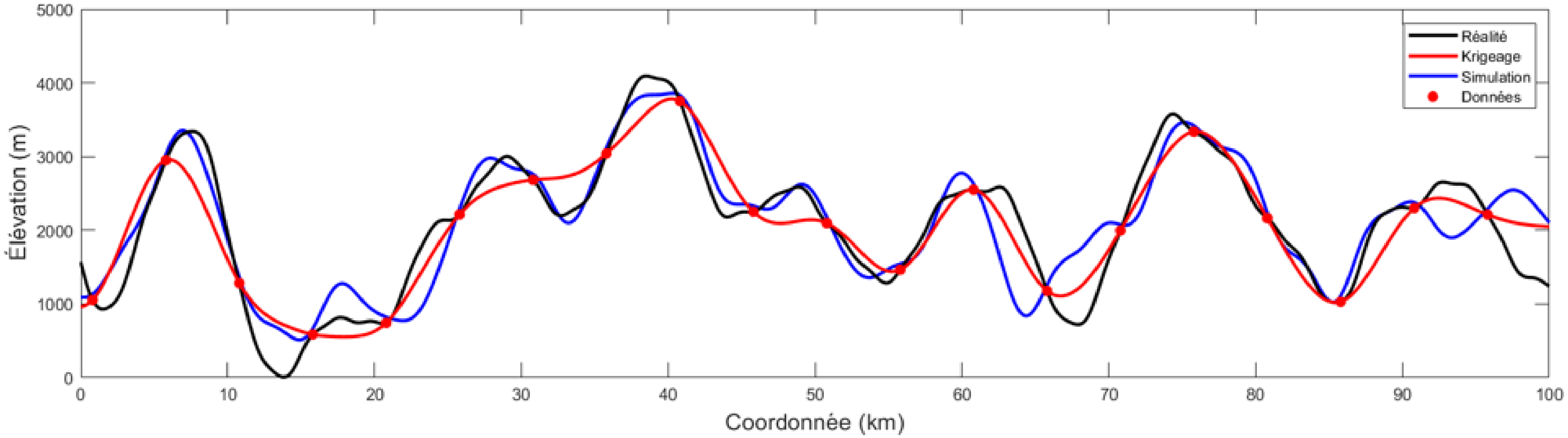

1) La longueur réelle est 105.6 km, le krigeage donne 102.6 km. On va manquer de câble !
2) La pente maximale réelle est de 43°, le krigeage donne 23°. Effet de lissage.

# 1. Introduction : contexte et problématique

## Mise en contexte : profil d'un fonds marin

Changement d'approche : les simulations géostatistiques conditionnelles

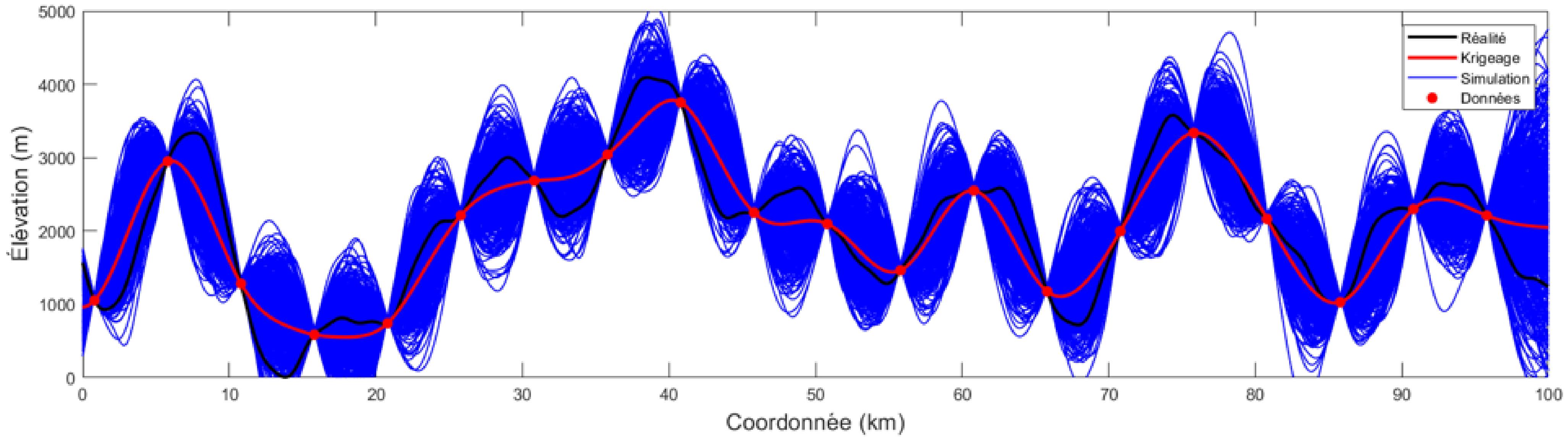


# 1. Introduction : contexte et problématique

## Mise en contexte : profil d'un fonds marin

Changement d'approche : les simulations géostatistiques conditionnelles



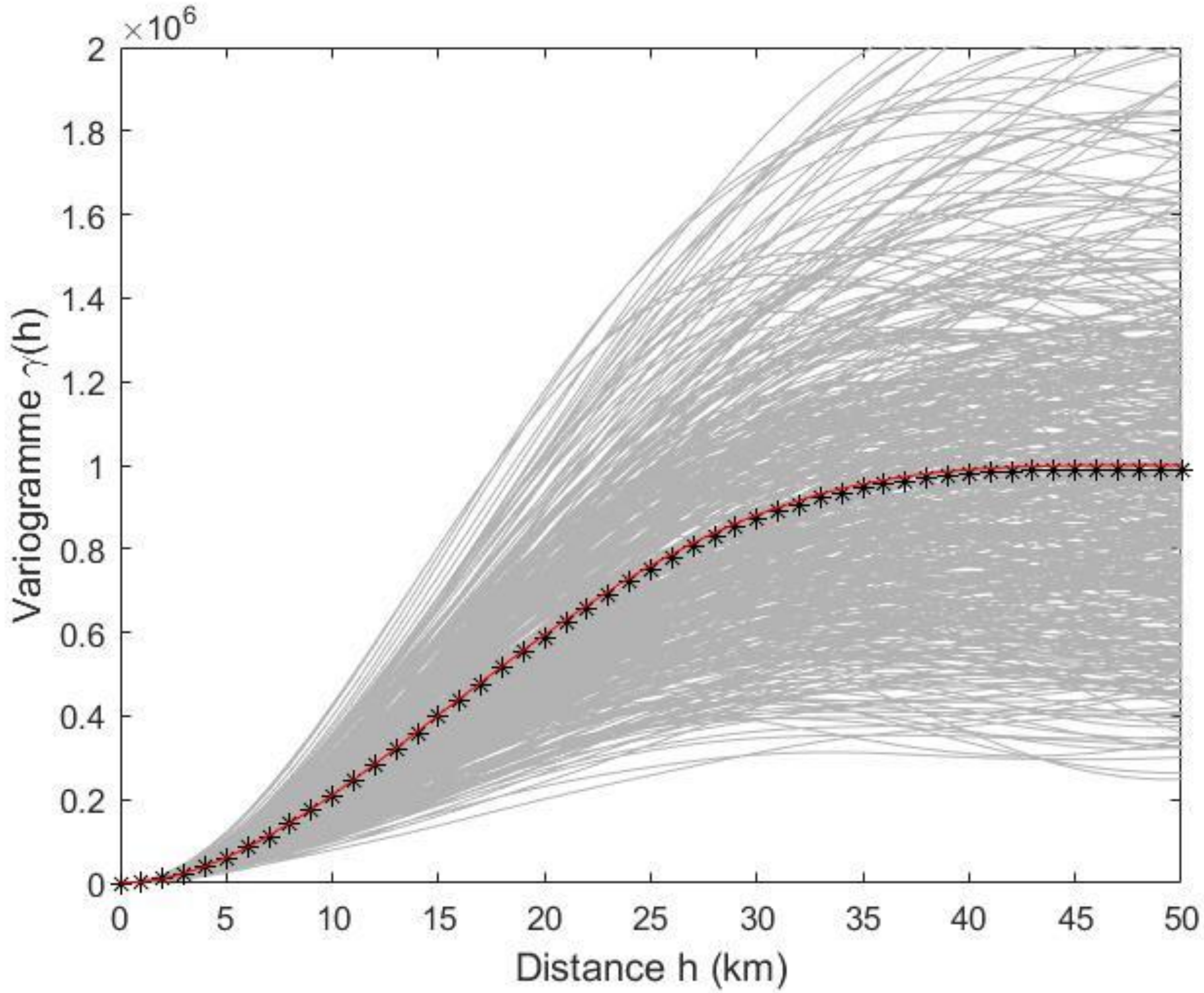

Longueur moyenne des câbles : 105.2km ± 1.6 km
Pente maximale moyenne des profils : 41.2°± 8.5°

# 1. Introduction : contexte et problématique

## Mise en contexte : profil d'un fonds marin

Changement d'approche : les simulations géostatistiques conditionnelles

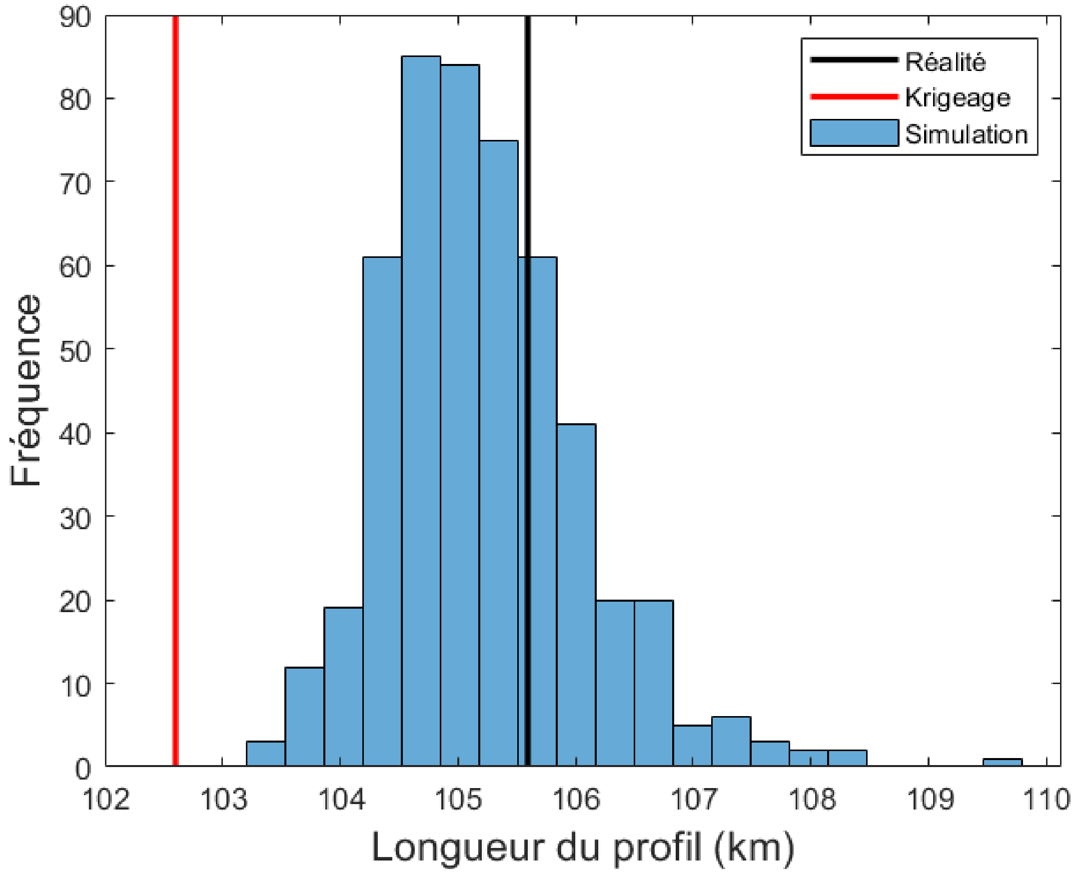
**Rouge** : Modèle théorique
**Gris** : Simulations
**Noir** : Moyenne des simulations

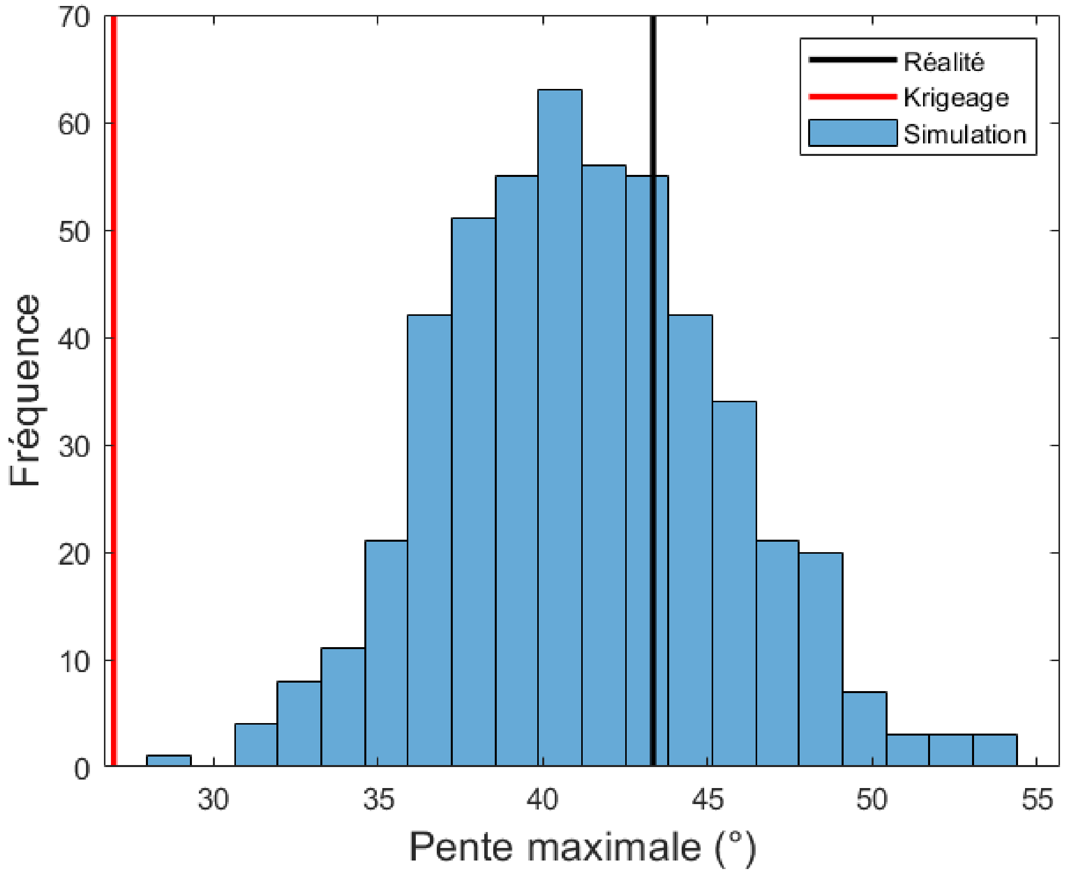


# 1. Introduction : contexte et problématique

## Mise en contexte : profil d'un fonds marin

Comparaison des résultats



# 1. Introduction : contexte et problématique

**Mise en contexte : profil d'un fonds marin**

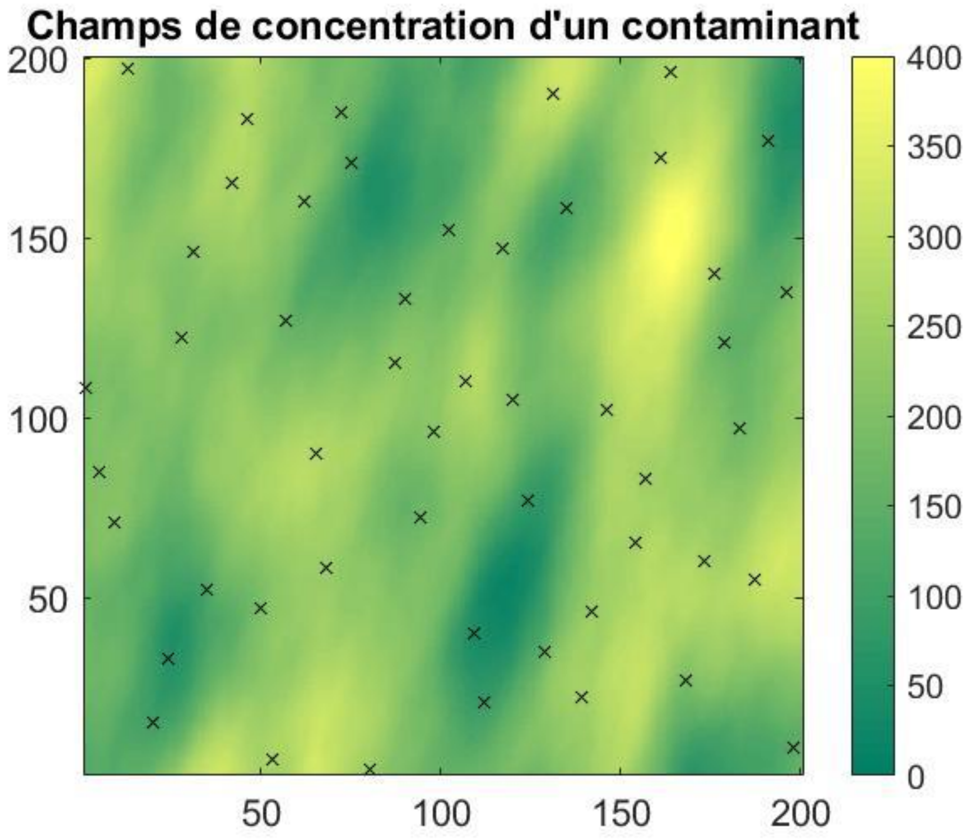
Fonction de transfert : relation non linéaire entre la variable principale (p.ex. profil du fond marin) et l'estimation désirée (p. ex. la longueur du câble, la pente maximale)

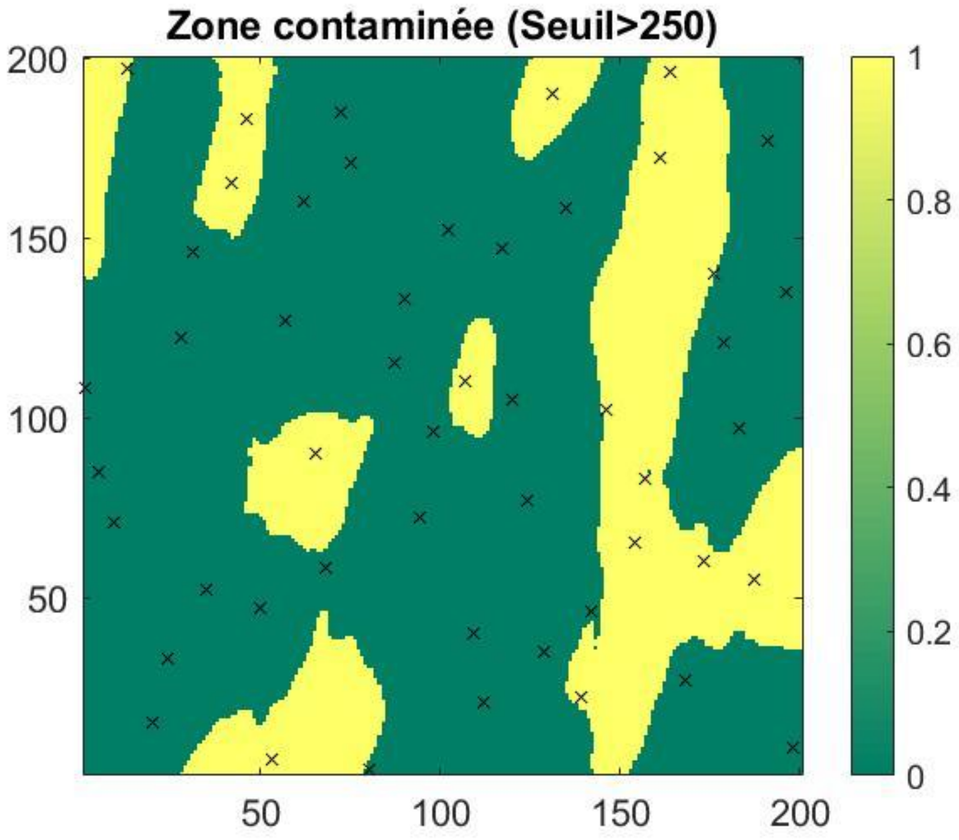
Dès que la fonction de transfert recherchée n'est pas linéairement reliée à la variable principale, on peut avoir intérêt à effectuer des simulations plutôt que des krigeages. C'est souvent le cas.

# 1. Introduction : contexte et problématique

## Mise en contexte : champs de concentration d'un contaminant

Vous disposez d'une cinquantaine d'observations d'un site contaminé. Le seuil de contamination est de 250 ppm. Quel est le volume de sol contaminé à excaver?

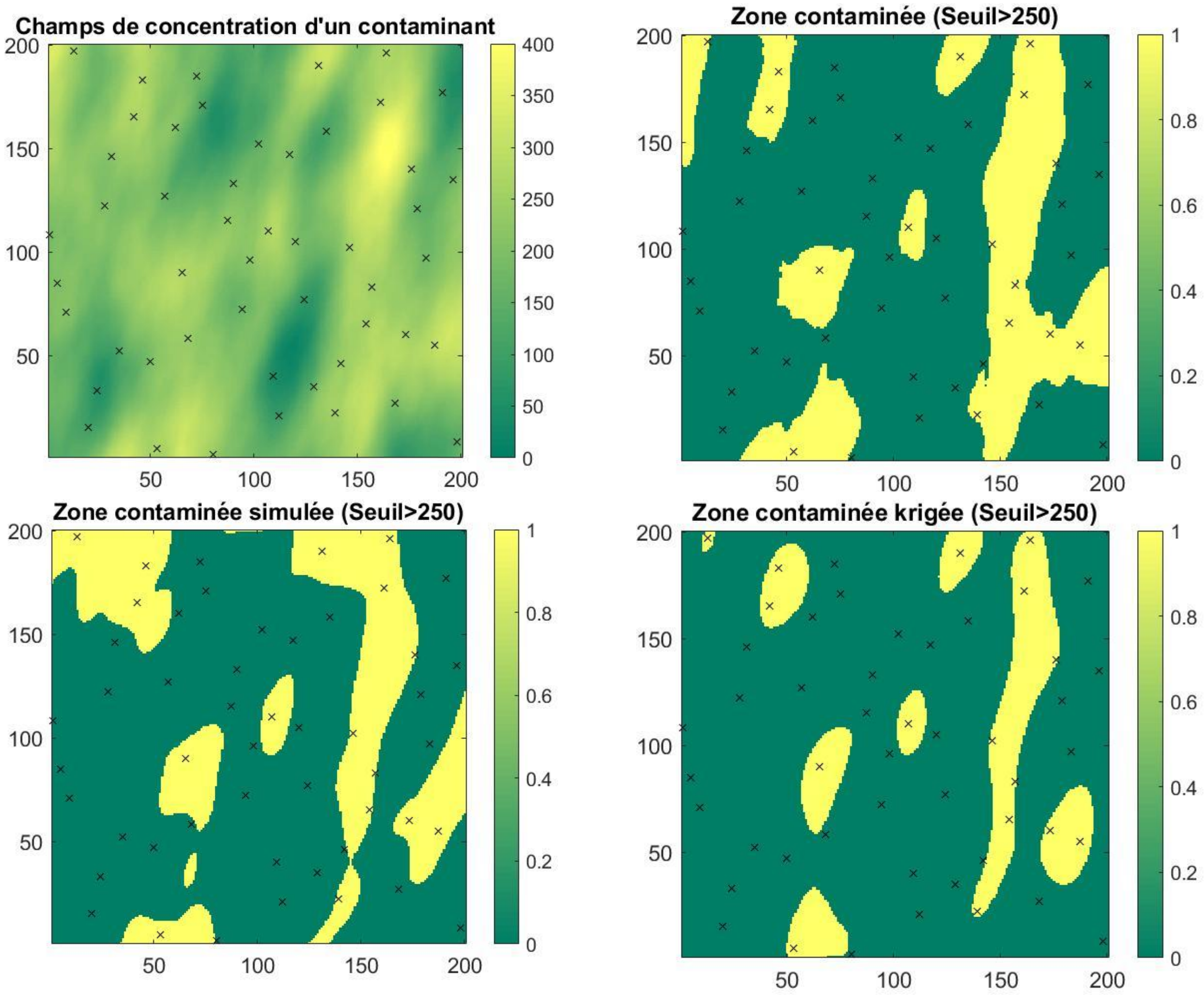


# 1. Introduction : contexte et problématique

## Mise en contexte : champs de concentration d'un contaminant

Les simulations permettent de bien estimer la fonction de transfert (non linéaire)




| Méthode d'estimation | Volume contaminé estimé (%) | Intervalle de confiance à 95% |
|---|---|---|
| Réalité | 25.70 | - |
| Estimation linéaire (Krigeage) | 15.75 | ? |
| Simulations géostatistiques (500 simulations) | 25.03 | [20.43 , 29.63] |

# 1. Introduction : contexte et problématique

**Réservoir pétrolier**

Variables observées

Sismique et forages
(géologie, porosité, perméabilité)

Krigeage
Cokrigeage

Modèle du réservoir

Fonction de transfert (simulateur d'écoulement)

Champs de pression théorique

Corresponds à l'historique de production?

Oui, parfait
Non, quoi faire?

# 1. Introduction : contexte et problématique

**Réservoir pétrolier**

Variables observées

Sismique et forages
(géologie, porosité, perméabilité)

Simulations :
plusieurs réalisations
possibles

Modèle du réservoir

Fonction de
transfert
(simulateur
d'écoulement)

Champs de pression théorique

Corresponds à l'historique
de production?

Conserver les modèles
compatibles, rejeter les autres

Bonus : idée de l'incertitude quant
au véritable modèle du réservoir

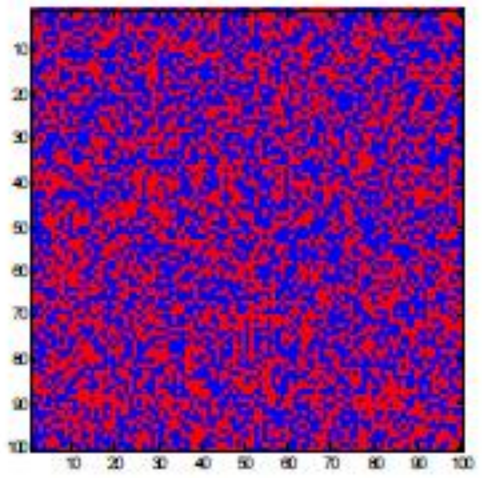
# 1. Introduction : contexte et problématique

## Hydrogéologie : changement d'échelle

Déterminer la transmissivité (T) d'un bloc

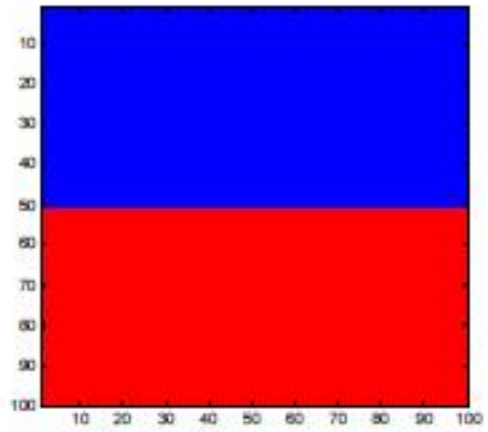
$$T_{m,point} \neq T_{bloc}$$

| Facies | T (cm²/s) |
|---|---|
| Bleu | $10^{-6}$ |
| Rouge | $10^{-2}$ |



| Échelle | T (cm²/s) |
|---|---|
| $T_{m,point}$ | $5 \times 10^{-3}$ |
| $T_{bloc}$ | $10^{-4}$ |

Moyenne géométrique



| Échelle | T (cm²/s) |
|---|---|
| $T_{m,point}$ | $5 \times 10^{-3}$ |
| $T_{bloc,horiz.}$ | $5 \times 10^{-3}$ |
| $T_{bloc,vertic.}$ | $2 \times 10^{-6}$ |

Moyenne arithmétique

Moyenne harmonique

# 1. Introduction : contexte et problématique

**Hydrogéologie : changement d’échelle**

Déterminer la transmissivité (T) d’un bloc

Variable principale : tests piézométriques (quasi ponctuel)
Fonction de transfert : Simulateur d’écoulement

1. Krigeage de blocs → moyenne des transmissivités des points dans le bloc

   Ce n’est pas ce dont on a besoin

2. Krigeage ponctuel et simulateur d’écoulement → les transmissivités krigées n’auront pas le bon variogramme

   Effet de lissage indésiré

3. Simulation des transmissivités ponctuelles et simulateur d’écoulement → transmissivité de bloc

   Bonne approche

# 1. Introduction : contexte et problématique

## Hydrogéologie : aires de protection d'un puits

Variables observées

Tests piézométriques, charges hydrauliques dans des puits d'observation, conditions frontières, recharge estimée, débit des puits

Simulations : plusieurs réalisations de la conductivité hydraulique

Modèle d'écoulement

Fonction de transfert (simulateur d'écoulement)

Champs de charges hydrauliques

Déterminer plusieurs aires de captage :

- Jamais captés
- Toujours captés
- Captés dans x% des cas

Bonus : se prémunir contre de mauvaises surprises, p. ex. un puits qui aurait dans sa zone de captage une zone potentiellement contaminée

*Procédure similaire pour :

- déterminer des distributions de temps de transport
- déterminer les trajets possibles d'une contamination

# 1. Introduction : contexte et problématique

**Mines : optimisation de fosses à ciel ouvert**

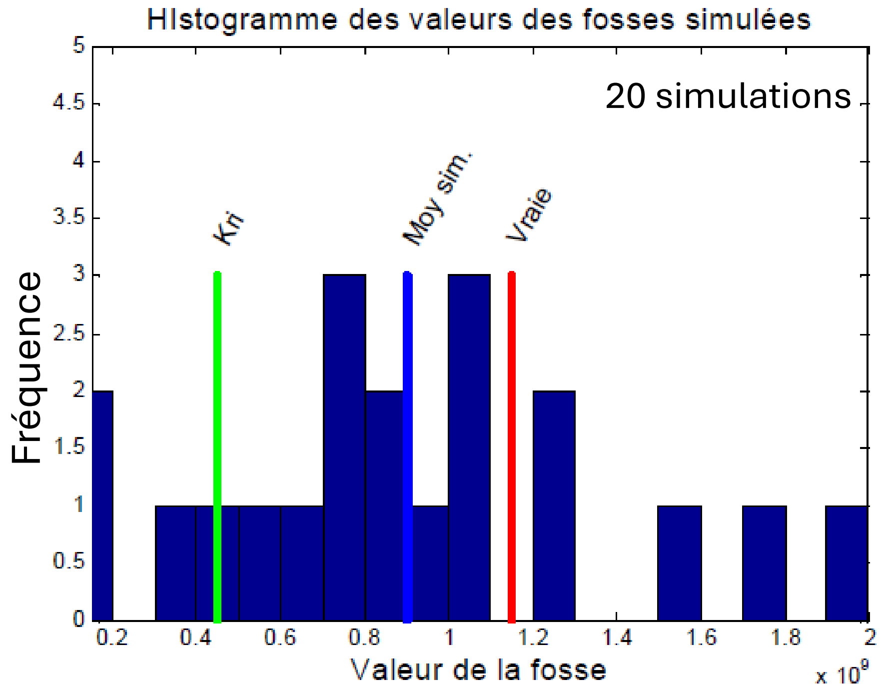
Variable principale : les teneurs quasi ponctuelles
Fonction de transfert : Teneurs de blocs et optimisation des contours de la fosse

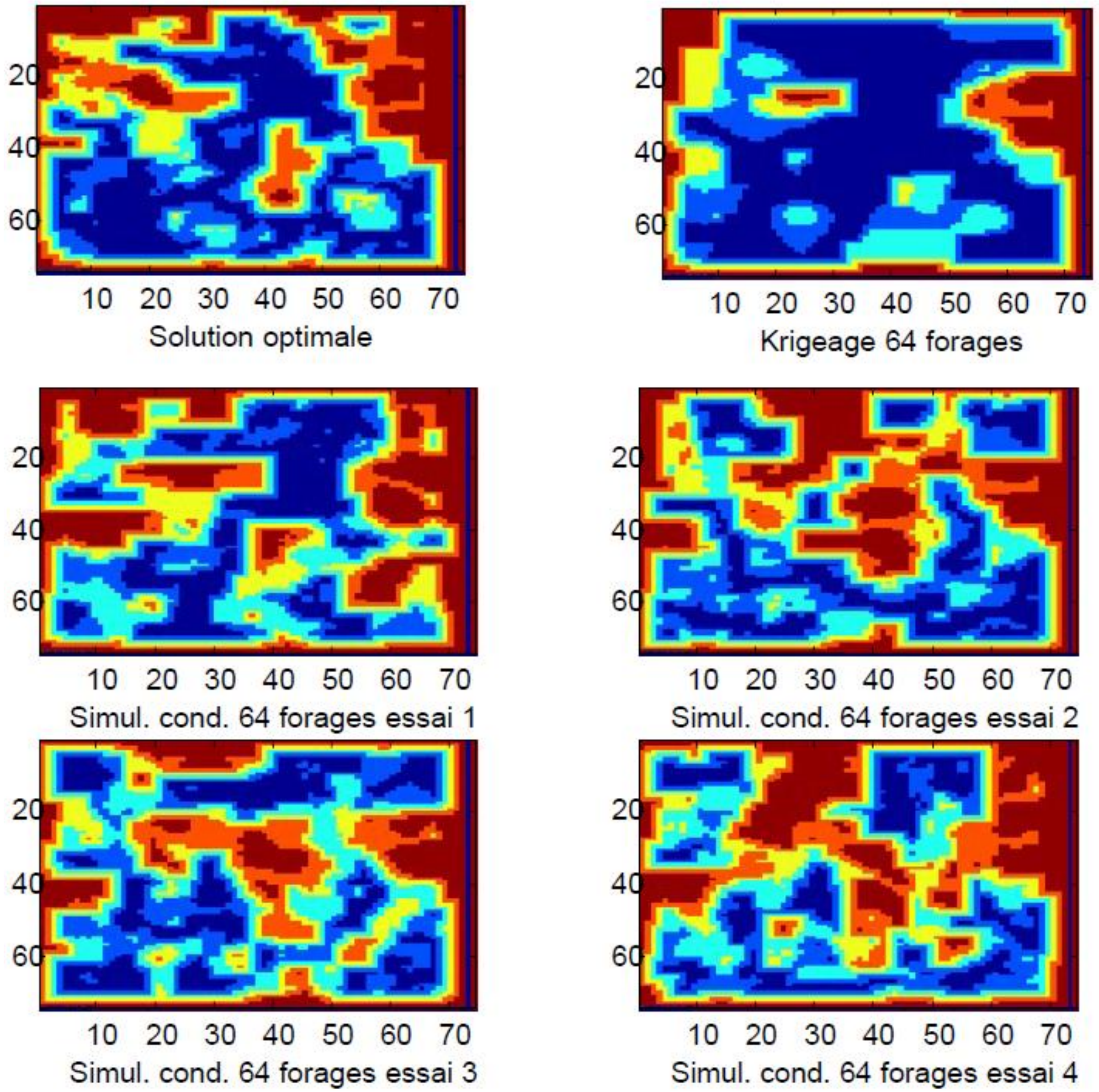
Lissage du krigeage → La fosse optimisée peut être très éloignée de la fosse qui sera réellement minée

Simulations géostatistique → Produire $n$ fosses optimisées qui encadrera la valeur de la fosse qui sera minée.

# 1. Introduction : contexte et problématique

## Mines : optimisation de fosses à ciel ouvert



# 1. Introduction : contexte et problématique

**Mines : précision sur les ressources et le profit conventionnel**

Chaque réalisation

- ressources estimées différentes, compatibles avec les observations actuelles

- distribution des ressources, intervalle de confiance

Approche valide pour toute situation où l’on applique un seuil de sélection

p. ex.

- Environnement : volume à excaver parce que contaminé
- Biologie : aires avec une biomasse suffisante pour nourrir un prédateur

# 2. Simulation conditionnelle et non-conditionnelle

## Idée principale des simulations géostatistiques

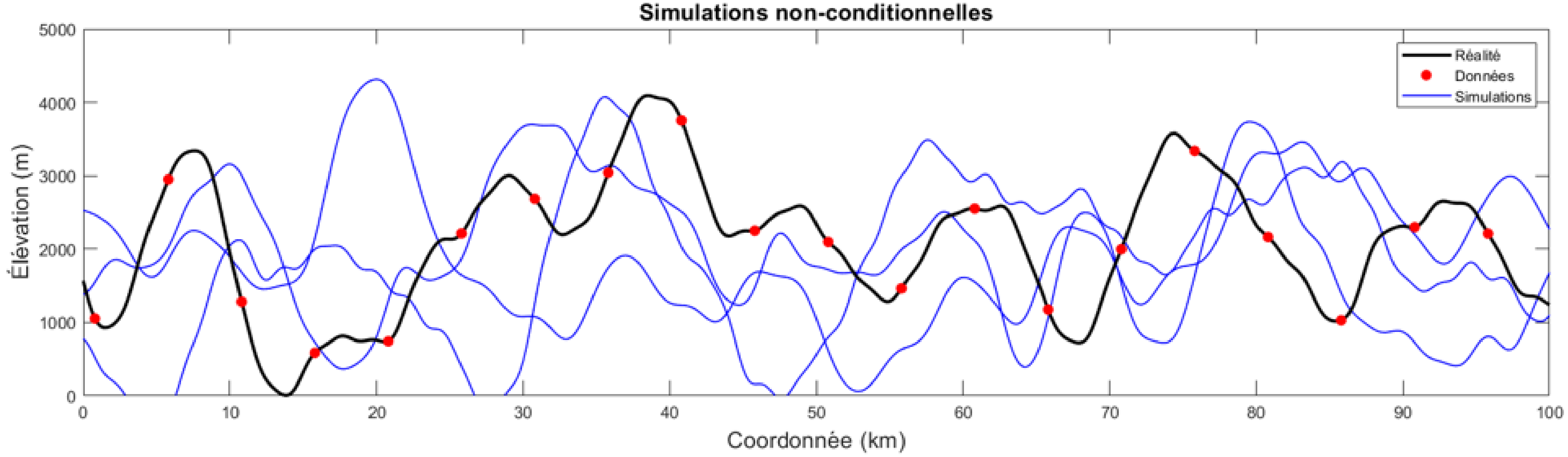
Proposer une solution à tout problème impliquant des **transformations non linéaires** des variables mesurées.

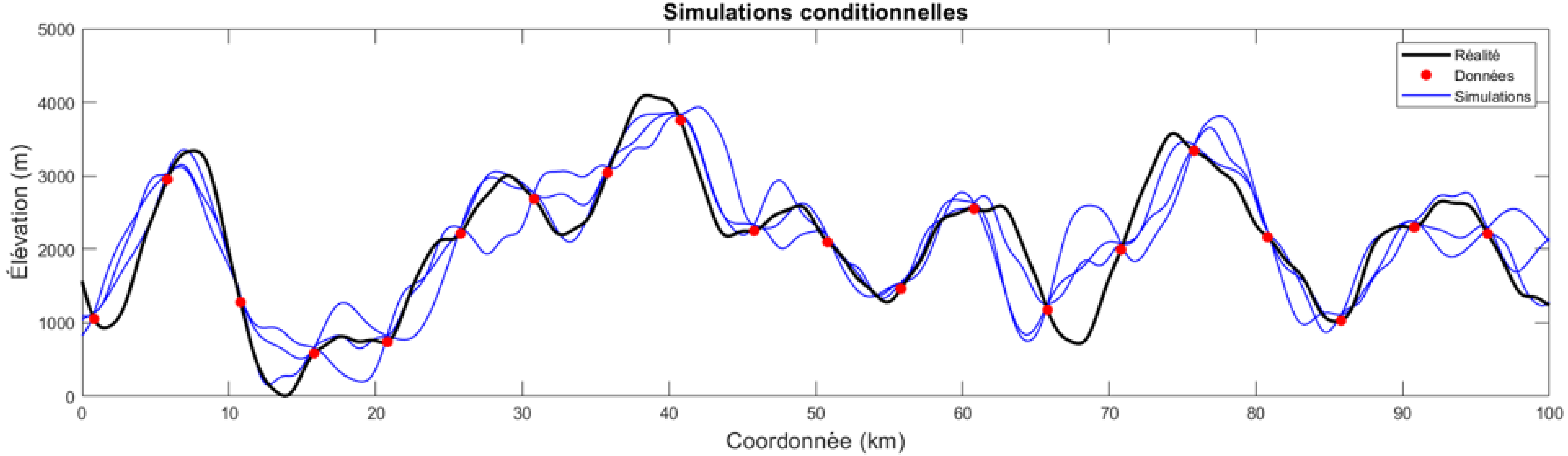
Simulation non conditionnelle : Produire des champs montrant la même structure spatiale (variogramme) et le même histogramme que ceux inférés à partir des données observées.

Simulation conditionnelle : Produire des champs montrant la même structure spatiale (variogramme), le même histogramme en plus de respecter les données observées.

# 2. Simulation conditionnelle et non-conditionnelle

**Exemple**



# 2. Simulation conditionnelle et non-conditionnelle

**En tableau**

| Reproduit ? | Non-conditionnelle | Conditionnelle |
|---|---|---|
| Histogramme | Oui | Oui |
| Variogramme | Oui (théorique) ? ⟷ | Oui ( mélange de théorique et expérimentale) |
| Données | Non | Oui |
| Fonction de transfert non linéaire | Oui, mais pas directement ? ⟷ | Oui, mais pas directement |

# 3. Méthode de simulations géostatistiques

## Différences entre les méthodes

Il existe un grand nombre de méthodes de simulations

Ce qui les distingue :

1. Variable <u>continue</u> ou catégorique
2. Utilise la <u>covariance</u> (bipoints) ou multipoints
3. Objets versus <u>pixels</u>
4. <u>Gaussien</u> versus distribution quelconque
5. Simulation conditionnelle <u>possible</u> ou non (directement)
6. -Limité en 1D ou <u>non</u>
7. -Grille régulière ou <u>quelconque</u>

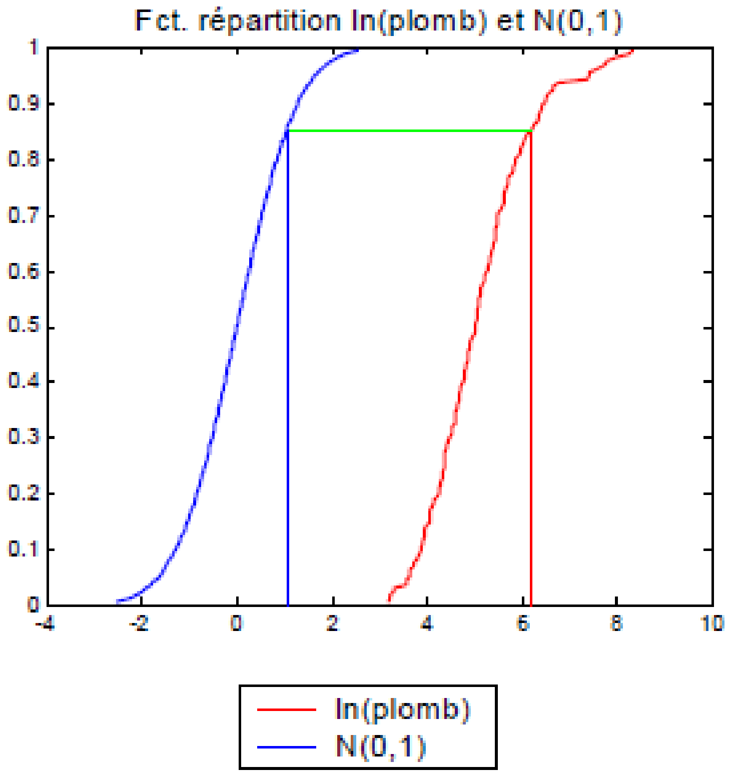
# 3. Méthode de simulations géostatistiques

**Normalité des observations :** Cas non gaussien

LU, SGS, FFTMA et les bandes tournantes nécessitent un champ $Z(x)$ gaussien

<u>Que faire si ce n'est pas gaussien?</u>

1. Transformer $Z(x)$ en $Y(x) = f^{-1}(Z(x))$;
2. Calculer et modéliser le variogramme de $Y(x)$;
3. Simuler (conditionnel ou non) $Y(x)$;
4. Effectuer la transformation inverse $Z(x) = f\big(Y(x)\big)$



# 3. Méthode de simulations géostatistiques

***<u>Méthode de Cholesky (LU) : Algorithme simulation non conditionnelle</u>***

Soit $Z(x)$ gaussien, de moyenne 0 et de covariance $C(h)$:

On cherche à simuler n points à des emplacements $x_{\mathrm{i}}, i = 1, \dots, n$ :

1. Construire la matrice des covariances entre les n points, nommée K. K est positive définie. (Il s'agit de la même matrice K que le krigeage simple.)

2. Effectuer la décomposition $K = LL'$. ($L$ est triangulaire inférieure $n \times n$.)

3. Générer un vecteur aléatoire $Y_{n \times 1}$ tiré d'une loi normale $N(0,1)$.

4. Calculer $Z = LY$.

$$\mathrm{Cov}(\mathrm{Z}, \mathrm{Z}') = \mathrm{E}[ZZ'] = \mathrm{E}[LYY'L'] = L\mathrm{E}[YY']L' = LIL' = K$$

## ***Méthode de Cholesky (LU) : Algorithme simulation conditionnelle***

Soit $Z(x)$ gaussien, de moyenne 0 et de covariance $C(h)$et soit $z^1(x_i), i = 1, \dots N$, N observations de $Z(X)$ :

On cherche à simuler $z^2(x_j)$ en n emplacements $x_j, j = 1, \dots, n$ , conditionnellement à $z^1(x_i), i = 1, \dots N$

1. Construire la matrice des covariances entre les $N + n$ points, nommée $K$. $K$ est positive définie.

2. Effectuer la décomposition $K = LL'$. ($L$ est triangulaire inférieure)

3. Partager $K$ et $L$ en 4 blocs :

$$K = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \qquad L = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix}$$

$K_{11}$ et $L_{11}$ sont $N \ x \ N$
$K_{22}$ et $L_{22}$ sont $n \ x \ n$
Indice 1 : points observés
Indice 2 : points à simuler

## ***Méthode de Cholesky (LU) : Algorithme simulation conditionnelle (suite)***

4. Déterminer les valeurs du vecteur $y^1_{N\times 1}$ qui assure la reproduction des données observées.

$$z^1 = L_{11}y^1 \ \rightarrow y^1 = L_{11}^{-1}z^1$$

5. Générer un vecteur aléatoire $y^2_{n\times 1}$ tiré d'une loi normale $N(0,1)$.

6. Calculer $Z = LY$ :

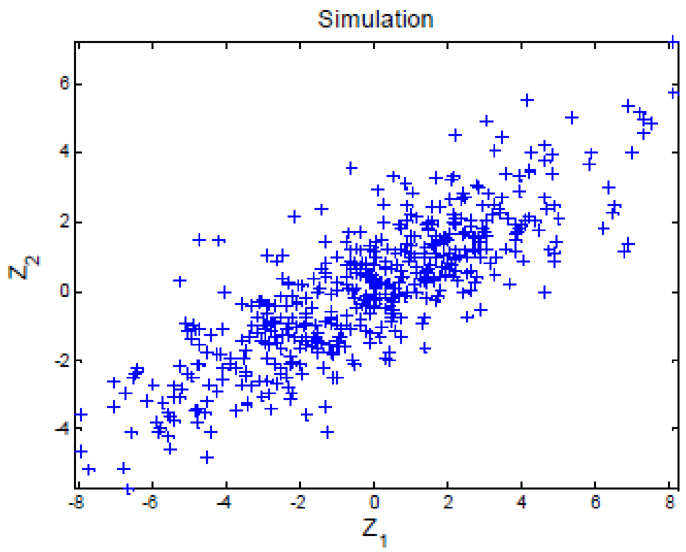
$$Z = LY \rightarrow \begin{bmatrix} z^1 \\ z^2 \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} y^1 \\ y^2 \end{bmatrix}$$

$$z^1 = L_{11}y^1$$

$$z^2 = L_{21}y^1 + L_{22}y^2 = (L_{21}L_{11}^{-1})z^1 + L_{22}y^2$$

## Méthode de Cholesky (LU) : Exemple

On cherche à simuler deux variables ($Z_1$ et $Z_2$) de variance respective de 9 et 4 avec une corrélation de 0.8 :




$$Cov(Z_1, Z_2) = \rho\sigma_{Z_1}\sigma_{Z_2} = 0.8 \times (9)^{0.5} \times (4)^{0.5} = 4.8$$

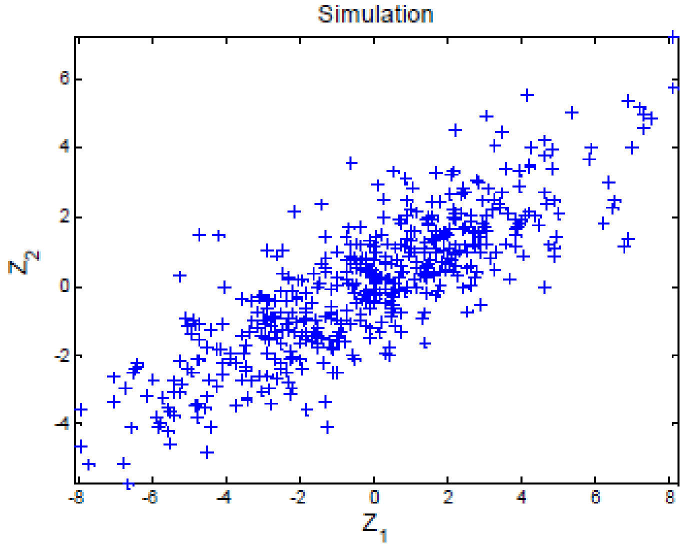
$$K = \begin{bmatrix} 9 & 4.8 \\ 4.8 & 4 \end{bmatrix} \qquad L = \begin{bmatrix} & ? & \end{bmatrix}$$

**Déterminer L :**

## Méthode de Cholesky (LU) : Exemple

On cherche à simuler deux variables ($Z_1$ et $Z_2$) de variance respective de 9 et 4 avec une corrélation de 0.8 :

$$Cov(Z_1, Z_2) = 0.8 \times (9 \times 4)^{0.5} = 4.8$$




$$K = \begin{bmatrix} 9 & 4.8 \\ 4.8 & 4 \end{bmatrix} \qquad L = \begin{bmatrix} 3 & 0 \\ 1.6 & 1.2 \end{bmatrix}$$

$$\text{Soit : } Y_1 = \begin{bmatrix} -0.12 \\ -0.5 \end{bmatrix} \rightarrow \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 1.6 & 1.2 \end{bmatrix} \begin{bmatrix} -0.12 \\ -0.5 \end{bmatrix} = \begin{bmatrix} -0.36 \\ -0.79 \end{bmatrix}$$

$$Y_2 = \begin{bmatrix} 1.47 \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 1.6 & 1.2 \end{bmatrix} \begin{bmatrix} 1.47 \\ 1 \end{bmatrix} = \begin{bmatrix} 4.41 \\ 3.55 \end{bmatrix}$$

# 3. Méthode de simulations géostatistiques

**Méthode de Cholesky (LU) : Avantages et inconvénients**

### Avantages

- Facile à comprendre et programmer
- Très rapide pour de petits champs (n petits)
- Cas conditionnel pose aucun problème
- Se généralise immédiatement au cas multivariable

### Inconvénients

- Limiter à la simulation de petits champs à cause de problèmes d'espace mémoire : (n+N)<10000
- $K$ peut ne pas être définie numériquement ( $K$ l'est théoriquement). Impossibilité d'effectuer la décomposition de Cholesky.

# 3. Méthode de simulations géostatistiques

## ***Méthode séquentielle gaussienne (SGS) : algorithme***

Soit $Z(x)$ gaussien, de moyenne 0 et de covariance $C(h)$et soit $\mathrm{Z}(x_i), i = 1, \dots N, N$ observations de $Z(X)$ :

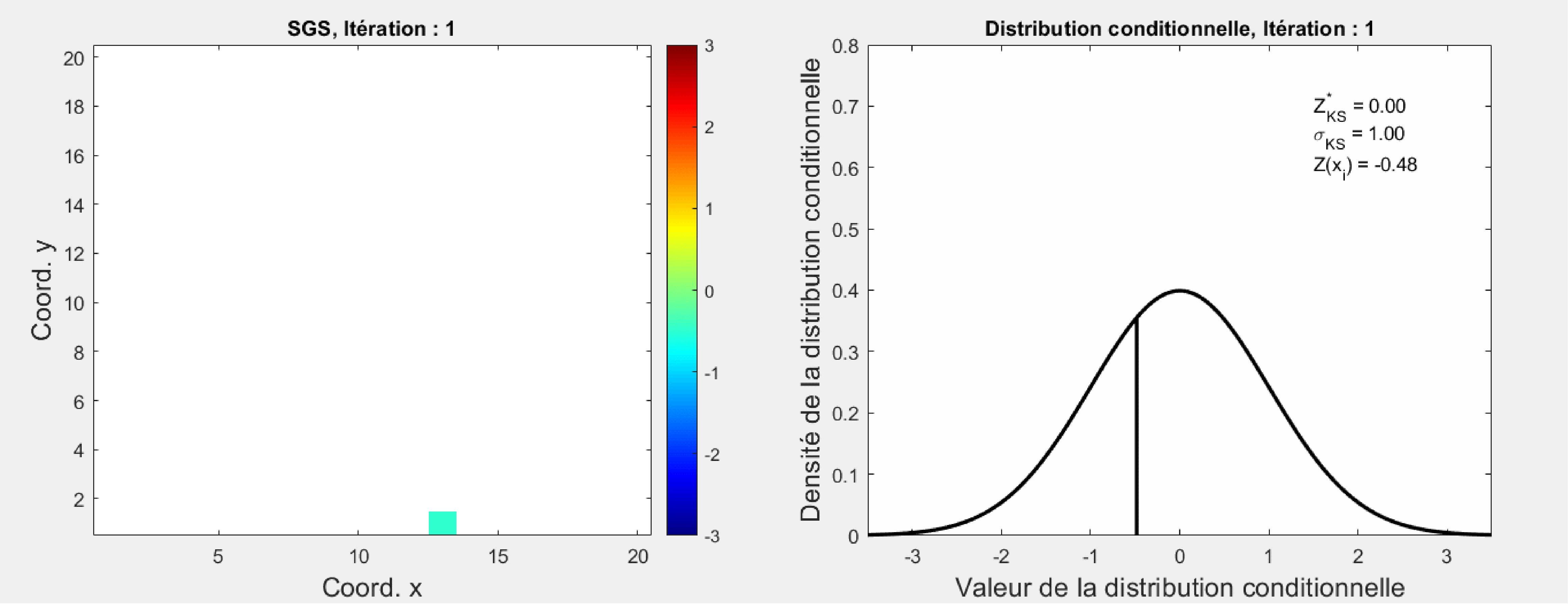
On cherche à simuler n points à des emplacements $x_{\mathrm{j}}, j = 1, \dots, n$ :

1. Choisir une coordonnée $x_{\mathrm{j}}$ aléatoirement;
2. Effectuer le krigeage simple à ce point en utilisant les observations $Z(x_i), i = 1, \dots N$. On obtient ainsi un estimé $Z_{KS}^*(x_j)$ et la variance conditionnelle $\sigma_{KS}^2$
3. Tirer aléatoirement une donnée de la distribution $N\left(Z_{KS}^*(x_j), \sigma_{KS}^2\right)$
4. Ajouter cette valeur aux valeurs observées
5. Répéter les étapes 1 à 4 pour les emplacements restants à simuler.

# 3. Méthode de simulations géostatistiques

## ***Méthode séquentielle gaussienne (SGS) :***

**Animation :** simulation d'un champ 20x20 avec modèle sphérique isotrope (a=5)



# 3. Méthode de simulations géostatistiques

**Méthode séquentielle gaussienne (SGS) : démonstration par induction**

Supposons que l'algorithme permet de simuler « n » valeurs normales de moyenne 0 et de covariance $C(h)$.

On a :

$$Cov\left(Z_n^s, {Z_n^s}'\right) = Cov(Z_n, Z_n') = K_{n\times n}$$

On réalise le krigeage simple :

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

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

On ajoute l'erreur :

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

# 3. Méthode de simulations géostatistiques

**Méthode séquentielle gaussienne (SGS) : démonstration par induction**

On calcule alors :

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

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

# 3. Méthode de simulations géostatistiques

## Méthode séquentielle gaussienne (SGS) : voisinage

Au fur et à mesure que l'algorithme progresse, le nombre de points disponible pour le krigeage augmente; pour N grand, deviens prohibitif.

→ Effectuer les krigeages en voisinages glissants (**effet d'écran**)

- Assure approximativement la reproduction de K, d'autant mieux que l'effet d'écran est important;
- Certaines covariances sont difficiles à reproduire par cet algorithme (ex. modèle gaussien), car dans ce cas l'effet d'écran est faible.

# 3. Méthode de simulations géostatistiques

## Méthode séquentielle gaussienne (SGS) : Avantages et inconvénients

### Avantages

- Facile à comprendre et programmer
- Cas conditionnel immédiat
- Se généralise au cas multivariable. Utiliser le CS au lieu de KS.

### Inconvénients

- Certaines covariances difficiles à bien simuler
- Assez lent pour de grands champs
- Limiter les voisinages peut mener à d’importantes distorsions sur la covariance simulée

# 3. Méthode de simulations géostatistiques

**Méthode séquentielle gaussienne (SGS) :**

Question

Le SGS consiste à tirer une valeur de la distribution conditionnelle obtenue par krigeage simple dans le cas gaussien

Le krigeage d'indicatrice permet d'estimer une distribution conditionnelle dans le cas où la variable n'est pas gaussienne

Comment pourrait-on combiner ces deux algorithmes pour développer une méthode séquentielle pour le cas non gaussien?

# 3. Méthode de simulations géostatistiques

## Méthode séquentielle gaussienne (SGS) : Variantes

SGS est une méthode de simulation assez versatile, car elle a besoin simplement d'une moyenne conditionnelle (l'estimation) et une variance conditionnelle. Donc toutes méthodes géostatistiques permettant d'obtenir ces deux paramètres peut être combiné à l'idéologie du SGS :

1. Krigeage simple avec moyenne stationnaire.
2. Krigeage ordinaire.
3. Cokrigeage simple.
4. Simulation colocalisée (cosimulation).
5. (Co)krigeage simple avec moyenne localement variable (LVM).
6. Cokrigeage colocalisé.
7. Cokrigeage avec le modèle de corrélation intrinsèque.
8. Mise à jour bayésienne.
9. Et autres (direct, d'indicatrice, avec krigeage de block,...)

# 3. Méthode de simulations géostatistiques

## Méthode séquentielle gaussienne (SGS) : Variantes

SGS est une méthode de simulation assez versatile, car elle a besoin simplement d'une moyenne conditionnelle (l'estimation) et une variance conditionnelle. Donc toutes méthodes géostatistiques permettant d'obtenir ces deux paramètres peut être combiné à l'idéologie du SGS :

1. Krigeage simple avec moyenne stationnaire.
2. Krigeage ordinaire.
3. Cokrigeage simple.
4. Simulation colocalisée (cosimulation).
5. (Co)krigeage simple avec moyenne localement variable (LVM).
6. Cokrigeage colocalisé.
7. Cokrigeage avec le modèle de corrélation intrinsèque.
8. Mise à jour bayésienne.
9. Et autres (direct, d'indicatrice, avec krigeage de block,...)