# Introduction à la Géostatistique

**Master DeSIgéo**

yann.meneroux@ign.fr
Service de Géodésie et de Métrologie
Institut National de l'Information Géographique et Forestière

**14-15 novembre 2023**

INSTITUT NATIONAL DE L'INFORMATION GÉOGRAPHIQUE ET FORESTIÈRE

Lab on Geographic Information Science for sustainable development and smart cities

# Interpolation stochastique

L'**interpolation stochastique** utilise des modèles probabilistes pour estimer la valeur prise par la variable régionalisée $z$ en tout point du champ $\mathcal{D}$ avec la variance de l'erreur associée $\mathbb{V}[\hat{z} - z]$.

# Interpolation stochastique

L'**interpolation stochastique** utilise des modèles probabilistes pour estimer la valeur prise par la variable régionalisée $z$ en tout point du champ $\mathcal{D}$ avec la variance de l'erreur associée $\mathbb{V}[\hat{z} - z]$.

Travail en deux temps :

# Interpolation stochastique

L'**interpolation stochastique** utilise des modèles probabilistes pour estimer la valeur prise par la variable régionalisée $z$ en tout point du champ $\mathcal{D}$ avec la variance de l'erreur associée $\mathbb{V}[\hat{z} - z]$.

Travail en deux temps :

1. Modélisation de la structure spatiale de la variable régionalisée

# Interpolation stochastique

L'**interpolation stochastique** utilise des modèles probabilistes pour estimer la valeur prise par la variable régionalisée $z$ en tout point du champ $\mathcal{D}$ avec la variance de l'erreur associée $\mathbb{V}[\hat{z} - z]$.

Travail en deux temps :

1. Modélisation de la structure spatiale de la variable régionalisée
2. Estimations globales et locales

# Interpolation stochastique

L'**interpolation stochastique** utilise des modèles probabilistes pour estimer la valeur prise par la variable régionalisée $z$ en tout point du champ $\mathcal{D}$ avec la variance de l'erreur associée $\mathbb{V}[\hat{z} - z]$.

Travail en deux temps :

1. Modélisation de la structure spatiale de la variable régionalisée
2. Estimations globales et locales

$\rightarrow$ Niveau d'abstraction plus élevé

# Interpolation stochastique

L'**interpolation stochastique** utilise des modèles probabilistes pour estimer la valeur prise par la variable régionalisée $z$ en tout point du champ $\mathcal{D}$ avec la variance de l'erreur associée $\mathbb{V}[\hat{z} - z]$.

Travail en deux temps :

1. Modélisation de la structure spatiale de la variable régionalisée
2. Estimations globales et locales

$\rightarrow$ Niveau d'abstraction plus élevé

$\rightarrow$ Précautions supplémentaires lors du retour au *monde physique*

# Interpolation stochastique

## Définition

La variable régionalisée est une réalisation de la fonction aléatoire (on parle aussi de processus stochastique) : $z(x) = Z(x, \omega)$

$$Z \ : \ \mathcal{D} \times \Omega \quad \longrightarrow \quad \mathbb{R}$$

.

# Interpolation stochastique

## Définition

La variable régionalisée est une réalisation de la fonction aléatoire (on parle aussi de processus stochastique) : $z(x) = Z(x, \omega)$

$$Z \ : \ \mathcal{D} \times \Omega \quad \longrightarrow \quad \mathbb{R}$$

.

La fonction aléatoire est définie sur 2 espaces :

# Interpolation stochastique

## Définition

La variable régionalisée est une réalisation de la fonction aléatoire (on parle aussi de processus stochastique) : $z(x) = Z(x, \omega)$

$$Z \ : \ \mathcal{D} \times \Omega \quad \longrightarrow \quad \mathbb{R}$$

.

La fonction aléatoire est définie sur 2 espaces :

1. L'espace géographique $\mathcal{D}$

# Interpolation stochastique

## Définition

La variable régionalisée est une réalisation de la fonction aléatoire (on parle aussi de processus stochastique) : $z(x) = Z(x, \omega)$

$$Z \ : \ \mathcal{D} \times \Omega \quad \longrightarrow \quad \mathbb{R}$$

.

La fonction aléatoire est définie sur 2 espaces :

1. L'espace géographique $\mathcal{D}$
2. L'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$

# Interpolation stochastique

## Définition

La variable régionalisée est une réalisation de la fonction aléatoire (on parle aussi de processus stochastique) : $z(x) = Z(x, \omega)$

$$Z \ : \ \mathcal{D} \times \Omega \quad \longrightarrow \quad \mathbb{R}$$

.

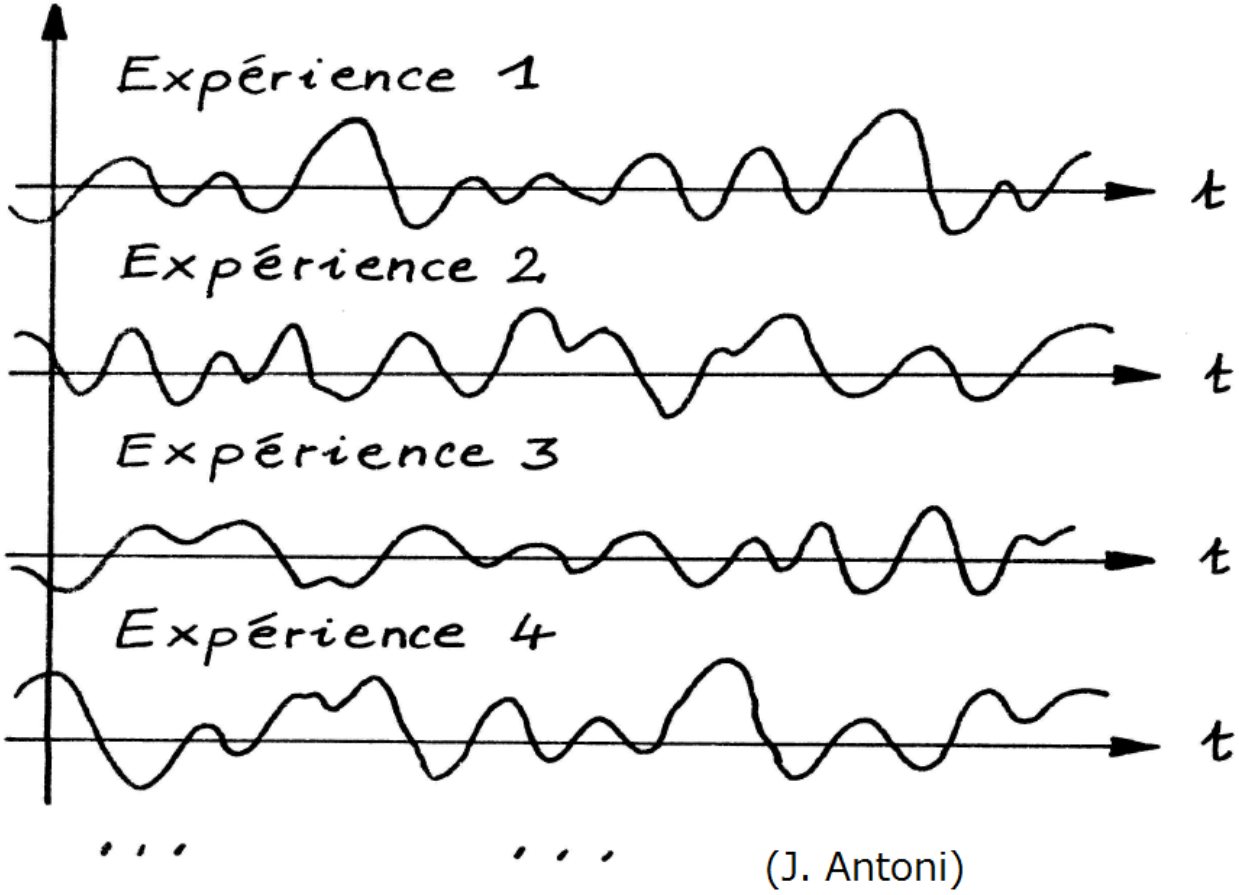
La fonction aléatoire est définie sur 2 espaces :

1. L'espace géographique $\mathcal{D}$
2. L'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$

$\rightarrow$ On parle alors d'un espace *topo-probabiliste*

# Interpolation stochastique

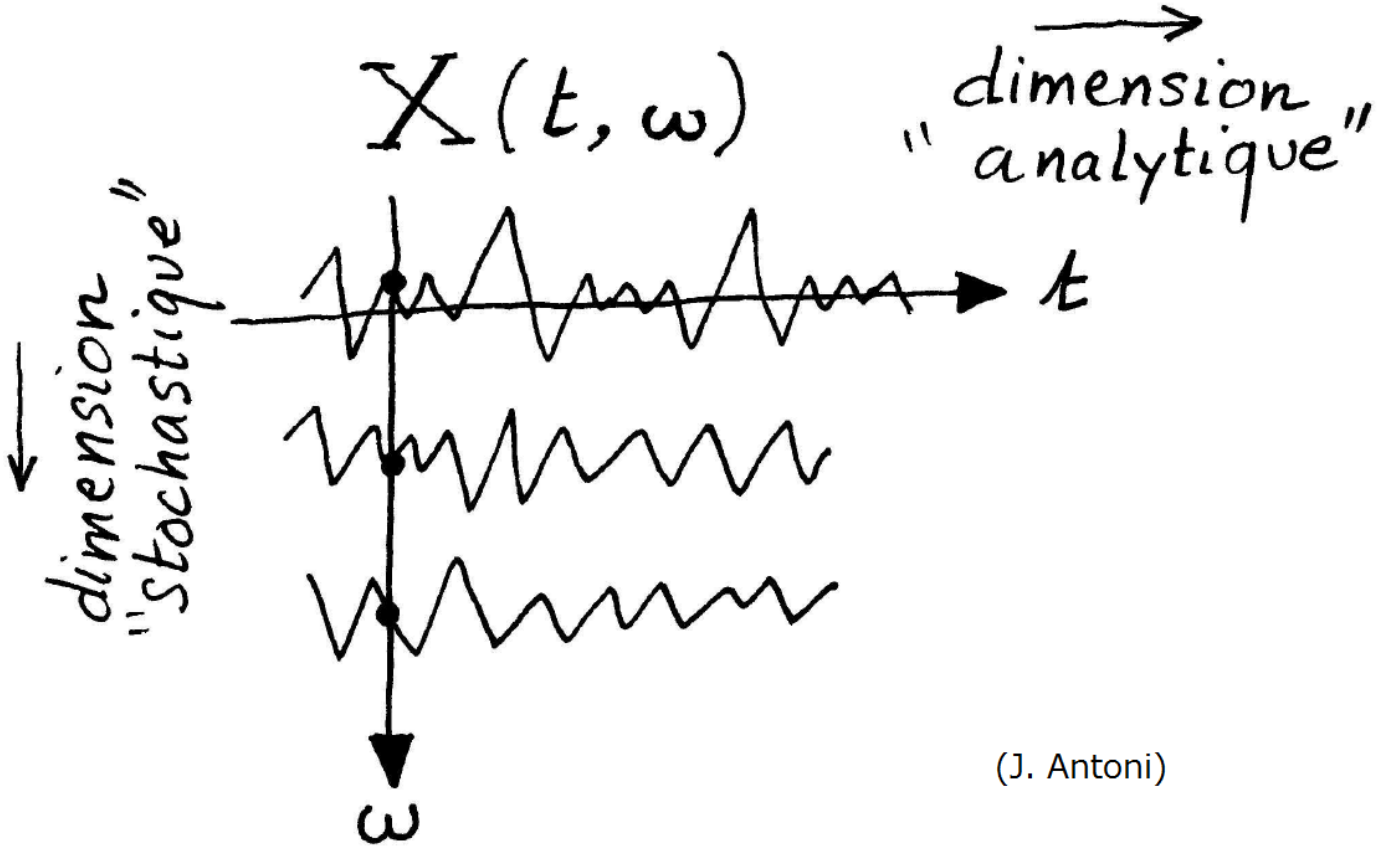
Espace topo-probabiliste




(J. Antoni)

# Interpolation stochastique

Espace topo-probabiliste




(J. Antoni)

# Interpolation stochastique

Le variable régionalisée $z_i$ est alors la réalisation d'une fonction aléatoire $Z_i$

# Interpolation stochastique

Le variable régionalisée $z_i$ est alors la réalisation d'une fonction aléatoire $Z_i$

Dans l'absolu, elle est caractérisée par une suite (infinie) de fonctions de répartition $(F_n)_{n\in\mathbb{N}}$ :

$$F_n : \quad \mathbb{R}^n \longrightarrow [0,1]$$

$$(z_1, z_2, ...z_n) \longrightarrow P(Z_1 \leqslant z_1,\ Z_2 \leqslant z_2 \ ...\ Z_n \leqslant z_n)$$

# Interpolation stochastique

Le variable régionalisée $z_i$ est alors la réalisation d'une fonction aléatoire $Z_i$

Dans l'absolu, elle est caractérisée par une suite (infinie) de fonctions de répartition $(F_n)_{n\in\mathbb{N}}$ :

$$F_n : \quad \mathbb{R}^n \longrightarrow [0,1]$$

$$(z_1, z_2, ...z_n) \longrightarrow P(Z_1 \leqslant z_1,\ Z_2 \leqslant z_2 \ ...\ Z_n \leqslant z_n)$$

Avec seulement une réalisation (de $n$ observations), il est impossible d'avoir une connaissance satisfaisante de la loi de $Z$.

# Interpolation stochastique

Le variable régionalisée $z_i$ est alors la réalisation d'une fonction aléatoire $Z_i$

Dans l'absolu, elle est caractérisée par une suite (infinie) de fonctions de répartition $(F_n)_{n\in\mathbb{N}}$ :

$$F_n : \quad \mathbb{R}^n \longrightarrow [0,1]$$

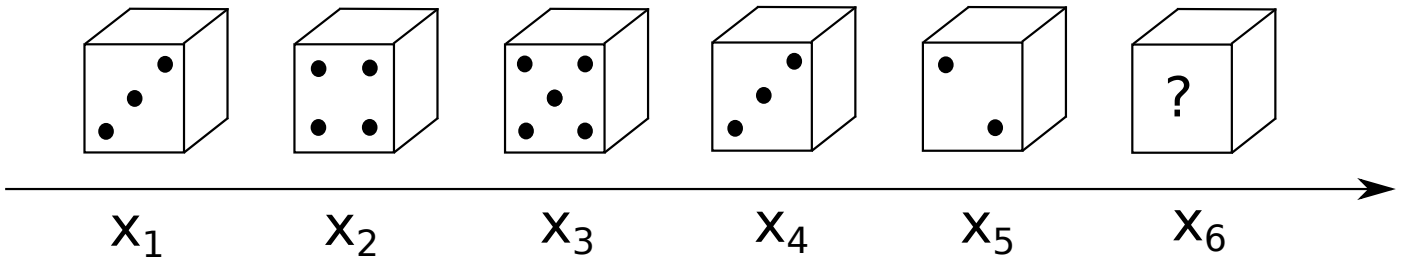
$$(z_1, z_2, ...z_n) \longrightarrow P(Z_1 \leqslant z_1,\ Z_2 \leqslant z_2 \ ...\ Z_n \leqslant z_n)$$

Avec seulement une réalisation (de $n$ observations), il est impossible d'avoir une connaissance satisfaisante de la loi de $Z$.

$\rightarrow$ Modélisation des 2 premiers moments (moyennes et covariances).

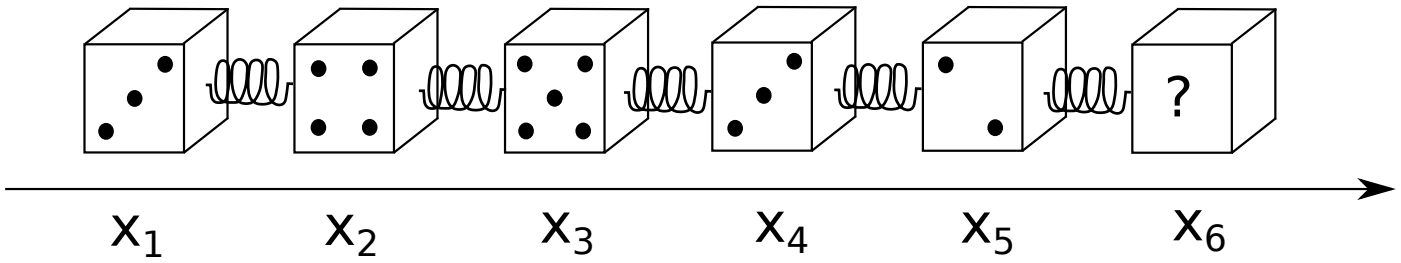
# Interpolation stochastique

(M. Arnaud, X. Emery)



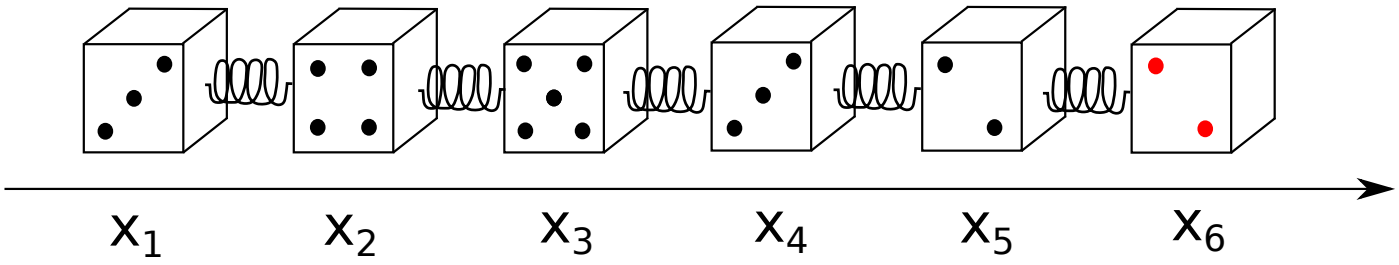
# Interpolation stochastique

(M. Arnaud, X. Emery)



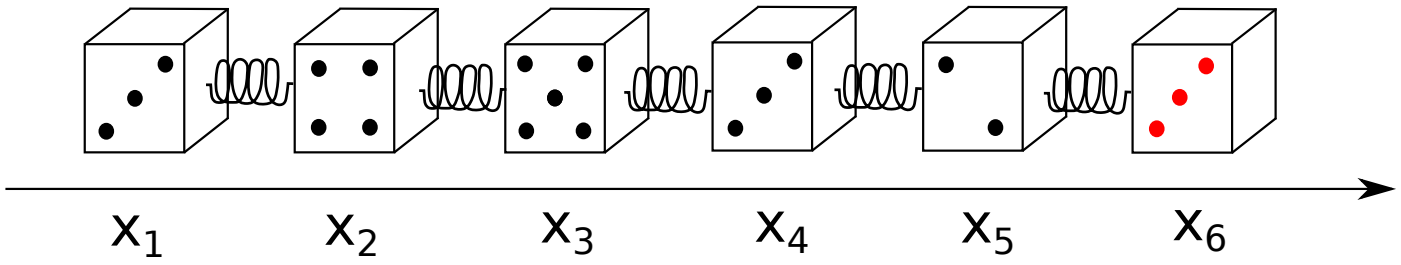
# Interpolation stochastique

(M. Arnaud, X. Emery)



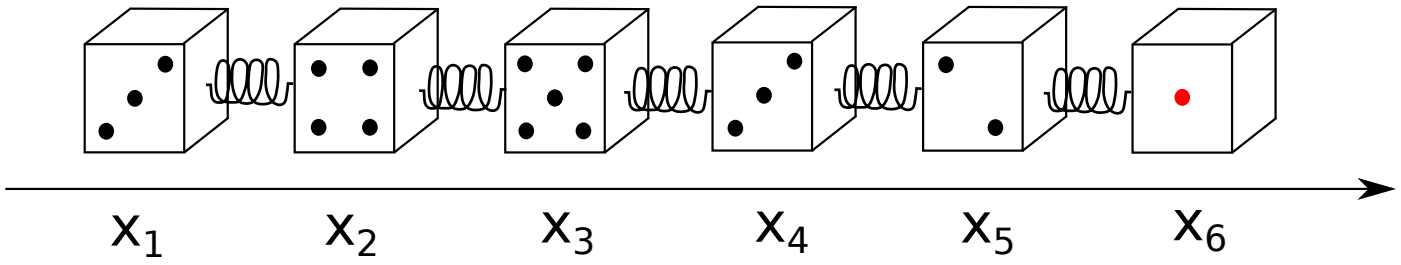
# Interpolation stochastique

(M. Arnaud, X. Emery)



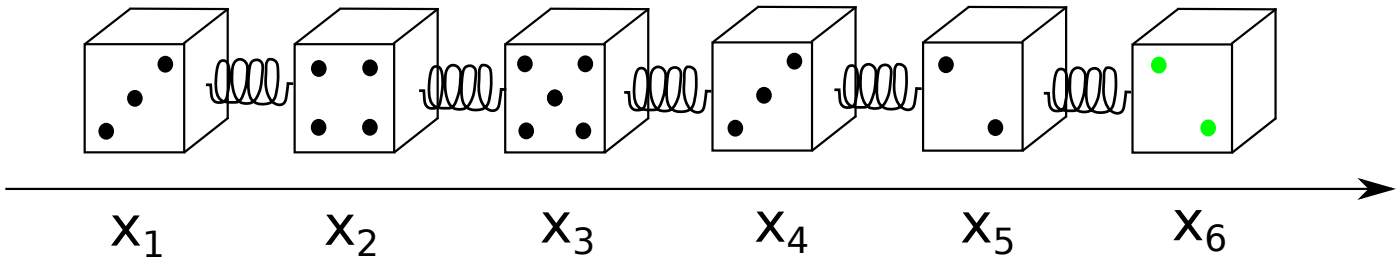
# Interpolation stochastique

(M. Arnaud, X. Emery)



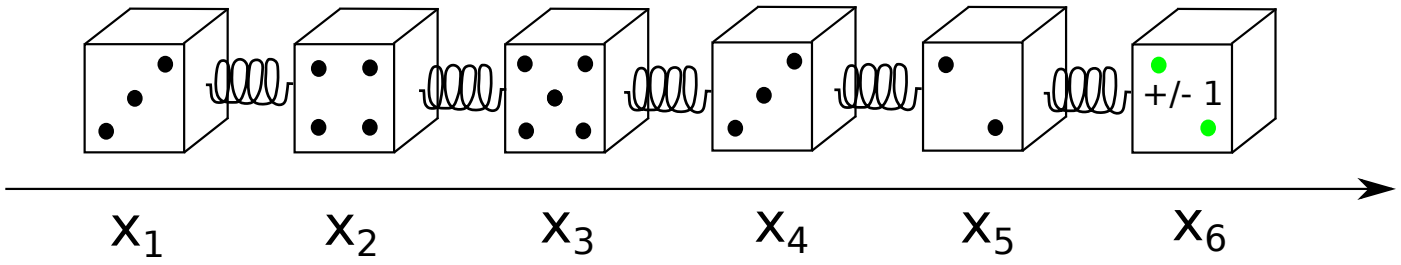
# Interpolation stochastique

(M. Arnaud, X. Emery)



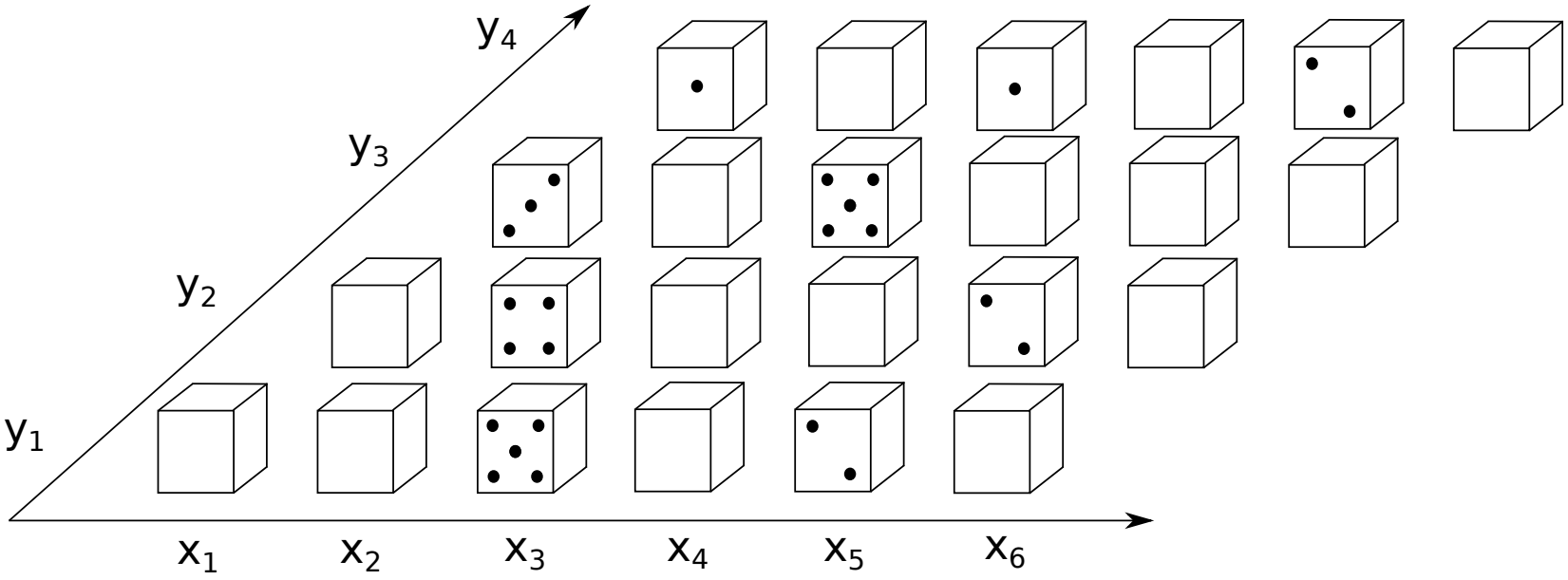
# Interpolation stochastique

(M. Arnaud, X. Emery)



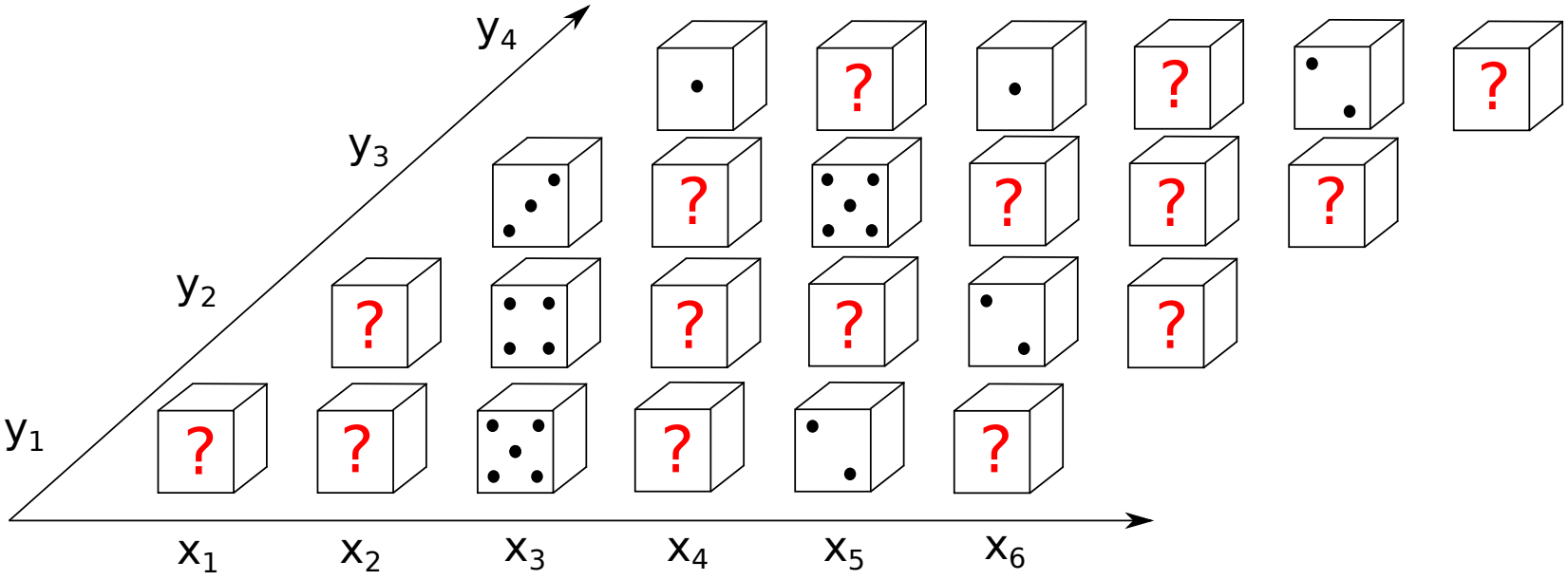
# Interpolation stochastique

(M. Arnaud, X. Emery)



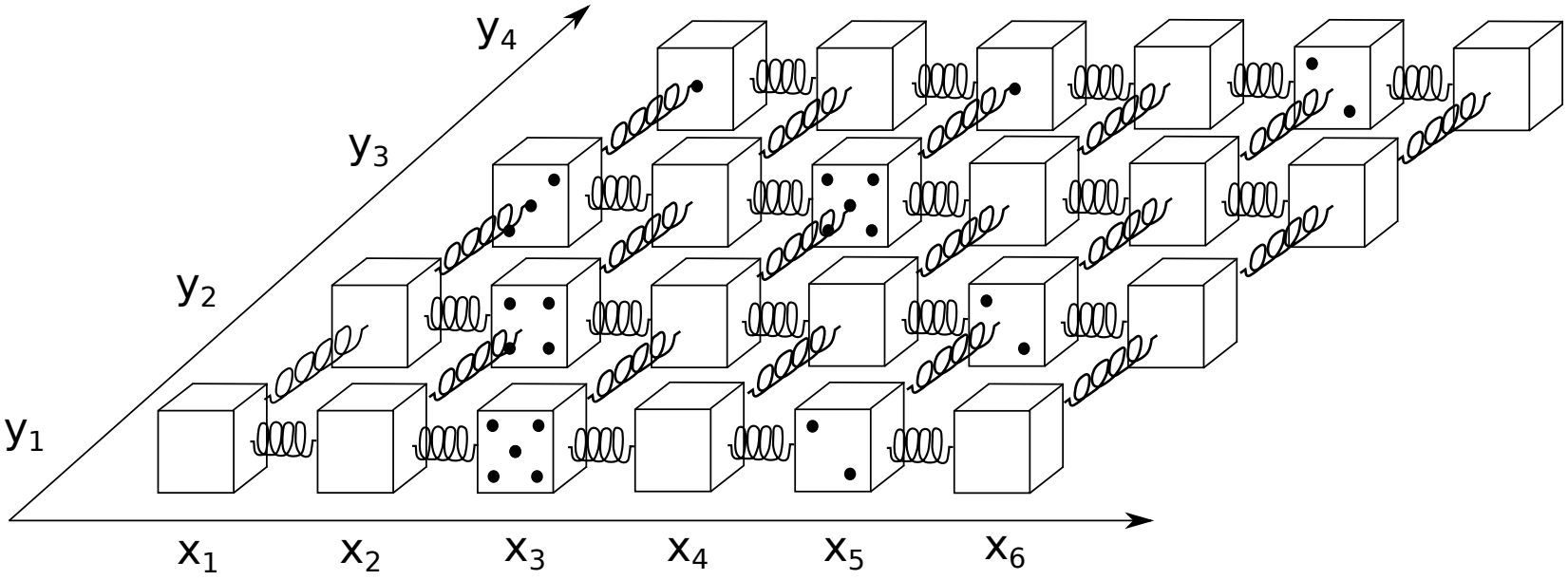
# Interpolation stochastique

(M. Arnaud, X. Emery)



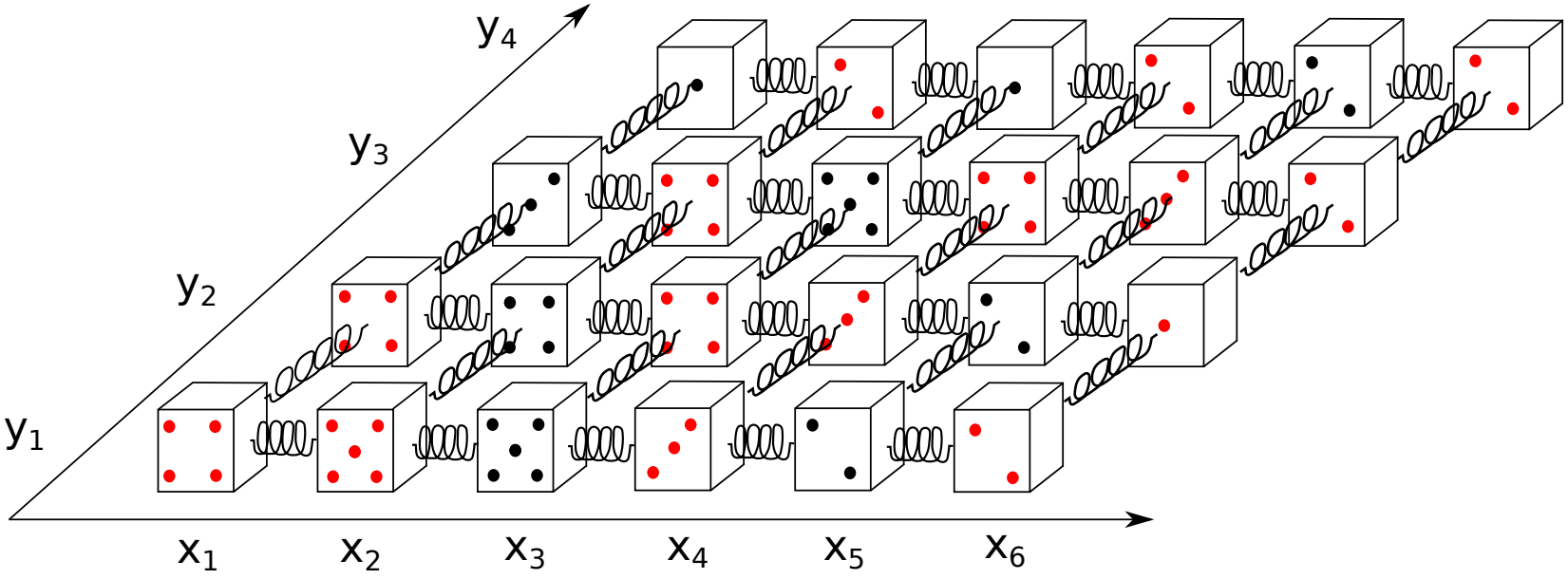
# Interpolation stochastique

(M. Arnaud, X. Emery)



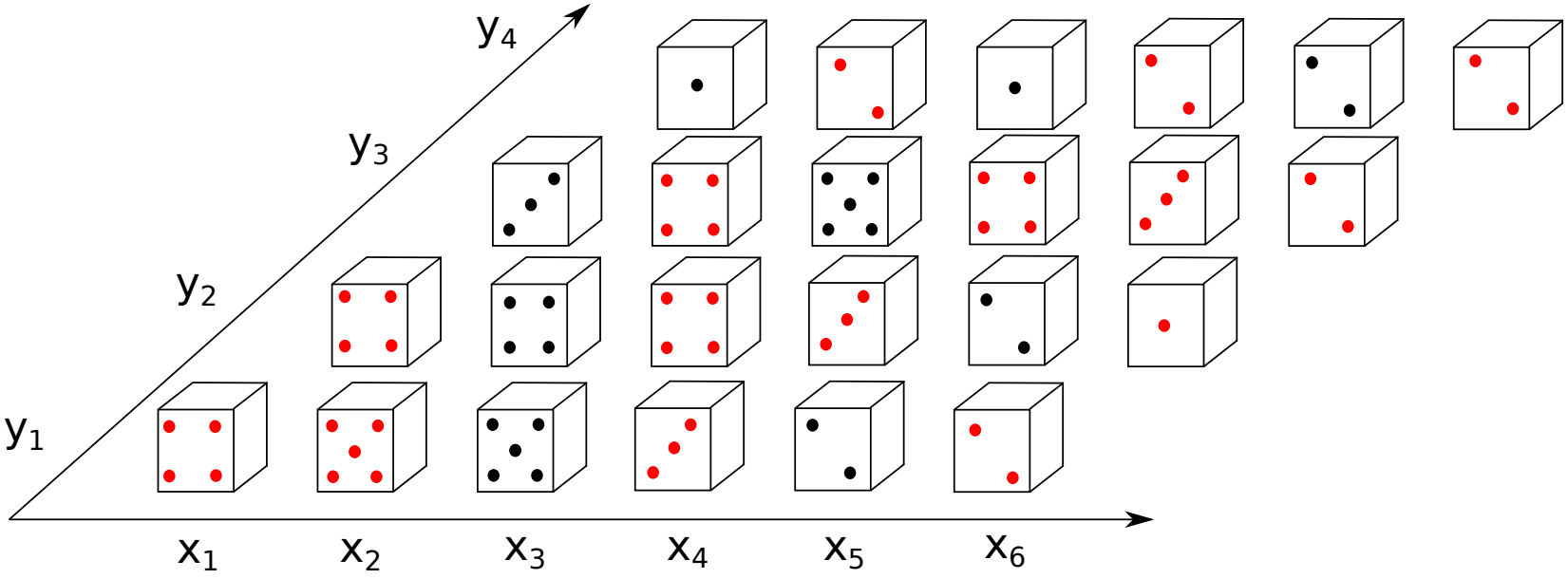
# Interpolation stochastique

(M. Arnaud, X. Emery)



# Interpolation stochastique

(M. Arnaud, X. Emery)



# Analyse variographique

**Inférence statistique :** utilisation de la réalisation, i.e. du semis d'observations, pour estimer les deux premiers moments de la fonction aléatoire. Pour simplifier le problème, on pose quelques hypothèses sur le processus que l'on cherche à modéliser.

# Analyse variographique

**Inférence statistique :** utilisation de la réalisation, i.e. du semis d'observations, pour estimer les deux premiers moments de la fonction aléatoire. Pour simplifier le problème, on pose quelques hypothèses sur le processus que l'on cherche à modéliser.

## Stationnarité

On dit qu'un signal aléatoire est *stationnaire au sens fort* (SSS) lorsque toutes ses lois jointes fini-dimensionnelles sont invariantes par translation $h \in \mathcal{D}$

$$\pi(z(x_1), z(x_2), ...z(x_k)) = \pi(z(x_1 + h), z(x_2 + h), ...z(x_k + h))$$

# Analyse variographique

**Inférence statistique :** utilisation de la réalisation, i.e. du semis d'observations, pour estimer les deux premiers moments de la fonction aléatoire. Pour simplifier le problème, on pose quelques hypothèses sur le processus que l'on cherche à modéliser.

## Stationnarité

On dit qu'un signal aléatoire est *stationnaire au sens fort* (SSS) lorsque toutes ses lois jointes fini-dimensionnelles sont invariantes par translation $h \in \mathcal{D}$

$$\pi(z(x_1), z(x_2), ...z(x_k)) = \pi(z(x_1 + h), z(x_2 + h), ...z(x_k + h))$$

En pratique, cette hypothèse est impossible à vérifier !

# Analyse variographique

## Stationnarité à l'ordre 2

On dit qu'un signal aléatoire est *stationnaire au sens faible* (WSS), ou *faiblement stationnaire*, lorsque que ses moments d'ordre 1 et 2 existent et sont invariants par translation :

(1) $\mathbb{E}[Z(x)] = m$
(2) $\mathrm{Cov}[Z(x_1), Z(x_2)] = \mathrm{Cov}[Z(x_1 + h), Z(x_2 + h)]$

# Analyse variographique

## Stationnarité à l'ordre 2

On dit qu'un signal aléatoire est *stationnaire au sens faible* (WSS), ou *faiblement stationnaire*, lorsque que ses moments d'ordre 1 et 2 existent et sont invariants par translation :

(1) $\mathbb{E}[Z(x)] = m$
(2) $\text{Cov}[Z(x_1), Z(x_2)] = \text{Cov}[Z(x_1 + h), Z(x_2 + h)]$

La condition (2) implique que la covariance ne dépend que du vecteur séparant les sites : $C(Z(x_1), Z(x_2)) = f(x_1 - x_2)$

# Analyse variographique

## Stationnarité à l'ordre 2

On dit qu'un signal aléatoire est *stationnaire au sens faible* (WSS), ou *faiblement stationnaire*, lorsque que ses moments d'ordre 1 et 2 existent et sont invariants par translation :

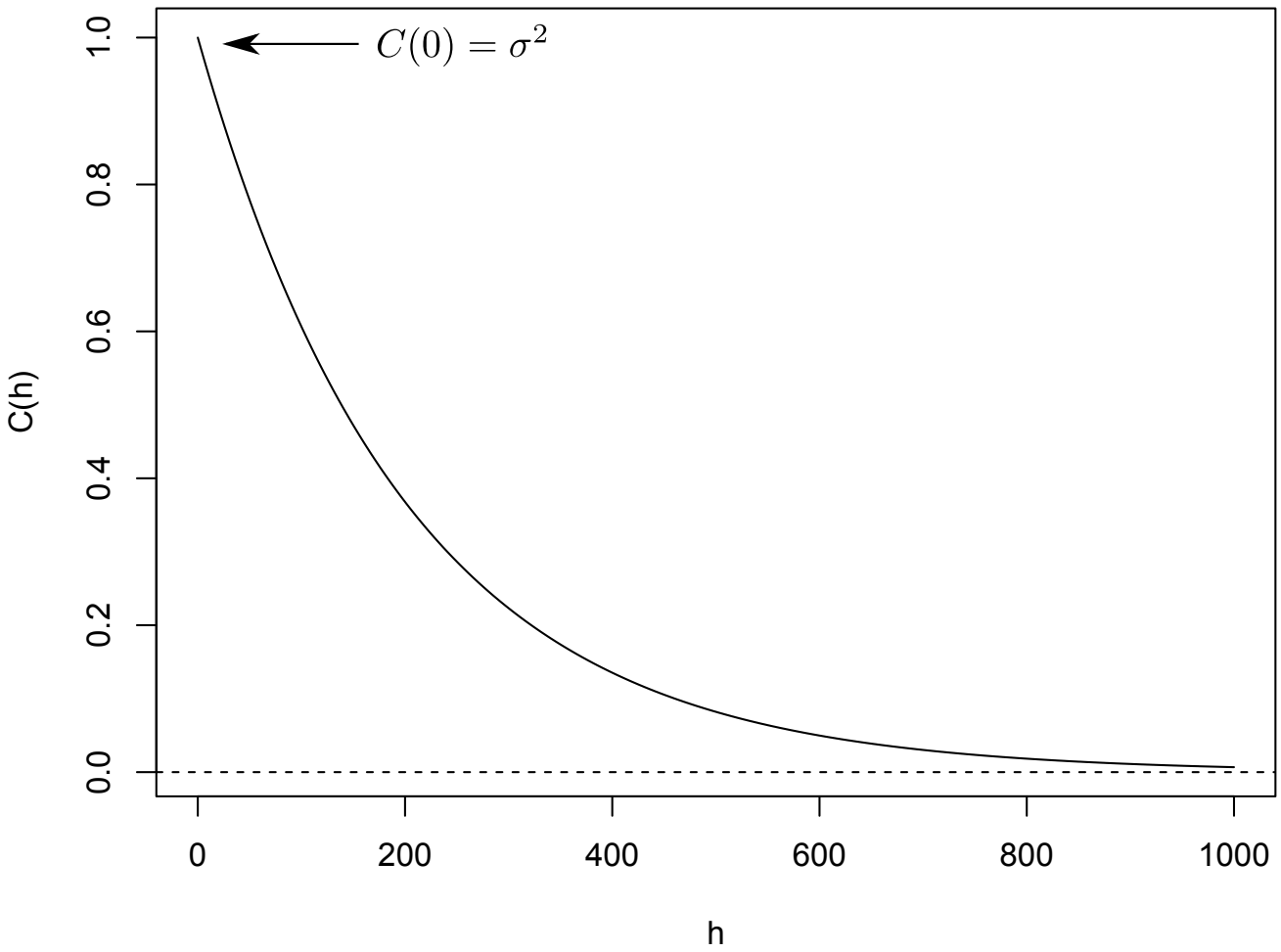
(1) $\mathbb{E}[Z(x)] = m$
(2) $\text{Cov}[Z(x_1), Z(x_2)] = \text{Cov}[Z(x_1 + h), Z(x_2 + h)]$

La condition (2) implique que la covariance ne dépend que du vecteur séparant les sites : $C(Z(x_1), Z(x_2)) = f(x_1 - x_2)$

En pratique, on stipule en plus que le phénomène est isotrope. Le covariogramme est une fonction de $\mathbb{R}^+$ dans $\mathbb{R}^+$ ne dépendant que de la distance séparant les deux sites : $C(Z_i, Z_j) = C(h_{ij})$.

# Analyse variographique

## Exemple de covariogramme



# Analyse variographique

Soit $z : \mathcal{D} \times \Omega \to \mathbb{R}$ un signal aléatoire stationnaire au sens fort

# Analyse variographique

Soit $z : \mathcal{D} \times \Omega \to \mathbb{R}$ un signal aléatoire stationnaire au sens fort

## Ergodicité (au sens fort)

L'hypothèse d'**ergodicité** assure que les moyennes spatiales du processus sont identiques aux moyennes statistiques. Pour toute fonction continue $g : \mathbb{R} \to \mathbb{R}$

$$\frac{1}{\mathcal{D}} \int_{\mathcal{D}} g(z(x,\omega))dx \underset{|\mathcal{D}| \to +\infty}{\longrightarrow} \int_{\Omega} g(Z(x,\omega))dP(\omega) = \mathbb{E}[g(Z)]$$

# Analyse variographique

Soit $z : \mathcal{D} \times \Omega \to \mathbb{R}$ un signal aléatoire stationnaire au sens fort

## Ergodicité (au sens fort)

L'hypothèse d'**ergodicité** assure que les moyennes spatiales du processus sont identiques aux moyennes statistiques. Pour toute fonction continue $g : \mathbb{R} \to \mathbb{R}$

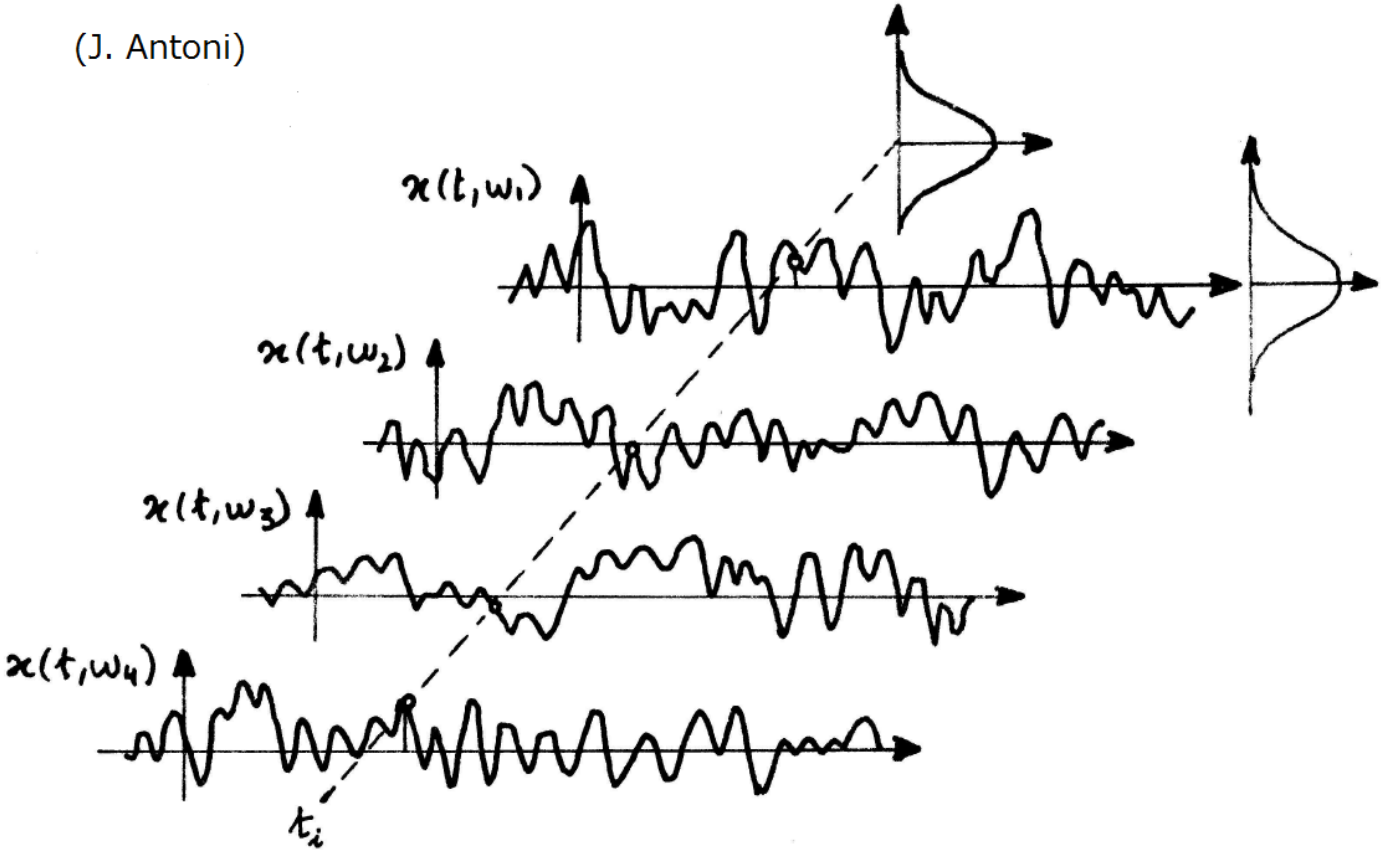
$$\frac{1}{\mathcal{D}} \int_{\mathcal{D}} g(z(x,\omega))dx \underset{|\mathcal{D}| \to +\infty}{\longrightarrow} \int_{\Omega} g(Z(x,\omega))dP(\omega) = \mathbb{E}[g(Z)]$$

Par suite, la distribution de probabilités de toutes les variables aléatoires $Z(x_i)$ peut être estimée à partir de la distribution des valeurs prises par une unique réalisation $\omega \in \Omega$ du processus.

# Analyse variographique

## Hypothèse d'ergodicité

(J. Antoni)



# Analyse variographique

Soit $z : \mathcal{D} \times \Omega \to \mathbb{R}$ un signal aléatoire stationnaire à l'ordre 1

## Ergodicité à l'ordre 1

L'hypothèse d'**ergodicité** à l'ordre 1 assure que la moyenne spatiale du processus est identique à la moyenne statistique :

$$\frac{1}{\mathcal{D}} \int_{\mathcal{D}} z(x,\omega)dx \underset{|\mathcal{D}| \to +\infty}{\longrightarrow} \int_{\Omega} Z(x,\omega)dP(\omega) = \mathbb{E}[Z]$$

La convergence est à comprendre au sens de la convergence presque sûre, i.e. la moyenne spatiale d'une réalisation converge presque sûrement (avec une probabilité égale à $1$) vers l'espérance de la valeur prise par la fonction aléatoire en chaque site.

# Analyse variographique

Soit $z : \mathcal{D} \times \Omega \to \mathbb{R}$ un signal aléatoire stationnaire à l'ordre 1

## Ergodicité à l'ordre 1

L'hypothèse d'**ergodicité** à l'ordre 1 assure que la moyenne spatiale du processus est identique à la moyenne statistique :

$$\frac{1}{\mathcal{D}} \int_{\mathcal{D}} z(x,\omega)dx \underset{|\mathcal{D}| \to +\infty}{\longrightarrow} \int_{\Omega} Z(x,\omega)dP(\omega) = \mathbb{E}[Z]$$

La convergence est à comprendre au sens de la convergence presque sûre, i.e. la moyenne spatiale d'une réalisation converge presque sûrement (avec une probabilité égale à $1$) vers l'espérance de la valeur prise par la fonction aléatoire en chaque site.

Un processus ergodique à l'ordre 1 est dit *à moyenne ergodique*.

# Analyse variographique

Soit $z : \mathcal{D} \times \Omega \to \mathbb{R}$ un signal aléatoire stationnaire à l'ordre 2

# Analyse variographique

Soit $z : \mathcal{D} \times \Omega \to \mathbb{R}$ un signal aléatoire stationnaire à l'ordre 2

## Ergodicité à l'ordre 2 : processus à covariance ergodique

L'hypothèse d'**ergodicité** à l'ordre 2 assure que la covariance spatiale du processus est identique à la covariance statistique :

$$\frac{1}{\mathcal{D}} \int_{\mathcal{D}} z(x+h,\omega) z(x,\omega) dx \underset{|\mathcal{D}| \to +\infty}{\longrightarrow} \mathbb{E}[Z(x+h)Z(x)]$$

# Analyse variographique

Soit $z : \mathcal{D} \times \Omega \to \mathbb{R}$ un signal aléatoire stationnaire à l'ordre 2

## Ergodicité à l'ordre 2 : processus à covariance ergodique

L'hypothèse d'**ergodicité** à l'ordre 2 assure que la covariance spatiale du processus est identique à la covariance statistique :

$$\frac{1}{\mathcal{D}} \int_{\mathcal{D}} z(x+h,\omega) z(x,\omega) dx \underset{|\mathcal{D}| \to +\infty}{\longrightarrow} \mathbb{E}[Z(x+h)Z(x)]$$

Un signal ergodique à l'ordre $k$ est nécessairement stationnaire à l'ordre $k$. La réciproque n'est pas vraie en général. Exemple : le signal aléatoire constant $Z(x,\omega) = A(\omega)$ est stationnaire au sens fort mais n'est pas ergodique (sauf si $A$ est constant presque sûrement).

# Analyse variographique

Exemples (signaux temporels 1D) : Stationnarité ? Ergodicité ?

1. $X_t = w_t$ avec $w_t \sim \mathcal{N}(\mu, \sigma)$ : bruit blanc gaussien

# Analyse variographique

Exemples (signaux temporels 1D) : Stationnarité ? Ergodicité ?

1. $X_t = w_t$ avec $w_t \sim \mathcal{N}(\mu, \sigma)$ : bruit blanc gaussien
2. $X_{t+1} = X_t + kw_t + b$ avec $(k, b) \in \mathbb{R}^2$ et $w_t \sim \mathcal{N}(0, 1)$ : marche aléatoire (à temps discret sur l'ensemble des réels)

# Analyse variographique

Exemples (signaux temporels 1D) : Stationnarité ? Ergodicité ?

1. $X_t = w_t$ avec $w_t \sim \mathcal{N}(\mu, \sigma)$ : bruit blanc gaussien
2. $X_{t+1} = X_t + kw_t + b$ avec $(k, b) \in \mathbb{R}^2$ et $w_t \sim \mathcal{N}(0, 1)$ : marche aléatoire (à temps discret sur l'ensemble des réels)
3. $X_t = \rho X_{t-1} + c + w_t$ avec $w_t \sim \mathcal{N}(0, 1)$ : processus AR(1)

# Analyse variographique

Exemples (signaux temporels 1D) : Stationnarité ? Ergodicité ?

1. $X_t = w_t$ avec $w_t \sim \mathcal{N}(\mu, \sigma)$ : bruit blanc gaussien
2. $X_{t+1} = X_t + kw_t + b$ avec $(k, b) \in \mathbb{R}^2$ et $w_t \sim \mathcal{N}(0, 1)$ : marche aléatoire (à temps discret sur l'ensemble des réels)
3. $X_t = \rho X_{t-1} + c + w_t$ avec $w_t \sim \mathcal{N}(0, 1)$ : processus AR(1)
4. $X_t = A\cos(wt)$, $Y_t = B\cos(wt)$ et $Z_t = X_t + Y_t$ avec $w \in \mathbb{R}$ et $A, B$ deux variables aléatoires i.i.d $\sim \mathcal{N}(0, 1)$

# Analyse variographique

Exemples (signaux temporels 1D) : Stationnarité ? Ergodicité ?

1. $X_t = w_t$ avec $w_t \sim \mathcal{N}(\mu, \sigma)$ : bruit blanc gaussien
2. $X_{t+1} = X_t + kw_t + b$ avec $(k, b) \in \mathbb{R}^2$ et $w_t \sim \mathcal{N}(0, 1)$ : marche aléatoire (à temps discret sur l'ensemble des réels)
3. $X_t = \rho X_{t-1} + c + w_t$ avec $w_t \sim \mathcal{N}(0, 1)$ : processus AR(1)
4. $X_t = A\cos(wt)$, $Y_t = B\cos(wt)$ et $Z_t = X_t + Y_t$ avec $w \in \mathbb{R}$ et $A, B$ deux variables aléatoires i.i.d $\sim \mathcal{N}(0, 1)$
5. $X_t = r\cos(wt + \phi)$, avec $(w, \phi) \in \mathbb{R}^2$ et $r \sim \mathcal{N}(0, 1)$

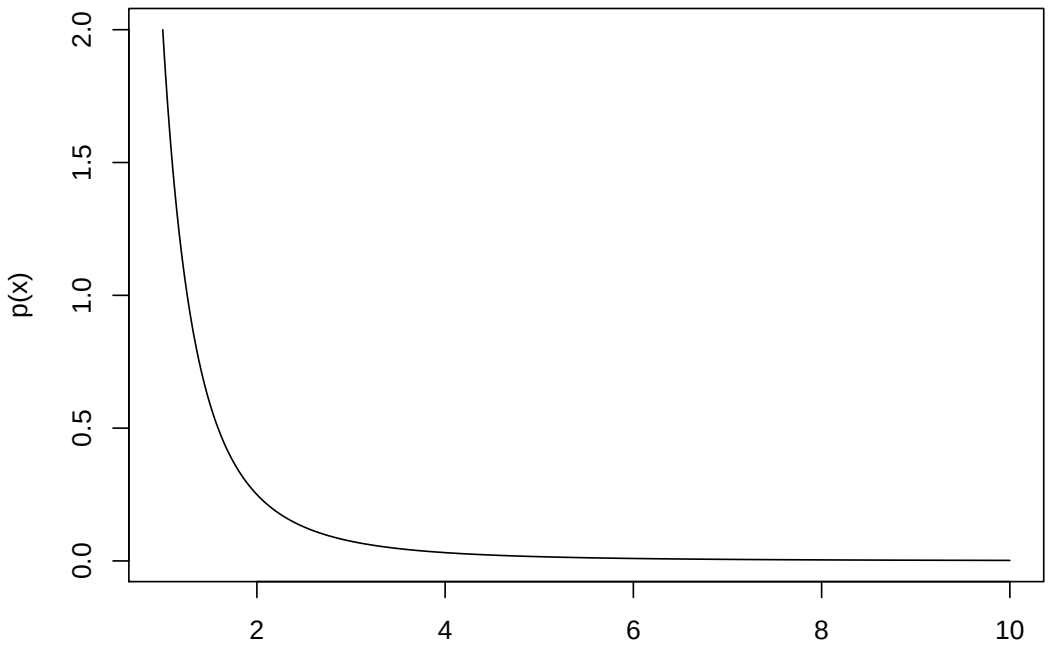
# Analyse variographique

Exemples (signaux temporels 1D) : Stationnarité ? Ergodicité ?

1. $X_t = w_t$ avec $w_t \sim \mathcal{N}(\mu, \sigma)$ : bruit blanc gaussien
2. $X_{t+1} = X_t + kw_t + b$ avec $(k, b) \in \mathbb{R}^2$ et $w_t \sim \mathcal{N}(0, 1)$ : marche aléatoire (à temps discret sur l'ensemble des réels)
3. $X_t = \rho X_{t-1} + c + w_t$ avec $w_t \sim \mathcal{N}(0, 1)$ : processus AR(1)
4. $X_t = A\cos(wt)$, $Y_t = B\cos(wt)$ et $Z_t = X_t + Y_t$ avec $w \in \mathbb{R}$ et $A, B$ deux variables aléatoires i.i.d $\sim \mathcal{N}(0, 1)$
5. $X_t = r\cos(wt + \phi)$, avec $(w, \phi) \in \mathbb{R}^2$ et $r \sim \mathcal{N}(0, 1)$
6. $X_t = r\cos(wt + \psi)$ avec $(r, w) \in \mathbb{R}^2$ et $\psi \sim \mathcal{U}[-\pi, \pi]$

# Analyse variographique

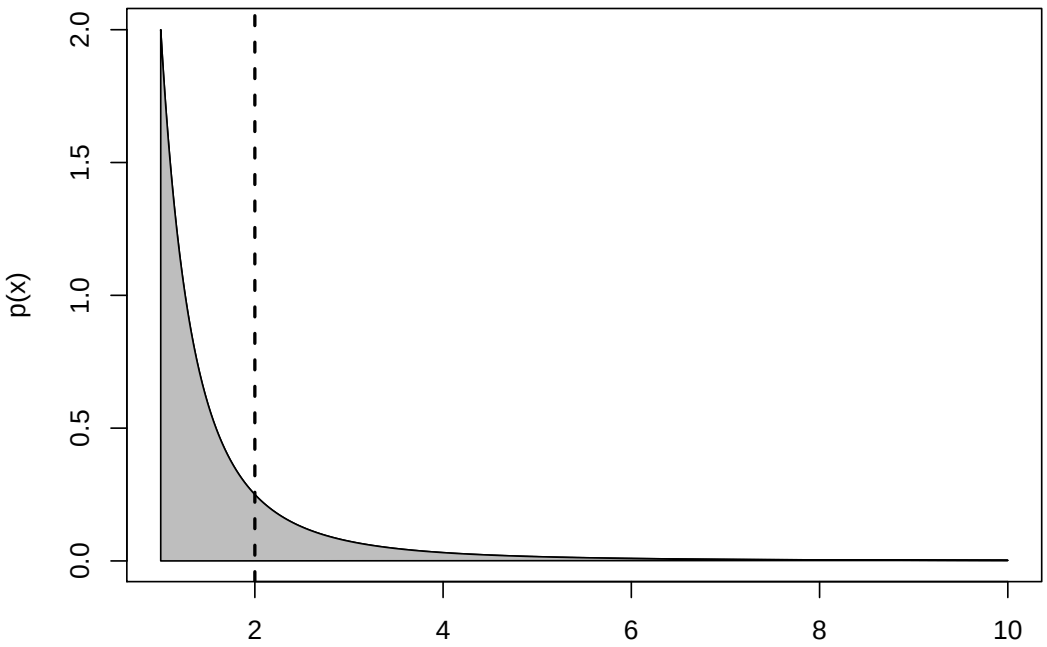
**Problème :** la fonction de covariance n'existe pas nécessairement.
Exemple : la distribution de probabilité $\pi(x) = 2/x^3 \; 1_{x \geqslant 1}$.




Montrer que $\pi$ est bien une densité de probabilité et que son espérance est finie. Quid de sa variance ?

# Analyse variographique

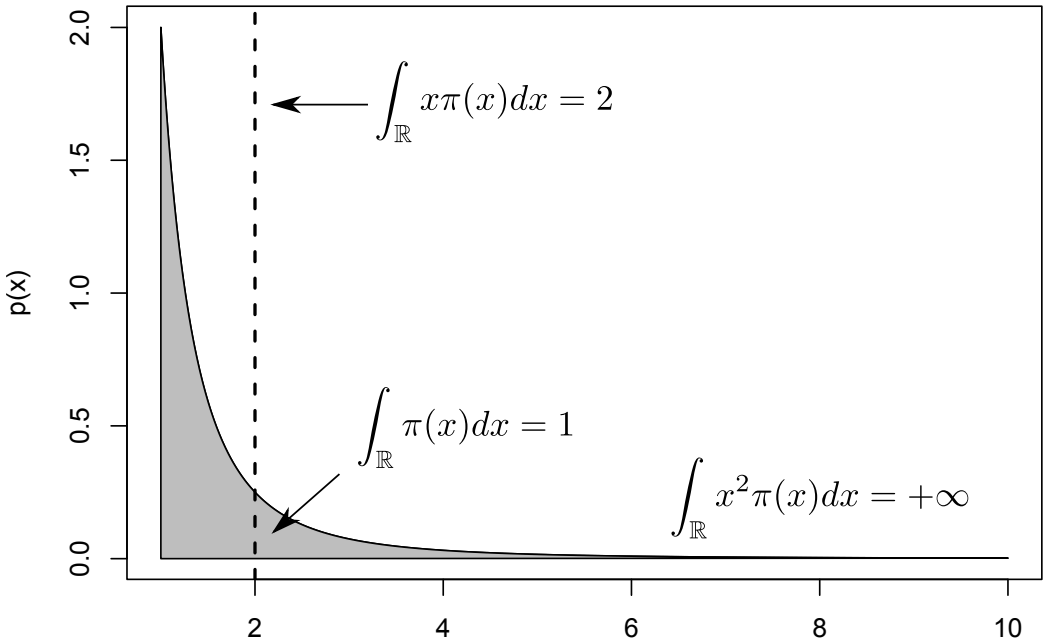
**Problème :** la fonction de covariance n'existe pas nécessairement.
Exemple : la distribution de probabilité $\pi(x) = 2/x^3 \; 1_{x \geqslant 1}$.




Montrer que $\pi$ est bien une densité de probabilité et que son espérance est finie. Quid de sa variance ?

# Analyse variographique

**Problème :** la fonction de covariance n'existe pas nécessairement.
Exemple : la distribution de probabilité $\pi(x) = 2/x^3 \; 1_{x \geqslant 1}$.




Montrer que $\pi$ est bien une densité de probabilité et que son espérance est finie. Quid de sa variance ?

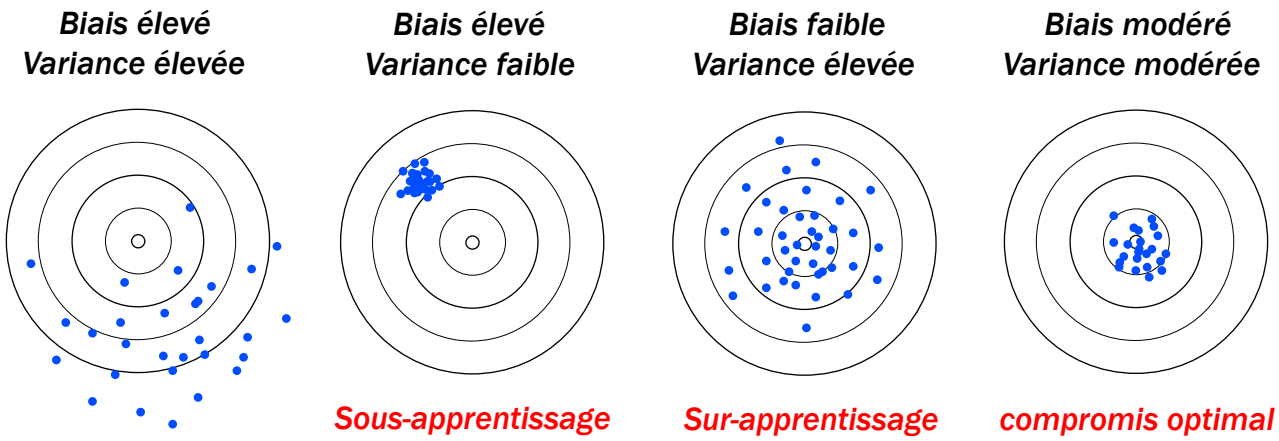
# Analyse variographique

Etude des *accroissements* : $\gamma(x_1, x_2) = V[Z(x_1) - Z(x_2)]$

# Analyse variographique

Etude des *accroissements* : $\gamma(x_1, x_2) = V[Z(x_1) - Z(x_2)]$

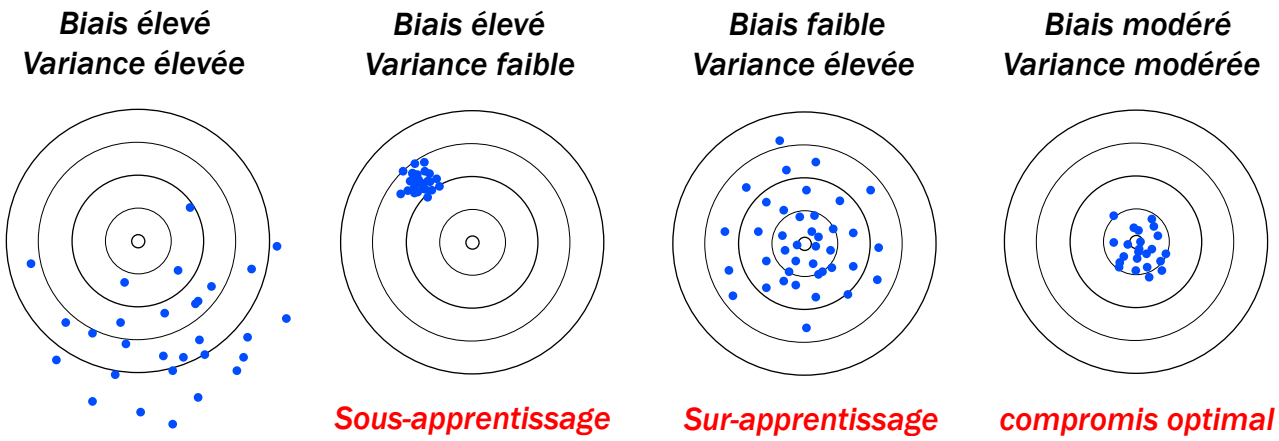
$$\mathbb{E}[X^2] = V[X] + \mathbb{E}[X]^2$$



# Analyse variographique

Etude des *accroissements* : $\gamma(x_1, x_2) = V[Z(x_1) - Z(x_2)]$

$$\mathbb{E}[X^2] = V[X] + \mathbb{E}[X]^2$$

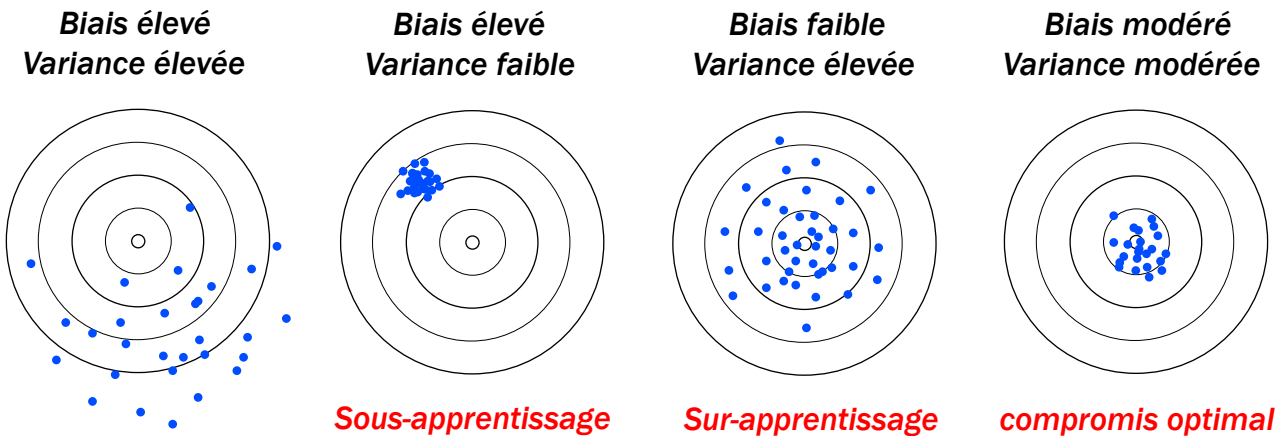



$$\mathbb{E}\Big[Z(x_1) - Z(x_2)\Big] = 0$$

# Analyse variographique

Etude des *accroissements* : $\gamma(x_1, x_2) = V[Z(x_1) - Z(x_2)]$

$$\mathbb{E}[X^2] = V[X] + \mathbb{E}[X]^2$$




$$\mathbb{E}\Big[Z(x_1) - Z(x_2)\Big] = 0 \quad \Rightarrow \quad \gamma(x_1, x_2) = \mathbb{E}\Big[\Big(Z(x_1) - Z(x_2)\Big)^2\Big]$$

# Analyse variographique

On utilise un objet plus générique que la fonction de covariance :

## Variogramme (ou semi-variogramme)

Pour deux sites donnés $x_1$ et $x_2$, le variogramme mesure la dissemblance des valeurs prises par $Z$ entre ces deux sites :

$$\forall\ (x_1, x_2) \in \mathcal{D}^2 \quad \gamma(x_1, x_2) = \frac{1}{2}\mathsf{Var}[Z(x_1) - Z(x_2)]$$

# Analyse variographique

On utilise un objet plus générique que la fonction de covariance :

## Variogramme (ou semi-variogramme)

Pour deux sites donnés $x_1$ et $x_2$, le variogramme mesure la dissemblance des valeurs prises par $Z$ entre ces deux sites :

$$\forall\ (x_1, x_2) \in \mathcal{D}^2 \quad \gamma(x_1, x_2) = \frac{1}{2}\mathsf{Var}[Z(x_1) - Z(x_2)]$$

Sous les hypothèses de **stationnarité** d'ordre 2 et d'**isotropie**, $\gamma$ ne dépend que des distances séparant les sites : $\gamma : \mathbb{R}^+ \to \mathbb{R}^+$. De plus, sous l'hypothèse d'**ergodicité**, on peut estimer $\gamma$ à partir d'une seule réalisation du processus stochastique.

# Analyse variographique

On utilise un objet plus générique que la fonction de covariance :

**Variogramme (ou semi-variogramme)**

Pour deux sites donnés $x_1$ et $x_2$, le variogramme mesure la dissemblance des valeurs prises par $Z$ entre ces deux sites :

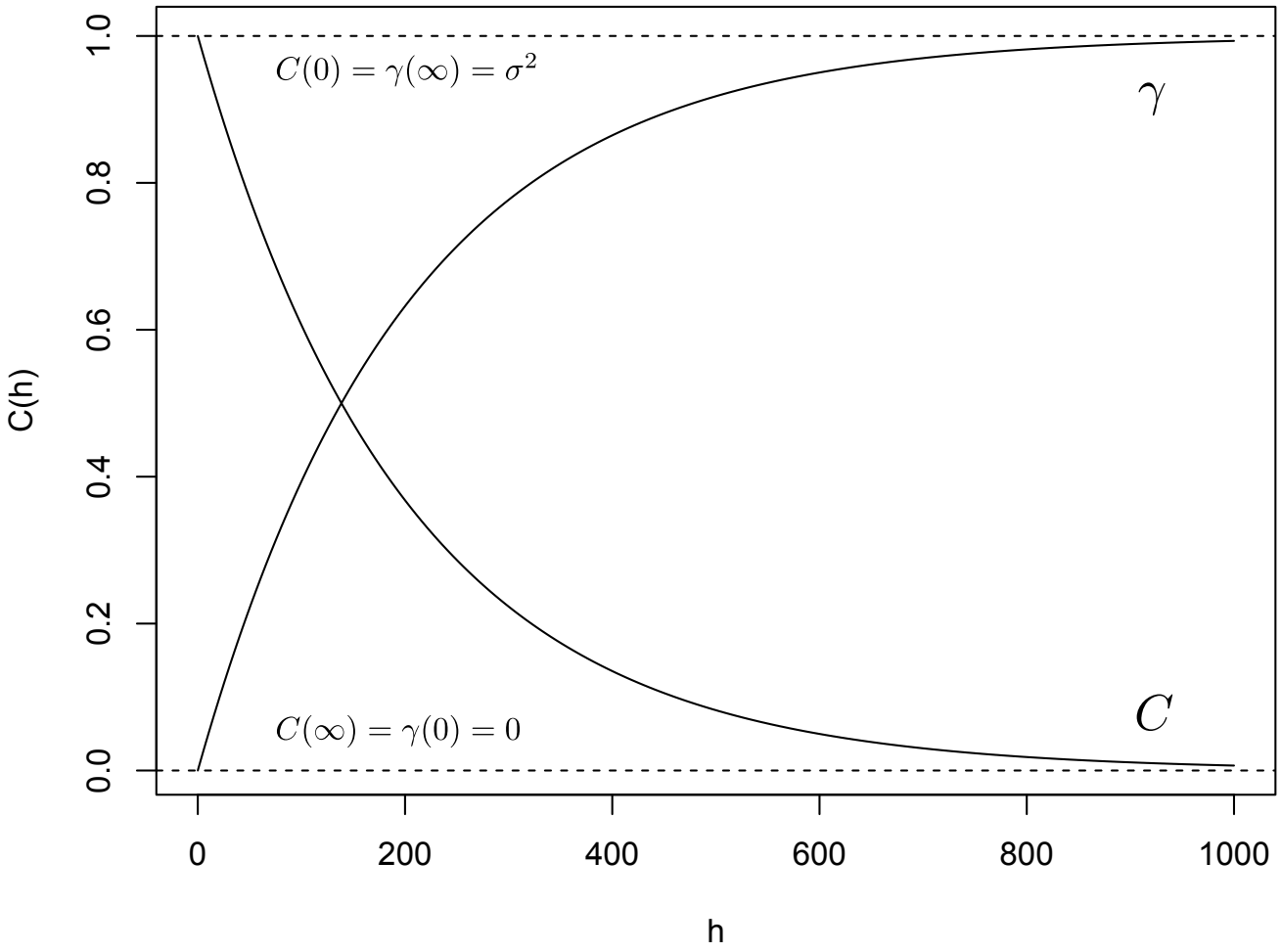
$$\forall\ (x_1, x_2) \in \mathcal{D}^2 \quad \gamma(x_1, x_2) = \frac{1}{2}\mathsf{Var}[Z(x_1) - Z(x_2)]$$

Sous les hypothèses de **stationnarité** d'ordre 2 et d'**isotropie**, $\gamma$ ne dépend que des distances séparant les sites : $\gamma : \mathbb{R}^+ \to \mathbb{R}^+$. De plus, sous l'hypothèse d'**ergodicité**, on peut estimer $\gamma$ à partir d'une seule réalisation du processus stochastique.

$\to$ Plus $\gamma(x_1, x_2)$ est grand, plus les deux sites *ont des chances* de prendre des valeurs $Z$ très différentes l'une de l'autre*.

# Analyse variographique

Exemple de couple variogramme / covariogramme



# Analyse variographique

Quelques remarques sur le variogramme :

# Analyse variographique

Quelques remarques sur le variogramme :

1. Le variogramme est plus générique que la covariance : $\gamma$ peut exister même lorsque $C$ n'est pas définie (on parle de stationnarité intrinsèque). L'inverse n'est pas vrai.

# Analyse variographique

Quelques remarques sur le variogramme :

1. Le variogramme est plus générique que la covariance : $\gamma$ peut exister même lorsque $C$ n'est pas définie (on parle de stationnarité intrinsèque). L'inverse n'est pas vrai.
2. Lorsque variogramme et covariance existent tous les deux et sont connus, on obtient théoriquement les mêmes résultats indépendamment de la fonction utilisée dans les calculs.

# Analyse variographique

Quelques remarques sur le variogramme :

1. Le variogramme est plus générique que la covariance : $\gamma$ peut exister même lorsque $C$ n'est pas définie (on parle de stationnarité intrinsèque). L'inverse n'est pas vrai.
2. Lorsque variogramme et covariance existent tous les deux et sont connus, on obtient théoriquement les mêmes résultats indépendamment de la fonction utilisée dans les calculs.
3. En pratique, l'estimateur de la covariance est plus robuste mais biaisé (sensibilité à la moyenne inconnue du processus).

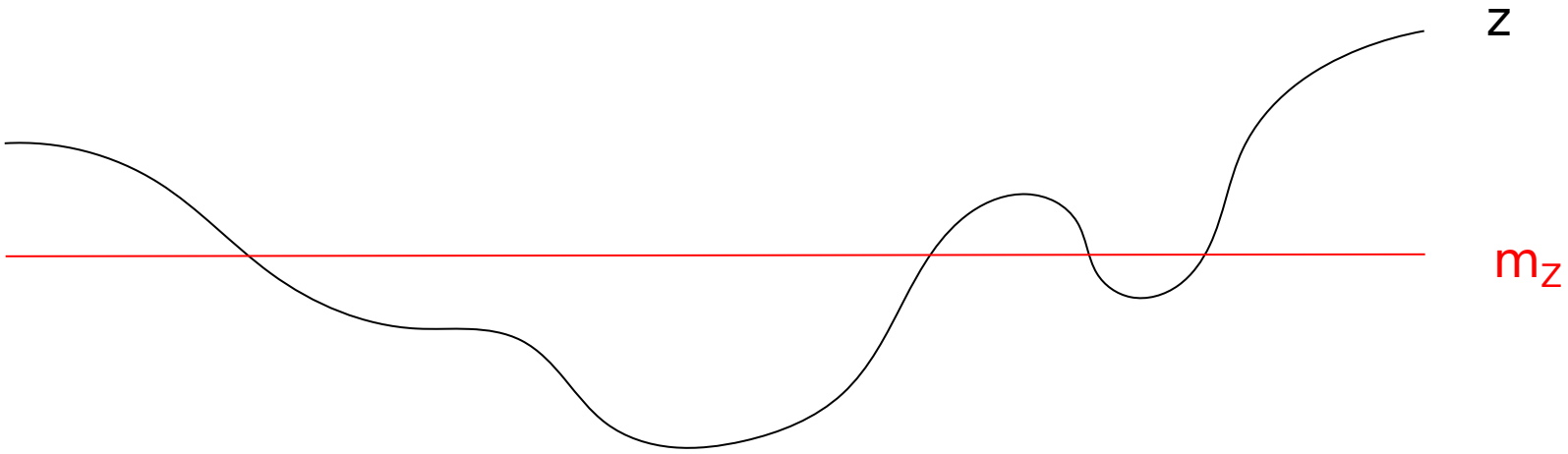
# Analyse variographique

Quelques remarques sur le variogramme :

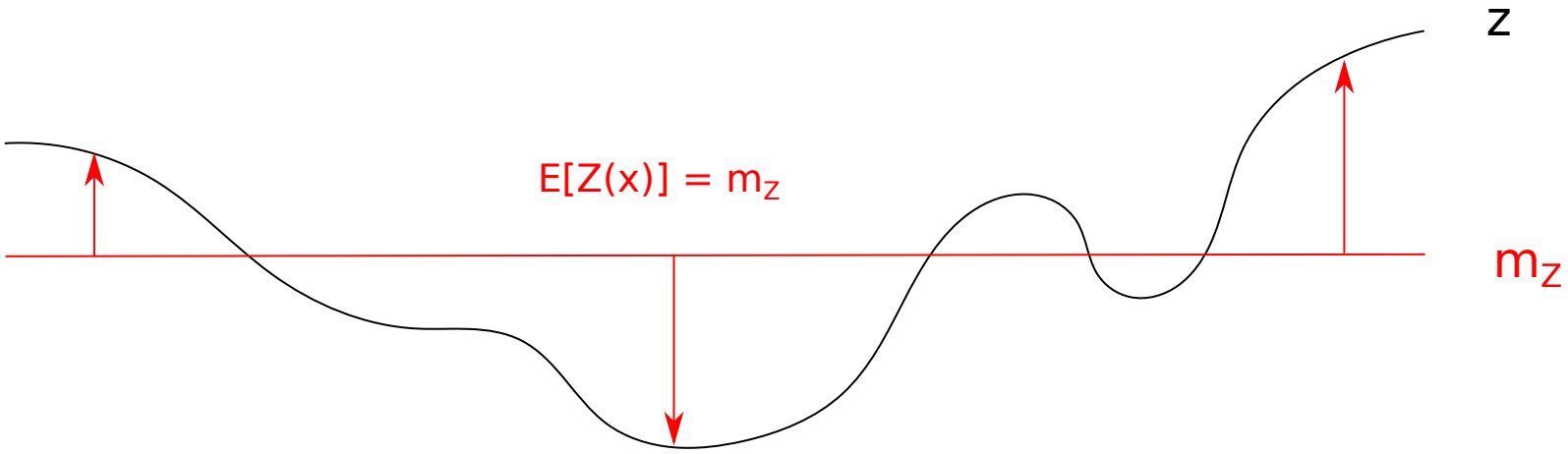
1. Le variogramme est plus générique que la covariance : $\gamma$ peut exister même lorsque $C$ n'est pas définie (on parle de stationnarité intrinsèque). L'inverse n'est pas vrai.
2. Lorsque variogramme et covariance existent tous les deux et sont connus, on obtient théoriquement les mêmes résultats indépendamment de la fonction utilisée dans les calculs.
3. En pratique, l'estimateur de la covariance est plus robuste mais biaisé (sensibilité à la moyenne inconnue du processus).

Par souci de généricité et d'efficacité statistique dans la phase d'inférence, on utilise le variogramme.

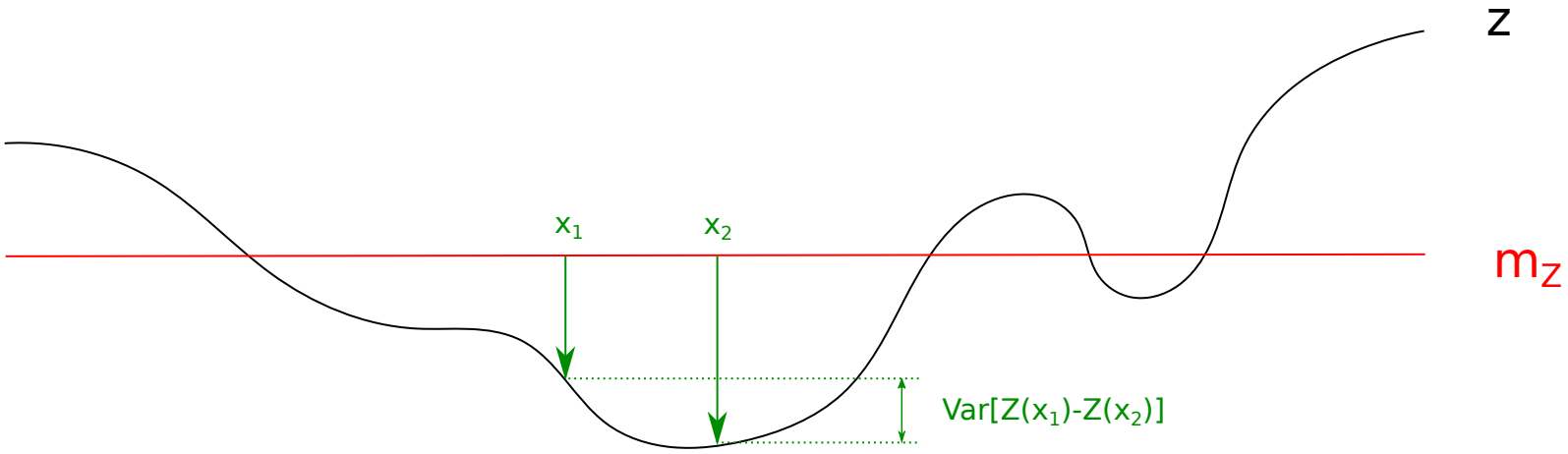
# Analyse variographique



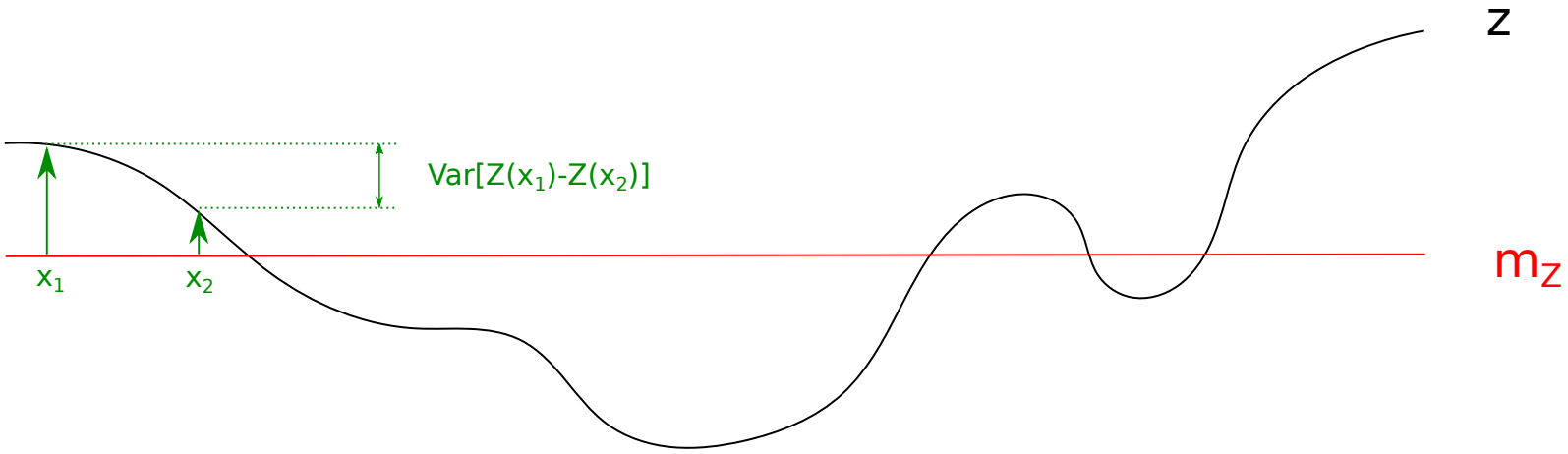
# Analyse variographique



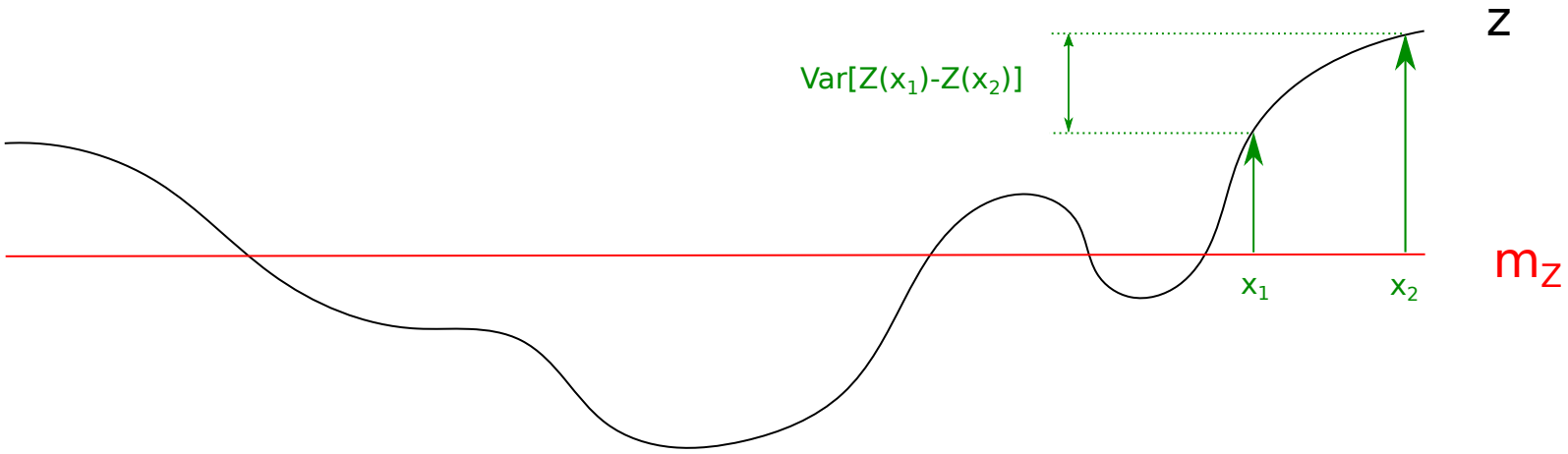
# Analyse variographique



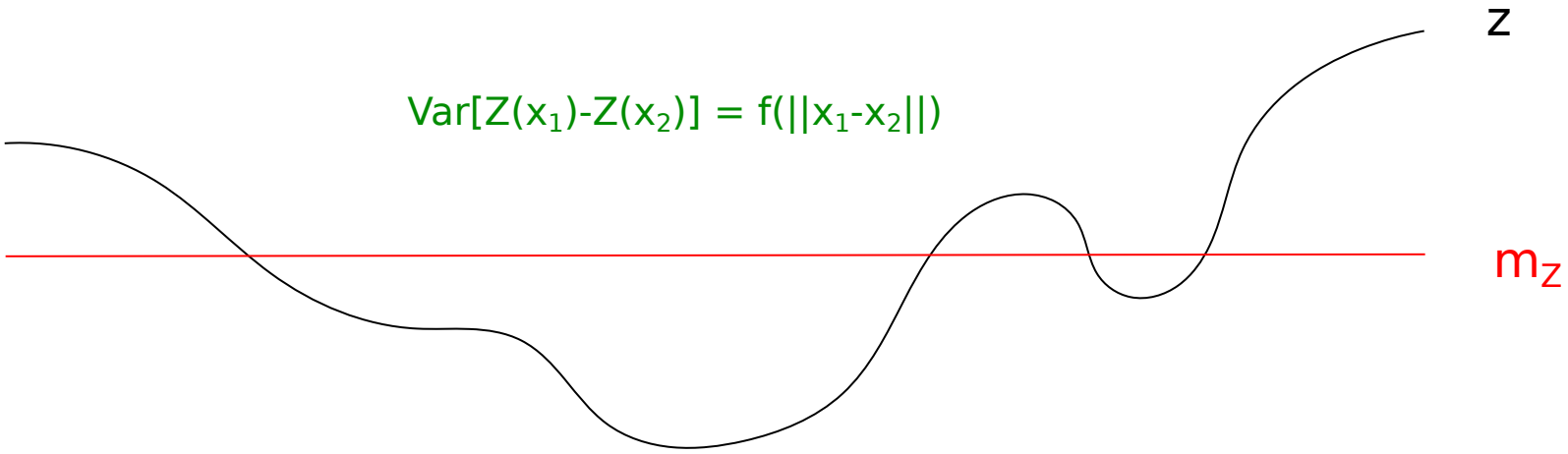
# Analyse variographique



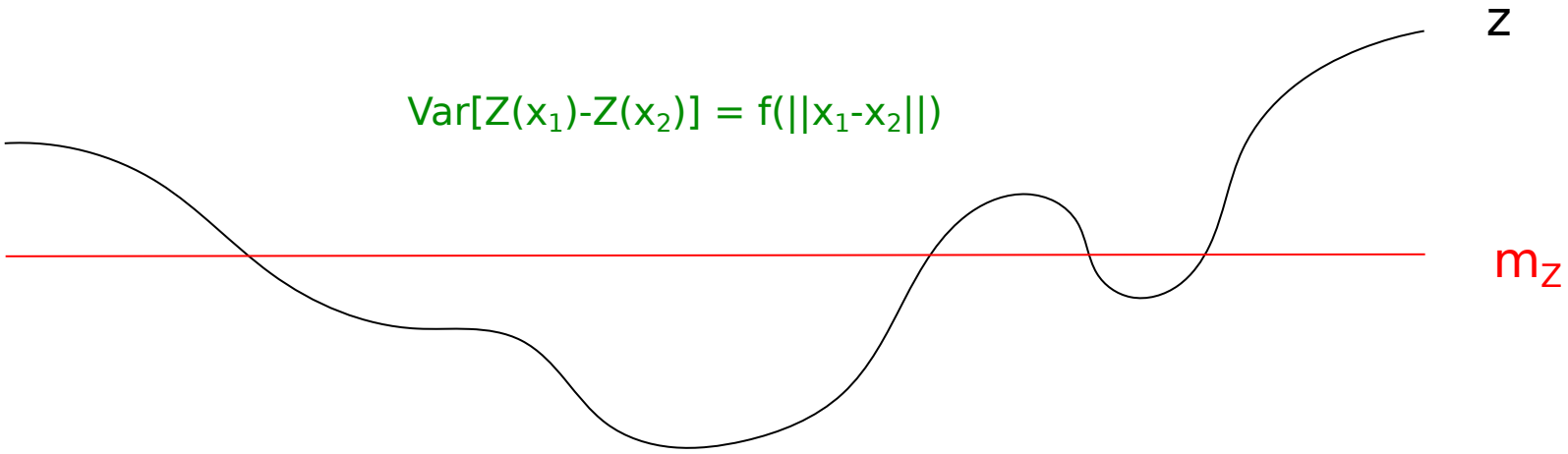
# Analyse variographique



# Analyse variographique

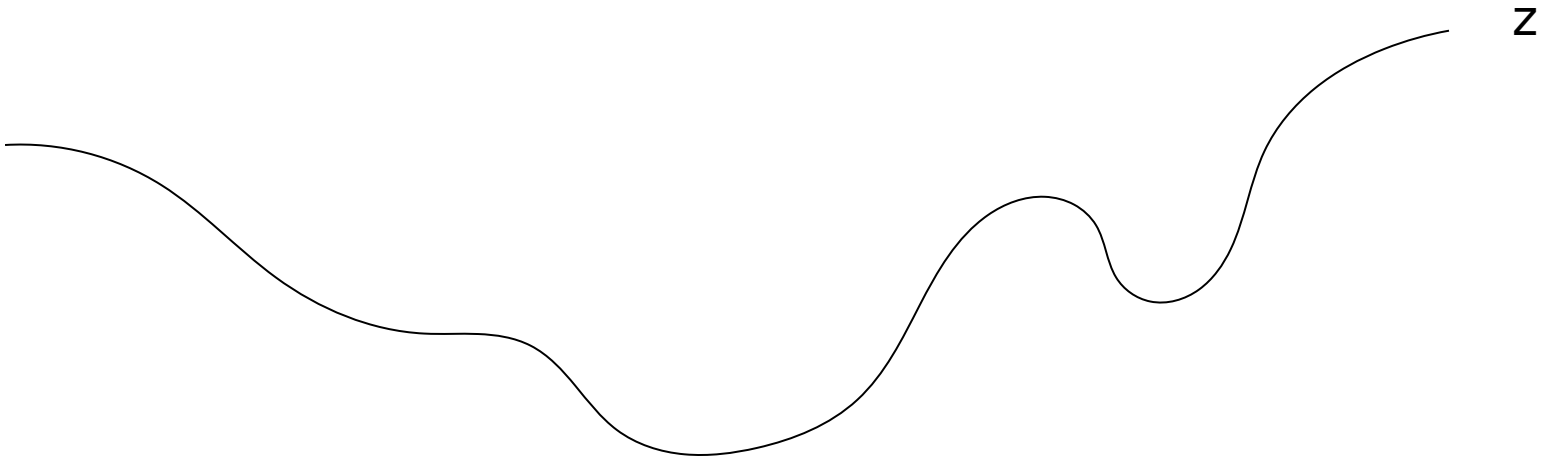


# Analyse variographique

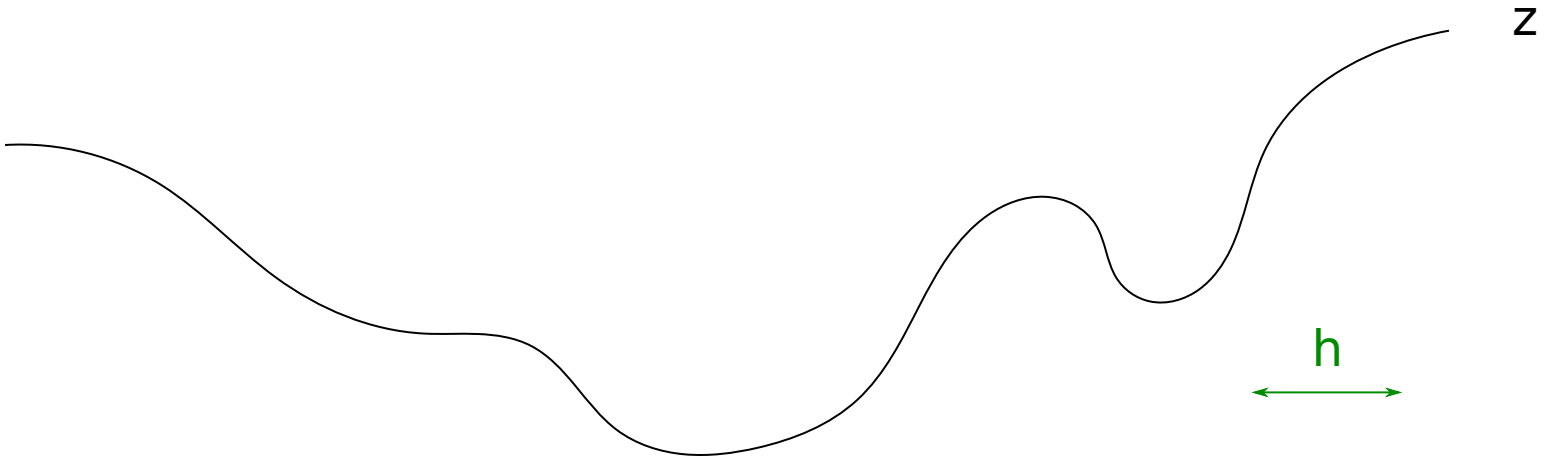



$$\gamma : h \in \mathbb{R} \quad \longrightarrow \quad \gamma(h) = \frac{1}{2}\mathrm{Var}[Z(x+h) - Z(x)]$$

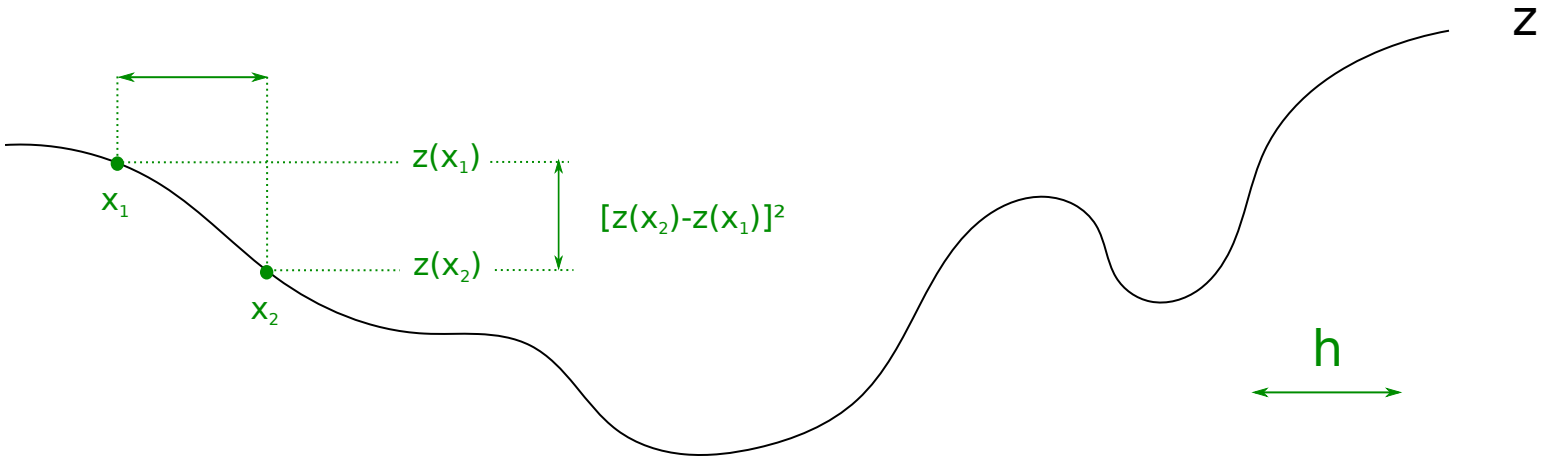
# Analyse variographique



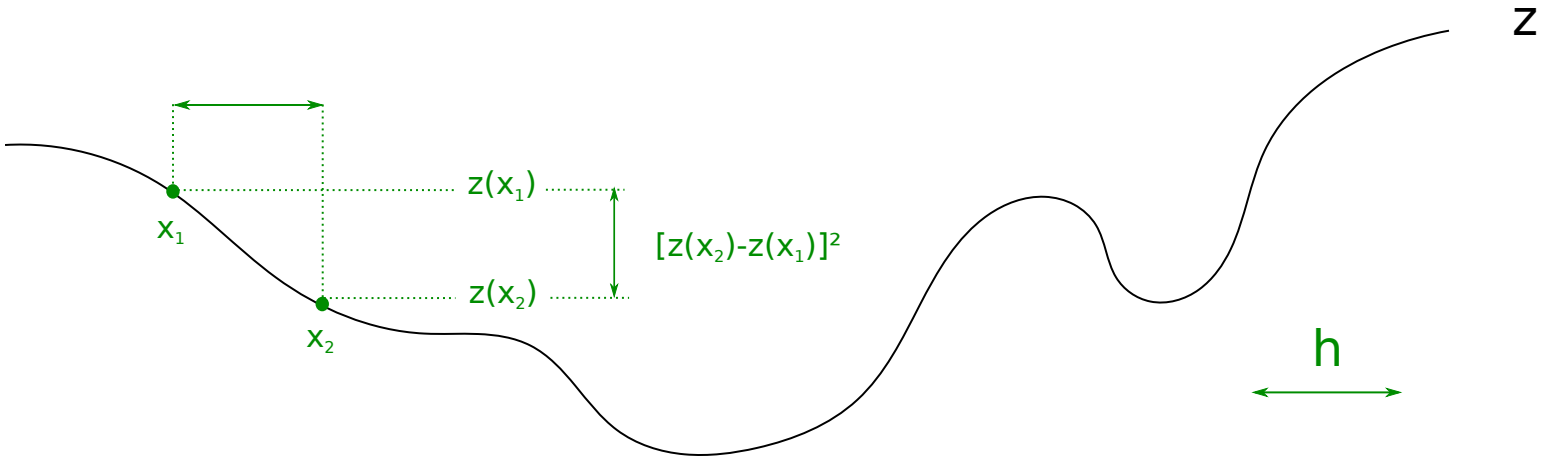
# Analyse variographique



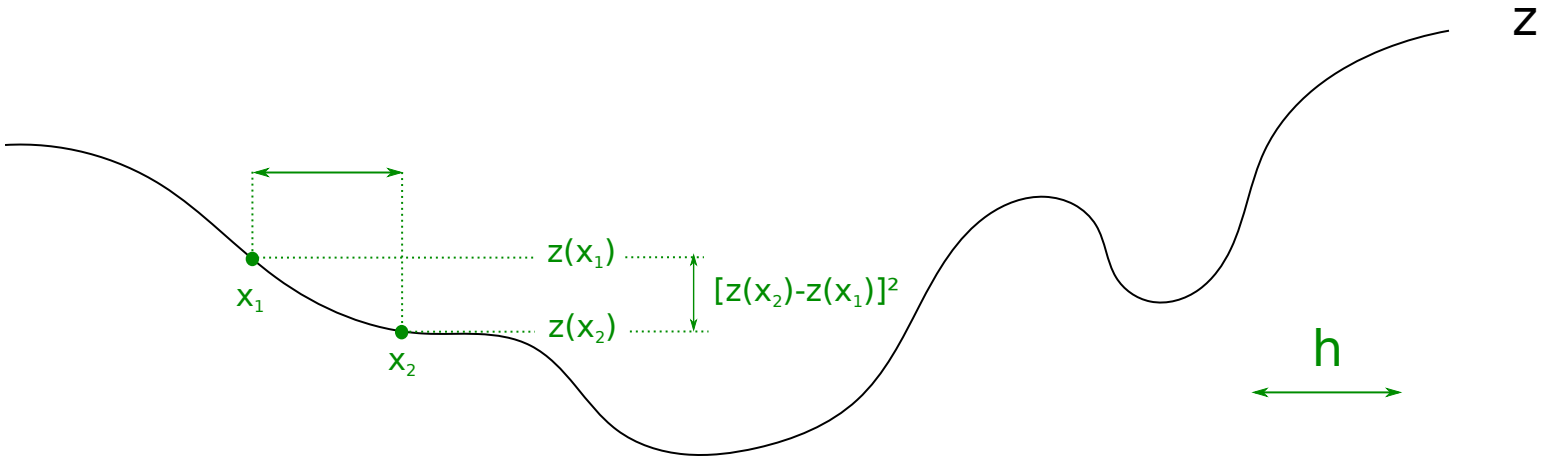
# Analyse variographique



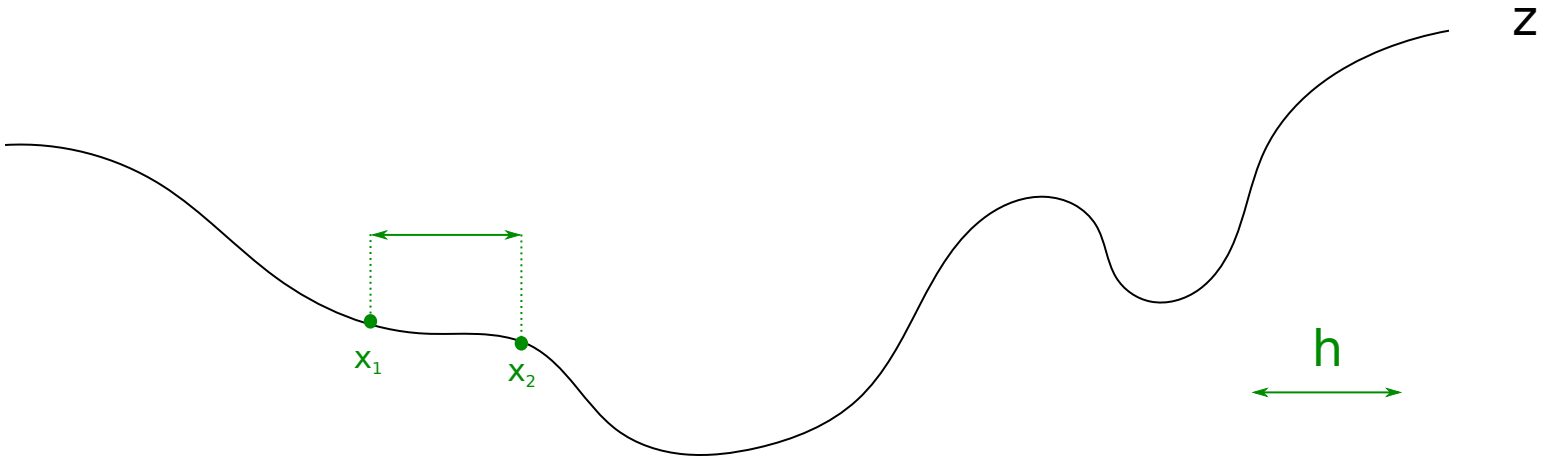
# Analyse variographique



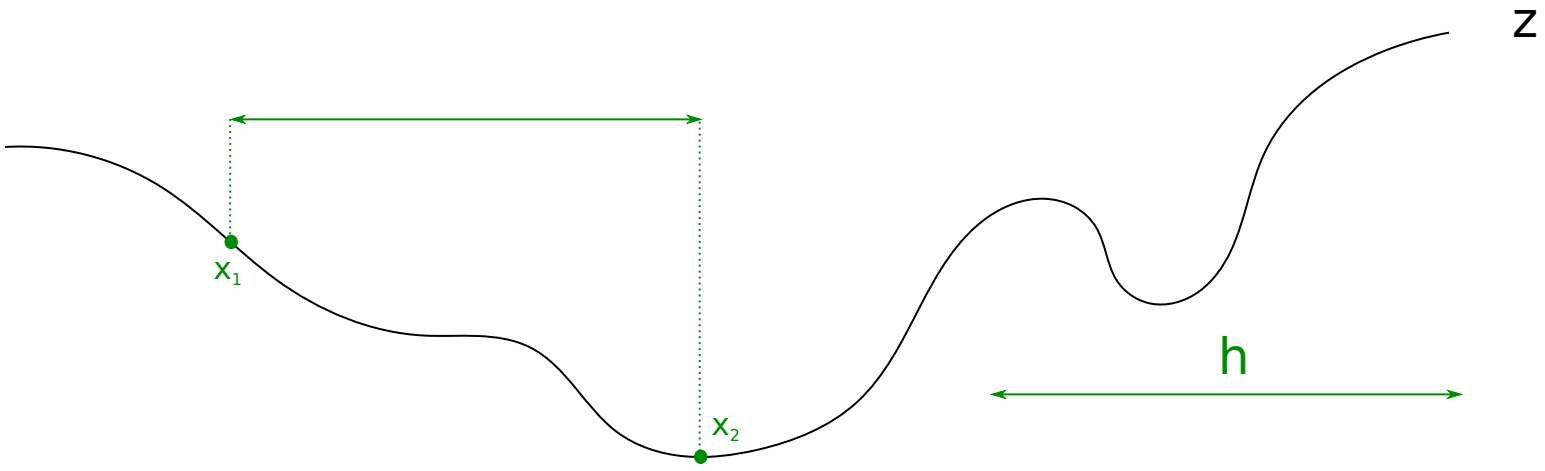
# Analyse variographique



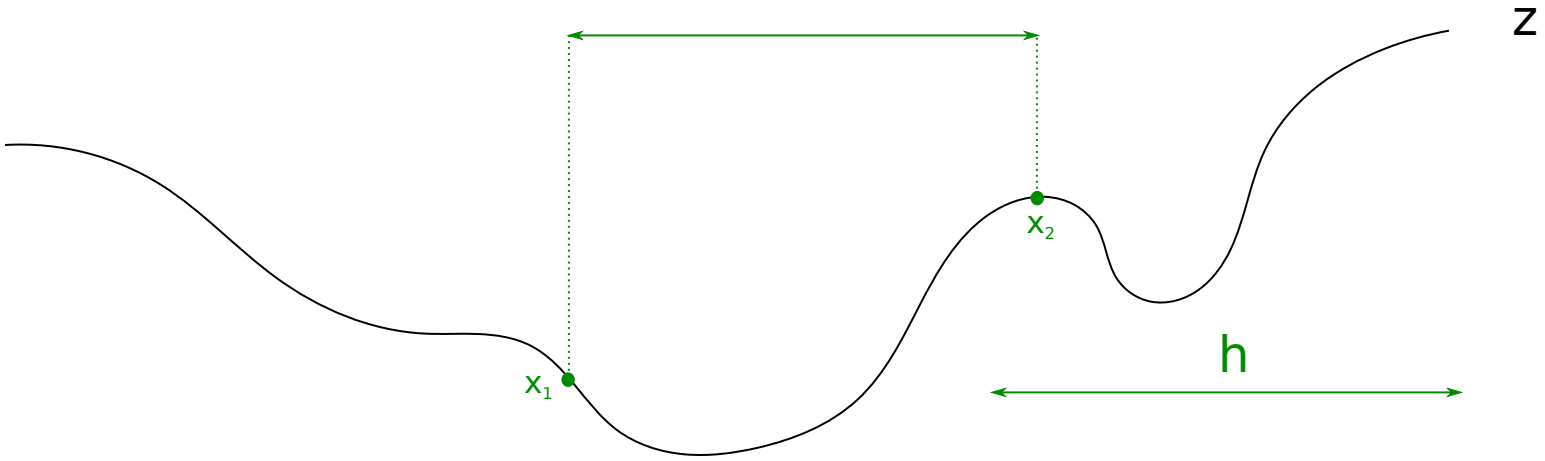
# Analyse variographique



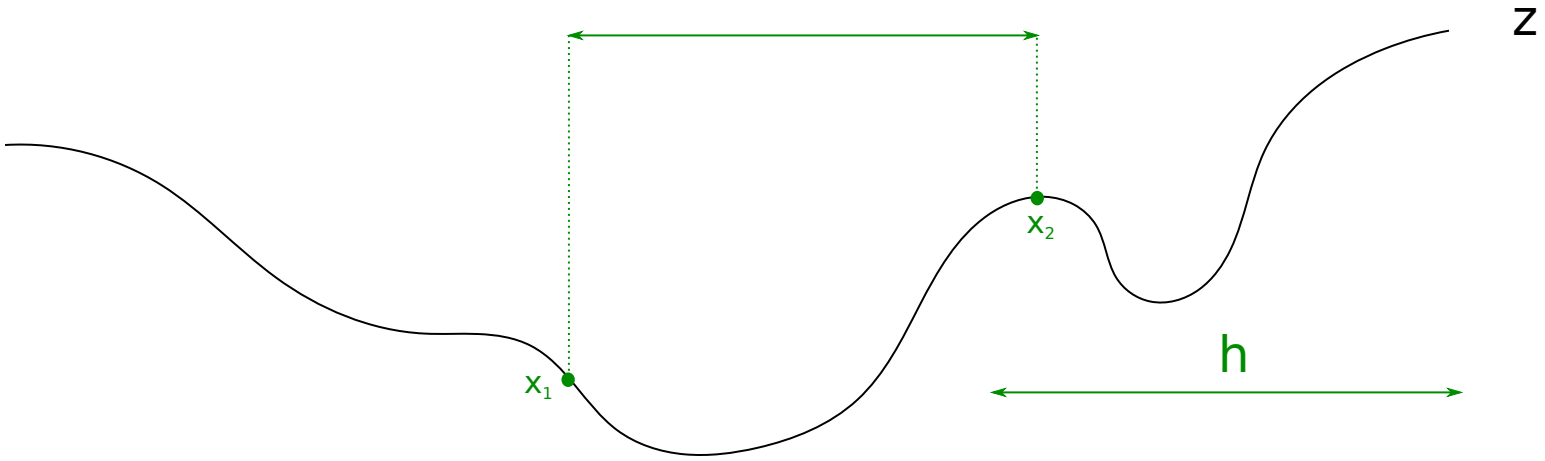
# Analyse variographique



# Analyse variographique



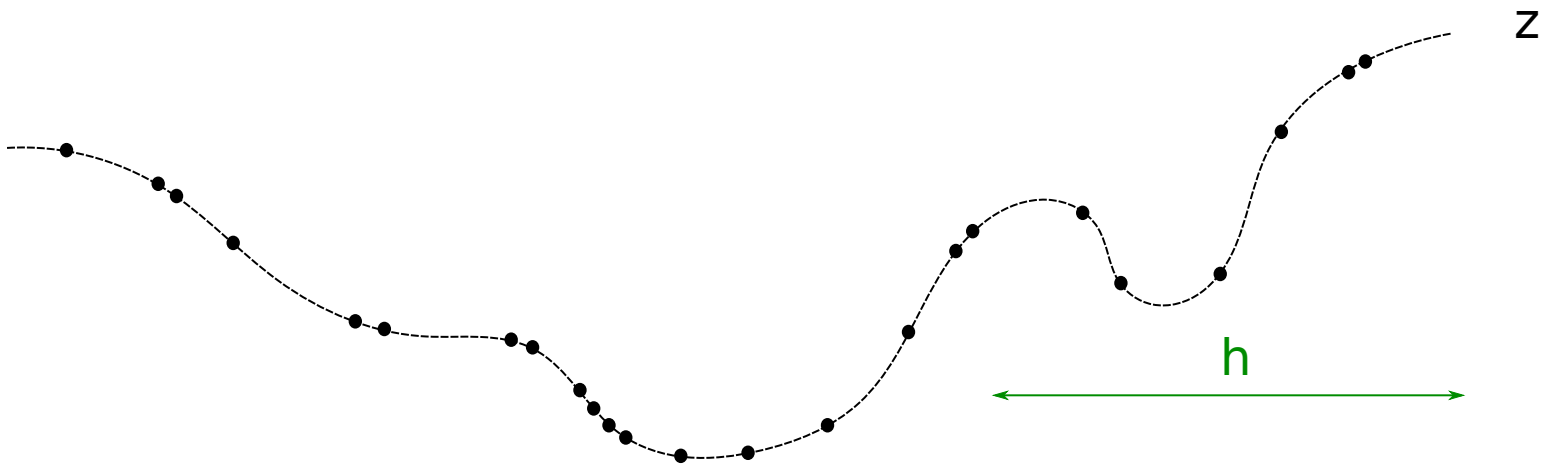
# Analyse variographique



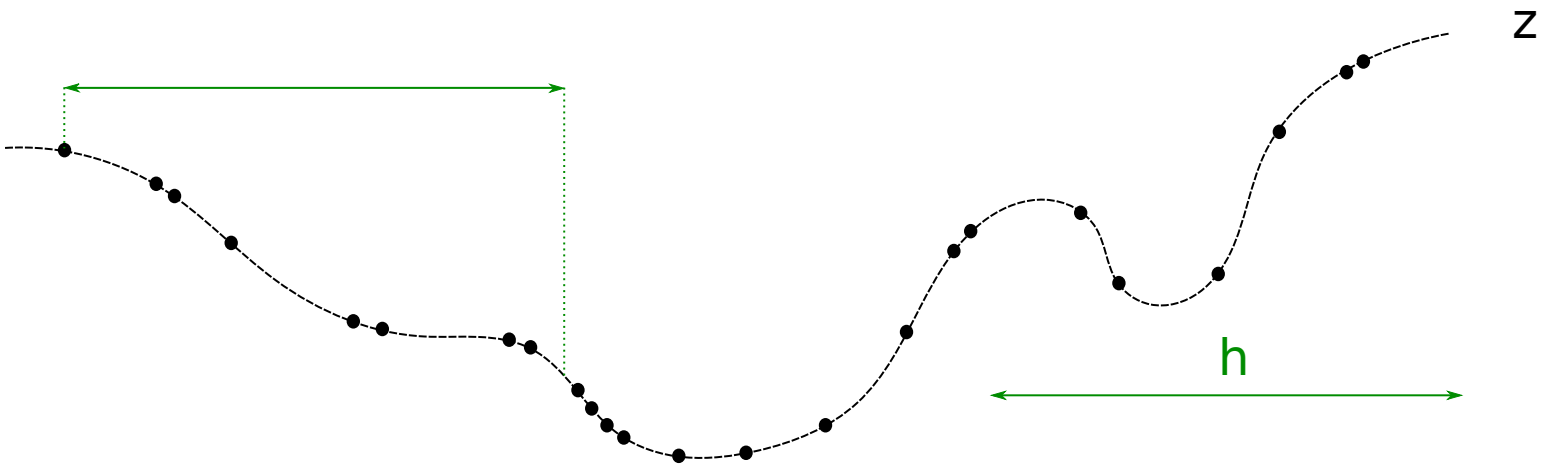

Variogramme théorique :

$$\gamma(h) = \frac{1}{2}\mathbb{E}[(Z(x) - Z(x+h))^2]$$

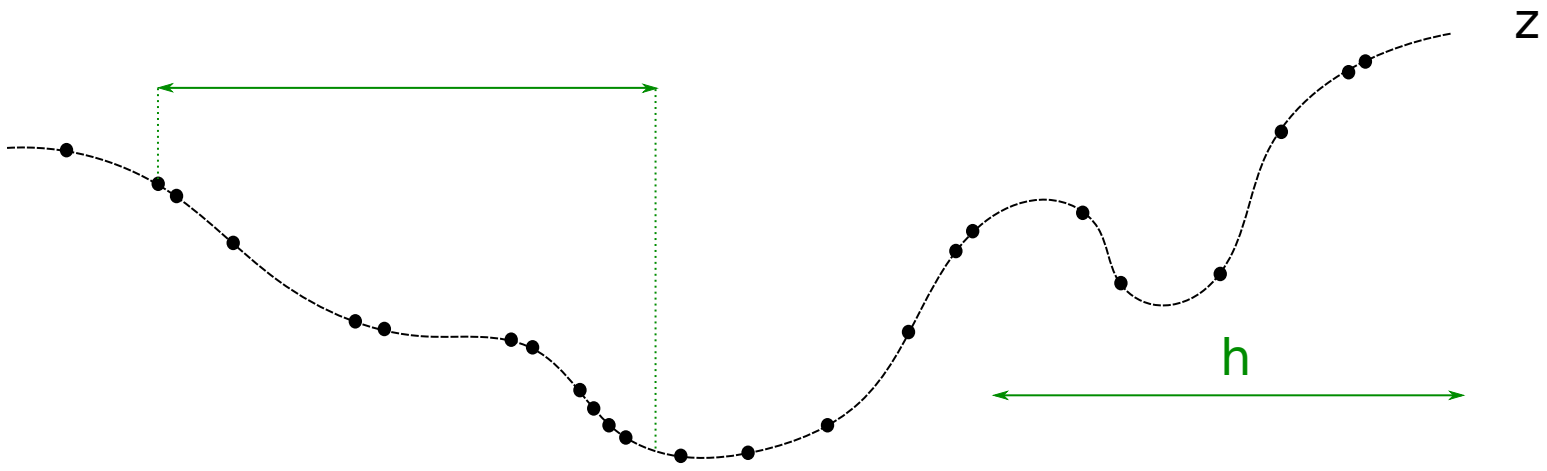
# Analyse variographique



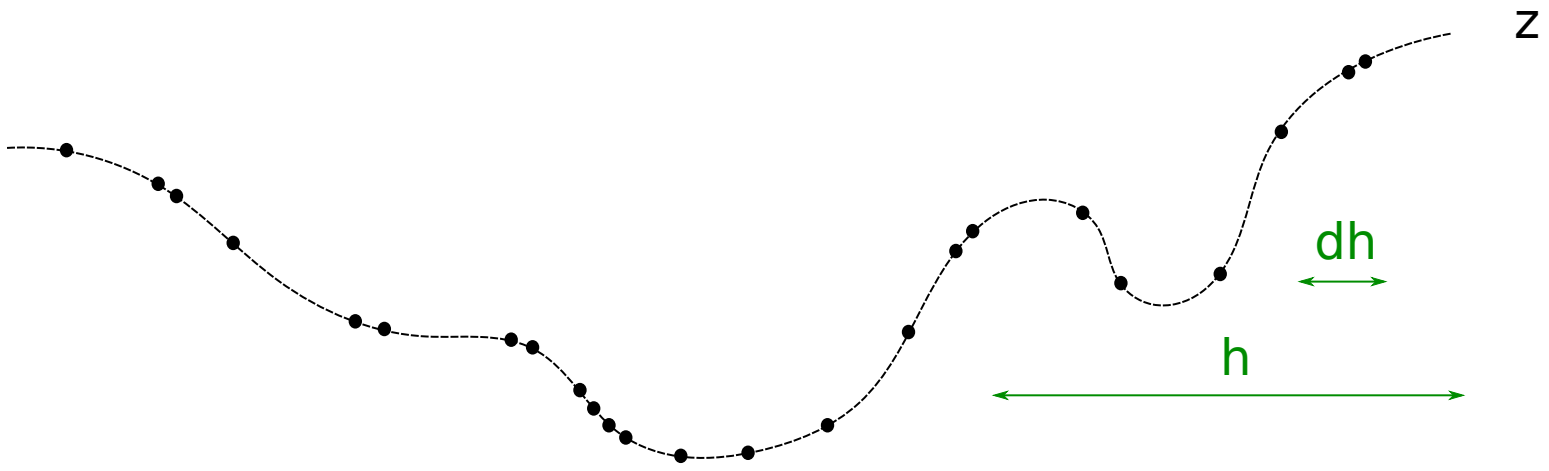
# Analyse variographique



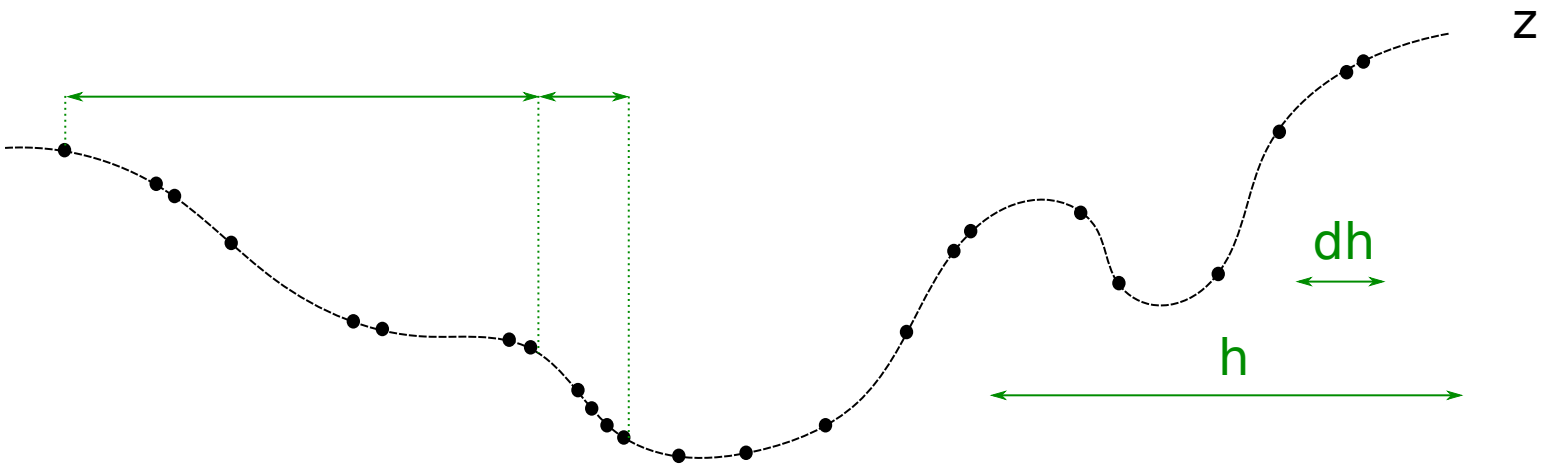
# Analyse variographique



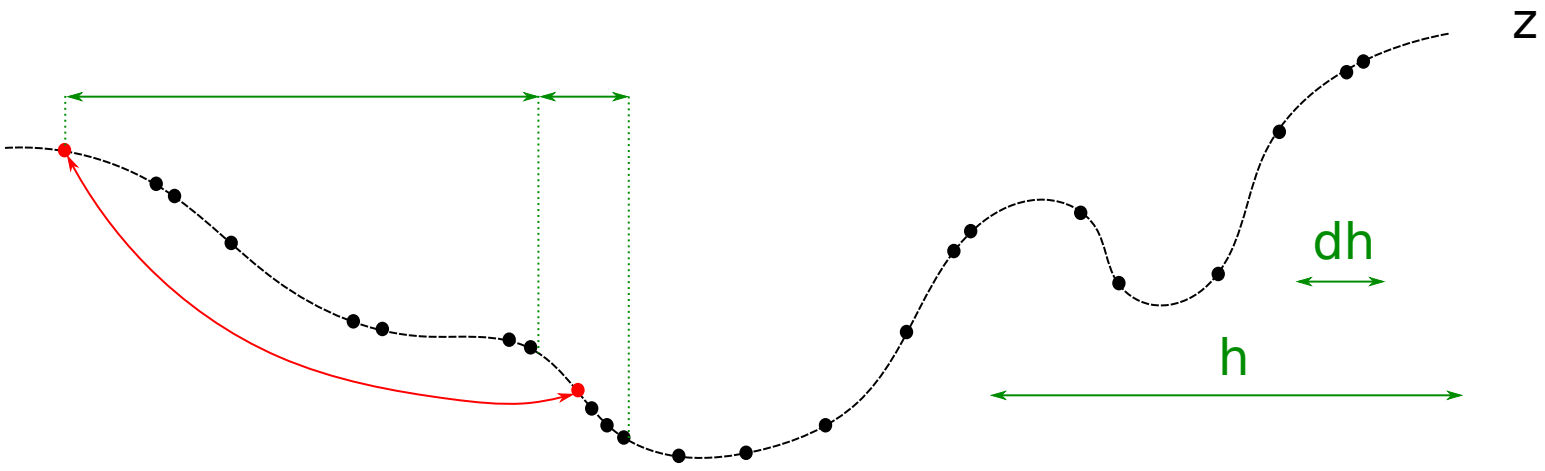
# Analyse variographique



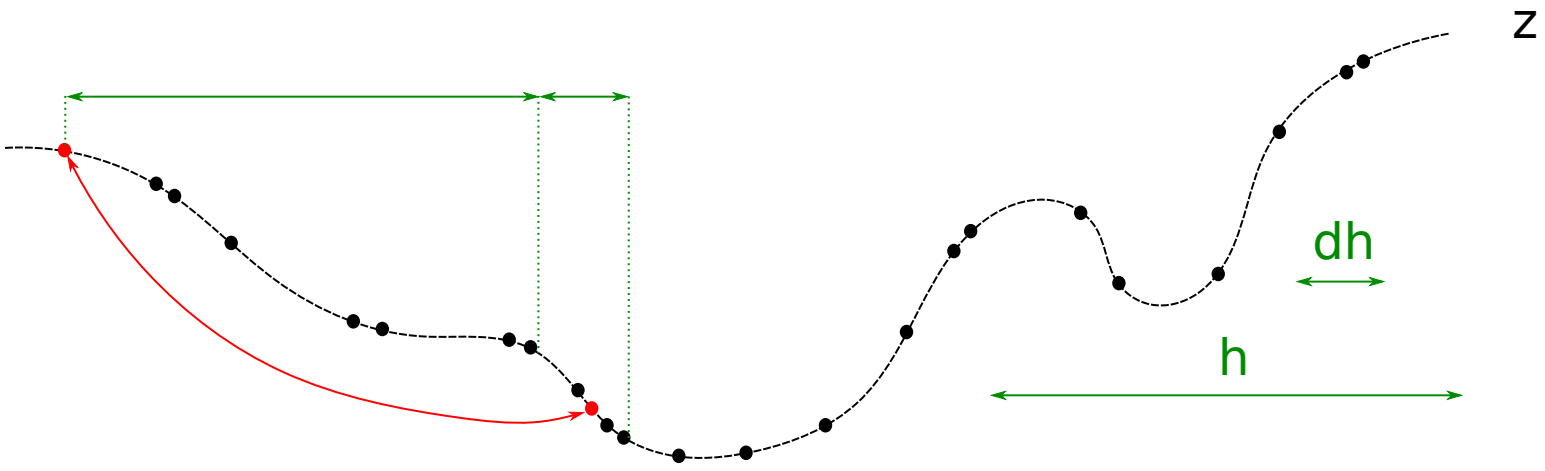
# Analyse variographique



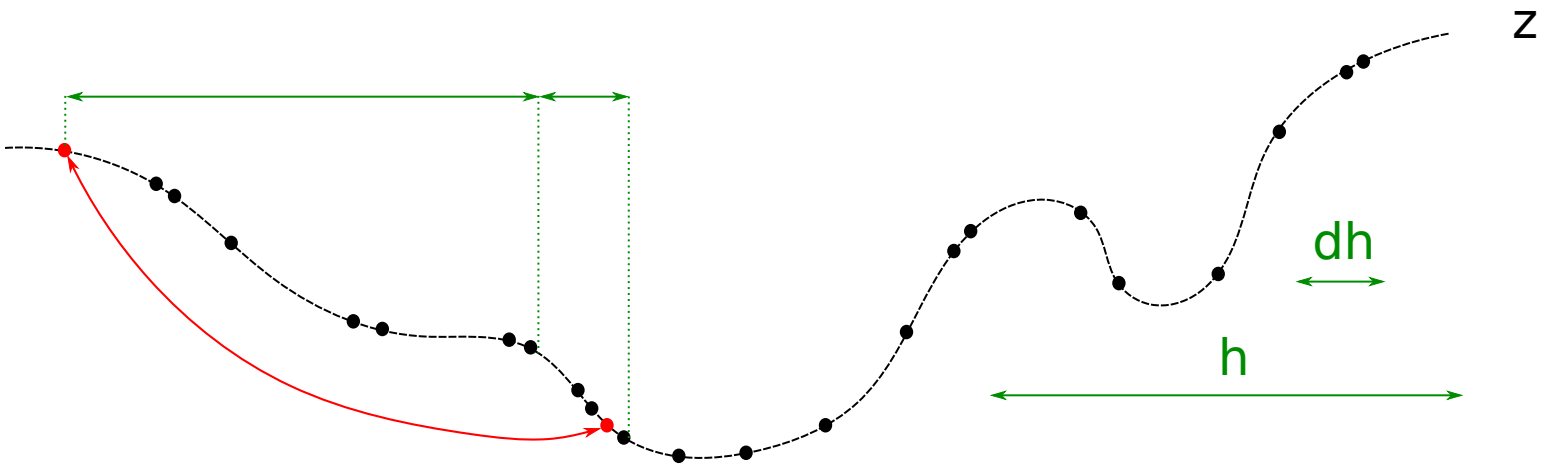
# Analyse variographique



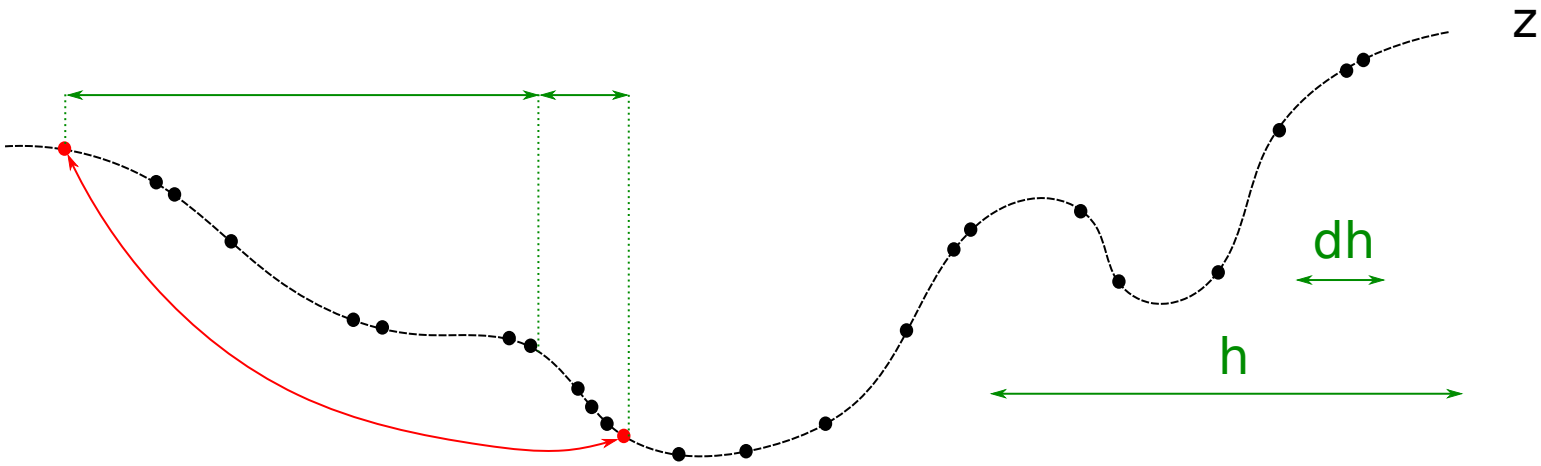
# Analyse variographique



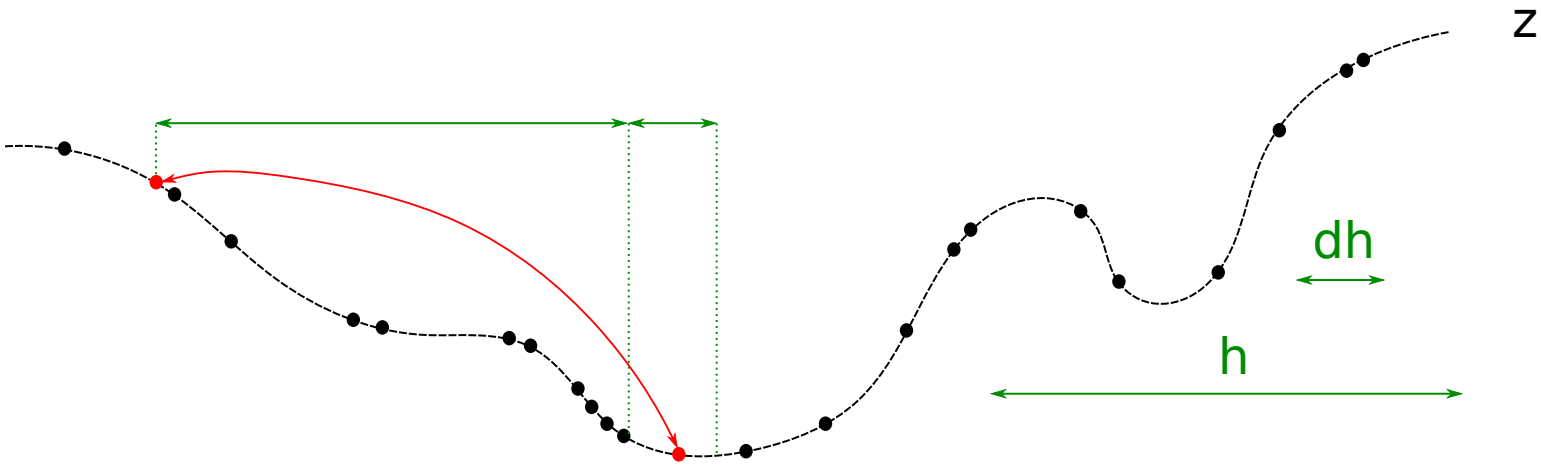
# Analyse variographique



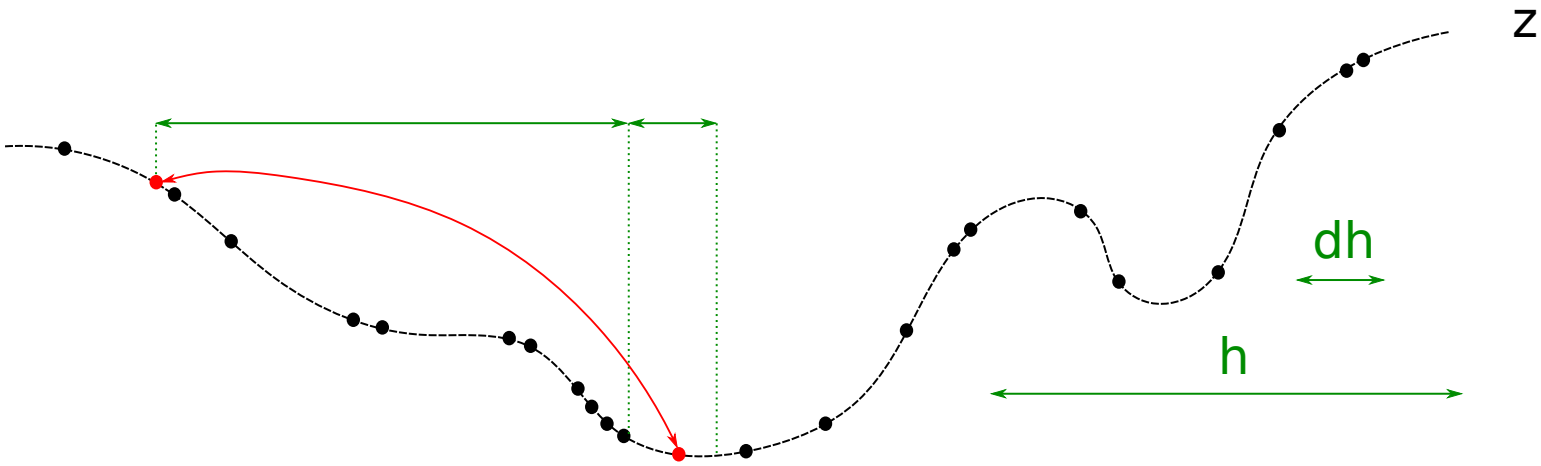
# Analyse variographique



# Analyse variographique



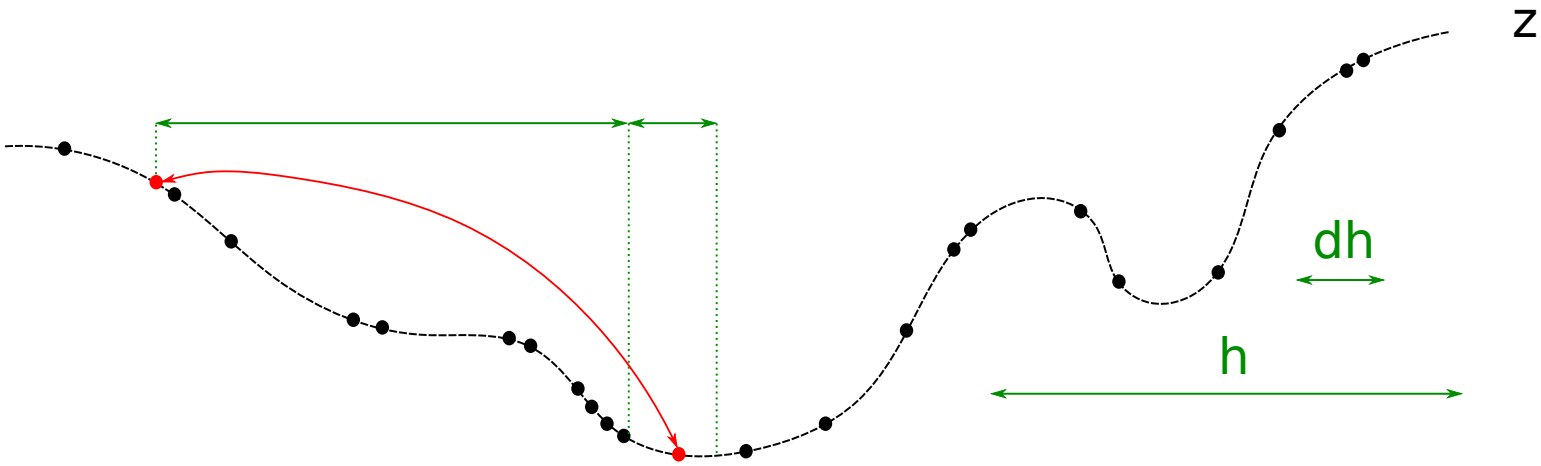
# Analyse variographique




Variogramme expérimental :

$$\hat{\gamma}(h) = \frac{1}{2|N(h)|} \sum_{(i,j)\in N(h)} [z(x_i) - z(x_j)]^2$$

# Analyse variographique

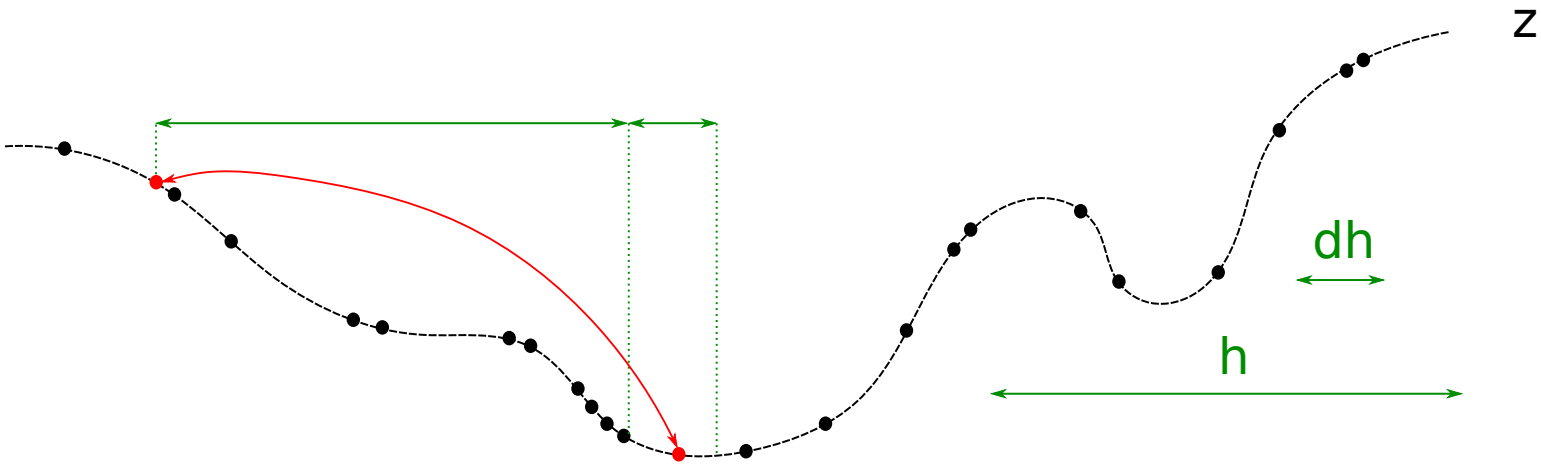



Variogramme expérimental :

$$\hat{\gamma}(h) = \frac{1}{2|N(h)|} \sum_{(i,j)\in N(h)} [z(x_i) - z(x_j)]^2$$

avec : $N(h) = \{(i,j) \text{ t.q. } h - \frac{dh}{2} \leqslant ||x_i - x_j|| \leqslant h + \frac{dh}{2}\}$

# Analyse variographique

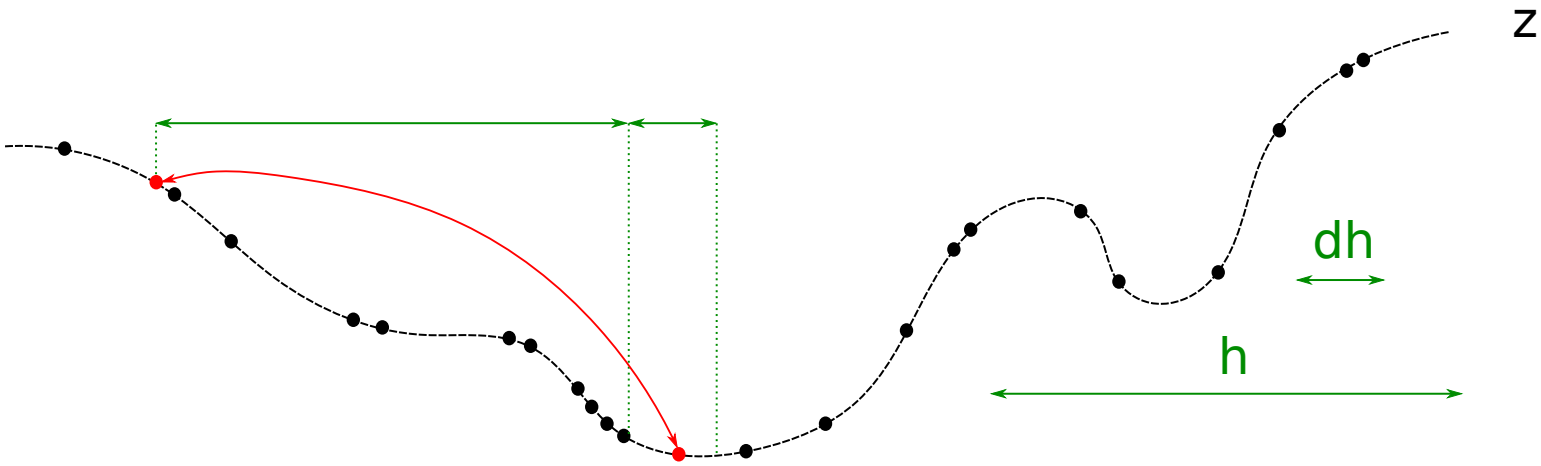



Variogramme expérimental :

$$\hat{\gamma}(h) = \frac{1}{2|N(h)|} \sum_{(i,j)\in N(h)} [z(x_i) - z(x_j)]^2$$

avec : $N(h) = \{(i,j)$ t.q. $h - \frac{dh}{2} \leqslant ||x_i - x_j|| \leqslant h + \frac{dh}{2}\}$ et $dh \in \mathbb{R}^{+*}$

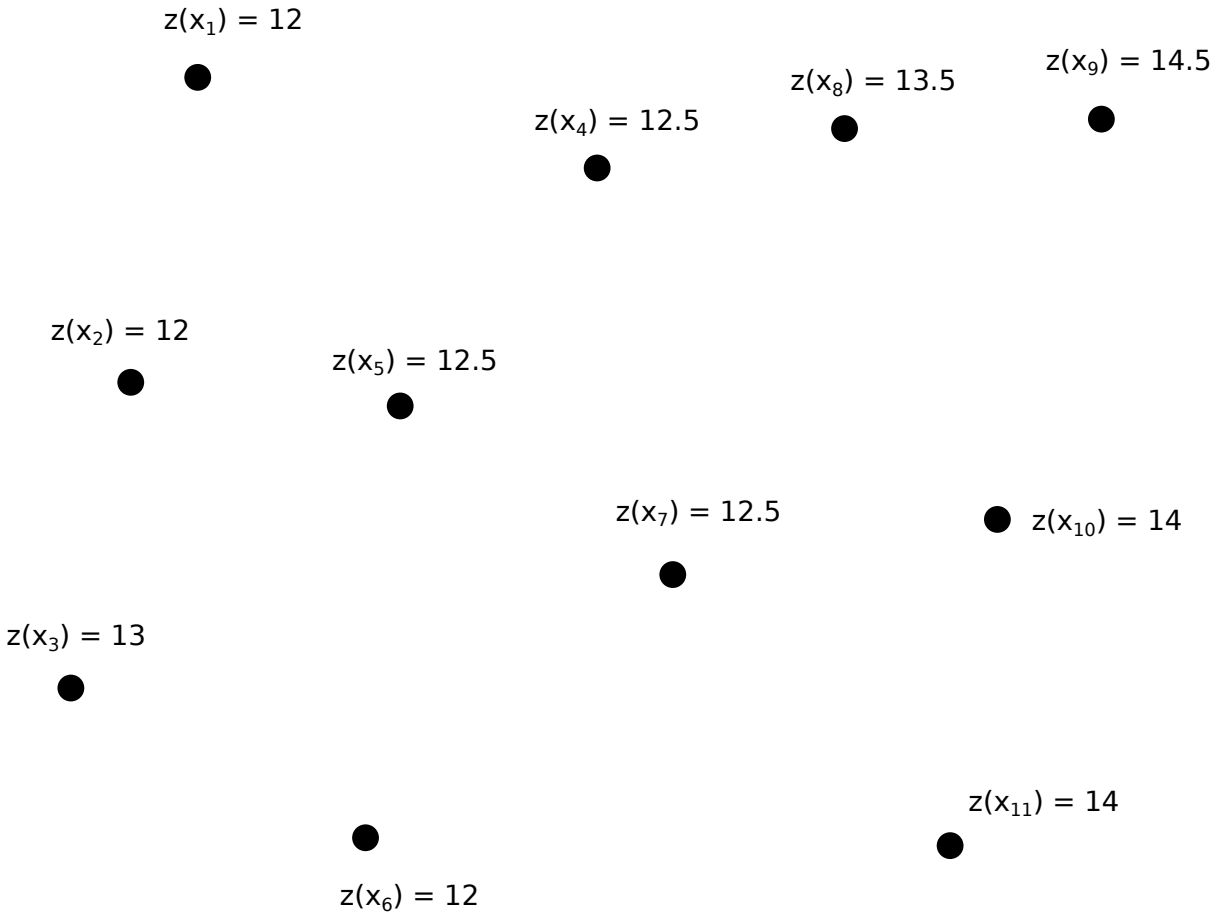
# Analyse variographique



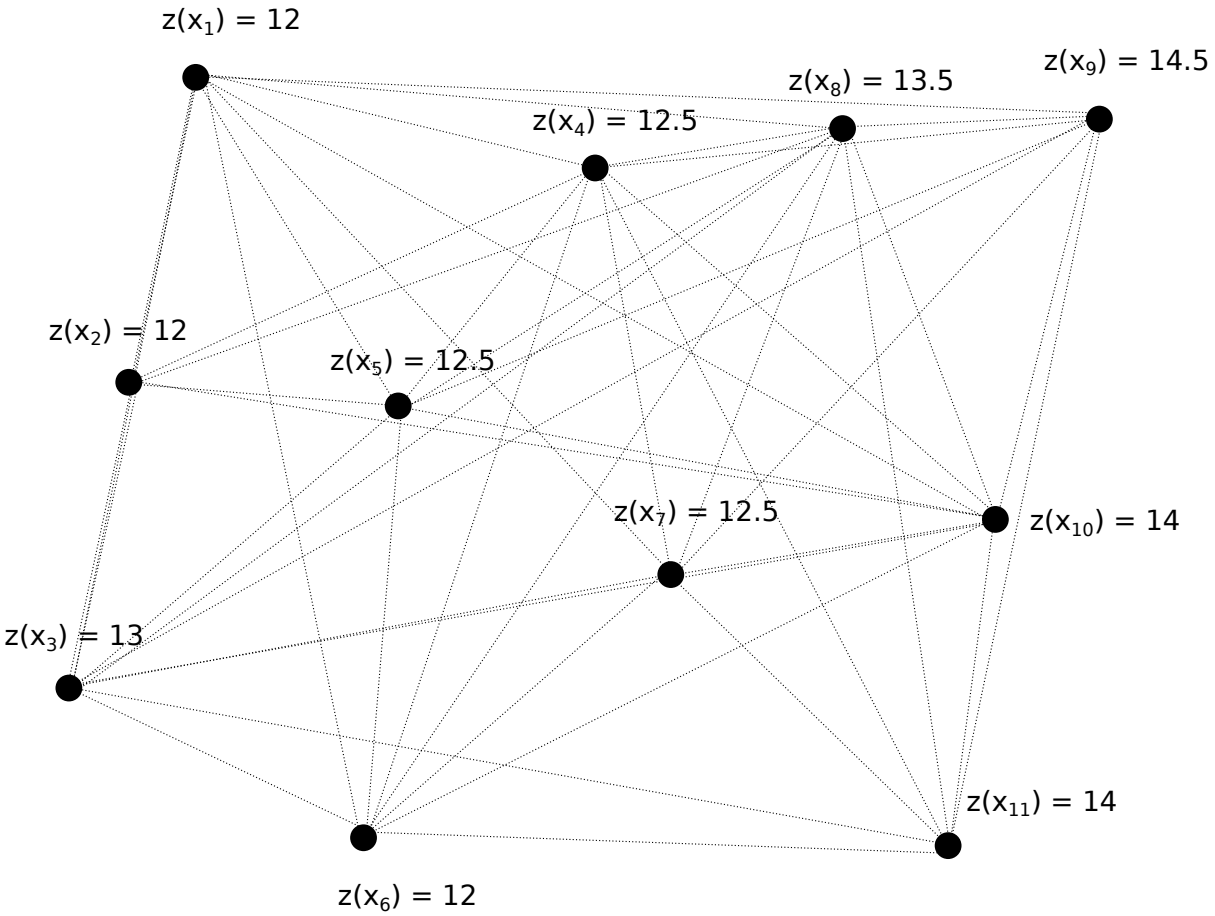

**Définition empirique :**

*Dans le cas imagé d'un semis de points observés sur un MNT, pour une distance horizontale $h$, le variogramme $\gamma(h)$ vaut la moitié de la moyenne des carrés des dénivelés entre les points séparés (approximativement) de $h$.*

# Analyse variographique



# Analyse variographique



# Analyse variographique

| - | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 28 | 56 | 37 | 35 | 71 | 62 | 59 | 82 | 83 | 97 |
| 2 | - | 0 | 28 | 47 | 24 | 46 | 53 | 69 | 91 | 79 | 86 |
| 3 | - | - | 0 | 67 | 39 | 30 | 56 | 86 | 99 | 86 | 81 |
| 4 | - | - | - | 0 | 28 | 64 | 38 | 22 | 46 | 48 | 69 |
| 5 | - | - | - | - | 0 | 39 | 28 | 48 | 68 | 55 | 64 |
| 6 | - | - | - | - | - | 0 | 37 | 78 | 94 | 64 | 53 |
| 7 | - | - | - | - | - | - | 0 | 43 | 57 | 30 | 35 |
| 8 | - | - | - | - | - | - | - | 0 | 23 | 39 | 66 |
| 9 | - | - | - | - | - | - | - | - | 0 | 38 | 68 |
| 10 | - | - | - | - | - | - | - | - | - | 0 | 30 |
| 11 | - | - | - | - | - | - | - | - | - | - | 0 |

# Analyse variographique

On fixe : $dh = 10$

# Analyse variographique

On fixe : $dh = 10$

On souhaite calculer $\hat{\gamma}(65)$

# Analyse variographique

On fixe : $dh = 10$

On souhaite calculer $\hat{\gamma}(65)$

$$N(65) = \{(i,j) \text{ t.q. } 60 \leqslant ||x_i - x_j|| \leqslant 70\}$$

# Analyse variographique

On fixe : $dh = 10$

On souhaite calculer $\hat{\gamma}(65)$

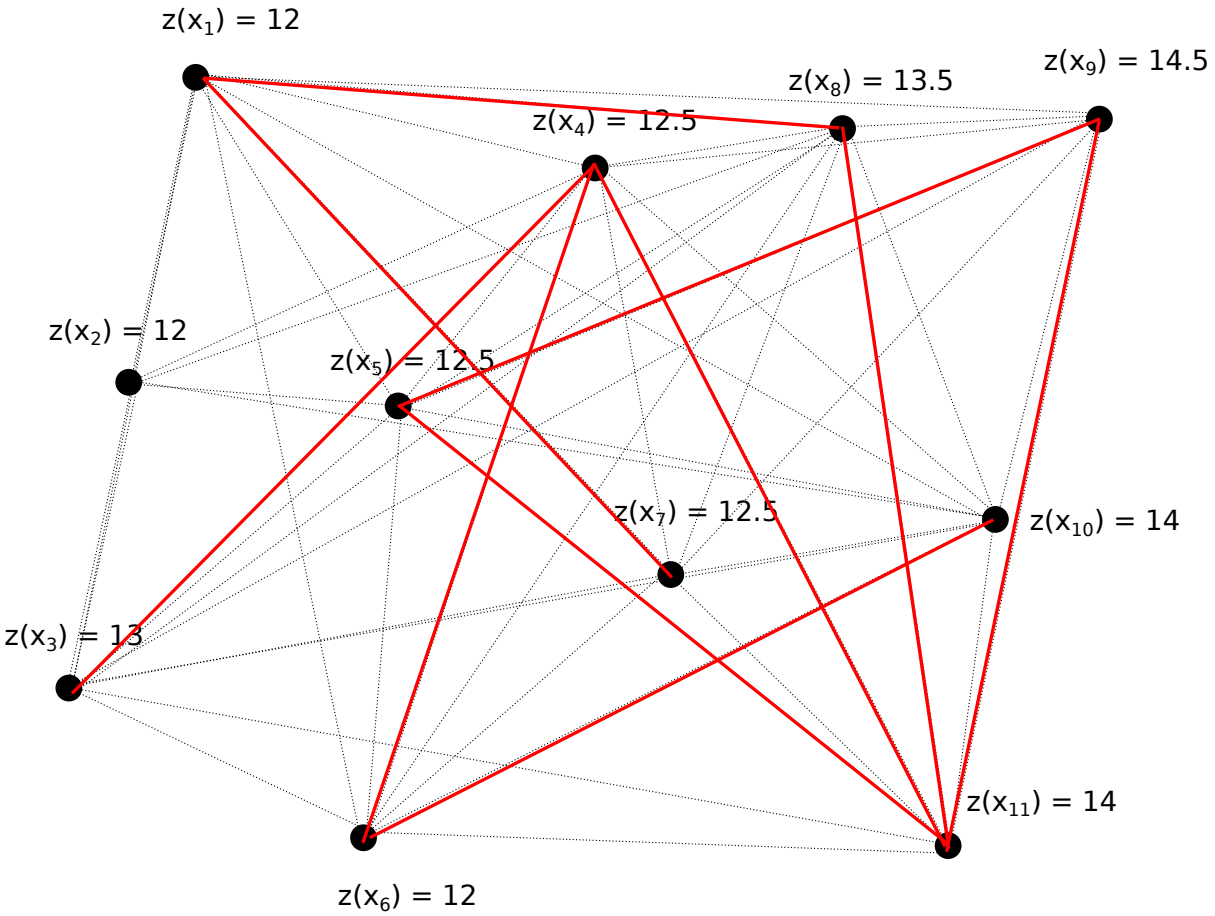
$$N(65) = \{(i,j) \text{ t.q. } 60 \leqslant ||x_i - x_j|| \leqslant 70\}$$

$$\hat{\gamma}(65) = \frac{1}{2|N(65)|} \sum_{(i,j)\in N(65)} [z(x_i) - z(x_j)]^2$$

# Analyse variographique

| - | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 28 | 56 | 37 | 35 | 71 | 62 | 59 | 82 | 83 | 97 |
| 2 | - | 0 | 28 | 47 | 24 | 46 | 53 | 69 | 91 | 79 | 86 |
| 3 | - | - | 0 | 67 | 39 | 30 | 56 | 86 | 99 | 86 | 81 |
| 4 | - | - | - | 0 | 28 | 64 | 38 | 22 | 46 | 48 | 69 |
| 5 | - | - | - | - | 0 | 39 | 28 | 48 | 68 | 55 | 64 |
| 6 | - | - | - | - | - | 0 | 37 | 78 | 94 | 64 | 53 |
| 7 | - | - | - | - | - | - | 0 | 43 | 57 | 30 | 35 |
| 8 | - | - | - | - | - | - | - | 0 | 23 | 39 | 66 |
| 9 | - | - | - | - | - | - | - | - | 0 | 38 | 68 |
| 10 | - | - | - | - | - | - | - | - | - | 0 | 30 |
| 11 | - | - | - | - | - | - | - | - | - | - | 0 |

# Analyse variographique



# Analyse variographique

$$N(65) = \{(1,7)(1,8)(3,4)(4,6)(4,11)(5,9)(5,11)(6,10)(8,11)(9,11)\}$$

# Analyse variographique

$$N(65) = \{(1,7)(1,8)(3,4)(4,6)(4,11)(5,9)(5,11)(6,10)(8,11)(9,11)\}$$

donc $|N(65)| = 10$

# Analyse variographique

$$N(65) = \{(1,7)(1,8)(3,4)(4,6)(4,11)(5,9)(5,11)(6,10)(8,11)(9,11)\}$$

donc $|N(65)| = 10$

$$\hat{\gamma}(65) = \frac{1}{2|N(65)|} \sum_{(i,j)\in N(65)} [z(x_i) - z(x_j)]^2$$

# Analyse variographique

$$N(65) = \{(1,7)(1,8)(3,4)(4,6)(4,11)(5,9)(5,11)(6,10)(8,11)(9,11)\}$$

donc $|N(65)| = 10$

$$\hat{\gamma}(65) = \frac{1}{2|N(65)|} \sum_{(i,j)\in N(65)} [z(x_i) - z(x_j)]^2$$

$$\hat{\gamma}(65) = \frac{1}{20}\Big[(z_1 - z_7)^2 + (z_1 - z_8)^2 + (z_3 - z_4)^2 + (z_4 - z_6)^2 + (z_4 - z_{11})^2 + (z_5 - z_9)^2 + (z_5 - z_{11})^2 + (z_6 - z_{10})^2 + (z_8 - z_{11})^2 + (z_9 - z_{11})^2\Big]$$

# Analyse variographique

$$N(65) = \{(1,7)(1,8)(3,4)(4,6)(4,11)(5,9)(5,11)(6,10)(8,11)(9,11)\}$$

donc $|N(65)| = 10$

$$\hat{\gamma}(65) = \frac{1}{2|N(65)|} \sum_{(i,j)\in N(65)} [z(x_i) - z(x_j)]^2$$

$$\hat{\gamma}(65) = \frac{1}{20}\Big[(12-12.5)^2 + (12-13.5)^2 + (13-12.5)^2 + (12.5-12)^2 + (12.5-14)^2 + (12.5-14.5)^2 + (12.5-14)^2 + (12-14)^2 + (13.5-14)^2 + (14.5-14)^2\Big]$$

# Analyse variographique

$$N(65) = \{(1,7)(1,8)(3,4)(4,6)(4,11)(5,9)(5,11)(6,10)(8,11)(9,11)\}$$

donc $|N(65)| = 10$

$$\hat{\gamma}(65) = \frac{1}{2|N(65)|} \sum_{(i,j)\in N(65)} [z(x_i) - z(x_j)]^2$$

$$\hat{\gamma}(65) = \frac{1}{20}\Big[0.25+2.25+0.25+0.25+2.25+4+2.25+4+0.25+0.25\Big]$$

# Analyse variographique

$$N(65) = \{(1,7)(1,8)(3,4)(4,6)(4,11)(5,9)(5,11)(6,10)(8,11)(9,11)\}$$

donc $|N(65)| = 10$

$$\hat{\gamma}(65) = \frac{1}{2|N(65)|} \sum_{(i,j)\in N(65)} [z(x_i) - z(x_j)]^2$$

$$\hat{\gamma}(65) = \frac{1}{20}\Big[0.25+2.25+0.25+0.25+2.25+4+2.25+4+0.25+0.25\Big]$$

$$\hat{\gamma}(65) = \frac{16}{20}$$

# Analyse variographique

$$N(65) = \{(1,7)(1,8)(3,4)(4,6)(4,11)(5,9)(5,11)(6,10)(8,11)(9,11)\}$$

donc $|N(65)| = 10$

$$\hat{\gamma}(65) = \frac{1}{2|N(65)|} \sum_{(i,j)\in N(65)} [z(x_i) - z(x_j)]^2$$

$$\hat{\gamma}(65) = \frac{1}{20}\Big[0.25+2.25+0.25+0.25+2.25+4+2.25+4+0.25+0.25\Big]$$

$$\hat{\gamma}(65) = 0.8$$

# Analyse variographique

$$N(65) = \{(1,7)(1,8)(3,4)(4,6)(4,11)(5,9)(5,11)(6,10)(8,11)(9,11)\}$$

donc $|N(65)| = 10$

$$\hat{\gamma}(65) = \frac{1}{2|N(65)|} \sum_{(i,j)\in N(65)} [z(x_i) - z(x_j)]^2$$

$$\hat{\gamma}(65) = \frac{1}{20}\Big[0.25+2.25+0.25+0.25+2.25+4+2.25+4+0.25+0.25\Big]$$

$\hat{\gamma}(65) = 0.8 \rightarrow$ calcul pour chaque pas $h$

# Analyse variographique

$$N(65) = \{(1,7)(1,8)(3,4)(4,6)(4,11)(5,9)(5,11)(6,10)(8,11)(9,11)\}$$

donc $|N(65)| = 10$

$$\hat{\gamma}(65) = \frac{1}{2|N(65)|} \sum_{(i,j)\in N(65)} [z(x_i) - z(x_j)]^2$$

$$\hat{\gamma}(65) = \frac{1}{20}\Big[0.25+2.25+0.25+0.25+2.25+4+2.25+4+0.25+0.25\Big]$$

$\hat{\gamma}(65) = 0.8 \rightarrow$ calcul pour chaque pas $h$

$\hat{\gamma}(5)$, $\hat{\gamma}(15)$, $\hat{\gamma}(25)$,... $\hat{\gamma}(95)$

# Analyse variographique

$$N(65) = \{(1,7)(1,8)(3,4)(4,6)(4,11)(5,9)(5,11)(6,10)(8,11)(9,11)\}$$

donc $|N(65)| = 10$

$$\hat{\gamma}(65) = \frac{1}{2|N(65)|} \sum_{(i,j)\in N(65)} [z(x_i) - z(x_j)]^2$$

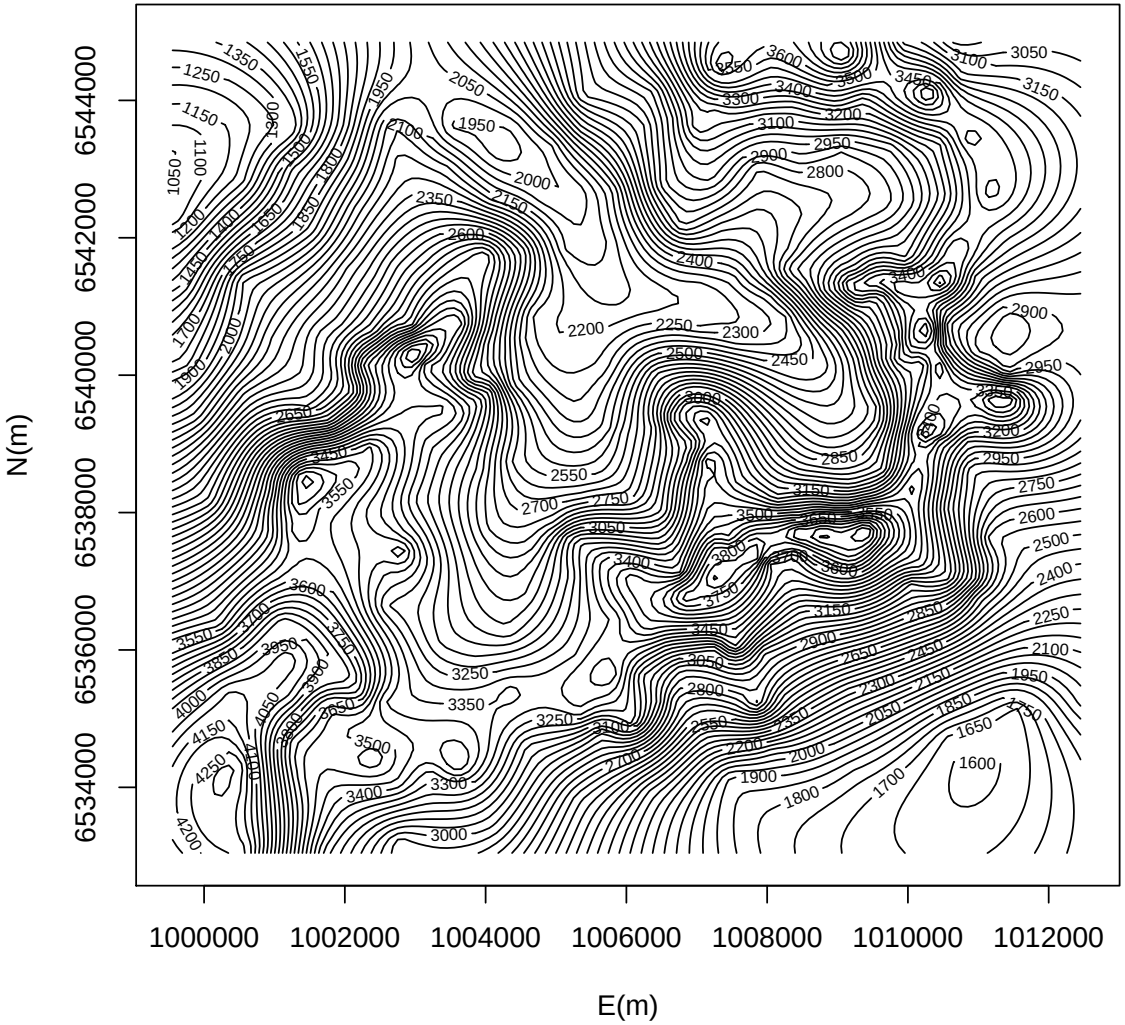
$$\hat{\gamma}(65) = \frac{1}{20}\Big[0.25+2.25+0.25+0.25+2.25+4+2.25+4+0.25+0.25\Big]$$

$\hat{\gamma}(65) = 0.8 \rightarrow$ calcul pour chaque pas $h$

$\hat{\gamma}(5)$, $\hat{\gamma}(15)$, $\hat{\gamma}(25)$,... $\hat{\gamma}(95) \rightarrow$ complexité en $\mathcal{O}(n^2)$

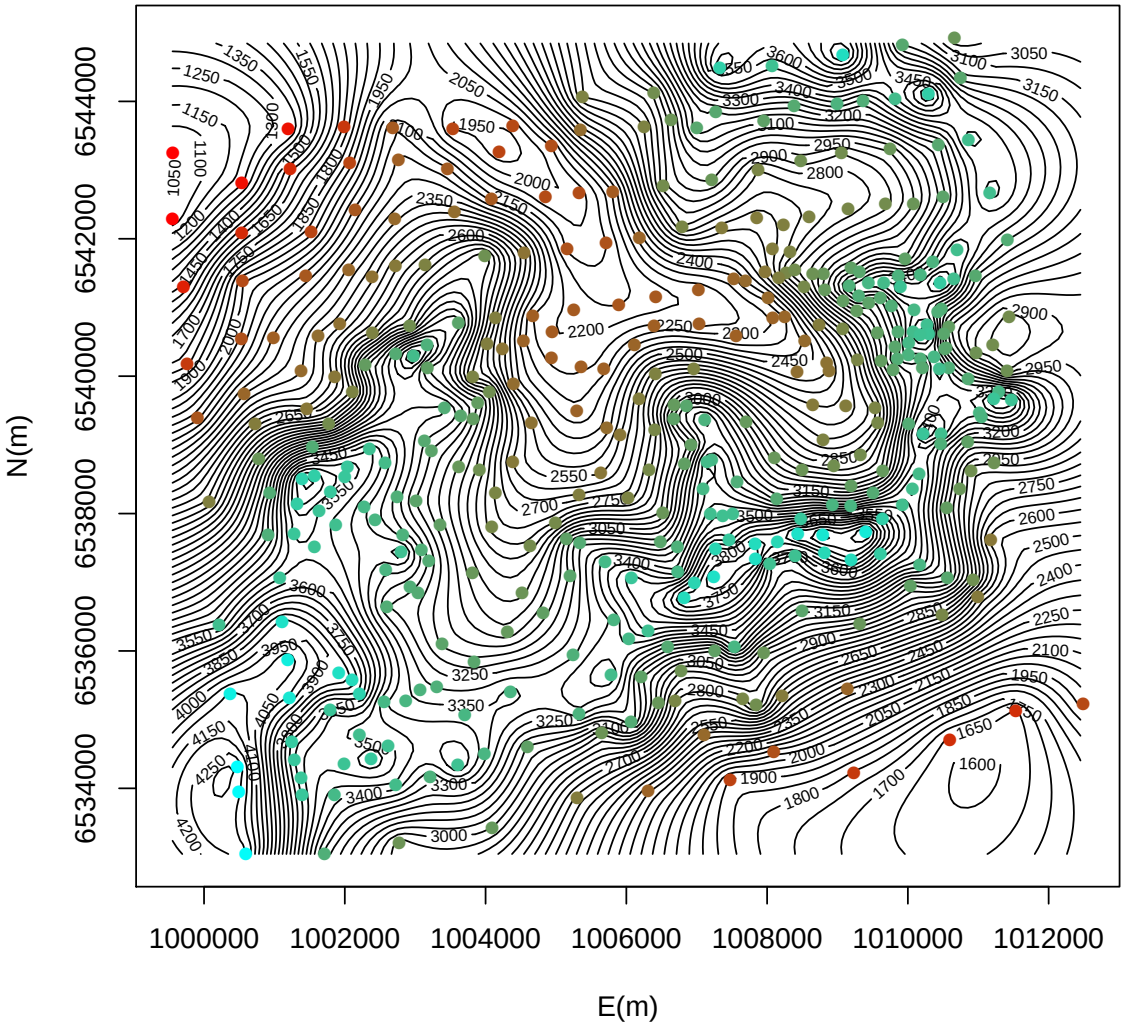
# Analyse variographique

Exemple de variogramme (pas $dh = 400$ m)



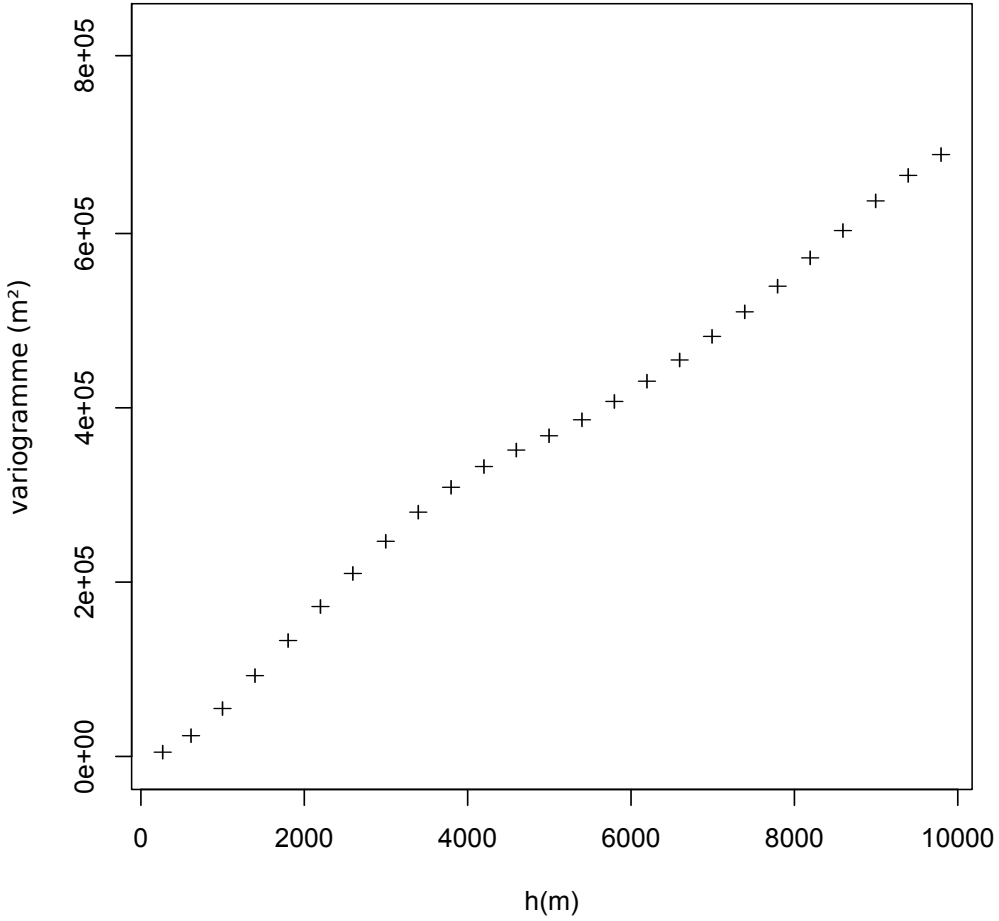
# Analyse variographique

Exemple de variogramme (pas $dh = 400$ m)



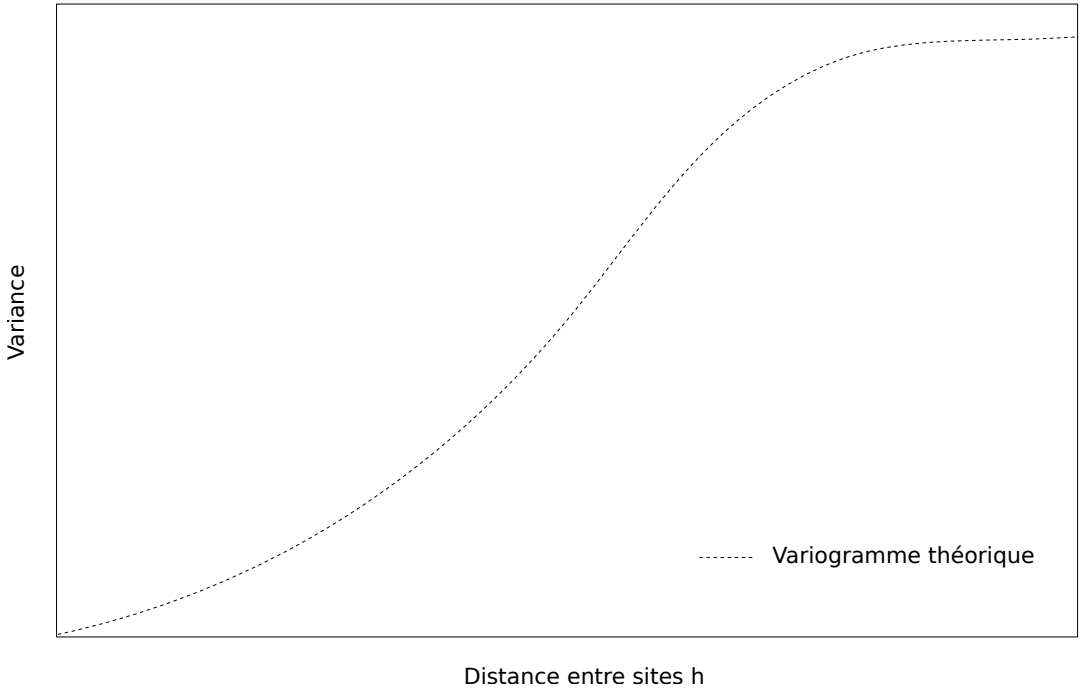
# Analyse variographique

Exemple de variogramme (pas $dh = 400$ m)



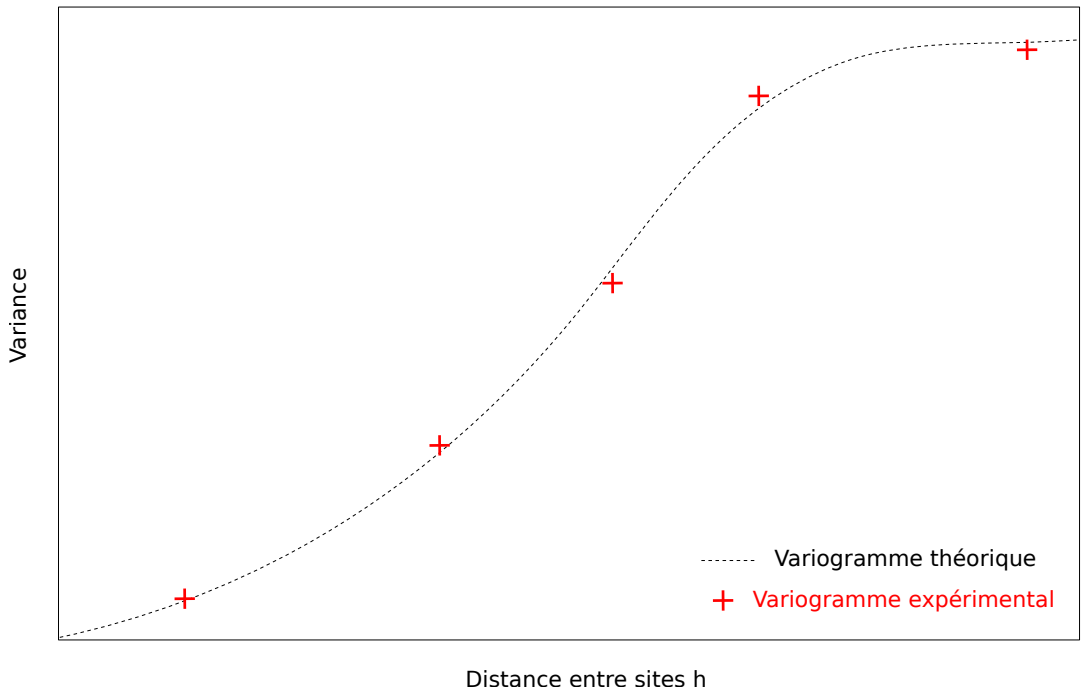
# Analyse variographique

Le pas $dh$ du variogramme expérimental, doit être un compromis :



# Analyse variographique

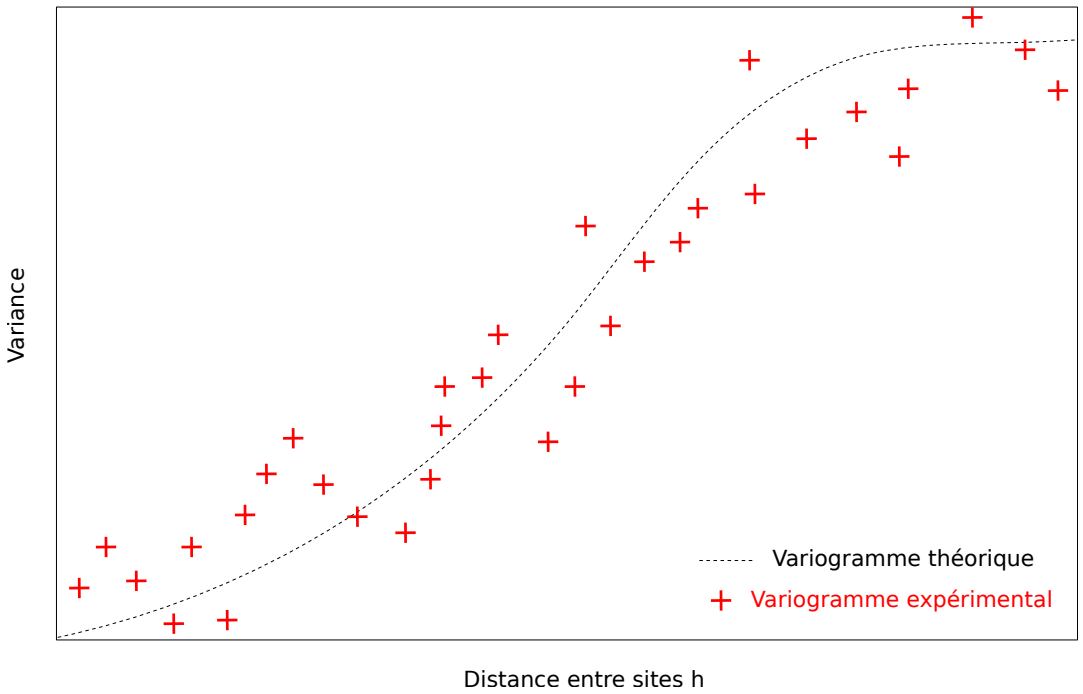
Le pas $dh$ du variogramme expérimental, doit être un compromis :




Si $dh$ est trop grand, on a trop peu de points pour modéliser le comportement de la variable régionalisée : le variogramme expérimental est très précis mais peu résolu

# Analyse variographique

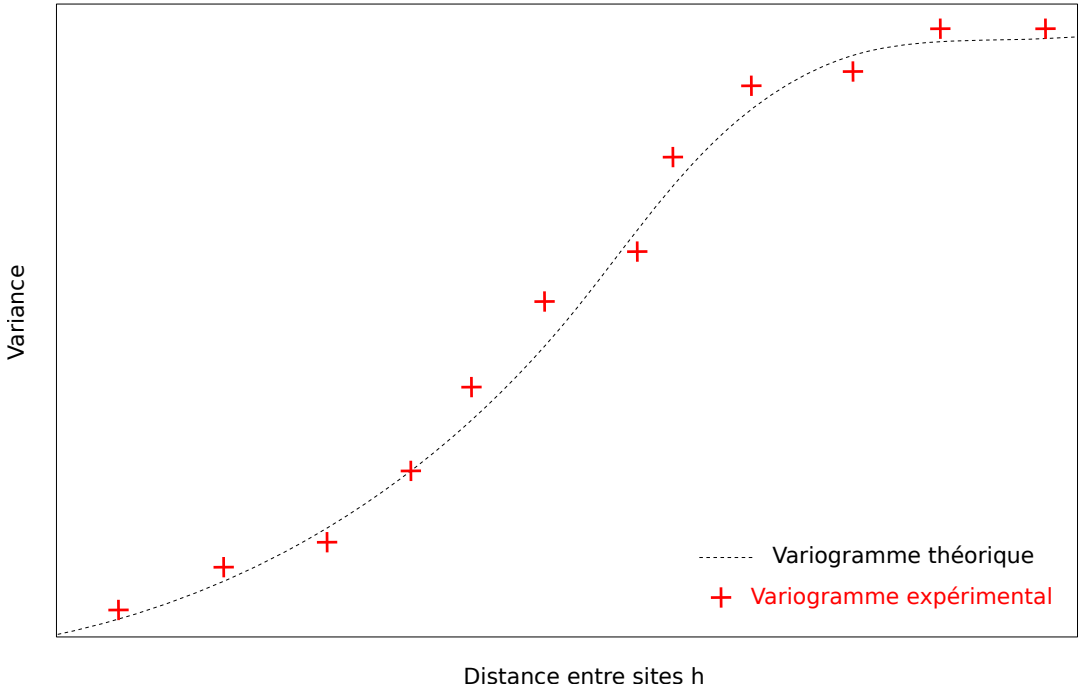
Le pas $dh$ du variogramme expérimental, doit être un compromis :




Si $dh$ est trop petit, chaque point du variogramme expérimental est calculé avec trop peu de couples $(x_i, x_j)$ : le variogramme expérimental est très résolu mais peu précis

# Analyse variographique

Le pas $dh$ du variogramme expérimental, doit être un compromis :




Le pas $dh$ est idéal quand il permet d'obtenir un bon compromis **robustesse** $\leftrightarrow$ **résolution**

# Analyse variographique

## Théorème

Le variogramme expérimental est un estimateur **non-biaisé** du variogramme théorique :

$$\forall\ h \in \mathbb{R}^+ \quad \mathbb{E}[\hat{\gamma}(h)] = \gamma(h)$$

# Analyse variographique

**Théorème**

Le variogramme expérimental est un estimateur **non-biaisé** du variogramme théorique :

$$\forall\ h \in \mathbb{R}^+ \quad \mathbb{E}[\hat{\gamma}(h)] = \gamma(h)$$

⚠ En règle générale, la dispersion du variogramme expérimental autour du variogramme théorique est d'autant plus grande que $h$ est grand : $\text{Var}[\hat{\gamma}(h)] = \hat{\gamma}(h)^2 N(h)^{-1}$

# Analyse variographique

**Théorème**

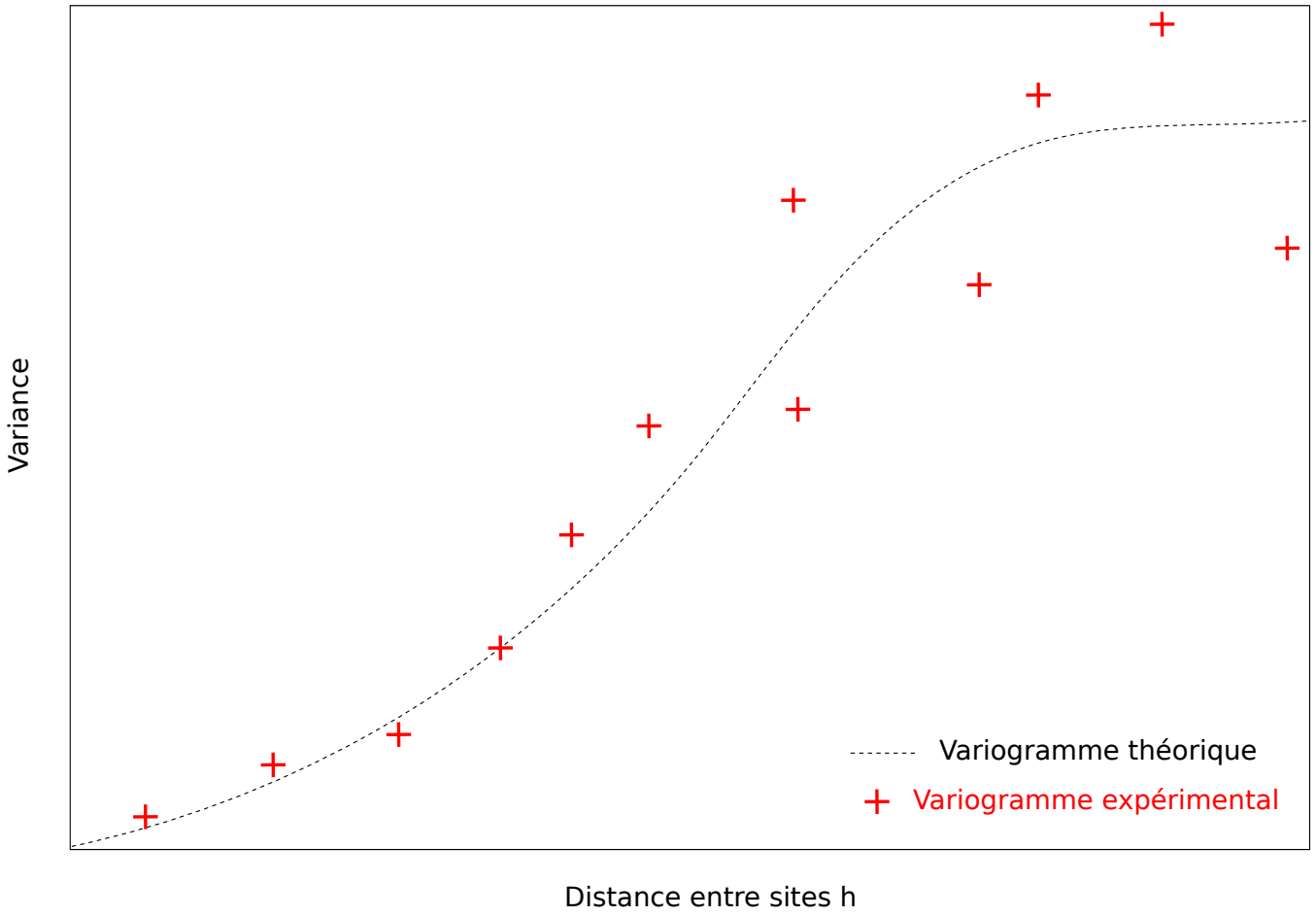
Le variogramme expérimental est un estimateur **non-biaisé** du variogramme théorique :

$$\forall\ h \in \mathbb{R}^+ \quad \mathbb{E}[\hat{\gamma}(h)] = \gamma(h)$$

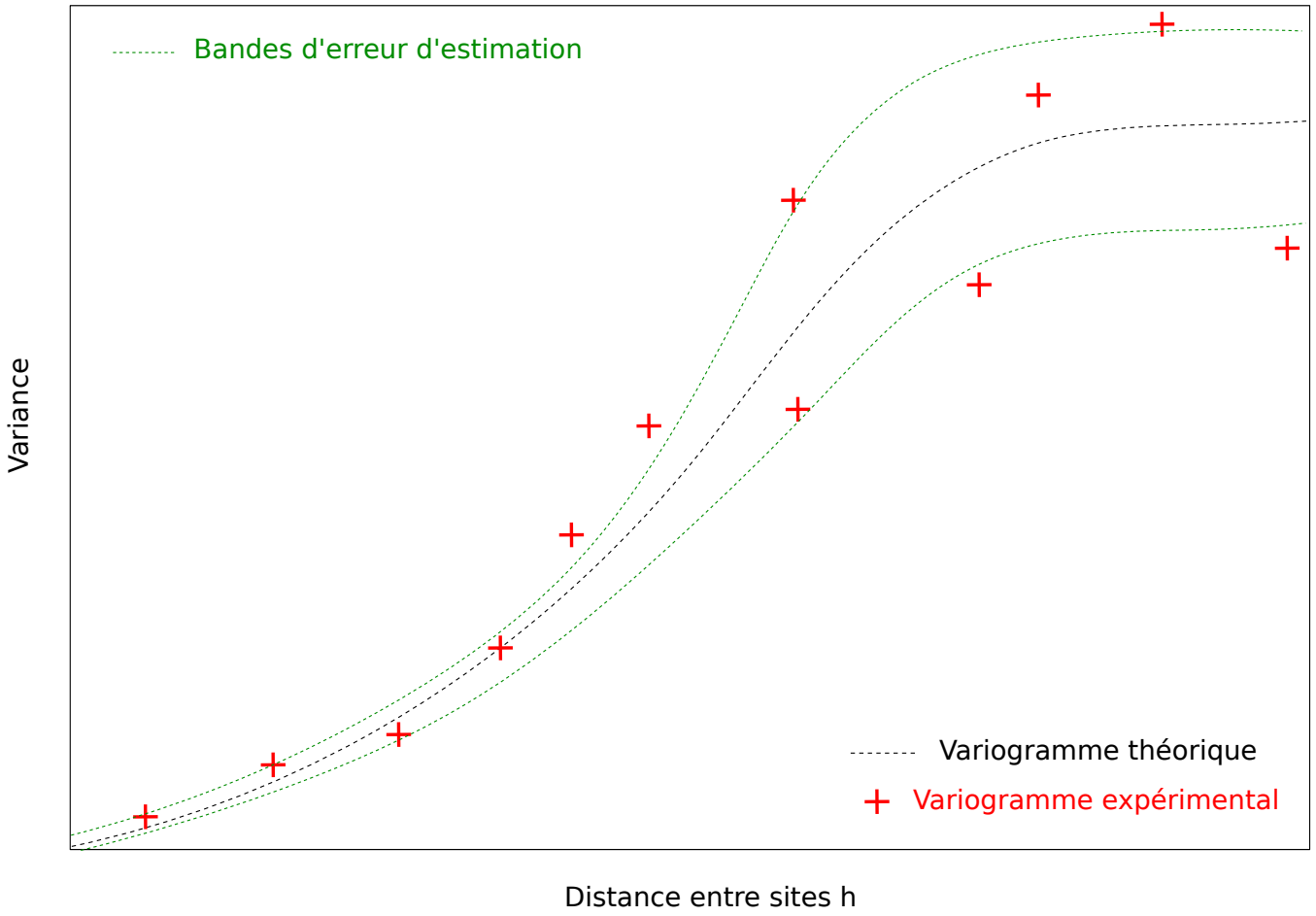
⚠ En règle générale, la dispersion du variogramme expérimental autour du variogramme théorique est d'autant plus grande que $h$ est grand : $\text{Var}[\hat{\gamma}(h)] = \hat{\gamma}(h)^2 N(h)^{-1}$

En pratique, on considérera que $\hat{\gamma}$ n'est plus un estimateur robuste de $\gamma$ pour des distances de séparation $h$ supérieures à la moitié du diamètre du champ $\mathcal{D}$.

# Analyse variographique



# Analyse variographique



# Analyse variographique

Propriétés du variogramme théorique :

# Analyse variographique

Propriétés du variogramme théorique :

**Propriété 1**

Le variogramme est nul à l'origine : $\gamma(0) = 0$

# Analyse variographique

Propriétés du variogramme théorique :

**Propriété 1**

Le variogramme est nul à l'origine : $\gamma(0) = 0$

$$\gamma(0) = \gamma(x, x) = \text{Var}[Z(x) - Z(x)] = \text{Var}[0] = 0$$

# Analyse variographique

Propriétés du variogramme théorique :

**Propriété 1**

Le variogramme est nul à l'origine : $\gamma(0) = 0$

$$\gamma(0) = \gamma(x, x) = \text{Var}[Z(x) - Z(x)] = \text{Var}[0] = 0$$

**Propriété 2 (pour un champ stationnaire)**

Variogramme et covariance sont liés par la relation :

$$\gamma(h) = C(0) - C(h)$$

avec $C(h)$ la covariance en 2 sites séparés d'une distance $h$.

# Analyse variographique

Propriétés du variogramme théorique :

**Propriété 1**

Le variogramme est nul à l'origine : $\gamma(0) = 0$

$\gamma(0) = \gamma(x, x) = \text{Var}[Z(x) - Z(x)] = \text{Var}[0] = 0$

**Propriété 2 (pour un champ stationnaire)**

Variogramme et covariance sont liés par la relation :

$$\gamma(h) = C(0) - C(h)$$

avec $C(h)$ la covariance en 2 sites séparés d'une distance $h$.

(valide ssi le variogramme atteint un palier : $\gamma(h) \underset{h \to +\infty}{\to} C(0)$)

# Analyse variographique

**Preuve :**

$$\gamma(h) = \frac{1}{2}\mathsf{Var}[Z(x) - Z(x+h)]$$

# Analyse variographique

**Preuve :**

$$\gamma(h) = \frac{1}{2}\mathsf{Var}[Z(x) - Z(x+h)]$$

$$\gamma(h) = \frac{1}{2}\big(\mathsf{Var}[Z(x)] + \mathsf{Var}[Z(x+h)] - 2\mathsf{Cov}[Z(x), Z(x+h)]\big)$$

# Analyse variographique

**Preuve :**

$$\gamma(h) = \frac{1}{2}\mathsf{Var}[Z(x) - Z(x+h)]$$

$$\gamma(h) = \frac{1}{2}\big(\mathsf{Var}[Z(x)] + \mathsf{Var}[Z(x+h)] - 2\mathsf{Cov}[Z(x), Z(x+h)]\big)$$

$$\gamma(h) = \frac{1}{2}\big(2\mathsf{Var}[Z(x)] - 2\mathsf{Cov}[Z(x), Z(x+h)]\big)$$

# Analyse variographique

**Preuve :**

$$\gamma(h) = \frac{1}{2}\mathsf{Var}[Z(x) - Z(x+h)]$$

$$\gamma(h) = \frac{1}{2}\big(\mathsf{Var}[Z(x)] + \mathsf{Var}[Z(x+h)] - 2\mathsf{Cov}[Z(x), Z(x+h)]\big)$$

$$\gamma(h) = \frac{1}{2}\big(2\mathsf{Var}[Z(x)] - 2\mathsf{Cov}[Z(x), Z(x+h)]\big)$$

$$\gamma(h) = \sigma_Z^2 - C(h) = C(0) - C(h)$$

avec : $C(h) = \mathsf{Cov}[Z(x), Z(x+h)]$

# Analyse variographique

Propriétés du variogramme théorique :

## Propriété 3 (admise)

Le comportement à l'origine du variogramme reflète le degré de régularité de la variable régionalisée :

**Discontinu** : la variable régionalisée est discontinue*

**Linéaire** : la variable régionalisée est continue mais non-dérivable*

**Parabolique** : la variable régionalisée est dérivable*

# Analyse variographique

Propriétés du variogramme théorique :

## Propriété 3 (admise)

Le comportement à l'origine du variogramme reflète le degré de régularité de la variable régionalisée :

**Discontinu :** la variable régionalisée est discontinue*

**Linéaire :** la variable régionalisée est continue mais non-dérivable*

**Parabolique :** la variable régionalisée est dérivable*

*Fonctions aléatoires $\rightarrow$ il faut introduire les notions probabilistes de continuité et de dérivabilité *en moyenne quadratique*.

# Analyse variographique

Propriétés du variogramme théorique :

## Propriété 4

Le variogramme est **conditionnellement défini-négatif** : pour tout $n$-uplet de points du champ : $(x_1, x_2, ..., x_n) \in \mathcal{D}^n$ et pour tout $n$-uplet de poids $(\lambda_1, \lambda_2..., \lambda_n) \in \mathbb{R}^n$ on a l'implication :

$$\sum_{i=1}^{n} \lambda_i = 0 \;\Rightarrow\; \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j \gamma(x_i - x_j) \leqslant 0$$

# Analyse variographique

Propriétés du variogramme théorique :

## Propriété 4

Le variogramme est **conditionnellement défini-négatif** : pour tout $n$-uplet de points du champ : $(x_1, x_2, ..., x_n) \in \mathcal{D}^n$ et pour tout $n$-uplet de poids $(\lambda_1, \lambda_2..., \lambda_n) \in \mathbb{R}^n$ on a l'implication :

$$\sum_{i=1}^{n} \lambda_i = 0 \;\Rightarrow\; \mathbf{\Lambda \Gamma \Lambda}^T \leqslant 0$$

# Analyse variographique

Propriétés du variogramme théorique :

## Propriété 4

Le variogramme est **conditionnellement défini-négatif** : pour tout $n$-uplet de points du champ : $(x_1, x_2, ..., x_n) \in \mathcal{D}^n$ et pour tout $n$-uplet de poids $(\lambda_1, \lambda_2..., \lambda_n) \in \mathbb{R}^n$ on a l'implication :

$$\sum_{i=1}^{n} \lambda_i = 0 \;\Rightarrow\; \mathbf{\Lambda \Gamma \Lambda}^T \leqslant 0$$

**Elément de preuve :** si la covariance existe, cette propriété découle directement du fait que la quantité $\text{Var}(\mathbf{\Lambda Z}^T) = \mathbf{\Lambda C \Lambda}^T$ est positive. Sinon, on parvient au résultat en étudiant la variable auxiliaire $Y(x) = Z(x) - Z(x_0)$ qui elle possède une covariance.

# Analyse variographique

Propriétés du variogramme théorique :

## Propriété 4

Le variogramme est **conditionnellement défini-négatif** : pour tout $n$-uplet de points du champ : $(x_1, x_2, ..., x_n) \in \mathcal{D}^n$ et pour tout $n$-uplet de poids $(\lambda_1, \lambda_2..., \lambda_n) \in \mathbb{R}^n$ on a l'implication :

$$\sum_{i=1}^{n} \lambda_i = 0 \;\Rightarrow\; \mathbf{\Lambda \Gamma \Lambda}^T \leqslant 0$$

Exemple de fonction ne respectant pas la propriété de défini-négativité conditionnelle : la fonction échelon $\gamma(h) = \mathbb{1}_{\{h \leqslant h_0\}}$

# Analyse variographique

Propriétés du variogramme théorique :

## Propriété 4

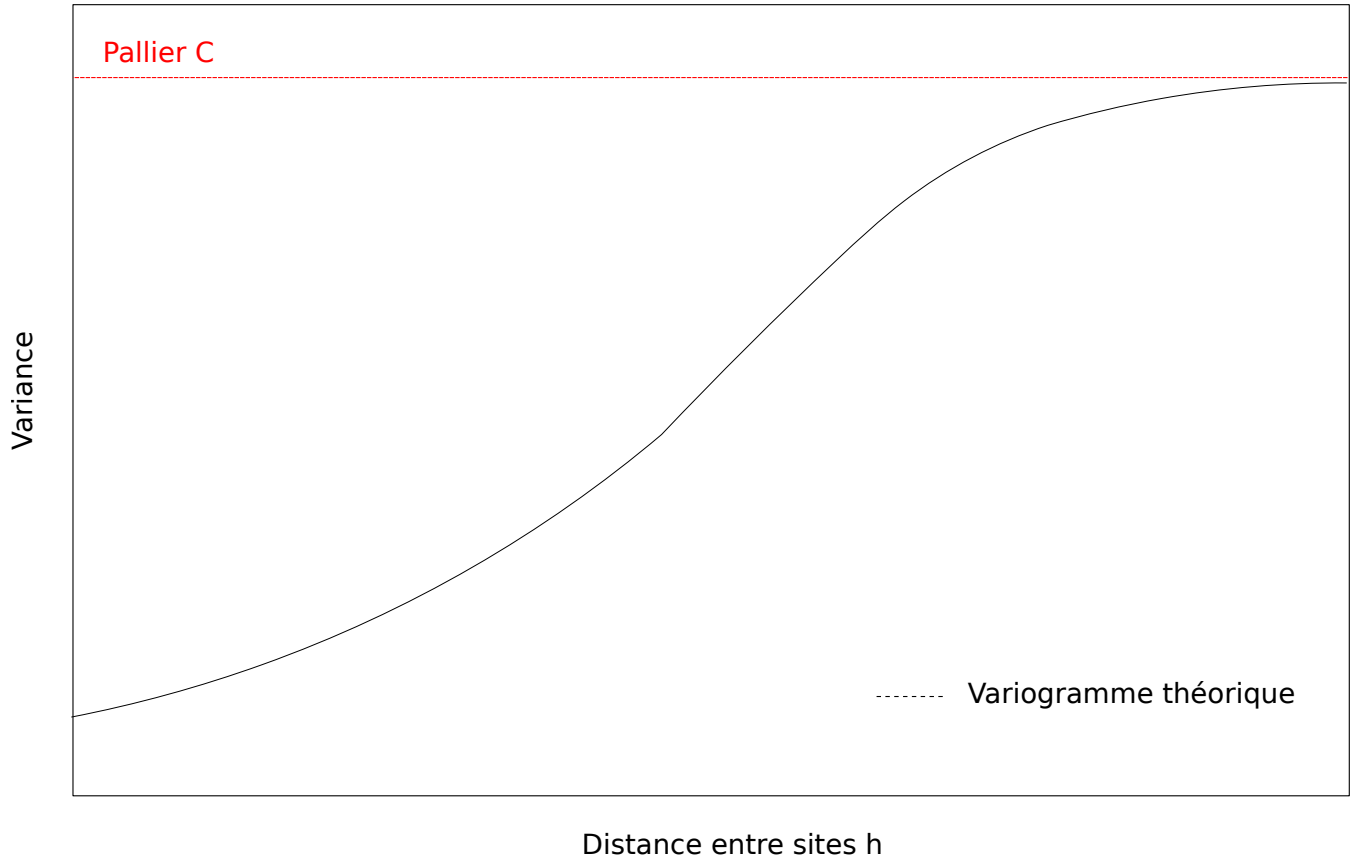
Le variogramme est **conditionnellement défini-négatif** : pour tout $n$-uplet de points du champ : $(x_1, x_2, ..., x_n) \in \mathcal{D}^n$ et pour tout $n$-uplet de poids $(\lambda_1, \lambda_2..., \lambda_n) \in \mathbb{R}^n$ on a l'implication :

$$\sum_{i=1}^{n} \lambda_i = 0 \;\Rightarrow\; \mathbf{\Lambda \Gamma \Lambda}^T \leqslant 0$$

$\rightarrow$ Le variogramme appartient à une classe de fonctions particulière

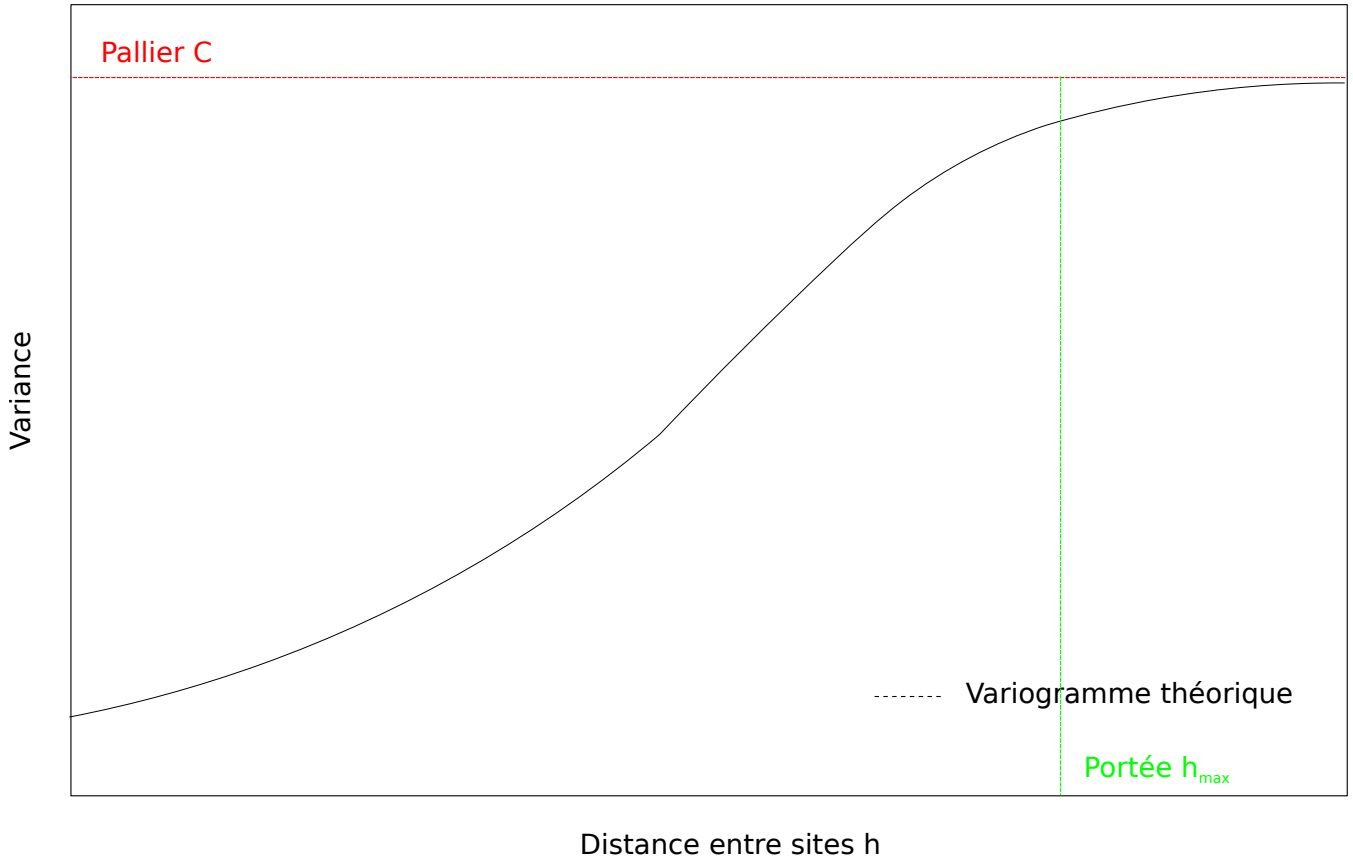
# Analyse variographique

## Structure du variogramme théorique



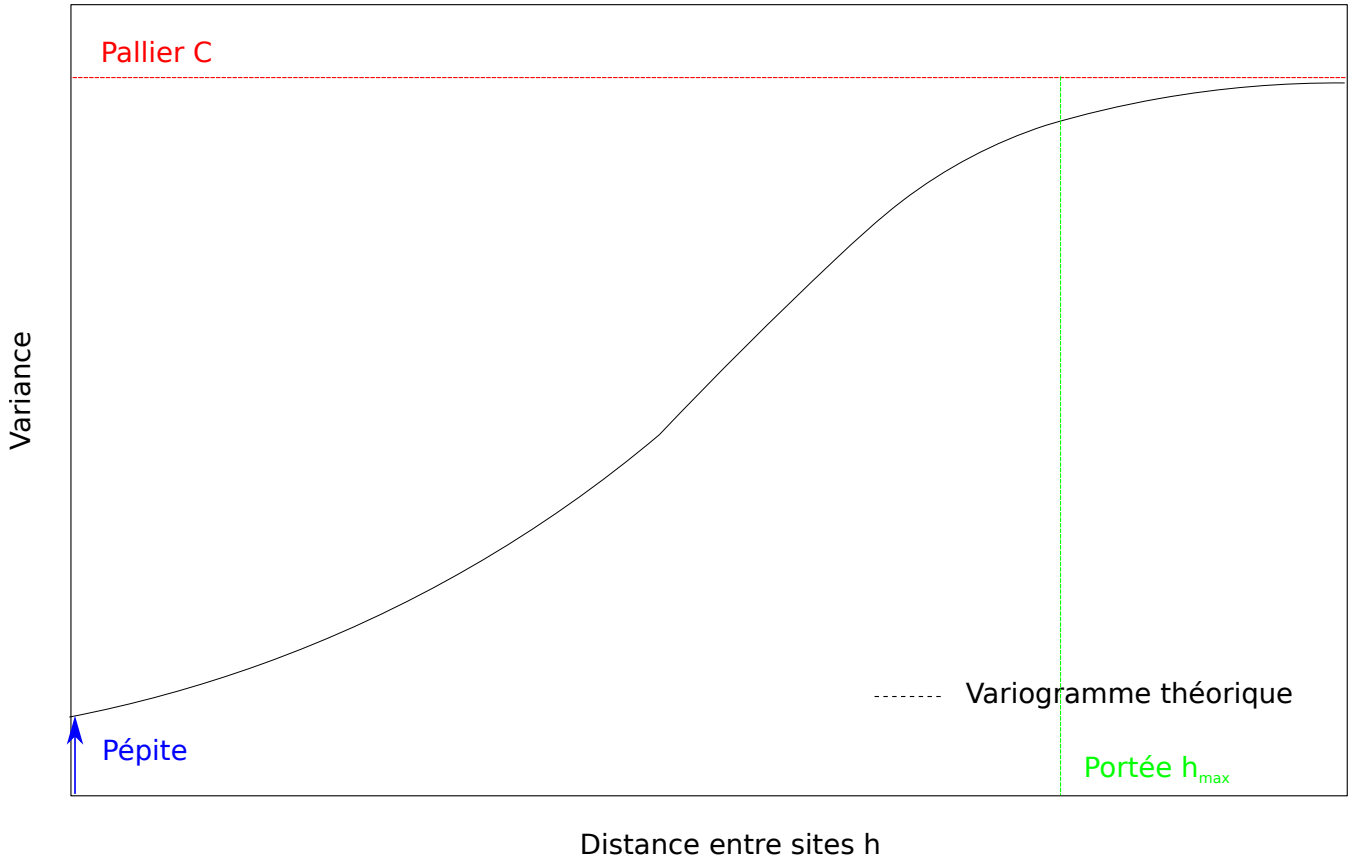
# Analyse variographique

## Structure du variogramme théorique



# Analyse variographique

## Structure du variogramme théorique



# Analyse variographique

## Définitions (pour un champ stationnaire)

Le **palier** : dissemblance maximale entre sites éloignés

$$\lim_{h \to +\infty} \gamma(h) = C(0) = C$$

# Analyse variographique

## Définitions (pour un champ stationnaire)

Le **palier** : dissemblance maximale entre sites éloignés

$$\lim_{h \to +\infty} \gamma(h) = C(0) = C$$

La **portée** : $h_{max}$ telle que $\gamma(h_{max}) \simeq C$, diamètre de la zone d'influence (zone de cohérence de la variable régionalisée).

# Analyse variographique

## Définitions (pour un champ stationnaire)

Le **palier** : dissemblance maximale entre sites éloignés

$$\lim_{h \to +\infty} \gamma(h) = C(0) = C$$

La **portée** : $h_{max}$ telle que $\gamma(h_{max}) \simeq C$, diamètre de la zone d'influence (zone de cohérence de la variable régionalisée).

La **pépite** : $\lim_{h \to 0^+} \gamma(h)$, discontinuité à l'origine.

# Analyse variographique

## Définitions (pour un champ stationnaire)

Le **palier** : dissemblance maximale entre sites éloignés

$$\lim_{h \to +\infty} \gamma(h) = C(0) = C$$

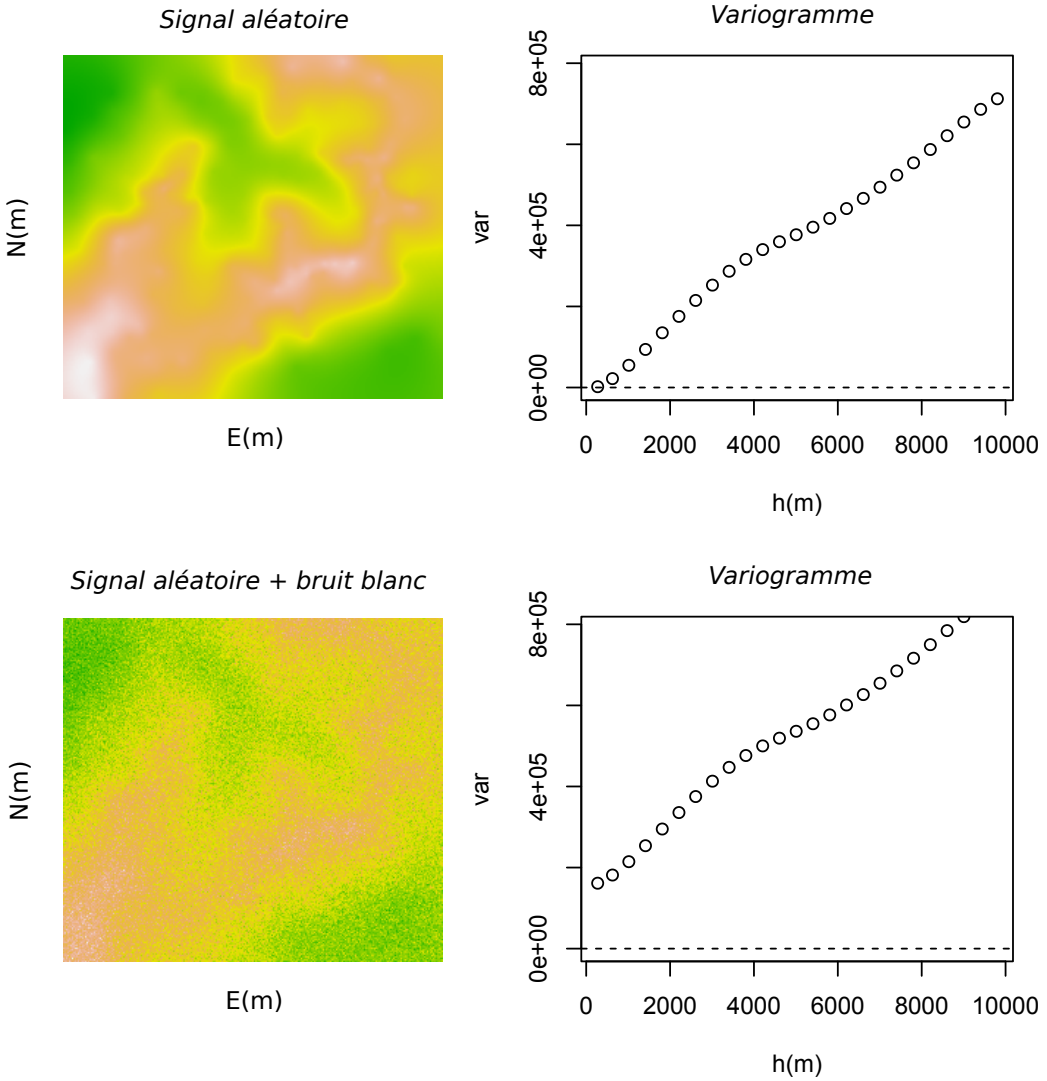
La **portée** : $h_{max}$ telle que $\gamma(h_{max}) \simeq C$, diamètre de la zone d'influence (zone de cohérence de la variable régionalisée).

La **pépite** : $\lim_{h \to 0^+} \gamma(h)$, discontinuité à l'origine.

Quand la pépite n'est pas nulle, on parle d'*effet de pépite*, se traduisant sur la variable régionalisée par un aspect localement erratique.

# Effet de pépite

pepite.r



# Les différents variogrammes

**Le variogramme théorique :** variogramme réel du phénomène que l'on cherche à modéliser. De par le nombre fini de points d'observation, il s'agit d'un objet conceptuel, inconnu en pratique, mais que l'on cherchera à estimer.

# Les différents variogrammes

**Le variogramme théorique :** variogramme réel du phénomène que l'on cherche à modéliser. De par le nombre fini de points d'observation, il s'agit d'un objet conceptuel, inconnu en pratique, mais que l'on cherchera à estimer.

**Le variogramme expérimental :** (ou variogramme empirique) estimation du variogramme théorique, calculée à l'aide des observations disponibles. C'est un estimateur bruité (variance statistique d'estimation) et fragmentaire (connu en quelques points seulement).

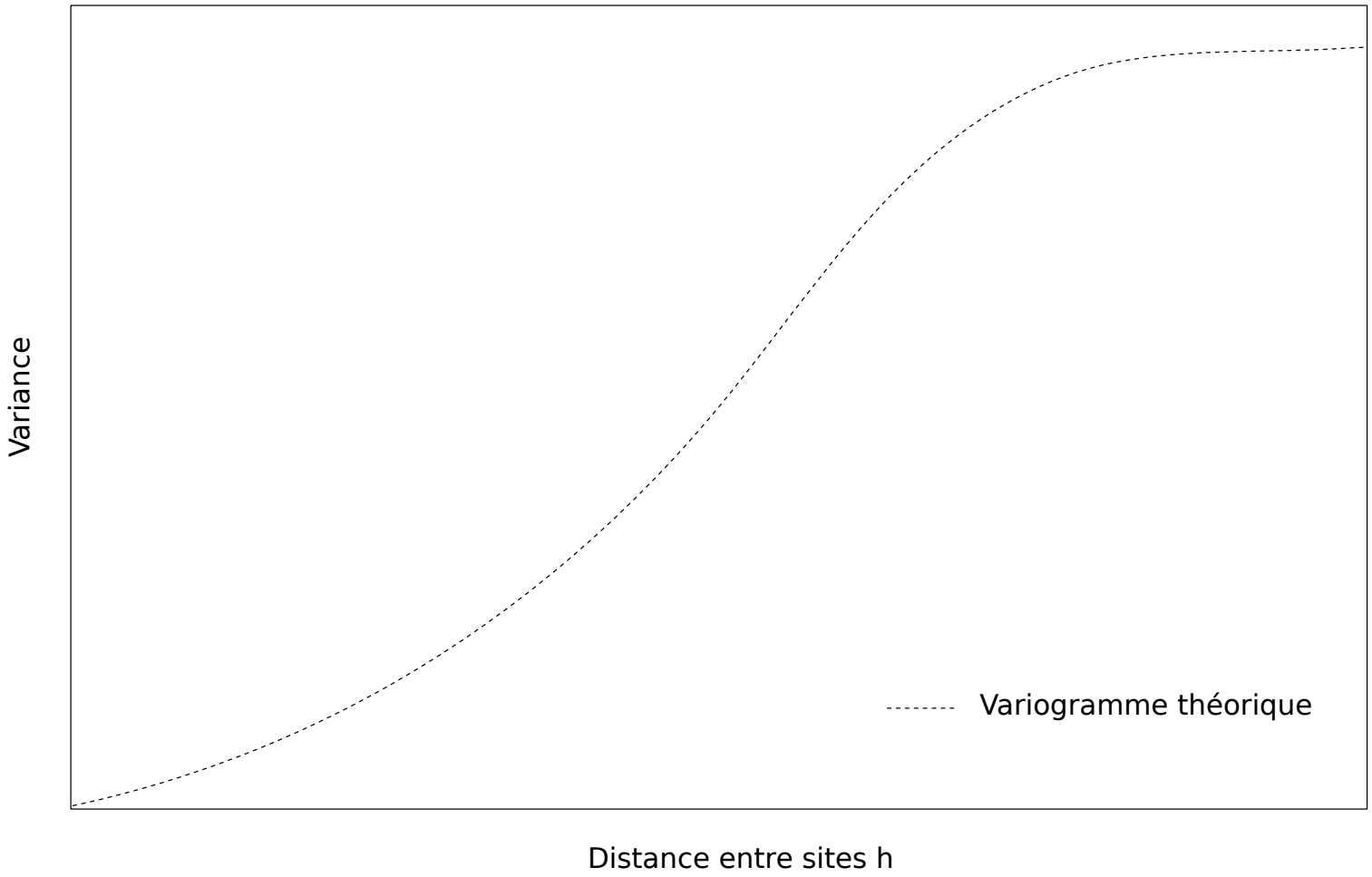
# Les différents variogrammes

**Le variogramme théorique :** variogramme réel du phénomène que l'on cherche à modéliser. De par le nombre fini de points d'observation, il s'agit d'un objet conceptuel, inconnu en pratique, mais que l'on cherchera à estimer.

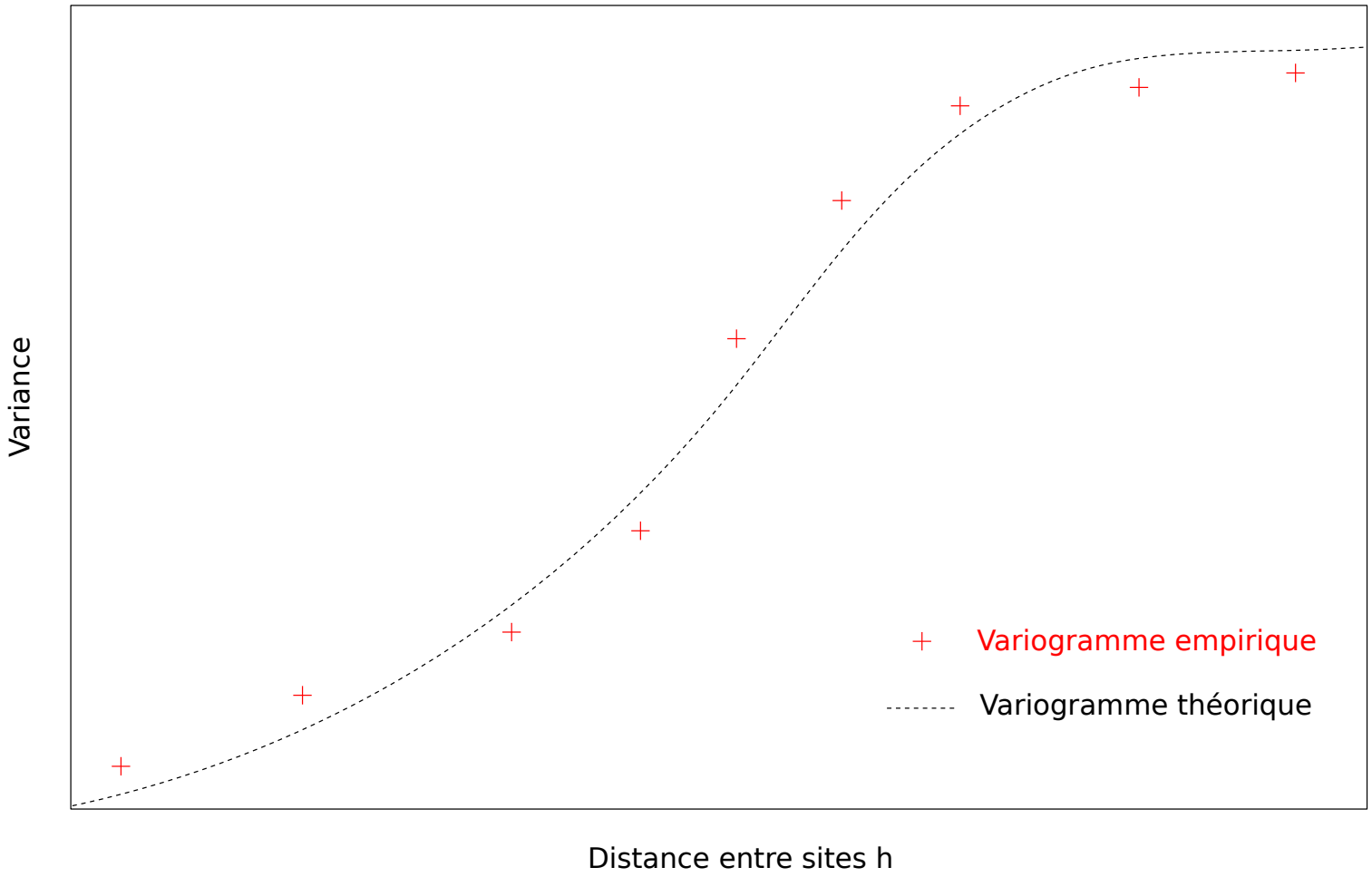
**Le variogramme expérimental :** (ou variogramme empirique) estimation du variogramme théorique, calculée à l'aide des observations disponibles. C'est un estimateur bruité (variance statistique d'estimation) et fragmentaire (connu en quelques points seulement).

**Les modèles de variogrammes :** bibliothèque de fonctions mathématiques satisfaisants les propriétés du variogramme théorique (défini-négativité conditionnelle), et dont les paramètres sont estimés à partir des observations. Les modèles de variogrammes permettent d'établir un pont entre le variogramme théorique *insaisissable* et son estimation bruitée, le variogramme expérimental.

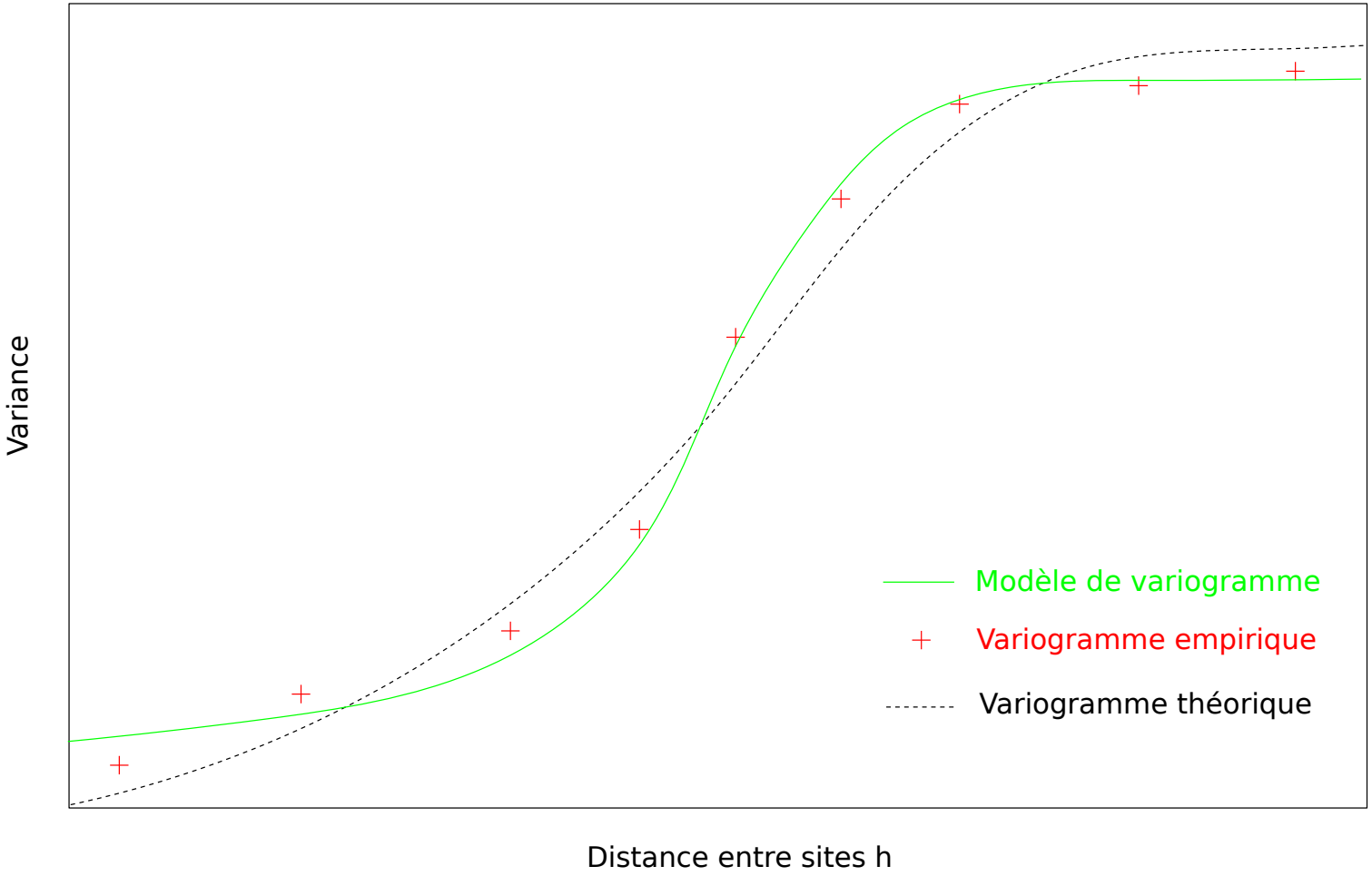
# Les modèles de variogramme



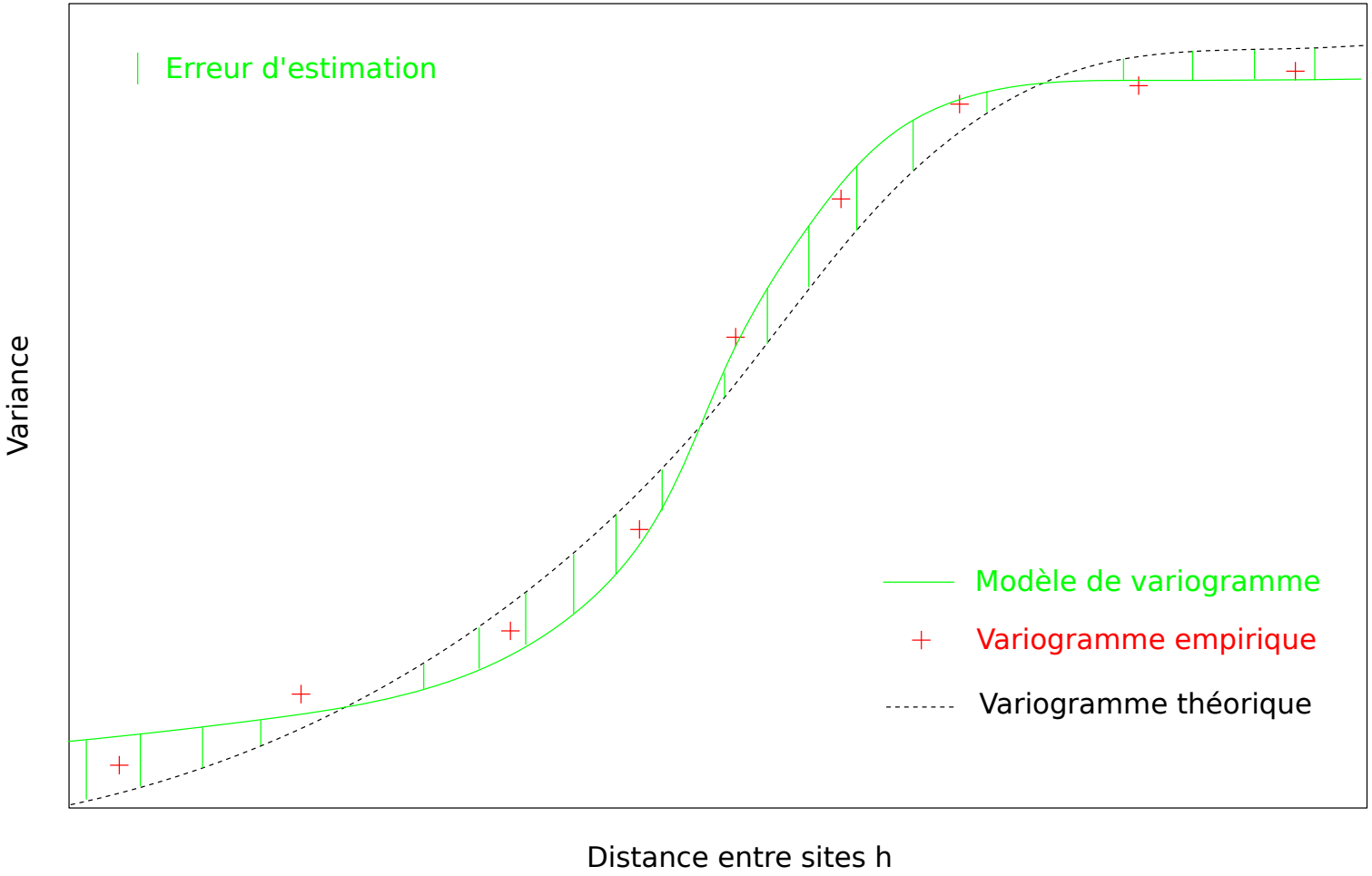
# Les modèles de variogramme



# Les modèles de variogramme



# Les modèles de variogramme

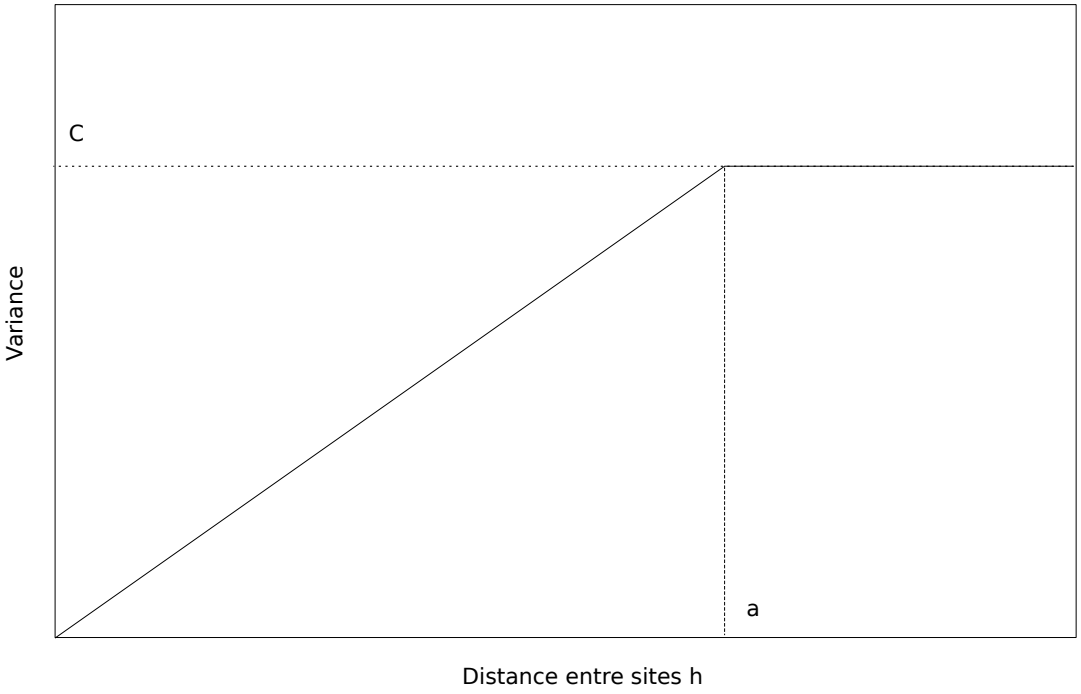


# Les modèles de variogramme

Modèle linéaire :

$$\gamma(h) = kh$$

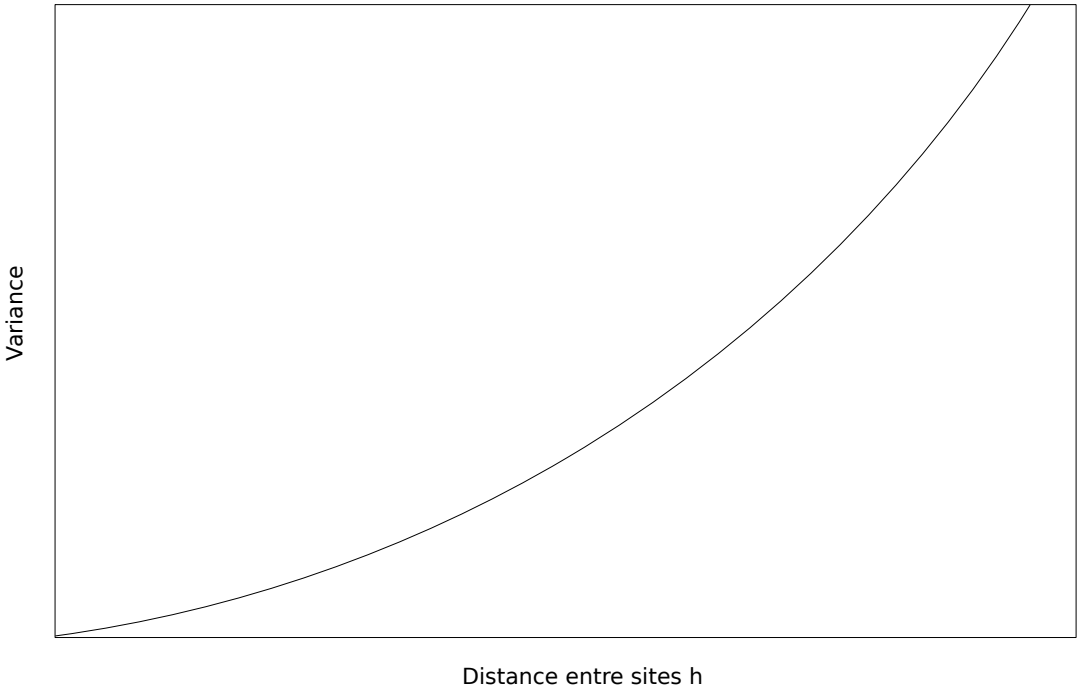
# Les modèles de variogramme




Modèle linéaire avec palier :

$$\gamma(h) = \begin{cases} \frac{C}{a}h & \text{si } h \leqslant a \\ C & \text{sinon.} \end{cases}$$

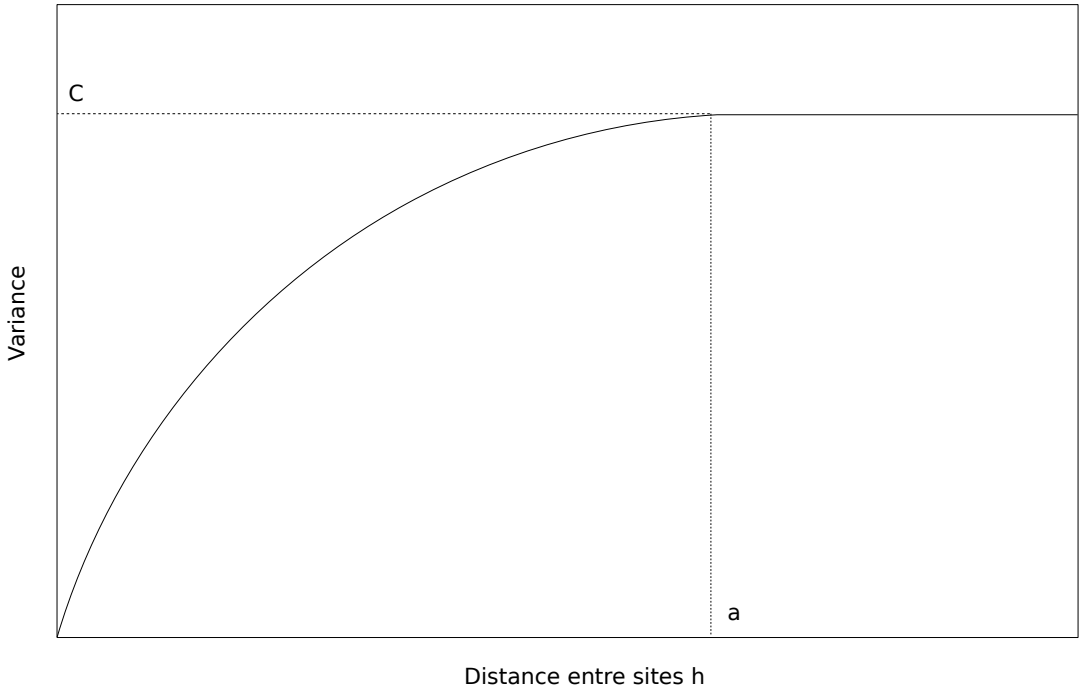
# Les modèles de variogramme




Modèle puissance :

$$\gamma(h) = kh^{\theta} \qquad \text{avec } 1 \leqslant \theta \leqslant 2$$

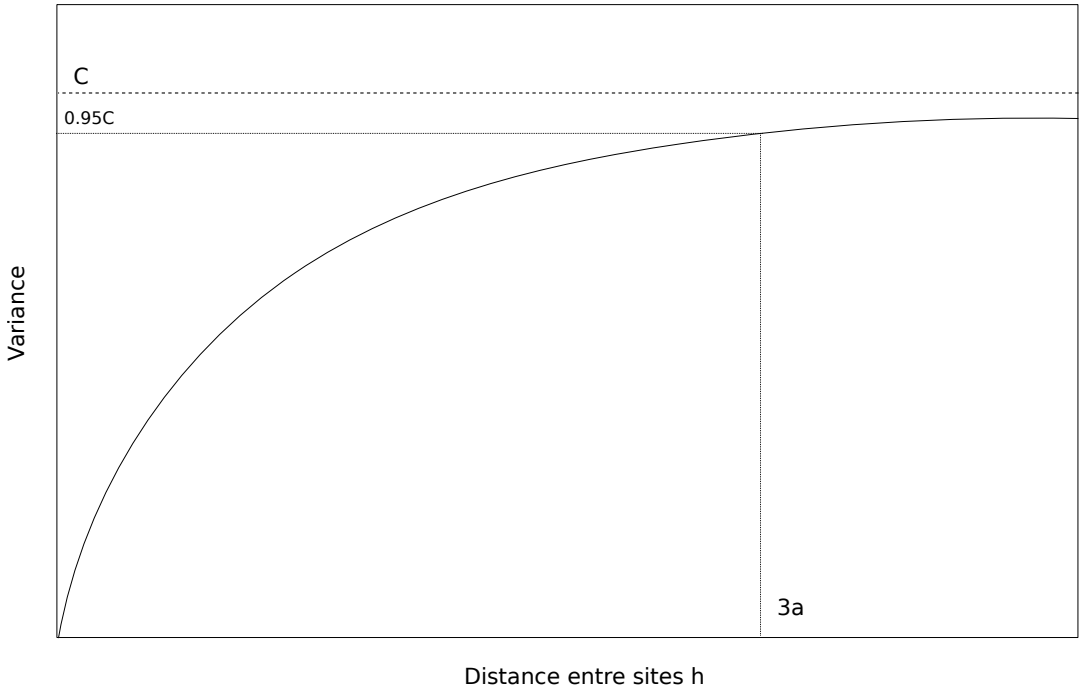
# Les modèles de variogramme




Modèle sphérique :

$$\gamma(h) = \begin{cases} C(\frac{3}{2}\frac{h}{a} - \frac{1}{2}\frac{h^3}{a^3}) & \text{si } h \leqslant a \\ C & \text{sinon.} \end{cases}$$

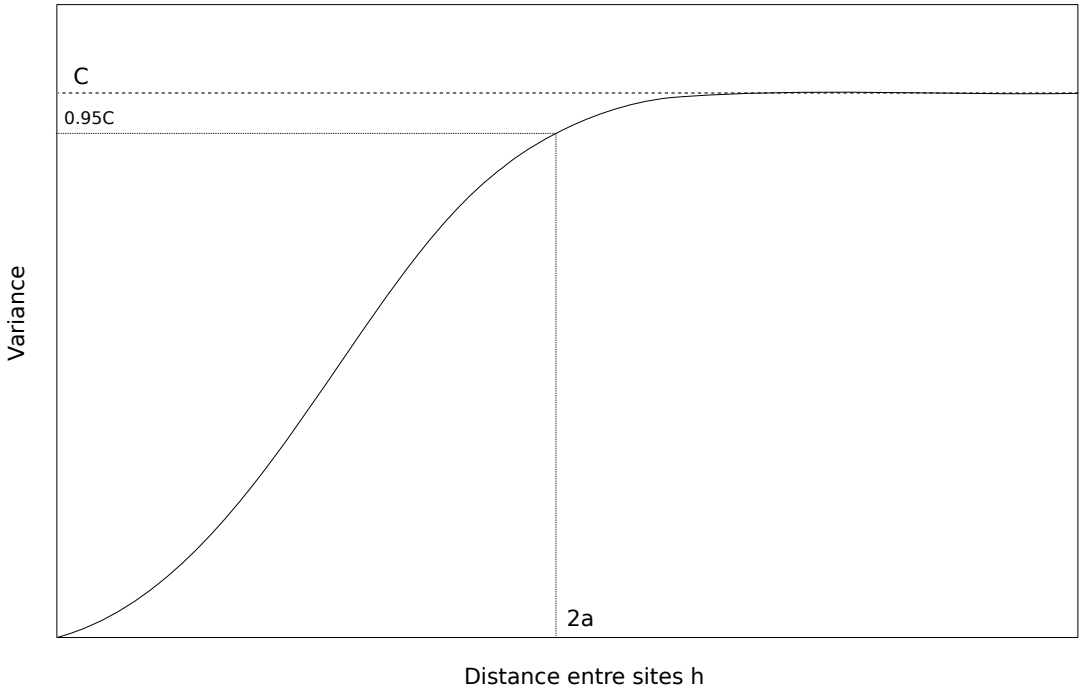
# Les modèles de variogramme




Modèle exponentiel :

$$\gamma(h) = C(1 - e^{-\frac{h}{a}})$$

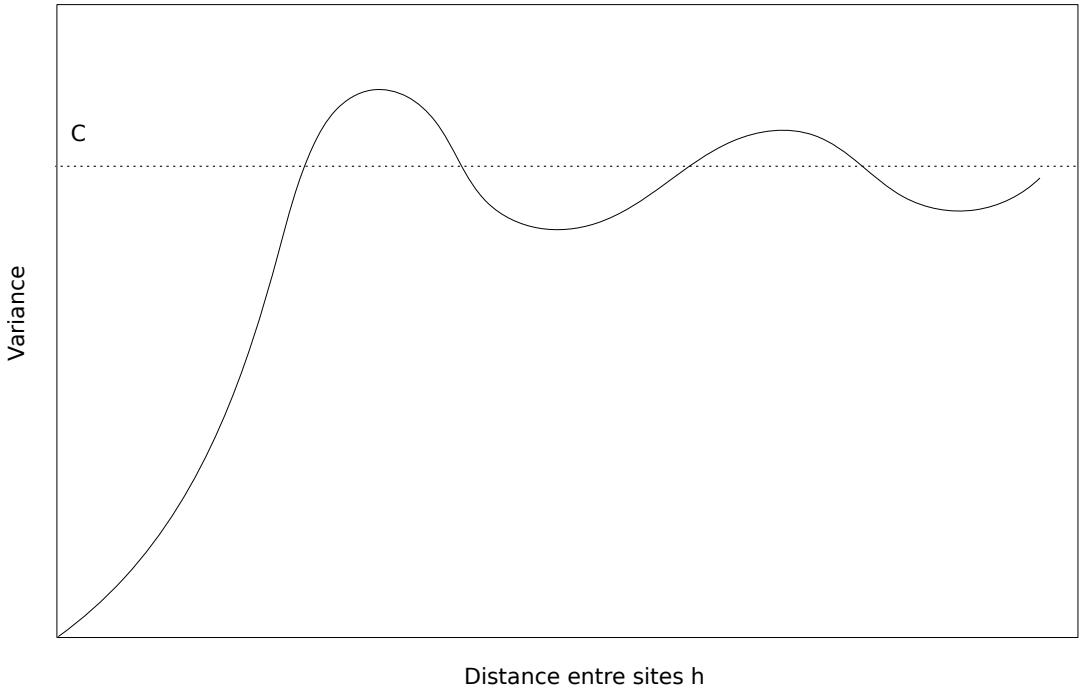
# Les modèles de variogramme




Modèle gaussien :

$$\gamma(h) = C(1 - e^{-\frac{h^2}{a^2}})$$

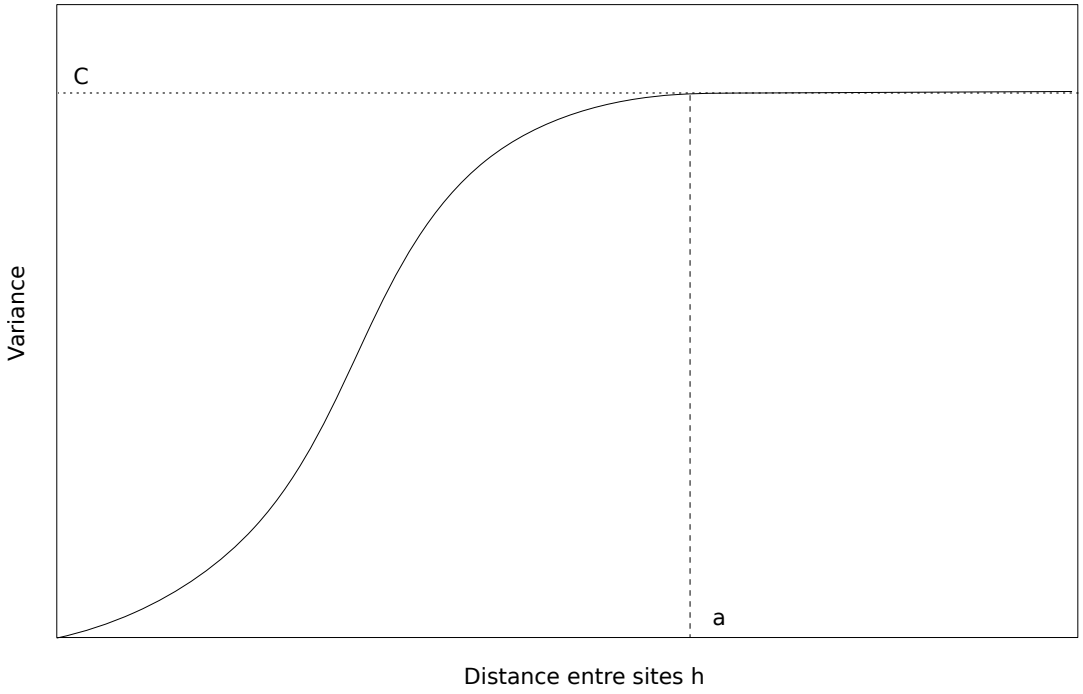
# Les modèles de variogramme




Modèle à effet de trou :

$$\gamma(h) = C(1 - \frac{sin(h/a)}{h/a})$$

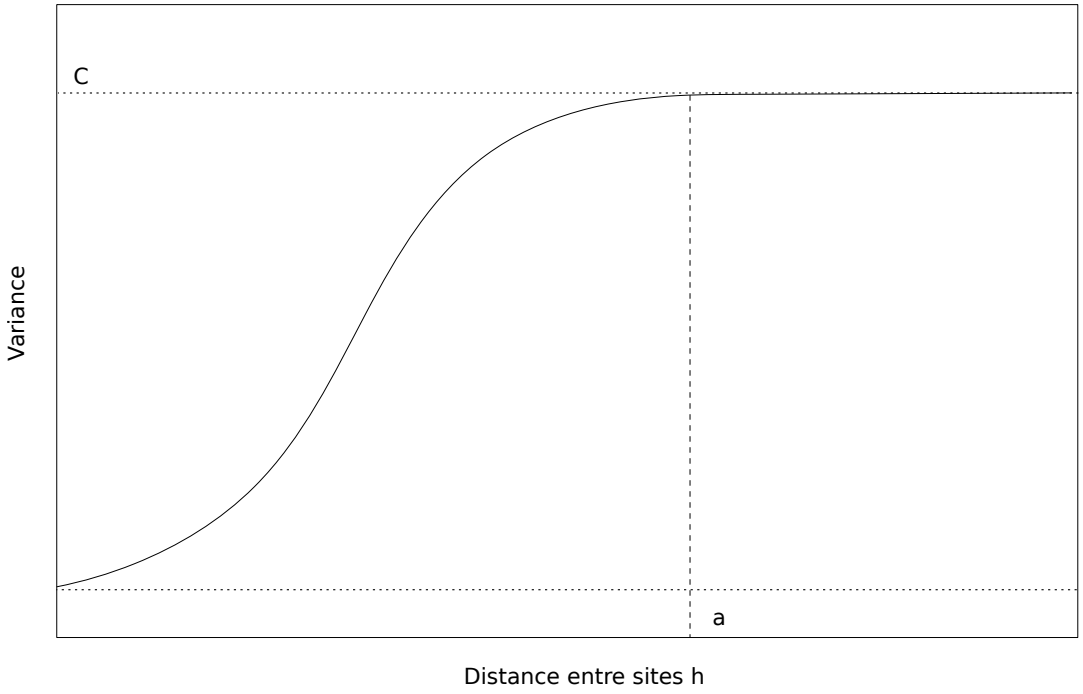
# Les modèles de variogramme




Modèle cubique :

$$\gamma(h) = \begin{cases} C(7\frac{h^2}{a^2} - \frac{35}{4}\frac{h^3}{a^3} + \frac{7}{2}\frac{h^5}{a^5} - \frac{3}{4}\frac{h^7}{a^7}) & \text{si } h \leqslant a \\ C & \text{sinon.} \end{cases}$$

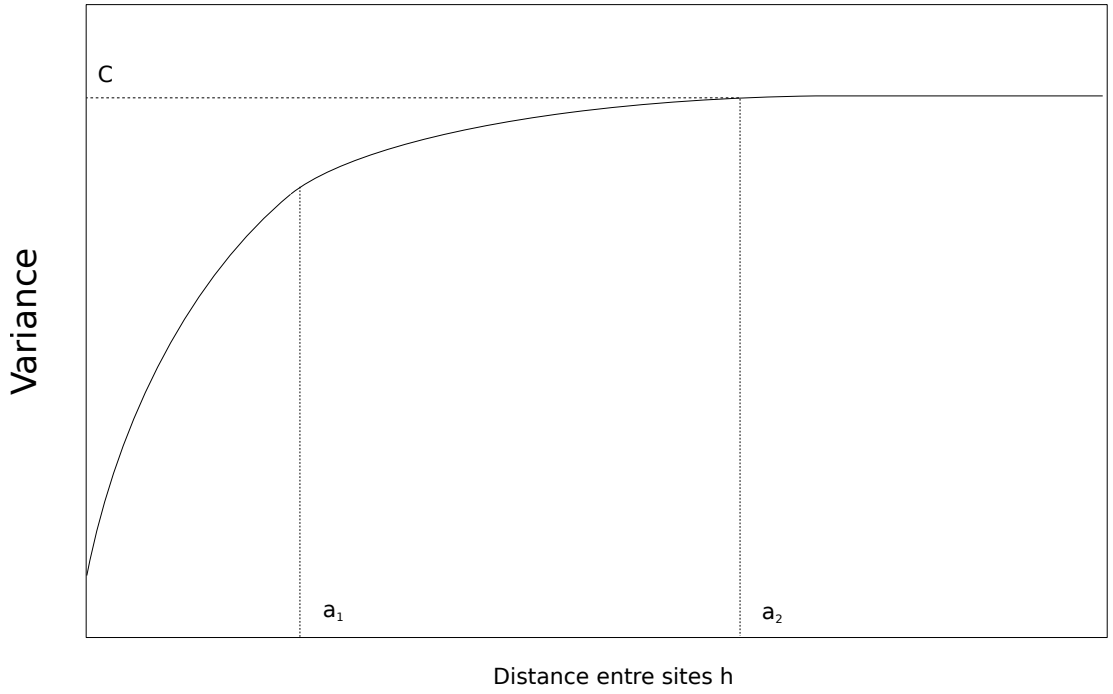
# Les modèles de variogramme




Modèle cubique avec pépite :

$$\gamma(h) = \begin{cases} C_0 + (C - C_0)(7\frac{h^2}{a^2} - \frac{35}{4}\frac{h^3}{a^3} + \frac{7}{2}\frac{h^5}{a^5} - \frac{3}{4}\frac{h^7}{a^7}) & \text{si } h \leqslant a \\ C & \text{sinon.} \end{cases}$$

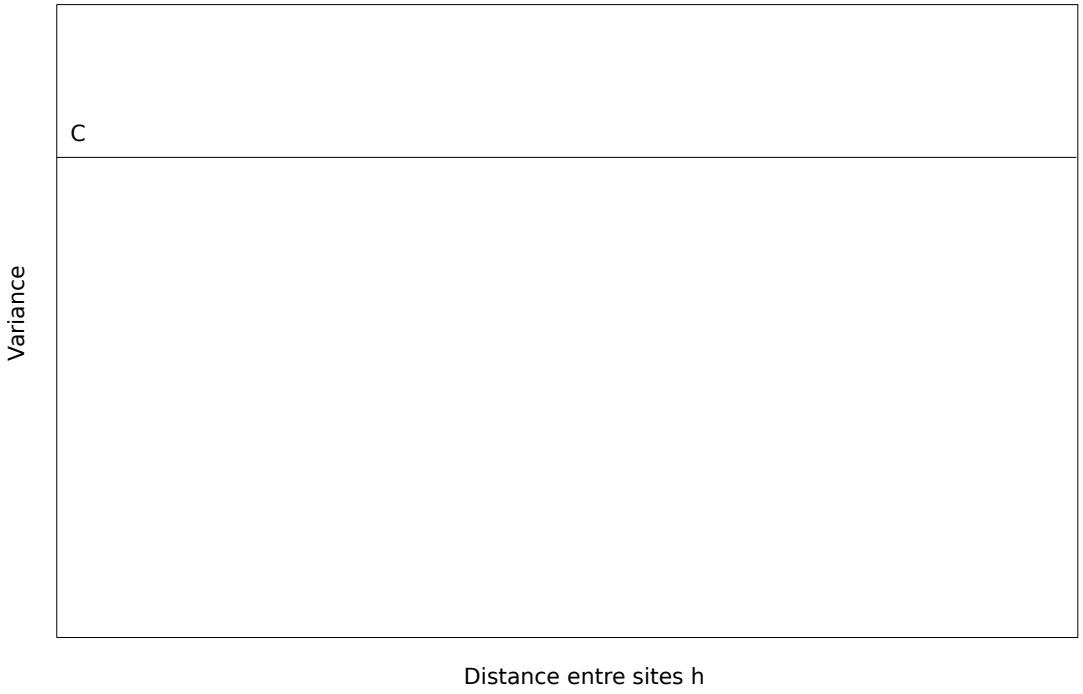
# Les modèles de variogramme




Modèle en gigogne :

$$\gamma(h) = \gamma_1(h) + \gamma_2(h) + ... + \gamma_n(h)$$

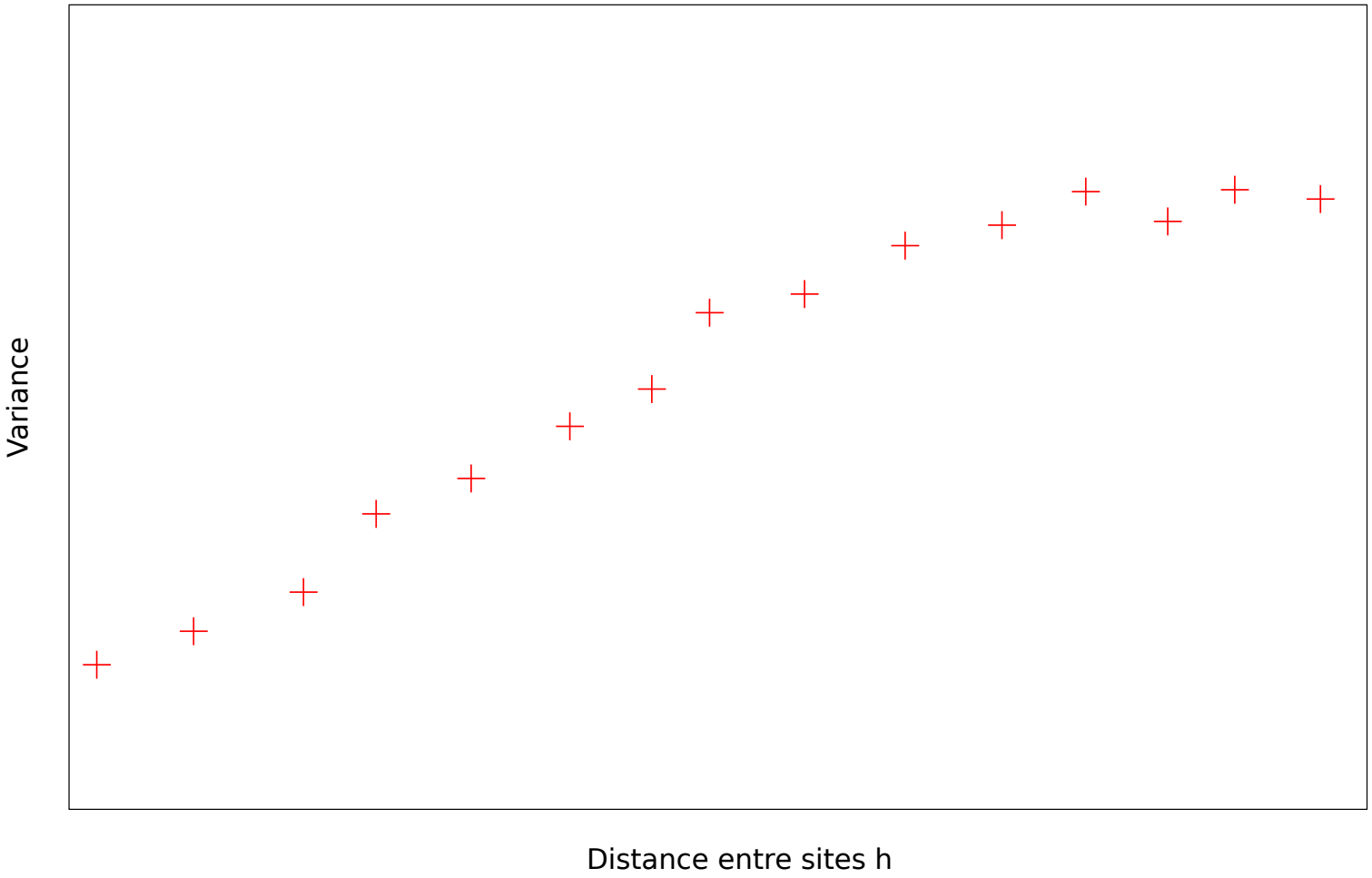
# Les modèles de variogramme



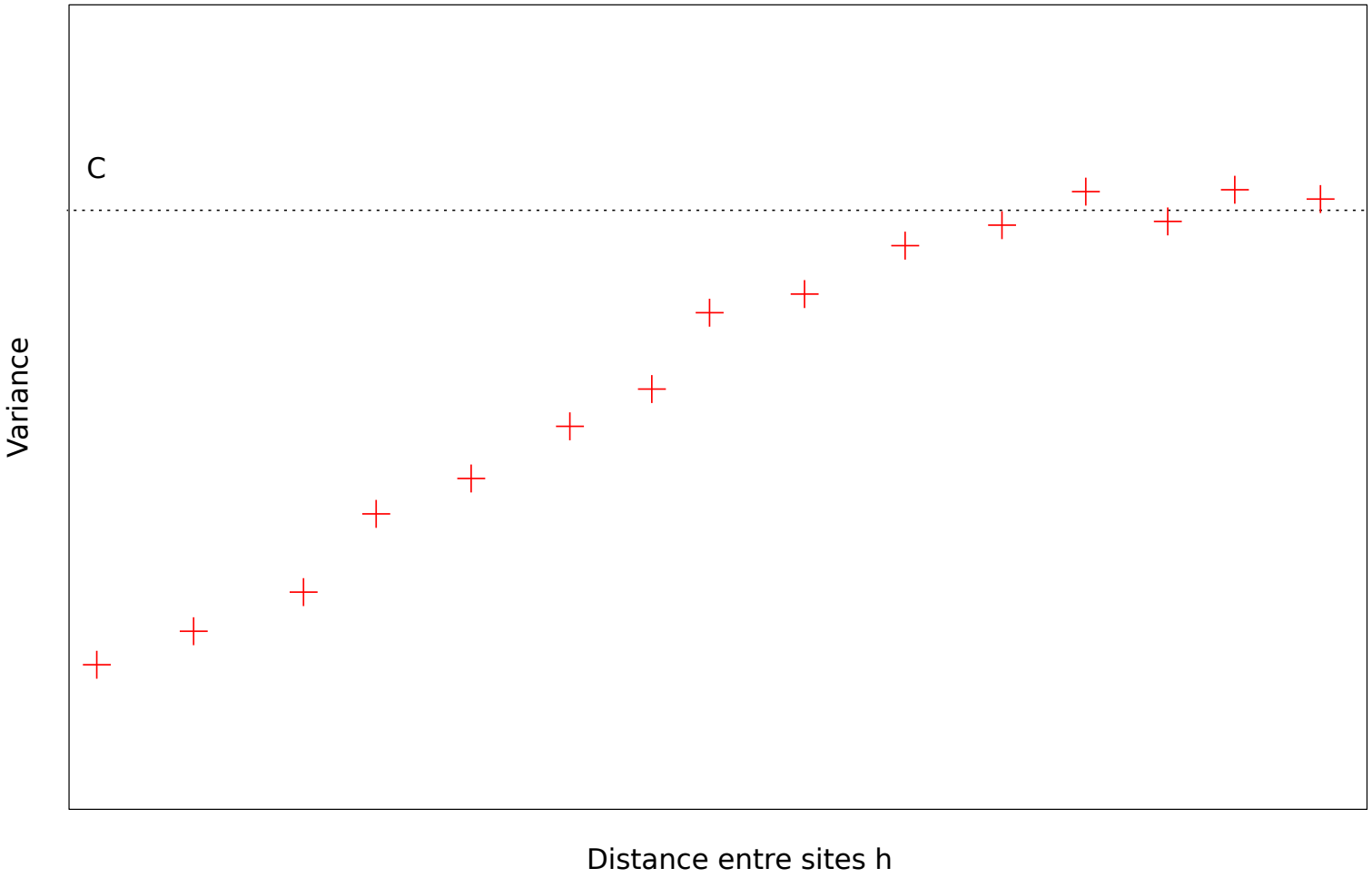

Modèle pépitique pur :

$$\gamma(h) = \begin{cases} 0 & \text{si } h = 0 \\ C & \text{sinon.} \end{cases}$$

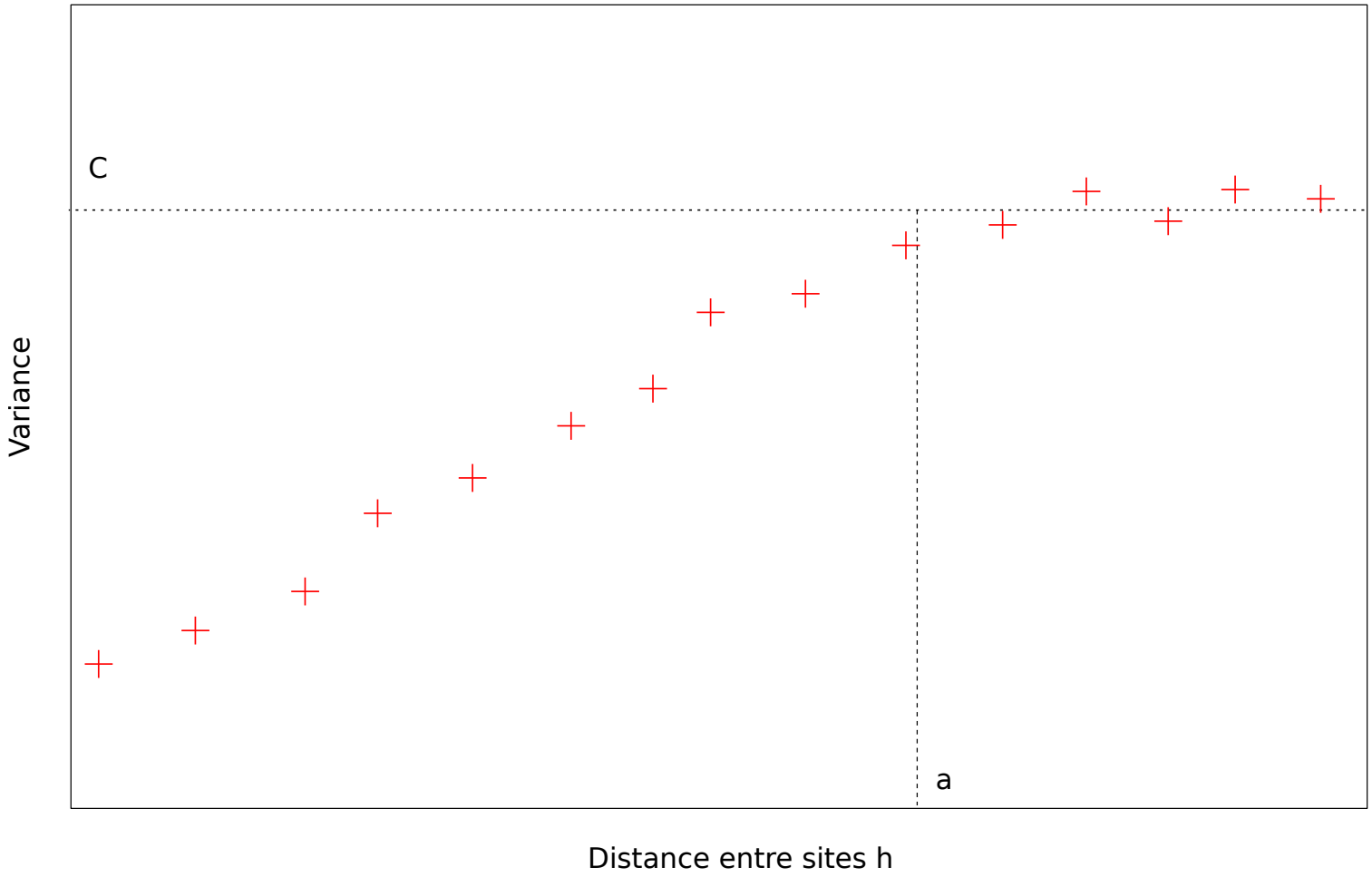
# Les modèles de variogramme



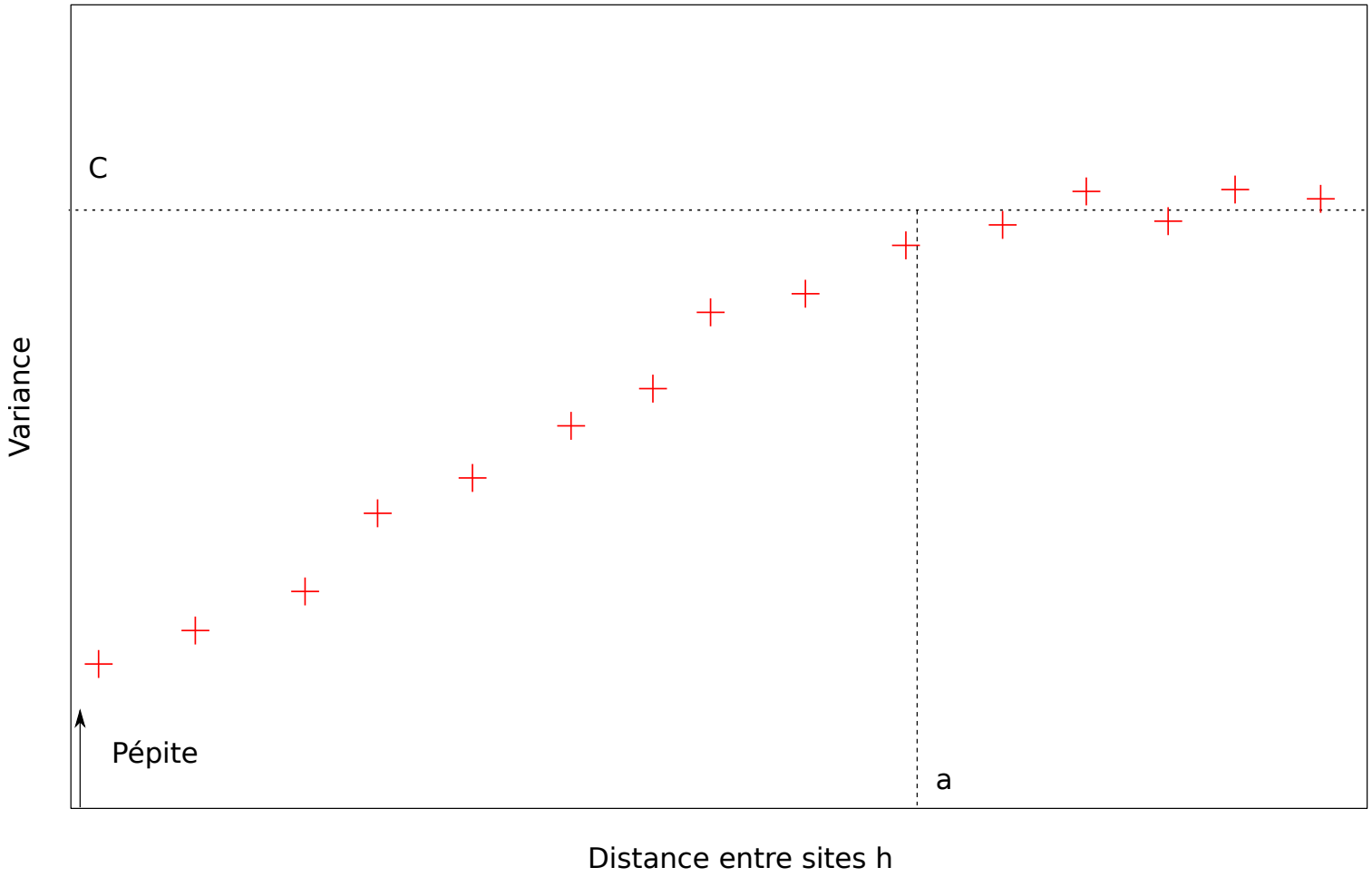
# Les modèles de variogramme



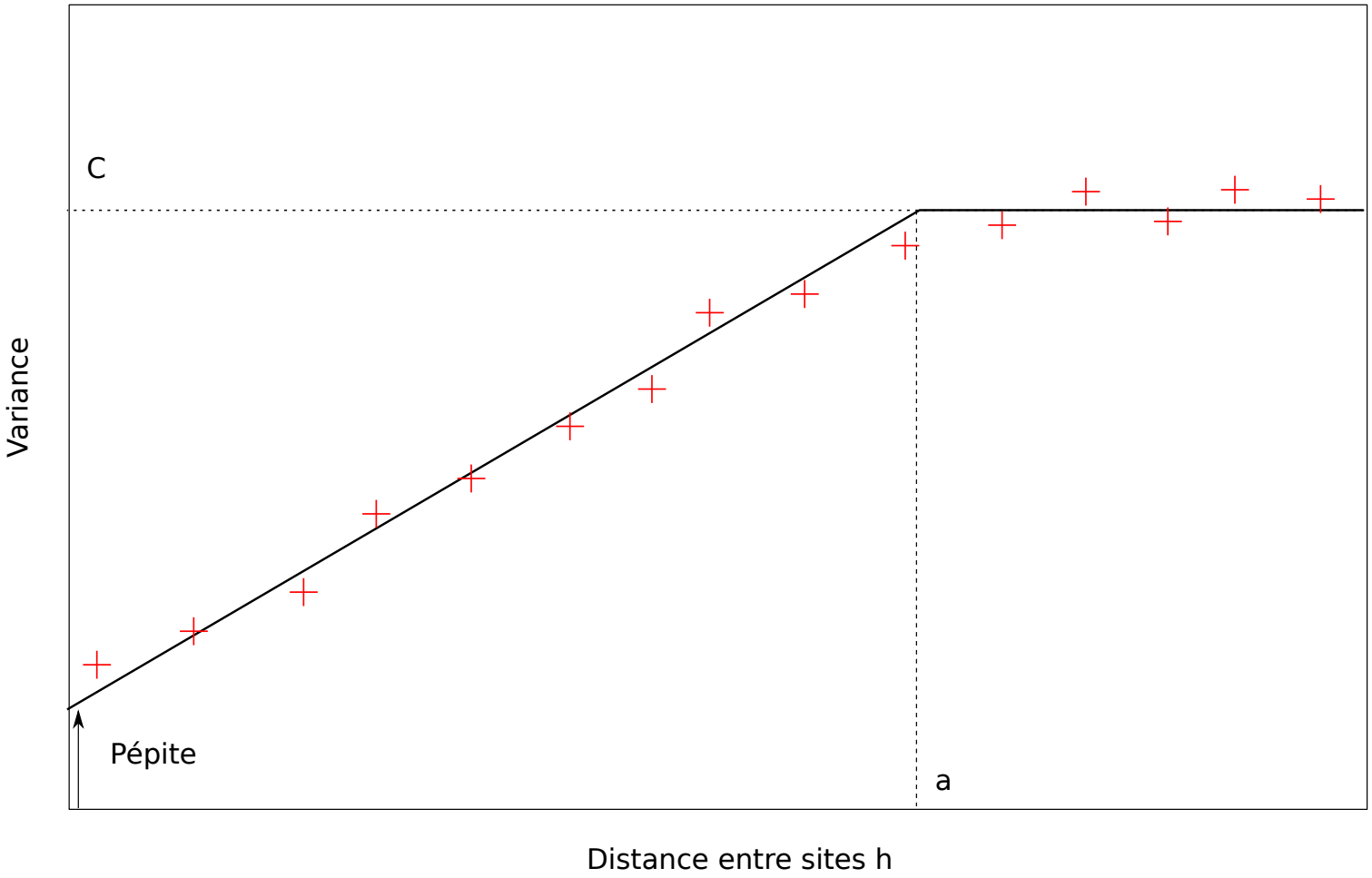
# Les modèles de variogramme



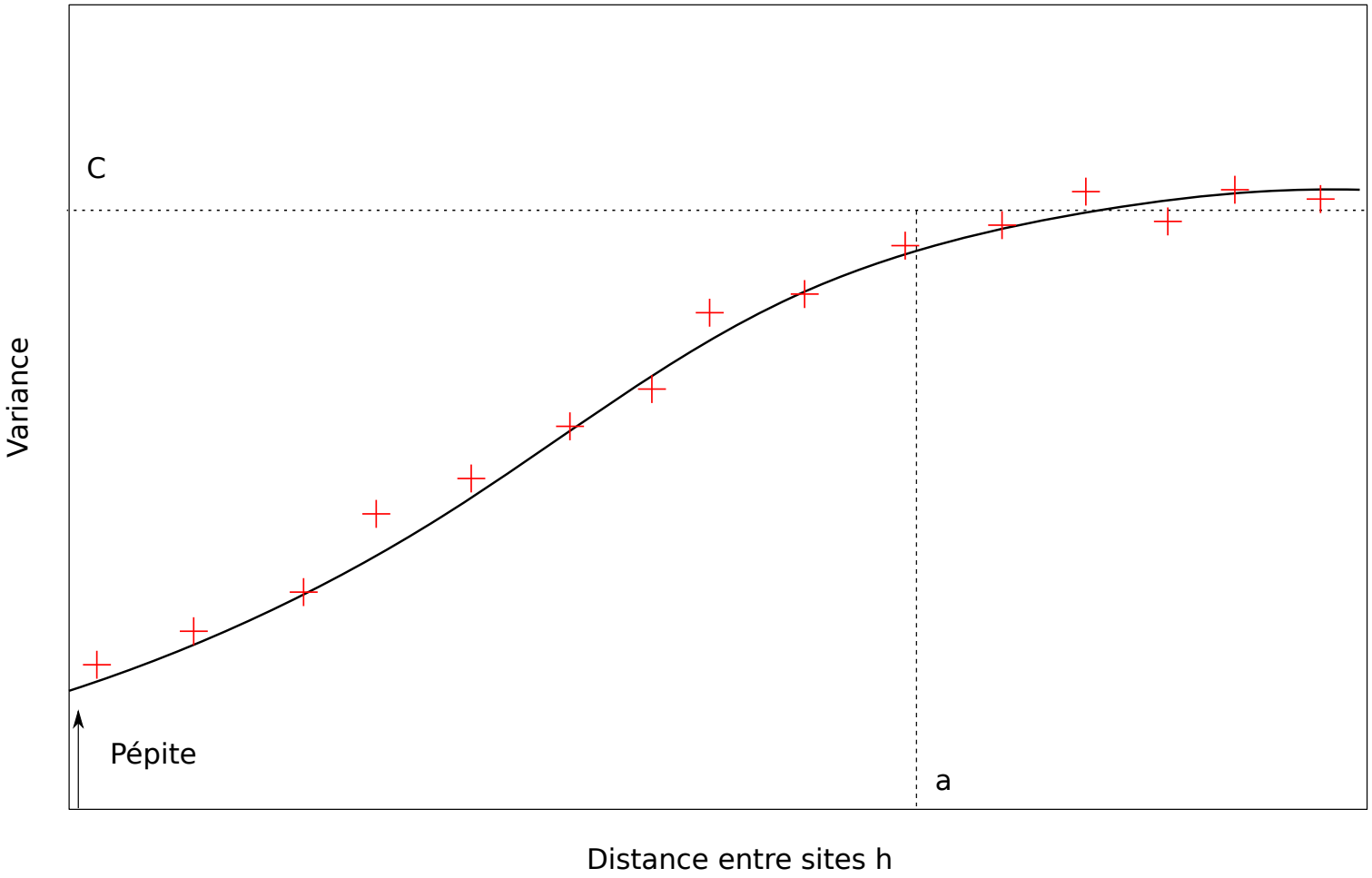
# Les modèles de variogramme



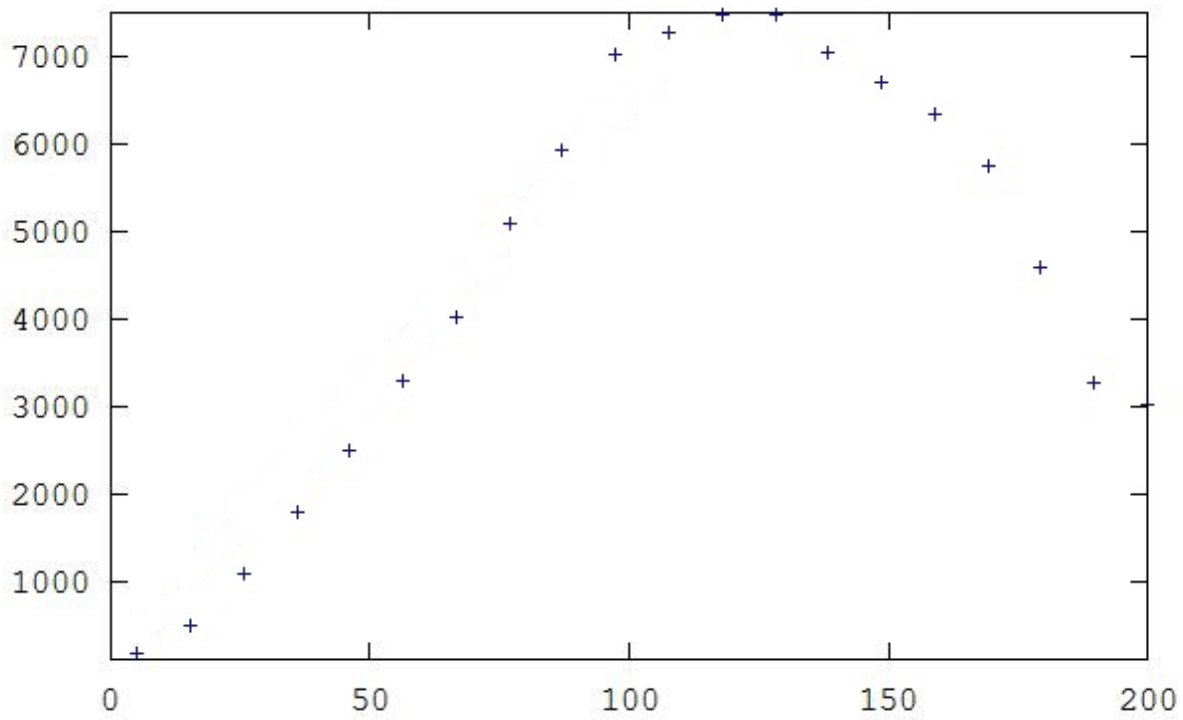
# Les modèles de variogramme



# Les modèles de variogramme

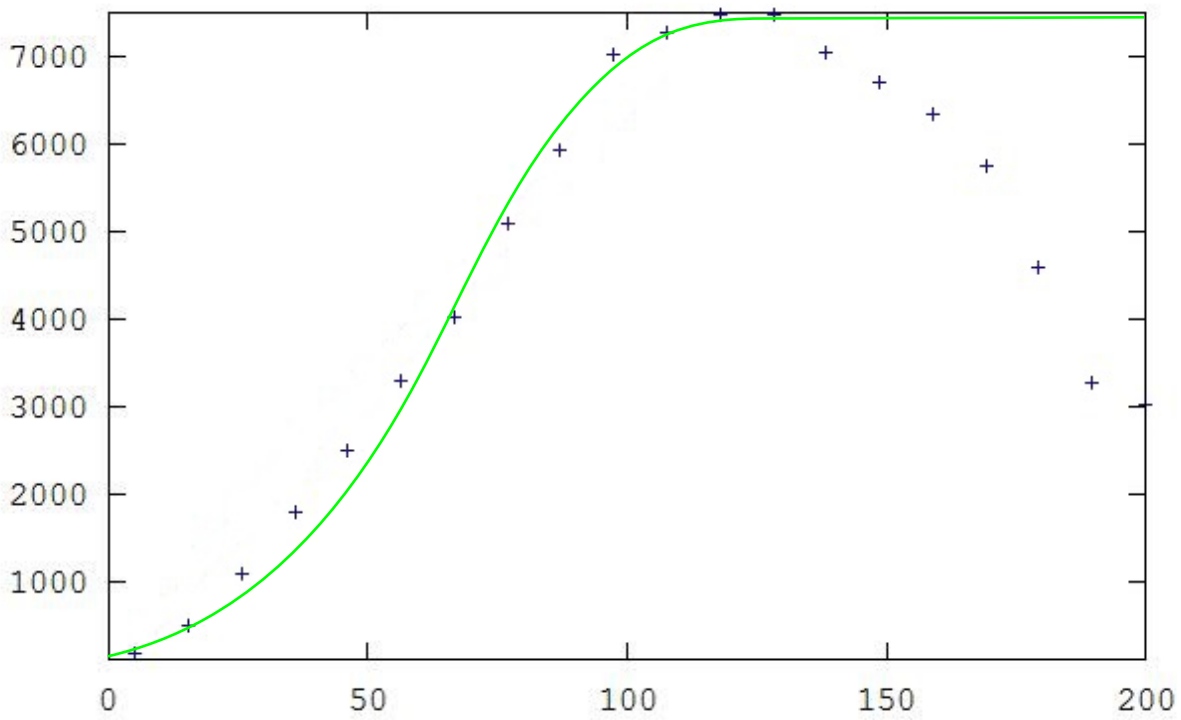


# Les modèles de variogramme



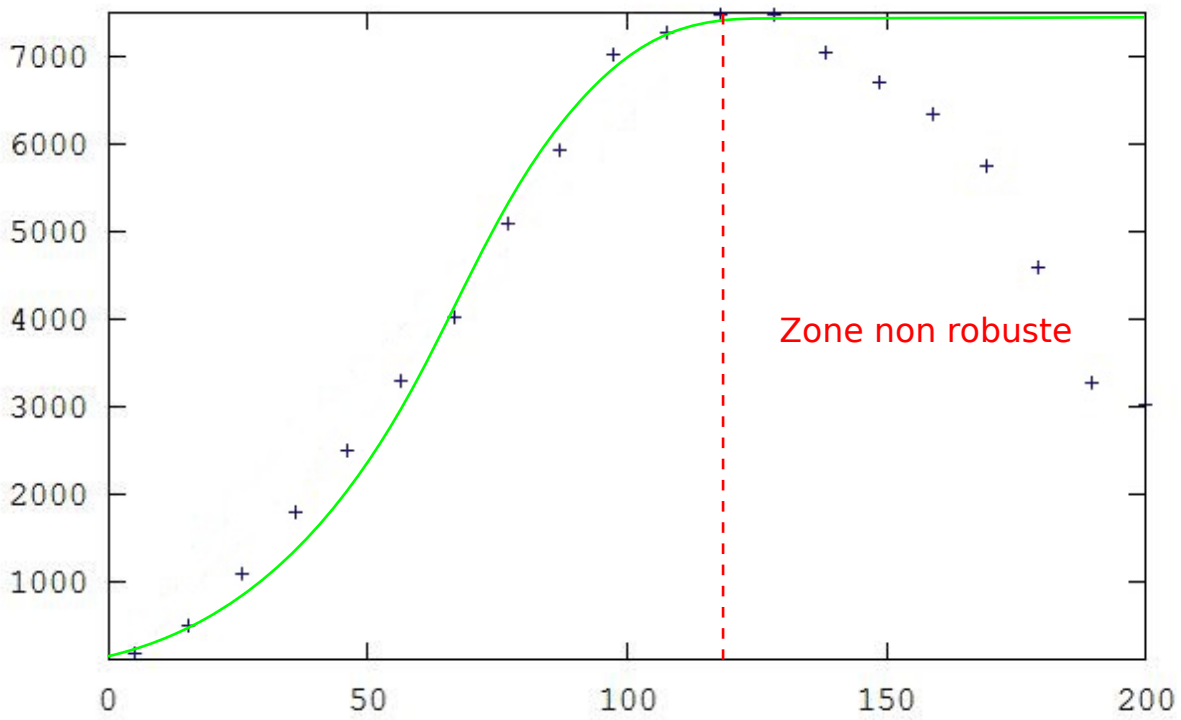

En pratique l'ajustement est calculé par les méthodes d'inférence paramétrique classiques (MCO, MCP, MV...).

# Les modèles de variogramme




En pratique l'ajustement est calculé par les méthodes d'inférence paramétrique classiques (MCO, MCP, MV...).

# Les modèles de variogramme




En pratique l'ajustement est calculé par les méthodes d'inférence paramétrique classiques (MCO, MCP, MV...).

# Analyse variographique

Comment s'assurer qu'une fonction $\gamma : \mathbb{R} \to \mathbb{R}^+$ est un modèle de variogramme ? Une condition suffisante sur $C(h) = C(0) - \gamma(h)$ :

## Théorème de Bochner (version édulcorée)

Une fonction $C : \mathbb{R} \to \mathbb{R}$ est un modèle de **covariogramme** si, et seulement si, elle s'exprime comme la **transformée de Fourier** d'une fonction **positive** $\pi : \mathbb{R} \to \mathbb{R}^+$ :

$$C(h) = \int_{\mathbb{R}} \pi(\nu) e^{-i\nu h} d\nu$$

La fonction $\pi$ est appelée la **densité spectrale de puissance** (DSP) du signal $Z$. La valeur $\pi(\nu)$ indique la contribution des fréquences $\nu$ dans la puissance totale du signal.

# Analyse variographique

Comment s'assurer qu'une fonction $\gamma : \mathbb{R} \to \mathbb{R}^+$ est un modèle de variogramme ? Une condition suffisante sur $C(h) = C(0) - \gamma(h)$ :

## Théorème de Bochner (version édulcorée)

Une fonction $C : \mathbb{R} \to \mathbb{R}$ est un modèle de **covariogramme** si, et seulement si, elle s'exprime comme la **transformée de Fourier** d'une fonction **positive** $\pi : \mathbb{R} \to \mathbb{R}^+$ :

$$C(h) = \int_{\mathbb{R}} \pi(\nu) e^{-i\nu h} d\nu$$

En effet, pour un signal stationnaire centré, on a $\int_{\mathbb{R}} \pi(\nu) d\nu = \int_{\mathbb{R}} \pi(\nu) e^{-0ih} d\nu = C(0) = \mathbb{E}[Z(x)Z(x+0)] = \mathbb{E}[Z(x)^2] = \sigma_Z^2$, la puissance moyenne du signal.

# Analyse variographique

Comment s'assurer qu'une fonction $\gamma : \mathbb{R} \to \mathbb{R}^+$ est un modèle de variogramme ? Une condition suffisante sur $C(h) = C(0) - \gamma(h)$ :

## Théorème de Bochner (version édulcorée)

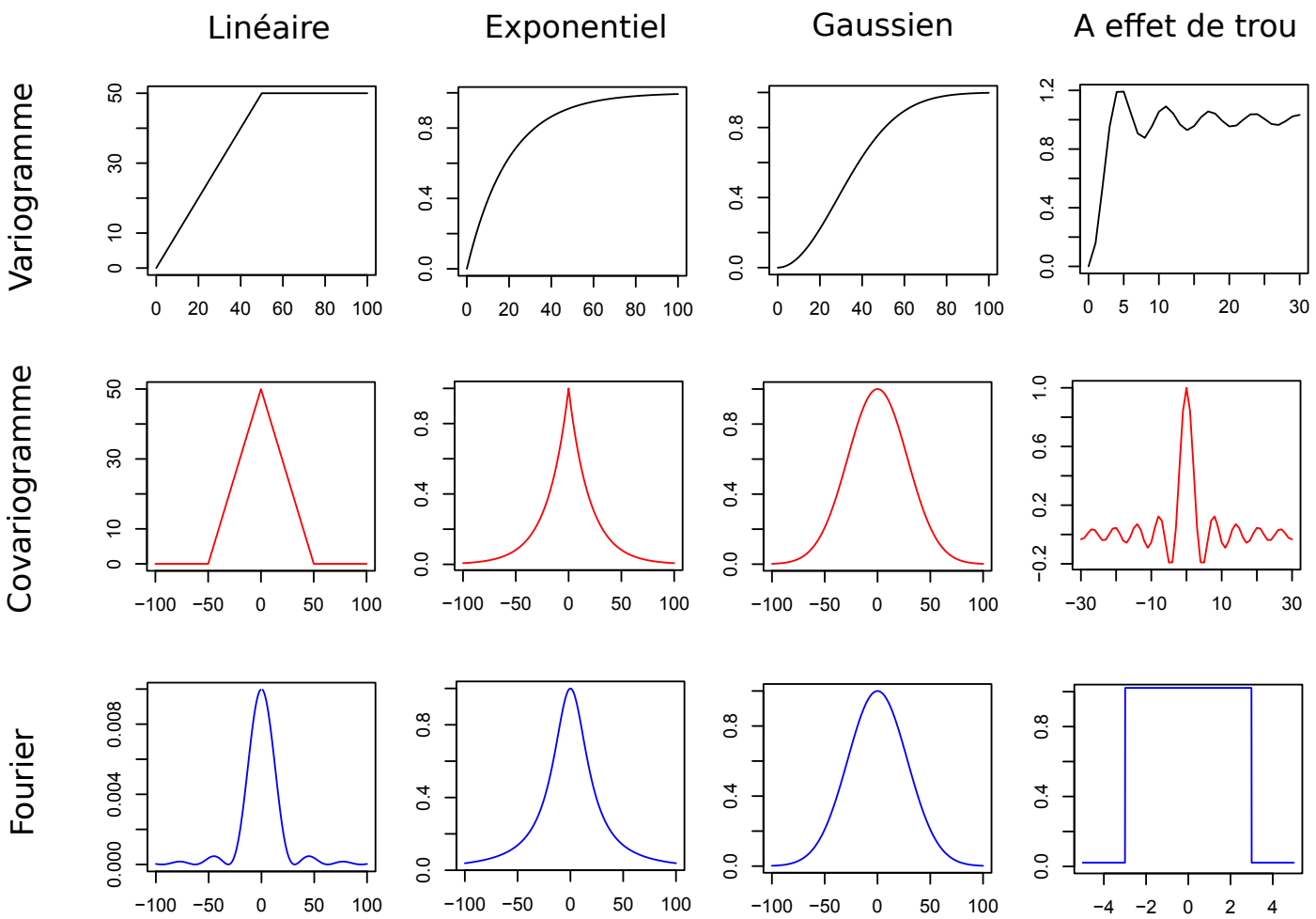
Une fonction $C : \mathbb{R} \to \mathbb{R}$ est un modèle de **covariogramme** si, et seulement si, elle s'exprime comme la **transformée de Fourier** d'une fonction **positive** $\pi : \mathbb{R} \to \mathbb{R}^+$ :

$$C(h) = \int_{\mathbb{R}} \pi(\nu) e^{-i\nu h} d\nu$$

**Bilan :** une CS pour savoir si $\gamma$ est un modèle de variogramme : $\gamma \to C = \gamma(\infty) - \gamma \to \pi = \mathcal{F}^{-1}[C] \to \forall\ \omega\ \pi(\omega) \geqslant 0$ ?

# Analyse variographique

Quelques exemples :



# Analyse variographique

La plupart des phénomènes géologiques possèdent un variogramme typique qui les caractérise (au second ordre) :

# Analyse variographique

La plupart des phénomènes géologiques possèdent un variogramme typique qui les caractérise (au second ordre) :

**Gisement d'or** : variogramme erratique, fort effet de pépite, faible portée (modèles exponentiel, sphérique...)

# Analyse variographique

La plupart des phénomènes géologiques possèdent un variogramme typique qui les caractérise (au second ordre) :

**Gisement d'or** : variogramme erratique, fort effet de pépite, faible portée (modèles exponentiel, sphérique...)

**Gisement de cuivre** : variogramme linéaire à l'origine, faible effet de pépite, grande portée (modèles linéaire, linéaire seuillé...)

# Analyse variographique

La plupart des phénomènes géologiques possèdent un variogramme typique qui les caractérise (au second ordre) :

**Gisement d'or** : variogramme erratique, fort effet de pépite, faible portée (modèles exponentiel, sphérique...)

**Gisement de cuivre** : variogramme linéaire à l'origine, faible effet de pépite, grande portée (modèles linéaire, linéaire seuillé...)

**Topographie** : variogramme très continu, absence d'effet de pépite, comportement parabolique à l'origine (modèles gaussien, cubique...)

# Analyse variographique

La plupart des phénomènes géologiques possèdent un variogramme typique qui les caractérise (au second ordre) :

**Gisement d'or** : variogramme erratique, fort effet de pépite, faible portée (modèles exponentiel, sphérique...)

**Gisement de cuivre** : variogramme linéaire à l'origine, faible effet de pépite, grande portée (modèles linéaire, linéaire seuillé...)

**Topographie** : variogramme très continu, absence d'effet de pépite, comportement parabolique à l'origine (modèles gaussien, cubique...)

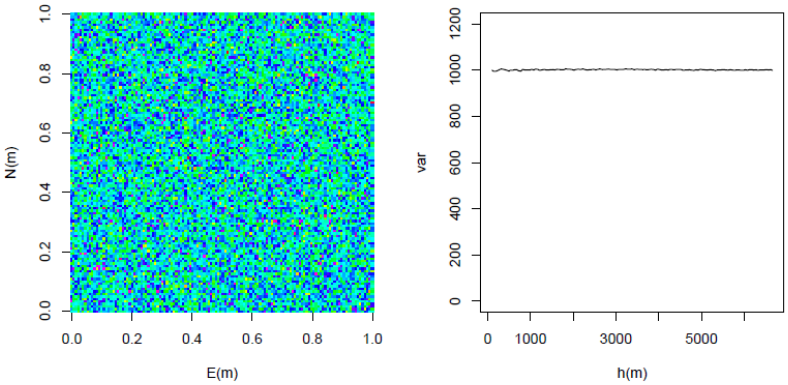
Pour aller plus loin : variogrammes anisotropiques $\gamma(||\vec{h}||, \arg(\vec{h}))$ et métriques ad hoc $h = \delta(x_1, x_2)$ avec $\delta : \mathcal{D}^2 \to \mathbb{R}_+$ non-euclidienne (e.g. strates géologiques, variétés riemaniennes...)

# Les modèles de variogramme

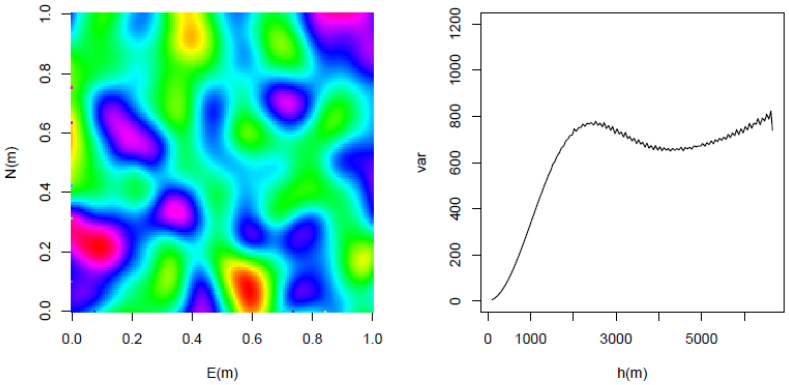
## Exemples d'associations : réalisation ↔ variogramme

class_variograms.r

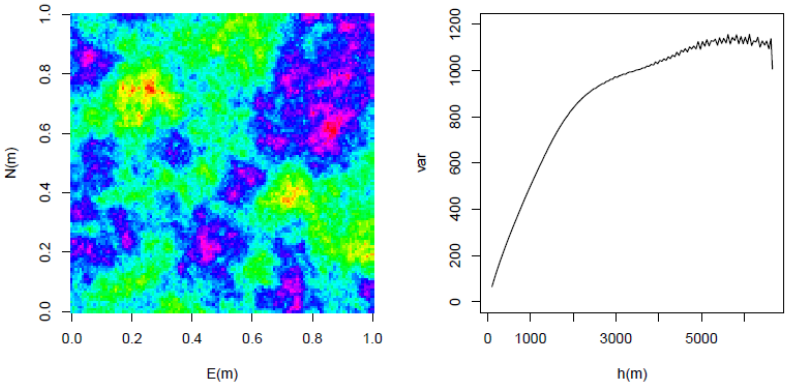
*Pépite pure (e.g. bruit blanc...)*



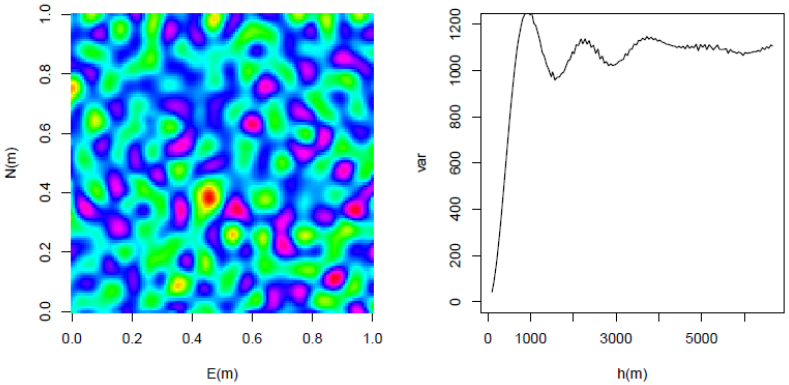

*Gaussien (e.g. MNT...)*




*Exponentiel (e.g. gisements miniers)*




*Modèle à effet de trou (e.g. harmoniques)*



# Analyse variographique : bilan

Le variogramme caractérise (à l'ordre 2) le phénomène à étudier

# Analyse variographique : bilan

Le variogramme caractérise (à l'ordre 2) le phénomène à étudier

La quantité d'observations étant généralement insuffisante, on régularise l'estimation avec une bibliothèque de fonctions paramétriques permettant d'assurer la définie-positivité de la fonction de covariance (lorsqu'elle existe).

# Analyse variographique : bilan

Le variogramme caractérise (à l'ordre 2) le phénomène à étudier

La quantité d'observations étant généralement insuffisante, on régularise l'estimation avec une bibliothèque de fonctions paramétriques permettant d'assurer la définie-positivité de la fonction de covariance (lorsqu'elle existe).

Le modèle de variogramme est choisi en fonction de la structure du variogramme expérimental, tandis que les paramètres sont fixés de sorte à minimiser l'erreur d'estimation (moindres carrés...).

# Analyse variographique : bilan

Le variogramme caractérise (à l'ordre 2) le phénomène à étudier

La quantité d'observations étant généralement insuffisante, on régularise l'estimation avec une bibliothèque de fonctions paramétriques permettant d'assurer la définie-positivité de la fonction de covariance (lorsqu'elle existe).

Le modèle de variogramme est choisi en fonction de la structure du variogramme expérimental, tandis que les paramètres sont fixés de sorte à minimiser l'erreur d'estimation (moindres carrés...).

Tous les calculs subséquents sont menés à partir du modèle de variogramme inféré (cohérence avec la théorie, stabilité numérique...).

# Le Krigeage

Le *Krigeage* est une technique géostatistique d'estimation locale : Danie G. **Krige** (1951), George **Matheron** (1962), Norbert **Wiener** (1949) et Andrey **Kolmogorov** (1941).

# Le Krigeage

Le *Krigeage* est une technique géostatistique d'estimation locale : Danie G. **Krige** (1951), George **Matheron** (1962), Norbert **Wiener** (1949) et Andrey **Kolmogorov** (1941).

L'interpolation par Krigeage est intrinséquement liée à l'**analyse variographique** du phénomène à étudier mais n'est pas sa seule application.

# Le Krigeage

Le *Krigeage* est une technique géostatistique d'estimation locale : Danie G. **Krige** (1951), George **Matheron** (1962), Norbert **Wiener** (1949) et Andrey **Kolmogorov** (1941).

L'interpolation par Krigeage est intrinséquement liée à l'**analyse variographique** du phénomène à étudier mais n'est pas sa seule application.

Le Krigeage permet d'estimer la valeur prise par la variable régionalisée en un site non-échantillonné, en tenant en compte des **corrélations** avec les valeurs prises en d'autres sites.

# Le Krigeage

Le *Krigeage* est une technique géostatistique d'estimation locale : Danie G. **Krige** (1951), George **Matheron** (1962), Norbert **Wiener** (1949) et Andrey **Kolmogorov** (1941).

L'interpolation par Krigeage est intrinséquement liée à l'**analyse variographique** du phénomène à étudier mais n'est pas sa seule application.

Le Krigeage permet d'estimer la valeur prise par la variable régionalisée en un site non-échantillonné, en tenant en compte des **corrélations** avec les valeurs prises en d'autres sites.

Contrairement aux méthodes d'interpolation vues précédemment, le Krigeage permet d'estimer également l'**incertitude formelle** de l'estimation en chaque point.

# Les contraintes du Krigeage

## ① Contrainte de linéarité

La valeur *krigée* est une combinaison linéaire des valeurs observées :

$$\hat{z}(x_0) = \sum_{i=1}^{N} w_i z(x_i) = \mathbf{W^T Z}$$

Estimer une valeur $\hat{z}(x_0)$ par Krigeage revient à déterminer les valeurs numériques des poids $w_i \in \mathbb{R}$ (en pratique fonctions de $x_0$).

# Les contraintes du Krigeage

## ① Contrainte de linéarité

La valeur *krigée* est une combinaison linéaire des valeurs observées :

$$\hat{z}(x_0) = \sum_{i=1}^{N} w_i z(x_i) = \mathbf{W^T Z}$$

Estimer une valeur $\hat{z}(x_0)$ par Krigeage revient à déterminer les valeurs numériques des poids $w_i \in \mathbb{R}$ (en pratique fonctions de $x_0$).

## ② Contrainte d'autorisation

L'erreur $\hat{z}(x_0) - z(x_0)$ doit être une **combinaison linéaire autorisée**, c'est-à-dire que l'on doit être capable de calculer ses deux premiers moments (moyennes et covariances).

**Stationnarité d'ordre 2** ($\Leftrightarrow$ variogramme borné) : toute combinaison linéaire est autorisée. **Hypothèse intrinsèque** : $\sum w_i = 1$.

# Les contraintes du Krigeage

## ③ Contrainte de non-biais

L'estimateur doit être non-biaisé, c'est-à-dire que l'espérance de l'erreur d'estimation doit être nulle :

$$\mathbb{E}[\hat{z}(x_0) - z(x_0)] = 0$$

$$\mathbb{E}[\hat{z}(x_0) - z(x_0)] = m\left(\sum_{i=1}^{N} w_i - 1\right) = 0 \iff \sum_{i=1}^{N} w_i = 1$$

$\rightarrow$ On cherche un vecteur de poids $\mathbf{W}$ de somme unitaire.

# Les contraintes du Krigeage

## ③ Contrainte de non-biais

L'estimateur doit être non-biaisé, c'est-à-dire que l'espérance de l'erreur d'estimation doit être nulle :

$$\mathbb{E}[\hat{z}(x_0) - z(x_0)] = 0$$

$$\mathbb{E}[\hat{z}(x_0) - z(x_0)] = m\Big(\sum_{i=1}^{N} w_i - 1\Big) = 0 \iff \sum_{i=1}^{N} w_i = 1$$

$\rightarrow$ On cherche un vecteur de poids $\mathbf{W}$ de somme unitaire.

## ④ Contrainte d'optimalité

L'estimateur doit être de variance minimale. On cherche donc :

$$\mathbf{W} = \underset{\mathbf{W}\in\mathbb{R}^N}{\text{argmin}} \ \text{Var}\Big[\sum_{i=1}^{N} w_i z(x_i) - z(x_0)\Big]$$

# Le Krigeage

L'estimation par Krigeage s'exprime alors sous forme d'un problème d'optimisation sous contrainte (contrainte d'égalité).

## Formulation du problème

On cherche : $\hat{z}(x_0) = \sum_{i=1}^{N} w_i z(x_i)$ avec :

$$\mathbf{W} = \underset{\mathbf{W} \in \mathbb{R}^N}{\text{argmin}} \ \text{Var}\Big[\sum_{i=1}^{N} w_i z(x_i) - z(x_0)\Big] \quad \text{s.t.} \ \sum_{i=1}^{n} w_i = 1$$

# Le Krigeage

L'estimation par Krigeage s'exprime alors sous forme d'un problème d'optimisation sous contrainte (contrainte d'égalité).

## Formulation du problème

On cherche : $\hat{z}(x_0) = \sum_{i=1}^{N} w_i z(x_i)$ avec :

$$\mathbf{W} = \underset{\mathbf{W} \in \mathbb{R}^N}{\text{argmin}} \ \text{Var}\Big[\sum_{i=1}^{N} w_i z(x_i) - z(x_0)\Big] \quad \text{s.t.} \ \sum_{i=1}^{n} w_i = 1$$

$\hat{z}(x_0)$ est le <u>meilleur estimateur linéaire non-biaisé</u> de $z(x_0)$

# Le Krigeage

Optimisation sous contraintes : **multiplicateurs de Lagrange**

On cherche à minimiser la quantité :

$$\mathcal{L}(w_1, w_2, ...w_n, \lambda) = \mathsf{Var}\Big[\sum_{i=1}^{N} w_i z_i - z_0\Big] + 2\lambda\Big(1 - \sum_{i=1}^{N} w_i\Big)$$

où $\lambda \in \mathbb{R}$ est un multiplicateur de Lagrange (variable latente inconnue du problème).

# Le Krigeage

Optimisation sous contraintes : **multiplicateurs de Lagrange**

On cherche à minimiser la quantité :

$$\mathcal{L}(w_1, w_2, ...w_n, \lambda) = \text{Var}\Big[\sum_{i=1}^{N} w_i z_i - z_0\Big] + 2\lambda\Big(1 - \sum_{i=1}^{N} w_i\Big)$$

où $\lambda \in \mathbb{R}$ est un multiplicateur de Lagrange (variable latente inconnue du problème).

On exprime $\mathcal{L}(\mathbf{W}, \lambda)$ à l'aide de la covariance (stationnaire) :

$$C(0) + \sum_i \sum_j w_i w_j C(h_{ij}) - 2\sum_i w_i C(h_{i0}) + 2\lambda\Big(1 - \sum_{i=1}^{N} w_i\Big)$$

# Le Krigeage

$$C(0) + \sum_i \sum_j w_i w_j C(h_{ij}) - 2 \sum_i w_i C(h_{i0}) + 2\lambda \left(1 - \sum_{i=1}^{N} w_i\right)$$

# Le Krigeage

$$C(0) + \sum_i \sum_j w_i w_j C(h_{ij}) - 2\sum_i w_i C(h_{i0}) + 2\lambda\left(1 - \sum_{i=1}^{N} w_i\right)$$

Relation fondamentale de la géostatistique :

$$\gamma(h) = C(0) - C(h)$$

# Le Krigeage

$$C(0) + \sum_i \sum_j w_i w_j C(h_{ij}) - 2\sum_i w_i C(h_{i0}) + 2\lambda\left(1 - \sum_{i=1}^{N} w_i\right)$$

Relation fondamentale de la géostatistique :

$$\gamma(h) = C(0) - C(h)$$

$$\mathcal{L}(\mathbf{W}, \lambda) = -\sum_i \sum_j w_i w_j \gamma_{ij} + 2\sum_i w_i \gamma_{i0} + 2\lambda\left(\sum_{i=1}^{N} w_i - 1\right)$$

avec : $\gamma_{ij} = \gamma(x_i, x_j) = \gamma(||x_i - x_j||)$

# Le Krigeage

$$C(0) + \sum_i \sum_j w_i w_j C(h_{ij}) - 2\sum_i w_i C(h_{i0}) + 2\lambda\left(1 - \sum_{i=1}^{N} w_i\right)$$

Relation fondamentale de la géostatistique :

$$\gamma(h) = C(0) - C(h)$$

$$\mathcal{L}(\mathbf{W}, \lambda) = -\sum_i \sum_j w_i w_j \gamma_{ij} + 2\sum_i w_i \gamma_{i0} + 2\lambda\left(\sum_{i=1}^{N} w_i - 1\right)$$

avec : $\gamma_{ij} = \gamma(x_i, x_j) = \gamma(||x_i - x_j||)$

Optimisation :

$$\frac{\partial \mathcal{L}}{\partial w_i} = 0 \qquad \forall\ i \in [1, N] \qquad \text{et} \qquad \frac{\partial \mathcal{L}}{\partial \lambda} = 0$$

# Le Krigeage

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial w_i} = 2\gamma_{i0} - 2\lambda - 2\sum_j w_j \gamma_{ij} = 0 & \forall\ i \in [1, N] \\ \frac{\partial \mathcal{L}}{\partial \lambda} = \sum_i w_i - 1 = 0 \end{cases}$$

$$\begin{cases} \sum_j w_j \gamma_{ij} + \lambda = \gamma_{i0} & \forall\ i \in [1, N] \\ \sum_i w_i = 1 \end{cases} \qquad (\lambda = 0 \Rightarrow \text{Wiener-Hopf})$$

$\rightarrow$ Système de $N+1$ équations à $N+1$ inconnues

# Le Krigeage

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial w_i} = 2\gamma_{i0} - 2\lambda - 2\sum_j w_j \gamma_{ij} = 0 & \forall\ i \in [1, N] \\ \frac{\partial \mathcal{L}}{\partial \lambda} = \sum_i w_i - 1 = 0 \end{cases}$$

$$\begin{cases} \sum_j w_j \gamma_{ij} + \lambda = \gamma_{i0} & \forall\ i \in [1, N] \\ \sum_i w_i = 1 \end{cases}$$

($\lambda = 0 \Rightarrow$ Wiener-Hopf)

$\rightarrow$ Système de $N+1$ équations à $N+1$ inconnues

Résolution par inversion matricielle

# Krigeage : résolution

On note $\mathbf{\Gamma} \in \mathbb{R}^{N \times N}$ la matrice des variances entre sites :

$$\Gamma_{ij} = \gamma_{ij} = \gamma(h_{ij}) = \gamma(||x_i - x_j||) \;\; \forall \; i, j \in N$$

$$\mathbf{\Gamma} = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1N} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ \gamma_{N1} & \gamma_{N2} & \cdots & \gamma_{NN} \end{bmatrix}$$

# Krigeage : résolution

On note $\mathbf{\Gamma} \in \mathbb{R}^{N \times N}$ la matrice des variances entre sites :

$$\Gamma_{ij} = \gamma_{ij} = \gamma(h_{ij}) = \gamma(||x_i - x_j||) \;\; \forall \; i, j \in N$$

$$\mathbf{\Gamma} = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1N} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ \gamma_{N1} & \gamma_{N2} & \cdots & \gamma_{NN} \end{bmatrix}$$

On note $\mathbf{\Gamma_0} \in \mathbb{R}^N$ le vecteur des variances entre les sites observés et le site à estimer $x_0$ :

$$\mathbf{\Gamma_0} = \begin{bmatrix} \gamma_{10} & \gamma_{20} & \cdots & \gamma_{N0} \end{bmatrix}^T$$

# Krigeage : résolution

L'équation matricielle s'exprime alors par :

$$\begin{bmatrix} \mathbf{\Gamma} & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{W} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{\Gamma_0} \\ 1 \end{bmatrix}$$

# Krigeage : résolution

L'équation matricielle s'exprime alors par :

$$\begin{bmatrix} \mathbf{\Gamma} & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{W} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{\Gamma_0} \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{W} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{\Gamma} & 1 \\ 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{\Gamma_0} \\ 1 \end{bmatrix}$$

# Krigeage : résolution

L'équation matricielle s'exprime alors par :

$$\begin{bmatrix} \mathbf{\Gamma} & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{W} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{\Gamma_0} \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{W} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{\Gamma} & 1 \\ 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{\Gamma_0} \\ 1 \end{bmatrix}$$

Remarque : la matrice à inverser ne dépend que des configurations des sites d'observation. Elle peut donc être inversée *une fois pour toutes* en amont de la procédure d'interpolation.

# Mise en oeuvre du Krigeage

On se donne une fonction de variogramme $\gamma : \mathbb{R}^+ \to \mathbb{R}^+$ (éventuellement calculée à partir du variogramme empirique $\hat{\gamma}$)

# Mise en oeuvre du Krigeage

On se donne une fonction de variogramme $\gamma : \mathbb{R}^+ \to \mathbb{R}^+$ (éventuellement calculée à partir du variogramme empirique $\hat{\gamma}$)

On calcule la distance $h_{ij} = ||x_i - x_j||$ entre couples de sites :

$$\mathbf{H} = \begin{bmatrix} 0 & h_{12} & ... & h_{1N} \\ h_{21} & 0 & ... & h_{2N} \\ ... & ... & ... & ... \\ h_{N1} & h_{N2} & ... & 0 \end{bmatrix}$$

# Mise en oeuvre du Krigeage

On se donne une fonction de variogramme $\gamma : \mathbb{R}^+ \to \mathbb{R}^+$ (éventuellement calculée à partir du variogramme empirique $\hat{\gamma}$)

On calcule la distance $h_{ij} = ||x_i - x_j||$ entre couples de sites :

$$\mathbf{H} = \begin{bmatrix} 0 & h_{12} & ... & h_{1N} \\ h_{21} & 0 & ... & h_{2N} \\ ... & ... & ... & ... \\ h_{N1} & h_{N2} & ... & 0 \end{bmatrix}$$

Par définition de la distance, la diagonale de $\mathbf{H}$ est nulle :

$$\forall\ i \quad h_{ii} = ||x_i - x_i|| = 0$$

# Mise en oeuvre du Krigeage

On calcule la matrice de variances entre sites à partir de la matrice des distances ($\gamma$ étant fonction de la distance entre les sites uniquement) : $\mathbf{\Gamma} = \gamma(\mathbf{H})$, complétée par des $1$ :

# Mise en oeuvre du Krigeage

On calcule la matrice de variances entre sites à partir de la matrice des distances ($\gamma$ étant fonction de la distance entre les sites uniquement) : $\mathbf{\Gamma} = \gamma(\mathbf{H})$, complétée par des $1$ :

$$\mathbf{A} = \begin{bmatrix} \gamma(0) & \gamma(h_{12}) & ... & \gamma(h_{1N}) & 1 \\ \gamma(h_{21}) & \gamma(0) & ... & \gamma(h_{2N}) & 1 \\ ... & ... & ... & ... & 1 \\ \gamma(h_{N1}) & \gamma(h_{N2}) & ... & \gamma(0) & 1 \\ 1 & 1 & ... & 1 & 0 \end{bmatrix}$$

# Mise en oeuvre du Krigeage

On calcule la matrice de variances entre sites à partir de la matrice des distances ($\gamma$ étant fonction de la distance entre les sites uniquement) : $\mathbf{\Gamma} = \gamma(\mathbf{H})$, complétée par des $1$ :

$$\mathbf{A} = \begin{bmatrix} \gamma(0) & \gamma(h_{12}) & ... & \gamma(h_{1N}) & 1 \\ \gamma(h_{21}) & \gamma(0) & ... & \gamma(h_{2N}) & 1 \\ ... & ... & ... & ... & 1 \\ \gamma(h_{N1}) & \gamma(h_{N2}) & ... & \gamma(0) & 1 \\ 1 & 1 & ... & 1 & 0 \end{bmatrix}$$

On inverse la matrice $\mathbf{A}$ et on la sauvegarde en mémoire

# Mise en oeuvre du Krigeage

Pour chaque point $x_0$ à interpoler, on calcule le vecteur des distances entre $x_0$ et les $N$ sites :

$$\mathbf{H_0} = \begin{bmatrix} h_{10} & h_{20} & ... & h_{N0} \end{bmatrix}^T$$

# Mise en oeuvre du Krigeage

Pour chaque point $x_0$ à interpoler, on calcule le vecteur des distances entre $x_0$ et les $N$ sites :

$$\mathbf{H_0} = \begin{bmatrix} h_{10} & h_{20} & ... & h_{N0} \end{bmatrix}^T$$

On calcule le vecteur des covariances entre le point à interpoler et les sites par $\mathbf{\Gamma_0} = \gamma(\mathbf{H_0})$, complété par un $1$ :

$$\mathbf{B} = \begin{bmatrix} \gamma(h_{10}) & \gamma(h_{20}) & ... & \gamma(h_{N0}) & 1 \end{bmatrix}^T$$

# Mise en oeuvre du Krigeage

Pour chaque point $x_0$ à interpoler, on calcule le vecteur des distances entre $x_0$ et les $N$ sites :

$$\mathbf{H_0} = \begin{bmatrix} h_{10} & h_{20} & ... & h_{N0} \end{bmatrix}^T$$

On calcule le vecteur des covariances entre le point à interpoler et les sites par $\mathbf{\Gamma_0} = \gamma(\mathbf{H_0})$, complété par un $1$ :

$$\mathbf{B} = \begin{bmatrix} \gamma(h_{10}) & \gamma(h_{20}) & ... & \gamma(h_{N0}) & 1 \end{bmatrix}^T$$

On détermine les pondération $\mathbf{W}$ en $x_0$ : $\begin{bmatrix} \mathbf{W} & \lambda \end{bmatrix}^T = \mathbf{A}^{-1}\mathbf{B}$

# Mise en oeuvre du Krigeage

Pour chaque point $x_0$ à interpoler, on calcule le vecteur des distances entre $x_0$ et les $N$ sites :

$$\mathbf{H_0} = \begin{bmatrix} h_{10} & h_{20} & ... & h_{N0} \end{bmatrix}^T$$

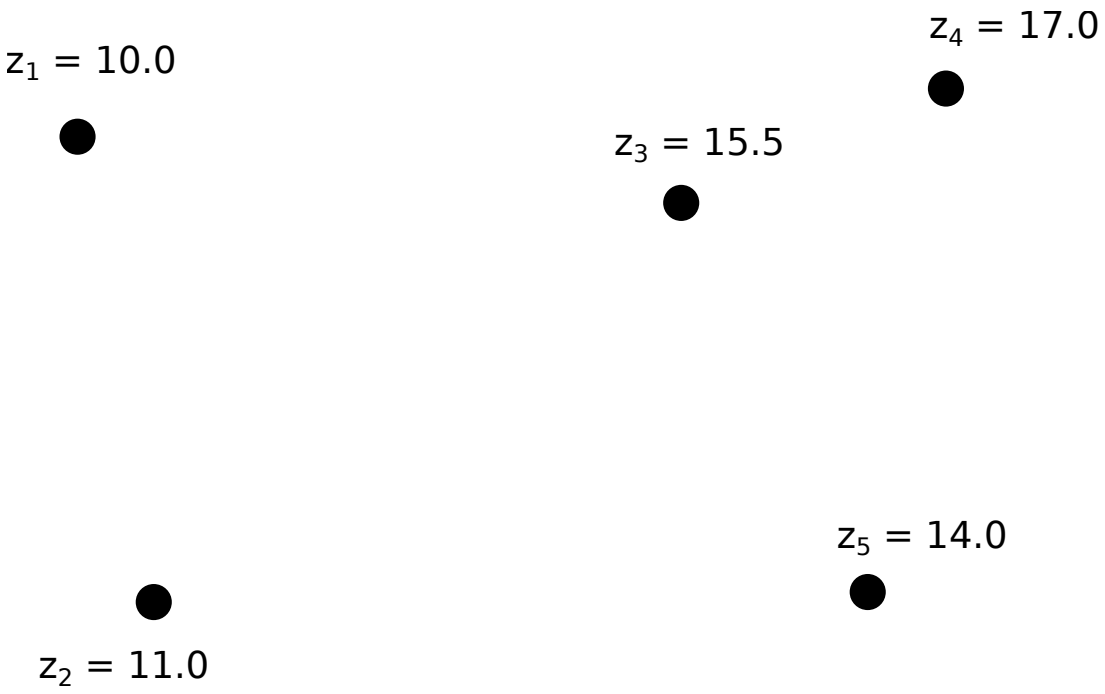
On calcule le vecteur des covariances entre le point à interpoler et les sites par $\mathbf{\Gamma_0} = \gamma(\mathbf{H_0})$, complété par un $1$ :

$$\mathbf{B} = \begin{bmatrix} \gamma(h_{10}) & \gamma(h_{20}) & ... & \gamma(h_{N0}) & 1 \end{bmatrix}^T$$

On détermine les pondération $\mathbf{W}$ en $x_0$ : $\begin{bmatrix} \mathbf{W} & \lambda \end{bmatrix}^T = \mathbf{A}^{-1}\mathbf{B}$

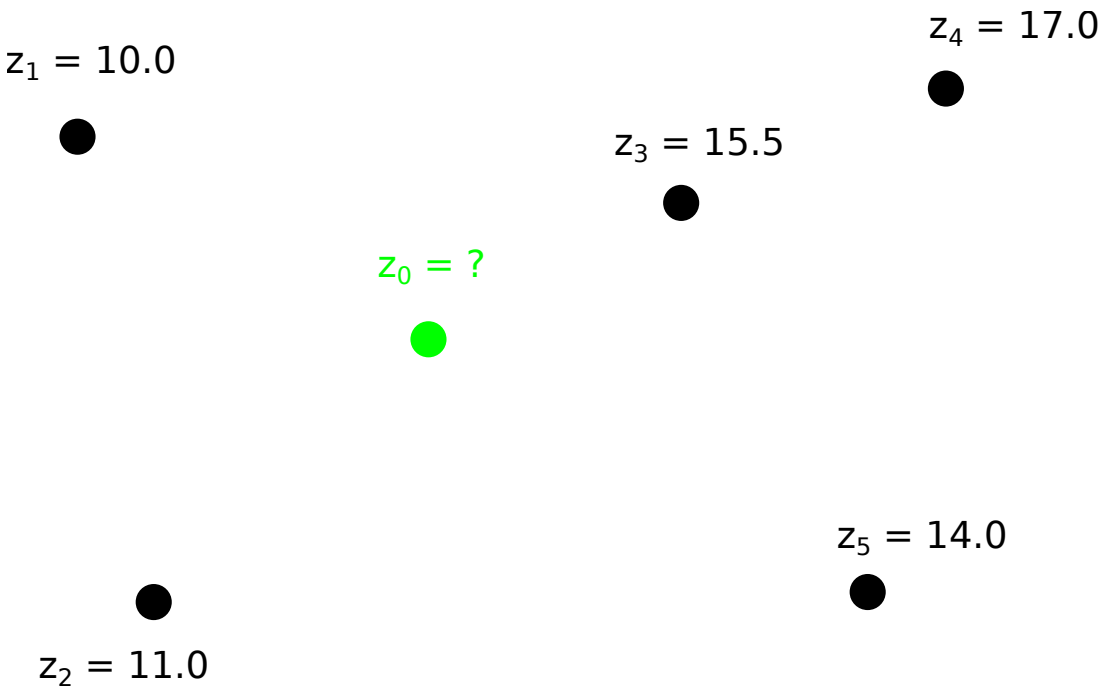
On en déduit la valeur interpolée recherchée : $\hat{z}(x_0) = \mathbf{W^T Z}$

# Exemple



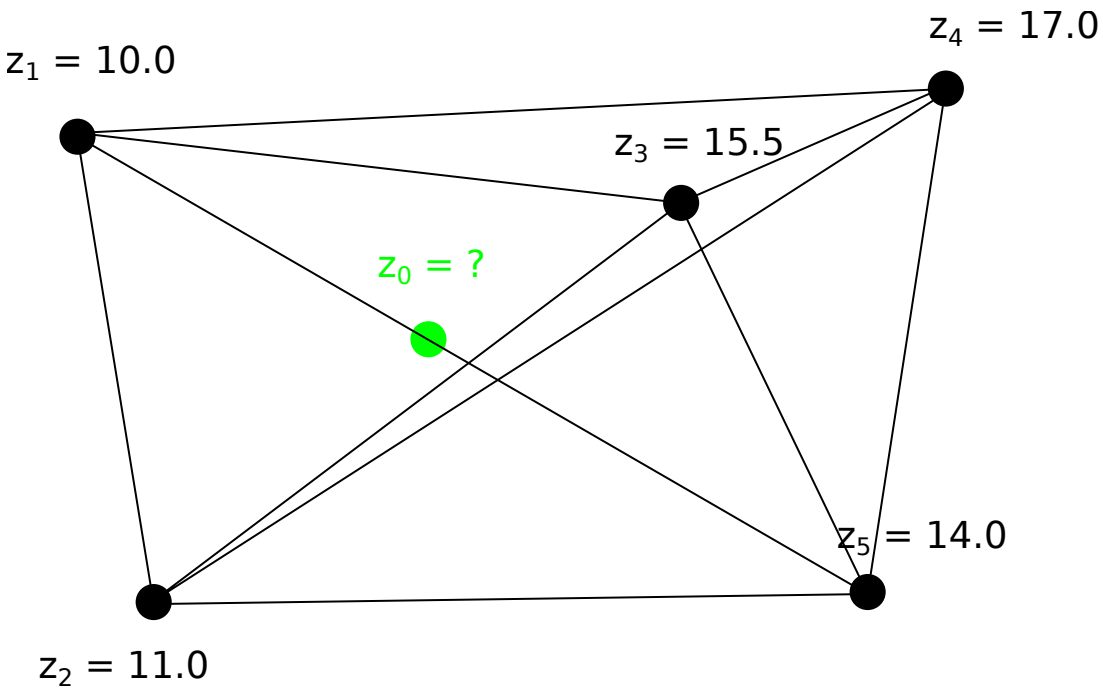

$N = 5$ points d'observations

# Exemple



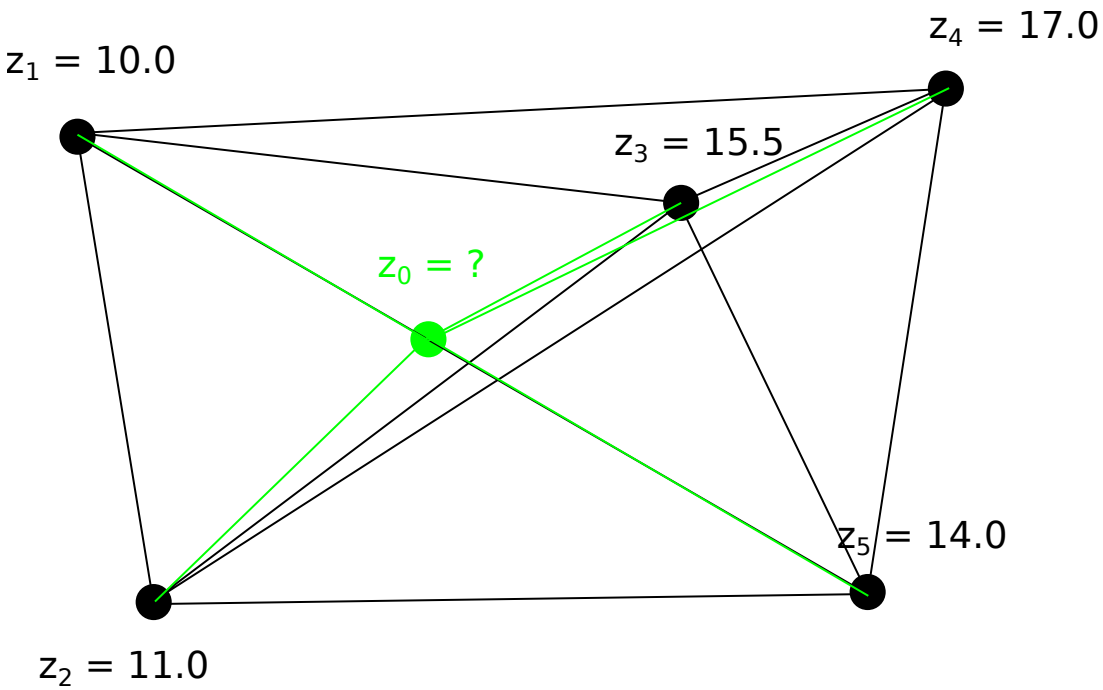

$N = 5$ points d'observations, on recherche $z(x_0)$

# Exemple



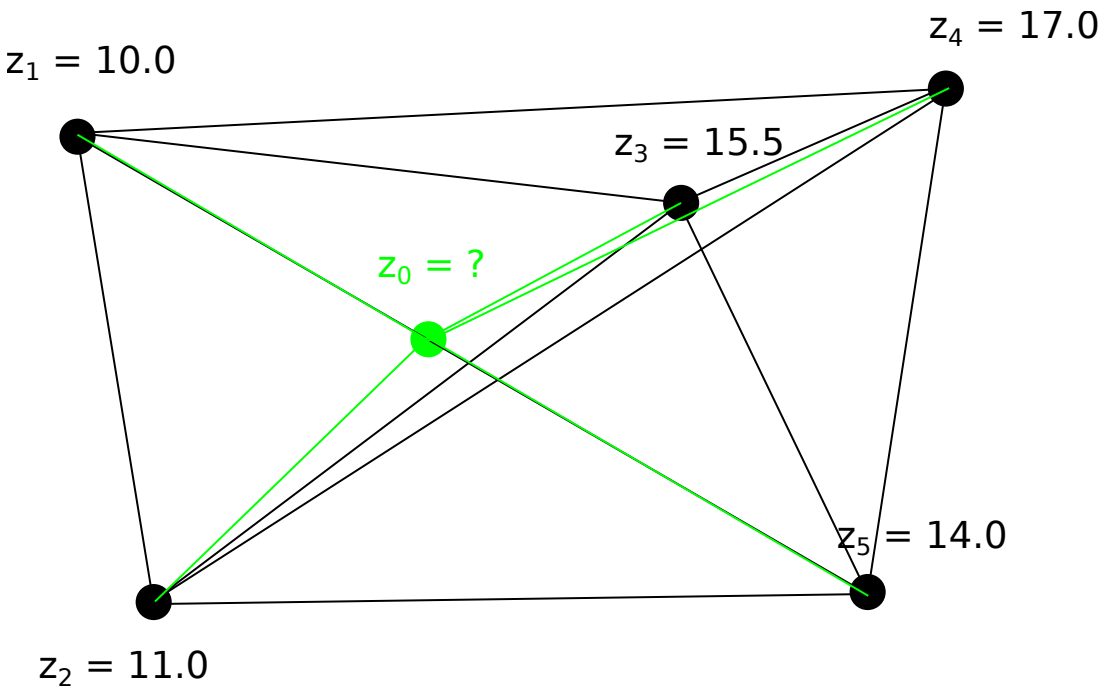

Calcul des distances entre les sites $\rightarrow$ matrice $\mathbf{H}$

# Exemple




Calcul des distances entre les sites et $x_0$ : $\rightarrow$ vecteur $\mathbf{H_0}$

# Exemple




On choisit un variogramme linéaire avec palier :
$\gamma(h) = h$ si $h \leqslant 40$, $\gamma(h) = 40$ sinon

# Exemple

Calcul des distances entre les sites :

$$\mathbf{H} = \begin{bmatrix} 0.0 & 24.2 & 30.5 & 44.0 & 46.3 \\ 24.2 & 0.0 & 33.0 & 47.4 & 36.2 \\ 30.5 & 33.0 & 0.0 & 14.8 & 22.1 \\ 44.0 & 47.4 & 14.8 & 0.0 & 25.7 \\ 46.3 & 36.2 & 22.1 & 25.7 & 0.0 \end{bmatrix}$$

# Exemple

Calcul des distances entre les sites :

$$\mathbf{H} = \begin{bmatrix} 0.0 & 24.2 & 30.5 & 44.0 & 46.3 \\ 24.2 & 0.0 & 33.0 & 47.4 & 36.2 \\ 30.5 & 33.0 & 0.0 & 14.8 & 22.1 \\ 44.0 & 47.4 & 14.8 & 0.0 & 25.7 \\ 46.3 & 36.2 & 22.1 & 25.7 & 0.0 \end{bmatrix}$$

Calcul des distances entre les sites et le point à interpoler :

$$\mathbf{H_0} = \begin{bmatrix} 20.6 & 19.0 & 14.4 & 29.1 & 25.8 \end{bmatrix}^T$$

# Exemple

Calcul des variances entre les sites :

$$\mathbf{A} = \begin{bmatrix} 0.0 & 24.2 & 30.5 & 40.0 & 40.0 \\ 24.2 & 0.0 & 33.0 & 40.0 & 36.2 \\ 30.5 & 33.0 & 0.0 & 14.8 & 22.1 \\ 40.0 & 40.0 & 14.8 & 0.0 & 25.7 \\ 40.0 & 36.2 & 22.1 & 25.7 & 0.0 \end{bmatrix}$$

Calcul des variances entre les sites et le point à interpoler :

$$\mathbf{B} = \begin{bmatrix} 20.6 & 19.0 & 14.4 & 29.1 & 25.8 \end{bmatrix}^T$$

# Exemple

Complétion avec des vecteurs de 1 :

$$\mathbf{A} = \begin{bmatrix} 0.0 & 24.2 & 30.5 & 40.0 & 40.0 & 1 \\ 24.2 & 0.0 & 33.0 & 40.0 & 36.2 & 1 \\ 30.5 & 33.0 & 0.0 & 14.8 & 22.1 & 1 \\ 40.0 & 40.0 & 14.8 & 0.0 & 25.7 & 1 \\ 40.0 & 36.2 & 22.1 & 25.7 & 0.0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

Complétion avec un 1 :

$$\mathbf{B} = \begin{bmatrix} 20.6 & 19.0 & 14.4 & 29.1 & 25.8 & 1 \end{bmatrix}^T$$

# Exemple

Inversion de la matrice de gauche :

$$\mathbf{A}^{-1} = \begin{bmatrix} -0.026 & 0.0155 & 0.009 & 0.000 & 0.000 & 0.274 \\ 0.015 & -0.0264 & 0.002 & 0.001 & 0.006 & 0.230 \\ 0.009 & 0.0027 & -0.051 & 0.026 & 0.011 & -0.049 \\ 0.000 & 0.0017 & 0.026 & -0.037 & 0.008 & 0.296 \\ 0.000 & 0.0064 & 0.011 & 0.008 & -0.027 & 0.248 \\ 0.274 & 0.2304 & -0.049 & 0.296 & 0.248 & -25.846 \end{bmatrix}$$

# Exemple

Inversion de la matrice de gauche :

$$\mathbf{A}^{-1} = \begin{bmatrix} -0.026 & 0.0155 & 0.009 & 0.000 & 0.000 & 0.274 \\ 0.015 & -0.0264 & 0.002 & 0.001 & 0.006 & 0.230 \\ 0.009 & 0.0027 & -0.051 & 0.026 & 0.011 & -0.049 \\ 0.000 & 0.0017 & 0.026 & -0.037 & 0.008 & 0.296 \\ 0.000 & 0.0064 & 0.011 & 0.008 & -0.027 & 0.248 \\ 0.274 & 0.2304 & -0.049 & 0.296 & 0.248 & -25.846 \end{bmatrix}$$

$$\mathbf{A}^{-1}\mathbf{B} = \begin{bmatrix} 0.195 & 0.304 & 0.553 & -0.147 & 0.093 & -1.501 \end{bmatrix}^T$$

# Exemple

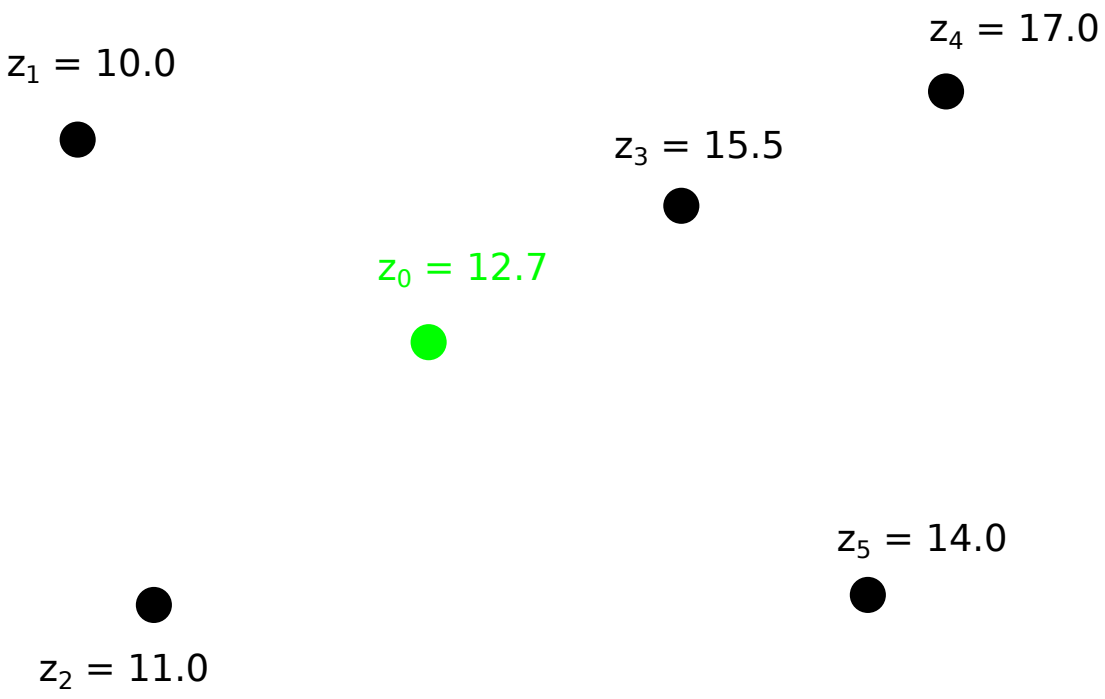
Inversion de la matrice de gauche :

$$\mathbf{A}^{-1} = \begin{bmatrix} -0.026 & 0.0155 & 0.009 & 0.000 & 0.000 & 0.274 \\ 0.015 & -0.0264 & 0.002 & 0.001 & 0.006 & 0.230 \\ 0.009 & 0.0027 & -0.051 & 0.026 & 0.011 & -0.049 \\ 0.000 & 0.0017 & 0.026 & -0.037 & 0.008 & 0.296 \\ 0.000 & 0.0064 & 0.011 & 0.008 & -0.027 & 0.248 \\ 0.274 & 0.2304 & -0.049 & 0.296 & 0.248 & -25.846 \end{bmatrix}$$

$$\mathbf{A}^{-1}\mathbf{B} = \begin{bmatrix} 0.195 & 0.304 & 0.553 & -0.147 & 0.093 & -1.501 \end{bmatrix}^T$$

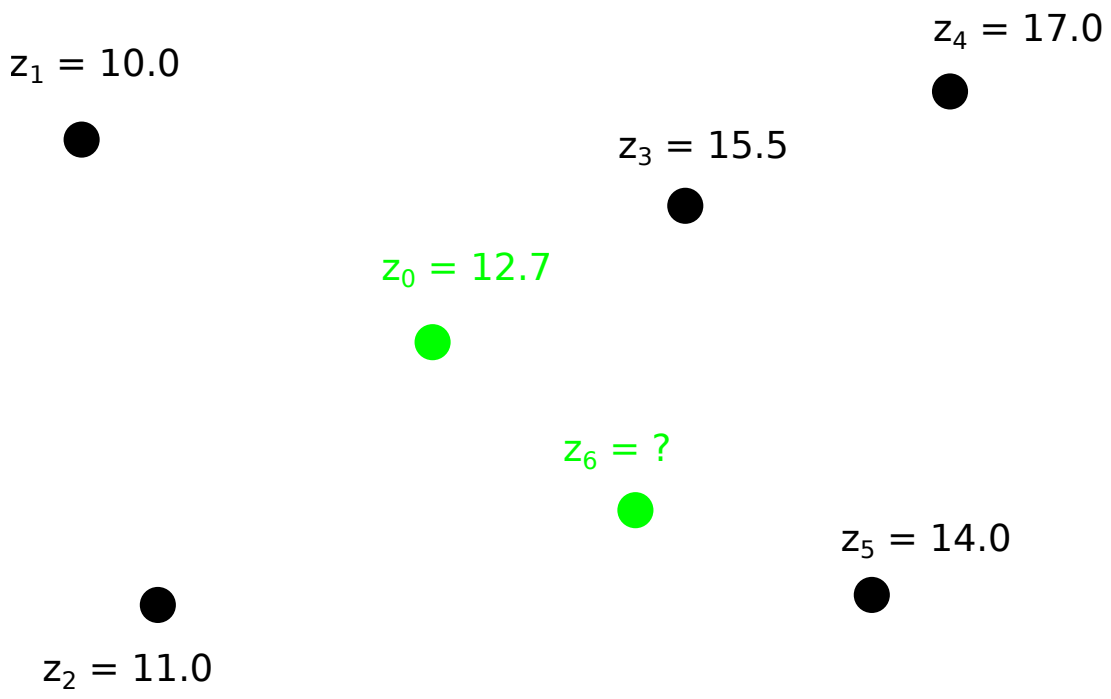
Pondérations : $\mathbf{W} = \begin{bmatrix} 0.195 & 0.304 & 0.553 & -0.147 & 0.093 \end{bmatrix}^T$

# Exemple



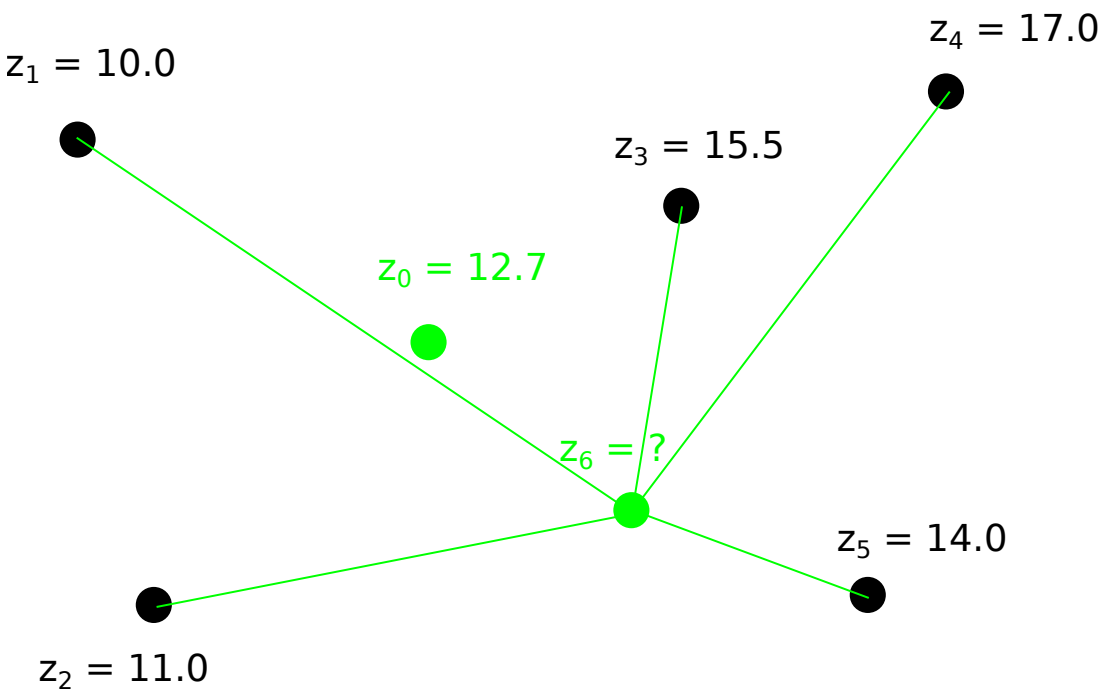

Interpolation : $\hat{z}(x_0) = \mathbf{W}^T\mathbf{Z} = 12.7$

# Exemple




Nouveau point $x_6$ à interpoler ?

# Exemple




Calculs de $H_6 \rightarrow \Gamma_6 \rightarrow B \rightarrow W(x_6) \rightarrow \hat{z}(x_6)$

# Exemple

$\rightarrow$ Interpolation $\forall\ x \in \mathcal{D}$

# Calcul de la variance

En un point donné $x_0$, la variance associée à l'estimation s'exprime par :

$$\sigma_z(x_0)^2 = \sum_{i=1}^{N} w_i \gamma_{i0} - \lambda$$

# Calcul de la variance

En un point donné $x_0$, la variance associée à l'estimation s'exprime par :

$$\sigma_z(x_0)^2 = \sum_{i=1}^{N} w_i \gamma_{i0} - \lambda$$

$\rightarrow$ Le multiplicateur de Lagrange est une variable latente du problème d'estimation, mais intervient dans l'évaluation de l'écart-type associé

# Calcul de la variance

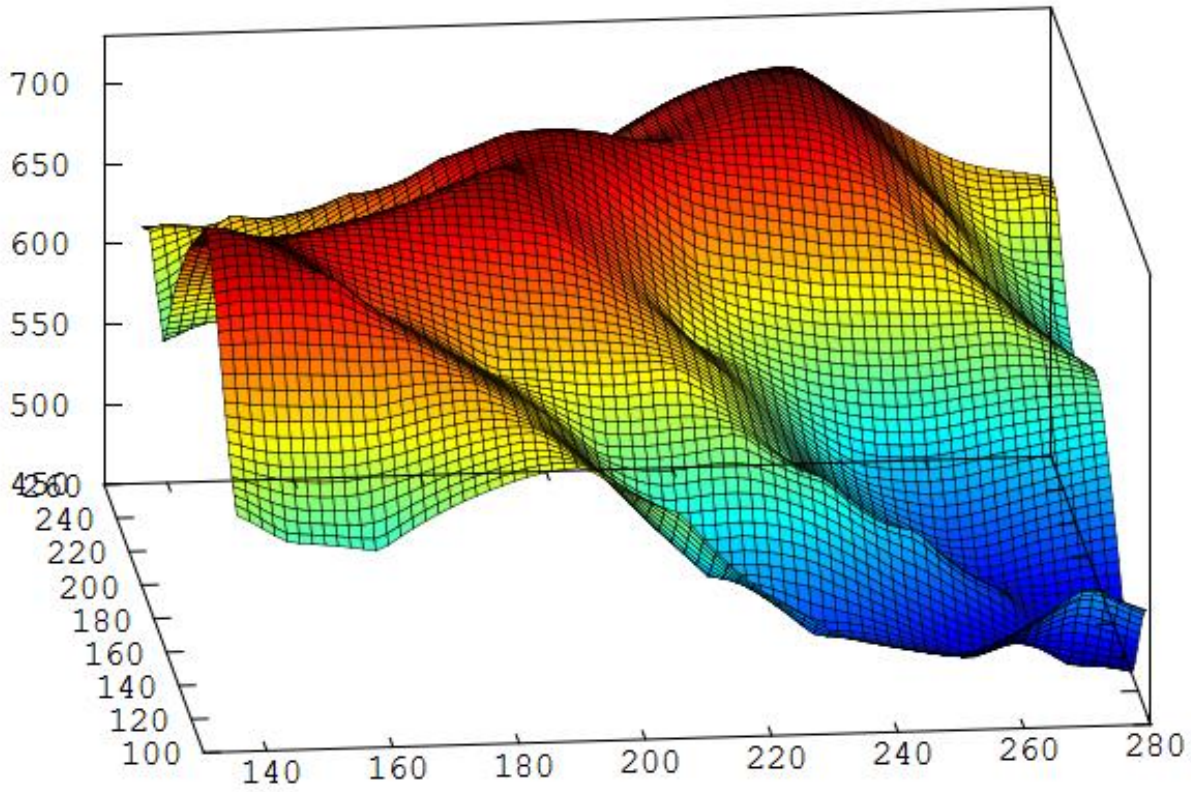
En un point donné $x_0$, la variance associée à l'estimation s'exprime par :

$$\sigma_z(x_0)^2 = \sum_{i=1}^{N} w_i \gamma_{i0} - \lambda$$

$\rightarrow$ Le multiplicateur de Lagrange est une variable latente du problème d'estimation, mais intervient dans l'évaluation de l'écart-type associé

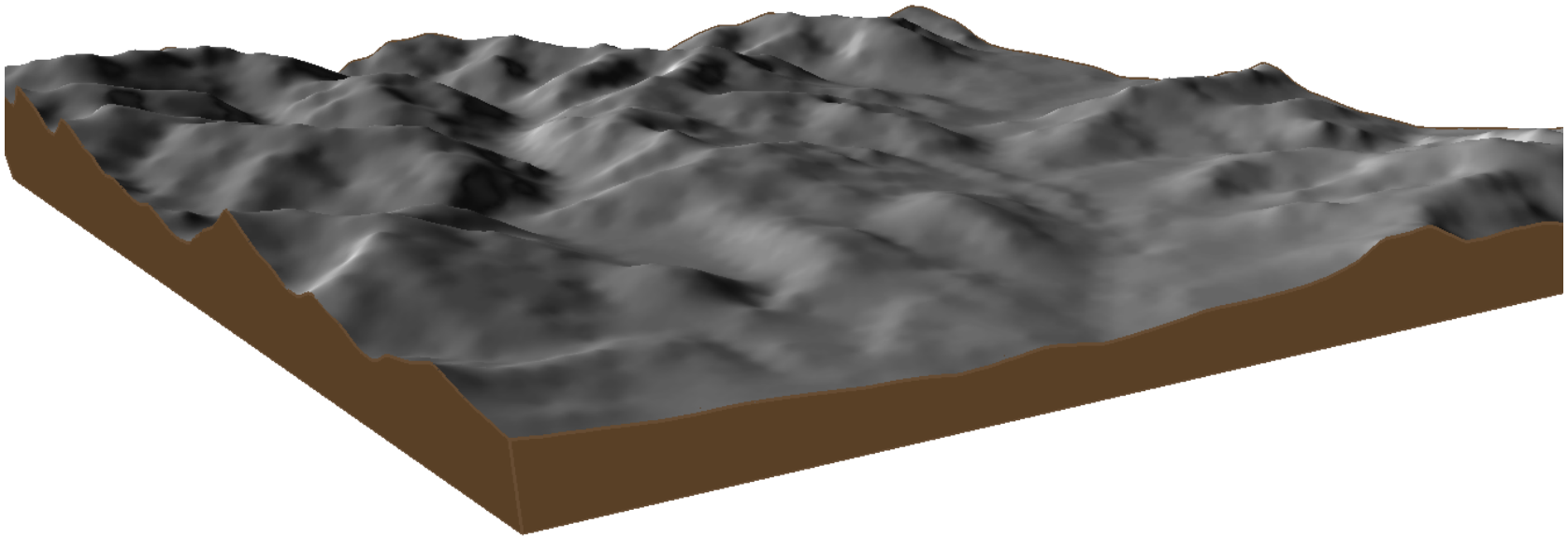
$\rightarrow$ La variance d'estimation en un site ne dépend que de sa configuration géométrique relativement aux autres sites, et non des valeurs observées

# Exemples



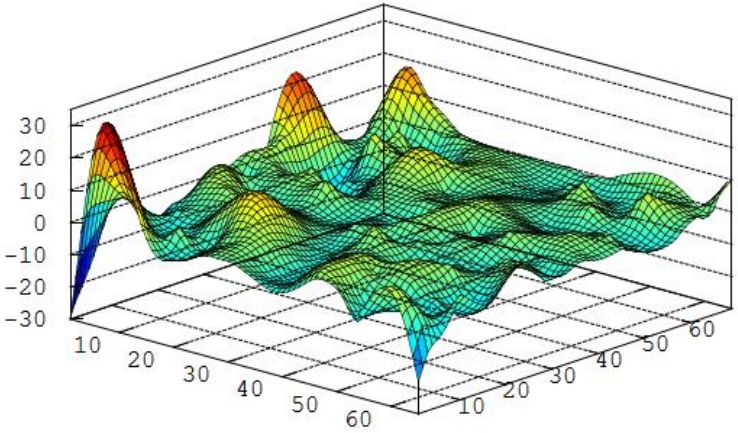

Krigeage avec variographie linéaire (Sartène)

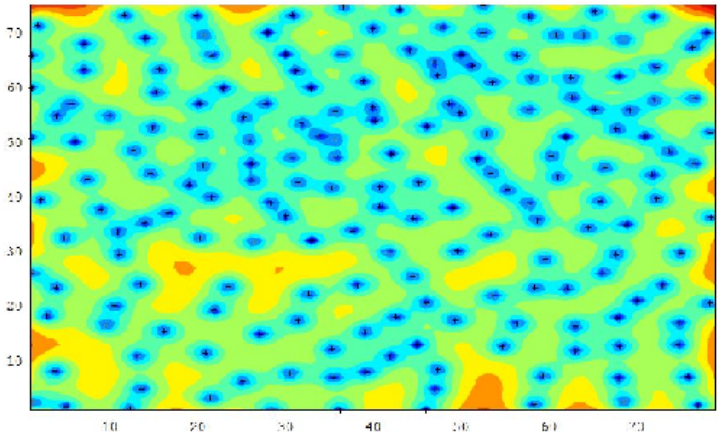
# Exemples



Krigeage avec variographie linéaire (Alpes du Sud)

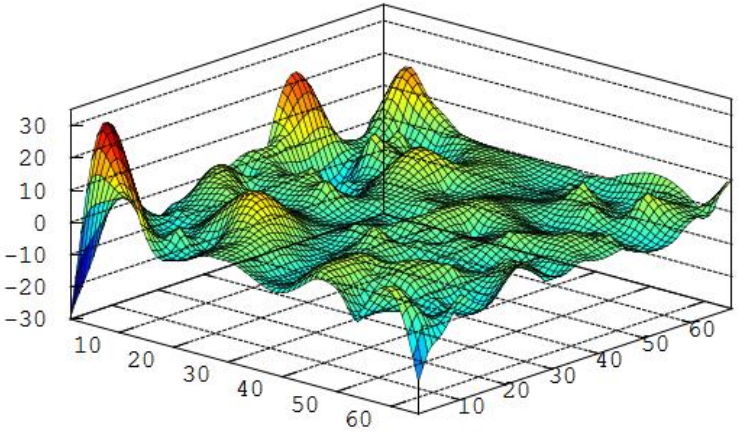
# Exemples

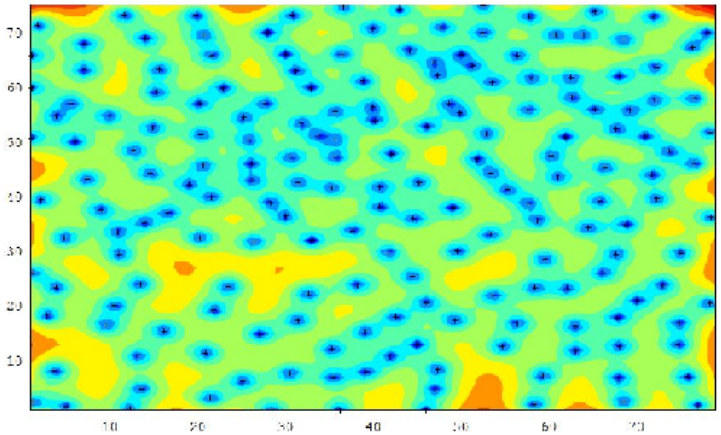


Variance d'estimation

# Exemples

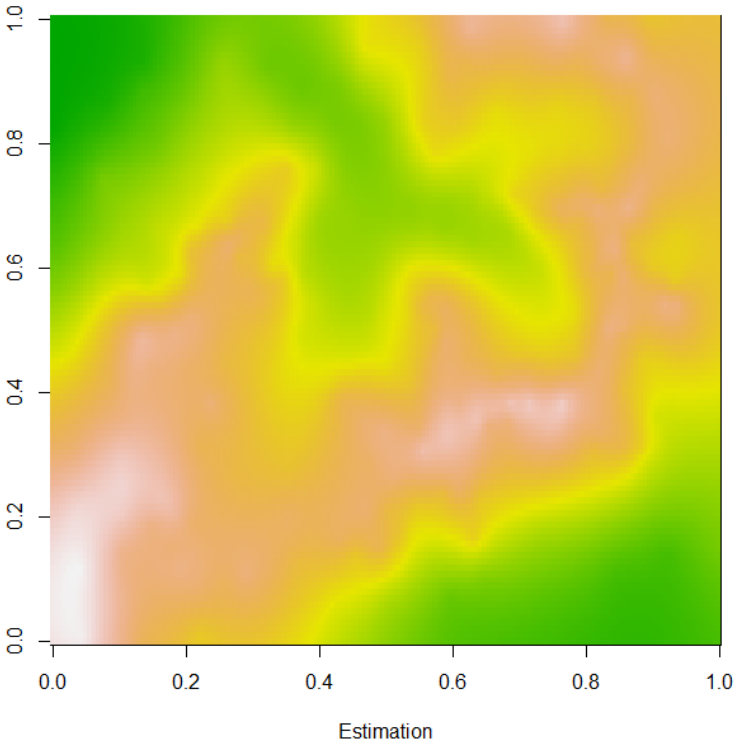


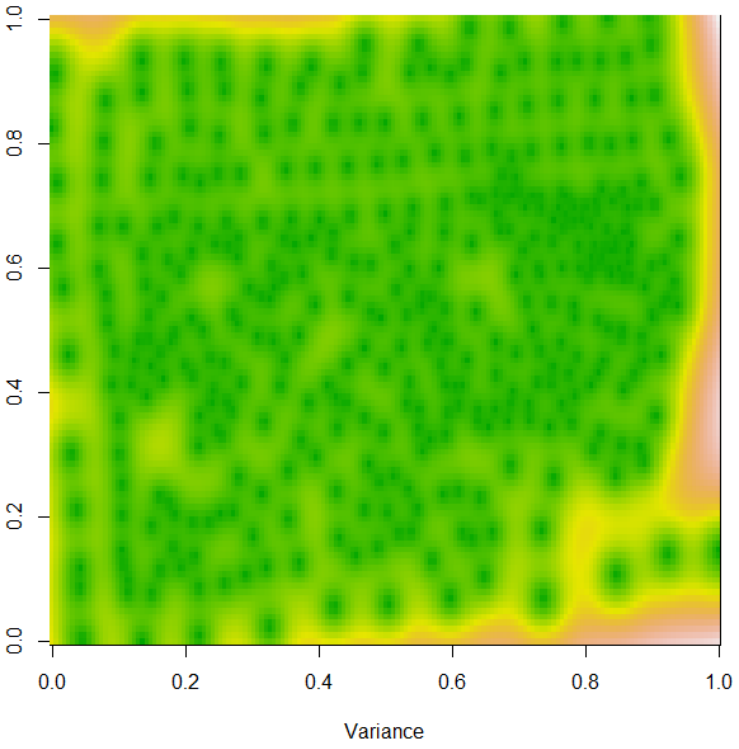



Variance d'estimation

→ Variance nulle au niveau des points d'observations

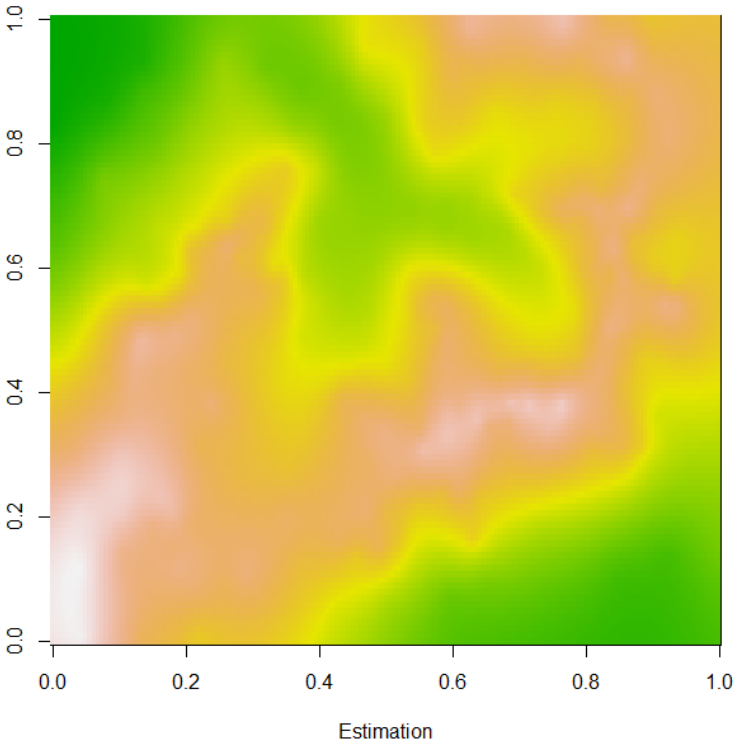
# Exemples

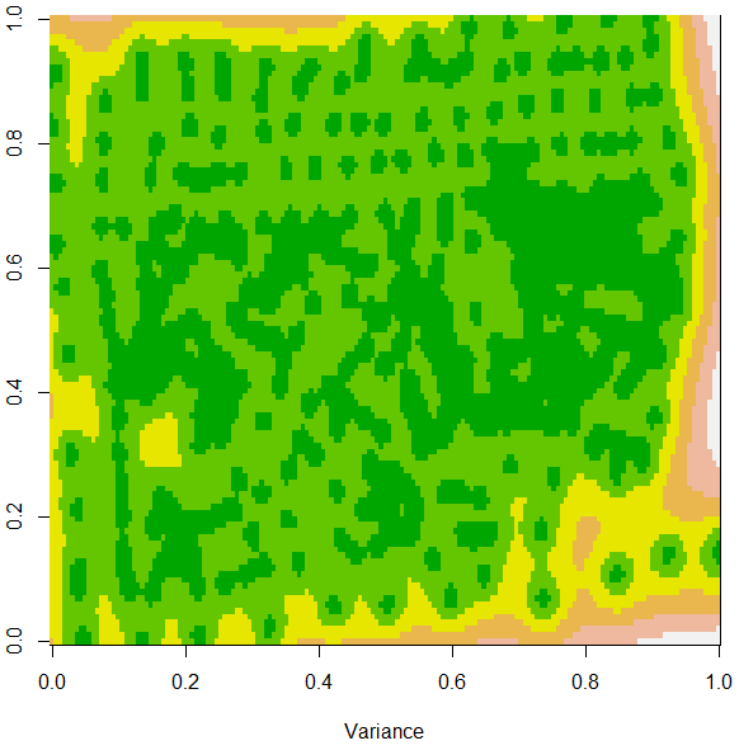



Exemple d'applications sur la Mer de Glace (Massif du Mont Blanc)

# Exemples




Exemple d'applications sur la Mer de Glace (Massif du Mont Blanc)

# Propriétés

## Propriété importante

Pour deux variogrammes $\gamma_1$ et $\gamma_2$ tels que $\gamma_1 = k\gamma_2$, en tout point $x$ du champ $\mathcal{D}$, les estimations de Krigeage menées avec ces deux variogrammes sont identiques :

$$\hat{z}(x, \gamma_1) = \hat{z}(x, \gamma_2) \quad \forall\, x \in \mathcal{D}$$

En revanche, les variances d'estimations seront reliées par :

$$\mathrm{Var}[\hat{z}(x, \gamma_1)] = k\mathrm{Var}[\hat{z}(x, \gamma_2)]$$

# Propriétés

## Propriété importante

Pour deux variogrammes $\gamma_1$ et $\gamma_2$ tels que $\gamma_1 = k\gamma_2$, en tout point $x$ du champ $\mathcal{D}$, les estimations de Krigeage menées avec ces deux variogrammes sont identiques :

$$\hat{z}(x, \gamma_1) = \hat{z}(x, \gamma_2) \quad \forall\, x \in \mathcal{D}$$

En revanche, les variances d'estimations seront reliées par :

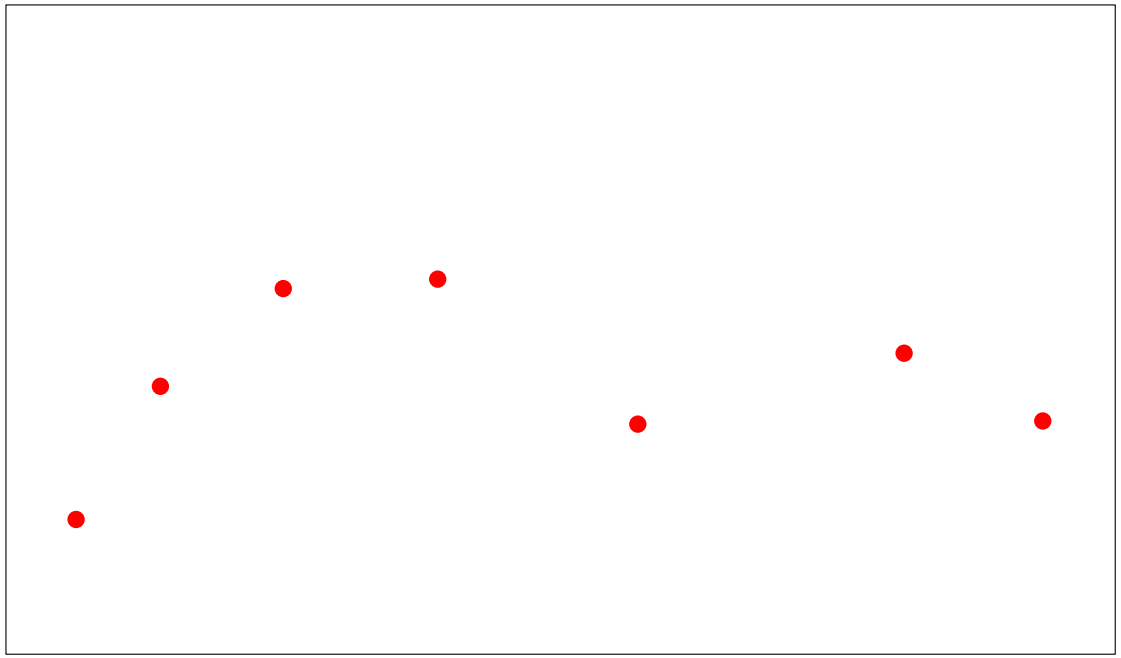
$$\mathrm{Var}[\hat{z}(x, \gamma_1)] = k\mathrm{Var}[\hat{z}(x, \gamma_2)]$$

Ce résultat se déduit immédiatement des équations de Krigeage : $\sum_j w_j \gamma_{ij} + \lambda = \gamma_{i0}$. Si $(w_i, \lambda)$ est solution pour le variogramme $\gamma_1$, alors $(w_i, k\lambda)$ est solution pour le variogramme $\gamma_2$.

# Propriétés

On a vu que les poids de Krigeage n'étaient pas nécessairement compris entre 0 et 1 $\rightarrow$ combinaison linéaire non-convexe des observations.

Implication

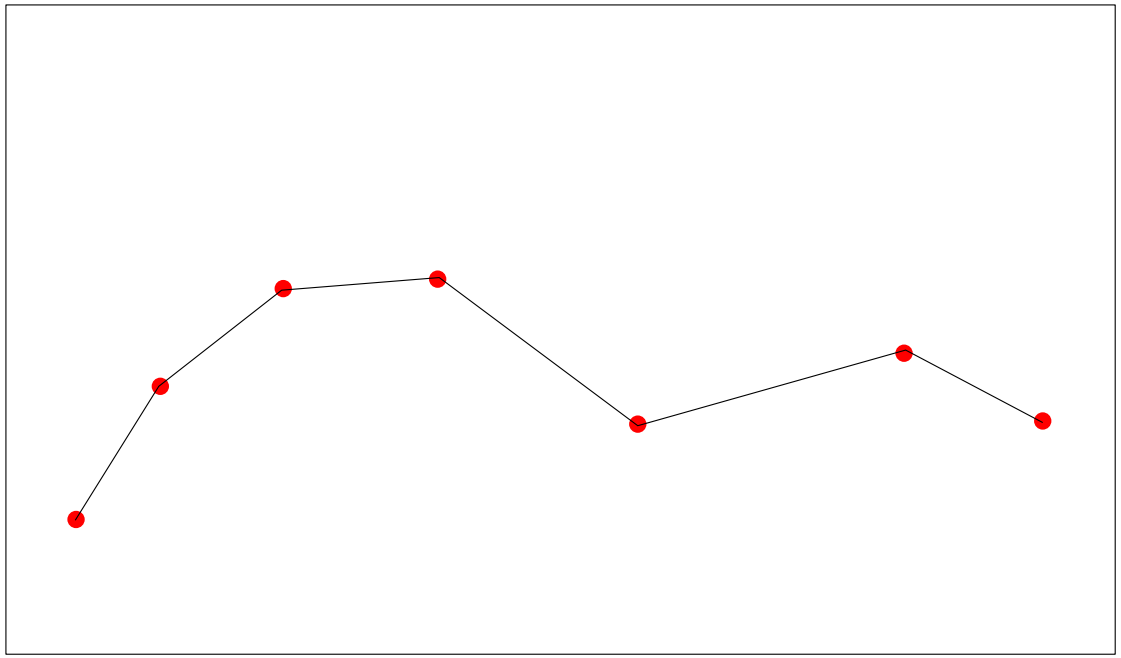


Semis de points d'observations

# Propriétés

On a vu que les poids de Krigeage n'étaient pas nécessairement compris entre 0 et 1 $\rightarrow$ combinaison linéaire non-convexe des observations.

Implication

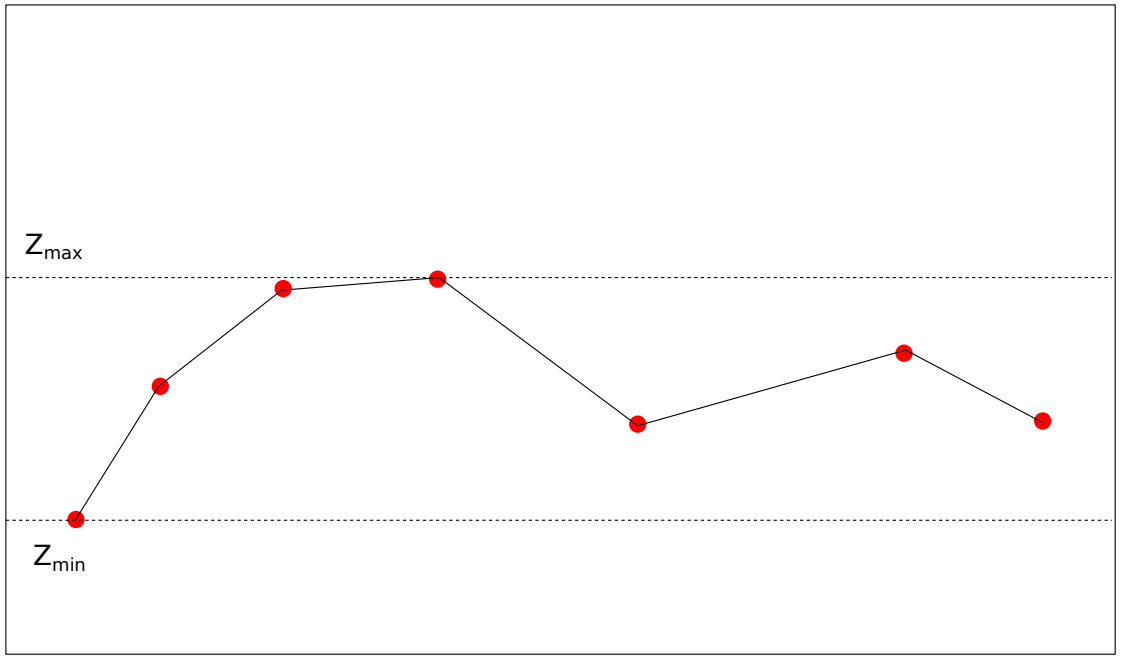


Interpolation linéaire

# Propriétés

On a vu que les poids de Krigeage n'étaient pas nécessairement compris entre 0 et 1 $\rightarrow$ combinaison linéaire non-convexe des observations.

Implication

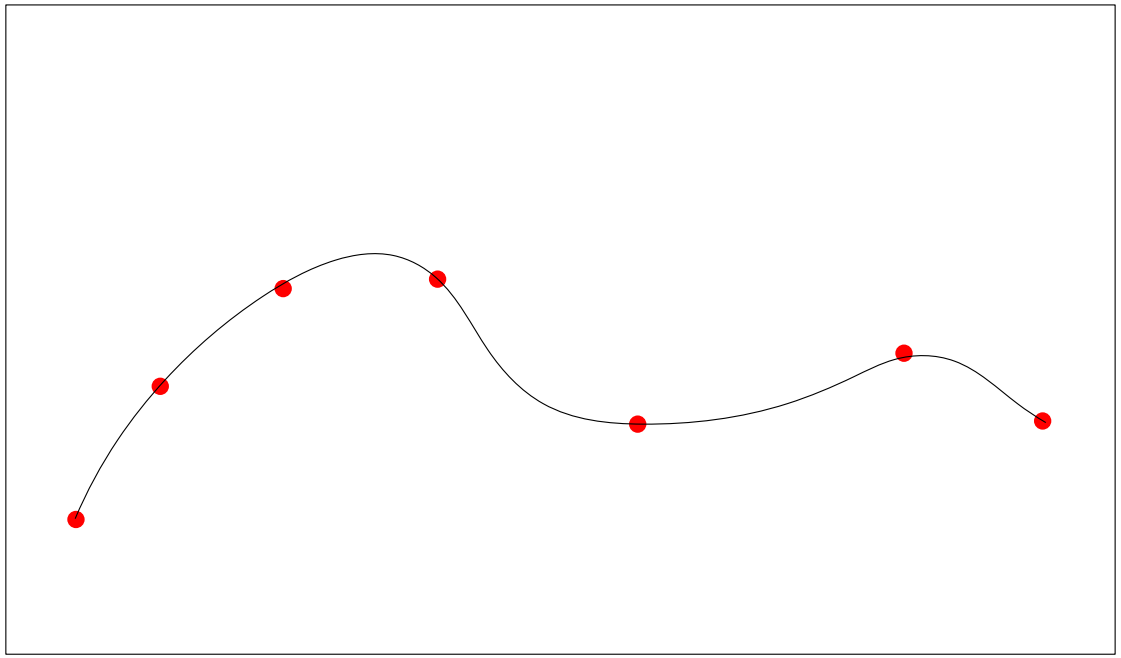



Interpolation linéaire non-convexe

# Propriétés

On a vu que les poids de Krigeage n'étaient pas nécessairement compris entre 0 et 1 $\rightarrow$ combinaison linéaire non-convexe des observations.

Implication

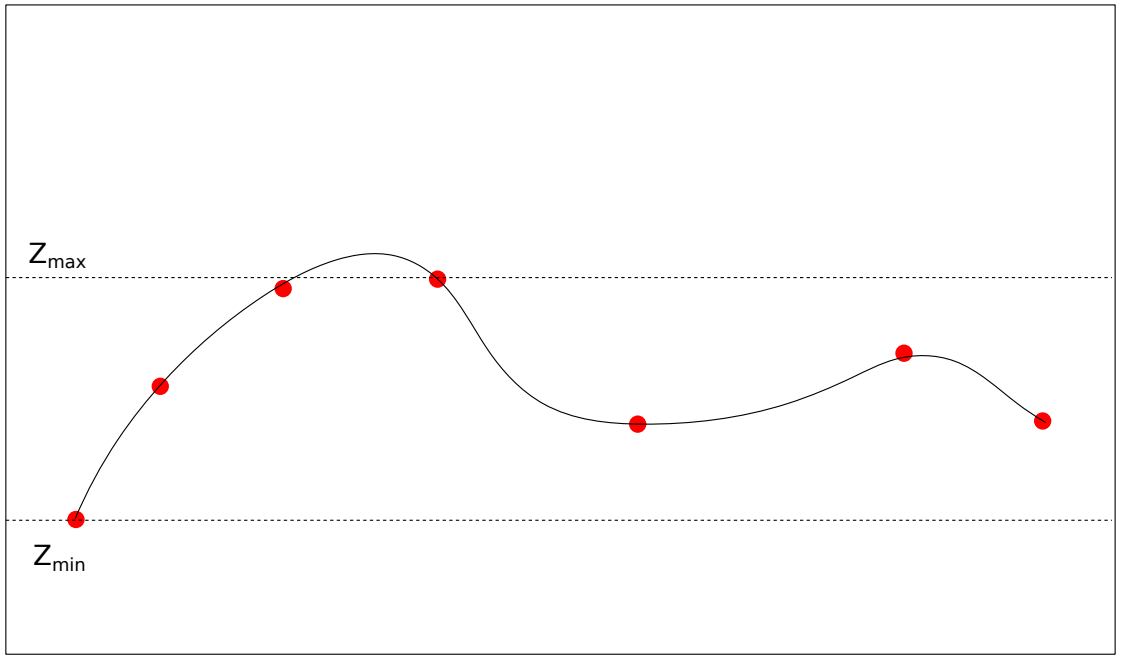


Interpolation par Krigeage

# Propriétés

On a vu que les poids de Krigeage n'étaient pas nécessairement compris entre 0 et 1 $\rightarrow$ combinaison linéaire non-convexe des observations.

Implication




Interpolation par Krigeage : non convexe

# Propriétés

$\oplus$ Le Krigeage est $\mathcal{C}^\infty$ sur $\mathcal{D}$ $\rightarrow$ Surfaces esthétiques

# Propriétés

- $\oplus$ Le Krigeage est $\mathcal{C}^\infty$ sur $\mathcal{D} \rightarrow$ Surfaces esthétiques
- $\oplus$ Prise en compte de la configuration géométrique des observations

# Propriétés

$\oplus$ Le Krigeage est $\mathcal{C}^\infty$ sur $\mathcal{D} \rightarrow$ Surfaces esthétiques

$\oplus$ Prise en compte de la configuration géométrique des observations

$\oplus$ C'est un interpolateur exact mais il est possible de gérer les erreurs de mesure grâce à l'effet pépitique du variogramme

# Propriétés

⊕ Le Krigeage est $\mathcal{C}^\infty$ sur $\mathcal{D} \rightarrow$ Surfaces esthétiques

⊕ Prise en compte de la configuration géométrique des observations

⊕ C'est un interpolateur exact mais il est possible de gérer les erreurs de mesure grâce à l'effet pépitique du variogramme

⊕ Toutes les observations sont a priori prises en compte dans le calcul d'estimation

# Propriétés

⊕ Le Krigeage est $\mathcal{C}^\infty$ sur $\mathcal{D} \rightarrow$ Surfaces esthétiques

⊕ Prise en compte de la configuration géométrique des observations

⊕ C'est un interpolateur exact mais il est possible de gérer les erreurs de mesure grâce à l'effet pépitique du variogramme

⊕ Toutes les observations sont a priori prises en compte dans le calcul d'estimation

⊕ Calcul relativement rapide (optimisation par Krigeage dual)

# Propriétés

⊕ Le Krigeage est $\mathcal{C}^\infty$ sur $\mathcal{D} \rightarrow$ Surfaces esthétiques
⊕ Prise en compte de la configuration géométrique des observations
⊕ C'est un interpolateur exact mais il est possible de gérer les erreurs de mesure grâce à l'effet pépitique du variogramme
⊕ Toutes les observations sont a priori prises en compte dans le calcul d'estimation
⊕ Calcul relativement rapide (optimisation par Krigeage dual)
⊕ Possibilité d'ajout d'information externe (co-krigeage)

# Propriétés

⊕ Le Krigeage est $\mathcal{C}^\infty$ sur $\mathcal{D} \rightarrow$ Surfaces esthétiques
⊕ Prise en compte de la configuration géométrique des observations
⊕ C'est un interpolateur exact mais il est possible de gérer les erreurs de mesure grâce à l'effet pépitique du variogramme
⊕ Toutes les observations sont a priori prises en compte dans le calcul d'estimation
⊕ Calcul relativement rapide (optimisation par Krigeage dual)
⊕ Possibilité d'ajout d'information externe (co-krigeage)
⊕ Calcul de la variance d'estimation

# Propriétés

⊕ Le Krigeage est $\mathcal{C}^\infty$ sur $\mathcal{D} \rightarrow$ Surfaces esthétiques
⊕ Prise en compte de la configuration géométrique des observations
⊕ C'est un interpolateur exact mais il est possible de gérer les erreurs de mesure grâce à l'effet pépitique du variogramme
⊕ Toutes les observations sont a priori prises en compte dans le calcul d'estimation
⊕ Calcul relativement rapide (optimisation par Krigeage dual)
⊕ Possibilité d'ajout d'information externe (co-krigeage)
⊕ Calcul de la variance d'estimation

⊖ Nécessité de validation de l'hypothèse de stationnarité d'ordre 2 (ou éventuellement sa relaxation : hypothèse intrinsèque)

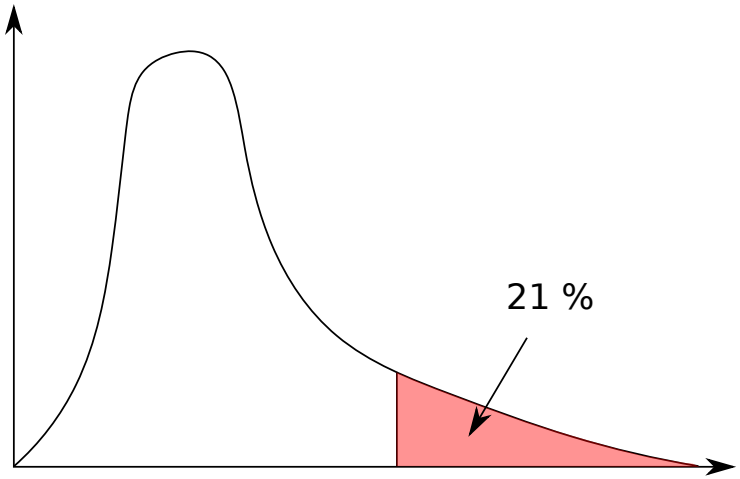
# Propriétés

⊕ Le Krigeage est $\mathcal{C}^\infty$ sur $\mathcal{D}$ → Surfaces esthétiques
⊕ Prise en compte de la configuration géométrique des observations
⊕ C'est un interpolateur exact mais il est possible de gérer les erreurs de mesure grâce à l'effet pépitique du variogramme
⊕ Toutes les observations sont a priori prises en compte dans le calcul d'estimation
⊕ Calcul relativement rapide (optimisation par Krigeage dual)
⊕ Possibilité d'ajout d'information externe (co-krigeage)
⊕ Calcul de la variance d'estimation

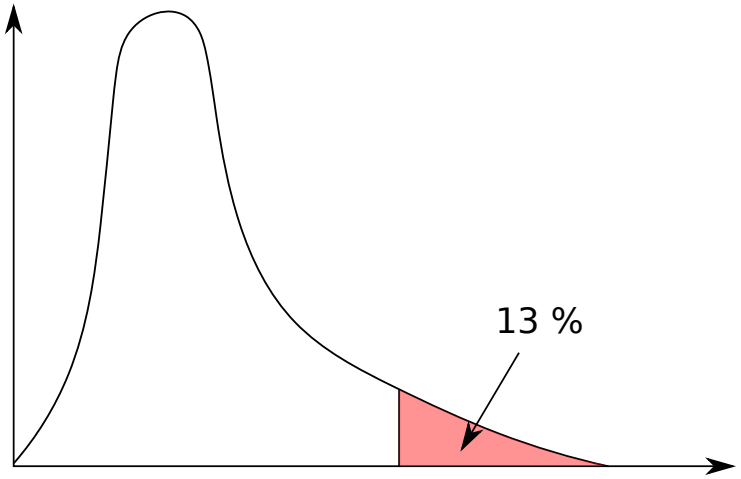
⊖ Nécessité de validation de l'hypothèse de stationnarité d'ordre 2 (ou éventuellement sa relaxation : hypothèse intrinsèque)
⊖ **Inadapté aux transformations non-linéaires** (seuillage, calculs de risques...) → géostatistique non linéaire (ou simulations)

# Propriétés

## Propriété de lissage du krigeage



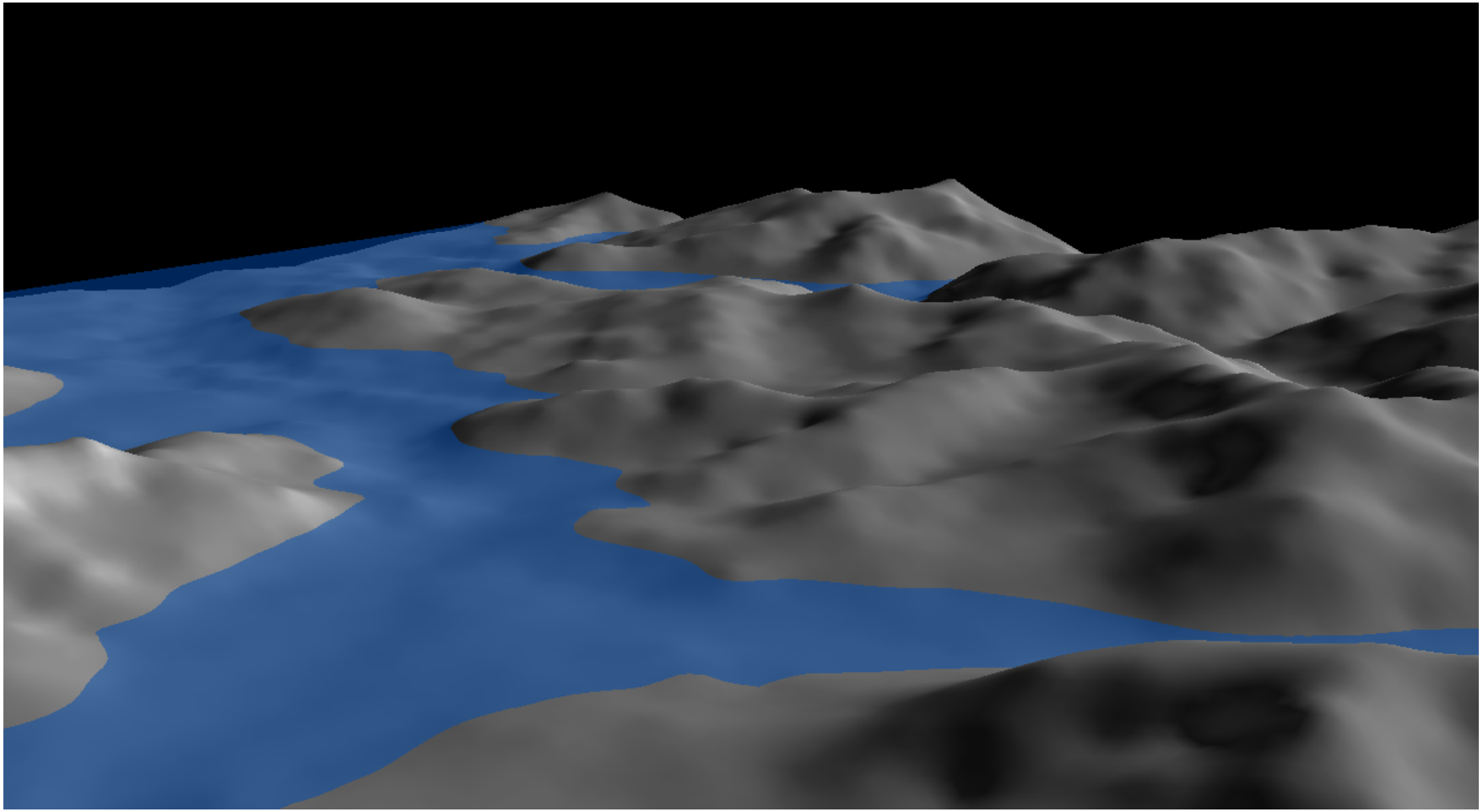

Distribution réelle des valeurs




Distribution des valeurs estimées par krigeage

# Propriétés

Propriété de lissage du krigeage



# Variantes du Krigeage

La version de Krigeage étudiée précedemment, est classiquement nommée **krigeage ordinaire**. On l'appelle aussi Krigeage stationnaire à moyenne inconnue. Il existe de nombreuses autres versions de la technique du krigeage.

→ Krigeage simple (moyenne connue)

→ Cokrigeage (multivarié)

→ Krigeage universel (ou à dérive externe)

→ Krigeage d'indicatrice, Krigeage de gradient, Krigeage de blocs, Krigeage avec contraintes d'inégalité...

# Variantes du krigeage

**Le krigeage simple** suppose la connaissance de la valeur moyenne $m$ de la variable régionalisée

$$\hat{z}(x_0) = \sum_{i=1}^{N} w_i z(x_i) + \left(1 - \sum_{i=1}^{N}\right) m$$

# Variantes du krigeage

**Le krigeage simple** suppose la connaissance de la valeur moyenne $m$ de la variable régionalisée

$$\hat{z}(x_0) = \sum_{i=1}^{N} w_i z(x_i) + \left(1 - \sum_{i=1}^{N}\right) m$$

L'hypothèse de somme des poids nulle n'est alors plus nécessaire pour obtenir un estimateur non-biaisé.

# Variantes du krigeage

**Le krigeage simple** suppose la connaissance de la valeur moyenne $m$ de la variable régionalisée

$$\hat{z}(x_0) = \sum_{i=1}^{N} w_i z(x_i) + \Big(1 - \sum_{i=1}^{N}\Big) m$$

L'hypothèse de somme des poids nulle n'est alors plus nécessaire pour obtenir un estimateur non-biaisé.

$\rightarrow$ La variance d'estimation du krigeage simple est toujours inférieure à celle du krigeage ordinaire (la connaissance supposée de la moyenne est une information additionnelle).

# Variantes du krigeage

**Le krigeage simple** suppose la connaissance de la valeur moyenne $m$ de la variable régionalisée

$$\hat{z}(x_0) = \sum_{i=1}^{N} w_i z(x_i) + \Big(1 - \sum_{i=1}^{N}\Big) m$$

L'hypothèse de somme des poids nulle n'est alors plus nécessaire pour obtenir un estimateur non-biaisé.

$\rightarrow$ La variance d'estimation du krigeage simple est toujours inférieure à celle du krigeage ordinaire (la connaissance supposée de la moyenne est une information additionnelle).

$\rightarrow$ La moyenne de la variable régionalisée est bien souvent inconnue en pratique $\rightarrow$ méthode souvent inopérante

# Variantes du krigeage

**Le cokrigeage :** (extension au cas multivarié) permet d'interpoler plusieurs variables en même temps losque celles-ci ne sont pas indépendantes (e.g. météo, température, pression etc).

# Variantes du krigeage

**Le cokrigeage :** (extension au cas multivarié) permet d'interpoler plusieurs variables en même temps losque celles-ci ne sont pas indépendantes (e.g. météo, température, pression etc).

Variogrammes croisés : $\gamma_{ij}(h) = \frac{1}{2}\text{Var}[z_i(x+h) - z_j(x)]$

# Variantes du krigeage

**Le cokrigeage :** (extension au cas multivarié) permet d'interpoler plusieurs variables en même temps losque celles-ci ne sont pas indépendantes (e.g. météo, température, pression etc).

Variogrammes croisés : $\gamma_{ij}(h) = \frac{1}{2}\text{Var}[z_i(x+h) - z_j(x)]$

La variance de cokrigeage est inférieure à la variance de krigeage (informations externes sur les liens de dépendances entre variables).

# Variantes du krigeage

**Le cokrigeage :** (extension au cas multivarié) permet d'interpoler plusieurs variables en même temps losque celles-ci ne sont pas indépendantes (e.g. météo, température, pression etc).

Variogrammes croisés : $\gamma_{ij}(h) = \frac{1}{2}\text{Var}[z_i(x+h) - z_j(x)]$

La variance de cokrigeage est inférieure à la variance de krigeage (informations externes sur les liens de dépendances entre variables).

Dans sa version simple, le cokrigeage permet d'affiner la précision d'estimation d'une variable $z_1$, à l'aide d'une seconde variable latente $z_2$ observée soit aux mêmes sites, soit en des sites différents.

# Variantes du krigeage

**Le cokrigeage :** (extension au cas multivarié) permet d'interpoler plusieurs variables en même temps losque celles-ci ne sont pas indépendantes (e.g. météo, température, pression etc).

Variogrammes croisés : $\gamma_{ij}(h) = \frac{1}{2}\text{Var}[z_i(x+h) - z_j(x)]$

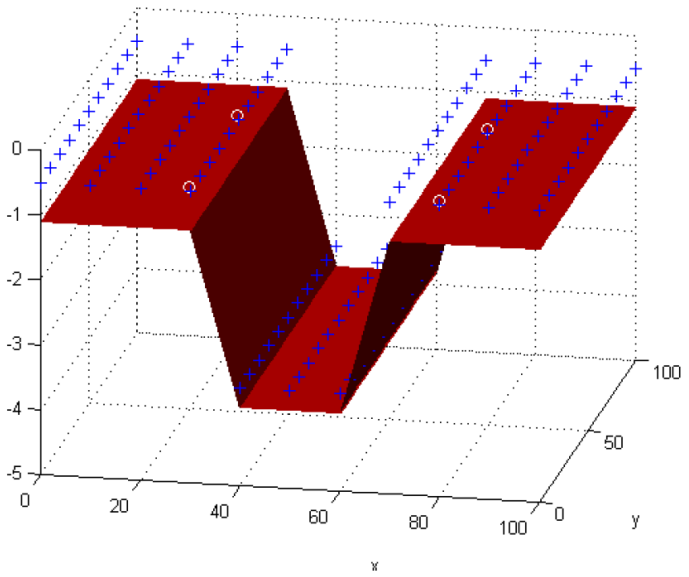
La variance de cokrigeage est inférieure à la variance de krigeage (informations externes sur les liens de dépendances entre variables).

Dans sa version simple, le cokrigeage permet d'affiner la précision d'estimation d'une variable $z_1$, à l'aide d'une seconde variable latente $z_2$ observée soit aux mêmes sites, soit en des sites différents.

On distingue de même le cokrigeage ordinaire, le cokrigeage simple, le cas mixte correspondant à la connaissance de la moyenne d'une partie seulement des variables

# Variantes du krigeage

Exemple de cokrigeage (Gloaguen, 2000) :

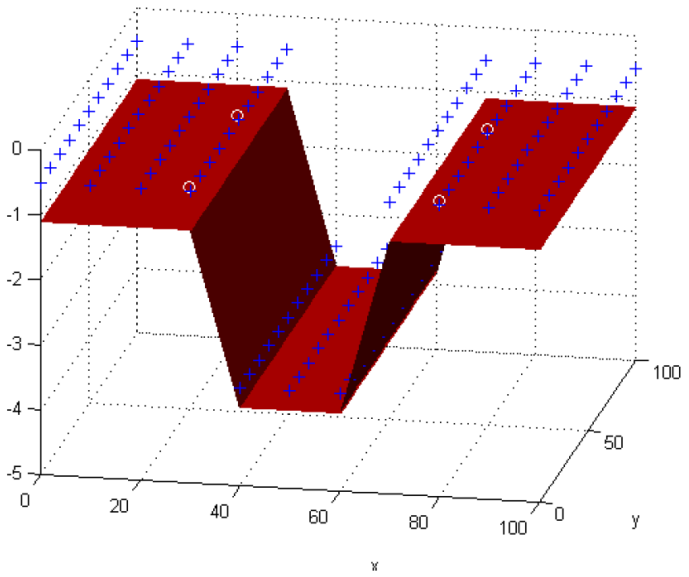



$z_1$ = forage en 4 points données
$z_2$ = observations radar sur une grille de 10 m de résolution

Objectif : déterminer le niveau de fond de la nappe

# Variantes du krigeage

Exemple de cokrigeage (Gloaguen, 2000) :

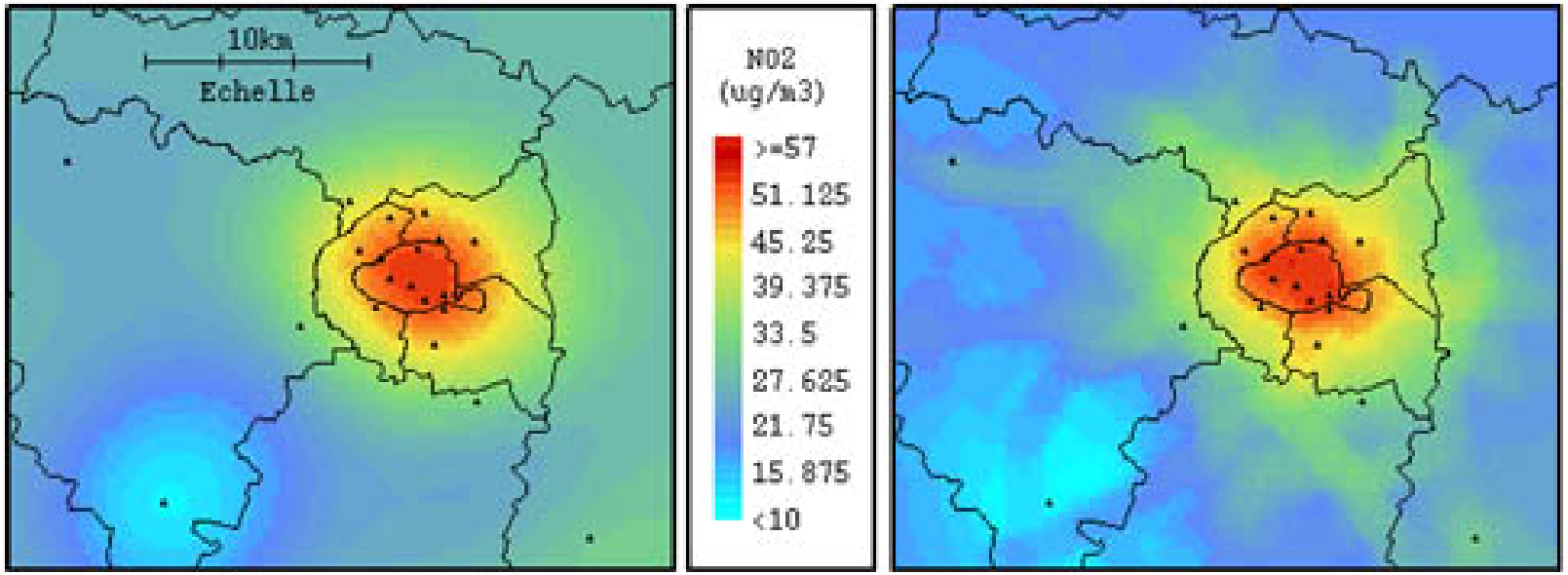



$z_1$ = forage en 4 points données
$z_2$ = observations radar sur une grille de 10 m de résolution

$$\hat{z}(x_0) = \sum_{i=1}^{2} \sum_{j=1}^{N} w_{ij} z_i(x_j)$$

# Variantes du krigeage

Exemple de cokrigeage (J.J. Peraudin, 2003) :




(a) Krigeage de NO2

(b) Co-krigeage avec les emissions

# Variantes du krigeage

**Le krigeage universel** permet de s'affranchir de l'hypothèse de stationnarité :

$$Z(x) = Y(x) + m(x)$$

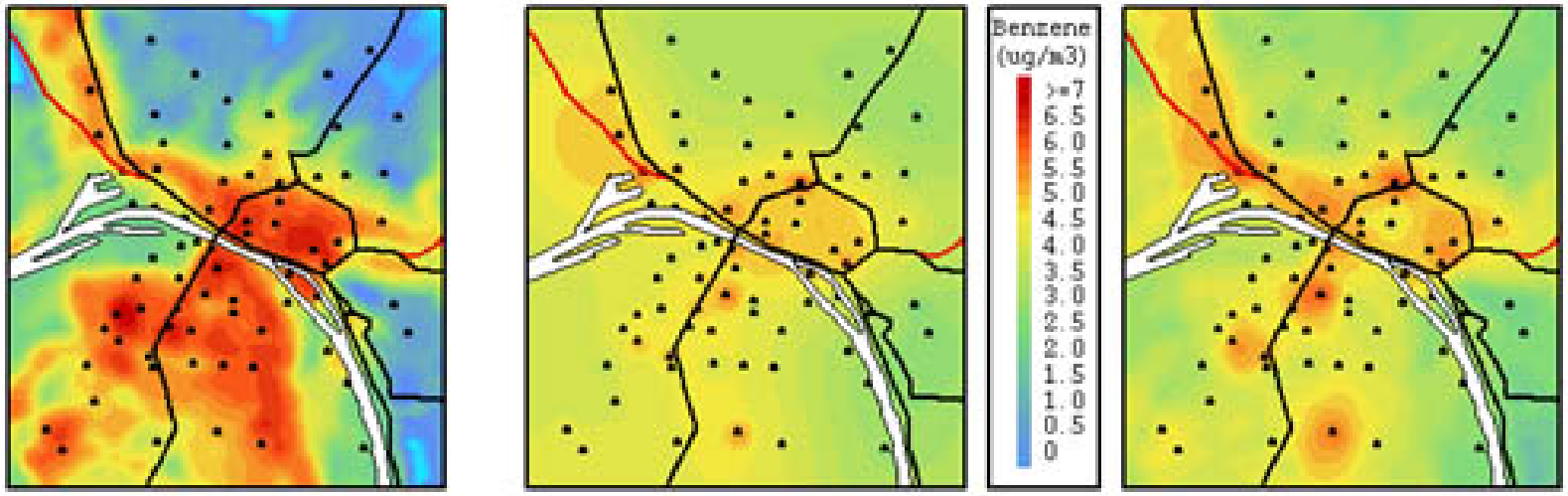
avec $m$ une fonction déterministe de $\mathcal{D} \to \mathbb{R}$ connue

**Le krigeage à dérive externe** est une version plus générale du krigeage universel :

$$Z(x) = Y(x) + am(x) + b$$

avec $a$ et $b$ deux constantes réelles à estimer.

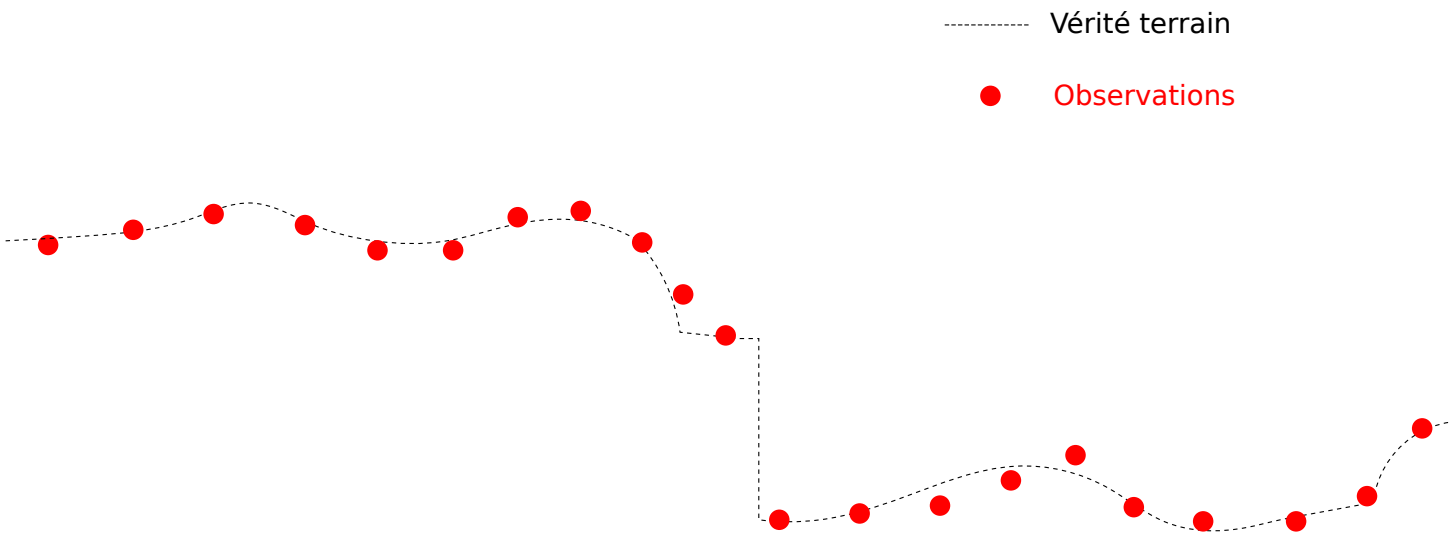
# Variantes du krigeage




(a) Derive externe (b) Krigeage ordinaire (c) Krigeage avec dérive externe

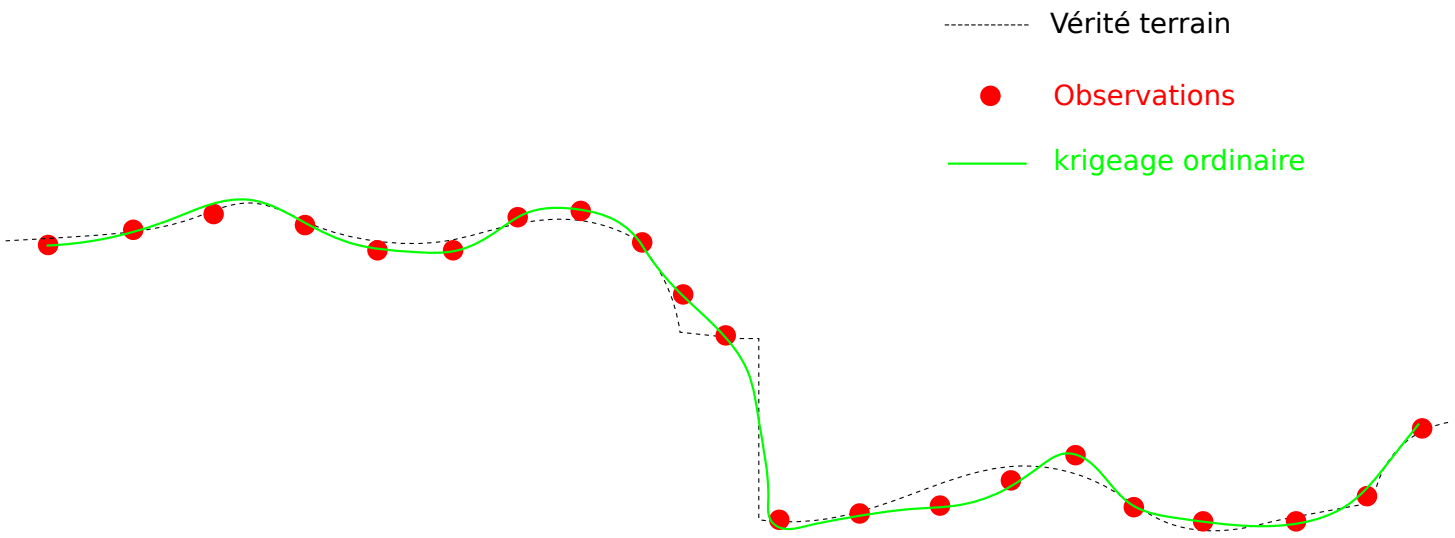
Exemple de krigeage à dérive externe (J.J. Peraudin, 2003) :

# Variantes du krigeage



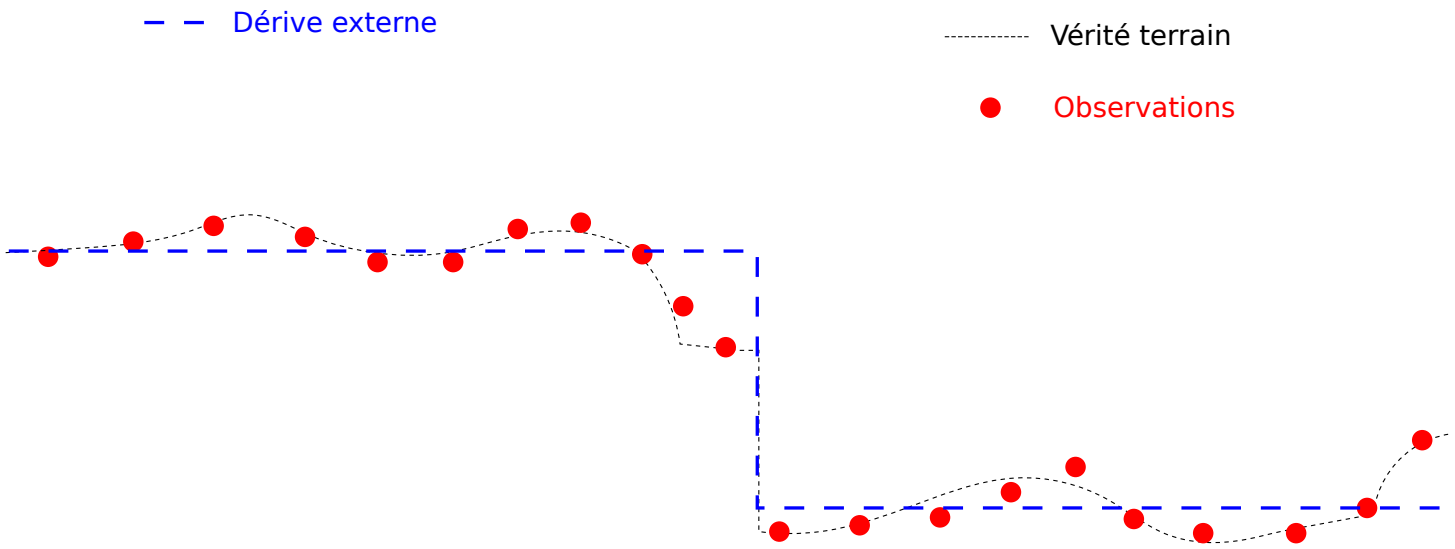

Variable régionalisée et observations

# Variantes du krigeage



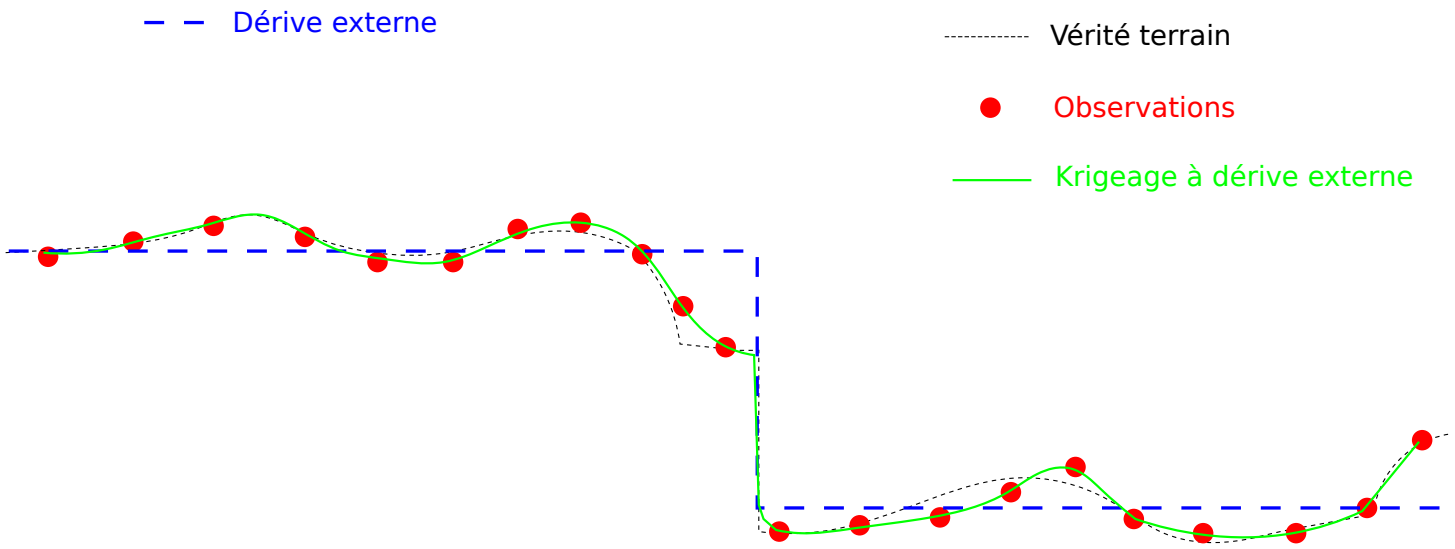

Krigeage ordinaire : difficulté de modélisation de la discontinuité

# Variantes du krigeage



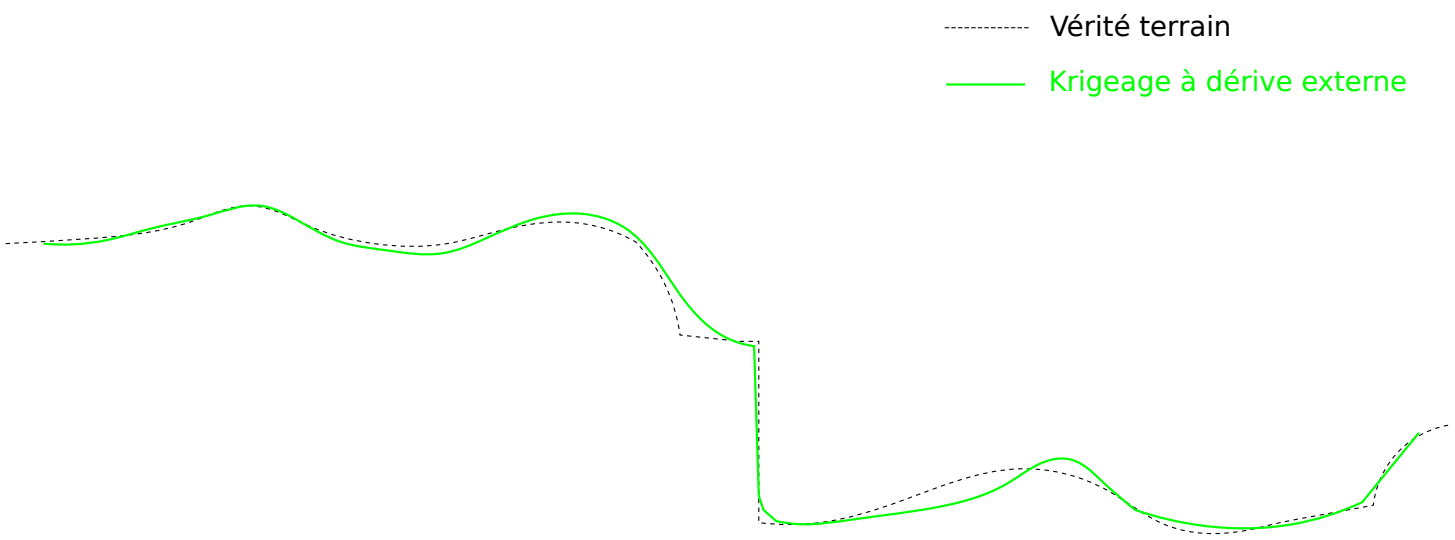

Introduction d'une dérive externe

# Variantes du krigeage




Krigeage avec dérive externe

# Variantes du krigeage




Krigeage avec dérive externe : modélisation fine de la discontinuité

# Variantes du krigeage

⚠ Attention, le krigeage dual est une méthode d'optimisation permettant de pré-calculer le motif matriciel pour accélerer les estimations en chaque point.

# Variantes du krigeage

⚠ Attention, le krigeage dual est une méthode d'optimisation permettant de pré-calculer le motif matriciel pour accélerer les estimations en chaque point.

La présentation duale du krigeage permet de calculer $\hat{z}(x)$ en tout point $x$ du champ en ne résolvant qu'un seul système d'équations.

# Variantes du krigeage

⚠ Attention, le krigeage dual est une méthode d'optimisation permettant de pré-calculer le motif matriciel pour accélerer les estimations en chaque point.

La présentation duale du krigeage permet de calculer $\hat{z}(x)$ en tout point $x$ du champ en ne résolvant qu'un seul système d'équations.

Il ne s'agit donc pas d'une approche différente.

# Variantes du krigeage

⚠ Attention, le krigeage dual est une méthode d'optimisation permettant de pré-calculer le motif matriciel pour accélerer les estimations en chaque point.

La présentation duale du krigeage permet de calculer $\hat{z}(x)$ en tout point $x$ du champ en ne résolvant qu'un seul système d'équations.

Il ne s'agit donc pas d'une approche différente.

Le krigeage dual présente l'inconvénient de ne pas être adapté à la fois au calcul rapide de l'estimation et à celui de la variance associée.

# Variantes du krigeage

⚠ Attention, le krigeage dual est une méthode d'optimisation permettant de pré-calculer le motif matriciel pour accélerer les estimations en chaque point.

La présentation duale du krigeage permet de calculer $\hat{z}(x)$ en tout point $x$ du champ en ne résolvant qu'un seul système d'équations.

Il ne s'agit donc pas d'une approche différente.

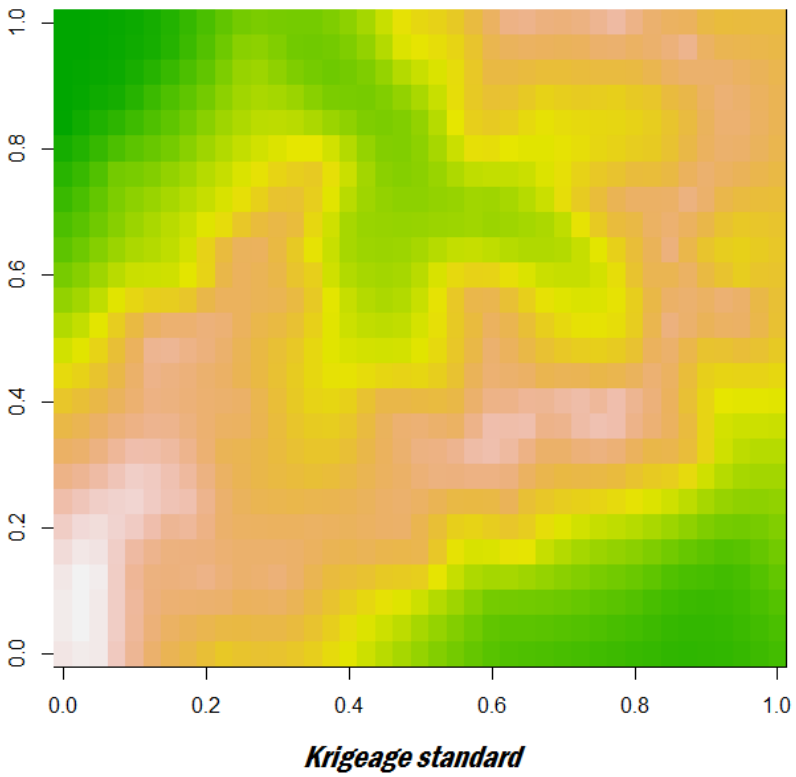
Le krigeage dual présente l'inconvénient de ne pas être adapté à la fois au calcul rapide de l'estimation et à celui de la variance associée.

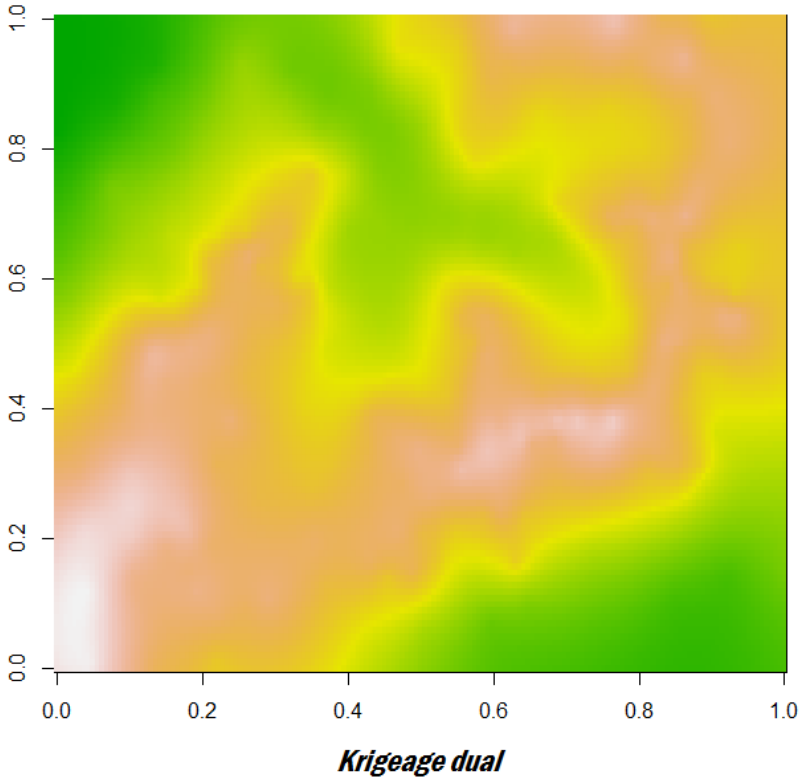
La présentation duale du krigeage permet de faire le lien avec le formalisme des splines de régression.

# Variantes du krigeage

Accélération par la technique du krigeage dual (comparaison pour un temps de calcul fixé à $t = 1.22$ s)

krigeage_dual.r



# Validation croisée

En utilisant la propriété d'invariance de l'estimation par multiplication du variogramme par un scalaire, on peut toujours trouver un variogramme $\gamma$, tel que la variance sur l'estimation $\hat{z}$ en tout point $x$ soit aussi petite qu'on le souhaite.

# Validation croisée

En utilisant la propriété d'invariance de l'estimation par multiplication du variogramme par un scalaire, on peut toujours trouver un variogramme $\gamma$, tel que la variance sur l'estimation $\hat{z}$ en tout point $x$ soit aussi petite qu'on le souhaite.

La carte de variance d'estimation retournée par la méthode de krigeage n'a de sens que dans l'hypothèse où le variogramme choisi modélise de manière satisfaisante le phénomène étudié.

# Validation croisée

En utilisant la propriété d'invariance de l'estimation par multiplication du variogramme par un scalaire, on peut toujours trouver un variogramme $\gamma$, tel que la variance sur l'estimation $\hat{z}$ en tout point $x$ soit aussi petite qu'on le souhaite.

La carte de variance d'estimation retournée par la méthode de krigeage n'a de sens que dans l'hypothèse où le variogramme choisi modélise de manière satisfaisante le phénomène étudié.

$\rightarrow$ Nécessité d'introduire une méthode objective de qualification des résultats : la validation croisée

# Validation croisée

La **validation croisée** permet de donner une estimation de la précision attendue sur l'interpolation.

1. On se donne un ratio $r$ (en %)

# Validation croisée

La **validation croisée** permet de donner une estimation de la précision attendue sur l'interpolation.

1. On se donne un ratio $r$ (en %)
2. L'interpolation est réalisée sur une fraction $r$ des observations, choisies aléatoirement et uniformément dans le champ

# Validation croisée

La **validation croisée** permet de donner une estimation de la précision attendue sur l'interpolation.

1. On se donne un ratio $r$ (en %)
2. L'interpolation est réalisée sur une fraction $r$ des observations, choisies aléatoirement et uniformément dans le champ
3. Les valeurs estimées au niveau des $1 - r$ sites restants sont comparées aux valeurs observées en ces mêmes sites $\rightarrow$ calcul de la variance de l'erreur.

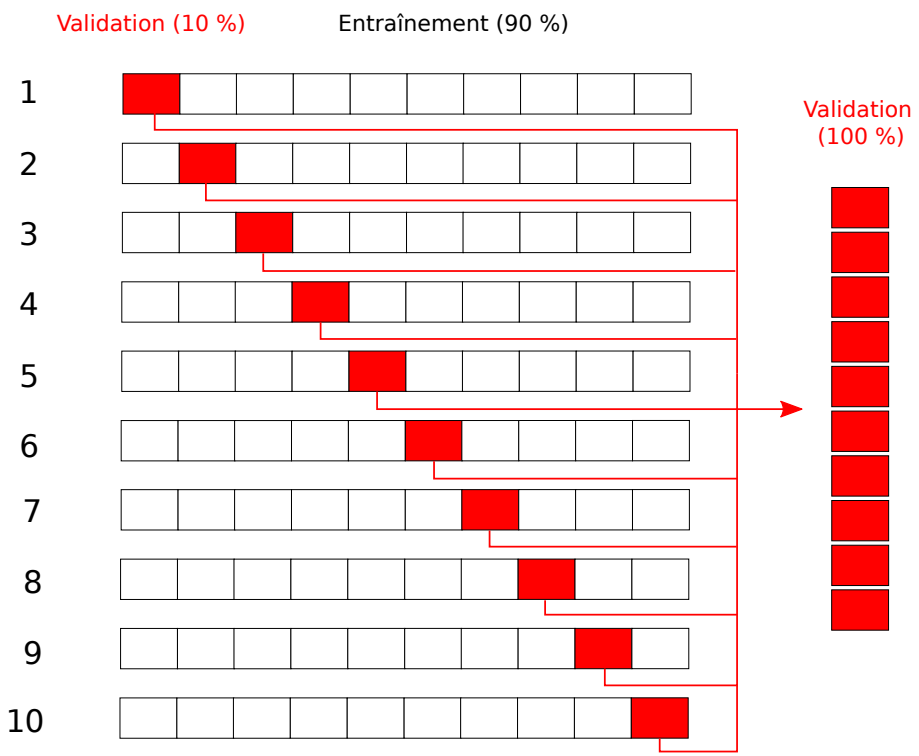
# Validation croisée

La **validation croisée** permet de donner une estimation de la précision attendue sur l'interpolation.

1. On se donne un ratio $r$ (en %)
2. L'interpolation est réalisée sur une fraction $r$ des observations, choisies aléatoirement et uniformément dans le champ
3. Les valeurs estimées au niveau des $1 - r$ sites restants sont comparées aux valeurs observées en ces mêmes sites $\rightarrow$ calcul de la variance de l'erreur.
4. On réitère les étapes 1 à 3 plusieurs fois de sorte à ce que les valeurs en chaque site soient comparées aux valeurs interpolées et les variances obtenues sont moyennées.

# Validation croisée

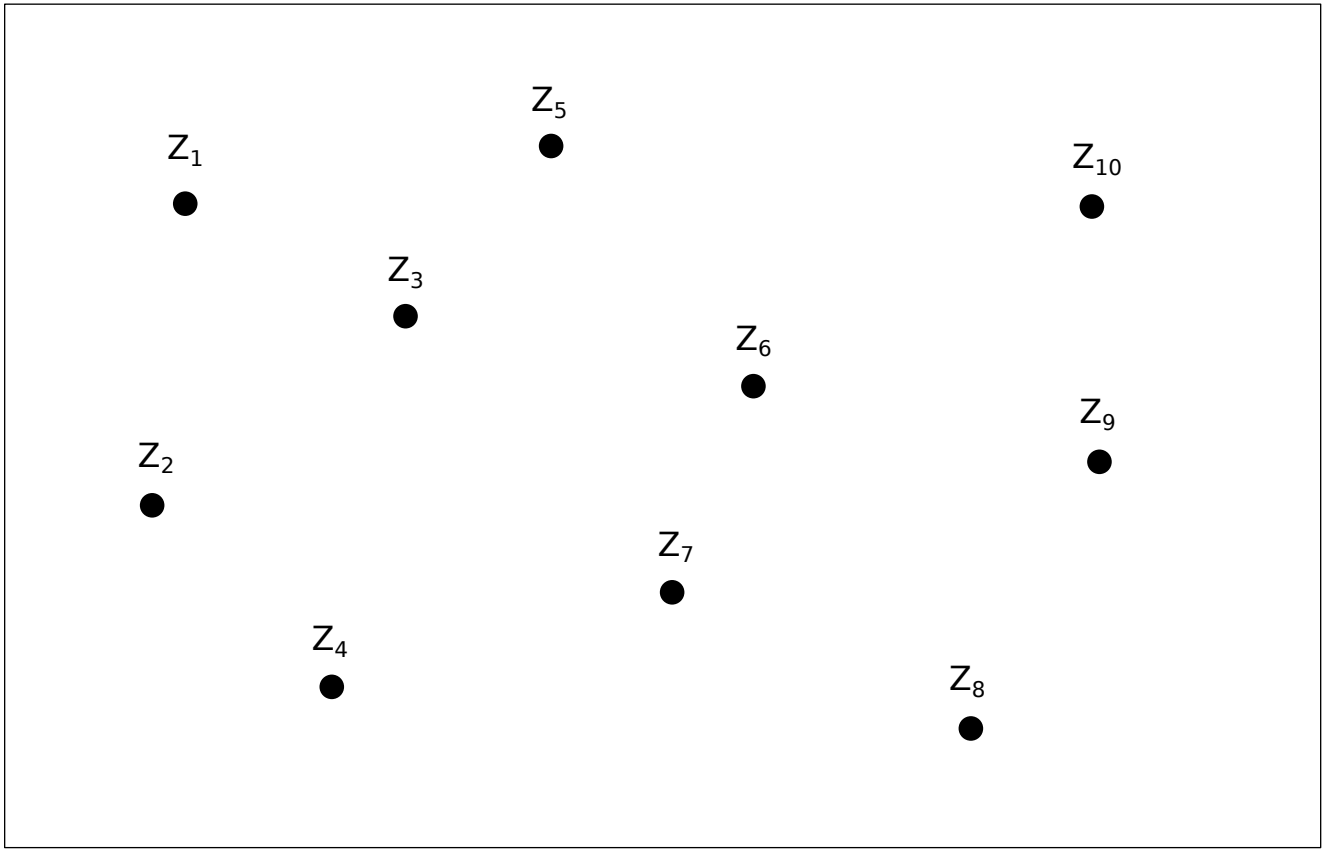
Exemple avec $r = 90\%$ :




Le ratio $r$ doit être suffisamment grand pour ne pas trop biaiser l'estimation de l'erreur finale.

# Validation croisée

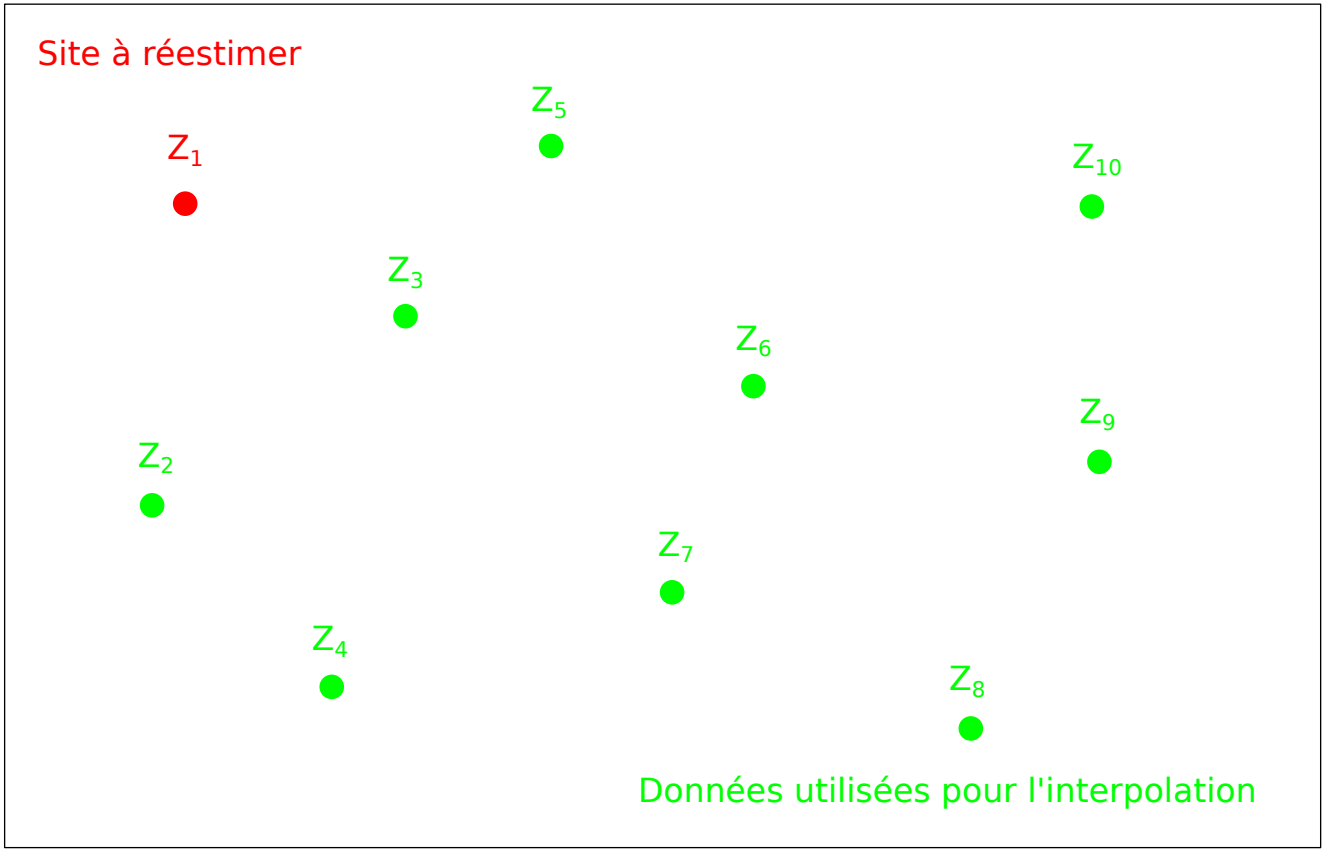
Cas extrême : $r = \frac{N-1}{N}$ (méthode *leave-one-out*)




Semis de $N$ sites observés

# Validation croisée

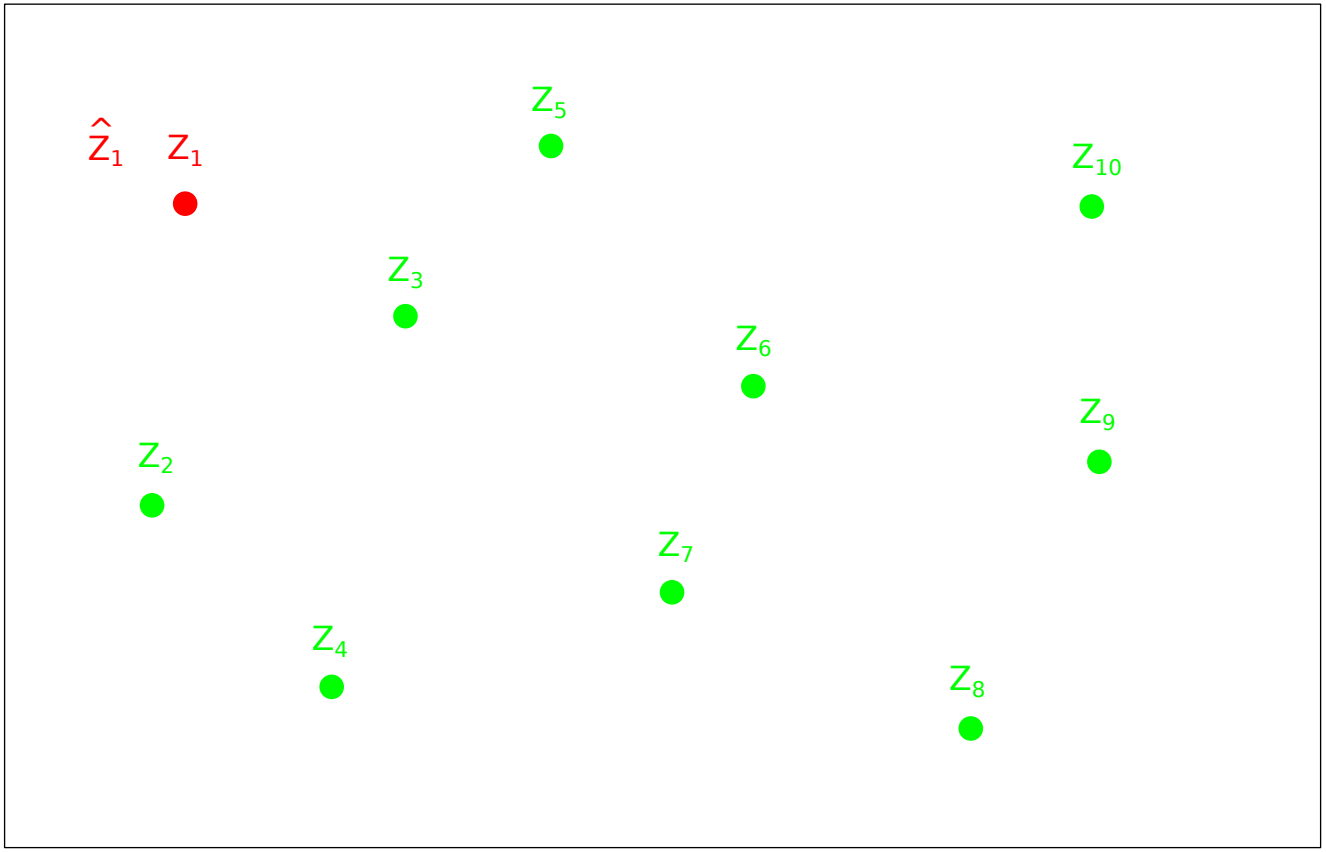
Cas extrême : $r = \frac{N-1}{N}$ (méthode *leave-one-out*)




Exclusion de l'observation $z_1$

# Validation croisée

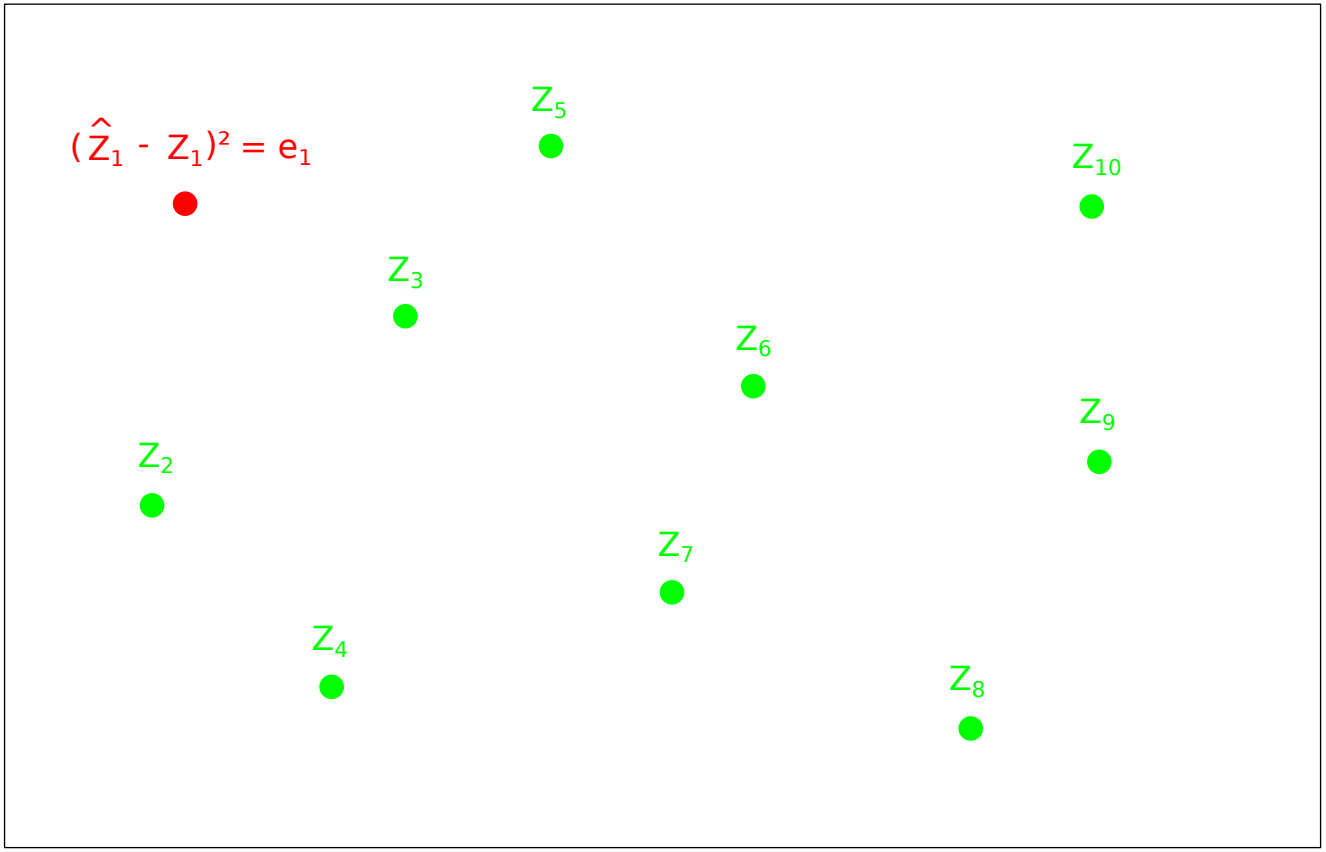
Cas extrême : $r = \frac{N-1}{N}$ (méthode *leave-one-out*)




Réestimation de $z_1$ à partir des autres valeurs observées : $\hat{z}_1$

# Validation croisée

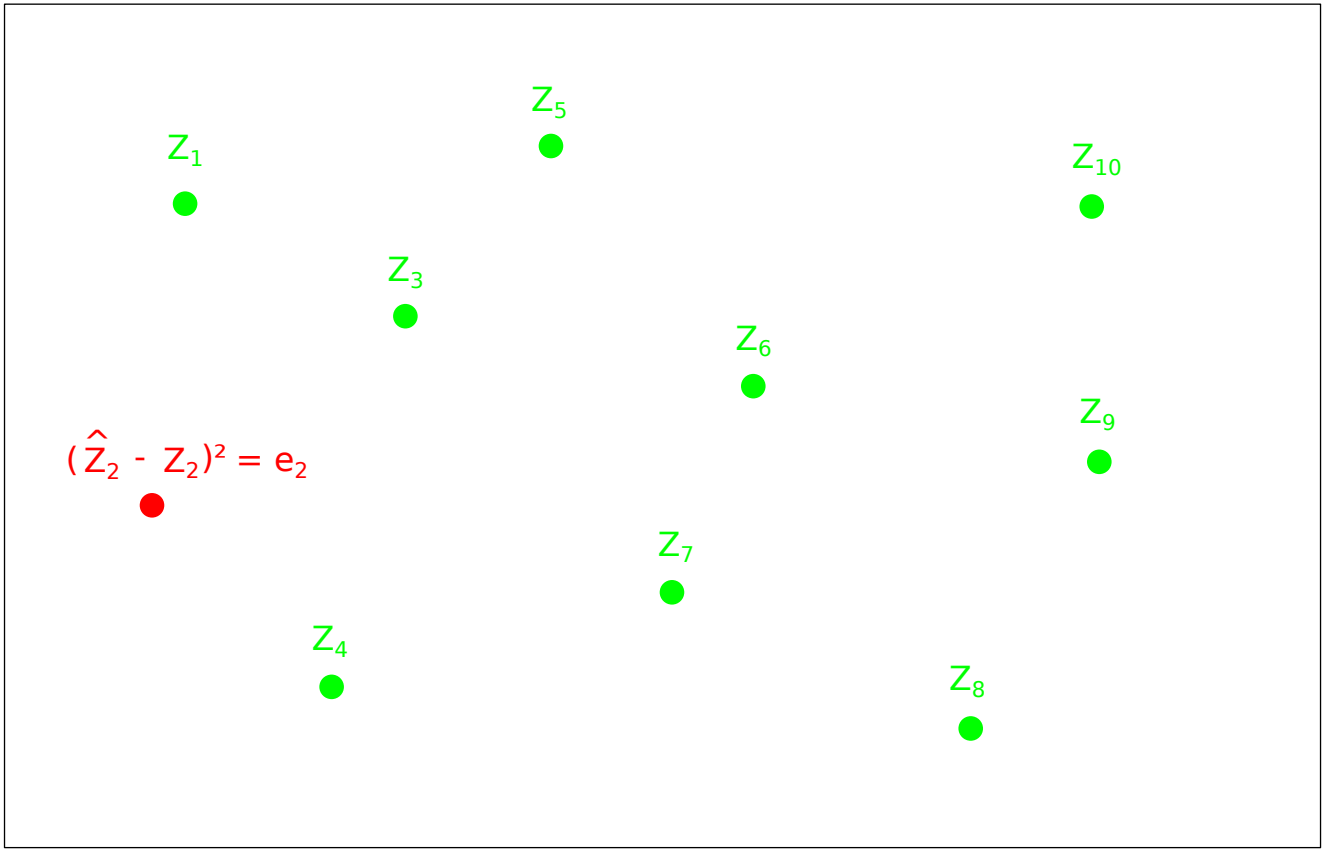
Cas extrême : $r = \frac{N-1}{N}$ (méthode *leave-one-out*)




Calcul de l'erreur commise au carré : $e_1$

# Validation croisée

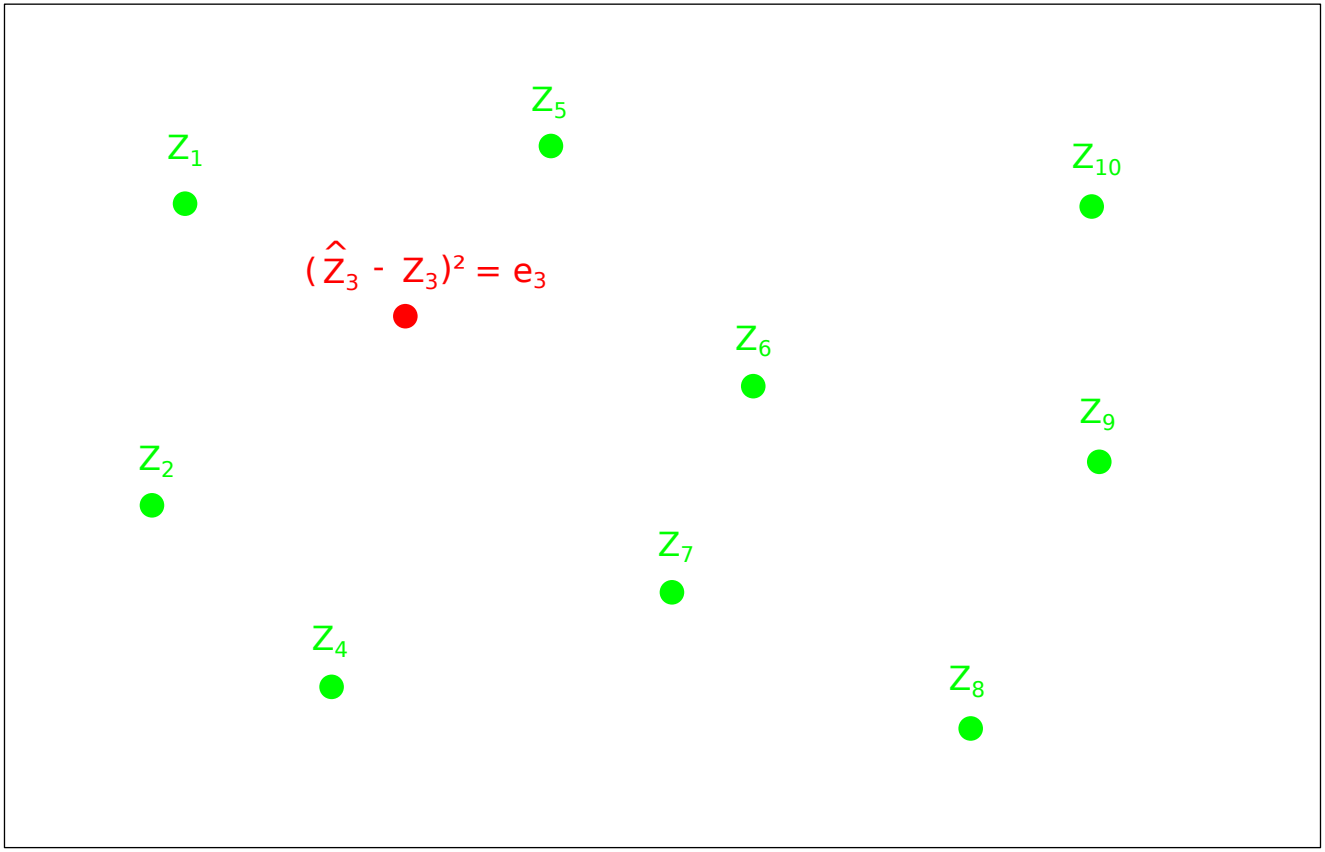
Cas extrême : $r = \frac{N-1}{N}$ (méthode *leave-one-out*)




Réitération sur $z_2$ : $e_2$

# Validation croisée

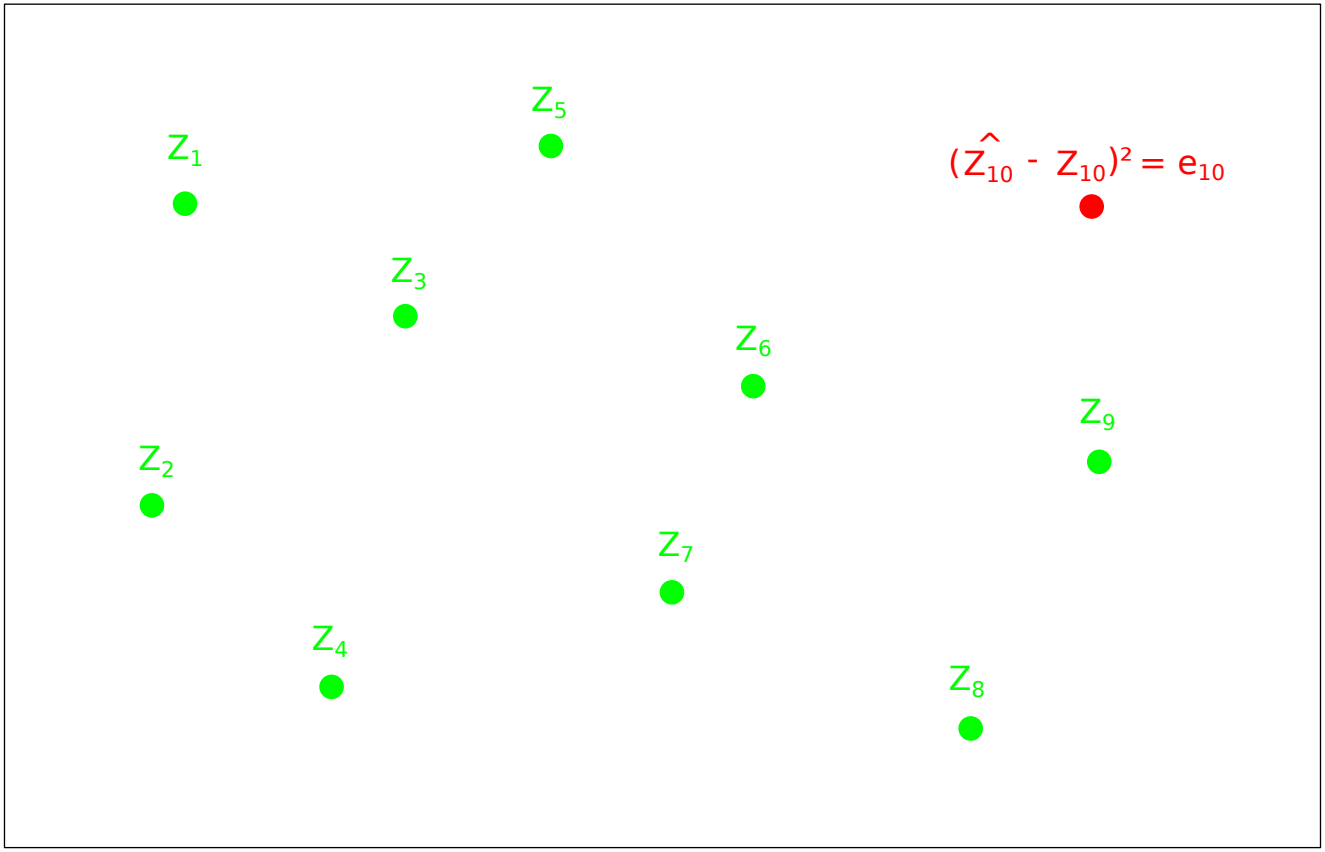
Cas extrême : $r = \frac{N-1}{N}$ (méthode *leave-one-out*)




Calcul de $e_3$...

# Validation croisée

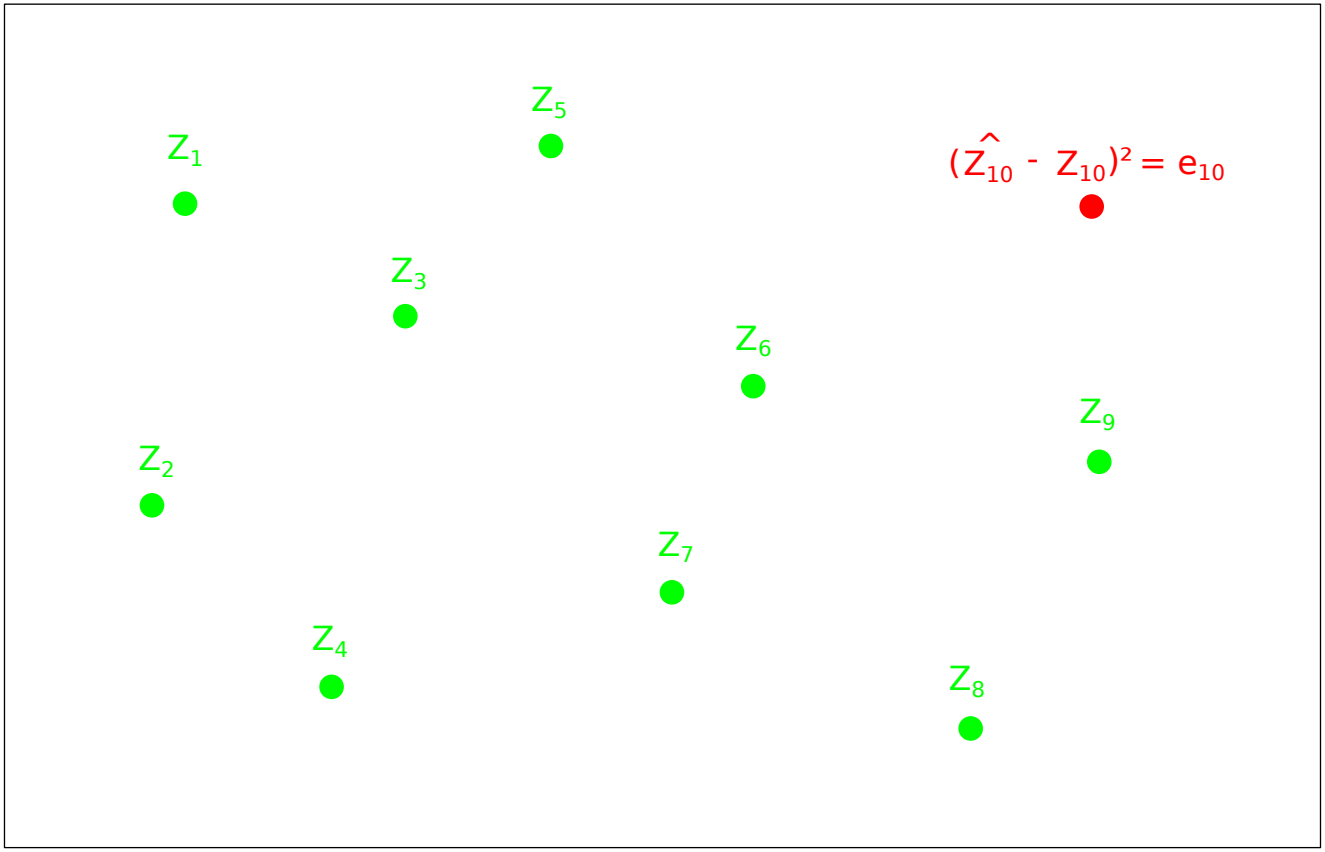
Cas extrême : $r = \frac{N-1}{N}$ (méthode *leave-one-out*)




Calcul de $e_{10}$

# Validation croisée

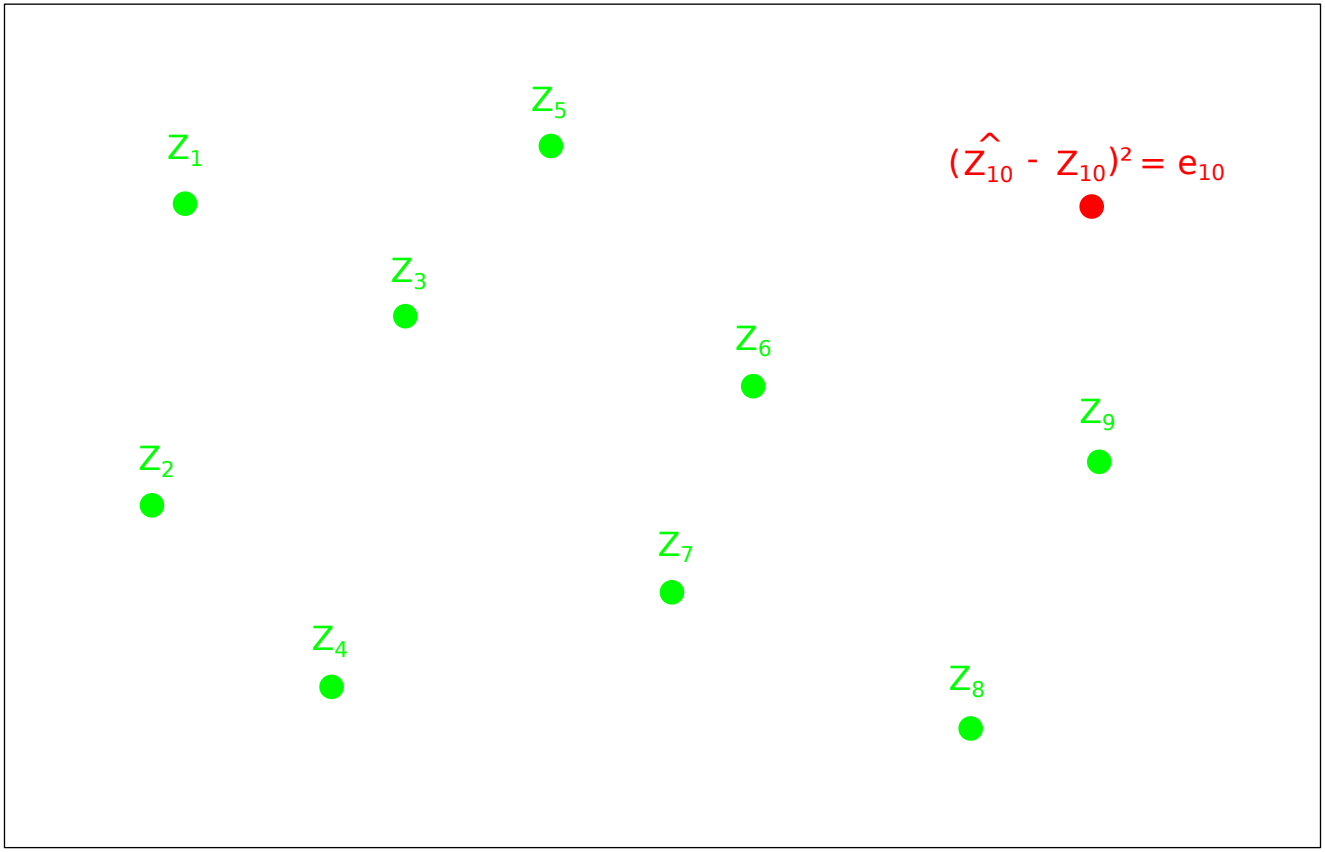
Cas extrême : $r = \frac{N-1}{N}$ (méthode *leave-one-out*)




Estimation de la variance d'interpolation : $\mathrm{Var}[\hat{z}] = \frac{1}{10}\sum_{i=1}^{10} e_i$

# Validation croisée

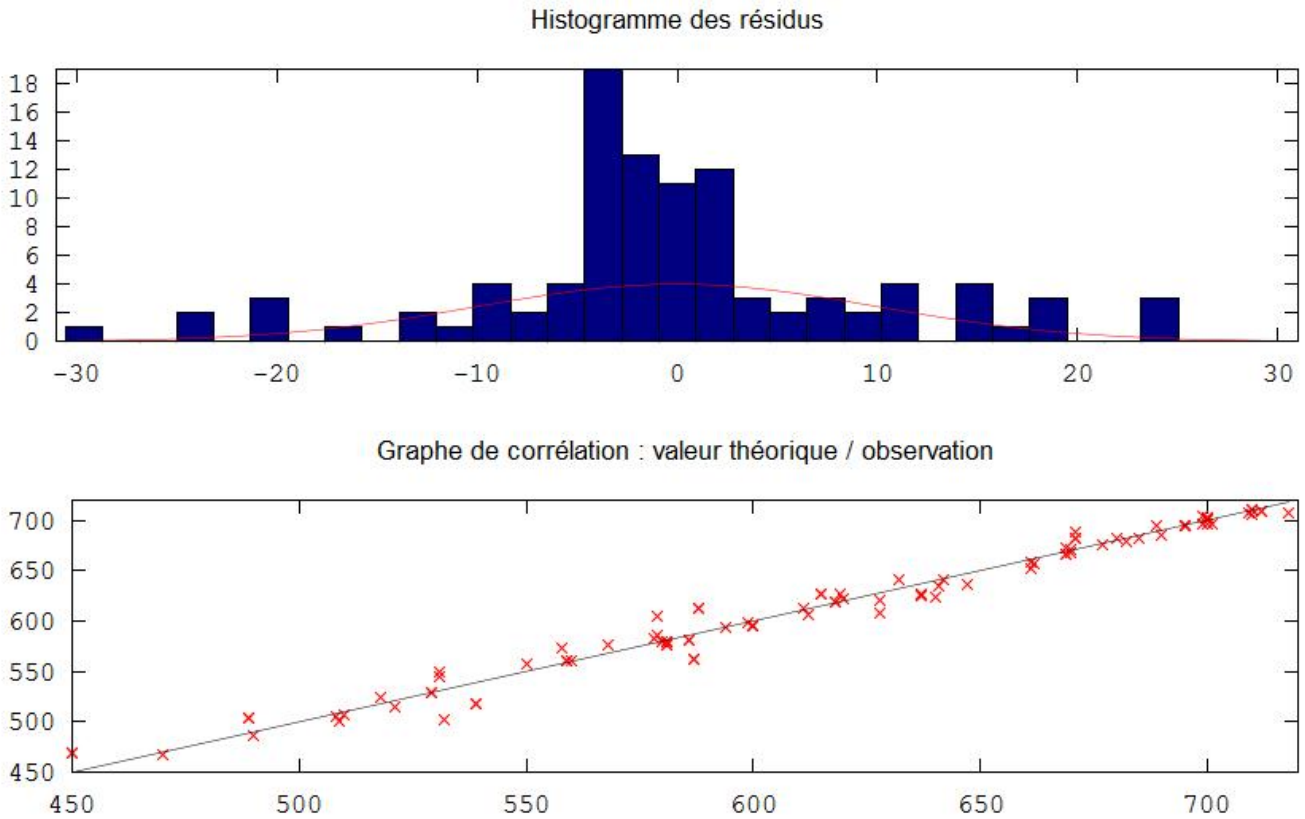
Cas extrême : $r = \frac{N-1}{N}$ (méthode *leave-one-out*)




Estimation de l'écart-type d'interpolation : $\sigma_{\hat{z}} = \sqrt{\frac{1}{10}\sum_{i=1}^{10} e_i}$

# Validation croisée

Exemple de validation croisée




Validation *leave-one-out* réalisée sur un estimateur par krigeage à variographie linéaire.

# Validation croisée : bilan

⊕ Méthode générique permettant d'obtenir une précision de l'interpolation, quelque soit le type de méthode employée (particulièrement utile pour évaluer la performance des méthodes non-stochastiques).

# Validation croisée : bilan

⊕ Méthode générique permettant d'obtenir une précision de l'interpolation, quelque soit le type de méthode employée (particulièrement utile pour évaluer la performance des méthodes non-stochastiques).

⊕ Réglage des paramètres de l'algorithme d'interpolation

# Validation croisée : bilan

⊕ Méthode générique permettant d'obtenir une précision de l'interpolation, quelque soit le type de méthode employée (particulièrement utile pour évaluer la performance des méthodes non-stochastiques).

⊕ Réglage des paramètres de l'algorithme d'interpolation

⊕ Comparaison des performances de plusieurs algorithmes

# Validation croisée : bilan

⊕ Méthode générique permettant d'obtenir une précision de l'interpolation, quelque soit le type de méthode employée (particulièrement utile pour évaluer la performance des méthodes non-stochastiques).

⊕ Réglage des paramètres de l'algorithme d'interpolation

⊕ Comparaison des performances de plusieurs algorithmes

⊕ Possibilité de représenter la distribution des erreurs

# Validation croisée : bilan

⊕ Méthode générique permettant d'obtenir une précision de l'interpolation, quelque soit le type de méthode employée (particulièrement utile pour évaluer la performance des méthodes non-stochastiques).

⊕ Réglage des paramètres de l'algorithme d'interpolation

⊕ Comparaison des performances de plusieurs algorithmes

⊕ Possibilité de représenter la distribution des erreurs

⊕ Possibilité de faire une carte des erreurs (interpolation spatiale des résidus)

# Validation croisée : bilan

⊕ Méthode générique permettant d'obtenir une précision de l'interpolation, quelque soit le type de méthode employée (particulièrement utile pour évaluer la performance des méthodes non-stochastiques).

⊕ Réglage des paramètres de l'algorithme d'interpolation

⊕ Comparaison des performances de plusieurs algorithmes

⊕ Possibilité de représenter la distribution des erreurs

⊕ Possibilité de faire une carte des erreurs (interpolation spatiale des résidus)

⊖ Estimation biaisée : sur-estimation de la variance d'erreur (l'exclusion de chaque point sous-échantillonne le semis d'observations à l'endroit précis où on souhaite faire l'estimation)

# Application à l'optimisation bayésienne

**Contexte** : on souhaite déterminer le maximum (ou le minimum) sur $\mathcal{D}$ d'une variable régionalisée dont on ne souhaite observer qu'un minimum de points.

# Application à l'optimisation bayésienne

**Contexte** : on souhaite déterminer le maximum (ou le minimum) sur $\mathcal{D}$ d'une variable régionalisée dont on ne souhaite observer qu'un minimum de points.

L'idée consiste donc à chaque étape, à choisir le point à observer en fonction des observations précédentes, de sorte à échantillonner de préférence les zones *prometteuses*, i.e. les zones où l'on a *le plus de chances* de trouver l'extremum de $z$.

# Application à l'optimisation bayésienne

**Contexte** : on souhaite déterminer le maximum (ou le minimum) sur $\mathcal{D}$ d'une variable régionalisée dont on ne souhaite observer qu'un minimum de points.

L'idée consiste donc à chaque étape, à choisir le point à observer en fonction des observations précédentes, de sorte à échantillonner de préférence les zones *prometteuses*, i.e. les zones où l'on a *le plus de chances* de trouver l'extremum de $z$.

Exemple d'application : on dispose d'un algorithme, dont la mise en oeuvre requiert de fixer un certain nombre $p$ de paramètres continus (flottants ou entiers avec suffisamment de modalités). L'execution de l'algorithme retourne un indice de performance $z$ que l'on souhaite optimiser avec un bon choix de valeurs pour les paramètres.

# Application à l'optimisation bayésienne

Lorsque $p$ est relativement grand (typiquement $p \geqslant 3$ ou $4$), la combinatoire du problème interdit d'échantillonner uniformément la zone, de sorte qu'il est difficile de connaître la fonction réponse $z$, même sur un maillage ($p$-dimensionnel) peu résolu. Cette affirmation est d'autant plus vraie que l'algorithme est coûteux en temps de calcul. L'optimisation bayésienne va nous permettre de trouver, avec un minimum d'appels à l'algorithme, un bon set de paramètres.

# Application à l'optimisation bayésienne

Lorsque $p$ est relativement grand (typiquement $p \geqslant 3$ ou $4$), la combinatoire du problème interdit d'échantillonner uniformément la zone, de sorte qu'il est difficile de connaître la fonction réponse $z$, même sur un maillage ($p$-dimensionnel) peu résolu. Cette affirmation est d'autant plus vraie que l'algorithme est coûteux en temps de calcul. L'optimisation bayésienne va nous permettre de trouver, avec un minimum d'appels à l'algorithme, un bon set de paramètres.

L'optimisation bayésienne repose sur la mise-à-jour des connaissances sur $z$, à l'aide d'un nombre limité d'observations (chaque nouvelle observation impose une mise-à-jour des connaissances sur $z$). En ce sens, cette technique d'optimisation est qualifiée de bayésienne.

# Application à l'optimisation bayésienne

Lorsque $p$ est relativement grand (typiquement $p \geqslant 3$ ou $4$), la combinatoire du problème interdit d'échantillonner uniformément la zone, de sorte qu'il est difficile de connaître la fonction réponse $z$, même sur un maillage ($p$-dimensionnel) peu résolu. Cette affirmation est d'autant plus vraie que l'algorithme est coûteux en temps de calcul. L'optimisation bayésienne va nous permettre de trouver, avec un minimum d'appels à l'algorithme, un bon set de paramètres.

L'optimisation bayésienne repose sur la mise-à-jour des connaissances sur $z$, à l'aide d'un nombre limité d'observations (chaque nouvelle observation impose une mise-à-jour des connaissances sur $z$). En ce sens, cette technique d'optimisation est qualifiée de bayésienne.

Une bonne connaissance de $z$ requiert de pouvoir calculer à chaque itération, une estimation $\hat{z}$ de $z$ avec sa variance associée $\mathrm{Var}[\hat{z}]$ toutes deux déduites des $n$ premières observations.

# Application à l'optimisation bayésienne

A chaque étape, on doit donc être capable de :

1. Utiliser l'estimation courante $\hat{z}$ et sa variance associée $\mathrm{Var}[\hat{z}]$ pour déterminer un nouveau point d'observation $x_{n+1} \in \mathcal{D}$. Ce nouveau point doit être *prometteur* dans la recherche du maximum.

# Application à l'optimisation bayésienne

A chaque étape, on doit donc être capable de :

1. Utiliser l'estimation courante $\hat{z}$ et sa variance associée $\mathrm{Var}[\hat{z}]$ pour déterminer un nouveau point d'observation $x_{n+1} \in \mathcal{D}$. Ce nouveau point doit être *prometteur* dans la recherche du maximum.

2. Calculer la valeur de $z$ en ce nouveau site : $z(x_{n+1}) = z_{n+1}$

# Application à l'optimisation bayésienne

A chaque étape, on doit donc être capable de :

1. Utiliser l'estimation courante $\hat{z}$ et sa variance associée $\mathrm{Var}[\hat{z}]$ pour déterminer un nouveau point d'observation $x_{n+1} \in \mathcal{D}$. Ce nouveau point doit être *prometteur* dans la recherche du maximum.

2. Calculer la valeur de $z$ en ce nouveau site : $z(x_{n+1}) = z_{n+1}$

3. Mettre à jour l'estimation sur $z$ et sa variance avec la connaissance de cette nouvelle observation $\rightarrow \hat{z}$ et $\mathrm{Var}[\hat{z}]$

# Application à l'optimisation bayésienne

A chaque étape, on doit donc être capable de :

1. Utiliser l'estimation courante $\hat{z}$ et sa variance associée Var$[\hat{z}]$ pour déterminer un nouveau point d'observation $x_{n+1} \in \mathcal{D}$. Ce nouveau point doit être *prometteur* dans la recherche du maximum.

2. Calculer la valeur de $z$ en ce nouveau site : $z(x_{n+1}) = z_{n+1}$

3. Mettre à jour l'estimation sur $z$ et sa variance avec la connaissance de cette nouvelle observation $\rightarrow \hat{z}$ et Var$[\hat{z}]$

4. Itérer les étapes 1 à 3 jusqu'à obtenir une valeur satisfaisante (en fonction du problème posé)

# Application à l'optimisation bayésienne

Il s'agit d'un algorithme de recherche, donc il faut obtenir un compromis entre :

1. **Intensification** : on cherche davantage dans les zones prometteuses (c'est-à-dire celles au voisinage desquelles on a trouvé des valeurs relativement hautes ou basses)
2. **Exploration** : on cherche à parcourir le champ de manière suffisamment exhaustive pour s'assurer qu'il n'existe pas d'autres zones prometteuses

L'optimisation bayésienne fournit un critère rigoureux de compromis entre ces deux tendances.

# Application à l'optimisation bayésienne

A chaque étape, on cherche à observer le point le plus prometteur, i.e. un point qui soit dans une zone à la fois maximale et mal connue

# Application à l'optimisation bayésienne

A chaque étape, on cherche à observer le point le plus prometteur, i.e. un point qui soit dans une zone à la fois maximale et mal connue

## Critère de sélection du site d'échantillonnage

**Probability of Improvement** :

$$x_{n+1} = \underset{x \in \mathcal{D}}{\text{argmax}} \;\; \Phi(g_n(x)) \quad \text{où} \quad g_n(x) = \frac{\hat{z}_(x_{best}) - \hat{z}(x)}{\sigma[\hat{z}(x)]}$$

avec $\Phi : \mathbb{R} \to [0,1]$ la fonction de répartition de la loi normale.

Dans l'expression de $g_n$, les estimations de $\mathcal{D}$ de $\hat{z}$ et $\sigma[\hat{z}(x)]$ sont faites par krigeage sur les $n$ premières observations.

# Application à l'optimisation bayésienne

**Probability of improvement :**

$$\eta(x \mid x_1, x_2, ...x_n) = \mathbb{P}[\hat{z}(x) \geqslant z^* \mid x_1, x_2, ...x_n]$$

# Application à l'optimisation bayésienne

**Probability of improvement :**

$$\eta(x \mid x_1, x_2, ...x_n) = \mathbb{P}[\hat{z}(x) \geqslant z^* \mid x_1, x_2, ...x_n]$$

**Expected improvement :**

$$\eta(x \mid x_1, x_2, ...x_n) = \mathbb{E}_{z|x_1,x_2,...x_n}[\hat{z}(x) - z^*]$$

# Application à l'optimisation bayésienne

**Probability of improvement :**

$$\eta(x \mid x_1, x_2, ...x_n) = \mathbb{P}[\hat{z}(x) \geqslant z^* \mid x_1, x_2, ...x_n]$$

**Expected improvement :**

$$\eta(x \mid x_1, x_2, ...x_n) = \mathbb{E}_{z|x_1,x_2,...x_n}[\hat{z}(x) - z^*]$$

**Upper confidence bound** $\alpha \in ]0, 1[$ **:**

$$\eta_\alpha(x \mid x_1, x_2, ...x_n) = M \quad \text{t.q.} \quad \mathbb{P}[M \geqslant \hat{z}(x) \mid x_1, x_2, ...x_n] = \alpha$$

# Application à l'optimisation bayésienne

**Probability of improvement :**

$$\eta(x \mid x_1, x_2, ...x_n) = \mathbb{P}[\hat{z}(x) \geqslant z^* \mid x_1, x_2, ...x_n]$$

**Expected improvement :**

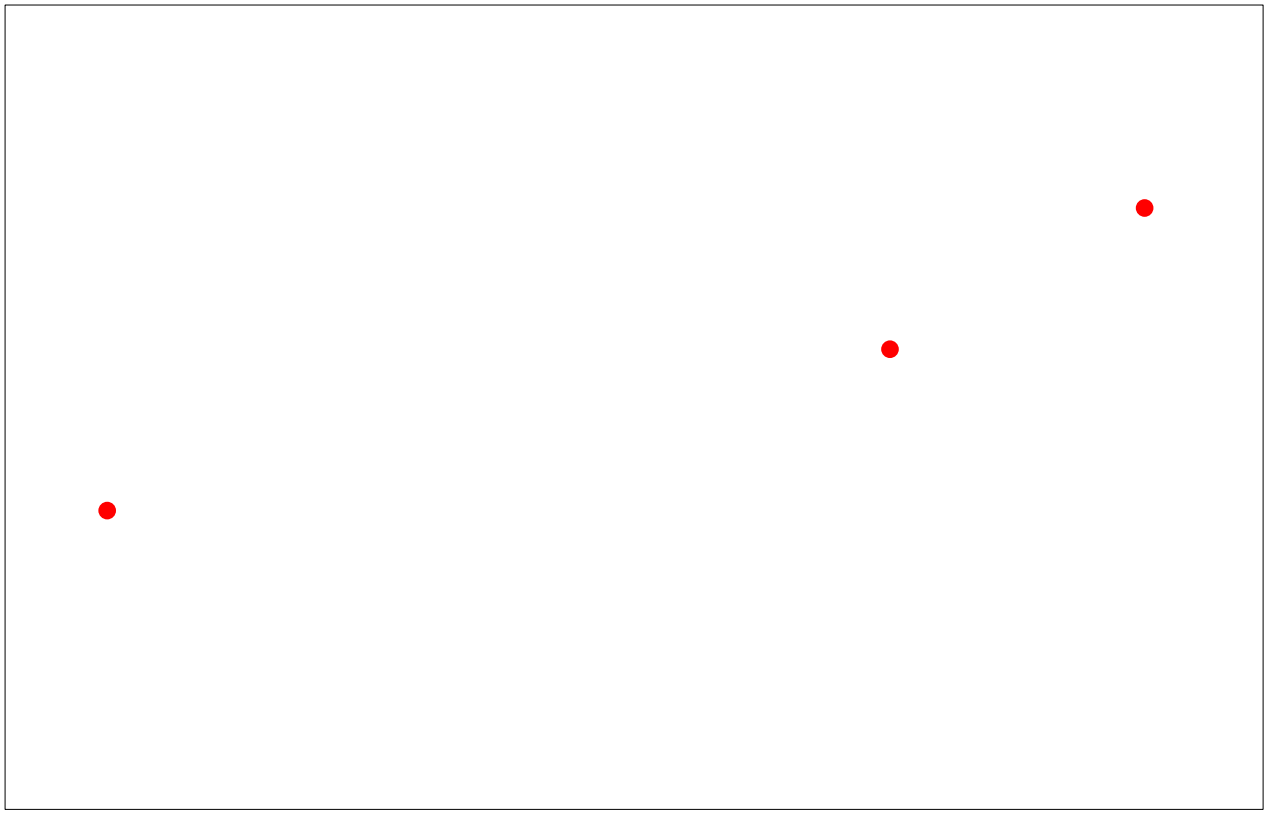
$$\eta(x \mid x_1, x_2, ...x_n) = \mathbb{E}_{z|x_1,x_2,...x_n}[\hat{z}(x) - z^*]$$

**Upper confidence bound** $\alpha \in ]0, 1[$ **:**

$$\eta_\alpha(x \mid x_1, x_2, ...x_n) = M \quad \text{t.q.} \quad \mathbb{P}[M \geqslant \hat{z}(x) \mid x_1, x_2, ...x_n] = \alpha$$

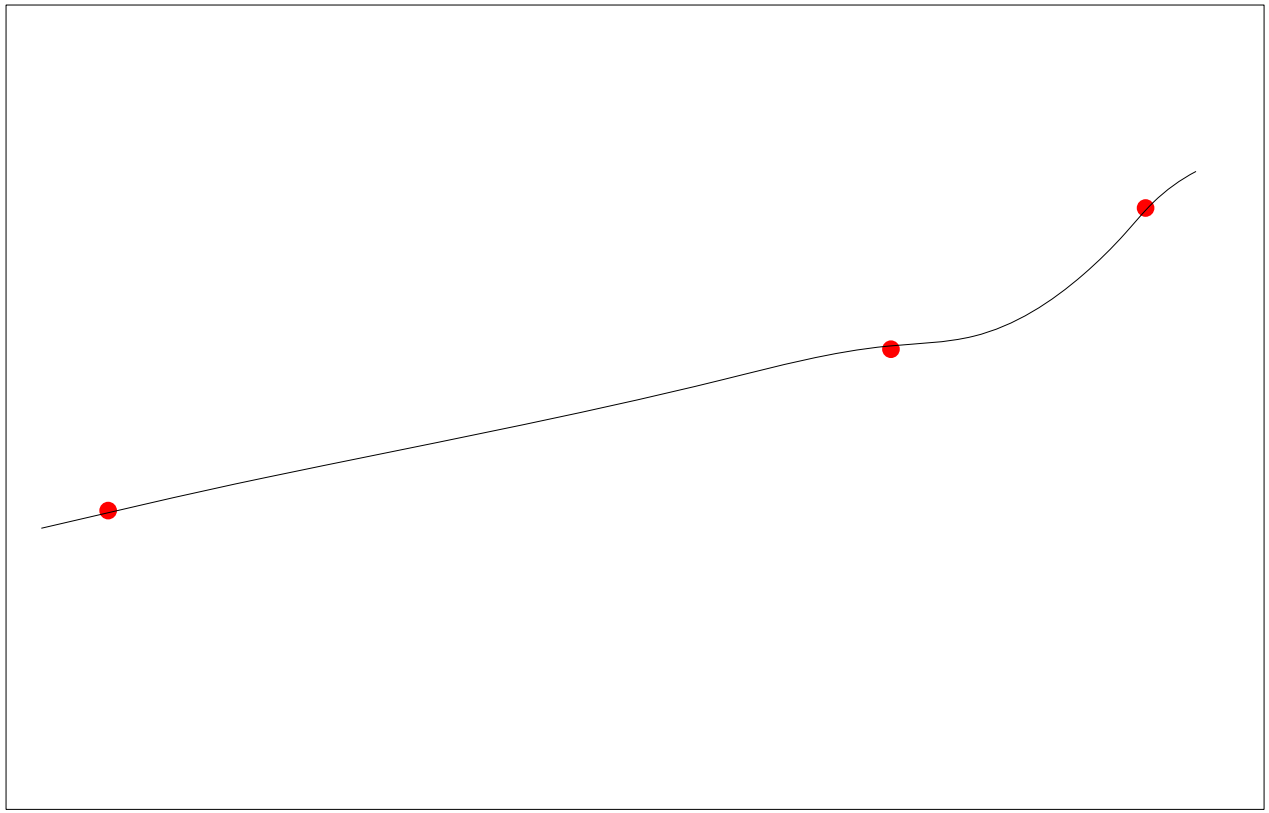
**Itération :** $$x_{n+1} = \underset{x \in \mathcal{D}}{\text{argmax}} \ \eta(x \mid x_1, x_2, ...x_n)$$

# Application à l'optimisation bayésienne



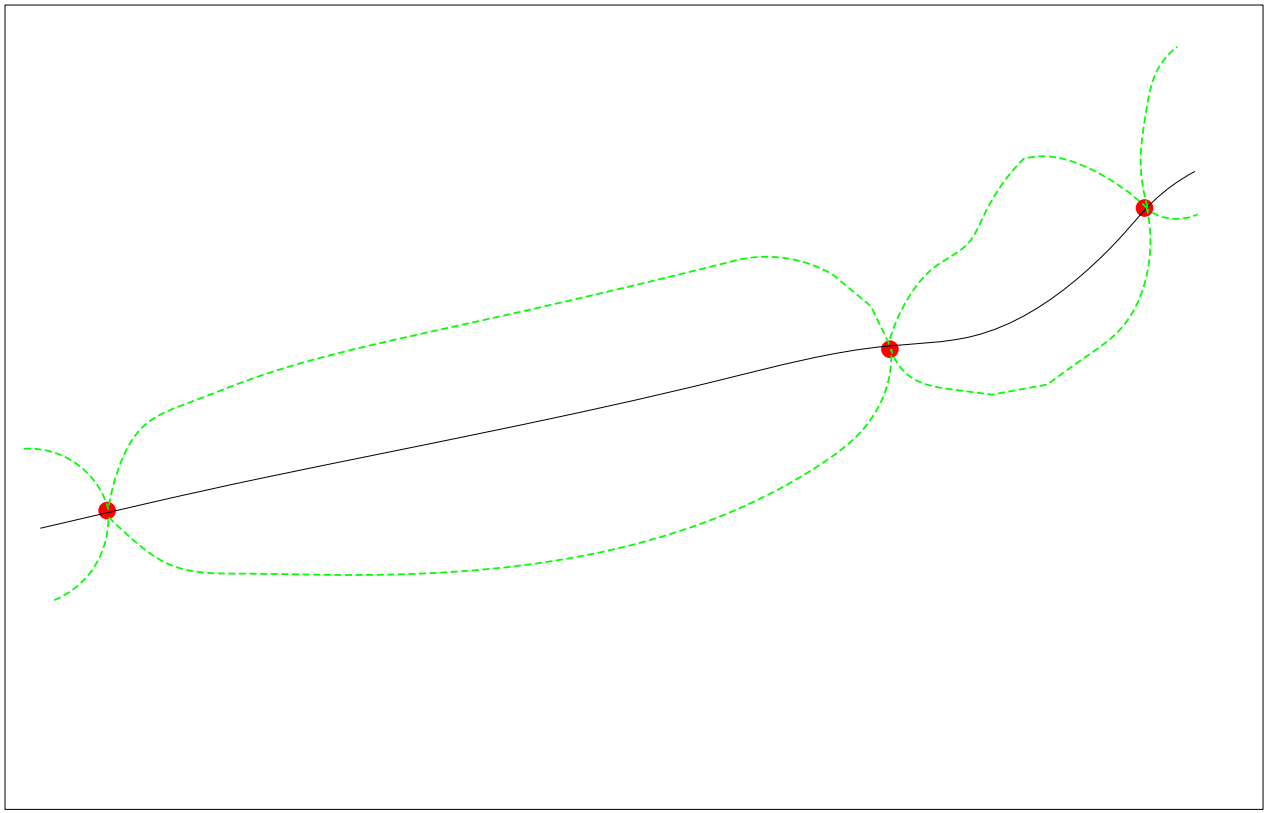
Semis de points observée à une étape donnée : $z(x_1)$, $z(x_2)$ et $z(x_3)$.

# Application à l'optimisation bayésienne



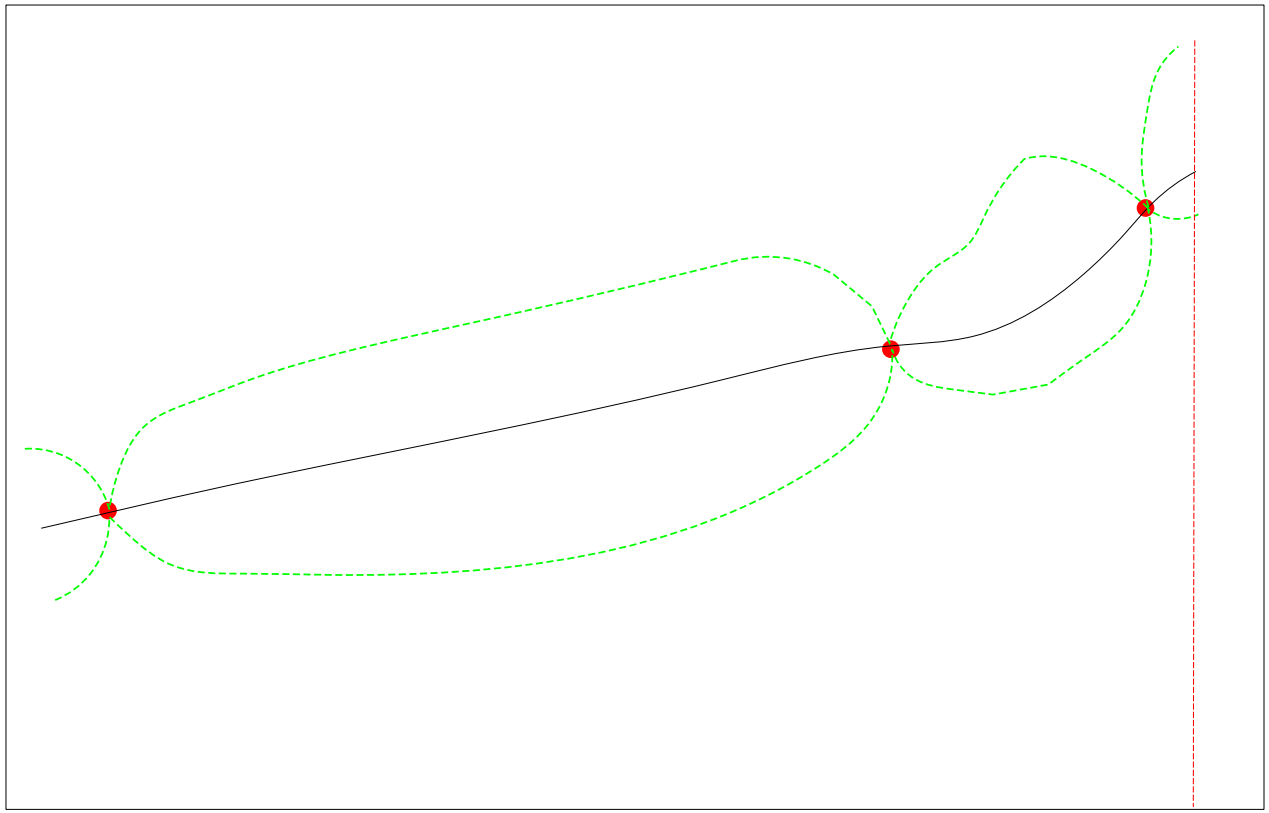
Estimation de $\hat{z}$ par krigeage

# Application à l'optimisation bayésienne



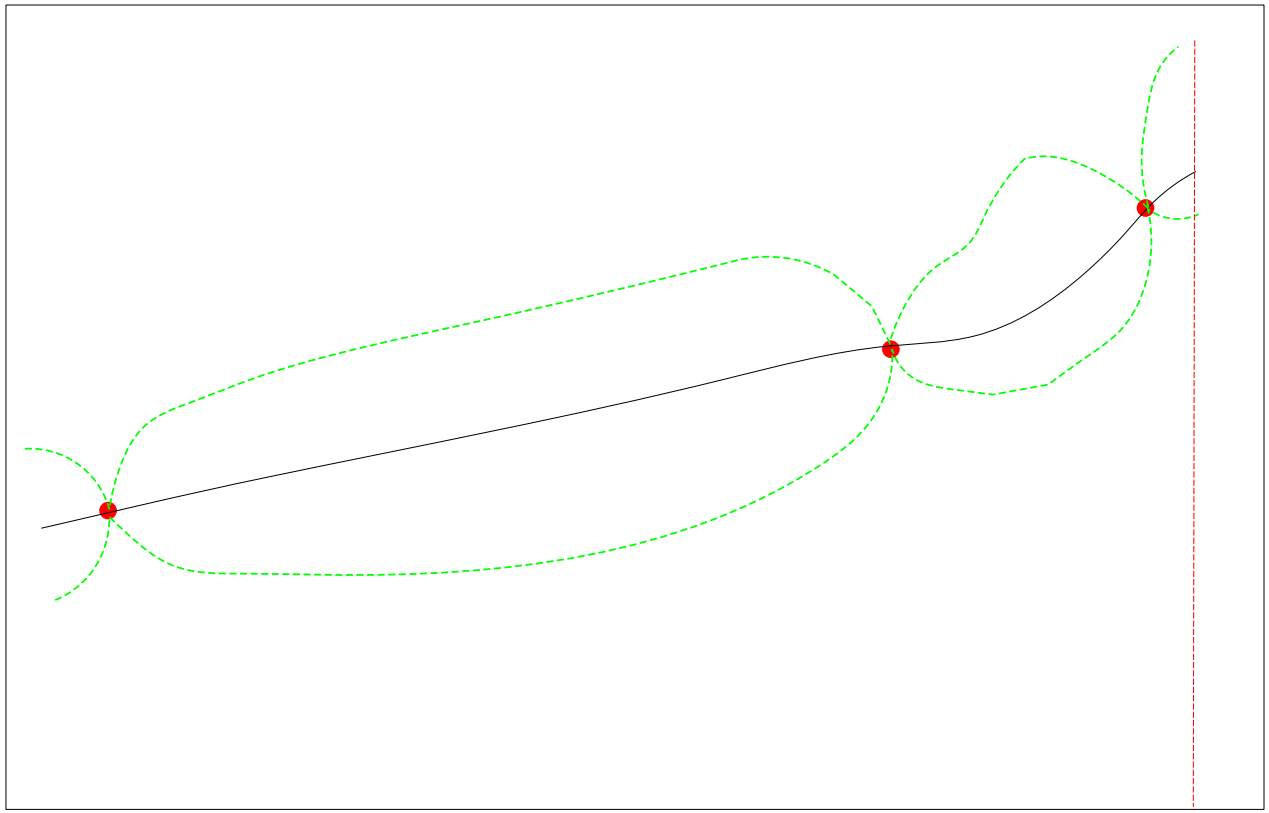
Estimation de la variance associée $\mathrm{Var}[\hat{z}]$

# Application à l'optimisation bayésienne



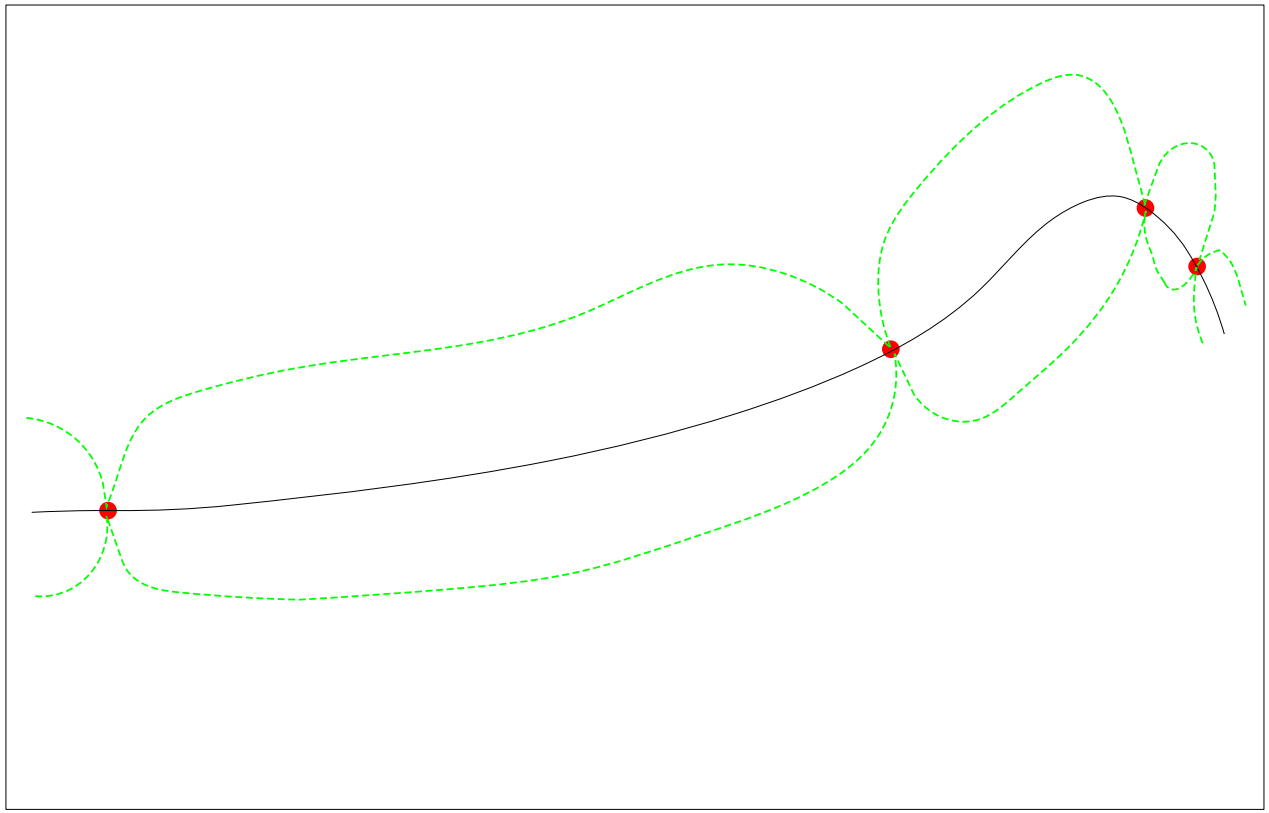
Sélection du site le plus prometteur $x_4$

# Application à l'optimisation bayésienne



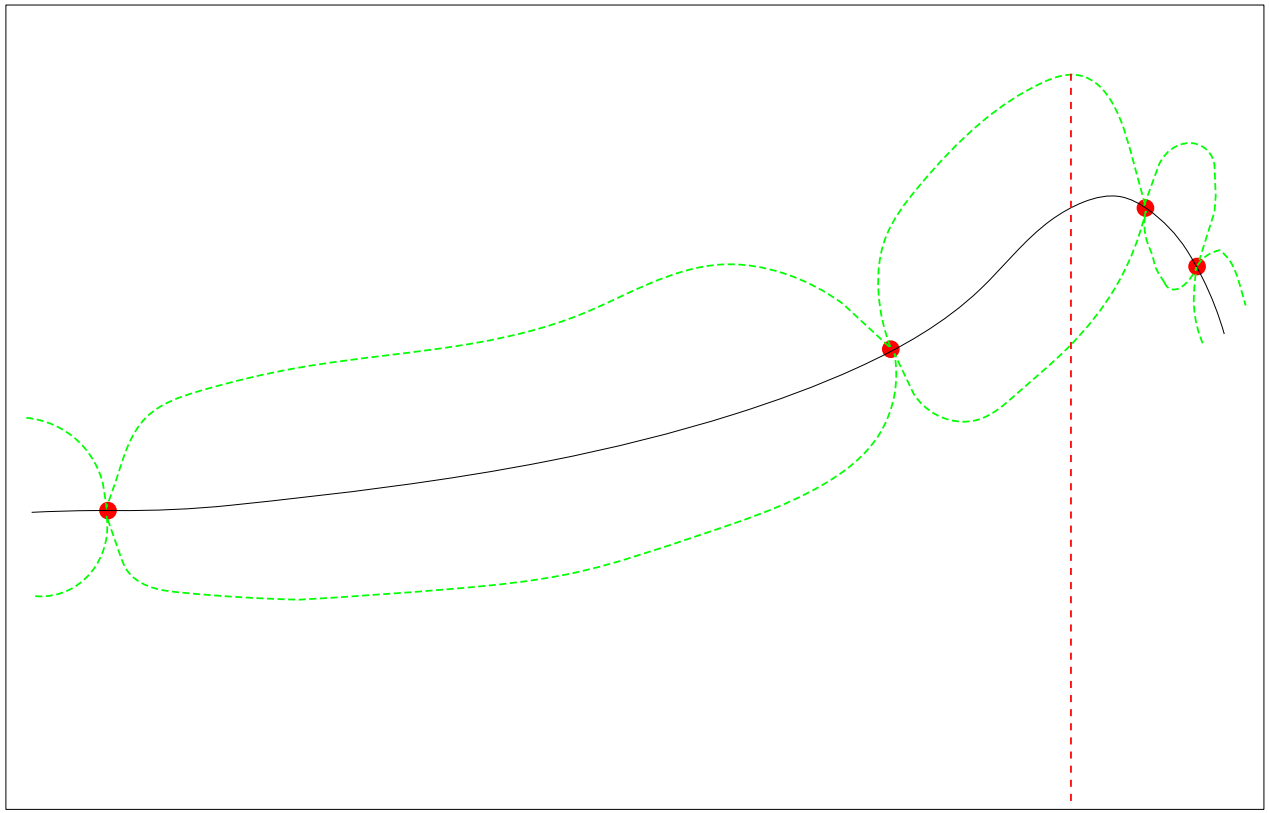
Observation $z_4 = z(x_4)$

# Application à l'optimisation bayésienne



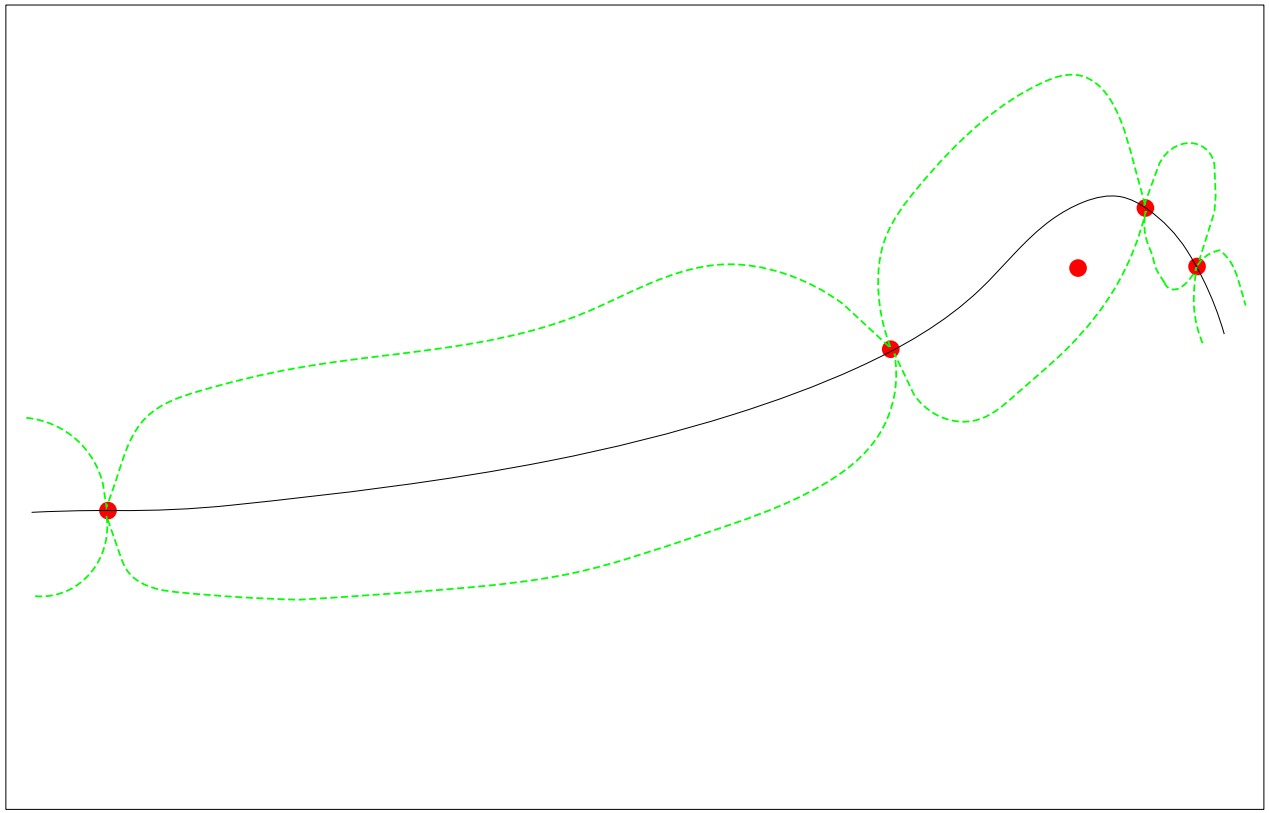
Observation $z_4 = z(x_4)$ et mises à jour de $\hat{z}$ et Var$[\hat{z}]$

# Application à l'optimisation bayésienne



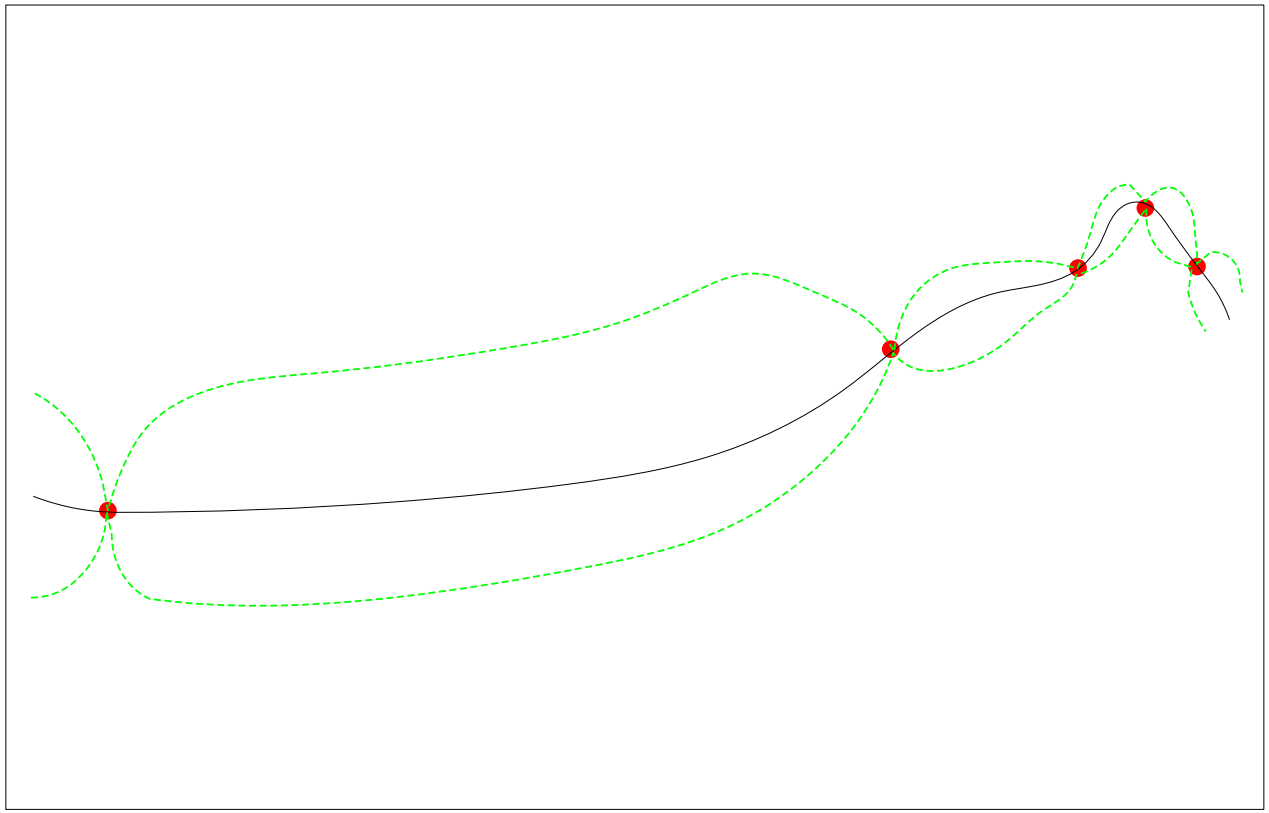
Sélection du site le plus prometteur $x_5$

# Application à l'optimisation bayésienne



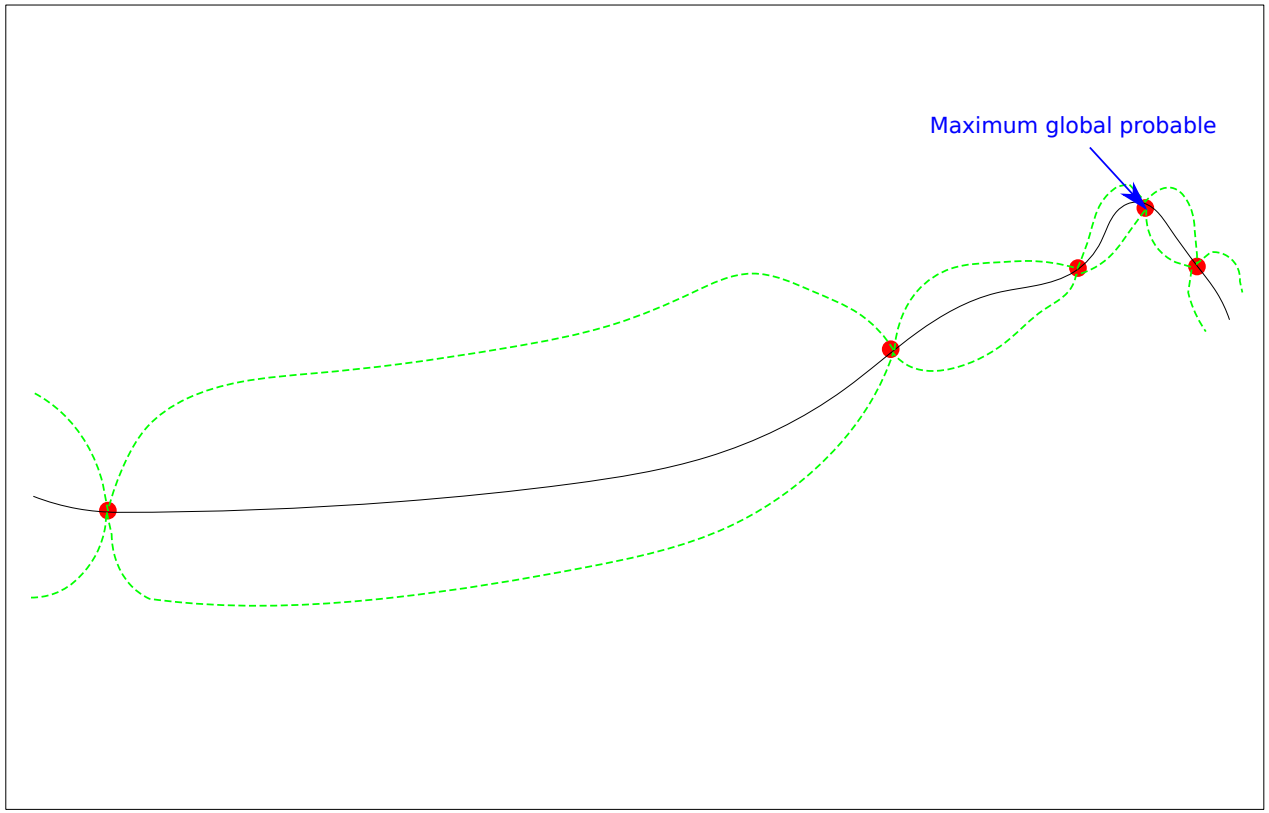
Observation $z_5 = z(x_5)$

# Application à l'optimisation bayésienne



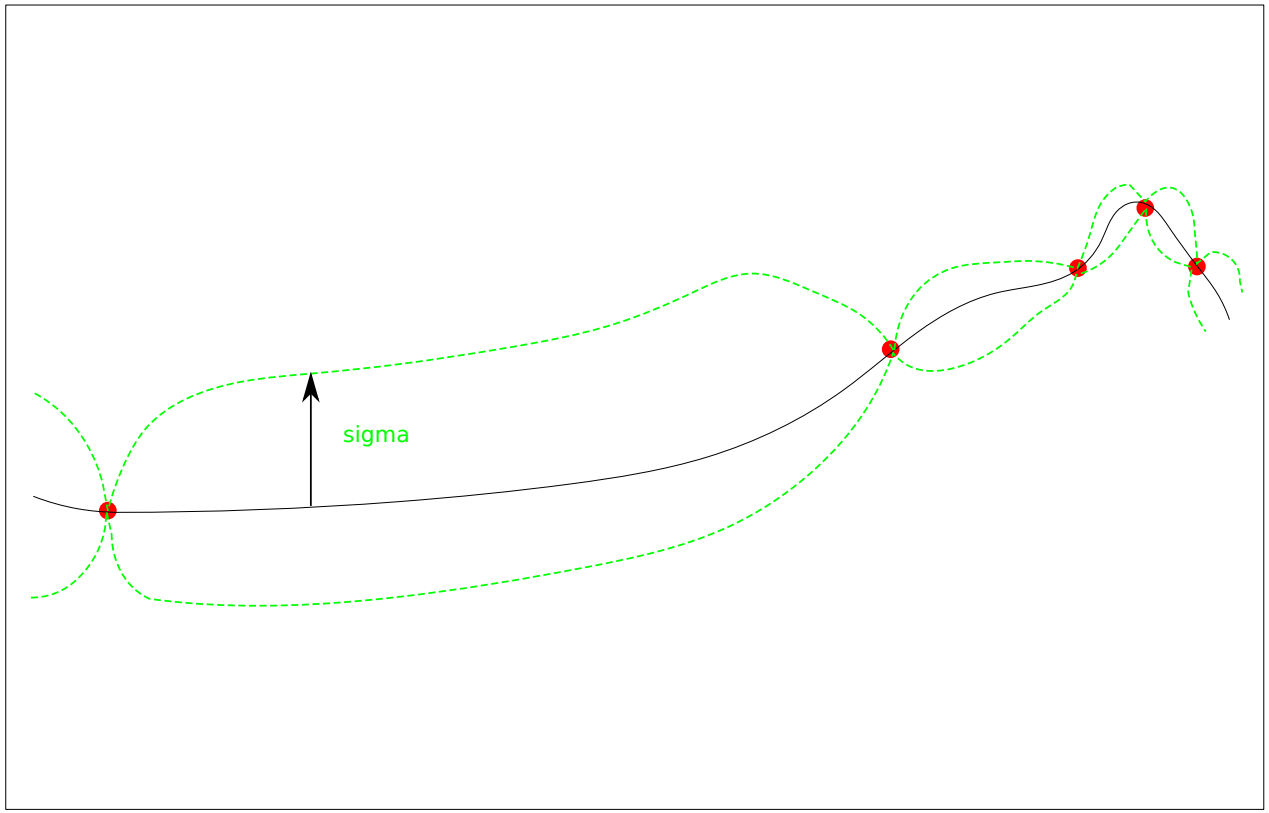
Mises à jour de $\hat{z}$ et $\mathrm{Var}[\hat{z}]$

# Application à l'optimisation bayésienne



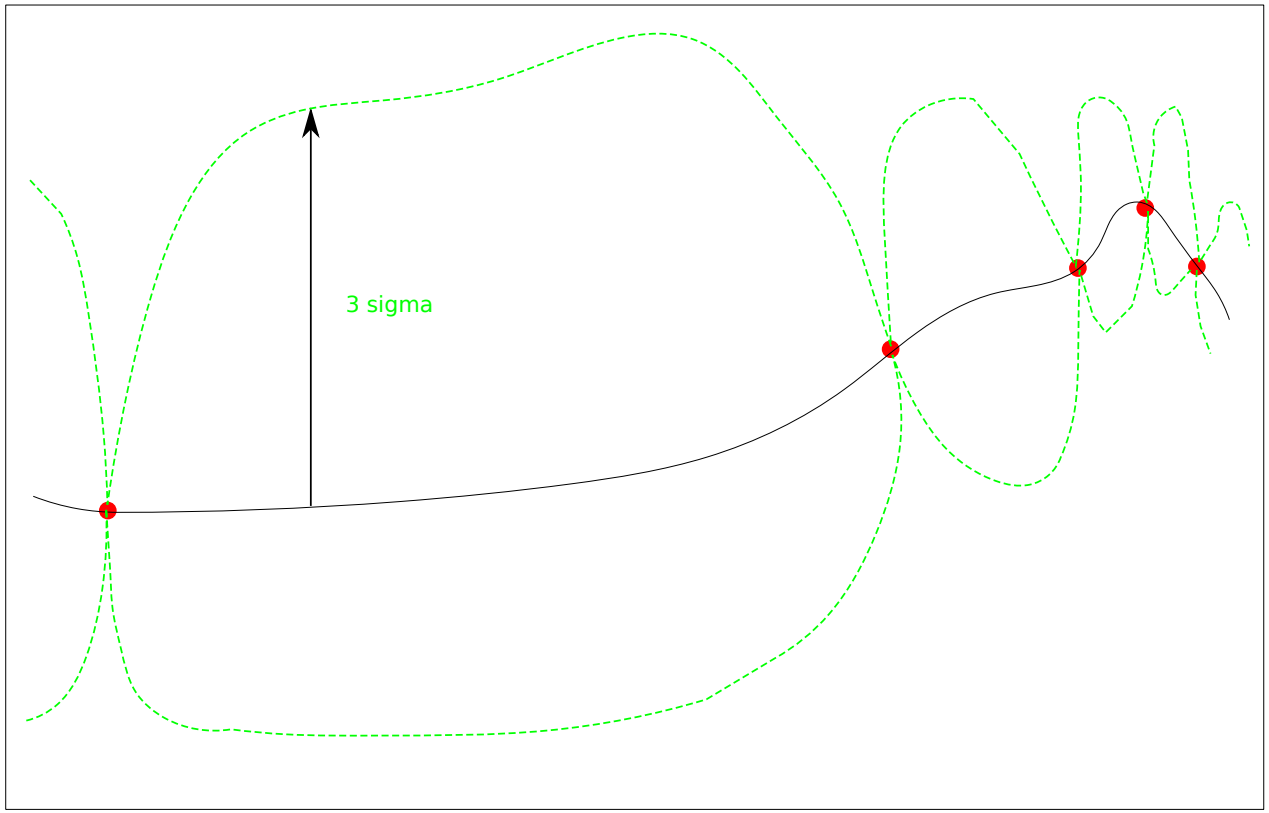

Après un certain nombre d'étapes : sélection du maximum global

# Application à l'optimisation bayésienne



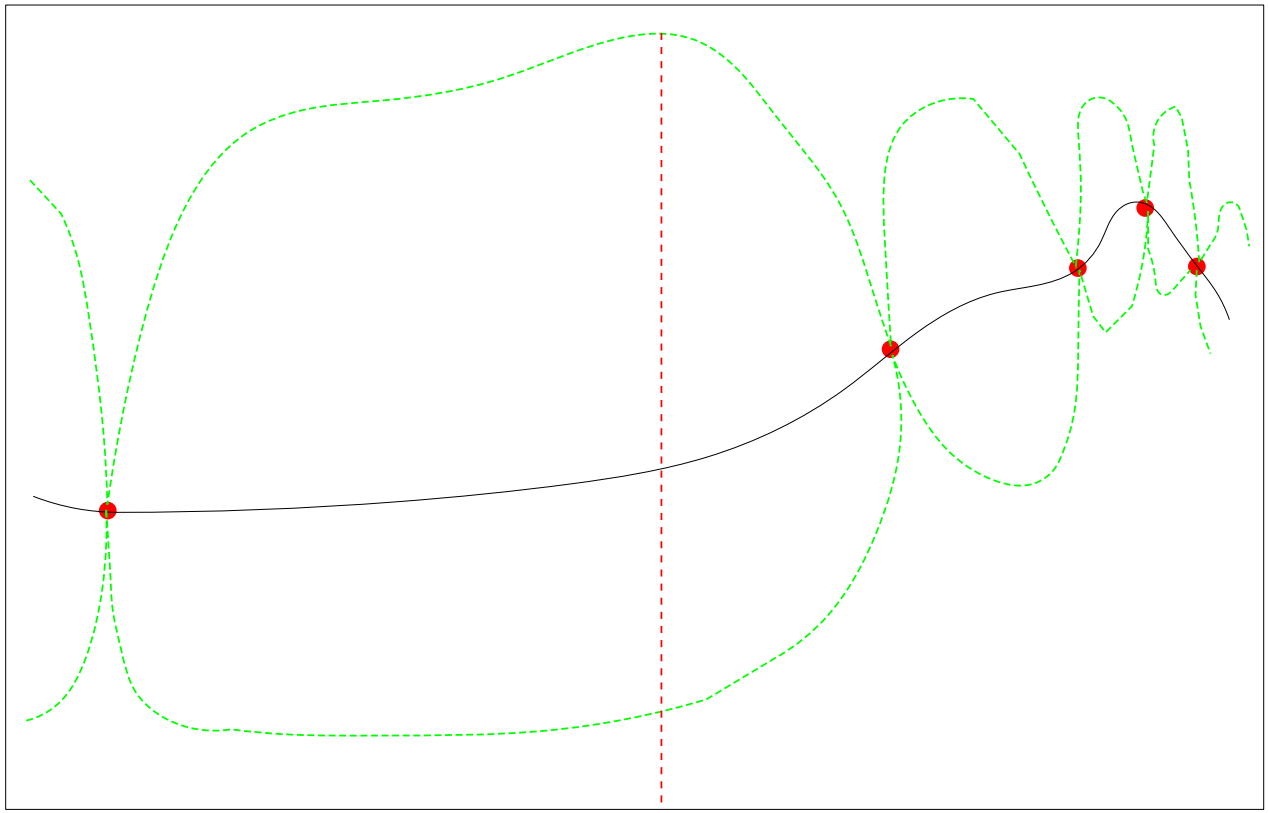

Les courbes vertes se calculent par $\hat{z} \pm k\sigma_{\hat{z}}$

# Application à l'optimisation bayésienne



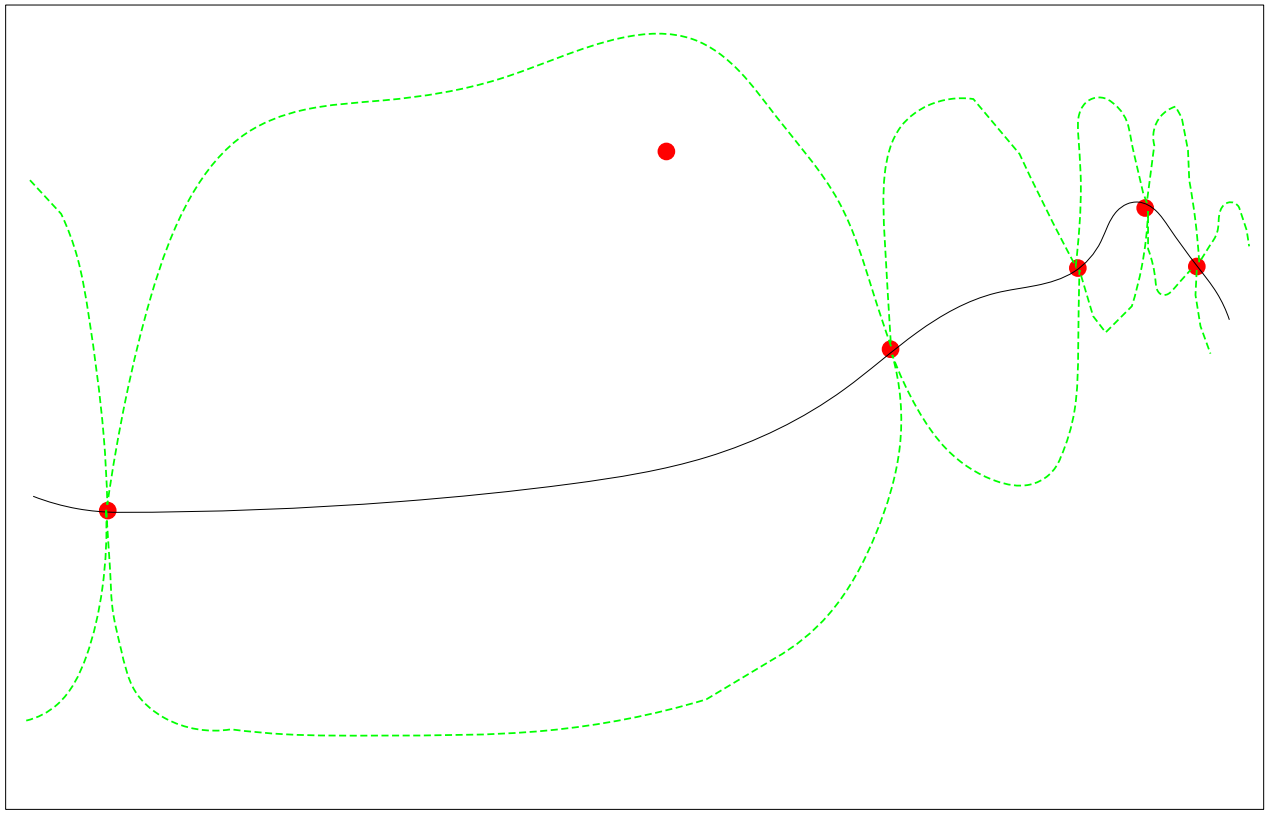

Le paramètre $k$ détermine le compromis exploration / intensification

# Application à l'optimisation bayésienne



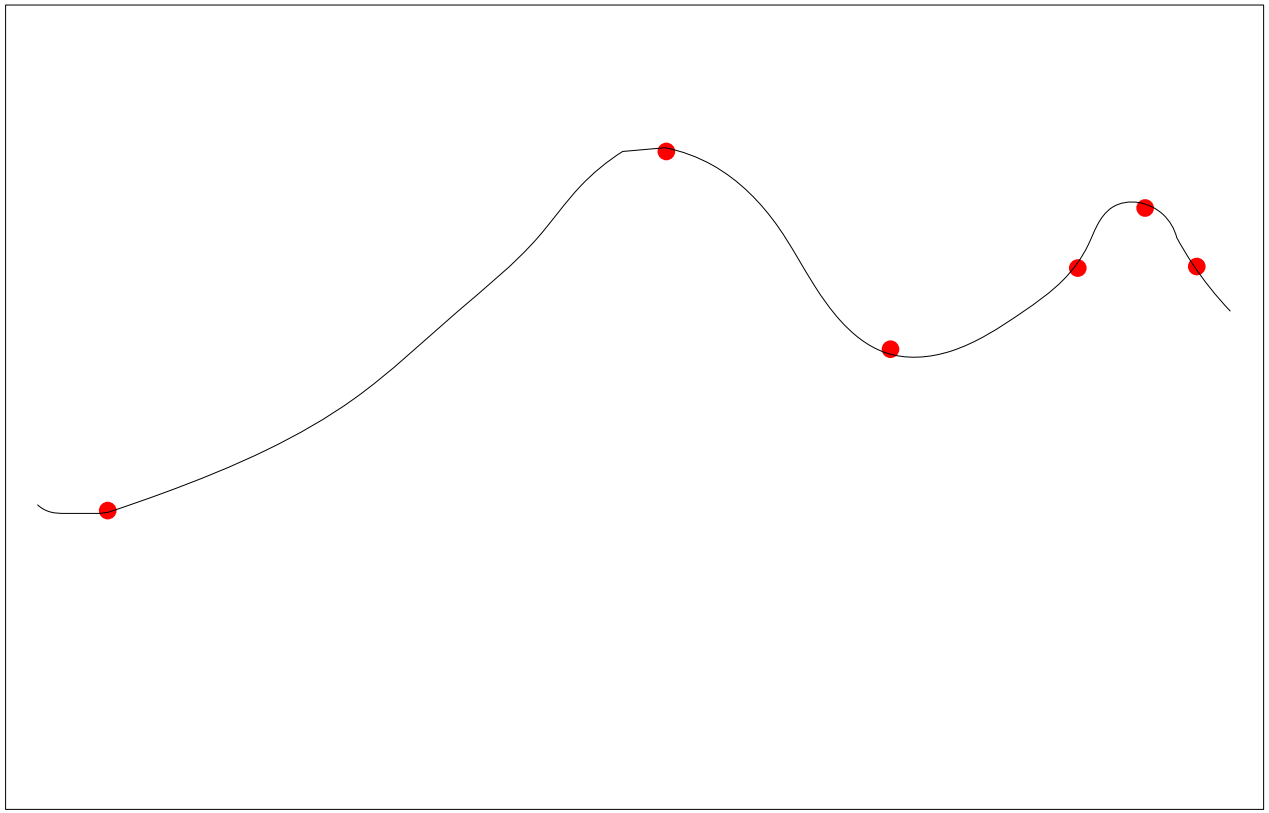
Sélection du site le plus prometteur $x_6$

# Application à l'optimisation bayésienne



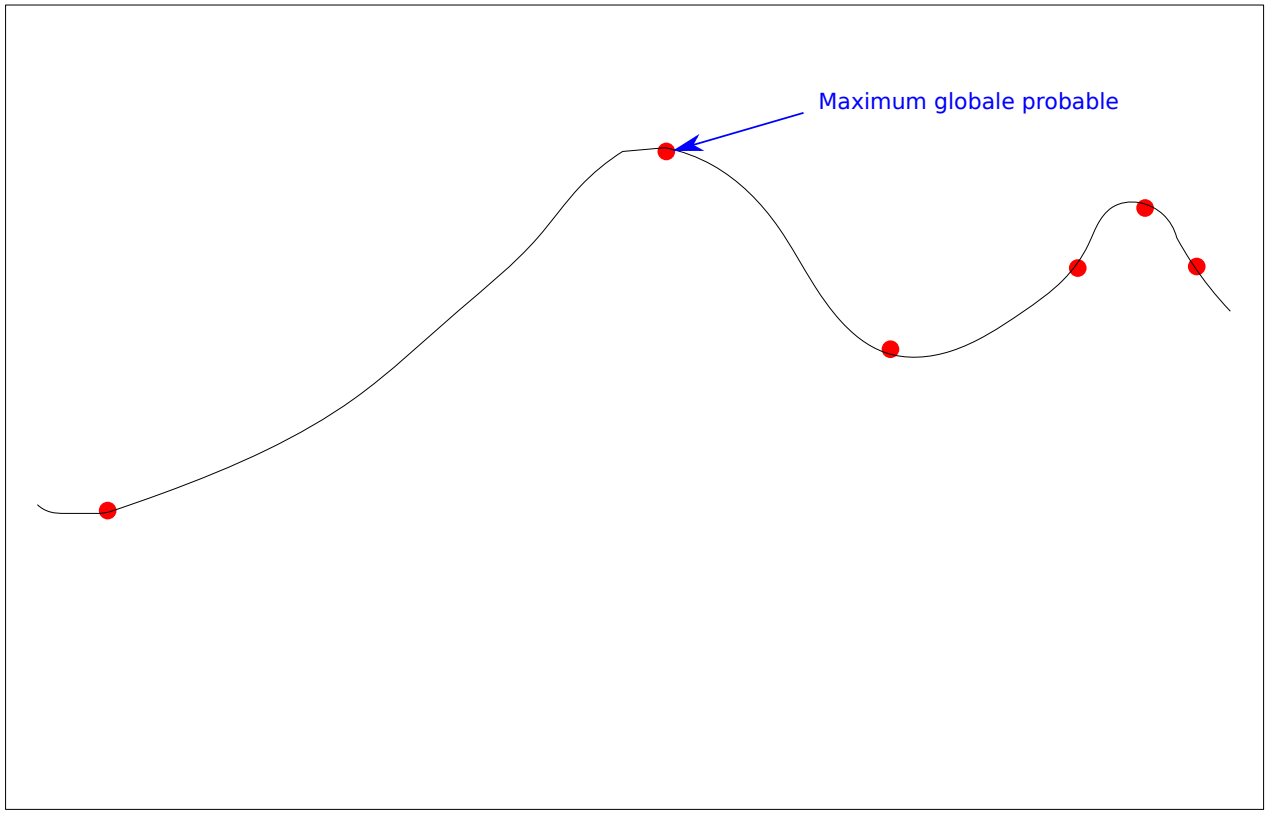
Observation $z_6 = z(x_6)$

# Application à l'optimisation bayésienne



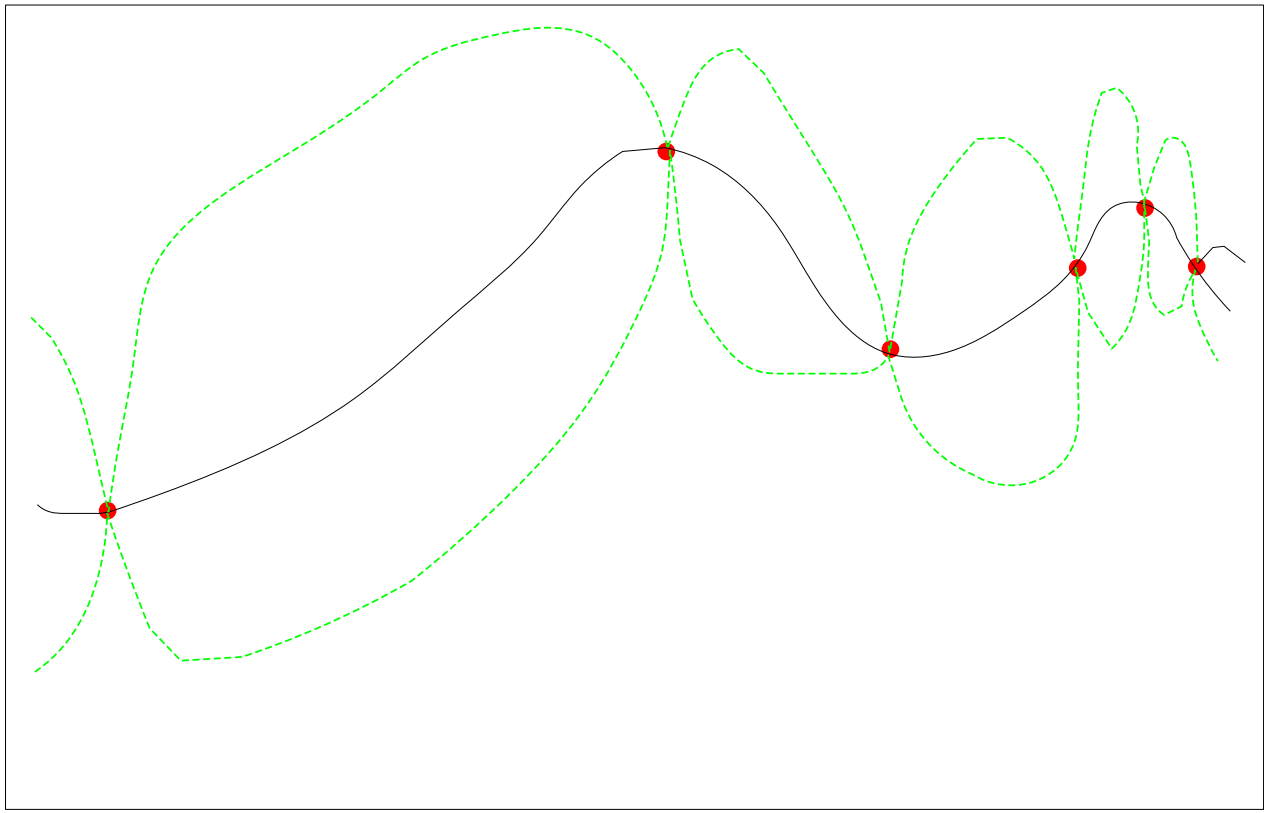
Estimation de $\hat{z}$ par krigeage

# Application à l'optimisation bayésienne



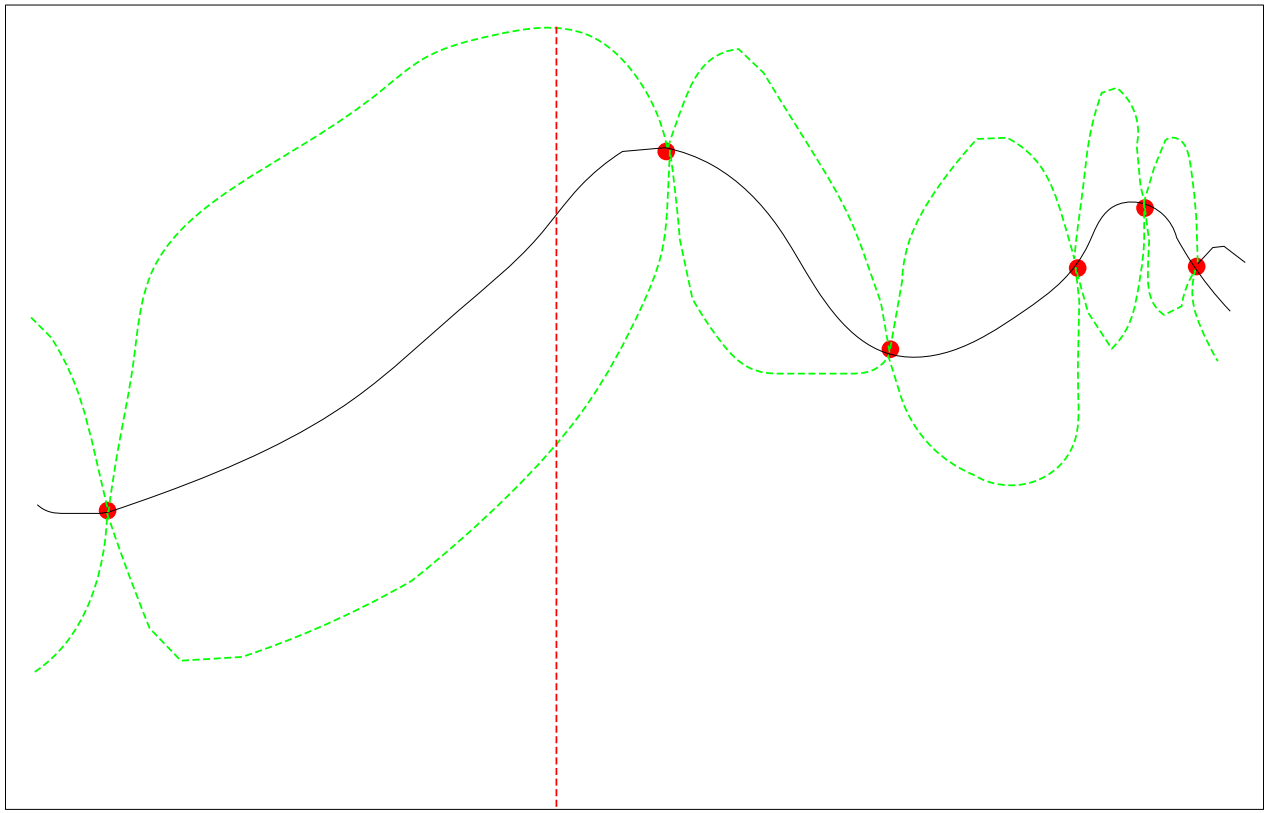

Nouveau maximum global (au moins aussi bon que le précédent)

# Application à l'optimisation bayésienne



Estimation de la variance associée $\mathrm{Var}[\hat{z}]$

# Application à l'optimisation bayésienne



Sélection du site le plus prometteur $x_7$ etc, etc

# La simulation

Etant donné un modèle de variogramme décrivant un phénomène doté d'une organisation spatiale, il est facile de générer aléatoirement un champ simulant le phénomène considéré.

# La simulation

Etant donné un modèle de variogramme décrivant un phénomène doté d'une organisation spatiale, il est facile de générer aléatoirement un champ simulant le phénomène considéré.

Soit $\mathbf{\Sigma}$ la matrice de variances-covariances du phénomène et $\mathbf{X} = (x_1, x_2, ...x_n)$ un vecteur contenant $n$ réalisations d'une distribution normale centrée réduite : $\mathcal{N}(0, 1)$.

# La simulation

Etant donné un modèle de variogramme décrivant un phénomène doté d'une organisation spatiale, il est facile de générer aléatoirement un champ simulant le phénomène considéré.

Soit $\mathbf{\Sigma}$ la matrice de variances-covariances du phénomène et $\mathbf{X} = (x_1, x_2, ...x_n)$ un vecteur contenant $n$ réalisations d'une distribution normale centrée réduite : $\mathcal{N}(0, 1)$.

Soit $\mathbf{A}$ une matrice carrée de dimension $n$. D'après la formule de propagation des incertitudes, la matrice de variances-covariances de la quantité $\mathbf{AX}$ est $\mathbf{A\Sigma_X A}^T = \mathbf{AA}^T$ puisque $\mathbf{\Sigma_X} = \mathbf{I_n}$.

# La simulation

Etant donné un modèle de variogramme décrivant un phénomène doté d'une organisation spatiale, il est facile de générer aléatoirement un champ simulant le phénomène considéré.

Soit $\mathbf{\Sigma}$ la matrice de variances-covariances du phénomène et $\mathbf{X} = (x_1, x_2, ...x_n)$ un vecteur contenant $n$ réalisations d'une distribution normale centrée réduite : $\mathcal{N}(0, 1)$.

Soit $\mathbf{A}$ une matrice carrée de dimension $n$. D'après la formule de propagation des incertitudes, la matrice de variances-covariances de la quantité $\mathbf{AX}$ est $\mathbf{A\Sigma_X A}^T = \mathbf{AA}^T$ puisque $\mathbf{\Sigma_X} = \mathbf{I_n}$.

Le vecteur $\mathbf{AX}$ a donc pour covariances $\mathbf{\Sigma}$ si et seulement si :

$$\mathbf{AA}^T = \mathbf{\Sigma}$$

# La simulation

**Une recette simple pour simuler un champ aléatoire :**

1) Calculer la matrice de variances-covariances $\mathbf{\Sigma}$ entre les $n$ sites sur lesquels on veut simuler le champ.

2) Calculer la décomposition de Cholesky de cette matrice : $\mathbf{\Sigma} = \mathbf{L}\mathbf{L}^T$ ($\mathbf{\Sigma}$ est définie-positive)

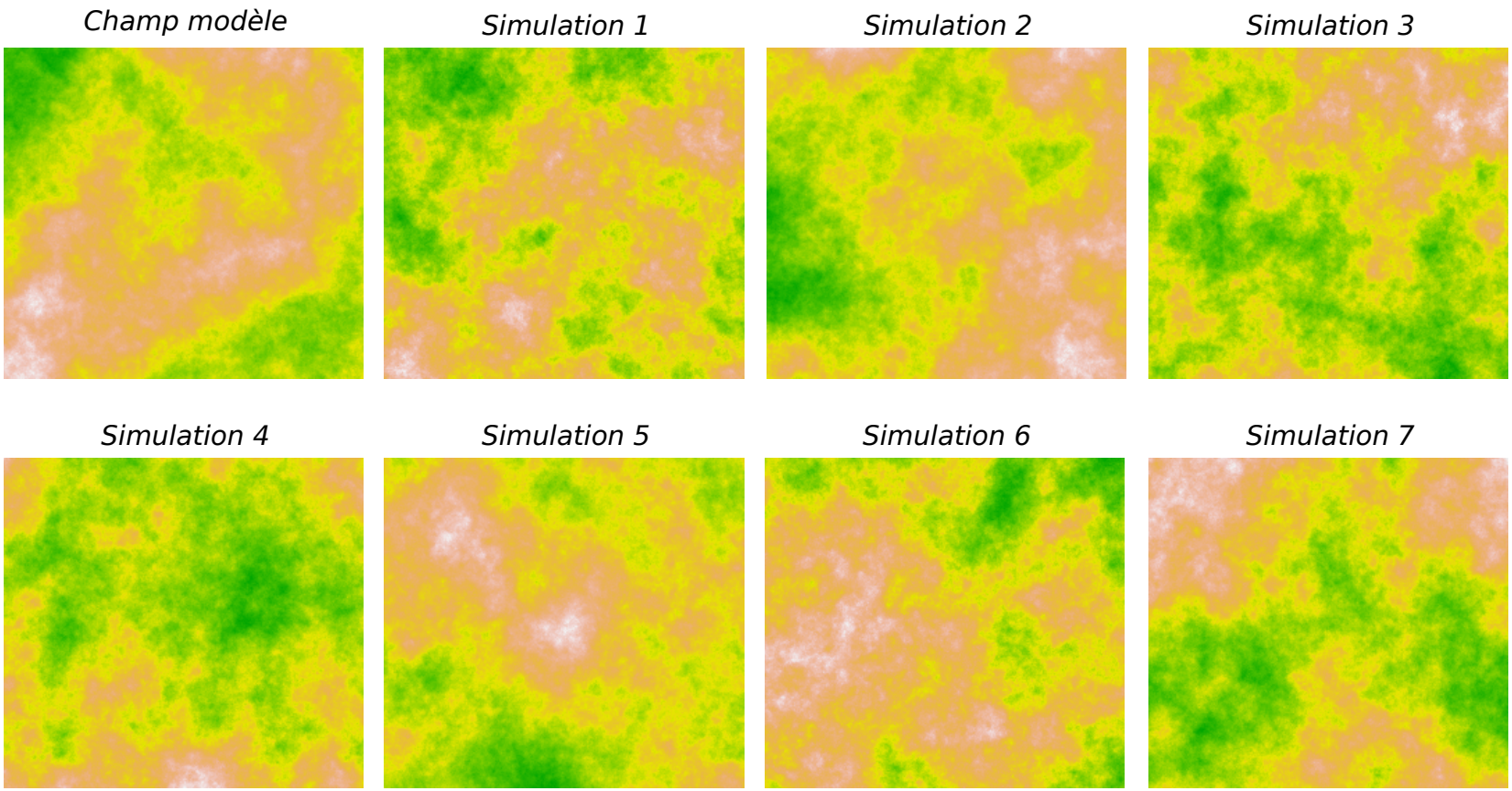
3) Générer $n$ variables aléatoires $\mathbf{X} = (x_1, x_2, ...x_n)$ suivant une loi normale standard : $\mathcal{N}(0_{\mathbb{R}^n}, \mathbf{Id}_{\mathbb{R}^n})$

4) Simuler le phénomène cible : $\mathbf{Z} = \mathbf{L}\mathbf{X}$

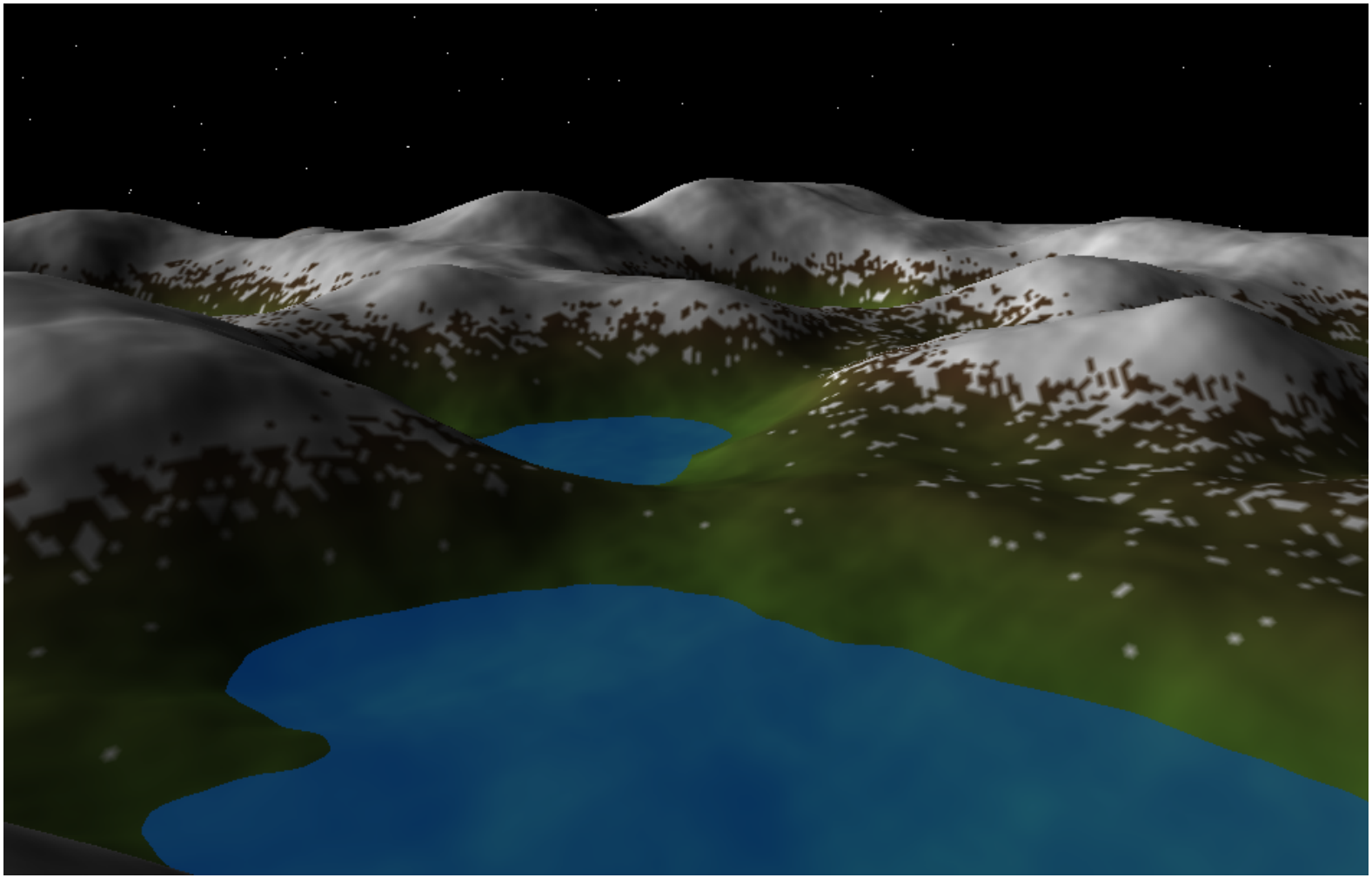
Note : il existe d'autres méthodes (bandes tournantes...)

# La simulation

## Exemple de simulation non-conditionnelle



# La simulation

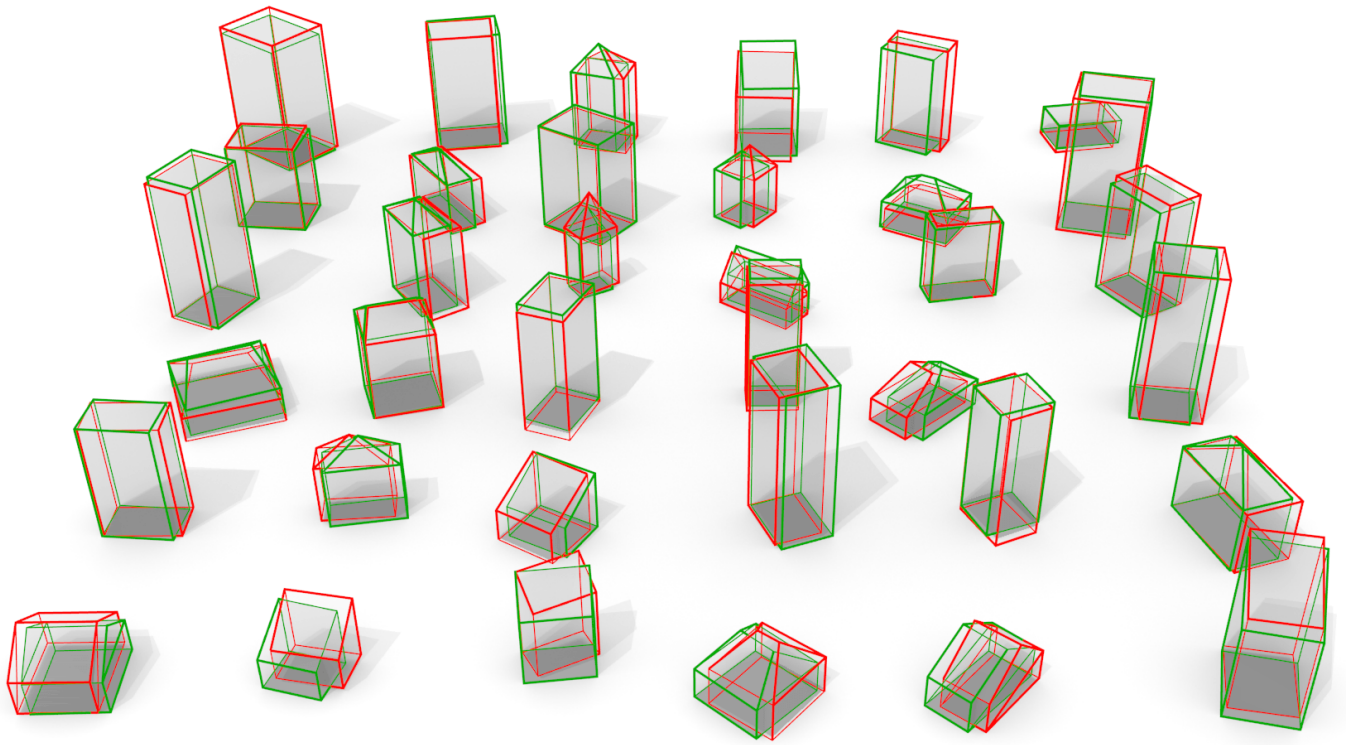


# La simulation

Biljecki et al : Propagation of positional error in 3D GIS : estimation of the solar irradiation of building roofs

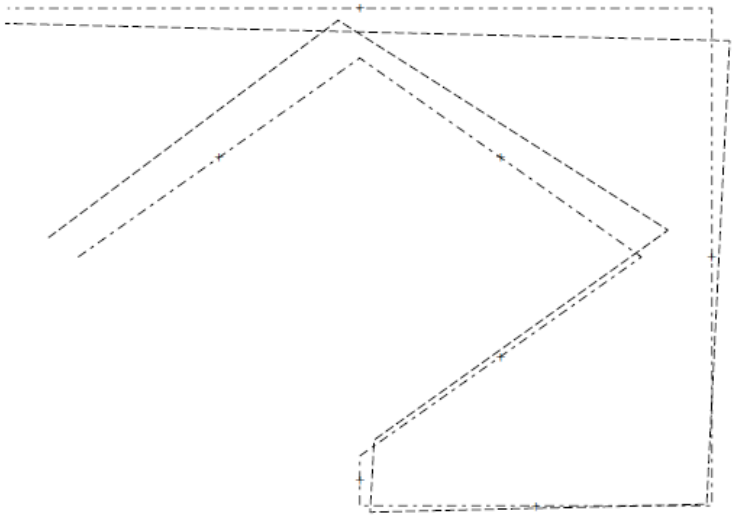
# La simulation

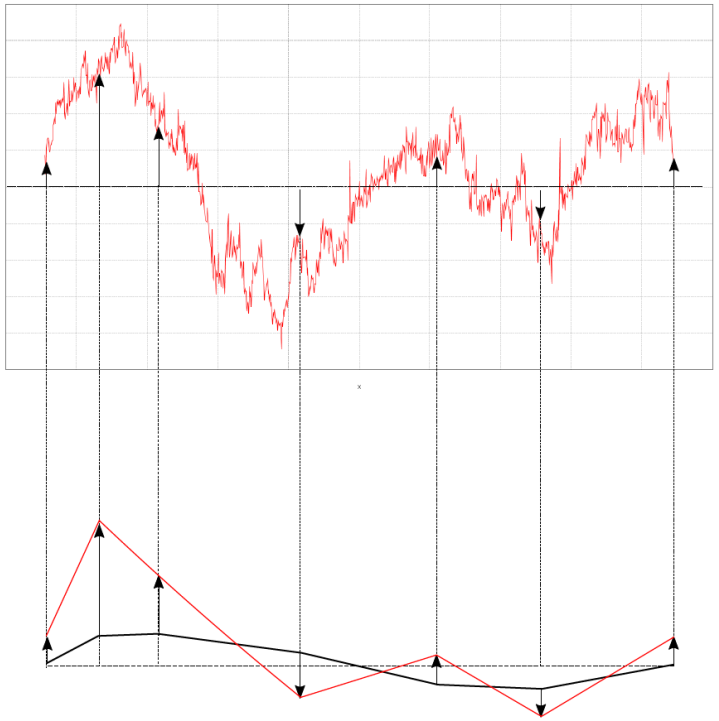
Biljecki et al : Propagation of positional error in 3D GIS : estimation of the solar irradiation of building roofs



# La simulation

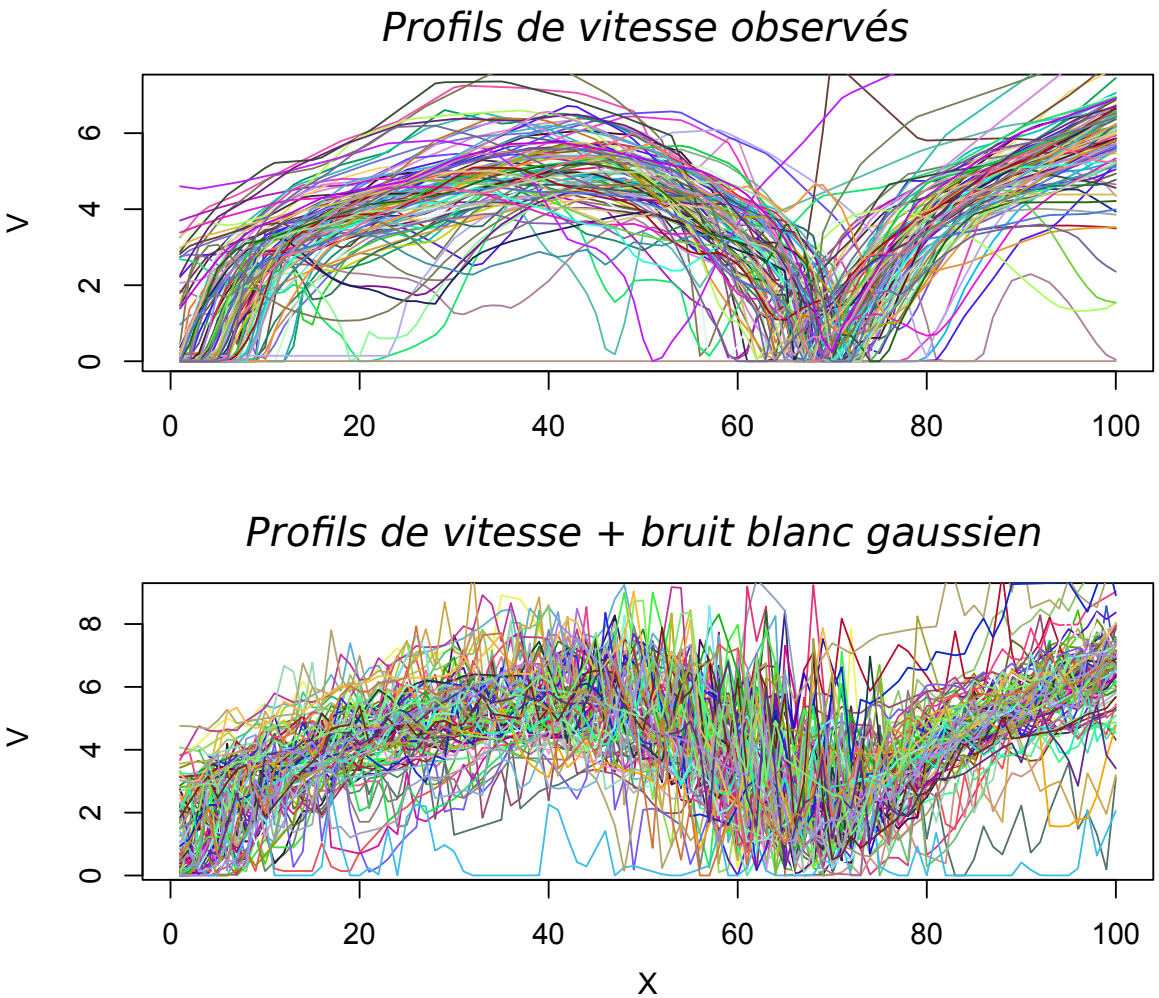
La simulation permet de générer des manière réaliste des champs dont on connaît un modèle de variogramme. Exemple : bruitage *réaliste* d'une polyligne pour de l'analyse de sensibilité.





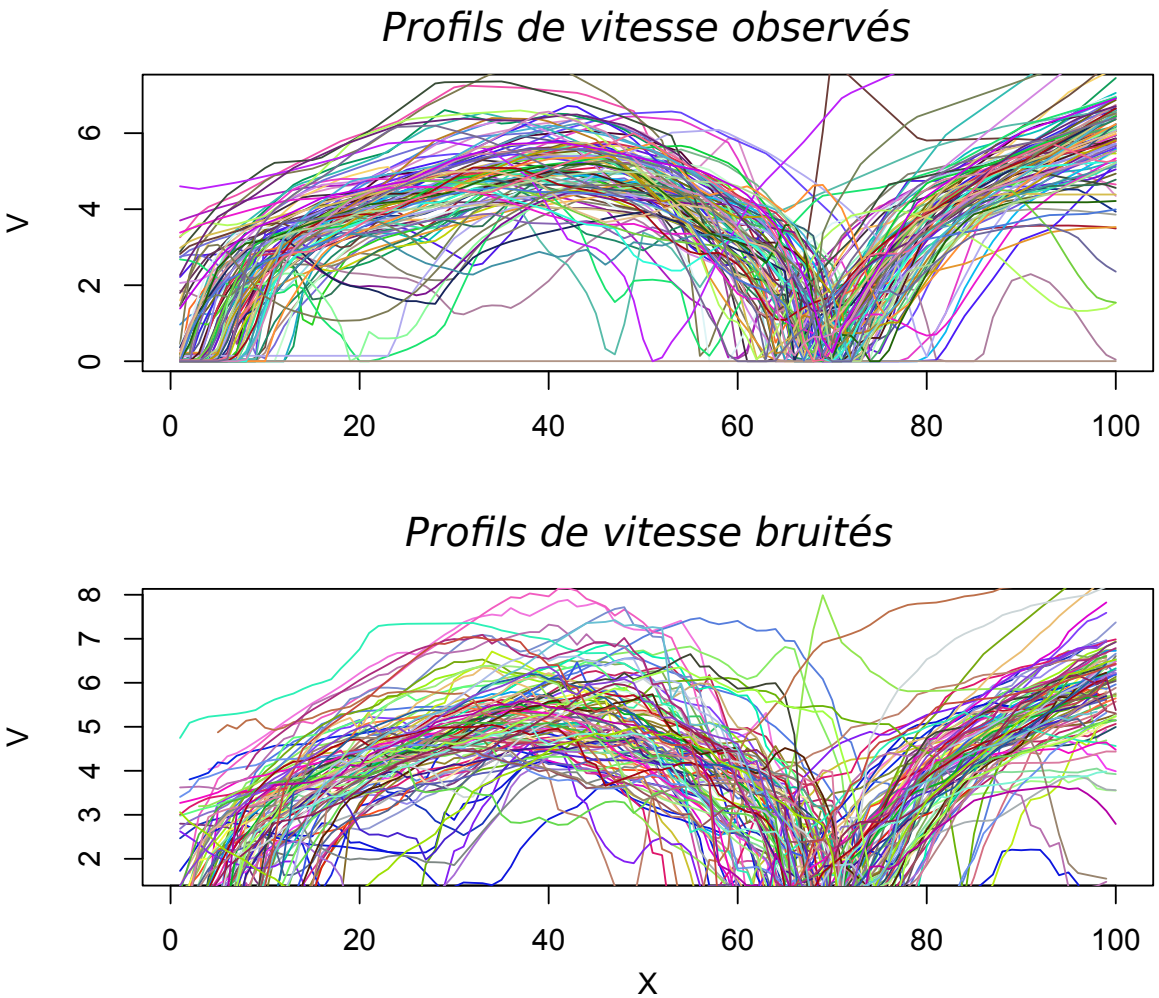
# La simulation

Exemple : bruitage de profils de vitesse issus de véhicules traceurs



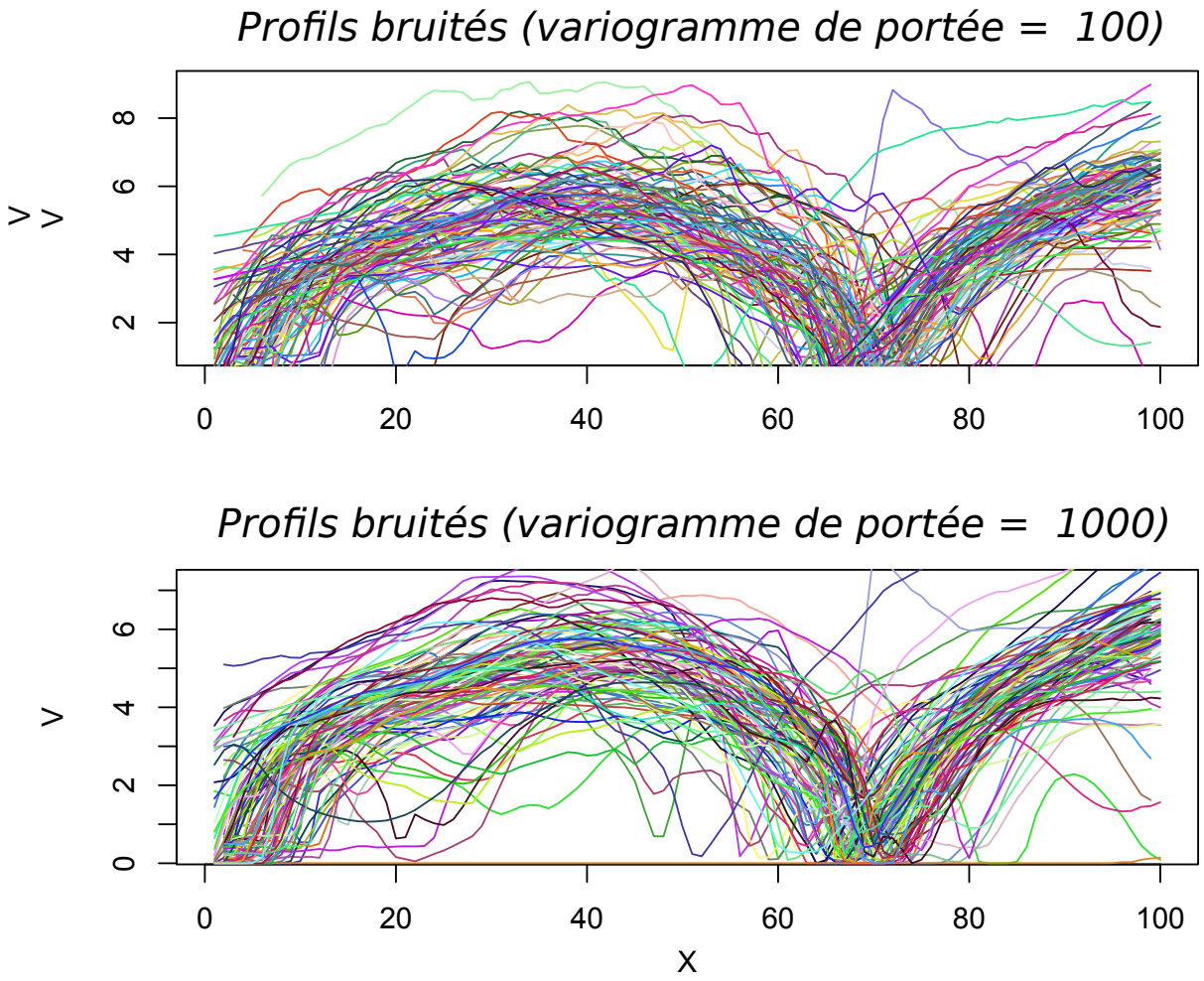
# La simulation

Exemple : bruitage de profils de vitesse issus de véhicules traceurs



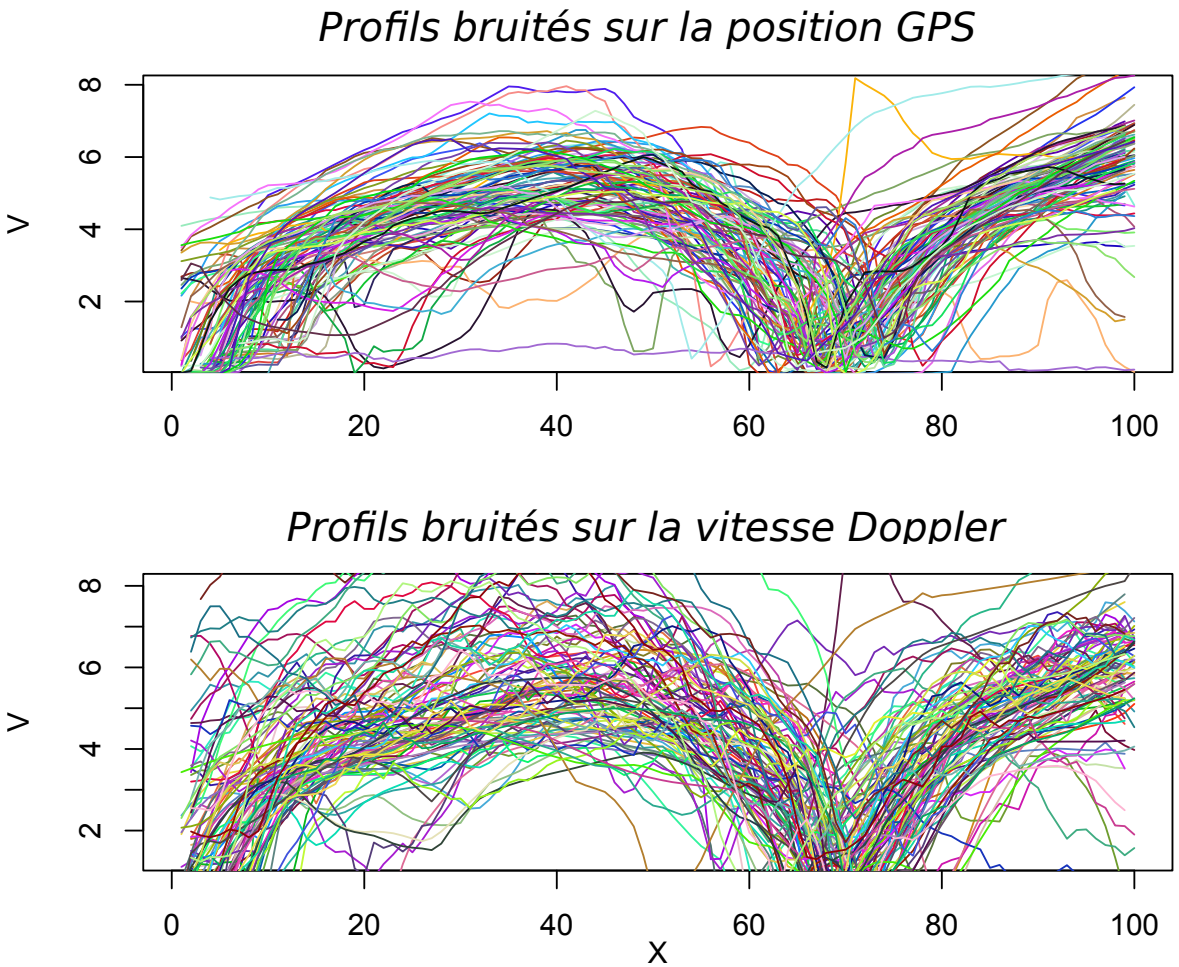
# La simulation

Exemple : bruitage de profils de vitesse issus de véhicules traceurs



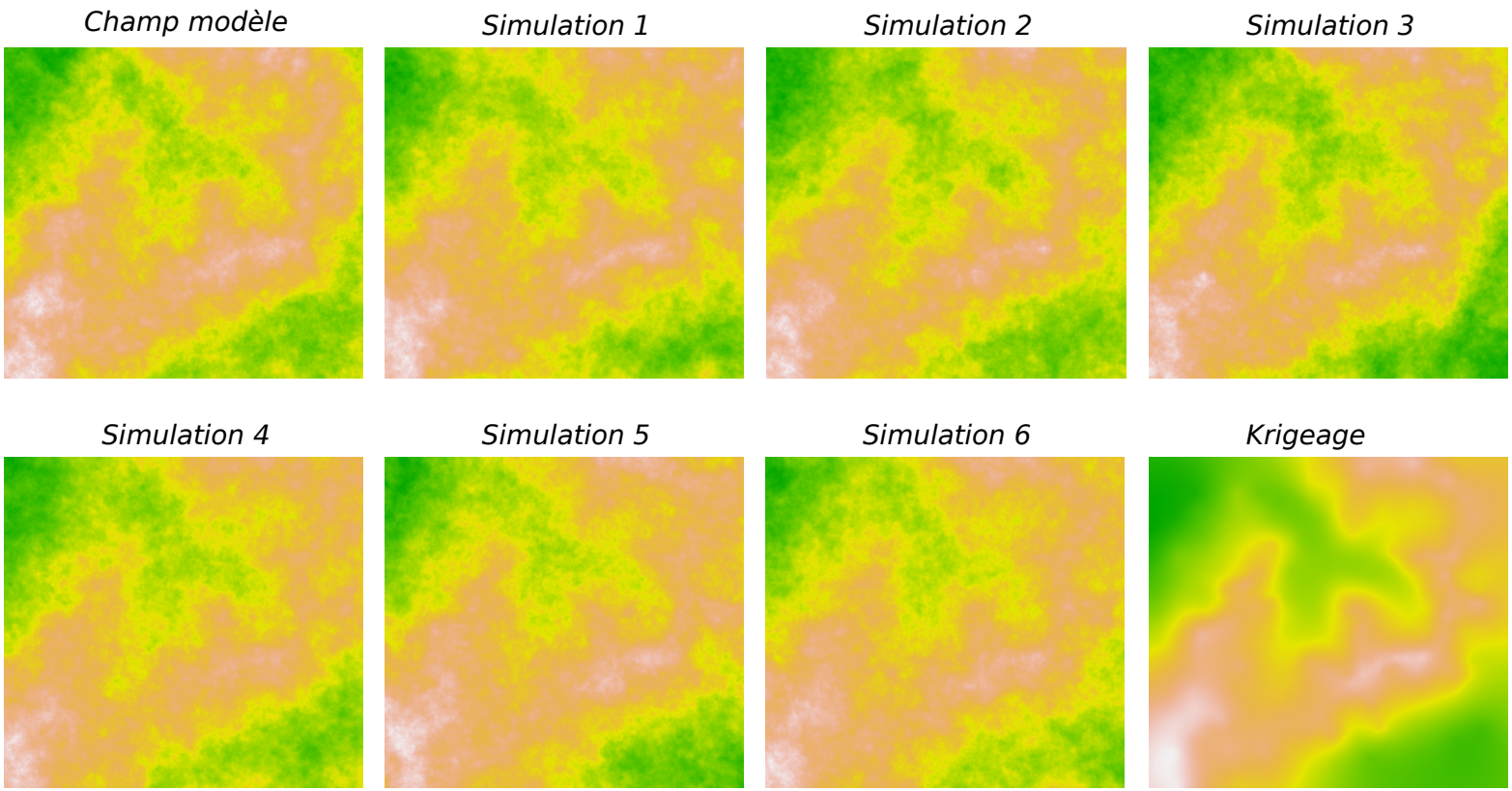
# La simulation

Exemple : bruitage de profils de vitesse issus de véhicules traceurs



# La simulation

On parle de **simulation conditionnelle** lorsque le champ est contraint de passer par un certain nombre de points.



# La simulation

**Intérêts de la simulation conditionnelle :**

# La simulation

**Intérêts de la simulation conditionnelle :**

Le résultat obtenu est donc un exemple (parmi une infinité d'autres) de champ respectant à la fois la structure du phénomène étudié ainsi que les observations effectuées.

# La simulation

**Intérêts de la simulation conditionnelle :**

Le résultat obtenu est donc un exemple (parmi une infinité d'autres) de champ respectant à la fois la structure du phénomène étudié ainsi que les observations effectuées.

Contrairement au krigeage, la solution obtenue (qui est différente à chaque tirage) ne possède pas de propriété de minimalité de la variance d'estimation.

# La simulation

**Intérêts de la simulation conditionnelle :**

Le résultat obtenu est donc un exemple (parmi une infinité d'autres) de champ respectant à la fois la structure du phénomène étudié ainsi que les observations effectuées.

Contrairement au krigeage, la solution obtenue (qui est différente à chaque tirage) ne possède pas de propriété de minimalité de la variance d'estimation.

Une simulation donnée est donc a priori moins *précise* que la solution du krigeage.

# La simulation

**Intérêts de la simulation conditionnelle :**

En revanche, la simulation conditionnelle possède l'avantage de fournir une solution différente à chaque tirage.

# La simulation

**Intérêts de la simulation conditionnelle :**

En revanche, la simulation conditionnelle possède l'avantage de fournir une solution différente à chaque tirage.

La moyenne des solutions obtenues est en pratique plus robuste que la solution du krigeage, en particulier en présence de fortes non-linéarités (e.g. discontinuités) et effets de seuillage.

# La simulation

**Intérêts de la simulation conditionnelle :**

En revanche, la simulation conditionnelle possède l'avantage de fournir une solution différente à chaque tirage.

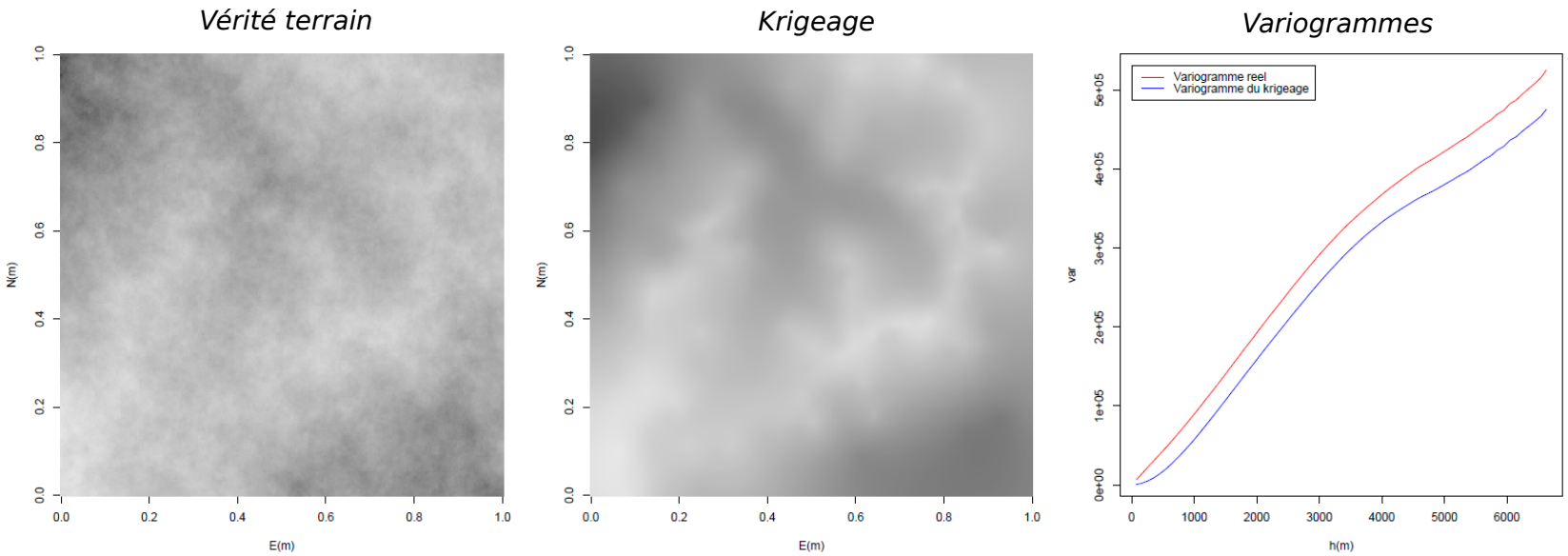
La moyenne des solutions obtenues est en pratique plus robuste que la solution du krigeage, en particulier en présence de fortes non-linéarités (e.g. discontinuités) et effets de seuillage.

D'autre part, on obtient une population de solutions, ce qui permet d'estimer les écarts-types et les intervalles de confiance des quantités à estimer.

# La simulation

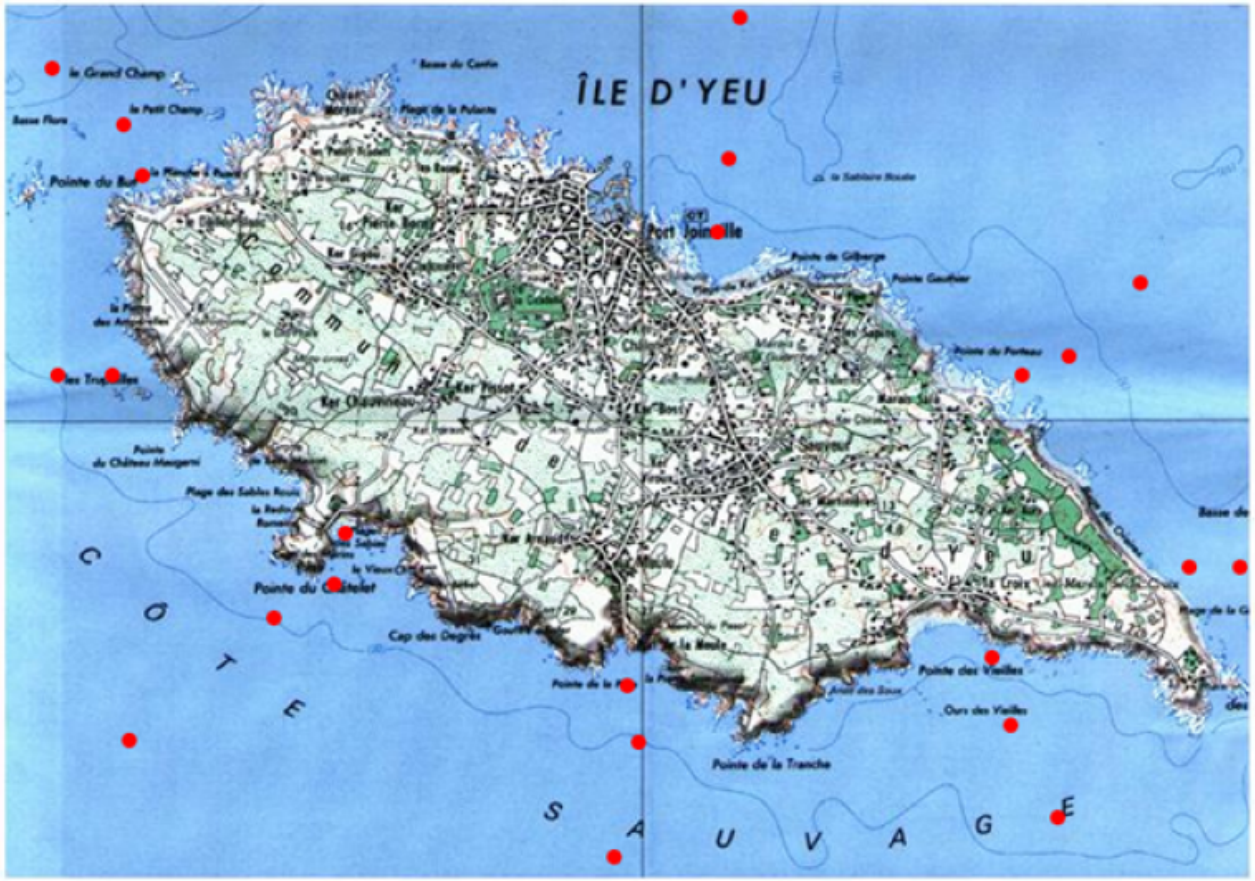
Le Krigeage a tendance à *lisser* les observations

effet_lissage.r



# La simulation

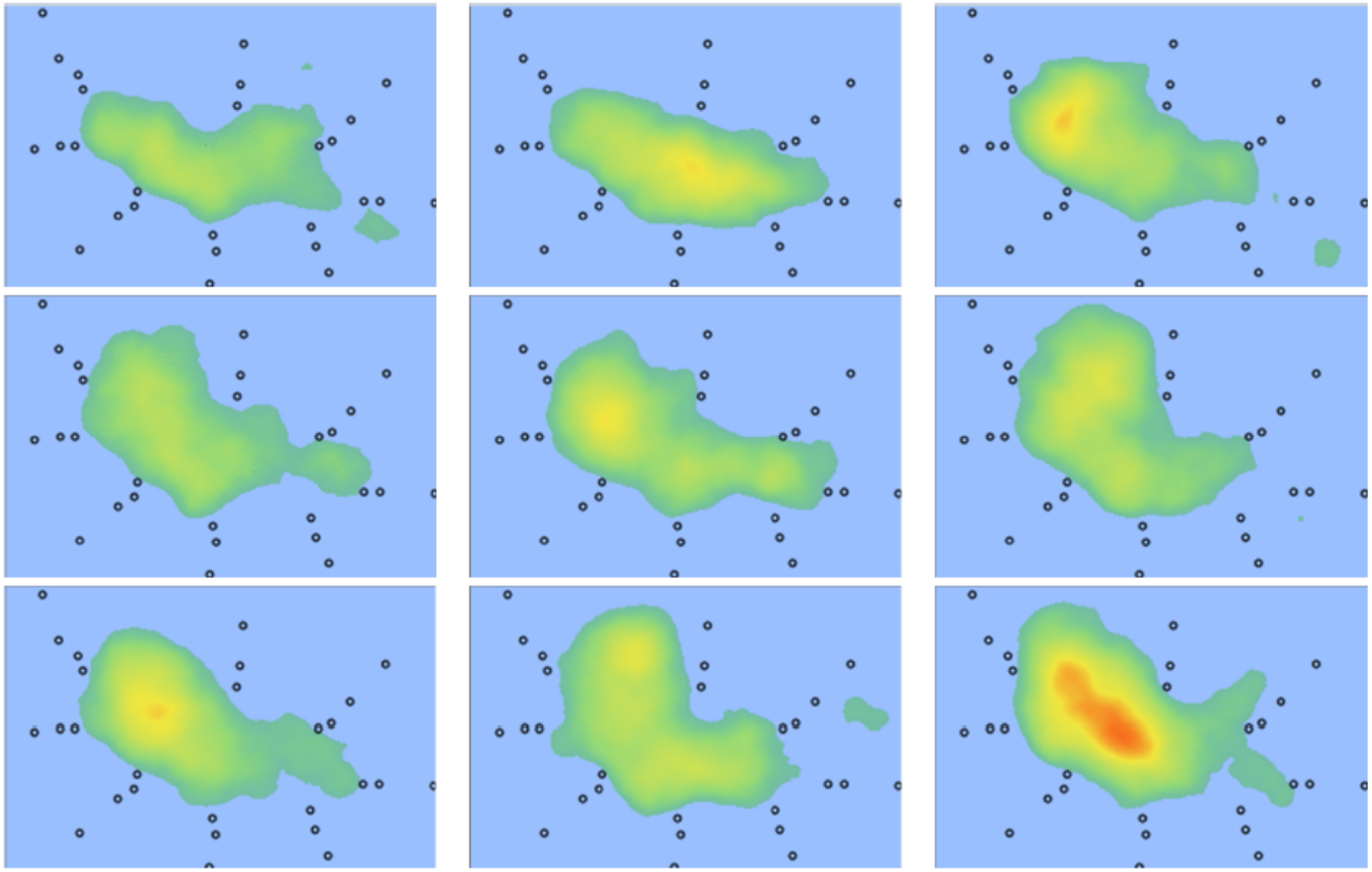
Calcul de la surface d'une île (Wackernagel, 2013)




Mesure de la profondeur bathymétrique en plusieurs points.

# La simulation

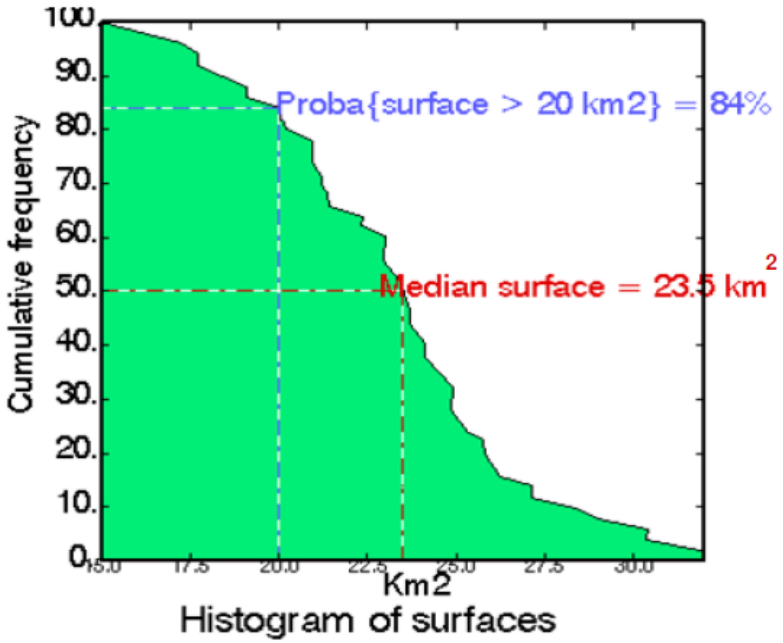
Calcul de la surface d'une île (Wackernagel, 2013)



Plusieurs simulations conditionnelles (les altitudes négatives sont seuillées à zéro)

# La simulation

Calcul de la surface d'une île (Wackernagel, 2013)

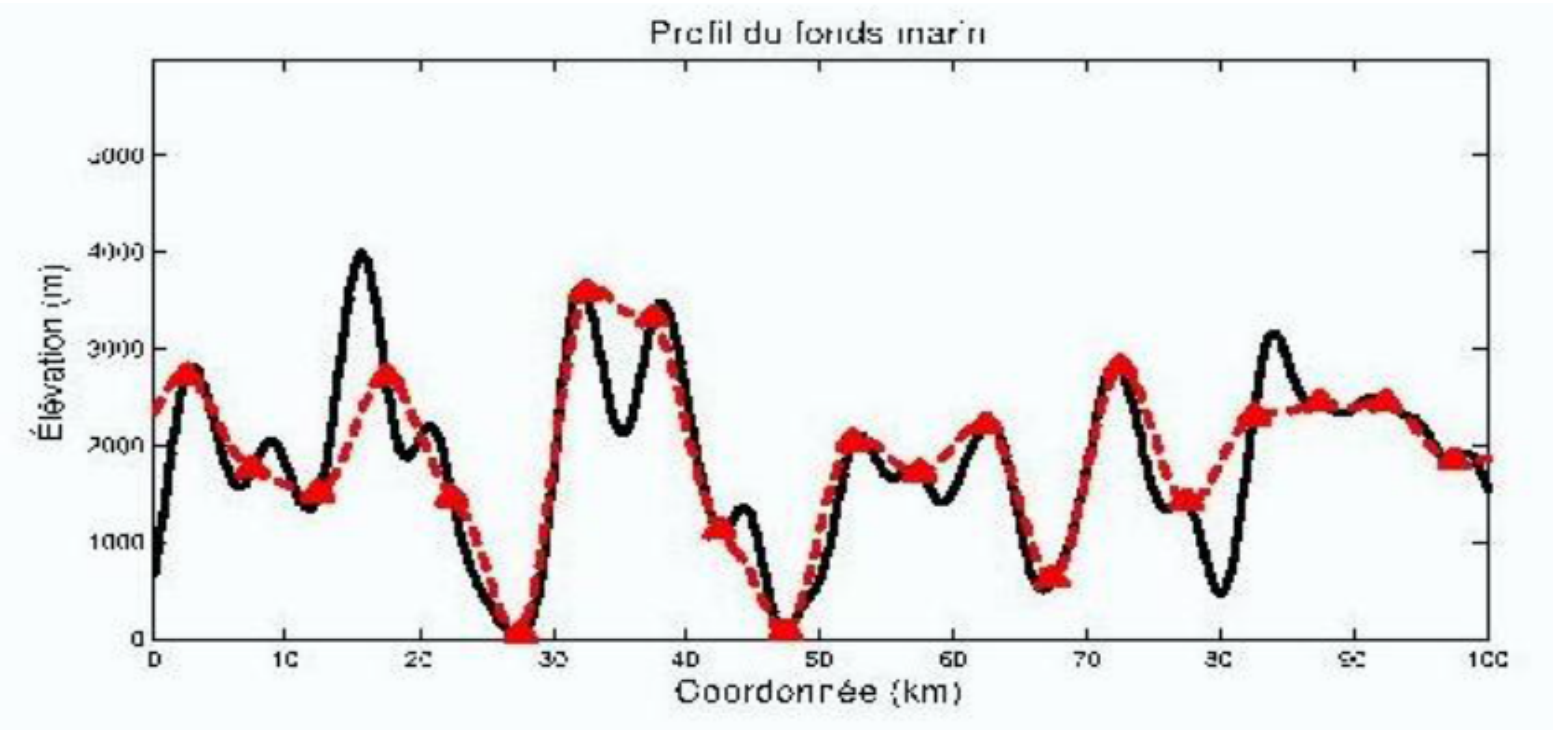



| | Real | From kriging | Sim. min | Sim. mean | Sim. max |
|---|---|---|---|---|---|
| Surface | **23.3** | 22.9 | 15.4 | 23.2 | 31.9 |

Calcul de la surface de l'île et de l'intervalle de confiance associé

# La simulation

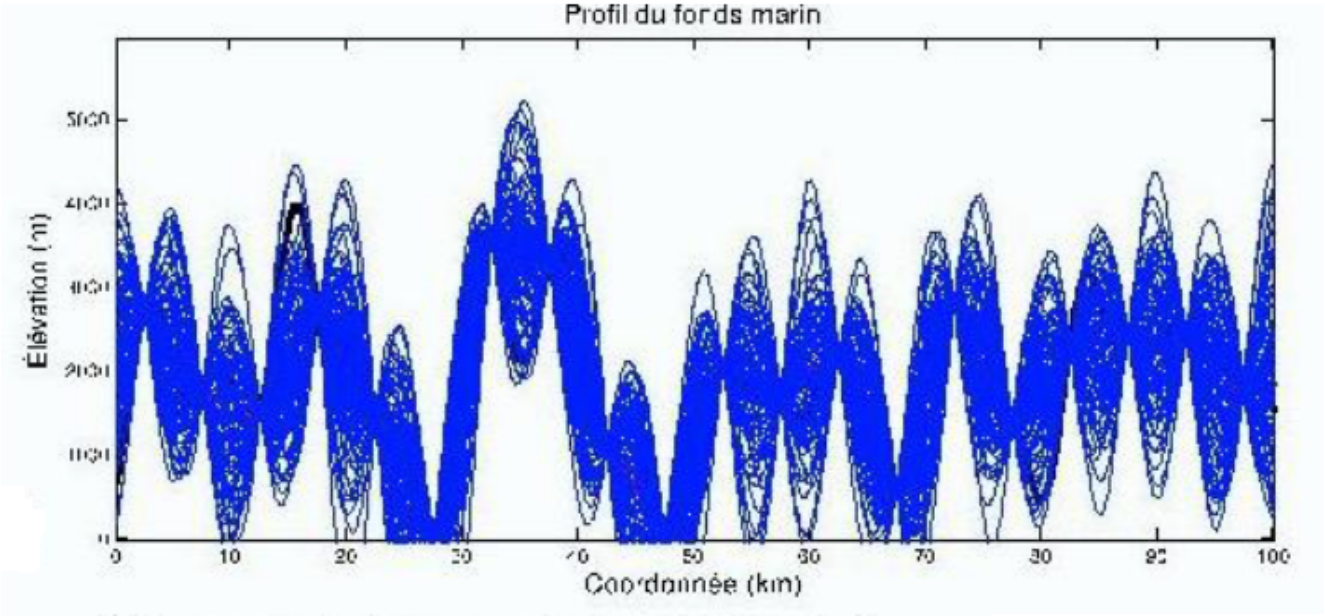
Calcul de longueur de câble (Marcotte, cours EPM)




Longueur réelle : 110 km. Estimation par krigeage : 104.2 km.

# La simulation

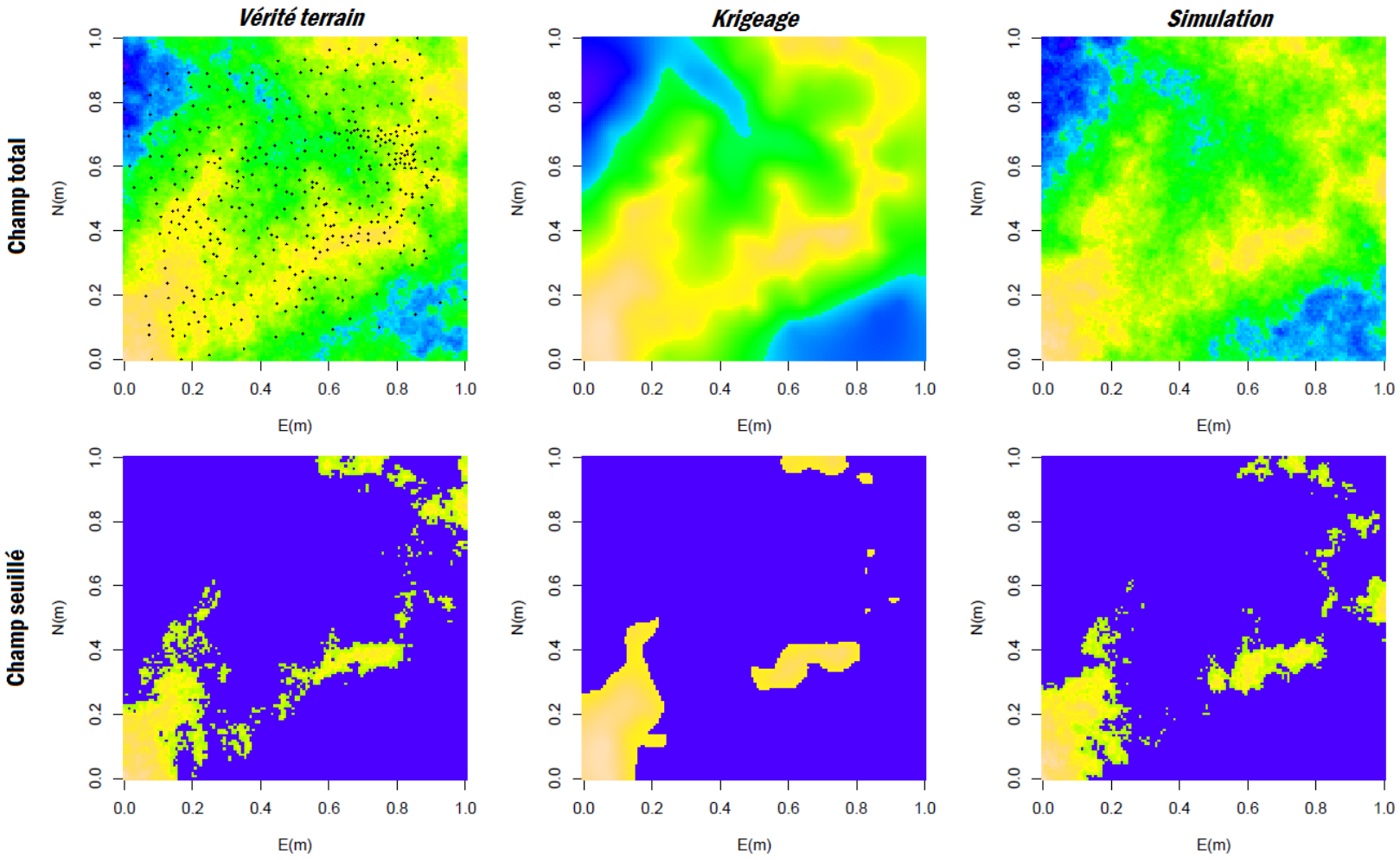
Calcul de longueur de câble (Marcotte, cours EPM)




Simulations conditionnelles : longueur moyenne 110 km. IC à 95 % : [108.8, 113.5]

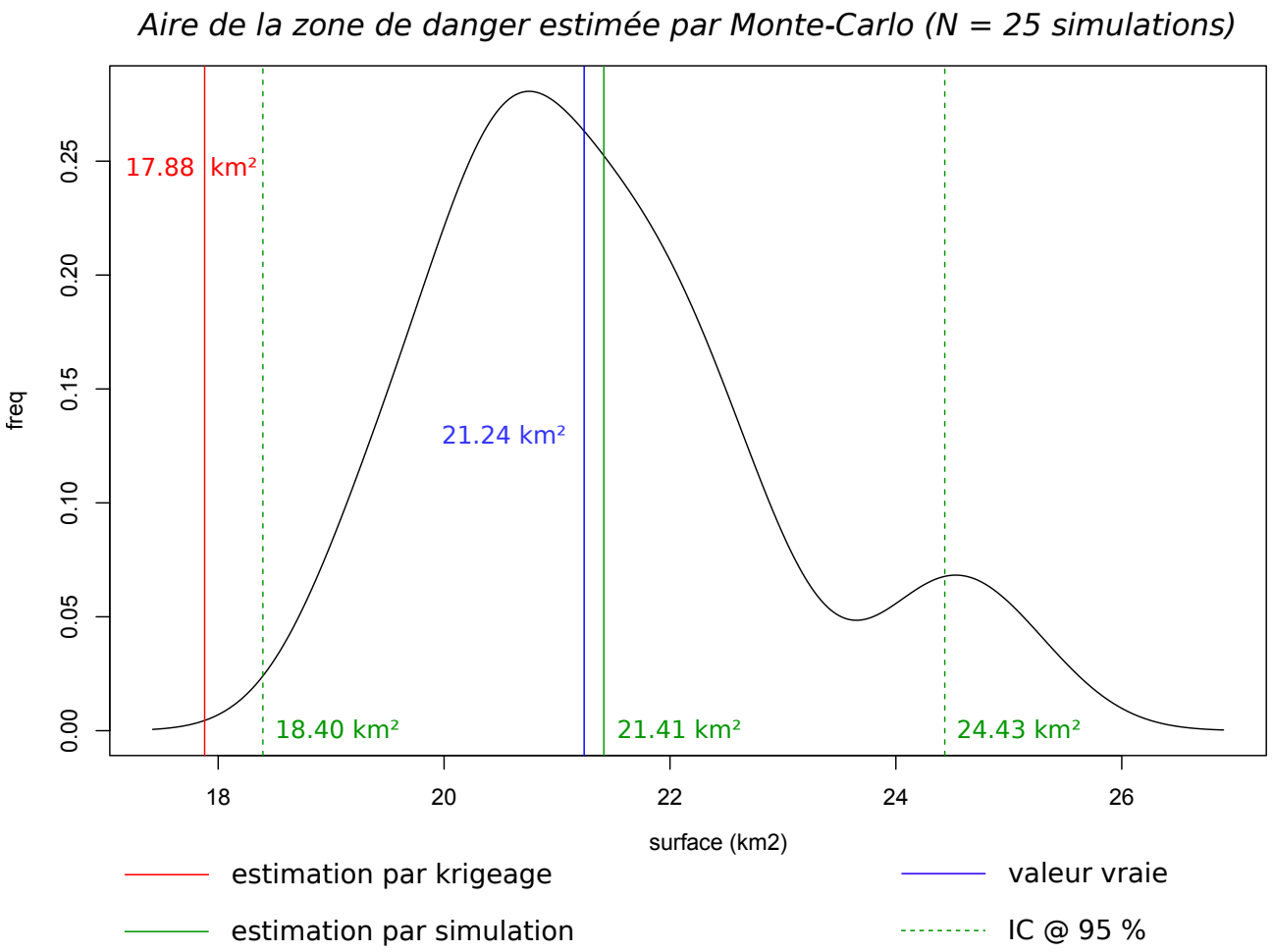
# La simulation

## Estimation de l'aire d'une zone de danger



# La simulation

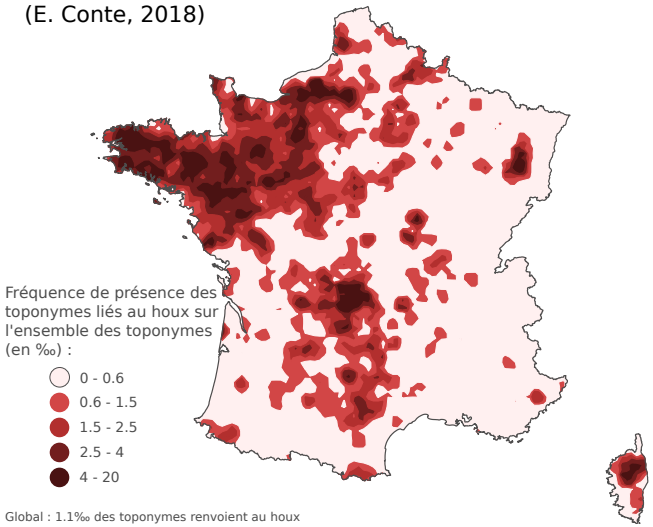
## Estimation de l'aire d'une zone de danger



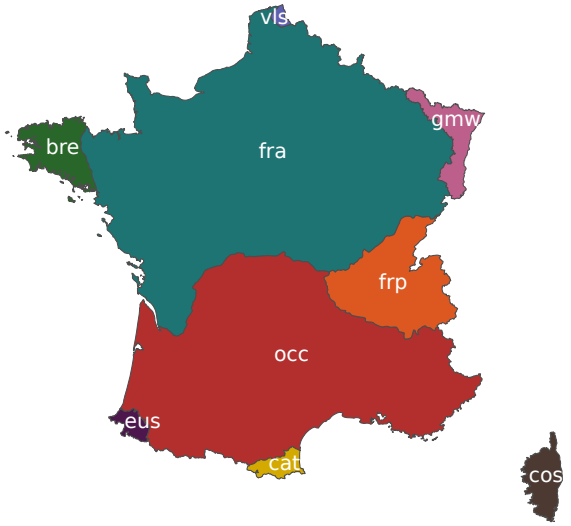
# La simulation

***Densité de toponymes***

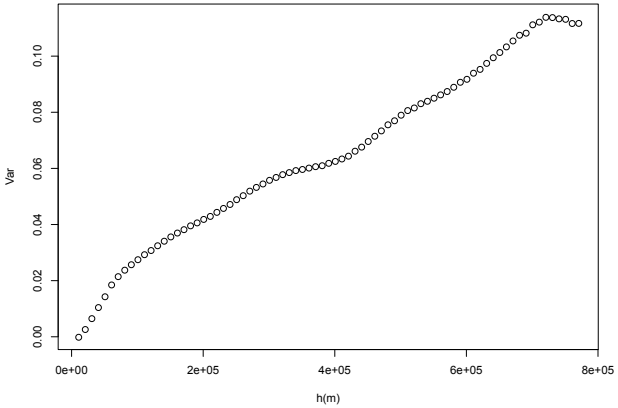
(E. Conte, 2018)



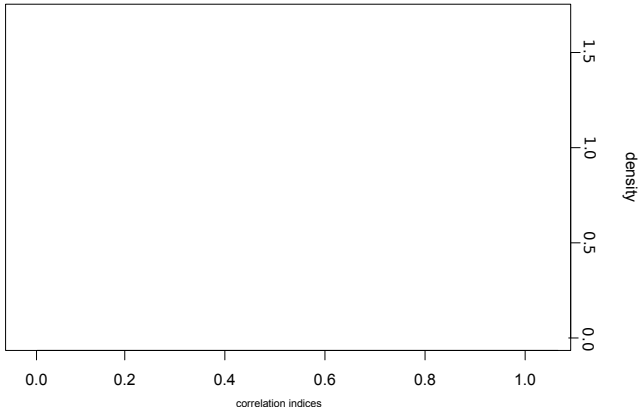

***Régions linguistiques***



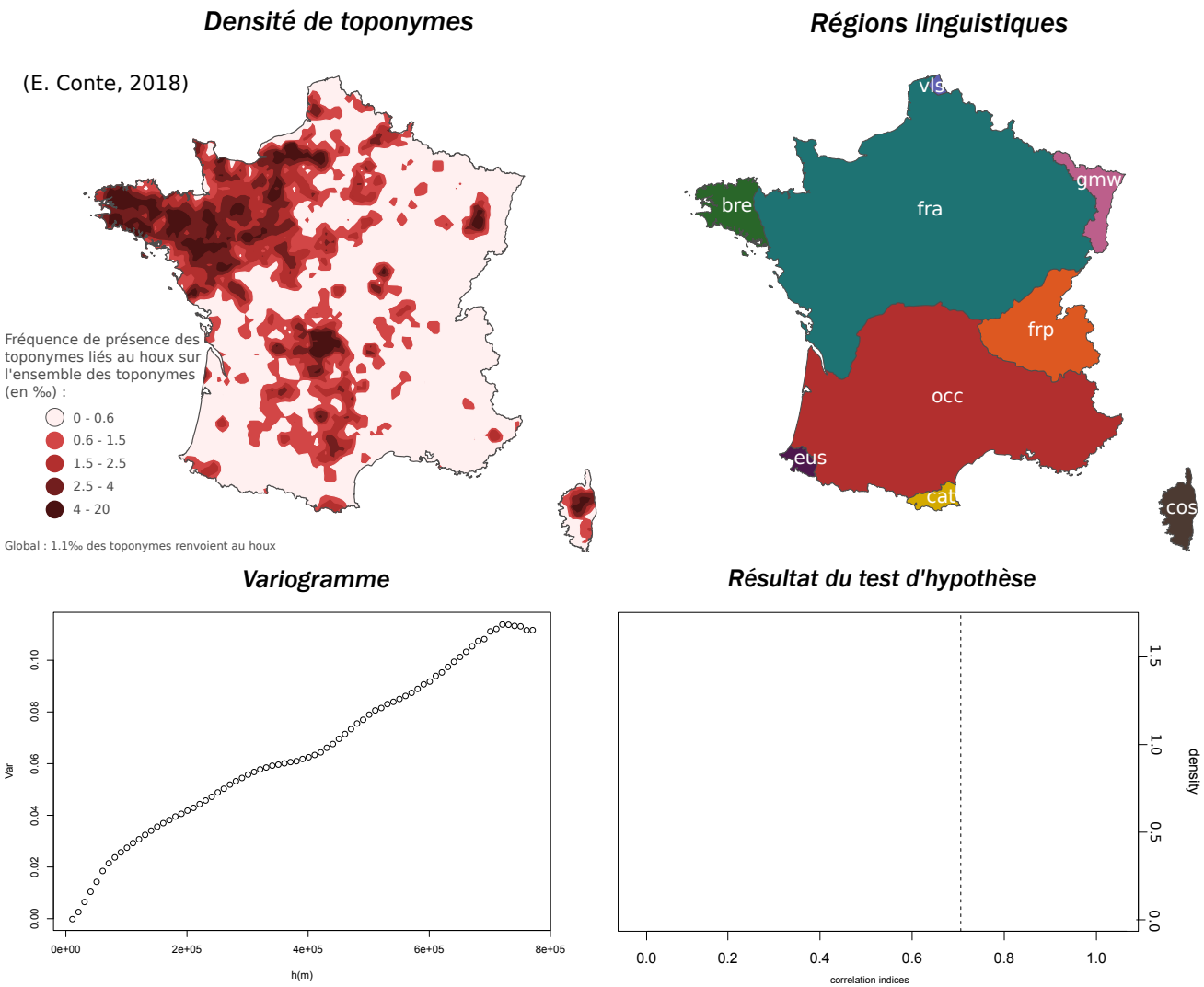

***Variogramme***




***Résultat du test d'hypothèse***



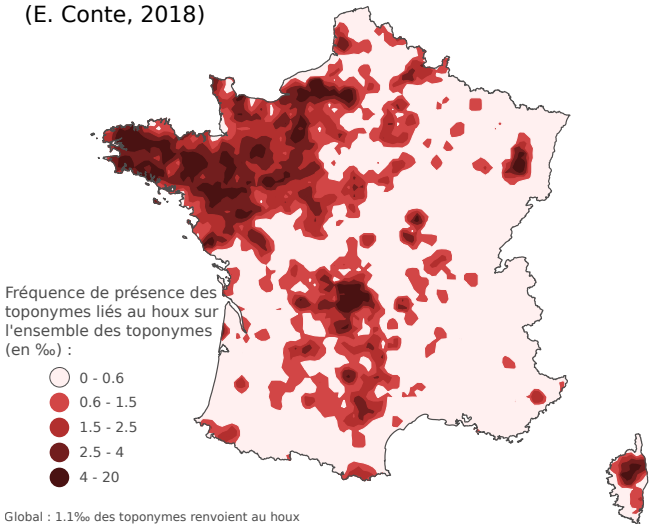
# La simulation



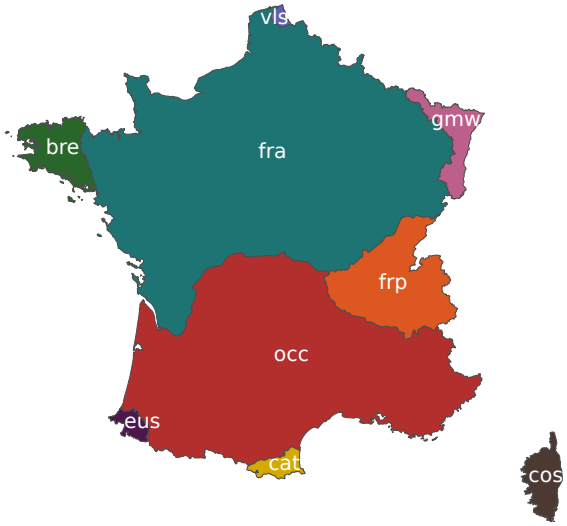
# La simulation

***Densité de toponymes***

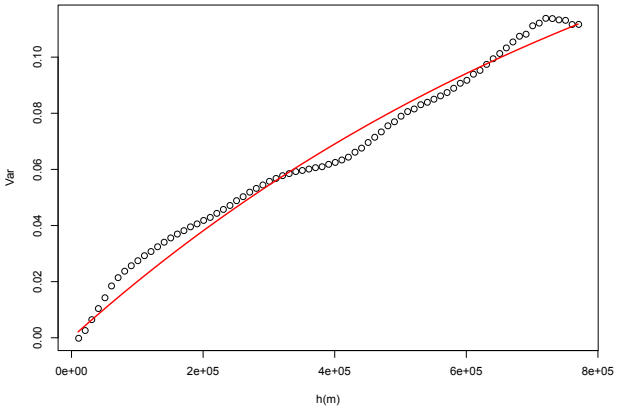
(E. Conte, 2018)



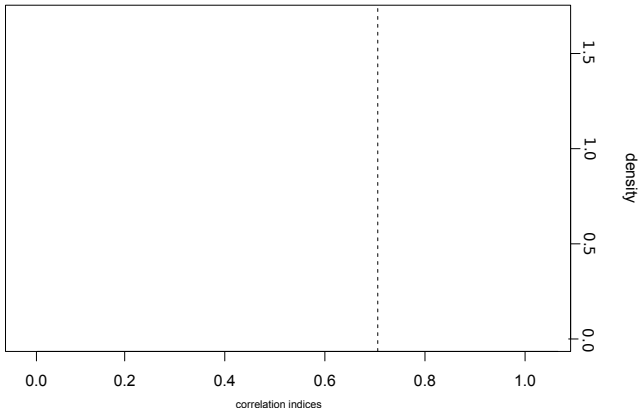

***Régions linguistiques***



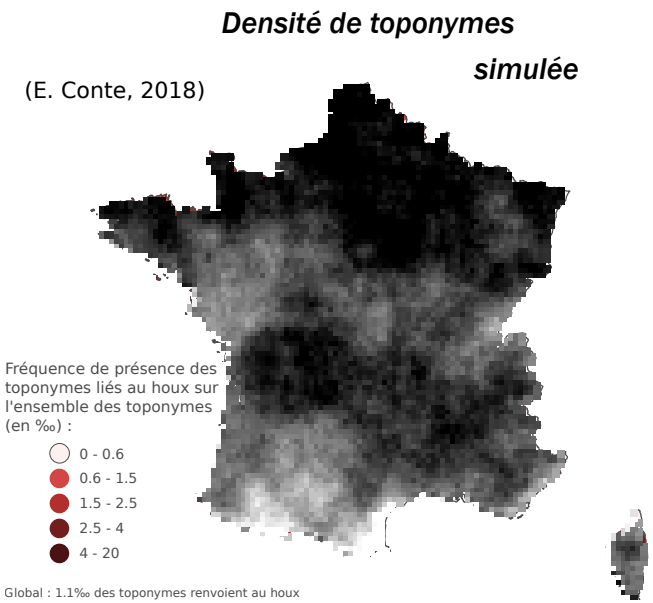

***Variogramme***

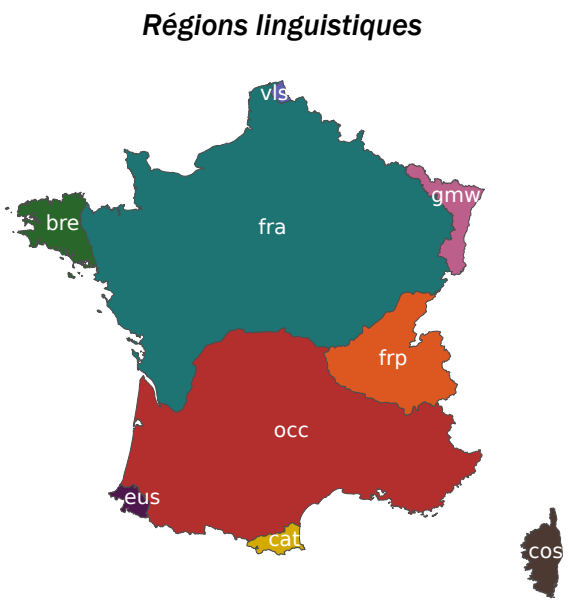


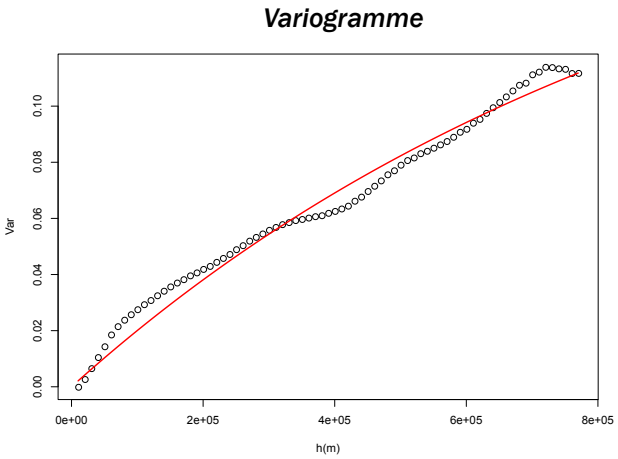

***Résultat du test d'hypothèse***

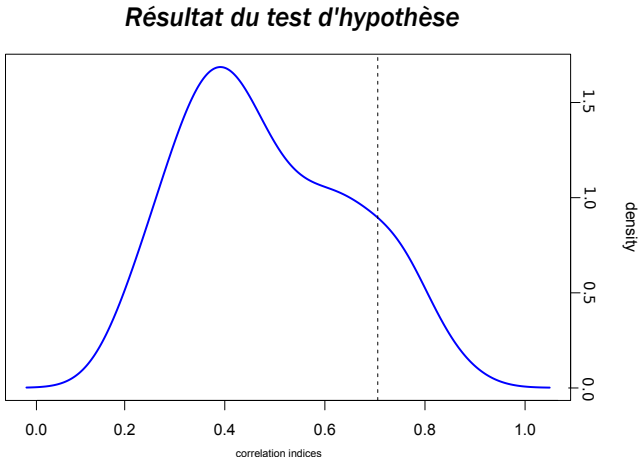


# La simulation



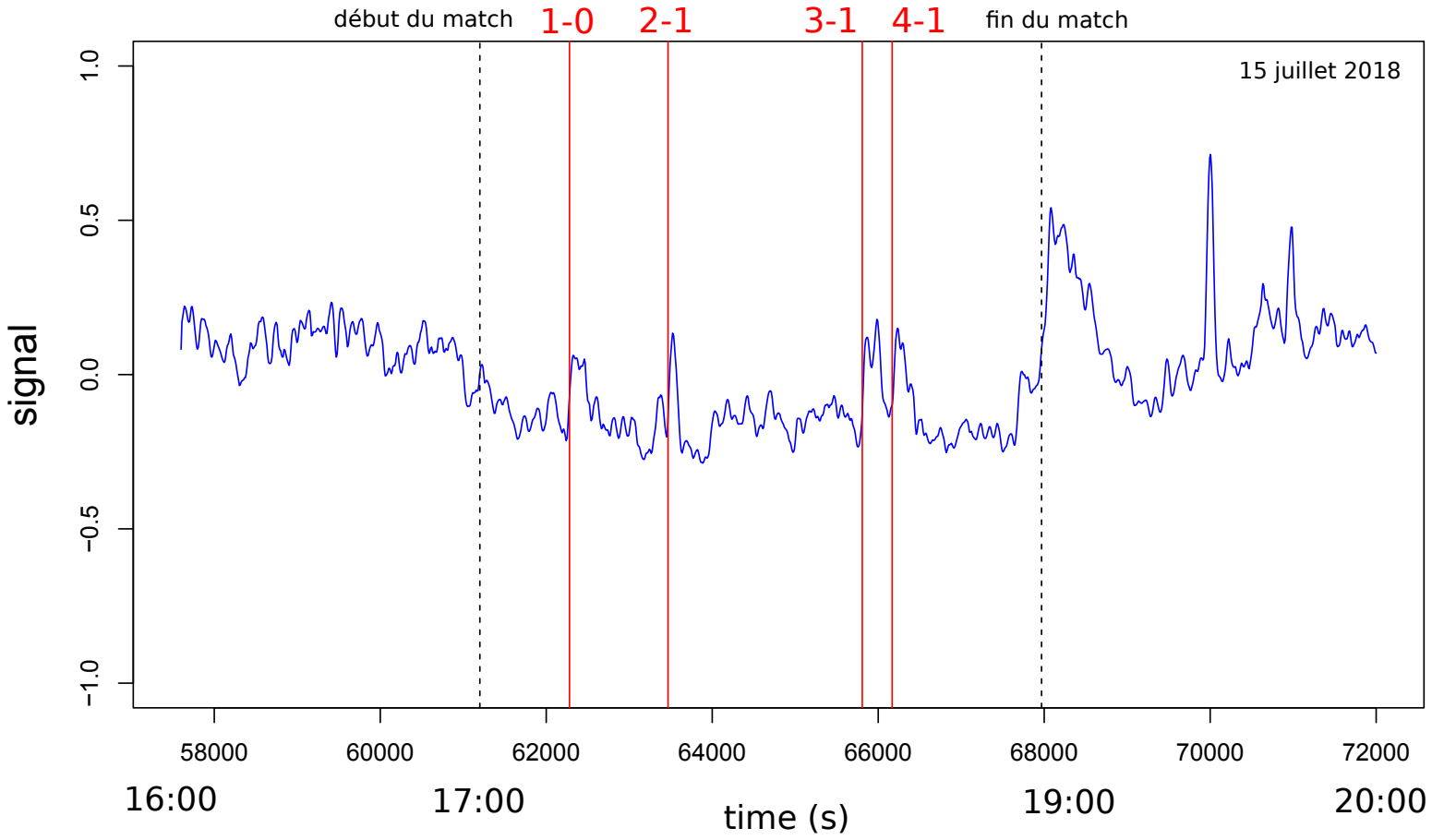



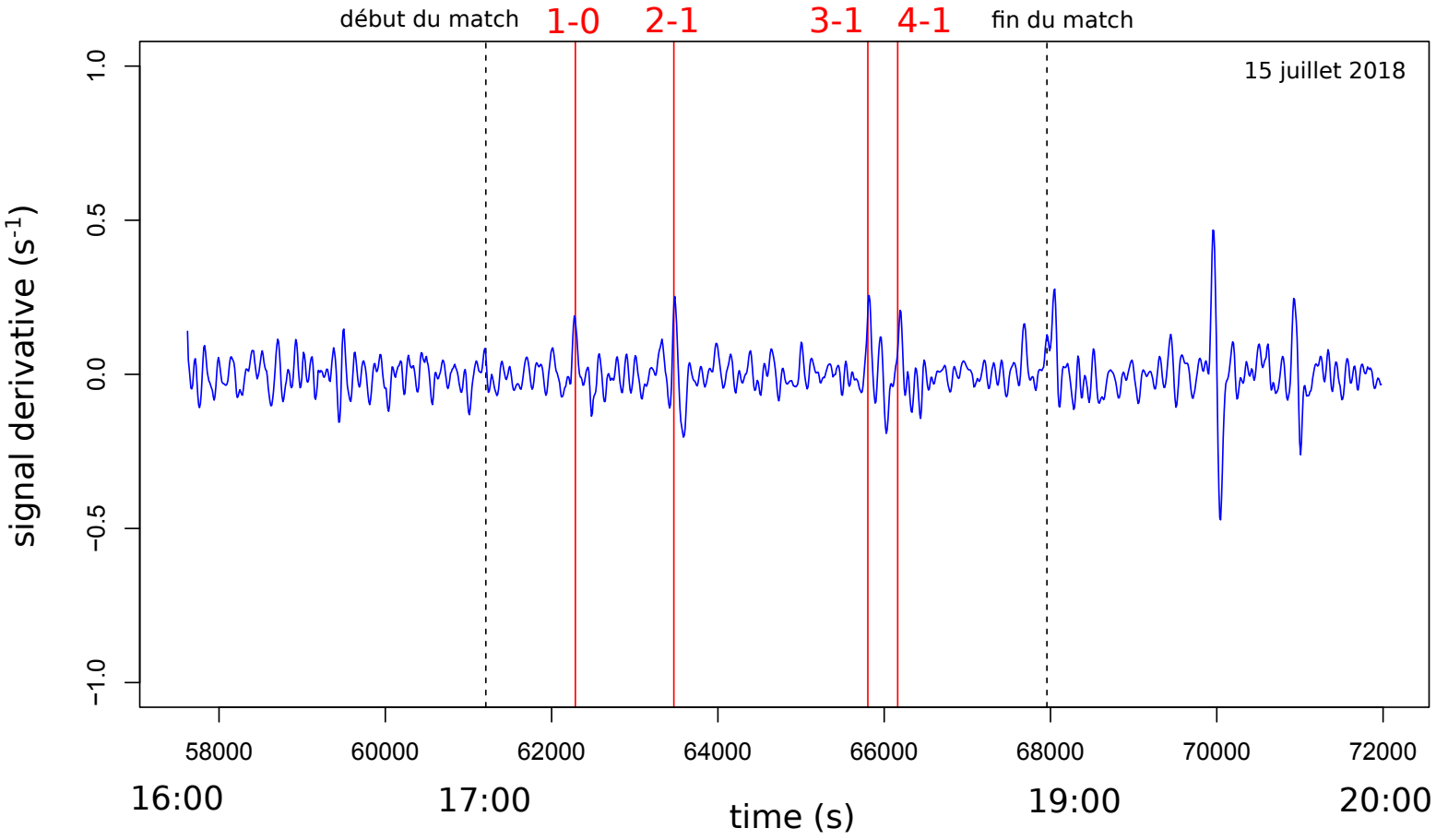
# La simulation

Relevés sismologiques pendant la finale de la Coupe du Monde de Football 2018 (d'après Jérôme Vergne) → à faire en TP



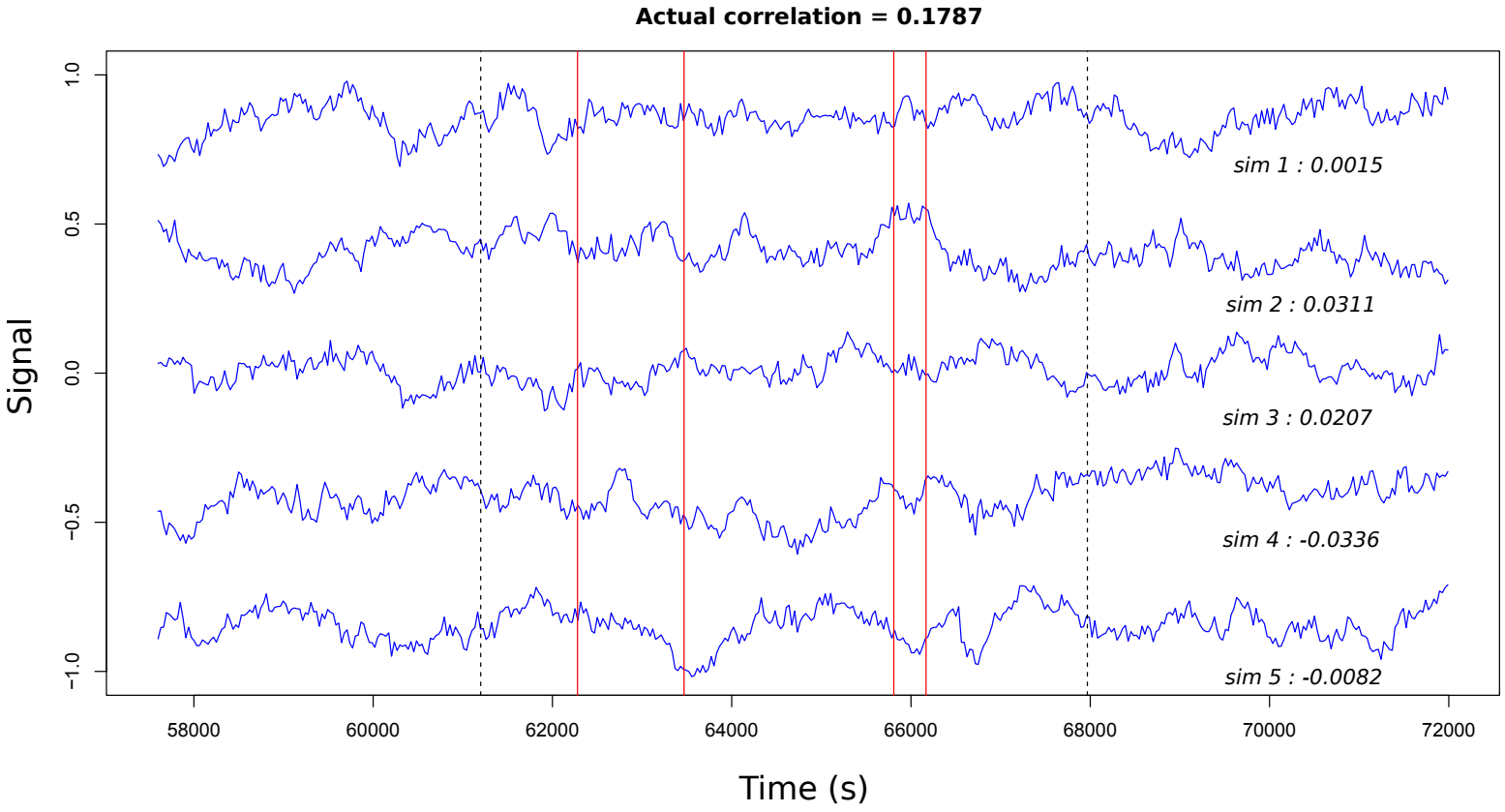
# La simulation

Relevés sismologiques pendant la finale de la Coupe du Monde de Football 2018 (d'après Jérôme Vergne) → à faire en TP



# La simulation

Relevés sismologiques pendant la finale de la Coupe du Monde de Football 2018 (d'après Jérôme Vergne) → à faire en TP



# Questions

DES QUESTIONS ?

# Références

*D. Allard, Statistiques spatiales : introduction à la géostatistique, Cours à l'Université de Montpellier, 2012*

*J. Antoni, Cours de signaux aléatoires, Master MEGA, INSA Lyon*

*M. Arnaud, X. Emery, Estimation et interpolation spatiale, méthodes déterministes et méthodes géostatistiques, éditions Hermes, 2000*

*S. Baillargeon, Le krigeage : revue de la théorie et application à l'interpolation spatiale de données de précipitations, mémoire de master de l'Université de Laval, 2005*

*Biljecki, F., Heuvelink, G. B. M., Ledoux, H., and Stoter, J. (2015). Propagation of positional error in 3D GIS : estimation of the solar irradiation of building roofs. International Journal of Geographical Information Science, vol. 29(12), December 2015, pp. 2269-2294*

# Références

*P. Bosser, Interpolation spatiale, cours de l'Ecole Nationale des Sciences Géographiques, 2012*

*P. Chauvet, Aide-mémoire de géostatistique linéaire, Edition 2008, Presses de mines*

*D. Chessel, J. Thioulouse, Analyse de données spatialisées, Cours de Biométrie et Biologie Evolutive de l'Université Lyon 1, 2003*

*X. Emery, Géostatistique linéaire, cours de l'Ecole des Mines de Paris, 2001*

*B. Flahaut, L'autocorrélation spatiale comme outil géostatistique d'identification des concentrations spatiales des accidents de la route, Revue Cybergeo, Système, modélisation & géostatistique num 185, fevrier 2001*

# Références

*C. Lajaunie, L'estimation géostatistique non linéaire, cours de l'Ecole des Mines de Paris, 1993*

*P. Sillard, Estimation par moindres carrés, collection ENSG-IGN, Editions Lavoisier, 2001*

*H. Wackernagel, Basics in Geostatistics 3. Geostatistical Monte-Carlo methods : Conditional simulation, April 2013*