GLQ3401/GLQ3651 : Deuxième partie

Cours 8 : Variance d’estimation et krigeage

POLYTECHNIQUE MONTRÉAL

UNIVERSITÉ D’INGÉNIERIE

GLQ3401/GLQ3651 : Deuxième partie

Cours 8a : Variance d’estimation

POLYTECHNIQUE MONTRÉAL

UNIVERSITÉ D’INGÉNIERIE

Objectifs

Variance d'estimation

- Comprendre la notion de variance d'estimation;
- Calculer des variances d'estimation pour une configuration d'estimation et le modèle de variogramme;
- Identifier le lien entre patron d'échantillonnage et anisotropie du variogramme.

Plan du cours

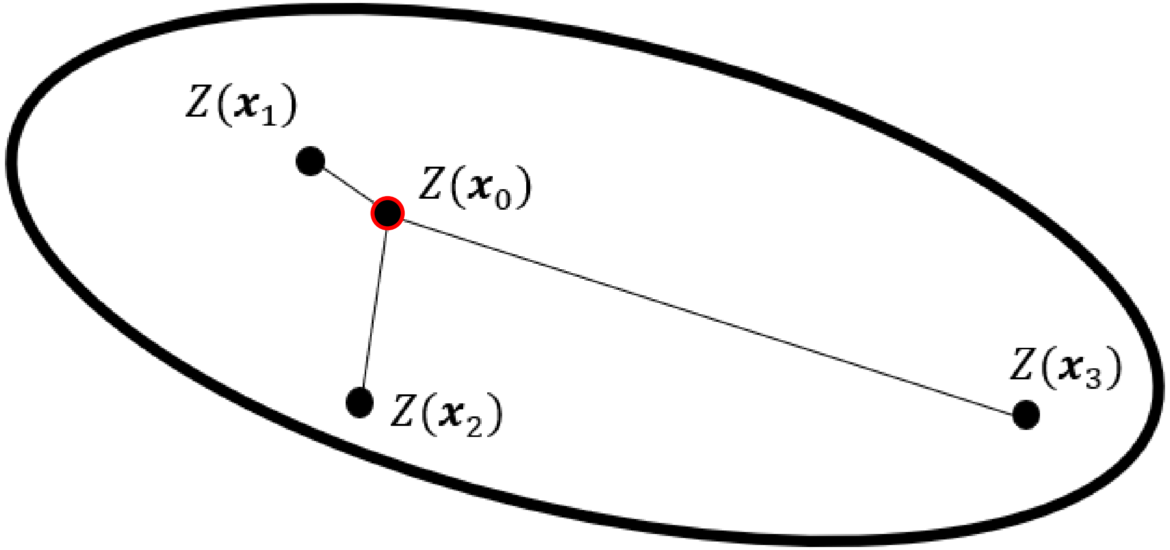
Cours : Variance d’estimation

1. Définition
2. Les trois composantes essentielles
3. Applications
4. Abaques
5. Principes de combinaison d’erreurs élémentaires

1. Définition

Mise en situation (2D)

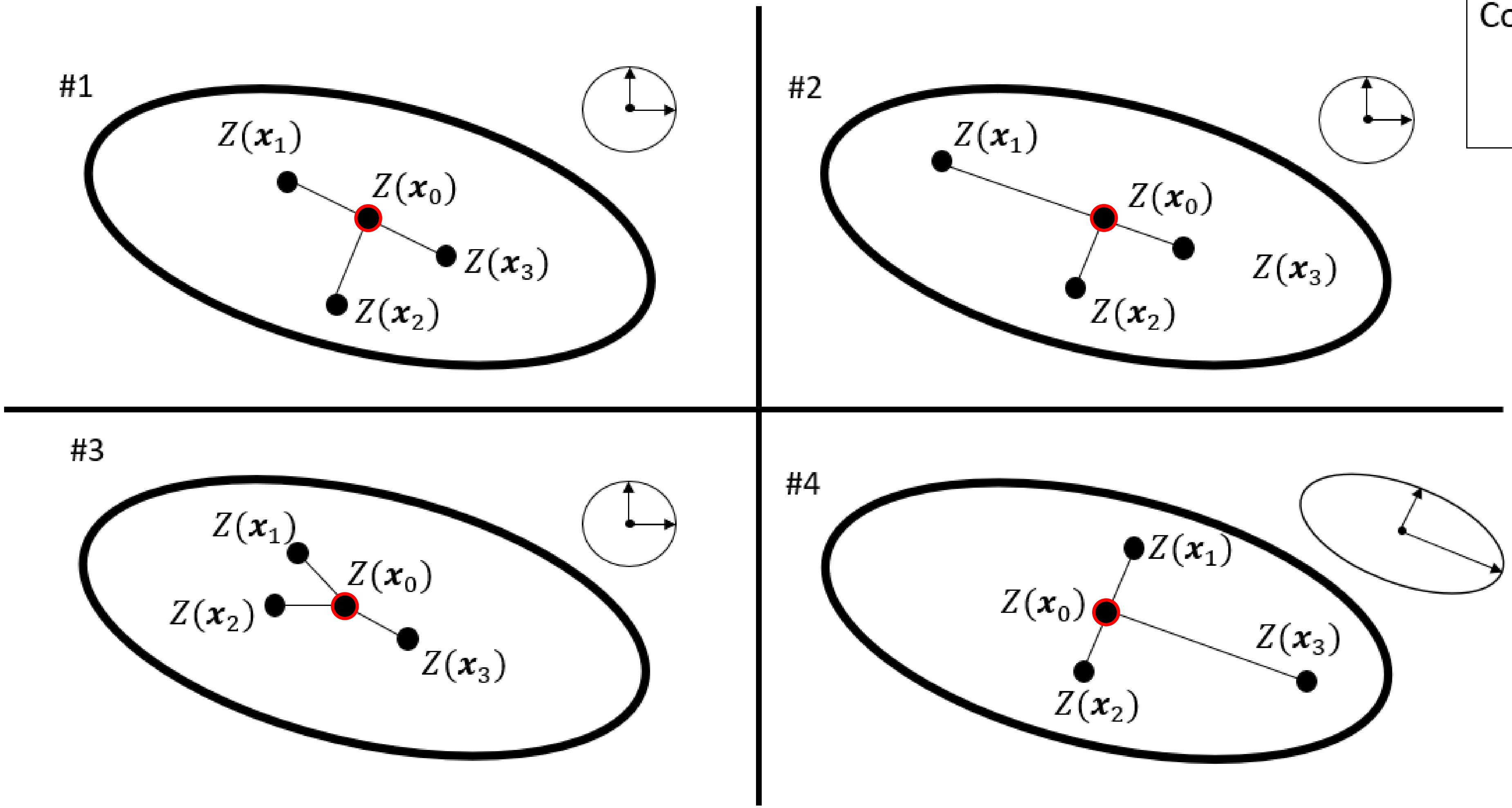
Comment déterminer l'erreur associée à l'estimé $Z^*(\boldsymbol{x}_0)$ à partir des données $Z(\boldsymbol{x}_1)$, $Z(\boldsymbol{x}_2)$ et $Z(\boldsymbol{x}_3)$?

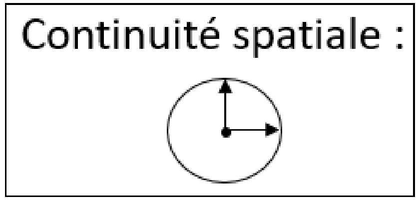


$$\mathrm{Z}_0^* = \sum_{i=1}^{3} \lambda_i Z_i \ ; \ \sum_{i=1}^{3} \lambda_i = 1$$

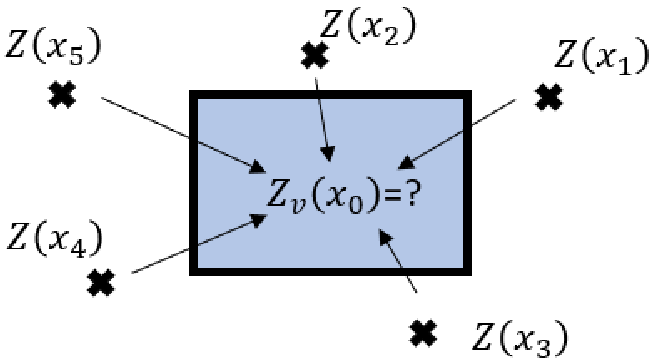
1. Définition

Mise en situation (2D)





1. Définition



Mise en situation (2D)

Estimation linéaire du bloc v :

$$Z_v^* = \sum_{i=1}^{n} \lambda_i Z_i$$

Erreur d'estimation du bloc v :

$$e = Z_v - Z_v^*$$

(erreur = réel – estimé)

Variance d'estimation:

$$Var(e) = Var(Z_v - Z_v^*)$$

2. Les trois composantes essentielles

Signification:

$$Var(e) = Var(Z_v - Z_v^*)$$

$$Var(e) = Var(Z_v) + Var(Z_v^*) - 2Cov(Z_v, Z_v^*)$$

$$\boxed{Z_v^* = \sum_{i=1}^{n} \lambda_i Z_i}$$

$$\sigma_e^2 = Var(Z_v) + \sum_i \sum_j \lambda_i \lambda_j \, Cov(Z_i, Z_j) - 2 \sum_i \lambda_i Cov(Z_i, Z_v)$$

Ce que l'on cherche à estimer est-il foncièrement variable ou non ?

Quel est le degré de redondance entre les observations ?

Les observations sont-elles bien placées par rapport à ce que l'on veut estimer ?

2. Les trois composantes essentielles

Signification:

Exercice en équipe

1) Comprendre l'importance des trois composantes essentielles de la variance d'estimation

$$\sigma_e^2 = Var(Z_v) + \sum_i \sum_j \lambda_i \lambda_j \, Cov(Z_i, Z_j) - 2 \sum_i \lambda_i Cov(Z_i, Z_v)$$

Premier terme	Deuxième terme	Troisième terme
Ce que l'on cherche à estimer est-il foncièrement variable ou non ?	Quel est le degré de redondance entre les observations ?	Les observations sont-elles bien placées par rapport à ce que l'on veut estimer ?

2. Les trois composantes essentielles

En termes de variogramme:

Si nous posons : $\sum_i \lambda_i = 1$ ⟵ Vrai pour la plupart des estimateurs

$$Var(e) = \sigma_e^2 = -\bar{\gamma}(v,v) - \sum_i \sum_j \lambda_i \lambda_j \gamma(x_i, x_j) + 2 \sum_i \lambda_i \bar{\gamma}(x_i, v)$$

Que remarquez-vous sur la structure de l'équation ?

$Var(e)$ **ne dépend que de la géométrie et du variogramme**. Indépendant des valeurs *observées* et estimées. Existe même si le variogramme ne présente pas un palier.

Note : Si la condition sur les poids n'est pas respectée, alors il n'est pas possible de calculer $var(e)$ lorsque le variogramme n'a pas de palier.

3. Applications

Utilité:

1. Comparer la précision obtenue par différentes méthodes d’interpolation
   - P. Ex. : Polygone vs Inverse de la distance vs Triangle vs Krigeage

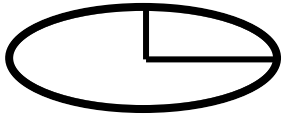
2. Déterminer le nombre d’observations requises pour atteindre une précision donnée
   - P. Ex. Planification des campagnes d’échantillonnage

3. Déterminer un patron d’échantillonnage optimal
   - P. Ex. Localisation des forages selon la continuité spatiale du gisement

3. Applications : patron d'échantillonnage

Choix de grille:

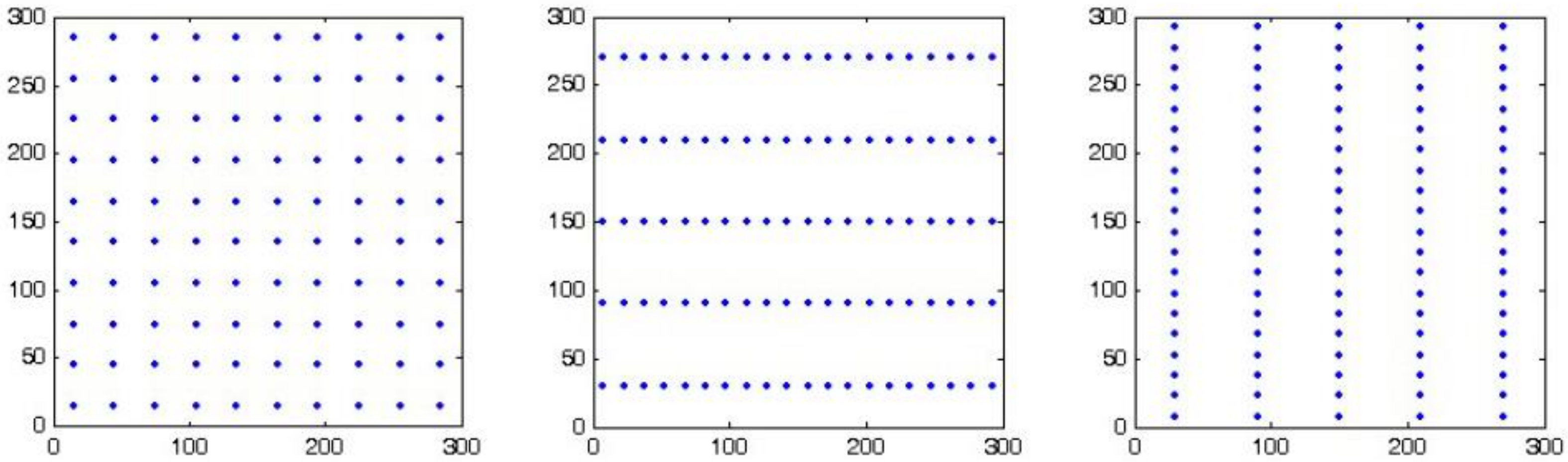
Portée variogramme



Variogramme sphérique $(a_x = 100, a_y = 25, C = 40, C_0 = 10)$

Domaine à estimer : 300×300

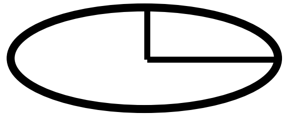
100 observations, 3 patrons différents. Lequel procure la plus grande précision pour l'estimation globale ?



3. Applications : patron d'échantillonnage

Choix de grille:

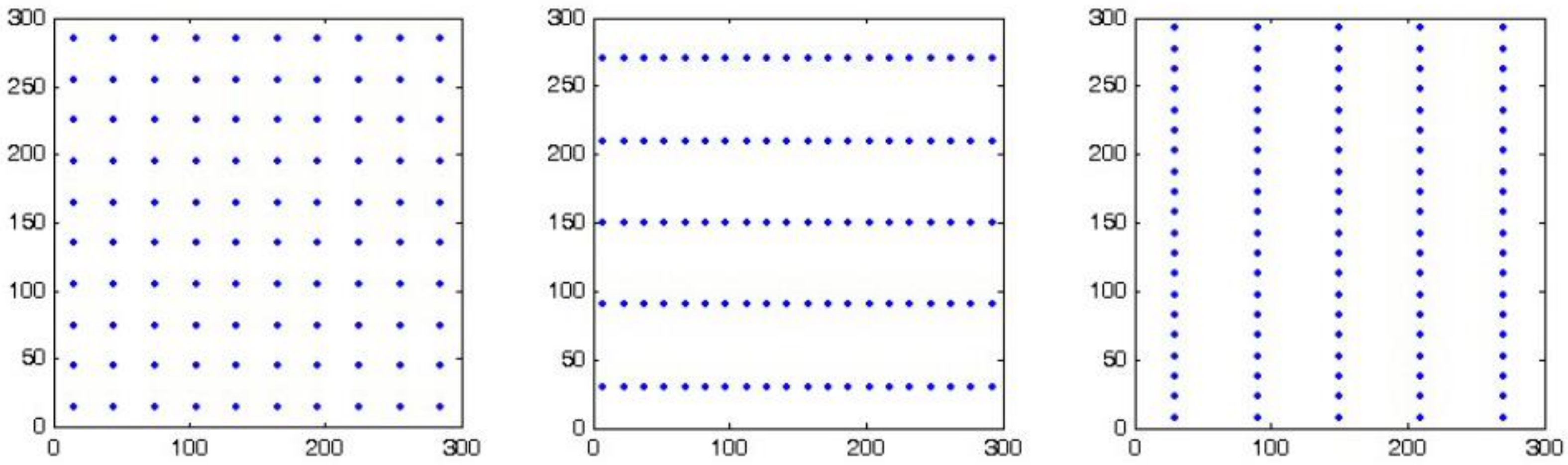
Portée variogramme



Variogramme sphérique $\left(a_x = 100, a_y = 25, C = 40, C_0 = 10\right)$

Domaine à estimer : 300×300

$Var(e) = 0.24$ $\quad$ $Var(e) = 0.36$ $\quad$ $Var(e) = 0.19$



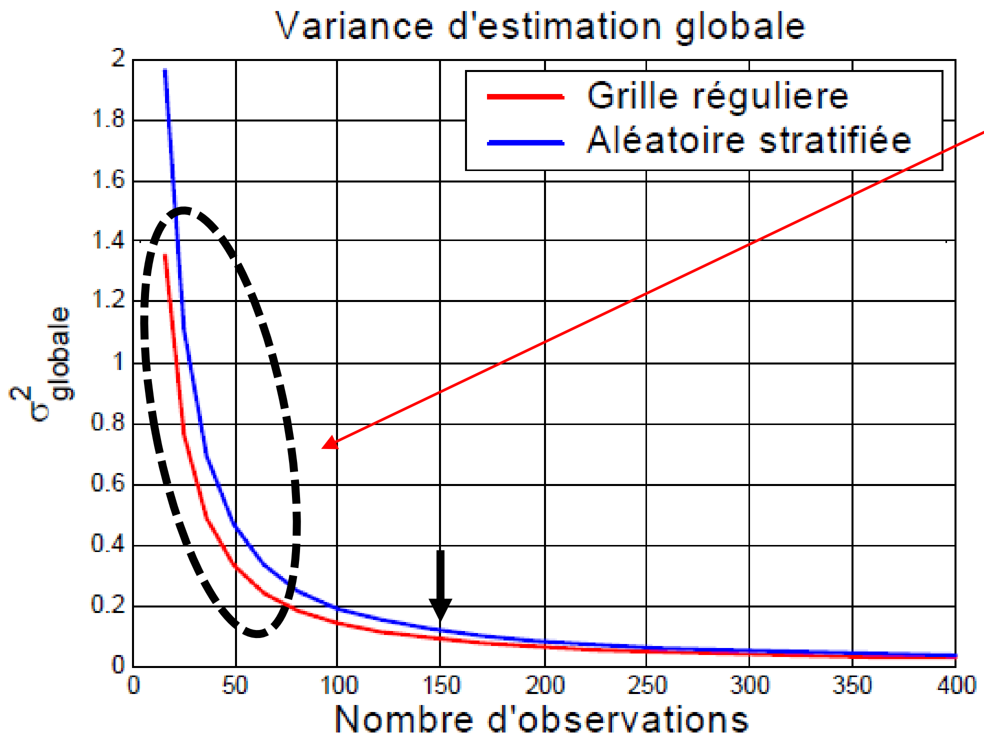
Densifier l'échantillonnage dans la direction de faible continuité spatiale

3. Applications : nombre d'observations requis et grille

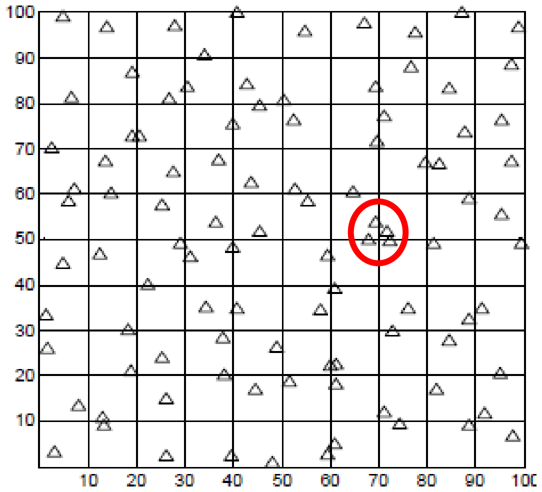
Grille aléatoire versus grille régulière:

Variogramme sphérique $(a = 100, C = 40, C_0 = 10)$

Domaine à estimer : 300×300



Que constatez-vous ?

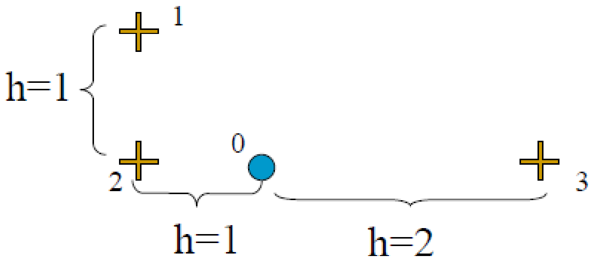


3. Applications : comparaison méthodes d’estimation

Estimateur linéaire:

On peut calculer $Var(e)$ pour tout estimateur linéaire

1) Polygone
   - $\lambda_i = 1, \lambda_j = 0 \; \forall \; j \neq i$
2) Inverse de la distance
   - $\lambda_i = \frac{1/d_i^b}{\sum_{j=1}^{n} 1/d_j^b}$
3) Triangle
   - $\lambda_i = \frac{w_i}{\sum_{j=1}^{3} w_j}, \forall \; i = 1, \dots, 3$
4) Krigeage
   - $\sum_{j=1}^{n} \lambda_j Cov(Z_i, Z_j) = Cov(Z_i, Z_v), \forall \; i$
   - $\sum_{j=1}^{n} \lambda_j = 1$



On a un outil pour juger de la précision d’une méthode versus une autre pour un gisement donné !

3. Applications : comparaison méthodes d'estimation

Exemple : Tableau récapitulatif des $var(e)$

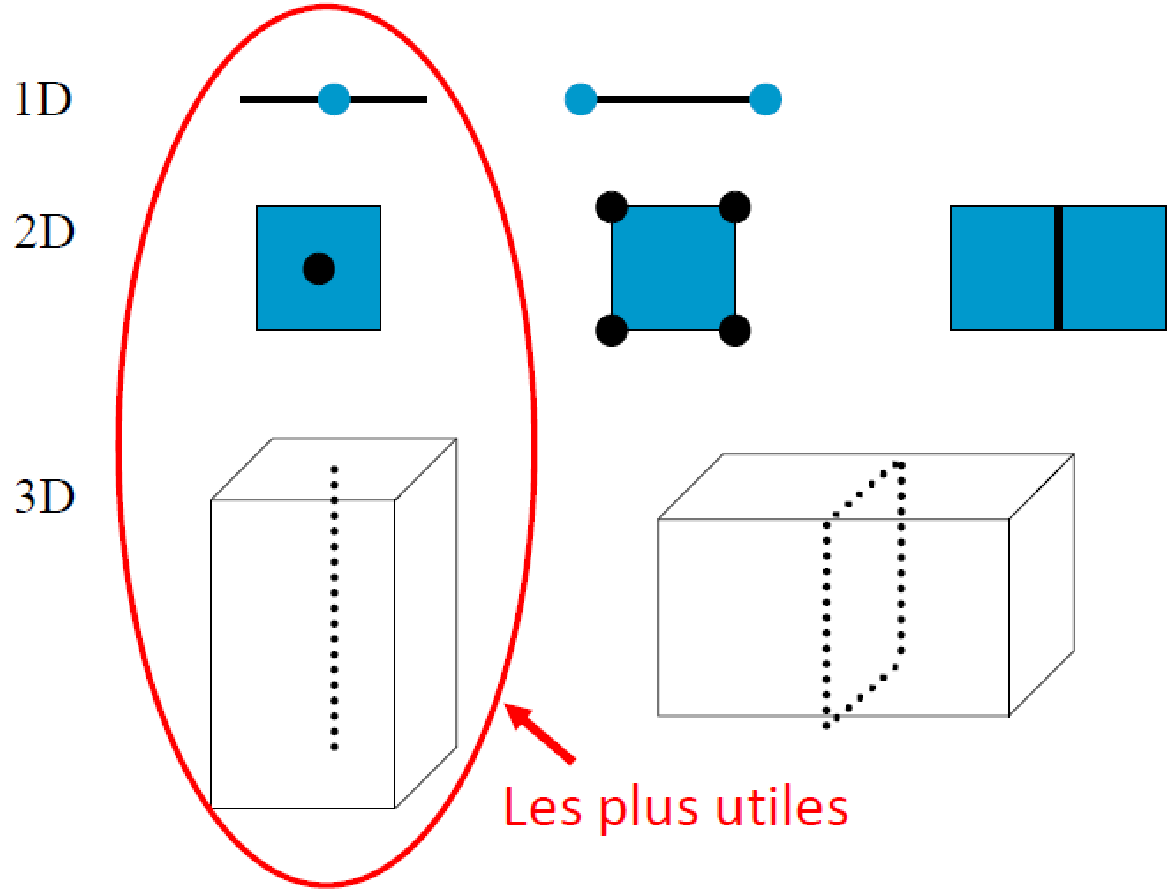
Méthode	Linéaire (pente=1,C_0=1)	Sphérique (a=10,C=1, C_0=0)	Pépite (C_0=1)	Gaussien (a=2,C=1, C_0=0)
Polygone	**4.00**	**0.30**	**2**	**0.44**
Inverse de la distance (b=1)	**2.72**	**0.21**	**1.36**	**0.36**
Inverse de la distance (b=2)	**2.88**	**0.22**	**1.43**	**0.37**
Triangle	**2.89**	**0.20**	**1.56**	**0.32**
Krigeage	**2.65**	**0.20**	**1.33**	**0.31**

* Le krigeage est la seule méthode dont les poids varient selon le variogramme

Utilisation des abaques

Seules des configurations très simples sont répertoriées

Supposent le variogramme ponctuel connu



4. Abaques

Effet de pépite

- Si le support des données est **quasi-ponctuel** :

  Remplacer C_0 par une composante sphérique de portée « epsilon » et utiliser les abaques avec cette composante

  Dans les abaques, cela revient à poser que $L/a = \infty$

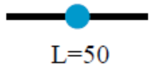
- Si le support des données **n'est pas quasi-ponctuel** :

  La contribution à Var(e) est $C_0\ /n$ où « n » est le nombre de fois que le support d'observation entre dans le support d'extension

  Par exemple, le support des données est des carottes de $5m$ dans un support d'extension de $\mathrm{L} = 25m$. Ainsi, $n = 25/5 = 5.$

4. Abaques

Exemples : variogramme sphérique ponctuel ($C_0 = 20, C = 40, a = 100$)



Étendre un point à un segment (abaque Fig.4)

$$\boldsymbol{Var(e)}$$

$$= C \times E\left(\frac{L}{a}\right) + C_0 \times E\left(\frac{L}{a}\right)$$

$$= C \times E\left(\frac{50}{100}\right) + C_0 \times E\left(\frac{\infty}{a}\right)$$

$$= 40 \times E(0.5) + 20 \times E(\infty)$$

$$= 40 \times 0.13 + 20 \times 1$$

$$= 25.2$$

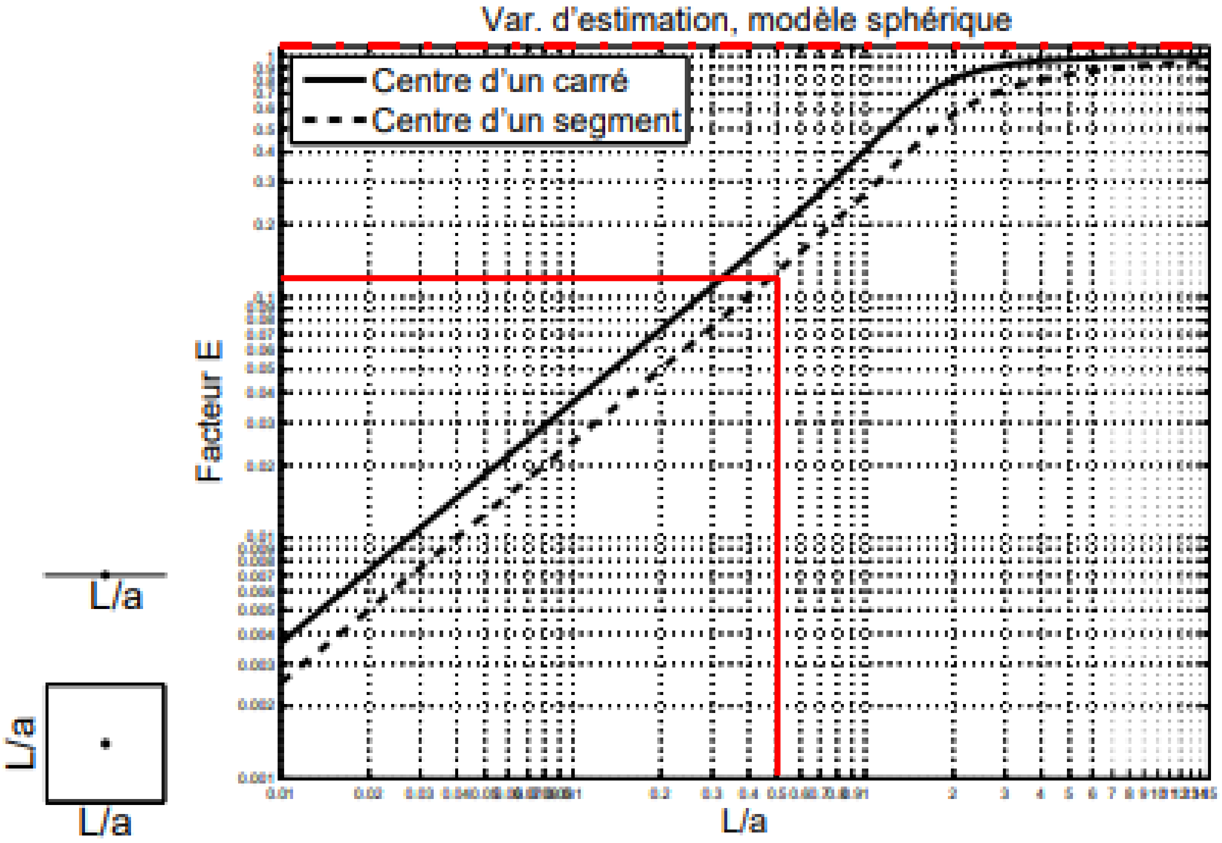
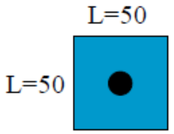


Fig. 4. Variance d'estimation: un segment ou un carré estimé par son point central

4. Abaques

Exemples : variogramme sphérique ponctuel ($C_0 = 20, C = 40, a = 100$)



Étendre un point à un carré (abaque Fig.4)

$\boldsymbol{Var(e)}$

$= C \times E\left(\frac{L}{a}\right) + C_0 \times E\left(\frac{L}{a}\right)$

$= C \times E\left(\frac{50}{100}\right) + C_0 \times E\left(\frac{\infty}{a}\right)$

$= 40 \times E(0.5) + 20 \times E(\infty)$

$= 40 \times 0.20 + 20 \times 1$

$= 28$

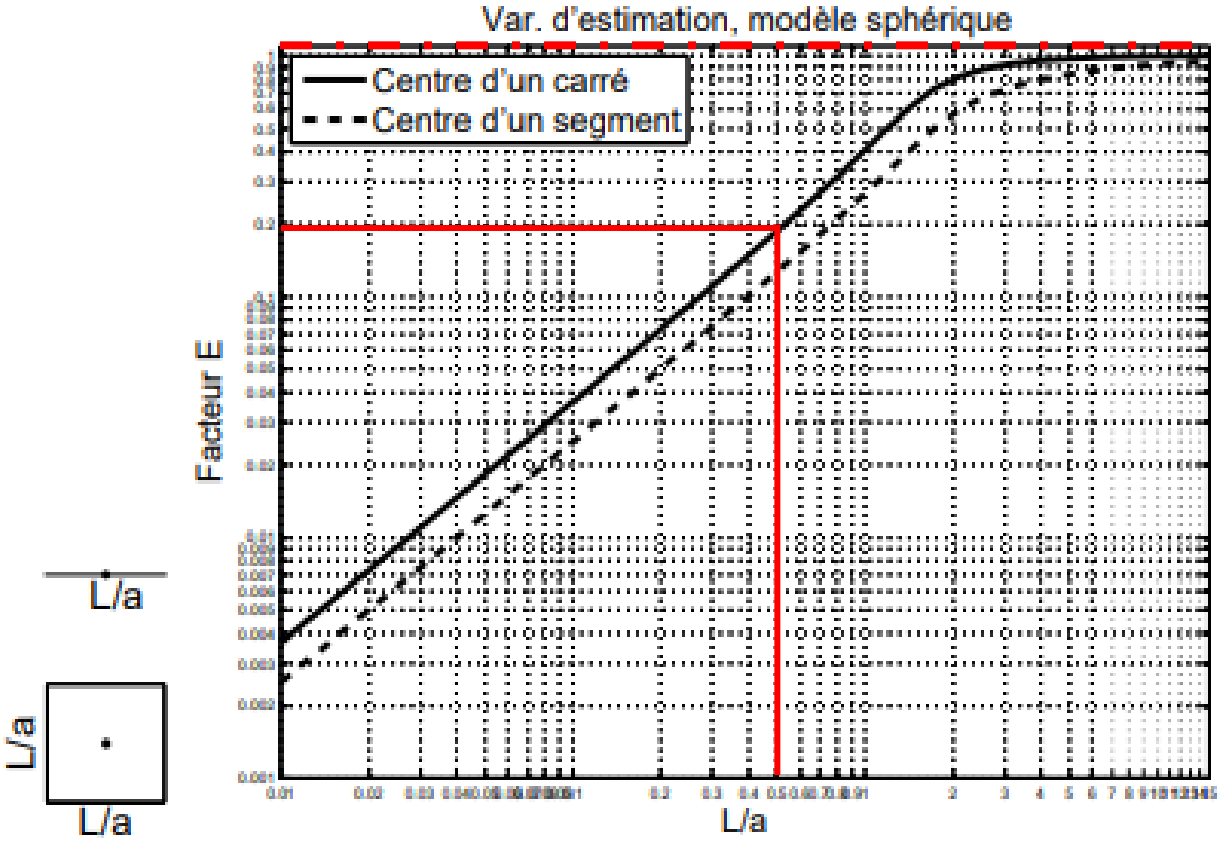
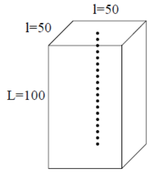


Fig. 4. Variance d'estimation: un segment ou un carré estimé par son point central

Exemples : variogramme sphérique ponctuel ($C_0 = 20, C = 40, a = 100$)



Étendre un segment à un parallélépipède à section carré (abaque Fig.8)

$\boldsymbol{Var(e)}$ Si le segment est très petit

$$= C \times E\left(\frac{L_x}{a_x}, \frac{L_z}{a_z}\right) + C_0 \times E\left(\frac{L_x}{a_x}, \frac{L_z}{a_z}\right)$$

$$= C \times E\left(\frac{50}{100}, \frac{100}{100}\right) + C_0 \times E\left(\frac{\infty}{a_x}, \frac{\infty}{a_z}\right)$$

$$= 40 \times E(0.5,1) + 20 \times E(\infty, \infty)$$

$$= 40 \times 0.047 + 20 \times 0$$

$$= 1.88$$

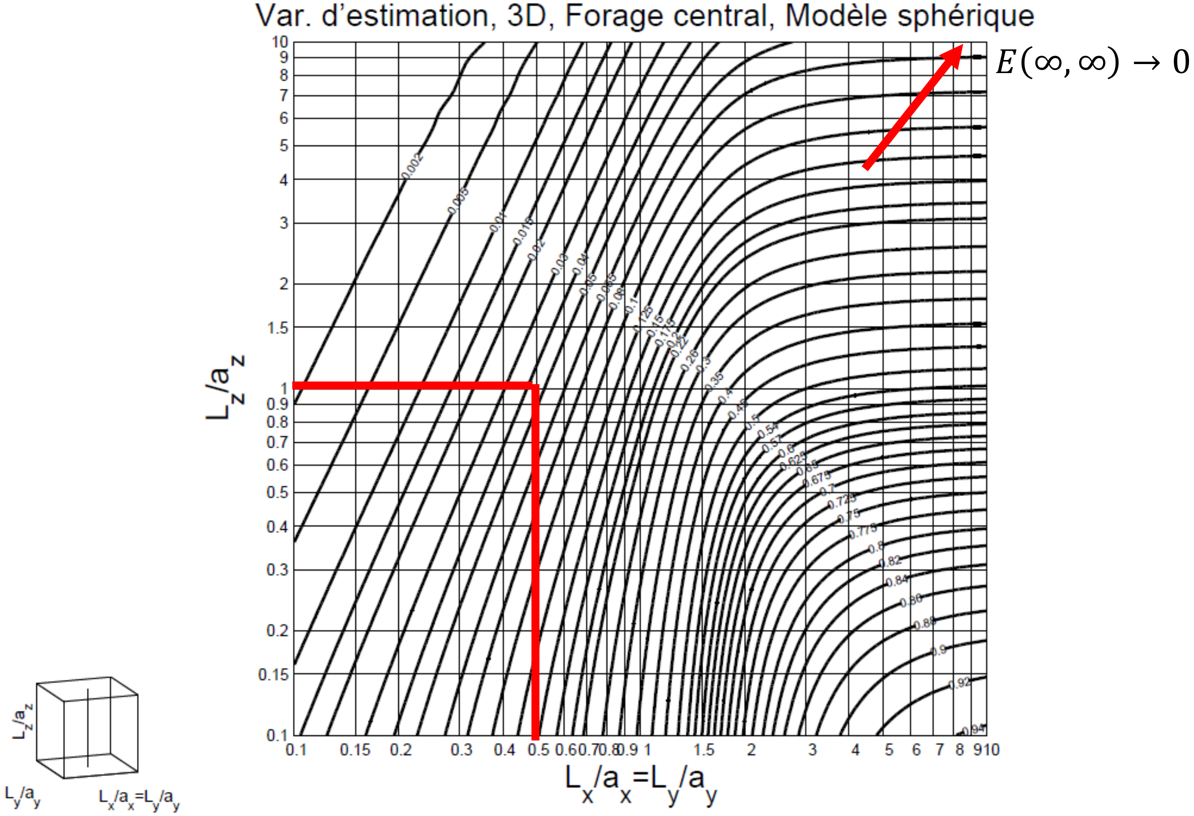
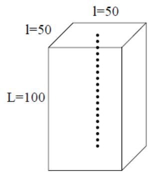


Fig. 8. Variance d'estimation: un bloc à section carrée de côté $L_x/a_x = L_y/a_y$ estimé par le segment central parallèle à L_z/a_z

4. Abaques

Exemples : variogramme sphérique ponctuel ($C_0 = 20, C = 40, a = 100$)



Étendre un segment à un parallélépipède à section carré (abaque Fig.8)

$\boldsymbol{Var(e)}$ Si carottes de 5m

$$= C \times E\left(\frac{L_x}{a_x}, \frac{L_z}{a_z}\right) + \frac{C_0}{n}$$

$$= C \times E\left(\frac{50}{100}, \frac{100}{100}\right) + \frac{C_0}{\left(\frac{L_Z}{longeur\ carotte}\right)})$$

$$= 40 \times E(0.5,1) + \frac{20}{\frac{100}{5}}$$

$$= 40 \times 0.047 + 1$$

$$= 2.88$$

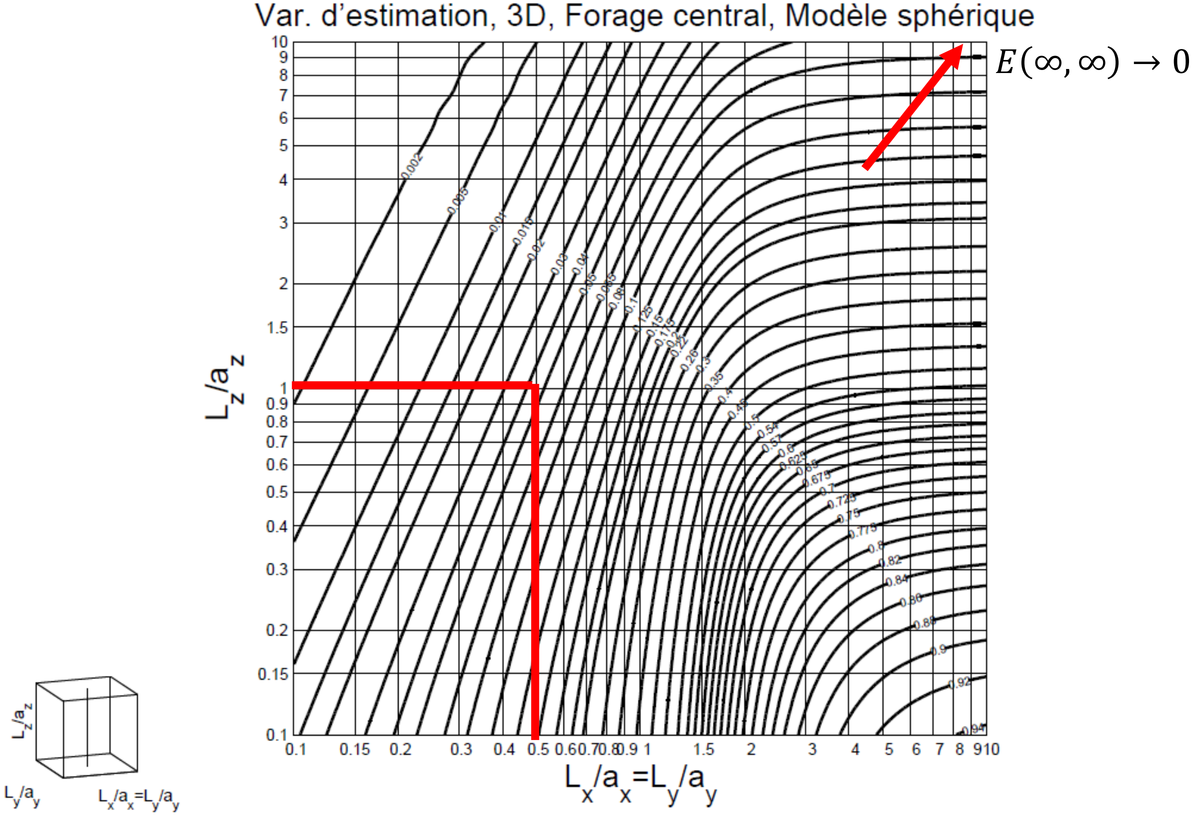
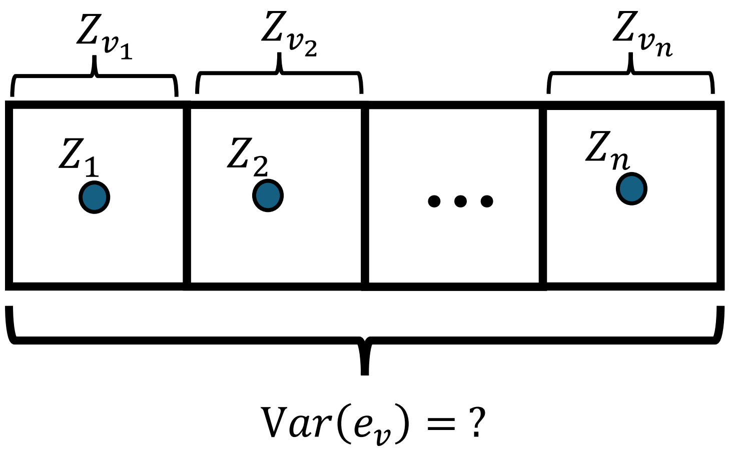


Fig. 8. Variance d'estimation: un bloc à section carrée de côté $L_x/a_x = L_y/a_y$ estimé par le segment central parallèle à L_z/a_z

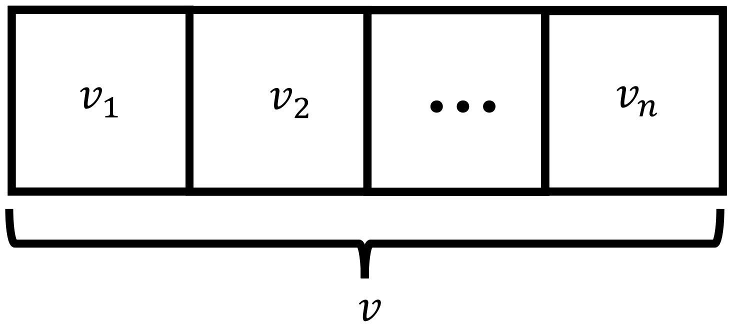
5. Principe de combinaison d'erreurs élémentaires

Contexte

Teneur :



Volume :



Réalité :

$$Z_v = \frac{1}{v}\sum_{i=1}^{n} v_i Z_{v_i}\,, \qquad v = \sum_{i=1}^{n} v_i$$

Estimateur :

$$Z_v^* = \frac{1}{v}\sum_{i=1}^{n} v_i Z_i\,,$$

Variance d'estimation du bloc v :

$$Var(e_v) = Var(Z_v - Z_v^*) = ?$$

5. Principe de combinaison d'erreurs élémentaires

Relation

$$\boxed{Var(e_v) \approx \frac{1}{v^2}\sum_{i=1}^{n} v_i^2 Var(e_i)}$$

Condition d'application

Chaque zone $(v_1, v_2, \ldots, v_n)$ est estimée uniquement avec les données contenues dans la zone

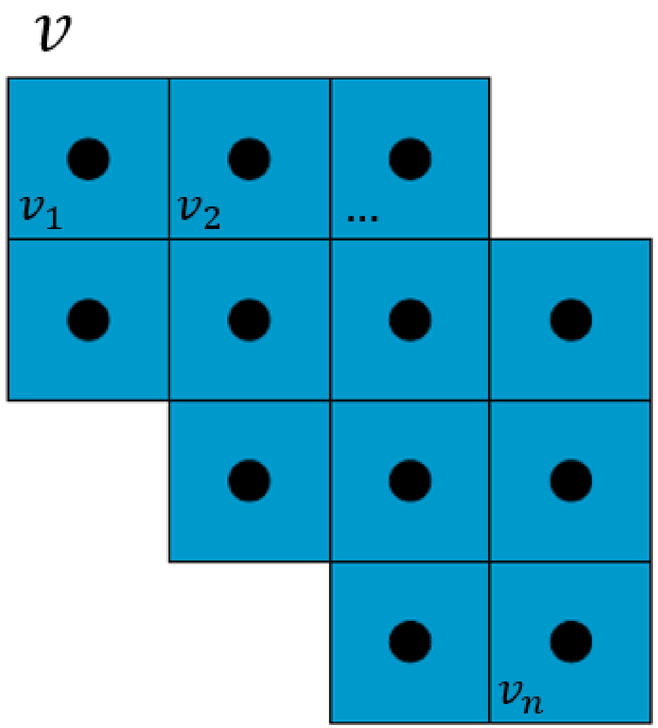
- Se généralise à autant de zones quelconques que nécessaire ;
- $Var(e_i)$ peut être obtenu soit directement soit par composition des erreurs élémentaires (échantillonnage régulier ou aléatoire stratifié) ;
- On peut toujours recourir au calcul exact de $Var(e)$, mais l'approche par composition des erreurs élémentaires allège les calculs.

5. Principe de combinaison d'erreurs élémentaires

Relation

$$\boxed{Var(e_v) \approx \frac{1}{v^2}\sum_{i=1}^{n} v_i^2 Var(e_i)}$$

Condition d'application



1. Les v_i sont tous égaux
2. Les $Var(e_i)$ sont tous égaux

Abaque

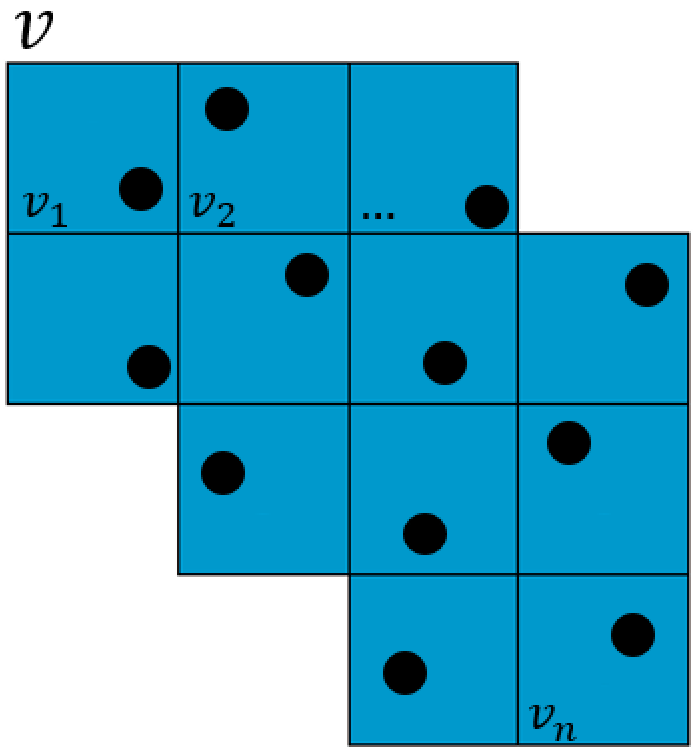
$$Var(e_v) = \frac{1}{n} Var(e_1)$$

5. Principe de combinaison d'erreurs élémentaires

Relation

$$\boxed{Var(e_v) \approx \frac{1}{v^2}\sum_{i=1}^{n} v_i^2 Var(e_i)}$$

Condition d'application



1. Les v_i sont tous égaux
2. Les $Var(e_i)$ ne sont plus égaux

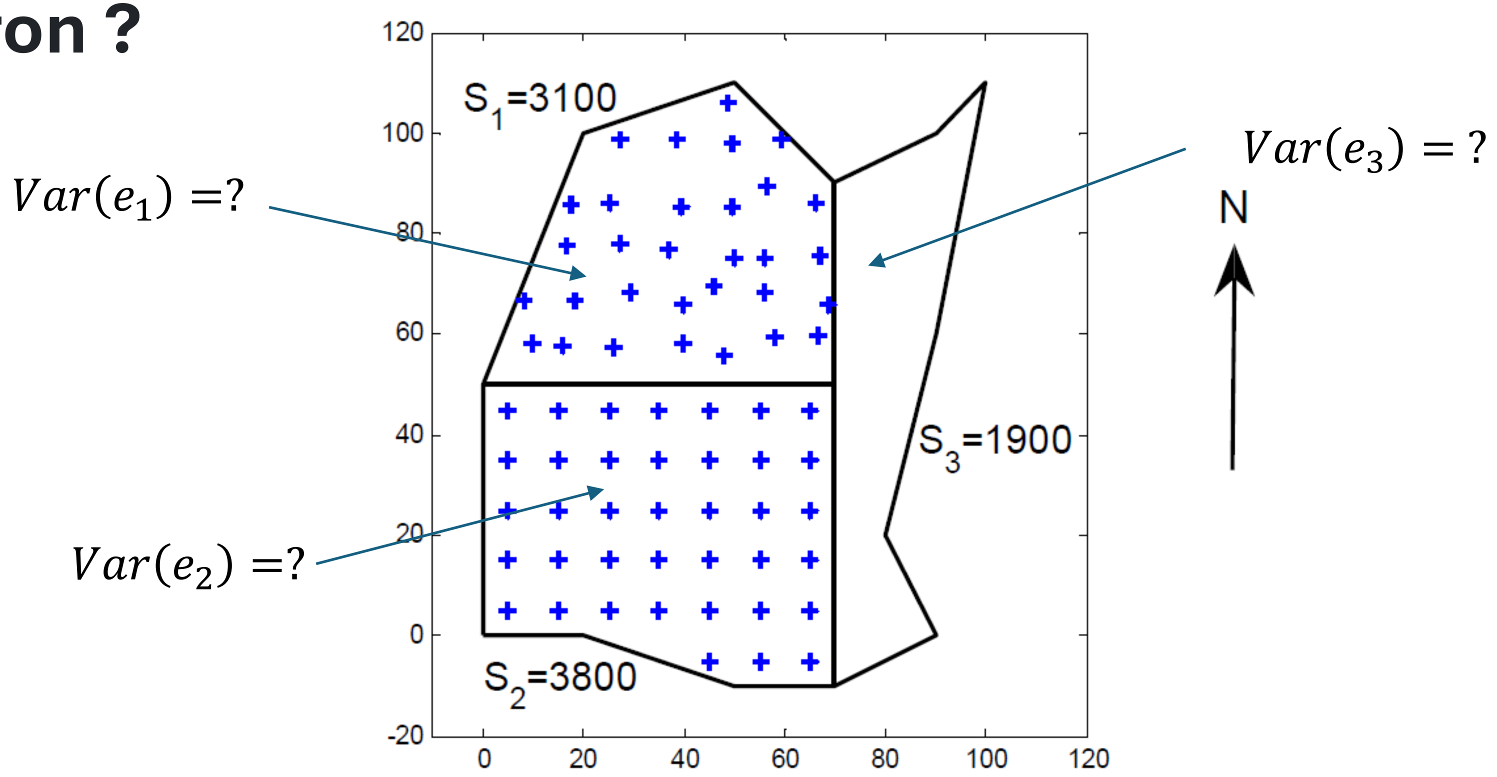
$$Var(e_v) \approx \frac{1}{n} D^2(\cdot\,|v_i)$$

5. Principe de combinaison d'erreurs élémentaires

Relation

$$\boxed{Var(e_v) \approx \frac{1}{v^2} \sum_{i=1}^{n} v_i^2 Var(e_i)}$$

Que faire avec ce type de patron ?



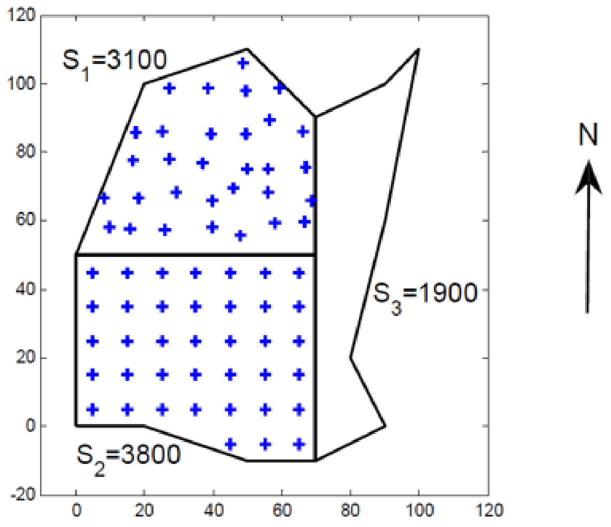
5. Principe de combinaison d'erreurs élémentaires

Relation

$$\boxed{Var(e_v) \approx \frac{1}{v^2}\sum_{i=1}^{n} v_i^2 Var(e_i)}$$

Que faire avec ce type de patron ?

$$Var(e) \approx \frac{1}{S^2}\left(S_1^2 Var(e_1) + S_2^2 Var(e_2) + S_3^2 Var(e_3)\right)$$



	Nature de la grille	Variance d'estimation
Zone 1 :	Grille aléatoire stratifiée	$Var(e_1) \approx \frac{1}{n_1} D^2(\cdot\,\vert v_1)$
Zone 2 :	Grille régulière	$Var(e_2) = \frac{1}{n_2} Var(e_{v_2})$
Zone 3 :	Estimé par krigeage	$Var(e_3) = \sigma_K^2$

GLQ3401/GLQ3651 : Deuxième partie

Cours 5b : Krigeage

POLYTECHNIQUE MONTRÉAL

UNIVERSITÉ D'INGÉNIERIE

Objectifs

Krigeage

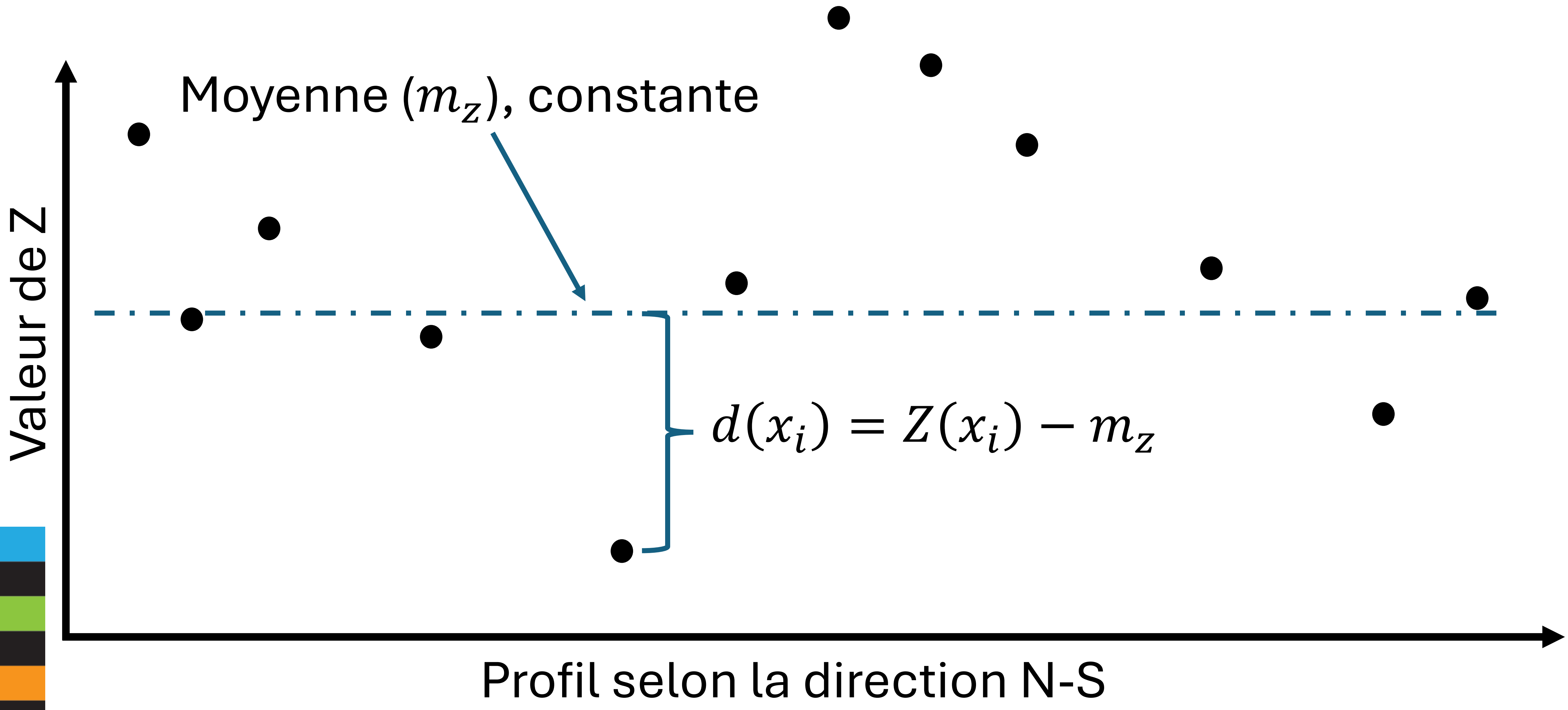
- Expliquer les différences entre krigeage simple et ordinaire;
- Être capable de dériver les équations du krigeage;
- Construire et résoudre les systèmes de krigeage simple et ordinaire, calculer la teneur estimée et la variance de krigeage;
- Expliquer les différentes propriétés du krigeage;
- Pouvoir utiliser et interpréter la validation croisée par krigeage en lien avec le modèle de variogramme.

Plan du cours

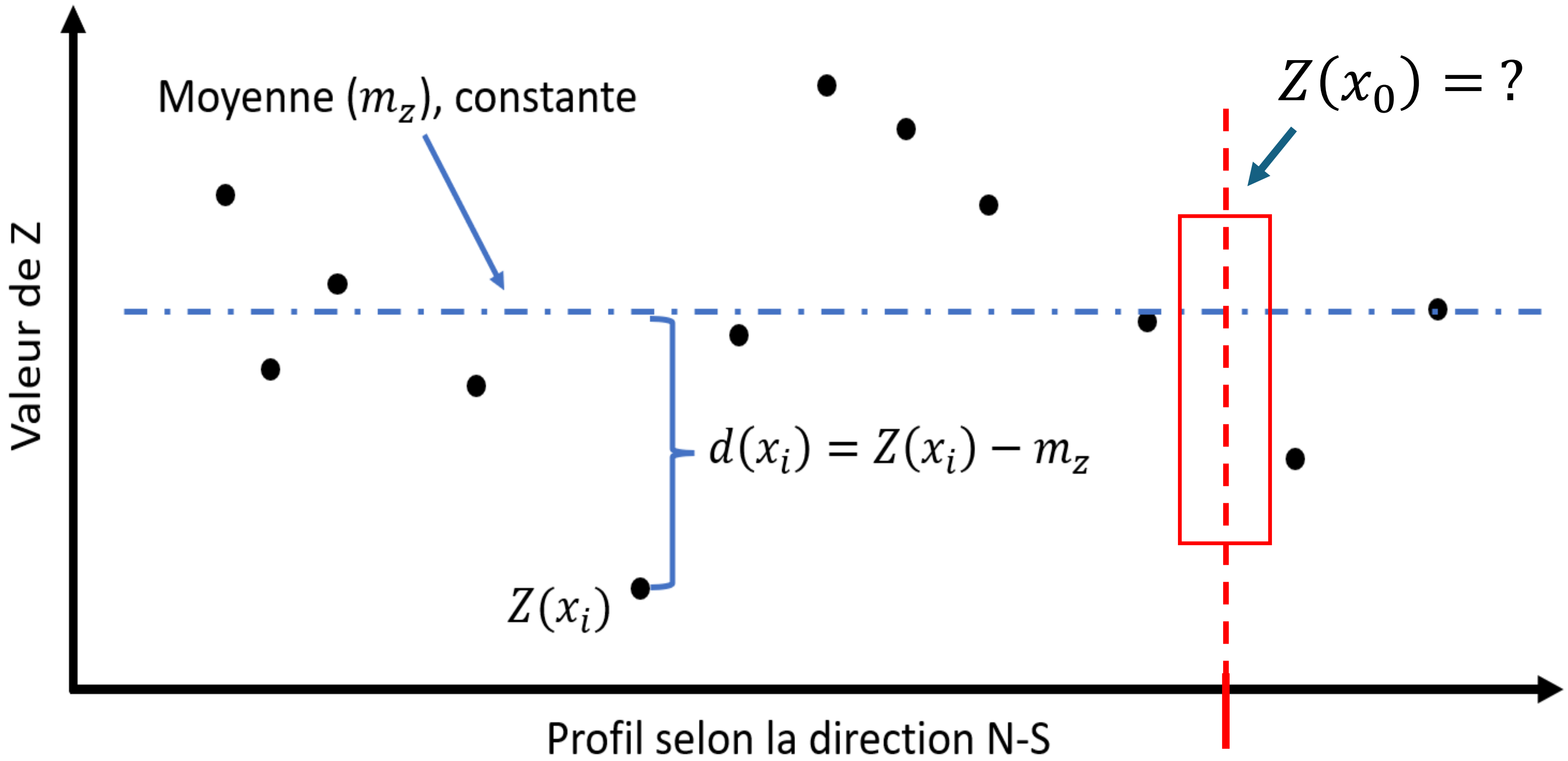
Cours : Krigeage

1. Estimation linéaire
2. Krigeage simple et krigeage ordinaire
   - Krigeage simple
   - Krigeage ordinaire
3. Interprétation
4. Exemple numérique
5. Propriétés du krigeage
6. Aspects pratiques
7. Validation croisée
8. Lien entre krigeage simple et krigeage ordinaire

1. Estimation linéaire



1. Estimation linéaire



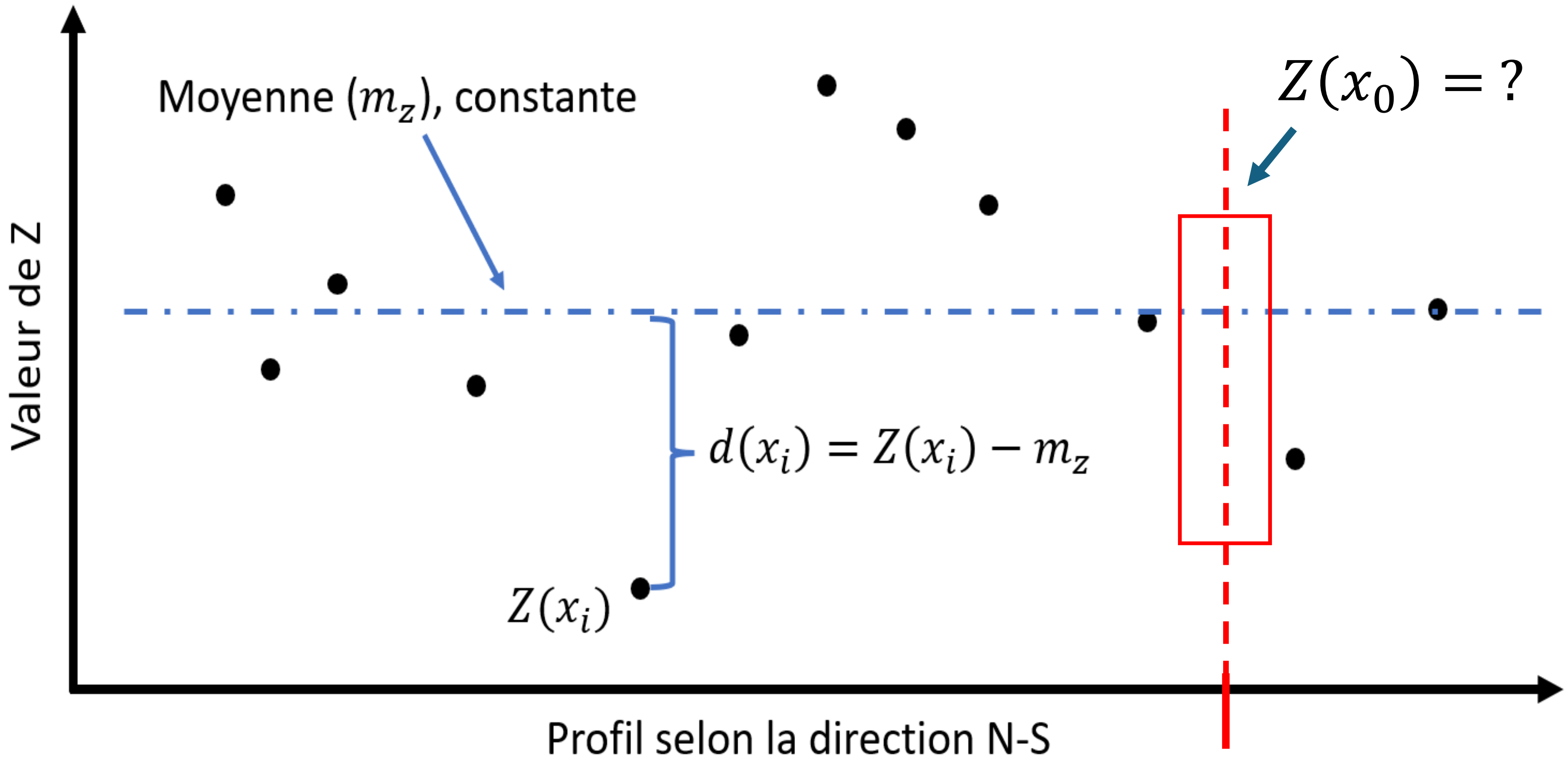
$Z(\boldsymbol{x}_0)$: à estimer
$Z(\boldsymbol{x}_i)$: Données connues

$Z^*(\boldsymbol{x}_0)$: Estimé de $Z(\boldsymbol{x}_0)$
λ_i : poids associé à $Z(\boldsymbol{x}_i)$
m_Z : Moyenne de Z

Contrainte : sans biais

$$Z^*(\boldsymbol{x}_0) = m_Z + \sum_{i=1}^{n} \lambda_i (Z(\boldsymbol{x}_i) - m_Z)$$

1. Estimation linéaire



$Z(\boldsymbol{x}_0)$: à estimer
$Z(\boldsymbol{x}_i)$: Données connues

$Z^*(\boldsymbol{x}_0)$: Estimé de $Z(\boldsymbol{x}_0)$
λ_i : poids associé à $Z(\boldsymbol{x}_i)$
m_Z : Moyenne de Z

Contrainte : sans biais

$$Z^*(\boldsymbol{x}_0) = \sum_{i=1}^{n} \lambda_i Z(\boldsymbol{x}_i) + \left(1 - \sum_{i=1}^{n} \lambda_i\right) m_Z \; ; \quad \sum_{i=1}^{n} \lambda_i = 1$$

2.1 Krigeage Simple

Lorsque la moyenne est connue :

Pour le krigeage simple, la moyenne ‘m’ est connue. Il n’y a aucune contrainte sur les poids.

$$Z_v^* = m + \sum_{i=1}^{n} \lambda_i (Z_i - m)$$

On a alors un problème de minimisation classique par dérivées partielles sur les poids.

$$\sigma_e^2 = Var(Z_v) + \sum_i \sum_j \lambda_i \lambda_j \, Cov(Z_i, Z_j) - 2 \sum_i \lambda_i Cov(Z_i, Z_v)$$

$$\frac{\partial \sigma_e^2}{\partial \lambda_i} = 0$$

2.1 Krigeage Simple

Dérivées partielles :

$$\sigma_e^2 = Var(Z_v) + \sum_i \sum_j \lambda_i \lambda_j \, Cov(Z_i, Z_j) - 2 \sum_i \lambda_i Cov(Z_i, Z_v)$$

Système d'équations à n inconnus :

$$\frac{\partial \sigma_e^2}{\partial \lambda_i} = 0 \rightarrow \quad \sum_{j=1}^{n} \lambda_j Cov(Z_i, Z_j) = Cov(Z_i, Z_v) \ , \qquad \forall i = 1, \dots, n$$

Variance du krigeage simple :

$$\sigma_{KS}^2 = Var(Z_v) - \sum_{i=1}^{n} \lambda_i Cov(Z_i, Z_v)$$

Estimée par krigeage simple :

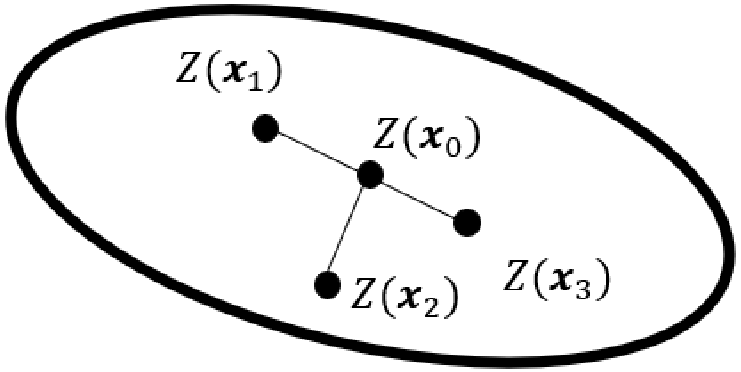
$$Z_v^* = m + \sum_{i=1}^{n} \lambda_i (Z_i - m)$$

2.1 Krigeage Simple

Forme matricielle :

$$K_S \lambda_S = k_S \rightarrow \lambda_S = K_S^{-1} k_S$$

$$\sigma_{KS}^2 = \sigma_v^2 - \lambda_s' k_s$$



$$\boxed{\begin{aligned} &cov(Z_n, Z_n) = var(Z_n) = \sigma^2 \\ &cov(Z_n, Z_m) = cov(Z_m, Z_n) \end{aligned}}$$

$$\underbrace{\overset{K_S}{\begin{bmatrix} Cov(Z_1,Z_1) & Cov(Z_1,Z_2) & Cov(Z_1,Z_3) \\ Cov(Z_2,Z_1) & Cov(Z_2,Z_2) & Cov(Z_2,Z_3) \\ Cov(Z_3,Z_1) & Cov(Z_3,Z_2) & Cov(Z_3,Z_3) \end{bmatrix}}}_{\textbf{Matrice de redondance}} \overset{\lambda_S}{\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix}} = \underbrace{\overset{k_S}{\begin{bmatrix} Cov(Z_1,Z_0) \\ Cov(Z_2,Z_0) \\ Cov(Z_3,Z_0) \end{bmatrix}}}_{\textbf{Vecteur de proximité}}$$

2.2 Krigeage Ordinaire

Lorsque la moyenne est inconnue :

Pour le krigeage ordinaire, la moyenne ‘m’ n’est pas connue. Il faut imposer une contrainte sur les poids pour obtenir un estimateur sans biais.

$$Z_v^* = \sum_{i=1}^{n} \lambda_i Z_i \qquad ; \qquad \sum_{i=1}^{n} \lambda_i = 1$$

On a alors un problème de minimisation sous contrainte. Méthode de Lagrange.

$$L(\lambda, \mu) = \sigma_e^2 + 2\mu \left(\sum_{i=1}^{n} \lambda_i - 1 \right)$$

2.2 Krigeage Ordinaire

Méthode de Lagrange :

$$L(\lambda,\mu) = \sigma_e^2 + 2\mu\left(\sum_{i=1}^{n} \lambda_i - 1\right)$$

$$L(\lambda,\mu) = Var(Z_v) + \sum_i \sum_j \lambda_i \lambda_j \, Cov\left(Z_i, Z_j\right) - 2\sum_i \lambda_i Cov(Z_i, Z_v) + 2\mu\left(\sum_{i=1}^{n} \lambda_i - 1\right)$$

<u>Système d'équations à $n+1$ inconnus</u>

$$\frac{\partial L(\lambda,\mu)}{\partial \lambda_i} = 0 \rightarrow \qquad \sum_{j=1}^{n} \lambda_j Cov\left(Z_i, Z_j\right) + \mu = Cov(Z_i, Z_v) \ , \qquad \forall i = 1, \dots, n$$

$$\frac{\partial L(\lambda,\mu)}{\partial \mu} = 0 \rightarrow \qquad \sum_{i=1}^{n} \lambda_i = 1$$

2.2 Krigeage Ordinaire

Estimateur et variance :

Variance du krigeage ordinaire :

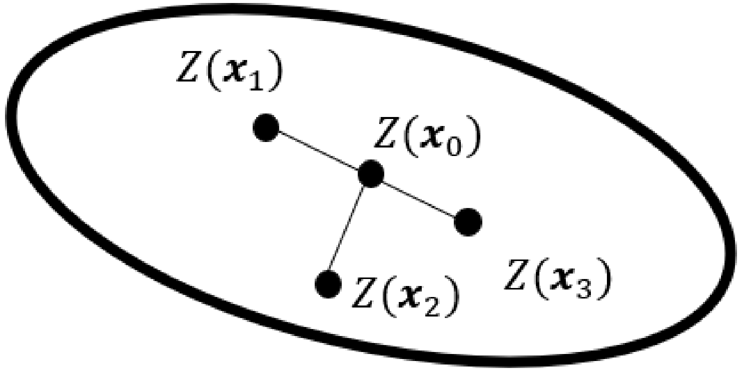
$$\sigma^2_{KO} = Var(Z_v) - \sum_{i=1}^{n} \lambda_i Cov(Z_i, Z_v) - \mu$$

Estimée par krigeage ordinaire :

$$Z_v^* = \sum_{i=1}^{n} \lambda_i \, Z_i$$

2.2 Krigeage Ordinaire

Forme matricielle : $$K_O \lambda_O = k_O \rightarrow \lambda_O = K_O^{-1} k_O$$



$$\sigma_{KO}^2 = \sigma_v^2 - \lambda_O k_O$$

$$\underbrace{\overset{\boldsymbol{K_O}}{\begin{bmatrix} \sigma^2 & cov(Z_1, Z_2) & \dots & cov(Z_1, Z_n) & 1 \\ cov(Z_2, Z_1) & \sigma^2 & \dots & cov(Z_2, Z_n) & 1 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ cov(Z_n, Z_1) & cov(Z_n, Z_2) & \dots & \sigma^2 & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix}}}_{\textbf{Matrice de redondance}} \overset{\boldsymbol{\lambda_O}}{\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \dots \\ \dots \\ \lambda_n \\ \mu \end{bmatrix}} = \underbrace{\overset{\boldsymbol{k_O}}{\begin{bmatrix} cov(Z_1, Z) \\ cov(Z_2, Z) \\ \dots \\ \dots \\ cov(Z_n, Z) \\ 1 \end{bmatrix}}}_{\textbf{Vecteur de proximité}}$$

3. Interprétation

Observations:

- Le krigeage minimise la variance d'estimation théorique calculée à partir du variogramme.

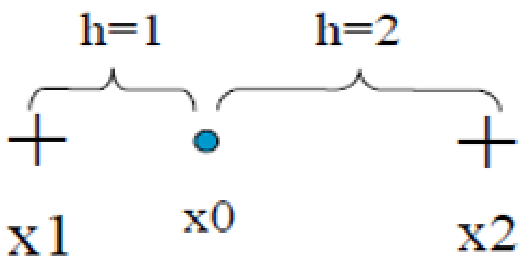
- Le krigeage est-il plus juste que tout autre estimateur linéaire ?
  - Oui, en moyenne (c.-à-d. sur un grand nombre de valeurs estimées), lorsque:
    - Hypothèse de stationnarité est valide
    - On a le bon modèle de variogramme
  - Pour un bloc particulier ou un point, on ne peut rien affirmer.

- En pratique, dans la plupart des cas, le krigeage est en moyenne au moins aussi juste que les autres estimateurs.

- **Les poids du krigeage peuvent être positifs ou négatifs.**

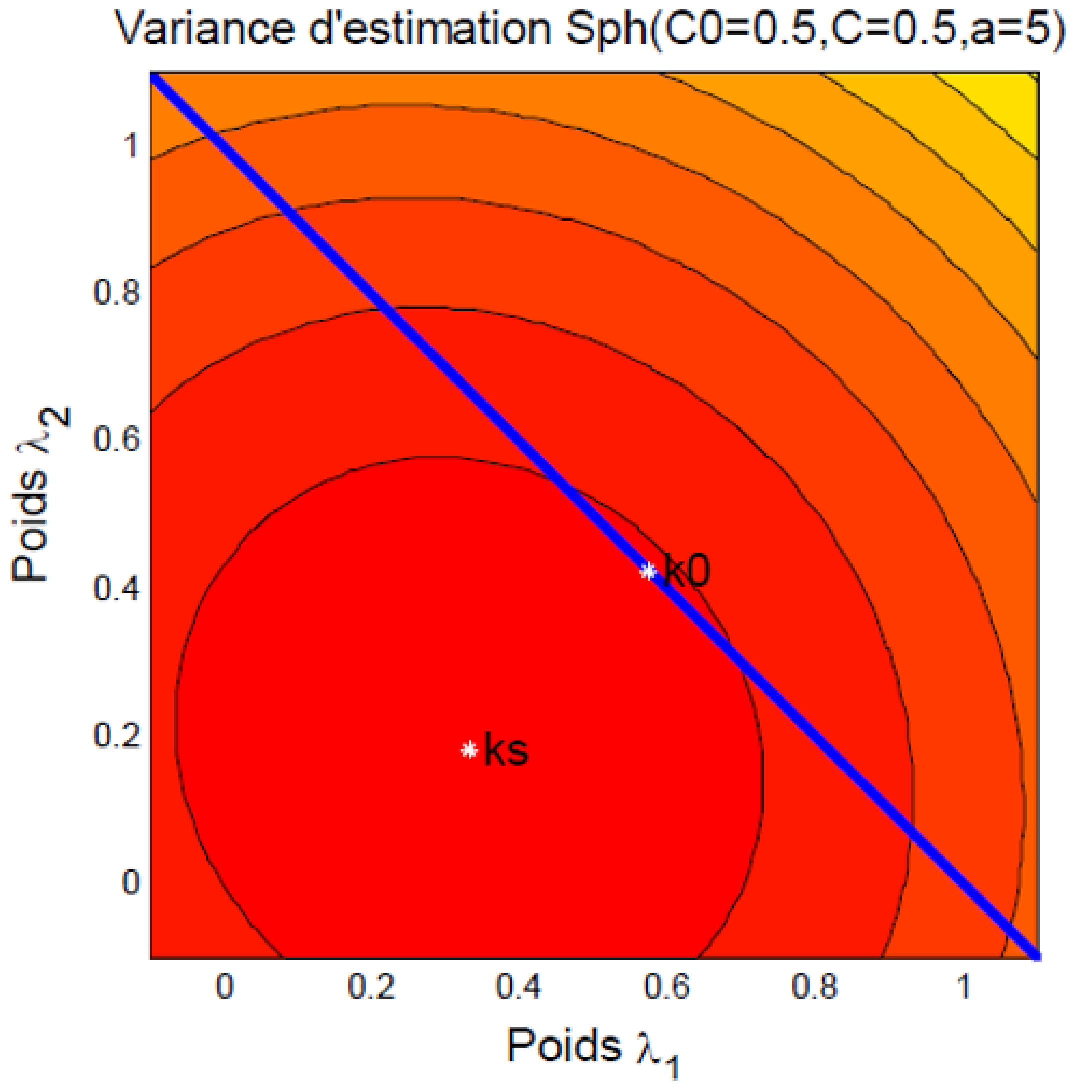
3. Interprétation

Liens entre KS et KO :



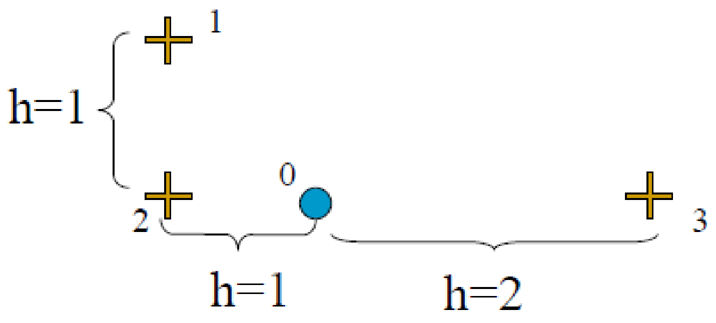
On a toujours :

$$\sigma_{KS}^2 \leq \sigma_{KO}^2$$



4. Exemple

Cas simple (3 observations et 1 estimation) :



Modèle de variogramme :

- Effet de pépite, $C_0 = 1$
- Modèle sphérique, $a = 3$ et $C = 10$

Données observées

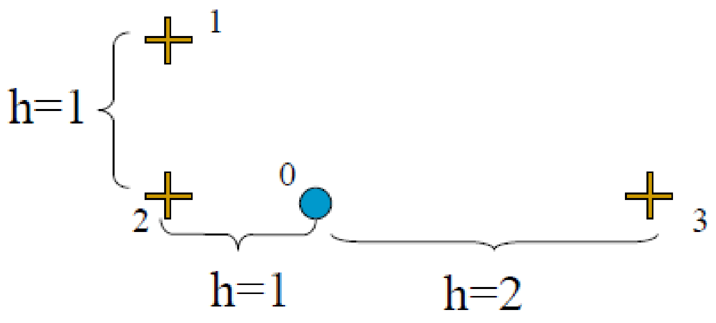
- $Z_1 = 9, Z_2 = 3, Z_3 = 4$

Comment solutionner un système de krigeage ?

Approche en 6 étapes

4. Exemple

Cas simple (3 observations et 1 estimation) :



Modèle de variogramme :

- Effet de pépite, $C_0 = 1$
- Modèle sphérique, $a = 3$ et $C = 10$

Données observées

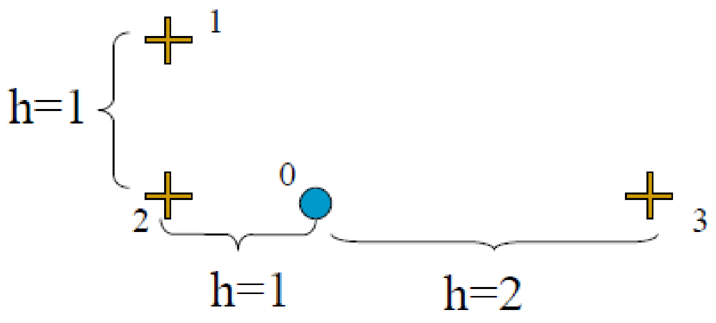
- $Z_1 = 9, Z_2 = 3, Z_3 = 4$

Comment solutionner un système de krigeage ?

Approche en 6 étapes

4. Exemple

Cas simple (3 observations et 1 estimation) :



Modèle de variogramme :

- Effet de pépite, $C_0 = 1$
- Sphérique, $a = 3$ et $C = 10$

Données observées

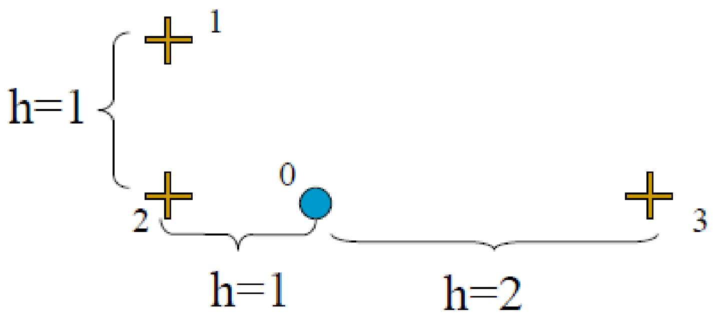
- $Z_1 = 9, Z_2 = 3, Z_3 = 4$

Étape 1) Calculer la matrice des distances

Cood.	x_0	x_1	x_2	x_3
x_0	0	1.4	1	2
x_1	1.4	0	1	3.2
x_2	1	1	0	3
x_3	2	3.2	3	0

4. Exemple

Cas simple (3 observations et 1 estimation) :



Modèle de variogramme :

- Effet de pépite, $C_0 = 1$
- Sphérique, $a = 3$ et $C = 10$

Données observées

- $Z_1 = 9, Z_2 = 3, Z_3 = 4$

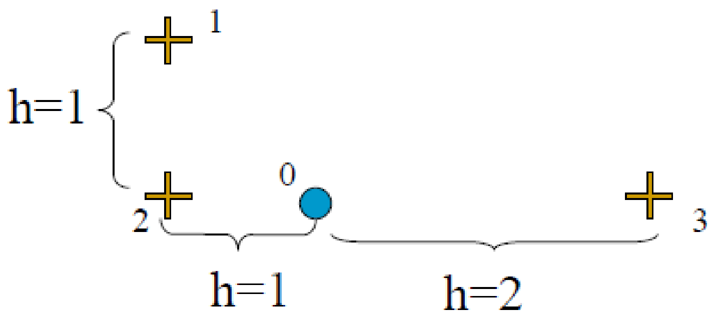
Étape 2) Transformer la matrice des distances en matrice de variogramme

$$h \rightarrow \gamma(h)$$

Cood.	x_0	x_1	x_2	x_3
x_0	0	7.55	5.81	9.52
x_1	7.55	0	5.81	11
x_2	5.81	5.81	0	11
x_3	9.52	11	11	0

4. Exemple

Cas simple (3 observations et 1 estimation) :



Modèle de variogramme :

- Effet de pépite, $C_0 = 1$
- Sphérique, $a = 3$ et $C = 10$

Données observées

- $Z_1 = 9, Z_2 = 3, Z_3 = 4$

Étape 3) Transformer la matrice de variogramme en matrice de covariance

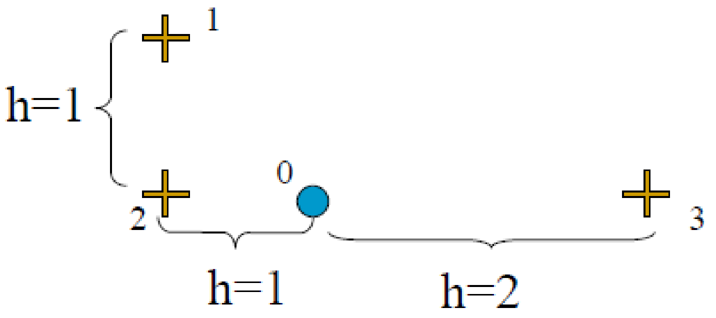
$$\gamma(h) \rightarrow C(h) = \sigma^2 - \gamma(h)$$

Cood.	x_0	x_1	x_2	x_3
x_0	11	3.45	5.19	1.48
x_1	3.45	11	5.19	0
x_2	5.19	5.19	11	0
x_3	1.48	0	0	11

Rouge : K
Bleu : k

4. Exemple

Cas simple (3 observations et 1 estimation) :



Modèle de variogramme :

- Effet de pépite, $C_0 = 1$
- Sphérique, $a = 3$ et $C = 10$

Données observées

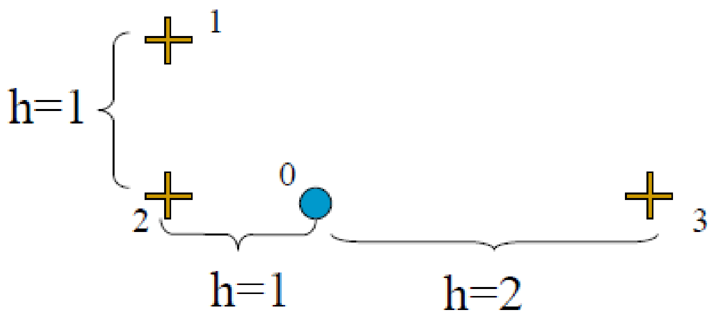
- $Z_1 = 9, Z_2 = 3, Z_3 = 4$

Étape 4) Construire le système de krigeage (ici KO)

$$
\begin{array}{ccccc}
K_O & & \lambda_O & & k_O \\
\begin{bmatrix} 11 & 5.19 & 0 & 1 \\ 5.19 & 11 & 0 & 1 \\ 0 & 0 & 11 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} & & \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \mu \end{bmatrix} & = & \begin{bmatrix} 3.45 \\ 5.19 \\ 1.48 \\ 1 \end{bmatrix}
\end{array}
$$

4. Exemple

Cas simple (3 observations et 1 estimation) :



Modèle de variogramme :

- Effet de pépite, $C_0 = 1$
- Sphérique, $a = 3$ et $C = 10$

Données observées

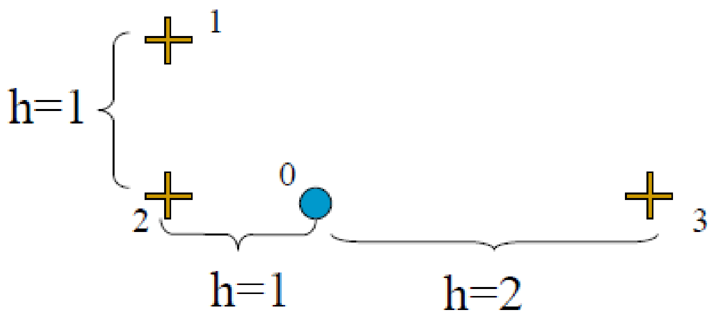
- $Z_1 = 9, Z_2 = 3, Z_3 = 4$

Étape 5) Résoudre le système d'équations

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \mu \end{bmatrix} = \begin{bmatrix} 0.21 \\ 0.51 \\ 0.28 \\ -1.55 \end{bmatrix}$$

4. Exemple

Cas simple (3 observations et 1 estimation) :



Modèle de variogramme :

- Effet de pépite, $C_0 = 1$
- Sphérique, $a = 3$ et $C = 10$

Données observées

- $Z_1 = 9, Z_2 = 3, Z_3 = 4$

Étape 6) Estimer Z_0^* et calculer la variance de krigeage

$$Z_0^* = \sum_{i=1}^{3} \lambda_i Z_i = (.21) \times 9 + (.51) \times 3 + (.28) \times 4 = 4.54$$

$$\sigma_{KO}^2 = \sigma^2 - \lambda' k_0 = 11 - (.21) \times 3.45 - (.51) \times 5.19 - (.28) \times 1.48 + 1.55 = 8.76$$

5. Propriétés du krigeage

Par construction : estimateur linéaire, sans biais, à variance d'estimation minimale

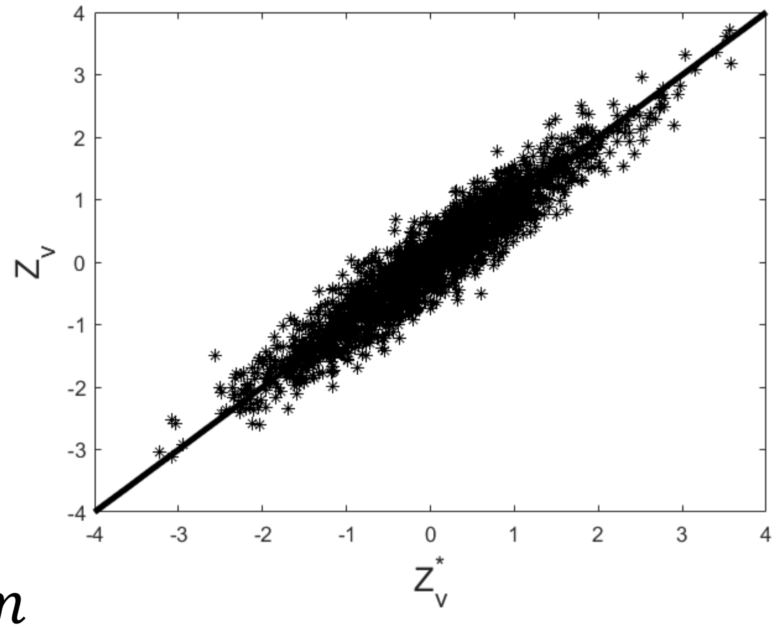
1. Presque sans biais conditionnel
2. Effet de lissage
3. Interpolateur exact
4. Effet d'écran
5. Tiens compte de la taille du champ à estimer et de la position des points entre eux.
6. Tiens compte de la continuité spatiale du phénomène étudié
7. Transitif (cohérence des estimations)

1. Presque sans biais conditionnel :

Définition : Biais conditionnel

Si Z_v et Z_v^* suivent une loi binormale de moyenne m, alors la régression linéaire de Z_v sur Z_v^* s'écrit :

$$E[Z_v|Z_v^*] = a + bZ_v^*$$



Avec

$$b = \frac{Cov(Z_v, Z_v^*)}{Var(Z_v^*)}$$

$$\mathrm{a} = \left(1 - \frac{Cov(Z_v, Z_v^*)}{Var(Z_v^*)}\right) m = (1 - b)m$$

5. Propriétés du krigeage

1. Presque sans biais conditionnel :

$$E[Z_v|Z_v^*] = a + bZ_v^* \; ; \quad b = \frac{Cov(Z_v, Z_v^*)}{Var(Z_v^*)} \; ; \quad \mathrm{a} = (1-b)m$$

<u>Krigeage simple, par construction :</u>

$$Var(Z_v^*) = Cov(Z_v, Z_v^*) \rightarrow b = 1 \, ; a = 0 \rightarrow \boxed{E[Z_v|Z_v^*] = Z_v^*}$$

- Le krigeage simple est sans biais conditionnel seulement dans le cas normal.

- Si on n'est pas dans un cas normal, alors on peut simplement dire que l'estimateur KS est approximativement sans biais conditionnel.

6. Aspects pratiques

1. Presque sans biais conditionnel :

$$E[Z_v|Z_v^*] = a + bZ_v^* \; ; \quad b = \frac{Cov(Z_v, Z_v^*)}{Var(Z_v^*)} \; ; \quad \mathrm{a} = (1-b)m$$

<u>Krigeage ordinaire, par construction :</u>

$$Var(Z_v^*) + \mu = Cov(Z_v, Z_v^*) \rightarrow b = 1 + \frac{\mu}{Var(Z_v^*)} \; ; a = \frac{-\mu}{Var(Z_v^*)}$$

$$\rightarrow E[Z_v|Z_v^*] = \boxed{Z_v^* + \frac{\mu}{Var(Z_v^*)}(Z_v^* - m)}$$

- Dans le cas normal, l'estimateur KO présente un biais conditionnel proportionnel à μ
- Si $\mu < 0 \; \rightarrow b < 1$, alors les fortes valeurs de KO surestiment les vraies teneurs des blocs
- Si $|\mu|$ tend vers 0, alors KO tend à être sans biais conditionnel

5. Propriétés du krigeage

2. Effet de lissage :

Krigeage simple : $Var(Z_v) = Var(Z_v^*) + \sigma_{KS}^2$

$$\boxed{Var(Z_v^*) \leq Var(Z_v)}$$

L'estimateur KS est ***toujours*** moins variable que la réalité qu'il cherche à estimer

Krigeage ordinaire : $Var(Z_v) = Var(Z_v^*) + \sigma_{KO}^2 + 2\mu$

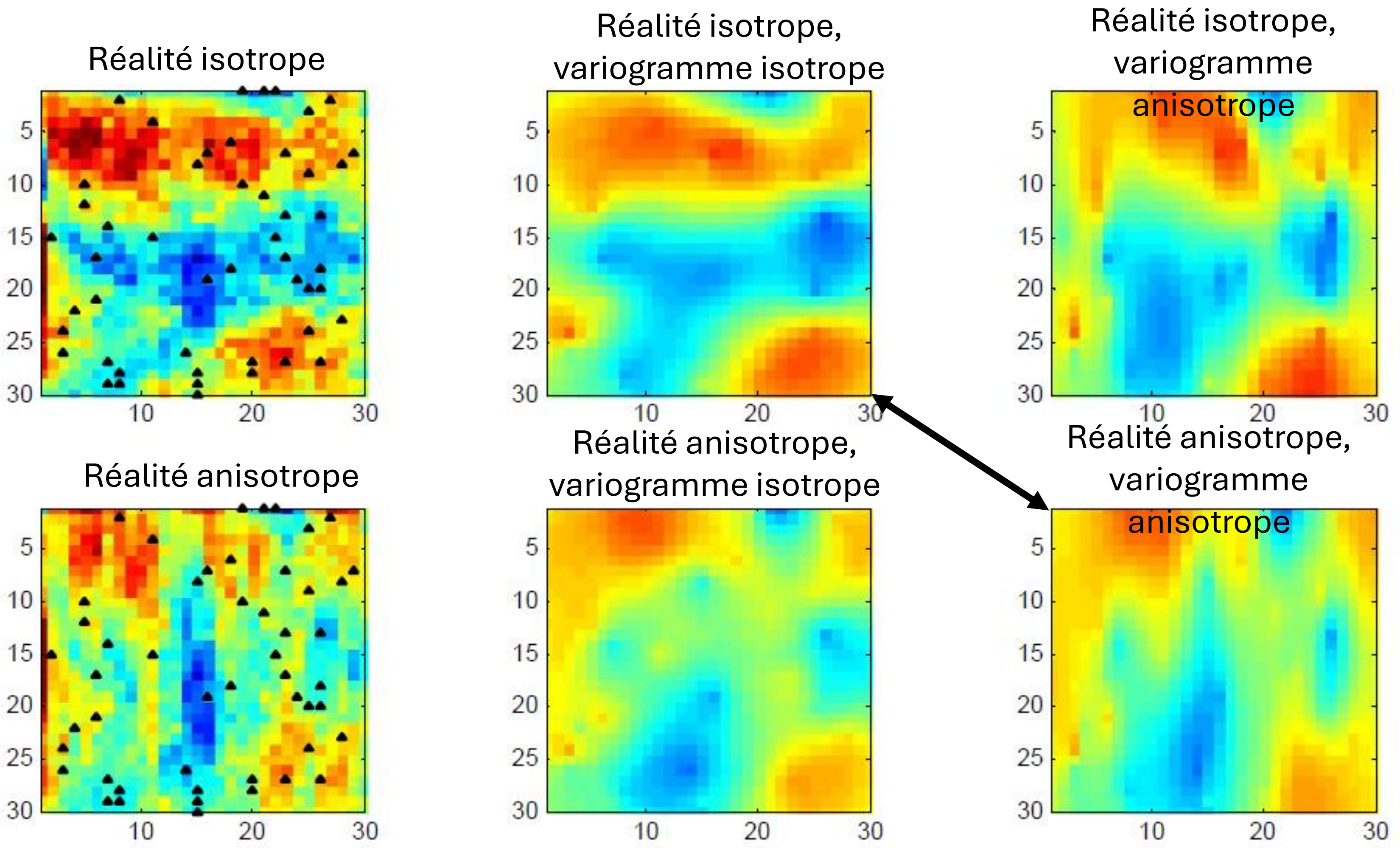
Habituellement, $\mu < 0$ et $2|\mu| < \sigma_{KO}^2$

$$\boxed{Var(Z_v^*) \leq Var(Z_v)}$$

L'estimateur KO est ***habituellement*** moins variable que la réalité qu'il cherche à estimer

5. Propriétés du krigeage

2. Effet de lissage :



5. Propriétés du krigeage

2. Effet de lissage :

<u>Lien entre le lissage et le biais conditionnel :</u>

$$b = \frac{Cov(Z_v, Z_v^*)}{Var(Z_v^*)} = \frac{\rho\sigma_v\sigma_v^*}{\sigma_v^{*2}} = \frac{\rho\sigma_v}{\sigma_v^*}$$

Une absence de biais conditionnel implique que $b = 1$, alors $\sigma_v^* \leq \sigma_v$

- Un estimateur sans lissage est nécessairement avec biais conditionnel.

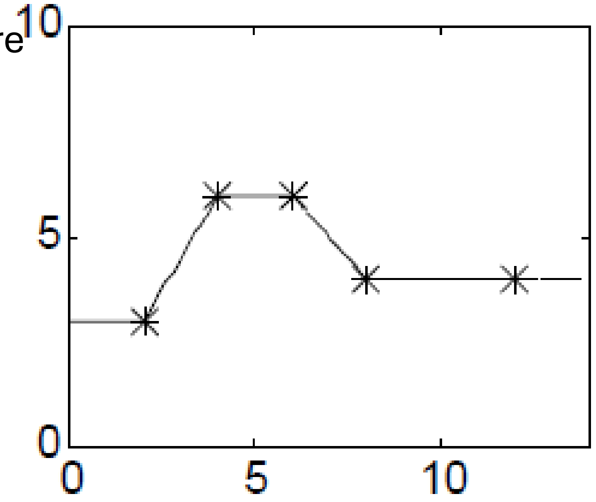
- Les valeurs estimées doivent montrer une variance inférieure aux vraies valeurs.

5. Propriétés du krigeage

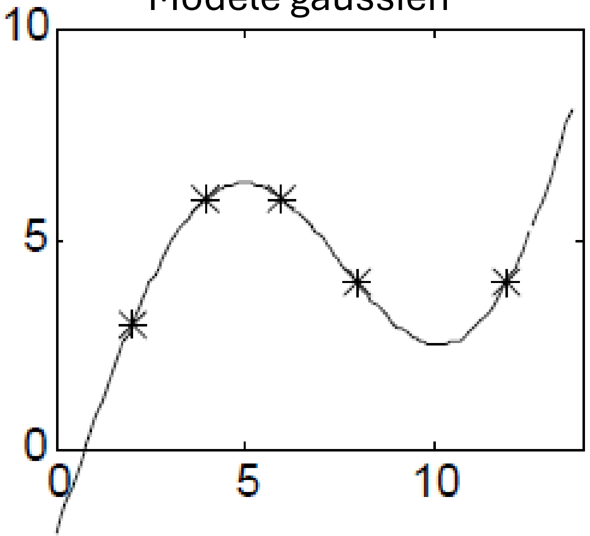
3. Interpolateur exact :

Estime les données observées avec exactitude

Modèle linéaire



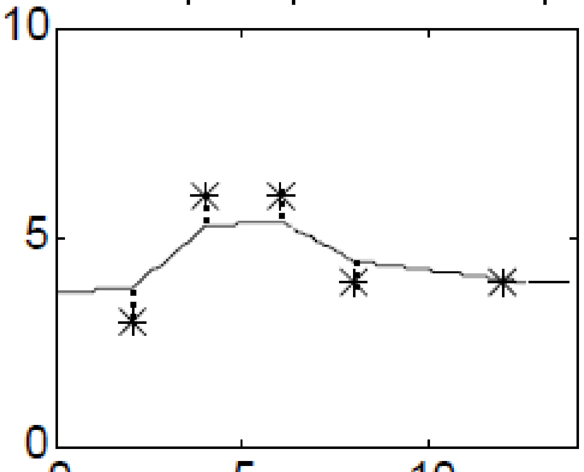
Modèle gaussien



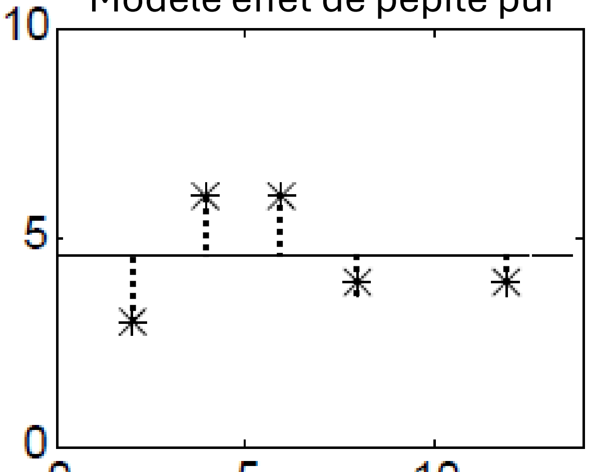
En présence d'effet de pépite, les valeurs interpolées sont discontinues

→ éviter d'estimer un point observé

Modèle sphérique et effet de pépite

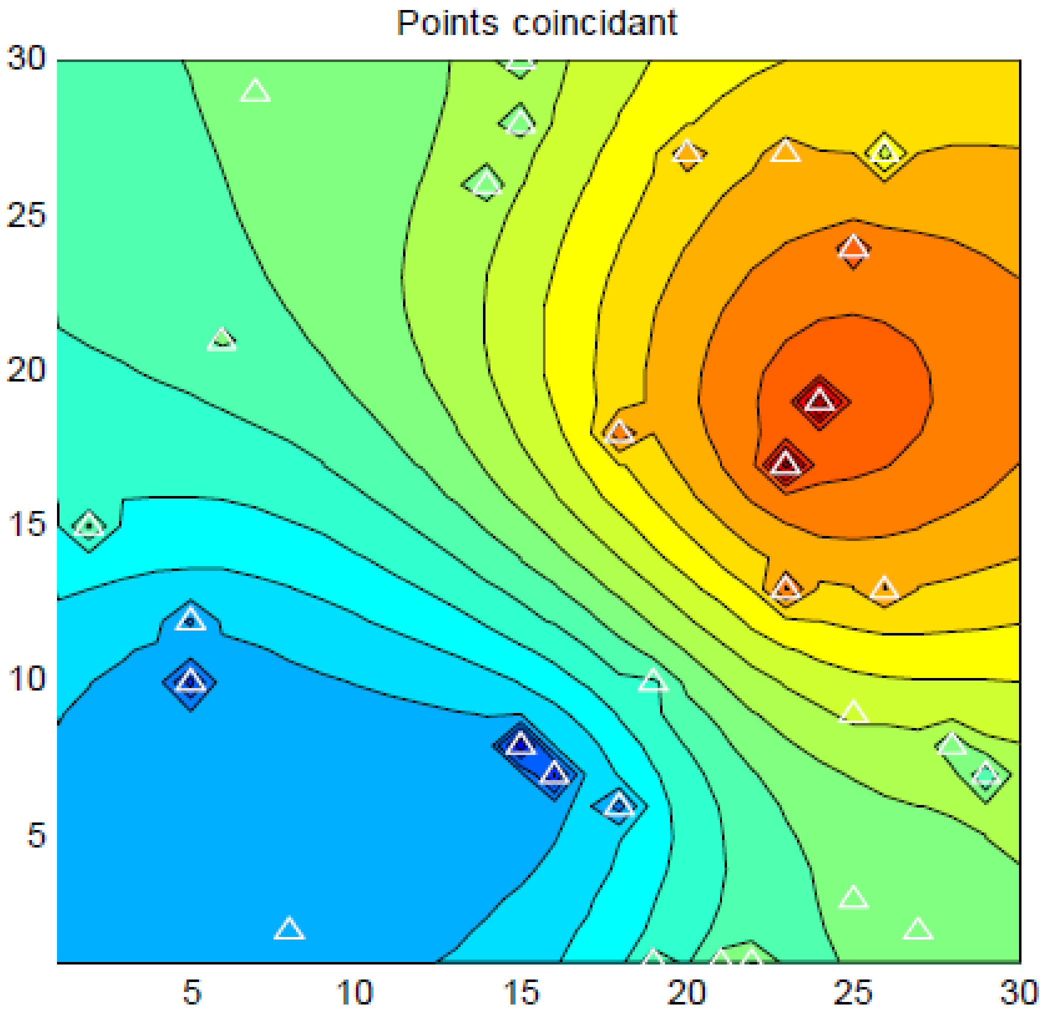


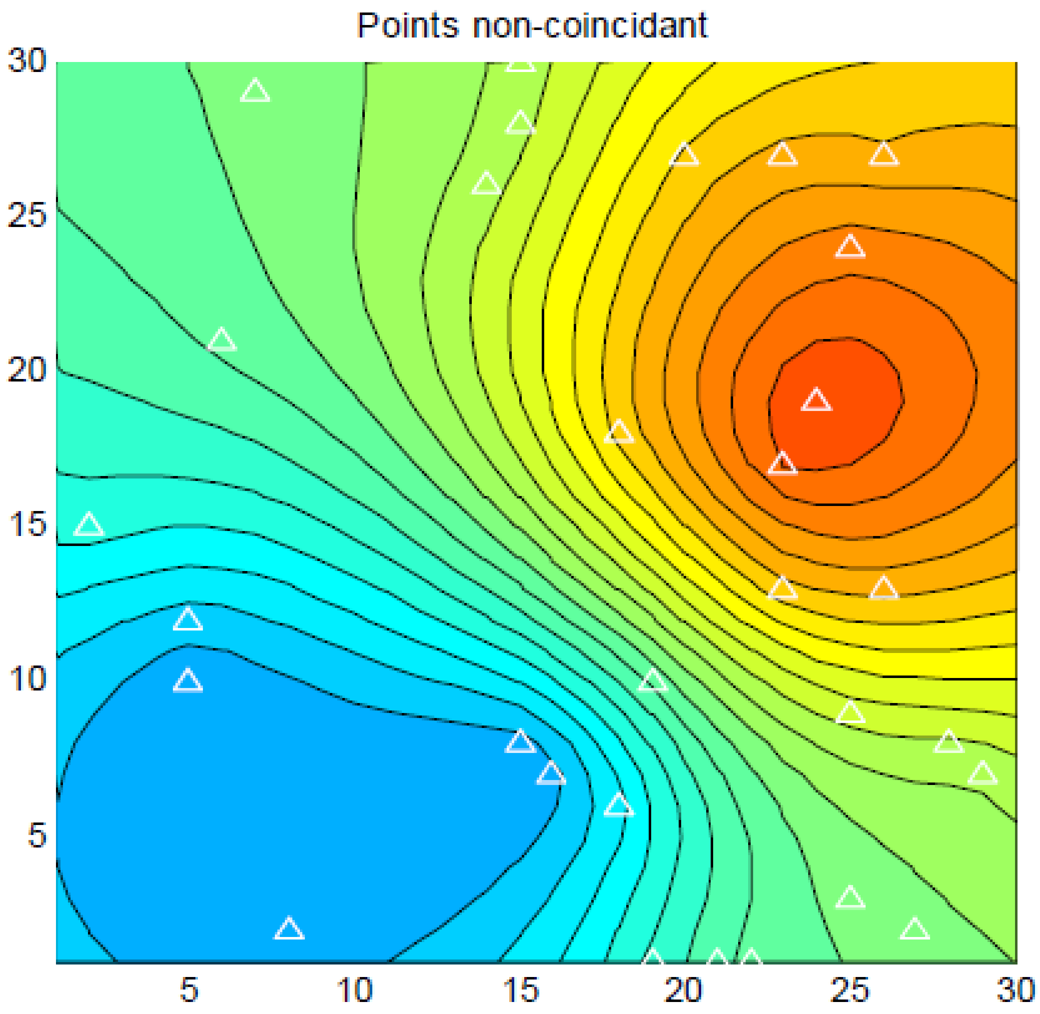
Modèle effet de pépite pur



5. Propriétés du krigeage

3. Interpolateur exact :



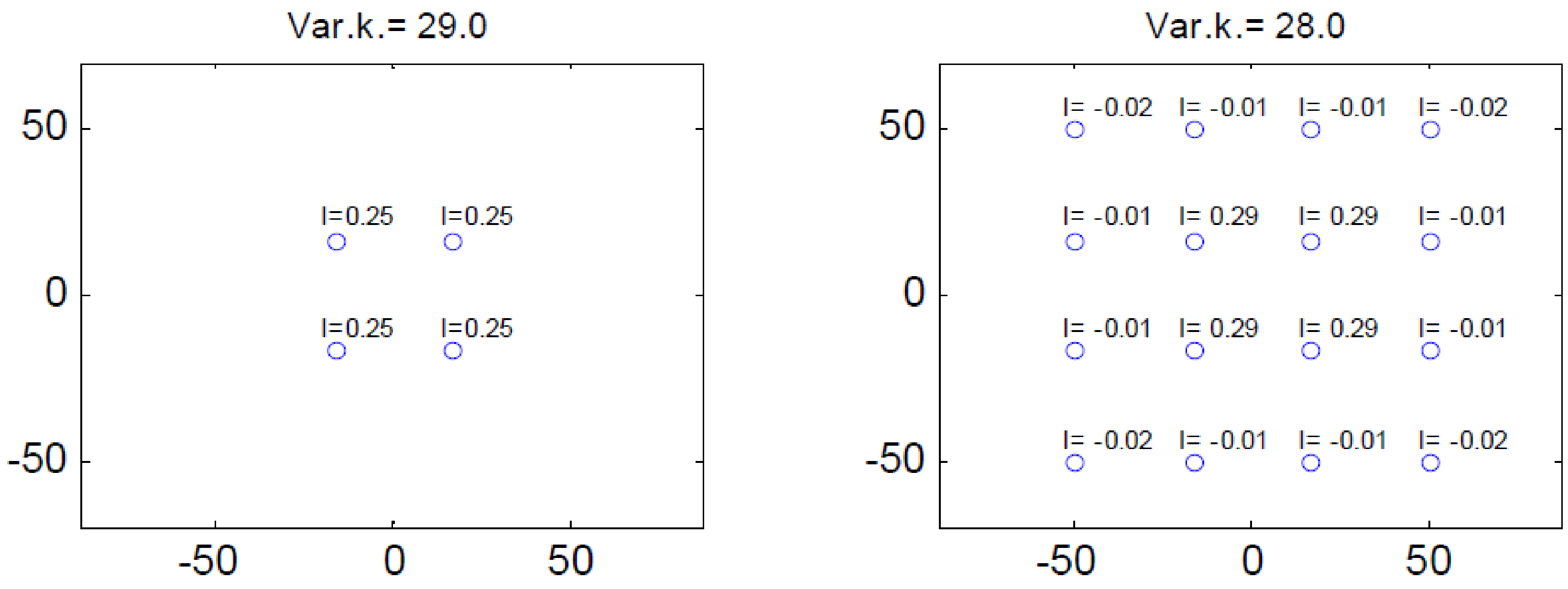


5. Propriétés du krigeage

4. Effet d'écran :

- Les poids des points périphériques sont faibles.

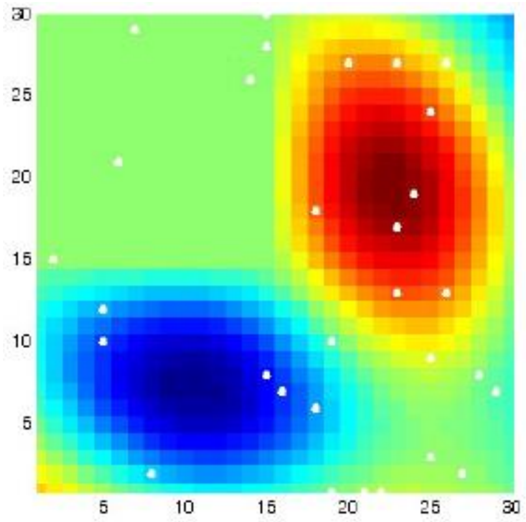
- Les poids des points proches sont forts.

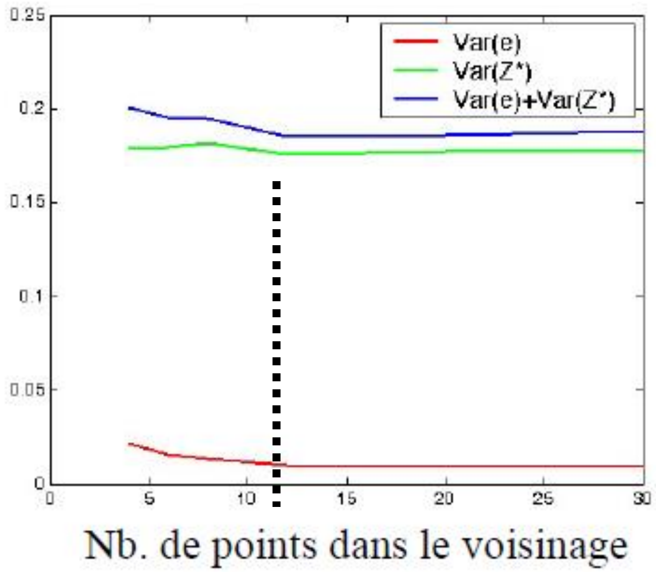


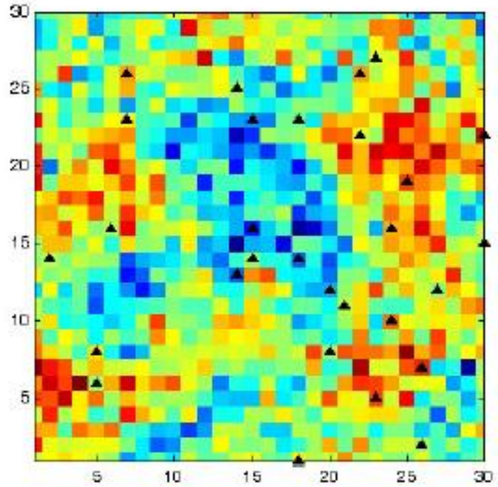
Limiter l'estimation d'un point par son voisinage

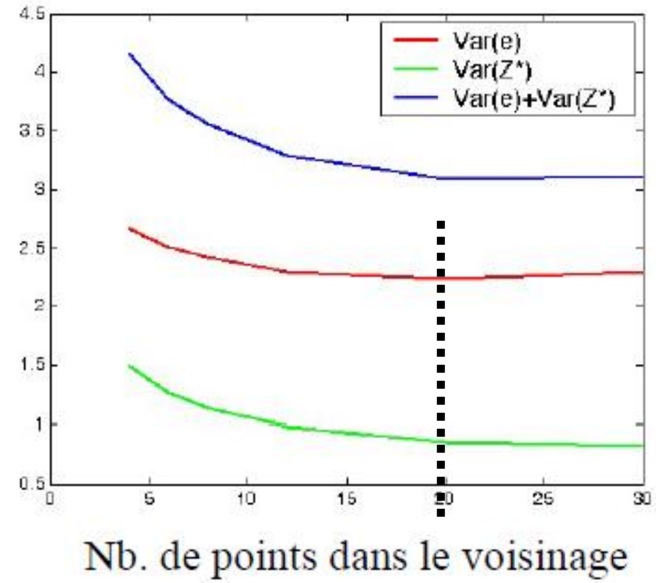
5. Propriétés du krigeage

4. Effet d'écran :

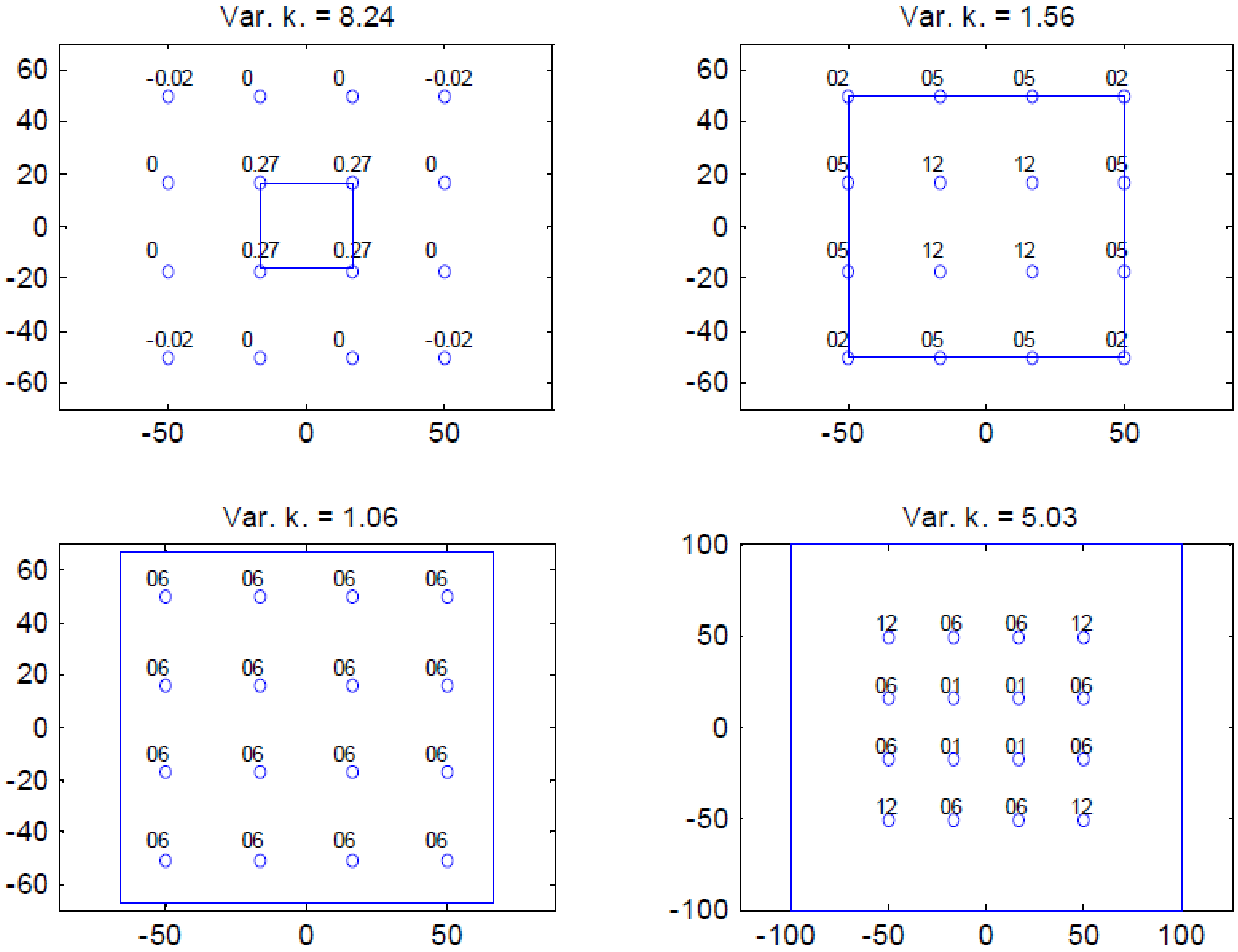






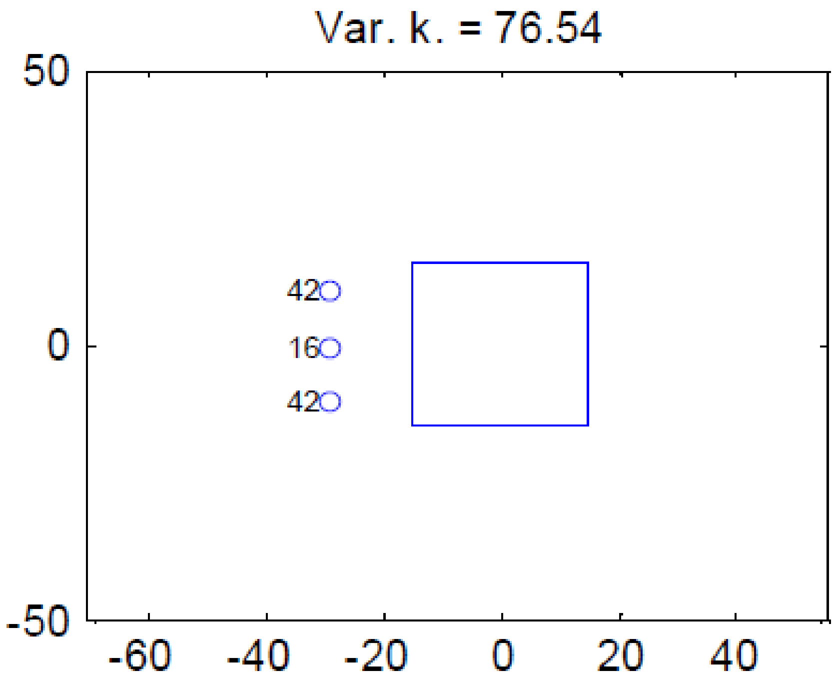


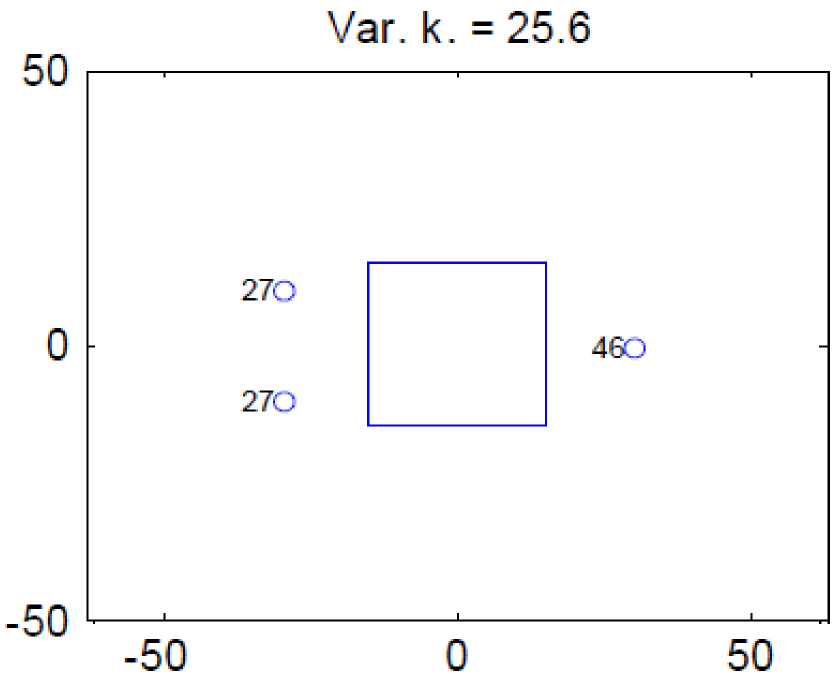
5. Tiens compte de la taille du champ à estimer et de la position des points entre eux :



5. Tiens compte de la taille du champ à estimer et de la position des points entre eux :

Redondance des données :

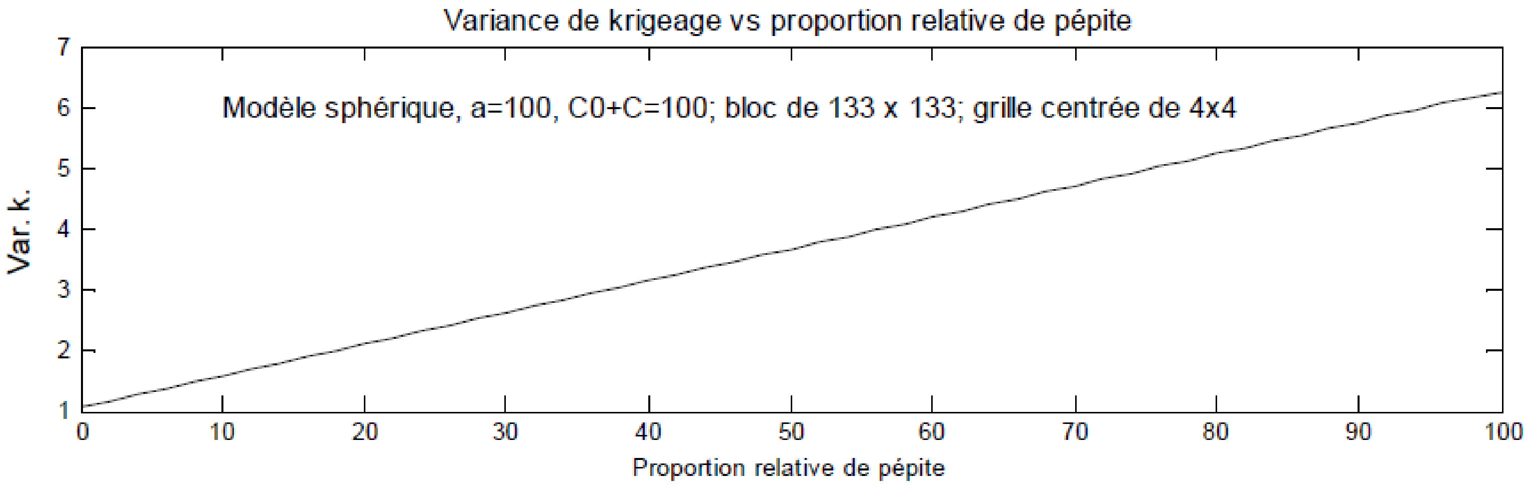




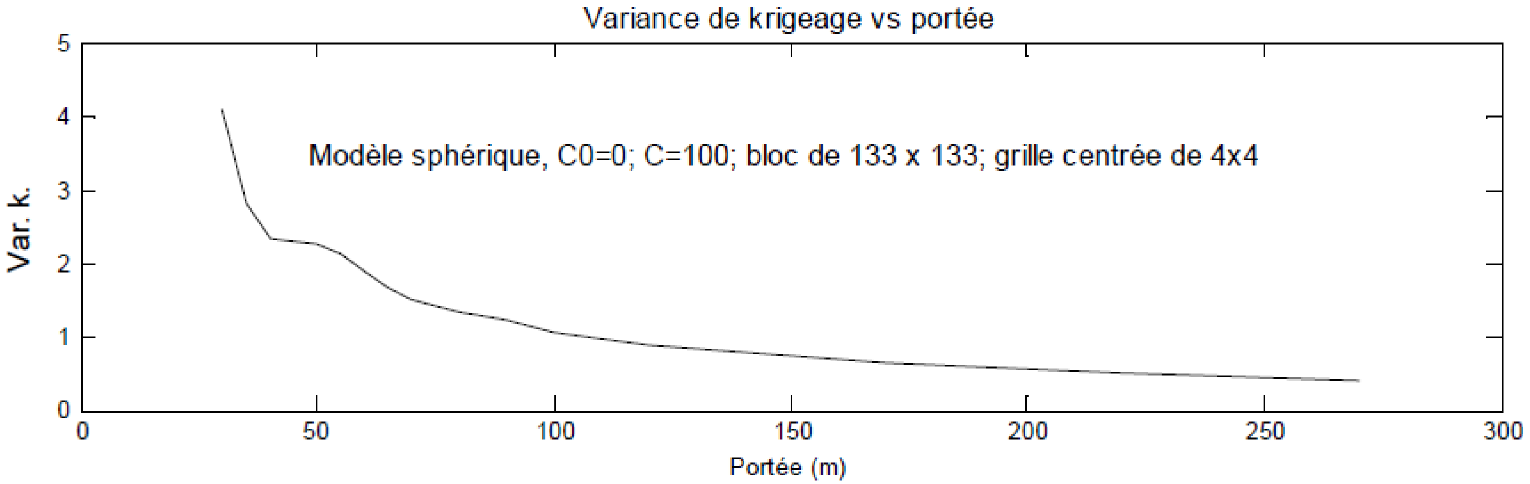
5. Propriétés du krigeage

6. Tiens compte de la continuité spatiale du phénomène étudié :

Influence de l'effet de pépite

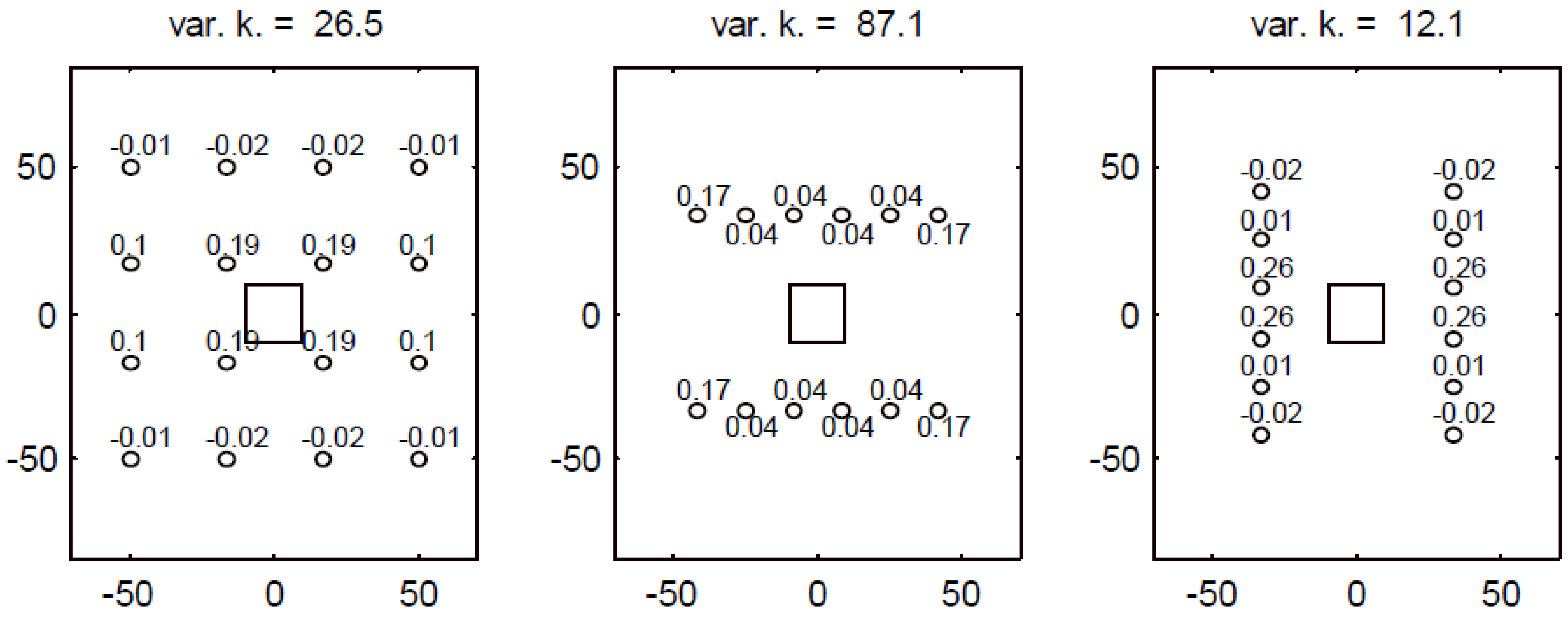


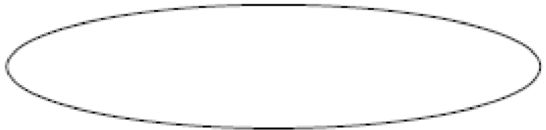
Influence de la portée



6. Tiens compte de la continuité spatiale du phénomène étudié :

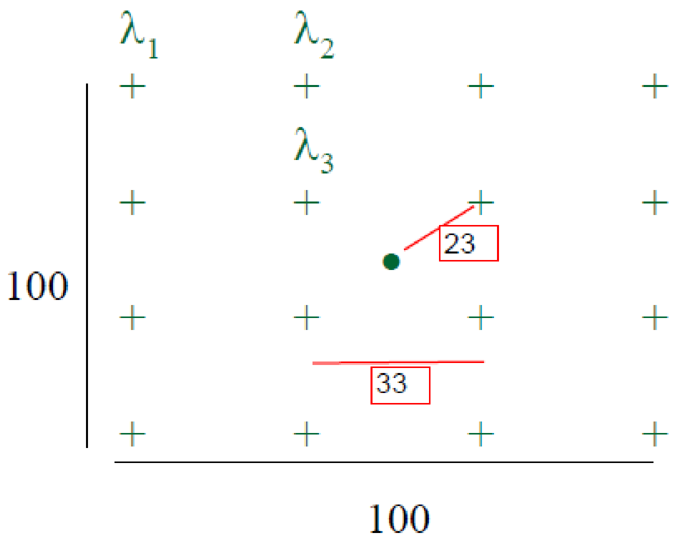
Influence de l'anisotropie :

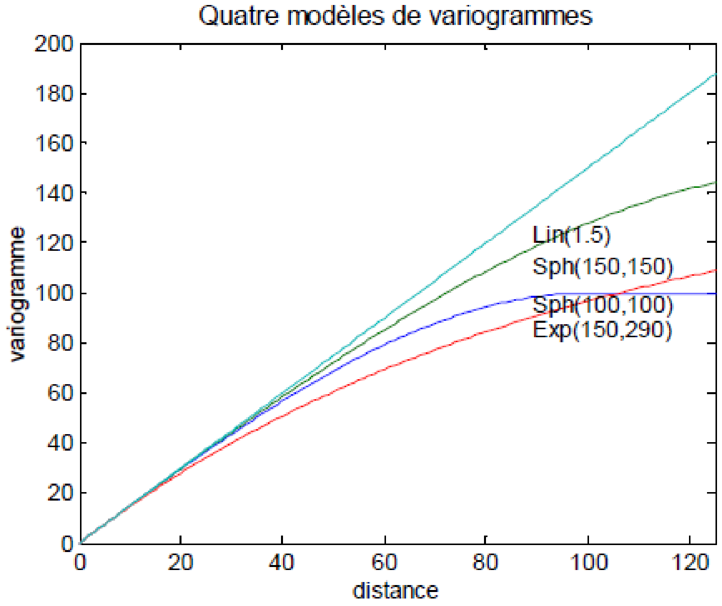


Var. anisotrope 

6. Tiens compte de la continuité spatiale du phénomène étudié :

Influence du modèle :





	λ_1	λ_2	λ_3	σ_k^2
Sphérique, C=100, a=100	-.02	-.01	.29	28.0
Sphérique, C=150, a=150	-.01	-.01	.29	27.8
Exp. C=150, a_{eff}=290	-.01	-.01	.28	28.2
Linéaire, pente=1.5	-.01	-.01	.28	27.6

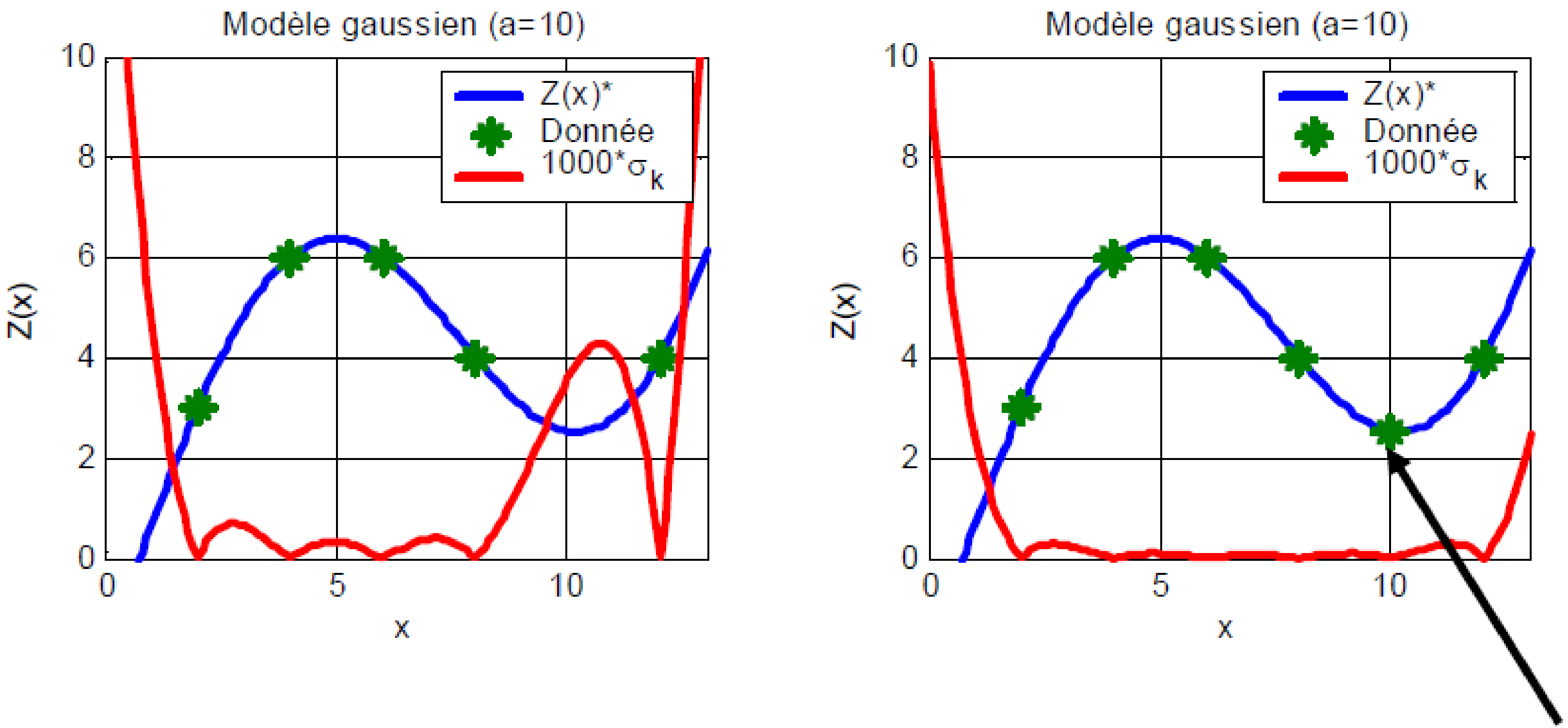
4 ajustements équivalents h<30

=> mêmes poids λ

=> même σ_k^2

5. Propriétés du krigeage

7. Transitif (cohérence des estimations) :



À droite, à x=10, on observe une donnée égale à la valeur krigée à gauche. Toutes les valeurs krigées demeurent inchangées. Seules les variances de krigeage sont réduites.

6. Aspects pratiques

- Krigeage ordinaire : syst. d'éq. linéaire (n+1) équations et (n+1) inconnues
  → limite pratique sur « n »

  « m » est estimée implicitement → voisinage glissant (ou local) permet de relaxer l'hypothèse de stationnarité (« m » peut fluctuer d'un voisinage à l'autre)

- Grille de krigeage: régulière ou non, points ou blocs.

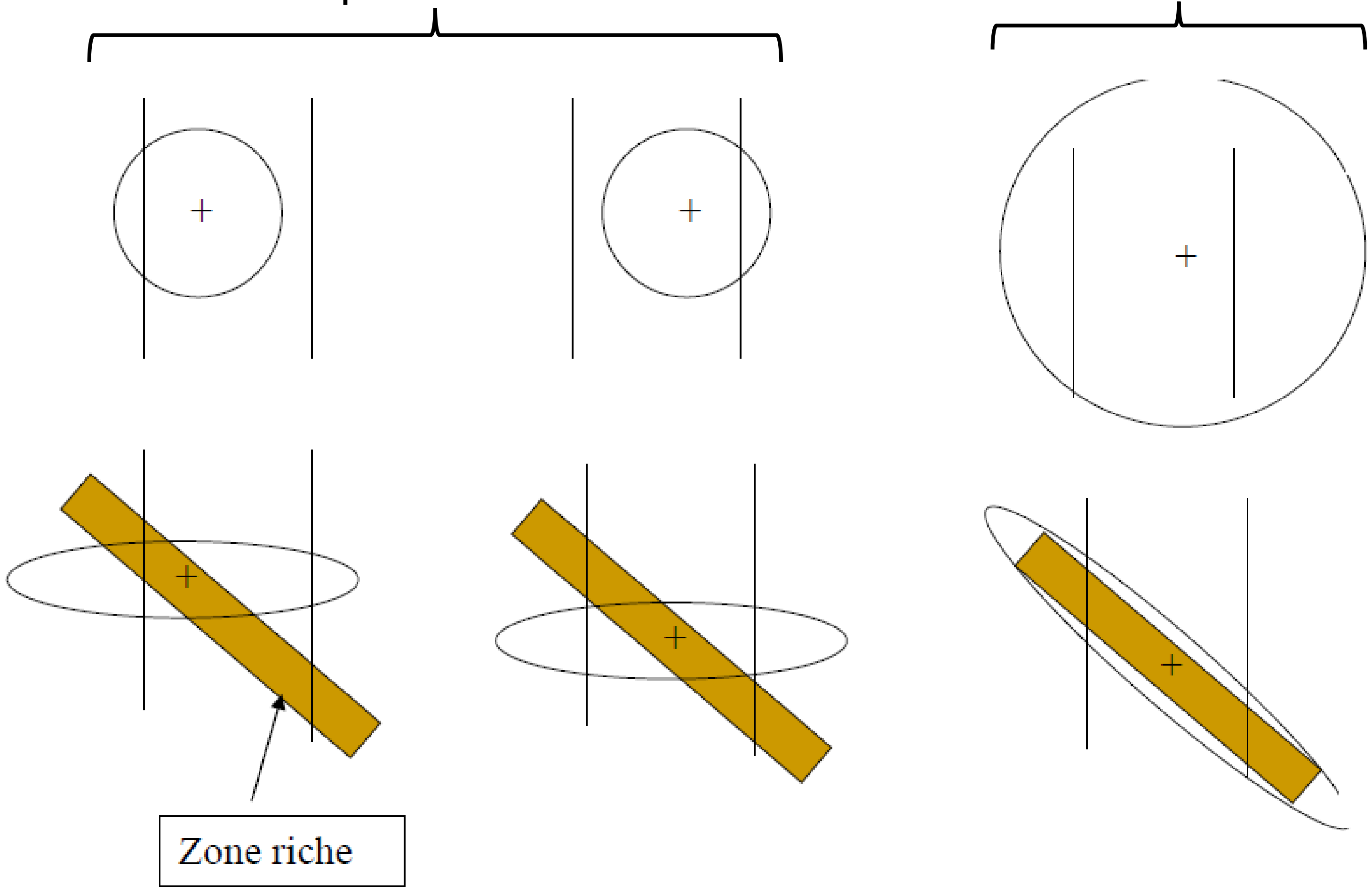
- Voisinage utilisé pour le krigeage:
  - Habituellement en voisinages glissants.
  - Nombre de points suffisant (>10; peut atteindre jusqu'à 50-100).
    - Zone de recherche assez grande pour assurer un minimum de points.
    - Recherche par quadrants (2D) ou octants (3D) (min 2 ou 3 points par quadrant/octant)

6. Aspects pratiques

Importance du voisinage :

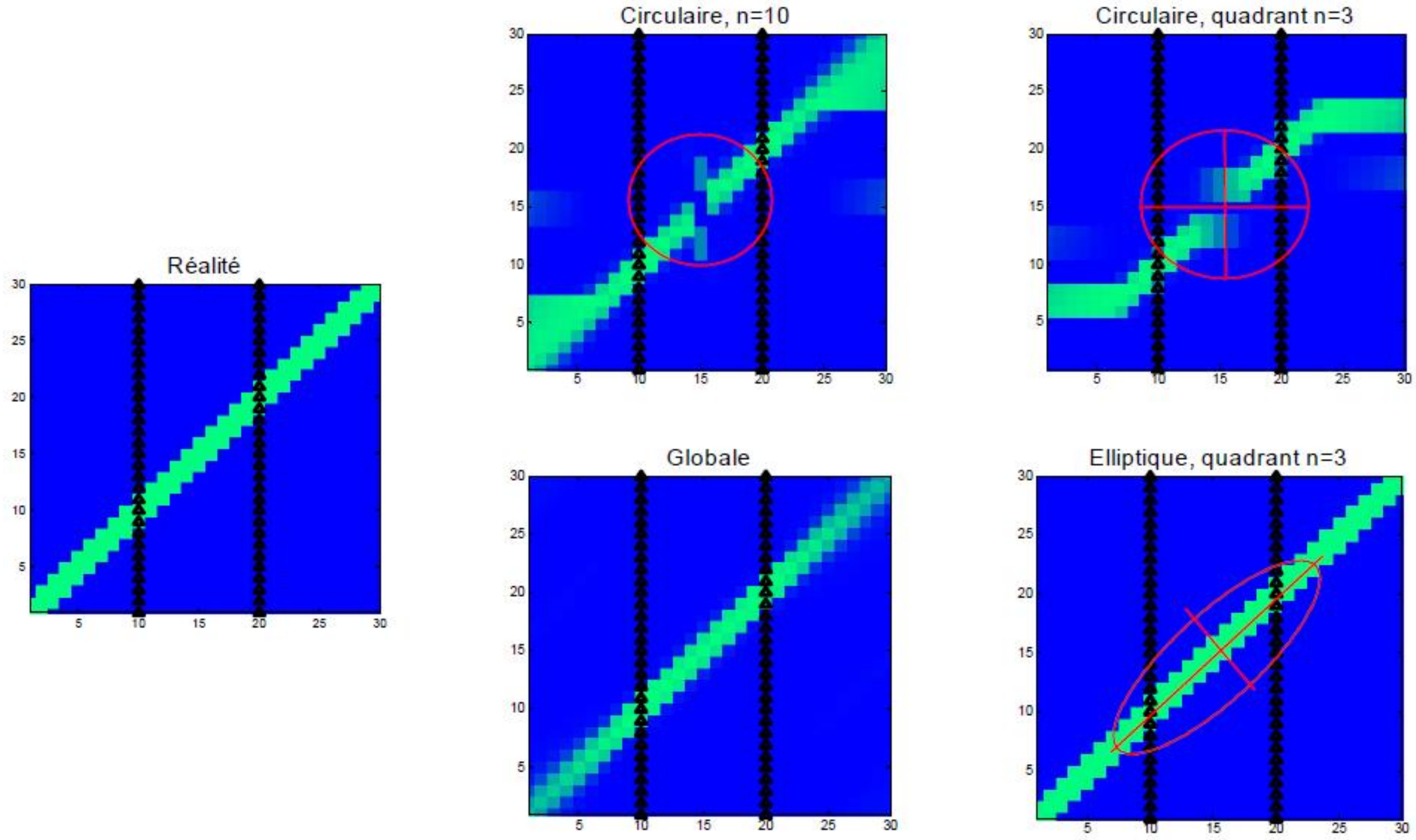
Estimation peu précise
Discontinuité prononcée sur la carte

Estimation précise
Pas de discontinuité



6. Aspects pratiques

Importance du voisinage : Exemple



7. Validation croisée

Objectifs :

Permets de valider un modèle de variogramme

Permets de valider le voisinage utilisé pour le krigeage

Technique du '*leave-one-out*' :

Consiste à retirer une observation pour ensuite l'estimer par krigeage à partir des autres données observées. Cela est répété pour tous les points.

$$e_i = Z_v(x_i) - Z_v^*(x_i) \;\; ; \;\; n_i = \frac{e_i}{\sigma_{K,i}}$$

Statistique pour comparaison :

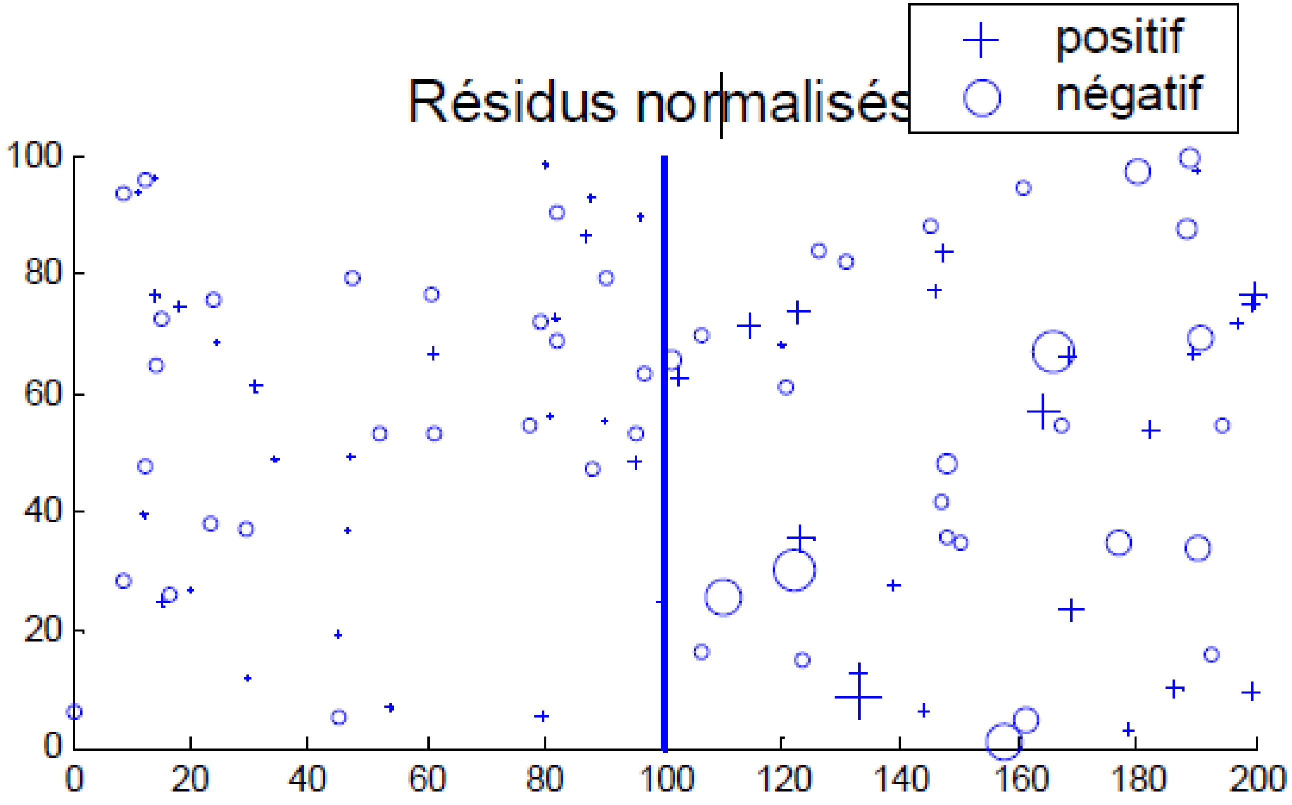
$$\sum_i e_i \approx 0 \;\; et \sum_i n_i \approx 0$$

$$\min\left(\sum_i |e_i|\right) ou \min\left(\sum_i e_i^2\right) \; ou \sqrt{\frac{1}{N}\sum_i n_i^2} \approx 1$$

7. Validation croisée

Comparaison visuelle :

- Histogramme des résidus et résidus normalisés
- Carte des résidus et résidus normalisés



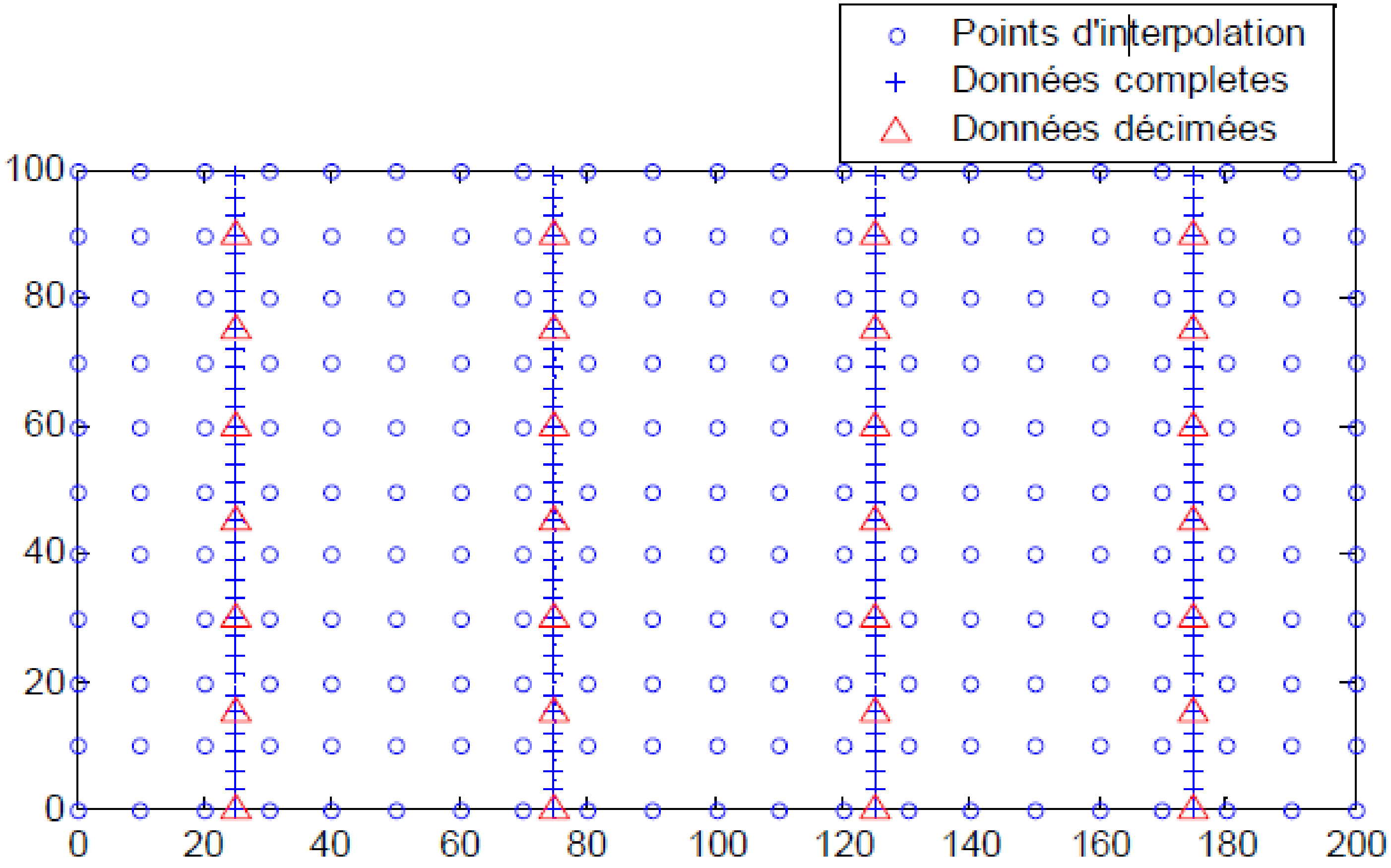
Diviser en deux zones distinctes ?

7. Validation croisée

Reproduire des situations réalistes d'estimation :

- Grille complète des données → valider le variogramme à petite échelle
- Grille décimée → valider le variogramme à des distances plus grandes

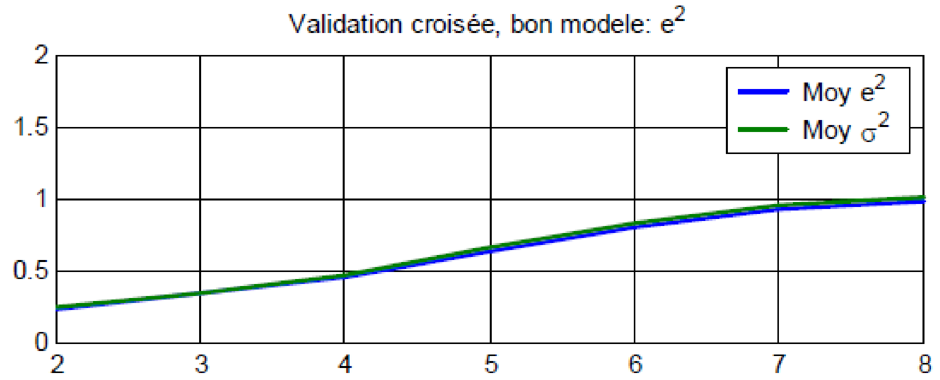
Pourquoi ?



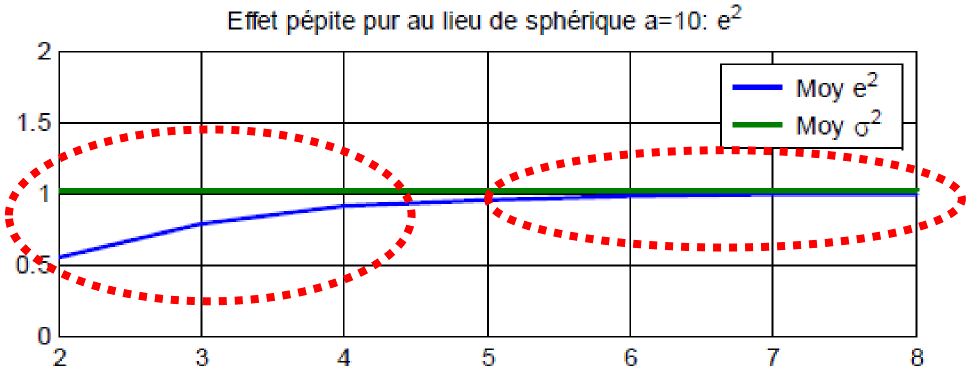
7. Validation croisée

Exemple de validation croisée:

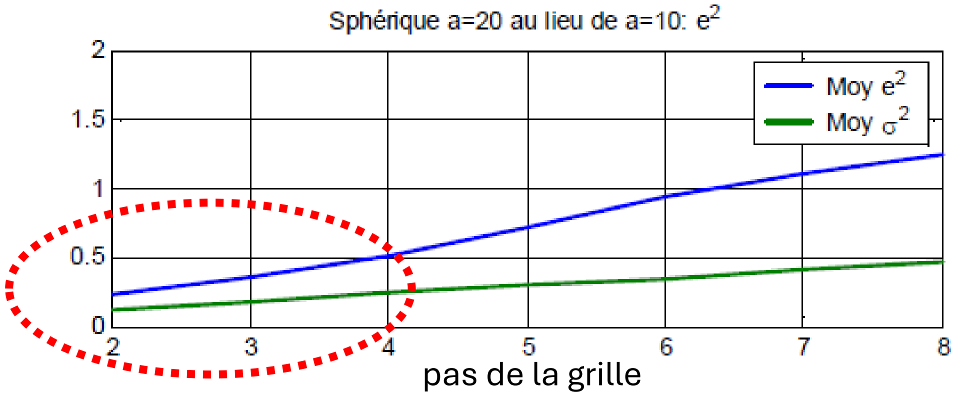
- Grille de dimension 40 x 40 (1 600 points)
- Modèle de variogramme : sphérique isotrope $(a = 10, C = 1, C_0 = 0)$
- Voisinage : 50 points



1) Modèle fourni : sphérique isotrope $(a = 10, C = 1, C_0 = 0)$



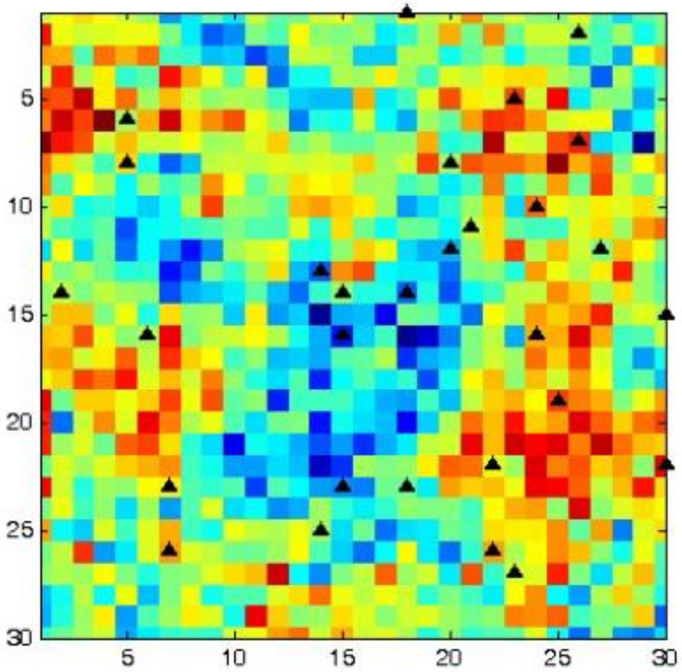
2) Modèle fourni : effet de pépite pur (C_0 =0)

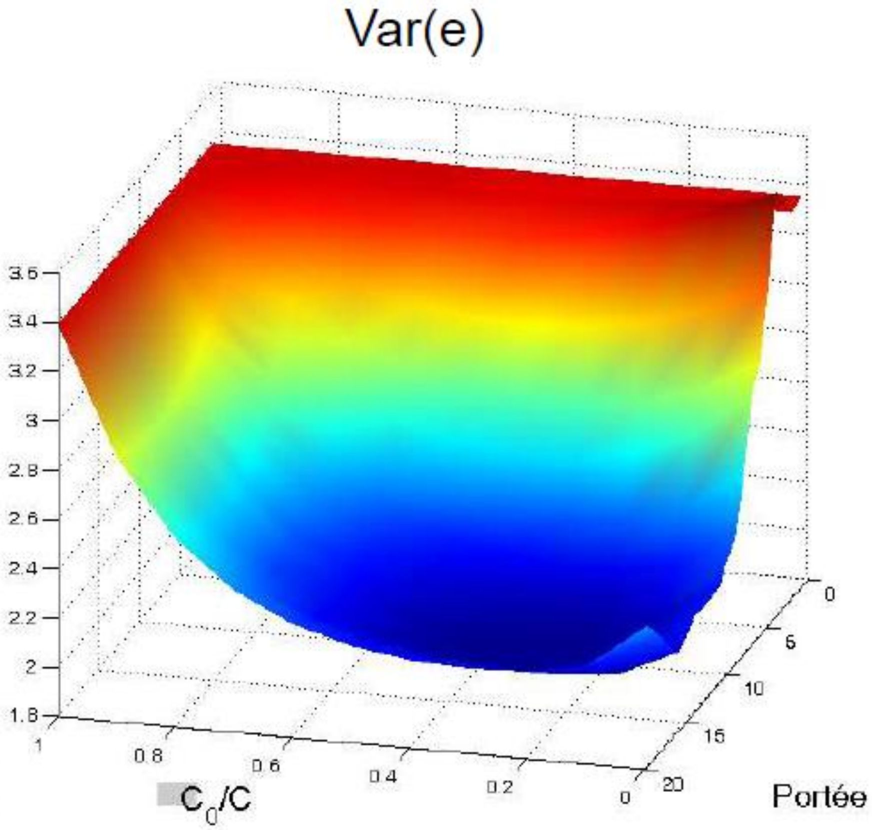


3) Modèle fourni : sphérique isotrope $(a = 20, C = 1, C_0 = 0)$

7. Validation croisée

Exemple simulé de validation croisée:





Minimum en a=15, C0/C=0.3; près des valeurs utilisées pour la simulation (C0/C=0.33; a=10)