

GML6402A : Géostatistique

Cours 1 : Variogramme



**POLYTECHNIQUE
MONTREAL**

UNIVERSITÉ
D'INGÉNIERIE

Objectifs

- Expliquer ce que représente le variogramme et en quoi il permet de mesurer la continuité spatiale des teneurs;
- Expliquer la différence entre variogramme expérimental et théorique et comprendre la nécessité d'un modèle théorique;
- Calculer un variogramme expérimental;
- Ajuster un modèle théorique à un variogramme expérimental;
- Calculer la covariance entre deux points à partir d'un modèle de variogramme;
- Tenir compte dans le calcul du variogramme ou de la covariance d'anisotropies;
- Connaître les principales caractéristiques des modèles de variogramme courants.

1. Point de vue de la géostatistique
 - 1.1 Rappels
 - 1.2 Effet de support
 - 1.3 Effet d'information
2. Géostatistique linéaire
 - 2.1 Hypothèse de stationnarité
 - 2.2 Variogramme expérimental
 - 2.3 Variogramme théorique
3. Modélisation du variogramme théorique
 - 3.1 Modèles de variogramme admissible
 - 3.2 Anisotropies
4. Problème et stratégie de modélisation



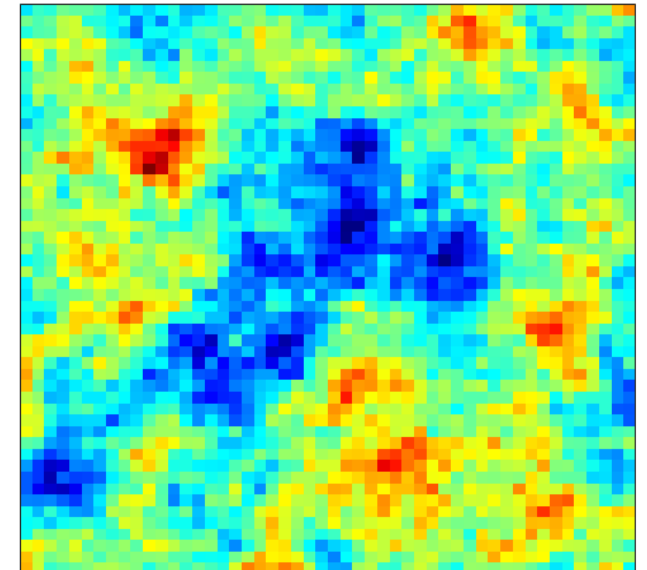
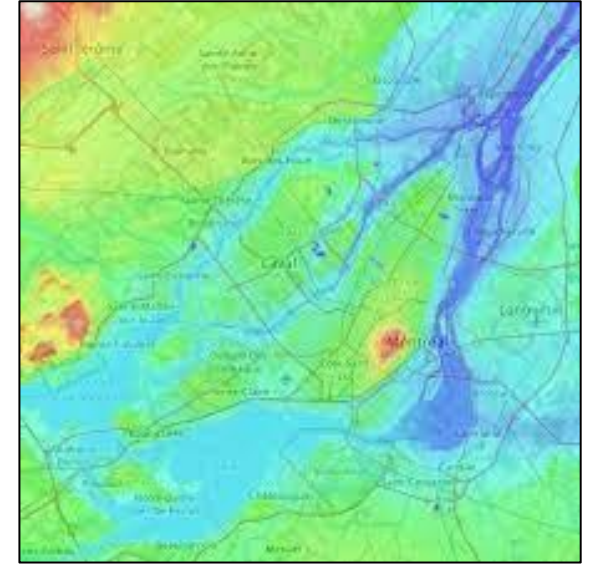
1. Point de vue de la géostatistique

Première loi de la géographie

« *Tout interagit avec tout, mais deux objets proches ont plus de chances de le faire que deux objets éloignés* », **Waldo Tobler**

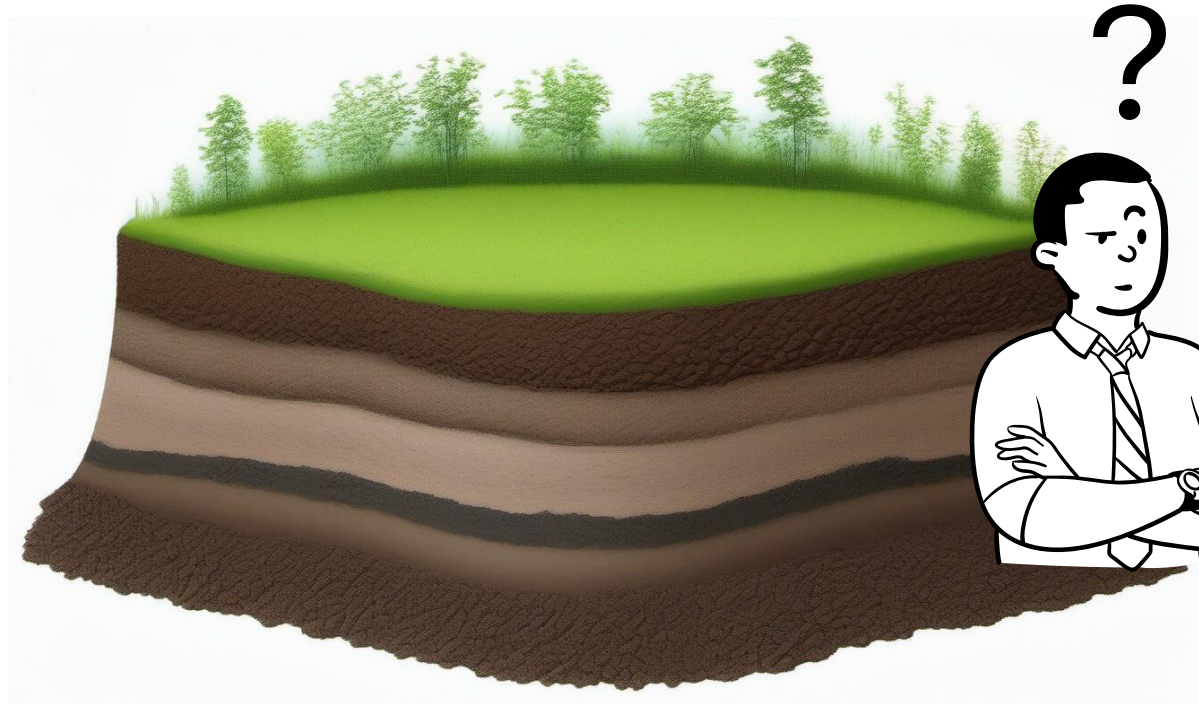
Aussi vrai en géologie.

Deux observations situées l'une près de l'autre devraient, en moyenne, se ressembler davantage que deux observations éloignées.



1. Point de vue de la géostatistique

Objectifs de la géostatistique



Géologie : résultante de processus géologiques complexes et temporelles.



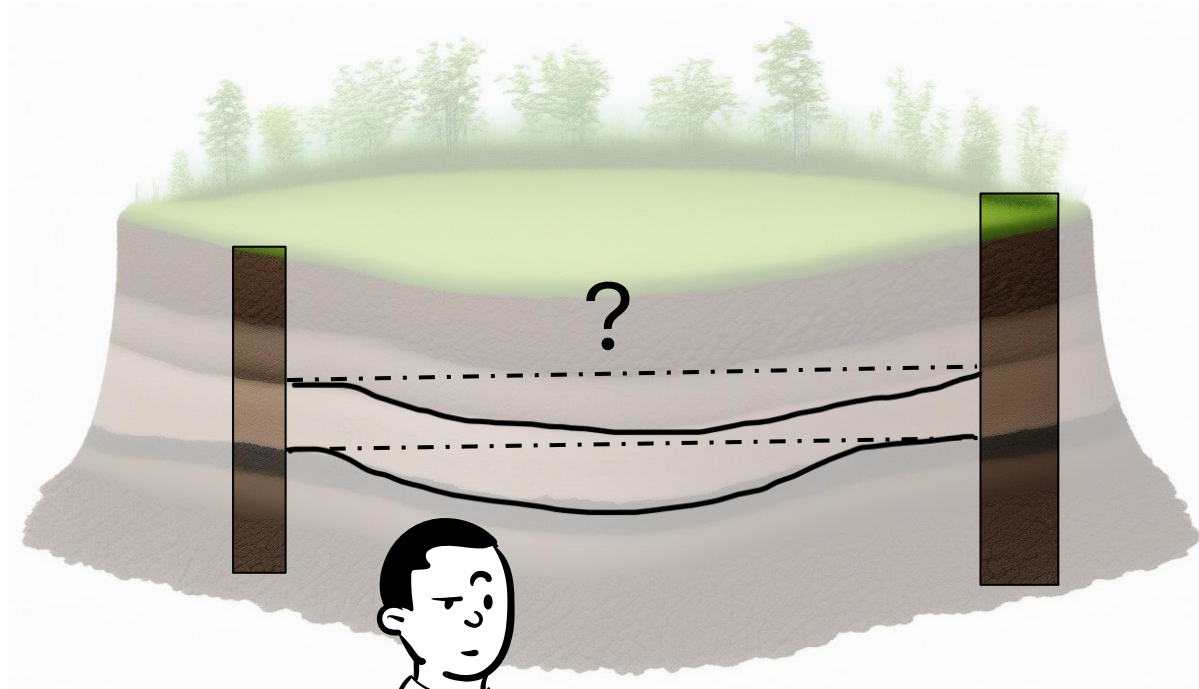
Impossible de connaître le sous-sol dans son entièreté.
Les incertitudes sont omniprésentes en géosciences.



Comprendre et identifier ces processus d'un point de vue statistique.

1. Point de vue de la géostatistique

Objectifs de la géostatistique



Forages (données ponctuelles);



Géophysiques;



Études antérieures;

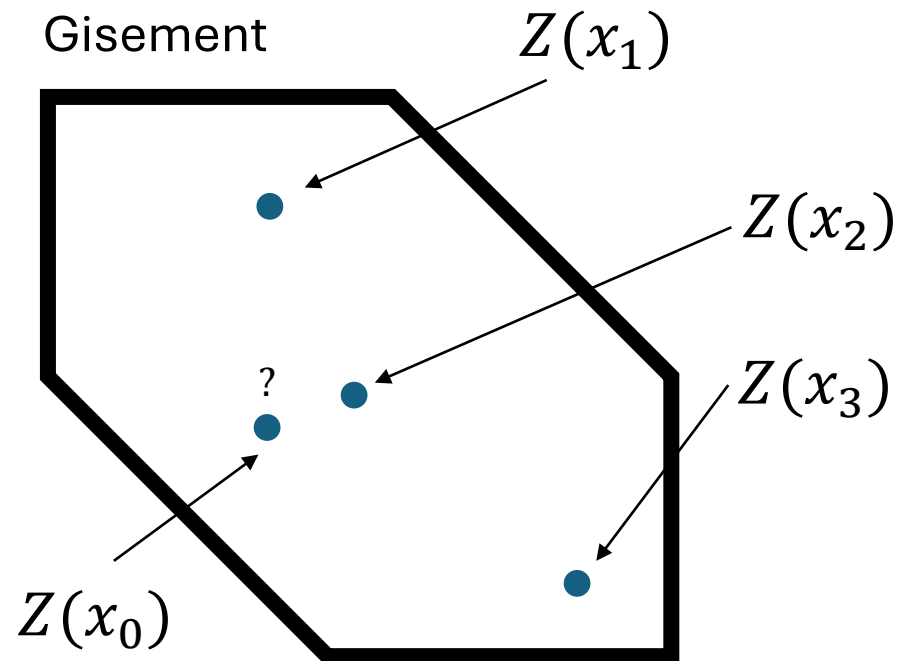


Base de données (SIGEOM, SIH, ...).

1. Point de vue de la géostatistique

Idée fondamentale : La nature n'est pas entièrement "imprévisible". Deux observations situées l'une près de l'autre devraient, en moyenne, se ressembler davantage que deux observations éloignées.

Attaquer le problème d'un point de vue de la géostatistique

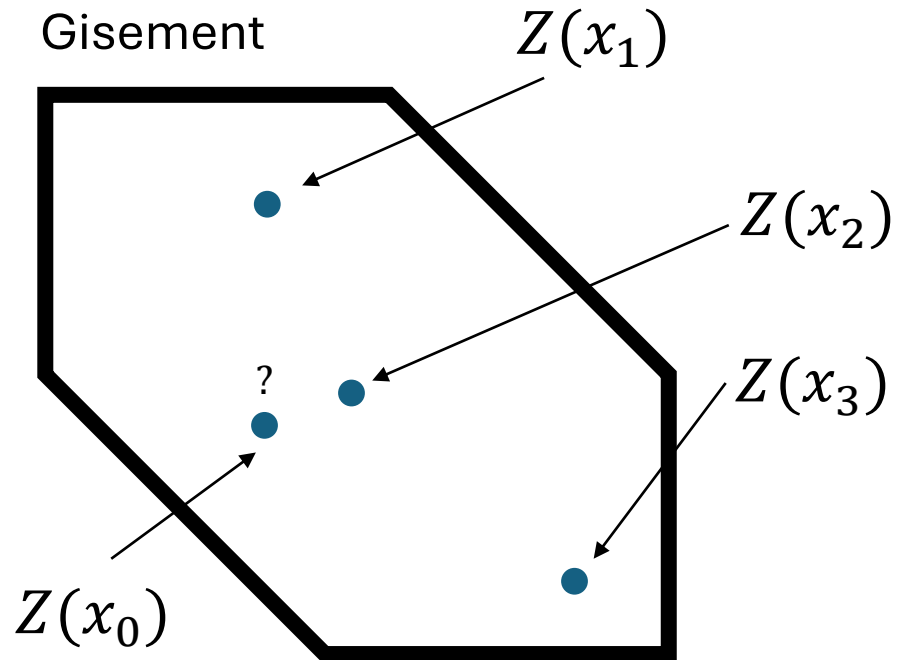


Un gisement est constitué de plusieurs points ('quasi-points')

- Chaque point représente une teneur
- Chaque teneur est une v.a. $Z_i(x_i)$
- L'ensemble des teneurs suit une loi multivariable, de dimension n , le nombre de points.

1. Point de vue de la géostatistique

Question : comment estimer $Z(x_0)$?

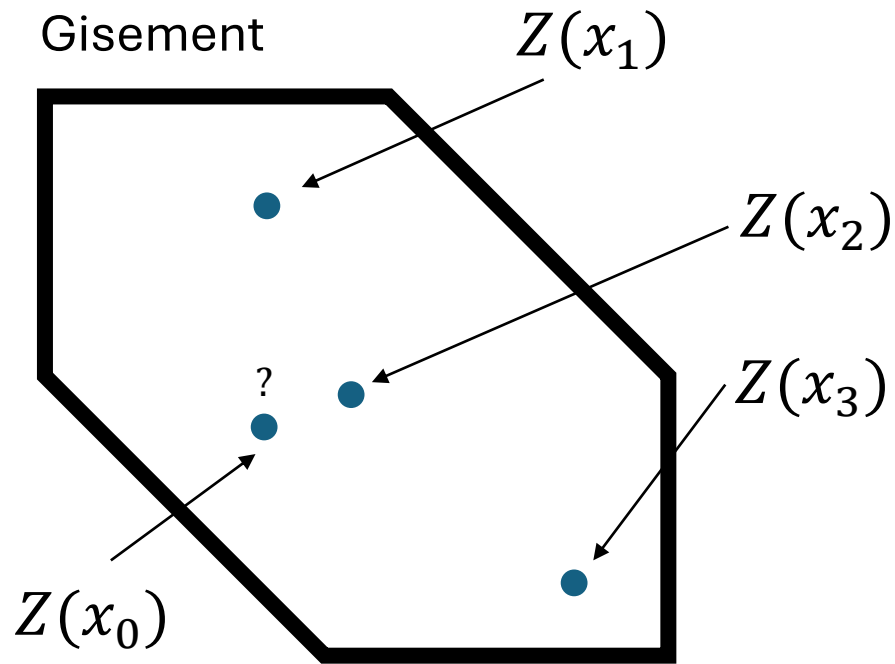


1. Estimer, à partir des données, la loi de densité conjointe
2. Estimer, à partir des données, la loi de densité multivariable
3. Estimer, à partir des données, la loi de densité d'une variable

Impossible, on ne connaît pas $Z(x_0)$ et on a qu'une observation des $Z(x_i), i = 1, \dots, 3$

1. Point de vue de la géostatistique

Question : comment estimer $Z(x_0)$?



Recours à des hypothèses → les statistiques **ne dépendent pas de x** . Elles **dépendent de la distance h** entre les points.

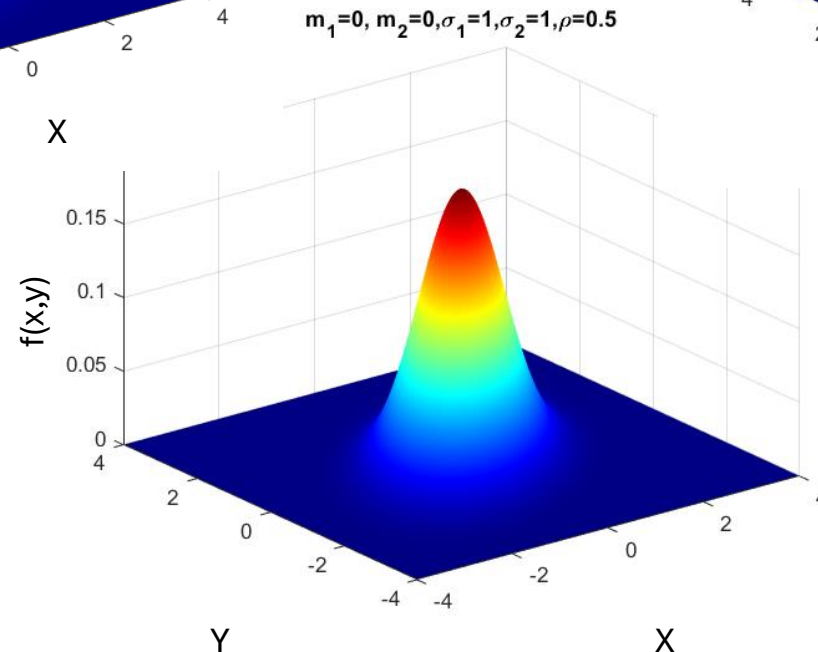
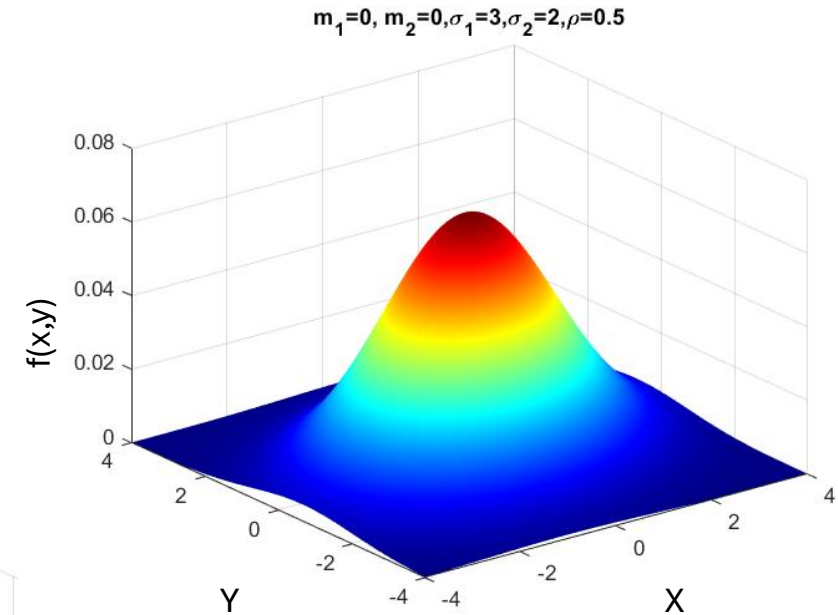
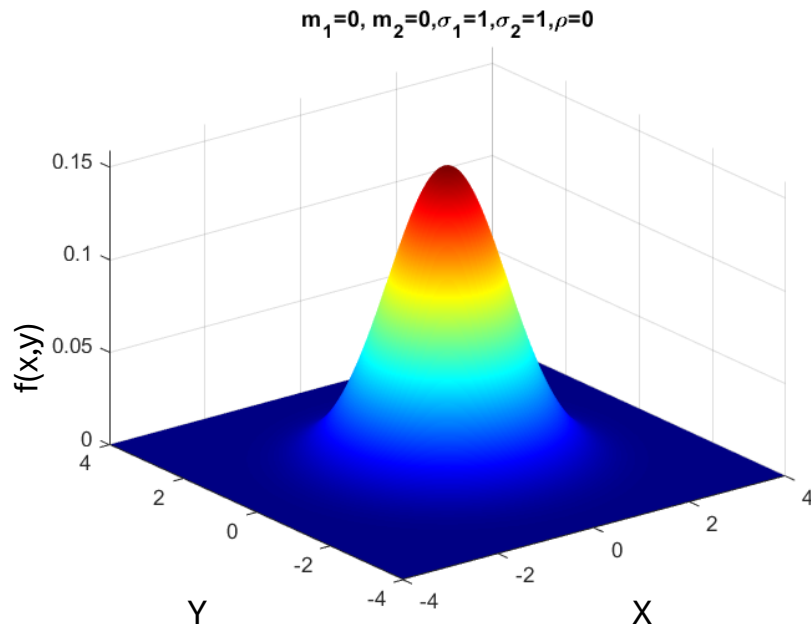
Stationnarité de second ordre

- $E[Z(x)] = m$
- $E[Z(x) - Z(x + h)] = 0$
- $Cov(Z(x), Z(x + h)) = C(h)$
- $\frac{1}{2} Var(Z(x) - Z(x + h)) = \gamma(h)$

* Ces hypothèses supposent une certaine régularité, une certaine homogénéité du gisement étudié. Si on peut reconnaître des zones très différentes géologiquement, on a habituellement intérêt à les traiter séparément.

1.1 Rappels

Deux variables aléatoires (X et Y) : caractérisé par la loi de densité conjointe $f(x,y)$



1.1 Rappels

Deux variables aléatoires (X et Y) : la distribution conjointe de deux v.a peut se résumer par

- La moyenne des deux variables (m_X et m_Y)
- L'écart-type (ou la variance) des deux variables (σ_X et σ_Y)
- La corrélation (ou covariance) entre les deux variables (ρ_{XY})

Covariance : mesure le degré d'association entre deux v.a

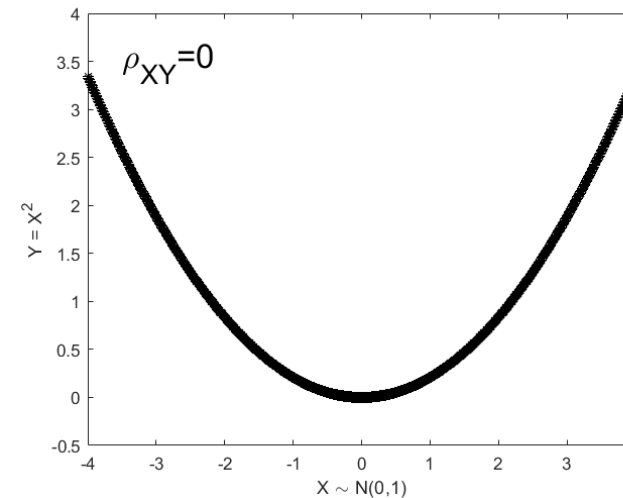
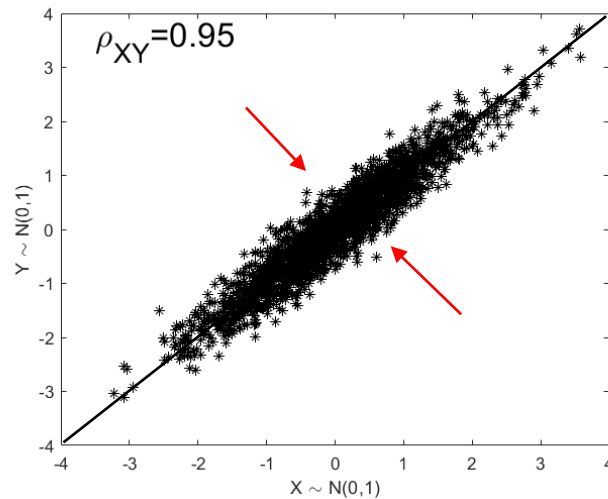
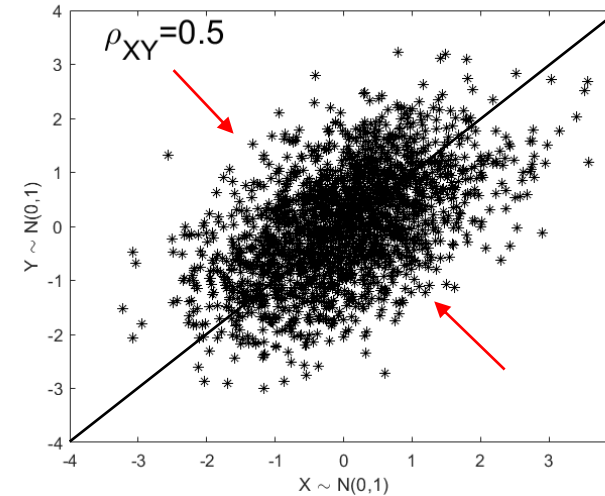
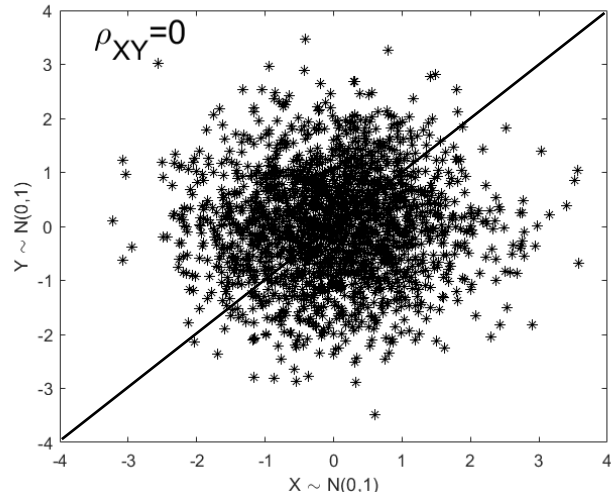
Corrélation (ρ) : covariance entre deux v.a normalisées pour présenter un écart-type de 1

$$\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} \in [-1, 1]$$



1.1 Rappels

Deux variables aléatoires (X et Y) : influence de la corrélation sur la distribution bivariable.



1.1 Rappels

Espérance mathématique

Définition :

- $E[X] = \int x f_X(x) dx$
- $E[g(X)] = \int g(x) f_X(x) dx$
- $E[g(X, Y)] = \int g(x, y) f_{XY}(x, y) dx dy$

Variance :

Si $g(X) = (X - m)^2$, alors $E[g(X)] = Var(X)$

Covariance :

Si $g(X, Y) = (X - m_x)(Y - m_y)$, alors $E[g(X, Y)] = Cov(X, Y)$

1.1 Rappels

Espérance mathématique et variance

Généralisation :

$$\text{Var} \left(c \int Z(x) dx \right) = c^2 \int \int \text{Cov}(Z(x), Z(y)) dx dy$$

← Très utile

$$\text{Cov} \left(a \int Z(x) dx, b \int Z(y) dy \right) = ab \int \int \text{Cov}(Z(x), Z(y)) dx dy$$

À noter :

- $\text{Cov}(X, X) = \text{Var}(X)$
- $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$ si X et Y sont indépendants
- $\text{Cov}(X, Y) = 0$ si X et Y sont indépendants

1.1 Rappels

Espérance mathématique et variance

Propriétés : opérateur linéaire

- $E[cg(X)] = cE[g(X)]$
- $E[g(X) + h(X)] = E[g(X)] + E[h(X)]$
- $E[g(X, Y) + h(X, Y)] = E[g(X, Y)] + E[h(X, Y)]$

En particulier, on obtient :

- $Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y)$
- $Var(aX + bY) = a^2Var(X) + b^2Var(Y) + 2abCov(X, Y)$

Généralisation :

$$Var\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n \sum_{j=1}^n a_i a_j Cov(X_i, X_j)$$



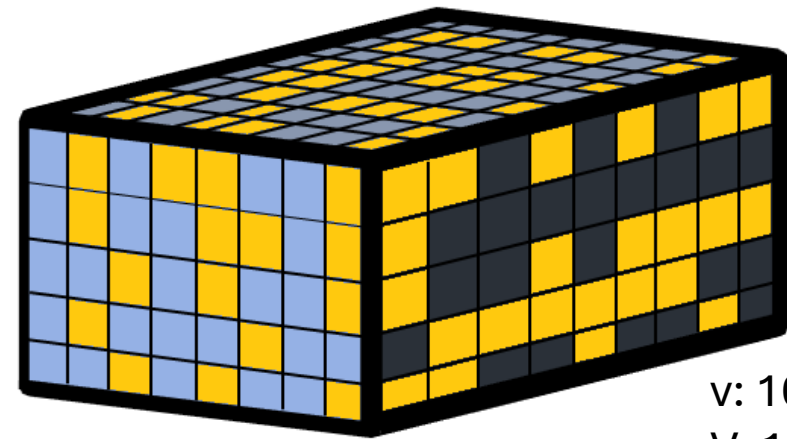
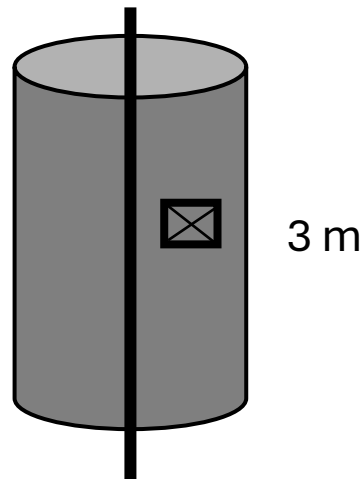
Très utile

1.1 Effet de support

Définition :

Support : est la **taille physique**, caractérisée par une **géométrie et une orientation**, du **volume** sur lequel une mesure de la variable régionalisée Z est effectuée

Variable régionalisée : est toute **fonction mathématique déterministe** destinée à **modéliser** un phénomène présentant une **structure** plus ou moins prononcée dans **l'espace et/ou le temps**



v: 10mx10mx5m

V: 120mx120x20m

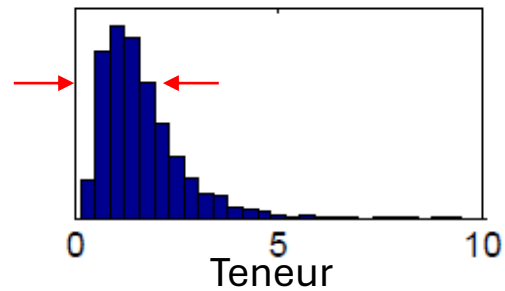
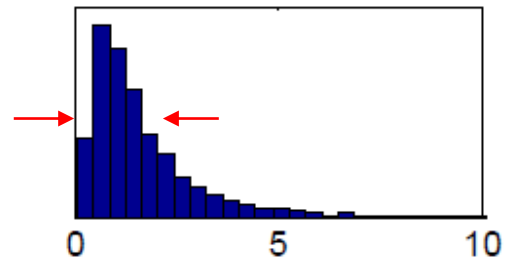
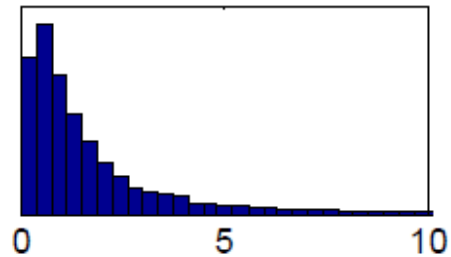
Exemple : Le support est crucial dans les travaux miniers, car il ne s'agit pas seulement d'identifier des points riches en minerais, mais de déterminer quels panneaux de taille fixe (définie par les techniques employées) ont probablement une teneur en minerais répondant à une valeur de coupure définie par des contraintes techniques et économiques.

1.1 Effet de support

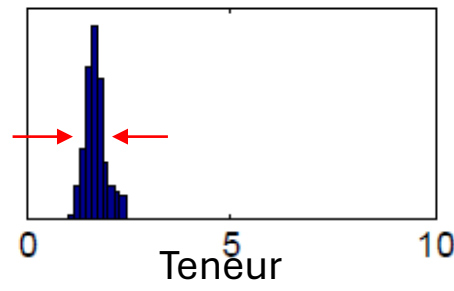
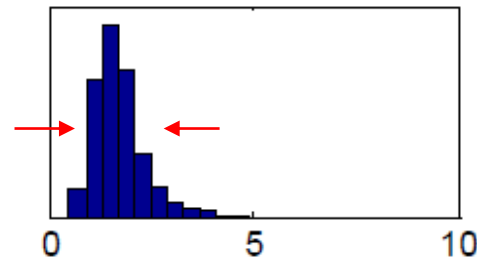
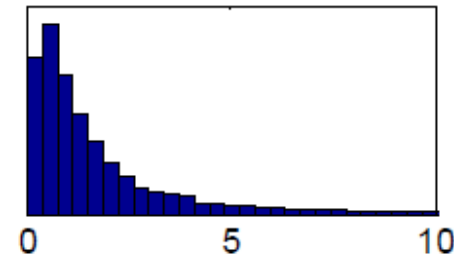
Effet de support



Gisement A



Gisement B



Qu'observe-t-on de particulier ?

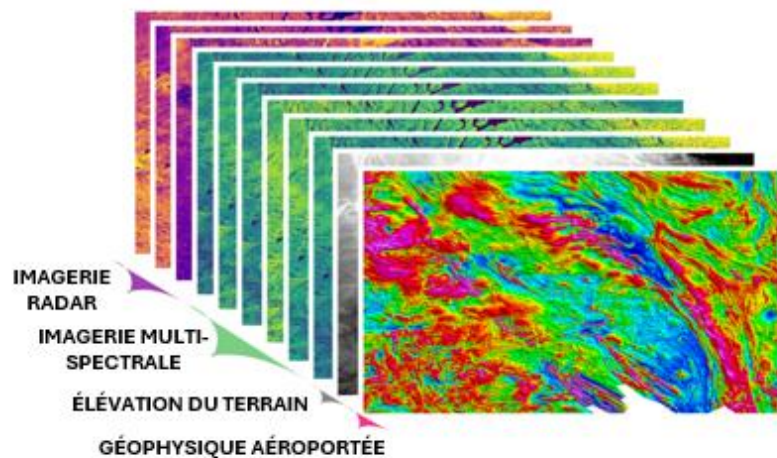
Pourquoi ? Explication ?

1.1 Effet d'information

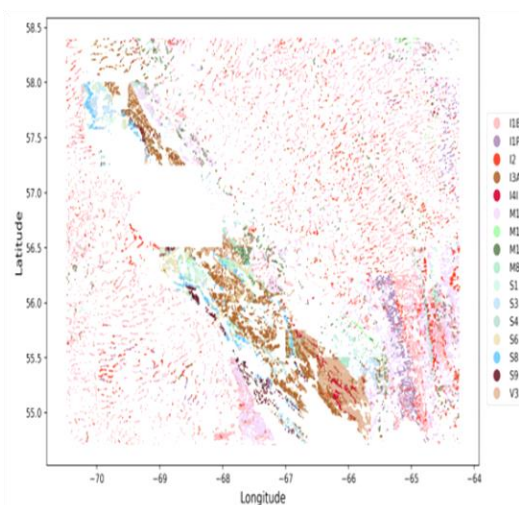
Définition :

Information : Indication, renseignement, précision que l'on donne ou que l'on obtient sur quelqu'un ou quelque chose

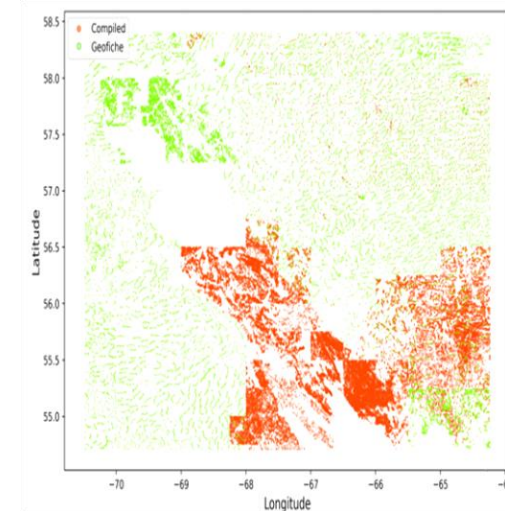
Variables auxiliaires



Données lithologiques

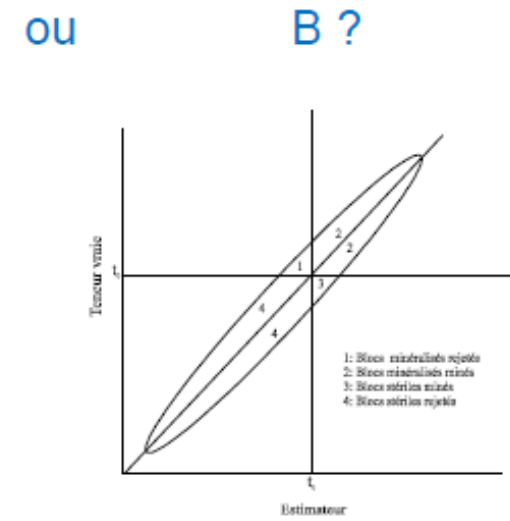
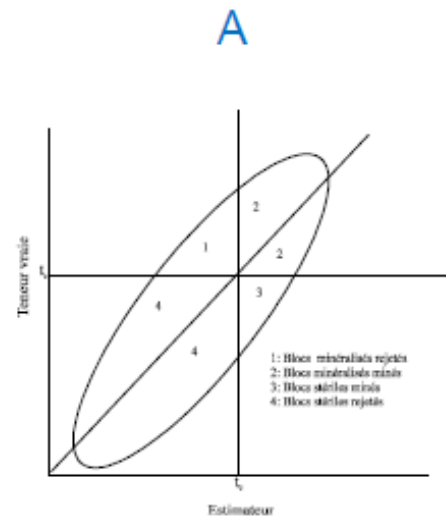
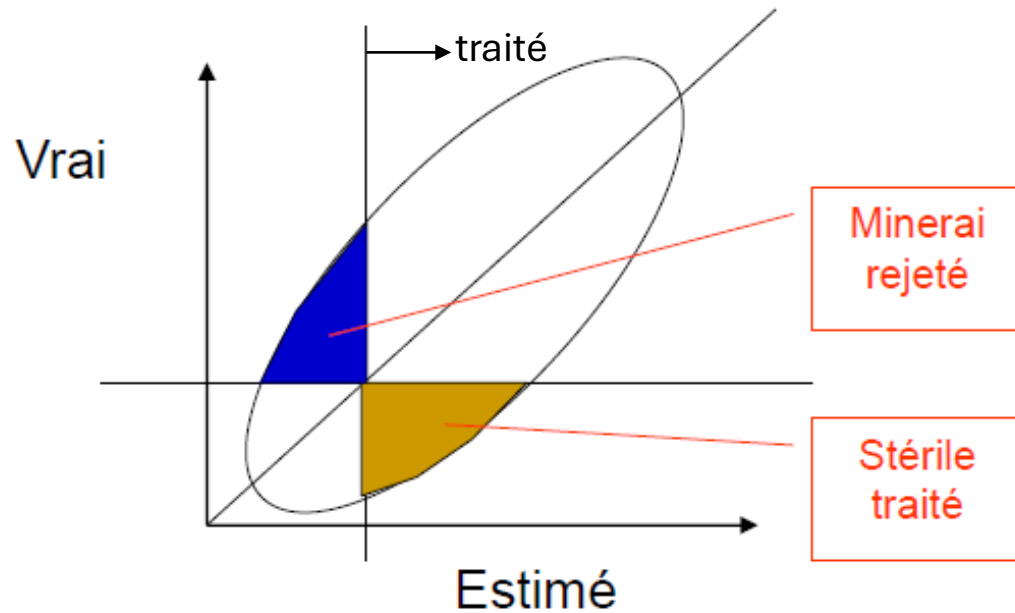


Données géochimiques



On prend nos décisions à partir d'estimations, mais on récolte des valeurs vraies.

1.2 Effet d'information



Peut-on prévoir les plages d'erreur et les pertes en \$?
Augmenter l'information disponible (\$) pour réduire les pertes ?

* La précision dépend du nombre de données, mais aussi des caractéristiques du gisement.

2. Géostatistique linéaire

Pour une même quantité de matière extraite, statistiquement parlant :

→ on récupère toujours moins de métal avec de gros blocs qu'avec des petits blocs (**effet support**)

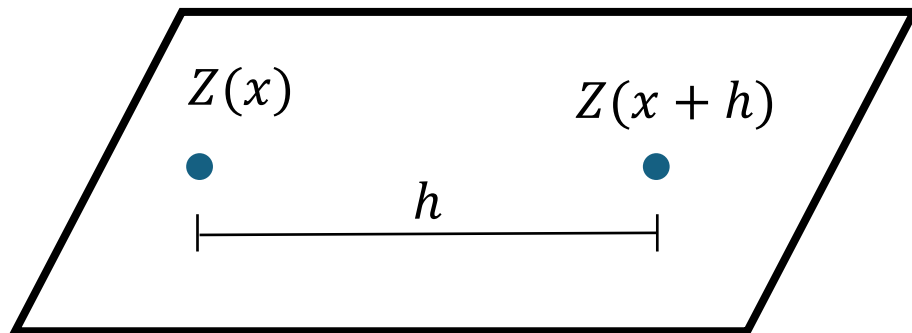
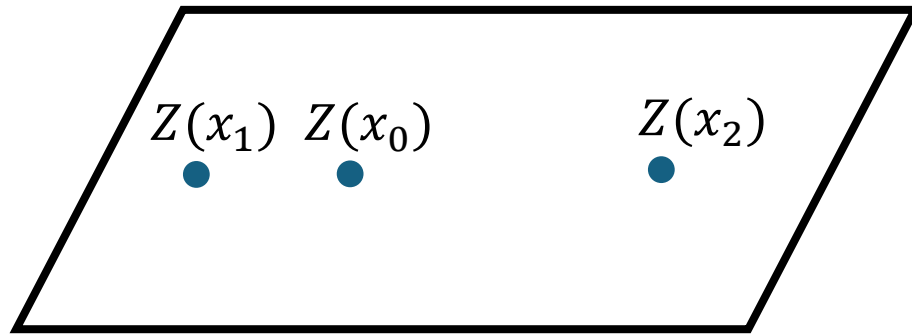
→ on récupère toujours moins de métal avec des estimations qu'avec les vraies valeurs (**effet information**)

La géostatistique permet théoriquement de

- prévoir l'ampleur de ces effets
- Minimiser ces effets
- Prendre des décisions éclairées au vu de ces effets

2.1 Hypothèse de stationnarité

Gisement



Stationnarité du second ordre

L'espérance mathématique ne dépend pas de la localisation x

- $E[Z(x)] = m$
- $E[Z(x) - Z(x + h)] = 0$

La covariance ne dépend que de h

- $Cov(Z(x), Z(x + h)) = C(h)$, covariogramme
- $\frac{1}{2} Var(Z(x) - Z(x + h)) = \gamma(h)$, variogramme
- $Var [Z(x)] = C(0)$, constante

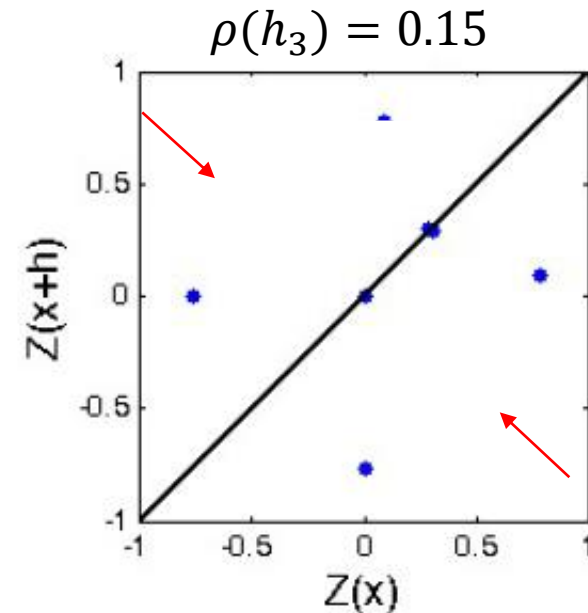
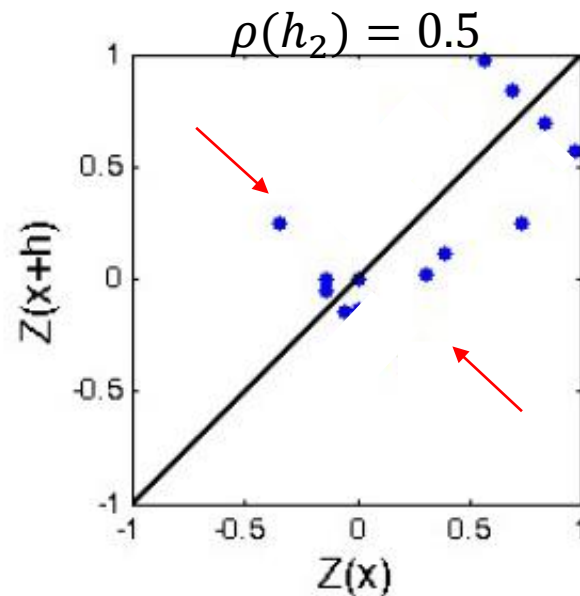
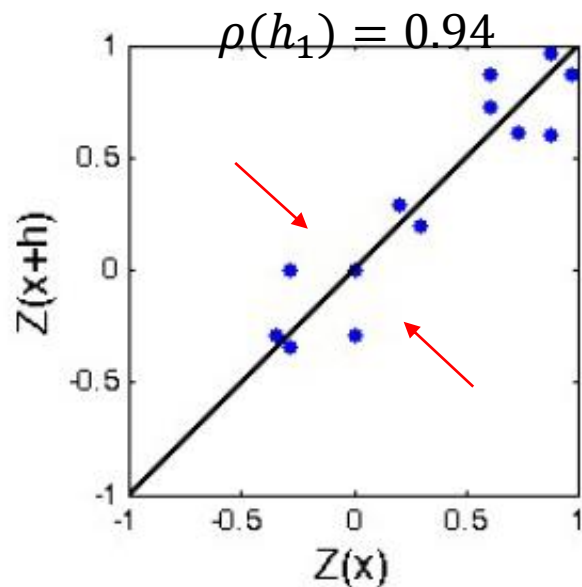
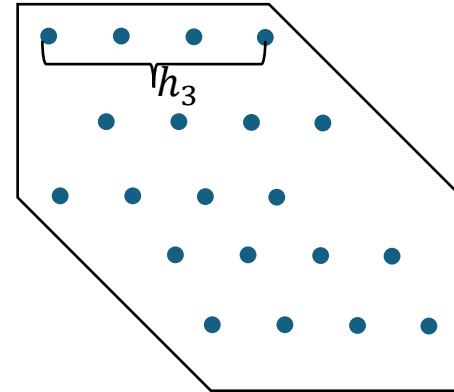
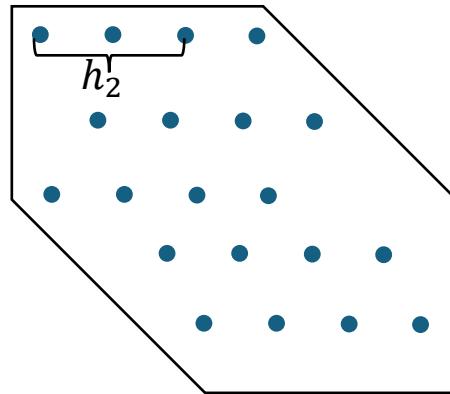
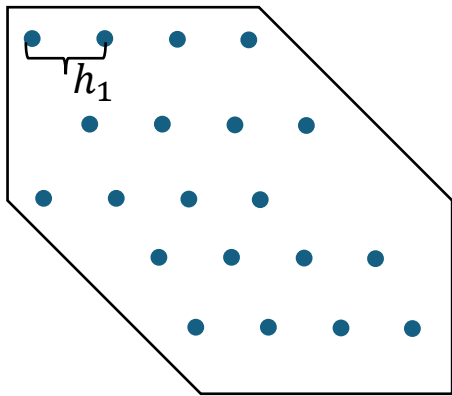
$\gamma(h) \rightarrow$ fonction la plus utilisée pour caractériser la continuité d'une minéralisation

2.1 Hypothèse de stationnarité

Covariance et corrélation

$$C(h) = \text{Cov}(Z(x), Z(x+h)) \\ = E[(Z(x) - m)(Z(x+h) - m)]$$

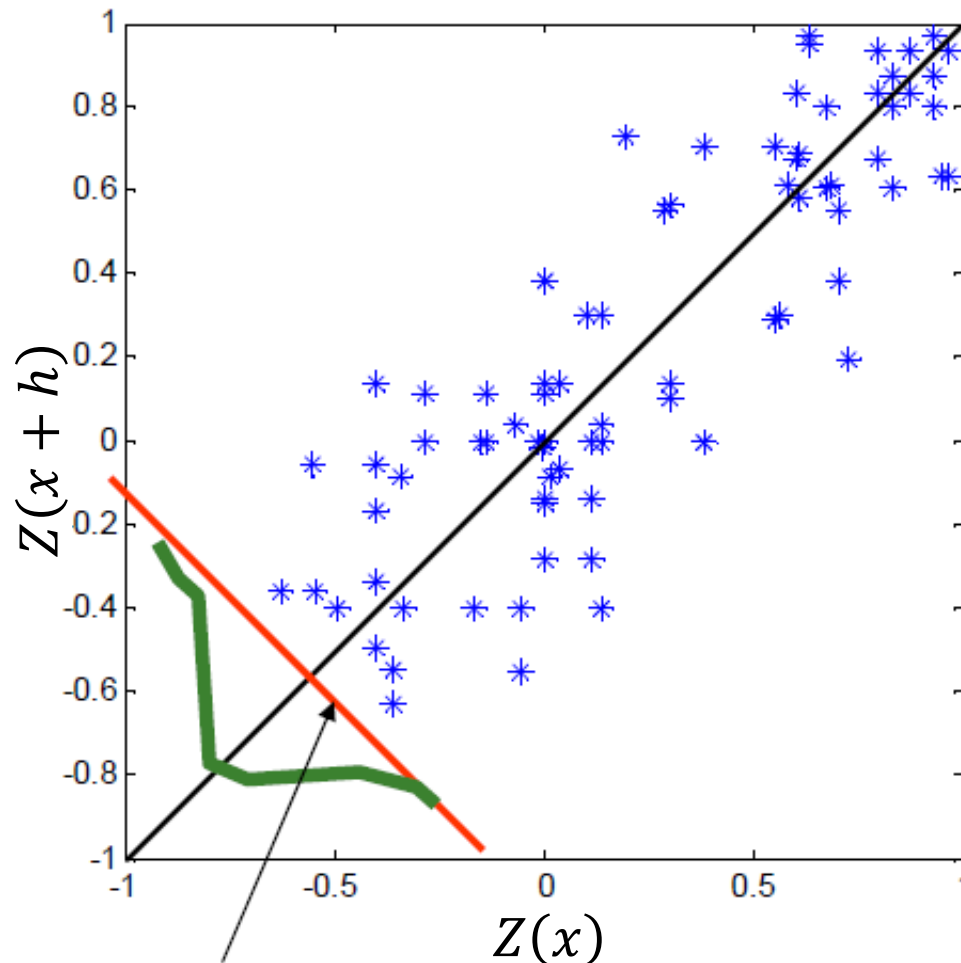
Gisement



Généralement
si $h \uparrow$, $C \downarrow$, $\rho \downarrow$

2.1 Hypothèse de stationnarité

Variogramme



Mesure la dispersion sur cette droite

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

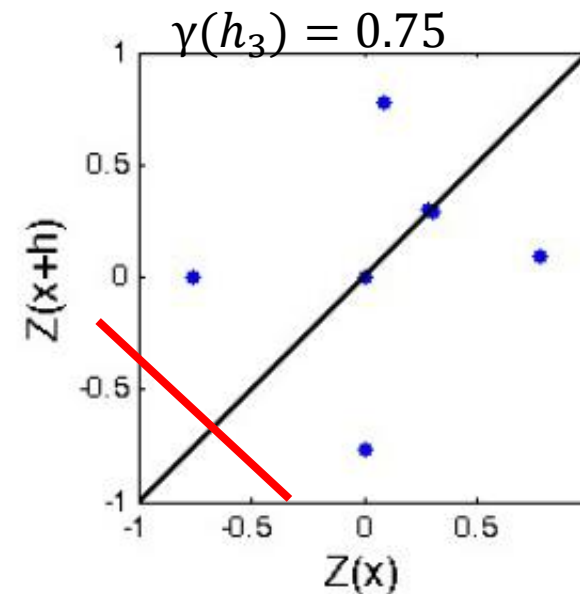
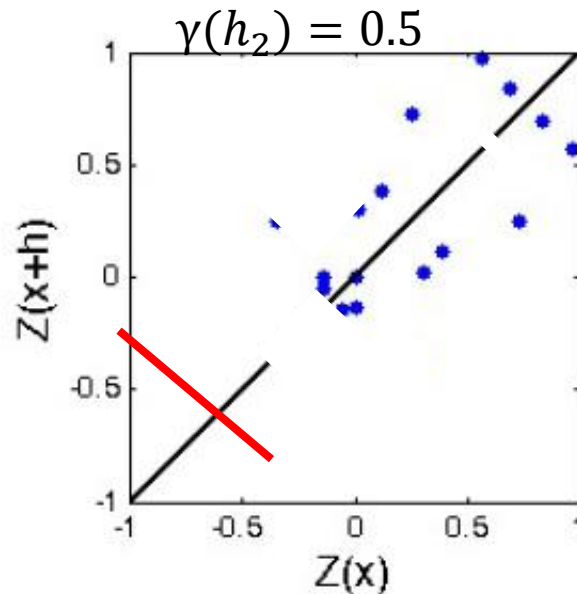
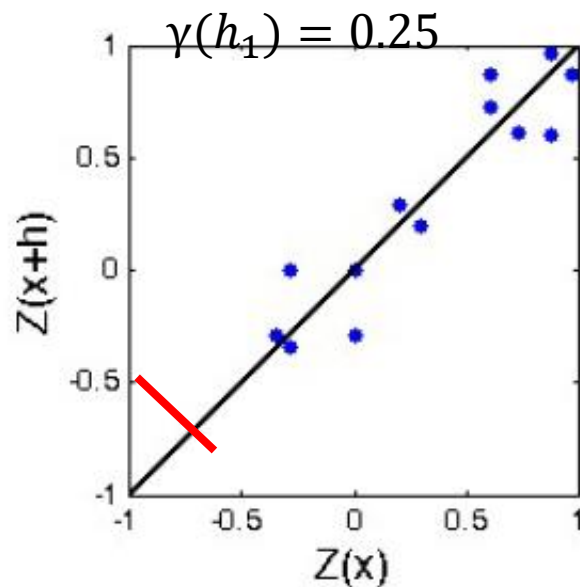
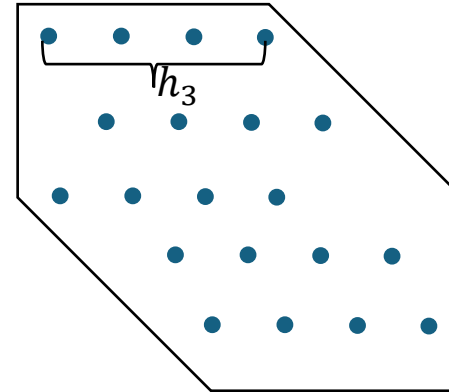
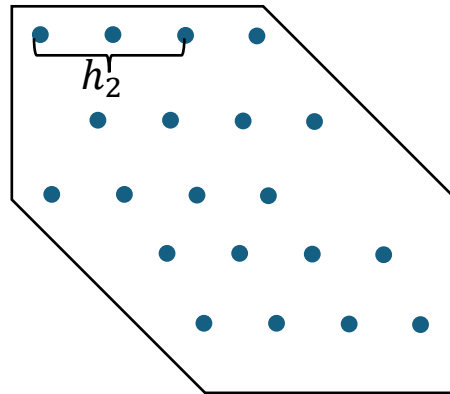
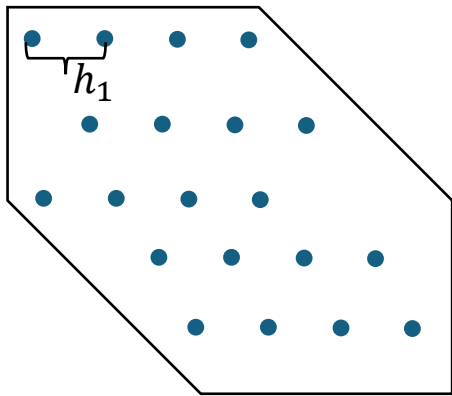
Généralement
si $h \uparrow, \gamma \uparrow$

2.1 Hypothèse de stationnarité

Variogramme

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

Gisement



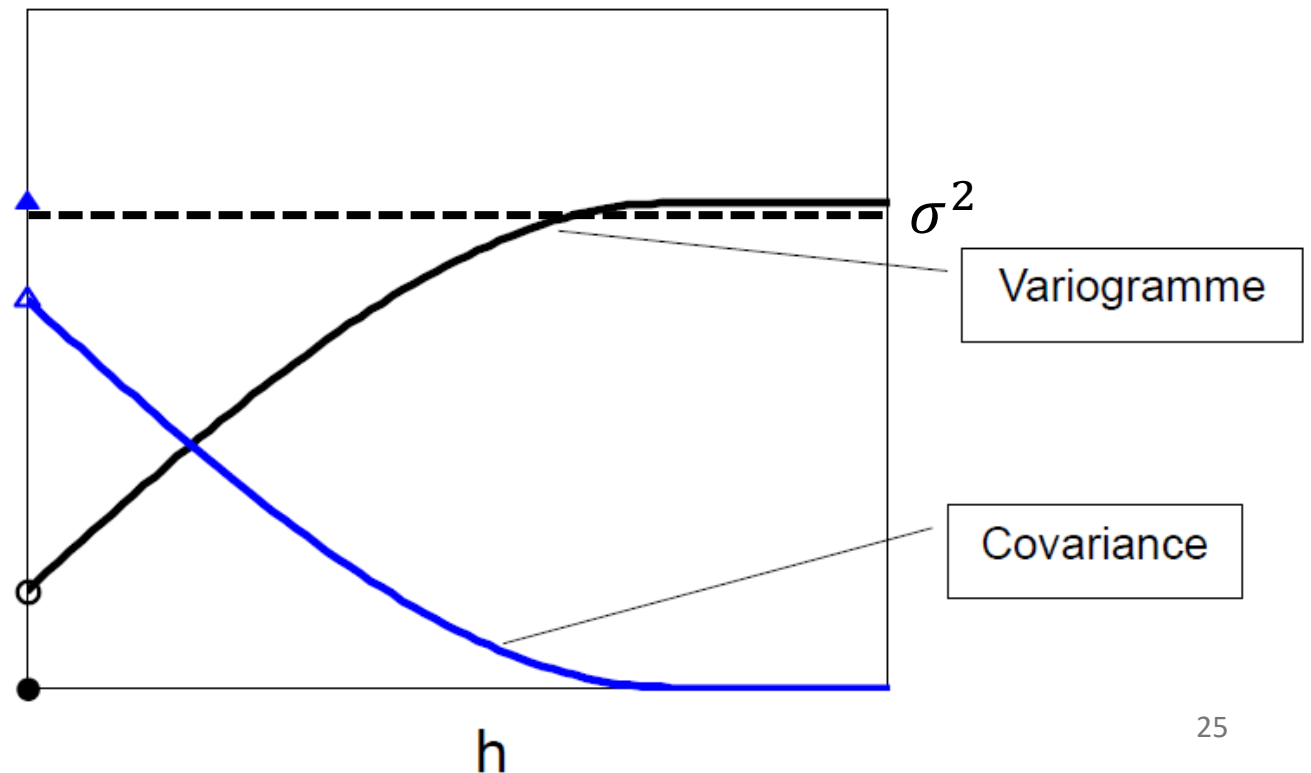
Généralement
si $h \uparrow, \gamma \uparrow$

2.1 Hypothèse de stationnarité

Liens entre covariance et variogramme

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

Uniquement vrai si le variogramme présente un palier, sinon σ^2 n'est pas défini



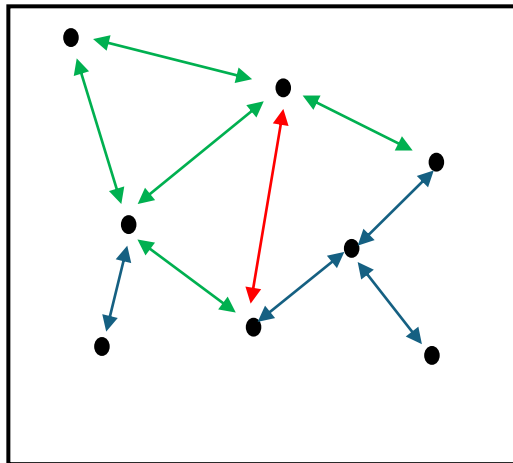
2.1 Variogramme expérimental

Objectifs de la géostatistique

La **géostatistique**, qui étudie les phénomènes continus et catégoriels, a permis le développement de méthodes spécifiques pour **étudier les relations spatiales entre les observations** et construire des **outils prédictifs**.

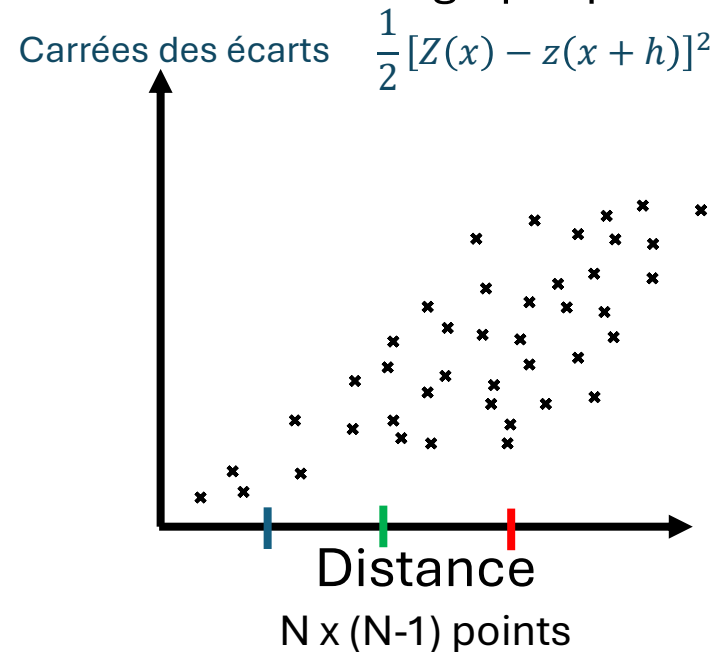
Nuée variographique

Plan de position

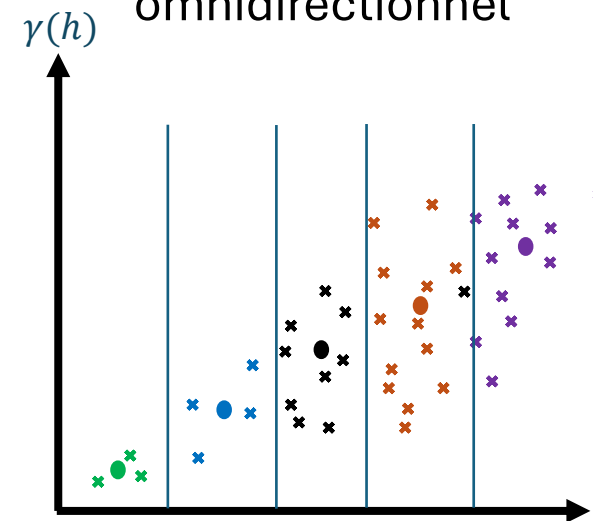


N points de données de la propriété Z

Nuée variographique



Variogramme omnidirectionnel



2.1 Variogramme expérimental

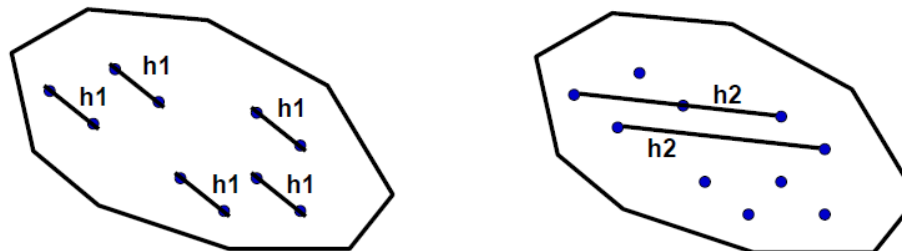
Variogramme expérimental

$$\gamma_e(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} (Z(x_i) - Z(x_i + h))^2$$

$N(h)$: nombre de paires de points dans la classe h considérée

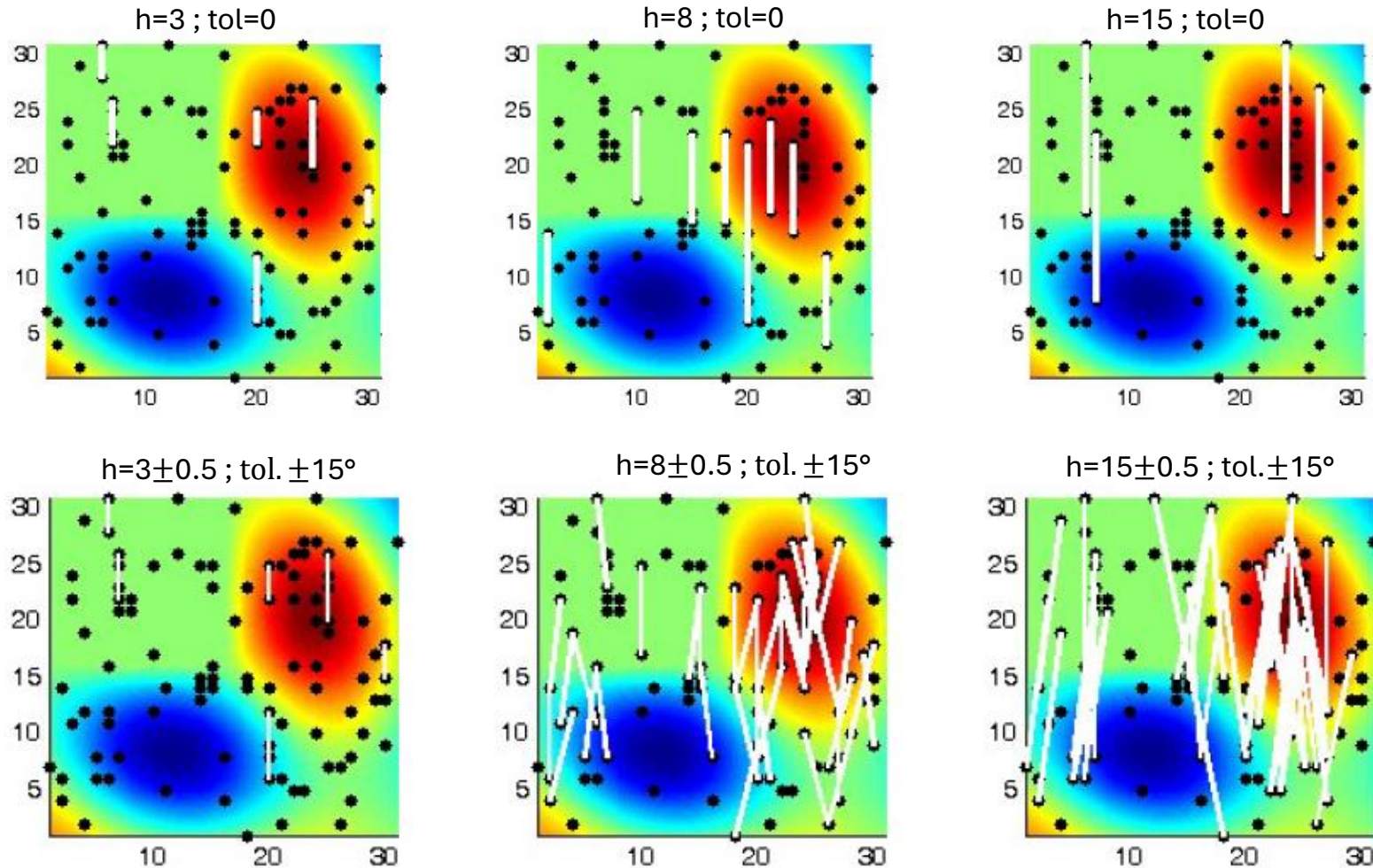
Pour calculer le variogramme expérimental, il faut :

1. Choisir une direction h et une tolérance angulaire
2. Discrétiser h en classes distinctes
3. Répartir les paires de points $(x_i, x_i + h)$ dans les classes



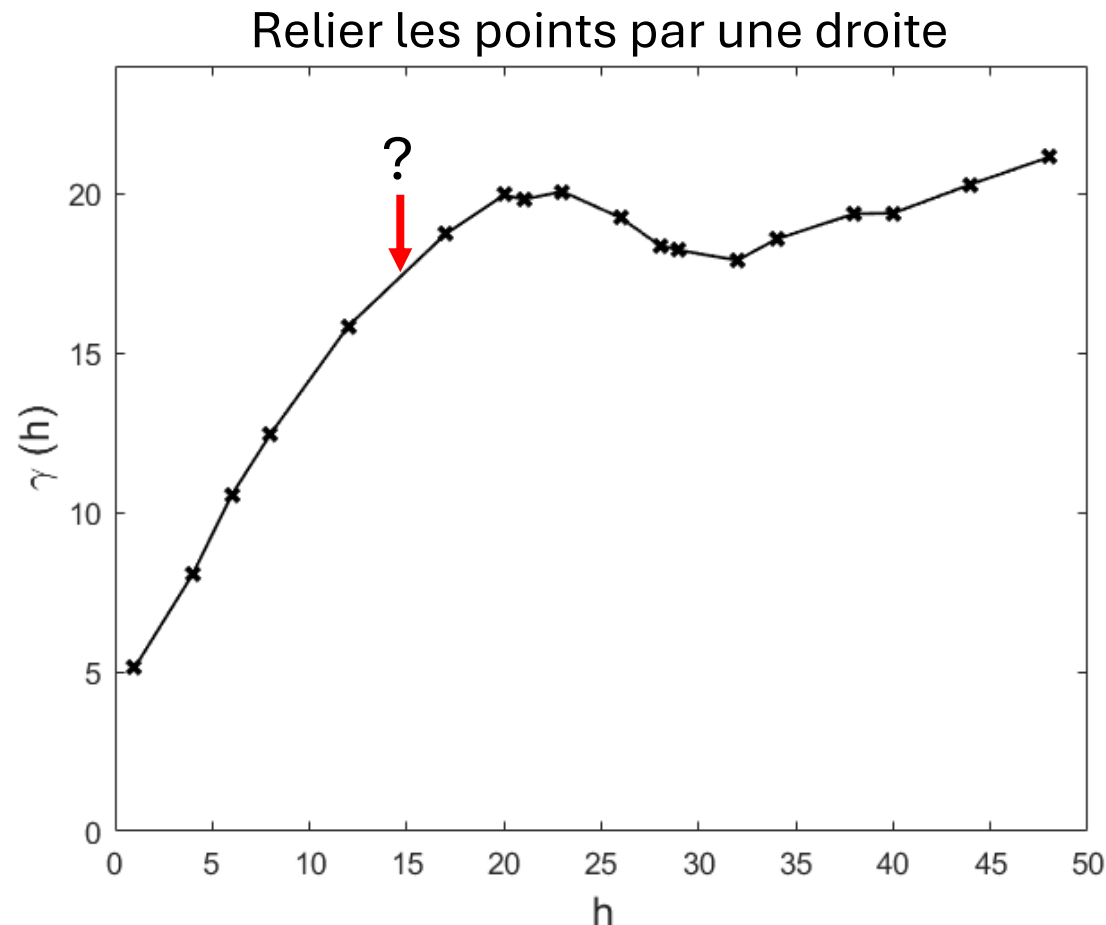
2.1 Variogramme expérimental

Variogramme expérimental : 2D



2.1 Variogramme expérimental

Comment déterminer le variogramme à des distances non observées ?



Une solution si simple,
mais malheureusement
erronée !!!

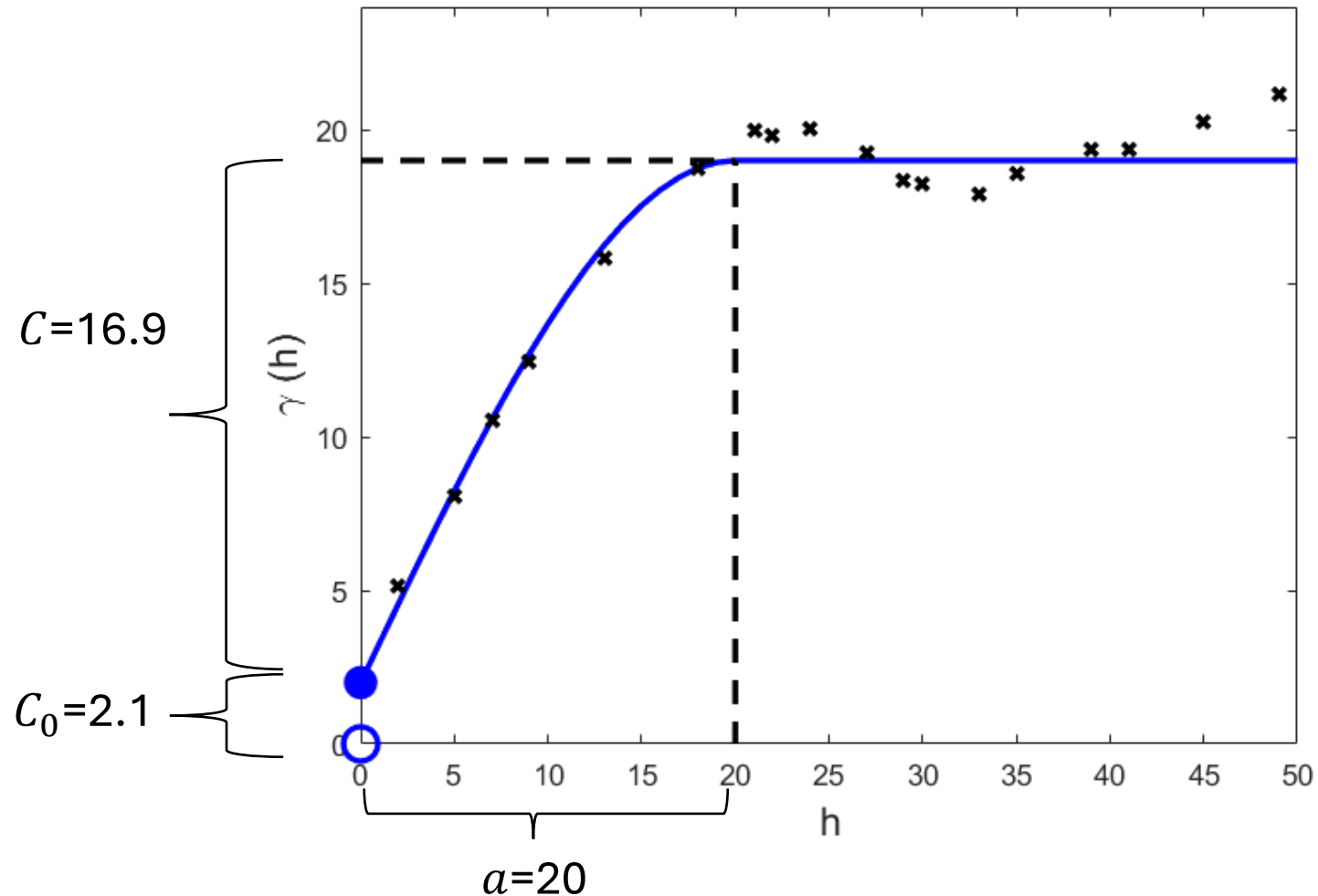
↓

Vers un modèle
théorique

2.2 Variogramme théorique

Il faut utiliser un modèle démontré admissible :

Assurer que toute variance calculée à partir du modèle soit positive



3. Modélisation du variogramme théorique

Un modèle théorique est requis afin d'assurer que toutes variances calculées à partir du modèle sont nulles ou positives.

Soit une combinaison linéaire $\sum_{i=1}^n \lambda_i Z_i$, il faut démontrer pour $\forall n, \forall \lambda_i$ et $\forall Z_i$ que :

$$\text{Var} \left(\sum_{i=1}^n \lambda_i Z_i \right) = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \text{Cov}(Z_i, Z_j) = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j C(h_{i,j}) \geq 0$$

OUCH!!!

FIOU !!

Utilisation de modèles éprouvés

Note 1 : C est nommé la matrice de covariance et elle fournit la covariance entre chaque paire d'éléments. Il suffit de démontrer que C est semi-défini positive pour vérifier l'admissibilité d'un modèle pour toutes les combinaisons linéaires possibles.

Note 2: La relation est valide seulement si le variogramme présente un palier. Sinon, l'équation à démontrer est légèrement plus complexe. Il faut stipuler que $\sum_{i=1}^n \lambda_i = 0$, et démontrer que $\text{Var}(\sum_{i=1}^n \lambda_i Z_i) = -\sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \gamma(h_{i,j}) \geq 0$.

3. Modélisation du variogramme théorique

Densité spectrale

Une **fonction de densité spectrale est admissible** si on peut démontrer qu'elle est **positive (ou nulle) pour chaque s** , garantissant ainsi que la **covariance directe associée est admissible**.

$$C(h) = \int_{\mathfrak{R}^d} e^{i(s \cdot h)} F(ds)$$

où \cdot est le produit scalaire et $F(ds)$ est la mesure spectrale.

Lorsque $C(h)$ est de carré intégrable, la mesure spectrale peut s'écrire sous la forme d'une densité spectrale, $F(ds) = f(s)ds$.

Si $C(h)$ est absolument intégrable, la densité spectrale peut être obtenue par

$$f(s) = \frac{1}{(2\pi)^d} \int_{\mathfrak{R}^d} e^{-i(s \cdot h)} C(h) dh$$

3. Modélisation du variogramme théorique

Lorsqu'un variogramme **présente un seuil**, il est défini par :

- **La portée** (a) : distance à partir de laquelle deux points de l'espace ne sont plus liés un à l'autre (covariance nulle, $C(h > a) = 0$). Pour un modèle asymptotique, la covariance tend vers zéro ($C(h > a) \rightarrow 0$).
- **Le palier** ($\sigma^2 = C_0 + \sum_i C_i$): variance de la variable aléatoire $Z(x)$.
- **L'effet de pépité** (C_0): variation à très courte échelle, erreurs de mesure, de localisation, de précision ou d'analyse

3. Modélisation du variogramme théorique

Construction de modèles admissibles :

- Toute somme (coefficients positifs) de modèles de variogramme est admissible
- Toute somme (coefficients positifs) de modèles de covariance est admissible
- Tout produit (coefficients positifs) de modèles de covariance est admissible
- Chaque modèle peut être isotrope ou anisotrope, les directions d'anisotropie peuvent varier d'un modèle à l'autre
- **IMPORTANT** : Un modèle peut être admissible en 1D et non-admissible en 2D, 3D,...,n-D. Par contre, s'il est admissible en n-D il l'est aussi en (n-1)-D jusqu'à 1D.

3.1 Modèle de variogramme admissible

Effet de pépité :

$$\gamma(h) = \begin{cases} 0 & \text{si } h = 0 \\ C_0 & \text{si } h > 0 \end{cases}$$

Linéaire:

$$\gamma(h) = b h$$

Sphérique :

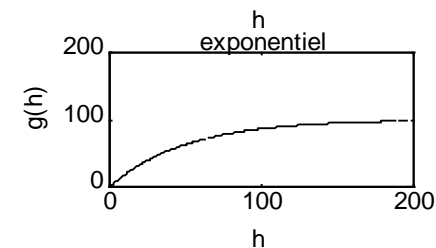
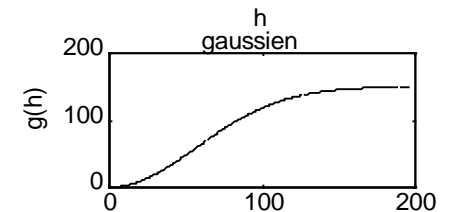
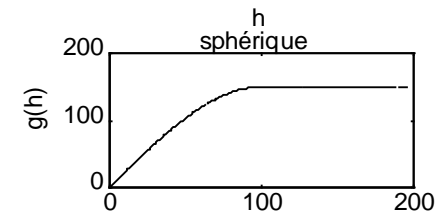
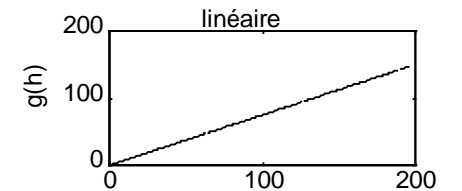
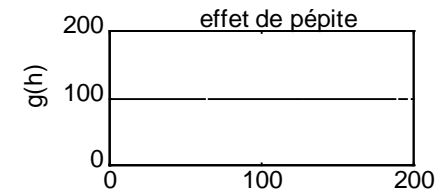
$$\gamma(h) = \begin{cases} C [1.5 h/a - 0.5 (h/a)^3] & \text{si } 0 < h < a \\ C & \text{si } h \geq a \end{cases}$$

Gaussien:

$$\gamma(h) = C [1 - \exp(-3(h/a)^2)]$$

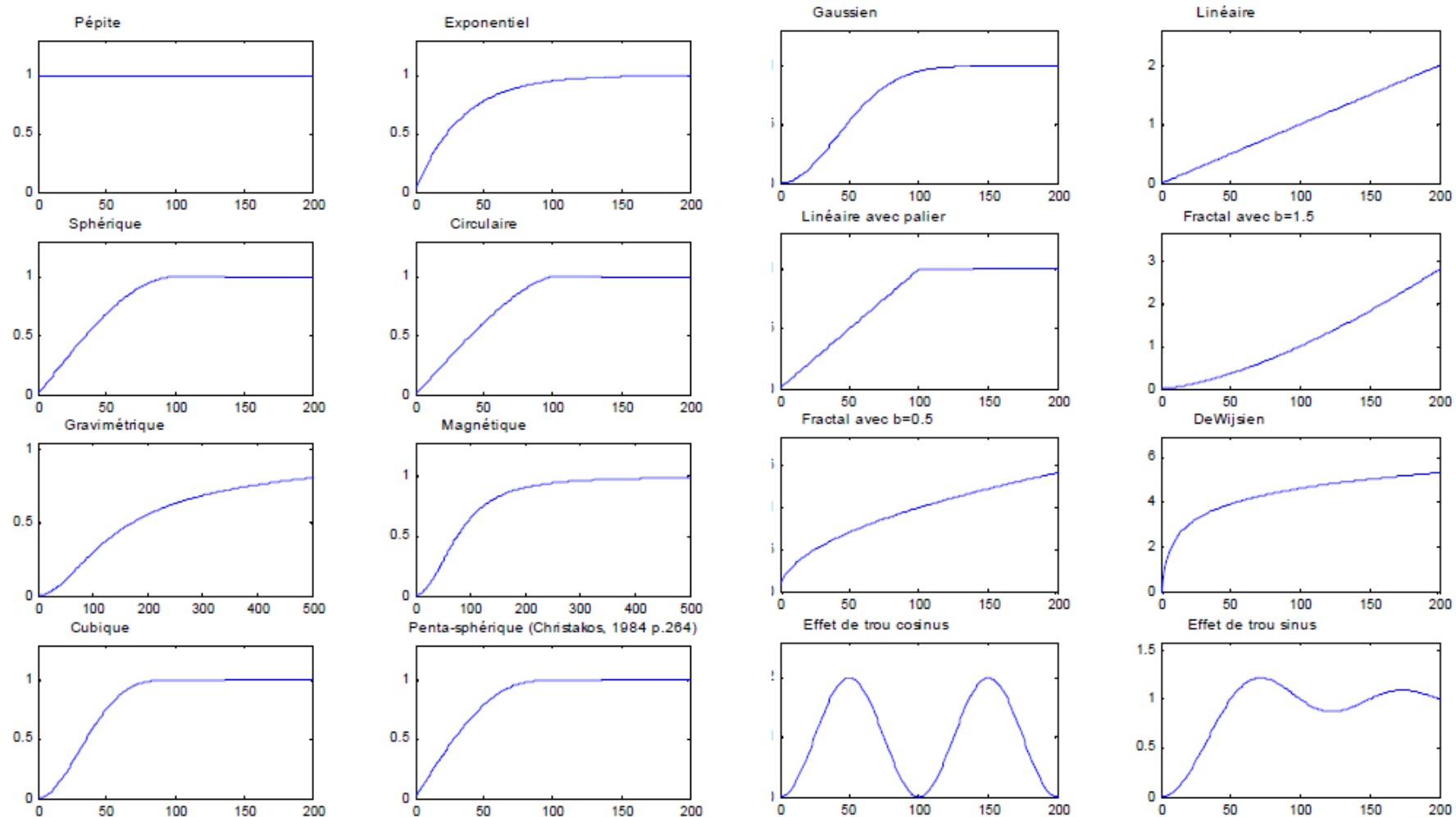
Exponentiel

$$\gamma(h) = C [1 - \exp(-3h/a)]$$



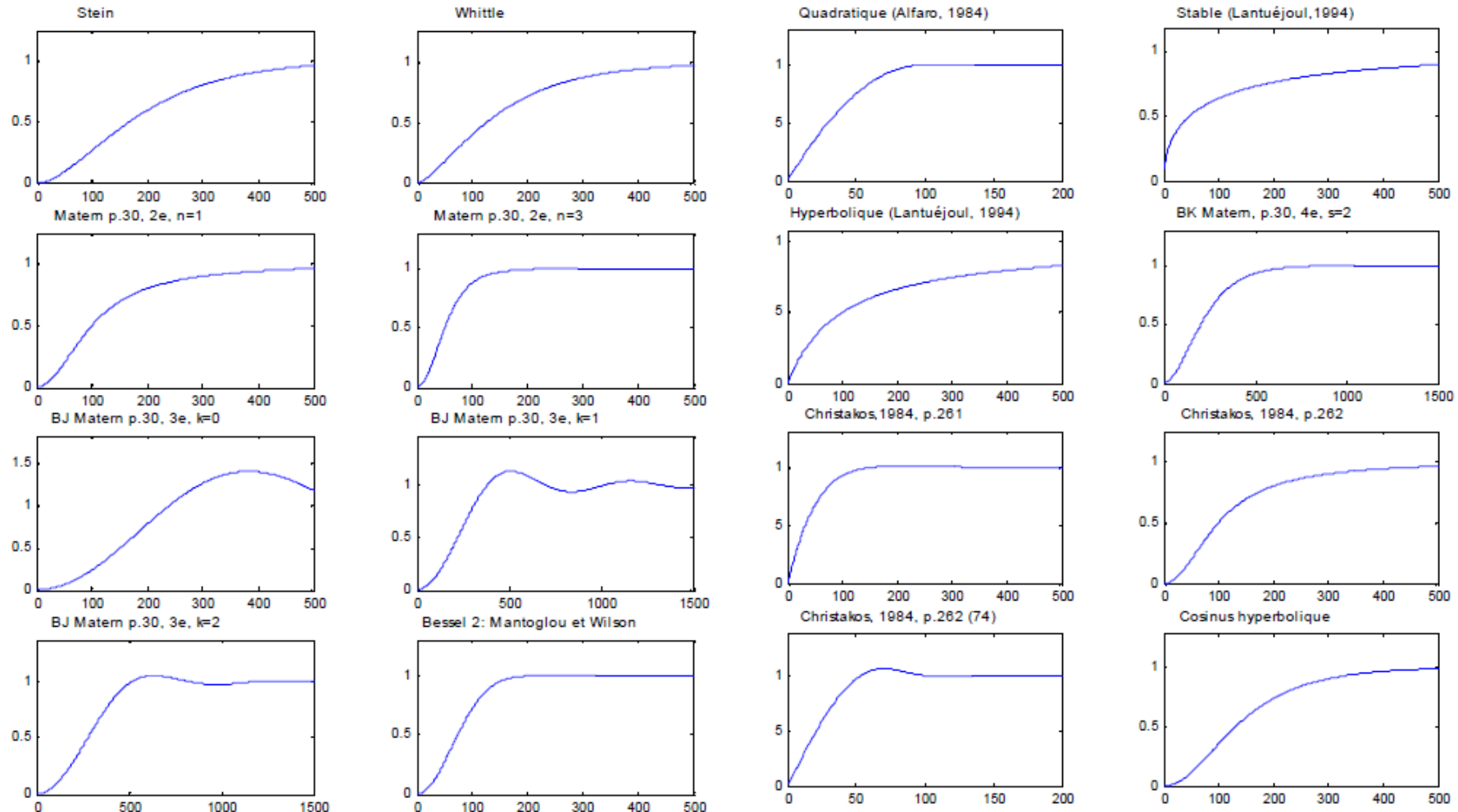
3.1 Modèle de variogramme admissible

Il existe plusieurs modèles admissibles. En voici une liste :



3.1 Modèle de variogramme admissible

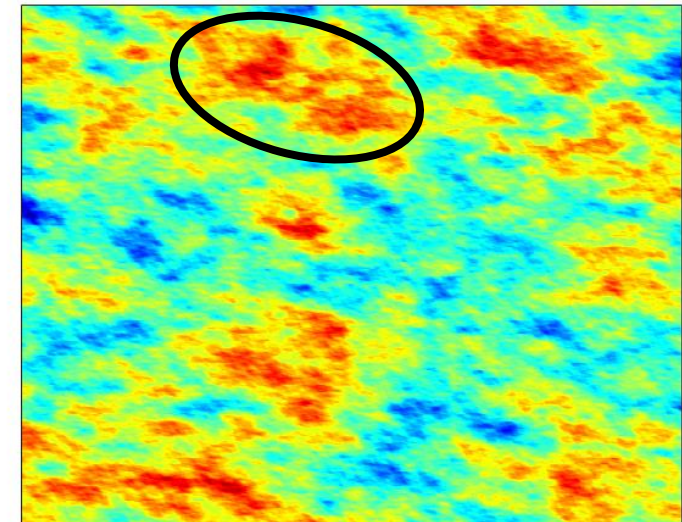
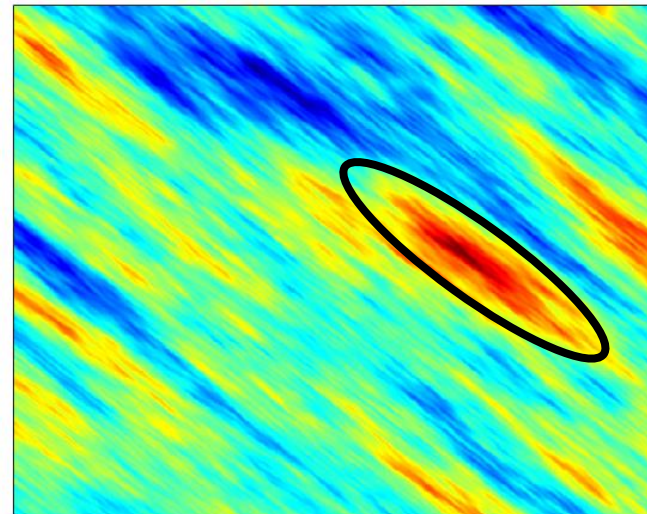
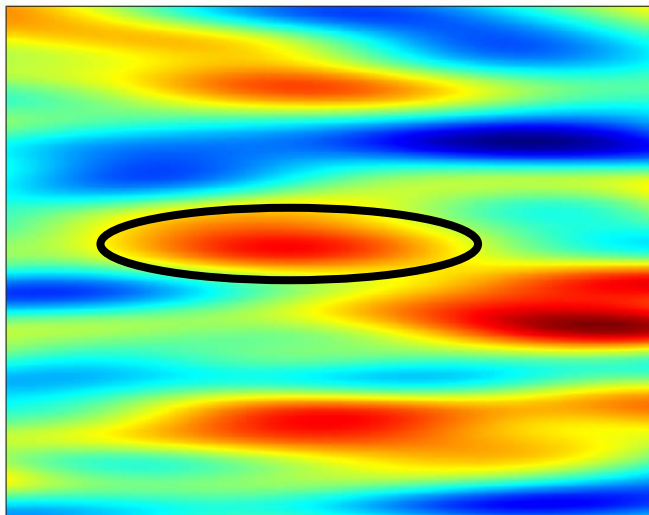
Il existe plusieurs modèles admissibles. Voici une autre liste :



3.2 Anisotropies

Anisotropies géométriques :

Il n'est pas rare que la continuité spatiale présente des directions préférentielles. Le contexte de formation des zones minéralisées favorise une direction au détriment d'une autre (veines, intrusions, sédimentations ...).

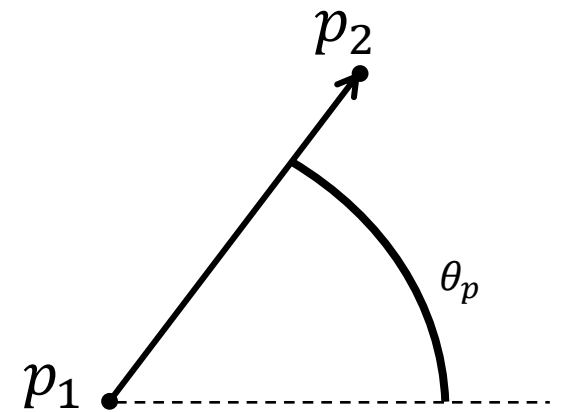
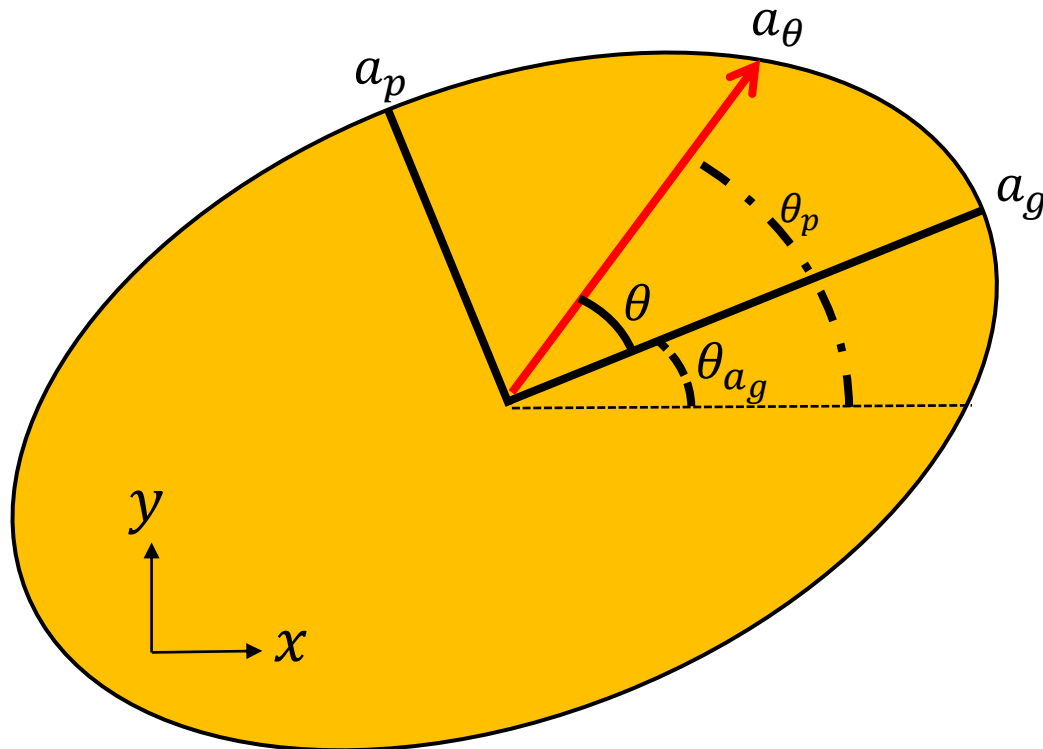


On utilise généralement un ellipsoïde pour caractériser l'anisotropie

3.2 Anisotropies

Anisotropies géométriques

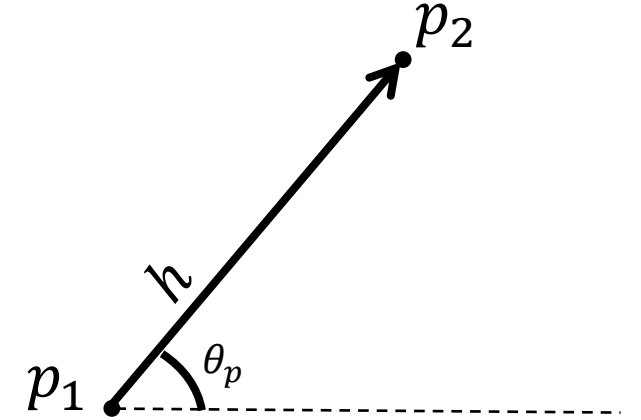
On définit la portée dans une direction $\overrightarrow{p_1 p_2}$ par :
$$a_\theta = \frac{a_g a_p}{\sqrt{a_p^2 \cos^2 \theta + a_g^2 \sin^2 \theta}}$$



3.2 Anisotropies

Anisotropies géométriques : calcul du variogramme pour deux points

1) déterminer la distance (h) et l'angle (θ_p) formé par les points p_1 et p_2 ;

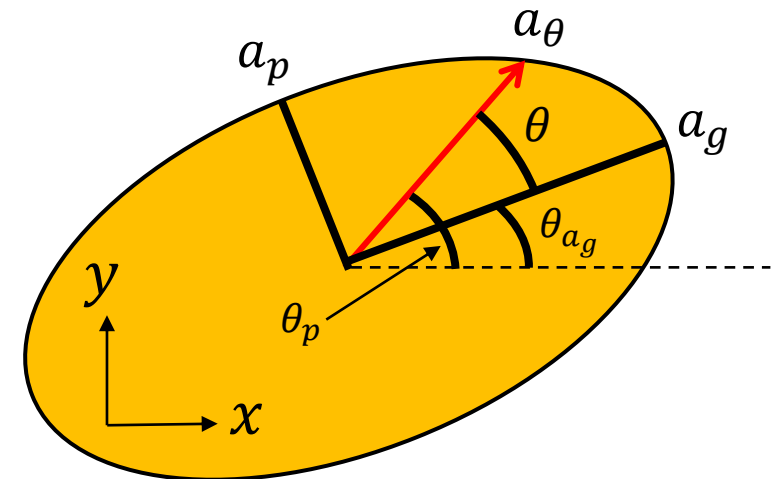


2) Calculer θ soit l'angle formé entre a_g et le vecteur constitué des points p_1 et p_2 ;

3) Calculer a_θ soit la portée effective dans la direction du vecteur formé par p_1 et p_2 ;

$$a_\theta = \frac{a_g a_p}{\sqrt{a_p^2 \cos^2 \theta + a_g^2 \sin^2 \theta}}$$

4) Calculer le variogramme $\gamma(h, a_\theta)$ avec la portée a_θ et la distance h .



3.2 Anisotropies

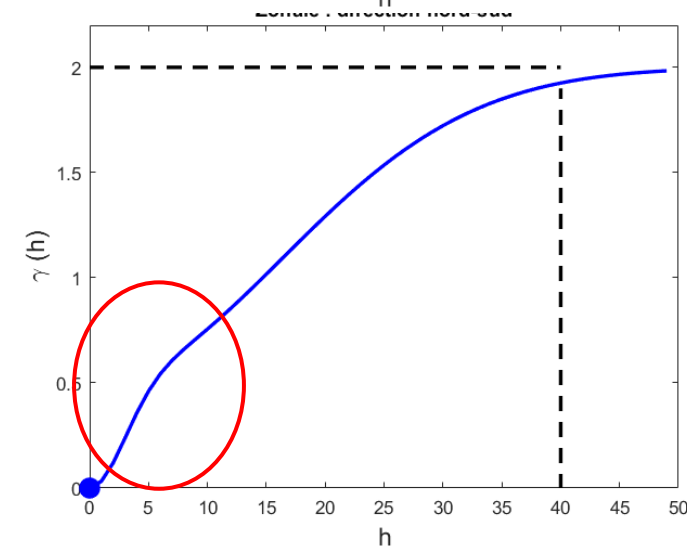
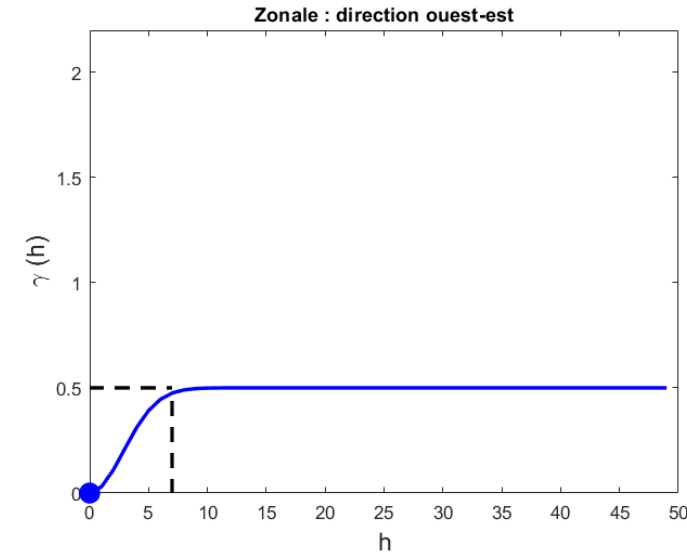
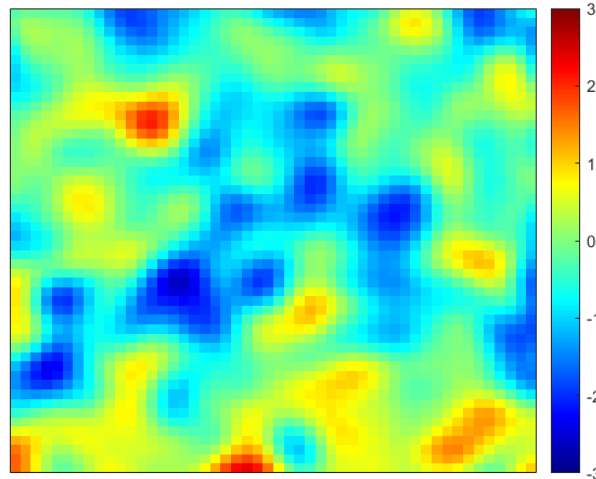
Anisotropies zonales : toute anisotropie qui n'est pas géométrique

→ somme de composantes isotropes et avec anisotropies géométriques

Modèle gaussien isotrope de portée effective $a = 7$

+

Modèle gaussien anisotrope de portée effective $a_p = 40$ et $a_g = \infty$



4. Problèmes et stratégies de modélisation

Le variogramme est une statistique d'ordre 2

Ce n'est pas suffisant pour caractériser tous les aspects d'une image ou d'un processus

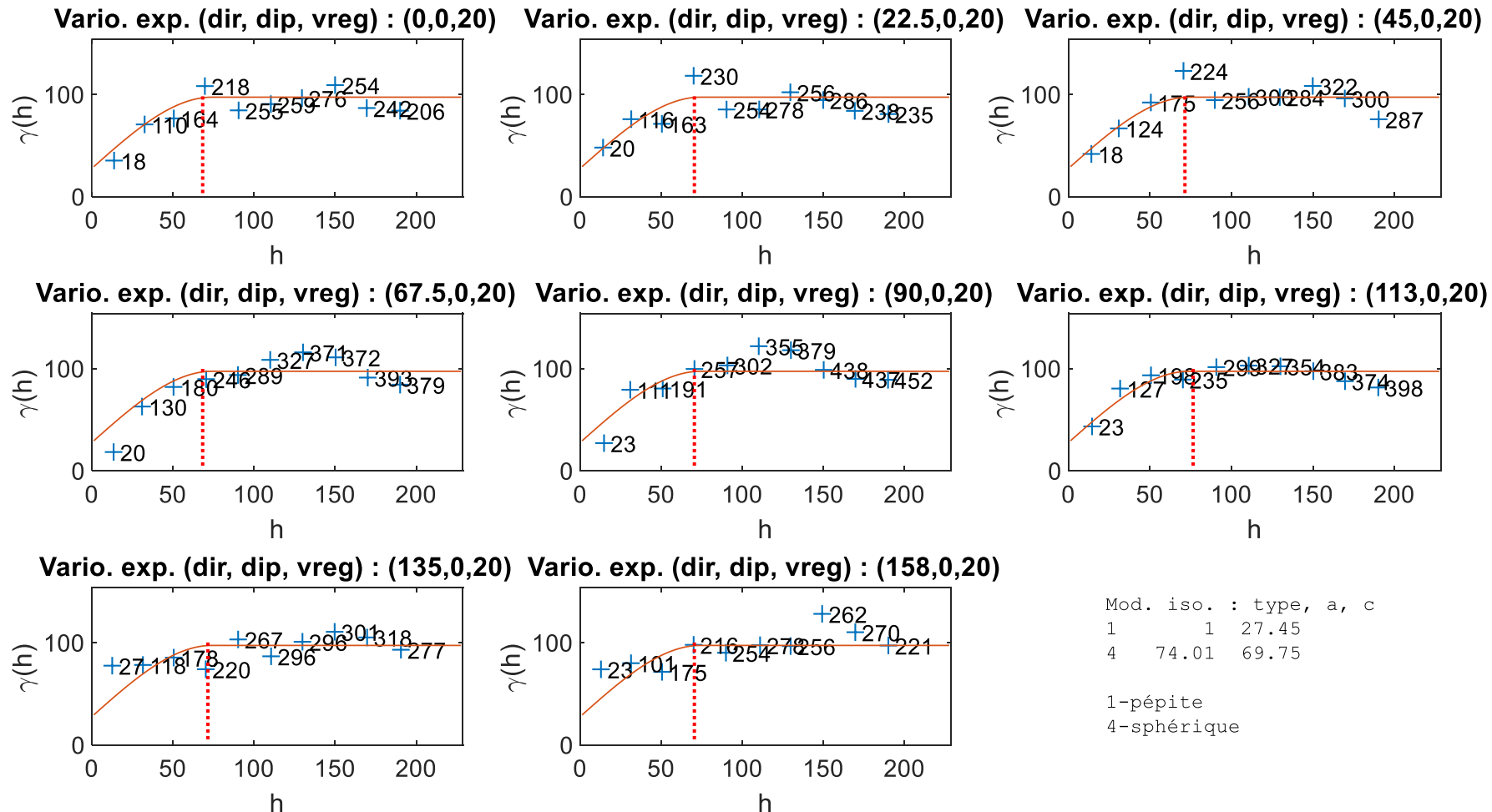
p. ex. on peut créer plusieurs images ayant même \underline{m} , même variogramme et présentant pourtant des textures très différentes

Aparté:

La géostatistique s'intéresse de plus en plus à des caractérisations plus complètes des champs étudiés (p. ex. les moments définis sur plus de deux points, statistiques d'ordre trois comme l'asymétrie directionnel).

4.1 Ajustement d'un variogramme expérimental

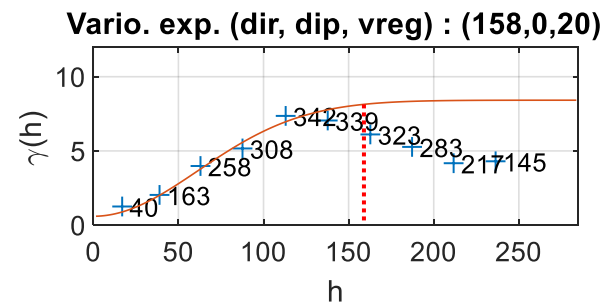
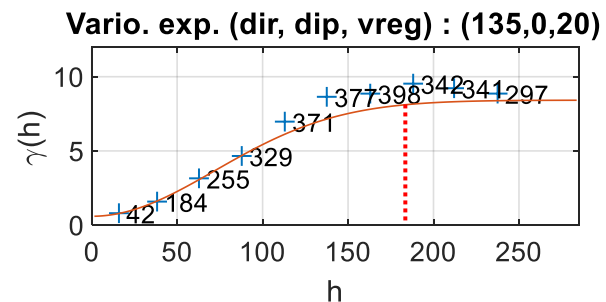
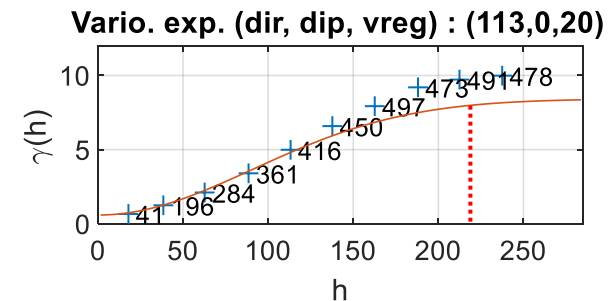
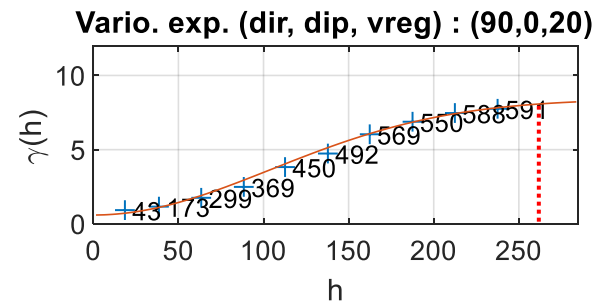
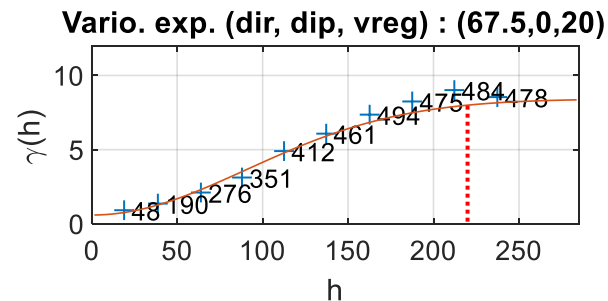
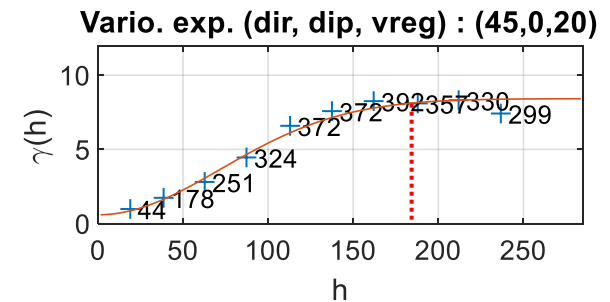
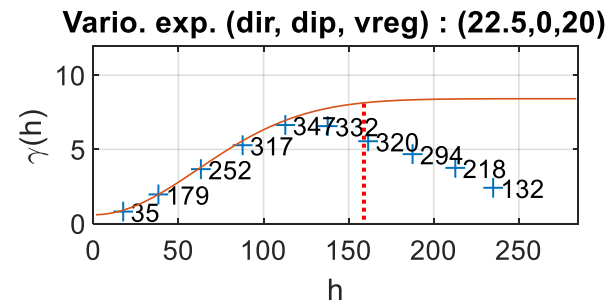
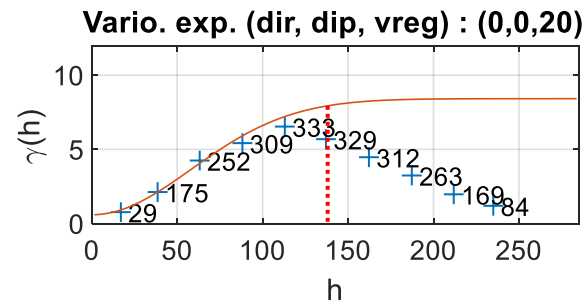
Exemple 1: Directionnel isotrope



Modèle effet de pépité ($C_0 = 27.45 \%m^2$) et modèle sphérique isotrope ($C = 69.75 \%m^2, a = 74.01$)

4.1 Ajustement d'un variogramme expérimental

Exemple 2: Directionnel anisotrope



```
Mod. aniso. 2D : type, ax ay, rot (trigo), c
1 1 1 1 0.5959
3 147.9 82.62 0 7.821
```

1-pépite
3-gaussien

Modèle effet de pépite ($C_0=0.6 \text{ m}^2$) et modèle gaussien avec anisotropie géométrique ($C = 7.8 \text{ m}^2, a_g = a_{90} = 256.2 \text{ m}, a_p = a_0 = 143.1 \text{ m}$)

4.1 Ajustement d'un variogramme expérimental

Stratégies de modélisation

1. Définition minutieuse du domaine (taille de blocs, frontières, etc.)
2. Examen des données, données extrêmes ?
 - Au besoin sous-échantillonnage des données pour éviter de surreprésenter des zones particulières (*Declustering and debiasing*, Pyrcz and Deutsch (2007))
3. Déterminer les directions géologiques principales
 - Variogramme omnidirectionnel → modèle isotrope candidat
 - Calculer les variogrammes directionnels (au moins 4 directions) attention aux paramètres de calculs (classes de distance et tolérance)
 - Comparer les variogrammes directionnels au modèle isotrope candidat
 - acceptable → terminé
 - Inacceptable → ajuster un modèle anisotrope (géométrique)
 - anisotrope (géométrique) acceptable → terminé
 - Inacceptable → anisotropie zonale ?

Dans tous les cas, il importe surtout d'ajuster les premiers points du variogramme. Éviter de «surajuster» les données à grande distance.

4.1 Ajustement d'un variogramme expérimental

Stratégies de modélisation : remarques concernant le calcul des variogrammes

- Au **minimum 30 paires de points** pour la plupart des classes du variogramme
- Au moins **4 à 6 points avant le palier** afin d'ajuster le modèle
- Limiter les distances **d'ajustement à de petit h** (i.e $h < \frac{h_{max}}{2}$)
 - Éviter les effets de bord
 - Assurer la présence de données impliquant le centre du phénomène
 - Le comportement géologique aux bordures peut différer de celui du centre
- Avoir **un minimum de données observées** sur le terrain
 - >30 pour un variogramme omnidirectionnel
 - >60 pour les variogrammes directionnels

4.1 Ajustement d'un variogramme expérimental


Il existe plusieurs stratégies différentes et chacun développe probablement ses propres habitudes. Pour ma part, je recommande la stratégie suivante:

- Calculer et modéliser le variogramme omnidirectionnel en respectant, pour le choix des classes, les règles énoncées ci-haut.
- Si le nombre de données le permet (i.e. au moins 50, préférablement 100) calculer les variogrammes directionnels selon différentes directions (ex. 0, 45, 90, et 135 degrés) en respectant les règles ci-haut (i.e. nb. paires >30, demi-longueur du champ comme distance maximale) et surimposer le modèle omnidirectionnel sur les variogrammes expérimentaux directionnels. La géologie peut apporter une information précieuse dans le choix des directions et la décision d'adopter ou non un modèle anisotrope.
- Déterminer s'il y a anisotropie: différences de palier, mais surtout de portées qui ne peuvent raisonnablement être imputées à des fluctuations aléatoires du variogramme. Cette évaluation concerne surtout le début du variogramme. Des divergences à distance plus grandes sont normales et sans importance. Si nécessaire, procéder à l'ajustement d'un modèle anisotrope.

4.1 Ajustement d'un variogramme expérimental

Ajustement visuel vs méthodes automatiques

1. Visuelle
2. Moindres carrés
3. Maximum de vraisemblances
4. Validation croisée
5. Machine Learning

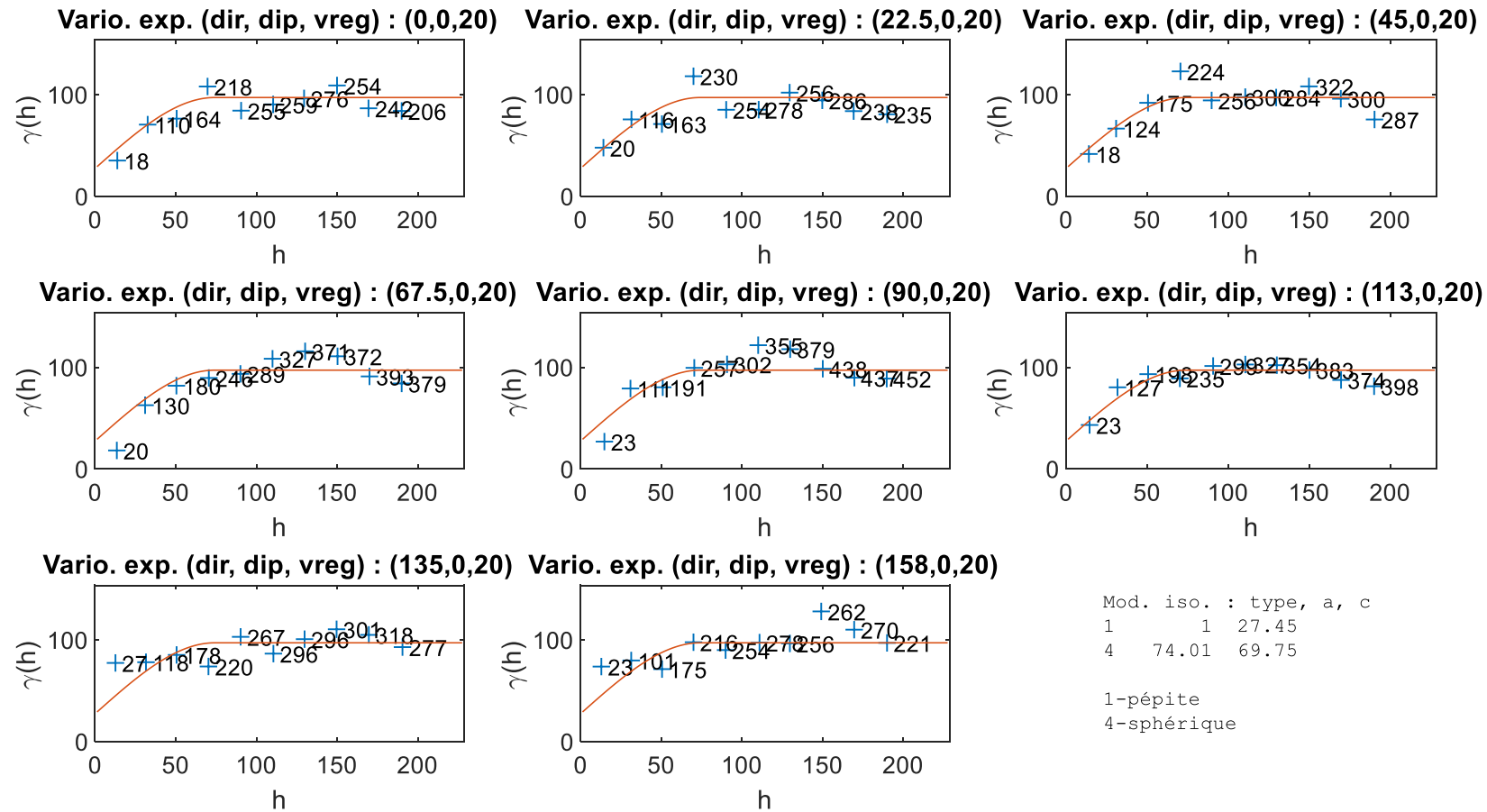


Opinion : Les méthodes entièrement automatiques sont probablement à éviter. Un *monitoring* des caractéristiques essentielles du modèle est souhaitable (type de modèle, anisotropie, direction d'anisotropie). Les connaissances de l'utilisateur concernant le phénomène doivent être mises à contribution. **La géostatistique n'est pas une boîte noire!**

4.1 Ajustement d'un variogramme expérimental

Ajustement visuel vs méthodes automatiques

1. Visuelle



4.1 Ajustement d'un variogramme expérimental

Ajustement visuel vs méthodes automatiques

2. Moindre carrés

Reproduisent l'approche manuelle courante de modélisation de semivariogrammes jusqu'au calcul de plusieurs semivariogrammes expérimentaux, puis automatisent la sélection des paramètres du semivariogramme pour minimiser l'écart avec les points expérimentaux.

$$f = \sum_{i=1}^K N(h_i) \left(\frac{\gamma_e(h_i) - \gamma_t(h_i)}{\gamma_t(h_i)} \right)^2$$

4.1 Ajustement d'un variogramme expérimental

Ajustement visuel vs méthodes automatiques

3. Maximum de vraisemblances

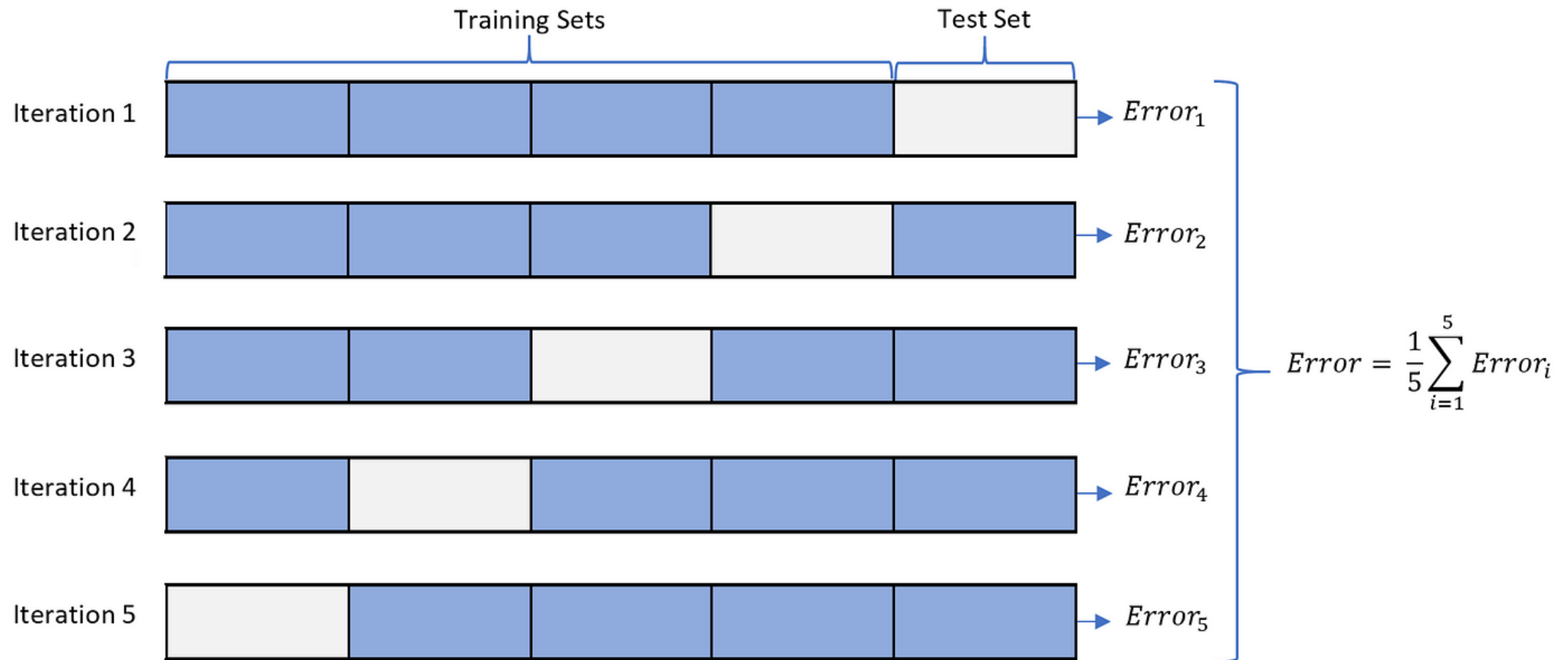
Définition : Une technique **statistique** utilisée pour **estimer les paramètres d'un modèle** statistique. Elle repose sur l'idée de trouver les valeurs des paramètres qui **maximisent la probabilité** (ou la vraisemblance) **d'observer les données** telles qu'elles ont été recueillies.

$$L(m, \theta) = \frac{n}{2} \log(2\pi) + \frac{1}{2} \log |\Sigma(\theta)| + \frac{1}{2} (Z - m)' \Sigma(\theta)^{-1} (Z - m)$$

4.1 Ajustement d'un variogramme expérimental

Ajustement visuel vs méthodes automatiques

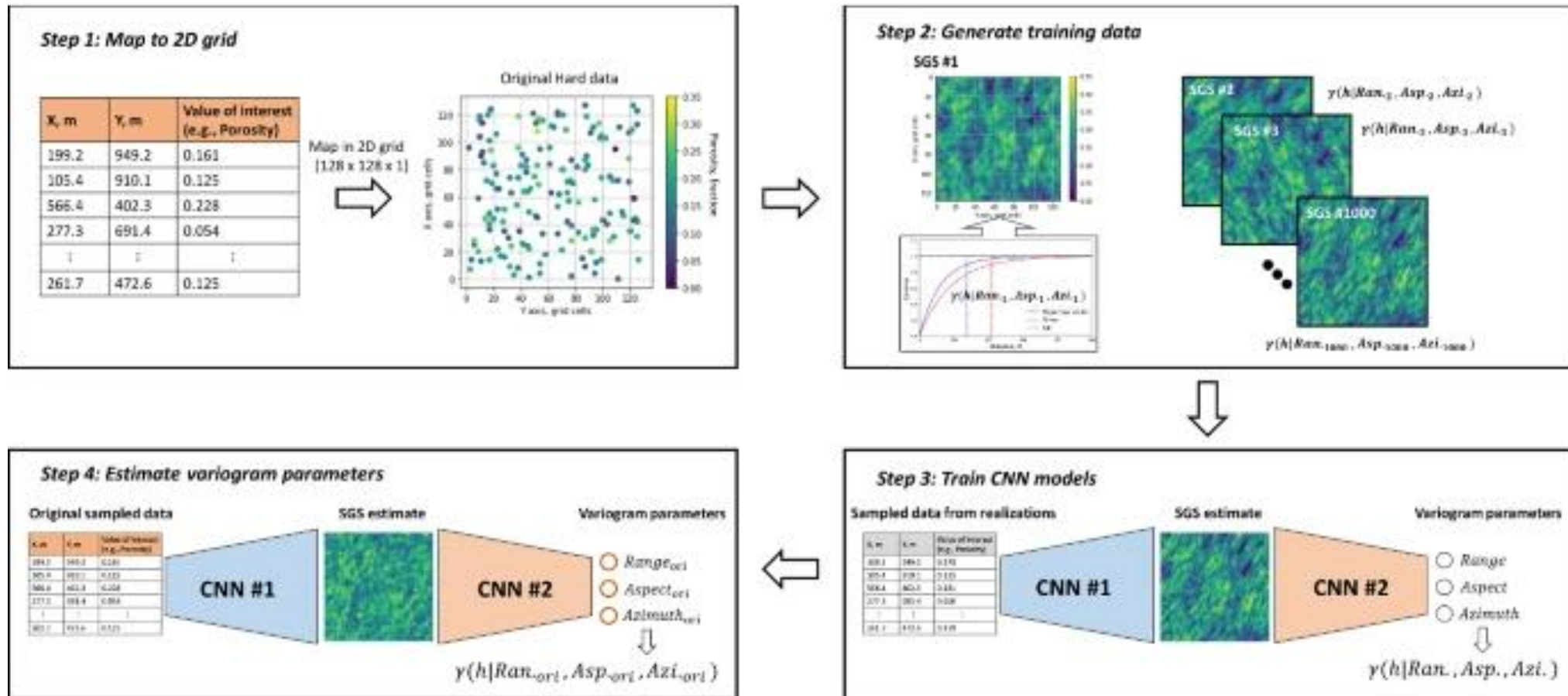
4. Validation Croisée



4.1 Ajustement d'un variogramme expérimental

Ajustement visuel vs méthodes automatiques

5. Machine Learning



5.5 Problèmes et stratégies de modélisation

Variogramme expérimental : 3D

Meilleures directions pour le calcul → **directions des forages**

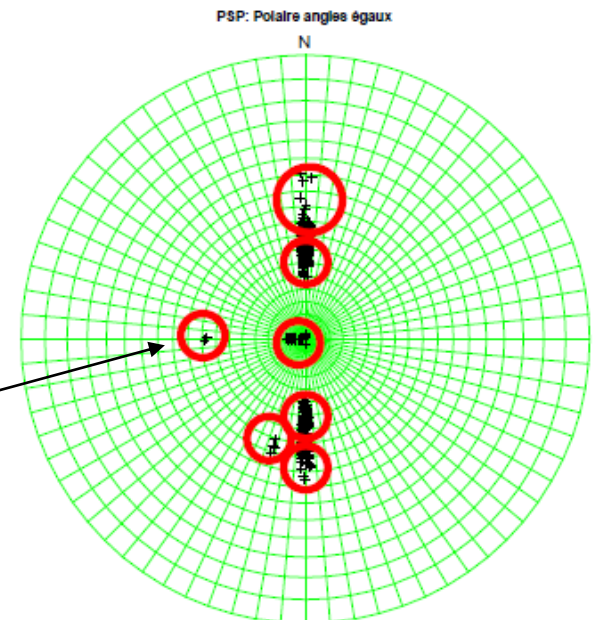
- Permet de bien estimer le variogramme à petite distance
- Erreurs de localisation et de direction ont moins d'impacts sur le variogramme, car les distances intercarottes demeurent inchangées

Problématique:

- Nécessite des forages ayant différentes directions pour pouvoir modéliser l'anisotropie

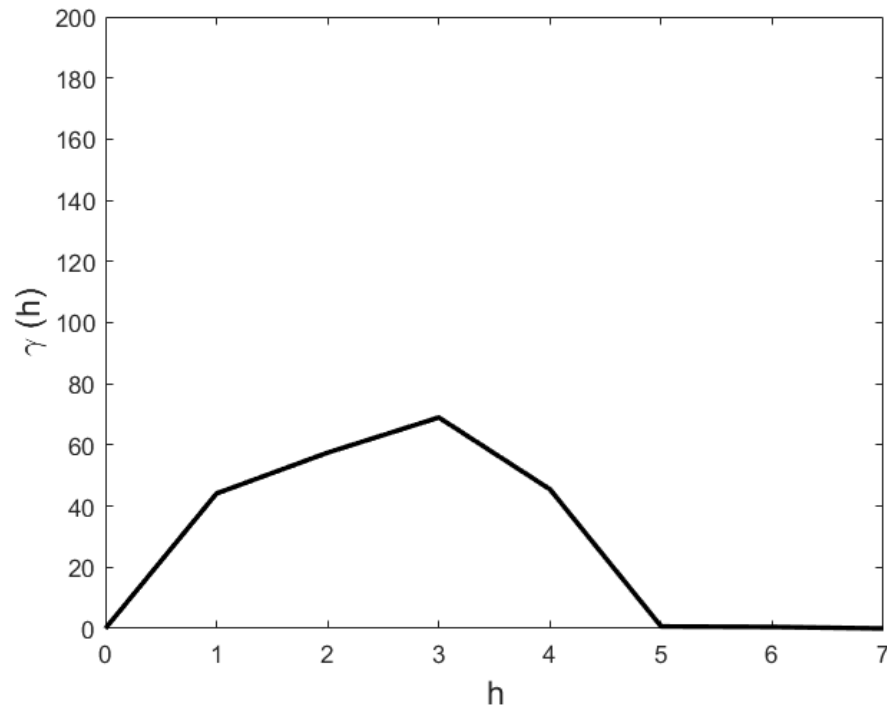
Utile de représenter les mesures de direction des forages sur un stéréonet

○ Direction de calcul et tolérance angulaire

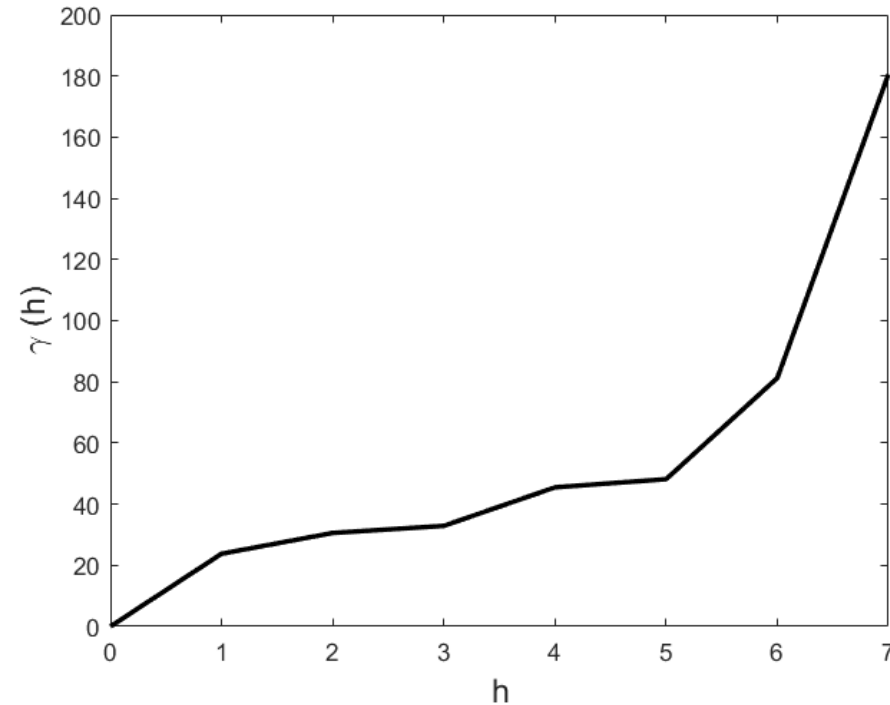


Problèmes : présence de données extrêmes

1	2	1	2	20	3	2	1
---	---	---	---	----	---	---	---



20	2	1	2	1	3	2	1
----	---	---	---	---	---	---	---



Solutions possibles :

- Transformer les données pour diminuer l'influence des valeurs extrêmes
- Éliminer la donnée s'il s'agit d'une erreur d'analyse
- Écrêtage (*capping*)

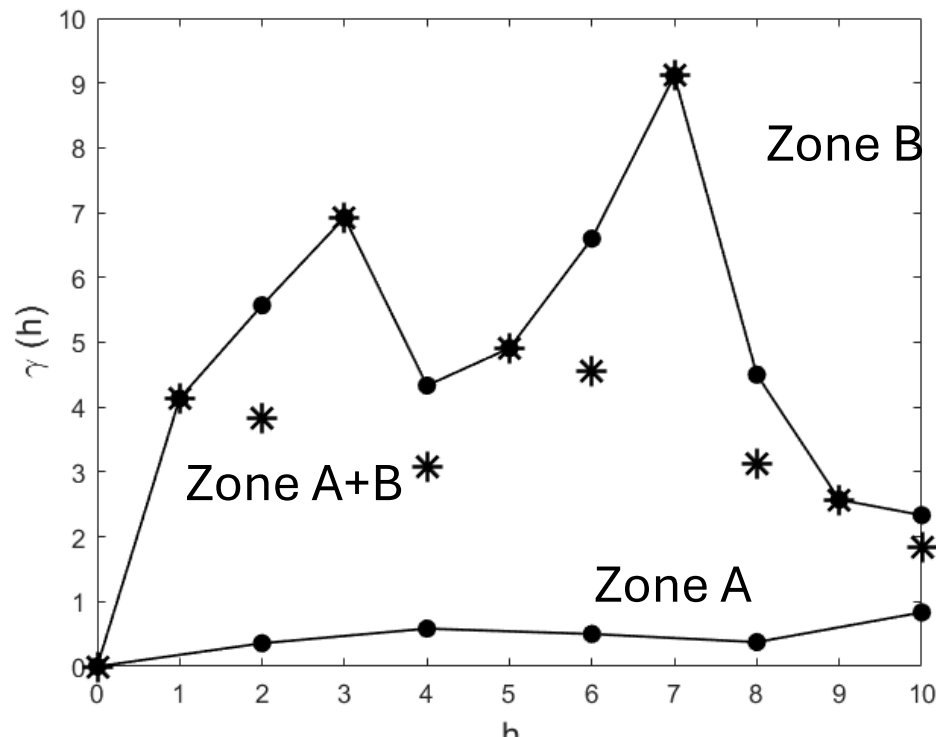
Problèmes : différence de variabilité entre deux zones

Zone A (coordonnées distantes de 2 m)

4	4	5	4	4	5	6	5
---	---	---	---	---	---	---	---

Zone B (coordonnées distantes de 1 m)

8	6	8	10	12	8	10	12	14	10	8	6	12	8	10	10
---	---	---	----	----	---	----	----	----	----	---	---	----	---	----	----



Solutions possibles:

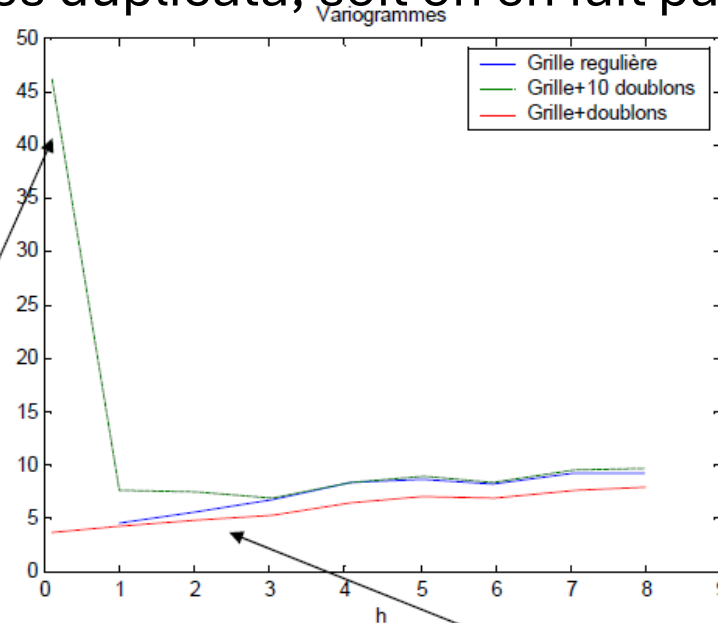
- Séparer en 2 zones si possible, sinon
- Uniformiser l'échantillonnage, i.e. 1 point sur 2 dans la zone B.

Problèmes : rééchantillonner seulement les teneurs riches (doublons)

Très mauvaise idée d'utiliser les données originales et quelques duplicata

→ explosion de l'effet de pépète.

Si on fait des duplicata, soit on en fait partout, soit on les choisit au hasard.



A pour effet de fournir beaucoup de paires à petite distance et montrant de très fortes variations

Si on rééchantillonne chaque point, pas de problème

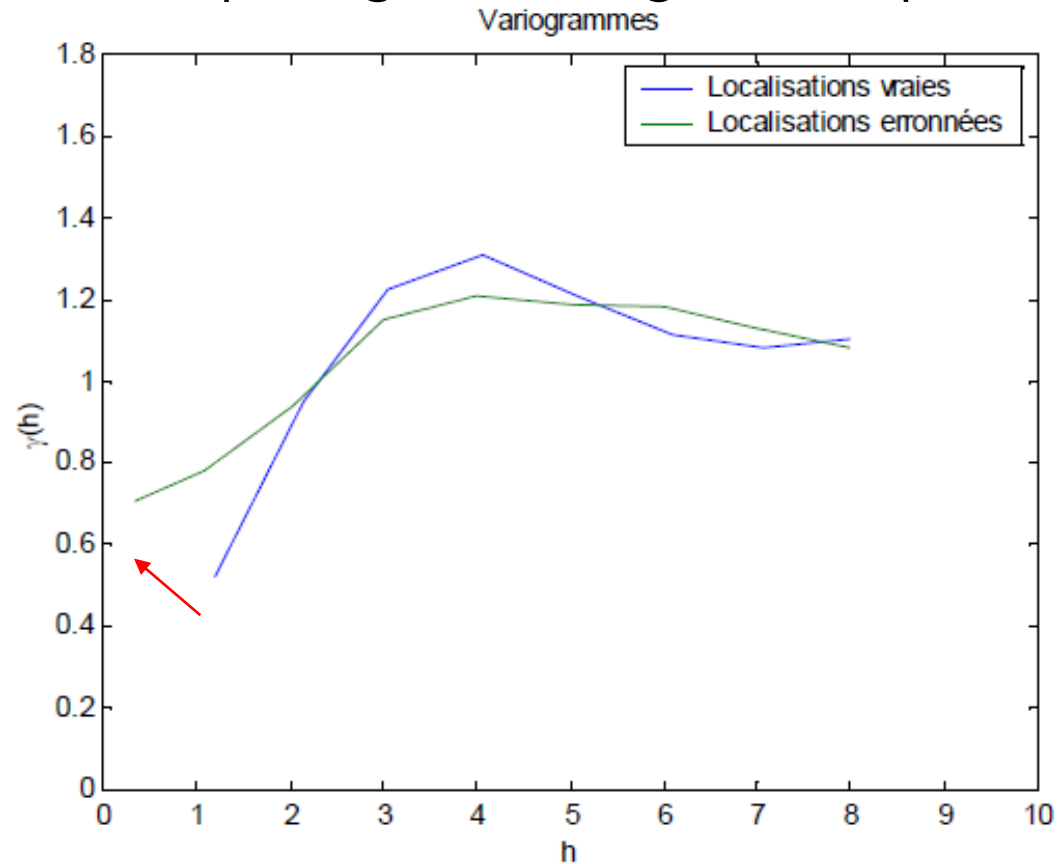
Solutions possibles:

- Éviter les stratégies d'échantillonnage biaisées vers les valeurs fortes
- Décimer l'échantillon pour assurer une couverture uniforme partout

Problèmes : erreurs de localisation

Peu induire un effet de pépite non désiré lors de la modélisation

L'arpentage des forages est important pour minimiser cet effet



Solutions possibles:

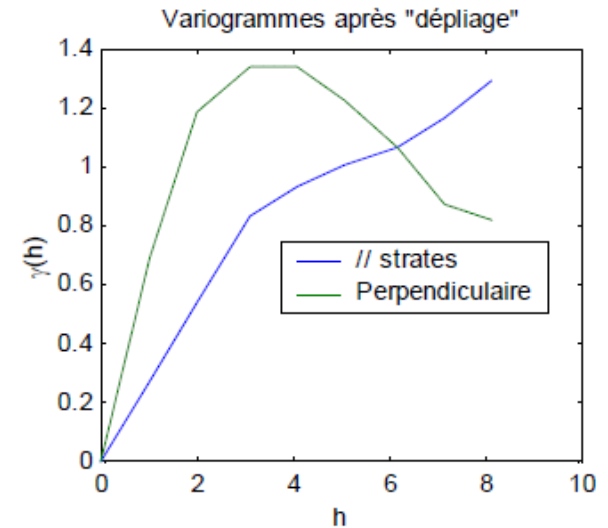
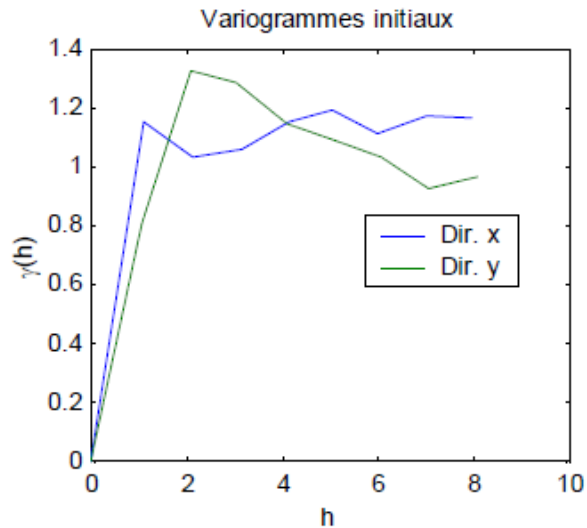
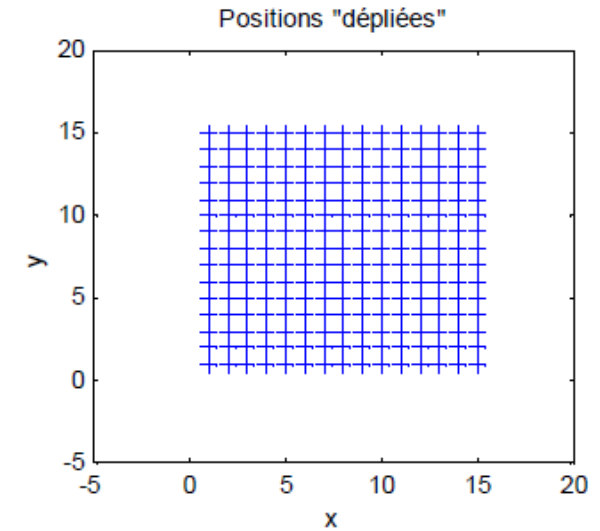
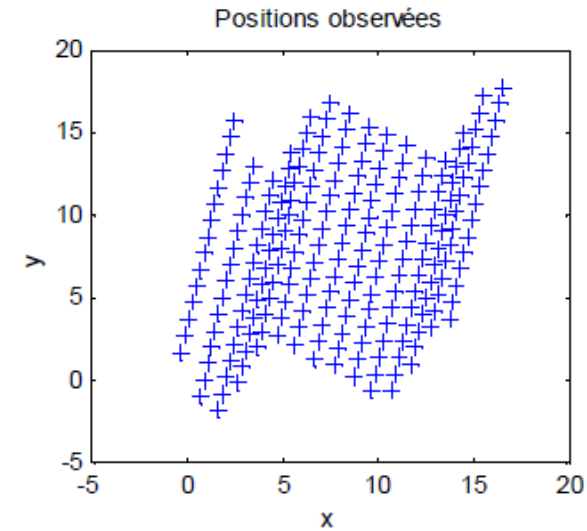
- Localiser avec soin les données
- Tenir compte des erreurs de localisation lors de l'estimation

Problèmes : géologie du gisement

Souvent les teneurs se mettent en place avant les évènements tectoniques.

Les corrélations spatiales initiales sont déformées dans cet espace déformé.

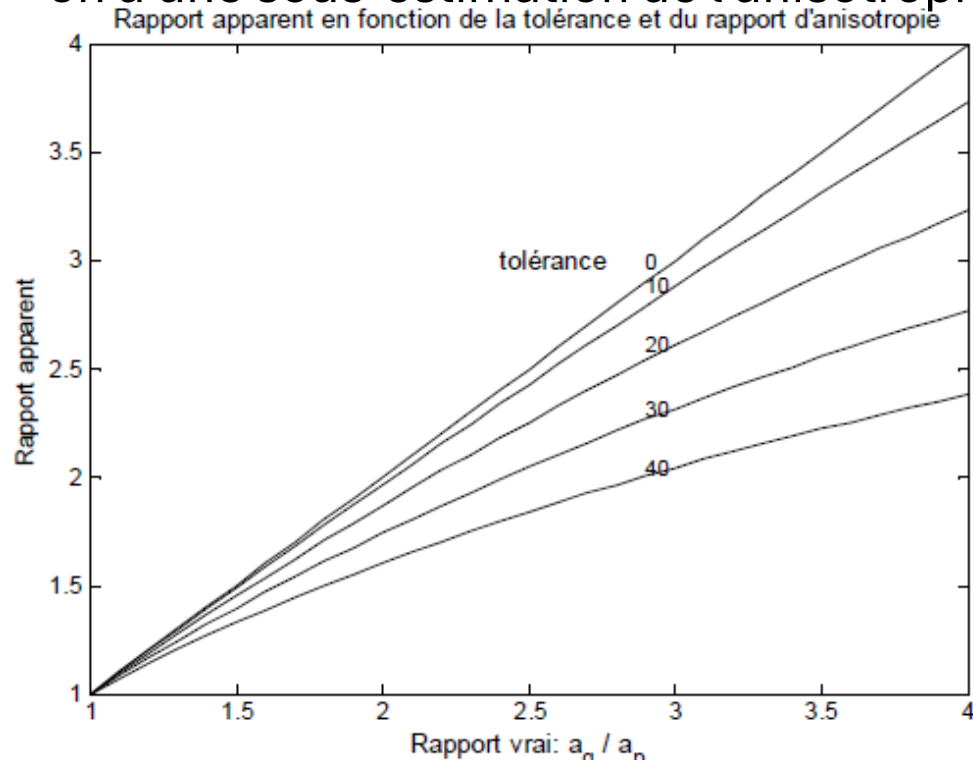
Dans certains cas **simples**, on peut déplier le tout pour retrouver les corrélations originales.



Problèmes : détection des anisotropies

- Tolérance maintenue faible : $< 22.5^\circ$, idéalement $0-10^\circ$
- Plus la tolérance augmente et plus :
 - ont augmentent le nombre de paires de points disponibles
 - les variogrammes deviennent moins directionnels
 - on a une sous-estimation de l'anisotropie

} Choix à faire

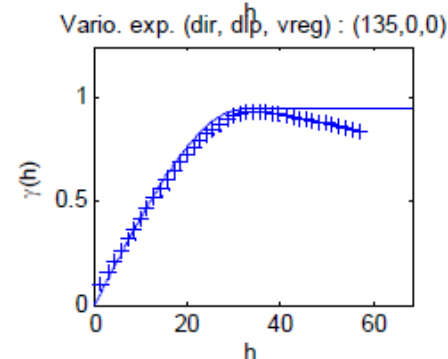
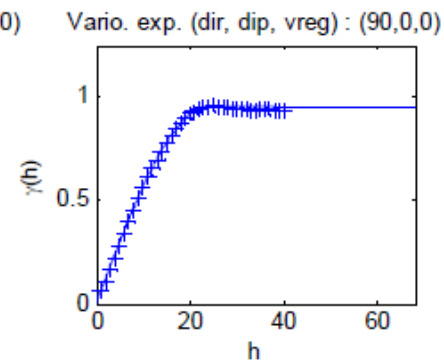
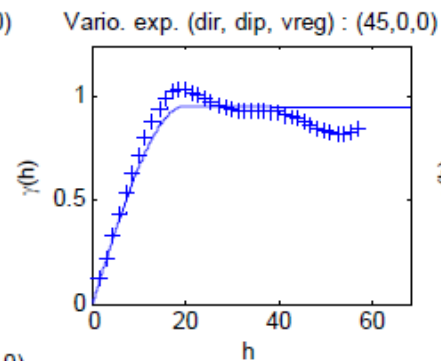
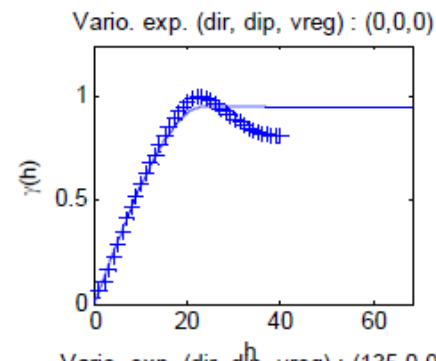
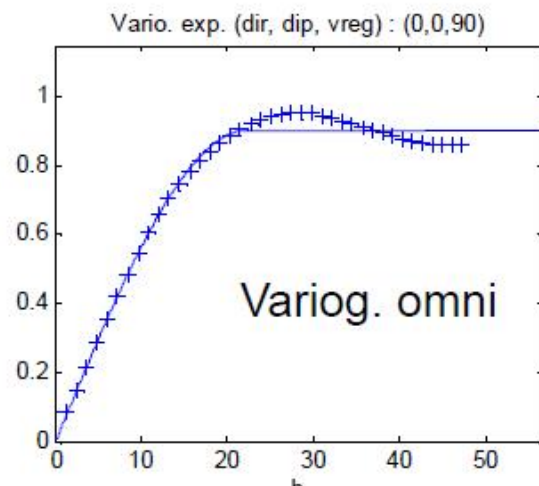
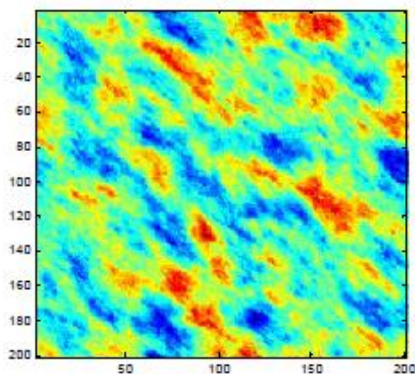


Solutions possibles:

- Discuter avec les spécialistes
Est-ce probable que le gisement présente une anisotropie ?
- Varier la tolérance et observer l'influence sur l'anisotropie

Problèmes : isotrope versus anisotrope

Piège : ce n'est pas parce qu'il y a une anisotropie géométrique que le variogramme expérimental omnidirectionnel n'est pas joli

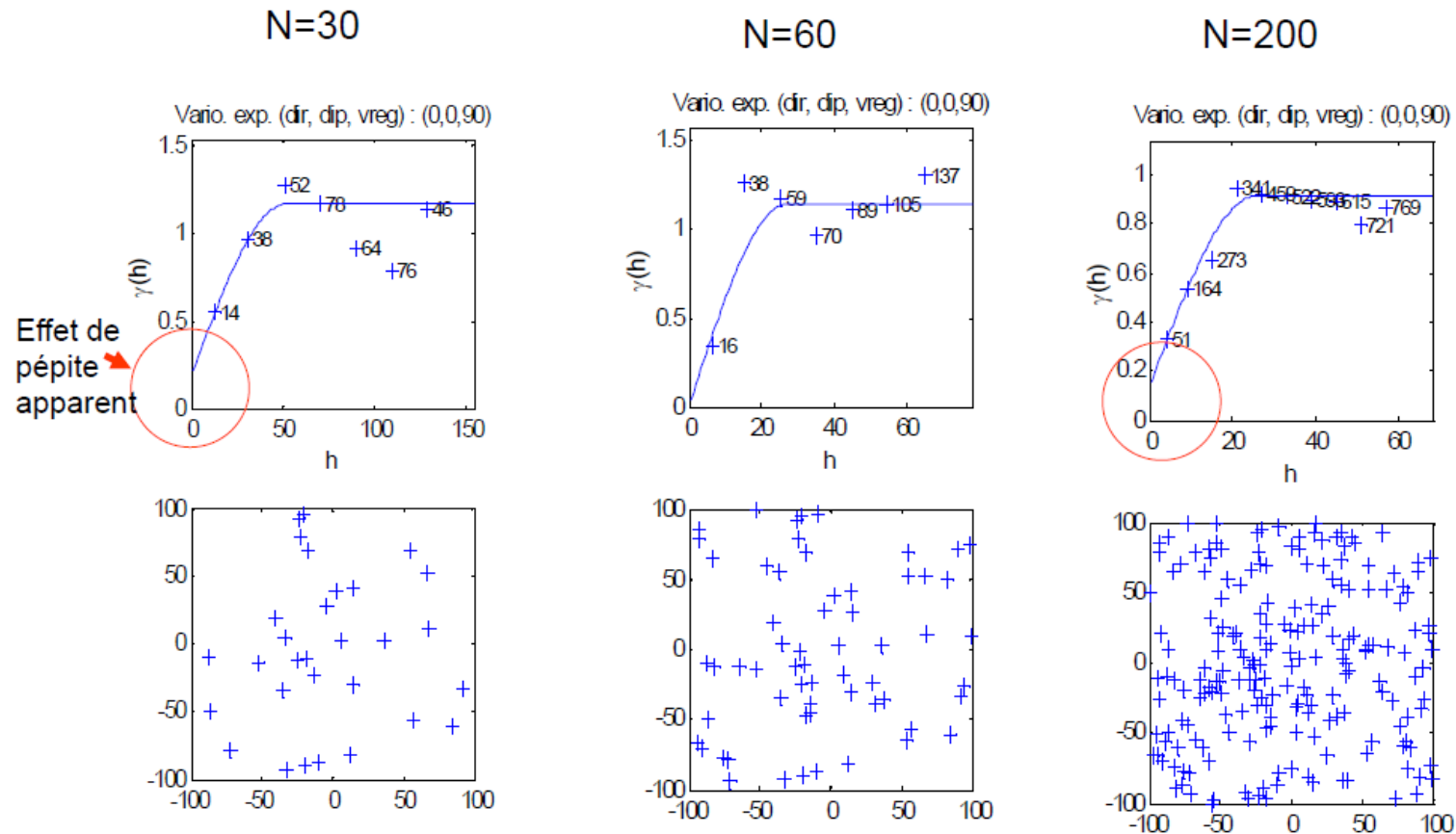


Variog. Directionnels

Modèle sphérique $a_{45}=20$, $a_{135}=35$

Problèmes : effet de pélite apparent

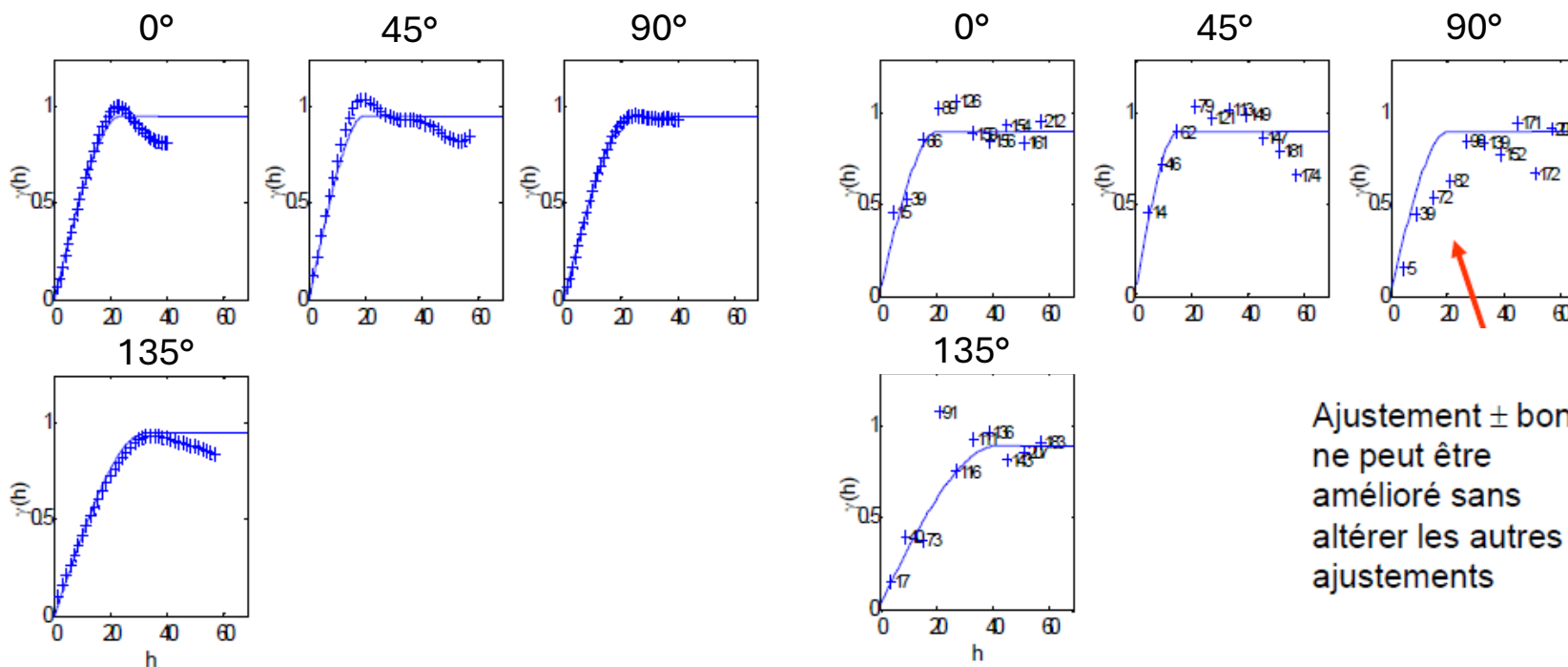
Avec moins de données, parfois on voit un effet de pélite apparent alors qu'en vrai il n'y en a pas.



Il faut savoir accepter une direction mal ajustée afin d'éviter des comportements encore bien pires ailleurs

N=40000

N=200



Ajustement \pm bon;
ne peut être
amélioré sans
altérer les autres
ajustements

Autres outils utiles

- Variogramme des **log(teneurs)** pour identifier les anisotropies possibles, la (les) portée, l'importance approximative de l'effet de pépite;
- Variogramme d'une **transformation des teneurs** (e.g., rang), même chose que les log;
- **Validation croisée** de modèles candidats (krigeage); p. ex. Tester un modèle isotrope vs anisotrope; tester un effet de pépite de 10% vs 30%;...
- Modèle permettant de prédire les **variances des composites de tailles différentes** ?
- Modèle permettant de prédire les **variances des valeurs krigées** ?