Géostatistique multivariable

Automne 2003

Plan

- Besoin et exemples
- Cokrigeage simple et ordinaire
- Covariance croisée et variogramme croisé
- Cas où le cokrigeage peut être utile
- Exemples

Contexte et problématique

Idée : améliorer les estimations obtenues par krigeage en utilisant l'information fournie par d'autres variables (variables secondaires) que la variable principale

Exemples :

a) Estimer la position du sommet d'un réservoir pétrolier ou gazier

Variable principale : position du sommet observé dans quelques forages

Variables secondaires : - positions interprétées par sismique

- attitude du contact dans les forages

=> Obtenir une estimation qui respecte les données de forage, la forme générale de la surface sismique et l'attitude observée dans les forages.

Exemples :

b) Produire une carte de charges hydrauliques qui soit réaliste

Variable principale : surface piézométrique dans quelques puits

Variables secondaires : - surface topographique (aquifère à nappe libre)

- frontières imperméables connues
- vecteurs gradients connus
- mesures géoradar

=> Obtenir une estimation qui respecte les données de puits et les frontières et la forme générale décrite par les mesures géoradar et la topographie.

Exemples :

c) Modélisation de la transmissivité d'un aquifère

Variable principale : transmissivités dans un modèle d'éléments finis ou de différences finies

Variables secondaires : - transmissivité ou conductivités hydrauliques obtenues par « slug-test », test de pompage, courbes granulométriques, capacité spécifique de puits.

- frontières connues

- charges hydrauliques connues (permanent et transitoire)

=> Obtenir une estimation de la transmissivité qui respecte les données de transmissivité à différentes échelles, les données piézométriques et les frontières connues.

Exemples :

d) Inversion de données gravimétriques

Variable principale : densité pour un modèle de blocs du sous-sol

Variables secondaires : - anomalie gravimétrique mesurée au sol

- anomalie gravimétrique aéroportée

- quelques mesures de densité en forage

=> Obtenir une estimation de la densité qui respecte les anomalies et les densités observées

Note : ici, souvent, on n'a aucune observation directe de la variable principale !

Exemples :

e) Inversion de tomographies radar

Variable principale : modèle de « lenteurs » pour un modèle de blocs du sous-sol dans le plan défini par 2 forages coplanaires

Variable secondaire : temps de parcours émetteur-récepteur

=> Obtenir une estimation des lenteurs qui respecte les temps de parcours

Note : souvent, on n'a aucune observation de la variable principale !

Exemples :

f) Cartographie de la température de l'eau en surface

Variable principale : température de l'eau mesurée directement (bateau)

Variable secondaire : température obtenue par interprétation du signal spectral d'un satellite

g) Cartographie de la bathymétrie

Variable principale : mesures directes (bateau)

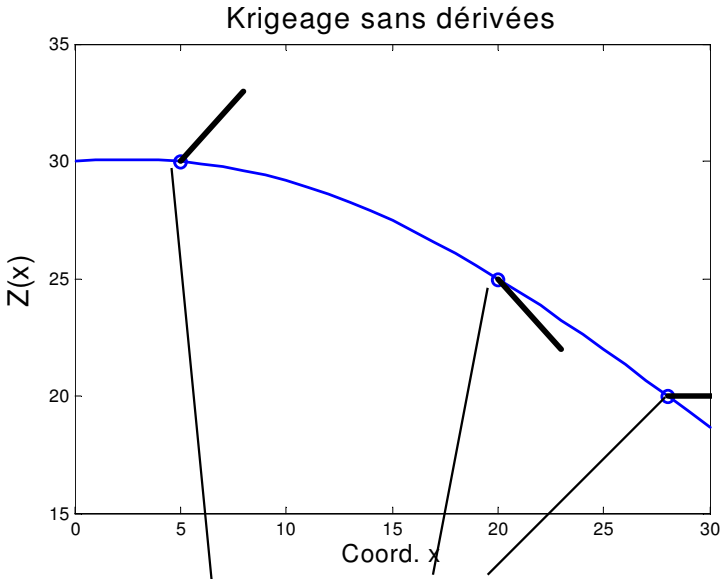
Variable secondaire : interprétation d'un levé hydro-acoustique

h) Cartographie de la contamination

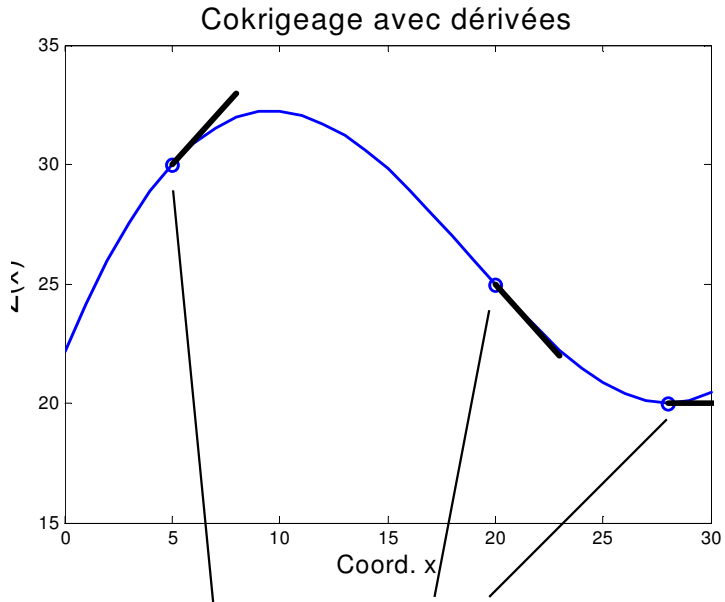
Variable principale : teneurs d'un polluant

Variables secondaires : topographie, surface piézométrique, …

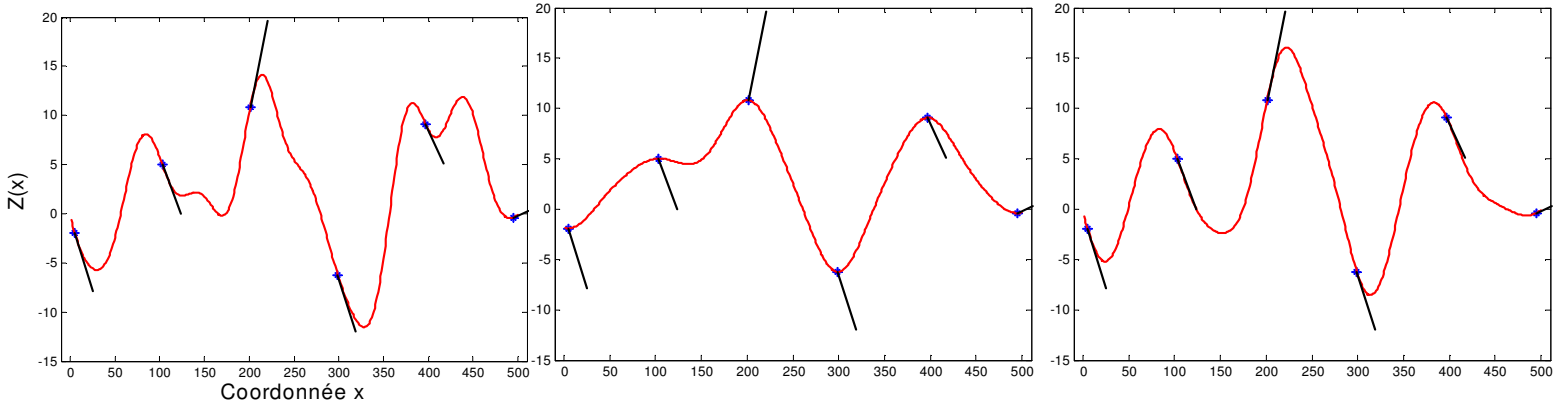
Exemple de cokrigeage



La dérivée de la surface krigée n'est pas égale à la dérivée observée



La dérivée de la surface cokrigée est égale à la dérivée observée



Réalité

Krigeage

Cokrigeage avec dérivées

MAE : krigeage 4.13

cokrigeage : 1.24

Autre exemple de cokrigeage

On mesure la base (B), le sommet (S) et l'épaisseur (E) d'une même formation en quelques forages.

On modélise les variogrammes de ces 3 variables

On effectue le krigeage de ces 3 variables => B*(x), S*(x), E*(x)

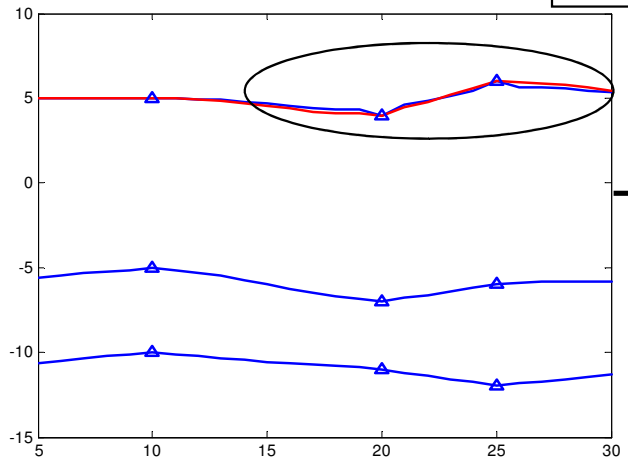
Aura-t-on : E*(x) = S*(x)-B*(x) ?

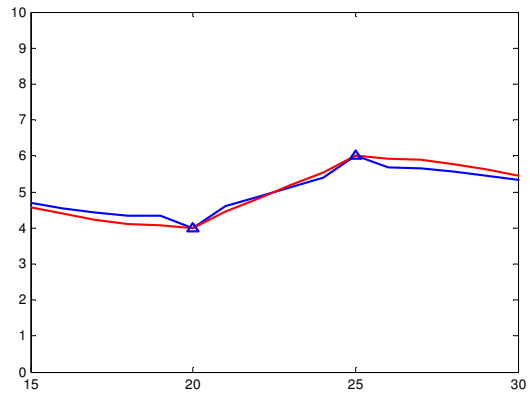
Non pour le krigeage

Oui pour le cokrigeage

(si les covariances reflètent la relation physique E=S-B)

Krigeage

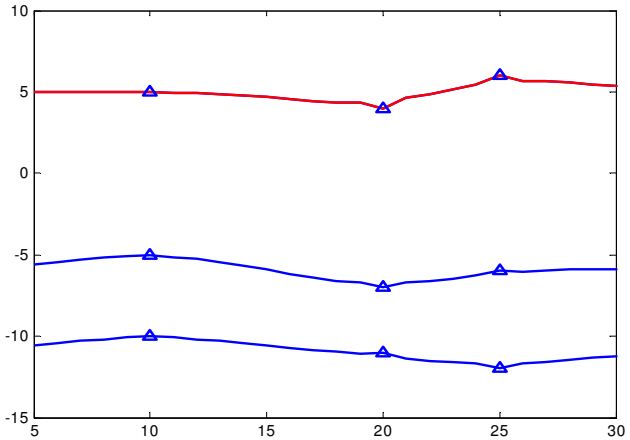




S*-B*

E*

Cokrigeage



Cokrigeage simple

Cas de 2 variables

Z(x) : variable principale : nz observations

Y(x) : variable secondaire : ny observations

Les moyennes de Z(x) et Y(x) sont connues m_Z m_y

$$Z_0^* = m_z + \sum_{i=1}^{nz} \lambda_i (Z_i - m_z) + \sum_{i=1}^{ny} \alpha_i (Y_i - m_y)$$

$$\mathrm{Var}(Z_0 - Z_0^*) = \mathrm{Var}(Z_0) + \sum_{i=1}^{nz}\sum_{j=1}^{nz} \lambda_i \lambda_j \mathrm{Cov}(Z_i, Z_j) + \sum_{i=1}^{ny}\sum_{j=1}^{ny} \alpha_i \alpha_j \mathrm{Cov}(Y_i, Y_j) + 2\sum_{i=1}^{nz}\sum_{j=1}^{ny} \lambda_i \alpha_j \mathrm{Cov}(Z_i, Y_j) - 2\sum_{i=1}^{nz} \lambda_i \mathrm{Cov}(Z_0, Z_i) - 2\sum_{i=1}^{ny} \alpha_i \mathrm{Cov}(Z_0, Y_i)$$

$$\sum_{j=1}^{nz} \lambda_j Cov(Z_i, Z_j) + \sum_{j=1}^{ny} \alpha_j Cov(Z_i, Y_j) = Cov(Z_0, Z_i) \qquad \forall \mathrm{i} = 1 \ldots \mathrm{nz}$$

$$\sum_{j=1}^{nz} \lambda_j Cov(Y_i, Z_j) + \sum_{j=1}^{ny} \alpha_j Cov(Y_i, Y_j) = Cov(Z_0, Y_i) \qquad \forall \mathrm{i} = 1 \ldots \mathrm{ny}$$

$$\sigma_{ck}^2 = Var(Z_0) - \sum_{i=1}^{nz} \lambda_i Cov(Z_0, Z_i) - \sum_{i=1}^{ny} \alpha_i Cov(Z_0, Y_i)$$

Cokrigeage ordinaire

Les moyennes de Z(x) et Y(x) sont inconnues

$$Z_0^* = \sum_{i=1}^{nz} \lambda_i Z_i + \sum_{i=1}^{ny} \alpha_i Y_i$$

$$\sum_{i=1}^{nz} \lambda_i = 1 \text{ et } \sum_{i=1}^{ny} \alpha_i = 0$$

$$\sum_{j=1}^{nz} \lambda_j Cov(Z_i, Z_j) + \sum_{j=1}^{ny} \alpha_j Cov(Z_i, Y_j) + \mu_z = Cov(Z_0, Z_i) \qquad \forall i = 1 \ldots nz$$

$$\sum_{j=1}^{nz} \lambda_j Cov(Y_i, Z_j) + \sum_{j=1}^{ny} \alpha_j Cov(Y_i, Y_j) + \mu_y = Cov(Z_0, Y_i) \qquad \forall i = 1 \ldots ny$$

$$\sum_{i=1}^{nz} \lambda_i = 1$$

$$\sum_{i=1}^{ny} \alpha_i = 0$$

$$\sigma_{ck}^2 = Var(Z_0) - \sum_{i=1}^{nz} \lambda_i Cov(Z_0, Z_i) - \sum_{i=1}^{ny} \alpha_i Cov(Z_0, Y_i) - \mu_z$$

Sous forme matricielle, cokrigeage ordinaire et simple s'écrivent:

$$K\lambda = k \text{ et } \sigma_{ck}^2 = Var(Z_0) - \lambda' k$$

Avec, pour le Cokrigeage simple

$$\begin{bmatrix} K_{zz} & K_{zy} \\ K_{yz} & K_{yy} \end{bmatrix} \begin{bmatrix} \lambda \\ \alpha \end{bmatrix} = \begin{bmatrix} k_{zz} \\ k_{yz} \end{bmatrix}$$

pour le Cokrigeage ordinaire

$$\begin{bmatrix} K_{zz} & K_{zy} & 1 & 0 \\ K_{yz} & K_{yy} & 0 & 1 \\ 1' & 0' & 0 & 0 \\ 0' & 1' & 0 & 0 \end{bmatrix} \begin{bmatrix} \lambda \\ \alpha \\ \mu_z \\ \mu_y \end{bmatrix} = \begin{bmatrix} k_{zz} \\ k_{yz} \\ 1 \\ 0 \end{bmatrix}$$

K_{zz} est nz x nz ; K_{zy} est nz par ny; K_{yz} est ny x nz; K_{yy} est ny par ny; k_{zz} est nz x 1; k_{yz} est ny x 1

Les matrices de cokrigeage sont *toujours* symétriques => $K_{zy} = K'_{yz}$

Toutefois, les *fonctions* de covariances croisées, elles, ne sont pas nécessairement symétriques, en effet :

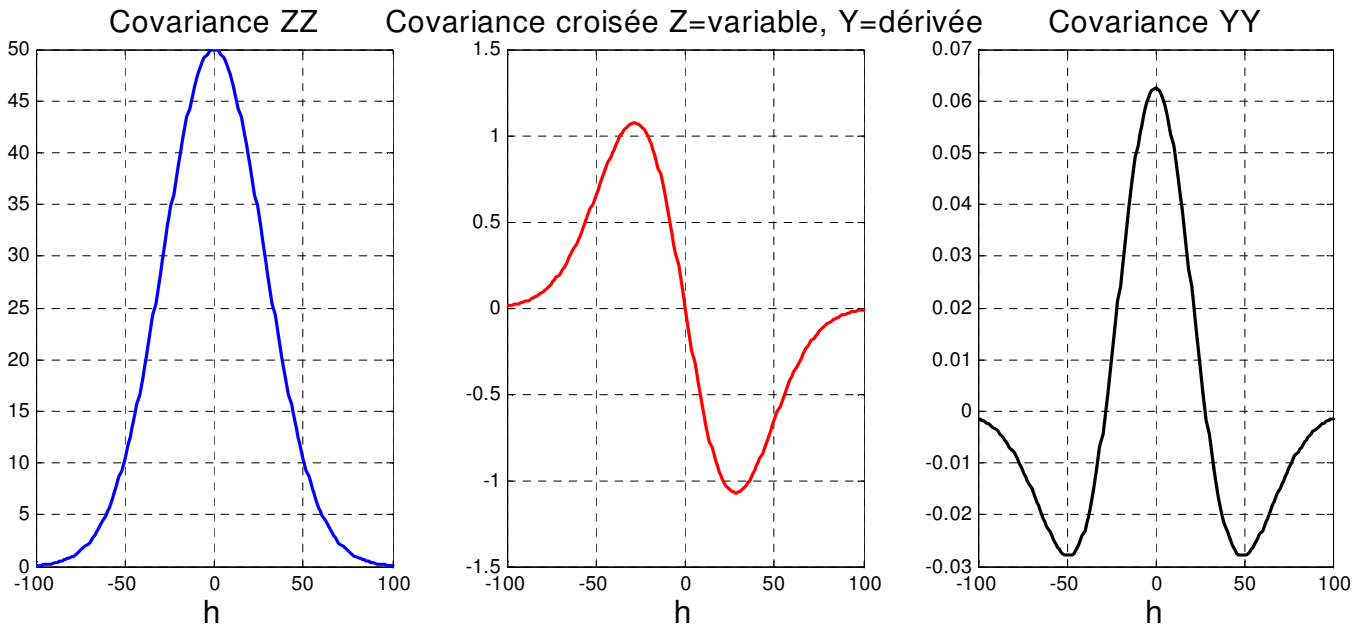
$$C_{zy}(h) = Cov(Z(x),Y(x+h)) = Cov(Y(x+h),Z(x)) = C_{yz}(-h)$$

mais, en général

$$C_{zy}(h) = Cov(Z(x),Y(x+h)) \neq Cov(Y(x),Z(x+h)) = C_{yz}(h) = C_{zy}(-h)$$

Exemple

Cas 1D d'une variable et de sa dérivée



Notes:

a) krigeage => + souvent krigeage ordinaire
cokrigeage => cokrigeage simple est souvent plus indiqué (e.g. cas où la variable principale n'est pas observée)

b) krigeage => variogramme est l'outil de base
cokrigeage => covariance est l'outil de base

Covariance croisée et variogramme croisé

Covariance croisée

$C_{zy}(h) = Cov(Z(x),Y(x+h))$

Variogramme croisé

$$\gamma_{zy}(h) = 0.5 * E\left[\left(Z(x) - Z(x+h)\right)\left(Y(x) - Y(x+h)\right)\right]$$
$$= 0.5 * Cov\left(\left(Z(x) - Z(x+h)\right), \left(Y(x) - Y(x+h)\right)\right)$$

=> Le variogramme croisé est toujours symétrique

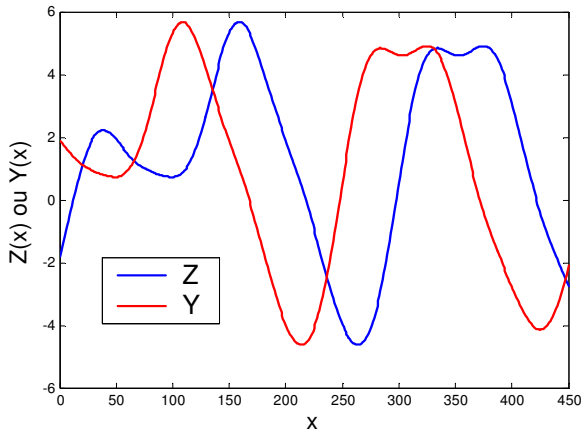
Lien

$$\gamma_{zy}(h) = C_{zy}(0) - 0.5 * \left(C_{zy}(h) + C_{zy}(-h)\right)$$

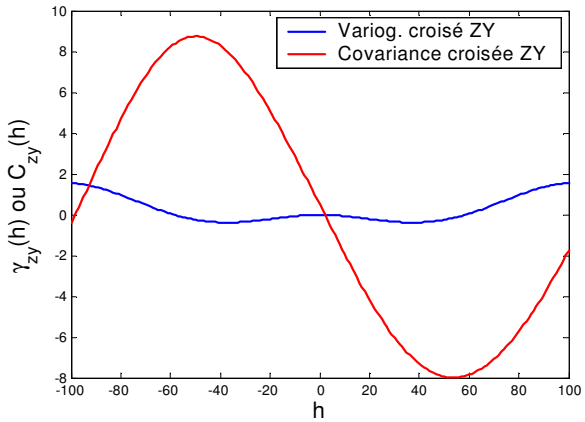
$\Rightarrow$ La covariance croisée ne peut être déduite du variogramme croisé que si elle est symétrique !

$\Rightarrow$ La covariance croisée est + générale que le variogramme croisé

Exemple



Y est rien d'autre qu'une copie décalée (de 50) vers la gauche de Z



Le variogramme croisé => pas de structure!
La covariance croisée => très forte structure
- asymétrique
- reconnaît le décalage

Comparaison

Covariance croisée	Variogramme croisée
m_Z et m_y connus	Pas besoin
Tous les points où l'une ou l'autre variable est connue peuvent être utilisés dans le calcul de la covariance	Seuls les points où les 2 variables sont connues peuvent être utilisés dans le calcul du variogramme croisé
La covariance peut être asymétrique	La covariance doit être symétrique

Modèles admissibles

But : assurer que toute combinaison linéaire de Z et Y présente une variance théorique positive.

Vérification délicate, on se limite à des cas certains

a) Modèle découlant de relations physiques ou mathématiques entre Z et Y. Si l'on connaît C_{zz} alors on peut déduire C_{zy} et C_{yy}.

Exemple 1D : $Y(x) = dZ(x)/dx \Rightarrow$

$$Cov(Z(x),Y(x+h)) = d\,[Cov(Z(x),Z(x+h))]/dh$$

$$Cov(Y(x),Y(x+h)) = -\,d^2[Cov(Z(x),Z(x+h))]/dh^2$$

b) Modèle de corégionalisation linéaire

$$\begin{bmatrix} C_{zz}(h) & C_{zy}(h) \\ C_{yz}(h) & C_{yy}(h) \end{bmatrix} = \begin{bmatrix} B_{1,zz} & B_{1,zy} \\ B_{1,yz} & B_{1,yy} \end{bmatrix} C_1(h) + \begin{bmatrix} B_{2,zz} & B_{2,zy} \\ B_{2,yz} & B_{2,yy} \end{bmatrix} C_2(h) + \ldots$$

$$= B_1 C_1(h) + B_2 C_2(h) + \ldots$$

$\Rightarrow$ **toutes les covariances et covariances croisées peuvent s'exprimer comme une *combinaison linéaire de quelques structures de base* (de palier 1, arbitrairement). Les matrices B sont des matrices de coefficients.**

Exemple: 2 structures de base : 1 effet de pépite + 1 sphérique, a=30

$C_1(h)$ représente l'effet de pépite => $C_1(0) = 1$; $C_1(h>0) = 0$
$C_2(h)$ représente la partie sphérique => $C_2(h) = 1-(1.5\ h/a-0.5\ (h/a)^3)$

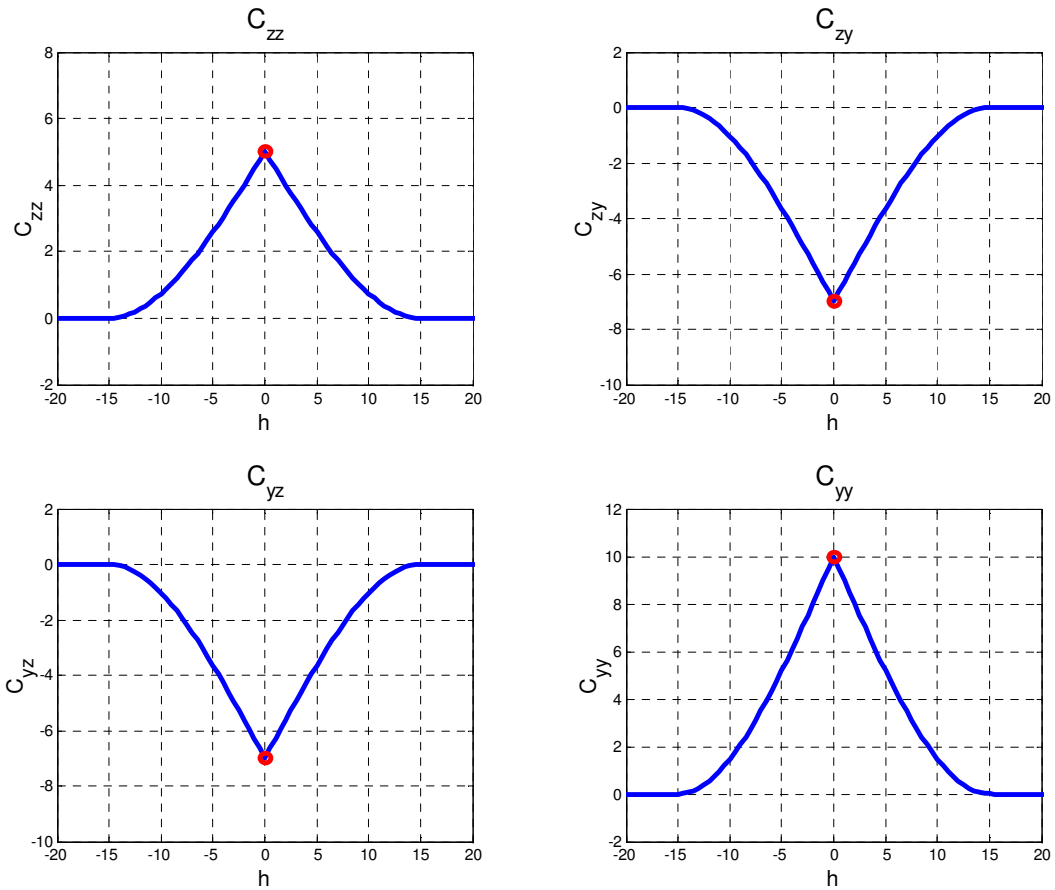
Si chaque matrice *B est positive semi-définie* (i.e. avec 2 variables det(B) ≥0) alors le modèle est admissible.
Si l'une des matrices n'est pas positive semi-définie, alors on ne peut rien dire sur l'admissibilité du modèle.

Note : lorsque toutes les matrices B *sont proportionnelles*, on peut réécrire le modèle linéaire de corégionalisation comme:

$$\begin{bmatrix} C_{zz}(h) & C_{zy}(h) \\ C_{yz}(h) & C_{yy}(h) \end{bmatrix} = \begin{bmatrix} B_{zz} & B_{zy} \\ B_{yz} & B_{yy} \end{bmatrix} C(h) = B\,C(h)$$

Ce modèle indique que toutes les covariances sont *proportionnelles* à un modèle de base unique de palier arbitraire 1 et pouvant comprendre plusieurs composantes élémentaires (e.g. effet pépite+sphérique avec a=30). Ce modèle est admissible *ssi* B est positive semi-définie.

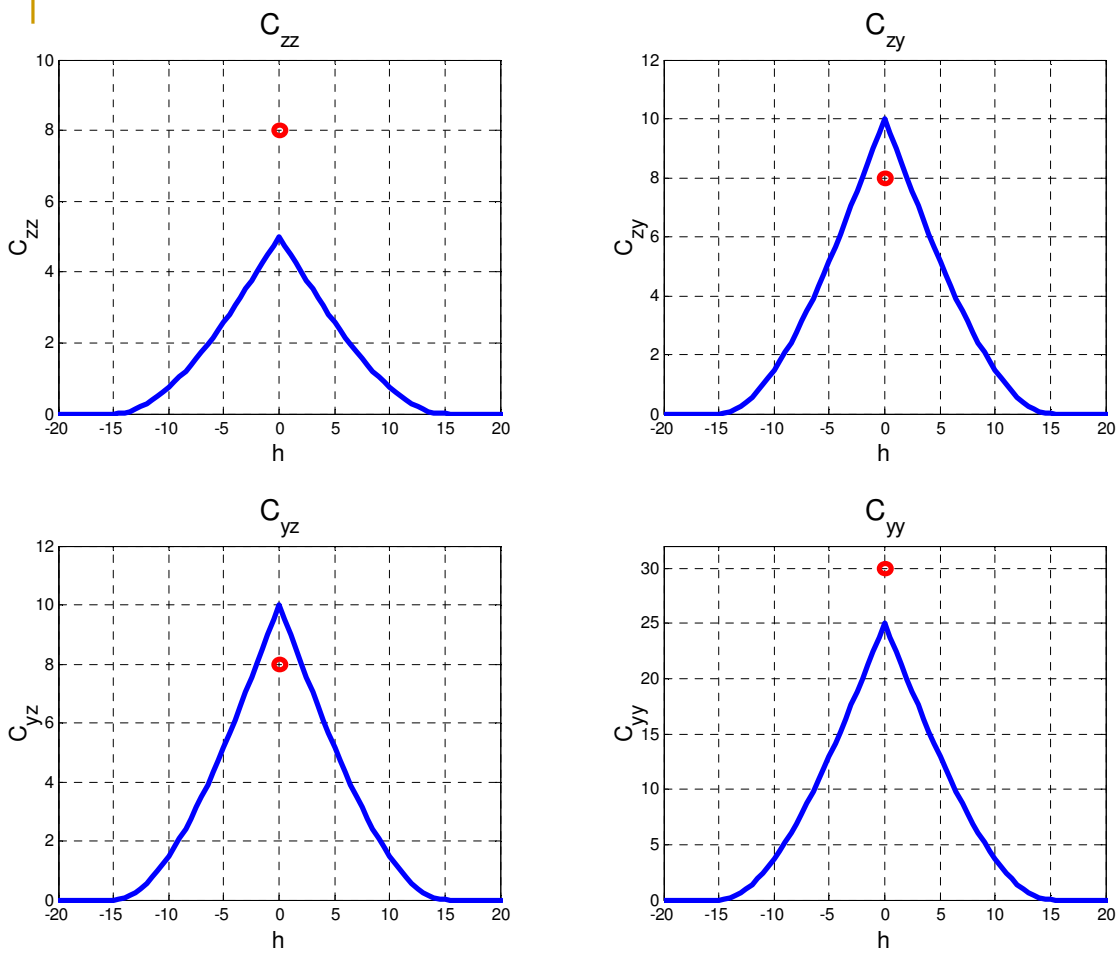
Exemples de modèle linéaire de corégionalisation



Modèle sphérique
a=15

$$B = \begin{bmatrix} 5 & -7 \\ -7 & 10 \end{bmatrix}$$

Det(B)=1 =>
modèle admissible



Effet pépite+

Modèle sphérique
a=15

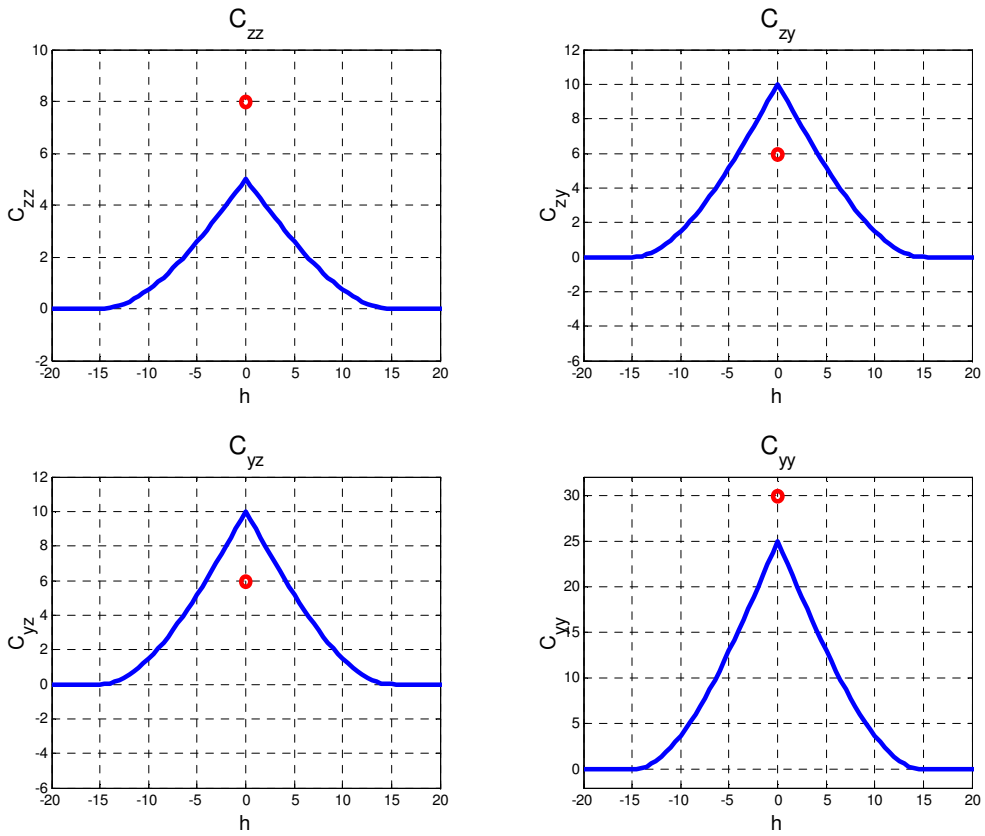
$$B_1 = \begin{bmatrix} 3 & -2 \\ -2 & 5 \end{bmatrix}$$

$$B_2 = \begin{bmatrix} 5 & 10 \\ 10 & 25 \end{bmatrix}$$

Det(B1)=11

Det(B2)=25

=> modèle admissible



Effet pépite+

Modèle sphérique
a=15

$$B_1 = \begin{bmatrix} 3 & -4 \\ -4 & 5 \end{bmatrix}$$

$$B_2 = \begin{bmatrix} 5 & 10 \\ 10 & 25 \end{bmatrix}$$

Det(B1)= -1

Det(B2)=25

=> modèle non admissible

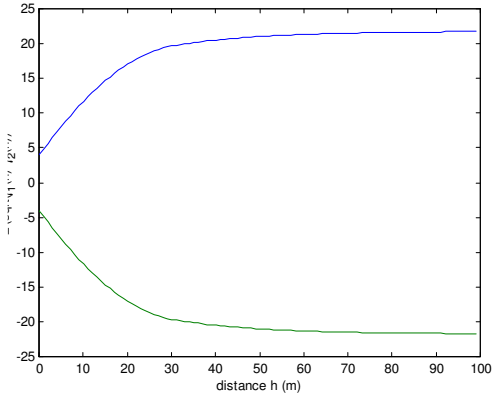
Condition de Cauchy-Schwartz

Variogramme croisé : covariance des incréments (Z(x)-Z(x+h)) et (Y(x)-Y(x+h))

Variogrammes simples: variance des incréments

Donc il faut que : $$\gamma_{ZY}(h) \leq \sqrt{\gamma_{ZZ}(h)\gamma_{YY}(h)}$$

Malheureusement, cette condition est nécessaire mais non suffisante!



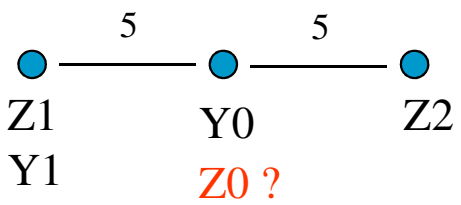
Condition C-S définit une enveloppe où doit se retrouver entièrement le modèle de variogramme croisé.

Exemple numérique détaillé

$$\begin{bmatrix} C_{zz}(h) & C_{zy}(h) \\ C_{yz}(h) & C_{yy}(h) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \delta(h) + \begin{bmatrix} 2 & 2.4 \\ 2.4 & 4 \end{bmatrix} Sphérique(a = 30m, C = 1)$$

$\delta(h)$ =1 si h=0, 0 si |h|>0.

Z_1 et Y_1 en x_1=0 , Z_2 en x_2=10 et Y_0 en x_0=5



Note: à h=0, corr. (Z,Y)= $2.4/(3*5)^{0.5}$ = 0.62

À h=5, corr(Z(x),Z(x+h))= 1.50/3 = 0.50

Par cokrigeage simple

$$
\begin{array}{c}
 \\ Z_1 \\ Z_2 \\ Y_1 \\ Y_0
\end{array}
\begin{array}{c}
\begin{array}{cccc} Z_1 & Z_2 & Y_1 & Y_0 \end{array} \\
\begin{bmatrix} 3 & 1.037 & 2.4 & 1.8056 \\ 1.037 & 3 & 1.2444 & 1.8056 \\ 2.4 & 1.2444 & 5 & 3.0093 \\ 1.8056 & 1.8056 & 3.0093 & 5 \end{bmatrix}
\end{array}
\begin{bmatrix} \lambda_{1,z} \\ \lambda_{2,z} \\ \lambda_{1,y} \\ \lambda_{0,y} \end{bmatrix}
=
\begin{array}{c} Z_0 \\ \begin{bmatrix} 1.5046 \\ 1.5046 \\ 1.8056 \\ 2.4 \end{bmatrix} \end{array}
$$

$$\lambda_{1,z} = 0.2294, \lambda_{2,z} = 0.2336, \lambda_{1,y} = 0.0072, \lambda_{0,y} = 0.3085 \quad \text{et } \sigma^2_{\text{ck}} = 1.5500$$

Poids important donné à la variable secondaire

Variance de cokrigeage < variance de krigeage simple (1.88)

Par cokrigeage ordinaire

$$
\begin{array}{c}
 \\ Z_1 \\ Z_2 \\ Y_1 \\ Y_0 \\ \\ \\
\end{array}
\begin{array}{c}
\begin{array}{cccccc} Z_1 & Z_2 & Y_1 & Y_0 & & \end{array} \\
\begin{bmatrix}
3 & 1.037 & 2.4 & 1.8056 & 1 & 0 \\
1.037 & 3 & 1.2444 & 1.8056 & 1 & 0 \\
2.4 & 1.2444 & 5 & 3.0093 & 0 & 1 \\
1.8056 & 1.8056 & 3.0093 & 5 & 0 & 1 \\
1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 0
\end{bmatrix}
\end{array}
\begin{bmatrix}
\lambda_{1,z} \\ \lambda_{2,z} \\ \lambda_{1,y} \\ \lambda_{0,y} \\ \mu_z \\ \mu_y
\end{bmatrix}
=
\begin{array}{c}
Z_0 \\
\begin{bmatrix}
1.5046 \\ 1.5046 \\ 1.8056 \\ 2.4 \\ 1 \\ 0
\end{bmatrix}
\end{array}
$$

$$\lambda_{1,z} = 0.5494, \lambda_{2,z} = 0.4506, \lambda_{1,y} = -0.1678, \lambda_{0,y} = 0.1678,$$

$$\mu_Z = -0.5111, \mu_y = 0.2603 \text{ et } \sigma^2_{ck} = 1.907$$

Variance de cokrigeage ≈ variance de krigeage ordinaire (2.01)

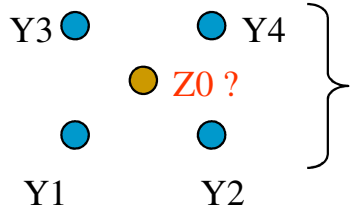
Cokrigeage simple (CoS) vs cokrigeage ordinaire (CoO)

Souvent le CoS est préférable

- CoS => ne nécessite pas d'observations de Z

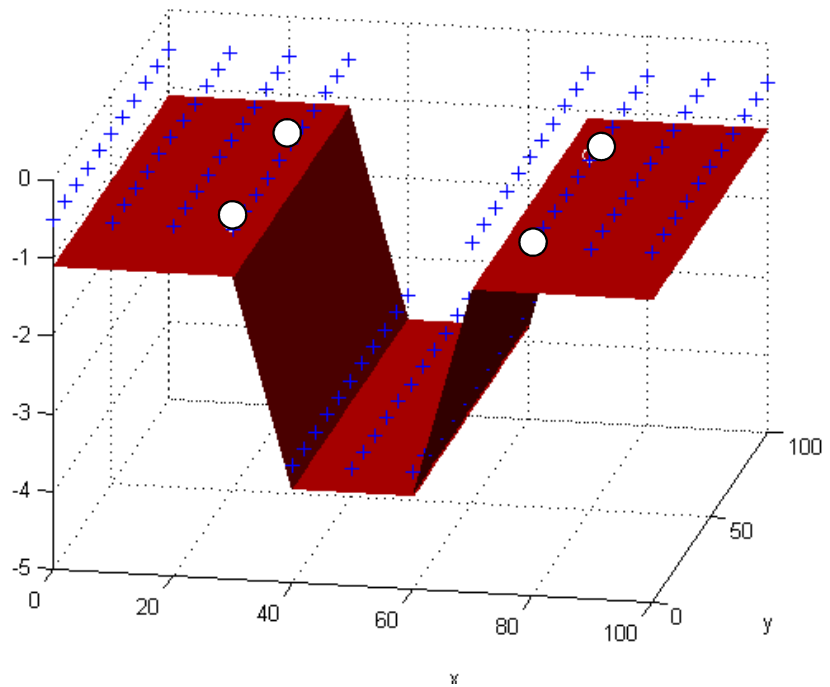
- CoO => dans certaines configurations symétriques, la contrainte sur les poids empêche de tenir compte de l'information de la variable secondaire

Ex.

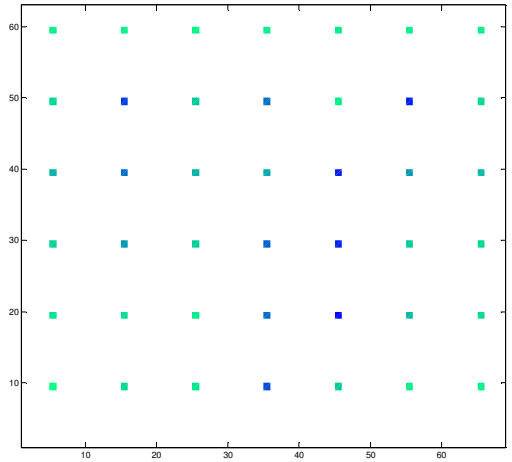


Dans cette configuration, tous les poids sur Y sont 0, peu importe la corrélation entre Z et Y !

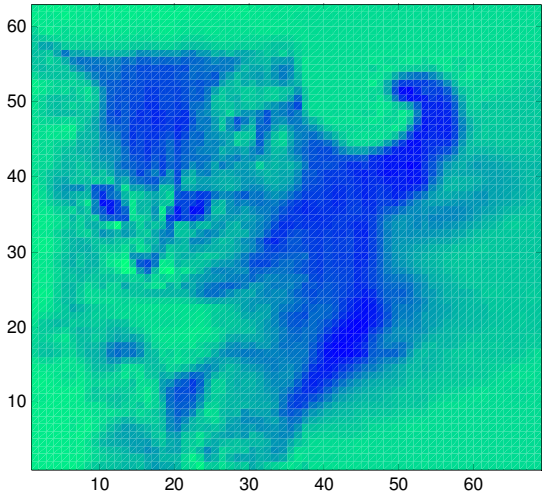
Exemple

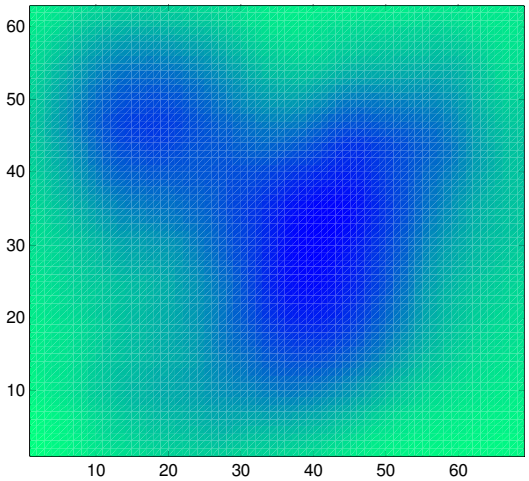


Exemple

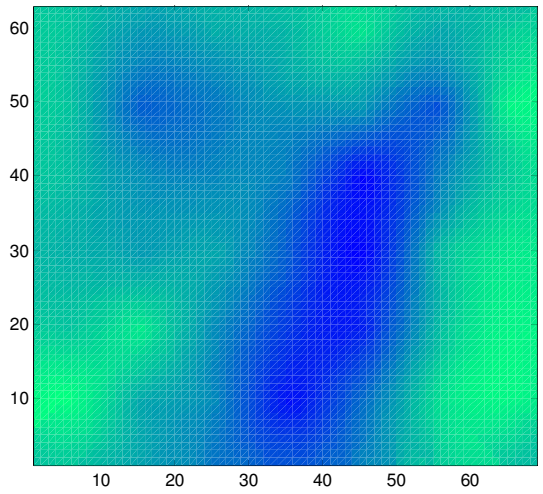


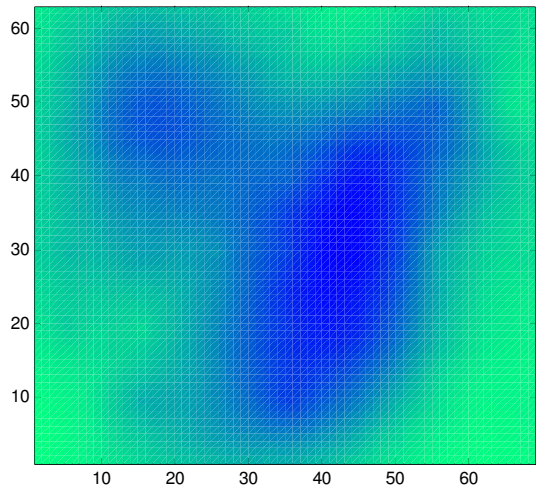
V.P.





V.S.





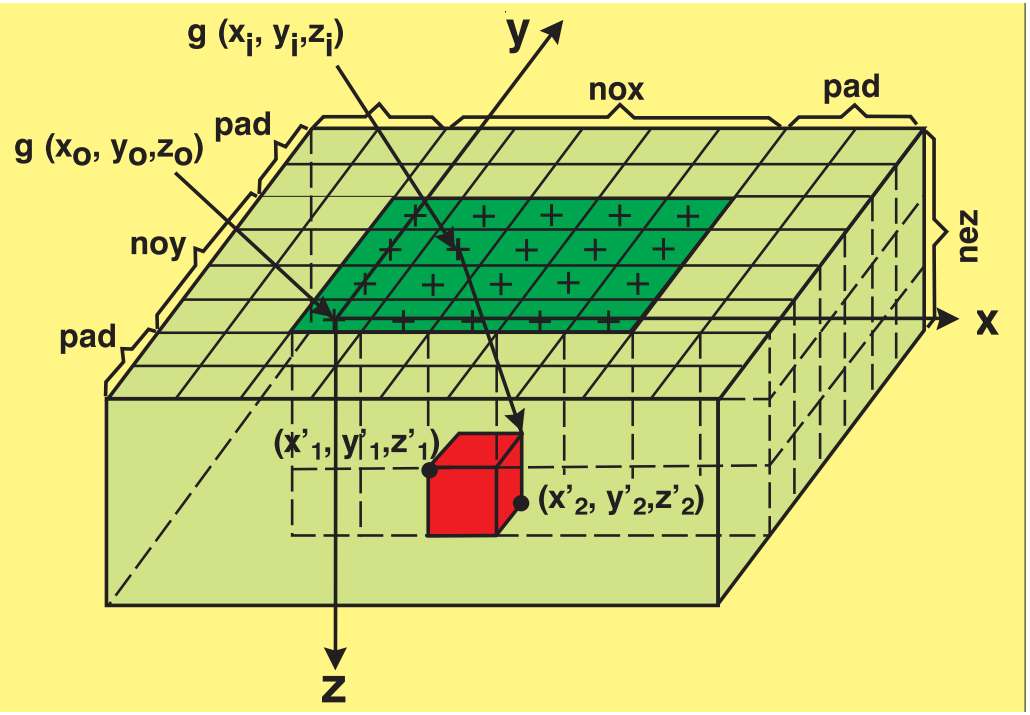
Krigeage : MAE = 26.3

Corr(Z,Z*) = 0.73

Cokrigeage simple : MAE = 22.1

Corr(Z,Z*) = 0.81

Inversion de données gravimétriques



$$\frac{g_i}{\rho_j} = -\gamma \sum_{p=1}^{2} \sum_{q=1}^{2} \sum_{r=1}^{2} \mu_{pqs} \left[x'_p \ln\left(y'_q + r_{pqs}\right) + y'_q \ln\left(x'_p + r_{pqs}\right) - z'_s \arctan\left(\frac{x'_p \, y'_q}{z'_s \, r_{pqs}} \right) \right]$$

Sous forme matricielle :

$$\boxed{g = G\rho}$$

où

g est le vecteur des anomalies gravimétriques (n x 1)
G est la matrice géométrique (n x m)
ρ est le vecteur des contrastes de densité de blocs (m x 1)

Les covariances s'écrivent alors :

$$C_{gg} = GC_{\rho\rho}G^T$$

$$C_{g\rho} = GC_{\rho\rho}$$

où

C_{gg} matrice des covariances gravité-gravité (n x n)

$C_{\rho\rho}$ matrice des covariances densité-densité (m x m)

$C_{g\rho}$ matrice des covariances gravité-densité (n x m)

G matrice géométrique (n x m)

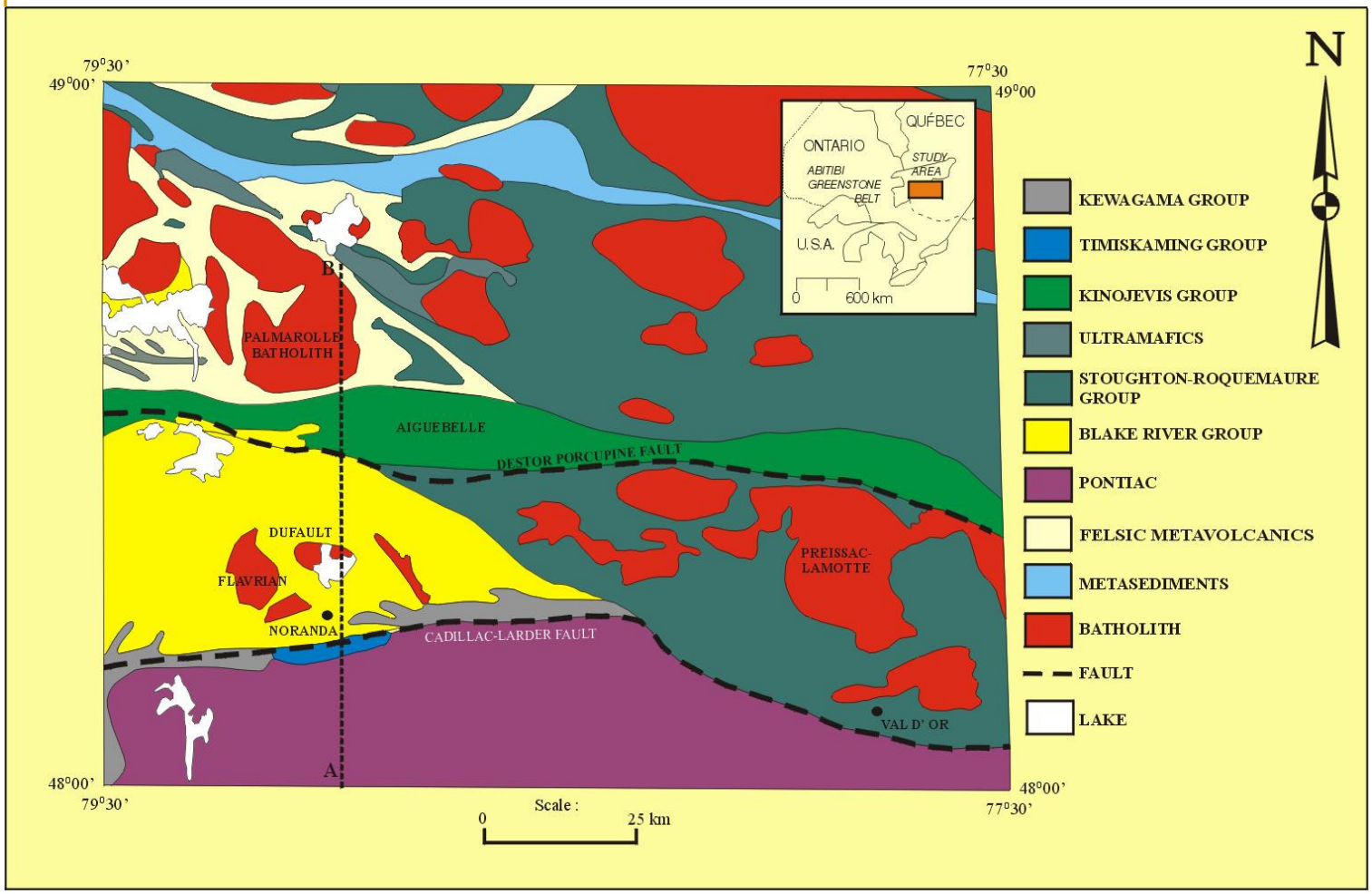
Le cokrigeage s'écrit :

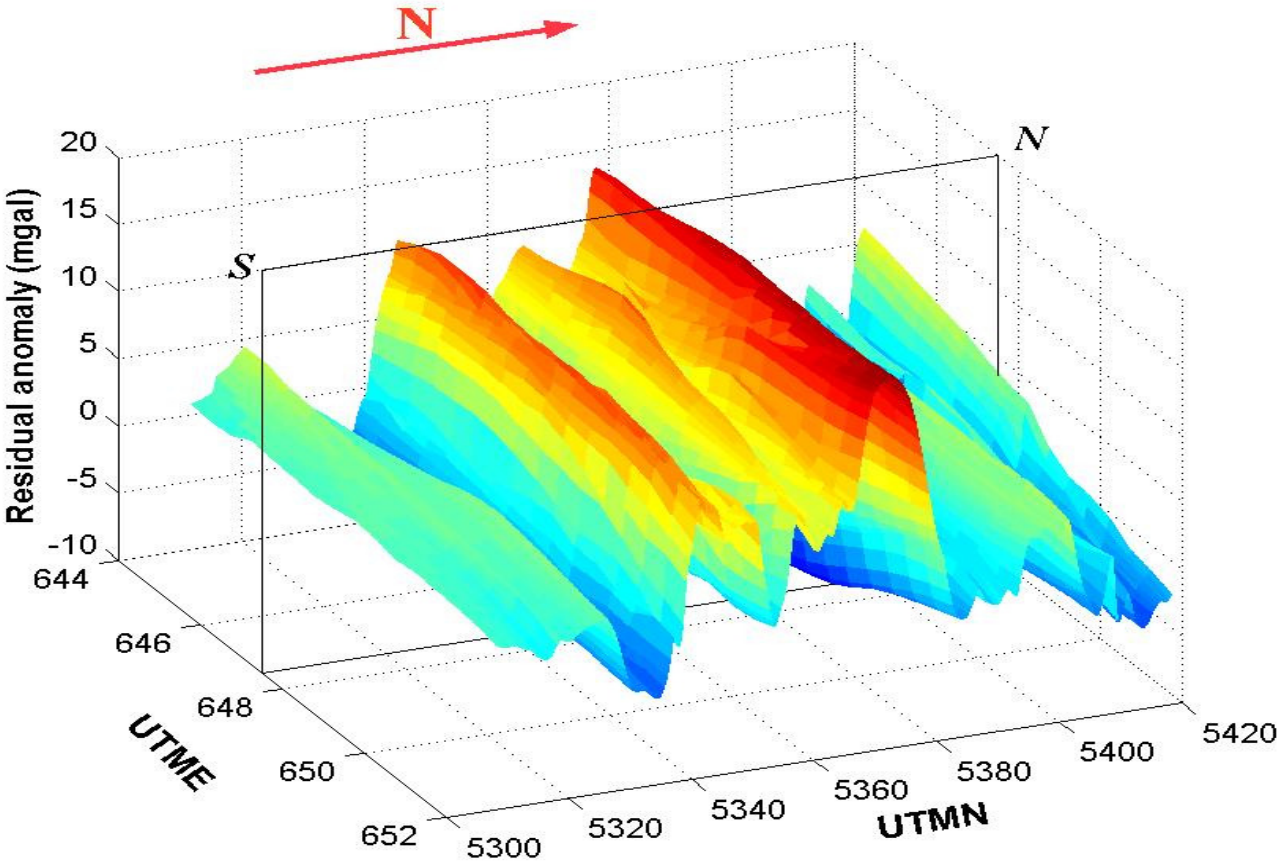
$$\boxed{\begin{aligned} \rho^* &= \Lambda' g \\ \Lambda &= C_{gg}^{-1} C_{g\rho} \end{aligned}}$$

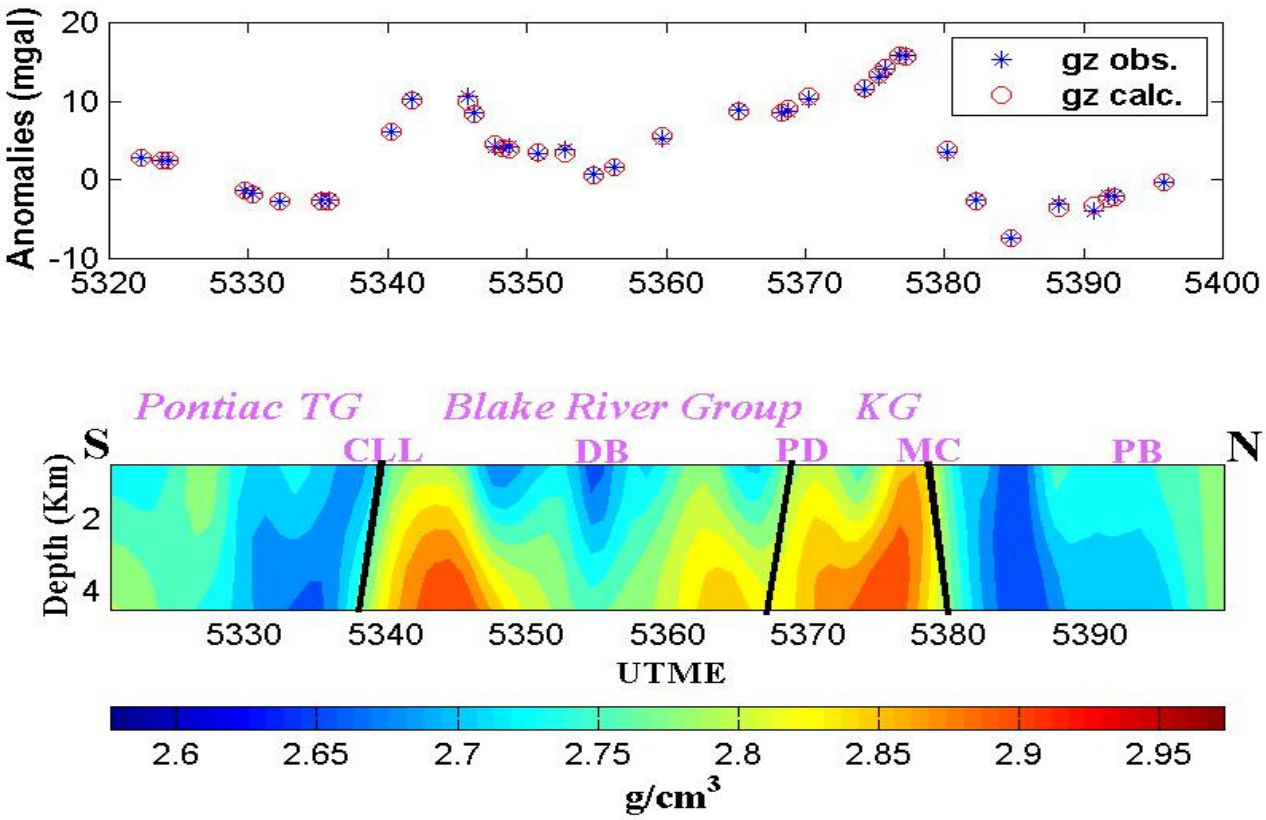
où Λ est la matrice contenant les poids de cokrigeage (n x m)

L'estimation par cokrigeage des densités reproduit *exactement* l'anomalie gravimétrique observée i.e:

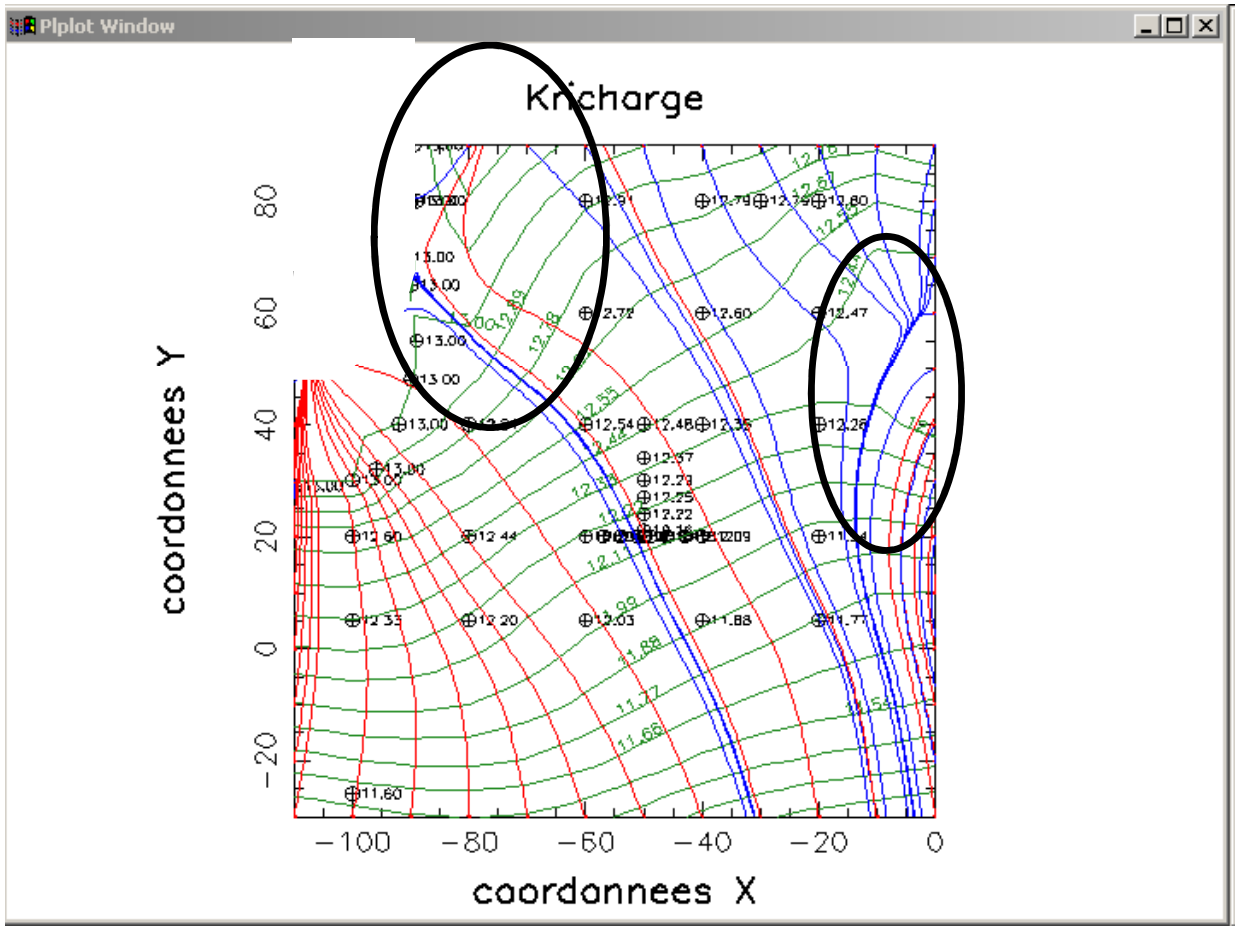
$$\boxed{g^* = G\rho^* = G\Lambda' g = GC_{\rho\rho}G'(GC_{\rho\rho}G')^{-1}g = g}$$


79°30'
49°00'
77°30
49°00
48°00'
79°30'
48°00'
77°30'
N
ONTARIO
QUÉBEC
ABITIBI GREENSTONE BELT
STUDY AREA
U.S.A.
0
600 km
B
A
PALMAROLLE BATHOLITH
AIGUEBELLE
DESTOR PORCUPINE FAULT
DUFAULT
FLAVRIAN
NORANDA
CADILLAC-LARDER FAULT
PREISSAC-LAMOTTE
VAL D' OR
Scale :
0
25 km
KEWAGAMA GROUP
TIMISKAMING GROUP
KINOJEVIS GROUP
ULTRAMAFICS
STOUGHTON-ROQUEMAURE GROUP
BLAKE RIVER GROUP
PONTIAC
FELSIC METAVOLCANICS
METASEDIMENTS
BATHOLITH
FAULT
LAKE

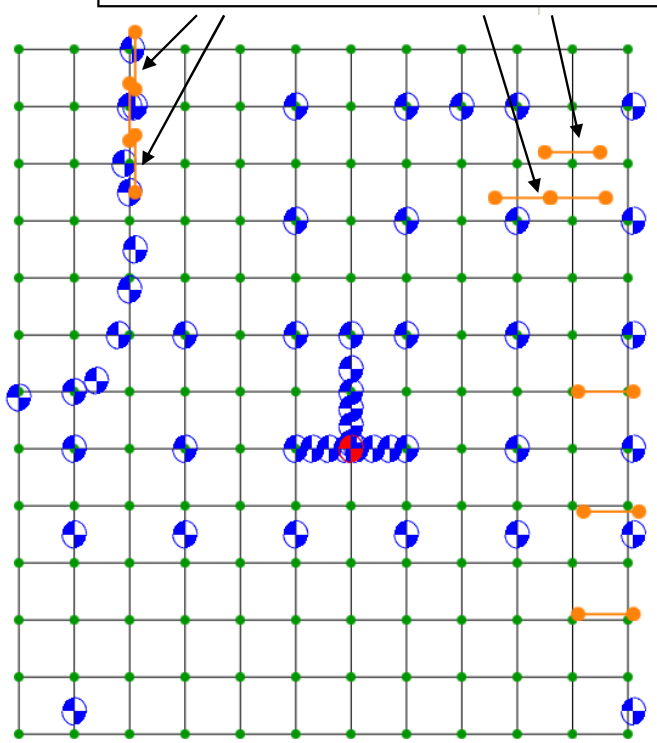

N
S
N
Residual anomaly (mgal)
20
15
10
5
0
-5
-10
644
646
648
650
652
UTME
5300
5320
5340
5360
5380
5400
5420
UTMN

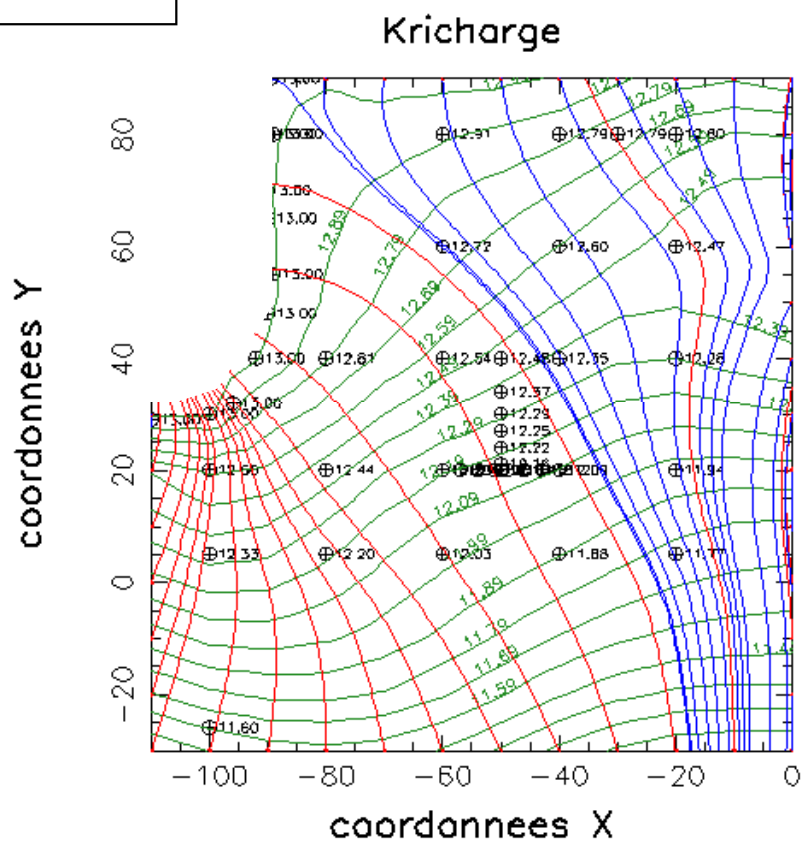

Anomalies (mgal)
20
10
0
-10
5320
5330
5340
5350
5360
5370
5380
5390
5400
gz obs.
gz calc.
Pontiac TG
Blake River Group
KG
S
N
CLL
DB
PD
MC
PB
Depth (Km)
2
4
5330
5340
5350
5360
5370
5380
5390
UTME
2.6
2.65
2.7
2.75
2.8
2.85
2.9
2.95
g/cm3

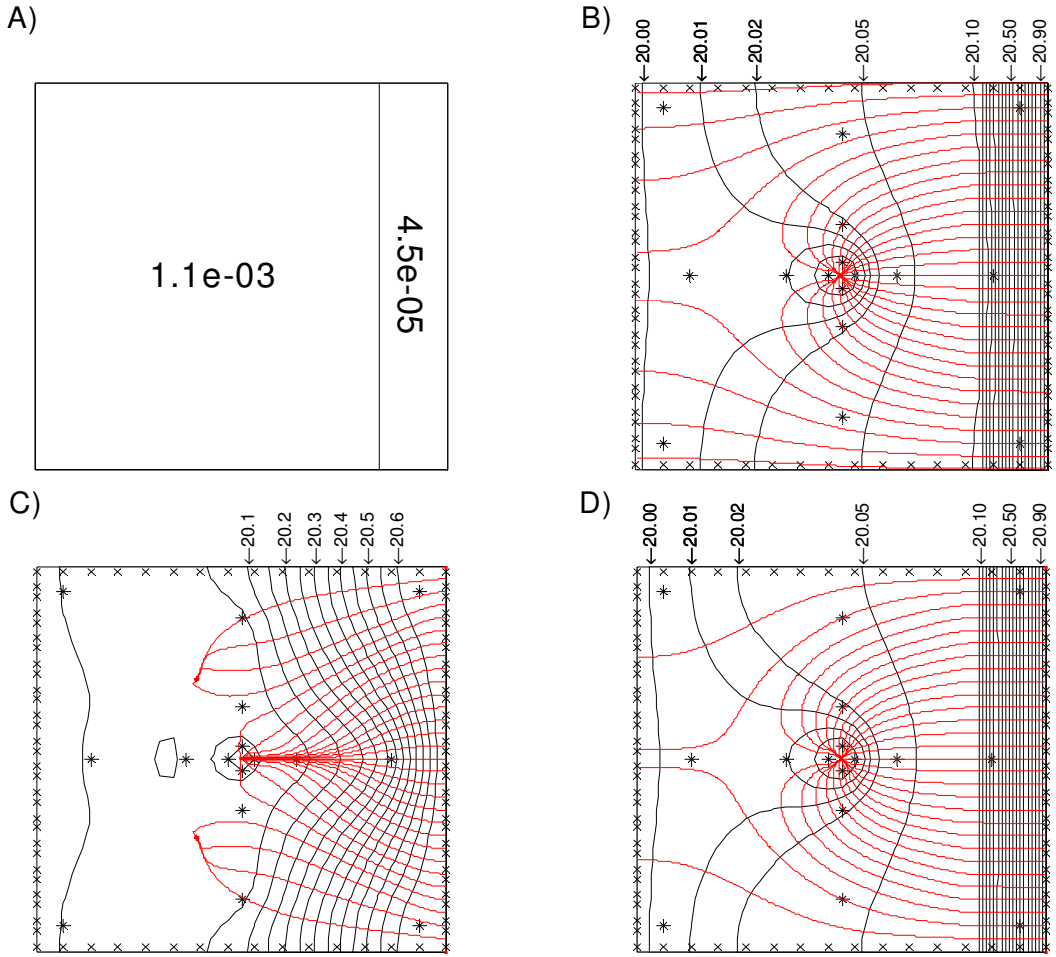
Exemple en hydrogéologie



Utilisation de points doublons : forcer localement une direction d'écoulement







Solution

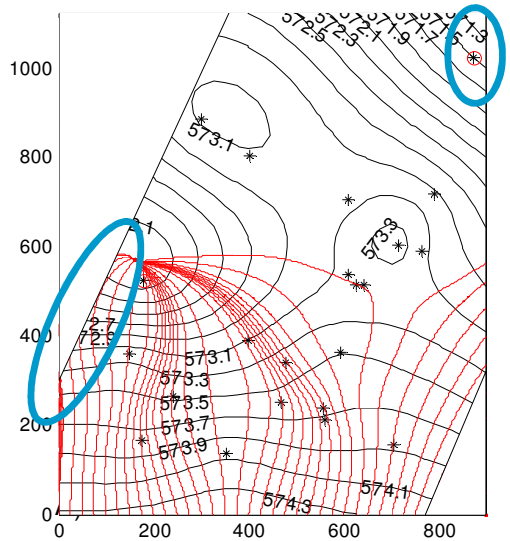
Éléments finis

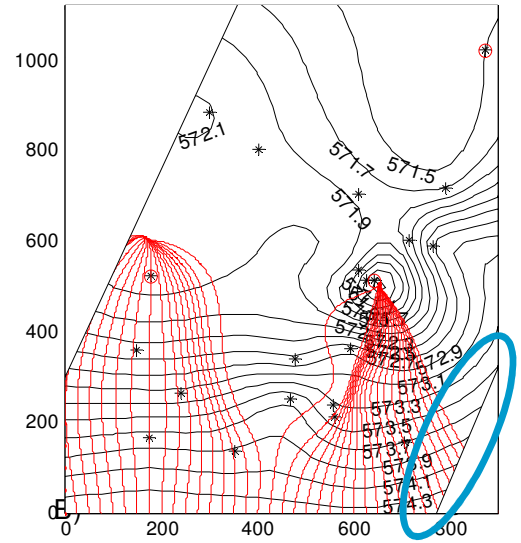
Cokrigeage 1

Cokrigeage 2

Dizy, repos

Sans doublons



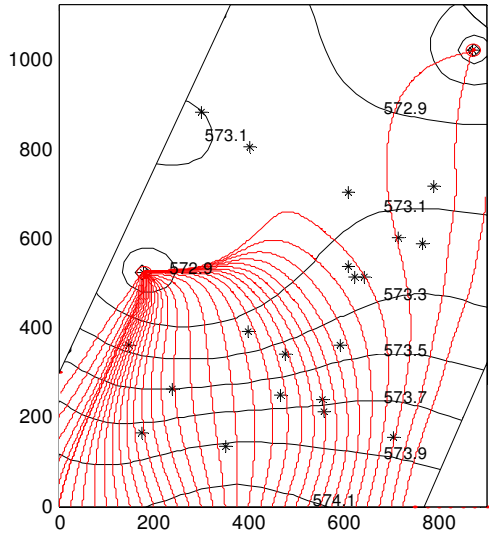


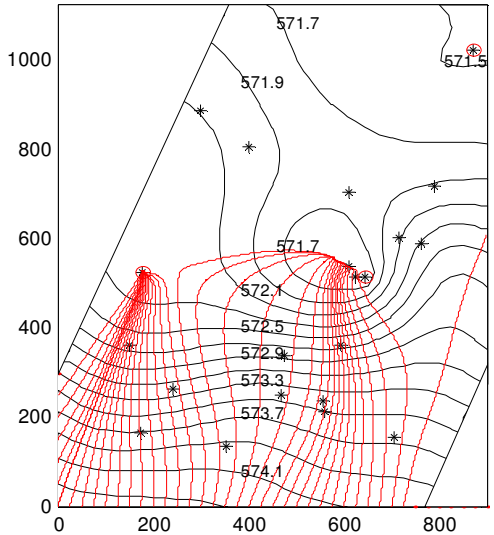
Dizy, pompage

Sans doublons

Dizy, repos

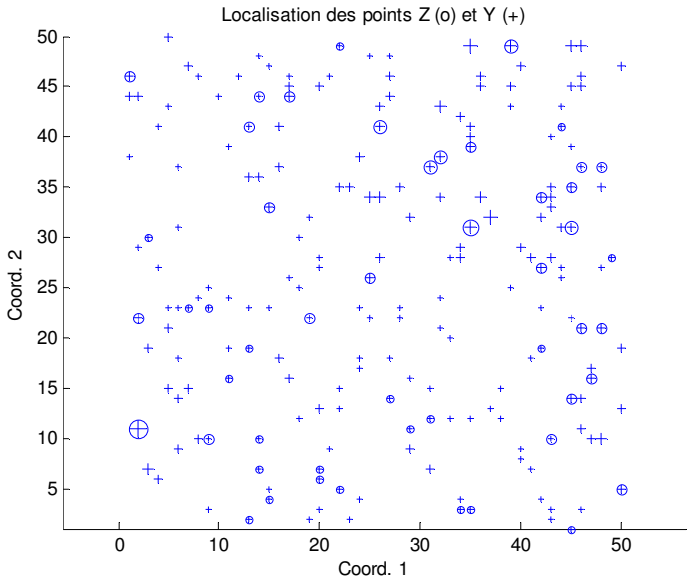
Avec doublons





Dizy, pompage

Avec doublons


Localisation des points Z (o) et Y (+)
Coord. 2
Coord. 1
50
45
40
35
30
25
20
15
10
5
0
10
20
30
40
50

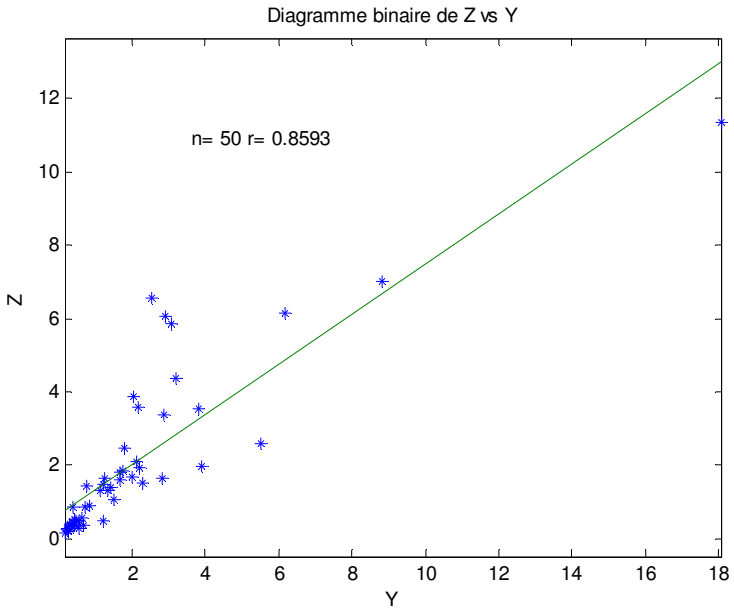
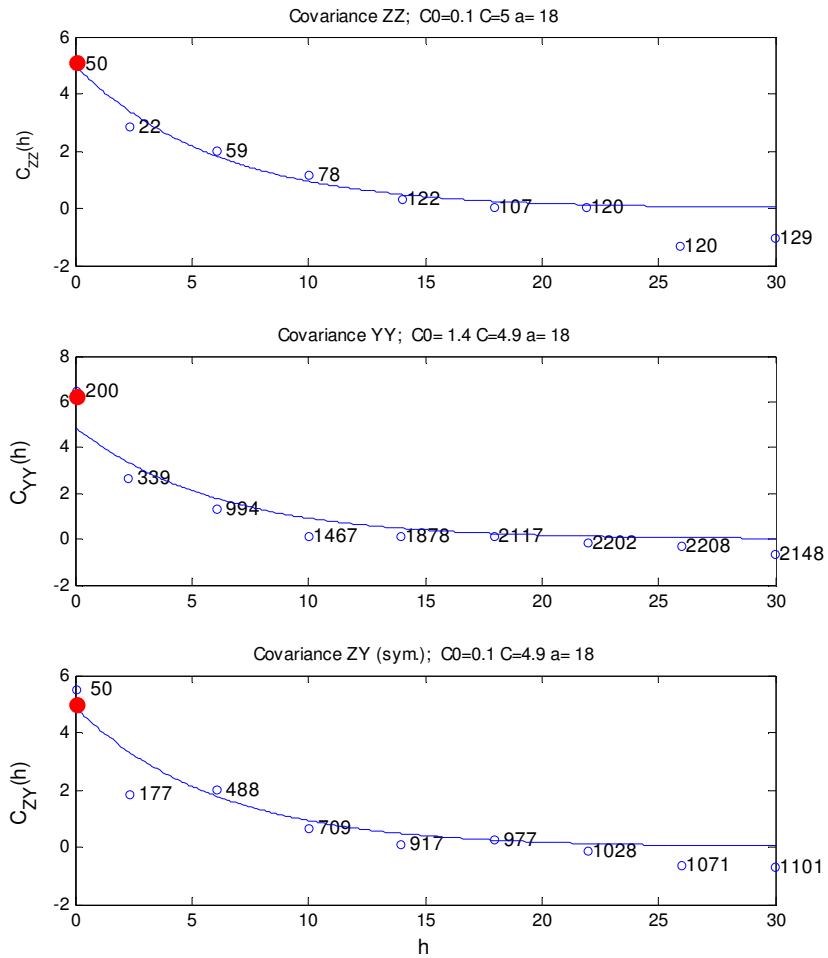
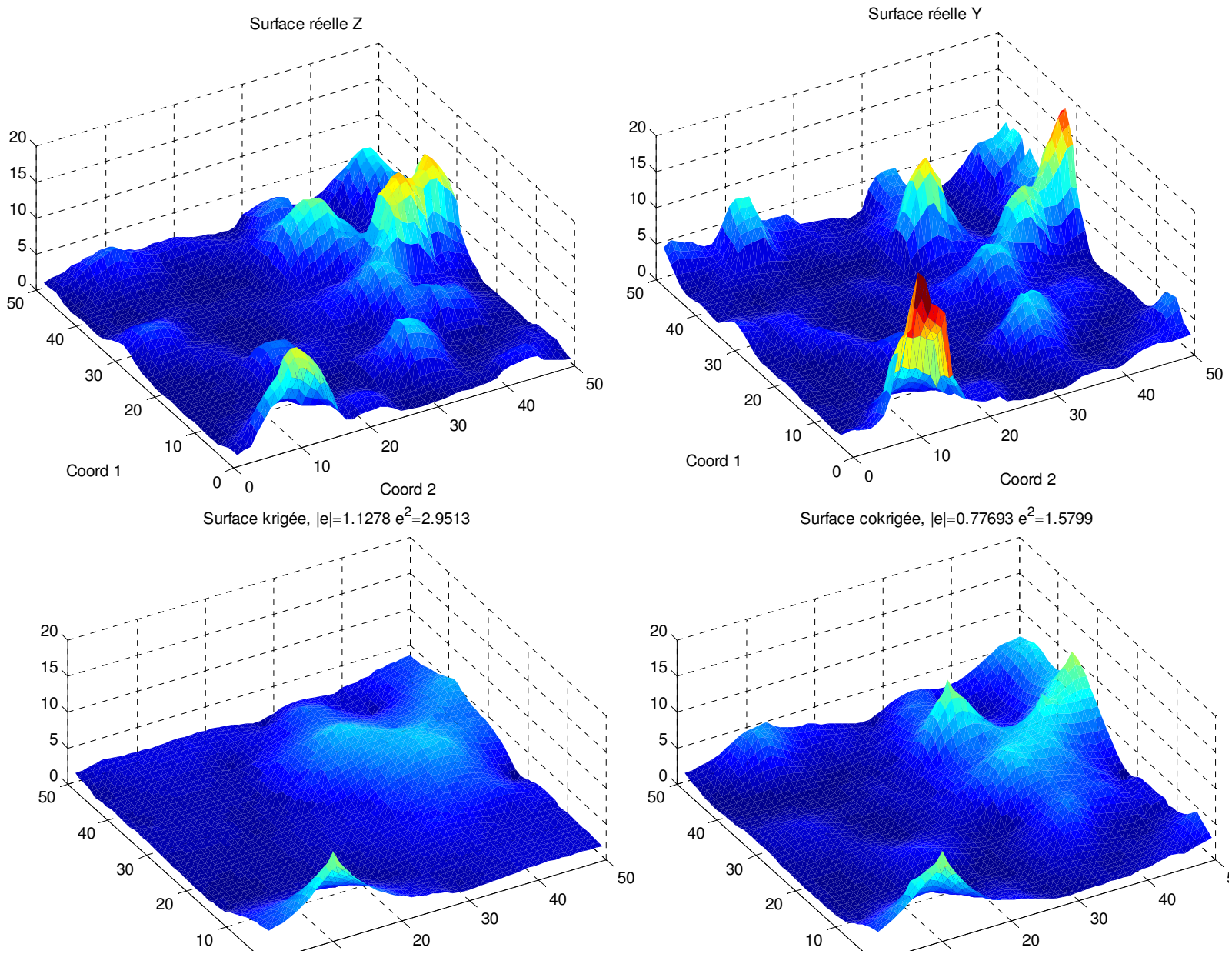
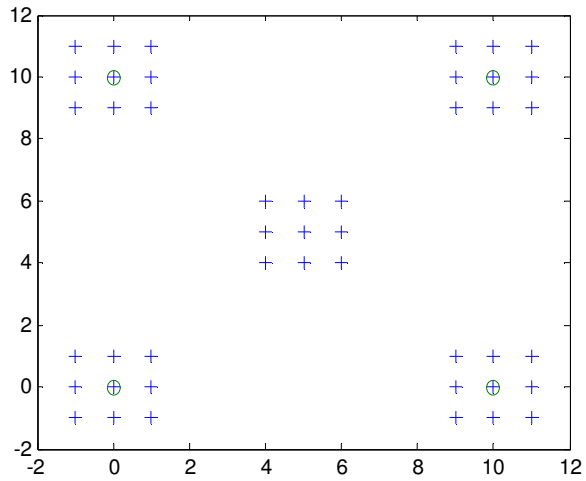


Diagramme binaire de Z vs Y
n= 50 r= 0.8593
Z
Y
12
10
8
6
4
2
0
2
4
6
8
10
12
14
16
18


Covariance ZZ; C0=0.1 C=5 a= 18
$C_{ZZ}(h)$
50
22
59
78
122
107
120
120
129
Covariance YY; C0= 1.4 C=4.9 a= 18
$C_{YY}(h)$
200
339
994
1467
1878
2117
2202
2208
2148
Covariance ZY (sym.); C0=0.1 C=4.9 a= 18
$C_{ZY}(h)$
50
177
488
709
917
977
1028
1071
1101
h


Surface réelle Z
Coord 1
Coord 2
Surface réelle Y
Coord 1
Coord 2
Surface krigée, |e|=1.1278 e2=2.9513
Surface cokrigée, |e|=0.77693 e2=1.5799



o : v. principale

+: v. secondaire

V1: 4 o et + en (5,5) (collocated)

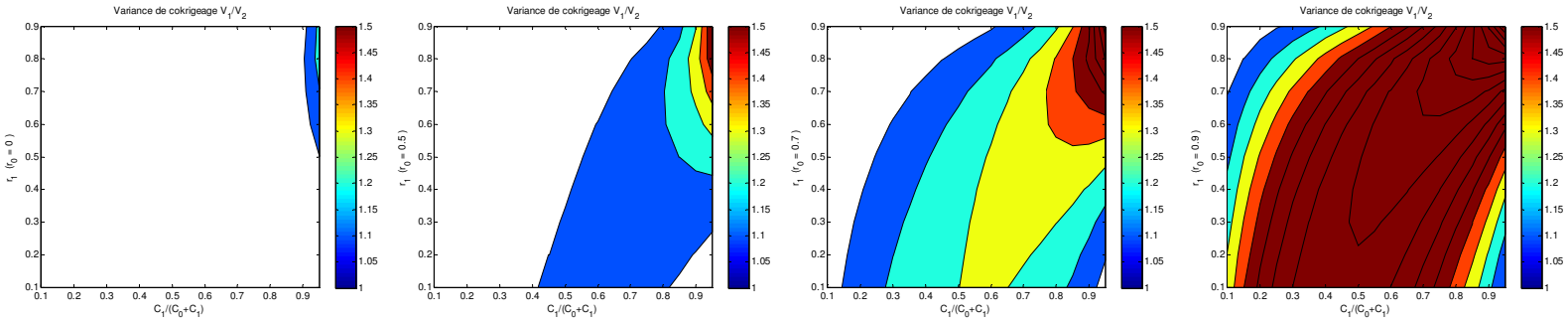
V2: 4 o et 5 + (aux 4 o et en (5,5))

V3: 4 o et 45 +

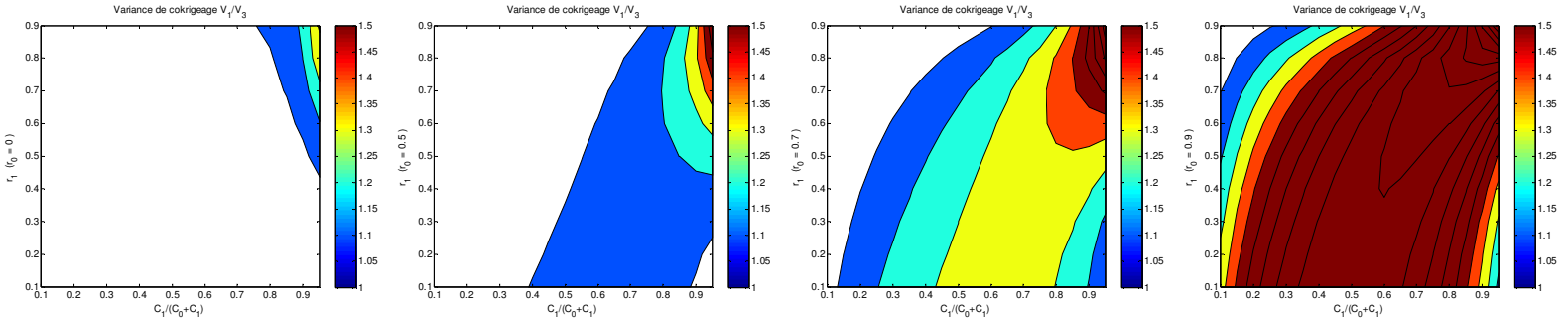
$$C(h) = \begin{bmatrix} 1 & r_0 \\ r_0 & 1 \end{bmatrix} \delta(h) + \begin{bmatrix} w & wr_1 \\ wr_1 & w \end{bmatrix} Sph(a = 50)$$

Influence du voisinage (cokrigeage simple)

V_1/V_2



$r_0=0$ $r_0=0.5$ $r_0=0.7$ $r_0=0.9$



V_1/V_3

L'exemple montre:

i. le cokrigeage colocalisé (V_1) s'avère un mauvais choix lorsque r_0 est non-nul; plus r_0 est élevé plus l'on perd en se limitant à V_1

ii. une fois incluse la v. s. aux points échantillons et au point à estimer, on gagne peu à inclure d'autres points (V1/V2 ~ V1/V3)