

PRISE EN COMPTE DE LA VARIANCE DE L'ERREUR DE MESURE

Les données expérimentales utilisées dans l'estimation des concentrations sont généralement entachées d'une erreur de mesure qu'il convient de prendre en compte dans le modèle géostatistique. Cette note a pour objet de décrire les méthodes applicables à cette fin et d'analyser les résultats obtenus pour chacune. Pour en faciliter la lecture, nous illustrons notre propos par des exemples de calcul réalisés avec des données réelles. Afin de ne pas alourdir les calculs, nous nous limitons au cas monovarié stationnaire d'ordre 2 ou de façon moins restrictive au cas intrinsèque.

Les données exploitées dans ce document ont été fournies par l'association AIR NORMAND. Elles sont issues d'une campagne de mesure de l'ozone (échantillonnage passif) conduite du 26 juin au 4 septembre 2000 par plusieurs AASQA dans le Nord de la France. La semaine du 10 au 16 juillet est traitée comme exemple.

1. Présentation des approches

1.1 Variabilité liée à la variance de l'erreur de mesure. Conséquence sur le variogramme

L'erreur de mesure peut être considérée comme une variable régionalisée de moyenne nulle, dotée d'une structure spatiale de covariance, et dont le champ se superpose à celui de la variable d'intérêt.

$$Z_\varepsilon(x_a) = Z(x_a) + \varepsilon(x_a)$$

Z_ε : variable régionalisée dont la réalisation aux points de mesure est l'ensemble des données expérimentales.

Z : variable régionalisée sous-jacente, non entachée d'erreur, que l'on cherche à estimer dans le champ à partir des données mesurées.

ε : erreur de mesure inconnue, vérifiant $E(\boldsymbol{\varepsilon}_a) = \mathbf{0}$. (On note désormais $\varepsilon(x_a) = \varepsilon_a$)

Avec l'hypothèse d'indépendance entre Z et $\boldsymbol{\varepsilon}$, les covariances de Z , Z_ε et $\boldsymbol{\varepsilon}$ vérifient l'égalité :

$$C^{Z_\varepsilon}(h) = C^Z(h) + C^\varepsilon(h)$$

En outre, on suppose le plus souvent que l'erreur de mesure n'est pas spatialement corrélée, soit $C^\varepsilon(h) = 0, \forall h \neq 0$.

Pour $h=0$, la variance moyenne de l'erreur de mesure joue le rôle d'un effet de pépite, noté σ_ε^2 , qui s'ajoute à la variabilité à très courte échelle (variabilité apparente, due à l'existence d'une microstructure qui n'est pas détectée par l'échantillonnage).

Les variogrammes de Z et de Z_ε sont ainsi liés par la relation suivante :

$$\gamma^Z_\varepsilon(h) = \gamma^Z(h) + \sigma_\varepsilon$$

Si à chaque point d'échantillonnage x_α est associée une variance de l'erreur de mesure $V_\alpha = \text{Var}(\varepsilon_\alpha)$, alors l'effet de pépite apparent causé par l'erreur de mesure est la moyenne des V_α soit $\sigma_\varepsilon^2 = \bar{V}_\alpha$.

D'un point de vue méthodologique, deux approches permettent d'intégrer la variance de l'erreur de mesure (VEM) dans le modèle géostatistique, suivant que **l'on considère une VEM moyenne dans le champ** ou que **l'on prend en compte distinctement toutes les valeurs V_α** .

1.2 Prise en compte de la variance de l'erreur de mesure. Conséquence sur la variance de krigeage

Le but est d'estimer la concentration Z par krigeage à partir des données mesurées :

$$Z^*(x_0) = \sum_i \lambda_i Z^\varepsilon(x_i)$$

Le krigeage rend minimale la variance d'estimation qui vaut :

$$\sigma_{K0}^2 = \text{Var}[Z(x_0) - Z^*(x_0)] = \text{Var}[Z(x_0) - \sum_i \lambda_i (Z(x_i) + \varepsilon_i)]$$

Dans le cadre stationnaire d'ordre 2, et moyennant l'hypothèse d'indépendance entre Z et ε , cette quantité s'écrit :

$$\begin{aligned} \sigma_{K0}^2 &= C^z(0) + \sum_{i,j} \lambda_i \lambda_j C_{ij}^z + \sum_i \lambda_i^2 \text{Var}(\varepsilon_i) - 2 \sum_i \lambda_i C_{0i}^z \\ &= C^z(0) + \sum_{i \neq j} \lambda_i \lambda_j C_{ij}^z + \sum_i \lambda_i^2 (C_{ii}^z + \text{Var}(\varepsilon_i)) - 2 \sum_i \lambda_i C_{0i}^z \end{aligned}$$

a) On suppose la variance de l'erreur de mesure constante dans le champ.

Cette variance est alors modélisée par un **effet de pépite σ_ε^2 égal à la VEM moyenne**. Pour tout point x_α , on substitue donc σ_ε^2 à $\text{Var}(\varepsilon_\alpha)$, ce qui donne :

$$\sigma_{K0}^2 = C^z(0) + \sum_{i \neq j} \lambda_i \lambda_j C_{ij}^{z\varepsilon} + \sum_i \lambda_i^2 [C_{ii}^{z\varepsilon}(0)] - 2 \sum_i \lambda_i C_{0i}^{z\varepsilon}$$

avec pour tout $i \neq j$, $C_{ij}^{z\varepsilon} = C_{ij}^z$ et $C^{z\varepsilon}(0) = C^z(0) + \sigma_\varepsilon^2$

$C^{z\varepsilon}$ est le modèle de covariance avec effet de pépite, construit à partir des données expérimentales (donc entachées d'erreur). En pratique, avec le logiciel Isatis, rien ne permet de préciser que cet effet de pépite est dû en totalité ou en partie à une erreur de mesure. Le logiciel ne fait pas de distinction entre Z et Z_ε . Tout se passe donc comme si l'on estimait la variable bruitée Z_ε , et l'on a alors :

$$\sigma_{K0}^2 = C^{z\varepsilon}(0) + \sum_{i,j} \lambda_i \lambda_j C_{ij}^{z\varepsilon} - 2 \sum_i \lambda_i C_{0i}^{z\varepsilon}$$

Cette variance intègre la variance de l'erreur de mesure σ_ε^2 dans son terme $C^{z\varepsilon}(0)$.

b) On spécifie la valeur de $\text{Var}(\epsilon_i)$ en tout point d'échantillonnage.

Dans cette approche, la variable que l'on cherche à estimer est la variable non bruitée Z . La variance de krigeage a pour expression – cas d'une estimation ponctuelle - :

$$\sigma_{K_0}^2 = C^z(0) + \sum_{i,j} \lambda_i \lambda_j C_{ij}^Z + \sum_i \lambda_i^2 \text{Var} \epsilon_i - 2 \sum_i \lambda_i C_{0i}^Z$$

En pratique, avec Isatis, le modèle C^Z est déduit du modèle $C^{Z\epsilon}$ par soustraction de la variance moyenne de l'erreur de mesure. Le terme $C^Z(0)$ n'inclut donc pas cette variance moyenne.

Si en tout point x_α , $\text{Var}(\epsilon_\alpha)$ a la même valeur σ_ϵ^2 , alors minimiser la variance de krigeage conduit au même système de krigeage que dans le cas précédent, et par suite à la même concentration estimée. En revanche la variance de krigeage calculée par Isatis vaut :

$$\begin{aligned} \sigma_{K_0}^2 &= C^z(0) + \sum_{i \neq j} \lambda_i \lambda_j C_{ij}^Z + \sum_i \lambda_i^2 (C_{ii}^Z + \sigma_\epsilon^2) - 2 \sum_i \lambda_i C_{0i}^Z \\ &= C^z(0) + \sum_{i,j} \lambda_i \lambda_j C_{ij}^{Z\epsilon} - 2 \sum_i \lambda_i C_{0i}^{Z\epsilon} \end{aligned}$$

Elle est donc inférieure de σ_ϵ^2 à la variance précédemment obtenue.

Dans une validation croisée, l'évaluation de la précision de l'estimation aux points expérimentaux implique de comparer les valeurs estimées et mesurées de la même variable, c'est-à-dire de la variable bruitée Z_ϵ . Il convient pour ce faire d'ajouter la VEM à l'écart-type de krigeage. L'erreur quadratique moyenne de prédiction vaut alors :

$$E[Z^\epsilon - \sum_i \lambda_i Z^\epsilon(x_i)]^2 = \sigma_{K_0}^2 + \text{Var}(\epsilon_0)$$

Ce calcul n'est possible qu'aux points d'échantillonnage. En dehors de ces points, la VEM est inconnue. Pour généraliser le calcul de l'erreur quadratique de prédiction à l'ensemble du domaine d'étude, on pourrait être tenté d'estimer en tout point la VEM par krigeage. Une telle procédure n'est pas acceptable en théorie, puisque la VEM n'est pas une variable additive.

Si toutefois la VEM varie peu d'un site de mesure à un autre, il est envisageable d'ajouter à l'écart-type de krigeage la variance moyenne de l'erreur de mesure, calculée sur l'ensemble des points expérimentaux.

Remarque : cas du krigeage de blocs

Dans un krigeage de blocs, la variance d'estimation avec variance de l'erreur de mesure s'écrit :

$$\sigma_{K_v}^2 = C^z(v, v) + \sum_{i \neq j} \lambda_i \lambda_j C_{ij}^Z + \sum_i \lambda_i^2 (C_{ii}^Z + \text{Var} \epsilon_i) - 2 \sum_i \lambda_i C^Z(v, x_i)$$

$$\text{Or } C^{Z\epsilon}(v, v) = \frac{1}{v^2} \iint_v C^{Z\epsilon}(x, y) dx dy = \frac{1}{v^2} \int_v C^Z(x, y) dx dy = C^Z(v, v)$$

Donc si en tout point x_α $\text{Var}(\epsilon_\alpha)$ vaut σ_ϵ^2 , alors

$$\begin{aligned} \sigma_{Kv}^2 &= C^z(v, v) + \sum_{i \neq j} \lambda_i \lambda_j C_{ij}^z + \sum_i \lambda_i^2 (C_{ii}^z + \sigma_\varepsilon^2) - 2 \sum_i \lambda_i C^z(v, x_i) \\ &= C^{z_\varepsilon}(v, v) + \sum_{i,j} \lambda_i \lambda_j C_{ij}^{z_\varepsilon} - 2 \sum_i \lambda_i C^{z_\varepsilon}(v, x_i) \end{aligned}$$

Dans ce cas, la variance d'estimation ne présente aucune différence avec celle que l'on obtient par un krigeage avec effet de pépite.

2. Mise en œuvre des différentes méthodes

Les deux approches décrites au paragraphe 1.2 sont appliquées à des données d'ozone (deuxième et troisième cas). En introduction (premier cas), nous présentons la situation où l'information sur l'erreur de mesure et la variabilité à très courte échelle est insuffisante on indisponible.

1.1 Premier cas : ajustement graphique d'un effet de pépite

L'effet de pépite est déterminé en extrapolant aux très courtes distances le variogramme expérimental (Figure 1). On ne cherche pas à distinguer la part de la variance effectivement due à l'erreur de mesure et celle qui a pour cause l'existence d'une microstructure non détectée par l'échantillonnage.

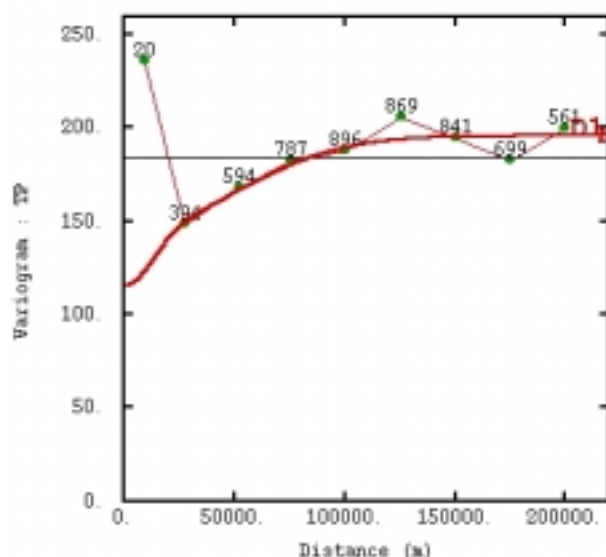


Figure 1 - Cas 1. Variogramme expérimental¹ et modélisé. Effet de pépite = 115

¹ Le variogramme expérimental est calculé uniquement à partir des données de type rural, qui sont supposées décrire la variabilité de l'ozone à l'échelle de la région.

Le variogramme théorique (courbe continue sur la figure1) a pour formule :

$$h > 0 : \gamma^a(h) = 115 + 29 \cdot (1 - \exp(-\frac{\delta \cdot h^2}{30000^2})) + 51,5 \cdot (1 - \exp(-\frac{\delta \cdot h^2}{120000^2}))$$

$$h = 0 : \gamma^a(0) = 0$$

Le modèle correspondant de covariance existe et s'écrit :

$$h > 0 : C^a(h) = \gamma^a(\infty) - \gamma^a(h) = (115 + 29 + 51,5) - \gamma^a(h) = 29 \cdot \exp(-\frac{\delta \cdot h^2}{30000^2}) + 51,5 \cdot \exp(-\frac{\delta \cdot h^2}{120000^2})$$

$$h = 0 : C^a(0) = 197$$

Soit x_0 , un point où l'on souhaite estimer la concentration par krigeage.

Soit x_1, \dots, x_n les points expérimentaux situés dans le voisinage d'estimation de ce point. Afin de ne pas alourdir la présentation des résultats, le voisinage adopté dans les calculs se limite à un domaine circulaire de 40 km de rayon centré sur x_0 . Il contient treize points de mesure.

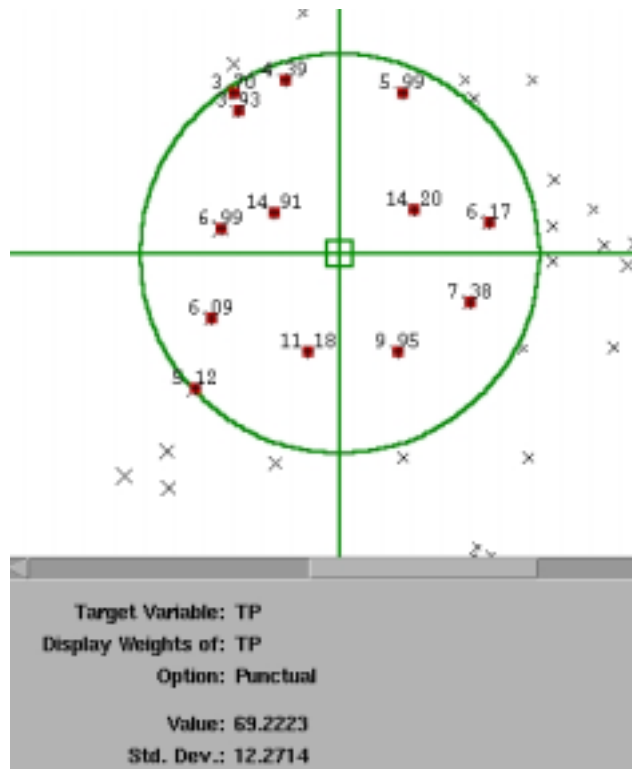


Figure 2 - Cas 1. Représentation du voisinage de krigeage et des poids de krigeage affectés aux 13 points expérimentaux situés dans ce voisinage.

L'estimateur recherché est de la forme :

$$Z_0^* = \sum_{i=1}^{13} \lambda_i^a Z_i^e$$

Le système de krigeage permettant de calculer les poids λ_i s'écrit :

- avec le variogramme

$$\begin{bmatrix} 0 & \gamma^a_{12} & \dots & \gamma^a_{1i} & \dots & \gamma^a_{113} & 1 \\ \gamma^a_{21} & 0 & \dots & \gamma^a_{2i} & \dots & \gamma^a_{213} & 1 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ \gamma^a_{i1} & \gamma^a_{i2} & \dots & 0 & \dots & \gamma^a_{i13} & 1 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ \gamma^a_{131} & \gamma^a_{132} & \dots & \gamma^a_{13i} & \dots & 0 & 1 \\ 1 & 1 & \dots & 1 & \dots & 1 & 0 \end{bmatrix} \times \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_i \\ \vdots \\ \lambda_{13} \\ \mu \end{bmatrix} = \begin{bmatrix} \gamma^a_{01} \\ \gamma^a_{02} \\ \vdots \\ \gamma^a_{0i} \\ \vdots \\ \gamma^a_{013} \\ \mu \end{bmatrix} \quad (\text{Equ.1})$$

$$\sigma_{K0}^2 = 2 \sum_i \lambda_i^a \gamma^a_{oi} - \sum_{i,j} \lambda_i^a \lambda_j^a \gamma^a_{ij}$$

- avec la covariance

$$\begin{bmatrix} C^a_{11} & C^a_{12} & \dots & C^a_{1i} & \dots & C^a_{113} & 1 \\ C^a_{21} & C^a_{22} & \dots & C^a_{2i} & \dots & C^a_{213} & 1 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ C^a_{i1} & C^a_{i2} & \dots & C^a_{ii} & \dots & C^a_{i13} & 1 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ C^a_{131} & C^a_{132} & \dots & C^a_{13i} & \dots & C^a_{1313} & 1 \\ 1 & 1 & \dots & 1 & \dots & 1 & 0 \end{bmatrix} \times \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_i \\ \vdots \\ \lambda_{13} \\ \mu \end{bmatrix} = \begin{bmatrix} C^a_{01} \\ C^a_{02} \\ \vdots \\ C^a_{0i} \\ \vdots \\ C^a_{013} \\ \mu \end{bmatrix} \quad (\text{Equ. 2})$$

$$C^a_{11} = \dots = C^a_{ii} = \dots = C^a_{nn} = C^a(0) = 197$$

$$\sigma_K^2 = C^1(0) - 2 \sum_i \lambda_i^a C^1_{oi} + \sum_{i,j} \lambda_i^a \lambda_j^a C^1_{ij}$$

Les poids de krigeage ont des valeurs relativement proches les unes des autres. En effet, à cause de l'effet de pépité élevé, les termes diagonaux de la matrice de covariance, qui décrivent à la variabilité à l'origine, sont prépondérants par rapport aux autres termes qui représentent la variabilité à moyenne échelle. En conséquence, les poids de krigeage sont peu influencés par les distances entre points et la concentration estimée s'approche de la moyenne expérimentale des données.

1.2 Deuxième cas : effet de pépité déterminé par la variance moyenne de l'erreur de mesure

On considère que l'effet de pépité procède entièrement de l'erreur de mesure. Sa valeur est prise égale à la variance moyenne de l'erreur de mesure (Figure 3).

Cette hypothèse s'applique surtout aux polluants qui varient de façon continue et régulière dans l'espace (ex : ozone). Pour les polluants qui varient de manière plus erratique, un échantillonnage resserré serait nécessaire afin de caractériser la variabilité aux très courtes

distances. Celle-ci pourrait être alors modélisée par un effet de pépite additionnel ou par une structure à très courte portée.

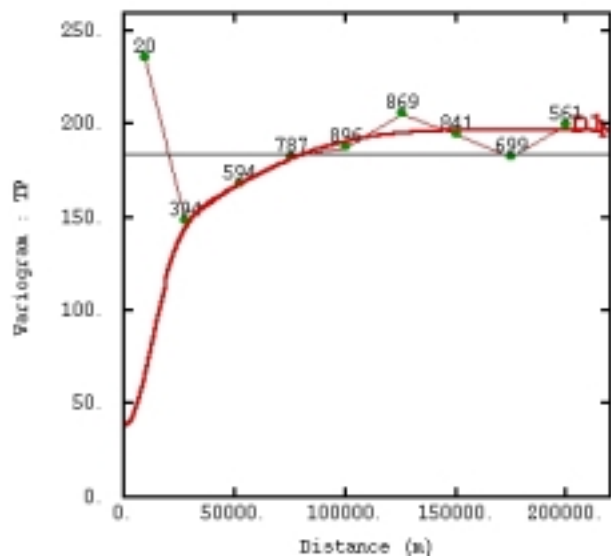


Figure 3 - Cas 2. Variogramme expérimental¹ et modélisé. Effet de pépite = 38

Le modèle de variogramme est :

$$h>0 : \gamma^b(h) = 38 + 107 \cdot (1 - \exp(-\frac{\delta \cdot h^2}{30000^2})) + 52 \cdot (1 - \exp(-\frac{\delta \cdot h^2}{120000^2})) = 38 + \gamma^b(h)$$

$$h=0 : \gamma^b(0)=0$$

Et le modèle correspondant de covariance :

$$h>0 : C^b(h) = C^b(0) - \gamma^b(h) = (38 + 107 + 52) - \gamma^b(h) = 197 - \gamma^b(h) = 107 \cdot \exp(-\frac{\delta \cdot h^2}{30000^2}) + 52 \cdot \exp(-\frac{\delta \cdot h^2}{120000^2})$$

$$h=0 : C^b(0)=197$$

Le système de krigeage est identique au précédent, en remplaçant γ^a par γ^b . Il s'écrit :

- avec le variogramme

$$\begin{bmatrix} 0 & \gamma^b_{12} & \dots & \gamma^b_{1i} & \dots & \gamma^b_{113} & 1 \\ \gamma^b_{21} & 0 & \dots & \gamma^b_{2i} & \dots & \gamma^b_{213} & 1 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ \gamma^b_{i1} & \gamma^b_{i2} & \dots & 0 & \dots & \gamma^b_{in} & 1 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ \gamma^b_{n1} & \gamma^b_{n2} & \dots & \gamma^b_{ni} & \dots & 0 & 1 \\ 1 & 1 & \dots & 1 & \dots & 1 & 0 \end{bmatrix} \times \begin{bmatrix} \lambda^b_1 \\ \lambda^b_2 \\ \vdots \\ \lambda^b_i \\ \vdots \\ \lambda^b_n \\ \mu \end{bmatrix} = \begin{bmatrix} \gamma^b_{01} \\ \gamma^b_{02} \\ \vdots \\ \gamma^b_{0i} \\ \vdots \\ \gamma^b_{013} \\ \mu \end{bmatrix} \quad (\text{Equ. 3})$$

$$Z_0^* = \sum_1^{13} \lambda_i^b Z_i$$

$$\sigma_{K0}^2 = 2 \sum_i \lambda_i \gamma^b_{0i} - \sum_{i,j} \lambda_i^b \lambda_j^b \gamma^b_{ij}$$

- avec la covariance

$$\begin{bmatrix} C^b_{11} & C^b_{12} & \dots & C^b_{1i} & \dots & C^b_{113} & 1 \\ C^b_{21} & C^b_{22} & \dots & C^b_{2i} & \dots & C^b_{213} & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ C^b_{i1} & C^b_{i2} & \dots & C^b_{ii} & \dots & C^b_{i13} & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ C^b_{131} & C^b_{132} & \dots & C^b_{13i} & \dots & C^b_{1313} & 1 \\ 1 & 1 & \dots & 1 & \dots & 1 & 0 \end{bmatrix} \times \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_i \\ \vdots \\ \lambda_{13} \\ \mu \end{bmatrix} = \begin{bmatrix} C^b_{01} \\ C^b_{02} \\ \vdots \\ C^b_{0i} \\ \vdots \\ C^b_{013} \\ \mu \end{bmatrix} \quad (\text{Equ. 4})$$

Avec $C^b_{11} = \dots = C^b_{ii} = \dots = C^b_{nn} = C^b(0) = 197$

$$\sigma_K^2 = C^b(0) - 2 \sum_i \lambda_i C^b_{oi} + \sum_{i,j} \lambda_i \lambda_j C^b_{ij}$$

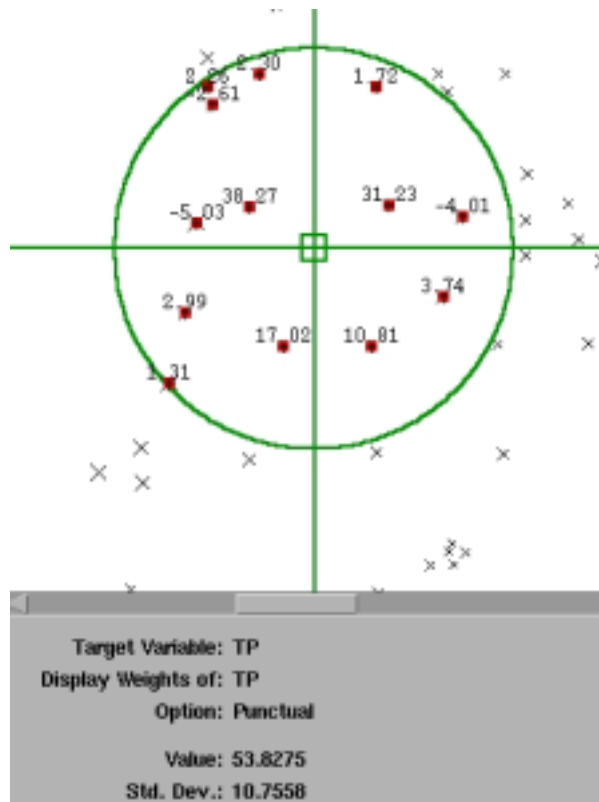


Figure 4 - Cas 2. Représentation du voisinage de krigage et des poids de krigage affectés aux 13 points expérimentaux situés dans ce voisinage.

Les poids de krigage sont plus contrastés (Figure 4), ce qui est la conséquence attendue d'un moindre effet de pépité.

1.3 Troisième cas : krigage avec variance de l'erreur de mesure

Cette méthode consiste à introduire dans l'algorithme de krigage la variance de l'erreur de mesure associée à chaque donnée expérimentale. Le modèle de variogramme est identique au modèle ajusté dans le deuxième cas (pour chaque structure, les valeurs des paramètres de

portée et de palier restent inchangées), mais l'effet de pépite est réduit exactement de σ_ϵ^2 (Figure 5).

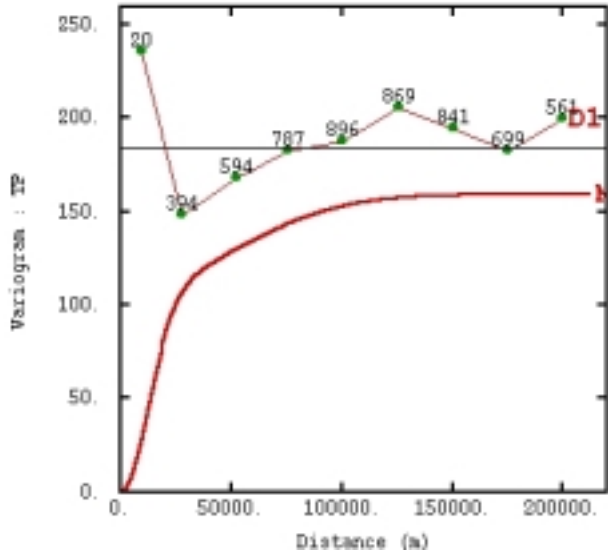


Figure 5 - Cas 3. La variabilité due à l'erreur de mesure n'est pas prise en compte dans le variogramme mais introduite dans la matrice de krigeage (cf. ci-dessous).

Le modèle de variogramme s'écrit :

$$\gamma^3(h) = \gamma^b(h) = 107 \cdot (1 - \exp(-\frac{\delta \cdot h^2}{30000^2})) + 52 \cdot (1 - \exp(-\frac{\delta \cdot h^2}{120000^2})) \text{ pour } h > 0$$

$$\gamma^3(0) = 0$$

Soit V_1, V_2, \dots, V_{13} les variances d'erreur de mesure associées aux données expérimentales.

Le système de krigeage devient :

- avec le variogramme

$$\begin{bmatrix} -V_1 & \gamma^b_{12} & \dots & \gamma^b_{1i} & \dots & \gamma^b_{113} & 1 \\ \gamma^b_{21} & -V_2 & \dots & \gamma^b_{2i} & \dots & \gamma^b_{213} & 1 \\ \vdots & & & & & & \\ \gamma^b_{i1} & \gamma^b_{i2} & \dots & -V_i & \dots & \gamma^b_{i13} & 1 \\ \vdots & & & & & & \\ \gamma^b_{131} & \gamma^b_{132} & \dots & \gamma^b_{13i} & \dots & -V_{13} & 1 \\ 1 & 1 & \dots & 1 & \dots & 1 & 0 \end{bmatrix} \times \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_i \\ \vdots \\ \lambda_{13} \\ \mu' \end{bmatrix} = \begin{bmatrix} \gamma^b_{01} \\ \gamma^b_{02} \\ \vdots \\ \gamma^b_{0i} \\ \vdots \\ \gamma^b_{0n} \\ 1 \end{bmatrix} \quad (\text{Equ. 5})$$

$$\sigma_K^2 = 2 \sum_i \lambda_i^b \gamma^b_{oi} - \sum_{i,j} \lambda_i^b \lambda_j^b \gamma^b_{ij} + \sum_i (\lambda_i^b)^2 V_i$$

- avec la covariance

$$\begin{bmatrix} C^{b'}_{11} + V_1 & C^{b'}_{12} & \dots & C^{b'}_{1i} & \dots & C^{b'}_{113} & 1 \\ C^{b'}_{21} & C^{b'}_{22} + V_2 & \dots & C^{b'}_{2i} & \dots & C^{b'}_{213} & 1 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ C^{b'}_{i1} & C^{b'}_{i2} & \dots & C^{b'}_{ii} + V_i & \dots & C^{b'}_{i13} & 1 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ C^{b'}_{131} & C^{b'}_{132} & \dots & C^{b'}_{13i} & \dots & C^{b'}_{1313} + V_{13} & 1 \\ 1 & 1 & \dots & 1 & \dots & 1 & 0 \end{bmatrix} \times \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_i \\ \vdots \\ \lambda_{13} \\ \mu \end{bmatrix} = \begin{bmatrix} C^{b'}_{01} \\ C^{b'}_{02} \\ \vdots \\ C^{b'}_{0i} \\ \vdots \\ C^{b'}_{013} \\ \mu \end{bmatrix}$$

$$\sigma_K^2 = C^{b'}(0) - 2 \sum_i \lambda_i^b C^{b'}_{oi} + \sum_{i,j} \lambda_i^b \lambda_j^b C^{b'}_{ij} + \sum_i (\lambda_i^b)^2 V_i$$

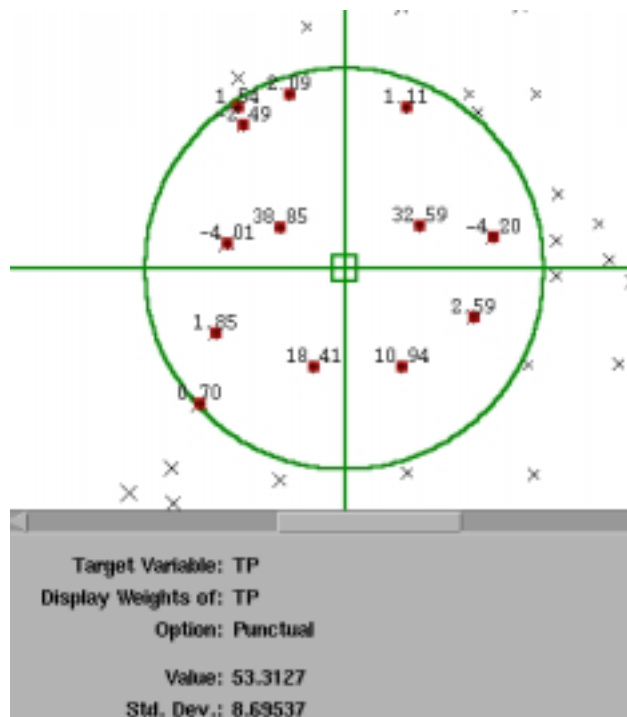


Figure 6 - Cas 3. Représentation du voisinage de krigeage et des poids de krigeage affectés aux 13 points expérimentaux situés dans ce voisinage.

Les poids de krigeage diffèrent des poids précédemment calculés. Aux quatre points les plus proches, où la variance de l'erreur de mesure est inférieure à la variance moyenne de $38 (\mu\text{g}/\text{m}^3)^2$, les poids de krigeage augmentent (Tableau 1).

Ils varient en sens inverse si, en ces mêmes points, on ajoute artificiellement $100 \mu\text{g}/\text{m}^3$ à la variance de l'erreur de mesure (Tableau 2). Cette dernière a donc un impact significatif sur la valeur des poids de krigeage. **Les concentrations assorties d'une plus forte variance de l'erreur de mesure contribuent avec un poids moindre à l'estimation.**

D'autre part, afin d'illustrer la remarque sur la variance de krigeage (cf. §1.2), le krigeage avec variance de l'erreur de mesure (VEM) est mis en œuvre avec une VEM égale à $38 (\mu\text{g}/\text{m}^3)^2$ en chaque point d'échantillonnage. Les poids de krigeage et la concentration estimée sont identiques à ceux qui résultent du deuxième modèle, mais la variance d'estimation associée au krigeage ponctuel diminue d'une valeur exactement égale à $\sigma_e^2=38 (\mu\text{g}/\text{m}^3)^2$. **Avec un krigeage de blocs, la variance d'estimation reste inchangée.**

Tableau 1 - Résultats de l'estimation ponctuelle par krigeage dans les trois cas présentés ci-dessus (krigeage au point x_0)

	Modèle 1	Modèle 2	Modèle 3	V_i	C_i
$\lambda_1 (.10^{-2})$	14.20	31.22	32.59	30.5	54.0
$\lambda_2 (.10^{-2})$	6.17	-4.01	-4.20	54.1	79.5
$\lambda_3 (.10^{-2})$	5.99	1.72	1.11	35.4	60.0
$\lambda_4 (.10^{-2})$	14.21	38.27	38.85	31.8	55.6
$\lambda_5 (.10^{-2})$	6.99	-5.03	-4.01	96.8	113.4
$\lambda_6 (.10^{-2})$	3.93	-2.61	-2.49	51.3	76.8
$\lambda_7 (.10^{-2})$	4.39	2.30	2.09	38.3	63.3
$\lambda_8 (.10^{-2})$	3.70	2.26	1.54	50.0	75.5
$\lambda_9 (.10^{-2})$	11.18	17.02	18.41	23.5	59.7
$\lambda_{10} (.10^{-2})$	6.09	2.99	1.85	54.1	79.5
$\lambda_{11} (.10^{-2})$	5.12	1.31	0.70	95.2	112.4
$\lambda_{12} (.10^{-2})$	9.95	10.81	10.94	26.6	48.9
$\lambda_{13} (.10^{-2})$	7.38	3.74	2.59	57.8	82.9
Z_0^*	69.22	53.83	53.31		
σ_{K0}^2	150.6	115.7	75.61		
σ_{K0}	12.27	10.76	8.70		

V_i : variance de l'erreur de mesure $(\mu\text{g}/\text{m}^3)^2$ C_i : concentration d'ozone mesurée $(\mu\text{g}/\text{m}^3)$

Tableau 2 - Evolution des poids de krigeage en fonction de la variance de l'erreur de mesure (krigeage au point x_0)

	Cas initial	$V_1 : 30,5 \rightarrow 130,5$	$V_1 : 30,5 \rightarrow 130,5$ $V_4 : 31,8 \rightarrow 131,8$	$V_1 : 30,5 \rightarrow 130,5$ $V_4 : 31,8 \rightarrow 131,8$ $V_9 : 23,51 \rightarrow 123,51$ $V_{12} : 26,6 \rightarrow 126,6$
$\lambda_1 (.10^{-2})$	32.59	18.15	19.28	19.99
$\lambda_2 (.10^{-2})$	-4.20	0.94	1.40	2.14
$\lambda_3 (.10^{-2})$	1.11	3.78	4.89	5.97
$\lambda_4 (.10^{-2})$	38.85	40.60	21.88	22.75
$\lambda_5 (.10^{-2})$	-4.01	-4.04	3.13	3.72
$\lambda_6 (.10^{-2})$	-2.49	-2.23	0.33	0.70
$\lambda_7 (.10^{-2})$	2.09	2.52	3.51	3.99
$\lambda_8 (.10^{-2})$	1.54	1.90	2.08	2.51
$\lambda_9 (.10^{-2})$	18.41	19.09	22.67	13.35
$\lambda_{10} (.10^{-2})$	1.85	2.21	2.72	5.38
$\lambda_{11} (.10^{-2})$	0.70	1.03	1.28	3.01
$\lambda_{12} (.10^{-2})$	10.94	12.10	12.13	9.56
$\lambda_{13} (.10^{-2})$	2.59	4.01	4.59	6.93
Z_0^*	53.31	55.63	61.16	63.95
σ_{K0}^2	75.61	81.53	89.98	94.46
σ_{K0}	8.70	9.03	9.49	9.72

2. Bilan. Comparaison entre les méthodes.

Pour chaque modèle (cas 1, cas 2 et cas 3), les concentrations d’ozone ont été estimées par krigeage ponctuel et par krigeage de blocs dans tout le domaine d’étude (maille de 5 km de côté). L’ensemble des données disponibles (sites ruraux, périurbains, urbains, littoraux) a été pris en compte dans l’estimation. Le voisinage de krigeage est un voisinage circulaire de 50 km de rayon.

Les cartes d’estimation et d’écart-type sont jointes à la fin de ce document.

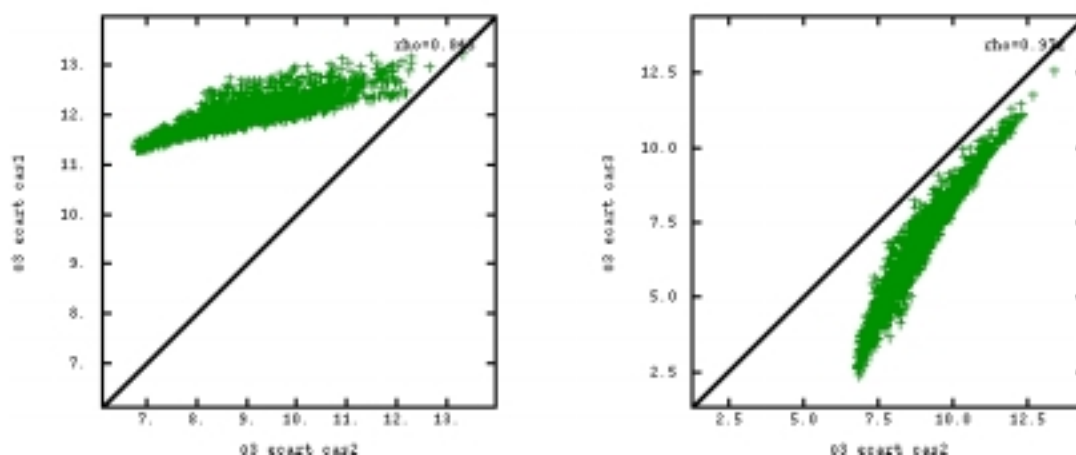


Figure 7 - Modèles 1, 2 et 3 : Comparaison entre écarts-types de krigeage. Krigeage ponctuel

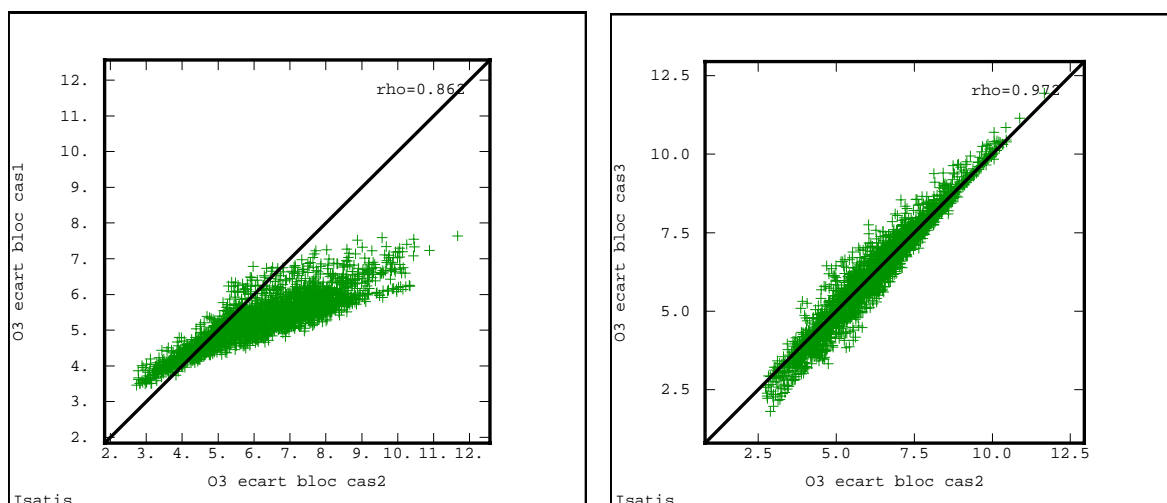


Figure 8 - Modèles 1, 2 et 3 : Comparaison entre écarts-types de krigeage. Krigeage de blocs

2.1 Comparaison entre les modèles 1 et 2

Krigeage ponctuel

Dans cet exemple, et pour une estimation ponctuelle, les écarts-types de krigeage diminuent sensiblement avec un effet de pépite réduit à 38 (cf. courbe de gauche, Figure 7).

Prendre un effet de pépite exactement égal à la variance moyenne de l’erreur de mesure est un choix délibéré qui doit reposer soit sur les informations contenues dans les

données, soit sur l'expérience de l'utilisateur. Il suppose en effet que l'erreur de mesure est l'unique source de variabilité aux très courtes distances. Cela signifie que la variabilité du polluant est parfaitement connue, ou bien que les structures à faible portée ont pu être identifiées grâce à un échantillonnage très resserré et modélisées en tant que telles.

Si le modèle ainsi ajusté s'accorde bien aux données expérimentales (à vérifier par la validation croisée), alors il peut conduire à une précision d'estimation accrue par rapport au premier modèle, et la carte d'estimation est beaucoup moins lissée.

Krigeage de blocs

Avec un krigeage de blocs et un effet de pépité plus faible, les écarts-types d'estimation diminuent pour les blocs qui contiennent des données et augmentent pour les blocs situés entre les points de mesure.

2.2 Comparaison entre les modèles 2 et 3

Krigeage ponctuel

Cette comparaison doit tenir compte du fait que dans le modèle 3, la variance de krigeage n'inclut pas la variance moyenne de l'erreur de mesure dans sa composante $C(0)$. Ainsi les valeurs de la variance de krigeage sont systématiquement plus faibles que dans le modèle 2, mais elles ne traduisent pas nécessairement une meilleure précision de l'estimation.

Krigeage de blocs

Avec un krigeage de blocs, le terme $C(v,v)$ de la variance de krigeage est identique avec ou sans VEM, et les écarts-types de krigeage sont très similaires.

L'avantage de la troisième méthode est que l'effet de pépité s'ajuste localement à la variance de l'erreur de mesure et que le poids de krigeage attribué à chaque donnée dépend de la VEM de cette donnée. Cet ajustement local n'a donc d'intérêt réel que si la VEM diffère significativement d'un point de mesure à un autre.

En outre, le krigeage avec VEM repose sur l'hypothèse restrictive que la concentration et l'erreur de mesure sont des variables indépendantes. Or cette hypothèse n'est pas toujours vérifiée. De fait, si l'erreur de mesure est inconnue, la VEM, en raison même de la méthode qui permet de la calculer, croît avec la concentration.. Aussi le krigeage avec VEM peut-il induire un lissage excessif de la carte, accordant un poids faible aux plus fortes concentrations (dotées d'une grande VEM) et un poids plus élevé aux valeurs les plus basses.

Ce lissage est visible sur les cartes d'estimation, dans la zone du littoral.

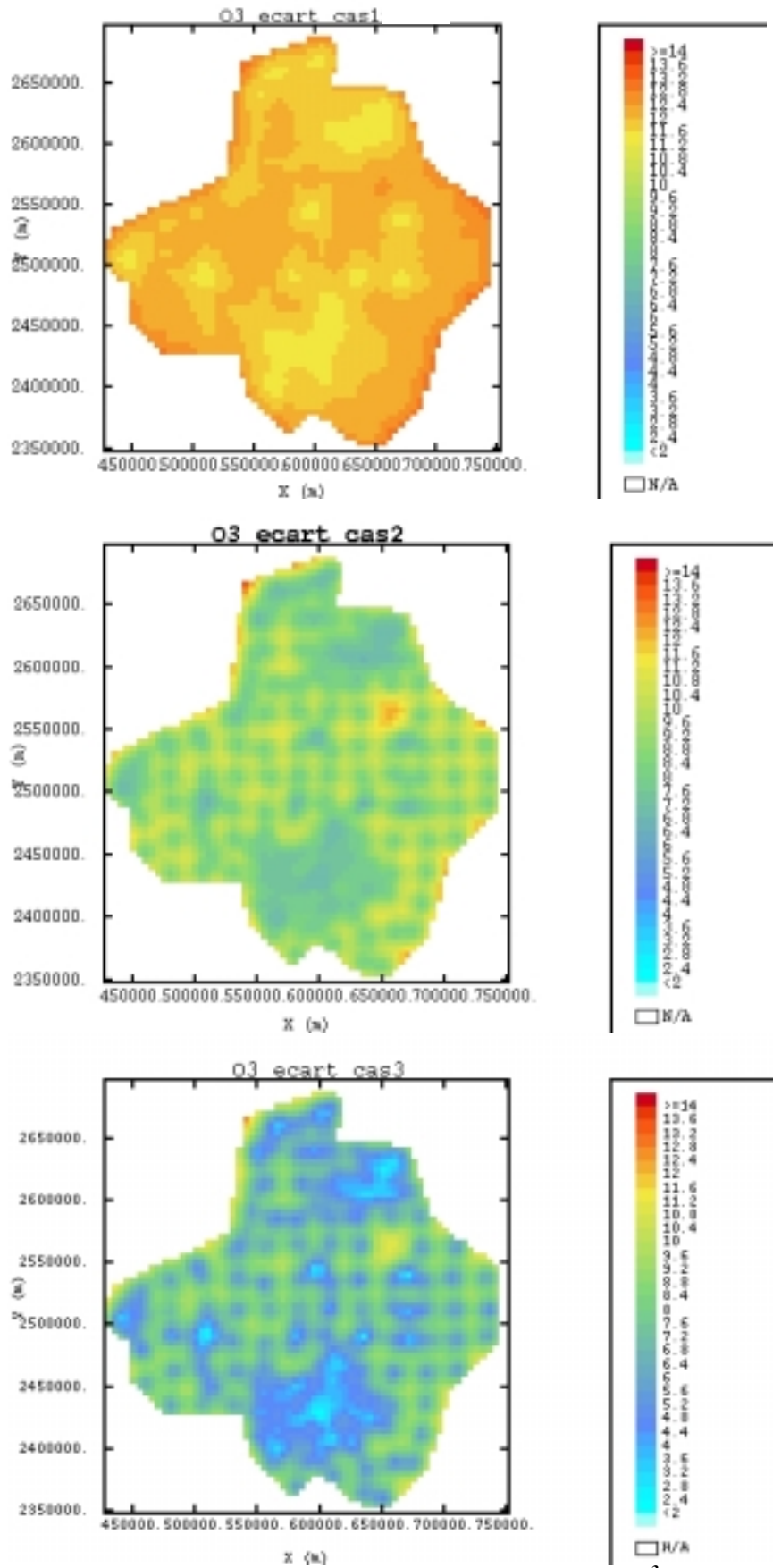


Figure 1 - Ecart-types de krigeage ($\mu\text{g}/\text{m}^3$). Krigeage ponctuel

Cas 1 (effet de pépite=120), cas 2 (effet de pépite=38), cas 3 (avec variance de l'erreur de mesure)

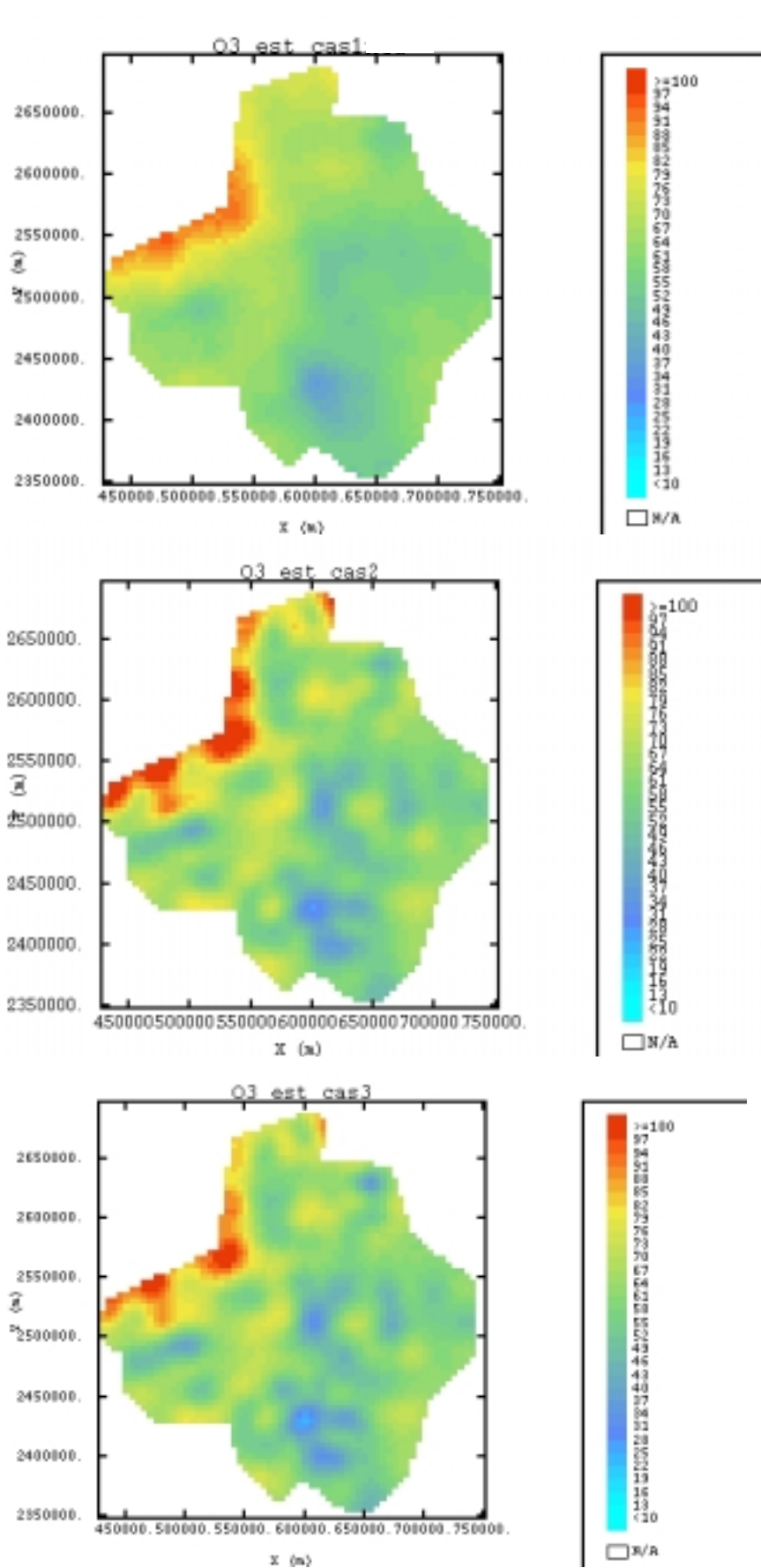


Figure 2 - Concentrations d'ozone ($\mu\text{g}/\text{m}^3$) estimées par krigeage ordinaire ponctuel
 Cas 1 (effet de pépite=120), cas 2 (effet de pépite=38), cas 3 (avec variance de l'erreur de mesure)

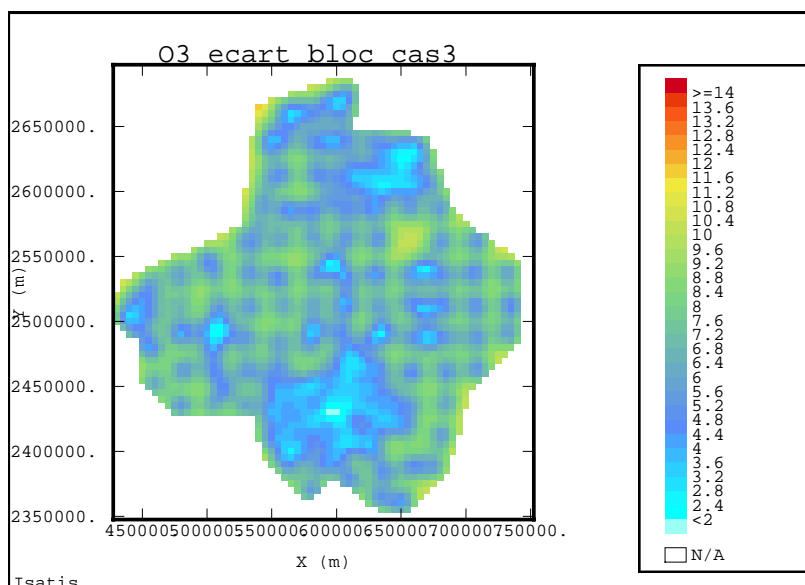
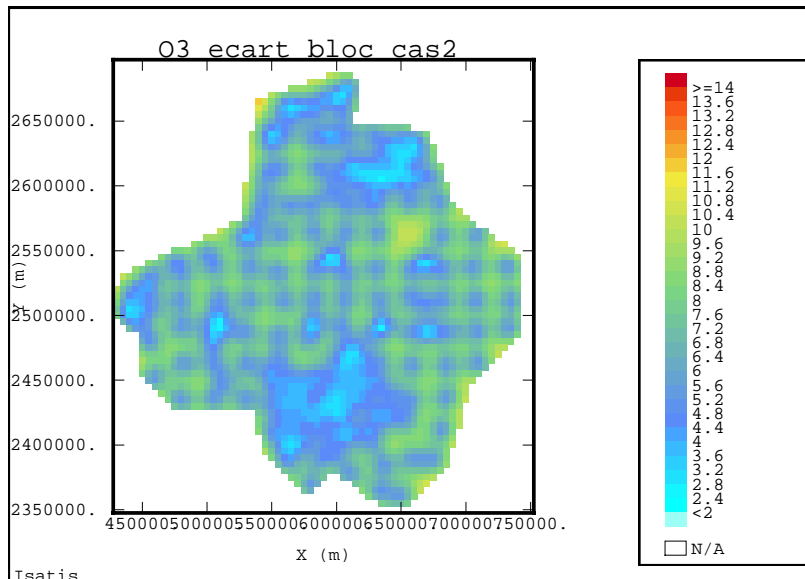
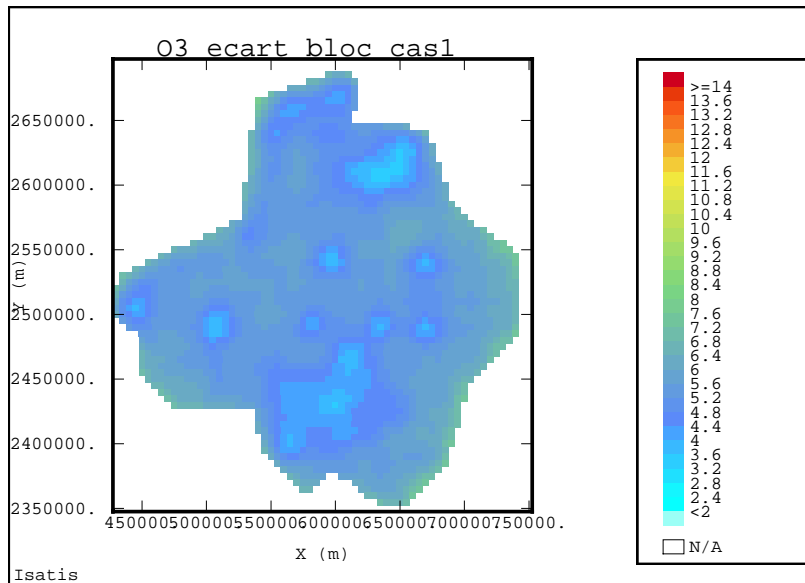


Figure 3 - Ecart-types de krigeage ($\mu\text{g}/\text{m}^3$). Krigeage de blocs

Cas 1 (effet de pépite=120,), cas 2 (effet de pépite=38), cas 3 (avec variance de l'erreur de mesure)

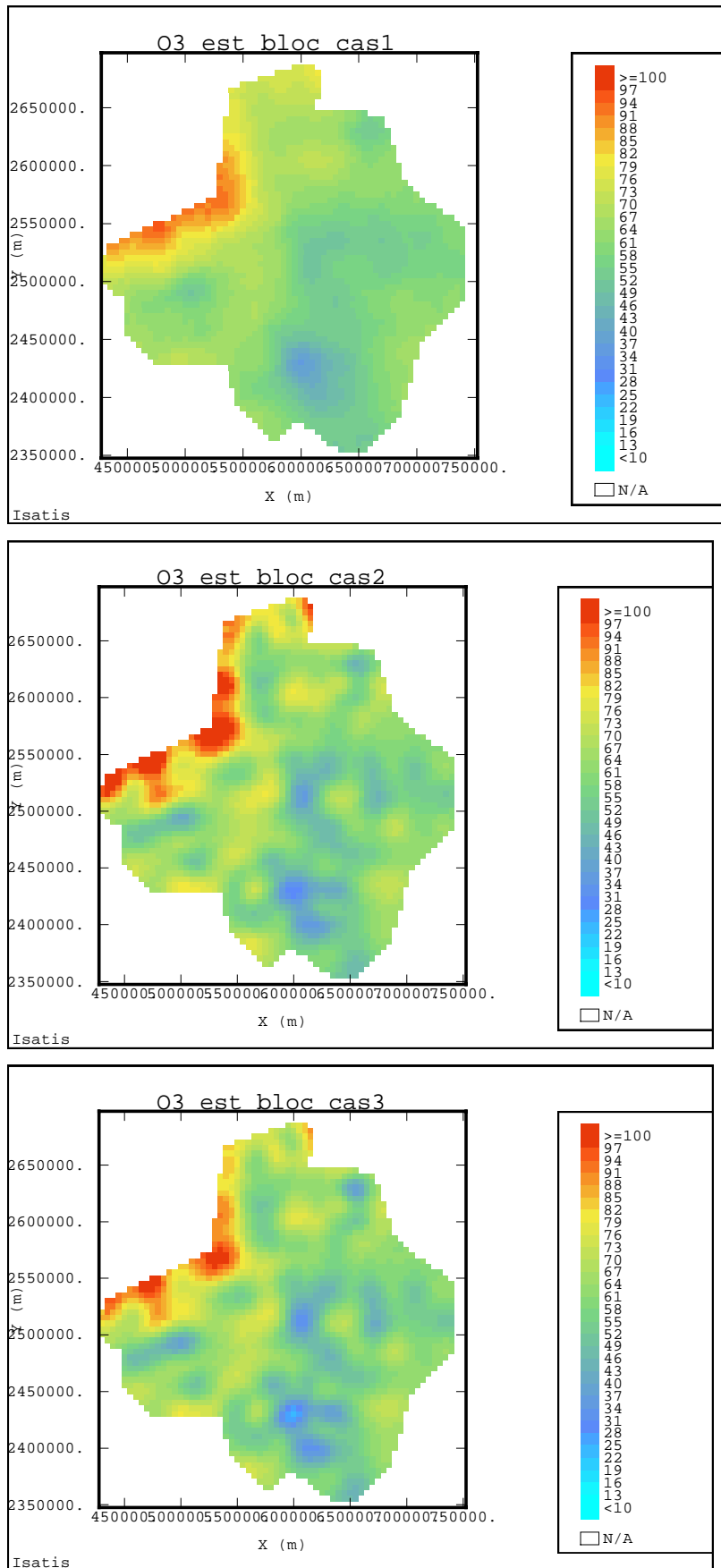


Figure 4 - Concentrations d'ozone ($\mu\text{g}/\text{m}^3$) estimées par krigeage ordinaire de blocs
 Cas 1 (effet de pépite=120), cas 2 (effet de pépite=38), cas 3 (avec variance de l'erreur de mesure)