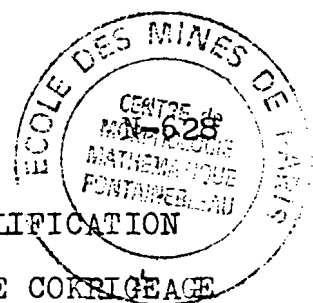


Fontainebleau/CG



RECHERCHE DE SIMPLIFICATION
DANS UN PROBLEME DE CORRIGEAGE

G. MATHERON

Novembre 1979

RECHERCHE DE SIMPLIFICATION DANS UN PROBLEME DE COKRIGEAGE

=====

Par

G. MATHERON

0 - INTRODUCTION.

L'objet de cette note est un problème très concret, qui se rencontre très fréquemment dans l'industrie pétrolière, celui de l'estimation simultanée des puissances d'un nombre assez grand d'horizons géologiques, numérotés de $i = 1$ à $i = N$. Ces puissances sont modélisées sous la forme de fonctions aléatoires $Z_i(x)$, x parcourant le plan horizontal. On s'intéresse également à la puissance totale :

$$Z(x) = \sum_{i=1}^N Z_i(x)$$

Les données expérimentales sont constituées des puissances $Z_{\alpha i} = Z_i(x_\alpha)$ observées en un certain nombre de points x_α ($\alpha = 1, 2, \dots, N$), et des puissances totales correspondantes $Z_\alpha = \sum_i Z_{\alpha i}$ (on n'envisage pas ici le cas où certaines données seraient manquantes, et on suppose toujours qu'en chaque point x_α toutes les puissances $Z_{\alpha i}$, $i = 1, \dots, N$ sont connues). A partir de ces données, on cherche à former des estimateurs linéaires $\hat{Z}_i(x)$ et $\hat{Z}(x)$ des puissances individuelles $Z_i(x)$ et de la puissance totale $Z(x)$ en un même point x . Pour des raisons évidentes de cohérence, la condition

$$(0-1) \quad \hat{Z}(x) = \sum_i \hat{Z}_i(x)$$

doit être impérativement vérifiée.

On sait que cette condition (0-1) est automatiquement vérifiée

si les estimateurs $\hat{Z}_i$ et $\hat{Z}$ sont formés par un cokrigeage complet sur les $Z_{\alpha j}$. Mais, le nombre N des horizons étant relativement élevé, le système des équations du cokrigeage est prohibitif en pratique. D'ailleurs, compte tenu des incertitudes qui entachent le plus souvent l'estimation des covariances croisées, cette méthode coûteuse, le cokrigeage, serait en grande partie illusoire. C'est pourquoi l'on doit s'orienter vers la recherche d'estimateurs beaucoup plus simples, assujettis pourtant à vérifier la condition de compatibilité (0-1).

1 - LA METHODE DE M. RIGUIDEL.

Voici, tout d'abord, une méthode simple et robuste, proposée par Michel Riguidel. De la covariance matricielle, on suppose connus les termes diagonaux $\sigma_{ii}(h)$, ainsi que la covariance

$$\sigma(h) = \sum_{i,j} \sigma_{ij}(h)$$

de la variable $Z(x) = \sum Z_i(x)$. Par contre, les termes rectangles $\sigma_{ij}(h)$ ($i \neq j$) restent inconnus (soit parce que leur estimation est entachée d'une trop grande incertitude, soit tout simplement parce que le nombre $N(N-1)$ de ces fonctions est trop élevé). Dans ces conditions, on peut kriger séparément chacune des variables Z_i et Z (avec les conditions d'universalité voulues). Soit Z_{Ki} le krigage propre de $Z_i(x)$ (à l'aide des seules variables $Z_{\alpha i}$ correspondant au même indice i), et soit de même Z_K le krigage propre de $Z(x)$ (à l'aide des seules variables Z_{α}). Ces estimateurs ne sont peut-être pas les meilleurs possibles (puisque'ils ne coïncident pas avec le cokrigeage), mais ce sont certainement de bons estimateurs, et faciles à former. Seulement, ils ne vérifient pas la condition de compatibilité (0-1). D'où l'idée d'imposer une correction simple, en prenant :

$$(1-1) \quad \hat{Z}_i = Z_{Ki} + a_i (Z_K - \sum_j Z_{Kj})$$

avec des coefficients a_i vérifiant la condition :

$$(1-2) \quad \sum a_i = 1$$

Malgré sa simplicité, l'expression (1-1) n'est nullement arbitraire. De fait, elle résulte obligatoirement des conditions d'universalité, dès lors que l'on choisit de s'imposer les deux conditions suivantes :

(i) - les estimateurs $\hat{Z}_i$ sont des combinaisons linéaires des Z_{Kj} et de Z_K

(ii) - $\sum \hat{Z}_i = Z_K$, autrement dit l'estimateur de la puissance totale $Z(x)$ doit coïncider avec le krigeage propre de celle-ci.

De fait, d'après (i), on doit prendre

$$\hat{Z}_i = B_i^j Z_{Kj} + a_i Z_K$$

D'après (ii), on doit de plus avoir

$$\sum a_i = 1 \quad \sum_i B_i^j = 0$$

Enfin, les conditions d'universalité imposent :

$$a_i + B_i^j = \delta_i^j$$

On a donc bien

$$\hat{Z}_i = Z_{Ki} + a_i (Z_K - \sum_j Z_{Kj})$$

$$\sum a_i = 1$$

De ces deux conditions, qui imposent la forme (1-1) de l'estimateur $\hat{Z}_i$, la première semble assez naturelle. En l'absence d'information sur les covariances croisées, il paraît, en effet, assez judicieux d'utiliser les krigeages propres comme point de départ, puisque ce sont les seules expressions possédant une certaine sorte

d'optimalité que nous soyons capables de former.

Par contre, la condition (ii) peut prêter à contestation. De fait, si les FA $Z_i(x)$ sont indépendantes (ou simplement orthogonales), les Z_{Ki} (krigeages propres des puissances Z_i) sont en fait des cokrigeages. Par suite, le cokrigeage de la puissance totale $Z(x) = \sum Z_i(x)$ est la somme $\sum Z_{Ki}$: le meilleur estimateur de cette puissance totale n'est donc pas le krigage propre Z_K , mais la somme $\sum Z_{Ki}$, qui n'a évidemment aucune raison de coïncider avec Z_K . Toutefois, en pratique, on ne sera pas le plus souvent en droit de déclarer négligeable l'influence des covariances croisées. Du fait que le nombre N des FA est élevé, la covariance $\sigma(h)$ de la puissance totale, qui est $\sum_{i,j} \sigma_{ij}(h)$ peut différer notablement de la somme $\sum_i \sigma_{ii}(h)$ des covariances propres, et cela même si chacun des termes rectangles $\sigma_{ij}(h)$, pris individuellement, n'est pas significativement différent de 0. Ainsi, et sauf peut-être dans le cas où l'on pourrait vérifier expérimentalement la relation approchée

$$\sigma(h) = \sum_i \sigma_{ii}(h)$$

il serait imprudent d'estimer la puissance totale par la somme $\sum Z_{Ki}$ des krigeages propres des puissances partielles.

Au total, donc, l'expression (1-1) fournit une solution acceptable. Il reste à choisir les coefficients correctifs a_i (assujettis à la relation (1-2)). On peut les déterminer en imposant, par exemple, la condition de minimiser la somme (pondérée) des variances d'estimation

$$\sum_i W_i \|Z_i - \hat{Z}_i\|^2$$

Toutefois, les poids W_i pouvant à leur tour être choisis arbitrairement, il est aussi simple de se donner directement les coefficients a_i . Nous verrons ci-dessous une méthode simple conduisant à un choix raisonnable de ces poids. Nous allons auparavant nous poser deux sortes de questions : tout d'abord, peut-on améliorer

l'estimation en remplaçant les Z_{Ki} par d'autres estimateurs Z_i^* (vérifiant les conditions d'universalité ? En second lieu, à quelle condition le cokrigeage de la puissance totale coïncide-t-il avec le krigeage propre Z_K ?

2 - CAS OU LES F.A. SONT ORTHOGONALES.

A la première de ces deux questions, on peut pragmatiquement répondre qu'on ne voit pas bien quelles autres combinaisons linéaires utiliser que les Z_{Ki} et Z_K . Toutefois, dans le cas où les covariances rectangles $\sigma_{ij}(h)$ ($i \neq j$) sont nulles, on peut donner une justification plus rigoureuse. Plus précisément, nous allons démontrer le théorème suivant :

Etant donnés des poids $w_i = 1/C_i$ (avec $C_i > 0$ et $\sum C_i = 1$), on considère l'ensemble des combinaisons linéaires du type :

$$Z_i^* = \sum_{\alpha} \lambda_{\alpha i}^{\alpha} Z_{\alpha i} \quad \text{avec} \quad \sum_{\alpha} \lambda_{\alpha i}^{\alpha} = 1$$

$$Z^* = \sum_{\alpha} \lambda^{\alpha} Z_{\alpha} \quad \text{avec} \quad \sum_{\alpha} \lambda^{\alpha} = 1$$

et des coefficients a_i tels que $\sum a_i = 1$. Alors, (les FA étant orthogonales) la forme quadratique

$$Q = \sum_i \frac{1}{C_i} \|Z_i - Z_i^* - a_i(Z^* - \sum_j Z_j^*)\|^2$$

est minimale lorsque $a_i = C_i$, $Z_i^* = Z_{Ki}$ et $Z^* = Z_K$.

Posons $\hat{Z}_i = Z_i^* - a_i(Z^* - \sum_j Z_j^*)$. Comme, dans ce modèle, les krigeages propres Z_{Ki} sont des cokrigeages, on a $\langle Z_i, \hat{Z}_i \rangle = \langle Z_{Ki}, \hat{Z}_i \rangle$ et on peut aussi, bien chercher le minimum de

$$Q' = \sum_i \frac{1}{C_i} \|Z_{Ki} - \hat{Z}_i\|^2$$

Equations en a_i . Prenons $Z_i^* = \delta_{Ki}$, $Z^* = Z_K$, $a_i = C_i$ et faisons varier les a_i , de sorte que $\delta \hat{Z}_i = \delta a_i (Z_K - \sum_j Z_{Kj})$. Alors :

$$\frac{1}{2} \delta Q' = \sum_i \frac{1}{C_i} \langle \delta \hat{Z}_i, \hat{Z}_i - Z_{Ki} \rangle = \sum_i \frac{a_i}{C_i} \langle \delta \hat{Z}_i, (Z_K - \sum_j Z_{Kj}^*) \rangle$$

Comme $a_i = C_i$, par hypothèse, il reste :

$$\frac{1}{2} \delta Q' = \sum_i \delta a_i \|Z_K - \sum_j Z_{Kj}\|^2 = 0$$

puisque $\sum \delta a_i = 0$: les équations en a_i sont donc vérifiées.

Equations en Z_{Ki} . Avec $a_i = C_i$, $Z^* = Z_K$, écrivons la variation de Q' lorsque $Z_i^* = Z_{Ki}$ est remplacée par $Z_{Ki} + \delta Z_i$ (où δZ_i est, pour chaque i , une combinaison linéaire autorisée des $Z_{\alpha i}$). Il vient :

$$\frac{1}{2} \delta Q' = \sum_i \frac{1}{C_i} \langle \delta \hat{Z}_i, \hat{Z}_i - Z_{Ki} \rangle = \sum_i \frac{a_i}{C_i} \langle \delta \hat{Z}_i, Z_K - \sum_j Z_{Kj} \rangle$$

Or $a_i = C_i$, par hypothèse, et

$$\delta \hat{Z}_i = \delta Z_i - a_i \sum_k \delta Z_k$$

Il vient donc :

$$\frac{1}{2} \delta Q' = \sum_i \langle \delta Z_i - a_i \sum_k \delta Z_k, Z_K - \sum_j Z_{Kj} \rangle$$

Compte tenu de $\sum a_i = 1$, on trouve bien $\delta Q' = 0$.

Equations en Z_K . Dans les mêmes conditions, écrivons la variation $\delta Q'$ lorsque Z^* passe de Z_K à $Z_K + \delta Z$, où δZ est une combinaison autorisée des Z_α . On a cette fois

$$\delta \hat{Z}_i = a_i \delta Z$$

et par suite (comme ci-dessus) :

$$\frac{1}{2} \delta Q' = \sum_i < \delta \hat{Z}_i, Z_K - \sum_j Z_{Kj} > = \sum_i a_i < \delta Z, Z_K - \sum_j Z_{Kj} >$$

Donc, comme $\sum a_i = 1$:

$$\frac{1}{2} \delta Q' = < \delta Z, Z_K - \sum_j Z_{Kj} >$$

Or, par définition du krigeage :

$$< Z_K, \delta Z > = < Z, \delta Z >$$

puisque Z est autorisée. D'autre part, δZ étant autorisée, et Z_{Kj} étant le cokrigeage de Z_j , on a

$$< \delta Z, \sum_j Z_{Kj} > = < \delta Z, \sum_j Z_j > = < \delta Z, Z >$$

On a donc bien $\delta Q' = 0$, ce qui achève la démonstration. (On remarque que l'orthogonalité intervient de manière essentielle dans cette démonstration, et que le théorème ne subsistera donc pas dans le cas général où les $Z_i(x)$ ne seraient pas orthogonales.)

3 - MODELE A PUISSANCE TOTALE AUTOKRIGEABLE.

Convenons de dire qu'une FA est autokrigeable si son krigeage propre coïncide avec son cokrigeage. Il reste à examiner sous quelles conditions la puissance totale Z est autokrigeable.

Plus généralement, donnons-nous une combinaison linéaire $Z(x) = \sum_j S_j^j Z_j(x)$ (il suffira de faire $S_j^j = 1$ pour que $Z(x)$ soit la puissance totale), et désignons par

$$\sigma(h) = S^i S^j \sigma_{ij}(h)$$

la covariance propre de $Z(x)$. Dans ces conditions : $Z(x)$ est auto-krigeable (pour toutes les configurations possibles de krigeage) si et seulement si :

$$(3-1) \quad S^j \sigma_{ij}(h) = \omega_i \sigma(h)$$

Cette condition est suffisante : si elle est vérifiée, on a $\langle L, Z_{\alpha i} \rangle = \omega_i \langle L, Z_{\alpha} \rangle$ pour toute combinaison linéaire L de la FA $Z(x)$. En particulier, Z_K désignant le krigeage propre de $Z(x)$, on a :

$$\langle Z(x) - Z_K, Z_{\alpha i} \rangle = \omega_i \langle Z(x) - Z_K, Z_{\alpha} \rangle$$

Par suite $Z(x) - Z_K$ est orthogonale à toutes les combinaisons autorisées $\sum_{\alpha} \lambda^{\alpha} Z_{\alpha i}$: ce qui entraîne que Z_K est le cokrigeage.

La réciproque est immédiate en l'absence de conditions d'universalité. Dans ce cas, en effet, (les variables étant supposées centrées) le krigeage de $Z(x)$ à partir d'un seul point $x_0 = 0$ est $Z_K = (\sigma(x)/\sigma(0)) Z(0)$. Si $Z(x)$ est autokrigeable, on a donc

$$\langle Z(x), Z_i(0) \rangle = \frac{\sigma(x)}{\sigma(0)} \langle Z(x), Z_i(0) \rangle$$

c'est-à-dire la relation (3-1), avec

$$\omega_i = \frac{\langle Z(0), Z_i(0) \rangle}{\sigma(0)} = \frac{S^j \sigma_{ij}(0)}{\sigma(0)}$$

Dans le cas où il y a des conditions d'universalité, la démonstration de la réciproque est un peu plus complexe, et nécessite l'hypothèse que le modèle vérifie la propriété ergodique :

$$\frac{1}{V} \int_V \sigma_{ij}(x) dx \rightarrow 0 \quad \text{lorsque volume } V \rightarrow \infty$$

En l'absence de cette propriété ergodique (vérifiée par tous les modèles usuels) on trouve seulement que la relation (3-1) est satisfaite à une constante près (dans le cas de la condition $\sum \lambda^{\alpha} = 0$)

ou à un polynome près de degré k (si λ^α doit annuler les polynômes de degré k).

Interprétation de la relation (3-1).

Interprétons (3-1) en raisonnant sur les variables centrées $Z'_i(x) = Z_i(x) - m_i(x)$, $Z'(x) = Z(x) - m(x)$. Cette relation s'écrit :

$$\langle Z'_i(x) Z'(x+h) \rangle = \omega_i \langle Z'(x) Z'(x+h) \rangle$$

Elle signifie donc que $Z'_i(x) - \omega_i Z'(x)$ est orthogonale à la FA $Z'(y)$. Autrement dit, posant :

$$Y_i(x) = Z'_i(x) - \omega_i Z'(x)$$

la formule

$$(3-2) \quad Z_i(x) = m_i + Y_i(x) + \omega_i (Z(x) - m)$$

décompose la FA $Z_i(x)$ en deux FA orthogonales, $Y_i(x)$ et $\omega_i (Z(x) - m)$ dont la première vérifie l'identité

$$(3-3) \quad S^1 Y_i(x) = 0$$

les coefficients ω_i , de leur côté, vérifiant :

$$(3-4) \quad S^1 \cdot \omega_i = 1$$

Exemple simple - Désignons par π_i la puissance relative de la couche i , soit

$$\pi_i(x) = \frac{Z_i(x)}{Z(x)} = \frac{Z_i(x)}{S^j Z_j(x)}$$

et considérons un modèle dans lequel la FA $Z(x)$ est indépendante des FA $\pi_i(y)$. Ainsi $Z_i(x)$ est le produit

$$Z_i(x) = \pi_i(x) Z(x)$$

de deux FA indépendantes. Prenant maintenant

$$\omega_i = E \Pi_i, \quad m_i = E Z_i, \quad m = E Z = S^i m_i$$

(donc, en particulier, $m_i = \omega_i m$), nous pouvons écrire

$$(3-5) \quad Z_i(x) = \omega_i Z(x) + [\Pi_i(x) - \omega_i] Z(x)$$

Comme, par hypothèse, Π_i est indépendante de $Z(x)$, on a

$$\langle Z(y), (\Pi_i(x) - \omega_i) Z(x) \rangle = \langle Z(y), Z(x) \rangle E[\Pi_i(x) - \omega_i] = 0$$

et la formule (3-5) décompose, comme en (3-2), la FA $Z_i(x)$ en deux FA orthogonales; soit (puisque ici $m_i = \omega_i m$)

$$Z_i(x) = \omega_i Z(x) + Y_i(x)$$

avec

$$Y_i(x) = (\Pi_i(x) - \omega_i) Z(x)$$

La relation $S^i Y_i(x) = 0$ est évidemment vérifiée, puisque, par construction, $S^i \Pi_i(x) = S^i \omega_i = 1$.

Intérêt du modèle (3-1).

Dans le cadre de ce modèle, la condition (ii) du paragraphe précédent (l'estimation de la puissance totale doit être le krigeage propre Z_K) est absolument justifiée, puisque, par construction, ce krigeage propre Z_K coïncide, dans ce modèle, avec le cokrigeage.

De plus, la question du choix des poids a_i intervenant dans la définition (1-1) des estimateurs $\hat{Z}_i$ se trouve automatiquement réglée.

De fait, donnons-nous N combinaisons linéaires

$$Z_i^* = \sum_{\alpha} \lambda_{\alpha}^i Z_{\alpha i} \quad ; \quad \sum \lambda_{\alpha}^i = 1$$

et cherchons le meilleur estimateur de Z_i parmi les combinaisons linéaires de Z_K et des Z_i^* vérifiant les conditions d'universalité. Autrement dit, $\hat{Z}_i$ doit être de la forme

$$\hat{Z}_i = Z_i^* + a_i(Z_K - S^j Z_j^*)$$

et a_i est déterminé par la condition :

$$(3-6) \quad \|Z_i - Z_i^* - a_i(Z_K - S^j Z_j^*)\|^2 \text{ minimum}$$

Bien que ce critère n'impose a priori aucune condition sur les a_i , la condition $S^i a_i = 1$ va se trouver automatiquement vérifiée. De fait, chaque $\hat{Z}_i$ étant la projection de Z_i sur la variété engendrée par les Z_j^* et Z_K , $S^i \hat{Z}_i$ est automatiquement la projection de $Z = S^i Z_i$ sur la même variété, et coïncide donc nécessairement avec le cokrigage Z_K (puisque celui-ci se trouve appartenir à cette variété).

Analytiquement, on doit minimiser en a_i l'expression

$$\|Z_i - Z_i^*\|^2 - 2 a_i \langle Z_i - Z_i^*, Z_K - S^j Z_j^* \rangle + a_i^2 \|Z_K - S^j Z_j^*\|^2$$

et on trouve donc

$$a_i \|Z_K - S^j Z_j^*\|^2 = \langle Z_i - Z_i^*, Z_K - S^j Z_j^* \rangle$$

En sommant en i , il vient ainsi :

$$\|Z_K - S^j Z_j^*\|^2 S^i a_i = \langle Z - S^j Z_j^*, Z_K - S^j Z_j^* \rangle$$

Mais $Z_K - S^j Z_j^*$ est une combinaison linéaire autorisée. On a donc

$$\langle Z, Z_K - S^j Z_j^* \rangle = \langle Z_K, Z_K - S^j Z_j^* \rangle$$

puisque Z_K est le cokrigage de Z . Il vient ainsi

$$\|Z_K - S^j Z_j^*\|^2 S^i a_i = \langle Z_K - S^j Z_j^*, Z_K - S^j Z_j^* \rangle$$

c'est-à-dire

$$S^i a_i = 1$$

Evaluation des poids a_i .

Les poids a_i , dans ce modèle, sont, comme on l'a vu, donnés par

$$(3-7) \quad a_i = \frac{\langle Z_i - Z_i^*, Z_K - S^j Z_j^* \rangle}{\|Z_K - S^j Z_j^*\|^2}$$

Pour les calculer effectivement, il faudrait connaître explicitement toutes les covariances $\sigma_{ij}(h)$, ce qui n'est pas le cas en général. Toutefois, ces coefficients a_i représentant des termes correctifs, il n'est pas nécessaire de les connaître avec une grande précision. On peut donc se contenter de les évaluer dans le cadre d'un modèle grossièrement approché, mais simple.

Choisissons, pour cette évaluation grossière, le modèle où les $Z_i(x)$ seraient indépendantes et admettraient, à un facteur près, la même covariance $\sigma(h)$ que $Z(x)$, soit

$$\sigma_{ij}(h) = \frac{\sigma_i^2}{\sigma(0)} \delta_{ij} \sigma(h)$$

où σ_i^2 est la variance de Z_i . Dans ce modèle, le cokrigage de Z_i serait $\lambda^\alpha Z_{\alpha i}$, et celui de Z serait $\lambda^\alpha Z_\alpha = Z_K$ (avec les mêmes coefficients λ^α). On trouverait alors :

$$\begin{aligned} \langle Z_i - Z_i^*, Z_K - S^j Z_j^* \rangle &= \langle Z_i - Z_i^*, (\lambda^\alpha - \lambda_i^\alpha) Z_{\alpha i} \rangle \\ &= \langle (\lambda^\beta - \lambda_i^\beta) Z_{\beta i}, (\lambda^\alpha - \lambda_i^\alpha) Z_{\alpha i} \rangle \\ &= \frac{\sigma_i^2}{\sigma(0)} (\lambda^\beta - \lambda_i^\beta)(\lambda^\alpha - \lambda_i^\alpha) \sigma_{\alpha\beta} \end{aligned}$$

Cette expression représente (dans ce modèle grossier) la variance de $(\lambda^\alpha - \lambda_i^\alpha) Z_{\alpha i}$: or cette quantité est la différence de deux

estimations de la puissance Z_i , l'une $\lambda^\alpha Z_{\alpha i}$ calculée avec les poids λ^α du krigeage de la puissance totale, l'autre étant l'estimation $Z_i^* = \lambda_i^\alpha Z_{\alpha i}$ choisie par nous. Ainsi nous sommes conduits à prendre :

$$(3-8) \quad a_i = \frac{\text{Var} (\lambda^\alpha Z_{\alpha i} - Z_i^*)}{\sum_j \text{Var} (\lambda^\alpha Z_{\alpha j} - Z_j^*)}$$

Il est d'ailleurs préférable d'évaluer $\text{Var} (\lambda^\alpha Z_{\alpha i} - Z_i^*)$ à l'aide de la vraie covariance propre $\sigma_{ii}(h)$ (plutôt qu'à l'aide du modèle grossièrement approché ci-dessus).

En pratique, Z_i^* est imposée : c'est le krigeage propre Z_{Ki} de Z_i (avec des poids λ_{Ki}^α calculées, évidemment, à partir de la vraie covariance propre $\sigma_{ii}(h)$). De même, désignons par

$$Z_i' = \lambda^\alpha Z_{\alpha i}$$

l'estimation obtenue en utilisant les poids λ^α du krigeage Z_K de la puissance totale. ($Z_K = \lambda^\alpha Z_\alpha = \sum Z_i'$).

Dans ces conditions, les poids a_i seront :

$$(3-9) \quad a_i = \frac{\text{Var} (Z_{Ki} - Z_i')}{\sum_j \text{Var} (Z_{Kj} - Z_j')} = \frac{(\lambda_i^\alpha - \lambda^\alpha) \sigma_{ii, \alpha\beta} (\lambda_i^\beta - \lambda^\beta)}{\sum_j (\lambda_j^\alpha - \lambda^\alpha) \sigma_{jj, \alpha\beta} (\lambda_j^\beta - \lambda^\beta)}$$

Ainsi les poids a_i seront proportionnels non pas aux variances propres ou aux variances de krigeage propres, mais à la variance des différences $(Z_{Ki} - Z_i')$ entre le krigeage propre et le pseudo-krigeage Z_i' obtenu avec les poids λ^α du krigeage de la puissance totale.

Il y a une certaine logique dans cette formule (3-9) : de fait, le but de la correction associée aux poids a_i est d'amener la somme $\sum Z_i^*$ à coïncider avec la somme $Z_K = \sum Z_i'$. Si pour un indice i on a déjà Z_{Ki} égal (ou presque égal) à Z_i' , il n'y a aucune raison de lui imposer une correction (ou celle-ci doit être très faible). Tel est exactement le sens de la formule (3-9).

4 - AUTRES MODELES.

Le modèle le plus simple est évidemment le modèle à corrélation intrinsèque, soit :

$$\sigma_{ij}(h) = \frac{\sigma_{ij}(0)}{\sigma(0)} \sigma(h)$$

dans lequel tous les krigages propres sont des cokrigages et se calculent avec les mêmes poids λ^α . Toutefois, ce modèle intrinsèque est inadmissible lorsque les covariances propres $\sigma_{ii}(h)$, seules connues, ne sont pas proportionnelles à $\sigma(h)$.

On aura un modèle un peu plus général (et pouvant être ajusté de manière à être compatible avec les covariances propres données) en prenant

$$(4-1) \quad \sigma_{ij}(h) = \varpi_i \varpi_j \sigma(h) + K_{ij} R(h)$$

avec les conditions :

$$(4-2) \quad S^i \varpi_i = 1, \quad S^j K_{ij} = 0$$

Ce modèle vérifie la condition (3-1). Les FA correspondantes sont donc de la forme :

$$Z_i(x) = Y_i(x) + \varpi_i Z(x)$$

avec $Z(x) = S^i Z_i(x)$ et $S^i Y_i(x) = 0$, les FA $Y_i(x)$ étant orthogonales à $Z(x)$. $\sigma(h)$ est la covariance de $Z(x)$, tandis que $K_{ij} R(h)$ est la covariance rectangle de $Y_i(x)$ et $Y_j(x+h)$. Il s'agit en somme du modèle (3-2) avec la condition supplémentaire que les $Y_i(x) = Z_i(x) - \varpi_i Z(x)$ sont en corrélation intrinsèque.

Dans un modèle de ce type, le cokrigage est très facile à réaliser. A cause de l'orthogonalité, en effet, le cokrigage de $Z_i(x)$ est la somme du cokrigage $\varpi_i Z_K$ de $\varpi_i Z(x)$ (qui coïncide

avec le krigeage propre) et du cokrigeage de $Y_i(x)$ à partir des $Y_{i\alpha} = Z_{\alpha i} - \omega_i Z_\alpha$. Mais, les $Y_i(x)$ étant en corrélation intrinsèque, ce cokrigeage est lui-même un krigeage propre, à savoir

$$Y_{Ki} = b^\alpha Y_{\alpha i} = b^\alpha (Z_{\alpha i} - \omega_i Z_\alpha)$$

avec des coefficients b^α (indépendants de l'indice i) donnés par le système habituel :

$$\begin{cases} b^\alpha R_{\alpha\beta} = R_{\beta x} + \mu \\ \sum b^\alpha = 1 \end{cases}$$

Au total, le cokrigeage de $Z_i(x)$ est de la forme

$$Z_i^* = \omega_i Z_K + b^\alpha (Z_{\alpha i} - \omega_i Z_\alpha)$$

soit

$$(4-3) \quad Z_i^* = \omega_i (\lambda^\alpha - b^\alpha) Z_\alpha + b^\alpha Z_{\alpha i}$$

La condition de compatibilité $S^i Z_i^* = Z_K$ est réalisée automatiquement, comme on le vérifie facilement : cela était du reste prévisible, puisque Z_K et les Z_i^* sont (dans ce modèle) des cokrigeages.

L'estimateur (4-3) est le meilleur possible si le modèle (4-1) est bon. Mais, en pratique, on n'est pas à même d'en juger, faute de connaître les covariances rectangles $\sigma_{ij}(h)$. Mais on peut formuler le problème de manière différente : si l'on décide, a priori, de chercher l'estimateur $\hat{Z}_i$ parmi les formes linéaires du type (4-3), avec $\sum \lambda^\alpha = \sum b^\alpha = 1$, on est conduit à un problème en un certain sens plus simple que celui que pose l'estimateur de Riguidel. Distinguons deux cas, selon que la relation (3-1) est (ou non) vérifiée.

Premier cas : relations (3-1) supposées vérifiées.

Ici, on doit prendre pour λ^α les coefficients du krigeage (= cokrigeage) de $Z(x)$, puisque Z_K est le meilleur estimateur possible

de la puissance totale. Il reste à choisir un jeu de coefficients b^α (tels que $\sum b^\alpha = 1$).

Or, dans le modèle (3-1), la variance d'estimation associée à l'estimateur (4-3) peut être calculée. De fait :

$$\|Z_i - Z_i^*\|^2 = \|Z_i - b^\alpha Z_{\alpha i}\|^2 - 2 \varpi_i \langle Z_i - b^\alpha Z_{\alpha i}, (\lambda^\alpha - b^\alpha) Z_\alpha \rangle + \varpi_i^2 \|(\lambda^\alpha - b^\alpha) Z_\alpha\|^2$$

Or, d'après (3-1), on trouve :

$$\langle Z_i - b^\alpha Z_{\alpha i}, (\lambda^\alpha - b^\alpha) Z_\alpha \rangle = \varpi_i \langle Z - b^\alpha Z_\alpha, (\lambda^\alpha - b^\alpha) Z_\alpha \rangle$$

De plus, $Z_K = \lambda^\alpha Z_\alpha$ étant le krigeage de Z , $Z - Z_K$ est orthogonale à la combinaison linéaire autorisée $(\lambda^\alpha - b^\alpha) Z_\alpha$. Par suite :

$$\langle Z_i - b^\alpha Z_{\alpha i}, (\lambda^\alpha - b^\alpha) Z_\alpha \rangle = \varpi_i \|(\lambda^\alpha - b^\alpha) Z_\alpha\|^2$$

et il vient

$$\|Z_i - Z_i^*\|^2 = \|Z_i - b^\alpha Z_{\alpha i}\|^2 - \varpi_i^2 \|(\lambda^\alpha - b^\alpha) Z_\alpha\|^2$$

Explicitement, cela donne :

$$(4-4) \quad \|Z_i - Z_i^*\|^2 = \|Z_i\|^2 - 2 b^\alpha \langle Z_i, Z_{\alpha i} \rangle + b^\alpha b^\beta \langle Z_{\alpha i}, Z_{\beta i} \rangle - \varpi_i^2 (\lambda^\alpha - b^\alpha)(\lambda^\beta - b^\beta) \sigma_{\alpha\beta}$$

Cette expression (4-4) peut donc se calculer à partir de la covariance propre $\sigma_{ii}(h)$, de la covariance $\sigma(h)$ de $Z(x)$ et des coefficients ϖ_i .

Pour choisir les b^α , donnons-nous des poids $W_i = 1/C_i$ et cherchons le minimum en b de l'expression

$$\sum \frac{1}{C_i} \|Z_i - Z_i^*\|^2 \quad \text{sous la condition} \quad \sum b^\alpha = 1$$

On trouve le système suivant

$$\sum \frac{1}{C_i} [b^\beta < z_{\alpha i}, z_{\beta i} > - < z_i, z_{\alpha i} > - \omega_i^2 (b^\beta - \lambda^\beta) \sigma_{\alpha\beta}] = \mu$$

Comme $\lambda^\beta \sigma_{\alpha\beta} = \sigma_{x\alpha}$ à une constante près, ce système est de la forme

$$b^\beta K_{\alpha\beta} = K_{\beta,x} + \mu$$

$$\sum b^\beta = 1$$

avec une fonction de covariance $K(h)$ égale à la moyenne pondérée des covariances propres $\sigma_{ii}(h)$ corrigées du terme en $\sigma(h)$:

$$(4-5) \quad K(h) = \sum_i \frac{1}{C_i} (\sigma_{ii}(h) - \omega_i^2 \sigma(h))$$

On note que, le cas échéant, cette formule (4-5) résout aussi le problème de l'estimation du terme $R(h)$ figurant dans le modèle (4-1) : comme $R(h)$ n'est défini qu'à un facteur près, on peut s'imposer $R(0) = 1$ et prendre $R(h) = K(h)/K(0)$.

Il reste à choisir les poids C_i . On peut les prendre arbitrairement, on peut aussi s'inspirer de ce qui suit.

Deuxième cas : les relations (3-1) ne sont pas supposées vérifiées.

Le calcul de $\|Z_i - Z_i^*\|^2$ nécessitant les covariances rectangles $\sigma_{ij}(h)$, plaçons-nous dans le cas où les FA $Z_i(x)$ sont orthogonales ($\sigma_{ij}(h) = \delta_{ij}, \sigma_{ii}(h)$). Dans ce cas simple, on a

$$\begin{aligned} \|Z_i - Z_i^*\|^2 &= \|Z_i - b^\alpha z_{\alpha i}\|^2 - 2 \omega_i < Z_i - b^\alpha z_{\alpha i}, (\lambda^\beta - b^\beta) z_{\beta i} > \\ &\quad + \omega_i^2 \sum_j \|(\lambda^\alpha - b^\alpha) z_{\alpha j}\|^2 \end{aligned}$$

et il faut minimiser la somme

$$\sum \frac{1}{C_i} \|Z_i - Z_i^*\|^2$$

Il serait intéressant que, dans cette somme, le terme rectangle soit égal au terme en ω_i^2 . Or cela est réalisé automatiquement si l'on fait le choix

$$C_i = \omega_i$$

car alors

$$\begin{aligned} \sum_i \frac{\omega_i}{C_i} \langle Z_i - b^\alpha Z_{\alpha i}, (\lambda^\beta - b^\beta) Z_{\beta i} \rangle &= \sum_i \langle Z_i - b^\alpha Z_{\alpha i}, (\lambda^\beta - b^\beta) Z_{\beta i} \rangle \\ &= \sum_i \langle Z_i - b^\alpha Z_{\alpha i}, (\lambda^\beta - b^\beta) Z_\beta \rangle \end{aligned}$$

(à cause de l'orthogonalité), donc encore :

$$= \langle Z - b^\alpha Z_\alpha, (\lambda^\beta - b^\beta) Z_\beta \rangle$$

Mais, dans cette expression, Z peut être remplacé par son krigeage $Z_K = \lambda^\alpha Z_\alpha$ (puisque $(\lambda^\beta - b^\beta) Z_\beta$ est une combinaison autorisée). Notre terme rectangle se réduit donc bien à

$$\|(\lambda^\alpha - b^\alpha) Z_\alpha\|^2 = \sum_i \|(\lambda^\alpha - b^\alpha) Z_{\alpha i}\|^2$$

Comme, d'autre part, $\sum \frac{\omega_i}{C_i} = \sum \omega_i = 1$, il reste simplement (avec $C_i = \omega_i$) :

$$\sum \frac{1}{\omega_i} \|Z_i - Z_i^*\|^2 = \sum \frac{1}{\omega_i} \|Z_i - b^\alpha Z_{\alpha i}\|^2 - \|(\lambda^\alpha - b^\alpha) Z_\alpha\|^2$$

En cherchant le minimum en b (sous la condition $\sum b^\alpha = 1$), on est, comme on le voit facilement, conduit au système :

$$b^\alpha K_{\alpha\beta} = K_{\beta x} + \mu$$

$$\sum b^\alpha = 1$$

avec la covariance $K(h)$ définie par

$$(4-6) \quad K(h) = \sum_i \frac{1}{\omega_i} \sigma_{ii}(h) - \sigma(h)$$

CONCLUSION :

Cette covariance est identique à celle de la relation (4-5), pourvu seulement que l'on choisisse $C_i = \omega_i$. Ceci suggère donc que le procédé doit être relativement robuste. Au total la démarche serait la suivante (et pourrait être appliquée sans hypothèse particulière sur le modèle).

Prerequisites : les covariances propres $\sigma_{ii}(h)$ (ou simplement les variogrammes correspondants), et la covariance $\sigma(h)$ de la puissance totale.

Il faut aussi connaître les ω_i : il faut, pour cela, déterminer expérimentalement la matrice $\sigma_{ij}(0)$ des covariances de $Z_i(x)$ et $Z_j(x)$ au même point x . On aura alors :

$$\omega_i = \frac{\sum_j \sigma_{ij}(0)}{\sigma(0)}$$

(vérifier que les ω_i sont ≥ 0).

Algorithmes : Calculer les poids λ^α et b^α par les systèmes habituels

$$\lambda^\alpha \sigma_{\alpha\beta} = \mu + \sigma_{\beta x}$$

$$b^\alpha K_{\alpha\beta} = \mu' + K_{\beta x}$$

$$\sum \lambda^\alpha = 1$$

$$\sum b^\alpha = 1$$

($K(h)$ donné par (4-6) : il faut donc avoir $\omega_i \geq 0$, et s'assurer que $K(h)$ est de type positif).

Estimateur
$$\hat{Z}_i = \omega_i (\lambda^\alpha - b^\alpha) Z_\alpha + b^\alpha Z_{\alpha i}$$

Ce procédé est plus simple (mais peut-être moins robuste) que celui de Rignidél, puisqu'il ne nécessite que deux krigeages (au lieu de $N+1$). Il sera meilleur, si les covariances réelles ne diffèrent pas trop du modèle (4-1), moins bon sans doute dans d'autres cas. Tout ceci mériterait peut-être d'être testé sur des exemples expérimentaux.