

Paradoxe de l'estimation des ressources

Approche distributionnelle : Connaissant le variogramme des données ponctuelles, l'on peut calculer la variance de bloc. On peut aussi tenir compte de l'effet information et calculer la variance de l'estimateur futur qui servira à sélectionner les blocs. Si l'on a une distribution ponctuelle, on peut chercher à corriger celle-ci (moyennant certaines hypothèses) pour obtenir une distribution de blocs présentant la bonne variance. Cette distribution sert à établir les courbes tonnage-teneur en fonction de la teneur de coupure. L'on peut donc prédire les ressources sans devoir les localiser.

Paradoxe : *Il n'est pas possible de localiser précisément les blocs qui permettent d'atteindre les ressources calculées par l'approche distributionnelle.*

Si l'on fait un krigeage l'on observera une variance qui est celle des estimés actuels (laquelle est habituellement inférieure à la variance des estimés futurs plus précis). On aura donc des tonnages et teneurs différents de ceux calculés par l'approche distributionnelle. En pratique, plusieurs rapports NI-43-101 proposent d'ajuster les paramètres de voisinage du krigeage de façon à rapprocher la variance des estimés de celle utilisée globalement pour le calcul des ressources avec l'approche distributionnelle. Ceci est fait en réduisant le voisinage jusqu'à ce que la variance des estimés soit voisine de la valeur recherchée. Cette façon de faire est discutable. En réduisant le voisinage, on augmente la variance des estimés essentiellement dû à des effets de variation de voisinage. Les estimés sont moins précis et localisent donc moins bien les ressources. De plus, les estimés ainsi obtenus ne sont plus sans biais conditionnel. Ainsi, si l'on sélectionnait les blocs en se basant sur les estimés actuels, on ne retrouverait pas la teneur prévue par nos estimations. En résolvant le problème de localisation, on a créé un autre problème peut-être pire. Dit autrement, on a fait coïncider distribution des estimés et ressources globales prévues, mais la localisation des ressources est erronée et ne peut être utilisée pour sélectionner les blocs.

Il existe deux types de méthodes disponibles pour tenter de contrer en partie ce paradoxe : les simulations conditionnelles et les modèles non-linéaires. Dans les deux cas, on s'affranchit de la localisation exacte des ressources pour simplement indiquer des zones plus favorables où l'on devrait retrouver davantage de ressources. Le cumul des ressources estimées localement devrait coïncider assez bien avec les ressources globales estimées, et ce peu importe la quantité de données disponibles.

Malheureusement, les simulations conditionnelles et les méthodes non-linéaires sont encore de nos jours peu utilisées dans les rapports NI-43-101. Parmi les méthodes non-linéaires, celle qui semble la plus couramment utilisée est le conditionnement uniforme (« uniform conditioning » en anglais). L'idée est de découper le gisement en panneaux (blocs de grande taille) pour lesquels l'estimation par krigeage ordinaire peut se faire sans trop de problème, puis d'estimer les distributions des Z_v (ou des estimés futurs) localement, conditionnellement à la valeur krigée du panneau. L'emplacement exact des blocs constituant la ressource n'est donc pas indiqué, mais l'emplacement des panneaux où le plus de ressources sont disponibles, l'est.

Le conditionnement uniforme utilise à la base un modèle de changement de support appelé modèle gaussien discret.

Modèle gaussien discret

Ce modèle suppose que l'on peut écrire $Z = \phi(Y)$ et $Z_v = \phi_v(Y_v)$ où Y et Y_v sont des variables $N(0,1)$ suivant une distribution bigaussienne avec corrélation r . Si l'on prend un point x au hasard dans un bloc v , on a que :

$$E[Z(x) | Z_v] = Z_v$$

Il découle de la loi binormale que $E[Y(x) | Y_v] = rY_v$ et $Var(Y(x) | Y_v) = 1 - r^2$. On peut donc écrire conditionnellement à $Y_v = y_v$ que : $Y(x) = ry_v + (1 - r^2)^{1/2}U$ où U est $N(0,1)$.

On a finalement que :

$$Z_v = \phi_v(Y_v) = E[Z(x) | Z_v] = \int \phi(rY_v + (1 - r^2)^{1/2}u)g(u)du \quad (1)$$

où $g(u)$ est la fonction de densité $N(0,1)$. L'intégrale peut être évaluée numériquement, la fonction ϕ pouvant être obtenue expérimentalement de diverses façons (e.g. avec des polynômes d'Hermite ou par transformation graphique). Le coefficient r compris entre 0 et 1 est déterminé de façon à s'assurer que :

$Var(Z_v) = Var(\phi_v(Y_v)) = Var(Z(x)) - \bar{\gamma}(v, v)$. La corrélation entre Y et Y_v joue donc le rôle d'un coefficient de changement de support. Plus le bloc v est gros, plus la correction de support sera importante et plus le coefficient r sera faible.

Notes :

1- Connaissant Z_v associé à chaque Y_v et sachant que Y_v est distribué suivant une $N(0,1)$, on obtient la distribution de Z_v .

2- Un aspect intéressant de ce modèle est que le coefficient de changement de support est déterminé grâce au variogramme des teneurs (Z) et non du variogramme des valeurs normales (Y) correspondantes. Avec les simulations conditionnelles au contraire, c'est le variogramme des Y qui est utilisé à la base, ce qui oblige à adopter une hypothèse de multinormalité.

Extension pour inclure l'effet information

On peut étendre le modèle pour prédire les ressources qui seront obtenues en se basant sur l'estimation future obtenue au moment de la sélection. On pose que Y et Y_v^* suivent une loi binormale de corrélation q . On veut estimer la distribution de Z_v^* . On posera donc :

$$Z_v^* = \phi_v(Y_v^*) = E[Z(x) | Z_v^*] = \int \phi(qY_v^* + (1 - q^2)^{1/2}u)g(u)du \quad (1b)$$

Le coefficient q tient compte du changement de support et de l'effet information. Il est déterminé de telle sorte que :

$$Var(Z_v^*) = Var(\phi_v(Y_v^*)) = Var(Z(x)) - \bar{\gamma}(v, v) - \sigma_k^2$$

où σ_k^2 est la variance de krigeage obtenue avec l'estimateur futur (supposé identique pour tous les blocs).

Conditionnement uniforme

On s'intéresse à la distribution des teneurs de petits blocs v dans un panneau de taille V avec $V > v$ (typiquement quelques dizaines de blocs v dans V). On fait les mêmes hypothèses du modèle gaussien discret pour les supports v et V . Appelons r_v et $r_V < r_v$ les coefficients de changement de support obtenus (déterminés par les variances de blocs v et V). Considérons un bloc v pris au hasard dans V . On a que :

$$E[Z_v | Z_V] = Z_V$$

(i.e. en moyenne la teneur d'un petit bloc pris au hasard dans le panneau est égale à la teneur du panneau). Si l'on suppose que Y_v et Y_V suivent une loi binormale de corrélation R , on a : $E[Y_v | Y_V] = RY_V$ et

$Var(Y_v | Y_V) = 1 - R^2$ on a donc une équation similaire à (1) :

$$Z_v = \phi_v(Y_v) = E[Z_v | Z_V] = \int \phi_v(RY_V + (1 - R^2)^{1/2}u)g(u)du \quad (2)$$

Si l'on suppose aussi que $Y(x)$ et Y_v suivent une loi binormale de corrélation r_v :

$$Z_v = \phi_v(Y_v) = E[Z(x) | Z_V] = \int \phi(r_v Y_V + (1 - r_v^2)^{1/2}u)g(u)du \quad (3)$$

L'équation (3) permet de calculer la distribution des teneurs des panneaux V à partir des données ponctuelles. L'équation (2) permet d'obtenir cette même distribution à partir de la distribution des teneurs de blocs v . On peut montrer que les équations (1) à (3) impliquent que $R = r_V / r_v$. Pour ce faire, il est pratique de recourir aux polynômes d'Hermite. La fonction de densité de $Z(x)$ peut être représentée par

$$\begin{aligned} Z &= \phi(Y) = \sum_{n=0} \phi_n H_n(Y) \\ Z_v &= \phi(Y_v) = \sum_{n=0} \phi_n r_v^n H_n(Y_v) \\ Z_V &= \phi(Y_V) = \sum_{n=0} \phi_n r_V^n H_n(Y_V) \\ Z_V &= \sum_{n=0} \phi_n r_v^n (r_V/r_v)^n H_n(Y_V) = \sum_{n=0} (\phi_n r_v^n) R^n H_n(Y_V) \end{aligned}$$

qui montre que l'anamorphose de V peut s'obtenir soit de l'anamorphose ponctuelle, soit de l'anamorphose des

blocs v . Les variables Y_v et Y_V sont binormales de corrélation $R = \frac{r_V}{r_v}$

Connaissant Z_v et donc Y_v , et supposant que Y_v et Y_V suivent une loi binormale (de coefficient de corrélation R) il est possible d'estimer toute fonction de Z_v (ou de Y_v) par son espérance puisque la distribution conditionnelle de Y_v étant donné $Y_V = y_V$ est $N(RY_V, (1-R^2))$.

En pratique, l'on ne connaît pas Z_v . Toutefois, on peut utiliser la valeur estimée par krigeage ordinaire pour le panneau V (l'estimation à l'échelle du panneau est normalement précise). Connaissant Z_V , on détermine y_v dans l'éq. 3 par inversion de ϕ_v . On a alors tous les éléments pour l'espérance de toute fonction $f(y_v)$, dont le tonnage, la quantité de métal et la teneur. On voit donc que l'on peut localiser les ressources à l'échelle des panneaux pour les blocs v . Ainsi, le tonnage au dessus du seuil $c = \phi_v(y_c)$ est donné par :

$$\left(1 - G\left(\frac{y_{c_v} - RY_v}{\sqrt{1 - R^2}} \right) \right)$$

Remarque : le nom conditionnement uniforme indique que chaque petit bloc v du panneau V est conditionné par une même valeur Z_v ou Z_v^* . On pourrait aussi vouloir conditionner chaque petit bloc séparément par la valeur krigée (krigeage simple) obtenue Y_v^* (ceci demande le variogramme de Y et donc une hypothèse de multinormalité). Cet estimateur est l'espérance conditionnelle de Z_v étant donné Y_v^* . Toutefois, étant basé sur le krigeage simple, il requiert une hypothèse de stationnarité plus forte. On ne peut à cette échelle utiliser le krigeage ordinaire Y_v^* car les variances de krigeage seraient trop élevées, ce qui impliquerait un problème de biais conditionnel dans les estimations.

Sélection sur une estimation future

Dans la pratique l'on ne connaîtra jamais exactement Z_v , même au moment de la sélection. On pourrait donc vouloir prédire les ressources qui seront récupérées en se basant sur l'estimateur obtenu par krigeage Z_v^* (supposé sans biais conditionnel) qui sera utilisé au moment de la sélection. On utilise à nouveau le modèle gaussien discret. Tout ce qui change dans ce qui précède, c'est le coefficient de changement de support r_v qui devra être déterminé de façon à assurer la variance de l'estimateur futur $\text{Var}(Z_v^*)$ laquelle est déduite des relation de lissage du krigeage comme étant $\text{Var}(Z_v^*) = \text{Var}(Z_v) - \sigma_k^2$. Pour pouvoir utiliser cette approche, il faut que chaque bloc soit estimé de façon très similaire de sorte que la variance de krigeage ne varie pas ou très peu pour chaque bloc v . Sinon, on doit calculer un coefficient de changement de support différent pour chaque valeur de variance de krigeage.

L'utilisation des polynômes d'Hermite

Les polynômes d'Hermite simplifient la description du modèle gaussien discret (du moins pour l'aspect théorique). Dans un contexte gaussien, on peut représenter la fonction ϕ par une série infinie de polynômes d'Hermite :

$$\phi(Y) = \sum_{n=0}^{\infty} f_n H_n(Y) \quad (4)$$

où les f_n sont des coefficients à déterminer et les $H_n(Y)$ sont des polynômes d'Hermite donnés (à $Y=y$) par :

$$H_0(y) = 1$$

$$H_1(y) = -y$$

$$H_2(y) = (y^2 - 1) / \sqrt{2}$$

$$H_{n+1}(y) = -\frac{1}{\sqrt{n+1}} y H_n(y) - \sqrt{\frac{n}{n+1}} H_{n-1}(y)$$

Les f_n sont calculés en évaluant :

$$f_n = E[\phi(y) H_n(y)] = \int \phi(y) H_n(y) g(y) dy \quad (5)$$

On a en particulier que :

$$E[\phi(y)] = f_0$$

$$Var(\phi(y)) = \sum_{i=1}^{\infty} f_n^2$$

Dans le cadre du modèle gaussien discret, on peut montrer que :

$$Z_v = \phi_v(Y_v) = \sum_{n=0}^{\infty} f_n r_v^n H_n(Y_v) \quad (6)$$

et

$$Z_v = \phi_v(Y_v) = \sum_{n=0}^{\infty} f_n r_v^n H_n(Y_v) \quad (7)$$

On obtient que :

$$Var(Z_v) = \sum_{i=1}^{\infty} f_n^2 r_v^{2n} \quad (8)$$

et

$$Var(Z_v) = \sum_{i=1}^{\infty} f_n^2 r_v^{2n} \quad (9)$$

On peut réécrire :

$$\phi_v(Y(x)) = \sum_{n=0}^{\infty} f_n r_v^n H_n(Y_v) = \sum_{n=0}^{\infty} f_n r_v^n \frac{r_v^n}{r_v^n} H_n(Y_v) = \sum_{n=0}^{\infty} f_n r_v^n R^n H_n(Y_v) \quad (10)$$

qui démontre que le changement de support point $\rightarrow V$ (avec coefficient r_v peut être obtenu par le changement de point $\rightarrow v$ avec coefficient r_v suivi d'un changement de support de $v \rightarrow V$ avec coefficient R .

Note :

- En pratique, l'amplitude des coefficients f_n ne décroît pas nécessairement rapidement avec n . Par contre, en raison de l'exposant $2n$ sur le coefficient de corrélation, la contribution à la variance des termes d'ordre supérieur devient rapidement négligeable. Ici, les polynômes sont normalisés pour assurer

$\int H_n^2(y) g(y) dy = 1$. Les valeurs de ces polynômes varient entre $[-8,8]$ pour y entre $[-3$ et $3]$ pour les ordres de 1 à 20 (voir Figure 1). Aux valeurs plus extrêmes de y (peu fréquentes pour une $N(0,1)$) les polynômes peuvent atteindre des valeurs beaucoup plus élevées.

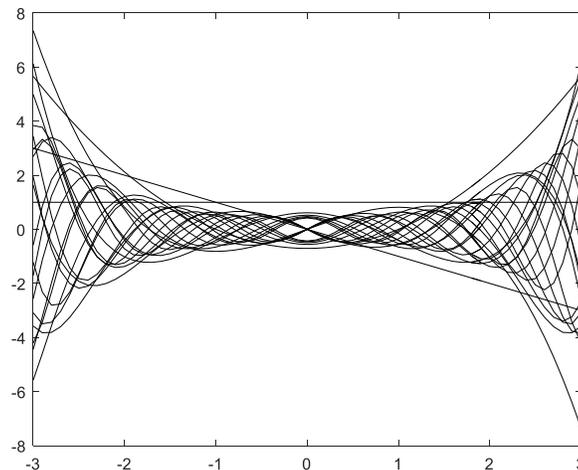


Figure 1 : valeurs des polynômes d’Hermite (n=0 à 20) en fonction de y

Pour bien ajuster une distribution donnée et diminuer les oscillations, il faut parfois atteindre des ordres très élevés. Exemple : même avec un ordre de 60, des oscillations persistent dans les valeurs prédites (points en rouge dans la figure 2- droite).

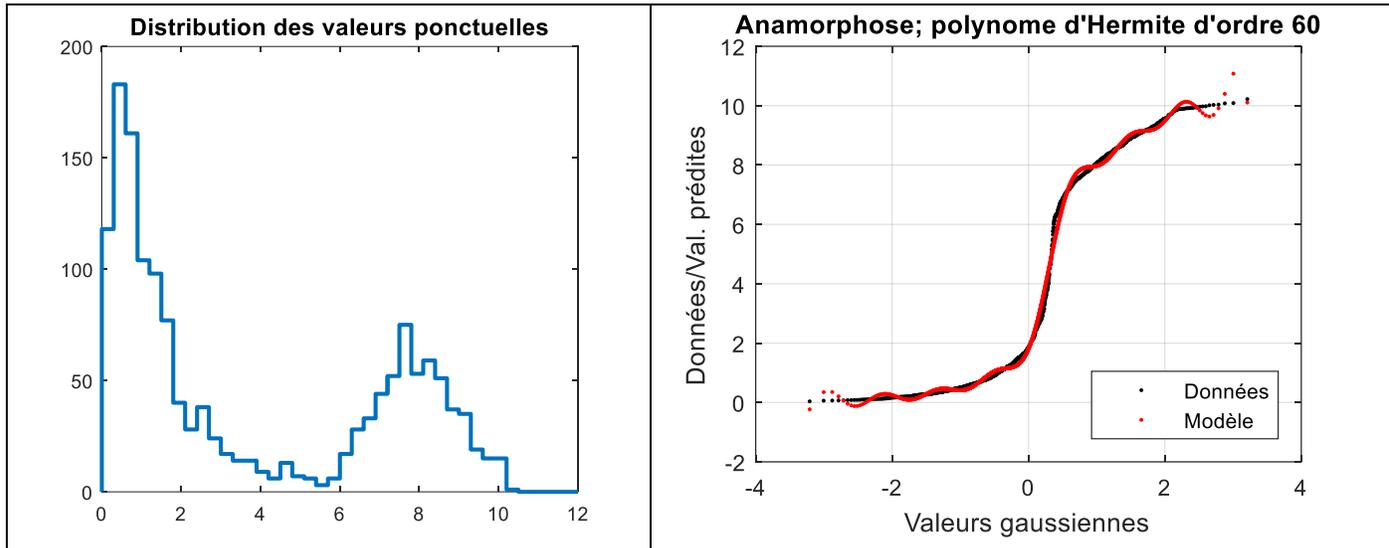


Figure 2 : Distribution ponctuelle bimodale (gauche) et courbes donnant les valeurs observées et prédites par l’expansion par polynôme d’Hermite à l’ordre 60 en fonction de la valeur gaussienne (droite).

On peut appliquer le modèle gaussien discret et le conditionnement uniforme sans avoir recours aux polynômes d’Hermite. Cependant, leur utilisation simplifie les calculs (au prix toutefois d’incohérences (e.g. fonction $\phi(y)$ non-croissante en fonction de y pour certains intervalles de y , particulièrement du côté des valeurs extrêmes). Ces incohérences peuvent être atténuées en augmentant l’ordre du polynôme.

Étapes de mise en œuvre du modèle gaussien discret et du conditionnement uniforme (avec les polynômes d’Hermite)

- 1- Choisir la zone d’étude. Calculer le variogramme des teneurs.
- 2- Choisir la taille des blocs (dépend de la méthode de minage). Choisir la taille des panneaux à considérer (fonction de la distance inter-forages).
- 3- Calculer les coefficients f_n (Eq. 5). Pour ce faire :
 - on classe les teneurs $z_i = \phi(y_i)$, on associe la valeur gaussienne correspondante y_i à chaque rang;
 - on choisit l’ordre de l’expansion N ;
 - on évalue les polynômes $H_n(y_i)$, $n=0\dots N$, à chaque valeur y_i
 - on évalue l’équation 5 numériquement par : $f_n = \frac{1}{nd} \sum_{i=1}^{nd} z_i H_n(y_i)$ (où nd est le nombre de données).
 - on compare les valeurs expérimentales z_i aux valeurs calculées par le modèle (équation 4). Si l’ajustement est satisfaisant, on arrête, sinon, on peut essayer d’augmenter l’ordre du polynôme (Note, les coefficients f_n , $n=0\dots N$ demeurent évidemment inchangés en passant à un ordre supérieur.

- 4- Utilisant le variogramme, calculer la variance de Z_v et de Z_V .
- 5- Déterminer r_v et r_V à partir des équations 8 et 9. Calculer $R=r_v/r_V$.
- 6- Estimer par krigeage ordinaire la teneur des panneaux. Considérer $Z_v \equiv Z_v^{KO}$ et déterminer la valeur y_v correspondante dans l'équation 7.
- 7- La fonction de répartition de Z_v pour ce panneau est obtenue en évaluant :

$$P(Z_v < c | Z_v = Z_v^{KO}) = P\left(N(0,1) < \frac{y_c - Ry_v}{(1-R^2)^{0.5}}\right)$$

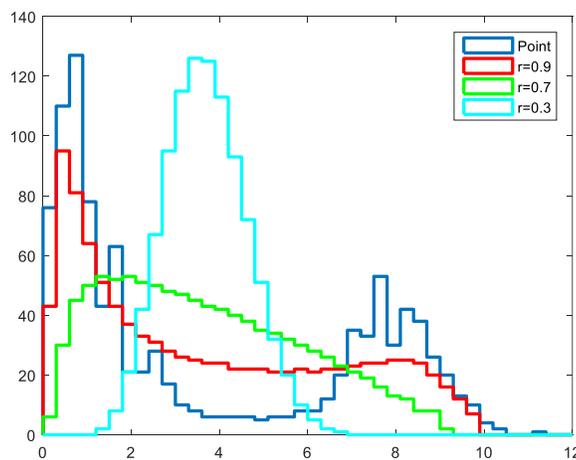
où y_c est obtenu à partir de l'équation 6 évaluée à la valeur $Z_v=c$, i.e.

$$c = \phi_v(y_c) = \sum_{n=0}^{\infty} f_n r_v^n H_n(y_c)$$

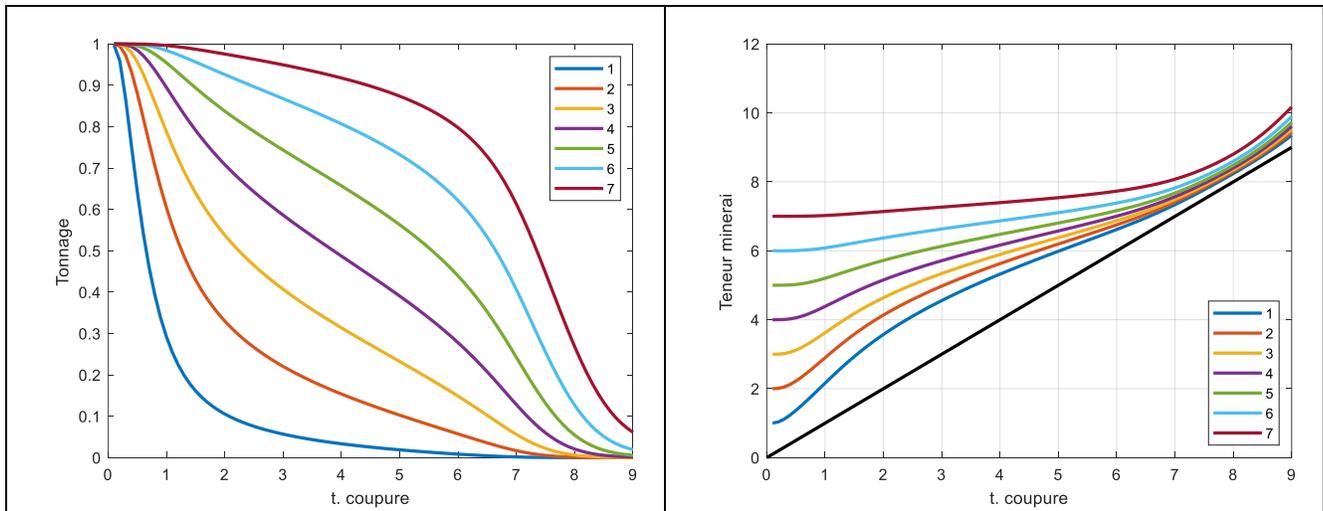
- 8- Ayant la fonction de répartition de Z_v , on peut estimer l'espérance de toute fonction de Z_v (tonnage, quantité de métal, etc.)

Quelques illustrations

- 1- Distributions obtenues en fonction du coefficient de changement de support r



- 2- Distributions du tonnage et teneur en fonction de la teneur de coupure obtenues pour différentes valeurs de krigeage du panneau (en légende dans la figure) : anamorphose précédente, variogramme sphérique avec $a=100$ $C_0=0$ $C=10$, panneau de $50 \times 50 \times 50$ ($r_v=0.78$) et smu de $10 \times 10 \times 10$ ($r_v=0.98$).



Méthodes globales de changement de support

Dans le cas où l'on s'intéresse uniquement au changement de support global, il existe aussi deux autres méthodes (en plus du modèle gaussien discret) couramment utilisées dans la littérature : la correction affine et la correction lognormale indirecte.

Correction affine

C'est la plus simple des corrections, elle consiste à poser

$$Z_v = m + b(Z_i - m) \quad (11)$$

avec

$$b = \left(\frac{\sigma_v^2}{\sigma^2} \right)^{1/2} \quad (12)$$

Cette méthode est à éviter car elle préserve sur la distribution des blocs tous les modes de la distribution ponctuelle et donc ne respecte pas le théorème centrale limite. Il procure des distributions irréalistes dès que le facteur b diminue notablement.

Correction lognormale indirecte

Cette méthode de changement de support suppose que l'on peut associer une valeur Z_v à chaque valeur ponctuelle Z_i par la relation :

$$Z_v = aZ_i^b \quad (13)$$

où les coefficients a et b sont choisis de façon à préserver la moyenne (du support ponctuel) et assurer une variance de bloc qui soit celle désirée et qui est calculée à partir du variogramme des Z_i .

La stratégie est d'abord de déterminer b puis a . Le coefficient b est choisi tel que :

Pour que Z_v soit sans biais, il faut choisir a tel que :

$$a = \frac{m}{E[Z_i^b]} \quad (14)$$

Puis, on calcule la variance de Z_v

$$\text{Var}(Z_v) = a^2 E[Z_i^{2b}] - (aE[Z_i^b])^2 = \sigma_v^2 \quad (15)$$

Substituant (11) dans (12) et simplifiant, on trouve :

$$\frac{E[Z_i^{2b}]}{E[Z_i^b]^2} - \left(1 + \frac{\sigma_v^2}{m^2}\right) = 0 \quad (16)$$

ce qui permet de déterminer b . Note : b est compris entre $b=0$ (où $\sigma_v^2 = 0$) et $b=1$ (où $\sigma_v^2 = \sigma^2$). La fonction 16 est strictement croissante sur l'intervalle 0-1.

On vérifie facilement que :

$$E[Z_v] = aE[Z_i^b] = m$$

et que :

$$\text{Var}(Z_v) = a^2 \text{Var}(Z_i^b) = a^2 (E[Z_i^{2b}] - E[Z_i^b]^2) = m^2 \left(\frac{E[Z_i^{2b}]}{E[Z_i^b]^2} - 1 \right) = m^2 \left(1 + \frac{\sigma_v^2}{m^2} - 1 \right) = \sigma_v^2$$

Ce modèle fonctionne assez bien lorsque la distribution de la variable Z_i n'est pas trop éloignée d'une distribution lognormale (donc variable continue, distribution unimodale avec asymétrie positive). Emery (2016) note toutefois plusieurs problèmes avec cette approche :

- i. pour de très grands blocs, la distribution ne tend pas vers une loi normale comme attendu avec le théorème centrale limite;
- ii. on retrouve la même proportion de valeurs nulles pour les blocs que pour les points, ce qui semble irréaliste;
- iii. la distribution des blocs peut être plus asymétrique que celle des points (comportement non-souhaitable);
- iv. la forme de la distribution des blocs change si l'on translate la distribution des points par une constante;
- v. le modèle est mal adaptée aux distributions discrètes puisqu'il produira aussi des distributions discrètes pour les valeurs des blocs;

Remarque : L'idée de corriger les distributions locales (e.g. obtenues par krigeage d'indicatrice) avec le même facteur que calculé globalement est très discutable. Emery (2016) montre que ceci peut mener à une sous-corrrection des distributions locales. On prévoit donc plus de profit conventionnel que le gisement peut en fournir. De plus, la cohérence entre les ressources calculées à l'échelle globale et à l'échelle locale n'est pas non plus assurée. Pour les corrections locales, il est probablement plus simple et plus avisé dans la majorité des cas de recourir à des simulations conditionnelles (gaussiennes ou autres).