III. Modélisation du variogramme

En pratique, le géostatiticien dégage les caracteriqstiques importants du variogramme expérimentale ; et peut alors lui ajuster une fonction mathématique (le variogramme théorique ou modèle de variogramme), qui peut considérée comme une estimation du vrai variogramme.

III.1 Modèles de variogrammes

A partir d'un variogramme expérimentale, on essaie de caler une fonction qui sera le modèle ajusté. Ce modèle sera utilisé dans l'interpolation. Lin et Chen (2004) ont montré que le calcul du variogramme expérimental a une grande influence sur la précision de l'estimation. Selon Schäfer-Neth et al. (2005), le krigeage exige un ajustement précis des paramètres des variogrammes.

Le variogramme expérimental est non biaisé , mais sa dispersion augmente fortement quand devient grand. C'est un estimateur peu robuste car sensible aux valeurs extrêmes ou aberrantes en particulier pour des pas faibles en comparaison des distances entre sites d'observation.

Le variogramme expérimental permet l'estimation du variogramme théorique pour un nombre défini de distances, c'est-à-dire uniquement des valeurs ponctuelles. De plus, il ne respecte pas nécessairement les propriétés théoriques du variogramme. L'idée est donc d'ajuster un modèle variographique théorique présentant les caractéristiques nécessaires et défini pour un ensemble de vecteurs.

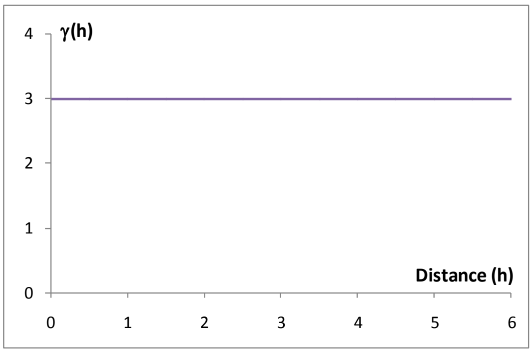
En fonction du variogramme expérimental, on choisira le modèle variographique qui s'en rapproche le plus. Parmi ces modèles on peut citer les exemples suivants:

III.1.1 Modèles isotropes avec palier

Dans un variogramme expérimental, on distingue deux comportements, le premier est à l'origine et le deuxième à l'infini. A l'origine, nous pouvons observer un comportement continu de type parabolique ou linéaire, ou un autre discontinu avec un effet de pépite ou pépite pure. La pépite est due à des erreurs de mesure ou bien à un phénomène microrégional à une échelle très inférieure au pas d'espace étudié. Un effet de pépite pur montre que la variable n'a pas de structure spatiale.

Modele pepitique de palier : comportement discontinu à l'origine. Il est caractéristique des gisements miniers, en particulier aurifères

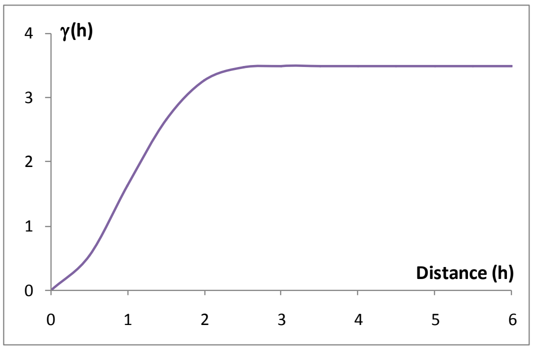
$$\gamma(h) = C_0 \qquad h \geq 0$$



Le comportement à l'infini est défini par un variogramme borné ou non borné. Le variogramme borné devient constant au-delà d'une distance de décorrélation (corrélation nulle) appelée portée, cette constante correspond à la variance de la variable Z appelée palier.

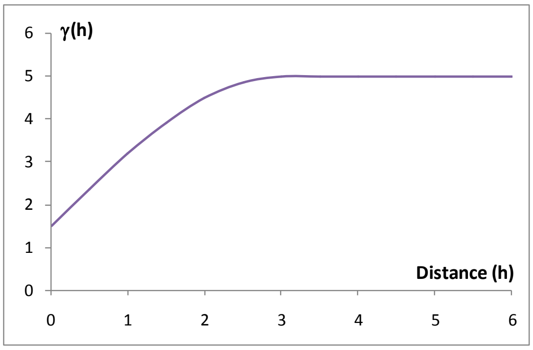
Modèle cubique de portée et de palier : parabolique à l'origine.

$$\gamma(h) = \begin{cases} C * \left[7\left(\frac{h}{a}\right)^2 - \frac{35}{4}\left(\frac{h}{a}\right)^3 + \frac{7}{2}\left(\frac{h}{a}\right)^5 - \frac{3}{4}\left(\frac{h}{a}\right)^7\right], 0 \le h \le a \\ C, \quad h > a \end{cases}$$



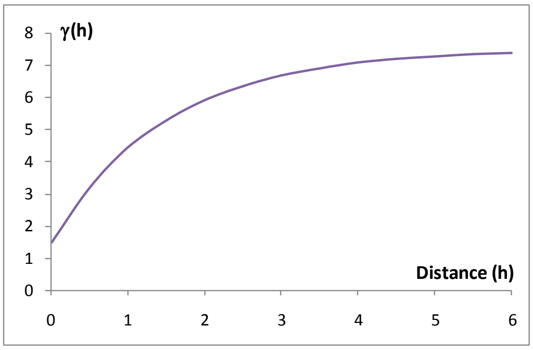
Modèle sphérique de portée et de palier : linéaire à l'origine.

$$\gamma(h) = \begin{cases} C * \left[1.5\left(\frac{h}{a}\right) - 0.5\left(\frac{h}{a}\right)^3\right], & 0 \le h \le a \\ C, & h > a \end{cases}$$



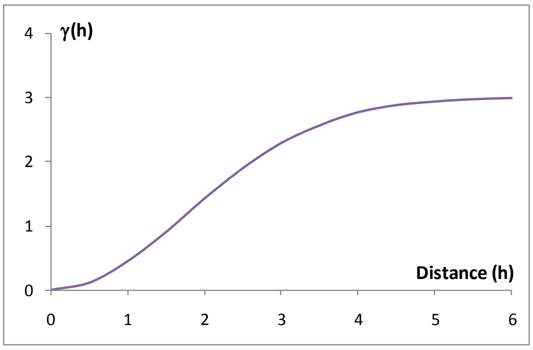
Modèle exponentiel de paramètre et de palier : linéaire à l'origine.

$$\gamma(h) = C * \left[1 - e^{-\left(\frac{h}{a}\right)}\right]$$



Modèle gaussien de paramètre et de palier : parabolique à l'origine

$$\gamma(h) = C * \left[1 - e^{-\left(\frac{h}{a}\right)^2}\right]$$



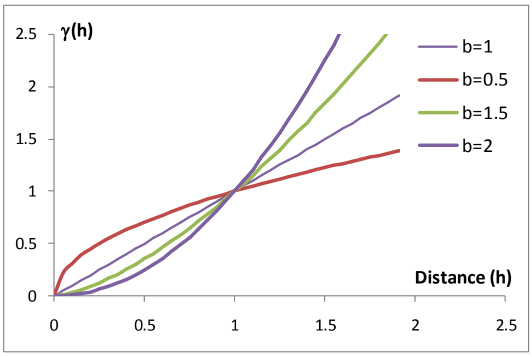
III.1.2 Modèles isotropes sans palier

Ce type de modèles dits non bornés ou sans palier ; ils sont utilisés dans le cadre non stationnaire d'ordre 2.

modèle puissance : ce modèle est sans palier, d'exposant b et de facteur d'échelle ω. Plus b est proche de 2, plus la variable régionalisée est régulière. Plus est proche de 0, plus son comportement est erratique.

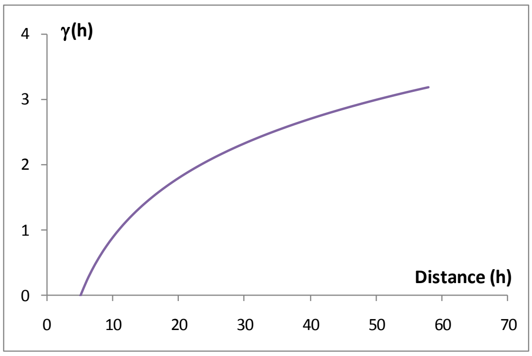
le modèle linéaire est un cas particulier du modèle puissance c'est le cas ou b=1.

$$\gamma(h) = \omega * h^b \qquad b = 0 \text{ à } 2$$



modèle logarithmique : ce modèle est sans palier.

$$\gamma(h) = C * log\left(\frac{h}{a}\right) \qquad h \geq a$$

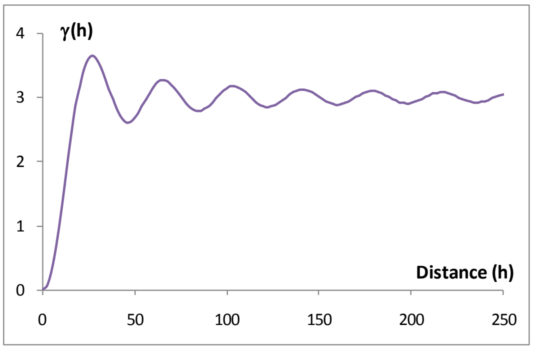


III.1.3 Modèles isotropes à effet de trou

On peut trouver aussi une chute brusque de la variance dans les dernières classes, due au nombre insuffisant de couples, ou bien à la mauvaise répartition des points dans l'espace, ce qu'on appelle "l'effet de trou". Dans le cas d'un variogramme non borné, la variable n'est pas stationnaire d'ordre 2, et la moyenne varie dans le champ de mesure. Cela est appelé une dérive.

Modèle a effet de trou : parabolique à l'origine ; il permet de modéliser un variogramme moins stable, présentant des fluctuations autour du palier : il est caractéristique de données plus hétérogènes.

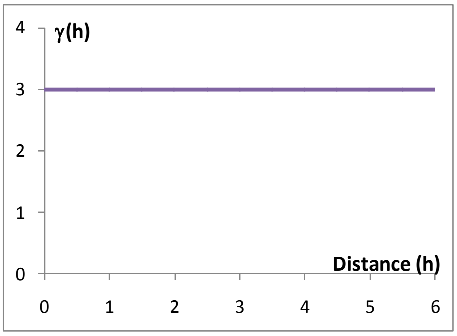
$$\gamma(h) = C * \left[1 - \frac{\sin\left(\frac{h}{a}\right)}{\frac{h}{a}}\right]$$



III.1.4 Comportement au voisinage de l'origine du variogramme

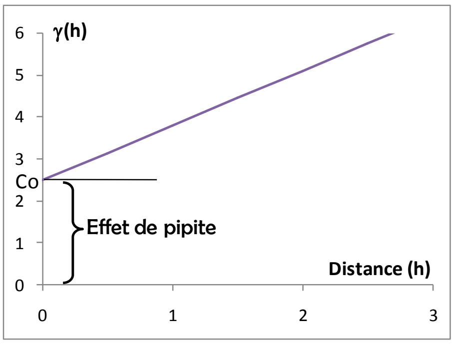
Il est très utile d'étudier le comportement du variogramme pour de petites valeurs de h, puisqu'il est relié à la continuité et à la régularité spatiale de la variable. La figure ci-dessous illustre quatre types de comportement au voisinage de l'origine :

a) **Plat: hasard pur ou bruit blanc.** Les valeurs mesurées sur deux échantillons distincts ne présentent aucune corrélation, quelle que soit la distance qui les sépare. Ceci constitue le cas limite d'une absence totale de structure.



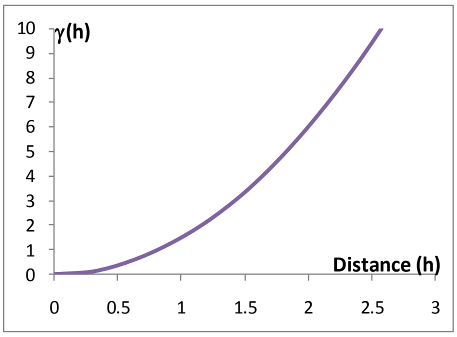
Aléatoire pur

b) **Discontinu à l'origine:** γ(h) ne tend pas vers 0 lorsque h tend vers 0. Ceci signifie que la variable est extrêmement irrégulière même à de petites distances. Ce saut à l'origine est appelé effet de pépite parce qu'il a été observé pour la première fois en Afrique du Sud où il était associé à la présence de pépites d'or dans le minerai. Il est commode d'utiliser ce terme d'effet de pépite pour désigner ce genre de fortes variations à courtes distances même lorsqu'on sait qu'il est dû à d'autres facteurs, par exemple des erreurs de mesure, ou une incertitude de localisation.



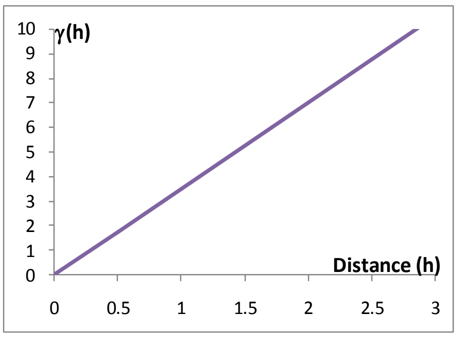
Discontinue

c) **Parabolique (γ(h)= h²):** un tel comportement indique que la variable régionalisée est très régulière: elle est continue et différentiable.



Continue

d) **Linéaire (γ(h)= h) :** dans ce cas la variable régionalisée est continue mais non différentiable, c'est à dire moins régulière que (a).



Continue

III.14 Anisotropies

Il existe certains phénomènes, dont la continuité spatiale n'est pas nécessairement la même dans toutes les directions. Lorsque le variogramme est calculé pour tout couple de points dans certaines directions comme Nord-Sud ou Est-Ouest, il révèle parfois des différences de comportement (c'est-à-dire une anisotropie). Si cela ne se produit pas, le variogramme ne dépend alors que de la distance entre les points, et on parle alors d'isotropie.

Lorsqu'on calcule un variogramme, il est important d'utiliser au minimum quatre directions. Si le variogramme était calculé pour seulement deux directions perpendiculaires, on risquerait de passer à côté de l'anisotropie.

Bien que dans la nature il existe une très grande variété d'anisotropies, en géostatistique, on ne peut modéliser aisément que les anisotropies géométriques. Les autres anisotropies peuvent être approchées en combinant plusieurs modèles isotropes ou avec anisotropie géométrique. Généralement on distingue deux principaux types d'anisotropie : l'anisotropie géométrique et l'anisotropie zonale.

1) Anisotropie géométrique

Appelée aussi anisotropie elliptique : ce type d'anisotropie est déterminer par des variogrammes directionnels qui ont les mêmes paliers dans toutes les directions même si leurs portées sont différentes, ou des pentes différentes lorsque il s'agit des variogrammes linéaires. dans ce cas l'anisotropie peut être corrigée par une transformation affine des coordonnées.

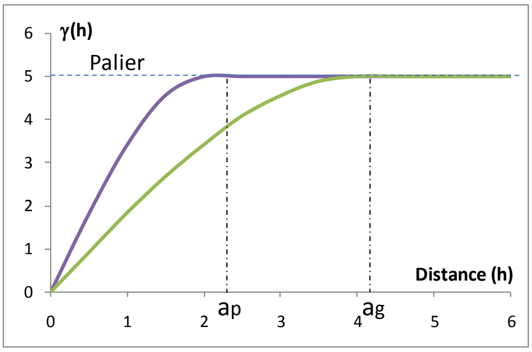


Fig. 01 : Anisotropie modèle sphérique

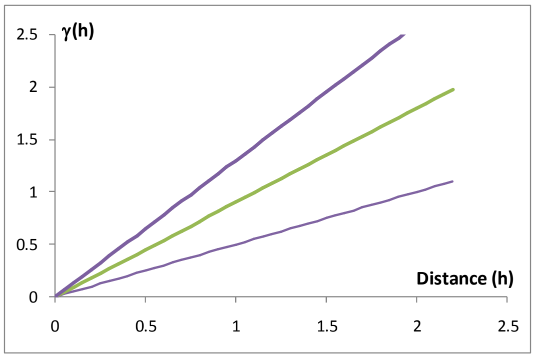


Fig. 02 : Anisotropie modèle linéaire

- On observe dans diverses directions des paliers et des composantes pépitiques identiques mais des portées différentes.
- Les portées maximales (ag) et minimales (ap) s'observent selon deux directions orthogonales.
- On peut rendre les portées identiques (et égales à ag suivant toutes les directions en multipliant la composante de la portée parallèle à ap par le facteur (ag/ap). Les portées décrivent une ellipse dont l'axe majeur est orienté parallèlement à ag.

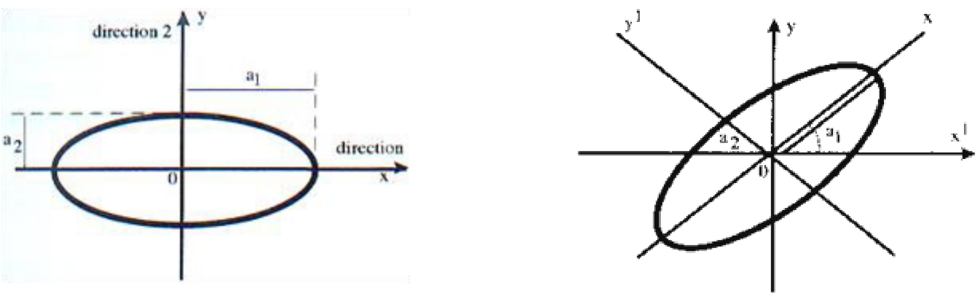


Fig. 03 : Deux exemples d'elipsses d'anisotropie

L'équation de l'ellipse est donnée par : $$\frac{(a_\theta Cos\theta)^2}{a_g^2} + \frac{(a_\theta Sin\theta)^2}{a_p^2} = 1$$

Ce type d'anisotropie peut être corrigé par une transformation des coordonnées. Connaissant a_g et a_p, on peut trouver a_θ, où θ désigne l'angle mesuré par rapport à la direction où est rencontrée ag.

$$a_\theta = \frac{a_g a_p}{\left[a_p^2\ Cos^2\theta + \ a_g^2 Sin^2\theta\right]^{1/2}}$$

Exemple :

Un gisement 2D est modélisé par un modèle avec anisotropie géométrique. Le modèle est sphérique avec C=17%^{2} et effet de pépite C0=13%^{2} et les portées sont de 100m dans la direction (convention trigonométrique) de plus grande continuité (30°) et 60m dans la direction de plus petite continuité (120°). Quelle est la valeur du variogramme entre deux observations situées aux coordonnées $(x_1,y_1)=(10,30)$ et $(x_2,y_2)=(40,20)$.

Réponse :

On calcule d'abord la distance séparant les deux points et la direction qu'ils définissent:

- La distance est donnée en appliquons le théorème de Pythagore par l'expression :

$$h = \sqrt{(\Delta x)^2 + (\Delta y)^2}\ \ avec\ \begin{cases}\Delta x = x_2 - x_1 \\ \Delta x = y_2 - y_1\end{cases}$$

d'où $$h = \sqrt{(20 - 30)^2 + (40 - 10)^2}\ = 31.62m$$

- La direction que définissent les deux points est :
  $\theta = arctan\left(\frac{y_2 - y_1}{x_2 - x_1}\right) = arctan\left(\frac{-10}{30}\right) = -18.4°$

- Cette direction forme un angle de (θ =48.4°) avec la direction de plus grande continuité (c à d suivant ag).

On calcule la portée dans cette direction en utilisant la formule donnant aθ :

$$a_\theta = \frac{100 \times 60}{\left[60^2\ Cos^2 48.4 + \ 100^2\ Sin^2 48.4\right]^{1/2}} = 70.81m$$

On calcule la valeur du variogramme en utilisant l'équation du modèle sphérique pour la distance h entre les deux points et avec la portée 70.81m:

Sachant que le modèle de variogramme c'est modèle sphérique avec un C=17%^{2} et effet de pépite C0=13%^{2}, et de portée a_θ = 70.81m calculée suivant la direction de plus grande continuité. L'expression du variogramme est :

$$\gamma(h) = \begin{cases} C_0 + C * \left[1.5\left(\frac{h}{a}\right) - 0.5\left(\frac{h}{a}\right)^3\right], & 0 \le h \le a \\ C_0 + C, & h > a \end{cases}$$

Pour h = 31.62 m

$$\gamma(h) = 13 + 17 * \left[1.5\left(\frac{31.62}{70.81}\right) - 0.5\left(\frac{31.62}{70.81}\right)^3\right] = 23.63\ \%^2$$

Pour la détection des anisotropies géométriques on doit vérifier certaines conditions :

a) Le facteur d'anisotropie géométrique obtenu avec les variogrammes expérimentaux sous estime en général le véritable facteur d'anisotropie en raison de l'utilisation d'une fenêtre angulaire et du fait que les variogrammes expérimentaux ne sont pas nécessairement orientés exactement selon les directions principales de l'ellipse d'anisotropie.
b) L'estimation correcte et à la limite, la détection, d'anisotropie géométrique n'est possible, en pratique, qu'à quatre conditions (fortement liées) devant être remplies simultanément:
   - Le nombre de données est suffisant (au moins 50)
   - Le facteur d'anisotropie est important (au moins 1.5)
   - Une des directions utilisées dans le calcul du variogramme est près de la direction de plus grande portée.
   - La fenêtre angulaire utilisée est suffisamment étroite.

2) Anisotropie zonale

Appelée aussi stratifiée, ce type d'anisotropie est très rare, elle n'est rencontrée que dans certain cas, lorsque il s'agit de la troisième dimension.

Dans un domaine à 3D la direction verticale joue souvent un rôle particulier parce qu'il y a plus de variations entre les strates qu'à l'intérieur des strates. Ce types d'anisotropie est plus complexe.

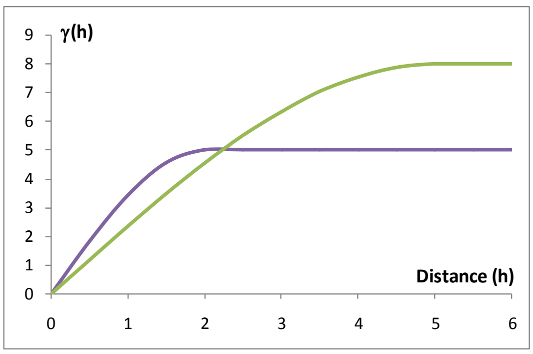


Fig. 04 : anisotropie zonale

Dans ces cas, la pratique courante est de séparer le variogramme en deux tenues, le premier étant isotrope, et le second ne dépendant que de la composante verticale :

$$\gamma(h) = \text{ } variogramme\ isotrope + composante\ verticale$$

$$\gamma_{Zonale}(h, \theta) = \gamma_{isotrope}(h) + \ \gamma_p(h.sin\theta)$$

Ou l'indice p représente le modèle anisotrope suivant la direction de portée minimale.

III.1.3 Choix d'un modèle

Dans la pratique, les logiciels de géostatistique proposent une liste de modèles d'ajustement aux variogrammes. L'utilisateur fournit donc d'abord le type de modèle (linéaire, exponentiel...) qu'il désire ajuster. Ensuite, il peut:

- Soit fournir lui-même par un ajustement visuel les paramètres du modèle (pépite, portée...);
- Soit demander au programme de trouver ces paramètres par une procédure d'ajustement automatique (moindres carrés,...)

Souvent, différents modèles peuvent a priori être ajustés à un même variogramme. Pour choisir le modèle optimal, deux techniques sont utilisées:

- On choisit le modèle s'ajustant le mieux au variogramme expérimental;
- On choisit le modèle fournissant les meilleures estimations aux points observés. Pour cela, on estime par krigeage (cf. Infra) la valeur en un point observé à partir des autres points observés. Le modèle retenu sera celui donnant les écarts entre valeurs estimées et observées les plus faibles. C'est la technique dite "Jackknife" ou "Option Thomas".

III.1.4 Conditions de calcul de variogrammes et l'ajustement de modèles

- On accorde plus de poids aux points du variogramme expérimental calculés avec beaucoup de paires.
- On essaie d'avoir $N(h) \geq 30$ pour chaque point expérimental du variogramme. Si ce n'est pas possible pour certaines classes, on accorde moins d'importance à ces points. Si le nombre de paires est très faible (<10), on ne considère plus du tout le point.
- On accorde plus de poids aux premiers points du variogramme (h petit) car ce sont ces valeurs qui ont le plus d'impact dans les calculs géostatistiques.
- Lorsque h dépasse environ $\frac{d_{max}}{2}$, on ne tient pas compte des valeurs du variogramme. dmax est la taille du phénomène étudié dans la direction considérée.
- On cherche à obtenir des modèles les plus simples possible qui rendent bien compte des valeurs expérimentales.

Toute modélisation d'un variogramme doit passer par les étapes suivantes :

- Calculer les variogrammes directionnels selon différentes directions (ex. 0°, 45°, 90°, 135°) ainsi que le variogramme omnidirectionnel (i.e. sans tenir compte de la direction). La géologie peut apporter une information précieuse dans le choix des directions et la présence ou non d'anisotropies.
- Vérifier les critères ci-dessus : $N(h) \geq 30$, h < dmax/2

- Si nécessaire, augmenter la tolérance angulaire ou le pas de calcul de façon à augmenter N(h).
- Déterminer s'il y a anisotropie (différences de palier ou de portées qui ne peuvent raisonnablement être imputées à des fluctuations aléatoires du variogramme).
- Procéder à l'ajustement d'un modèle anisotrope ou isotrope selon le cas (habituellement par essai et erreur, bien que l'on puisse aussi obtenir ces ajustements de façon automatique par régression (pondérée, et souvent, non linéaire).
- Chercher à respecter la règle de la parcimonie: adopter les modèles les plus simples possibles qui permettent un ajustement adéquat. Comparer des modèles concurrents à l'aide de la technique de validation croisée.