--- Polycopie du cours Géostatistique ---

Master M2 Eau et Environnement

Enseignant: Pr. B. Azouzi

Partie 2

- Généralités
- Théorie des variables régionalisées
- Variographie
- TD

Année universitaire 2020/21

I. Généralités : Théorie des Variables régionnalisées

I-1 INTRODUCTION:

L'approche géostatistique se manifeste à travers la théorie des variables régionnalisées qui se fixe comme objectifs de caractériser la structure spatiale des phénoménes naturels sous une forme mathématique approprié et de résoudre le probléme d'éstimation à partir d'un échantillonnage donné (Delhomme, 1976)

Voyons maintenant, qu'est ce qu'une Variable régionnalisée et comment est-elle étudiée et interpretée ?

La variable régionnalisée peut être conque comme une variable qui est susceptible de caractériser un phénomène se répétant dans l'espace et/ou temps et y manifestant une certaine structure (par le biais d'une correlation régionnalisée).

Notons que le terme "Regionnalisée" a été utilisé pour la première fois par G.MATHERON (1965).

Une telle régionnalisation ou comportement spatiale est trop complexe pour être étudié par les méthodes déterministes; c'est pourquoi on préfère utiliser les approches stochastiques (probabilistes) des fonctions aléatoires (F.A.) pour aborder les V.R (JOURNEL,76). Par conséquent une V.R est supposée être une réalisation de fonction aléatoire (F.A) c'est à dire comme le résultat d'un tirage au sort effectué selon une loi de probabilité qu'on doit déterminer.

La reconstitution de la loi de probabilité d'une F.A (qui est l'ensemble des V.A) par le biais de ses premiers moments (Espérance mathématique et Variance...) fait l'objet de l'inférence statistique qui se heurte à une impossibilité qui se résume comme suit : la détermination de la loi de probabilité d'une F.A suppose

qu'on a un nombre important de réalisations, or en pratique une seule réalisation de la F.A est disponible puisque le phénomene naturel étant généralement unique. Que faire dans ce cas ?.

Pour pouvoir lever cette impossibilité, certaines hypothèses de travail ont été proposées.

1-2 HYPOTHESE DE STATIONNARITE :

Statistiquement parlant, la stationnarité d'un phénomène exprime son homogénité spatiale qui fait qu'il se répète lui même en quelque sorte dans l'éspace (DELHOMME,76).

Une F.A est dite stationnaire si sa loi de probabilité ou du moins ses deux premiers moments (Moyenne et Variance, Covariance ou Variogramme) sont invariants par translation.

On parle alors de stationnarité d'ordre 2 si :

- L'espérance mathématique est constante et indépendante de X.

$$E I Z(X) I = m$$

- La covariance entre deux points X1 et $\,$ X2 $\,$ dépend $\,$ seulement $\,$ du vecteur $\,$ X1-X2 $\,$

E I
$$Z(X_1)$$
-m I I $Z(X_2)$ -m I = $C(X_1 - X_2)$

et la variance indépendante de X

$$E I (Z(X)-m)^2 I = VAR I Z(X) I = C(0)$$

Dans ce cas l'inférence statistique devient possible parceque tous les couples [Z(Xk), Z(Xk')]§ dont l'interdistance Xk-Xk' est égale à h pourraient être considérés comme des réalisations différents du couple de V.R [Z(X1), Z(X2)]§.

Mais cette hypothèse de stationnarité est trés forte, elle se vérifie trés rarement en réalité étant donné la capacité de dispersion très grande des phénomènes naturels (Cas de dérive ou "Trend") (G.BRIERE, M.RAZACK, BRGM.1982).

C'est pourquoi au lieu d'exiger la stationnarité de la F.A, on impose seulement aux accroissements de cette F.A d'être stationnaires d'ordre 2, C'est ce qu'on appelle :HYPOTHESE INTRINSEQUE.

Mais avant d'aborder l'hypothèse intrinsèque, notons qu'une autre hypothèse appellée : HYPOTHESE D'ERGODICITE a été imposée pour faciliter le travail.

Elle consiste à supposer que la réalisation unique étudiée suffit à rendre compte de l'ensemble des réalisations éventuelles (G.BRIERE, M.RAZACK, BRGM.1982).

ELle suppose également que des moyennes spatiales peuvent être utilisées pour estimer des moyennes dans l'ensemble des réalisations. (A.DIEULIN, B.BEAUDUIN, G.de Marsily,1980 in AMMARA.,1990)

L'hypothèse d'ergocité permet alors de déterminer la loi de probabilité de la F.A. par la connaissance d'une seule réalisation.

I-3 Hypothèse intrinsèque:

elle se traduit par le fait que pour tout vecteur h, l'accroissement I Z(X+h)-Z(X) I possède une espérance mathématique et une variance indépendante du point X. On a :

E I Z(X+h)-Z(X) I = 0 Var I Z(X+h)-Z(X) I = 2 $\delta(h)$ Dans ce cas , la F.A correspondante est dite fonction aléatoire intrinsèque F.A.I

La fonction **d**(h) s'appelle " **DEMI-VARIOGRAMME** " mais appellée communément " **VARIOGRAMME** "

Il faut noter aussi que parfois l'hypothèse intrinsèque n'est pas vérifié sur tout l'étendu de l'espace du phénomène étudié, alors on se contente uniquement d'une hypothèse quasi-intrinsèque oû la stationnarité des accroissements est localement vérifié pour des distances I h I inférieur à une limite " a " , et on définit ainsi des voisinages glissants à l'interieur desquels l'espérance et covariance (ou variogramme) peuvent être considérés comme stationnaires et oû l'information disponible est suffisante pour en faire l'inférence (JOURNEL.76).

On voit ainsi comment on a passé en affaiblissant l'hypothèse de stationnarité à l'hypothèse intrisèque puis à hypothèse quasi-intrinsèque, tout cela pour faciliter l'inférence statistique.

II. VARIOGRAMME:

De l'équation précédente (1), δ (h) se déduit ègale à :

$$\delta(h) = 1/2 \text{ VarI } Z(X+h)-Z(X) \text{ I}$$

Compte tenu de E I Z(X+h)-Z(X) I = 0

On peut écrire que :

$$\delta(h) = 1/2 E I (Z(X+h)-Z(X))^2 I$$
.

Donc $\delta(h)$ est le demi accroissement quadratique moyen entre deux points distant de h.

Cette fonction $\delta(h)$ est calculée à partir de l'ensemble des couples des points expérimentaux , étant donné que $2\delta(h)$ est considérée comme la moyenne arithmitique des écarts quadratiques entre deux mesures exprémentales I Z(Xi) , Z(Xi+h) I implantées en 2 points quelconques distant du vecteur h (JOURNEL ,76):

$$2\delta(h) = 1 / N(h) \sum I Z(Xi) - Z(Xi+h) I^{2}$$

N(h) étant le nombre de couples expérimentaux I Z(Xi), Z(Xi+h) I distant de h .

La technique du variogramme dévellopée par G.MATHERON et son école dans le cadre des relations de voisinage consiste en fait à évaluer et à discuter la correlation au sens statistique du terme entre les éléments voisins : On cherche à observer et interpréter l'évolution de cette correlation lorsque la distance entre éléments associés croit progressivement (M.DUPUY, 1969). En fait on essaie de caractériser la structure spatiale d'un phénomène naturel donné par l'outil variogramme et c'est la première étape indispensable à toute étude géostatistique.

Il faut noter que le variogramme calculé pour différentes classes de directions géographiques, permet de mettre en évidence l'isotropie ou l'anisotropie du phénomène : par exemple la portée d'un variogramme qui représente la distance au dela de laquelle le variogramme se stabilise et au dela de laquelle les mesures ne sont plus corrélées; cette portée peut changée d'une direction à une autre manifestant ainsi une anisotropie quelconque.

Remarque:

Notons avant de passer à l'estimation du variogramme et son interprétation graphique, que toutes les opérations géostatistiques depuis la caractérisation structurale (autocorrelation) jusqu'à l'estimation, s'effectuent aussi bien avec l'outil variogramme que l'outil Covariance.

Cette dernière s'applique dans le cadre de la stationnarité d'ordre 2 qui suppose l'existence d'une covariance et donc d'une variance à priori finie.

Quand la covariance existe, le variogramme existe, et ces 2 outils structraux équivalents se déduisent l'un de l'autre par la relation suivante :

$$\delta(h) = C(0) - C(h)$$

C(0) : Variance = $Var(Z(X) = \delta(\omega))$

C(h): Covariance $\delta(h)$: Variogramme

Fig 1 COVARIACE ET VARIOGRAMME $\delta(\infty) = C(\circ)$ $C(\circ)$ $C(\infty) = \circ$

Si nous connaissons la covariance , le variogramme est simplement sa symetrie avec une translation (BEUCHEUR, DELHOMME, MARSILY, 1981)

II- -1 DETERMINATION DU VARIOGRAMME :

Rappelons encore une fois la formule du variogramme :

$$\delta(h) = 1 / 2N(h) \sum I Z(Xi) - Z(Xi+h) I^2$$
avec :

N(h) le nombre de couples.

Dans le cadre de l'hypothèse intrinsèque, le calcul du variogramme se fait à partir d'un échantillonnage des points de mesures Zi, en définissant un certain nombre de classes de distances entre ces points. Puis on y range tous les couples formés par les points échantillonnés c'est à dir n (n-1) / 2 couples).

On calcule alors pour chaque classe =

- La distance moyenne h
- Le variogramme δ

- et le nombre de couples présents dans la classe et qui doit être le plus élevé possible pour que les moyennes (variogramme) estimées aient une précision suffisante (AIDAOUI 85 in AMMARA, 90). Mais en général, il y a plus de couples aux petites distances qu'aux grandes distances et le variogramme deviendra plus incertain au fur et à mesure que la distance h s'accroit (DELHOMME, 1981),

Par conséquent la présentation graphique de ce variogramme, indique jusqu'à quelle distance un point expérimental peut influencer sur l'ensemble des autres informations (points exprementaux).(Beucher, 81)

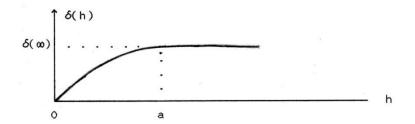
II · 2 Comportement du variogramme :

En général quand il n'y a pas de pépite, le variogramme croit à partir de l'origine, il prend la valeur 0 pour h=0 et à mesure que h grandit l'écart quadratique moyen $\delta(h)$ entre les 2 variables z(X) et z(X+h) tend à croite.

Mais dans un grand nombre de cas le variogramme au dela d'une certaine limite "a" appelée portée se stabilise autour d'une valeur limite $\delta(\infty)$ appelée "palier" qui se rapproche fortement de la variance $\delta(\infty)$ = var I Z(X) I = C(0)

Le variogramme qui prend un tel comportement caractérise ce qu'on appelle un phénoméne de transition du fait que la portée "a" caracterise la transition entre l'état de correlation spatiale (I h I \langle a) et l'état d'absence de corrélation (I h I \rangle a).

d'après Journel 76, un tel phénomène correspond à une F.A non seulement intrinsèque mais aussi stationnaire d'ordre 2



II- 2-1 Comportement à l'origine du variogramme:

le comportement à l'origine du variogramme définit la continuité et la régularité de la F.A Z(x). Par ordre de régularité décroissante on distingue classiquement quatre types de comportements à l'origine de $\delta(h)$:

a) Comportement parabolique :

$$\delta(h) \longrightarrow A I h I^2$$
 quand $h \longrightarrow 0$

Cette allure paraboloque à l'origine caractérise une variable avec une variabilité spatiale hautement régulière.

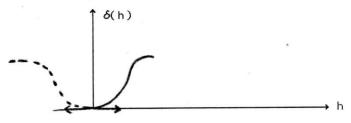


fig 3 Comportement parabolique (dérivable)

 $\delta(h)$ est 2 fois dérivable à l'origine.

b) Comportement linéaire :

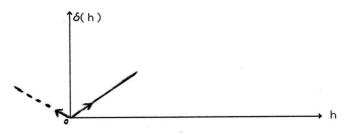


Fig 4 Comportement linéaire (Continue)

Remarque :

- Les 2 dérivées à droite et à gauche $\delta(0^+)$ et $\delta(0^-)$ éxistent mais sont de signes opposés.

- Une F.A Z(x) est dite continue en moyenne quadratique si lim $E(I Z(x+h)-Z(x) I^2)$ 0 quand 0

c) Discontinuité à l'origine :

Théoriquement on devrait avoir $\delta(h)=0$ pour h=0, mais on constate souvent que le variogramme présente une ordonnée à l'origine non nulle C ,ce qu'on appelle effet de pépite qui correspond à une forte concentration locale d'une teneur (en l'occurance l'Or)

Une fonction aléatoire F.A est dite discontinue en moyenne quadratique lorseque :

$$\lim E (IZ(x+h)-Z(x)I^{2}) \longrightarrow Cste quand h \longrightarrow 0$$

On dit alors que la variable est trés irrégulière car entre deux points expérimentaux trés proches il y a un écart de variance au moins égale à l'effet de pépite C (DELHOMME,76), qui est dû à la fois aux erreurs de mesure et aux micro-régionalisation (ou microvariabilité) que l'échantillonnage n'a pas pu cerner.

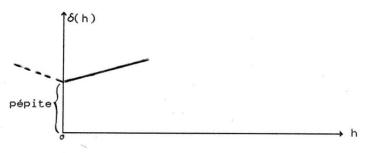


Fig.5 : Effet de pépite (partiel)

b) Effet de pépite pur:

Dans ce cas, il y a une absence totale de structure qui correspond au cas aléatoire pur c'est à dire qu'il n'ya aucune correlation entre 2 points expérimentaux quelque soit leur distance h.

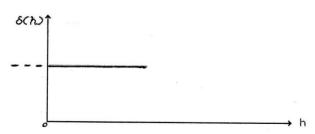


Fig.5: Effet de pépite pur (Aléatoire)

II- 2-2 COMPORTEMENT A L'INFINI :

aux grandes distances le variagrame peut présenter deux aspects à savoir:

a) Variogramme non borné: où il peut continue à croître indéfiniment.

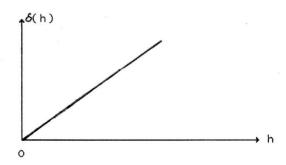


Fig 7 Variogramme non borné

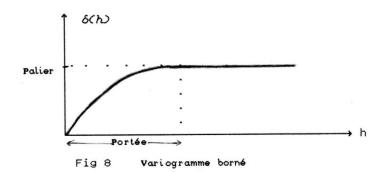
Dans ce cas on assiste à une variance infinie et la covariance n'existe pas, seule l'hypothèse intrinsèque est ici acceptable .

b) Variogramme borné:

On assiste ici à un phénomène de transition parceque le variagramme se stabilisé autour d'une valeur limite appelée "palier " qui corréspond à la variance.

On a déja dit que la distance à laquelle cette limite est atteinte s'appelle "portée" et elle indique la zone d'influence d'une information sur son entourage. Au dela de cette portée les points experentaux ne son plus corrélés.

Rappelons encore une fois, un tel variogramme presente une F.A non seulement intrinsèque comme le cas du variogramme non bornés, mais aussi stationnaire d'ordre 2.

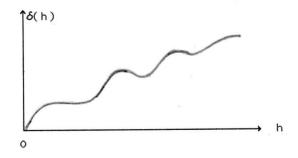


- II-2-3 Comportement particuliere de variagramme :

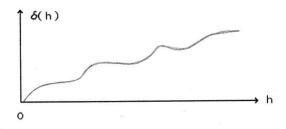
Notons qu'en dehors de ce qu'on a cité, le variogramme peut prendre d'autres comportements subordonnés aux differentes fluctuations apparents sur le graphe qui peuvent être dûes aux manque de couples par classe.

On peut citer :

a) Comportement periodique: Fig 9



b) Structures gigognes: Ou superposition de variations d'échelles differentes Fig 10



c) Effet de trou:

existance de correlation négative dûe à la présence de zones "riches" d'information entourées de zones pauvres et par conséquant la croissance du variogramme n'est pas monôtone (journel, 76).

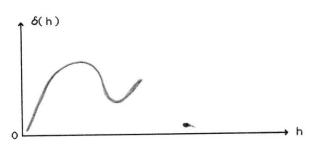


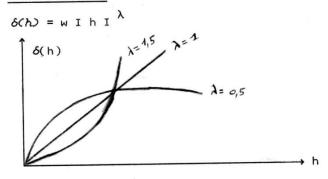
Fig 11 Effet de trou

II- ·3 Ajustement à un modèle :

Pour pouvoir bien exploiter le variogramme et d'en tirer le maximum d'informations sur la structure du phénomène, il est nécéssaire d'approximer le graphe du variogramme par un modèle théorique connu.

Il y a 5 types de modèles d'ajustement du variogramme à savoir :

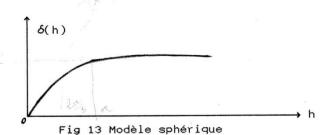
a) Modèle linéaire: Fig 12



c'est un schéma sans palier

b) Modèle Sphérique : A A A A Common

$$\delta(h) = M I (3/2 h/2 - 1/2(h/2)^3 I$$
 pour h (a $\delta(h) = M$ pour h) a



C'est un schéma de transition à palier avec un comportement linéaire à l'origine.

c) Modèle exponentiel:
$$\delta(h) = \text{WI 1 - e}^{-h/\alpha} \text{ I}$$

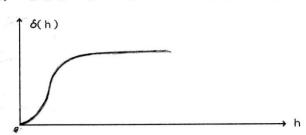
$$\delta(h)$$

O Fig 14 Modèle exponentiel

C'est un schéma de transition à palier avec un comportement linéaire à l'origine.

d) Modèle gaussien : Fig 15

$$\delta(h) = w I 1 - e^{-(h/\alpha)} I$$



C'est un schéma de transition à palier avec un comportement parabolique à l'origine.

e) Modèle Cubique :

$$\delta(h) = w I 7.(h)^2/a - 8,75.(h)^8/a + 3,5.(h)^5/a - 0,75.(h)^7/a I$$

 $\delta(h) = w$ pour h > a

II- -4 Vérification de la validité du modèle :

Une fois que le graphe expérimental du variogramme est ajusté à un modèle théorique, il est plus commode de s'assurer qu'un tel modèle est adéquat au phénomène étudié pour être utilisé par la suite dans la résolution du problème d'estimation par Krigeage.

Une vérification consiste à estimer par krigeage chaque valeur connue Z(Xi) par la combinaison linéaire des autres valeurs disponibles, on dispose ainsi à la fois de la variance d'estimation théorique du krigeage $\sigma^2(Xi)$ et de l'erreur K

expérimentale $[Z(Xi)-Z^*(Xi)]$ en chaque point Xi d'implantation d'une donnée.

On dit qu'on a un bon ajustement si :

-
$$1/n \Sigma (Z(xi)-Z^*(xi) \longrightarrow 0$$

-
$$1/n \Sigma (Z(x)-Z^*(xi)/\sigma)^2 \longrightarrow 1$$

TD-Variographie

Exercice 1

On a un variogramme sphérique avec Co=1 et C=10, a=20.

Soit les points x1=(0,0) et x2=(10,0)

- a) Quelle est la variance de Z(x1)? de Z(x2)?
- b) Quelle est la covariance entre Z(x1) et Z(x2)?
- c) On forme $Z_3=0.8*Z(x1)+0.2*Z(x2)$. Quelle est la variance de Z_3 ? Quelle est la covariance de Z_3 avec Z(x1)?
- d) On forme $Z_4=0.4*Z(x_1)+0.6*Z(x_2)$. Quelle est la covariance entre Z_3 et Z_4 ?

Exercice 2

Un site contaminé au Pb montre un modèle de variogramme comprenant 3 composantes différentes :

Composante	C (ppm ₂)	$a_g\left(m\right)$, θ_g (azimut)	$a_p(m)$, θ_p (azimut)
Effet pépite	120	-	-
Sphérique	1 580	1000, 87°	300, 177°
Sphérique	2 1200	400, 42°	200, 132°

Où θ_g est la direction de meilleure continuité et θ_P est la direction de moindre continuité. Selon l'azimut 30°, à quelle distance séparant deux points peut-on considérer les teneurs en Pb non corrélées?

Exercice 3

On vous donne 2 portions de forage sur lesquelles sont indiquées les teneurs en Au (ppm) pour des carottes de 3m. les 2 forages sont espacés de 9m de centre à centre (le dessin n'est pas à l'échelle).

9m			
5.2	5.2		
2.3	4.6		
3.7	6.3		
9.1	8.1		
4.3	2.7		

Calculez le variogramme expérimental omnidirectionnel à la distance h=9m exactement en prenant soin d'indiquer toutes les paires utilisées.

Exercice 4

Dans une mine 2D de Cu, un géologue vous dit qu'il a observé les faits suivants :

- Il a prélevé plusieurs échantillons de même taille côte à côte et il a obtenu, pour ces échantillons voisins, une différence carré des teneurs égale, en moyenne, à 2%².
- Selon son expérience, la zone d'influence d'un échantillon est de l'ordre de 70m dans la direction est-ouest. Il affirme que cette zone d'influence est approximativement deux fois plus courte dans la direction nord-sud. La direction est-ouest correspond à l'orientation générale des fractures dans le gisement.
- Il a calculé une variance de 10%² pour la teneur en Cu avec l'ensemble des échantillons couvrant une très grande surface par rapport à la zone d'influence d'un échantillon.

Suggérez un modèle de variogramme (type et paramètres) pouvant tenir compte de ces informations.

Exercice 5

Un gisement de Zn présente un variogramme sphérique (2D) avec effet de pépite de 4 %2, C=20%2 et des portées suivant les différentes directions décrites par un modèle d'anisotropie géométrique avec les axes principaux en « x » et « y » et ax=10m et ay=20m. On désire estimer la teneur au point x0 (situé en (0,0)) en se servant des teneurs mesurées aux points x1 à x3. Les coordonnées des points x1

à x_3 sont respectivement (-10,0), (0,20) et (5,22). On a observé à ces points les teneurs respectivement $Z_1=2\%$, $Z_2=3.3\%$ et $Z_3=3\%$.

Quelle est la covariance entre les teneurs aux points x2 et x3?

Exercice 6

Soit le plan de localisation représenté par les points suivants :

x1 (-30, 0) x2(0, 20) x3(10, 20) x4(30, 0) x0(0, 0)

Le variogramme est sphérique anisotrope avec $a_x=30m$ et $a_y=50m$ et effet de pépite de $5\%^2$. Le « C » du sphérique est de $50\%^2$ (palier total $55\%^2$).

Quelle est la covariance entre les teneurs aux points x₀ et x₃?

Corrigé

```
1- a) Var(Z(x1))=Var(Z(x2))=C_0+C=11
b) \gamma (h)=1+10*(1.5(10/20)-0.5*(10/20)<sup>3</sup>)=7.875
Cov(Z(x1),Z(x2))=11-7.875=3.125
c) Var(Z_3)=0.8^2Var(Z(x1))+0.2^2Var(Z(x2))+2*0.8*0.2*Cov(Z(x1),Z(x2))
=0.64*11+0.04*11+2*0.8*0.2*3.125=8.48
Cov(Z_3,Z(x1))=0.8*Var(Z(x1))+0.2*Cov(Z(x1),Z(x2))
=0.8*11+0.2*3.125=9.425
d) Cov(Z_3,Z_4)=Cov(0.8*Z(x1)+0.2*Z(x2),0.4*Z(x1)+0.6*Z(x2))=0.8*0.4*Var(Z(x1))+0.8*0.6*Cov(Z(x1),Z(x2))+0.2*0.4*Cov(Z(x1),Z(x2))+0.2*0.6*Var(Z(x2))=11*(0.8*0.4+0.2*0.6)+3.125*(0.8*0.6+0.2*0.4)=6.59
```

2- On cherche la portée selon la direction 30°. On calcule la portée pour chacune des 2 composantes sphériques et on retient la valeur maximale.

```
Sphérique 1 : angle avec a_0: 57^{\circ}, portée à 30^{\circ} = 1000^*300/[1000^{2*}sin(57)^2 + 300^{2*}cos(57)^2]^{0.5} = 351.1 Sphérique 2 : angle avec a_0: 12_{\circ}, portée à 30^{\circ} = 400^*200/[400^{2*}sin(12)^2 + 200^{2*}cos(12)^2]^{0.5} = 376.3 La portée dans la direction 30^{\circ} est donc 376.3 m
```

3- à 9m paires : (5.2, 9.1), (2.3,4.3) (5.2, 6.3) (4.6, 2.7) (5.2,5.2) (2.3,4.6) (3.7,6.3) (9.1,8.1), (4.3,2.7) N(h)=9 paires Somme des diff₂=46.84 γ (h)= 0.5*46.84/9=2.60

4- Variogramme sphérique ou exponentiel avec anisotropie géométrique, portée principale selon estouest de 70m, portée la plus petite de 35m selon nord-sud, C₀=0.5*2%²=1%², C=10-1=9%²

```
5- Cov(Z_2,Z_3) -> direction : azimut=atan((5-0)/(22-20))=68.2^0 portée dans cette direction : 20*10/((20sin68.2)<sub>2</sub>+(10*cos(68.2))<sub>2</sub>)<sup>0.5</sup>=10.56 h entre les deux points : (5_2+2_2)^{0.5}=5.385 C(h)=20*(1-(1.5*5.385/10.56-0.5*(5.385/10.56)_3))=6.02879
```

6- la distance est $(100+400)^{0.5}$ =22.36 l'angle avec l'axe des y : atan(10/20)=26.6° La portée dans cette direction= $50*30/(502\sin 2(26.6)+302\cos 2(26.6))^{0.5}$ =42.9m C(h)= $50*(1-1.5*22.36/42.9+0.5*(22.36/42.9)^3)$ =14.5