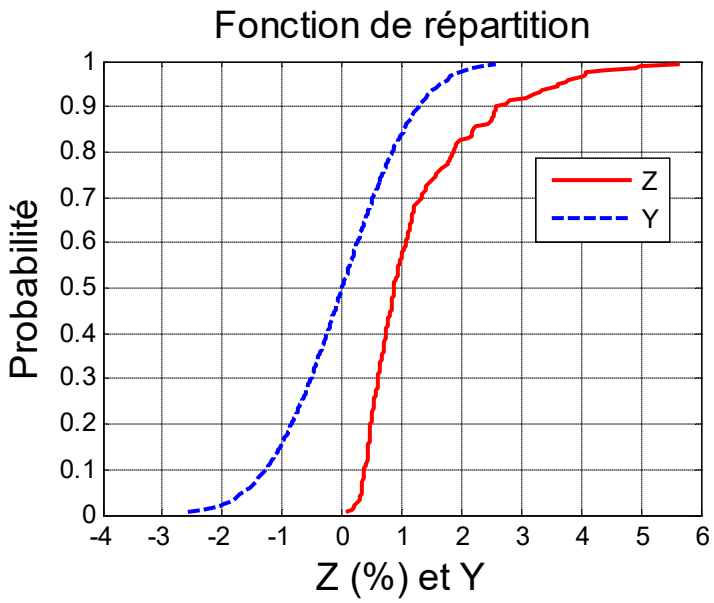
Examen final 2016
Question 1 (11 points)

La méthode de simulation SGS consiste à :

i. choisir au hasard un point à simuler;
ii. estimer par krigeage simple la distribution conditionnelle de la variable en ce point compte tenu des observations connues et des points déjà simulés;
iii. tirer aléatoirement une valeur de la distribution conditionnelle et ajouter cette valeur à l'ensemble des données observées et des données déjà simulées.

Lorsque les données (Z) ne suivent pas une distribution gaussienne, on doit au préalable les transformer (Y) pour les rendre gaussiennes. La figure suivante montre graphiquement une telle transformation.



2 pts *a) Une donnée valant Z(x)=3% aurait quelle valeur gaussienne après transformation?*

2 pts *b) On applique l'algorithme. À un point x_0 donné, le krigeage simple retourne la valeur -1.2 avec un écart-type de krigeage de 0.3. Quelle est la distribution conditionnelle de $Y(x_0)$?*

Question 1 (suite)

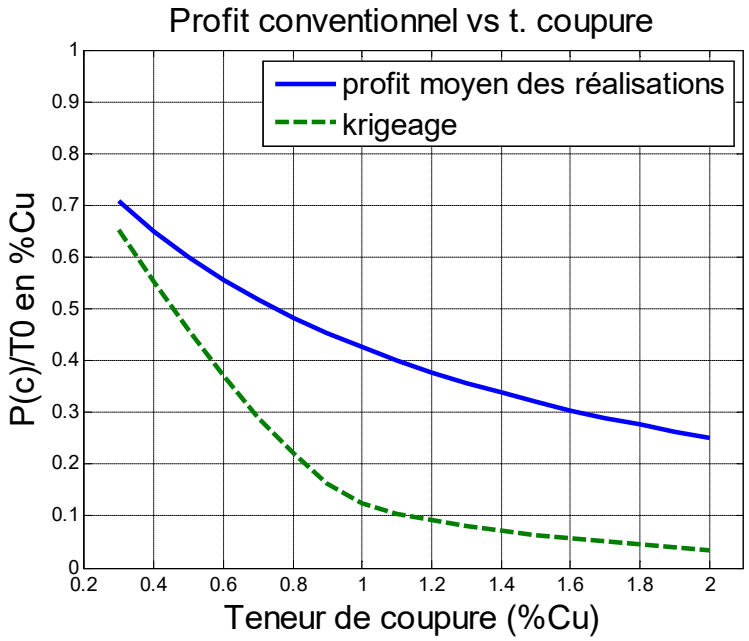
3 pts *c) Toujours au point x_0, on tire aléatoirement une valeur entre 0-1 interprétée comme la valeur de la fonction de répartition conditionnelle. Supposons que la valeur 0.9 est tirée. Quelles valeurs $Y(x_0)$ et $Z(x_0)$ simule-t-on?* (Note : une table N(0,1) est fournie en annexe).

2 pts *d) Pour le krigeage simple effectué au prochain point à simuler, doit-on utiliser $Y(x_0)$ ou $Z(x_0)$?*

2 pts *e) Suggérez et décrivez brièvement une approche qui permettrait d'utiliser le principe séquentiel du SGS, mais sans devoir transformer vers la loi normale et supposer la distribution multigaussienne de la variable transformée.*

Question 2 (10 points)

Soit la figure suivante montrant le profit conventionnel en fonction de la teneur de coupure obtenue par krigeage et le profit obtenu, en moyenne, pour les différentes réalisations d'une simulation.



4 pts *a) Si l'on sélectionnait les blocs à partir des estimés actuels de krigeage, quel profit conventionnel obtiendrait-on approximativement à une teneur de coupure de 1% ?*

4 pts *b) Si l'on connaissait parfaitement les teneurs réelles des blocs, quel profit pourrait-on espérer obtenir à la teneur de coupure 1% ?*

2 pts *c) Selon les courbes présentées, devrait-on chercher à améliorer l'estimation des blocs si la teneur de coupure est 1% ? Justifiez.*

Question 3 (12 points)

On vous fournit le tableau suivant donnant, le long d'un profil, les coordonnées, les valeurs observées, les valeurs d'une simulation non-conditionnelle, les valeurs krigées avec d'une part, les données observées et d'autre part, les valeurs simulées.

Coordonnées	Valeurs observées	Valeurs krigées avec les données	Simulation non-conditionnelle	Valeurs krigées avec les valeurs simulées aux points des données	Simulation conditionnelle *(à calculer)*
0	8	8	10	10	
50	-	8.9	7.5	9.7	
100	-	10	9	9.2	
150	-	11.1	11.8	8.7	
200	12	12	8	8	

6 pts *a) Transformez par post-conditionnement par krigeage la simulation non-conditionnelle en une simulation conditionnelle (inscrivez vos réponses directement dans le tableau)*

Question 3 (suite)

3 pts *b) Le signal simulé représente la topographie du fond marin. Supposons que l'on veuille déterminer l'espérance de la pente maximale susceptible d'être rencontrée le long du profil. Est-il préférable de :*

A) générer plusieurs réalisations, calculer la pente maximale pour chaque réalisation et faire la moyenne des pentes maximales sur toutes les réalisations

ou

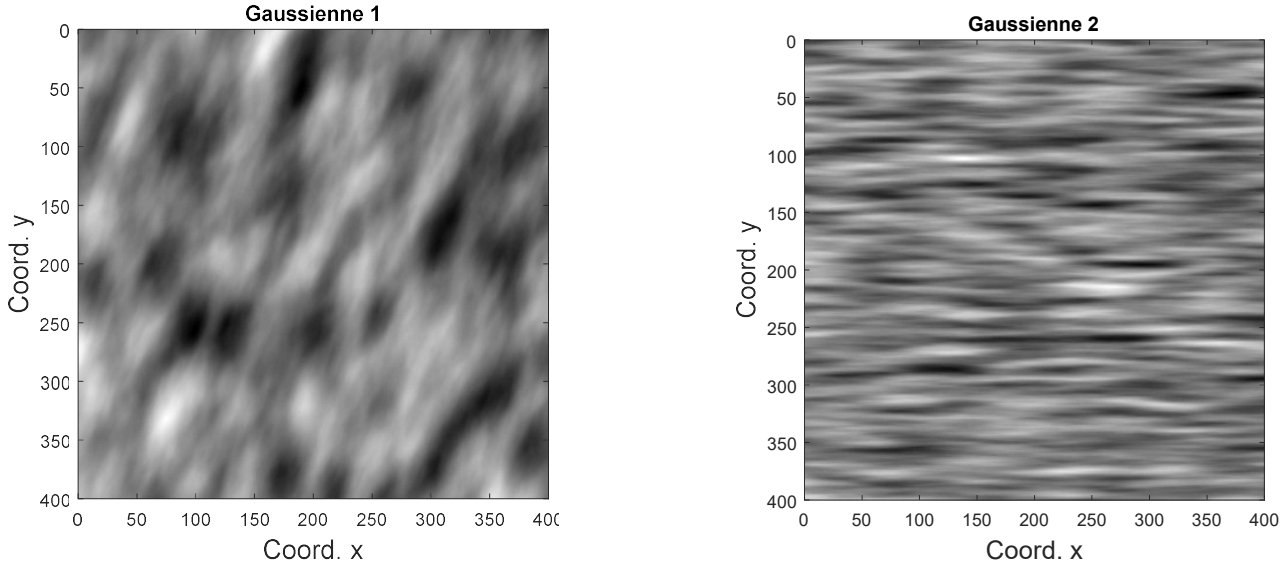
B) générer plusieurs réalisations, calculer la moyenne des réalisations et calculer la pente maximale sur cette « réalisation » moyenne.

Justifier.

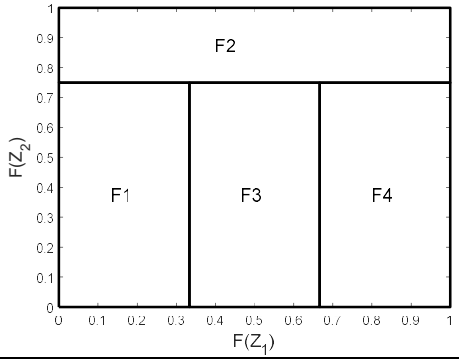
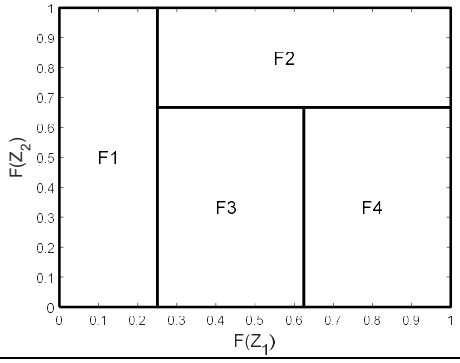
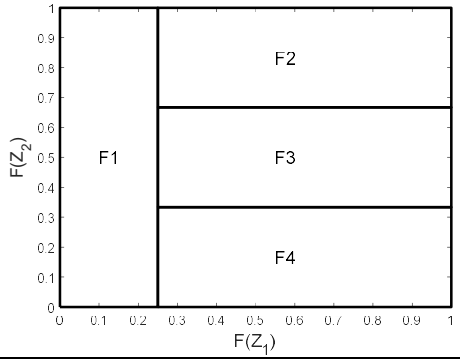
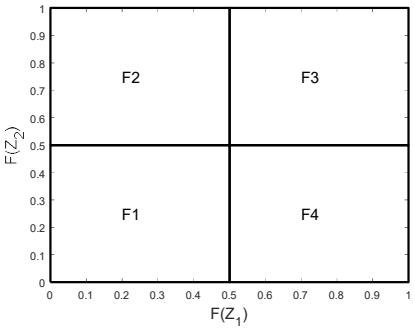
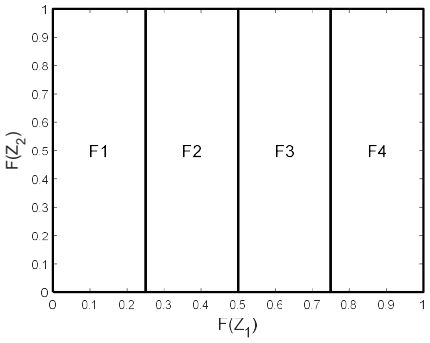
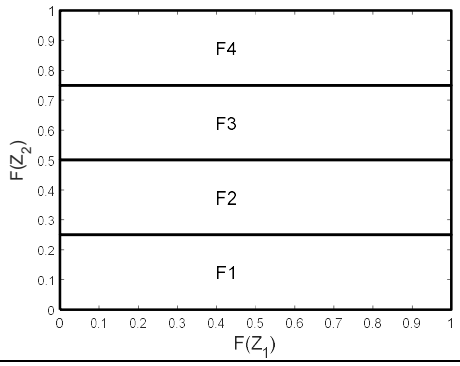
3 pts *c) Le modèle de variogramme utilisé pour la simulation est sphérique avec C=25 et a=230. La variance de krigeage au point x=100 vaut 17.6. Quel devrait être l'écart-type des réalisations au point x=100?*

Question 4 (16 points)

La figure 1 montre les réalisations de deux variables gaussiennes indépendantes ayant chacune une structure spatiale différente.

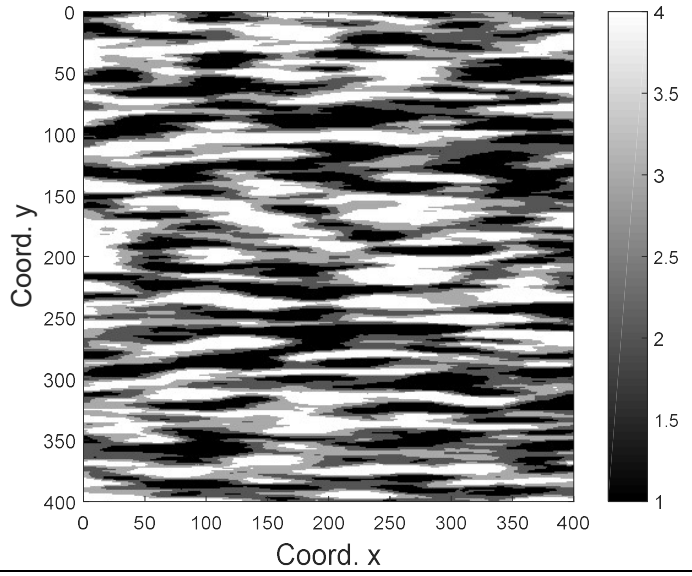
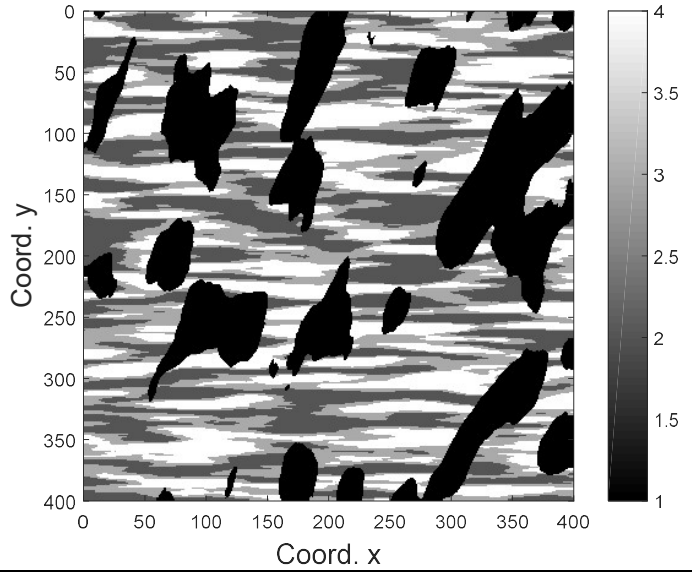


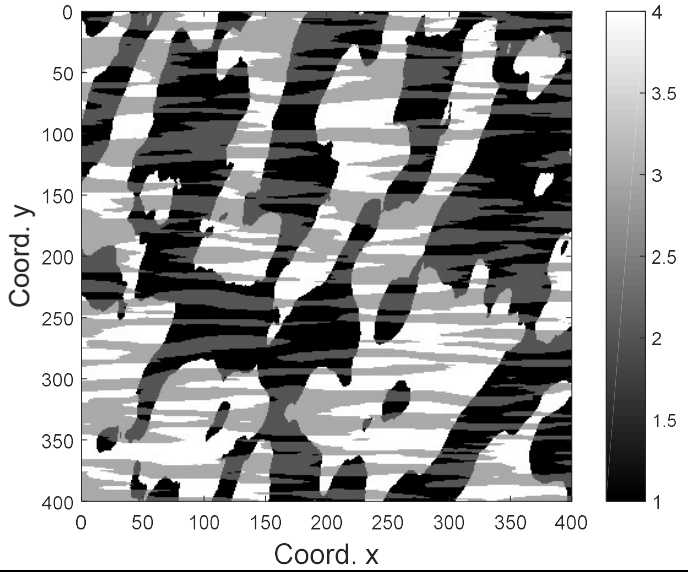
Le tableau 1 montre différents drapeaux de codages appliqués aux variables gaussiennes précédentes. L'axe des x est la fonction de répartition de la première gaussienne, l'axe des y est la fonction de répartition de la seconde.

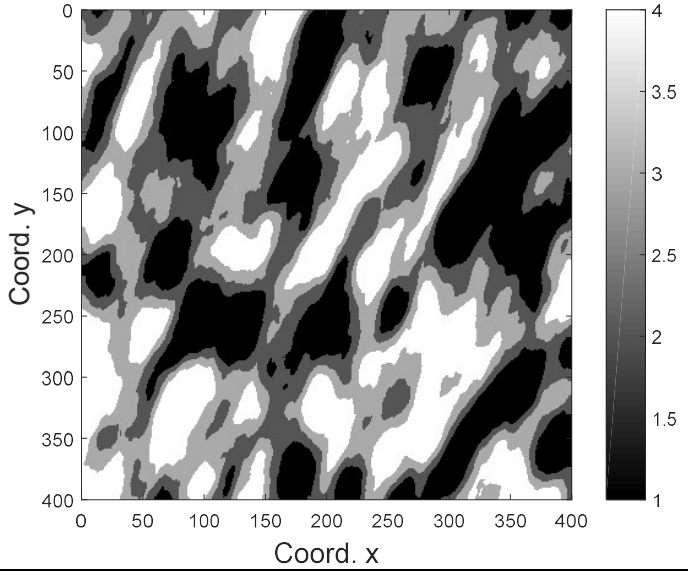
Drapeau A	Drapeau B	Drapeau C
		
Drapeau D	**Drapeau E**	**Drapeau F**
		

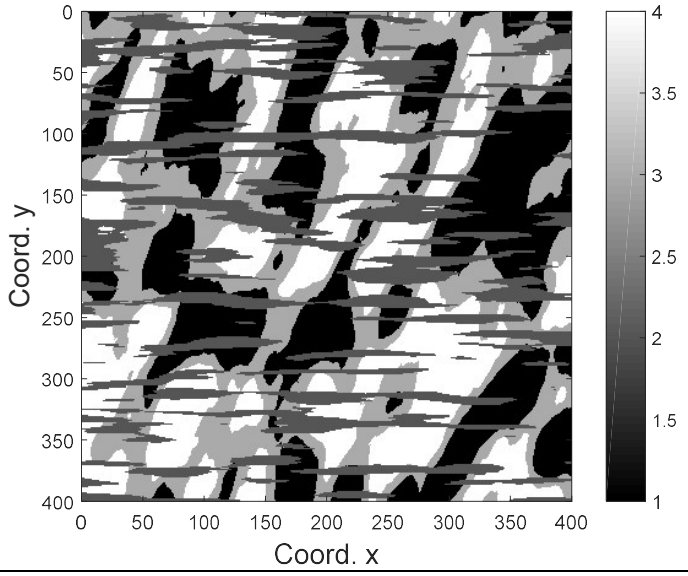
Question 4 (suite)

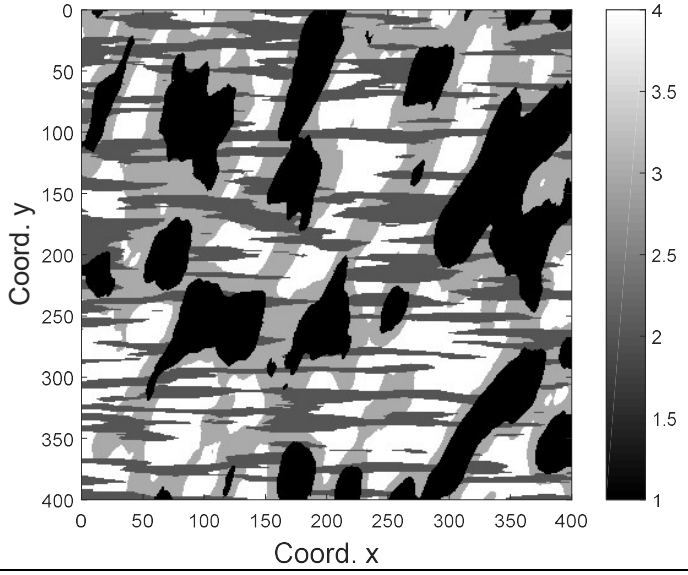
12 pts *a) Associez à chacune des réalisations suivantes un des drapeaux parmi ceux illustrés à la figure 1. L'échelle de gris identifie le faciès (1->F1, 2->F2, 3->F3, 4->F4).*

Réalisation	Drapeau #
	
	


Coord. y
Coord. x
0
50
100
150
200
250
300
350
400
1
1.5
2
2.5
3
3.5
4


Coord. y
Coord. x
0
50
100
150
200
250
300
350
400
1
1.5
2
2.5
3
3.5
4


Coord. y
Coord. x
0
50
100
150
200
250
300
350
400
1
1.5
2
2.5
3
3.5
4


Coord. y
Coord. x
0
50
100
150
200
250
300
350
400
1
1.5
2
2.5
3
3.5
4

Question 4 (suite)

2 pts *b) En un point donné on a les valeurs gaussiennes Z_1=1.2 et Z_2=2.2. Utilisant le drapeau B, quel faciès est simulé à ce point ?*

2 pts *c) Quel algorithme doit-on utiliser pour imposer les faciès observés aux points d'observation? Décrivez brièvement son fonctionnement.*

Question 5 (12 points, soit 2 points par sous-question)

On a un champ en 2D de taille D pour lequel on génère un très grand nombre de réalisations non-conditionnelles sur une grille serrée de points. Les simulations sont effectuées avec moyenne théorique « m » et variance théorique σ^2.

a) *On se place en un point donné x_0 et l'on examine les valeurs obtenues pour les différentes réalisations. Que valent la moyenne et la variance des valeurs simulées en ce point?*

b) *Pour la 1ère réalisation, on calcule la variance expérimentale sur l'ensemble du domaine. On répète l'exercice pour la 2e réalisation, la 3e et ainsi de suite. Que vaut la moyenne (calculée sur l'ensemble des réalisations) des variances expérimentales obtenues sur chaque réalisation?*

c) *Pour chaque réalisation, on regroupe les données en blocs de taille V (i.e. on calcule la moyenne des valeurs à l'intérieur du bloc). Pour un bloc donné, quelles seront la moyenne et la variance des teneurs de ce bloc sur l'ensemble des réalisations?*

d) *Pour la 1ère réalisation, on calcule la variance expérimentale des teneurs des blocs de taille V sur l'ensemble du domaine (D). On répète l'exercice pour la 2e réalisation, la 3e et ainsi de suite. Que vaut la moyenne (sur les différentes réalisations) des variances expérimentales ainsi obtenues?*

Question 5 (suite)

On considère cette fois un très grand nombre de réalisations d'une simulation conditionnelle sur une grille très serrée de points.

e) On se place en un point donné x_0 et l'on examine les valeurs obtenues pour les différentes réalisations. Que valent la moyenne et la variance des valeurs simulées en ce point?

f) Pour chaque réalisation, on regroupe les valeurs simulées en blocs de taille V (i.e. on calcule la moyenne des valeurs à l'intérieur du bloc). Pour un bloc donné, quelles seront la moyenne et la variance des teneurs de ce bloc sur l'ensemble des réalisations?

Question 6 (14 points)

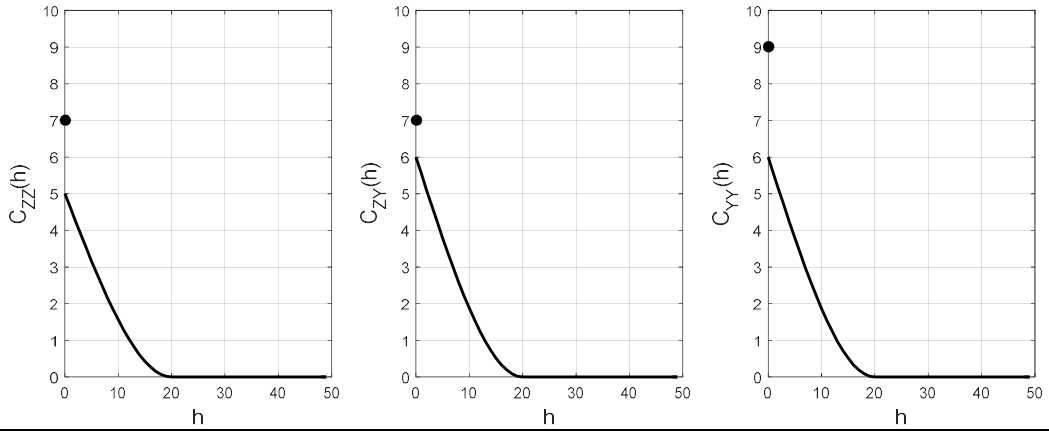
On vous présente différents modèles de covariances simples et croisées pour deux variables selon la direction x.

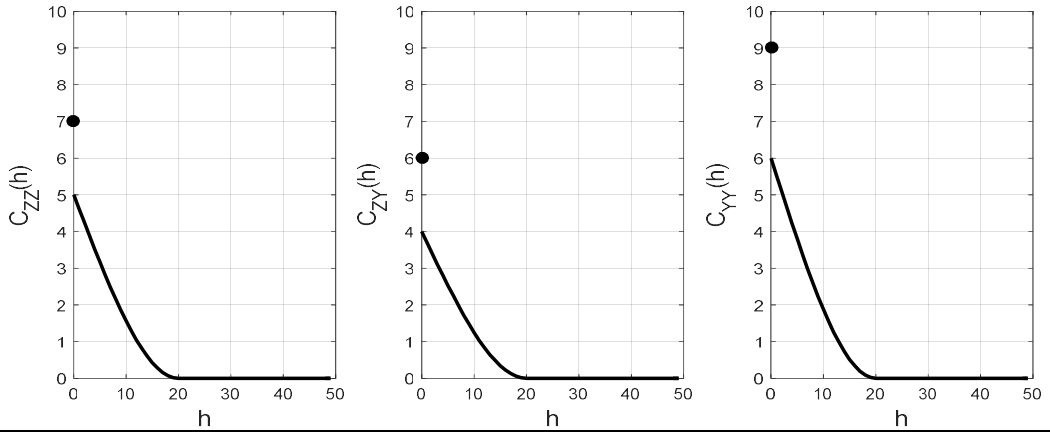
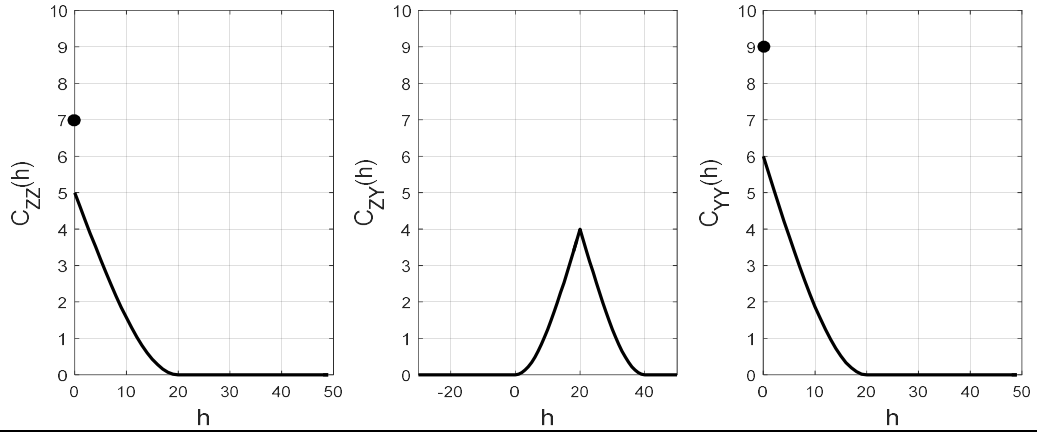
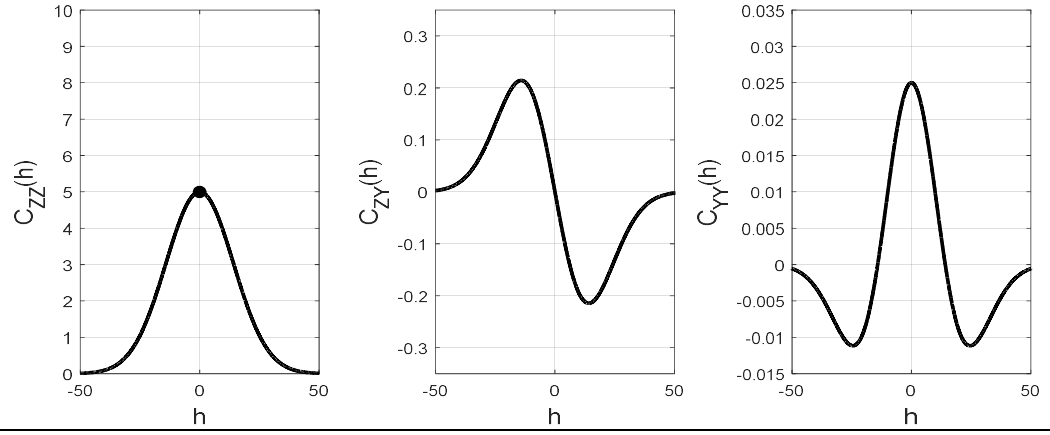
Associez à chaque cas du Tableau 2 l'énoncé parmi les énoncés A à G du Tableau 1 qui décrit le mieux le modèle illustré. Un même énoncé peut décrire plus d'un modèle et un énoncé peut ne décrire aucun modèle. À la fin du tableau 2 vous devez justifier vos choix.

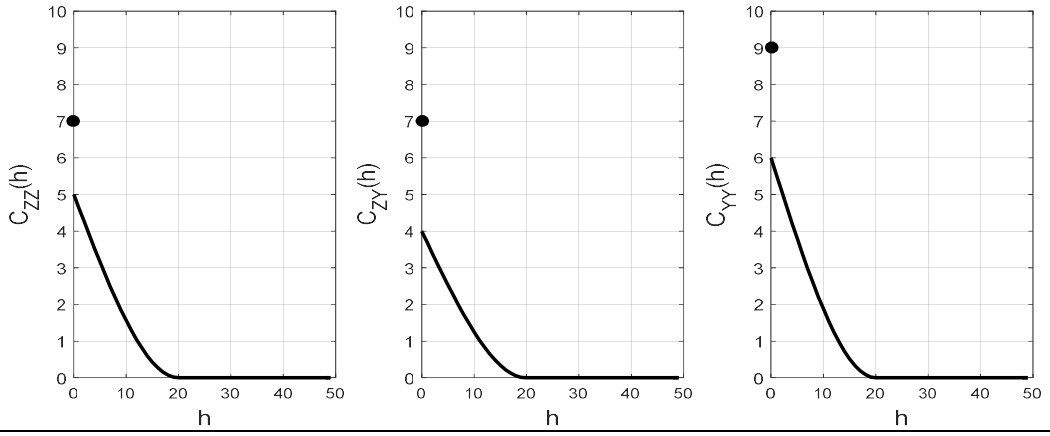
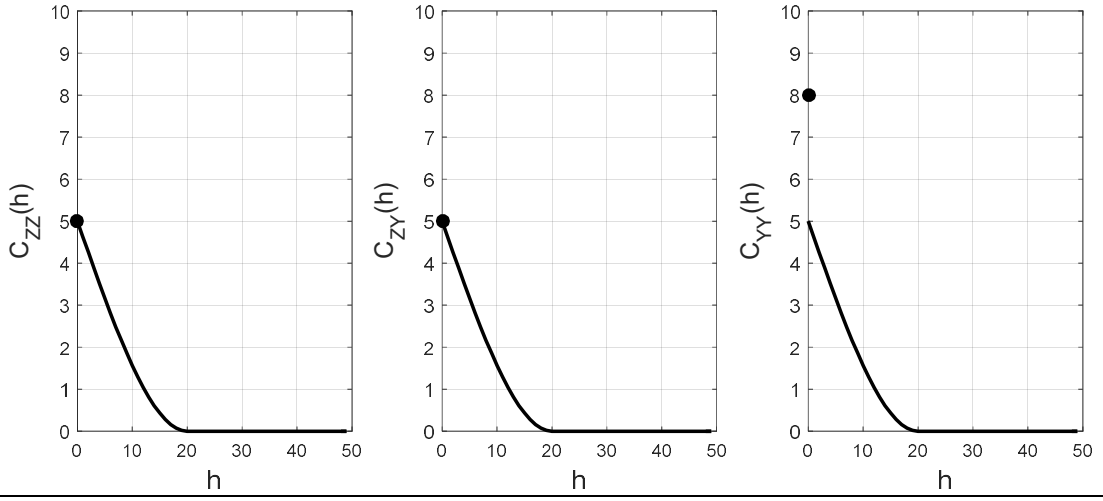
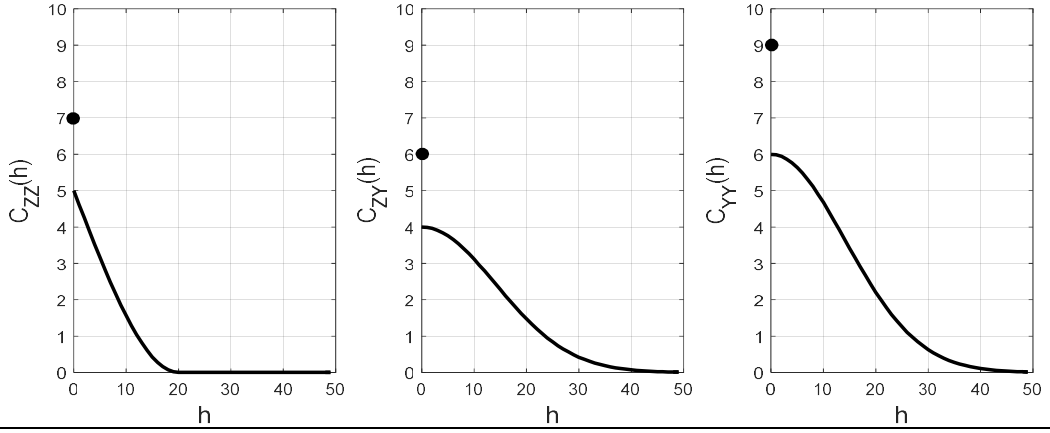
Tableau 1 : énoncés

A	Les variables Y et Z semblent être décalées spatialement l'une par rapport à l'autre.
B	La variable Y semble être corrélée à la dérivée de la variable Z selon la direction x.
C	La variable Y semble être une version bruitée de la variable Z. Z et Y définissent un modèle linéaire de corégionalisation admissible.
D	Il s'agit d'un modèle linéaire de corégionalisation admissible.
E	Il ne s'agit certainement pas d'un modèle linéaire de corégionalisation. On ne sait pas si ce modèle est admissible.
F	Le modèle a les apparences d'un modèle linéaire de corégionalisation mais après vérification il semble qu'il ne soit pas admissible.
G	Le modèle a les apparences d'un modèle linéaire de corégionalisation mais après vérification on est certain qu'il n'est pas admissible.

Tableau 2 : modèles. Pour chaque cas de gauche à droite : C_{zz}, C_{zy}, C_{yy}.

Cas	Modèle	Énoncé
1		

2		
3		
4		

5		
6		
7		

Question 6 (suite)

Justifiez brièvement vos réponses du tableau 2.

Question 7 (10 points)

Le tableau suivant montre les données disponibles dans le voisinage du point x_0 de coordonnées (0,0). Deux variables sont observées, Cu et Zn. On veut estimer par cokrigeage ordinaire le Cu au point x_0. Un tiret indique que la valeur n'est pas disponible.

Point	Coord. X	Coord. Y	Cu(x) (en %)	Zn(x) (en %)
x_1	-20	-20	3.1	2.7
x_2	10	5	2.5	4.3
x_3	20	0	4.1	--
x_4	-10	10	--	2.1
x_5	0	15	2.9	3.6
x_0 (à estimer)	0	0	--	3.3

Le modèle de corégionalisation linéaire est décrit par :

$$\begin{bmatrix} C_{zz}(h) & C_{zy}(h) \\ C_{yz}(h) & C_{yy}(h) \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \delta(h) + \begin{bmatrix} 4 & 4.5 \\ 4.5 & 6 \end{bmatrix} \exp(-|h|/50)$$

où $\delta(h)$ = 1 si h = 0, 0 sinon,

3 pts *a) Est-ce que le modèle obtenu et les données utilisées laissent présager une meilleure performance du cokrigeage ordinaire du Cu au point x_0 par rapport au krigeage ordinaire à ce même point? Justifiez.*

Question 7 (10 points)

7 pts *b) Complétez le système de cokrigeage ordinaire en indiquant les valeurs correspondantes pour chaque lettre de A à N.*

	Cu(x1)	Zn(x1)	Cu(x2)	Zn(x2)	Cu(x3)	Zn(x4)	Cu(x5)	Zn(x5)	Zn(x0)	uz	uy		Cu(x0)
Cu(x1)	G	I	1.83	2.06	1.64	2.39	1.79	2.01	2.56	A	B	λ_1	2.27
Zn(x1)	I	H	2.06	2.75	1.84	3.19	2.01	2.68	3.41	C	D	α_1	2.56
Cu(x2)	1.83	2.06	7.00	5.50	3.20	2.98	3.01	3.39	3.60	E	F	λ_2	3.20
Zn(x2)	2.06	2.75	5.50	8.00	3.60	3.97	3.39	4.52	4.80	0	1	α_2	3.60
Cu(x3)	1.64	1.84	3.20	3.60	7.00	2.39	2.43	2.73	3.02	1	0	λ_3	J
Zn(x4)	2.39	3.19	2.98	3.97	2.39	8.00	3.60	4.80	4.52	0	1	α_4	3.39
Cu(x5)	1.79	2.01	3.01	3.39	2.43	3.60	7.00	5.50	3.33	1	0	λ_5	2.96
Zn(x5)	2.01	2.68	3.39	4.52	2.73	4.80	5.50	8.00	4.44	0	1	α_5	K
Zn(x0)	2.56	3.41	3.60	4.80	3.02	4.52	3.33	4.44	8.00	0	1	α_0	L
uz	A	C	E	0	1	0	1	0	0	0	0	uz	M
uy	B	D	F	1	0	1	0	1	1	0	0	uy	N

Question 8 (15 points)

La figure 1 montre la fonction de répartition obtenue pour l'ensemble d'un gisement de Cu (en %). La figure 2 montre les variogrammes obtenus pour les indicatrices codées par rapport à différents quantiles de la distribution du Cu (10%, 50%, 90%).

Figure 1

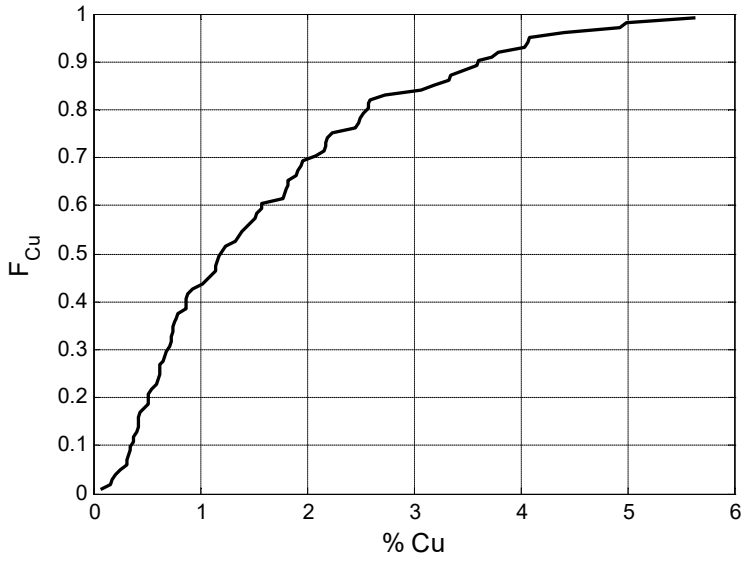
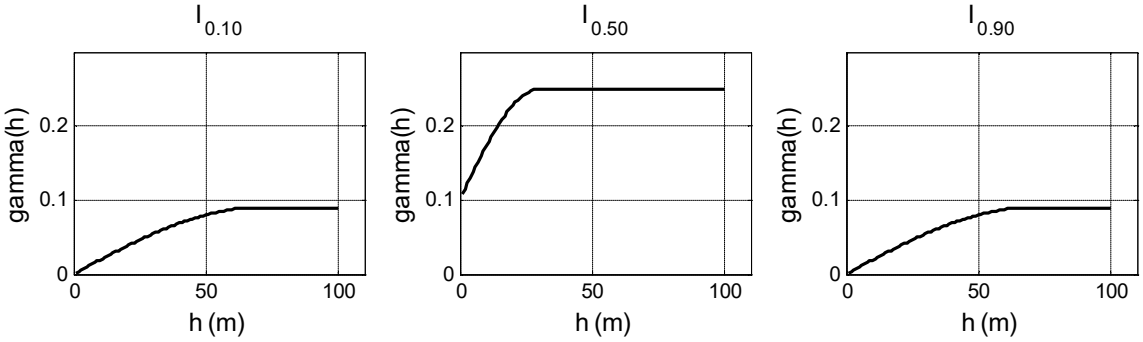


Figure 2



3 pts *a) L'hypothèse d'une distribution multinormale (après transformation vers la loi normale du %Cu) vous semble-t-elle réaliste à première vue ? Justifiez votre réponse.*

Question 8 (suite)

3 pts *b) Quel lien y a-t-il entre les indicatrices codées par rapport aux quantiles 10%, 50% et 90% de la distribution du %Cu et les indicatrices codées par rapport aux mêmes quantiles, mais cette fois de la distribution de Y, où Y est la variable normale obtenue par transformation graphique du %Cu ?* (Note on suppose que toutes les valeurs du %Cu sont distinctes).

On veut effectuer le krigeage simple d'indicatrices au point x_0 en utilisant les points x_1 à x_3. Les poids de krigeage simple obtenus pour différents seuils sont donnés au tableau suivant.

Point	% Cu au point x_i	poids Ks, au seuil 0.5% Cu	poids Ks, au seuil 1.0 % Cu	poids Ks, au seuil 2.0% Cu
x_1	1.2%	0.3	0.2	0.21
x_2	3%	0.45	0.3	0.4
x_3	0.3%	0.15	0.2	0.23

6 pts *c) Estimez par krigeage simple d'indicatrices au point x_0 la fonction de répartition conditionnelle du %Cu.* (Note : Il n'y a pas de correction d'ordre à effectuer).

3 pts d) Selon les résultats en c) quelle est P(%Cu au point $x_0 \leq 1.5\%$) ?

Annexe :

Fonction de répartition de la N(0,1), i.e. $P(Z \leq z)$. L'entier et la 1ère décimale de « z » sont lus en ligne, la 2e décimale est lue en colonne.

Exemple d'utilisation :

i. Trouver p=F(z) à « z » spécifié : F(1.37)=P(N(0,1)<1.37)=0.9147

ii. Trouver z à « p » spécifié : F(z)=0.9904 => z=2.34

Fonction de répartition N(0,1)

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998

Corrigé

Q1
a) environ 1.3
b) $Y(x_0) \sim N(-1.2, (0.3)^2)$
c) On cherche y tel que $P(Y_0<y)=0.9 \rightarrow P(N(0,1)<(y+1.2)/0.3)=0.9$. La table N(0,1) indique que cette valeur est 1.28, donc $y=1.28*0.3-1.2=-0.82$ La valeur de Z correspondante est environ 0.5.
d) $Y(x_0)$
e) Séquentielle d'indicatrice. On garle la même approche qu'en SGS sauf que l'on estime la distribution conditionnelle au point x_0 par un krigeage d'indicatrices. On visite un point à simuler choisi au hasard. On effectue le krigeage d'indicatrice pour obtient la distribution conditionnelle. On tire de cette distribution conditionnelle et on ajoute cette valeur aux données et points déjà simulés, et ainsi de suite jusqu'à ce que tous les points à simuler aient été visités.

Q2
a) Environ 0.12
b) environ 0.42
c) oui, la simulation démontre que l'on multiplierait le profit actuel par un facteur allant jusqu'à 3.5.

Q3
a) Dans l'ordre : 8, 6.7, 9.8, 14.2, 12
b) option A. La fonction de transfert est ici non-linéaire. Il faut donc l'appliquer à chaque réalisation avant de faire la moyenne.
c) $17.6^{0.5} =4.2$

Q4
a) Dans l'ordre : F, C, D, E, A, B
b) On trouve F(1.2)= 0.885 et F(2.2)=0.986. Ceci nous donne F2.
c) L'échantillonneur de Gibbs. On commence par prendre des valeurs de gaussiennes compatibles avec les faciès observés. Ensuite, un très grand nombre de fois, et pour chaque gaussienne séparément ou conjointement (si elles sont corrélées) on visite un point au hasard, on le retire et on l'estime chaque gaussienne par krigeage ou cokrigeage simple. Ceci définit une distribution conditionnelle pour chaque gaussienne. On tire une valeur de cette distribution. Si elle respecte le faciès observé, cette nouvelle valeur ou ces nouvelles valeurs remplacent les anciennes. Sinon, on ne fait rien, i.e. on conserve les anciennes valeurs et l'on passe au point suivant.

Q5
a) m et σ^2
b) $D^2(.|D)$
c) m, σ_V^2
d) $D^2(V|D)$
e) La valeur obtenue par krigeage simple et la variance de krigeage simple.
f) La valeur obtenue par krigeage simple de bloc et la variance de krigeage simple de bloc.

Q6
1-F, 2-D, 3-A, 4-B, 5-G, 6-C, 7-E
Justifications

1-F le modèle est [2 1;1 3] pépite + [5 6;6 6]spher(a=20). La 2e matrice montre un déterminant négatif, donc la condition suffisante n'est pas respectée. Le modèle est probablement non admissible. La réponse on est certain qu'il est non-admissible est aussi acceptée.

2-D le modèle est [2 2;2 3]pépite + [5 4;4 6]spher(a=20). les deux matrices ont un déterminant positif, le modèle est admissible.
3-A La covariance croisée est décalée de 20.
4-B La covariance croisée est anti-symétrique et la covariance simple montre des covariances négatives comme on attend pour la dérivée d'une variable.
5-G le modèle est [2 3;3 3]pépite + [5 4;4 6]spher(a=20). La première matrice a un déterminant négatif. Comme c,est l'effet de pépite, la condition de Cauchy-Schwartz n'est pas respectée et donc on est certain que le modèle n'est pas admissible.
6-C. On a la même structure sur le croisé et la variable secondaire sauf pour une ffet de pépite supplémentaire correspondant au bruit.
7-E. Ce n'est pas un modèle linéaire de corégionalisation car la structure gaussienne n'apparaît pas sur la variable Z. La vérification de l'admissibilité ne peut être faite facilement.

Q7
a) oui car les données Cu et Zn ne sont pas disponibles aux mêmes points. Entre autres Zn est observé au point x_0. De plus Zn montre une corrélation de $5.5/(7*8)^{0.5} = 0.73$, ce qui est assez élevé.
b)
A=1, B=0, C=0, D=1, E=1, F=0, G=7, H=8, I=5.5, J=4exp(-20/50)=2.68, K=4.5exp(-15/50)=3.33, L=5.5, M=1, N=0 .

Q8
a) Non car la médiane apparaît moins structurée que les quantiles plus extrêmes 0.1 et 0.9 (par l'effet de pépite relatif plus élevé et par la portée plus courte. Dans un champ gaussien, c'est à la médiane que la structure devrait être maximale.

b) Elles sont identiques car la transformation monotone ne change pas l'ordre.

c)De la fonction de répartition, on trouve 0.2, 0.44 et 0.7 pour les trois seuils.
Pour 0.5% Cu, les indicatrices sont 0 0 1 => estimé par KS vaut 0.15+ (1-0.9)*0.2 = 0.17
Pour 1% Cu, les indicatrices sont 0 0 1=> estimé vaut 0.2+(1-0.7)*0.44 = 0.33
Pour 2% Cu, les indicatrices sont 1,0,1=> estimé vaut 0.44+ (1-0.84)*0.7 = 0.55

d) Par interpolation linéaire, (0.33+0.55)/2 = 0.44

Examen 2015
Question 1 (15 points)

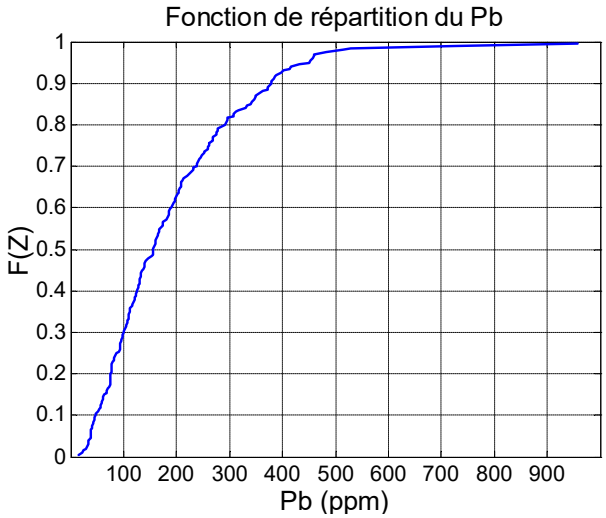
a) Expliquez pourquoi le post-conditionnement par krigeage des simulations non-conditionnelles peut être effectué beaucoup plus rapidement que le SGS sachant que l'on simule dans les deux cas le même nombre de points, et ce, en utilisant les mêmes données.

b) Par recuit simulé, on peut simuler des données dont le variogramme expérimental colle presque parfaitement au modèle souhaité. Expliquez en quoi ceci pourrait être en fait problématique?

c) La méthode SGS permet la simulation de variables continues gaussiennes. Proposez une modification à la méthode SGS qui permette la simulation avec de données non-gaussiennes (comme pour les données de contamination au Pb de Dallas).

Question 2 (12 points)

Un sol de la région de Montréal est contaminé au plomb dû à la présence d'une usine de recyclage de batteries automobiles. Deux cents analyses de la concentration en plomb sont disponibles et la fonction de répartition est illustrée à la figure suivante :



3 pts *a) Quel devrait être approximativement le palier pour l'indicatrice correspondant au seuil 300 ppm?*

Lors de l'estimation d'un point x_0, le krigeage simple de l'indicatrice correspondant au seuil 200ppm retourne les poids suivants :

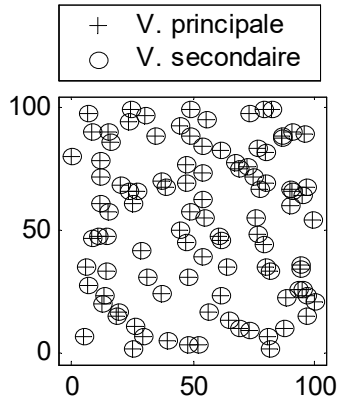
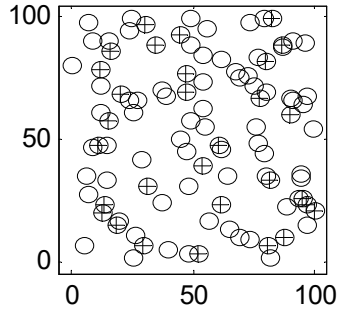
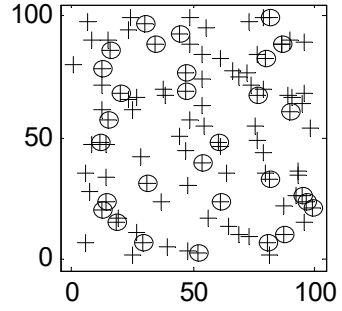
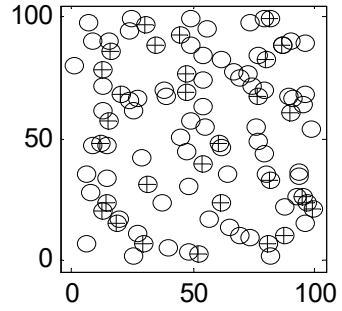
$$\lambda_1 = 0, \lambda_2 = -0.11, \lambda_3 = 0.22, \lambda_4 = 0.70$$

Les valeurs observées en ces points sont respectivement (en ppm) : Z_1=127, Z_2= 210, Z_3=185, Z_4=233

9 pts *b) Quelle est la probabilité (estimée par KI simple) que la concentration au point x_o soit supérieure à 200 ppm ?*

Question 3 (15 points)

Le tableau suivant présente, dans la colonne de gauche, le plan de localisation d'observations d'une variable principale (Z) et d'une variable secondaire (Y). Dans la colonne de droite, on donne le modèle linéaire de corégionalisation ajusté à ces données. Dans les matrices, Z apparaît en premier, Y en second. $\delta(h)$ représente la covariance effet de pépite.

A) 	$C(h)=\begin{bmatrix}1 & 0.7\\ 0.7 & 0.9\end{bmatrix}\delta(h)+\begin{bmatrix}8 & 8\\ 8 & 9\end{bmatrix}Sph(a=50, C=1)$
B) 	$C(h)=\begin{bmatrix}1 & 0.7\\ 0.7 & 0.9\end{bmatrix}\delta(h)+\begin{bmatrix}8 & 8\\ 8 & 9\end{bmatrix}Sph(a=50, C=1)$
C) 	$C(h)=\begin{bmatrix}1 & 0.7\\ 0.7 & 0.9\end{bmatrix}\delta(h)+\begin{bmatrix}8 & 8\\ 8 & 9\end{bmatrix}Sph(a=50, C=1)$
D) 	$C(h)=\begin{bmatrix}1 & 0.7\\ 0.7 & 0.9\end{bmatrix}\delta(h)+\begin{bmatrix}8 & 0\\ 0 & 9\end{bmatrix}Sph(a=50, C=1)$

Question 3 (suite)

12 pts *a) Indiquez pour chaque cas A) à D) si le cokrigeage de Z(x) (par Z et Y) est susceptible d'améliorer de façon importante la précision des estimations par rapport au seul krigeage de Z(x). Justifiez vos réponses.*

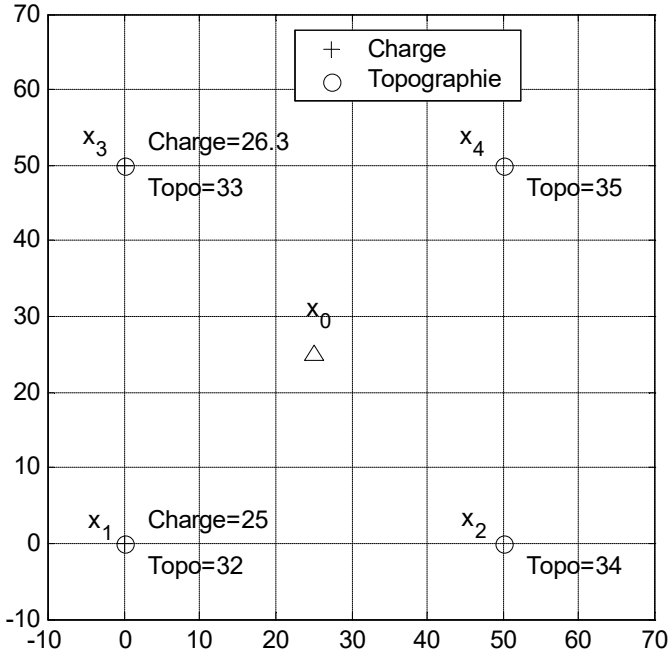
3 pts *b) Indiquez quelle méthode vous permettrait, avec les mêmes données, de confirmer ou d'infirmer vos réponses en a) ?*

Question 4 (18 points)

On désire estimer la charge hydraulique d'un aquifère à nappe libre. On observe une excellente corrélation entre la charge hydraulique et la topographie. Le modèle linéaire de corégionalisation pour ces deux variables (dans l'ordre : charge, topographie) est :

$$C(h) = \begin{bmatrix} 1 & 0.8 \\ 0.8 & 1.2 \end{bmatrix} \delta(h) + \begin{bmatrix} 7 & 9 \\ 9 & 12 \end{bmatrix} Sph(a = 100m, C = 1)$$

La figure suivante (coordonnées en m) montre les quatre points ou les données (charge et/ou topographie, les deux en m) sont disponibles pour l'estimation de la charge au point x_0.



3 pts *a) Quelle est la corrélation (à h=0) entre charge et topographie selon le modèle de corégionalisation?*

Question 4 (suite)

10 pts *b) Construisez (sans le résoudre) le système de cokrigeage ordinaire pour l'estimation de la charge hydraulique au point x_0. Identifiez clairement les entrées de la matrice et des vecteurs. Le tableau suivant pourrait vous éviter des calculs :*

Distance (m)	0	$25*2^{0.5}$	50	$50*2^{0.5}$
C(h), Sph(a=100m, C=1)	1	0.4918	0.3125	0.1161

Les poids de cokrigeage et les multiplicateurs de Lagrange obtenus pour la question b) sont :

$$\lambda_{1,ch\,arg\,e} = \lambda_{3,ch\,arg\,e} = 1/2 \quad \lambda_{1,topo} = \lambda_{3,topo} = -0.19 \quad \lambda_{2,topo} = \lambda_{4,topo} = 0.19 \quad \mu_{ch\,arg\,e} = -0.03 \quad \mu_{topo} = 0.31$$

5 pts *c) Quelle est la valeur estimée et la variance de cokrigeage pour la charge au point x_0?*

Question 5 (12 points)

On effectue une simulation de moyenne 0 par méthode de Choleski. La matrice L obtenue est :

$$L = \begin{bmatrix} 7.75 & & & & \\ 3.82 & 6.74 & & & \\ 4.59 & 3.23 & 5.34 & & \\ 4.36 & 1.45 & 2.20 & 5.84 & \\ 4.77 & 1.79 & 2.42 & 2.86 & 4.48 \end{bmatrix}$$

3 pts *a) Quelle est la covariance entre les 2^e et 3^e variables aléatoires?*

3 pts *b) On a observé Z_1=5. Que vaut Y_1?*

Question 5 (suite)

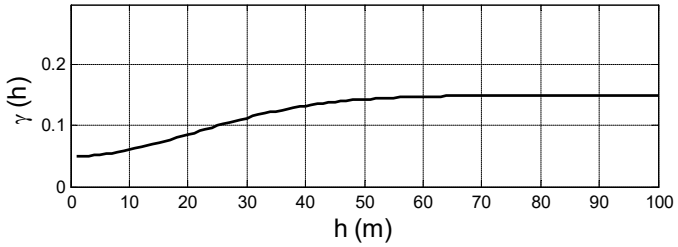
3 pts *c) Quelle est l'espérance conditionnelle de Z_2 sachant que $Z_1=5$?*

3 pts *d) En pratique, quelle est la principale limitation touchant la méthode de Choleski ?*

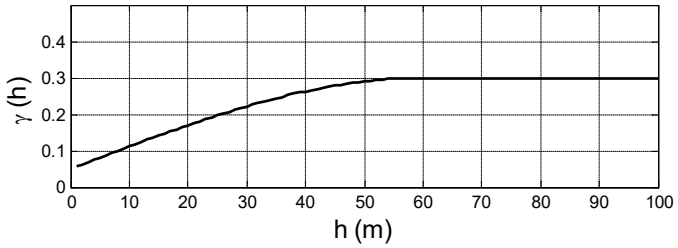
Question 6 (10 points)

La figure suivante montre des modèles de variogramme sensés être ajustés aux variogrammes des <u>indicatrices</u> correspondant à des seuils A et B pour la teneur en Pb (mesurée en ppm) du site de Dallas.

A



B

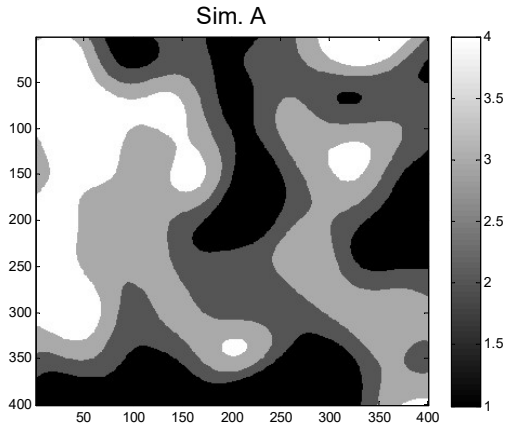


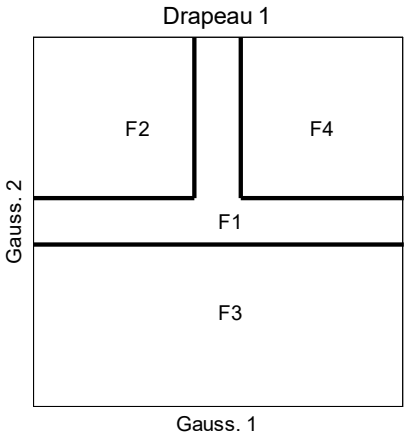
8 pts *a) Quelles erreurs flagrantes ont été commises dans l'ajustement des variogrammes (justifier)?*

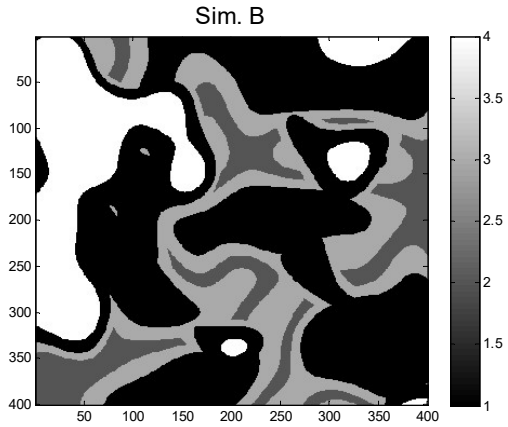
2 pts *b) Quelles sont les unités des variogrammes illustrés?*

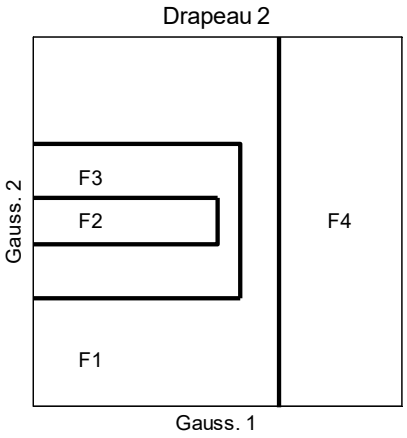
Question 7 (18 points)

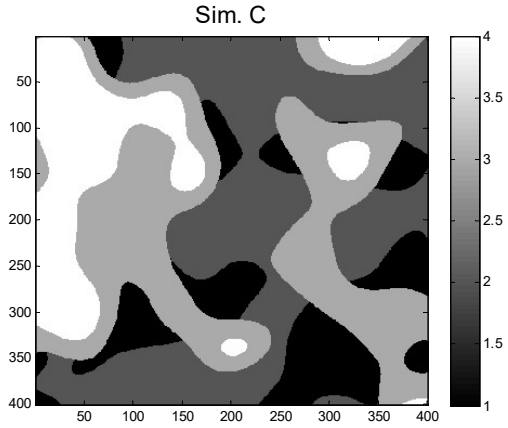
Les figures suivantes montrent 5 différentes simulations plurigaussiennes (A à E) et 5 différents drapeaux de codage (1 à 5) illustrés dans l'espace gaussien (entre -2 et 2). Dans chaque simulation, on retrouve 4 faciès représentés par des niveaux de gris différents (F1=1, F2=2, F3=3, F4=4). Les drapeaux ont été placés dans un ordre **quelconque**. On désire associer le bon drapeau à chaque image simulée.

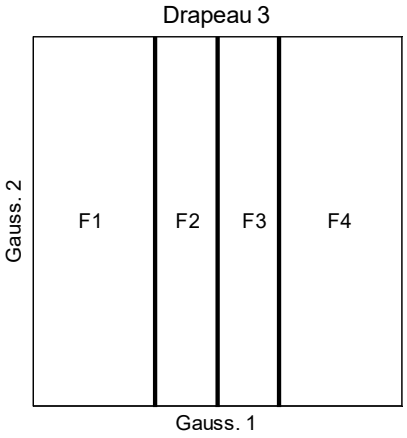




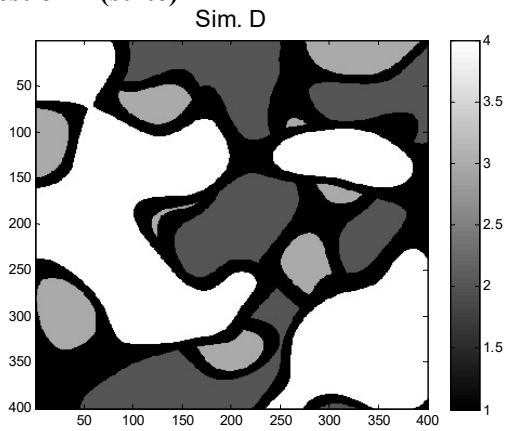


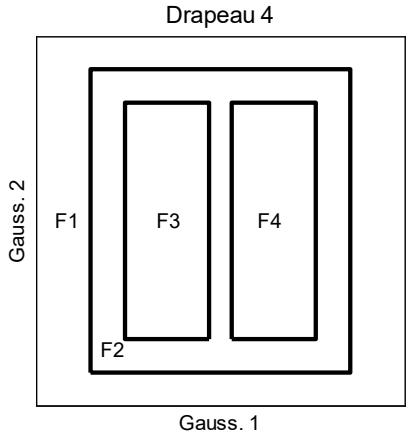


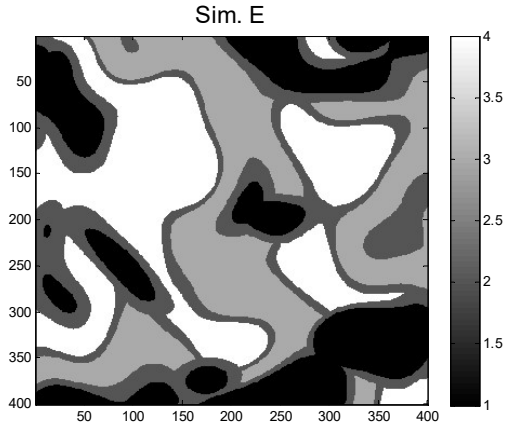


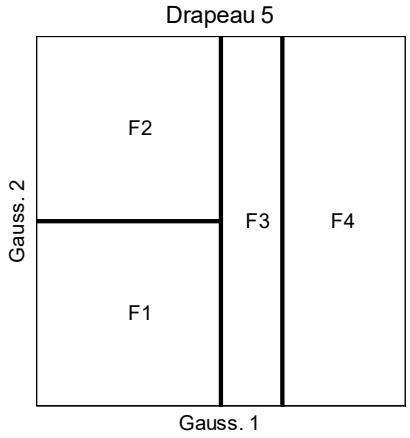


Question 7 (suite)









10 pts *a) Associez à chaque simulation le drapeau de codage (de 1à 5) correspondant. Indiquez « aucun » si aucun drapeau de codage ne correspond selon vous à la simulation considérée.*

Simulation	A	B	C	D	E
Drapeau					

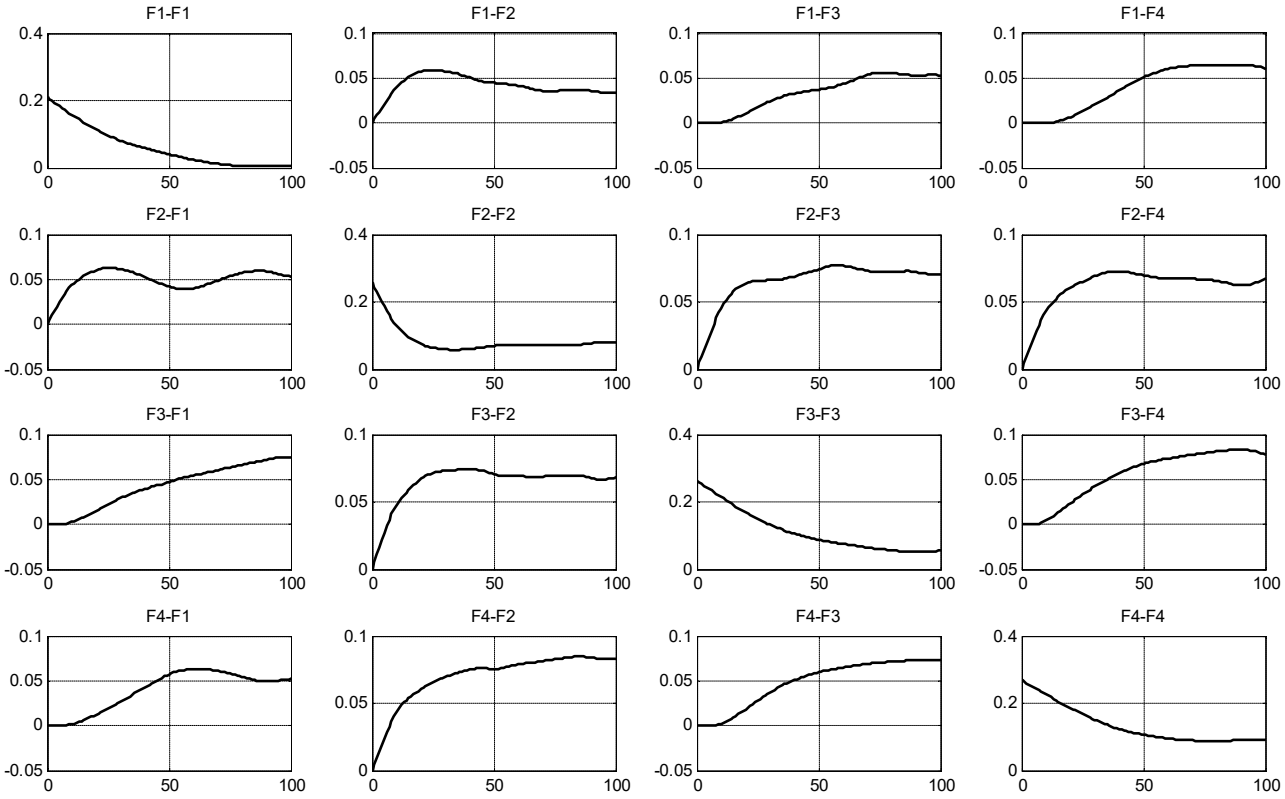
2 pts *b) Une seule simulation de deux variables gaussiennes a été utilisée pour produire les simulations de faciès précédentes. Le modèle de variogramme utilisé pour les deux variables est le même. Identifiez le modèle utilisé parmi les modèles suivants (justifiez votre choix) :*

i. Gaussien isotrope avec portée de 100 pixels
ii. Sphérique isotrope avec portée de 100 pixels
iii. Gaussien anisotrope avec portées 100 horizontalement et 10 verticalement
iv. Sphérique anisotrope avec portées 100 horizontalement et 10 verticalement

Question 7 (suite)

3 pts c) Identifiez la simulation (parmi A à E) qui aurait pu être obtenue par la méthode gaussienne tronquée plutôt que par la méthode plurigaussienne?

Les covariances non centrées expérimentales des indicatrices de faciès sont calculées pour l'une des simulations A à E. La figure suivante montre le résultat obtenu selon la direction verticale (positive vers le bas). On a considéré séparément Ii(x) Ij(x+h) et Ij(x) Ii(x+h).



3 pts d) Identifiez la simulation (parmi A à E) ayant fourni ces covariances non-centrées d'indicatrices. Justifiez votre choix.

Corrigé

Question 1

a) le post-condtionnement par krigeage est effectué avec toujours les mêmes données, si le nombre n'est pas trop grand, on peut faire le krigeage globalement et donc inverser la matrice de krigeage une seule fois. Avec SGS, le nombre de données inclut le nombre de points déjà simulés. Ce nombre normalement devient trop grand pour permettre l'inversion d'un seul coup.

b) Comme on l'a vu au cours, on doit s'attendre à voir des fluctuations du variogramme expérimental par rapport au modèle. Ne pas voir de fluctuations implique que statistiquement quelque chose n'est pas respecté.

c) au lieu de faire le KS, on effectue le krigeage d'indicatrices (seuils) et l'on tire une valeur de la distribution conditionnelle ainsi obtenue. On ajoute cette valeur aux données simulées et l'on poursuit.

Question 2

a) Au seuil 300 ppm la fonction de répartition vaut approximativement 0.82. Le palier devrait donc être assez près de 0.82*0.18=0.15.

b) On code les valeurs par rapport au seuil 200, on a $I_1=1$, $I_2=0$, $I_3=1$, $I_4=0$. La somme des λ donne 0.81. La valeur estimée vaut donc 0.22+ (1-0.81)*0.63=0.34. $P(Z_0<200)=0.34$, donc $P(Z_0>200)=1-0.34=0.66$.

Question 3

4- A) Même si la corrélation entre les 2 variables est forte, le fait que la variable secondaire soit connue aux mêmes emplacements devrait faire que le gain de précision sera assez limité.
B) La variable secondaire est plus abondante et l'on a une forte corrélation selon le modèle (0.92). Le cokrigeage devrait améliorer beaucoup la précision des prédictions.
C) La variable principale est connue en plus d'endroits que la variable secondaire, le cokrigeage apportera très peu;
D) La disposition est comme en B) mais ici la corrélation est seulement de 0.07, ce qui indique que la variable secondaire n'est pas utile pour déterminer la variable principale.

b) En effectuant une validation croisée où l'on enlèverait les deux informations (var. principale et var. secondaire) au point d'estimation, on pourrait apprécier le gain procuré par le cokrigeage.

Question 4

6- a) corrélation : $9.8/(8*13.2)^{0.5}=0.95$.
b) La matrice de cokrigeage est :

	Z_1	Y_1	Y_2	Z_3	Y_3	Y_4		
Z_1	8	9.8	2.8125	2.1875	2.8125	1.045	1	0
Y_1	9.8	13.2	3.75	2.8125	3.75	1.3934	0	1
Y_2	2.8125	3.75	13.2	1.045	1.3934	3.75	0	1
Z_3	2.1875	2.8125	1.045	8	9.8	2.8125	1	0
Y_3	2.8125	3.75	1.3934	9.8	13.2	3.75	0	1
Y_4	1.045	1.3934	3.75	2.8125	3.75	13.2	0	1
	1	0	0	1	0	0	0	0
	0	1	1	0	1	1	0	0

Le vecteur de droite est :

	Z_0
Z_1	3.4424
Y_1	4.4259
Y_2	4.4259
Z_3	3.4424
Y_3	4.4259
Y_4	4.4259
	1
	0

c) valeur estimée : (25+26.3)/2 –0.19*(33+32)+0.19*(34+35)= 26.41 m
variance de krigeage : 8-(3.44*.5-.19*.4.43+.19*4.43+.5*3.44-.19*4.43+.19*4.43-.03)= 4.59

Question 5

a) on multiplie la 2e ligne de L avec la 3e colonne de L' soit : 3.82*4.59+6.74*3.23=39.30

b) Y_1=5/7.75=0.65

c) 3.82*0.65=2.48

d) On ne peut simuler de grands champs pour des raisons d'espace mémoire. La matrice de covariance est en effet de taille n x n, où n est le nombre de points à simuler.

Question 6

a) erreurs flagrantes : A : pour une variable indicatrice, l'on doit ajuster un modèle avec comportement linéaire à l'origine, jamais un gaussien
B : le palier maximal ne devrait pas dépasser 0.25 en théorie.

b) sans unité puisque ce sont des variogrammes d'indicatrice.

Question7
a) Associez à chaque simulation le drapeau de codage correspondant.

Simulation	A	B	C	D	E
Drapeau	*3*	*2*	*5*	*1*	*4*

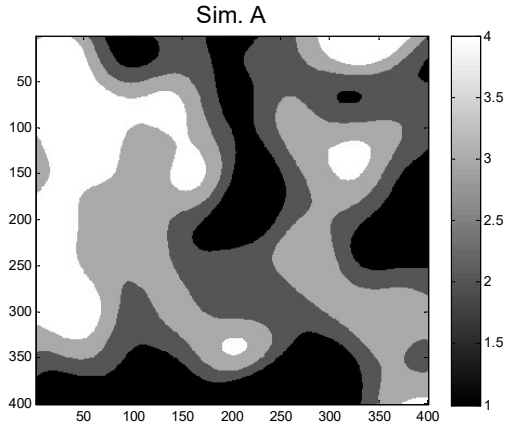
b) Gaussien isotrope. On ne voit aucun étirement indicateur d'anisotropie et la texture est clairement celle d'un gaussien en raison de la régularité des contours des facies.

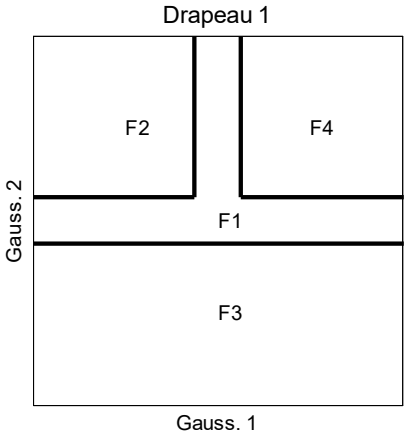
c) Celle correspondant au drapeau 3, donc A.

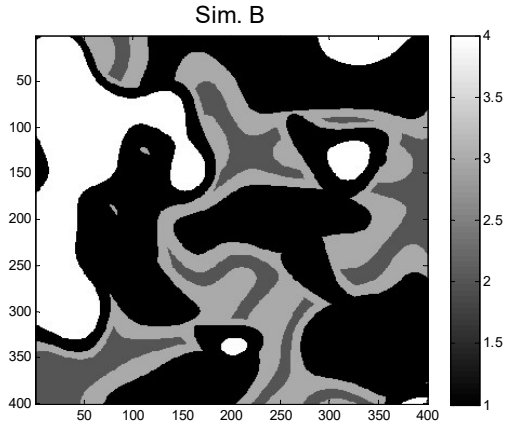
d) Pas de contact F3-F4, F1-F3 et F1-F4 => simulation E (drapeau 4)

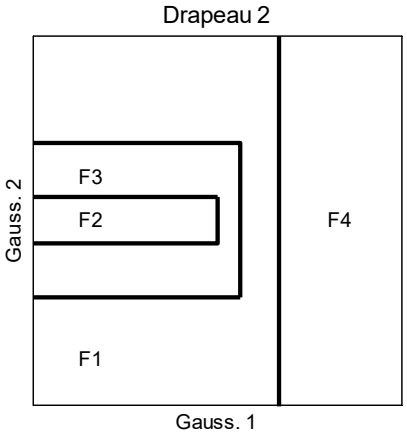
Examen 2014
Question 1 (12 points)

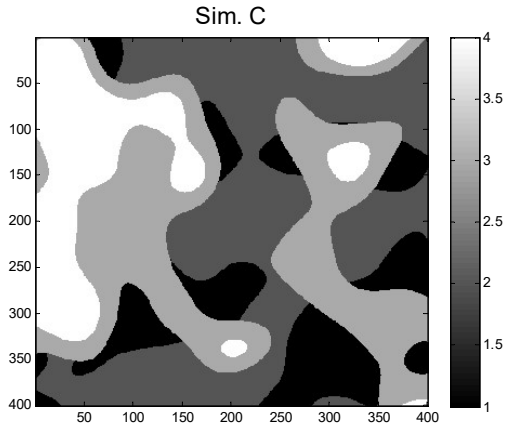
Les figures suivantes montrent 5 différentes simulations plurigaussiennes (A à E) et 5 différents drapeaux de codage (1 à 5) placés dans un ordre quelconque. Dans chaque simulation, on retrouve 4 faciès représentés par des niveaux de gris différents (F1=1, F2=2, F3=3, F4=4). On désire associer un drapeau à chaque image simulée.

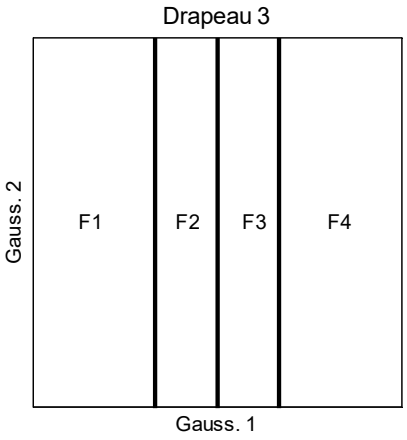




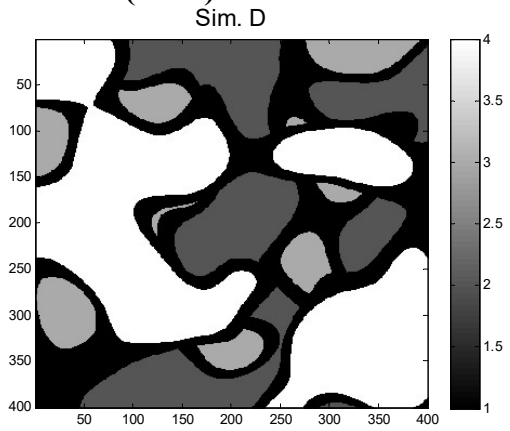


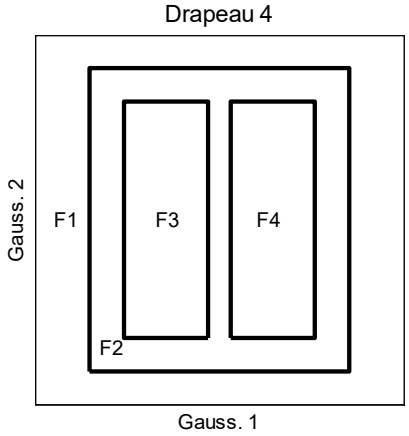


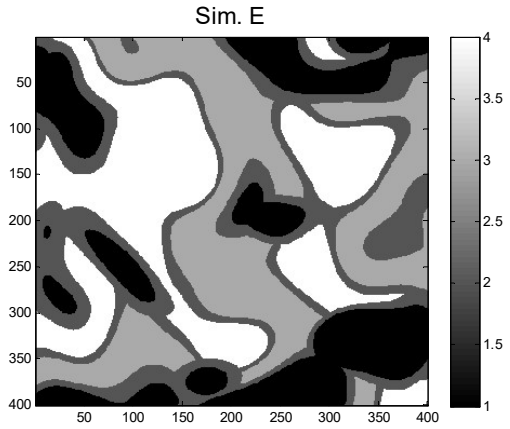


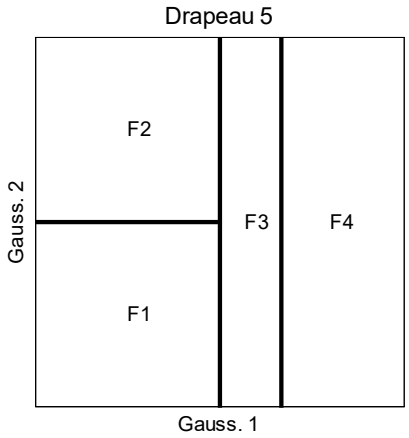


Question 1 (suite)









10 pts *a) Associez à chaque simulation le drapeau de codage correspondant. Indiquez « aucun » si aucun drapeau de codage ne correspond à la simulation considérée.*

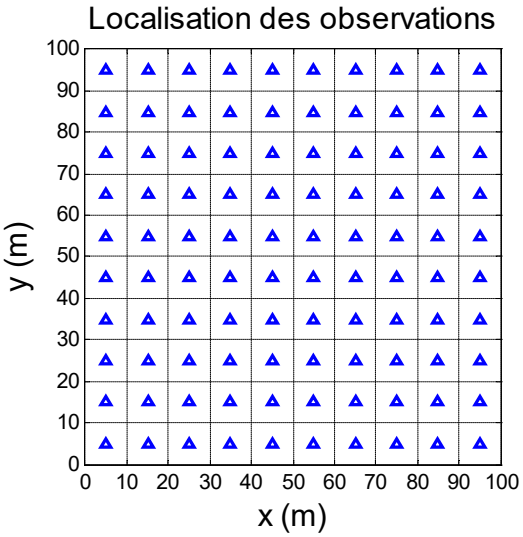
Simulation	A	B	C	D	E
Drapeau					

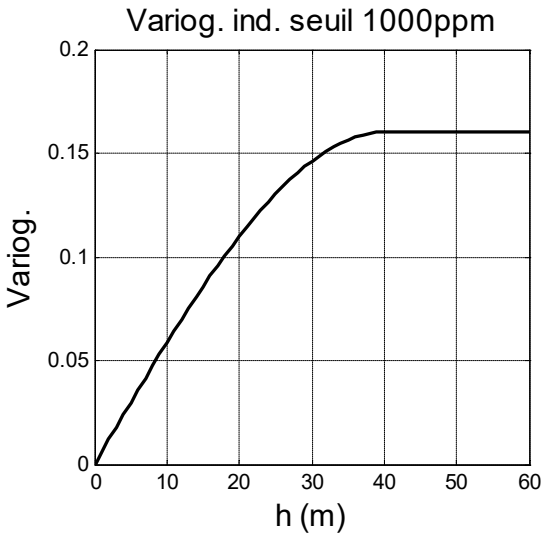
2 pts *b) Une seule simulation de deux variables gaussiennes a été utilisée pour produire les simulations de faciès précédentes. Le variogramme utilisé pour les deux variables est le même. Identifiez le modèle utilisé parmi les modèles suivants (justifiez votre choix) :*

i. Gaussien isotrope avec portée de 100 pixels
ii. Sphérique isotrope avec portée de 100 pixels
iii. Gaussien anisotrope avec portées 100 horizontalement et 10 verticalement
iv. Sphérique anisotrope avec portées 100 horizontalement et 10 verticalement

Question 2 (12 points)

À l'aide d'une grille régulière de points, on veut estimer la proportion de sols contaminés aux hydrocarbures sur la zone indiquée à la figure suivante. Un sol est considéré contaminé si sa valeur est supérieure à 1000 ppm. La figure montre aussi le variogramme (isotrope) de l'indicatrice définie au seuil 1000 ppm d'hydrocarbures.





3 pts *a) Quel est le modèle de variogramme de l'indicatrice ?*

3 pts *b) Considérant seulement le variogramme de l'indicatrice, approximativement quelle proportion de l'ensemble de la zone étudiée est contaminée (aide : deux réponses sont possibles) ?*

6 pts *c) Trouvez l'écart-type de l'erreur d'estimation de la proportion de sols contaminés sur l'ensemble de la zone (un abaque est fourni en annexe).*

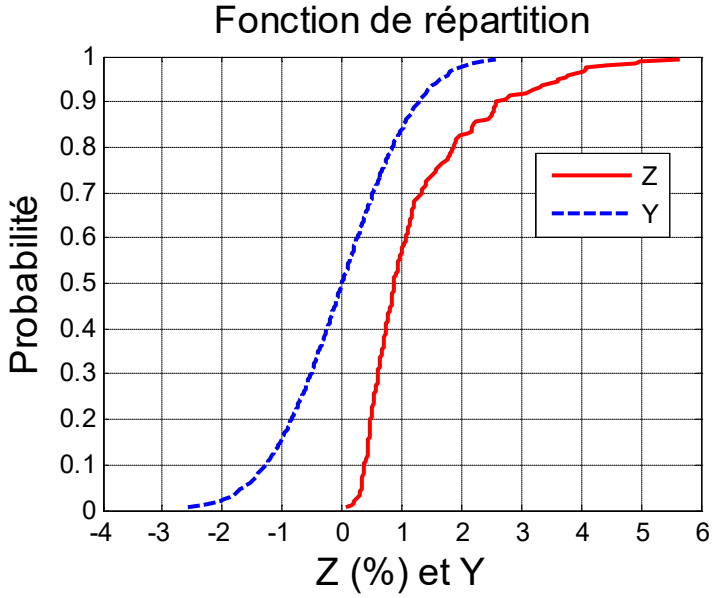
Question 3 (14 points)

La méthode de simulation SGS consiste à :

iv. choisir un point à simuler;
v. estimer par krigeage simple la distribution conditionnelle de la variable en ce point compte tenu des observations connues;
vi. tirer aléatoirement une valeur de la distribution conditionnelle et ajouter cette valeur à l'ensemble des données observées et des données déjà simulées.

3 pts *a) Expliquez pourquoi cet algorithme assure que chaque réalisation respectera les données observées aux points échantillons.*

Lorsque les données (Z) ne suivent pas une distribution gaussienne, on doit au préalable les transformer (Y) pour les rendre gaussiennes. La figure suivante montre graphiquement une telle transformation.



Question 3 (suite)

3 pts *b) Une donnée valant Z(x)=3% aurait quelle valeur gaussienne après transformation?*

5 pts *c) On applique l'algorithme. À un point donné, le krigeage simple retourne la valeur -1.2 avec un écart-type de krigeage de 0.3. Le générateur de nombres aléatoires distribués suivant une loi uniforme (0,1) retourne la valeur 0.9. Que vaut Z(x) (simulé) à ce point ?*

3 pts *d) Suggérez une approche qui permettrait d'utiliser le principe séquentiel du SGS mais sans devoir transformer vers la loi normale et supposer la distribution multigaussienne de la variable transformée.*

Question 4 (15 points)

On vous fournit le tableau suivant donnant, le long d'un profil, les coordonnées, les valeurs observées, les valeurs d'une simulation non-conditionnelle, les valeurs krigées avec d'une part, les données observées et d'autre part, les valeurs simulées.

Coordonnée	Valeurs observées	Simulation non-conditionnelle	Valeur krigée avec les données	Valeurs krigées avec les valeurs simulées aux points des données	Simulation conditionnelle *(à calculer)*
0	10	6	10	6	
100	-	7.5	9	8	
200	-	7	8	10	
300	-	10	7	12	
400	6	14	6	14	

6 pts *a) Transformez par post-conditionnement la simulation non-conditionnelle en une simulation conditionnelle (inscrivez vos réponses dans le tableau)*

3 pts *b) Le signal simulé représente la topographie du fond marin. Supposons que l'on veuille déterminer la pente moyenne susceptible d'être rencontrée le long du profil. Est-il préférable de :*

C) générer plusieurs réalisations, calculer la pente moyenne pour chaque réalisation et faire la moyenne sur toutes les réalisations de ces pentes moyennes

ou

D) générer plusieurs réalisations, calculer la moyenne des réalisations et calculer la pente moyenne sur cette « réalisation » moyenne.

Justifier.

Question 4 (suite)

3 pts *c) L'on effectue un grand nombre de réalisations et l'on calcule la moyenne des valeurs conditionnelles simulées au point x=100. Quelle devrait être cette valeur moyenne ?*

3 pts *d) Le modèle de variogramme utilisé pour la simulation est sphérique avec C=25 et a=30. La variance de krigeage au point x=100 vaut 12. Quel devrait être l'écart-type des réalisations au point x=100?*

Question 5 (12 points)

On simule par la méthode de Cholesky Z_2 et Z_3 conditionnellement à la valeur Z_1=2 et suivant un variogramme donné. La matrice triangulaire inférieure issue de la décomposition de la matrice de covariance est (les variables sont placées dans l'ordre Z_1, Z_2 et Z_3) :

$$L = \begin{bmatrix} 2.236 & 0 & 0 \\ 1.342 & 1.789 & 0 \\ 0.447 & 1.901 & 1.09 \end{bmatrix}$$

3 pts *a) Quel est le palier du variogramme simulé ?*

3 pts *b) Que vaut la valeur Y_1 ?*

3 pts *c) On tire Y_2 = -1 et Y_3=1.5. Quelles sera la valeur simulée pour Z_2 ?*

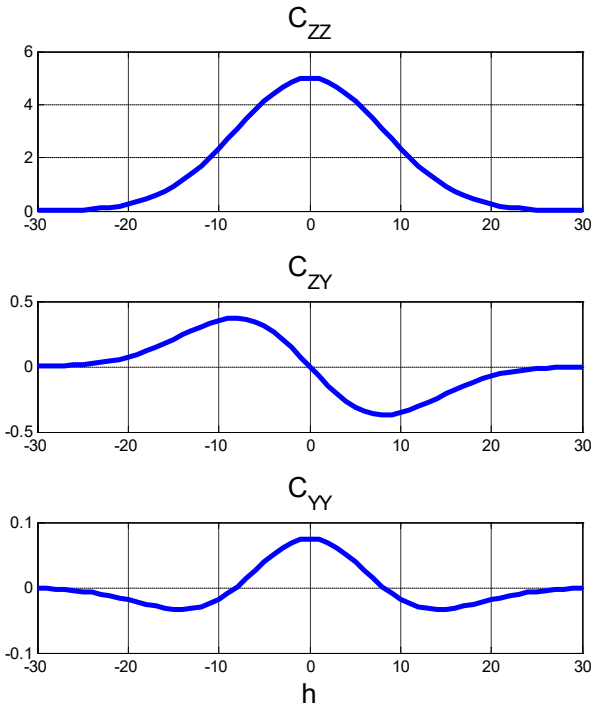
3 pts *d) Quelle est la covariance théorique entre Z_2 et Z_3, conditionnelle à Z_1=2 ? (Aide : Y_1 n'est pas aléatoire; exprimez Z_2 et Z_3 en fonction de Y_2 et Y_3; en déduire la covariance recherchée)*

Question 6 (15 points)

Le long d'un profil (axe « x »), on veut estimer l'élévation du sommet Z(x) d'une formation géologique. Soit la dérivée Y(x)=dZ(x)/dx décrivant l'accroissement de l'élévation en se déplaçant le long du profil. La covariance de Z est un modèle gaussien de portée effective 20 et de palier C=5. On a alors :

$$C_{ZZ}(h) = 5\exp(-3h^2/400)$$
$$C_{ZY}(h) = -3/40\ h\ \exp(-3h^2/400)$$
$$C_{YY}(h) = 3/40\ \exp(-3h^2/400) - 9/8000\ h^2\ \exp(-3h^2/400)$$

Ces trois fonctions de covariance sont présentées à la figure suivante :



Le tableau suivant fournit les valeurs des fonctions de covariance pour quelques valeurs particulières de « h ».

	h						
	-10	-7	-3	0	3	7	10
$C_{ZZ}(h)$	2.362	3.462	4.674	5	4.674	3.462	2.362
$C_{ZY}(h)$	0.354	0.364	0.210	0	-0.210	-0.364	-0.354
$C_{YY}(h)$	-0.018	0.014	0.061	0.075	0.061	0.014	-0.018

Question 6 (suite)

8 pts *a) On connaît la valeur de Z et de sa dérivée Y aux points $x_1=0$ et $x_2=10$. On désire effectuer l'estimation, par cokrigeage simple, au point $x_0=3$. Construisez le système de cokrigeage simple sous forme matricielle (ne pas résoudre). Indiquez clairement la signification des colonnes et lignes de la matrice de gauche et du vecteur de droite de ce système.*

Question 6 (suite)

3 pts *b) Est-ce un contexte où l'on s'attend à ce que le cokrigeage permette d'améliorer sensiblement l'estimation? Justifier.*

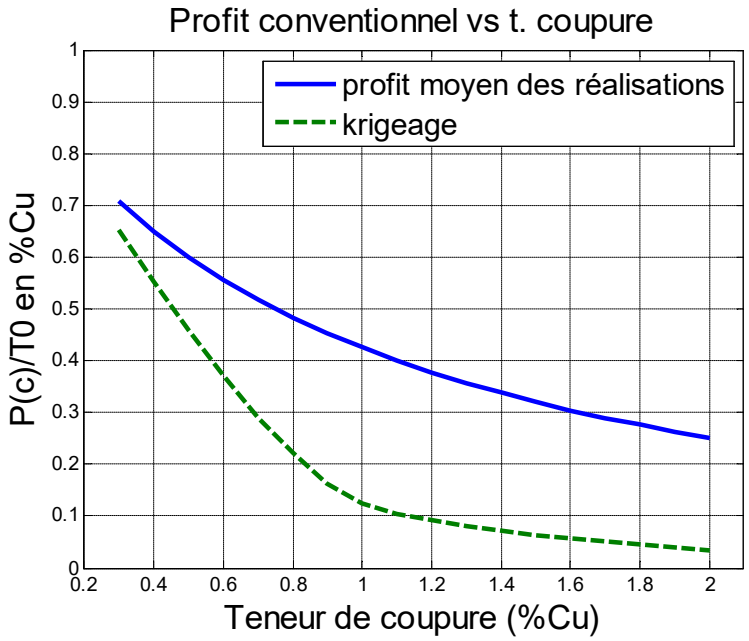
On a résolu le système de krigeage simple et celui de cokrigeage simple. On a obtenu :

$$\text{KS}: \lambda_1 = 0.782 \quad \lambda_2 = 0.323$$
$$\text{CO_KS}: \lambda_{1,Z} = 0.789 \quad \lambda_{2,Z} = 0.200 \quad \alpha_{1,Y} = 1.691 \quad \alpha_{2,Y} = -0.719$$

4 pts *c) Calculez la variance de krigeage simple et la variance de cokrigeage simple. Discutez.*

Question 7 (10 points)

Soit la figure suivante montrant le profit conventionnel, en fonction de la teneur de coupure, obtenu par krigeage et aussi le profit obtenu, en moyenne, pour les différentes réalisations d'une simulation.



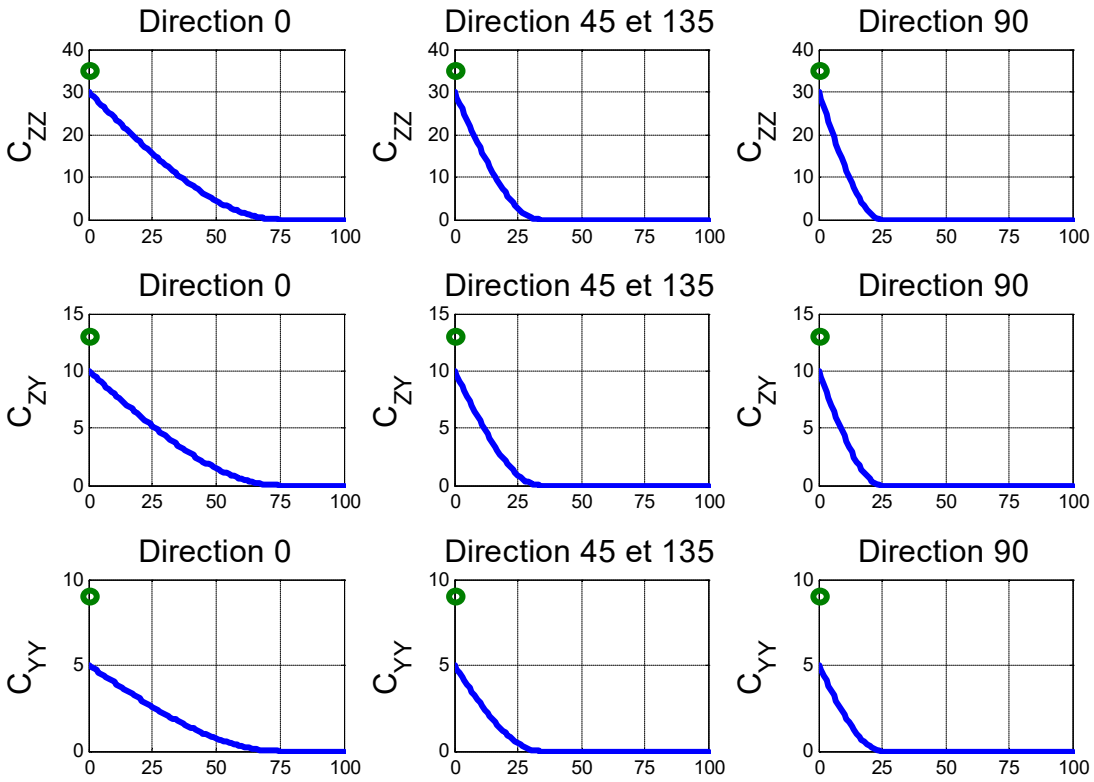
4 pts *a) Si l'on sélectionnait les blocs à partir des estimés actuels de krigeage, quel profit conventionnel obtiendrait-on approximativement à une teneur de coupure de 1% ?*

4 pts *b) Si l'on connaissait parfaitement les teneurs réelles des blocs, quel profit pourrait-on espérer obtenir à la teneur de coupure 1% ?*

2 pts *c) Selon les courbes présentées, devrait-on cherche à améliorer l'estimation des blocs si la teneur de coupure est 1% ? Justifier.*

Question 8 (10 points)

Soit le modèle linéaire de corégionalisation indiqué aux figures suivantes.



Décrivez entièrement le modèle linéaire de corégionalisation (en 2D) et déterminez s'il est admissible.

Annexe :

Fonction de répartition de la N(0,1), i.e. $P(Z \leq z)$. L'entier et la 1ère décimale de « z » sont lus en ligne, la 2e décimale est lue en colonne.

Exemple d'utilisation :

i. Trouver p=F(z) à « z » spécifié : F(1.37)=P(N(0,1)<1.37)=0.9147

ii. Trouver z à « p » spécifié : F(z)=0.9904 => z=2.34

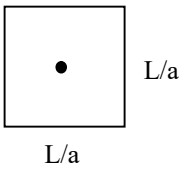
Fonction de répartition N(0,1)

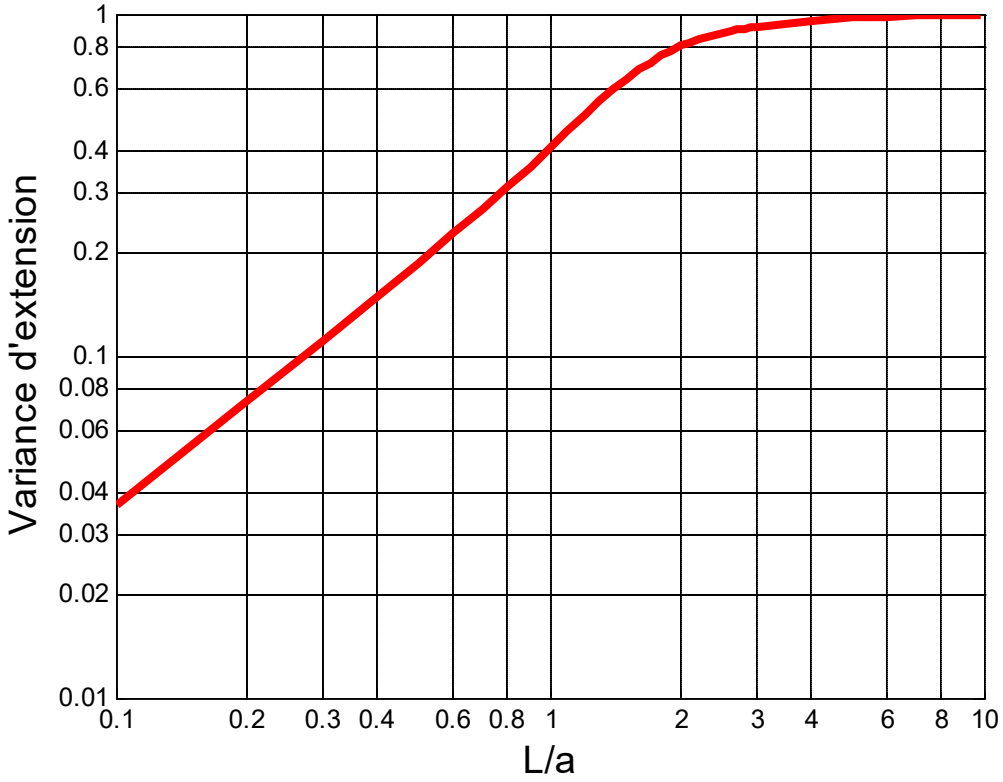
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998

Annexe :

Abaque

Variance d'extension d'un point à un carré, modèle sphérique avec C=1 (abaque 7, courbe E3)





Corrigé

Q 1*a) Associez à chaque simulation le drapeau de codage correspondant.*

Simulation	A	B	C	D	E
Drapeau	*3*	*2*	*5*	*1*	*4*

b) Gaussien isotrope. On ne voit aucun étirement indicateur d,anisotropie et la texture est clairement celle d'un gaussien en raison de la régularité des contours des facies.

Q2-
a) Sphérique, sans effet de pépite, avec portée de 40m et C=0.16

b) comme C=p*(1-p), la proportion peut donc être de 0.2 ou de 0.8.

c) On utilise l'abaque 7, courbe à 10/40, on trouve un facteur de 0.093, donc 0.093*0.16=0.0149 pour l'ensemble de la zone, la variance d'estimation sera : 0.0149/100=0.000149.

L'écart-type d'estimation est donc 0.012.

3- a) si le point que l'on veut simuler est une observation, le krigeage simple retourne la valeur observée et une variance de krigeage nulle (interpolateur exact). La valeur que l'on tire est donc nécessairement la valeur krigée, soit la valeur de l'observation.

b) p(3%)=0.91 Y(0.91)=1.3 environ

c) On cherche la valeur t correspondant à p=0.6 pour une distribution normale de moyenne -1.2 et d'écart-type 0.3. P((Ys-(-1.2))/0.3<t)=0.9 . Dans la table N(0,1), on trouve que t vaut 1.28. Donc Ys (x) est : 1.28*0.3-1.2= -0.816. Du graphique, on trouve un p de 0.21 environ et une valeur correspondante pour Z de 0.5% environ.

d) Il suffirait de remplacer le krigeage simple par un krigeage d'indicatrice pour estimer la distribution conditionnelle, on tirerait une valeur de cette distribution et l'on ajouterait cette valeur aux données avant de passer à la donnée suivante.

4- a) On fait $Z_{sc}=Z^*+(Z_s-Z_s^*)$

Coordonnée	Valeurs observées	Simulation non-conditionnelle	Valeur krigée avec les données	Valeurs krigées avec les valeurs simulées aux points des données	Simulation conditionnelle
0	10	6	10	6	10+(6-6)=10
100	-	7.5	9	8	9+(7.5-8)=8.5
200	-	7	8	10	8+(7-10)=5
300	-	10	7	12	7+(10-12)=5

400	6	14	6	14	6+(14-14)=6

b) La méthode A. La méthode B revient à appliquer une fonction de transfert non-linéaire aux valeurs krigées, ce que l'on sait être incorrect.

c) La valeur krigée, soit 9.

d) La racine de l'écart-type de krigeage $12^{0.5}$

Q5- a) On calcule K=L*L'=$\begin{bmatrix} 5 & 3 & 1 \\ 3 & 5 & 4 \\ 1 & 4 & 5 \end{bmatrix}$. Sur la diagonale, on a le palier du variogramme soit 5.

b) Comme Z1=2, Y1=Z1/L_{11}=2/2.236=0.894

c) Z_2=1.342*0.894+1.789*(-1)= -0.589

d) Z_2 =cte+1.789*Y_2; Z_3=1.901*Y_2+1.09*Y_3. Cov(Z_2,Z_3|Z_1)=1.789*1.901=3.40

Q6
a)

$$\begin{bmatrix} 5 & 2.362 & 0 & -0.354 \\ 2.362 & 5 & 0.354 & 0 \\ 0 & 0.354 & 0.075 & -0.18 \\ -0.354 & 0 & -0.18 & 0.075 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 4.674 \\ 3.462 \\ 0.210 \\ -0.364 \end{bmatrix}$$

b) Oui car il y a une relation déterministe entre les variables principales et secondaires.

c) pour ks : 5-0.782*4.674-0.323*3.462 = 0.2267
pour coks : 5-0.789*4.674 -0.2*3.462 -1.691*0.210 +0.364*(-0.719) = 0.003

On a réduit la variance d'estimation d'un facteur 7!

Q7-
a) ~0.12% Cu (valeur prédite par le krigeage)
b) 0.42% Cu (valeur moyenne prédite par les simulations).
c) On a évidemment intérêt à améliorer les estimations des blocs afin d,accroître nos profits. Les estimés actuels sont beaucoup trop imprécis.

Q8- C(h) = $\begin{bmatrix} 5 & 3 \\ 3 & 4 \end{bmatrix}$ effet de pépite (avec C=1) + $\begin{bmatrix} 30 & 10 \\ 10 & 5 \end{bmatrix}$ Sphérique avec anisotropie géométrique (a_0=75 et a_{90}=25; C=1)
les deux déterminants valent 11 et 50 > 0 donc le modèle est admissible.

Examen final 2013
Question 1 (10 points)

On vous donne le tableau suivant :

	Point x_1	Point x_2	Point x_0
Observé	20	10	-
Réalisation non-conditionnelle 1	6	12	9
Réalisation non-conditionnelle 2	15	9	12
Réalisation non-conditionnelle 3	10	17	10

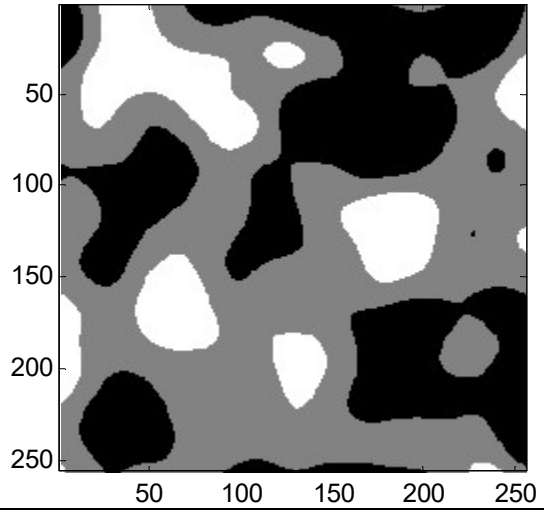
Lorsqu'on effectue le krigeage simple au point x_0, on obtient : $\lambda_1 = 0.4$, $\lambda_2 = 0.5$. Les moyennes du champ simulé et du champ réel sont égales à 0.

a) Transformez par post-conditionnement par krigeage simple les trois réalisations non-conditionnelles au point x_0 en trois réalisations conditionnelles aux valeurs observées aux points x_1 et x_2.

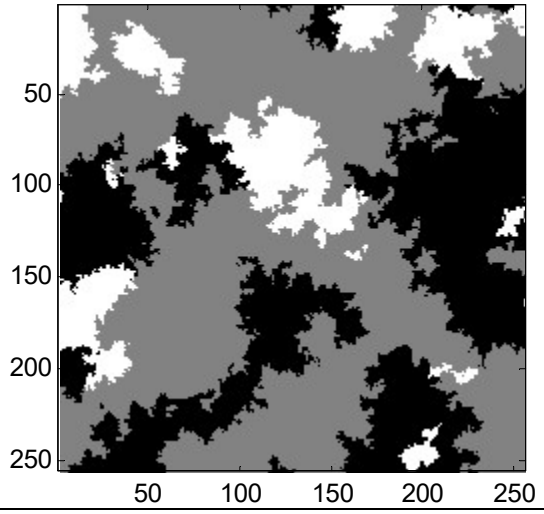
b) Démontrez, à l'aide du post-conditionnement par krigeage simple, qu'une simulation conditionnelle présente une variance d'estimation théorique égale au double de celle obtenue par krigeage simple.

Question 2 (14 points)

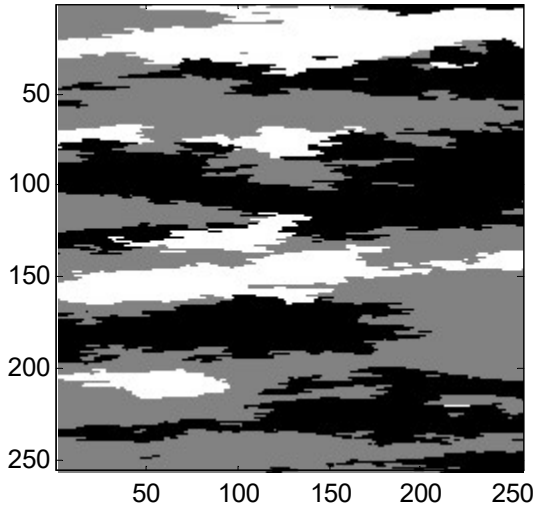
A



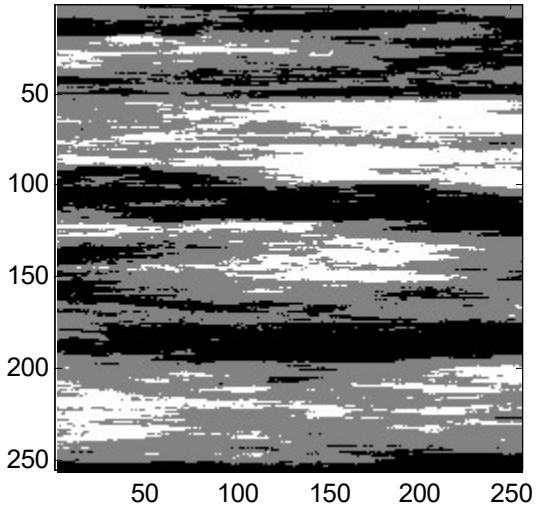
B



C

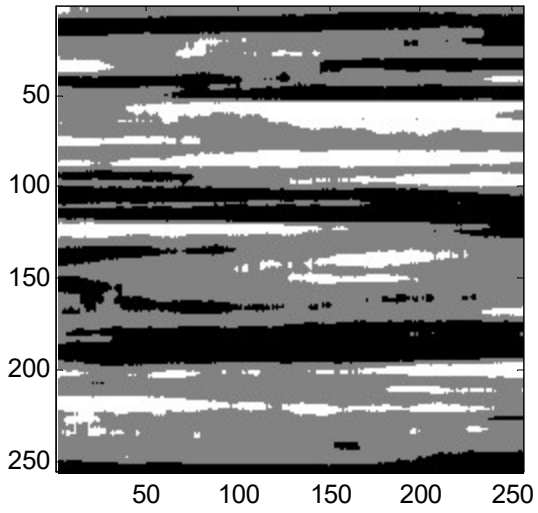


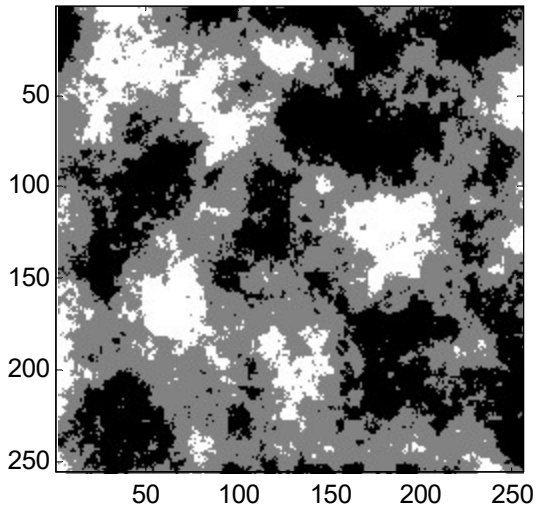
D



E

F





a) Complétez le tableau suivant en marquant d'un X, pour chaque image, les énoncés qui s'appliquent. (Aide : chaque colonne devrait avoir deux X au total)

Énoncé	**A**	**B**	**C**	**D**	**E**	**F**
L'image est obtenue par la méthode de simulation séquentielle d'indicatrices						
L'image est obtenue par la méthode « variable gaussienne tronquée »						
Le variogramme utilisé est gaussien isotrope						
Le variogramme utilisé est sphérique isotrope						
Le variogramme utilisé est gaussien anisotrope						
Le variogramme utilisé est sphérique anisotrope						

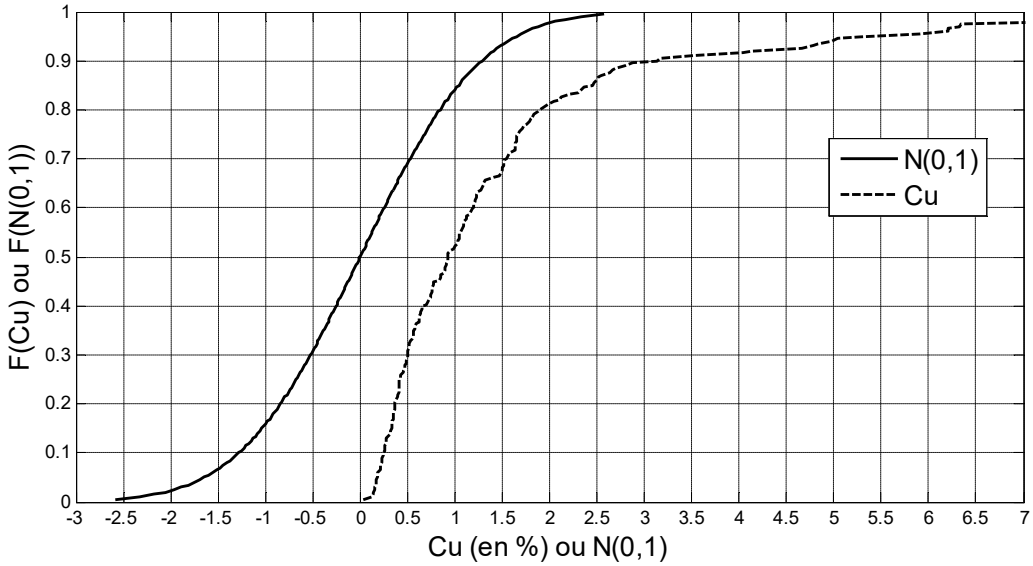
b) Dans un modèle gaussien tronqué, quelles proportions de facies F1, F2 et F3 seront simulées théoriquement si l'on adopte les seuils $c_1 = -0.2$ et $c_2 = 1.1$ pour la variable N(0,1) simulée ?

Question 2 (suite)

c) À quoi sert l'échantillonneur de Gibbs? Décrire succinctement l'algorithme.

Question 3 (10 points)

On effectue une simulation des teneurs de Cu (en %) d'un gisement par méthode séquentielle gaussienne (SGS). La teneur en Cu n'étant pas distribuée suivant une loi normale, on doit au préalable procéder à une transformation gaussienne (voir figure). À une certaine étape de l'algorithme, on effectue le krigeage simple de la variable transformée au point x_0 et l'on obtient une valeur krigée de 0.8 et une variance de krigeage de 0.3. On appelle le générateur de nombres aléatoires (distribution uniforme entre 0 et 1) et celui-ci retourne la valeur 0.33.



a) Quelle valeur gaussienne (N(0,1)) sera simulée au point x_0 ?

b) Quelle teneur de Cu sera simulée en ce même point ?

Question 3 (suite)

c) Que fait-on avec la valeur simulée pour poursuivre l'algorithme?

d) On applique plusieurs fois l'algorithme SGS. Pour obtenir une estimation de l'espérance conditionnelle et de l'écart-type conditionnel du Cu au point x_0 doit-on :

i- calculer la moyenne et l'écart-type des valeurs gaussiennes simulées en ce point et transformer la moyenne et l'écart-type à l'aide de la figure précédente ou

ii- transformer chaque valeur simulée au point x_0 et ensuite calculer la moyenne et l'écart-type des valeurs transformées ?

Justifier votre réponse.

Question 4 (14 points)

3- On simule une variable Z en 3 emplacements par la méthode de décomposition de Cholesky.

Les emplacements sont placés dans l'ordre 1,2,3 dans la matrice K. La matrice L obtenue est la suivante :

$$L = \begin{bmatrix} 3 & 0 & 0 \\ 2 & 2.2 & 0 \\ 1.9 & 1.34 & 1.9 \end{bmatrix}$$

a) Quelle est la covariance entre les v.a. Z correspondant aux 2^e et 3^e emplacements ?

b) On a observé Z_1=1. Exprimez Z_2 et Z_3 sous la forme a+b Y_2 + c Y_3.

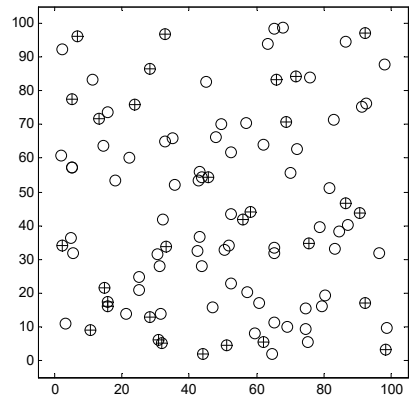
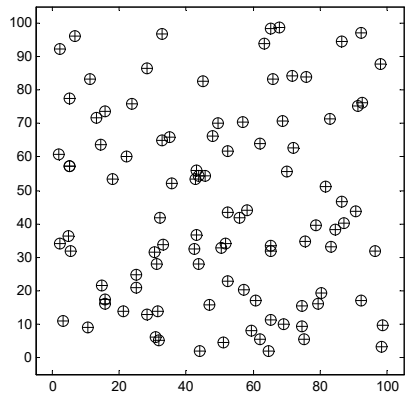
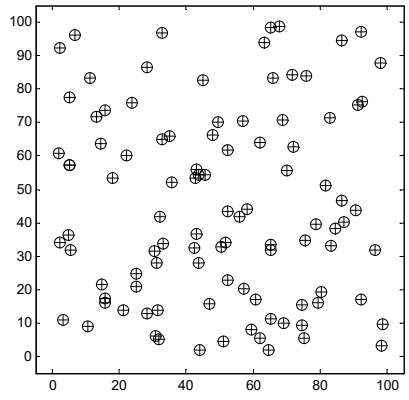
Question 4 (suite)

c) Déduisez de l'expression en b) la covariance (conditionnelle à Z_1) entre les v.a. Z_2 et Z_3.

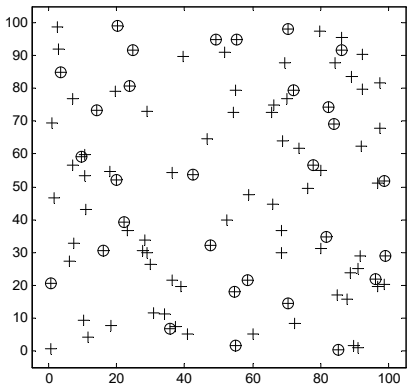
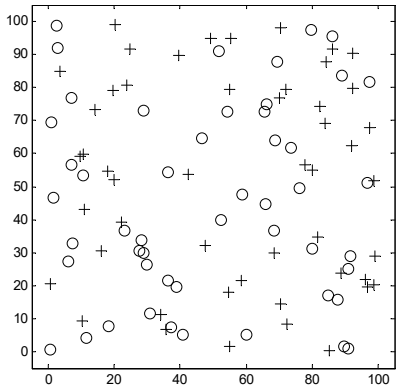
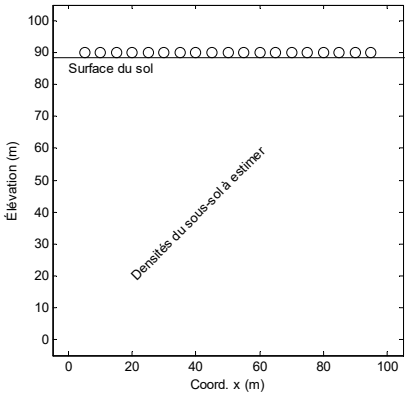
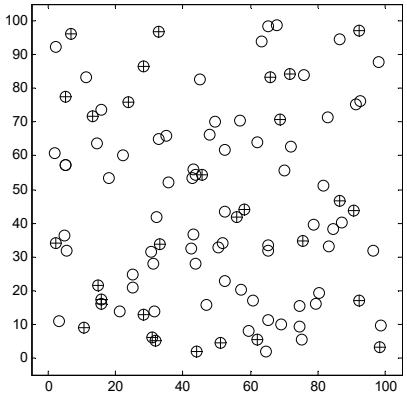
d) Quelle est la variance non-conditionnelle au 2e emplacement?

Question 5 (14 points)

Chaque cas (A à G) présente le modèle de covariance et la localisation des données. Z est la variable principale (les +) et Y la variable secondaire (les o). $\delta(h) = 1$ si |h| = 0, et 0 si |h|>0.

Cas	Modèle de covariance	Localisation
A	Modèle linéaire de corégionalisation $C(h) = \begin{bmatrix} 1.2 & 0 \\ 0 & 0.3 \end{bmatrix} \delta(h) + \begin{bmatrix} 7.1 & 1.3 \\ 1.3 & 10.2 \end{bmatrix}$ Sphér. (a=50 m)	
B	Z(x,y) a un variogramme de type gaussien , $a_{effectif}$ = 100 m. $Y(x,y) = \frac{\partial Z(x,y)}{\partial(x)}$ les covariances $C_{YY}(h)$ et $C_{ZY}(h)$ sont calculées en tenant compte de cette relation.	
C	Modèle linéaire de corégionalisation $C(h) = \begin{bmatrix} 7.1 & 8.1 \\ 8.1 & 10.2 \end{bmatrix}$ Sphér. (a=50 m)	

Question 5 (suite)

D	Modèle linéaire de corégionalisation $C(h) = \begin{bmatrix} 1.1 & 0.3 \\ 0.3 & 0.5 \end{bmatrix} \delta(h) + \begin{bmatrix} 6.9 & 7.9 \\ 7.9 & 9.2 \end{bmatrix}$ Sphér. (a=75 m)	
E	Modèle linéaire de corégionalisation $C(h) = \begin{bmatrix} 1.1 & 0.3 \\ 0.3 & 0.5 \end{bmatrix} \delta(h) + \begin{bmatrix} 6.9 & 7.9 \\ 7.9 & 9.2 \end{bmatrix}$ Sphér. (a=75 m)	
F	Z est le contraste de densité de blocs Y est l'anomalie gravimétrique mesurée en surface La relation physique liant la gravité aux densités est entièrement considérée dans le modèle de corégionalisation, i.e. $Y=GZ \Rightarrow C_{YY} = GC_{ZZ}G^T$ et $C_{YZ} = GC_{ZZ}$	
G	$C(h) = \begin{bmatrix} 7.1 & 8.6 \\ 8.6 & 10.2 \end{bmatrix}$ Sphér. (a=50 m)	

Question 5 (suite)

Complétez le tableau suivant en indiquant O (oui) ou N (non) selon que l'énoncé s'applique ou non au cas considéré (A à G). Vous pouvez apporter des justifications à vos réponses sous le tableau si vous le désirez. Une bonne réponse vaut 0.5 point, une mauvaise réponse vaut -0.5 point et une absence de réponse vaut 0 point.

	Cas						
Énoncé	A	B	C	D	E	F	G
Selon le modèle décrit, le cokrigeage devrait améliorer notablement l'estimation par rapport au krigeage (note : pour pouvoir répondre oui, il faut d'abord que le modèle soit admissible)							
Le modèle de corégionalisation décrit est admissible							
Indépendamment du modèle décrit et en supposant qu'un modèle admissible soit utilisé, seul le cokrigeage simple peut être effectué.							
Il est possible de calculer le variogramme croisé Z-Y **expérimental** avec les données indiquées.							

Question 6 (10 points)

On vous présente les observations suivantes donnant l'élévation d'un horizon repère.

Point	Coord x (m)	Coord. y (m)	Élévation (m)
x_1	30	100	29
x_2	90	160	37
x_3	220	80	52
x_4	300	350	47
x_5	40	200	$20 \le Z(x_5) \le 49$

En x_5, on n'a aucune information sur les probabilités pour $Z(x_5)$ à l'intérieur de l'intervalle indiqué [20, 49].

On a modélisé les variogrammes des différentes indicatrices et on a effectué le krigeage ordinaire au point x_0 de coordonnées (50,150). Les variogrammes des indicatrices étant ici tous proportionnels, les poids de krigeage sont les mêmes pour tous les seuils.

Lorsqu'on effectue le krigeage ordinaire avec les 5 points, on obtient les poids :

$$\lambda_1 = 0.35, \lambda_2 = 0.29, \lambda_3 = 0.07, \lambda_4 = 0.04, \lambda_5 = 0.25$$

Si l'on effectue le krigeage ordinaire avec les quatre premiers points, on obtient plutôt les poids :

$$\lambda_1 = 0.31, \lambda_2 = 0.39, \lambda_3 = 0.14, \lambda_4 = 0.16$$

a) Quelle est la probabilité que l'élévation de l'horizon repère au point x_0 soit inférieure à 50m?

b) Quelle est la probabilité que l'élévation de l'horizon repère au point x_0 soit inférieure à 40m?

Question 7 (14 points)

On a un champ gaussien 2D (de taille 200 m x 200 m) pour lequel on génère un très grand nombre de réalisations non-conditionnelles. Les simulations sont effectuées avec moyenne théorique 15. Le variogramme utilisé est sphérique avec C_0=2, C=8 et a=50 m.

a) L'on se place en un point donné x_0 et l'on examine les valeurs obtenues pour un très grand nombre de réalisations. Que valent la moyenne et la variance des valeurs simulées en ce point ?

b) Pour la 1ère réalisation, on calcule la variance expérimentale des valeurs ponctuelles sur l'ensemble du domaine. On répète l'exercice pour la 2^{e} réalisation, la 3^{e} et ainsi de suite. Que vaut la moyenne (calculée sur l'ensemble des réalisations) des variances expérimentales obtenues sur chaque réalisation?

c) On regroupe les valeurs ponctuelles simulées en blocs de taille 10 m x 10 m (i.e. on calcule la moyenne des valeurs ponctuelles à l'intérieur de chaque bloc). Si l'on considère les valeurs obtenues, pour un bloc donné, sur l'ensemble des réalisations, quelle moyenne et quelle variance observera-t-on?

Question 7 (suite)

d) Pour la 1ère réalisation, on calcule la variance expérimentale des teneurs des blocs de taille 10 m x 10 m sur l'ensemble du domaine de 200 m x 200 m. On répète l'exercice pour la 2e réalisation, la 3e et ainsi de suite. Que vaut la moyenne (sur les différentes réalisations) des variances expérimentales obtenues?

e) Qu'est ce qui justifie en général de faire des simulations? Fournir un exemple simple, en lien avec l'estimation des ressources minérales, ou la simulation est absolument nécessaire.

Question 8 (14 points)

On désire effectuer un cokrigeage simple d'indicatrice pour estimer la probabilité qu'un point montre une contamination en plomb dépassant la norme 500 ppm. La variable indicatrice vaut 1 si la teneur de l'échantillon est <u>supérieure</u> à 500 ppm et 0 sinon. La variable secondaire est donnée par $U(x_i)=\text{rang}(Z(x_i))/(n+1)$ où n est le nombre d'échantillons et le rang est la position occupée lorsque les données sont triées par ordre <u>croissant</u>. On a les données suivantes disponibles pour effectuer le cokrigeage au point x_0.

Point x_i	Coordonnée 1	Coordonnée 2	$I(x_i ; 500)$	$U(x_i)$
x_1	-50	0	1	0.9
x_2	50	0	0	0.3
x_0	0	0	?	-

On a le modèle linéaire de corégionalisation :

$$C(h) = \begin{bmatrix} 0.02 & 0 \\ 0 & 0.01 \end{bmatrix} \delta(h) + \begin{bmatrix} 0.23 & 0.1 \\ 0.1 & 0.083 \end{bmatrix} \exp(-|h|/200)$$

où $\delta(h)=1$ lorsque $|h|=0$ et $\delta(h)=0$ sinon.

a) Construisez le système de cokrigeage simple pour estimer la probabilité que la teneur au point x_0 soit supérieure à 500 ppm (ne pas le résoudre). Indiquez clairement la signification des valeurs fournies dans votre réponse.

Question 8 (suite)

b) Calculez la corrélation simple (à |h| = 0) entre I et U selon ce modèle.

Les poids obtenus par cokrigeage sont :

$$\lambda_{1,I} = \lambda_{2,I} = 0.44$$
$$\lambda_{1,U} = \lambda_{2,U} = 0.05$$

c) Calculez la probabilité estimée par cokrigeage simple que la concentration en plomb au point x_0 soit supérieure à 500 ppm, sachant que E[I;500]=0.3 et E[U]=0.5.

d) Calculez la variance de cokrigeage simple.

Annexes

Fonction de répartition N(0,1) : $F_X(x) = \int_{-\infty}^{x} f_X(x)dx$

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998

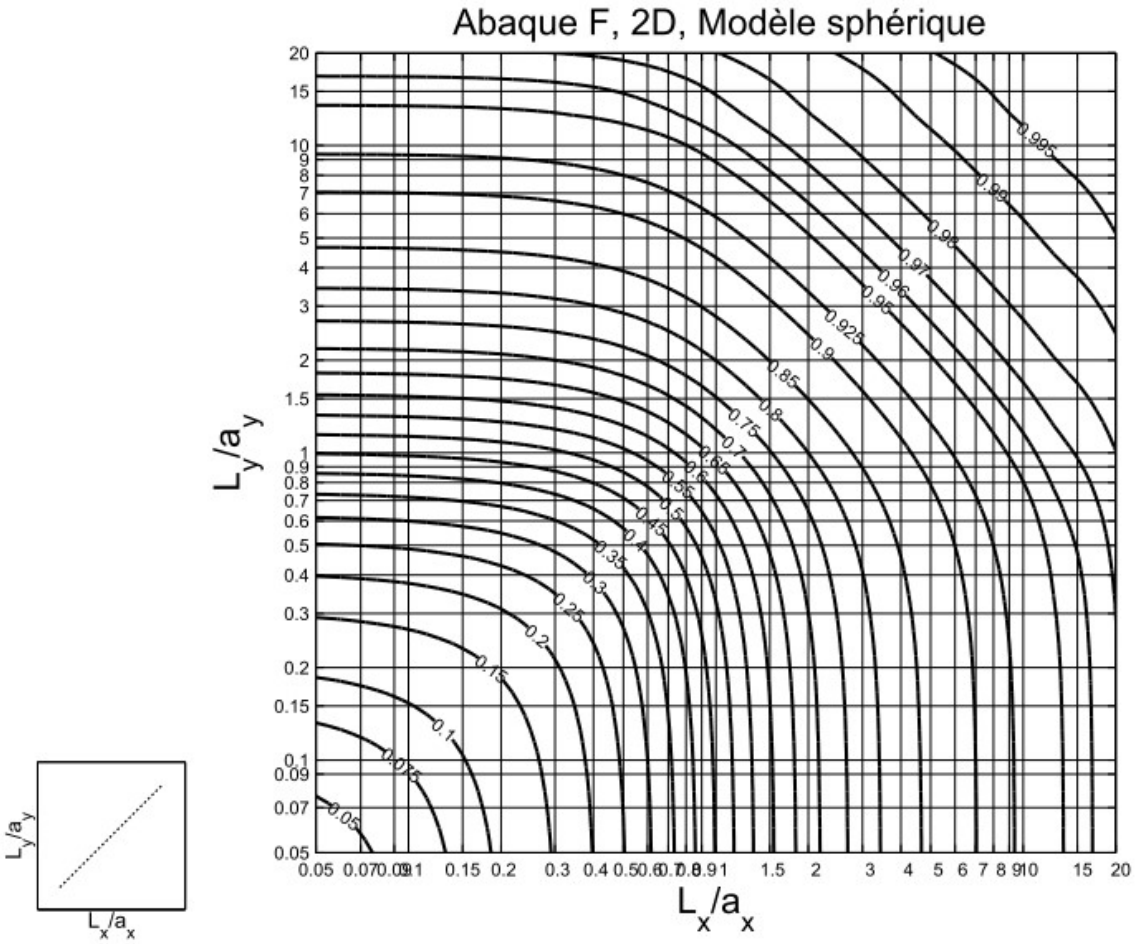


Fig. 2. Variance d'un point dans un rectangle, variogramme sphérique de palier C=1; $F(L_x/a_x, L_y/a_y) = D^2(\bullet|v) = \bar{\gamma}(v,v) = (1 - \bar{C}(v,v))$

Corrigé

Q1- a) $Z_0^* = 0.4*20+0.5*10 = 13$;
$Z_0^{s1*} = 0.4*6+0.5*12 = 8.4$
$Z_0^{s2*} = 0.4*15+0.5*9 = 10.5$
$Z_0^{s3*} = 0.4*10+0.5*17 = 12.5$

$Z_0^{sc1*} = 13+(9-8.4) = 13.6$
$Z_0^{sc2*} = 13+(12-10.5) = 14.5$
$Z_0^{sc3*} = 13+(10-12.5) = 10.5$

b) $\text{Var}(e) = \text{Var}(Z_0-Z_0^{sc}) = \text{Var}((Z_0 - Z_o^*) + (Z_0^{s*}-Z_0^s)) = 2\,\sigma_{ks}^2$

Q2-

Énoncé	**A**	**B**	**C**	**D**	**E**	**F**
L'image est obtenue par la méthode de simulation séquentielle d'indicatrices		X	X			
L'image est obtenue par la méthode « variable gaussienne tronquée »	X			X	X	X
Le variogramme utilisé est gaussien isotrope	X					
Le variogramme utilisé est sphérique isotrope		X				X
Le variogramme utilisé est gaussien anisotrope					X	
Le variogramme utilisé est sphérique anisotrope			X	X		

b)
Pour $F_1 = P(N(0,1)< -0.2) = 1-0.5793 = 0.4207$
Pour $F_2 = P(-0.2 < N(0,1) < 1.1) = 0.8643-0.4207 = 0.4436$
Pour $F_3 = P(N(0,1)>1.1) = 1-0.8643 = 0.1357$

c) À fournir des valeurs gaussiennes aux points échantillons où les faciès ont été observés.
Algorithme :
0- On initialise les valeurs en choisissant au hasard des valeurs gaussiennes dans l'intervalle correspondant à chaque faciès.
itération
1- On choisit un point au hasard. On le retire et on l'estime par KS en utilisant les autres points. On obtient ainsi une distribution conditionnelle normale de moyenne Z* et de variance σ_{ks}^2.

2- On tire une valeur au hasard de cette distribution et on code le faciès selon la valeur tirée. Si le faciès simulé correspond au faciès observé à ce point, on garde la valeur simulée et cette valeur remplace la valeur précédente en ce point, sinon on ne fait rien.
3- On évalue le critère de convergence et soit on arrête, soit on retourne à 1.

Q3- a) $Z_0 \sim N(0.8,0.3)$ => $W= (Z_0-0.8)/0.3^{0.5} \sim N(0,1)$
$F^{-1}(0.33) = W = -0.44 \Rightarrow Z_0 = -0.44*0.3^{0.5} +0.8 = 0.559$

b) Par lecture sur la figure environ 1.6%

c) On l'ajoute aux données observées et aux points déjà simulés.

d) option ii. La transformation étant non-linéaire, on ne peut l'appliquer sur une moyenne, on doit l'appliquer sur chaque valeur individuellement avant de prendre la moyenne.

Q4-
a) $Cov(Z_2,Z_3) = 2*1.9 + 2.2*1.34 = 6.748$

b) $Y_1 = 1/3$, $Z_2 = 2* 1/3 + 2.2* Y_2$ et $Z_3 = 1.9*1/3+1.34*Y_2+1.9*Y_3$

c) $Cov(Z_2,Z_3 \mid Z_1) = 2.2*1.34 = 2.948$

d) $Var(Z_2) = 2^2+2.2^2 = 8.84$

Q 5

	Cas						
Énoncé	A	B	C	D	E	F	G
<u>Selon le modèle décrit</u>, le cokrigeage devrait améliorer notablement l'estimation par rapport au krigeage (note : pour pouvoir répondre oui, il faut d'abord que le modèle soit admissible)	N	O	N	N	O	O	N
Le modèle de corégionalisation décrit est admissible	O	O	O	O	O	O	N
Indépendamment du modèle décrit et en supposant qu'un modèle admissible soit utilisé, seul le cokrigeage simple peut être effectué.	N	N	N	N	N	O	N
Il est possible de calculer le variogramme croisé Z-Y **<u>expérimental</u>** avec les données indiquées.	O	O	O	O	N	N	O

Q 6-
a) $I_0^* = 0.35*1+0.29*1+0.07*0+0.04*1+0.25*1 = 0.93$
b) $I_0^* = 0.31*1+0.39*1+0.14*0+0.16*0 = 0.7$

Q7- a) m=15 et var=C_0+C = 10

b) On obtient $D^2(.|G) = C_0$+CF(200/50,200/50) = 2+8* 0.967 = 9.736

c) m=15, Var(Z_v) = C(1-F(10/50,10/50))=8* (1-0.16) = 6.72

d) $D(Z_v|G)$ = 8(F(200/50,200/50)-F(10/50,10/50)) = 6.456

e) Il faut que l'on s'intéresse à une fonction non-linéaire de la variable d'intérêt. En mine par exemple si on veut estimer la précision de la ressource estimée sur des blocs d'une certaine taille, on doit recourir à des simulations (l'estimation de la ressource, elle, peut être calculée soit sur les valeurs krigées, soit à partir de la distribution des teneurs de blocs).
Aussi : si l'on veut intégrer des contraintes minières dans l'estimation des ressources.

Q8-
a)
exp(-100/200) = 0.6065et exp(-50/200)=0.7788

En plaçant dans l'ordre : I1 U1, I2 U2, on obtient :
k* λ =k_0
avec k=

0.2500	0.1000	0.1395	0.0607
0.1000	0.0930	0.0607	0.0503
0.1395	0.0607	0.2500	0.1000
0.0607	0.0503	0.1000	0.0930

k0=

0.1791
0.0779
0.1791
0.0779

λ les poids de cokrigeage dans l'ordre indiqué.

b) $0.1/(0.25*0.093)^{0.5}$ = 0.656

c) 0.3+0.44((1-0.3)+ (0-0.3)) + 0.05* ((0.9-0.5)+ (0.3-0.5)) = 0.489

d) 0.25- (0.44 0.05 0.44 0.05)*k_0 = 0.0846

GLQ3401 - Final 2012

Question 1 (8 points)

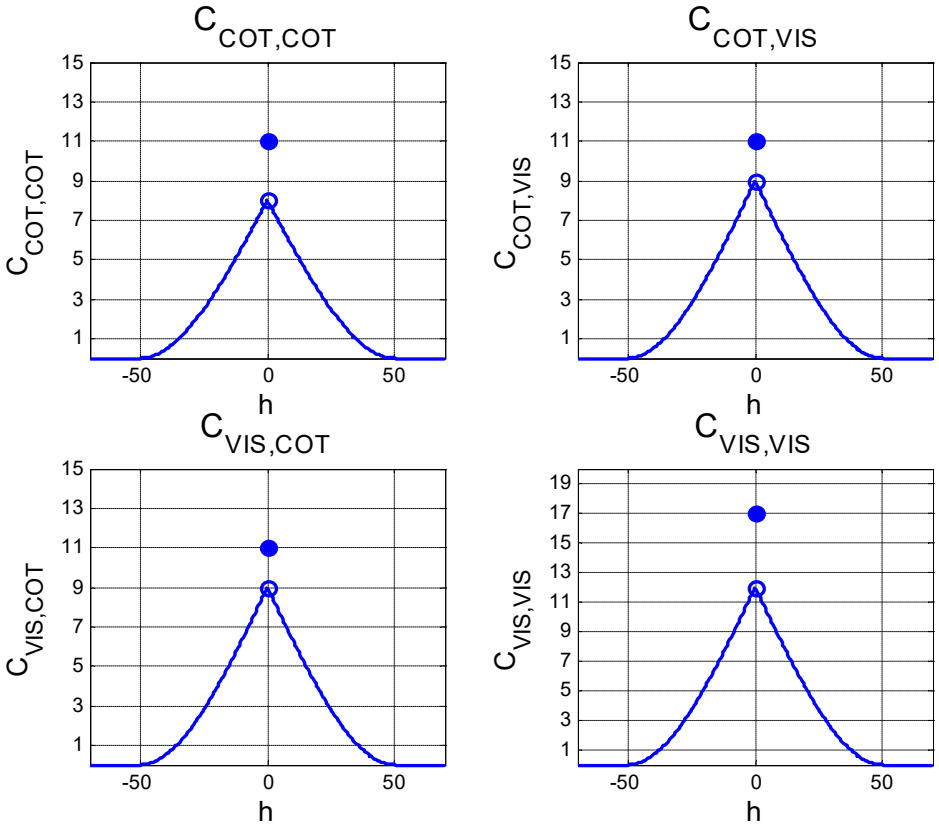
On considère la méthode de recuit simulé définie par une certaine fonction objectif O à minimiser et l'équation décrivant la probabilité d'accepter un changement donnant une hausse de la fonction objectif : $\exp\{(O_k - O_{candidat})/t_k\}$, où t_k est la « température » à l'itération k, donnée par une fonction décroissante du nombre d'itérations k.

a) Décrivez l'influence du choix de la fonction t_k sur le comportement du recuit simulé.

b) À l'itération k actuelle, l'on a t_k=10, O_k=100 et une modification proposée montre $O_{candidat}$ = 108. Le générateur de nombre aléatoire U(0,1) retourne la valeur 0.3. Doit-on accepter la modification proposée?

Question 2 (16 points)

Sur un site contaminé, on dispose d'analyses chimiques du contenu organique total (COT) du sol en 50 emplacements (la matière organique n'est pas d'origine naturelle, elle est liée à la présence de contaminants organiques dans le sol). En 1000 emplacements (incluant les 50 où COT est analysé), une appréciation visuelle (VIS) sur une échelle de 0 à 7 du contenu en matière organique a été obtenue. La figure suivante montre les modèles de covariance pour COT et VIS et la covariance croisée entre les deux variables. L'objectif est de fournir une estimation précise du COT en tout point du domaine étudié.



a) Donnez les paramètres du modèle linéaire de corégionalisation.

b) Ce modèle est-il admissible?

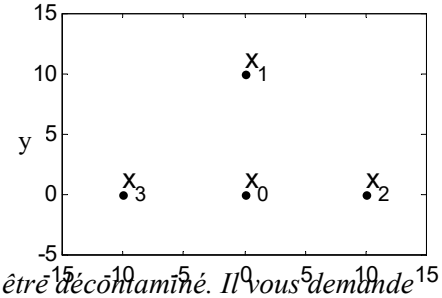
Les valeurs suivantes de covariance pourraient vous permettre de sauver quelques calculs à la question c) (la portée effective est 50m et C=1 pour chacun des 3 modèles) :

h	Sphérique	Exponentiel	Gaussien
10	0.704	0.549	0.887
14.1	0.587	0.428	0.787
20	0.432	0.301	0.619

c) Le tableau suivant donne l'emplacement de 3 données voisines du point x_0 de coordonnée (0,0) pour lequel on veut obtenir une estimation de COT par cokrigeage avec VIS comme variable secondaire. Construisez, sous forme matricielle, le système de cokrigeage simple. (Ne pas le résoudre; identifiez clairement les éléments de la matrice et des vecteurs)

Les points sont situés en :

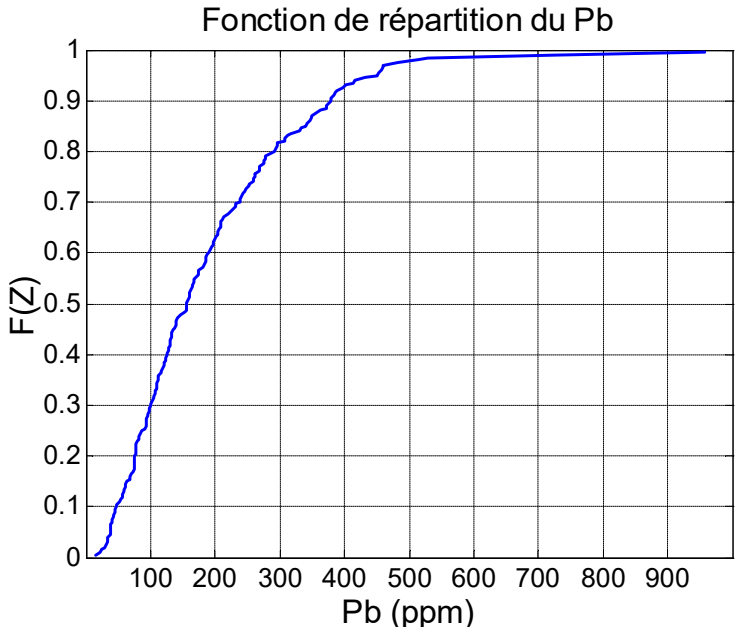
Point	Coordonnée x (m)	Coordonnée y (m)	COT	VIS
x_1	0	10	5%	2
x_2	10	0	-	1
x_3	-10	0	-	2



d) Le client vous informe que tout point du sol ayant COT>1% devra être décontaminé. Il vous demande de fournir une estimation du volume contaminé sur l'ensemble de la zone étudiée ainsi qu'une mesure de précision sur le volume estimé. Pouvez-vous répondre à cette demande à partir des résultats du cokrigeage? Si oui, indiquez comment. Si non, suggérez une approche qui permettrait d'y répondre.

Question 3 (12 points)

Un sol de la région de Montréal est contaminé au plomb dû à la présence d'une usine de recyclage de batteries automobiles. Deux cents analyses de la concentration en plomb sont disponibles et la fonction de répartition est illustrée à la figure suivante :



a) Indiquez quels sont les paliers attendus pour les indicatrices correspondant aux seuils suivants : 100 ppm, et 400 ppm.

b) Quelles sont les unités des variogrammes d'indicatrices ?

Le tableau suivant donne l'emplacement et les valeurs de Pb (en ppm) de 4 observations.

Point	Coord. x (m)	Coord. y (m)	Pb ppm
x_1	200	400	120
x_2	600	400	55
x_3	500	650	253
x_4	550	450	437

Le tableau suivant donne les poids obtenus par KI simple pour les divers seuils utilisés lors de l'estimation au point x_0 de coordonnée (500,500); la dernière colonne donne la distance entre les points x_i (i=1...4) et le point x_0 : (note : I(x,c) = 1 si $Z(x) \leq c$)

	100 ppm	200 ppm	300 ppm	400 ppm	500 ppm	distance x_i-x_0 (m)
λ_1	0	0.08	0	0	0	316.2
λ_2	-0.01	-0.02	-0.01	-0.01	0	141.4
λ_3	0.22	0.28	0.22	0	0	150
λ_4	0.70	0.60	0.70	0.12	0	70.7

c) Quelle est la probabilité (estimée par KI simple) que la concentration au point x_o soit supérieure à 200 ppm ? (Ne cherchez pas à effectuer les corrections pour relation d'ordre).

d) Que peut-on dire de la portée du variogramme d'indicatrice utilisé dans le krigeage pour le seuil 500 ppm?

Question 4 (14 points)

On désire simuler, le long d'un profil (1D), la profondeur du sommet d'une lentille minéralisée de Cu. Le tableau suivant présente une série de valeurs simulées et/ou krigées le long du profil. Les unités sont des mètres et la variable profondeur peut être considérée distribuée suivant une loi normale. La technique de simulation non-conditionnelle utilisée est la méthode FFT-ma. On désire obtenir des valeurs simulées conditionnellement aux observations en utilisant la technique de post-conditionnement par krigeage. Les krigeages effectués sont tous des krigeages simples.

Point	Coordonnée (m)	Profondeur obtenue par simulation non conditionnelle (m)	Krigeage avec les valeurs simulées aux points des données (m)	Krigeage avec les données observées (m)
x_1	0	240	230	165
x_2	100	220	240	190
x_3	200	250	250	180
x_4	300	270	260	170
x_5	400	260	265	175

a) Indiquez quelles sont les valeurs simulées conditionnellement pour les points x_1 à x_5.

b) Un des points sur le profil simulé correspond à une donnée observée. Pouvez-vous identifier ce point?

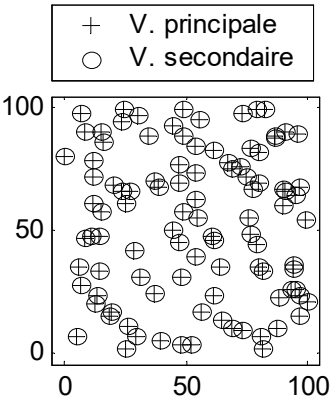
c) Au point x_5 la variance de krigeage avec les données observées vaut 30 m^2. Quelle est la variance de krigeage obtenue avec le krigeage des valeurs simulées aux mêmes points que les données ?

d) Donnez un intervalle devant contenir 95% des réalisations obtenues par simulation conditionnelle au point x_5 (rappel pour une loi normale un intervalle de confiance de niveau 95% correspond à $\pm$ 1.96 fois l'écart-type).

Question 5 (10 points)

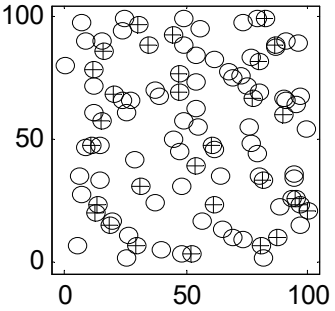
Le tableau suivant présente, dans la colonne de gauche, le plan de localisation d'observations d'une variable principale (Z) et d'une variable secondaire (Y). Dans la colonne de droite, on donne le modèle linéaire de corégionalisation ajusté à ces données. Dans les matrices, Z apparaît en premier, Y en second.

A)



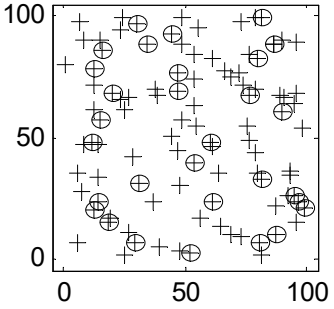
$$C(h)=\begin{bmatrix}1 & 0.7\\ 0.7 & 0.9\end{bmatrix}\delta(h)+\begin{bmatrix}8 & 8\\ 8 & 9\end{bmatrix}Sph(a=50,C=1)$$

B)



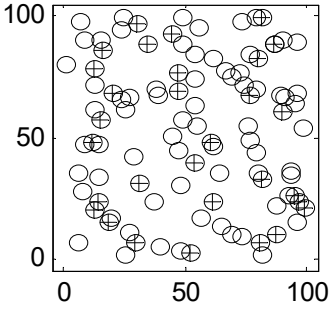
$$C(h)=\begin{bmatrix}1 & 0.7\\ 0.7 & 0.9\end{bmatrix}\delta(h)+\begin{bmatrix}8 & 8\\ 8 & 9\end{bmatrix}Sph(a=50,C=1)$$

C)



$$C(h)=\begin{bmatrix}1 & 0.7\\ 0.7 & 0.9\end{bmatrix}\delta(h)+\begin{bmatrix}8 & 8\\ 8 & 9\end{bmatrix}Sph(a=50,C=1)$$

D)



$$C(h)=\begin{bmatrix}1 & 0.7\\ 0.7 & 0.9\end{bmatrix}\delta(h)+\begin{bmatrix}8 & 0\\ 0 & 9\end{bmatrix}Sph(a=50,C=1)$$

a) Indiquez pour chaque cas A) à D) si le cokrigeage de Z(x) (par Z et Y) est susceptible d'améliorer de façon importante la précision des estimations par rapport au seul krigeage de Z(x). Justifiez chaque réponse.

b) Indiquez quelle méthode vous permettrait, utilisant les mêmes données, de confirmer ou d'infirmer vos réponses en a) ?

Question 6 (12 points)

On vous présente les observations suivantes donnant l'élévation d'un horizon repère.

Point	Coord x (m)	Coord. y (m)	Élévation (m)
x_1	30	100	29
x_2	90	160	45
x_3	220	80	52
x_4	300	350	47
x_5	40	200	$28 \leq Z(x_5) \leq 39$

En x_5, on n'a aucune information sur les probabilités pour $Z(x_5)$ à l'intérieur de l'intervalle indiqué [28, 39].

On a modélisé les variogrammes des différentes indicatrices et on a effectué le krigeage ordinaire au point x_0 de coordonnées (50,150). Les variogrammes des indicatrices étant ici tous proportionnels, les poids de krigeage sont les mêmes pour tous les seuils.

Lorsque l'on effectue le krigeage ordinaire avec les 5 points, on obtient les poids :

$$\lambda_1 = 0.32, \lambda_2 = 0.33, \lambda_3 = 0.03, \lambda_4 = 0.02, \lambda_5 = 0.30$$

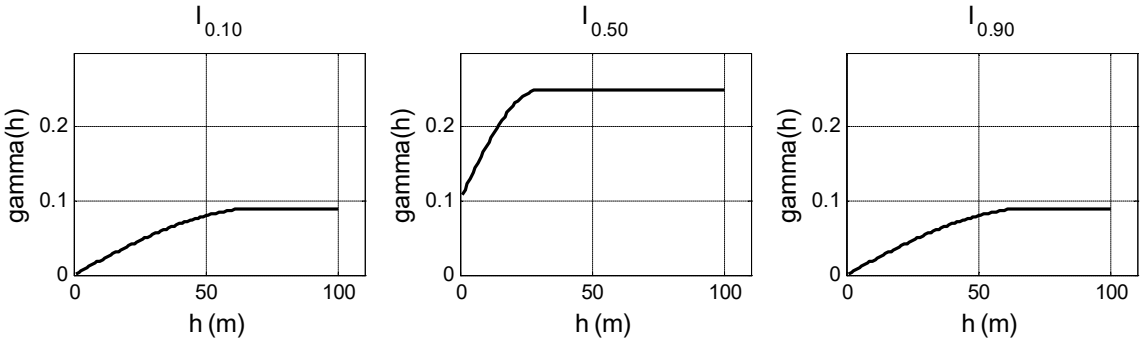
Si l'on effectue le krigeage ordinaire avec les quatre premiers points, on obtient plutôt les poids :

$$\lambda_1 = 0.41, \lambda_2 = 0.49, \lambda_3 = 0.04, \lambda_4 = 0.06$$

a) Quelle est la probabilité que l'élévation de l'horizon repère au point x_0 soit inférieure à 40m?

b) Quelle est la probabilité que l'élévation de l'horizon repère au point x_0 soit inférieure à 30m?

Les variogrammes des indicatrices obtenus pour les quantiles 10%, 50% et 90% de l'élévation sont donnés :



c) L'hypothèse d'une distribution bigaussienne pour l'élévation vous semble-t-elle réaliste ? Justifiez votre réponse.

Question 7 (18 points)

On effectue une simulation de moyenne 0 par méthode de Cholesky. La matrice L obtenue est :

$$L = \begin{bmatrix} 7.75 & & & & \\ 3.82 & 6.74 & & & \\ 4.59 & 3.23 & 5.34 & & \\ 4.36 & 1.45 & 2.20 & 5.84 & \\ 4.77 & 1.79 & 2.42 & 2.86 & 4.48 \end{bmatrix}$$

a) Quelle est la covariance (à priori) entre les 2^e^ et 3^e^ variables aléatoires?

b) On a observé Z_1=5. Que vaut Y_1?

c) Quelle est l'espérance conditionnelle de Z_2 sachant que Z_1=5?

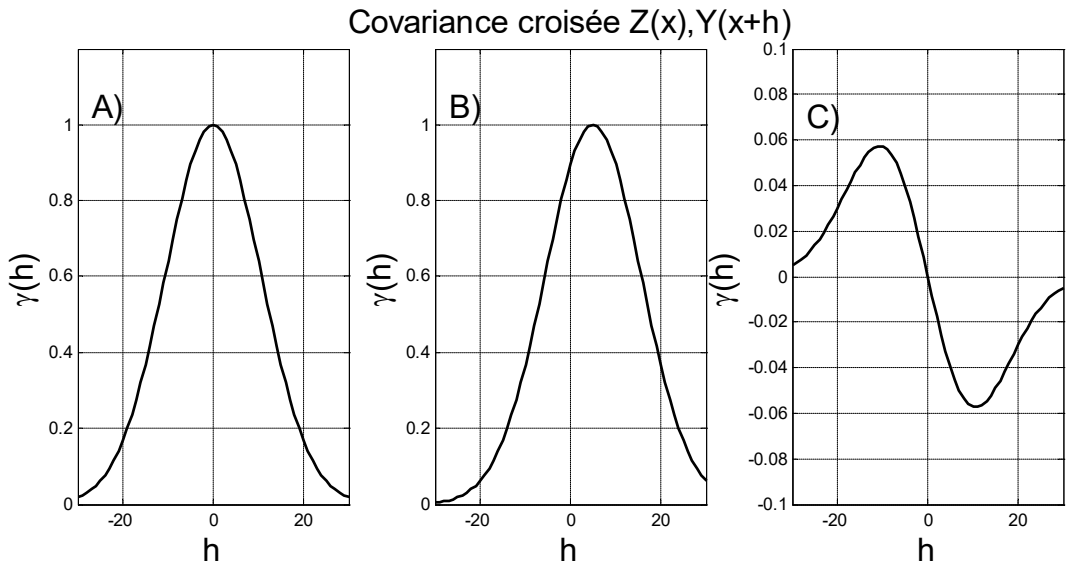
d) Quelle est la variance conditionnelle de Z_2 sachant Z_1=5 ?

e) Au plan pratique, quelle est la principale limitation touchant la méthode de Cholesky ?

f) En réalité, on voudrait simuler un processus de moyenne théorique 3. Que faut-il faire pour que les Z simulés aient cette moyenne théorique?

Question 8 (10 points)

1- On vous présente trois covariances croisées différentes (A-B-C) pour un problème 1D :



a) Identifiez la covariance croisée correspondant à Y(x)=dZ(x)/dx.

b) Identifiez la covariance croisée montrant Y(x) décalée par rapport à Z(x). Indiquez la valeur du décalage.

c) Identifiez la covariance croisée pouvant correspondre à un modèle de corégionalisation linéaire.

d) On effectue le cokrigeage avec ces modèles, indiquez celui ou ceux permettant d'obtenir une matrice de cokrigeage symétrique.

Corrigé

1a) Si t_k décroît trop rapidement, on risque de rester pris dans un minimum local, ce qui fait que l'on ne pourra pas rencontrer les objectifs inclus dans la fonction objectif.

Si tk décroît trop lentement, on aura beaucoup d'oscillations de la fonction objectif et la convergence sera donc lente.

b) On calcule exp((100-108)/10) = exp(-0.8) = 0.45. Comme le générateur retourne le nombre 0.3 et que 0.3<0.45, on accepte la modification.

2- a) C(h) = [3 2; 2 5] delta(h) + [8 9 ; 9 12] Sphérique (C=1, a=50).

b) les déterminants sont 11 et 15 donc le modèle est admissible.

c) Le système de cokrigeage s'écrit :
k*a=k0 avec

k=	11	11	5.2834	5.2834
	11	17	7.0446	7.0446
	5.2834	7.0446	17	5.184
	5.2834	7.0446	5.184	17

et k0= 5.632
6.336
6.336
6.336

d) Non on ne peut pas utiliser le cokrigeage pour cela car la sélection est une fonction de transfert non-linéaire appliquée aux blocs. On doit recourir à des simulations conditionnelles.

3- a) à 100 ppm , on voit F(100)~0.3 donc palier : 0.3*0.7 = 0.21

à 400 ppm, on lit F(400)~0.93 donc palier à 0.93*0.07 = 0.065

b) sans unité

c) les indicatrices par rapport au seuil 200 valent 1, 1, 0, 0.
$P(Z_0>200) = 1 - P(Z_0 \leq 200)$
$P(Z_0 \leq 200)) =$ 0.08 – 0.02 + (1-.94)*0.63 = 0.098

$P(Z_0>200)$ = 1-0.098 = 0.902

d) la portée est inférieure à la plus petite distance observée avec x_0 soit 70.7 m.

4- a) On fait la somme de la dernière colonne + la différence entre les 3e et 4e colonnes ,
donc : 165 + (240-230) = 175, 170, 180, 180, 170

b) x_3 car Z* car c'est le seul point où $Z^s{}^*=Z^s$

c) c'est la même chose soit 30 m^2 car on fait le krigeage avec les mêmes points et le même modèle.

d) $175 \pm 1.96\sqrt{30}$

5- a) A non car Z et Y sont échantillonnés aux mêmes points
B oui car Y est plus échantillonné que Z et la corrélation est forte à 0.92
C non car la v. secondaire est moins échantillonnée que Z.
D non car la corrélation Y-Z est trop faible à $0.7/(9*9.9)^{0.5} = 0.074$

b) On peut fair eune validation croisée où l'on compare les erreurs en valeur absolues ou au carré obtenues par krigeage et par cokrigeage.

6- a) à 40 m, I5 = 1, I1=1, I2=0, I3=0, I4=0 donc 0.32+0.3 = 0.62

b) à 30 m, I1=1, I2=0, I3=0 et I4=0, I5 n'est pas défini donc : 0.41

c) Non car le variogramme le moins structuré se retrouve à la médiane alors que dans le cas bigaussien, on devrait retrouver le maximum de structure `à la médiane.

7-
a) On multiplie la 2e ligne par la 3e ligne transposée => 3.82*4.59 + 6.74*3.23 = 39.304
b) $Y1= L_{11}^{-1}Z_1 = 5/7.75 = 0.645$
c) Z2= 3.82* 0.645+ 6.74 Y2. Comme Y2 est de moyenne 0 => 3.82*0.645 = 2.46
d) $6.74^2 = 45.43$
e) La matrice est de taille n x n. On ne peut simuler des champs ayant plusieurs milliers de points.
f) on ajoute 3 aux valeurs simulées.

8- a) C
b) B, le décalage est approximativement de 5 (Y est décalé à l'est par rapport à Z.
c) A
d) Les 3 modèles fournissent des matrices de cokrigeage symétriques.

GLQ3401 - Final 2011

Question 1 (18 points)

Les teneurs de Cu et Zn mesurées en quelques points sont fournies dans le tableau suivant :

Point	x	y	Cu (%)	Zn (%)
x_0	0	0	--	--
x_1	10	5	2	5
x_2	0	7	3	--
x_3	-5	0	--	3

La corégionalisation entre les 2 variables est décrite par un modèle linéaire de corégionalisation :

$$C(h) = \begin{matrix} \\ \text{Cu} \\ \text{Zn} \end{matrix}\begin{bmatrix} 3 & 2 \\ 2 & 4 \end{bmatrix}\delta(h) + \begin{matrix} \\ \text{Cu} \\ \text{Zn} \end{matrix}\begin{bmatrix} 4 & 3 \\ 3 & 6 \end{bmatrix} Sphérique(a = 30, isotrope)$$

a) Complétez le système de cokrigeage ordinaire suivant en substituant les valeurs appropriées pour a1 à a6.

$$\begin{matrix} \text{Cu}(x_1) & \text{Zn}(x_1) & \text{Cu}(x_2) & \text{Zn}(x_3) & & \end{matrix}$$
$$\begin{bmatrix} a2 & 5 & a1 & 0.8479 & 1 & 0 \\ 5 & 10 & 1.5292 & 1.6958 & 0 & 1 \\ a1 & 1.5292 & 7 & 1.745 & a3 & a4 \\ 0.8479 & 1.6958 & 1.745 & 10 & 0 & 1 \\ 1 & 0 & a3 & 0 & 0 & 0 \\ 0 & 1 & a4 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0.4983 \\ -0.1696 \\ 0.5017 \\ 0.1696 \\ -1.9392 \\ -0.4495 \end{bmatrix} = \begin{matrix} \text{Cu}(x_0) \\ \\ \\ \\ \\ \\ \end{matrix} \begin{bmatrix} 1.8675 \\ 1.4006 \\ 2.6254 \\ 2.2569 \\ a5 \\ a6 \end{bmatrix}$$

b) Quelle est la teneur estimée au point x_0 ?

Un 2e modèle a été testé par validation croisé et a fourni des estimations légèrement plus précises. Ce modèle est :

$$C(h) = \begin{matrix} \\ \text{Cu} \\ \text{Zn} \end{matrix}\begin{bmatrix} 4 & 5 \\ 5 & 6 \end{bmatrix} Sphérique(a = 30, isotrope)$$

c) Devrait-on utiliser ce modèle ? Justifiez.

d) On vous apprend que, bien qu'inconnue, la teneur moyenne du Zn sur l'ensemble du champ est égale à la teneur moyenne du Cu. Indiquez quelles modifications vous pourriez apporter au système de cokrigeage pour tenir compte de cette information tout en réduisant la variance théorique de cokrigeage..

Question 2 (7 points)

La figure 1 montre un plan de localisation d'une variable principale (+) et d'une variable secondaire (o). La figure 2 montre le diagramme binaire des 2 variables pour les points où les 2 variables ont été mesurées. L'objectif est d'estimer la variable principale.

Figure 1: Plan de localisation

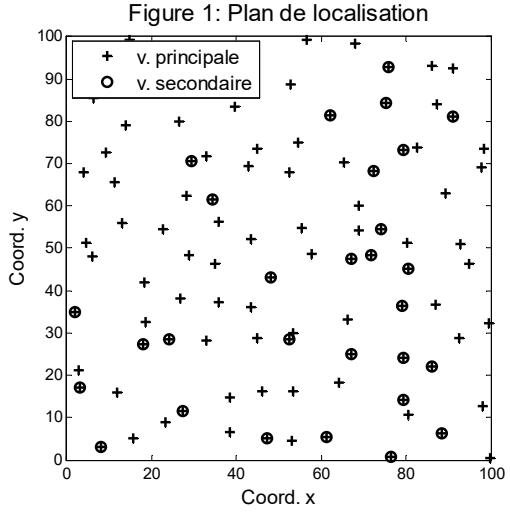
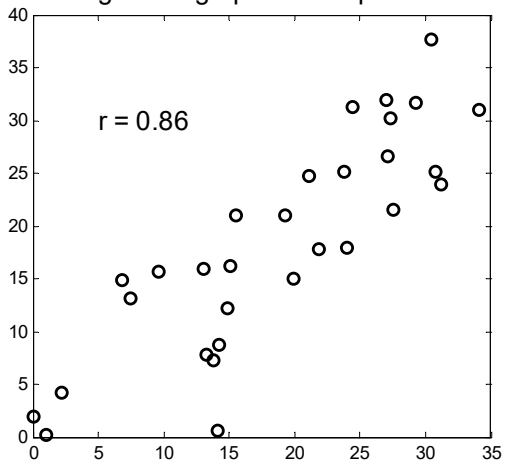


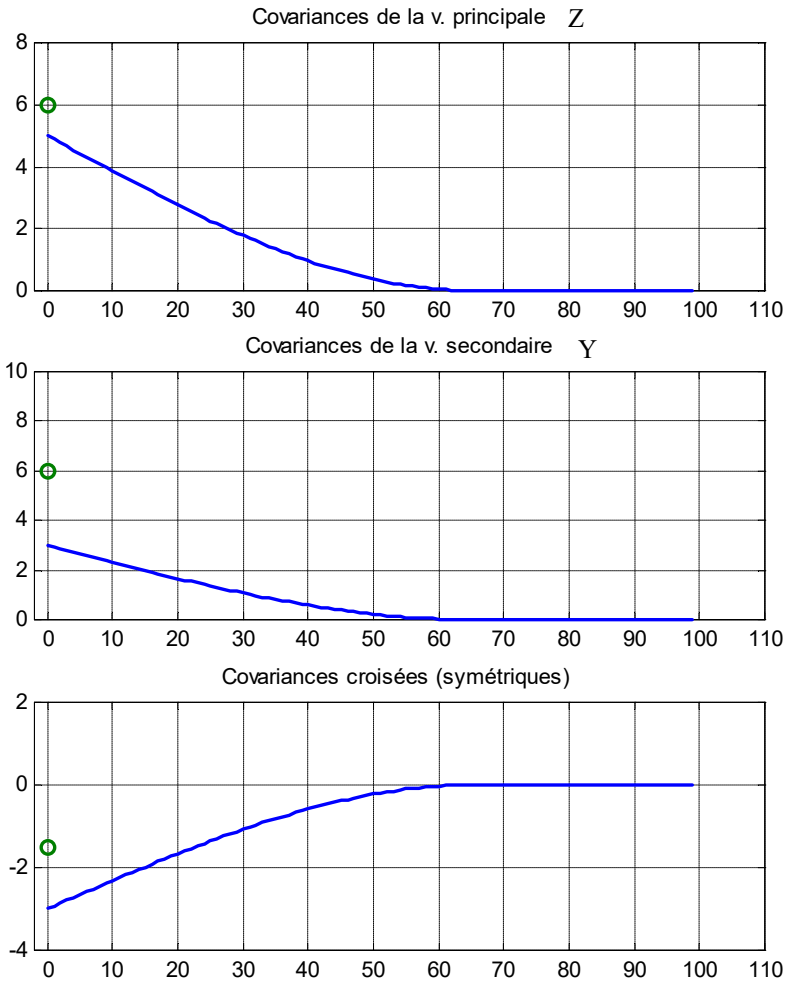
Figure 2 : graphe de dispersion



Croyez-vous que le cokrigeage réussira à améliorer les estimations de la variable principale ? Justifiez votre réponse.

Question 3 (10 points)

La figure suivante montre les fonctions de covariances simples et croisée de deux variables Z (principale) et Y (secondaire).



a) Décrivez le modèle linéaire de corégionalisation correspondant.

b) Calculez la corrélation entre la variable principale et la variable secondaire décrite par le modèle à h=0 et à h=0+.

c) En vue d'augmenter la corrélation entre Z et Y peut-on transformer la variable Y (en prenant le logarithme par exemple) ?

Question 4 (16 points)

Soit la matrice L suivante obtenue d'une décomposition de Cholesky (LU). On a observé Z_1=1.

$$L = \begin{bmatrix} 1.732 & 0 & 0 \\ 1.4434 & 0.9574 & 0 \\ 1.1547 & -0.4874 & 1.1954 \end{bmatrix}$$

a) Quelle est la covariance entre Z_2 et Z_3 ?

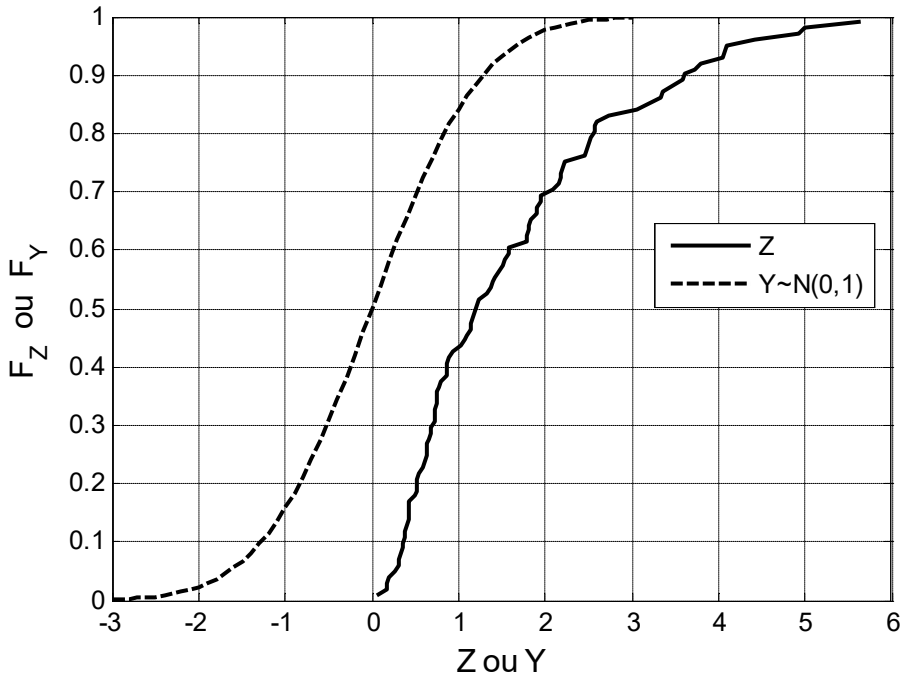
b) Quelle est l'espérance conditionnelle de Z_2 étant donné Z_1=1 ?

c) Quelle est la variance conditionnelle de Z_3 étant donné que Z_1=1 ?

d) Quelle est la covariance entre Z_2, et Z_3 étant donné que Z_1=1 ?

Question 5 (15 points)

La figure suivante vous montre la fonction de répartition expérimentale des données Z et la fonction de répartition de Y ~ N(0,1)



a) Est-ce que Z suit une loi normale? Pourquoi?

b) Pour effectuer la simulation par la méthode SGS, doit-on fournir le modèle de variogramme de Z ou celui de Y ?

c) Après transformation inverse vers Z est-ce que, en un point x_0 donné, la variance observée à travers les réalisations sera uniquement fonction du variogramme et de la position des observations ou dépendra-t-elle aussi des valeurs observées dans le voisinage du point x_0? Justifiez.

d) Quelle est la valeur Y associée à Z=1 ?

e) Quelle est la valeur Z associée à Y=1 ?

Question 6 (16 points)

Le tableau suivant présente 20 valeurs Z_s d'une simulation non-conditionnelle aux coordonnées x=1 à x=20, présentant un variogramme sphérique avec C=10, a=30. Des données sont observées aux points x_5, x_{10}, et x_{20}.

$Z^s(x)$	-3.55	-2.74	-1.64	-1.89	-1.37	-1.74	-3.16	-3.39	-3.51	-2.74	-3.60	-2.80	-1.39	-0.33	-1.05	-0.04	0.25	1.23	2.53	2.24
Z(x)	-	-	-	-	0.02	-	-	-	-	-0.53	-	-	-	-	-	-	-	-	-	-2.80
x	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20

Le tableau suivant donne les poids de krigeage simple obtenus pour l'estimation aux points x=7 et x=18.

Observation	Point estimé : x_7	Point estimé : x_{18}
Z_5	0.6	-0.1
Z_{10}	0.4	0.21
Z_{20}	0	0.80

a) Effectuez le post-conditionnement par krigeage de la simulation aux points x=7 et x=18.

b) Lorsque Z n'est pas normal et que l'on procède par transformation gaussienne Y=g(Z) où Y~N(0,1), doit-on effectuer le conditionnement par krigeage sur la variable Y, avant transformation inverse de la simulation non-conditionnelle, ou sur Z, après transformation inverse de la simulation non-conditionnelle de Y ? Justifiez votre réponse.

Question 7 (18 points)

La figure 1 montre la fonction de répartition obtenue pour l'ensemble d'un gisement de Cu (en %). La figure 2 montre les variogrammes obtenus pour les indicatrices codées par rapport à différents quantiles de la distribution du Cu (10%, 50%, 90%).

Figure 1

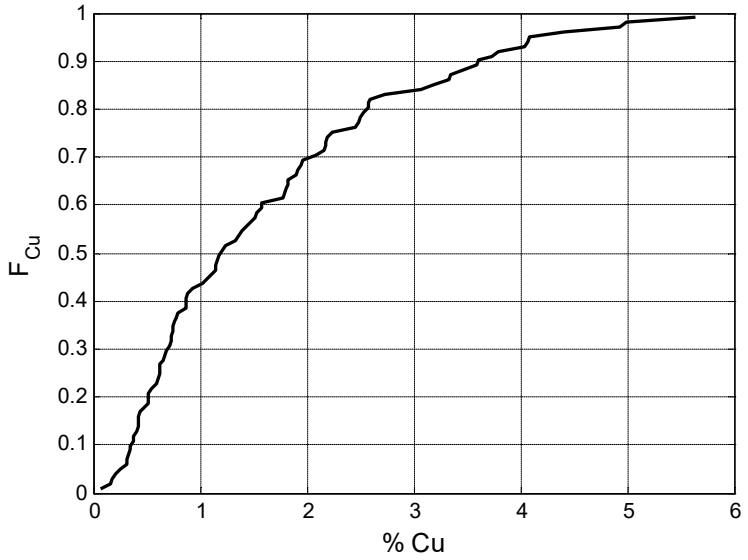
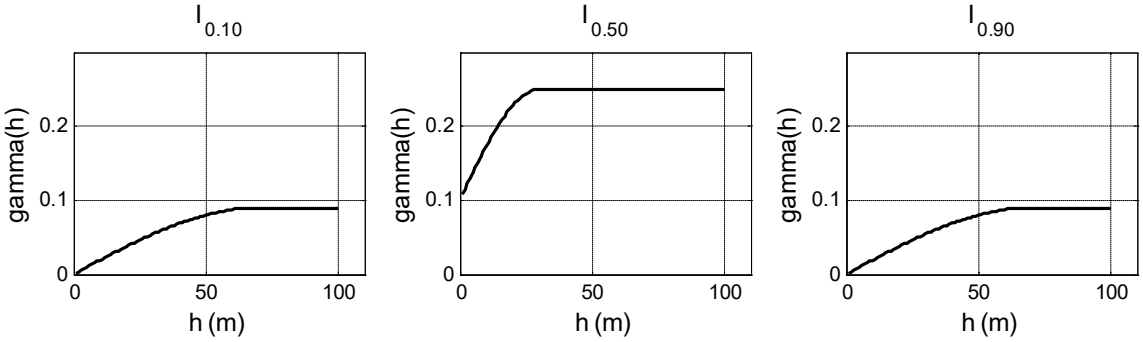


Figure 2



a) L'hypothèse d'une distribution multinormale (après transformation vers la loi normale du %Cu) vous semble-t-elle réaliste à première vue ? Justifiez votre réponse.

b) Quel lien y a-t-il entre les indicatrices codées par rapport aux quantiles 10%, 50% et 90% de la distribution du %Cu et les indicatrices codées par rapport aux mêmes quantiles, mais cette fois de la distribution de Y, où Y est la variable normale obtenue par transformation graphique du %Cu ?

On veut effectuer le krigeage simple d'indicatrices au point x_0 en utilisant les points x_1 à x_3. Les poids de krigeage simple obtenus pour différents seuils sont donnés au tableau suivant.

Point	% Cu au point x_i	poids Ks, au seuil 0.5% Cu	poids Ks, au seuil 1.0 % Cu	poids Ks, au seuil 2.0% Cu
x_1	1.2%	0.3	0.2	0.21
x_2	3%	0.45	0.3	0.4
x_3	0.3%	0.15	0.2	0.23

c) Estimez par krigeage simple d'indicatrices $P(Z_0 < 1.5\%$ Cu) (Z_0 : % Cu en x_0; il n'y a pas de correction d'ordre à effectuer).

Corrigé :

1 a) distance entre x1et x2 $(10^2+2^2)^{0.5}$ =10.198
a1=4*(1-(1.5*10.198/30-0.5*(10.198/30)^3))=2.039
a2=7
a3=1
a4=0
a5=1
a6=0

b) Cu(x0)*= 0.4983*2 -0.1696*5 0.5017*3 0.1696*3 =2.1625%

c) Non, ce modèle n'est pas admissible car le déterminant de la matrice des coefficients est -1.

d) il suffit d'imposer une seule contrainte sur la somme de tous les poids (Cu et Zn) =1. Conséquemment, le système sera celui présenté en a) mais la dernière ligne et la dernière colonne seront la somme des 2 dernières lignes et des 2 dernières colonnes, soit:

$$\begin{array}{c} \begin{matrix} Cu(x1) & Zn(x1) & Cu(x2) & Zn(x3) & \end{matrix} \\ \begin{bmatrix} 7 & 5 & 2.039 & 0.8479 & 1 \\ 5 & 10 & 1.5292 & 1.6958 & 1 \\ 2.039 & 1.5292 & 7 & 1.745 & 1 \\ 0.8479 & 1.6958 & 1.745 & 10 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix} \end{array} \begin{bmatrix} 0.2962 \\ 0.023 \\ 0.4174 \\ 0.2634 \\ -1.3953 \end{bmatrix} = \begin{array}{c} Cu(x0) \\ \begin{bmatrix} 1.8675 \\ 1.4006 \\ 2.6254 \\ 2.2569 \\ 1 \end{bmatrix} \end{array}$$

2- Le cokrigeage n'apportera rien car la variable principale est plus échantillonnée que la variable secondaire.

3- a) $$C(h) = \begin{array}{c} \begin{matrix} Cu & Zn \end{matrix} \\ \begin{bmatrix} 1 & 1.5 \\ 1.5 & 3 \end{bmatrix} \end{array} \delta(h) + \begin{array}{c} \begin{matrix} Cu & Zn \end{matrix} \\ \begin{bmatrix} 5 & -3 \\ -3 & 3 \end{bmatrix} \end{array} Sphérique(a = 30, isotrope)$$

b) en h=0: -1.5/(6*6)$^{0.5}$=-0.25
en h=0+: -3/(6*6)$^{0.5}$=-0.5

4-
a) Cov(Z2,Z3)= 1.4434*1.1547+0.9574*(-0.4874) = 1.20
b) z1=1 $\Rightarrow$ y1=1/1.732=0.5774
z2=1.4434*0.5774+0.9574*Y_2 = 0.83342

c) $(-0.4874)^2+1.1954^2$ = 1.665
d) -0.4874*0.9574 = -0.46664

5- a) Si Z était normal, sa fonction de répartition aurait la même forme que celle de Y
b) Celui de Y
c) Elle dépendra de Y*, et de la variance de krigeage. Comme Y* dépend des valeurs voisines, elle variera. Plus Y* est faible, moins les réalisations seront variables selon la forme de la fonction de répartition.
d) Y=-0.2 par lecture directe.
e) Z=3 par lecture directe.

6- $Z_7^* = 0.6*0.02+0.4*(-0.53) = -0.2$
$Z_s^{s*} = 0.6*(-1.37)+0.4*(-2.74) = -1.918$
$Z_7^{sc} = -0.2 + (-3.16-(-1.918)) = -1.442$

$Z_{18}^* = -0.1*0.02+.21*(-0.53) +.8*(-2.8) = -2.353$
$Z_{18}^{s*} = -0.1*(-1.37)+.21*(-2.74) +.8*(2.24) = 1.3536$
$Z_{18}^{sc} = -2.353 + (1.23-1.3536) = -2.4766$

b) On doit effectuer le conditionnement sur la valeur Y. Sinon, on ne reproduira pas bien l'histogramme puisque l'on ajoutera une erreur simulée à une distribution qui a déjà le bon histogramme. Ainsi, on peut facielemnet se retrouver avec des valeurs négatives même si Z est positif par définition.

7-
a) Non pas réaliste car la corrélation spatiale est plus forte aux quantiles 0.1 et 0.9 qu'à la médiane. Normalement, on attend le contraire.

b) Elles sont identiques. La transformation graphique ne change pas l'ordre des observations dans la séquence ordonnée.

c) On calcule par KS $P(Z_0 <1\%)$ et $P(Z_0<2\%)$ et on interpole linéairement.
$P(Z_0 <1\%) = 0.2+(1-0.7)*0.43 = 0.329$
$P(Z_0<2\%) = 0.21+0.23+(1-0.84)*0.7 = 0.553$
$P(Z_0<1.5\%) = 0.5(-0.329+0.553) = 0.4405$