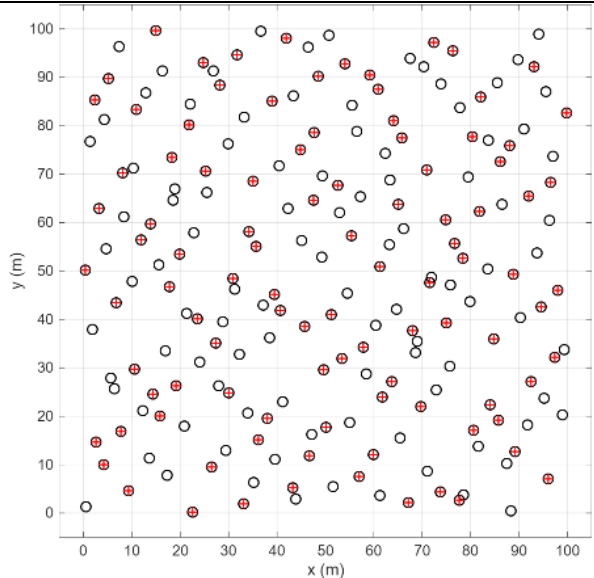
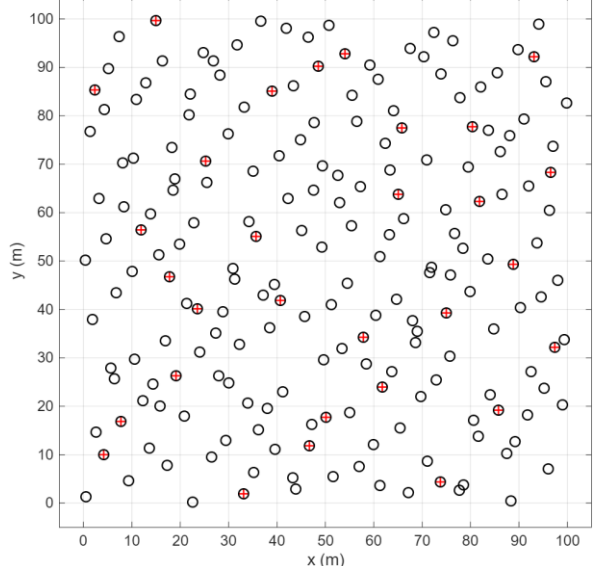
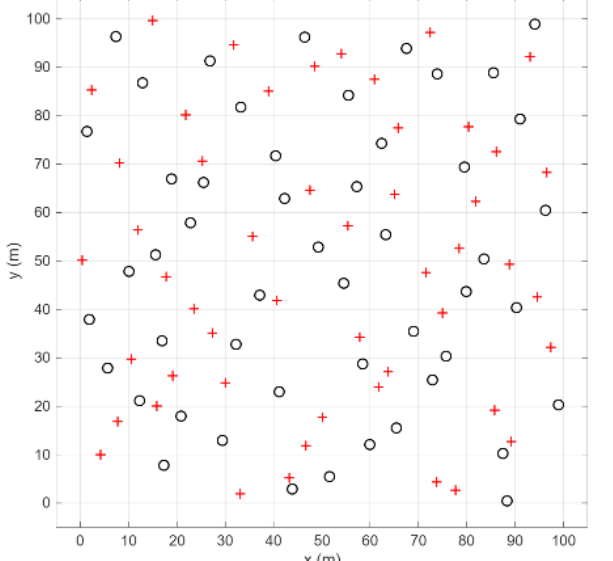
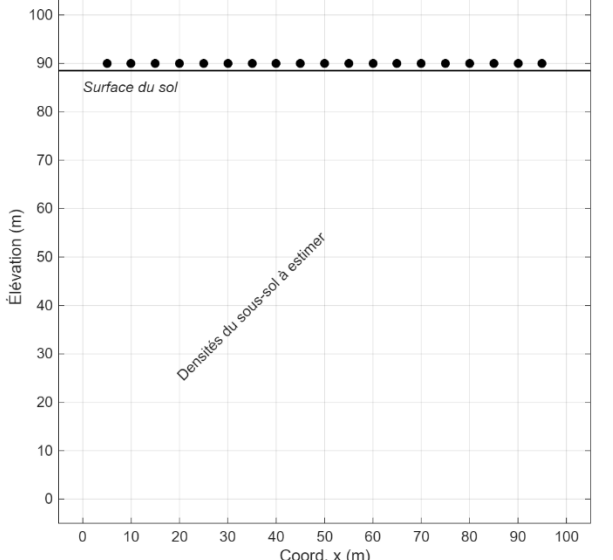


Question 1 (12 points) – Admissibilité et Utilité du Cokrigage

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

Complétez le tableau de la page 3 pour les cas A à D.

Cas	Modèle de covariance	Localisation
A	Modèle linéaire de corégionalisation $C(h) = \begin{bmatrix} 2 & 0.5 \\ 0.5 & 1 \end{bmatrix} \delta(h) + \begin{bmatrix} 4 & 1.2 \\ 1.2 & 6 \end{bmatrix} sph(a = 25)$ $det(B_1) = 2(1) - 0.5(0.5) = 2 - 0.25 = 1.75$ $det(B_2) = 4(6) - 1.2(1.2) = 24 - 1.44 = 22.56$ Admissible, car les deux sont positifs. $\rho(0) = \frac{1.7}{\sqrt{6 \cdot 7}} = 0.2623$ Corrélation trop faible.	
B	Modèle linéaire de corégionalisation $C(h) = \begin{bmatrix} 1 & 0.7 \\ 0.7 & 1.5 \end{bmatrix} \delta(h) + \begin{bmatrix} 3 & 2.1 \\ 2.1 & 4 \end{bmatrix} sph(a = 25)$ $det(B_1) = 1(1.5) - 0.7(0.7) = 1.5 - 0.49 = 1.01$ $det(B_2) = 3(4) - 2.1(2.1) = 12 - 4.41 = 7.59$ Admissible, car les deux sont positifs. $\rho(0) = 2.8 / (\sqrt{4 \cdot 5.5}) = 2.8 / \sqrt{22} \approx 0.597$ Corrélation modérée à forte. Bon pour le cokrigage, car la valeur de Y est beaucoup plus nombreuse que celle de Z.	

C	Modèle linéaire de corégionalisation $C(h) = \begin{bmatrix} 1.2 & 1.3 \\ 1.3 & 1.4 \end{bmatrix} \delta(h) + \begin{bmatrix} 4 & 2 \\ 2 & 5 \end{bmatrix} sph(a = 25)$ $det(B_1) = 1.2(1.4) - 1.3(1.3) = 1.68 - 1.69 = -0.01$ $det(B_2) = 4(5) - 2(2) = 20 - 4 = 16$ Non admissible, car le premier déterminant est négatif. $\rho(0) = 3.3/(\sqrt{(5.2 \cdot 6.4)}) = 3.3/\sqrt{33.28} \approx 3.3/5.77$ ≈ 0.572 Corrélation assez forte (mais le modèle n'est pas admissible).	
D	Z est le contraste de densité de blocs (aucune observation) Y est l'anomalie gravimétrique mesurée en surface La relation physique liant la gravité aux densités est entièrement considérée dans le modèle de corégionalisation, c.-à-d. $Y=GZ \Rightarrow C_{YY} = GC_{ZZ}G^T$ et $C_{YZ} = GC_{ZZ}$ Par construction, le modèle est admissible. Le modèle améliora l'estimation de la variable principale, car nous n'avons pas de mesure.	

Question 1 (suite)

Complétez chacune des cases du tableau suivant en indiquant O (oui) ou N (non) selon que l'énoncé s'applique ou non au cas considéré (A à D).

Note. Une bonne réponse vaut 0.75 point, une mauvaise réponse vaut -0.25 point et l'absence de réponse vaut 0.25 point. Il n'est pas possible d'avoir un pointage inférieur à 0.

Énoncé	Cas			
	A	B	C	D
1) Le modèle de corégionalisation décrit est admissible	O	O	N	O
2) Il est possible de calculer le variogramme croisé Z-Y <u>expérimental</u> à partir des données indiquées.	O	O	N	N
3) Selon le modèle décrit, le cokrigage devrait améliorer <u>notablement</u> l'estimation par rapport au krigeage.	N	O	N	O
4) Indépendamment du modèle décrit et en supposant qu'un modèle admissible soit utilisé, seul le cokrigage simple peut être effectué. (Note : les moyennes des deux variables sont supposées connues)	N	N	N	O

Question 2 (12 points) – Théorie sur le Cokrigage

On considère deux variables régionalisées, $Z(x)$ et $Y(x)$. On suppose que la moyenne de la variable Y est connue et notée m_Y , alors que la moyenne de Z est inconnue.

- 3 pts a) Formez un estimateur non biaisé de la variable Z en un point x_0 , qui tienne compte à la fois des données observées de Z et de Y .

$$Z^*(x_0) = \sum \lambda_i Z_i + \sum \alpha_i (Y_i - m_Y)$$

- 2 pts b) Quels seront les multiplicateurs de Lagrange requis, ainsi que les contraintes associées, pour assurer que l'estimateur de cokrigage obtenu en a) soit non biaisé ?

$$\sum \lambda_i = 1$$

- 3 pts c) En supposant que l'on dispose de 3 observations au voisinage du point x_0 , noté Z_1 , Y_1 , et Y_2 , présentez le système matriciel de cokrigage correspondant sous forme matricielle.

$$\begin{pmatrix} \text{Var}(Z_1) & \text{Cov}(Z_1, Y_1) & \text{Cov}(Z_1, Y_2) & 1 \\ \text{Cov}(Y_1, Z_1) & \text{Var}(Y_1) & \text{Cov}(Y_1, Y_2) & 0 \\ \text{Cov}(Y_2, Z_1) & \text{Cov}(Y_2, Y_1) & \text{Var}(Y_2) & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \alpha_1 \\ \alpha_2 \\ \mu_Z \end{pmatrix} = \begin{pmatrix} \text{Cov}(Z_1, Z_0) \\ \text{Cov}(Y_1, Z_0) \\ \text{Cov}(Y_2, Z_0) \\ 1 \end{pmatrix}$$

Question 2 (suite)

- 2 pts d) La variance de cokrigage obtenue avec ce système (intégrant Y comme variable secondaire) sera-t-elle inférieure, égale ou supérieure à la variance d'un krigeage ordinaire utilisant seulement Z ? Justifiez votre réponse sans faire de calculs.

Elle sera inférieure. La variance de cokrigage ordinaire est inférieure ou égale à celle du krigeage ordinaire.

La variance obtenue avec notre estimateur se situe entre celles du cokrigage simple et du cokrigage ordinaire. Ainsi, on peut conclure qu'elle sera inférieure.

- 2 pts e) Comment sera modifié le système matriciel de cokrigage de la question b, si on veut estimer la variable Y en un point x_0 ?

Changer le terme de droite seulement.

$$\begin{pmatrix} \text{Var}(Z_1) & \text{Cov}(Z_1, Y_1) & \text{Cov}(Z_1, Y_2) & 1 \\ \text{Cov}(Y_1, Z_1) & \text{Var}(Y_1) & \text{Cov}(Y_1, Y_2) & 0 \\ \text{Cov}(Y_2, Z_1) & \text{Cov}(Y_2, Y_1) & \text{Var}(Y_2) & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \alpha_1 \\ \alpha_2 \\ \mu_Z \end{pmatrix} = \begin{pmatrix} \text{Cov}(Z_1, Y_0) \\ \text{Cov}(Y_1, Y_0) \\ \text{Cov}(Y_2, Y_0) \\ 1 \end{pmatrix}$$

Question 3 (10 points) – Krigeage Ordinaire d'Indicatrices avec Contrainte d'Inégalité

Quatre puits d'exploration ont été forés. Les trois premiers ont recoupé le sommet de la minéralisation à des profondeurs de 305 m, 333 m et 309 m (les profondeurs étant mesurées positivement vers le bas). Le quatrième forage a dû être arrêté à 318 m, avant d'atteindre le sommet de la minéralisation, et aucune autre information n'est disponible à ce stade.

Le variogramme de la profondeur est sphérique avec $c_0 = 7 \text{ m}^2$, $c_1 = 29 \text{ m}^2$ et $a = 132 \text{ m}$. Le variogramme du log(profondeur) est proportionnel à celui-ci, de même que les variogrammes des différentes indicatrices, ce qui implique que les poids de krigeage restent invariants. Lorsque l'on effectue le krigeage ordinaire de la profondeur en utilisant les points x_1 à x_4 , d'une part, et x_1 à x_3 d'autre part, on obtient les poids suivants :

Point	Poids de krigeage de x_0 utilisant x_1 à x_4	Poids de krigeage de x_0 utilisant x_1 à x_3
x_1	0.2	0.25
x_2	0.15	0.20
x_3	0.3	0.55
x_4	0.35	--

Utilisez toutes les informations disponibles (incluant le forage stoppé en x_4) pour prédire, par krigeage ordinaire d'indicatrices, l'espérance conditionnelle de la profondeur du sommet de la minéralisation au point x_0 ainsi que la variance conditionnelle associée. Considérez les seuils 300 m, 310 m, 320 m, 330 m et 340 m dans votre krigeage ordinaire d'indicatrices et utilisez les milieux des classes comme valeurs représentatives. Rappel. On a : $Z(x_1) = 305 \text{ m}$; $Z(x_2) = 333 \text{ m}$; $Z(x_3) = 309 \text{ m}$ et $Z(x_4) > 318 \text{ m}$. (Il y a une page vierge de l'autre côté pour continuer les calculs).

On commence par coder les variables relatives aux différences de seuil.

	300	310	320	330	340
z1-305	0	1	1	1	1
z2-333	0	0	0	0	1
z3-309	0	1	1	1	1
z4-318	0	0	-	-	-
Fz	0	.5	.8	.80	1

Par la suite, les milieux de la classe seront 305, 315, 325 et 335.

Ainsi l'espérance conditionnelle est : $E[X] = x_1 p_1 + x_2 p_2 + \dots + x_n p_n = 0.5 * 305 + 0.3 * 315 + 0.2 * 335 = 314 \text{ m}$

Et la variance conditionnelle est : $Var[X] = (x_1^2 p_1 + x_2^2 p_2 + \dots + x_n^2 p_n) - E[x]^2 = 0.5 * 305^2 + 0.3 * 315^2 + 0.2 * 335^2 - 314^2 = 129 \text{ m}$

Question 4 (16 points) – Décomposition de Cholesky (Méthode LU)

On simule une variable Z à 4 emplacements à l'aide de la décomposition de Cholesky. Les emplacements sont placés dans l'ordre 1, 2, 3, 4 dans la matrice de covariance. La matrice L obtenue est la suivante :

$$L = \begin{bmatrix} 2.2361 & 0 & 0 & 0 \\ 1.7889 & 1.3416 & 0 & 0 \\ 1.3416 & 0.4472 & 1.7321 & 0 \\ 0.8944 & 1.0435 & 1.3472 & 1.1386 \end{bmatrix}$$

2 pts a) Quelle est la variance du processus stochastique Z ?

$$\text{Var}(Z) = 2.2361 * 2.2361 = 5$$

2 pts b) Quelle est la covariance entre Z_3 et Z_4 ?

$$\text{Cov}(Z_3, Z_4) = 1.3416 * 0.8944 + 0.4472 * 1.0432 + 1.7321 * 1.3472 + 0 * 1.1386 = 4$$

3 pts c) On a observé $Z_1 = 1$ ainsi que $Z_2 = 3$. Déterminez les valeurs de Y_1 et Y_2 qui permettent d'assurer un conditionnement exact par la méthode de Cholesky.

$$Y_1 = \frac{Z_1}{L_{11}} = \frac{1}{2.2361} = 0.44721$$

$$Y_2 = \frac{Z_2 - L_{12}Y_1}{L_{21}} = \frac{3 - 1.7889 * 0.44721}{1.3416} = 1.6466$$

Question 4 (suite)

4 pts d) Quelle est l'espérance conditionnelle de Z_4 sachant que $Z_1 = 1$ et $Z_2 = 3$? (Posez $Y_1 = 1$ et $Y_2 = -1$, si vous n'avez pas répondu à la question 4c.)

$$\begin{aligned}
 E[Z_4|Z_1, Z_2] &= E[L_{14}Y_1 + L_{24}Y_2 + L_{34}Y_3 + L_{44}Y_4] \\
 &= E[L_{14}Y_1] + E[L_{24}Y_2] + E[L_{34}Y_3] + E[L_{44}Y_4] \\
 &= L_{14}E[Y_1] + L_{24}E[Y_2] + L_{34}E[Y_3] + L_{44}E[Y_4] \\
 &= 0.8944 * 0.44721 + 1.0435 * 1.6398 + 1.3472 * 0 + 1.1386 * 0 \\
 &= 2.1111
 \end{aligned}$$

3 pts e) Quel est la covariance conditionnelle entre Z_3 et Z_4 ?

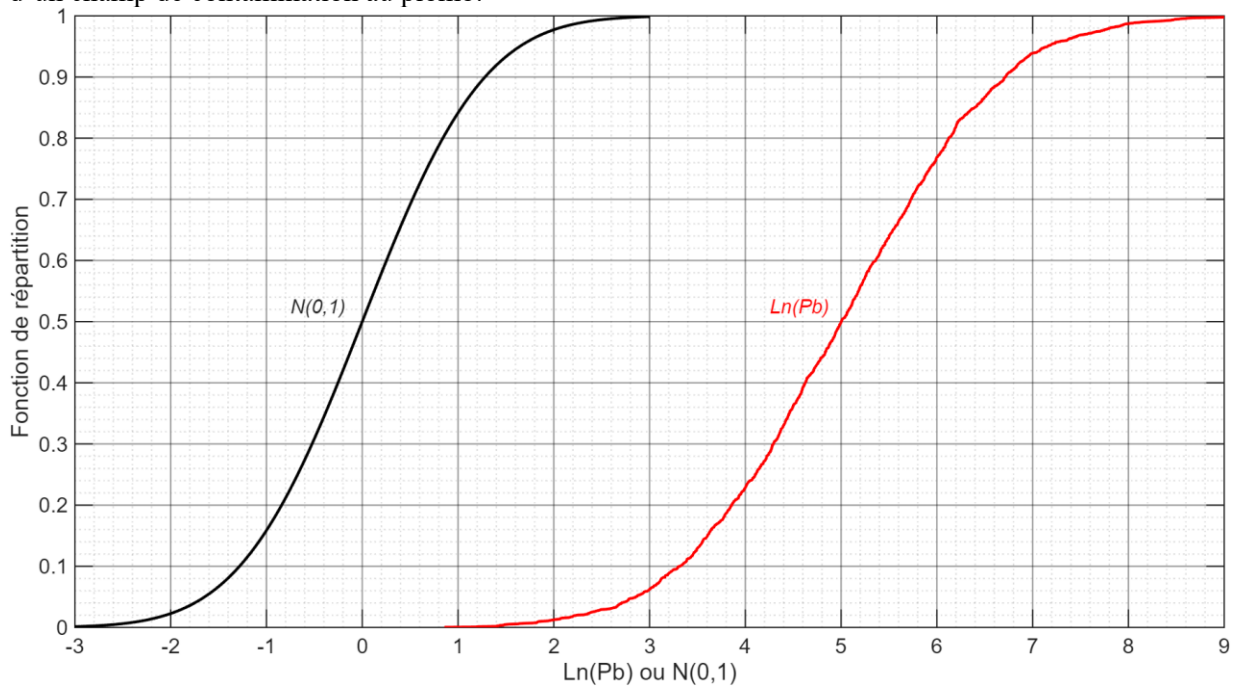
$$Cov(Z_3, Z_4|Z_1, Z_2) = 1.7321 * 1.3472 + 0 * 1.1386 = 2.3335$$

2 pts f) On a généré le vecteur aléatoire $[Y_3 \ Y_4] = [0.48, -0.19]$. Que vaut Z_4 (conditionnelle à Z_1 et Z_2)? (Posez $Y_1 = 1$ et $Y_2 = -1$, si vous n'avez pas répondu à la question 4c.)

$$\begin{aligned}
 Z_4 &= L_{14}Y_1 + L_{24}Y_2 + L_{34}Y_3 + L_{44}Y_4 \\
 &= 0.8944 * 0.44721 + 1.0435 * 1.6398 + 1.3472 * 0.48 + 1.1386 * -0.19 \\
 &= 2.5414
 \end{aligned}$$

Question 5 (12 points) – Simulation Séquentielle Gaussienne

La figure ci-bas montre les fonctions de répartition d'une $N(0,1)$ et du $\ln(\text{pb})$ obtenu à partir d'échantillon d'un champ de contamination au plomb.



2 pts a) On a observé trois données ayant les localisations et valeurs suivantes. Complétez le tableau suivant:

Donnée	Coord. X (m)	Coord. Y (m)	Pb (ppm)	$\ln(\text{Pb})$	$N(0,1)$
1	425	425	221.41	5.4	0.25
2	575	425	44.70	3.8	-0.8
3	625	625	403.43	6	0.75

La variable $Z \sim N(0,1)$ correspondant aux teneurs de Pb montre un variogramme sphérique isotrope de portée 400 m et avec $C_0 = 0.1$ et $C_1 = 0.9$.

2 pts b) Quelles sont les unités de C_0 et de C_1 ?

Il n'y a aucune unité

On effectue le krigeage simple au point $x_0 = (600, 600)$ et l'on obtient le système de krigeage suivant:

$$\begin{bmatrix} 1 & 0.4175 & 0.1045 \\ 0.4175 & 1 & 0.2658 \\ 0.1045 & 0.2658 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} = \begin{bmatrix} 0.1713 \\ 0.3422 \\ 0.7810 \end{bmatrix}, \text{ ce qui donne } \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} = \begin{bmatrix} 0.0403 \\ 0.1280 \\ 0.7428 \end{bmatrix}$$

Question 5 (suite)

2 pts c) Quelles seront la valeur krigée et la variance de krigeage simple en ce point x_0 ? (Si vous n'avez pas répondu à la question 5a, posez que respectivement les valeurs suivantes 0.25, -0.75, 0.5, pour les données 1, 2 et 3 dans l'espace normale.)

$$Z^* = \sum \lambda_i Z_i = 0.0403 * 0.25 + 0.1280 * -0.8 + 0.7428 * 0.75 = 0.4647$$

$$\begin{aligned} \sigma_{KS}^2 &= \sigma^2 - \sum \lambda_i \text{Cov}(Z_i, Z_0) = 1 - (0.0403 * 0.1713 + 0.1280 * 0.3422 + 0.7428 * 0.7810) \\ &= 0.3691 \end{aligned}$$

4 pts d) On tire d'une distribution uniforme la valeur $0.7 \sim \mathcal{U}(0,1)$. Quelle est la valeur gaussienne simulée à ce point?

Par lecture de la table de la loi normale, on a pour $F_Z(z) = 0.7 \rightarrow z = 0.525$

$$z^* = z\sigma_{KS} + Z^* = 0.525 * \sqrt{0.3691} + 0.4647 = 0.7836$$

2 pts e) Quelle est la concentration de plomb que l'on simule à ce point ? (Posez la valeur gaussienne simulée à 1.4 si vous n'avez pas pu répondre à la question 5d.)

Par lecture graphique, on a environ $\ln(\text{Pb})=6$. Donc la concentration de plomb est 403.42 ppm.

Question 6 (12 points) – Propriétés des Simulations Géostatistiques

Pour les questions a) à f), écrivez uniquement, sous forme mathématique, la propriété en jeu. Par exemple : $E[Z_{\text{sim}}] = \text{_____}$.

Pour a) à d), on a un champ en 2D de taille G pour lequel on génère un très grand nombre de **simulations non-conditionnelles** sur une grille serrée de points. Les simulations sont effectuées avec moyenne théorique « m » et la variance théorique « σ^2 ».

- 2 pts a) On se place à un point donné x_0 et on examine les valeurs obtenues pour les différentes réalisations. Quelles sont la moyenne et la variance des valeurs simulées à ce point?

$$E[Z] = m ; \text{Var}(Z) = \sigma^2$$

- 2 pts 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 pour chaque réalisation?

$$D^2(\cdot | G)$$

- 2 pts c) Pour chaque réalisation, on regroupe les données en blocs de taille V (c.-à-d. 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?

$$E[Z] = m ; \text{Var}(Z) = \sigma_V^2$$

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

$$D^2(V|G)$$

Question 6 (suite)

Pour e) et f), on considère, cette fois, un très grand nombre de **simulations conditionnelles** sur une grille de points très serrée toujours sur un champ en 2D de taille G .

- 2 pts e) On se place à un point donné x_0 et on examine les valeurs obtenues pour les différentes réalisations. Quelles sont la moyenne et la variance des valeurs simulées à ce point?

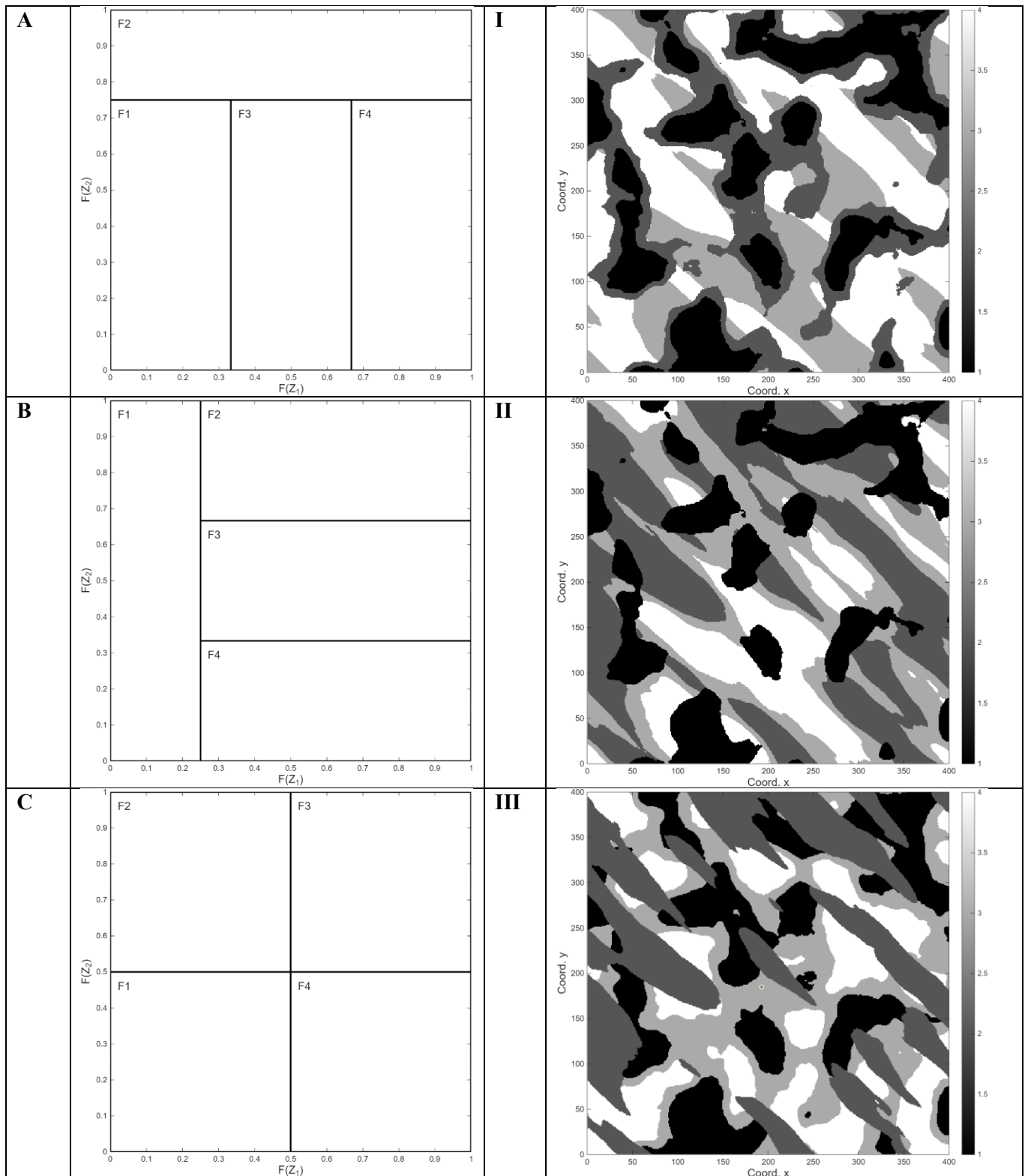
$$E[Z] = Z^* ; \text{Var}(Z) = \sigma_{KS}^2$$

- 2 pts f) Pour chaque réalisation, on regroupe les valeurs simulées en blocs de taille V (c.-à-d. 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?

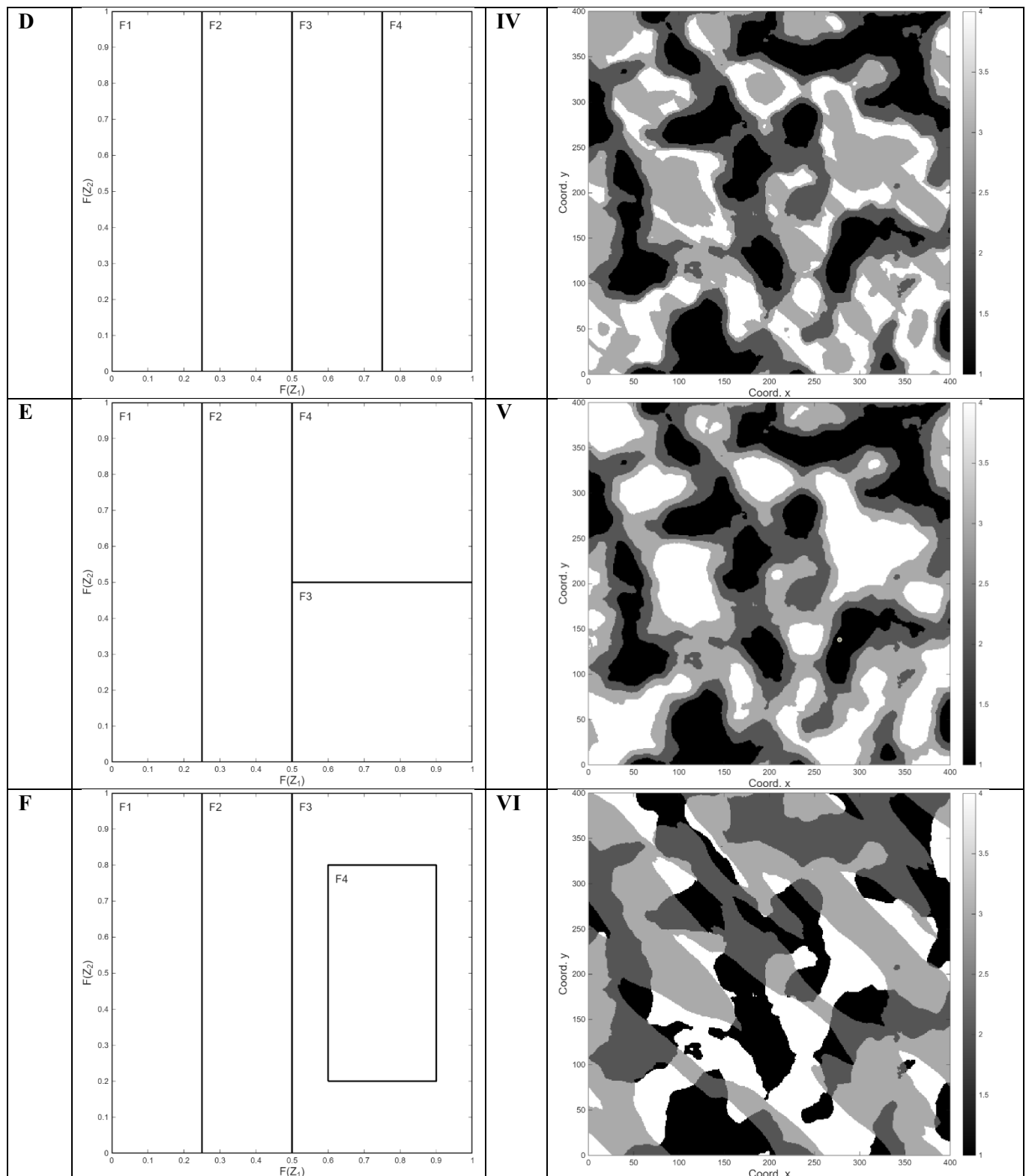
$$E[Z] = Z^* ; \text{Var}(Z) = \sigma_{V,KS}^2$$

Question 7 (10 points) – Simulations Plurigaussiennes

Complétez le tableau de la page 16 pour les cas A à F.



Question 7 (suite)



Question 7 (suite)

8 pts a) Associez chaque réalisation au drapeau correspondant.

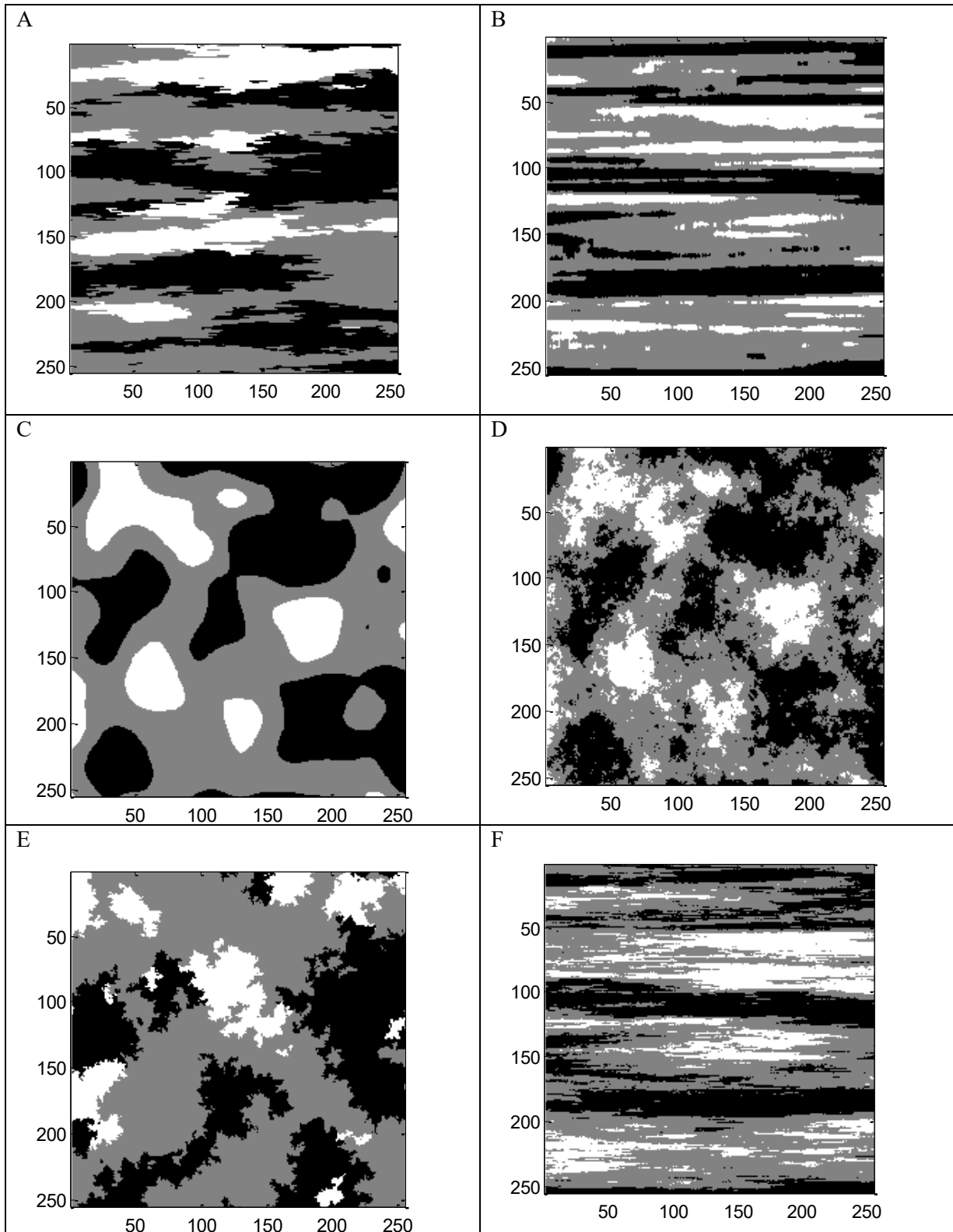
Réalisations	A	B	C	D	E	F
Drapeau	III	II	VI	V	I	IV

2 pts b) Le variogramme utilisé est cubique pour les deux variables. Lequel des énoncés représente la structure spatiale des gaussiennes simulées (ne pas justifiez votre choix) :

Énoncé	Type modèle gaussienne 1	Type modèle gaussienne 2
I)	Isotrope ($a=75$ m)	Isotrope ($a=75$ m)
II)	Isotrope ($a=75$ m)	Anisotrope ($ag=200$ m, $ap=50$ m, $\theta_g=135^\circ$)
III)	Anisotrope ($ag=200$ m, $ap=50$ m, $\theta_g=135^\circ$)	Isotrope ($a=75$ m)
IV)	Anisotrope ($ag=200$ m, $ap=50$ m, $\theta_g=135^\circ$)	Anisotrope ($ag=200$ m, $ap=50$ m, $\theta_g=135^\circ$)
V)	Isotrope ($a=75$ m)	Anisotrope ($ag=50$ m, $ap=25$ m, $\theta_g=135^\circ$)
VI)	Anisotrope ($ag=50$ m, $ap=25$ m, $\theta_g=135^\circ$)	Isotrope ($a=75$ m)

θ_g : azimute de la grande portée, ag : grande portée, ap : petite portée

Question 8 (10 points) – Simulation d'Indicatrices et Simulations Gaussiennes Tronquées
Complétez le tableau de la page 18 pour les cas A à F.



Question 8 (suite)

6 pts a) Complétez le tableau suivant en marquant d'un X, pour chaque image identifiée de A à F, les énoncés qui s'appliquent. (Aide : chaque colonne doit comporter deux "X" au total : un en haut de la double barre et un en bas de celle-ci)

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				X

2 pts b) Dans un modèle gaussien tronqué, quelles proportions de facies F_1 , F_2 et F_3 seront simulées théoriquement si l'on adopte les seuils $c_1 = -0.85$ et $c_2 = 0.78$ pour une variable $N(0,1)$ simulée ?

Par lecture graphique de la table de la loi normale :

- Pour c_1 , on lit la valeur à 0,85 et on obtient une probabilité de 0,8023. On prend ensuite le complément : $1 - 0,8023 = 0,1976$.
- Pour c_2 , on lit la valeur à 0,78 et on obtient 0,7823.

Les proportions des faciès sont alors :

- **F1** : 0,1976 (soit 19,76 %)
- **F2** : $0,7823 - 0,1976 = 0,5847$ (soit 58,47 %)
- **F3** : $1 - 0,7823 = 0,2177$ (soit 21,77 %)

Question 9 (8 points) – Application de la Géostatistique dans une Étude de Cas Réel

Vous devez appuyer une équipe d'ingénierie environnementale dans l'évaluation du risque lié à la présence d'arsenic (As) autour d'une fonderie en activité. Selon les normes en vigueur, toute zone dont la concentration en As dépasse 50 ppm est considérée comme contaminée par l'As, car cette valeur ne respecte pas le critère C du *Guide d'intervention – Protection des sols et réhabilitation des terrains contaminés*.

Une campagne d'échantillonnage a permis de mesurer les concentrations d'As dans 148 forages environnementaux répartis de manière non uniforme autour de la zone d'étude. Les résultats montrent une forte hétérogénéité spatiale, avec plusieurs valeurs élevées à proximité de la fonderie.

L'objectif du projet est de fournir au gouvernement, par le mandat confié à votre firme, des estimations réalistes et quantifiables de la contamination afin d'identifier l'étendue de la zone contaminée et d'évaluer les budgets associés à la décontamination du site.

Plus précisément, on vous demande de :

- Cartographier la probabilité que la concentration dépasse 50 ppm en tout point du domaine d'étude ;
- Fournir une estimation de la proportion totale de la surface où le seuil de 50 ppm est dépassé, accompagnée d'une mesure d'incertitude conditionnelle (par exemple, un intervalle de confiance à 95 %).

6 pts a) Votre travail consiste à proposer une méthodologie, basée sur les outils géostatistique vus en classe, qui permet de répondre à ces questions. Répondez par une liste à puce.

Voici un exemple pour des simulations géostatistiques

- **1. Transformer les données**
 - Transformer les données (par exemple, en log) pour obtenir une variable presque gaussienne ou gaussienne par une transformation graphique.
- **2. Modéliser le variogramme**
 - Calculer le variogramme expérimental de la variable transformée.
 - Ajuster un modèle (ex. sphérique) à utiliser pour la simulation.
- **3. Faire des simulations géostatistiques**
 - Utiliser la Simulation Séquentielle Gaussienne (SSG) ou la Cholesky (LU) pour produire plusieurs réalisations conditionnelles (ex. 100 simulations).
 - Chaque simulation représente un scénario possible du champ d'As autour de la fonderie.
- **4. Réaliser la transformation inverse**
 - Transformer les données vers l'espace original.
- **5. Calculer la probabilité de dépassement du seuil de 50 ppm**
 - Pour chaque point de la grille :
 - compter combien de simulations dépassent 50 ppm ;
 - diviser par le nombre total de simulations.
 - Cela donne une carte de probabilité de contamination.
- **6. Estimer la proportion de surface contaminée**
 - Pour chaque simulation :
 - calculer la portion de la surface où $As > 50$.
 - Faire une moyenne sur toutes les simulations.
 - Utiliser les percentiles (2.5 % – 97.5 %) pour obtenir un intervalle de confiance à 95 %.

Voici un exemple par le krigeage d'indicateurs

1. Choisir plusieurs seuils de teneur en As

Exemple: 10 ppm, 25 ppm, 50 ppm, 75 ppm, 100 ppm, 150ppm, etc

Ces seuils découpent la distribution en classes. Déterminer selon l'histogramme des données

2. Construire une indicatrice pour chaque seuil

Pour chaque seuil z_k , définir :

$$I_k(x) = \begin{cases} 1 & \text{si } As(x) < z_k \\ 0 & \text{sinon} \end{cases}$$

3. Modéliser un variogramme d'indicateur pour chaque seuil

- Pour chaque I_k , calculer un variogramme expérimental.
- Ajuster un modèle théorique

4. Faire le krigeage d'indicateurs pour chaque seuil

- Pour chaque point de la grille et pour chaque seuil :

$$\hat{I}_k(x) \approx P(As < z_k \mid \text{données})$$

- On obtient donc une série de probabilités monotones :

$$P(As < z_1), P(As < z_2), \dots, P(As < z_K)$$

5. Reconstruire la distribution locale complète

- Les probabilités cumulatives estimées servent à reconstruire une CDF locale.
- Entre les seuils, on interpole linéairement pour obtenir une CDF continue.
- Ne pas oublier de corriger pour les relations d'ordres

6. En déduire ce qui est demandé dans le projet

- La probabilité $As > 50$ ppm est directement $1 - \hat{I}_{50}(x)$.
- La proportion totale de surface contaminée est la moyenne spatiale de cette probabilité :

$$\text{Proportion contaminée} = \frac{1}{N} \sum_x (1 - \hat{I}_{50}(x))$$

7. Incertitude

- On peut déduire l'intervalle de confiance directement à partir de la CDF locale : il suffit d'identifier les seuils pour lesquels la probabilité cumulée atteint 2,5 % et 97,5 %. Ces deux valeurs donnent alors l'intervalle de confiance à 95 %.

Par ailleurs, la concentration en cuivre (Cu) a également été caractérisée en 248 localisations (148 données colocalisées avec l'As), et l'on observe une forte corrélation spatiale et statistique entre les concentrations en Cu et en As.

- 2 pts b) Comment modifieriez-vous votre algorithme ou votre démarche pour tenir compte de cette variable secondaire (Cu) lors de l'estimation des probabilités de dépassement et de l'incertitude associée ? Modifiez seulement la puce de la question a.

Selon la méthodologie proposée, il est nécessaire d'adapter l'algorithme géostatistique au cas multivariable. Dans tous les cas, il suffit de remplacer les matrices de krigeage par les matrices de cokrigeage. Ainsi, pour les indicatrices, on utilisera le cokrigeage d'indicatrices, tandis que pour LU et SGS, on utilisera la matrice de cokrigeage simple.

Annexe

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