

# **Caractérisation géostatistique de pollutions industrielles de sols**

**Cas des hydrocarbures aromatiques polycycliques  
sur d'anciens sites de cokeries**

**Nicolas Jeannée**



**Soutenance de thèse  
- Fontainebleau, 30 mai 2001 -**



## Contexte

- En France, 200 000 à 300 000 anciens sites industriels et d'activités de services (potentiellement pollués)
- Quantification de la pollution : nécessaire à l'audit de site et la réhabilitation
- Estimations empreintes d'empirisme, basées sur l'information historique

## Objectif

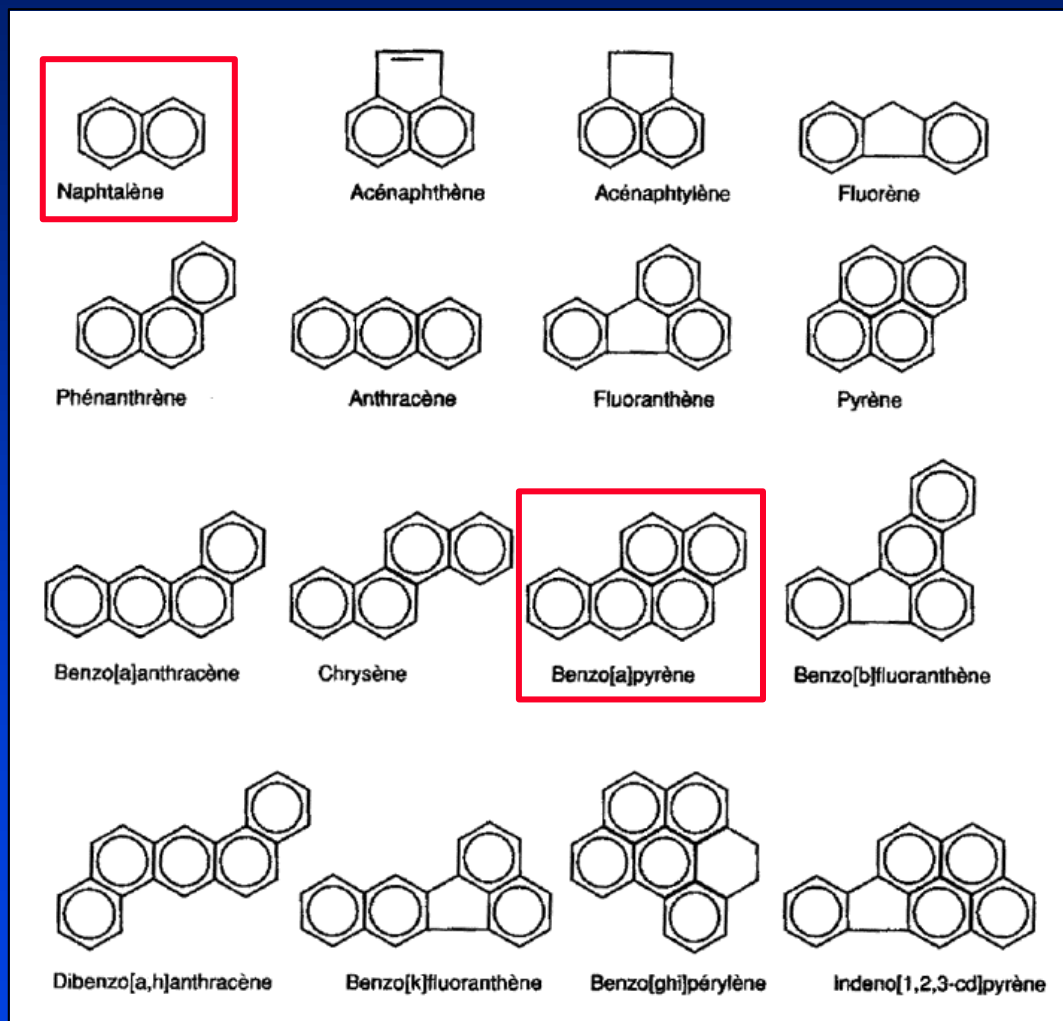
**Qu'apporte la géostatistique à la caractérisation de pollutions industrielles de sols ?**



**Recommandations et méthodes pour l'investigation d'un site industriel pollué**

# Problèmes rencontrés

- **Site** : zone d'épandage d'une ancienne cokerie
- Pollution par des Hydrocarbures Aromatiques Polycycliques (HAP) (45-70 % des goudrons de houille)
- 16 HAP analysés (liste US EPA)



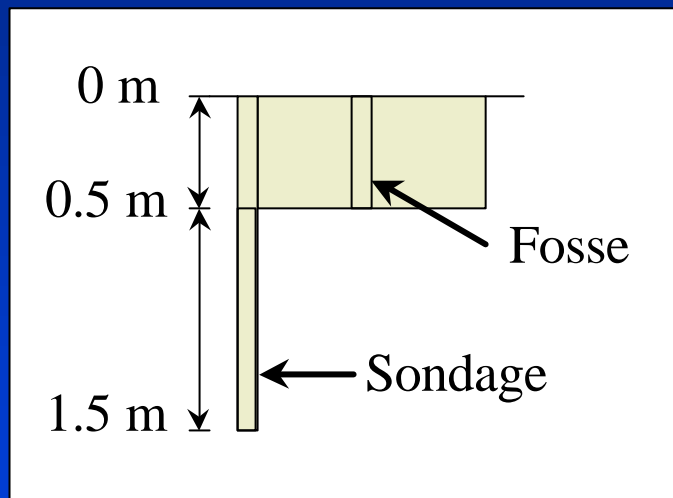
# Echantillonnage

- Echantillonnage non préférentiel

- Deux prélèvements par point :

Fosses

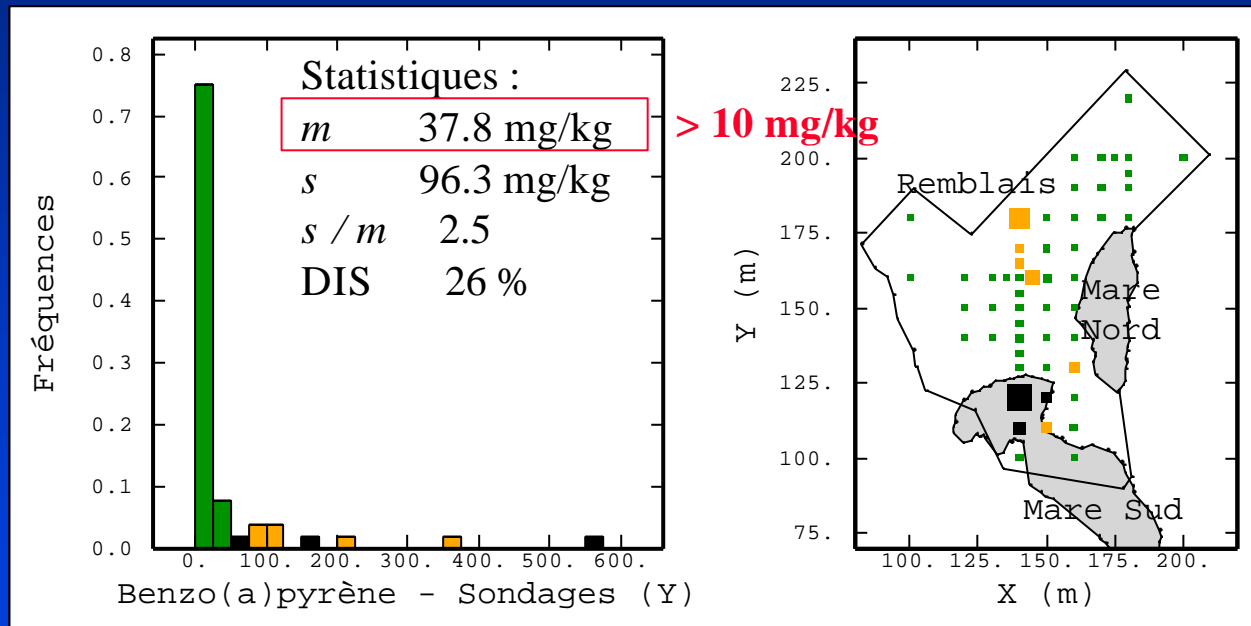
Sondages



- Relevé d'indices qualitatifs et organoleptiques

# Exemple : le benzo(a)pyrène (sondages)

- HAP à 5 cycles
- Histogramme et implantation
- Liens avec l'étude historique



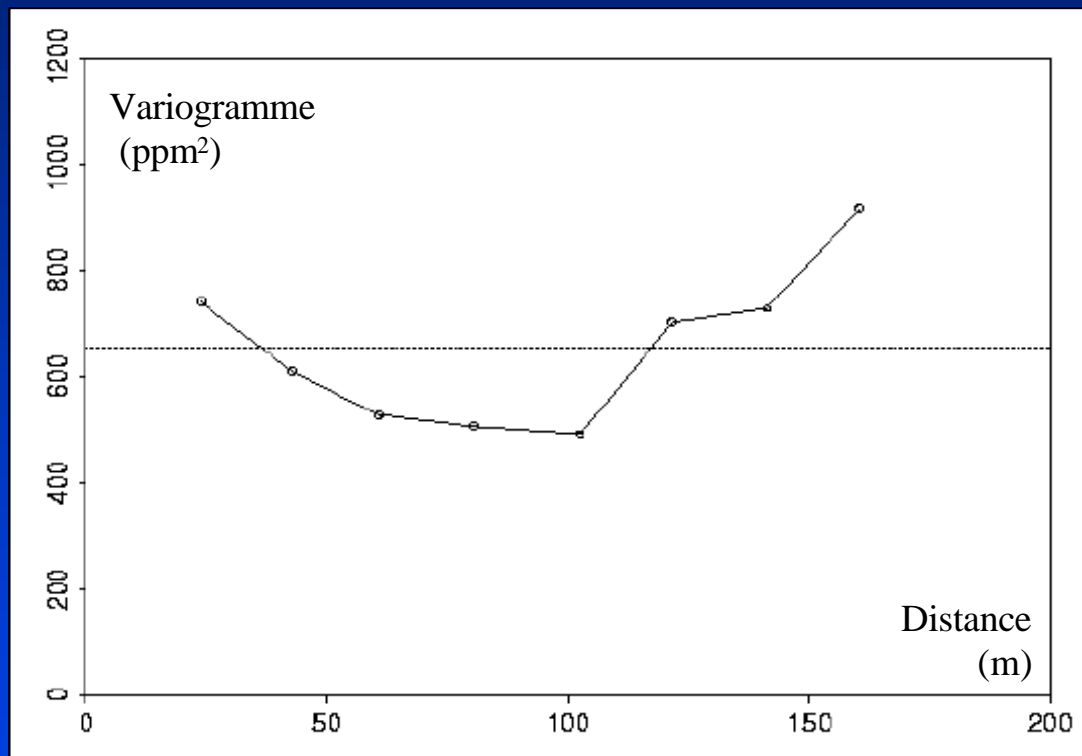
## Problèmes :

- Contraste des teneurs
- Répartition hétérogène des concentrations

→ Détection non exhaustive des zones de forte concentration

# Exemple : le benzo(a)pyrène (sondages)

- Structure spatiale :

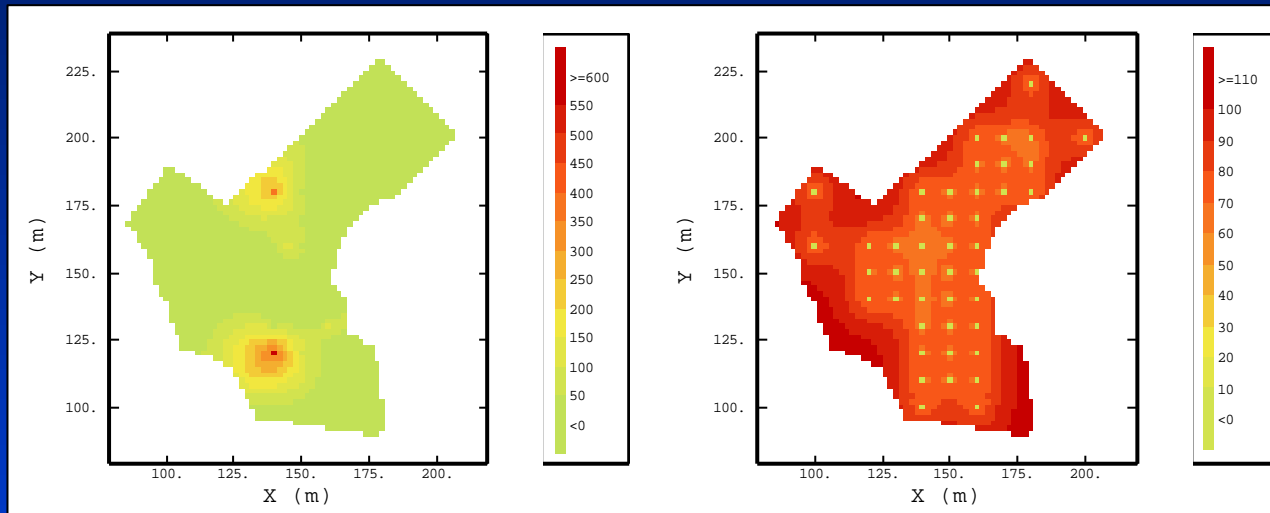


## Problèmes :

- Contraste des teneurs
- Répartition hétérogène des concentrations
- Peu de robustesse du variogramme classique
- Origines de l'effet de pépité ?

# Exemple : le benzo(a)pyrène (sondages)

- Estimation des concentrations en place :



Krigage ordinaire

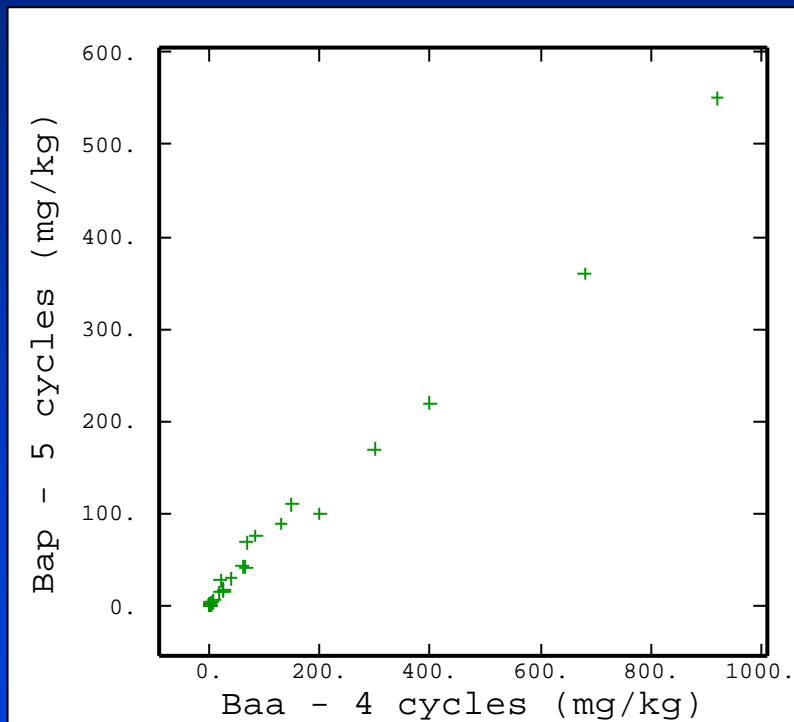
Ecart-types d'estimation

## Problèmes :

- Contraste des teneurs
- Répartition hétérogène des concentrations
- Peu de robustesse du variogramme classique
- Origines de l'effet de pépité ?
- Comment améliorer l'estimation ?

## Exemple : le benzo(a)pyrène (sondages)

- Estimation des concentrations en place :
- Corrélations entre les teneurs en HAP



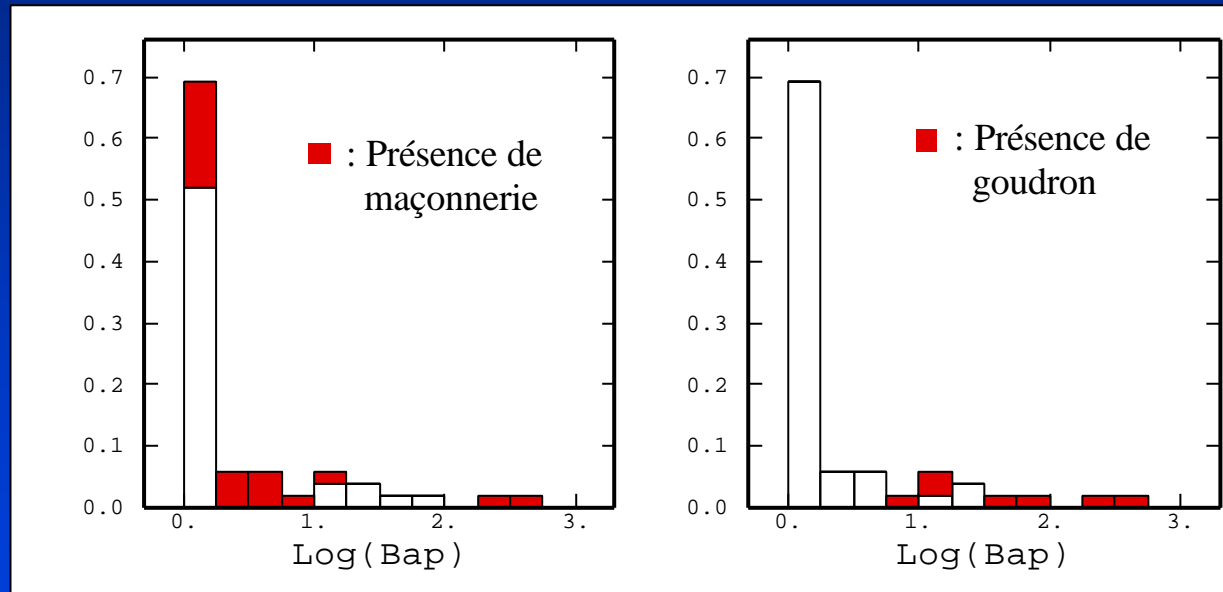
### Problèmes :

- Contraste des teneurs
- Répartition hétérogène des concentrations
- Peu de robustesse du variogramme classique
- Origines de l'effet de pépité ?
- Comment améliorer l'estimation ?



# Exemple : le benzo(a)pyrène (sondages)

- Estimation des concentrations en place :
  - Corrélations entre les teneurs en HAP
  - Corrélations avec des indices qualitatifs

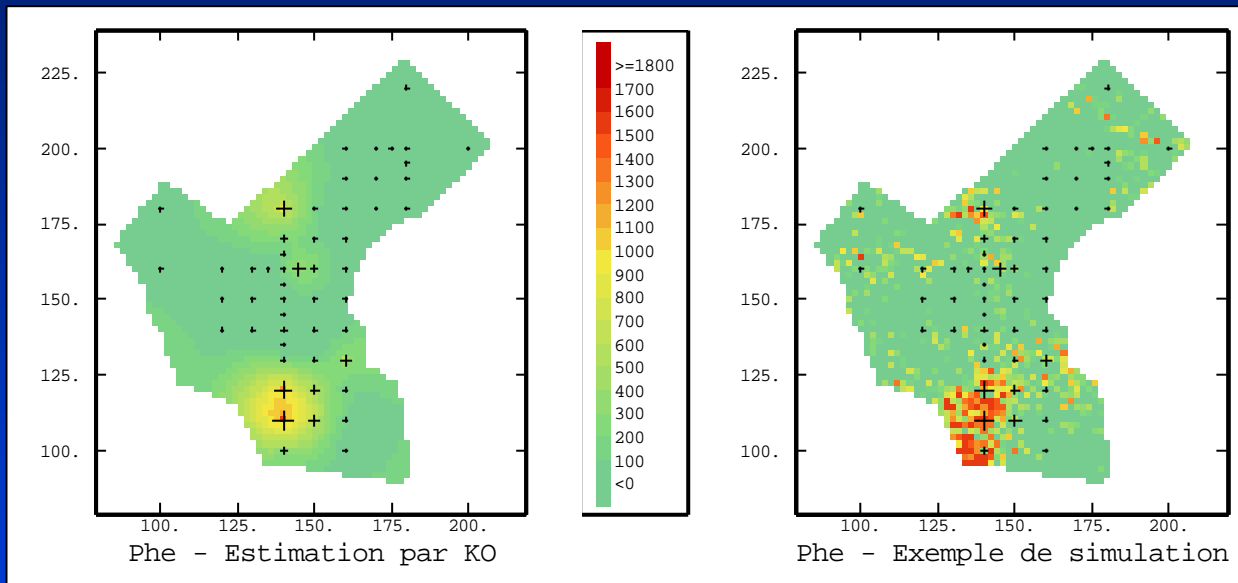


## Problèmes :

- Contraste des teneurs
- Répartition hétérogène des concentrations
- Peu de robustesse du variogramme classique
- Origines de l'effet de pépité ?
- Comment améliorer l'estimation ?

# Exemple : le phénanthrène (sondages)

- Sélection de zones à dépolluer :

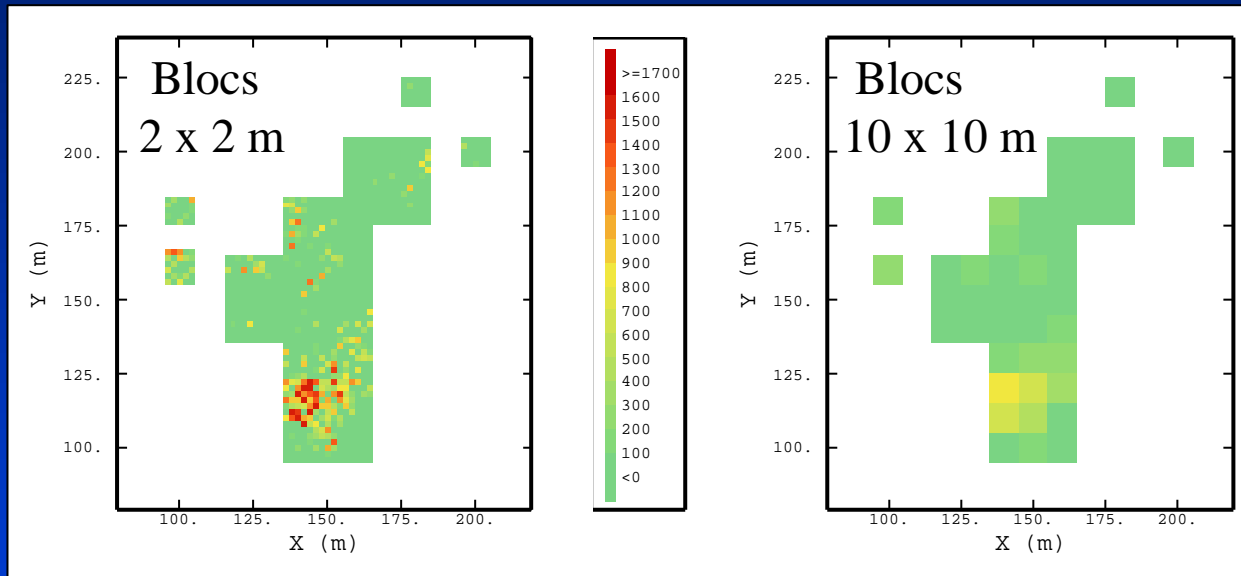


## Problèmes :

- Contraste des teneurs
- Répartition hétérogène des concentrations
- Peu de robustesse du variogramme classique
- Origines de l'effet de pépité ?
- Comment améliorer l'estimation ?
- Choix d'une méthode d'estimation de  $P[\text{HAP} > \text{seuil}]$

# Exemple : le phénanthrène (sondages)

- Sélection de zones à dépolluer :

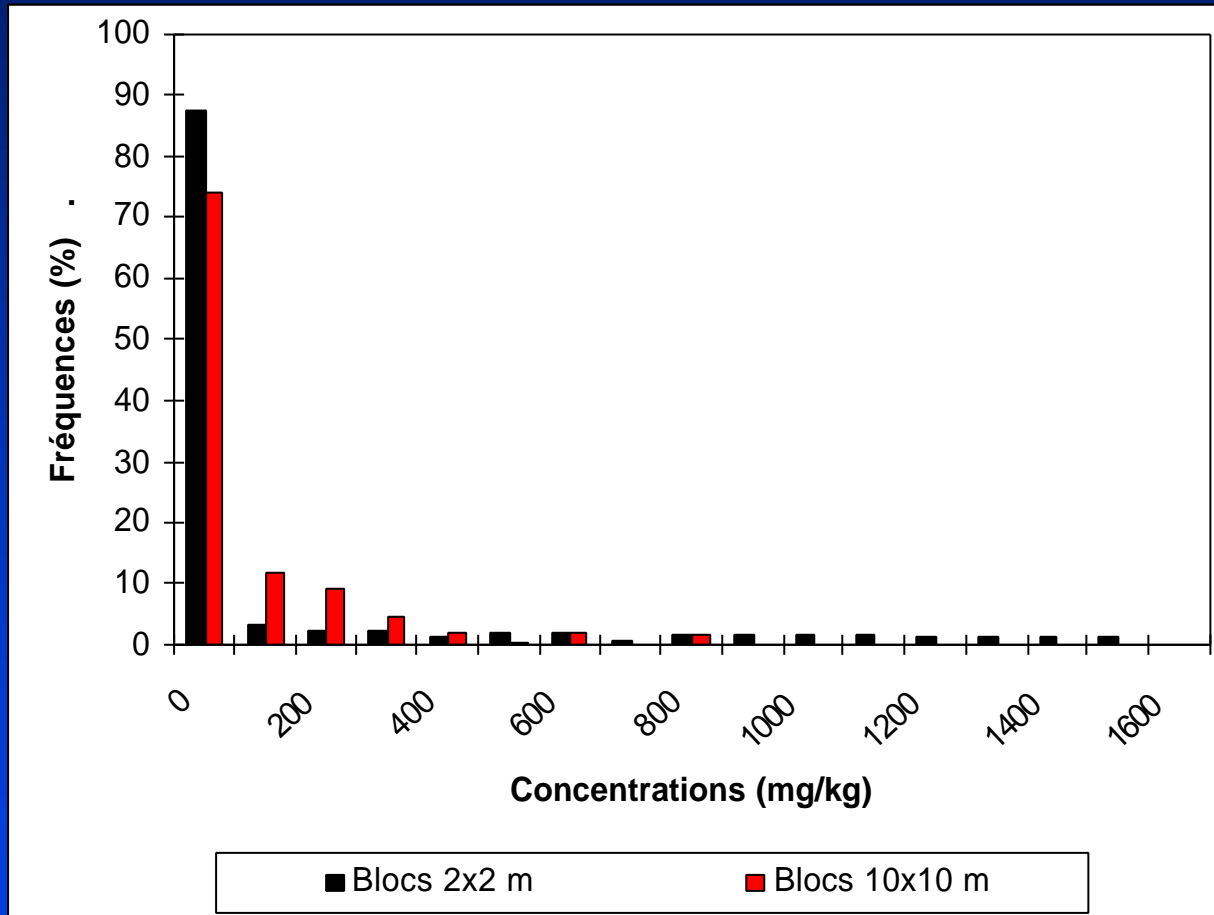


## Problèmes :

- Contraste des teneurs
- Répartition hétérogène des concentrations
- Peu de robustesse du variogramme classique
- Origines de l'effet de pépité ?
- Comment améliorer l'estimation ?
- Choix d'une méthode d'estimation de  $P[\text{HAP} > \text{seuil}]$
- Importance du support

# Exemple : le phénanthrène (sondages)

- Sélection de zones à dépolluer :



## Problèmes :

- Contraste des teneurs
- Répartition hétérogène des concentrations
- Peu de robustesse du variogramme classique
- Origines de l'effet de pépité ?
- Comment améliorer l'estimation ?
- Choix d'une méthode d'estimation de  $P[\text{HAP} > \text{seuil}]$
- Importance du support

# Plan

- Mise en évidence d'une structure spatiale
- Utilisation d'informations auxiliaires
- Sélection de zones à dépolluer

## Problèmes :

- Contraste des teneurs
- Répartition hétérogène des concentrations
- Peu de robustesse du variogramme classique
- Origines de l'effet de pépité ?
- Comment améliorer l'estimation ?
- Choix d'une méthode d'estimation de  $P[\text{HAP} > \text{seuil}]$
- Importance du support

# Mise en évidence d'une structure spatiale

**Difficulté : Concentrations très contrastées**

→ **Manque de robustesse du variogramme classique**

$$g^*(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(x_i + h) - z(x_i)]^2$$

→ **Question** : un variogramme classique expérimental pépitique est-il indicateur d'une absence de structure (à l'échelle de reconnaissance) ou d'une non adaptation de cet outil ?

→ **Recherche d'outils variographiques plus robustes**

# Variogramme déduit de la covariance non centrée

- Variogramme expérimental :

$$g^*(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(x_i + h) - z(x_i)]^2 \quad (\text{influence du contraste entre valeurs fortes et faibles})$$

- Covariance non centrée expérimentale (si stationnarité d'ordre deux) :

$$C^*(h) = \frac{1}{N(h)} \sum_{i=1}^{N(h)} z(x_i) \cdot z(x_i + h) \quad (\text{influence des valeurs fortes})$$

- Relation :

$$g^*(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(x_i + h)^2 + z(x_i)^2] - C^*(h)$$

$s^{2*}(h) = s^{2*}(0) = C^*(0)$  (si stationnarité)

→ Estimateur non biaisé du variogramme expérimental :

$$C^*(0) - C^*(h)$$

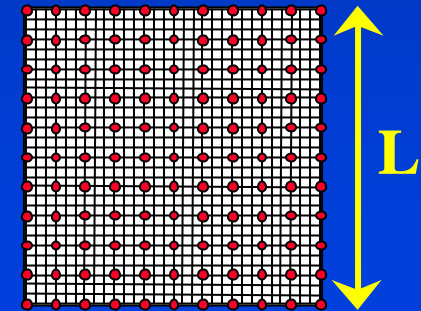
# Comparaison de $g_{classique}(h)$ et $C(0) - C(h)$

- **Objectif** : par **simulation** sur un champ fixé, comparer  $g_{classique}(h)$  et  $C(0) - C(h)$  pour un modèle sphérique et deux types de modèles :
  - loi spatiale gaussienne
  - loi spatiale lognormale

- **Critères de comparaison**, pour  $N$  simulations :

Variances d'estimation du variogramme expérimental :

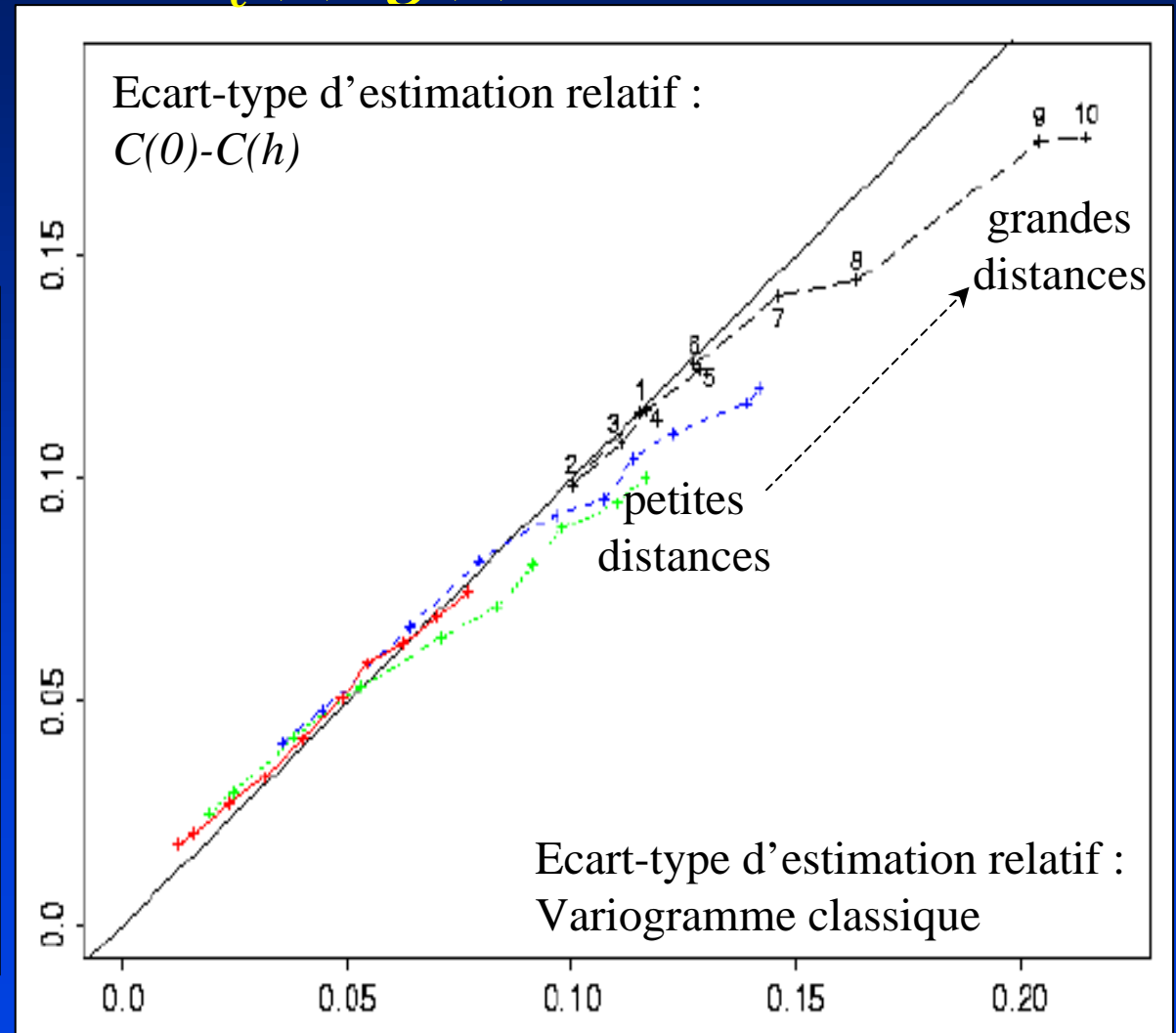
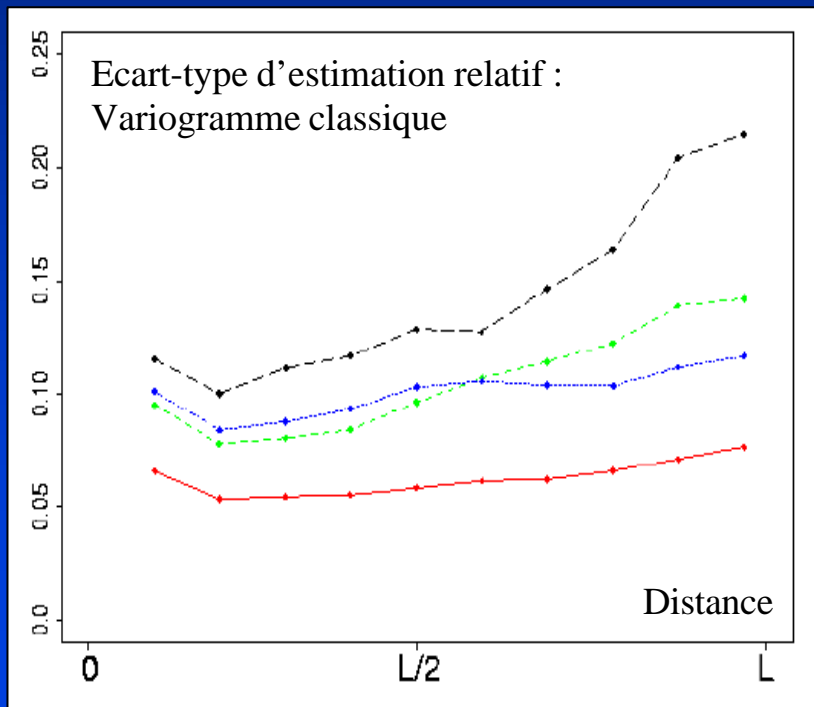
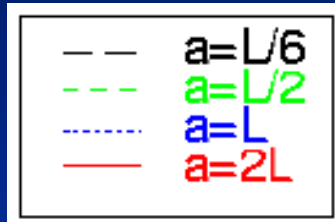
$$s_e^2(h) = \text{Var} \left[ \underbrace{\Gamma^*(h)}_{\substack{\text{Vrg expérimental} \\ (\bullet)}} - \underbrace{\Gamma_R(h)}_{\text{Vrg régional}} \right]$$





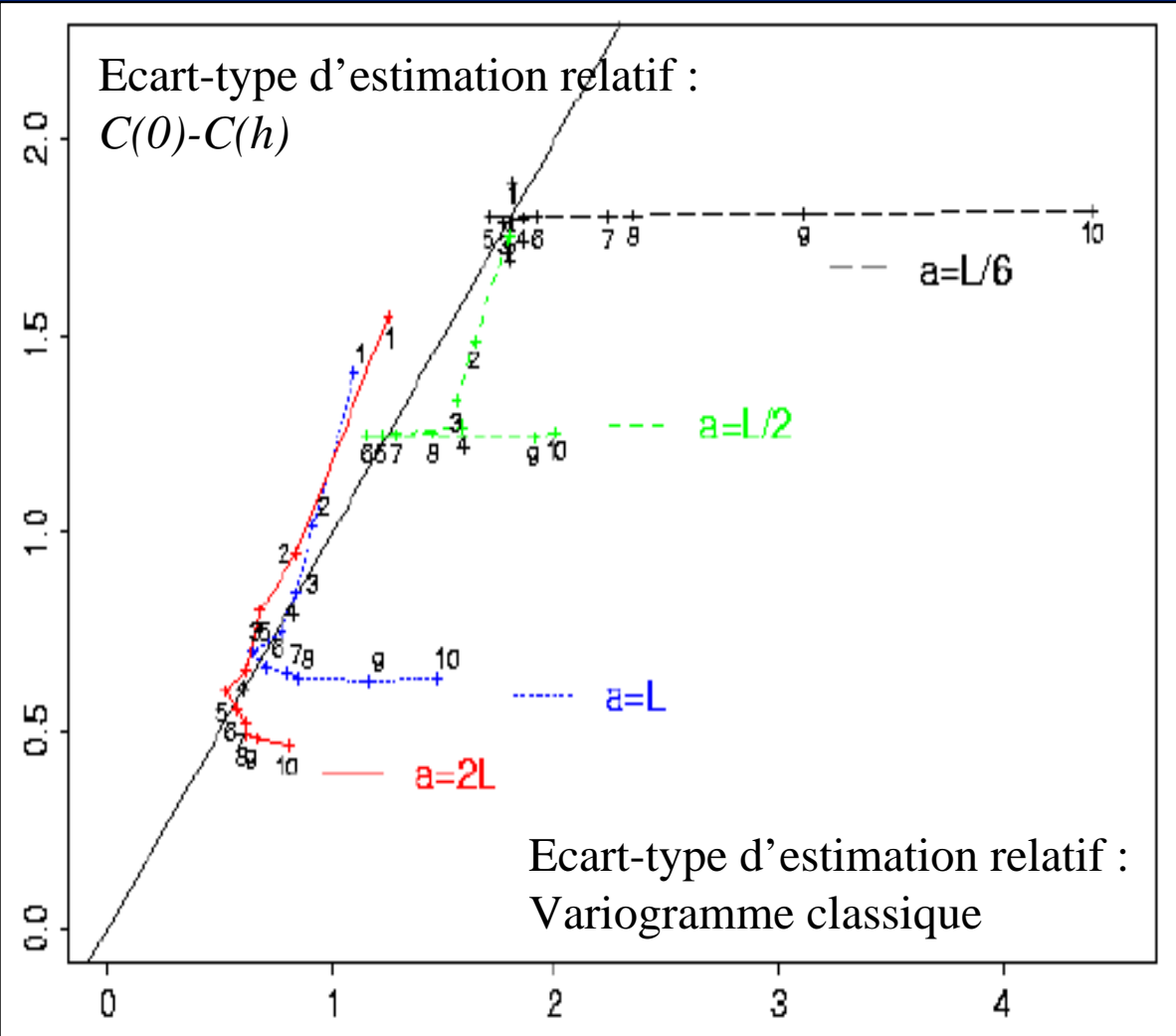
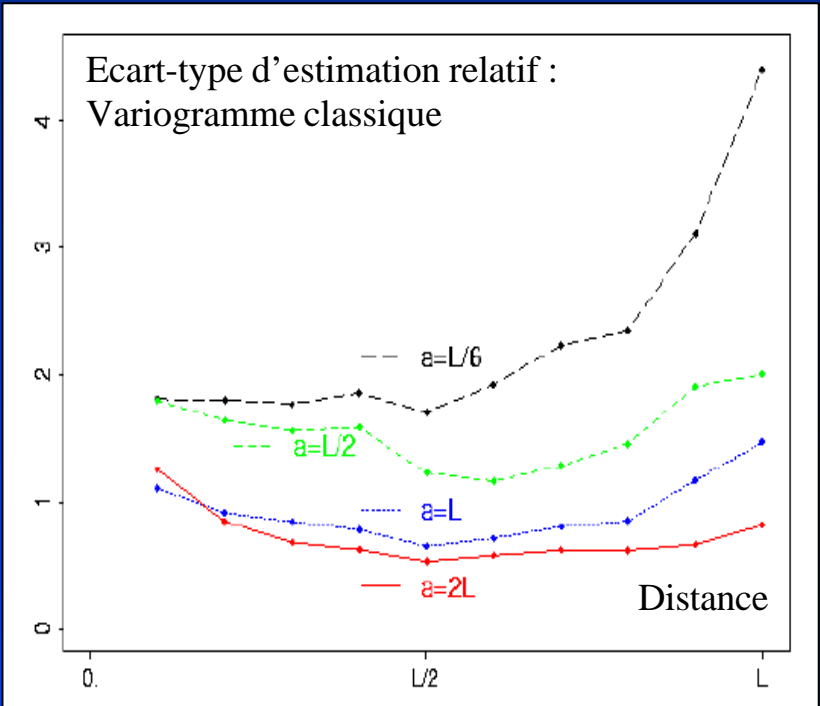
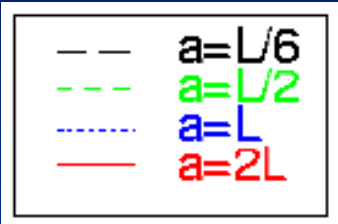
# Comparaison de $g_{classique}(h)$ et $C(0) - C(h)$

- Ecart-types d'estimation relatifs ( $s_e(h) / g(h)$ ) : cas gaussien



# Comparaison de $g_{classique}(h)$ et $C(0) - C(h)$

- Ecart-types d'estimation relatifs : cas lognormal (CV=2.9)



# Comparaison de $g_{classique}(h)$ et $C(0) - C(h)$

## - Conclusions -

- $g_{classique}(h)$  reste nécessaire pour tester l'hypothèse de stationnarité d'ordre deux.
- Si l'hypothèse est admissible :
  - **Fonction aléatoire gaussienne** : meilleure précision de  $C(0)-C(h)$  aux grandes distances
    - ➔  $C(0)-C(h)$  utile pour l'estimation du palier
  - **Fonction aléatoire lognormale** : meilleure précision de  $C(0)-C(h)$  également aux petites distances
    - ⚠ Variances d'estimation des variogrammes très élevées !

# Variogramme moyen par échantillon

## Construction :

(i) Calcul du variogramme classique associé à chaque  $x_i$

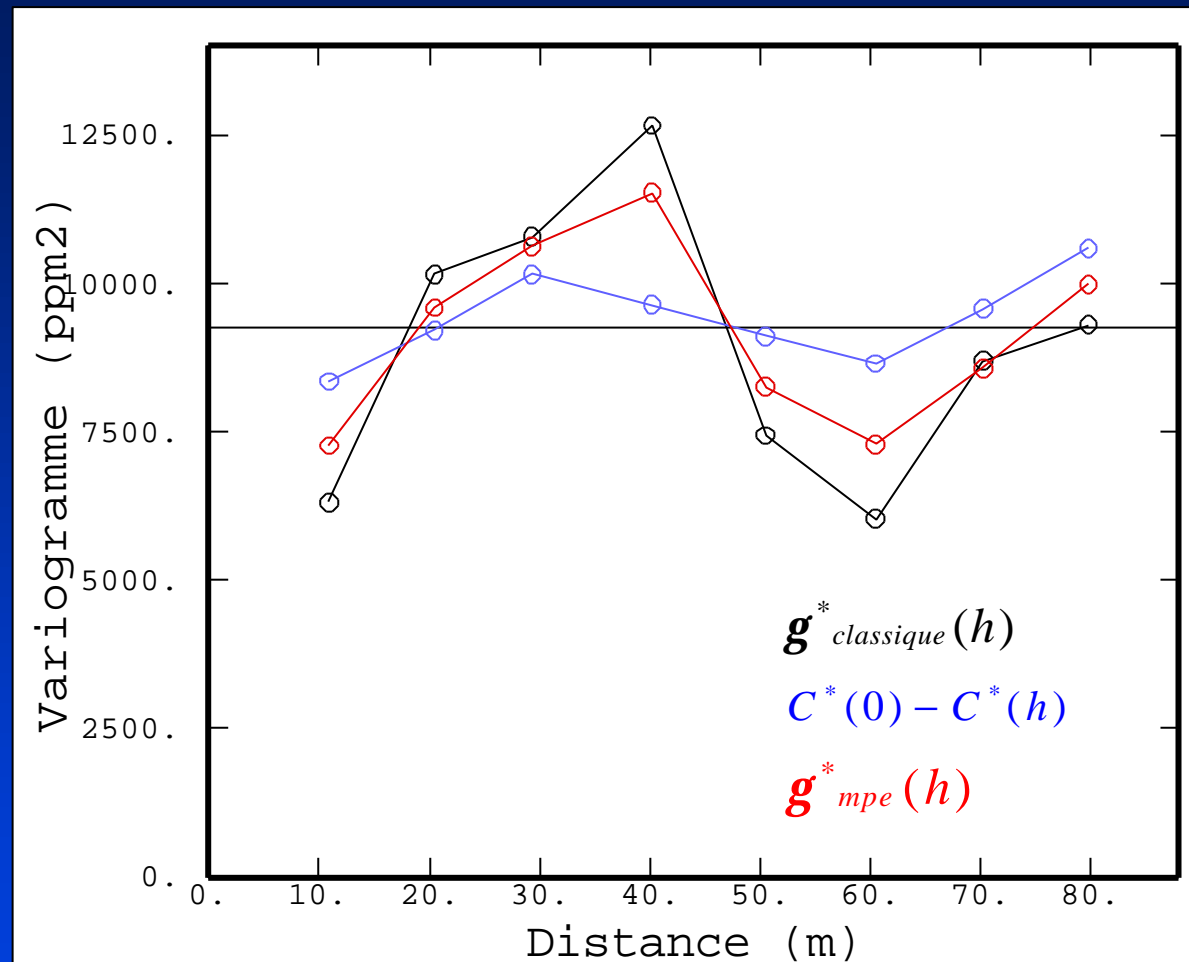
$$\mathbf{g}_i(h) = \frac{1}{2N_i(h)} \sum_{x_i - x_j \sim h} [z(x_i) - z(x_j)]^2 \quad (N_i(h) = \#\{x_j \text{ t.q. } x_i - x_j \sim h\})$$

(ii) Moyenne de ces « variogrammes par échantillon »

$$\mathbf{g}_{mpe}^*(h) = \frac{1}{n(h)} \sum_{i=1}^N \mathbf{g}_i(h) \quad (n(h) = \#\{x_i \text{ intervenant à la distance } h\})$$

➔ Mesure la variabilité spatiale de façon cohérente par rapport à l'ensemble des données

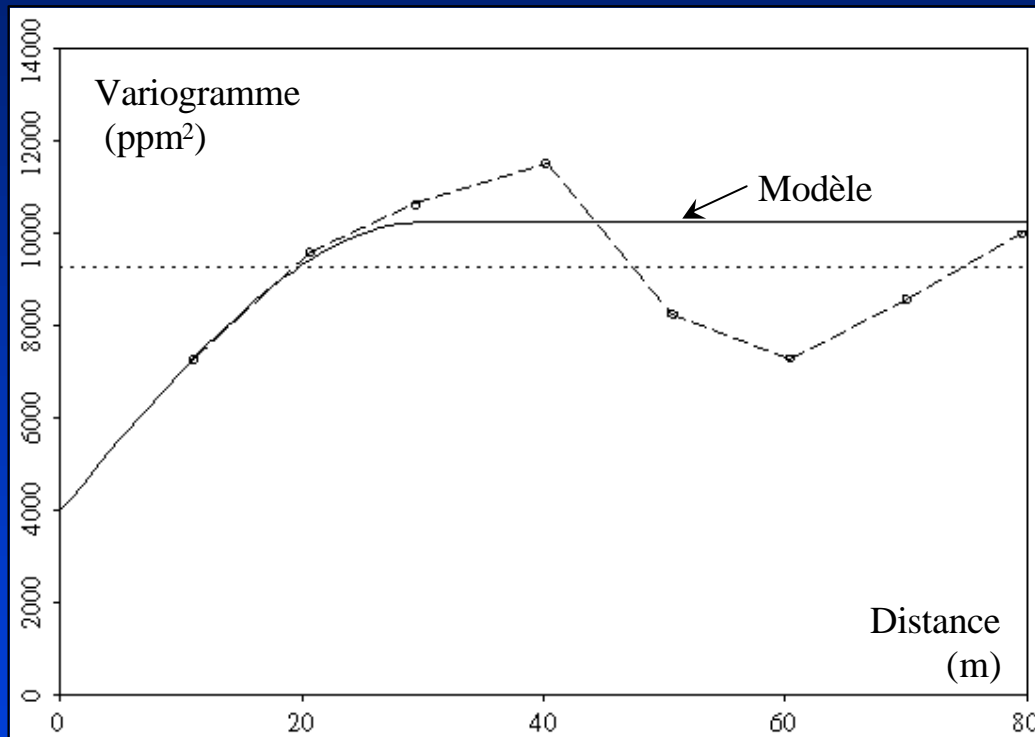
# Application : benzo(a)pyrène



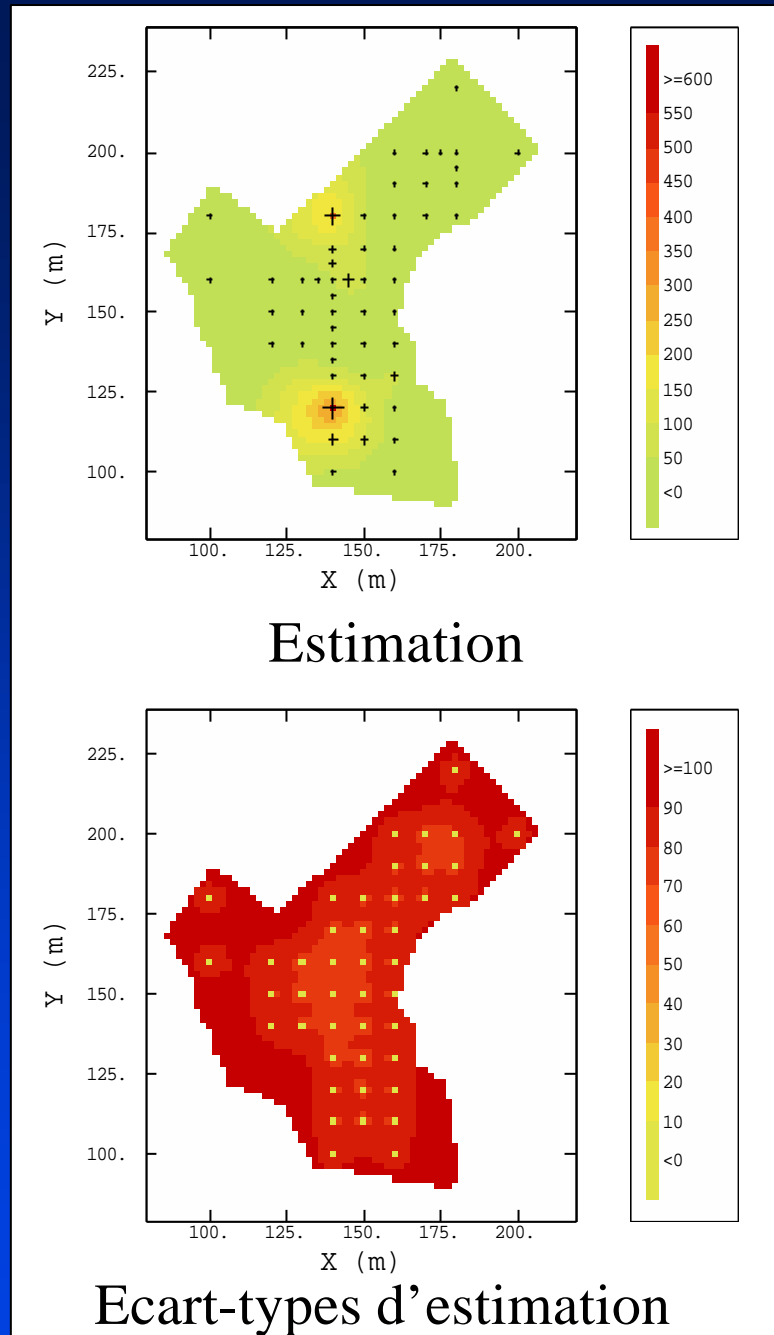
➔ Variogramme m.p.e. intéressant.

Existence d'une structure spatiale, même si effet de pépité élevé !

# Estimation du benzo(a)pyrène (sondages)



Variogramme moyen par échantillon

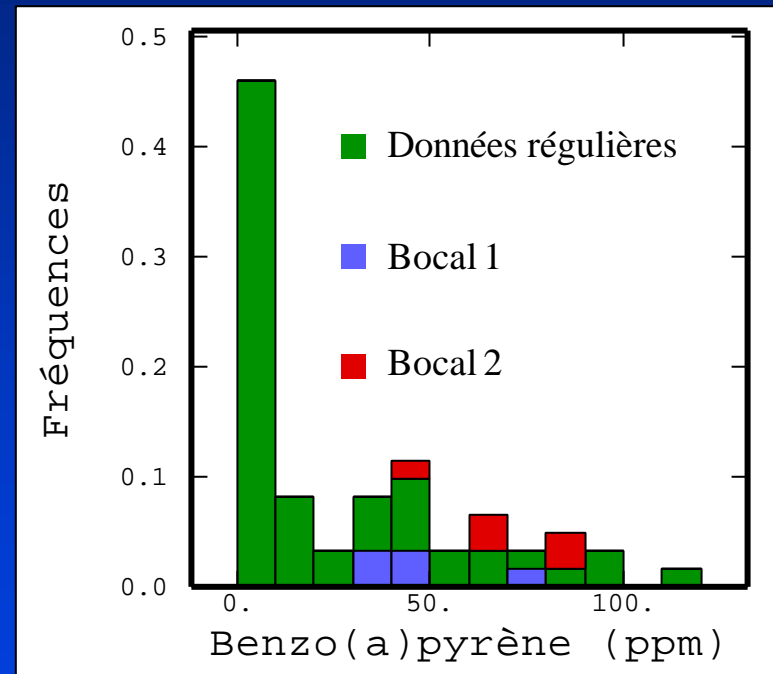


# Sources d'hétérogénéité des concentrations (1/2)

## - Préparation des échantillons-

- Prélèvement **ponctuel** de quelques kgs de sol
- **Homogénéisation**
- Séparation en deux bocaux
- 5 analyses par bocal

→ **Homogénéisation insuffisante !**

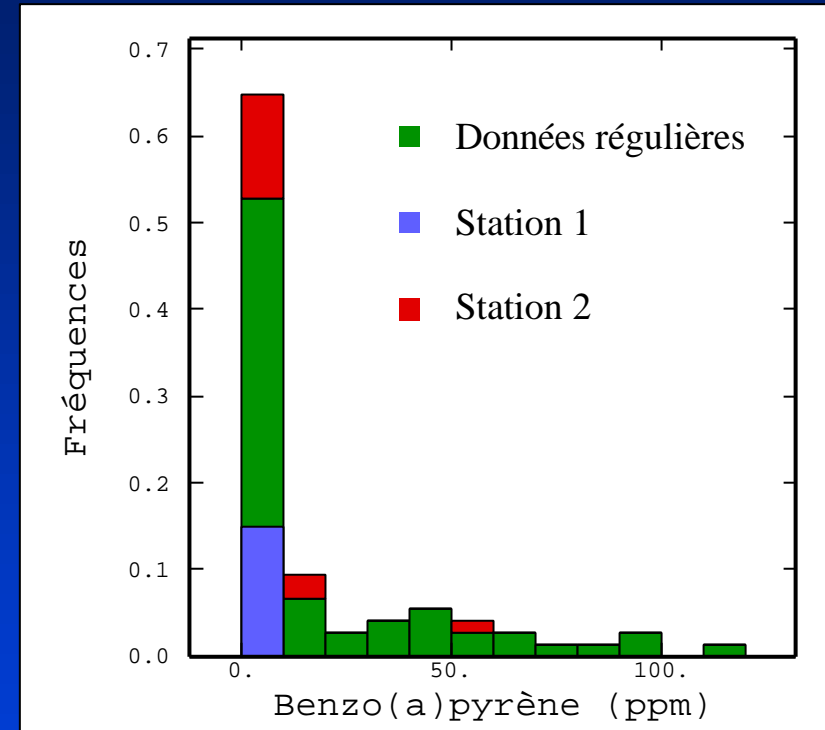
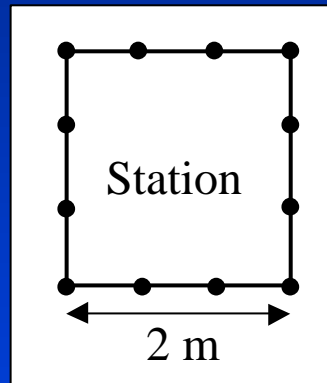


# Sources d'hétérogénéité des concentrations (2/2)

## - Variabilité à petite distance -

Deux stations (pas d'éch. < 1 m)

- 1 : zone de faibles concentrations
- 2 : à proximité de la concentration la plus élevée



➔ **Importante variabilité à petite échelle !**

Importance de resserrer localement l'échantillonnage pour pouvoir la prendre en compte ... mais jusqu'où ?



# Mise en évidence d'une structure spatiale

## - Conclusions -

- Importance de l'hypothèse de stationnarité
- Intérêt du variogramme moyen par échantillon
- Relation très éloignée entre structure expérimentale et régionale dans le cas lognormal !
- Plusieurs sources d'hétérogénéité des concentrations :
  - Homogénéisation insuffisante des échantillons
  - Echantillonnage ponctuel discutable
- ... mais existence d'une structure spatiale dans certains cas !

# Plan

- Mise en évidence d'une structure spatiale
- **Utilisation d'informations auxiliaires**
- Sélection de zones à dépolluer
- Conclusions

## Problèmes :

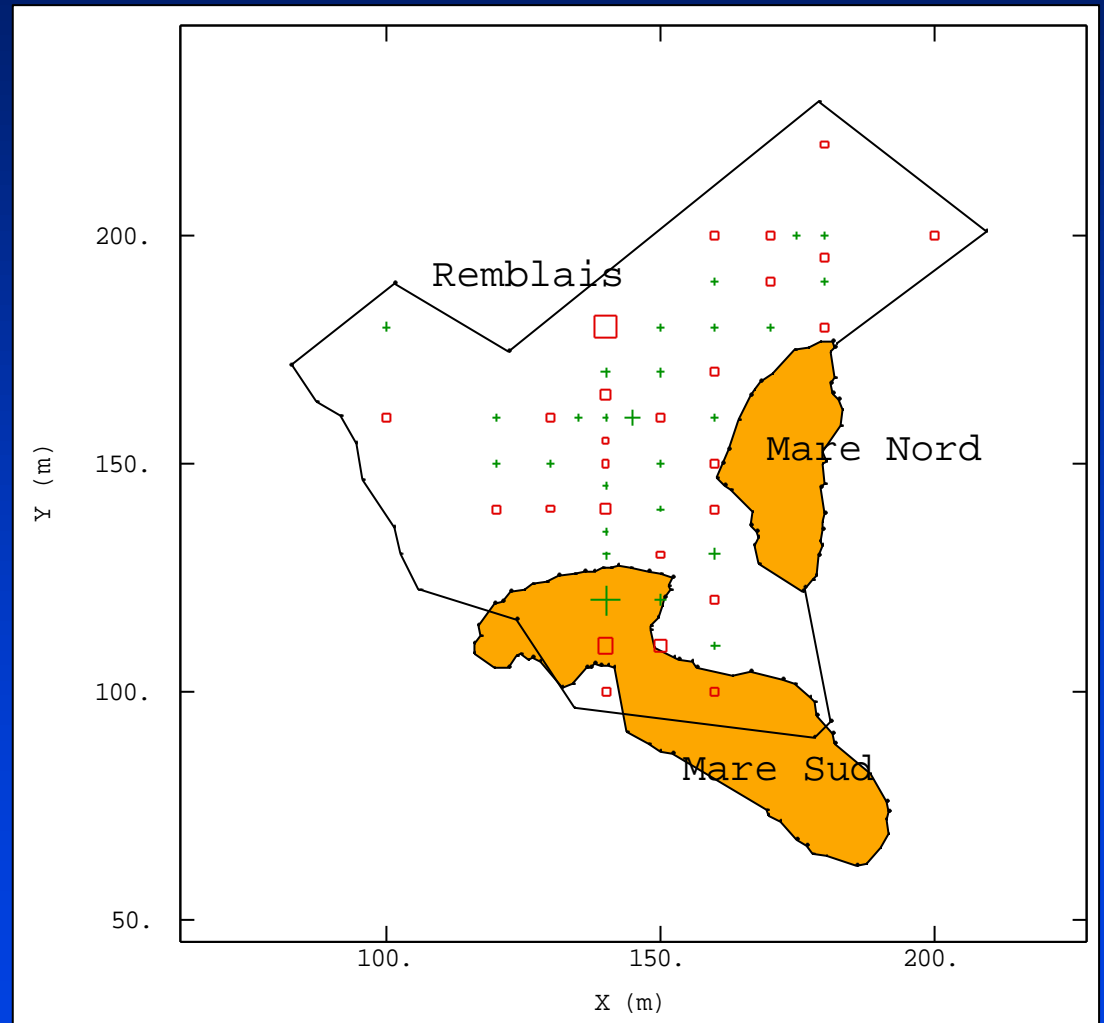
- Contraste des teneurs
- Répartition hétérogène des concentrations
- Peu de robustesse du variogramme classique
- Origines de l'effet de pépité ?
- **Comment améliorer l'estimation ?**
- Choix d'une méthode d'estimation de  $P[\text{HAP} > \text{seuil}]$
- Importance du support

# Informations auxiliaires et estimation multivariable

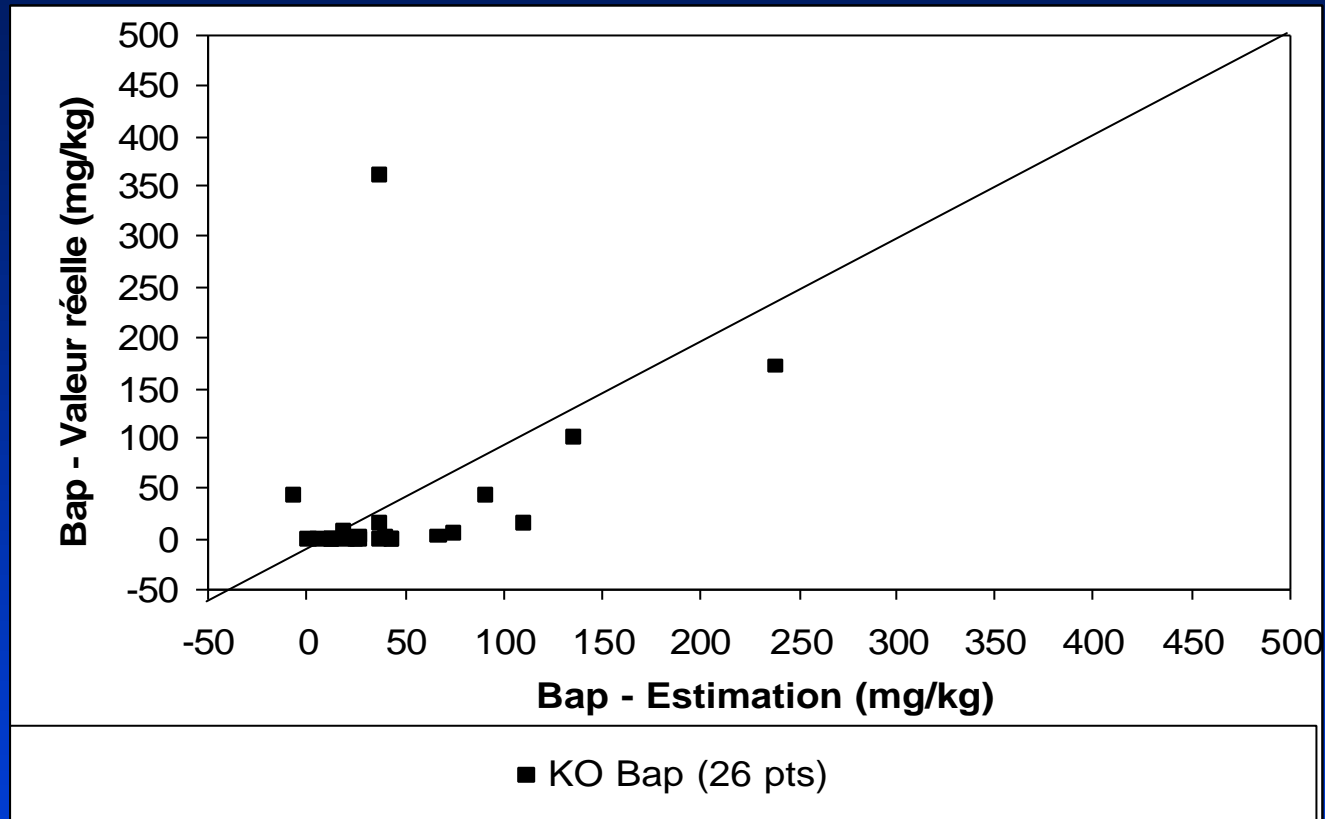
- Modèle linéaire de corégionalisation entre :
  - HAP de poids proche
  - HAP et variable qualitative
- Faible amélioration (cas **homotopique**)

**Si échantillonnage plus dense des variables auxiliaires ?**

- Supposons que Bap ne soit connu qu'en 50 % des points (croix vertes +)
- Estimation de la concentration en Bap aux points non utilisés (.) et calcul de l'erreur réalisée

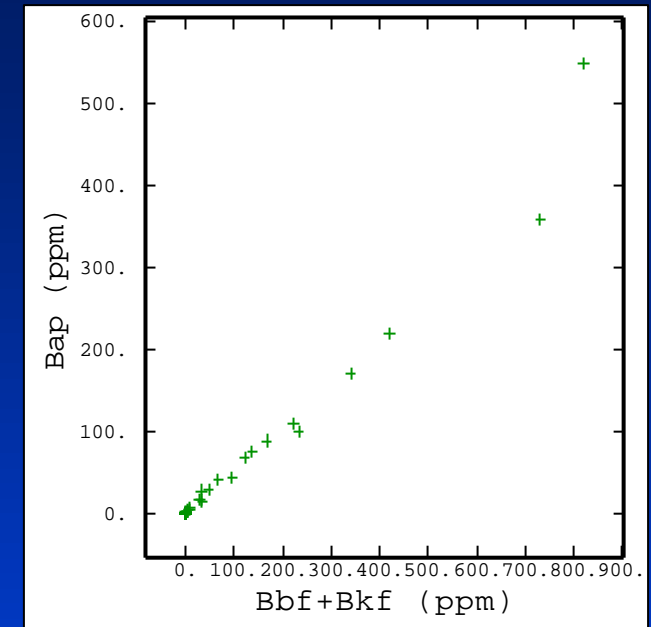
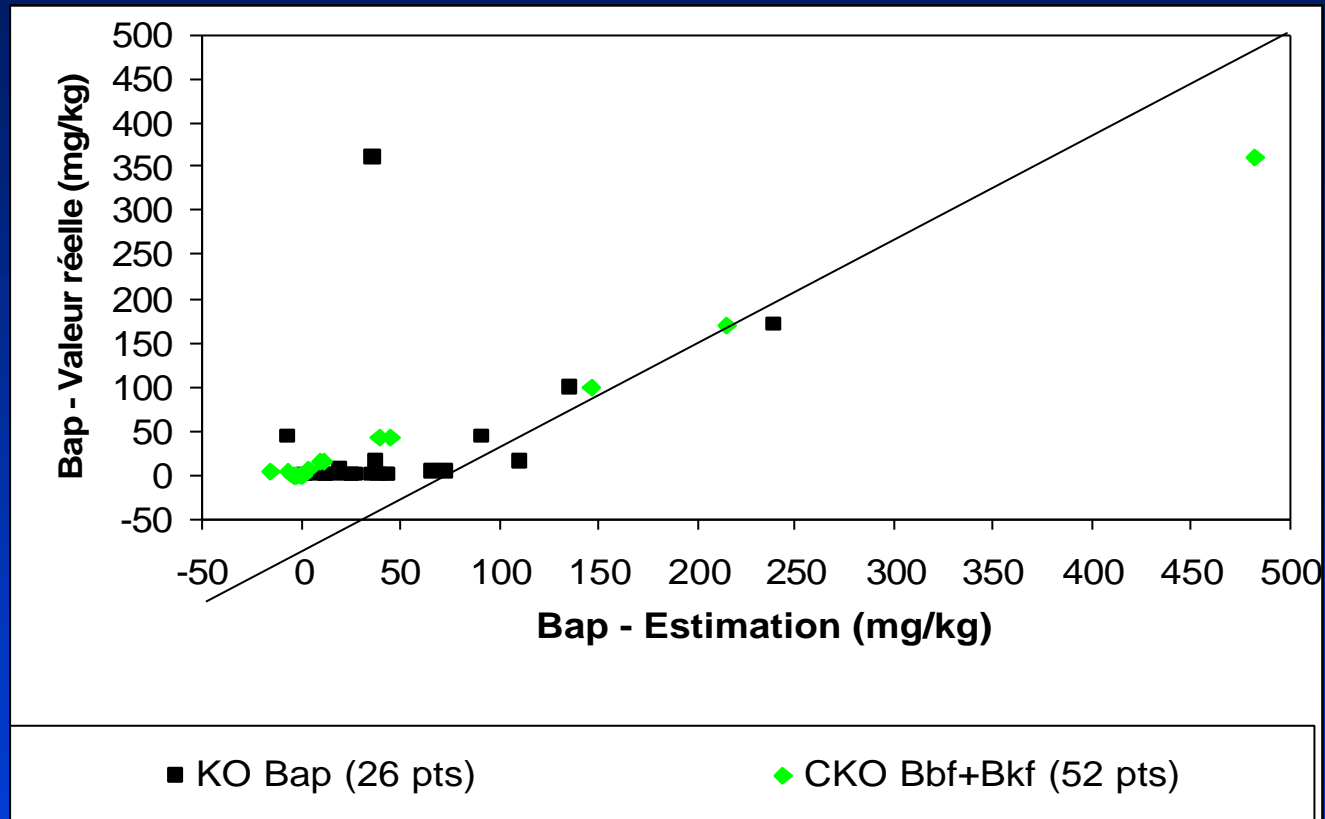


# Apport pour l'estimation d'un HAP



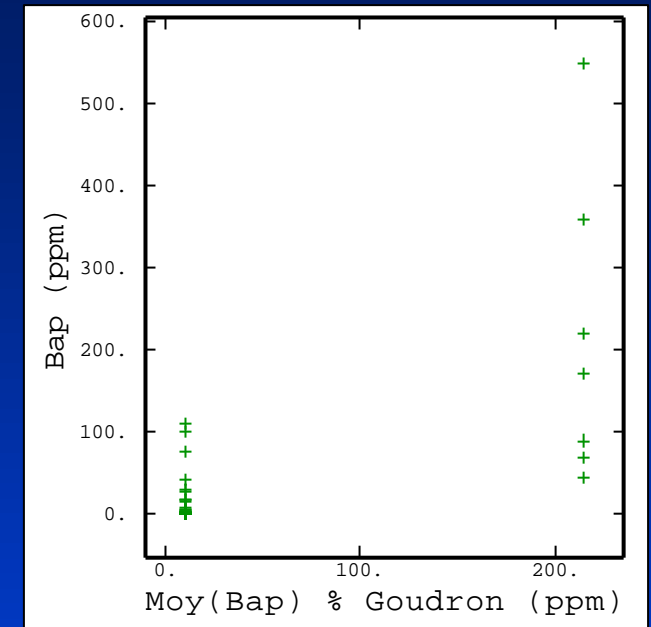
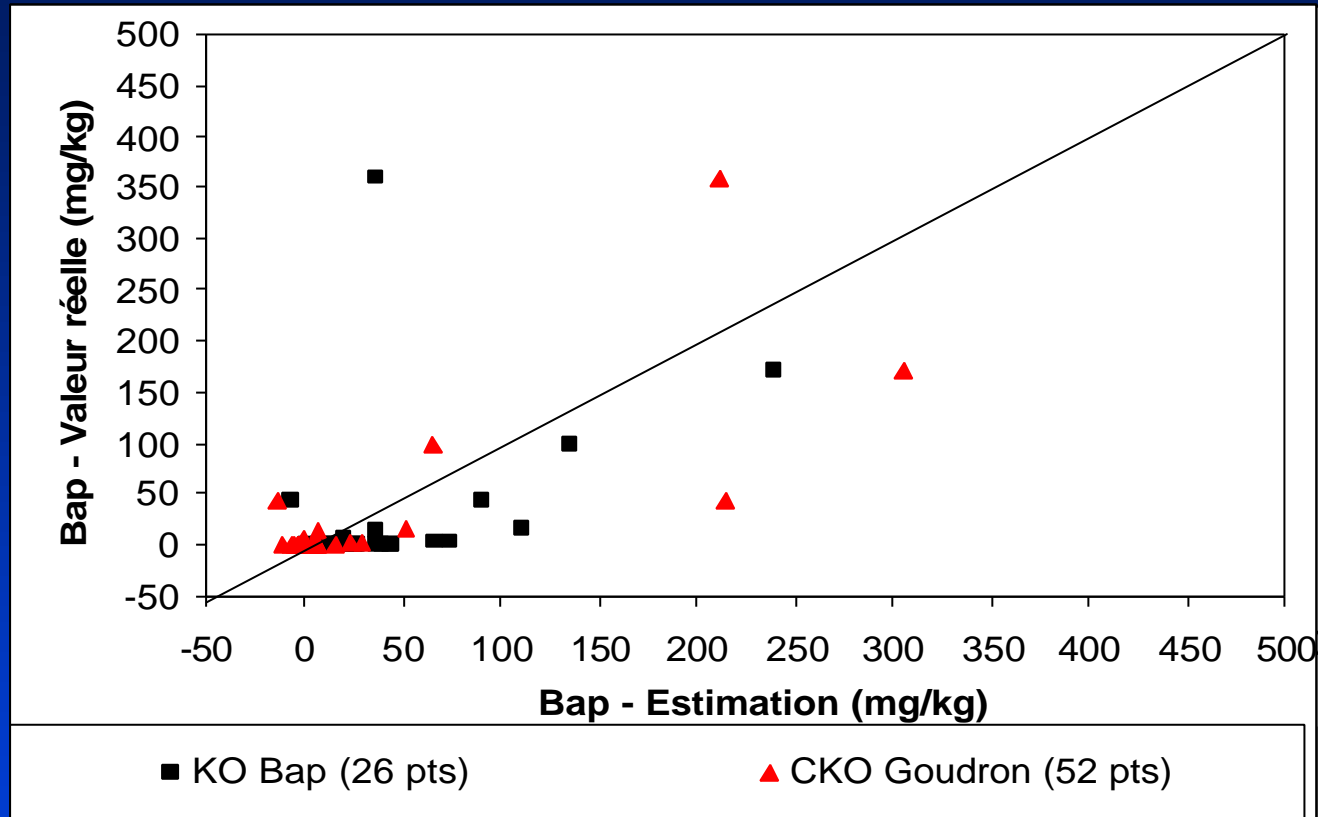
Erreur quadratique moyenne : • KO Bap : **6253**

# Apport pour l'estimation d'un HAP



Erreur quadratique moyenne : • KO Bap : **6253**  
• CKO Bbf+Bkf : **858**

# Apport pour l'estimation d'un HAP



Erreur quadratique moyenne :

- KO Bap : **6253**
- CKO Bbf+Bkf : **858**
- CKO Goudron : **3309**

# Utilisation d'informations auxiliaires

## - Conclusions -

- Existence de variables corrélées à la teneur en HAP :
  - Informations historiques  
*(permettent de guider un échantillonnage complémentaire)*
  - Variables qualitatives  
*(échantillonnage peu coûteux)*
  - Autres polluants
- Nette amélioration des estimations lorsque les informations auxiliaires sont **échantillonnées de façon plus dense**

# Plan

- Mise en évidence d'une structure spatiale
- Utilisation d'informations auxiliaires
- **Sélection de zones à dépolluer**
- Conclusions

## Problèmes :

- Contraste des teneurs
- Répartition hétérogène des concentrations
- Peu de robustesse du variogramme classique
- Origines de l'effet de pépite ?
- Comment améliorer l'estimation ?
- **Choix d'une méthode d'estimation de  $P[\text{HAP} > \text{seuil}]$**
- **Importance du support**



# Choix et validation d'un modèle

- Existence d'effets de bord  $\longrightarrow$  Modèle de diffusion

## Modèle gaussien anamorphosé :

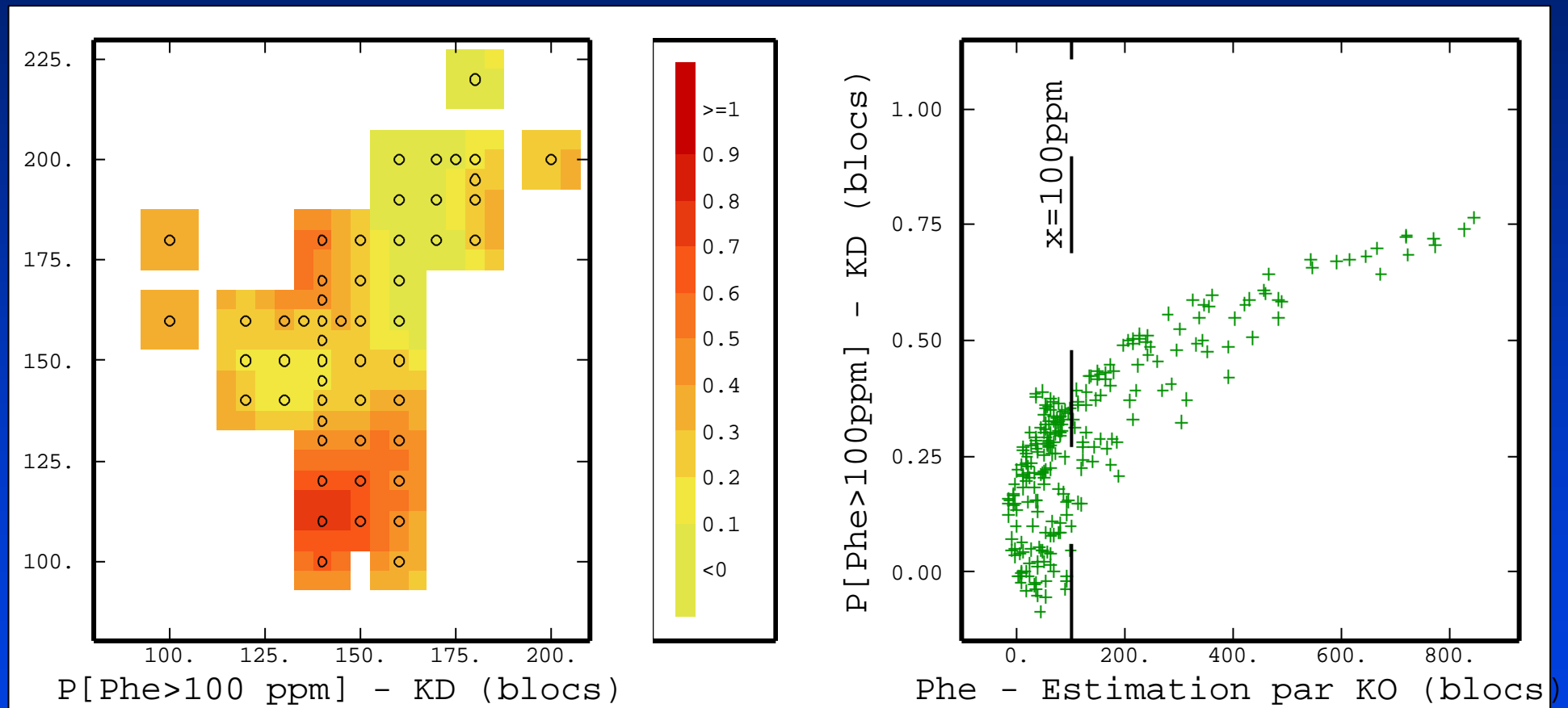
- Soit un seuil  $z_c$ , alors pour  $y_c = \Phi^{-1}(z_c)$  on a par bijectivité de  $\Phi$  :

$$Z(x) \geq z_c \Leftrightarrow Y(x) \geq y_c$$

- **Hypothèse du modèle** : lois  $(Y(x), Y(x+h))$  bigaussiennes
- Dépollution effectuée sur des blocs de taille cohérente avec la **technique de dépollution** et les **objectifs fixés**.  
 $\longrightarrow$  **Nécessité d'un modèle de changement de support**

# Krigeage disjonctif en modèle gaussien discret

- Application : phénanthrène pour un seuil égal à 100 mg/kg (blocs 5x5m) :



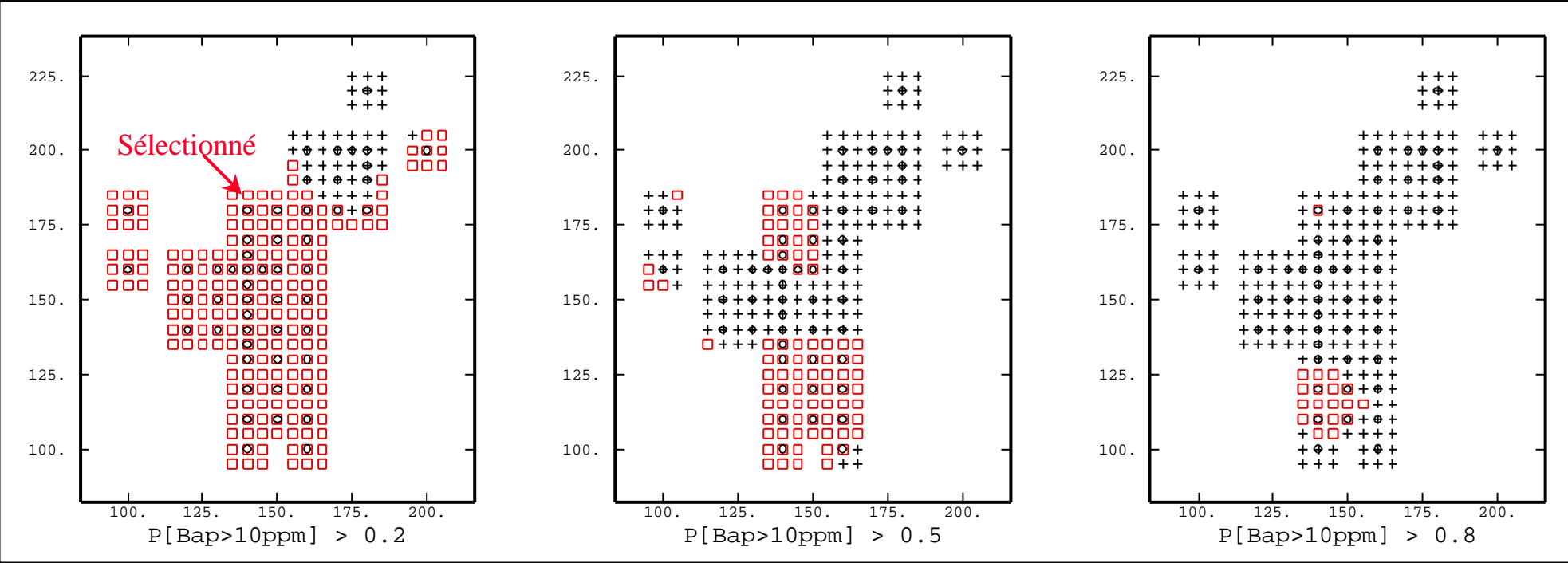
# Application au benzo(a)pyrène (sondages)

• Probabilité de dépassement de 10 mg/kg estimée supérieure à :

0.2

0.5

0.8



# Sélection de zones à dépolluer

## - Conclusions -

- Choix du modèle de diffusion le plus simple
- Importance du support
- Prise en compte de la probabilité de laisser en place des blocs de teneur supérieure à un seuil d'intervention

# Plan

- Mise en évidence d'une structure spatiale
- Utilisation d'informations auxiliaires
- Sélection de zones à dépolluer
- **Conclusions**

## Problèmes :

- Contraste des teneurs
- Répartition hétérogène des concentrations
- Peu de robustesse du variogramme classique
- Origines de l'effet de pépité ?
- Comment améliorer l'estimation ?
- Choix d'une méthode d'estimation de  $P[\text{HAP} > \text{seuil}]$
- Importance du support

## Conclusions (1/2)

- **Recommandations pour l'échantillonnage d'un site pollué**
  - Non préférentiel, risque d'une utilisation a priori de l'historique
  - Représentativité des échantillons :
    - homogénéisation de matériaux hétérogènes délicate
    - variabilité à très petite échelle
  - Informations auxiliaires :
    - variables qualitatives, mesures semi-quantitatives
    - autres polluants
  - Echantillonnage complémentaire, tenant compte des conclusions du premier échantillonnage, éventuellement orienté spatialement

## Conclusions (2/2)

- Malgré différentes sources d'hétérogénéité, mise en évidence dans certains cas d'une **structure spatiale**
- **Si existence d'une structure spatiale, des outils :**
  - **Estimation des concentrations en place**
    - Utilité des variables corrélées à la concentration (cas hétérotopique)
  - **Sélection de zones de dépassement de seuils d'intervention**
    - Risque de laisser en place des concentrations supérieures au seuil
    - Prise en compte du support

# Perspectives

- Mise en œuvre et validation des conseils méthodologiques
- Utilisation de photographies aériennes
- Homogénéisation, échantillonnages composites ?
- Méthodes d'investigation plus représentatives ?