

Traitement des incertitudes en simulation numérique

Cours 3 : Modélisation d'expériences numériques => métamodèle processus gaussien

Bertrand Iooss
Amandine Marrel

Module INSA Toulouse/GMM 5
Planification, risque et incertitudes

13 décembre 2014



DE LA RECHERCHE À L'INDUSTRIE

cea

edf

CHANGER L'ÉNERGIE ENSEMBLE

Depuis une trentaine d'années, l'industrie a développé des processus et des codes de calcul parfois très lourds pour modéliser des phénomènes complexes !

La plupart des ingénieurs sont amenés à manipuler ces codes & processus

1) Il est nécessaire d'optimiser leur utilisation pour prendre des décisions !

=> *Analyse de sensibilité, planification d'expérience, développement de modèles réduits*

2) La validation de leurs résultats est un problème crucial lorsqu'ils sont utilisés dans des cycles industriels (conception, sûreté, prévision, etc.)

=> *Gestion des incertitudes, calculs fiabilistes*

3 cours de 3h15 pour INSA GMM 5 & Master Pro 2 UPS

1.Cours 1 : Introduction, modélisation et propagation d'incertitudes

2.Cours 2 : Planification et analyse d'expériences numériques

3.Cours 3 : Modélisation d'expériences numériques, métamodèle processus gaussien

3 séances de TP pour INSA GMM 5

1.TP 1 : Exercices en R

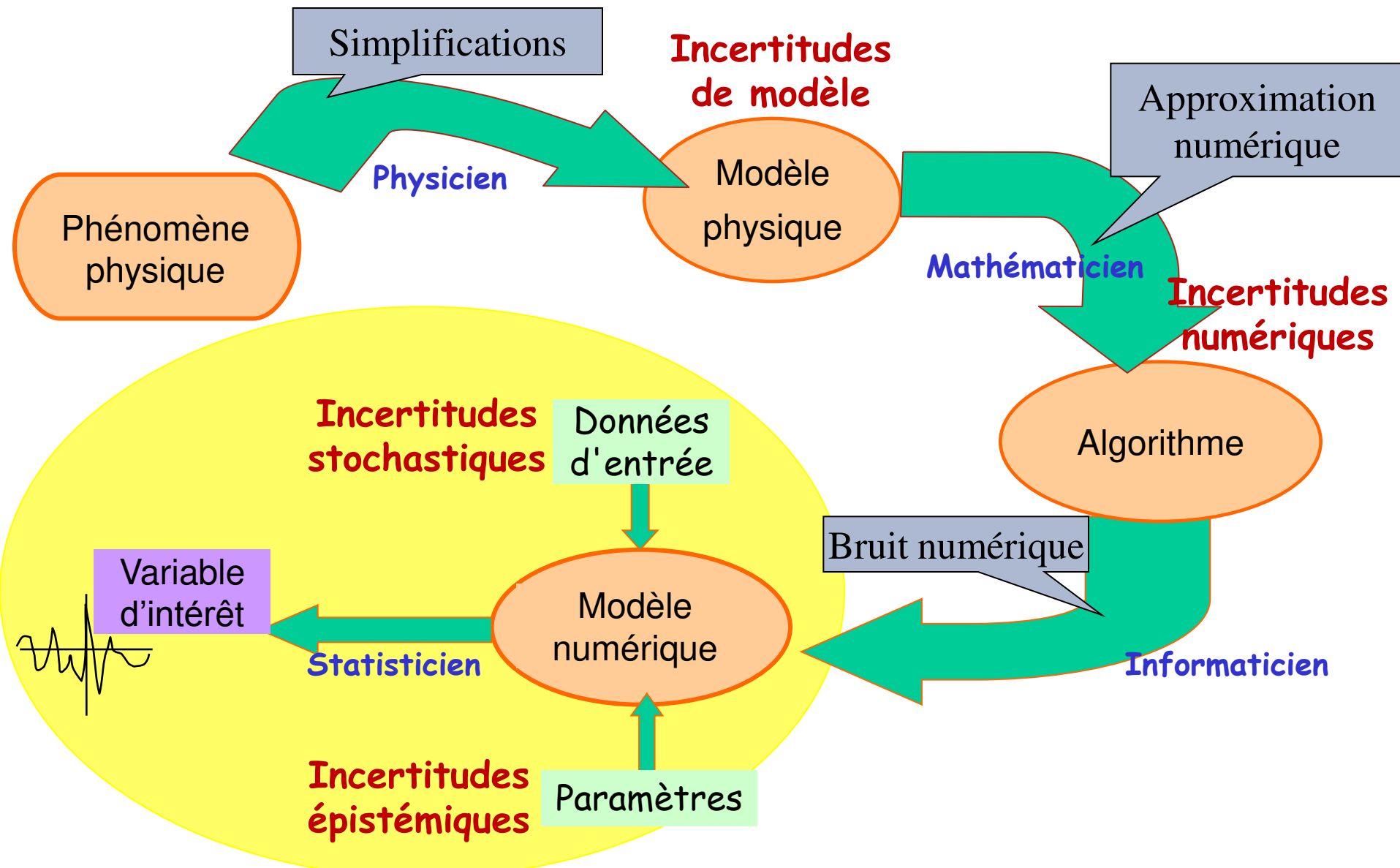
2.TP 2 : Exercices en R

3.TP 3 : Exercices en R

Une note sera délivrée via un compte-rends réalisés à l'issue du dernier TP

Cours 3 : Modélisation d'expériences numériques

⇒ Métamodèle processus gaussien



Étape C : Propagation des sources d'incertitude

Étape A : Spécification du problème

Variables d'entrée

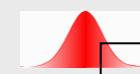
Incertaines : $\underline{x}$
Fixées : $\underline{d}$

Modèle
(ou processus de mesure)
 $G(\underline{x}, \underline{d})$

Variables d'intérêt
 $Z = G(\underline{x}, \underline{d})$

Quantité d'intérêt

Ex: variance, probabilité ..



Étape B: Quantification des sources d'incertitudes

Modélisation par des distributions



Étape C' : Analyse de sensibilité, Hiérarchisation

Rebouclage
(feedback)

Critère de décision
Ex: Probabilité < 10^{-b}

➤ Incertitudes sur les paramètres d'entrée des modèles numériques



➤ Gestion des incertitudes :

- Comment propager les incertitudes sur la sortie du code ?
- Quelles sont les incertitudes les plus préjudiciables ?
- Quelles sont les variables les plus influentes ?
- Incertitude au final sur la sortie ?
- Estimation de la marge de confiance sur la prise de décision ?



➤ Problème :

- Code souvent complexe & « coûteux »
- Grand nombre de variables d'entrée
- Grand nombre de simulations nécessaires pour réaliser les études de sensibilité et de propagation d'incertitude
- Etude souvent « multi-objectif »



Exploitation directe du code difficile



Utilisation d'un métamodèle

➤ Métamodèle :

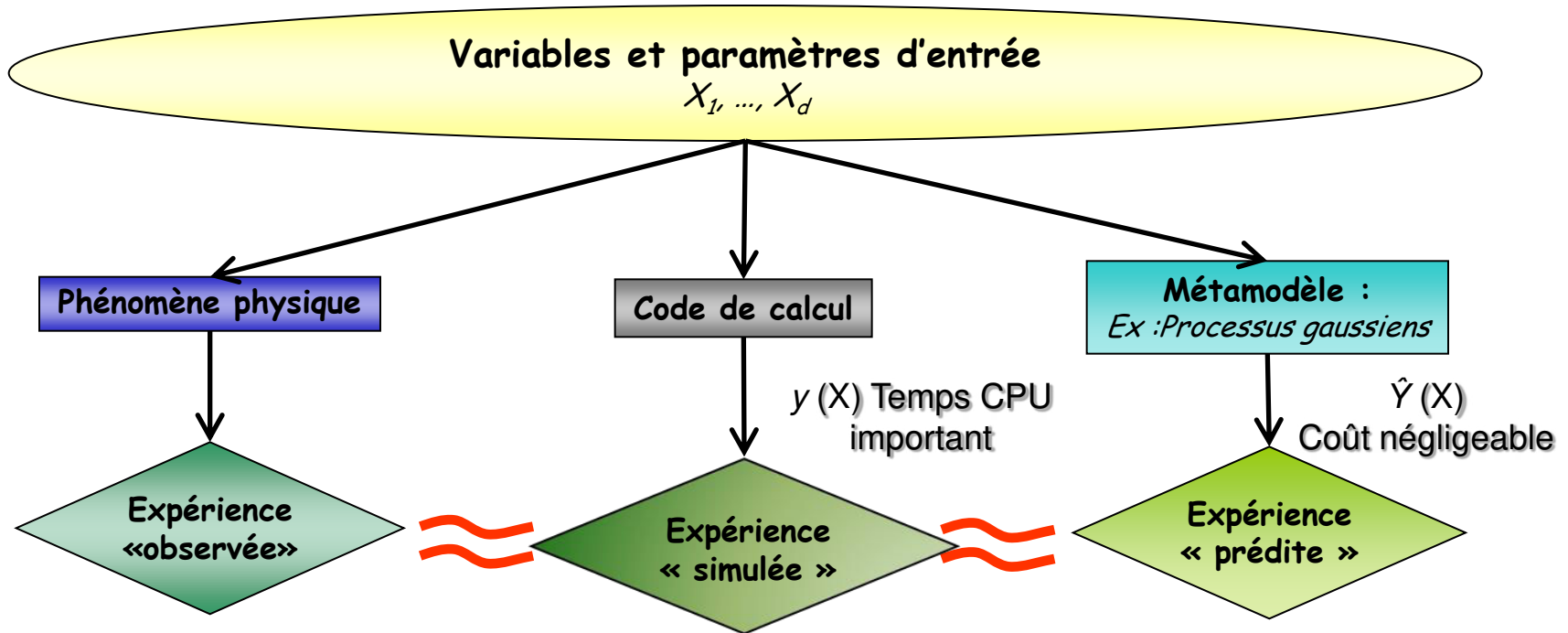
- Fonction statistique représentative du code de calcul
- Construction à partir de n simulations du code
- Temps d'évaluation négligeable par rapport à celui du code
- Approximant les réponses du code
- Permettant de prédire avec une « bonne précision » de nouvelles réponses dans le domaine de variation des paramètres incertains



Contrôle de la qualité d'approximation et de prédiction

- Outil multifonctionnel : analyse de sensibilité, propagation d'incertitudes, calibration...

Autres appellations : *surface de réponse , émulateur, surrogate model, modèle simplifié,...*

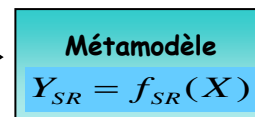
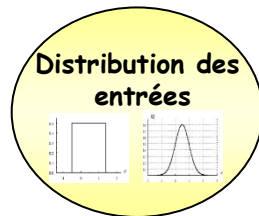
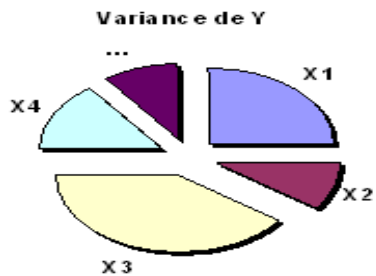


➤ Exploitation du métamodèle

▪ Analyse de sensibilité

▪ Propagation d'incertitudes

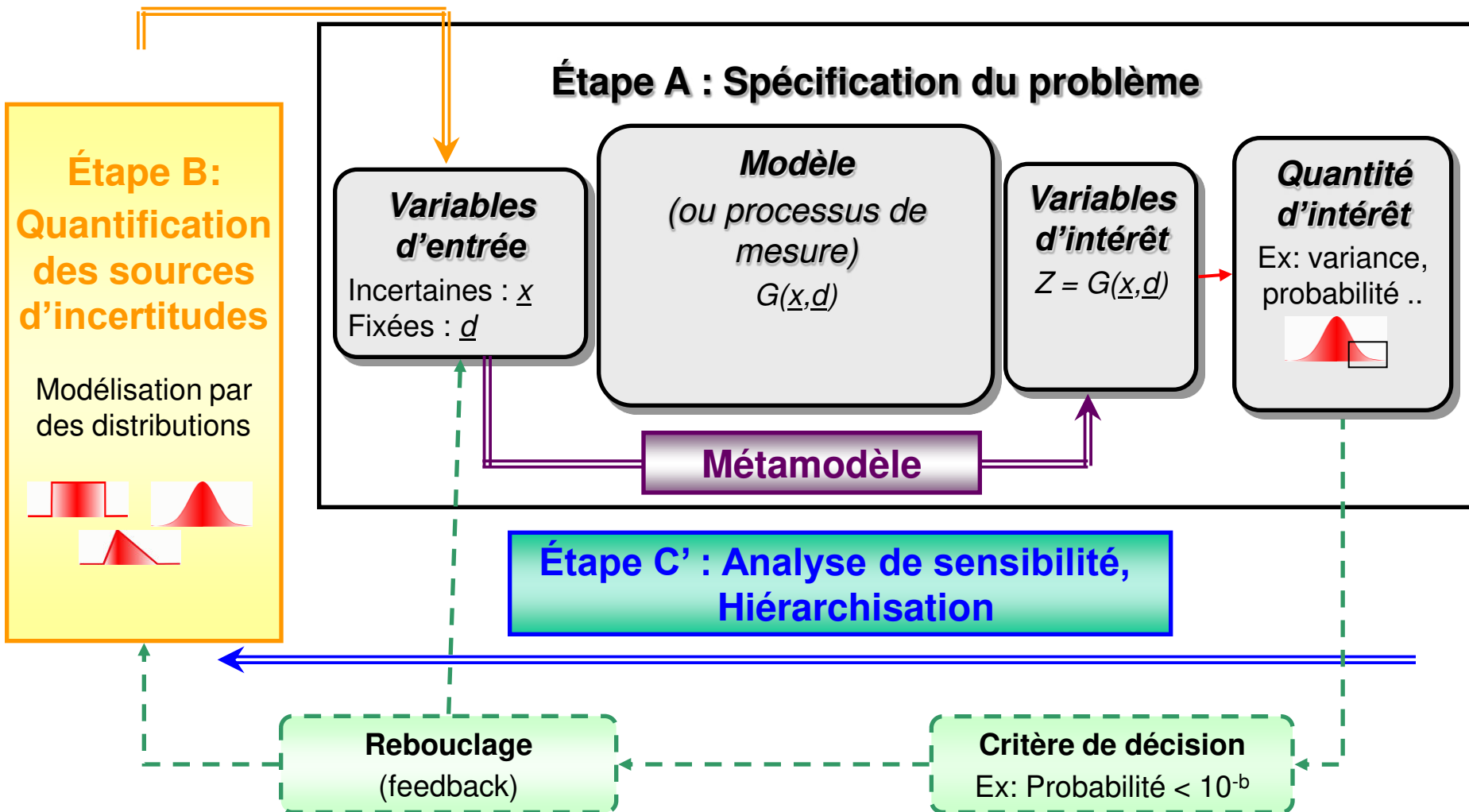
▪ Calibration



Détermination des paramètres

Adéquation expériences simulées et observées

Étape C : Propagation des sources d'incertitude



Planification d'expériences :

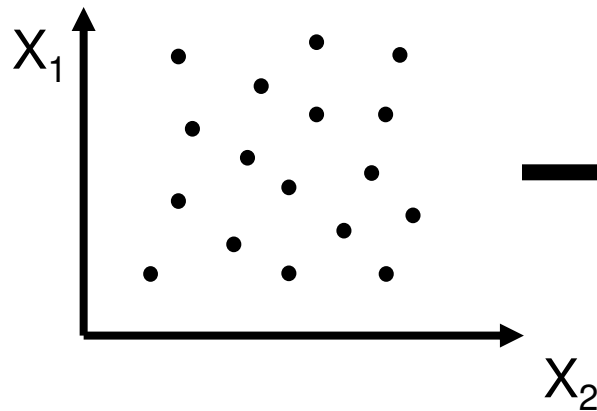
points où l'on va effectuer les simulations (Space Filling Design)

Simulation :

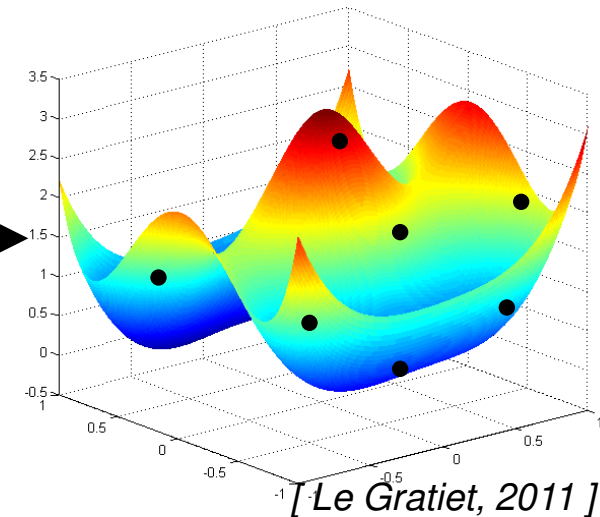
On effectue les simulations.

Méta-modélisation :

Approximation de la surface de réponse du code.



Code de calcul
Expérience



Méthode : On va choisir parmi une **famille de fonctions**, la fonction qui est la plus proche de la fonction objectif

➤ Différents types de métamodèles :

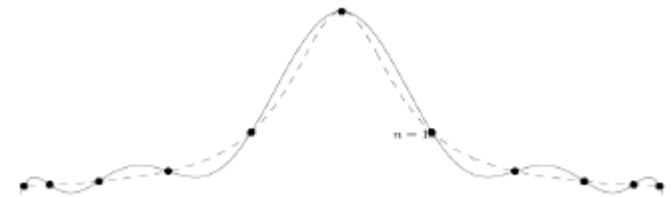
- Modèle linéaire, GLM

- Polynômes



- Splines

$$\hat{G}(\mathbf{x}) = \sum_{k=1}^K \hat{\beta}_k B_k(\mathbf{x}) \text{ avec } K \text{ le nb de noeuds}$$



- Modèles additifs, GAM

$$\hat{G}(\mathbf{x}) = \sum_{i=1}^p s_i(x_i) + \sum_{i < j} s_{ij}(x_i, x_j) + \dots$$

- Arbres de régression

- Réseaux de neurones

- **Krigeage ou processus gaussiens conditionnels**

- Polynômes de chaos

Rappels statistiques et théorie

- Variable, vecteur et processus gaussien

Définition du métamodèle processus gaussien (PG)

- Du krigeage à la régression par PG
- Formulation analytique
- Analogie avec le krigeage
- Simulation conditionnelles

Construction du métamodèle PG

- Choix de la fonction moyenne
- Choix de la fonction de covariance
- Estimation des paramètres

Utilisation du métamodèle PG

- Validation du métamodèle PG
- Illustration sur des exemples analytiques
- Plans d'expériences et planification adaptative

Application sur des cas industriels

➤ De la variable au champ aléatoire :

- **Variable aléatoire réelle** : fonction définie sur l'ensemble des éventualités Ω et qui associe à chaque éventualité ω un élément de $\mathbb{R}$

$$X : \omega \rightarrow X(\omega)$$

- **Vecteur aléatoire réel ou variable aléatoire vectorielle** : à chaque ω est associée n éléments de $\mathbb{R}$

$$\omega \rightarrow [X_1(\omega), X_2(\omega), \dots, X_n(\omega)]$$

- **Fonction aléatoire réelle** : variable aléatoire vectorielle avec une infinité de composantes

- **Dénombrable** : $Z_1(\omega), Z_2(\omega), \dots, Z_i(\omega), \dots$

- **Non dénombrable** : fonction $Z(t)$, fonction $Z(x)$

- **Champ aléatoire réel** : fonction aléatoire dans un espace à plusieurs dimensions : $Z(x,y)$, $Z(x,y,z)$, $Z(x,t)$...

➤ Variable aléatoire vectorielle :

▪ **Variable aléatoire réelle à 2 composantes** : $\omega \rightarrow (Z_1(\omega), Z_2(\omega)) \in \mathbb{R}^2$

▪ **Fonction de répartition conjointe :**

$$F_{Z_1, Z_2} : \mathbb{R}^2 \rightarrow [0, 1]$$

$$(z_1, z_2) \rightarrow F_{Z_1, Z_2}(z_1, z_2) = P(Z_1 \leq z_1, Z_2 \leq z_2)$$

▪ **Densité de probabilité conjointe :**

$$F_{Z_1, Z_2}(z_1, z_2) = \int_{-\infty}^{z_2} \int_{-\infty}^{z_1} f_{Z_1, Z_2}(u, v) du dv$$

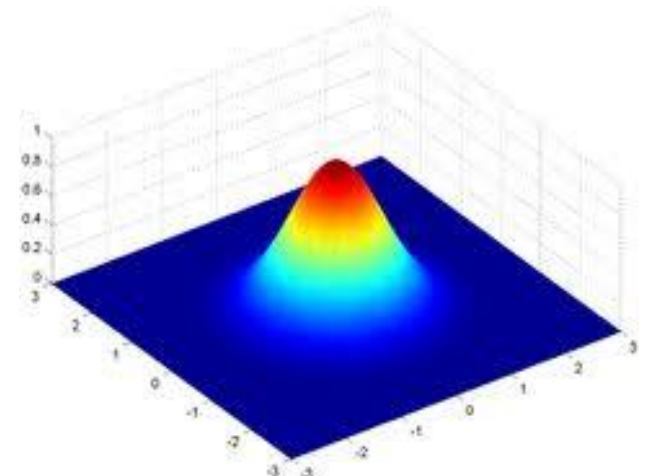
▪ **Indépendance entre Z_1 et Z_2** $\iff f_{Z_1, Z_2}(z_1, z_2) = f_{Z_1}(z_1) f_{Z_2}(z_1)$

▪ **Covariance et coefficient de corrélation :**

$$\begin{aligned} \text{cov}(Z_1, Z_2) &= E[Z_1 Z_2] - E[Z_1]E[Z_2] \\ &= E[(Z_1 - E[Z_1])(Z_2 - E[Z_2])] \end{aligned}$$

$$\rho_{1,2} = \frac{\text{Cov}(Z_1, Z_2)}{\sqrt{\text{Var}(Z_1)\text{Var}(Z_2)}} = \frac{\text{Cov}(Z_1, Z_2)}{\sigma_1 \sigma_2}$$

$$-1 \leq \rho_{1,2} \leq 1$$



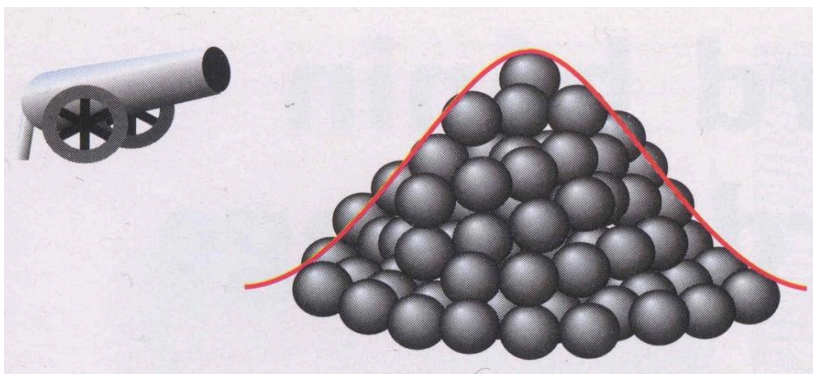
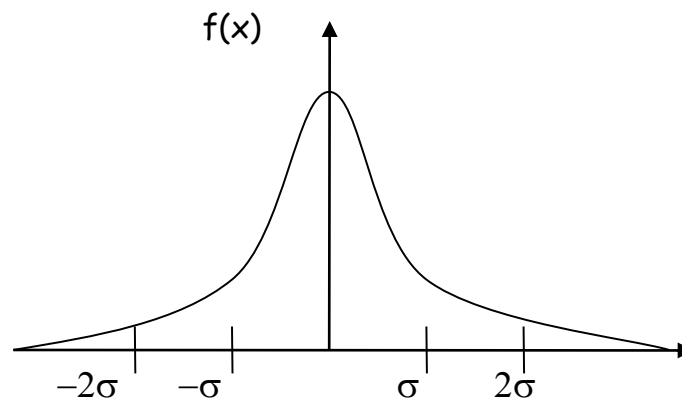
➤ Loi gaussienne ou loi normale $\mathcal{N}(\mu, \sigma^2)$

▪ Densité

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

▪ Moments

$$E(X) = \mu ; \text{var}(X) = \sigma^2$$



$$\begin{aligned} P(\mu - \sigma < X < \mu + \sigma) &= 0.68 \\ P(\mu - 1.64\sigma < X < \mu + 1.64\sigma) &= 0.90 \\ P(\mu - 1.96\sigma < X < \mu + 1.96\sigma) &= 0.95 \\ P(\mu - 3.09\sigma < X < \mu + 3.09\sigma) &= 0.998 \end{aligned}$$

Exemples : impacts des boulets de canon (Jouffret, 1872)

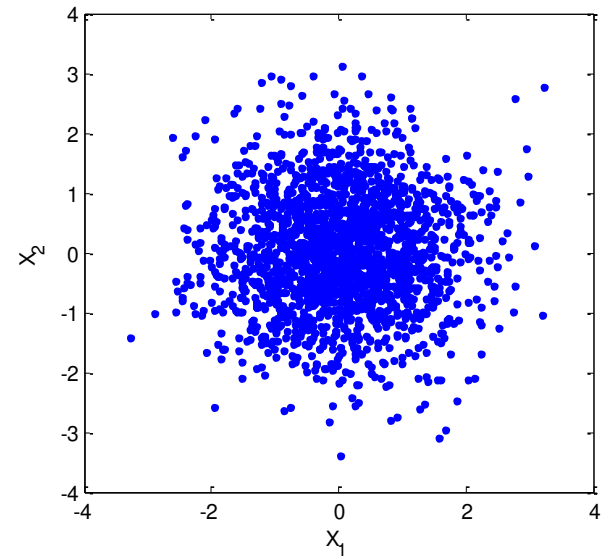
➤ Vecteur gaussien

▪ Loi $\mathcal{N}(0, I_n)$

$$\mathbf{Z} = (Z_1, Z_2, \dots, Z_n)^T \sim \mathcal{N}(0, I_n)$$



tous les Z_i ($i=1, \dots, n$) suivent indépendamment des lois gaussiennes centrées réduites $\sim \mathcal{N}(0, 1)$



▪ Vecteur gaussien d-dimensionnel $\mathbf{Z}$

- Si toute combinaison linéaire de ses composantes suit une loi gaussienne
- Si sa fonction caractéristique est de la forme :

$$\forall u \in \mathbb{R}^n, \phi_{\mathbf{Z}}(u) = \mathbb{E}[e^{i\langle u, \mathbf{Z} \rangle}] = e^{-i\langle u, m \rangle - \frac{1}{2}\langle u, Au \rangle}$$

avec $m \in \mathbb{R}^n$ et $A \in M^n(\mathbb{R})$ une matrice symétrique positive

- Si $\mathbf{Z}$ s'écrit comme une transformation affine d'une $\mathcal{N}(0, I_n)$:

$$\mathbf{Z} = m + A\mathbf{N} \quad \text{où } \mathbf{N} \sim \mathcal{N}(0, I_n)$$

avec $m \in \mathbb{R}^n$ et $A \in \mathcal{M}_n(\mathbb{R})$

➤ Vecteur gaussien

▪ Propriétés :

Transformation affine

- Si $A \in \mathcal{M}_{n,k}(\mathbb{R})$, $b \in \mathbb{R}^n$ et $Z \sim \mathcal{N}_n(m, \Sigma)$, alors

$$AZ + b \sim \mathcal{N}(Am + b, A \Sigma A^t)$$

Utilisation pour simuler un vecteur gaussien

$Z \sim \mathcal{N}(0, I_n)$ et Σ matrice symétriquedéfinie positive

alors $m + \Sigma^{\frac{1}{2}} Z \sim \mathcal{N}(m, \Sigma)$



Simulation à partir d'une $\mathcal{N}(0, I_n)$

➤ Vecteur gaussien $Z = (Z_1, Z_2, \dots, Z_n)^T \sim \mathcal{N}(m, \Sigma)$

- Vecteur de la moyenne

$$m = E(Z) = [\mu_1, \dots, \mu_n]^T = [E(Z_1), \dots, E(Z_n)]^T$$

- Matrice de covariance (hypo : $\text{Var}[Z_i] = \sigma_i^2 > 0$)

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \text{Cov}(Z_1, Z_2) & \dots & \text{Cov}(Z_1, Z_n) \\ \text{Cov}(Z_1, Z_2) & \sigma_2^2 & \dots & \text{Cov}(Z_2, Z_n) \\ \dots & \dots & \dots & \dots \\ \text{Cov}(Z_1, Z_n) & \text{Cov}(Z_2, Z_n) & \dots & \sigma_n^2 \end{pmatrix}$$

$$\text{avec } \begin{cases} \text{Cov}(Z_i, Z_j) = E[(Z_i - E[Z_i])(Z_j - E[Z_j])] \\ \sigma_i^2 = \text{Cov}(Z_i, Z_i) = \text{Var}[Z_i] \end{cases}$$

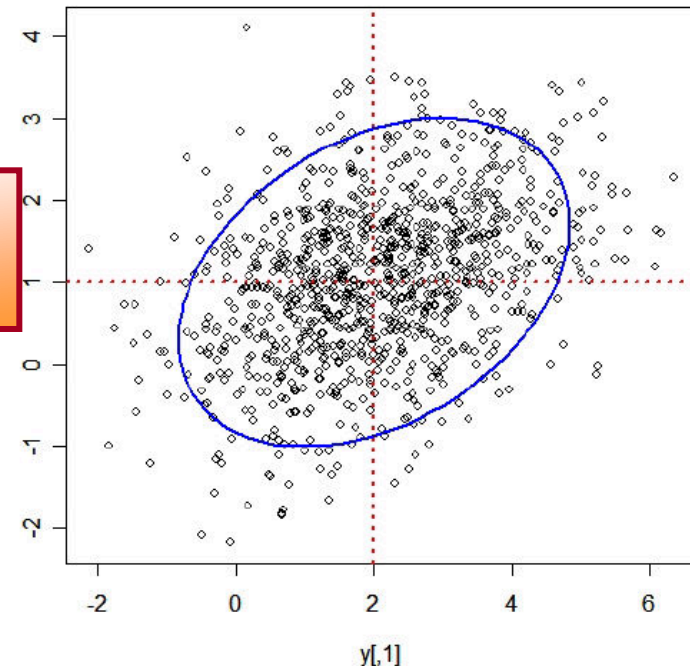
➤ Vecteur gaussien

- Loi bi-gaussienne : $\mathbf{Z} = (Z_1, Z_2)^T \sim \mathcal{N}(\mathbf{m}, \Sigma)$

$$\begin{cases} \mathbf{m} = [\mu_1, \mu_2]^T \\ \Sigma = \begin{pmatrix} \sigma_1^2 & \text{Cov}(Z_1, Z_2) \\ \text{Cov}(Z_1, Z_2) & \sigma_2^2 \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \sigma_{1,2} \\ \sigma_{1,2} & \sigma_2^2 \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \end{cases}$$

Simulation d'un échantillon de 1000 points
avec $\mathbf{m} = (2, 1)^T$; $\sigma_1^2 = 2$; $\sigma_2^2 = 1$ et $\rho = 0.35$

Ellipse de confiance à 95%



➤ Vecteur gaussien

▪ Propriétés :

- Les composantes d'un vecteur gaussien sont gaussiennes (Attention : la réciproque n'est pas vraie)
- La somme de 2 vecteurs indépendants est un vecteur gaussien

▪ **Dépendance $\Leftrightarrow$ corrélation**

▪ Représentation de loi multi-gaussienne : $Z \sim \mathcal{N}(m, \Sigma)$

- Cas non-dégénéré : Σ symétrique définie positive

$$\Rightarrow Y = \Sigma^{-\frac{1}{2}}(Z - m) \sim N(0, I_n)$$

$$\Rightarrow Z = m + \Sigma^{\frac{1}{2}}Y$$



**Simulation à partir
d'une $\mathcal{N}(0, I_n)$**

Z admet une densité :

$$\forall z \in R^n, f_Z(z) = \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} e^{-\frac{1}{2} \left[\Sigma^{-\frac{1}{2}}(z-m) \right]^T \left[\Sigma^{-\frac{1}{2}}(z-m) \right]}$$

➤ Vecteur gaussien

- **Propriétés de projection** : Théorème de Cochran
- **Statistiques des échantillons gaussiens**: moyenne et variance empiriques indépendantes et resp. de loi normale et khi2

- **Conditionnement** : $Z = (Z_1, Z_2)^T$

Loi jointe : $Z \sim \mathcal{N}(m, \Sigma)$ avec $m = (m_1, m_2)^T$ et $\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{1,2} \\ \sigma_{1,2} & \sigma_2^2 \end{pmatrix}$

Loi conditionnelle Z_2 sachant $Z_1 = z_1$:

- $Z_2|Z_1=z_1$ suit une loi gaussienne moyenne et variance :

$$\begin{cases} E[Z_2|Z_1 = z_1] = m_2 + \sigma_{1,2}\sigma_1^{-2}(z_1 - m_1) \\ \text{Var}[Z_2|Z_1 = z_1] = \sigma_2^2 - \sigma_{1,2}\sigma_1^{-2}\sigma_{1,2} \end{cases}$$

Rappel conditionnement : $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$.

Pour des variables continues, densité de $Z_2|Z_1=z_1$: $f_{Z_2|Z_1=z_1}(z_2) = \frac{f_{Z_2, Z_1}(z_2, z_1)}{f_{Z_1}(z_1)}$

➤ Vecteur gaussien

▪ Conditionnement :

Extension au cas général $Z'_1 = (Z_1, \dots, Z_p)^T$ et $Z'_2 = (Z_{p+1}, \dots, Z_n)^T$

$$(Z'_1, Z'_2)^T \sim N \left(\begin{pmatrix} m'_1 \\ m'_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{1,2} \\ \Sigma_{1,2}^T & \Sigma_{22} \end{pmatrix} \right)$$

avec m'_1 vecteur $p \times 1$, m'_2 vecteur $(n-p) \times 1$,

Σ_{11} matrice $p \times p$, Σ_{22} matrice $(n-p) \times (n-p)$

et $\Sigma_{1,2}$ matrice $(p) \times (n-p)$

⇒ loi conditionnelle de $Z'_2 | Z'_1 = z'_1$:

$$Z'_2 | Z'_1 = z'_1 \sim N \left(m'_2 + \Sigma_{1,2}^T \Sigma_{11}^{-1} (z'_1 - m'_1), \Sigma_{22} - \Sigma_{1,2}^T \Sigma_{11}^{-1} \Sigma_{1,2} \right)$$

➤ Processus Gaussiens (PG)

- **Définition** : un processus gaussien est un processus aléatoire réel $\{Z(x)\}_{x \in D \subset \mathbb{R}^d}$ dont toutes les lois finies-dimensionnelles $(Z(x_1), \dots, Z(x_n))$ sont gaussiennes.
- **Propriétés** : entièrement caractérisé par sa moyenne et sa fonction de covariance

$$Z(x) \sim PG(m(x), C(x, x'))$$

$$\text{où } m(x) = E[Z(x)]$$

$$\text{et } C(x, x') = \text{Cov}(Z(x), Z(x')) = E[(Z(x) - m(x)) (Z(x') - m(x'))]$$

- **Processus centré** : $m(x) = 0$
- **Covariance stationnaire** : $C(x, x') = C(x - x')$
- **Covariance isotrope** : $C(x, x') = C(\|x - x'\|)$
- **Fonction de corrélation R** : $C(x, x') = \sigma^2 R(x - x')$ avec $\sigma^2 = \text{Var}[Z(x)]$

Rappels statistiques et théorie

-  Variable, vecteur et processus gaussien

Définition du métamodèle processus gaussien (PG)

-  Du krigage à la régression par PG
-  Formulation analytique
-  Analogie avec le krigage
-  Simulation conditionnelles

Construction du métamodèle PG

-  Choix de la fonction moyenne
-  Choix de la fonction de covariance
-  Estimation des paramètres

Utilisation du métamodèle PG

-  Validation du métamodèle PG
-  Illustration sur des exemples analytiques
-  Plans d'expériences et planification adaptative

Application sur des cas industriels

➤ Du krigeage à la régression par processus gaussiens

- **Géostatistique/statistique spatiale** : application de la théorie des fonctions aléatoires à des données localisées dans un espace géographique,
=> Application de méthodes probabilistes pour l'étude de phénomènes corrélés dans l'espace
- **1955 : Daniel Krige**, mines d'or sud-africaines
Approche empirique pour corriger les problèmes de biais lors de l'estimation de la teneur d'un bloc de minerai à partir d'un nombre limité d'échantillons pris autour du bloc
- **1960-1970 : Georges Matheron**
Formalisation théorique et introduction d'un outil pour analyser la continuité spatiale des variables d'intérêt : le variogramme
=> Méthode d'estimation basée sur le variogramme : **krigeage**

➤ Du krigeage à la régression par processus gaussiens

- **1970-actuellement** : Extension des méthodes géostatistiques à d'autres domaines : pétrole, environnement, hydrologie, océanographie, foresterie, **codes de calcul**, ...

L'idée du krigeage pour les codes est d'interpoler les réponses du code en dimension p comme pour une cartographie spatiale

- **Différents points de vue** :
 - Weight space view => Krigeage => estimateur BLUP
 - Function space view => vecteurs gaussiens, loi jointe et loi conditionnelle => métamodèle processus Gaussien

2 approches différentes mais qui conduisent au même métamodèle



Métamodèle krigeage ou métamodèle processus gaussien

➤ Formulation analytique du métamodèle PG :

- Entrées du code : $\mathbf{x} \in \mathbf{D} \subset \mathbf{R}^d$

- Sortie du code : $\mathbf{z}(\mathbf{x}) \in \mathbf{R}$

- Hypothèse :

la sortie déterministe $\mathbf{z}(\mathbf{x})$ est considérée comme la réalisation d'un champ aléatoire $\mathbf{Z}(\mathbf{x})$:

$$\mathbf{Z}(\mathbf{x}) = f_o(\mathbf{x}) + \mathbf{W}(\mathbf{x})$$

avec :

- f_o la fonction moyenne
- $\mathbf{W}(\mathbf{x})$ un processus gaussien centré stationnaire de variance σ^2 et de fonction de corrélation $\mathbf{R}$:

$$\text{Cov}(\mathbf{W}(\mathbf{x}), \mathbf{W}(\mathbf{x}')) = c(\mathbf{x} - \mathbf{x}') = \sigma^2 \mathbf{R}(\mathbf{x} - \mathbf{x}')$$

➤ Formulation analytique du métamodèle PG :

▪ Base d'apprentissage (BA) :

- N points de simulations du code de calcul $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$

$$\mathbf{X}_s = [\mathbf{x}_1^T, \dots, \mathbf{x}_n^T]^T$$

- N réalisations/observations du code associées :

$$\mathbf{Z}_s = \mathbf{Z}(\mathbf{X}_s) = [\mathbf{Z}(\mathbf{x}_1), \dots, \mathbf{Z}(\mathbf{x}_n)]^T$$

▪ Loi jointe de la BA: $\mathbf{Z}_s \sim \mathcal{N}(\mathcal{F}_s, \Sigma_s)$

avec $\mathcal{F}_s = f_o(\mathbf{X}_s) = [f_o(\mathbf{x}_1), \dots, f_o(\mathbf{x}_n)]^T$ et Σ_s la matrice de covariance
 $(\Sigma_s)_{i,j} = c(\mathbf{x}_i - \mathbf{x}_j) = \sigma^2 R(\mathbf{x}_i - \mathbf{x}_j)$

▪ Point de prédiction : nouveau point non échantillonné $\mathbf{x}^*$

➤ Formulation analytique du métamodèle PG :

- Loi jointe $\{BA+x^*\}$:

$$[Z(X_s), Z(x^*)] \sim N \left(\begin{pmatrix} F_s \\ f_0(x^*) \end{pmatrix}, \begin{pmatrix} \Sigma_s & k(x^*) \\ k(x^*) & \sigma^2 \end{pmatrix} \right)$$

- Loi conditionnelle : $Z(x^*) | Z(X_s)$

$$Z(x^*)_{|Y(X_s)=Y_s} \sim N(\mu(x^*), \tilde{\sigma}^2(x^*))$$

$$\text{avec} \begin{cases} \mu(x^*) = E[Z(x^*) | Z(X_s) = Z_s] = f_0(x^*) + k(x^*)^T \Sigma_s (Z_s - F_s) \\ \tilde{\sigma}^2(x^*) = \text{Var}[Z(x^*) | Z(X_s) = Z_s] = \sigma^2 - k(x^*)^T \Sigma_s k(x^*) \end{cases}$$

➤ Formulation analytique du métamodèle PG :

- Prédicteur du métamodèle PG :

$$\hat{Z}(x^*) = \mu(x^*) = f_0(x^*) + k(x^*)^T \Sigma_S (Z_S - F_S)$$

« Tendance déterministe »
Terme identique à celui des MC classiques

« Partie permettant l'interpolation »
Prise en compte de la configuration des données

=> Interpolateur exact des points de la base d'apprentissage :

$$\hat{Z}(X_S) = Z_S$$

=> Meilleur prédicteur linéaire sans biais (BLUP)

➤ Propriétés du métamodèle PG :

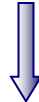
- Variance du prédicteur => Erreur du prédicteur

$$MSE[\hat{Z}(x^*)] = \tilde{\sigma}^2(x^*) = \sigma^2 - k(x^*)^T \Sigma_s k(x^*)$$

⇒ Erreur indépendante de la valeur des données, dépend uniquement de la covariance et du schéma d'échantillonnage

⇒ Visualisation des régions où l'estimation est imprécise et où il conviendrait de placer des nouveaux points de mesure

- Prédicteur et variance identiques à ceux du krigeage obtenus avec l'approche « Weight space view »



Krigeage universel

➤ Analogie avec le krigeage :

▪ Hypothèses :

- Processus spatial Z de covariance c connue
- n sites de mesures $\{s_1, \dots, s_n\} \Rightarrow$ observations de $\{Z(s_1), \dots, Z(s_n)\}$
- Z centré de moyenne nulle

▪ Objectif : Prédire le processus pour un nouveau site non observé s^*

▪ Idée : Construire un estimateur qui soit une combinaison linéaire des observations :

$$\hat{Z}(s^*) = \sum_{i=1}^n \lambda_i Z(s_i)$$

▪ Sans biais : $E[\hat{Z}(s^*) - Z(s^*)] = 0$ (1)

▪ De variance minimale : $Var[\hat{Z}(s^*) - Z(s^*)] = \min_{\lambda \in R^n} Var\left[\sum_{i=1}^n \lambda_i Z(s_i) - Z(s^*)\right]$ (2)

Estimateur BLUP

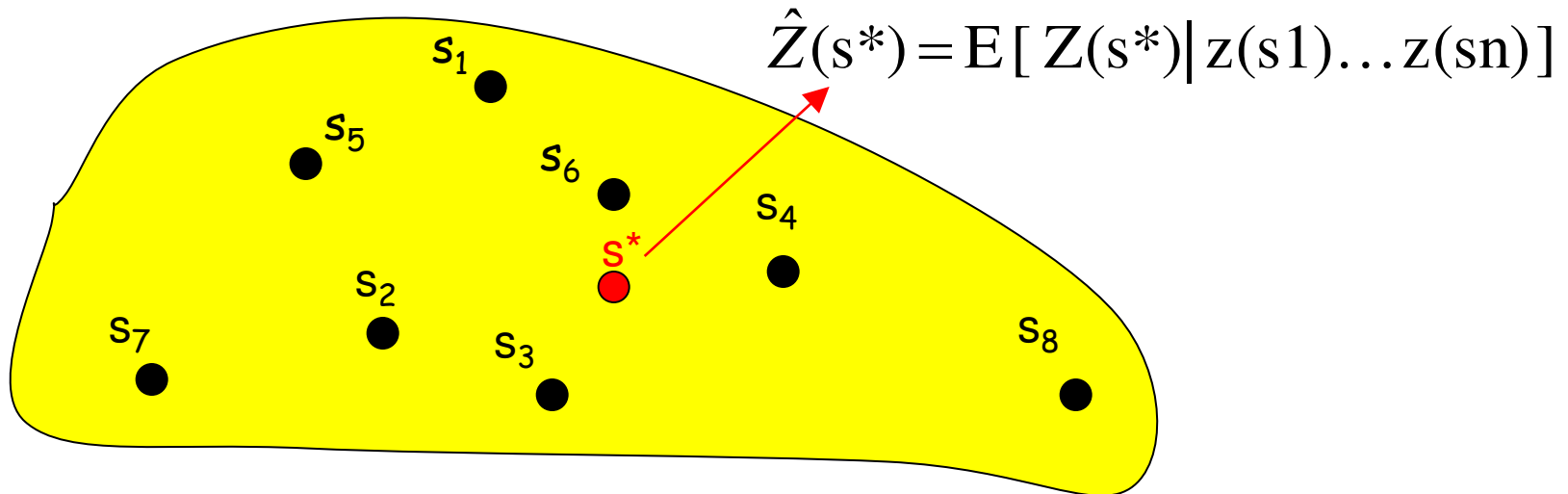
Détermination des λ_i

➤ Analogie avec le krigeage :

- Hypothèses : Meilleur prédicteur linéaire sans biais (BLUP)

Estimateur sans biais : $E [Z^*(s) - Z(s)] = 0$
la moyenne des erreurs est nulle

$Z^*(u)$ est optimal : $\text{Var} [Z^*(s) - Z(s)]$ est minimale
la dispersion des erreurs est réduite



➤ Analogie avec le krigeage :

▪ Détermination des λ_i :

Condition (2) implique q équations:

$$\text{Var}\left[\sum_i \lambda_i Z(s_i) - Z(s^*)\right] = \text{Var}[Z(s^*)] + \sum_i \sum_j \lambda_i \lambda_j \text{Cov}[Z(s_i), Z(s_j)] - 2 \sum_i \lambda_i \text{Cov}[Z(s_i), Z(s^*)]$$

Dérivation par rapport aux λ_i

=> système linéaire avec n équations à n inconnues

$$\sum_j \lambda_j \text{Cov}[Z(s_i), Z(s_j)] = \text{Cov}[Z(s_i), Z(s^*)] \quad \forall i = 1..n$$

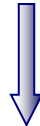
=> Réécriture sous forme matricielle du système linéaire :

$$\underbrace{\begin{bmatrix} \sigma^2 & \text{Cov}(Z_1, Z_2) & \bullet & \text{Cov}(Z_1, Z_n) \\ \text{Cov}(Z_2, Z_1) & \sigma^2 & \bullet & \text{Cov}(Z_2, Z_n) \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \text{Cov}(Z_n, Z_1) & \text{Cov}(Z_n, Z_2) & \bullet & \sigma^2 \end{bmatrix}}_{K_s} \underbrace{\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \bullet \\ \bullet \\ \lambda_n \end{bmatrix}}_{\lambda_s} = \underbrace{\begin{bmatrix} \text{Cov}(Z_1, Z_v) \\ \text{Cov}(Z_2, Z_v) \\ \bullet \\ \bullet \\ \text{Cov}(Z_n, Z_v) \end{bmatrix}}_{k_s}$$

➤ Analogie avec le krigage :

- Détermination des λ_i : Résolution du système linéaire :

$$K_s \lambda_s = k_s \quad \Rightarrow \quad \lambda_s = K_s^{-1} k_s$$



$$\begin{cases} \hat{Z}(s^*) = {}^t k_s K_s^{-1} Z \\ \text{Var}[\hat{Z}(s^*) - Z(s^*)] = \sigma^2 - {}^t k_s K_s^{-1} k_s \end{cases}$$

- Généralisation aux processus non centré :
 - Moyenne m constante et connue => krigage simple
 - Moyenne m constante et inconnue => krigage ordinaire
 - Moyenne fonction de s , $m(s)$ inconnue (présence d'une « dérive »)
=> krigage universel ou krigage avec dérive

➤ Formulation analytique du métamodèle PG/krigeage:

- **Loi conditionnelle : $Z(x^*) | Z(X_s)$**

$$Z(x^*)_{|Y(X_s)=Y_s} \sim N(\mu(x^*), \tilde{\sigma}^2(x^*))$$

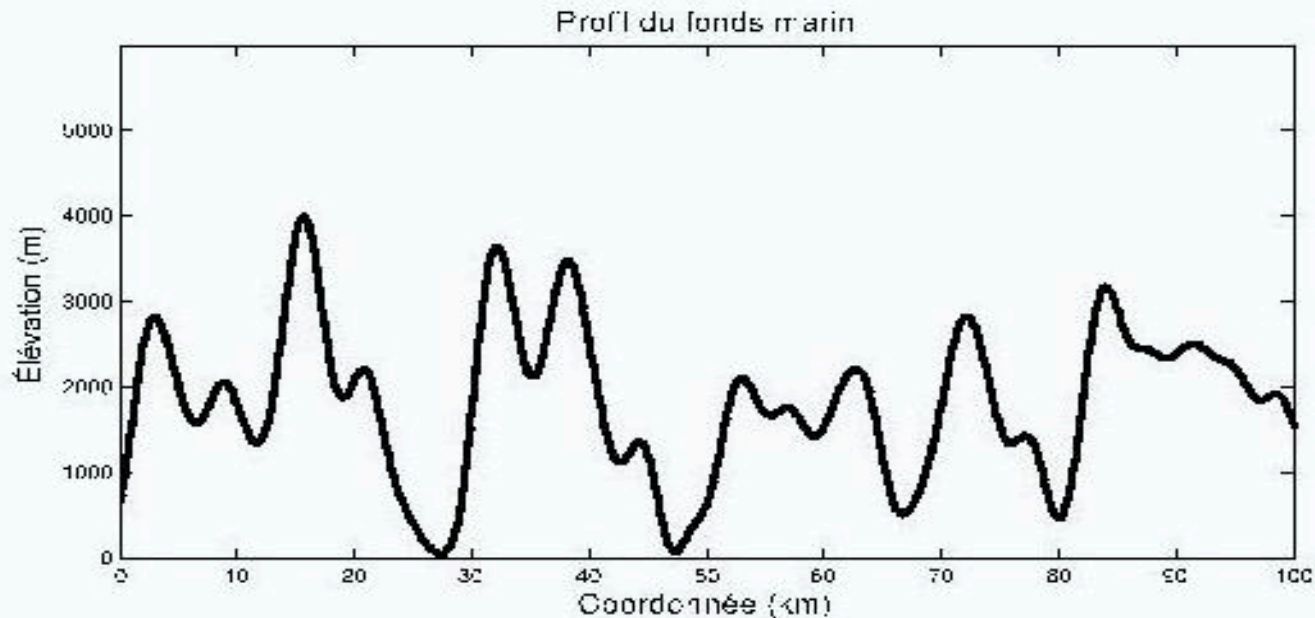
$$\text{avec} \begin{cases} \mu(x^*) = E[Z(x^*) | Z(X_s) = Z_s] = f_0(x^*) + k(x^*)^T \Sigma_s (Z_s - F_s) \\ \tilde{\sigma}^2(x^*) = \text{Var}[Z(x^*) | Z(X_s) = Z_s] = \sigma^2 - k(x^*)^T \Sigma_s k(x^*) \end{cases}$$

- **Le prédicteur du PG cherche l'estimation optimale** (sans biais, variance d'erreur minimale) de la variable en tout point, à partir de données expérimentales

- **Utilisation de la loi conditionnelle :**

- **Construction d'intervalle de confiance**
- **Réalisation de simulations conditionnelles :** une simulation de la loi conditionnelle représente une réalisation possible du phénomène réel. Elle cherche à reproduire sa **variabilité** tout en respectant les données expérimentales (conditionnement).

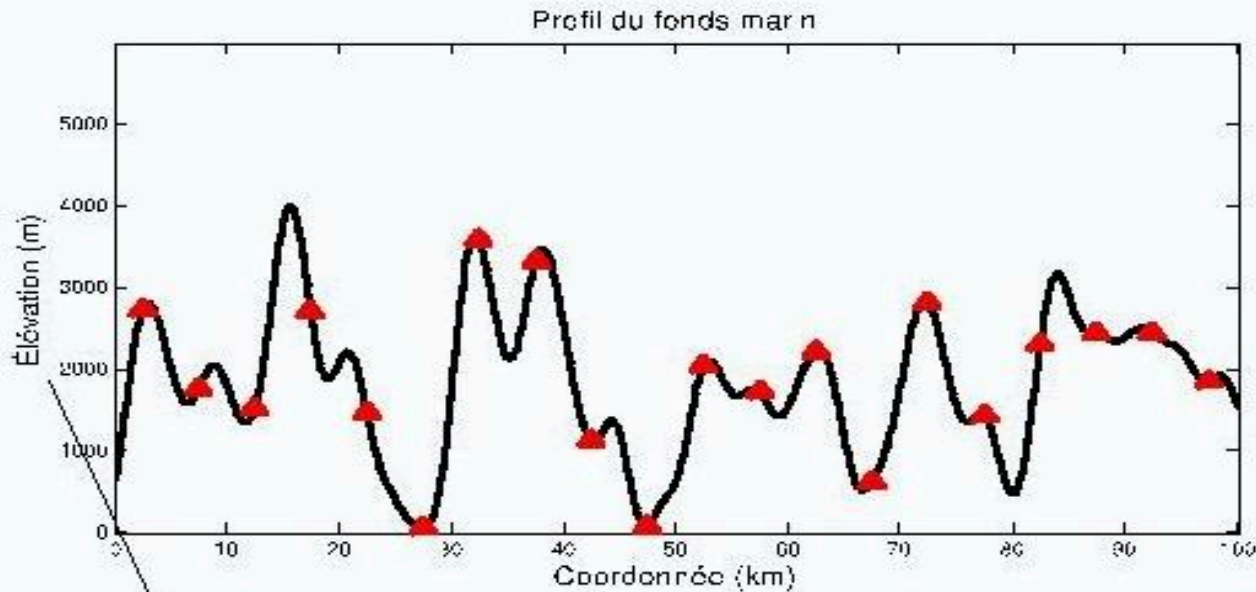
➤ Exemple de simulations : profil du fond marin



**Vous déposez un câble de fibre optique sur le fonds marin.
Quelle longueur devrait avoir ce câble ?**

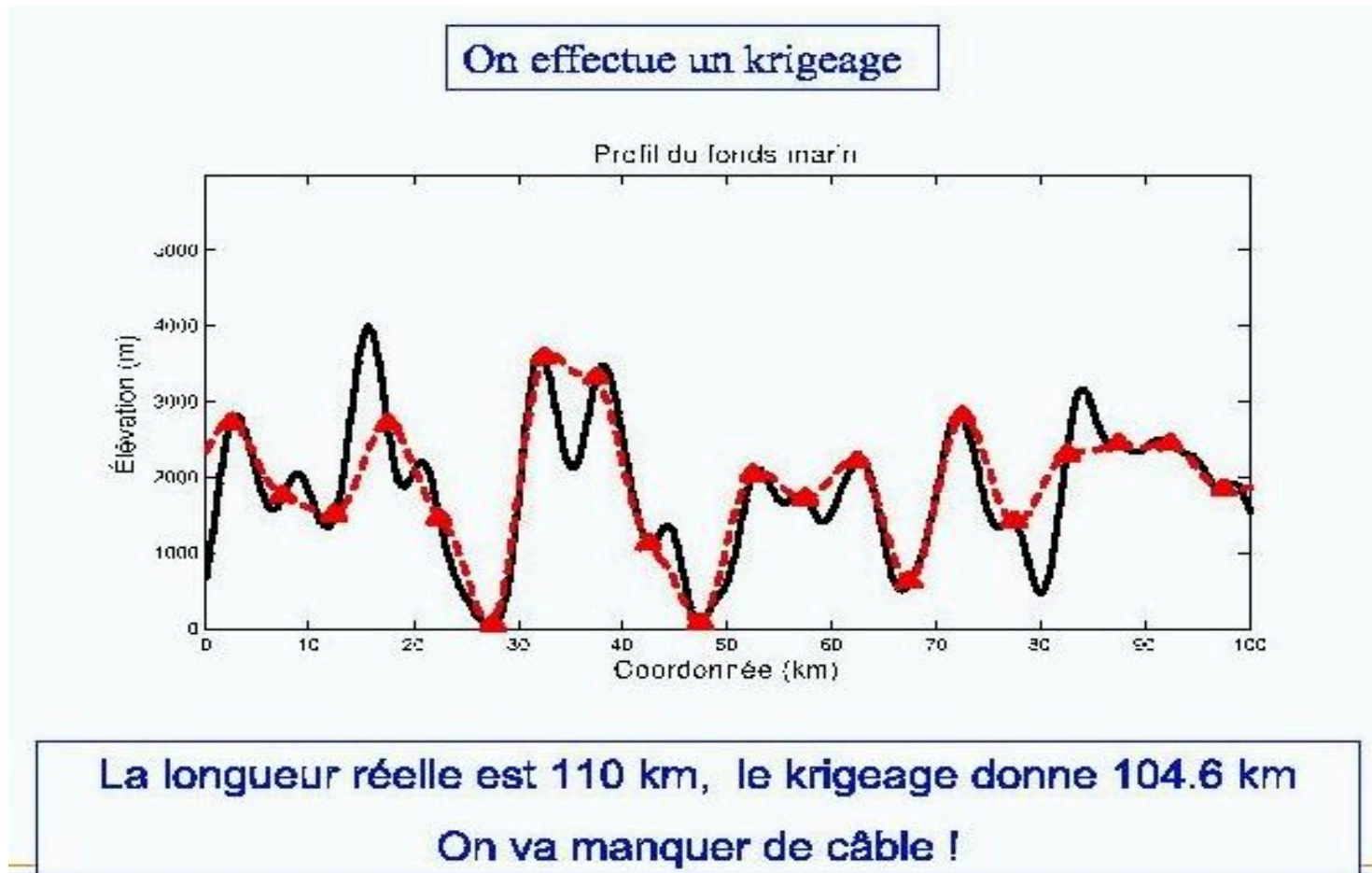
➤ Exemple de simulations : profil du fond marin

La profondeur exacte est connue uniquement aux points observations



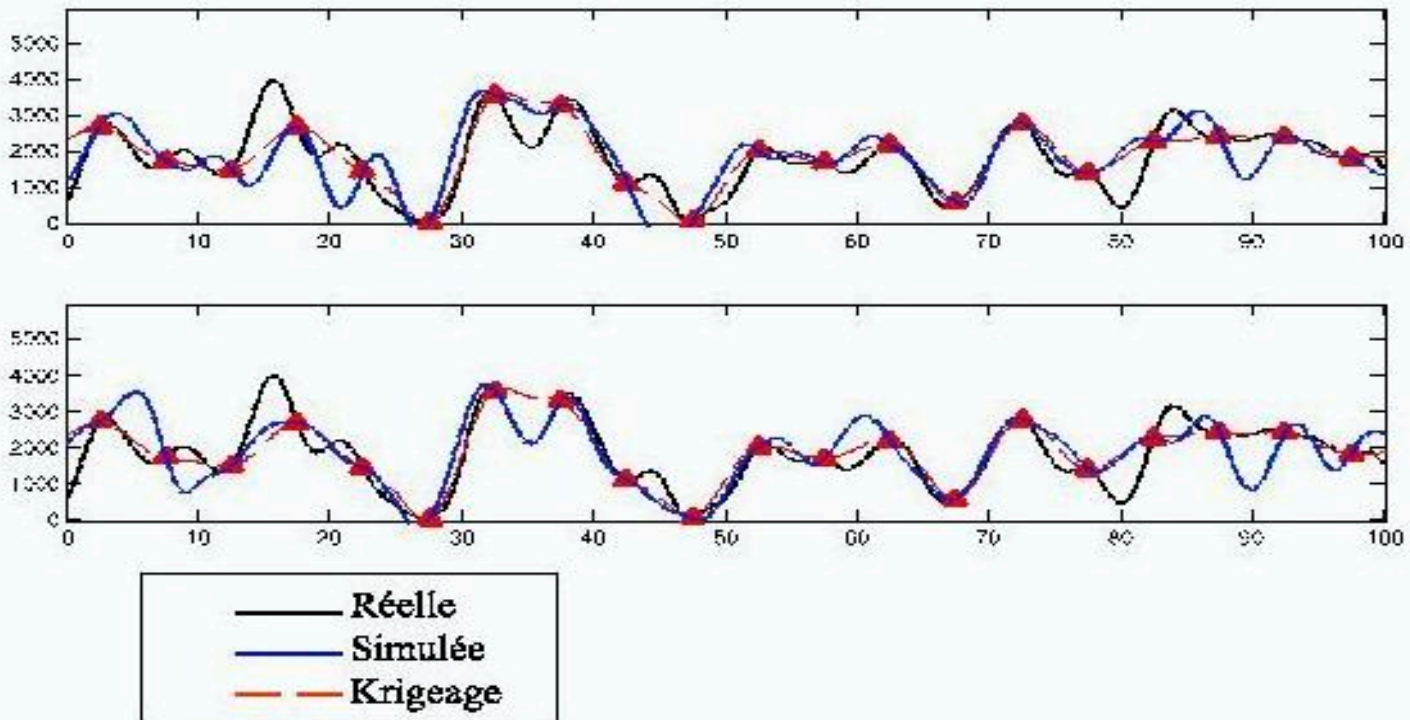
Exagération verticale environ 10

➤ Exemple de simulations : profil du fond marin

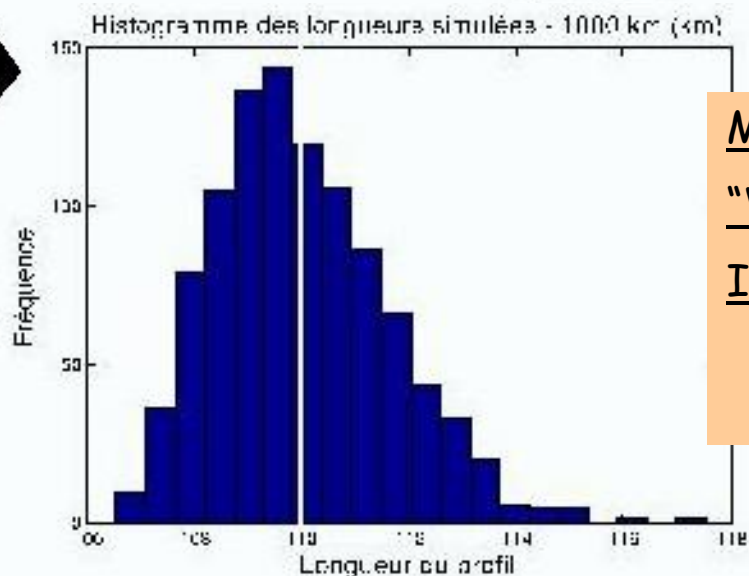
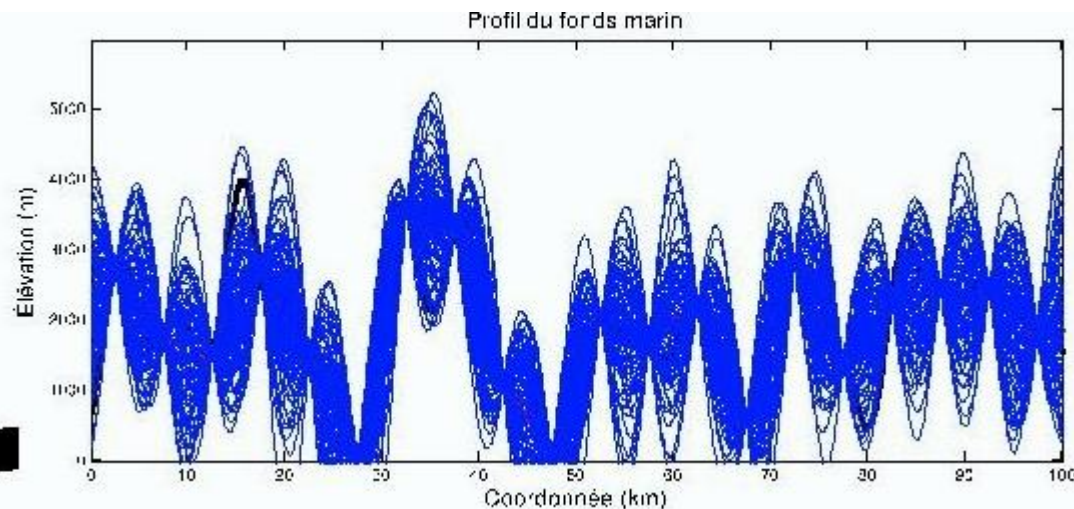


➤ Exemple de simulations : profil du fond marin

Une autre approche : des simulations conditionnelles



➤ Exemple de simulations : profil du fond marin



Moyenne des 1000 réalisations : 110

"Vraie" valeur : 110

Intervalle de confiance à 95 % : [108.8 ; 113.5]

=> prédicteur de krigeage à 104.2, non compris dans l'IC 95 % !!

➤ Exemple de simulations : profil du fond marin

Longueur de câble = f (prof. du fonds marin) avec f la fonction de transfert
Ici f est non linéaire

Or le prédicteur du PG est un estimateur linéaire (qui gomme les variabilités locales)

Pour l'estimation d'une moyenne, on peut utiliser le prédicteur du PG car la fonction de transfert est linéaire

Comment calculer la probabilité de dépasser un seuil ou un quantile ?

- **à partir des simulations** (hypothèse gaussienne et calculs lourds)
- **Krigeage d'indicatrices** : estimation d'indicatrices $1_{Z(x) \geq s}$

Rappels statistiques et théorie

- Variable, vecteur et processus gaussien

Définition du métamodèle processus gaussien (PG)

- Du krigeage à la régression par processus gaussien
- Formulation analytique
- Analogie avec le krigeage
- Simulation conditionnelles

Construction du métamodèle PG

- Choix de la fonction moyenne
- Choix de la fonction de covariance
- Estimation des paramètres

Utilisation du métamodèle PG

- Validation du métamodèle PG
- Illustration sur des exemples analytiques
- Plans d'expériences et planification adaptative

Application sur des cas industriels

▪ **Modèle** : $\mathbf{x} \in \mathcal{D} \subset \mathbb{R}^d$

$$\mathbf{Z}(\mathbf{x}) = f_o(\mathbf{x}) + W(\mathbf{x})$$

avec :

▪ f_o la fonction moyenne

▪ $W(\mathbf{x})$: PG centré stationnaire de variance σ^2 et de f_x de corrélation R

➤ Choix de la fonction « moyenne » f_o

▪ **Constante** $f_o(\mathbf{x}) = c$

▪ **Polynôme de degré 1** $f_o(\mathbf{x}) = \beta_0 + \sum_{i=1}^d \beta_i x_i$

▪ **Polynôme d'ordre supérieur**

▪ **Modèle additif généralisé**

$$f_o(\mathbf{x}) = \beta_0 + \sum_{i=1}^d \beta_i f_i(x_i) + \sum_{\substack{i,j=1 \\ i \neq j}}^d \beta_{i,j} f_{i,j}(x_i, x_j) + \dots$$

▪ **Modèle** : $x \in D \subset \mathbb{R}^d$

$$Z(x) = f_o(x) + W(x)$$

avec :

▪ f_o la fonction moyenne

▪ $W(x)$: PG centré stationnaire de variance σ^2 et de fx de corrélation R

➤ Choix de la fonction de covariance/corrélation

▪ **Propriétés par définition:**

- Symétrique $c(x, x') = c(x', x)$

- Semi-positive $Var[\sum_{i=1}^n \lambda_i Z(s_i)] \geq 0 \Rightarrow \begin{cases} \sum_{i,j=1}^n \lambda_i \lambda_j c(s_i, s_j) \geq 0 \quad \forall \lambda_j, \lambda_i \text{ réels} \\ c \text{ est une fonction de type positif} \end{cases}$

▪ **Covariance stationnaire :**

$$c(x, x') = c(x - x')$$

▪ **Covariance isotrope :**

$$c(x, x') = c(\|x - x'\|)$$

Rque : différentiabilité de c directement liée à la régularité du processus

➤ Choix de la fonction de covariance/corrélation

▪ Modèles paramétriques 1D:

- Covariance exponentielle $c(h) = \sigma^2 \exp\left(-\frac{h}{\theta}\right)$
- Covariance gaussienne $c(h) = \sigma^2 \exp\left(-\left(\frac{h}{\theta}\right)^2\right)$
- Covariance exponentielle généralisée $c(h) = \sigma^2 \exp\left(-\left(\frac{h}{\theta}\right)^p\right)$ avec $p > 0$
- Covariance de Matern : $c(h) = \sigma^2 \frac{1}{2^{\alpha-1} \Gamma(\alpha)} \left(\frac{h}{\theta}\right)^\alpha K_\alpha\left(\frac{h}{\theta}\right)$ avec $\alpha > 0$
 K_α est la fonction de Bessel

Cas particulier $\alpha = p + 1/2$, $p \in \mathbb{N} \Rightarrow$ exponentiel ($p=0$), Matern3/2 ($p=1$), Matern5/2 ($p=2$)

$$p = 1 \Rightarrow c(h) = \sigma^2 \left(1 + \frac{\sqrt{3}h}{\theta}\right) \exp\left(-\frac{\sqrt{3}h}{\theta}\right)$$

$$p = 2 \Rightarrow c(h) = \sigma^2 \left(1 + \frac{\sqrt{5}h}{\theta} + \frac{5h^2}{3\theta^2}\right) \exp\left(-\frac{\sqrt{5}h}{\theta}\right)$$

Rque : p directement liée à la régularité du processus (différentiabilité de c)

➤ Choix de la fonction de covariance/corrélation

■ Généralisation en dimension supérieure:

- Covariance isotrope : $c(h) = c(\|h\|)$

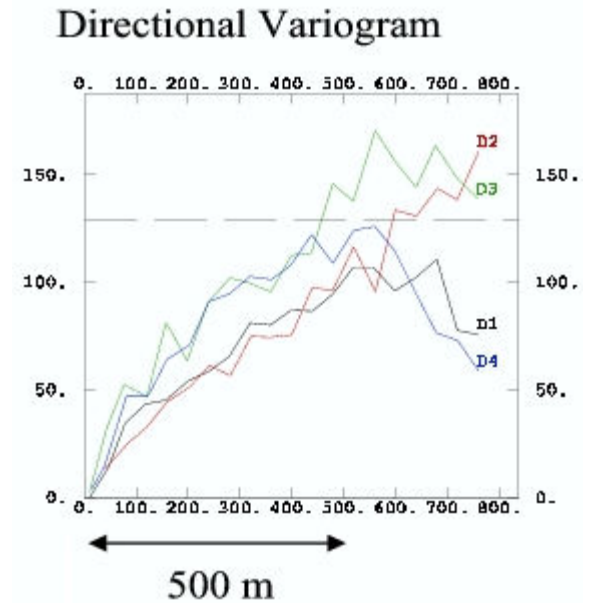
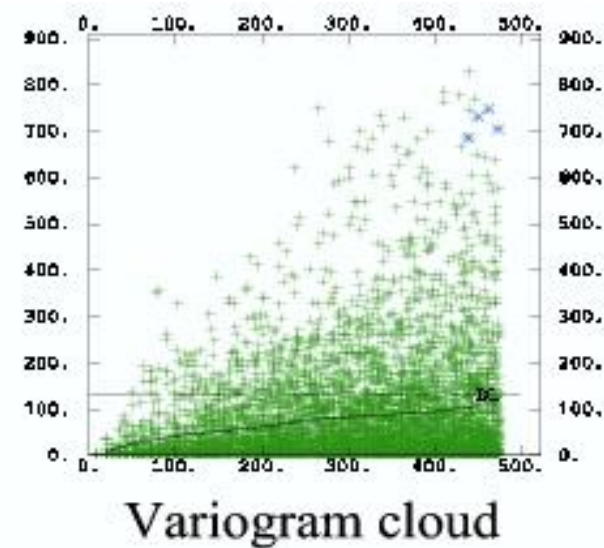
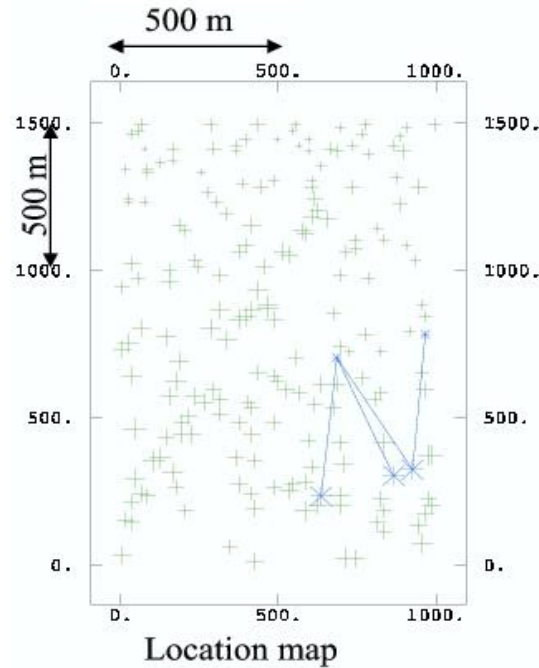
- Covariance anisotrope : $c(h) = \prod_{i=1}^d c_i(h_i)$

- Exponentielle $c(h) = \sigma^2 \prod_{i=1}^d \exp\left(-\frac{h_i}{\theta_i}\right) = \sigma^2 \exp\left(-\sum_{i=1}^d \frac{h_i}{\theta_i}\right)$

- Gaussienne $c(h) = \sigma^2 \exp\left(-\sum_{i=1}^d \left(\frac{h_i}{\theta_i}\right)^2\right)$

- ...

➤ Analogie avec le variogramme en géostatistique



[Géovariances]

Nuée variographique : pour tous les couples de point (x_1, x_2) , on calcule

$$\frac{1}{2}[Z(x_1) - Z(x_2)]^2 \quad \text{et on le reporte en fonction de } x_1 - x_2$$

➤ Analogie avec le variogramme en géostatistique

Hypothèse localement stationnaire (ou intrinsèque) : les incréments de $Z(x)$ sont stationnaires

$$\text{Variogramme : } \gamma(h) = 0.5 \text{ Var}[Z(x+h) - Z(x)]$$

$E[Z(x+h) - Z(x)]$ et $\gamma(h)$ caractérisent entièrement Z

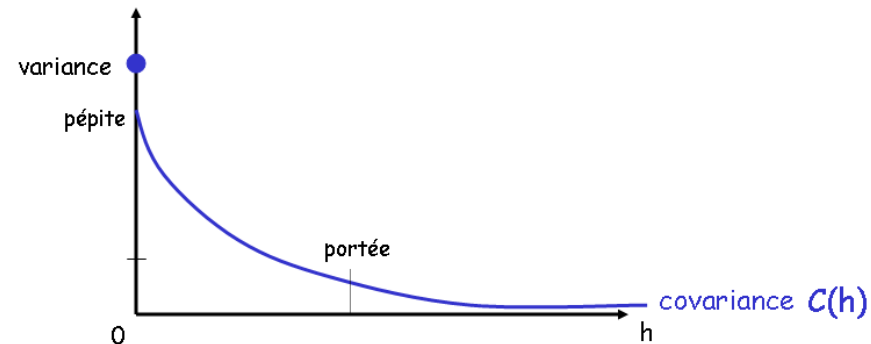
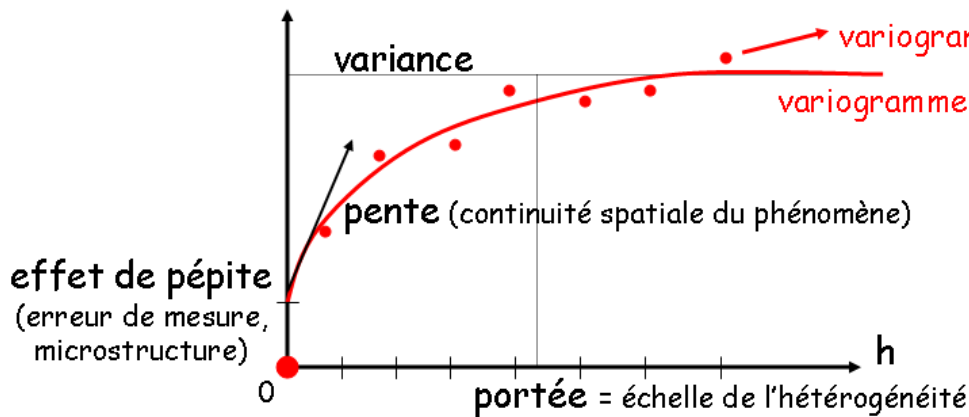
Stationnarité, $C(h)$

Moyenne et variance
de Z sont constantes

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

Stationnarité locale, $\gamma(h)$

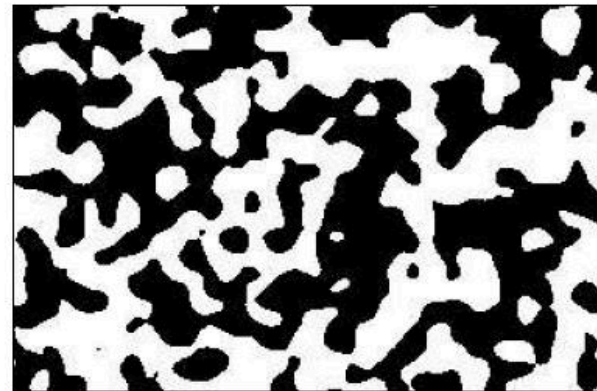
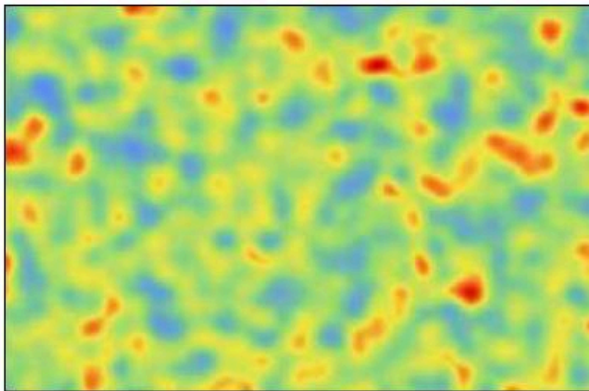
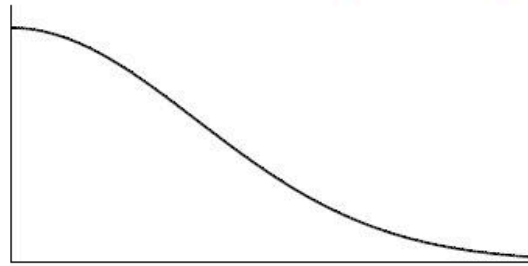
Moyenne et variance
de Z sont linéaires



➤ Choix de la fonction de covariance/corrélation

- Covariance gaussienne

$$C(h) = \exp \left\{ -\frac{|h|^2}{a^2} \right\}$$

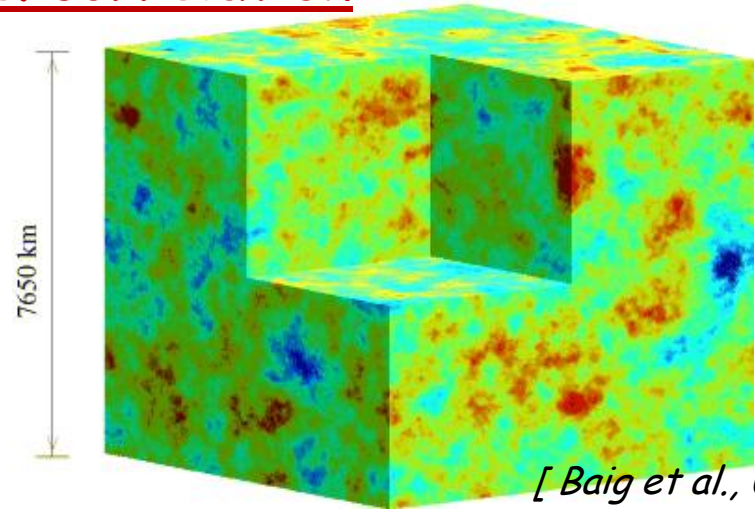
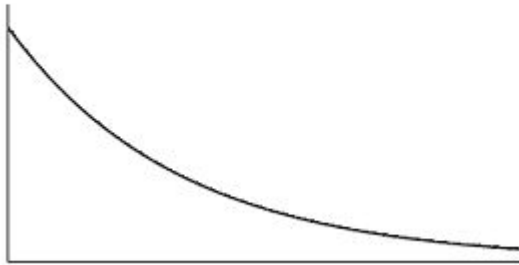


Simulation d'un champ 2D isotrope

➤ Choix de la fonction de covariance/corrélation

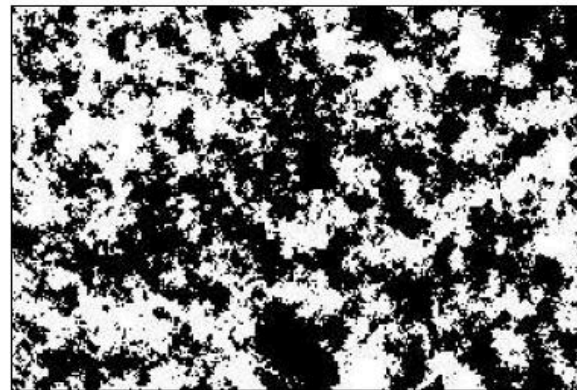
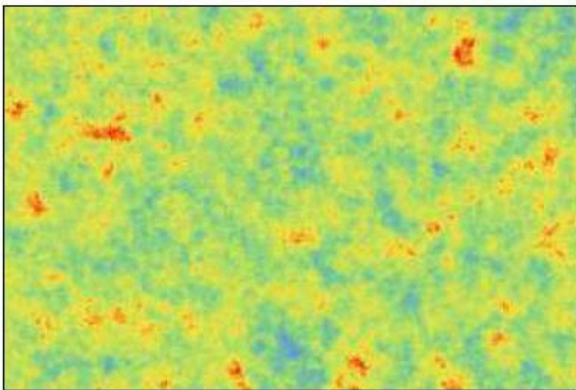
▪ Covariance exponentielle

$$C(h) = \exp \left\{ -\frac{|h|}{a} \right\}$$



[Baig et al., 03]

Exemples : milieux sédimentaires
(champs de porosité)



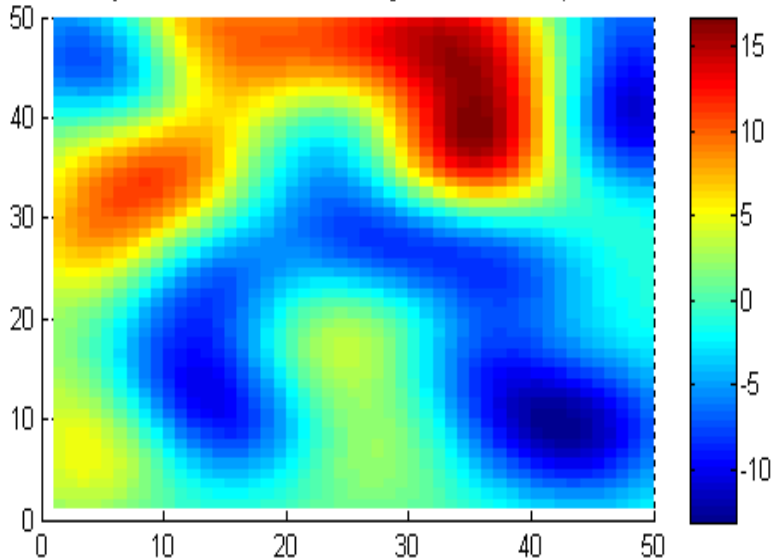
Simulation d'un champ 2D isotrope

➤ Choix de la fonction de covariance/corrélation

▪ Covariance exponentielle anisotrope

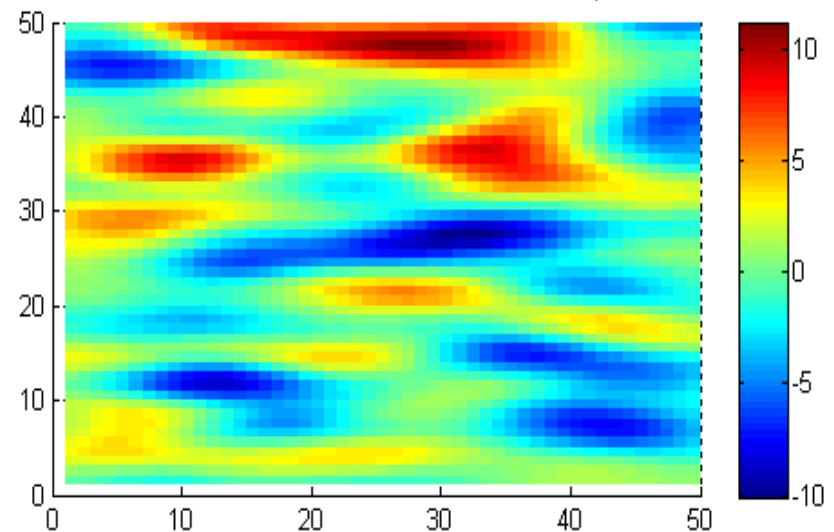
$$c(h) = \sigma^2 \exp \left(- \left(\frac{h_1}{\theta_1} \right)^2 - \left(\frac{h_2}{\theta_2} \right)^2 \right) \text{ avec } \theta_1 \neq \theta_2$$

Simulation trajectoires PG de covariance gaussienne isotrope avec $\theta = 0.1$



Simulation d'un champ 2D isotrope
(covariance gaussienne)

Simulation trajectoires PG de covariance gaussienne avec $\theta_1 = 0.1$ et $\theta = 0.03$



Simulation d'un champ 2D anisotrope
(covariance gaussienne)

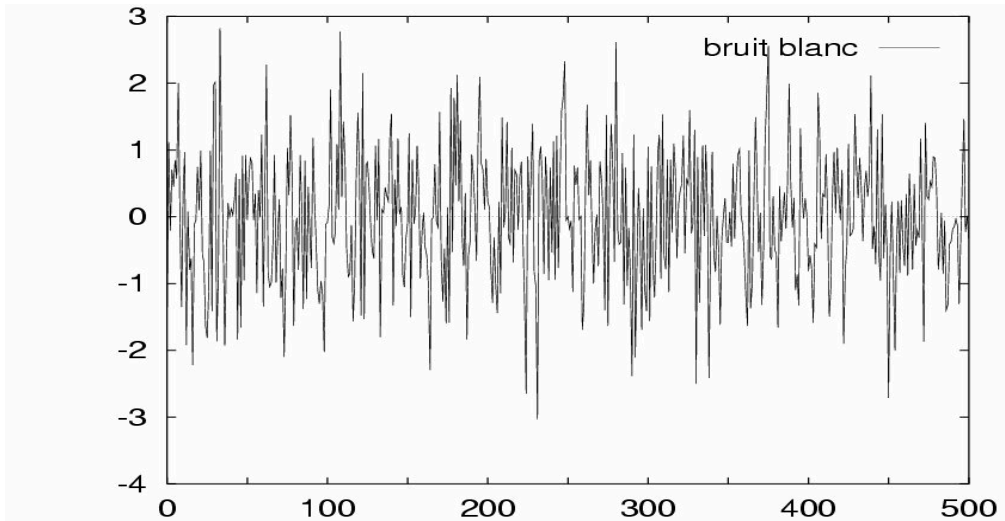
➤ Choix de la fonction de covariance/corrélation

- Effet de pépité

$$\tilde{c}(h) = \tilde{c}(h) + \varepsilon^2 1_{h=0}$$

- Discontinuité de la covariance à l'origine
- Bruit blanc additionnel
- Géostatique : modélisation d'hétérogénéités locales
- Pour le métamodèle PG :
 - Relaxation de la contrainte d'interpolateur exact des données
 - Amélioration du conditionnement de la matrice de covariance (rajout d'un terme sur la diagonale)

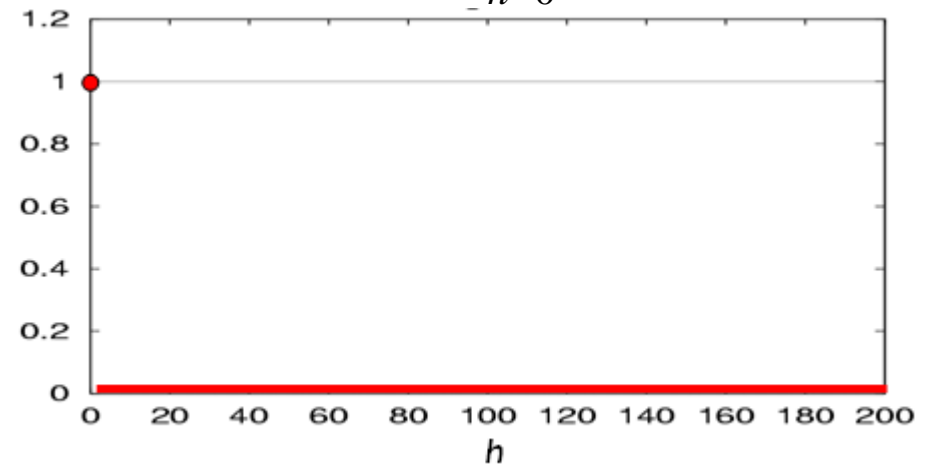
Effet de pépite pur => bruit blanc



Moyenne = 0

Variance = 1

$$c(h) = 1_{h=0}$$



➤ Estimation des paramètres du métamodèle PG

- Paramètres de la moyenne β
- Paramètre de variance σ^2
- Paramètres de la covariance θ : appelés « hyperparamètres »



Estimation par maximum de vraisemblance

- Log Vraisemblance sur la BA :

$$\text{Ln}L(Z_s, \beta, \sigma, \theta) = -\frac{N}{2} \text{Ln}(2\pi\sigma^2) - \frac{1}{2} \text{Ln}(\det \Sigma_s) - \frac{1}{2} \sigma^2 (Z_s - F_s)^T \Sigma_s^{-1} (Z_s - F_s)$$

- Maximisation :

$$(\beta^*, \sigma^*, \theta^*) = \underset{(\beta, \sigma, \theta)}{\text{Argmax}} \text{Ln}L(Z_s, \beta, \sigma, \theta)$$

➤ Estimation des paramètres du métamodèle PG

- Si f_0 de forme linéaire $f_0(x) = \beta_0 + \sum_{i=1}^d \beta_i x_i \Rightarrow F_S = \tilde{X}_S \beta$
avec $\tilde{X}_S = [1, x_1^T, \dots, x_n^T]^T$ et $\beta = [\beta_0, \beta_1, \dots, \beta_d]^T$

=> Solutions analytiques pour β^* et σ^{2*} en fonction de θ

$$\beta^*(\theta) = \left(\tilde{X}_S^T \Sigma_s^{-1} \tilde{X}_S \right)^{-1} \Sigma_{s,\theta}^{-1} \left(\tilde{X}_S^T \Sigma_s^{-1} Z_S \right)$$

$$\sigma^{2*}(\theta) = \frac{1}{N} \left(Z_S - \tilde{X}_S \beta^*(\theta) \right)^T \Sigma_{s,\theta}^{-1} \left(Z_S - \tilde{X}_S \beta^*(\theta) \right)$$

=> Estimation des hyperparamètres θ^* (injection de σ^{2*} dans LnL)

$$\theta^* = \underset{\theta}{\text{Argmax}} \varphi(\theta) \quad \text{avec} \quad \varphi(\theta) = \left| \Sigma_{s,\theta} \right|^{-\frac{1}{N}} \sigma^{2*}(\theta)$$



Optimisation numérique pour l'estimation des θ^*

➤ Etapes de construction du métamodèle PG :

- Choix de la fonction moyenne
- Choix de la fonction de covariance
- Estimation des paramètres :
 - paramètres de la moyenne et hyperparamètres
 - par maximum de vraisemblance sur la base d'apprentissage
- Construction métamodèle PG conditionnel aux points de la BA :
 - Moyenne = prédicteur
 - Variance = MSE du prédicteur

$$\hat{Z}(x^*) = f_0(x^*) + k(x^*)^T \Sigma_S (Z_S - F_S)$$

$$MSE[\hat{Z}(x^*)] = \hat{\sigma}^2 - k(x^*)^T \Sigma_S k(x^*)$$

Rappels statistiques et théorie

- Variable, vecteur et processus gaussien

Définition du métamodèle processus gaussien (PG)

- Du krigage à la régression par processus gaussien
- Formulation analytique
- Analogie avec le krigage
- Simulation conditionnelles

Construction du métamodèle PG

- Choix de la fonction moyenne
- Choix de la fonction de covariance
- Estimation des paramètres

Utilisation du métamodèle PG

- Validation du métamodèle PG
- Illustration sur des exemples analytiques
- Plans d'expériences et planification adaptative

Application sur des cas industriels

➤ Validation du métamodèle PG

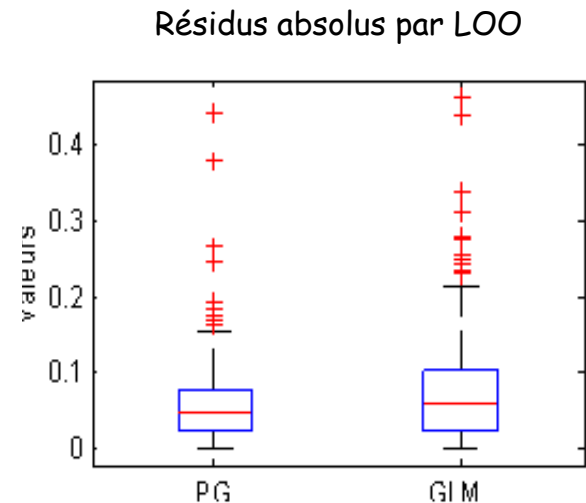
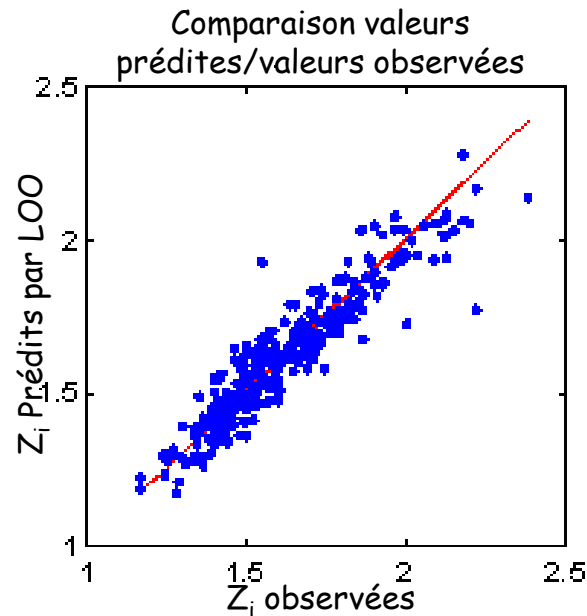
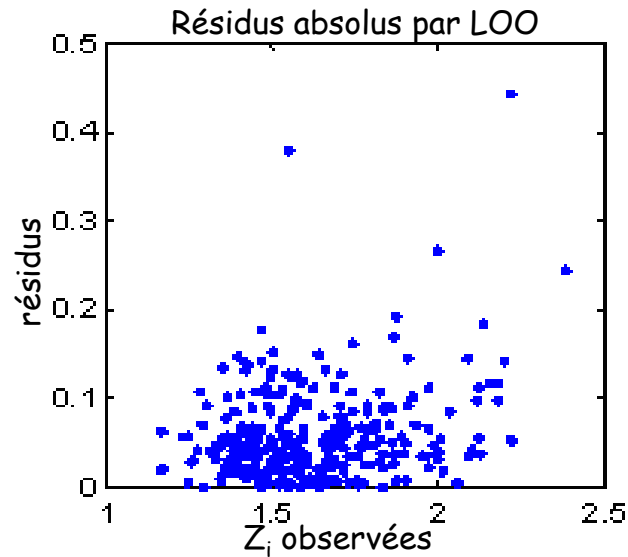
- **Qualité d'ajustement parfaite des données de la base d'apprentissage (BA):**
car interpolateur exact des points de la base d'apprentissage
- **Qualité de prédiction pour un nouveau point ?? => prédictivité ?**
 - **2 possibilités :**
 - **Utilisation d'une base de test :** points différents de la BA et non utilisés pour la construction du métamodèle
 - **Approche par validation croisée sur la BA :**
 - « leave one out » (LOO) (Hastie et al. 1990) : on enlève un point de la BA, on construit le métamodèle avec les points restants, puis on compare la prédiction du métamodèle pour le point enlevé avec sa vraie valeur. Répétition pour les n points de la BA => n prédictions et donc n résidus
 - Validation croisée par blocs



**Comparaison des valeurs observées et des valeurs
prédites: calcul des résidus**

➤ Validation du métamodèle PG

- ✓ Estimation des prédictions $\hat{Z}_i$: sur une base de test ou par **validation croisée**
- ✓ Validation graphique : Graphe des résidus, comparaison valeurs prédites/valeurs observées



➤ Validation du métamodèle PG

✓ Validation quantitative :

▪ Calcul du coefficient de prédictivité : Q_2

≈ 1 - écart quadratique normalisé entre prédictions et valeurs réelles

$$Q_2(Z, \hat{Z}) = 1 - \frac{\sum_{i=1}^N (Z_i - \hat{Z}_i)^2}{\sum_{i=1}^N (\bar{Z} - Z_i)^2} \quad \text{avec} \quad \begin{cases} Z_i : \text{valeurs observées} \\ \hat{Z}_i : \text{valeurs prédites par le métamodèle} \\ \bar{Z} : \text{moyenne des valeurs observées} \end{cases}$$

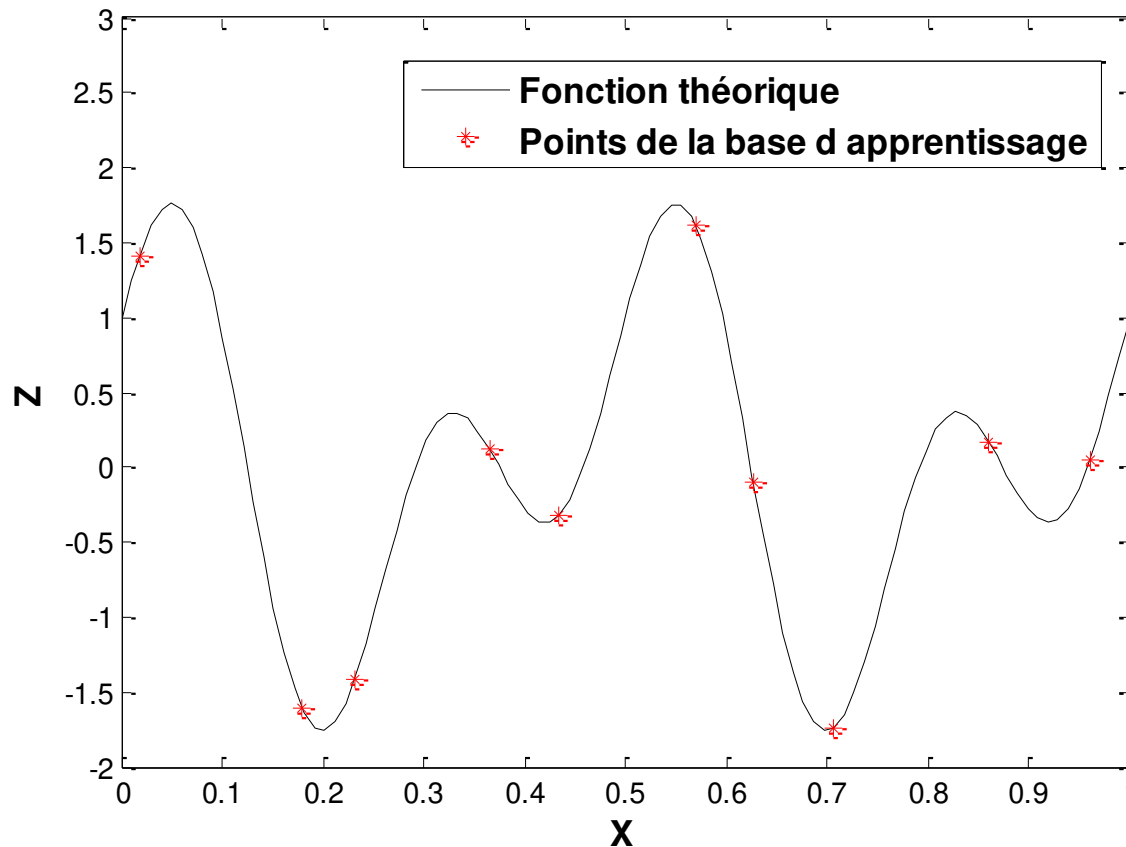
▪ Interprétation du Q_2 :

« Plus il est proche de 1, meilleur est l'ajustement du modèle aux observations ».

= Part de la variance de Y expliquée par le métamodèle.

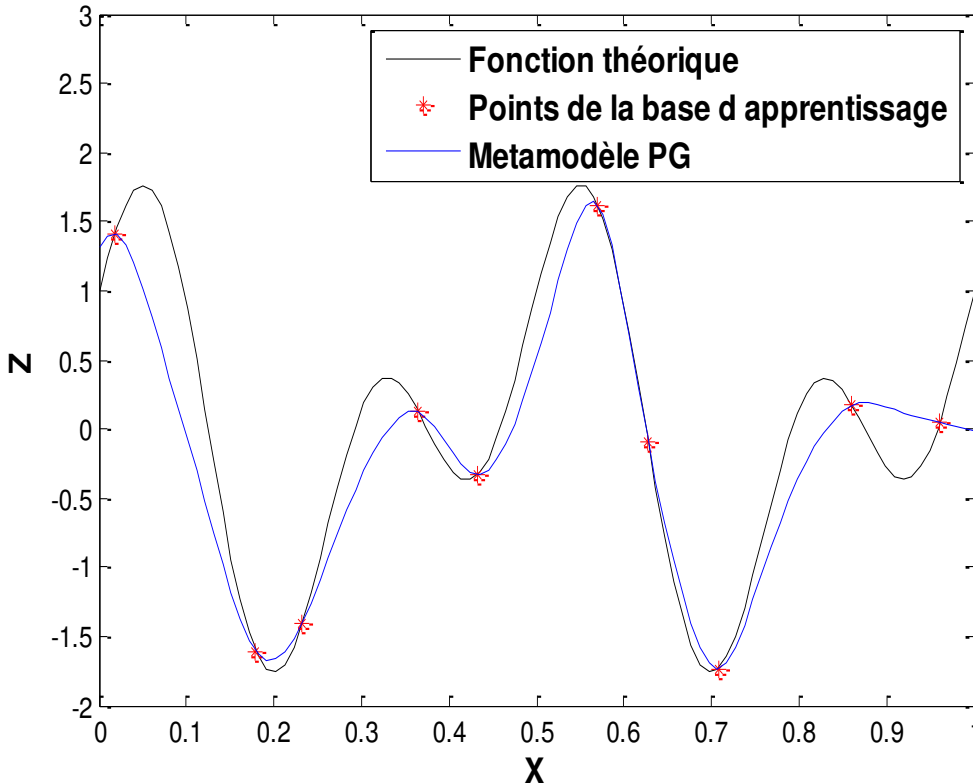
➤ Illustration sur des exemples analytiques

- **Fonction théorique** : $f(x) = \cos(4\pi X) + \sin(8\pi X)$
- **Base d'apprentissage** : tirage aléatoire de 10 points



➤ Illustration sur des exemples analytiques $f(x) = \cos(4\pi X) + \sin(8\pi X)$

- Ajustement d'un métamodèle PG : avec covariance Matern 5/2

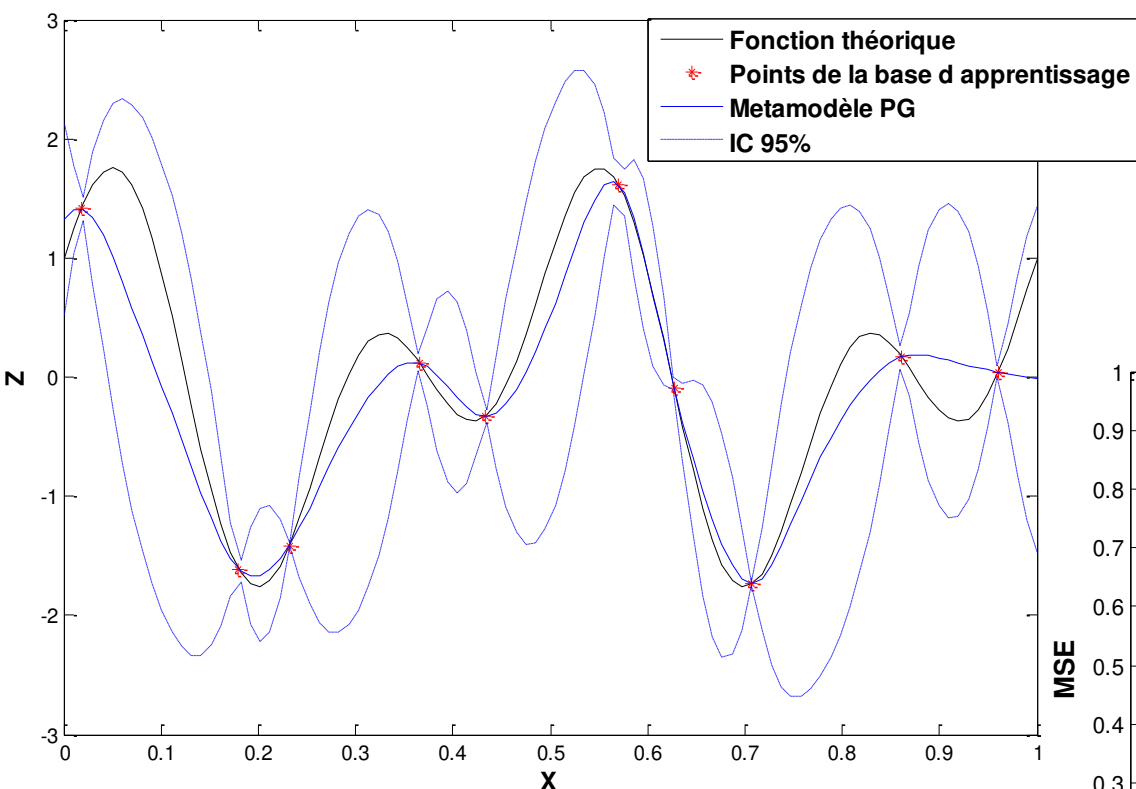


Calcul du Q^2 sur une base de test de 100 points :

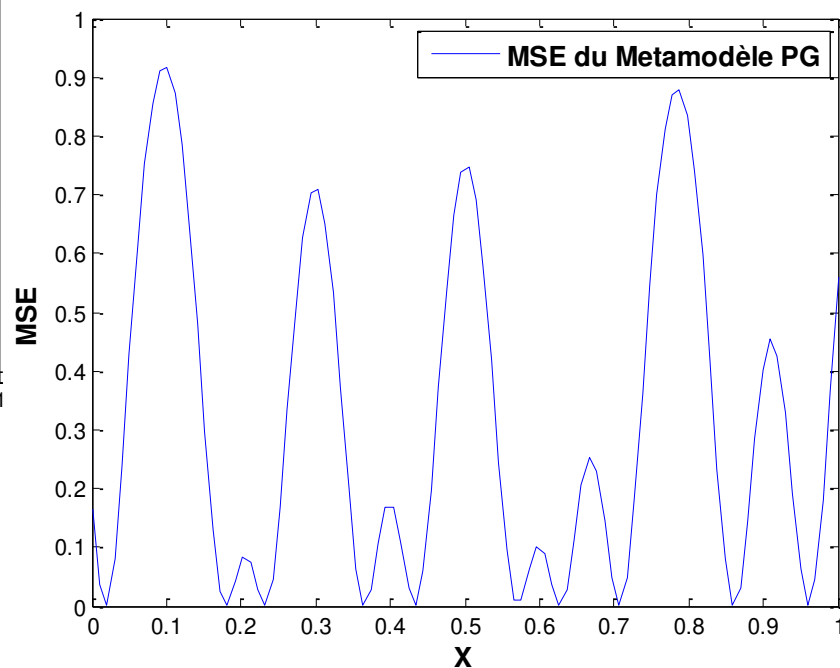
$$Q^2 = 0.85$$

➤ Illustration sur des exemples analytiques $f(x) = \cos(4\pi x) + \sin(8\pi x)$

▪ Utilisation du MSE du métamodèle PG

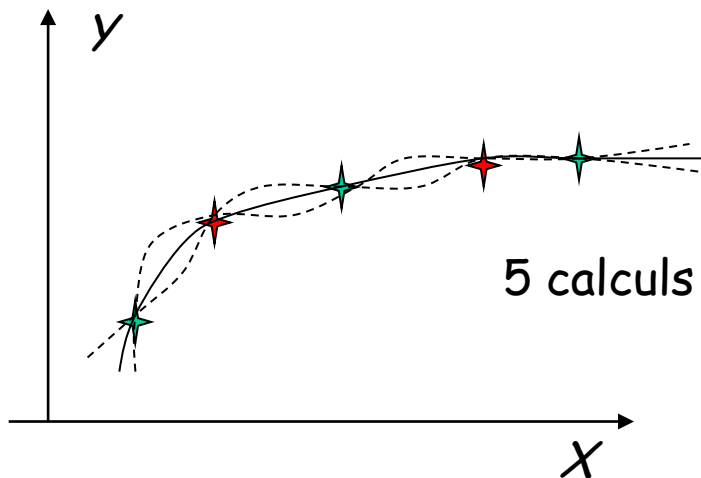
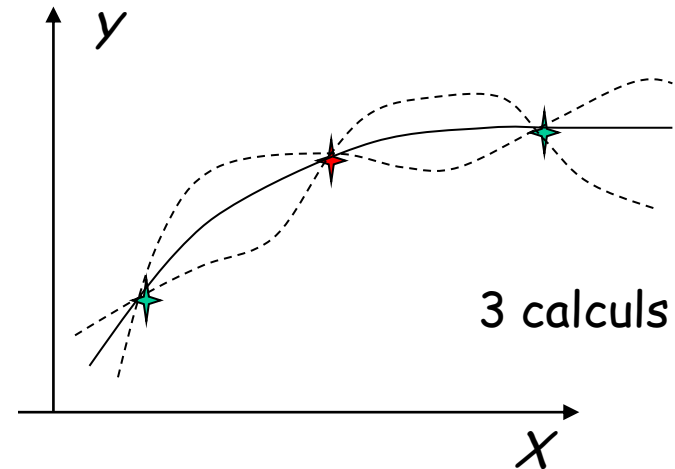
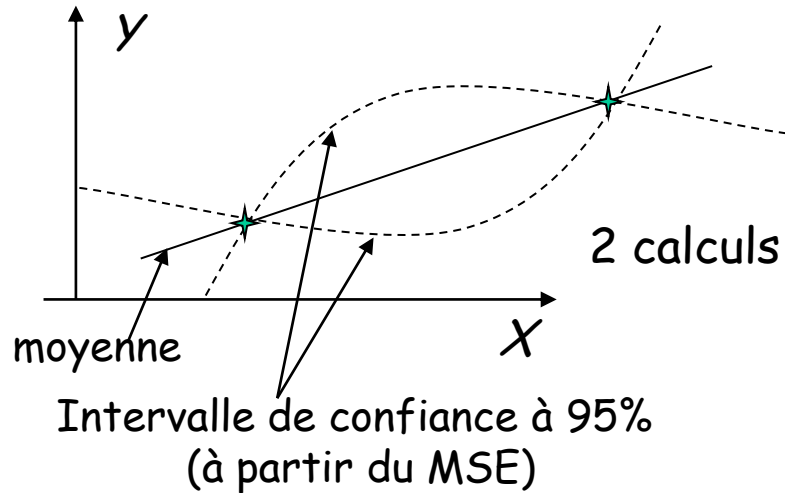


Construction d'intervalles de confiance pour les prédictions



➤ Illustration sur des exemples analytiques

▪ Utilisation du MSE du métamodèle PG



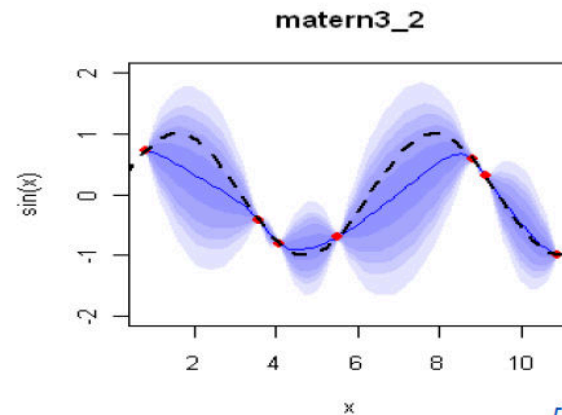
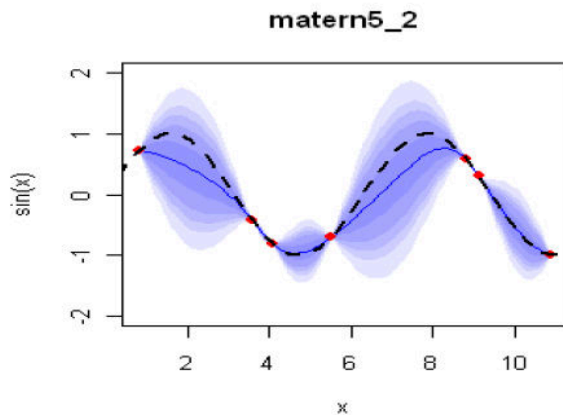
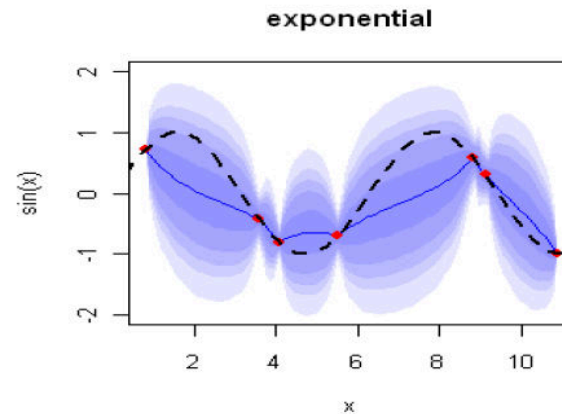
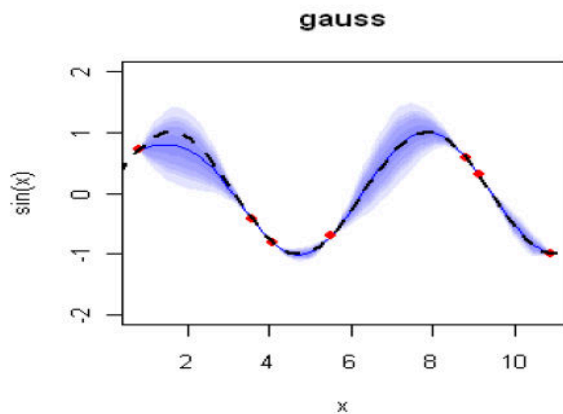
Conclusion : à partir d'un certain nb de points, le métamodèle devient précis

Problème : fléau de la dimension p (dans l'optimisation des hyperparamètres)

➤ Illustration sur des exemples analytiques

- Fonction théorique : $Y = \sin(X)$
- Base d'apprentissage : 7 points tirés aléatoirement

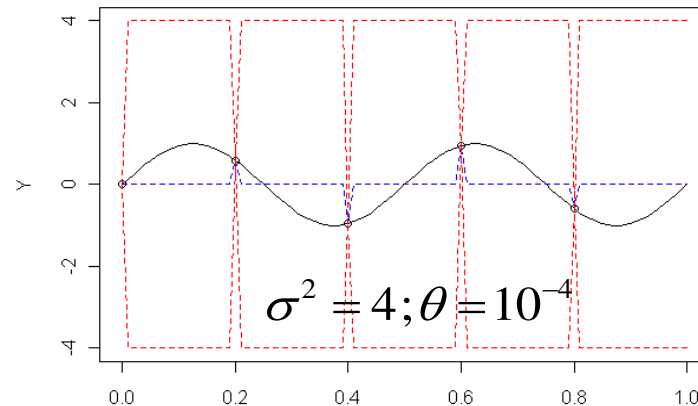
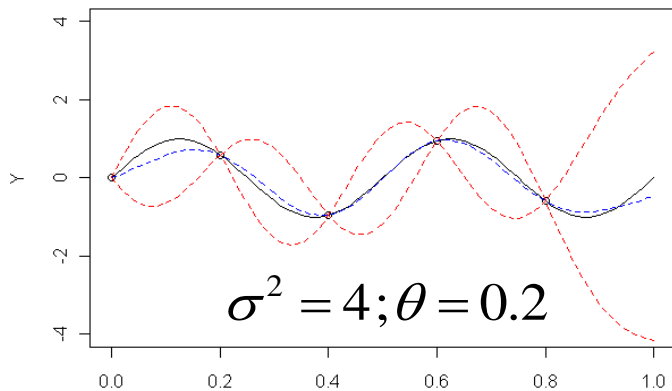
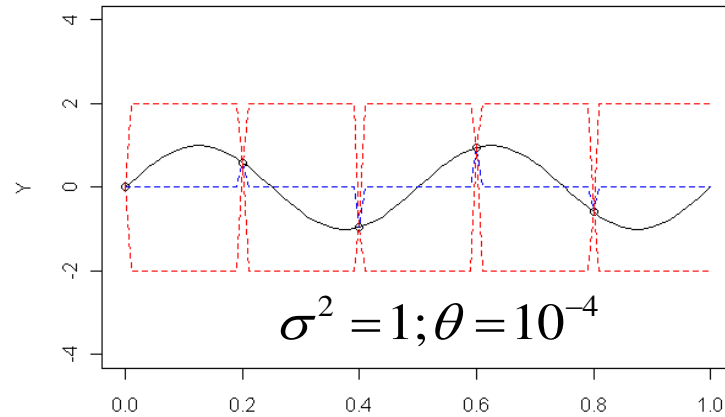
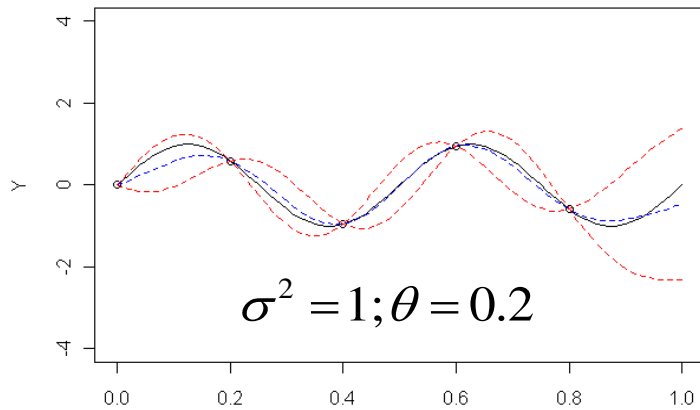
⇒ Impact du choix de la fonction de covariance



➤ Illustration sur des exemples analytiques

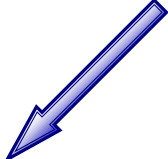
- **Fonction théorique** : $Y = \sin(4\pi X)$
- **Base d'apprentissage** : 5 points équirépartis

⇒ **Impact des hyperparamètres**

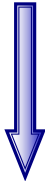


➤ Avantages du métamodèle PG

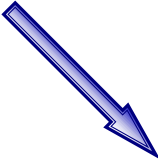
- **Interpolateur exact** des points de la base d'apprentissage (relaxation possible si ajout d'un effet de pépité)
- **Possibilité de « jouer » sur la forme de la fonction moyenne**
=> moyenne du processus devient une fonction de x (krigeage universel)
=> prise en compte d'avis d'expert
- **Possibilité d'estimer les paramètres par maximum de vraisemblance**
=> extension + facile en grande dimension
- **Formulation analytique : exploitation de la loi du prédicteur**
=> Simulation de la loi gaussienne conditionnelle
- **Rapidité d'évaluation du prédicteur**



Analyse de
sensibilité



Propagation
d'incertitude

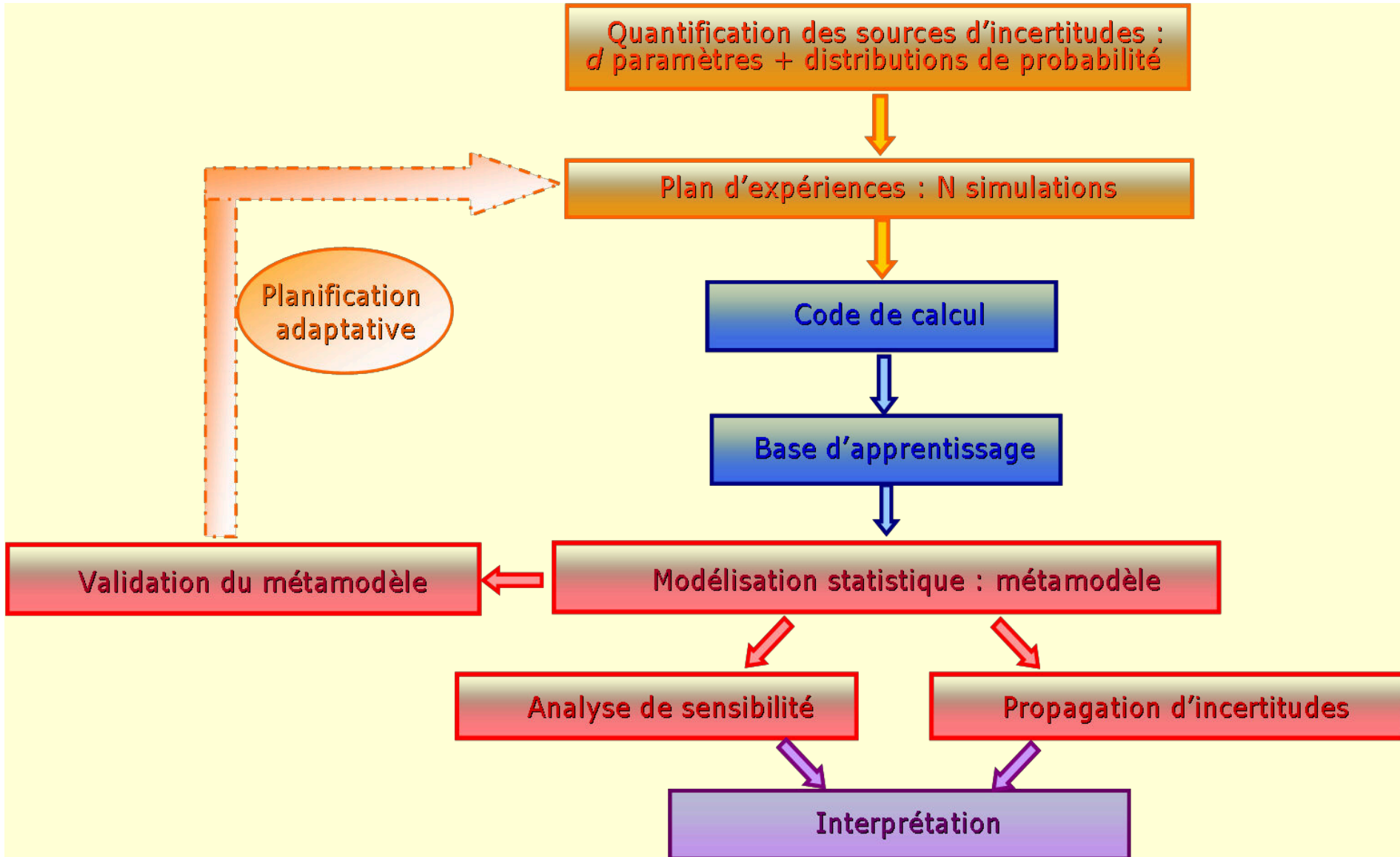


Calibration

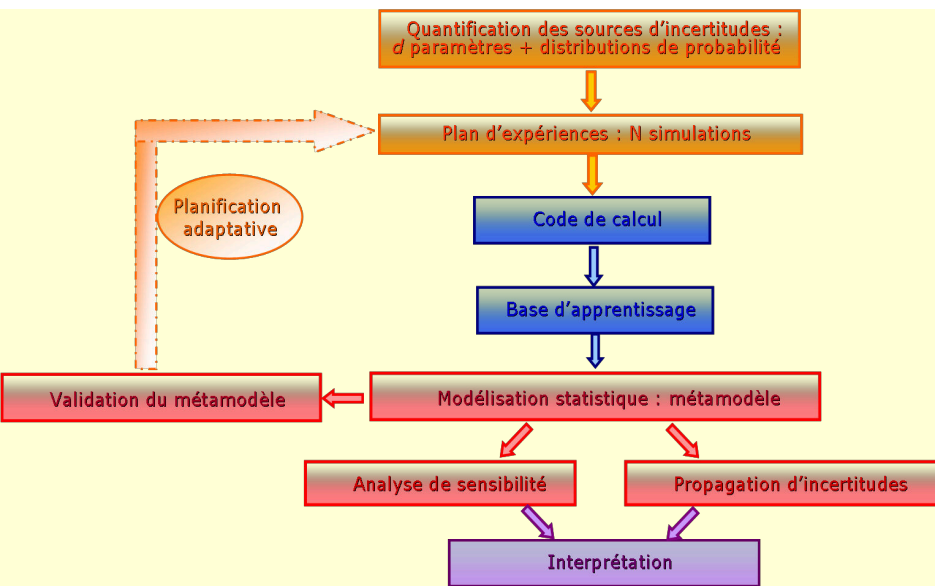
➤ Inconvénients du métamodèle PG

- **Choix des fonctions moyenne et de covariance**
- **Estimation des paramètres de régression et de covariance**
=> Estimation en grande dimension
- **Pas adapté quand beaucoup de données**
inversion de la matrice de covariance coûteuse => optimisation de la vraisemblance difficile
- **Prédicteur à tendance à lisser**
=> pas adapté à l'estimation de valeurs extrêmes, de seuil...

➤ Utilisation du métamodèle PG dans une démarche « incertitudes »



➤ Utilisation du métamodèle PG dans une démarche « incertitudes »



1. Détermination du domaine de variation des variables d'entrée influentes (et éventuellement de leur loi de proba)
2. Choix d'un type de métamodèle
3. Choix d'un type de plan d'expériences numériques
4. Évaluation du code pour le plan d'expériences
5. Construction du métamodèle à partir des expériences simulées
6. Validation du métamodèle
7. Exploitation du métamodèle

Les étapes 2 et 3 sont guidées par le problème traité (*analyse d'incertitudes, calcul de sensibilité, outil de prédiction, évaluation d'événements rares, optimisation, ...*)

A l'issue de l'étape 6, on peut revenir à l'étape 3 => **plan adaptatif**

Nota Bene : l'un des intérêts importants du métamodèle est de donner la possibilité d'étudier l'impact du choix de la distribution des entrées

➤ Choix du plan d'expériences pour la base d'apprentissage

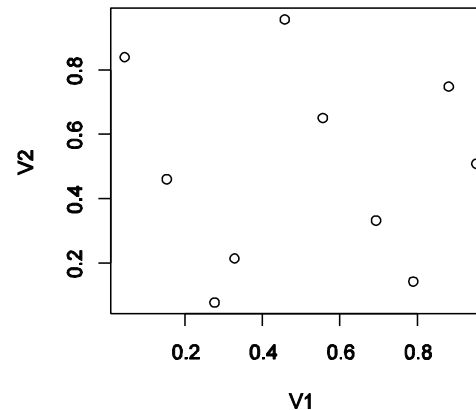
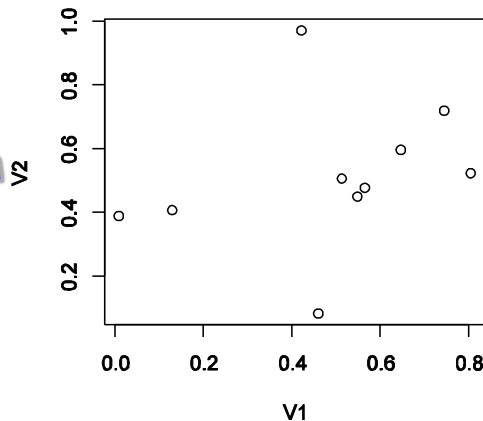
▪ Métamodèle PG adapté :

- au cas des modèles numériques $\Rightarrow$ expériences « déterministes »
- quand quelques dizaines d'entrée et quelques centaines de simulations possibles $\Rightarrow$ points d'expériences dispersés $\Rightarrow$ « *sparsity* »

▪ Plan d'expériences recommandés dans ces cas :

- **Space-filling design** : couverture optimale du domaine de variation des paramètres incertains

Simple
Random
Sample



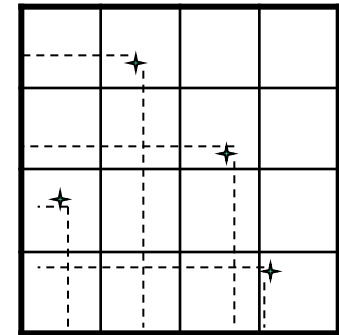
Space
Filling
Design

➤ Choix du plan d'expériences pour la base d'apprentissage

- **Space-filling design de type hypercubes latins (LHS)** : projections uniformes sur les marginales

⇒ compromis entre "étalement" des points et répartition uniforme

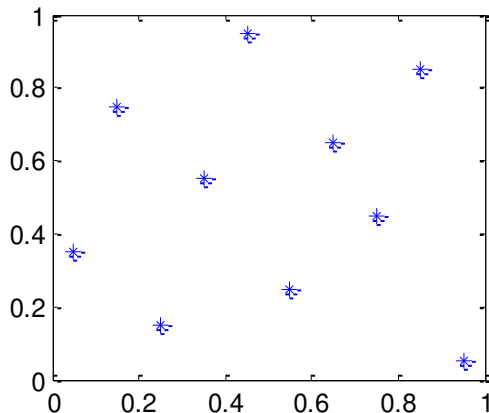
On divise chaque dimension en N intervalles
et on prend un point dans chaque intervalle



plan LHS de 4 points en dimension 2

- **Plans LHS optimisés** : optimisation d'un critère d'occupation de l'espace

ex : critère maximin, minimax, discrédance, ...



*plan LHS maximin de 10
points en dimension 2*

➤ Planification adaptative avec le métamodèle PG

■ Principe :

- Plan d'expériences initial
- Construction d'un premier métamodèle
- Ajout de nouveaux points de manière séquentielle en utilisant le métamodèle ⇒ **planification adaptative**

■ Mise en œuvre avec le métamodèle PG:

■ Utilisation d'un critère basé sur le MSE du métamodèle PG

- Objectif : choisir comme nouveau point celui qui améliore le plus la qualité du métamodèle
 - Exemple : point où le MSE du métamodèle PG est maximal => point où l'erreur de prédiction est potentiellement la + grande
 - Autres critères utilisant le MSE du PG : point qui minimise l'intégrale du MSE sur le domaine, ajout par blocs de points, ...

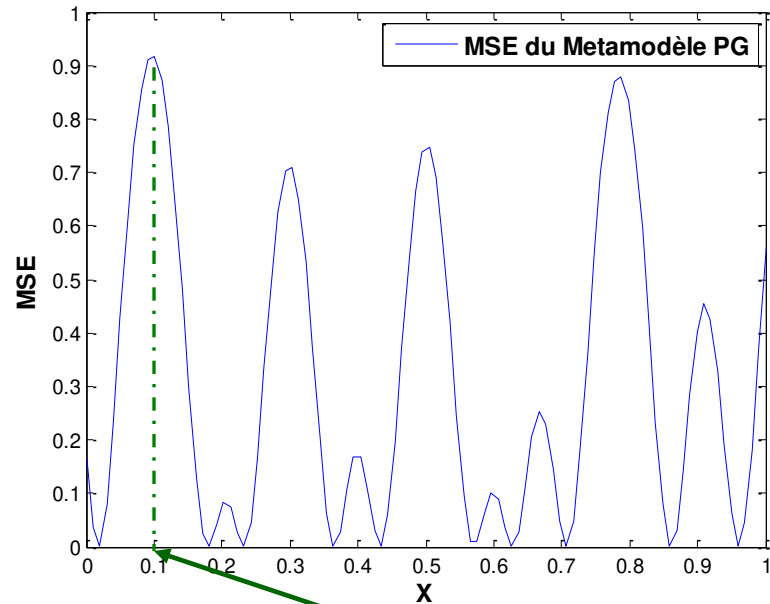
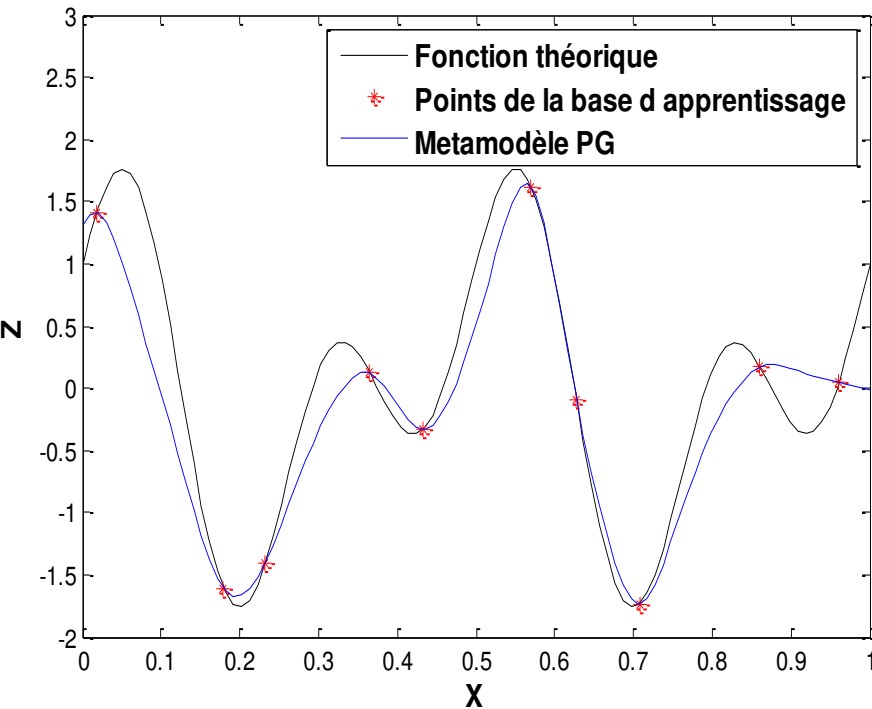
■ Autres critères liés à l'objectif de l'étude : recherche de minima,...

➤ Planification adaptative avec le métamodèle PG

■ Illustration : Utilisation d'un critère basé sur le MSE

Nouveau point = celui où le MSE du métamodèle PG est maximal

$$x_{new} = \underset{x}{\operatorname{Argmax}} \operatorname{MSE}[\hat{Z}(x)] = \underset{x}{\operatorname{Argmax}} \left(\sigma^2 - k(x)^T \Sigma_s k(x) \right)$$



Nouveau point : $x_{new} = 0.0972$

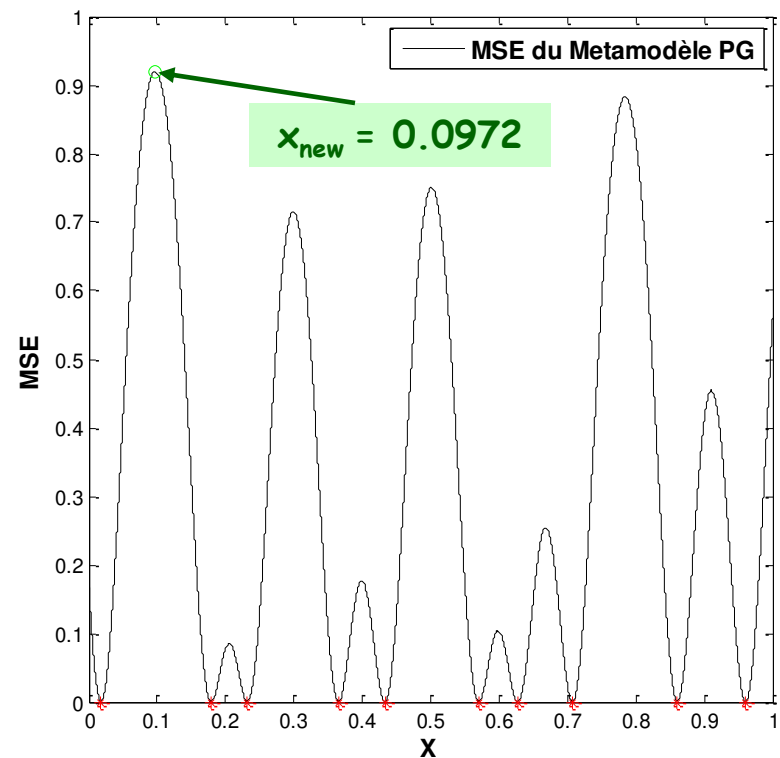
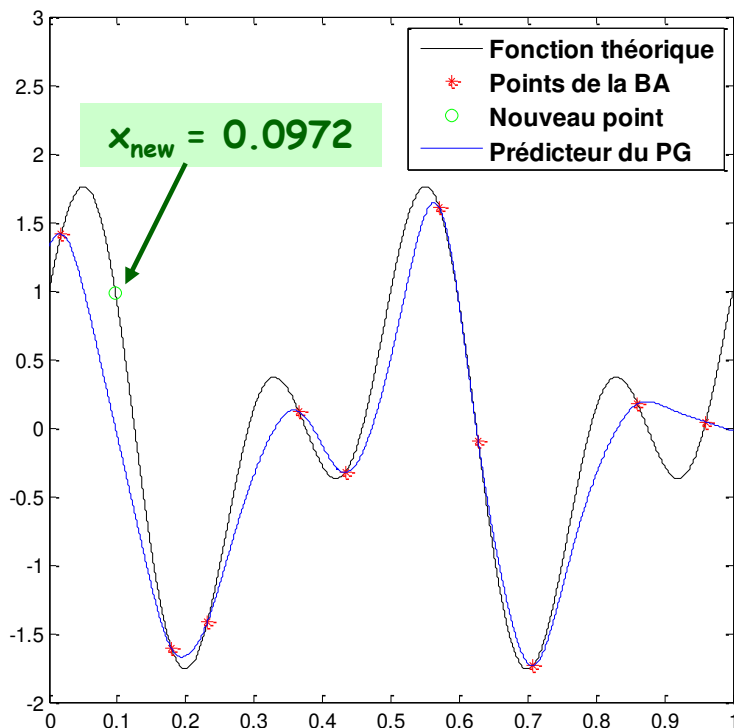
➤ Planification adaptative avec le métamodèle PG

■ Illustration : Utilisation d'un critère basé sur le MSE

Nouveau point = celui où le MSE du métamodèle PG est maximal

$$x_{new} = \underset{x}{\operatorname{Argmax}} \operatorname{MSE}[\hat{Z}(x)] = \underset{x}{\operatorname{Argmax}} (\sigma^2 - k(x)^T \Sigma_s k(x))$$

Etape 1 : Recherche de x_{new} (maximisation du MSE => pbe d'optimisation)

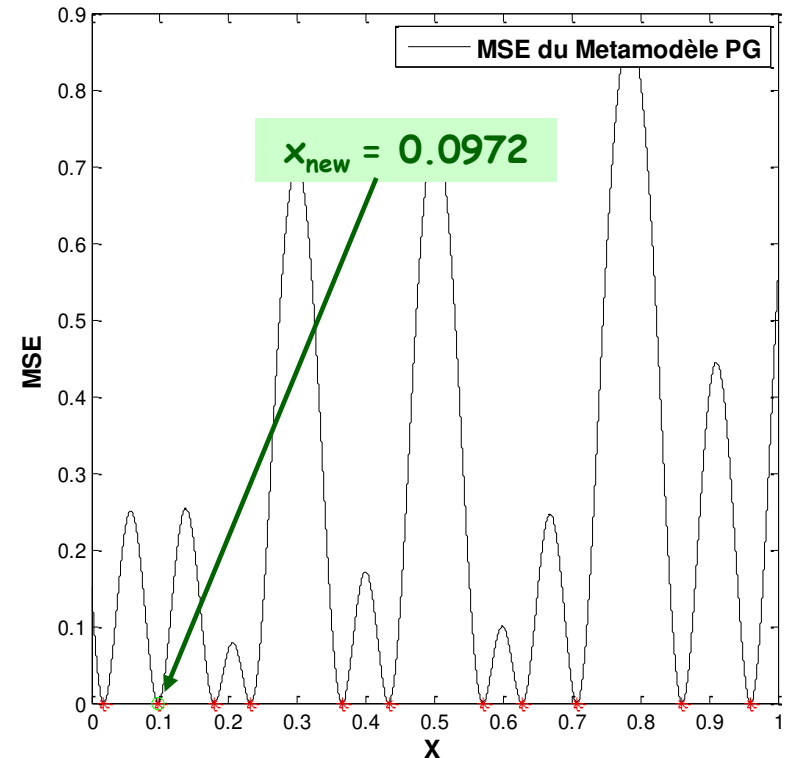
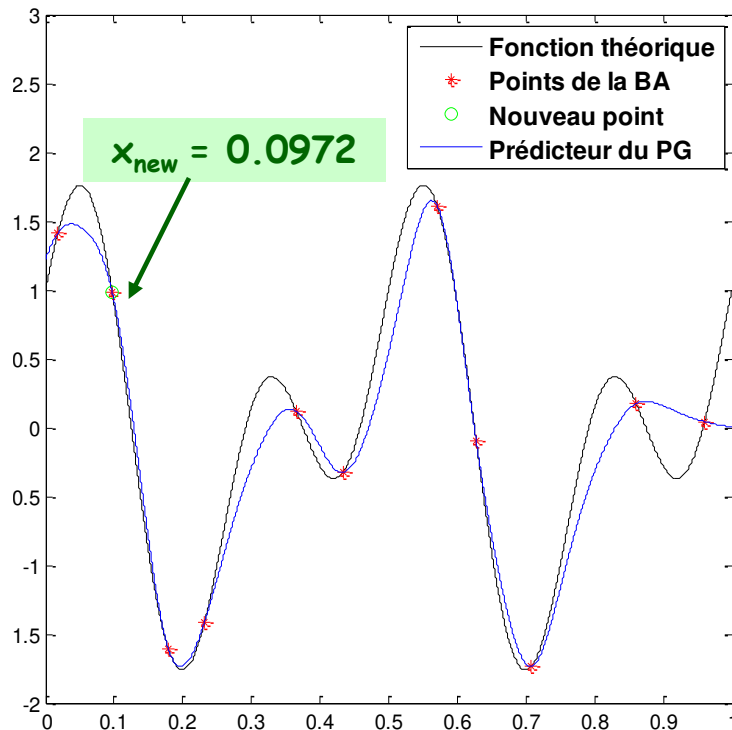


➤ Planification adaptative avec le métamodèle PG

■ Illustration : Utilisation d'un critère basé sur le MSE

Etape 2 : Simulation de $x_{\text{new}} \Rightarrow z_{\text{new}}$

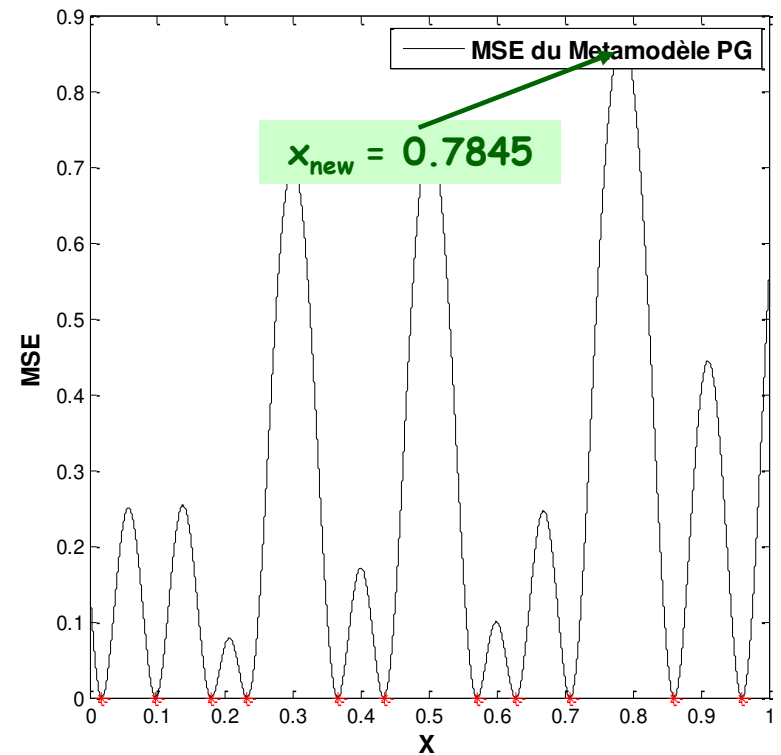
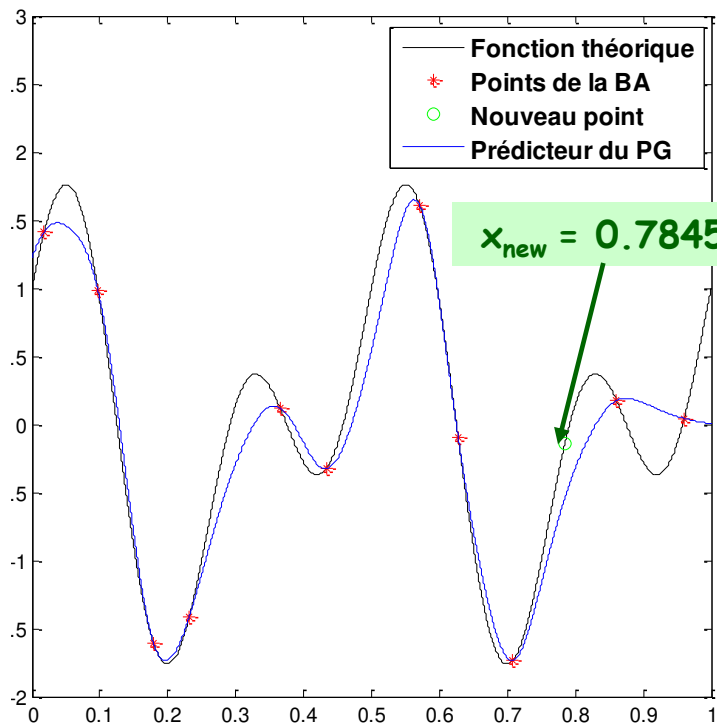
Ajout de $(x_{\text{new}}, z_{\text{new}})$ à la BA et mise à jour du métamodèle PG



➤ Planification adaptative avec le métamodèle PG

- **Illustration : Utilisation d'un critère basé sur le MSE**

Etape 1 : Recherche de x_{new} (maximisation du MSE => pbe d'optimisation)

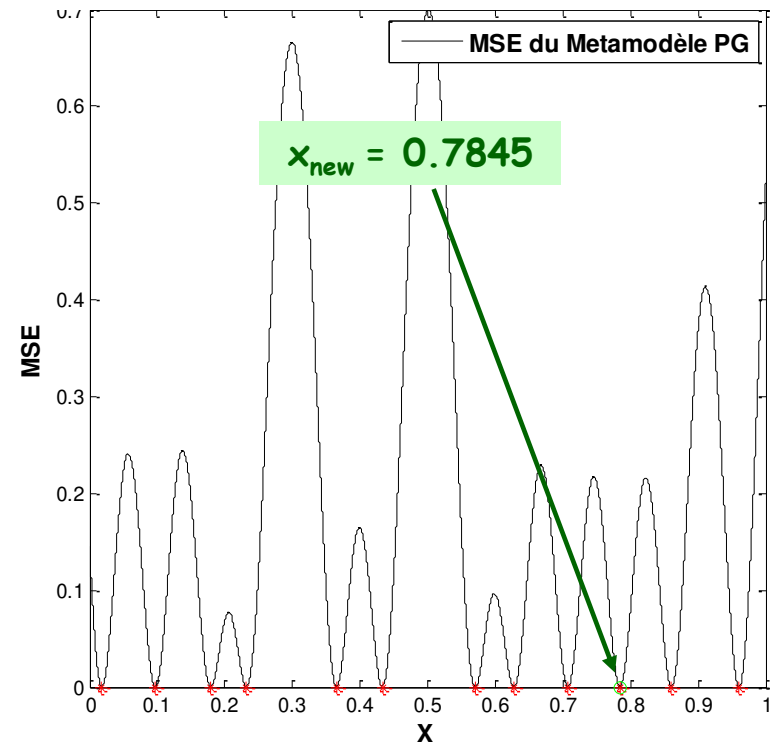
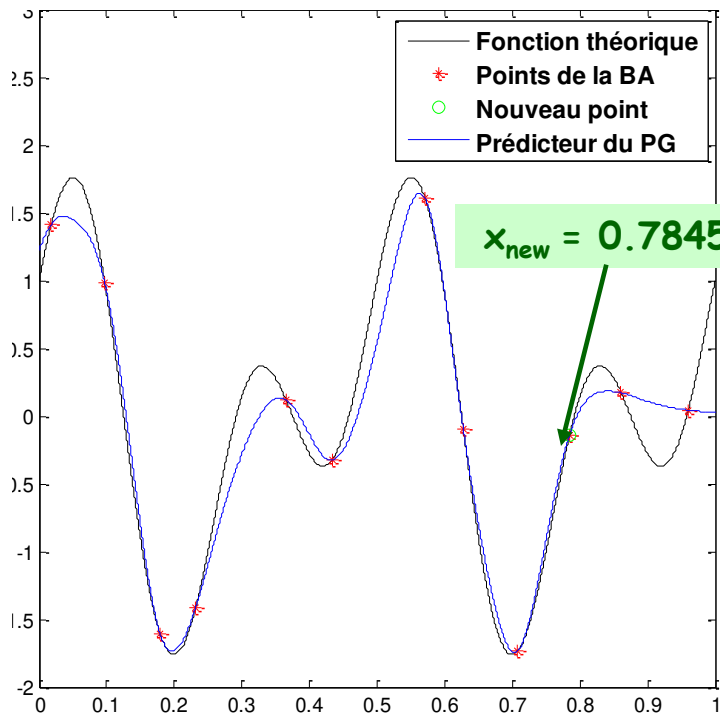


➤ Planification adaptative avec le métamodèle PG

- **Illustration : Utilisation d'un critère basé sur le MSE**

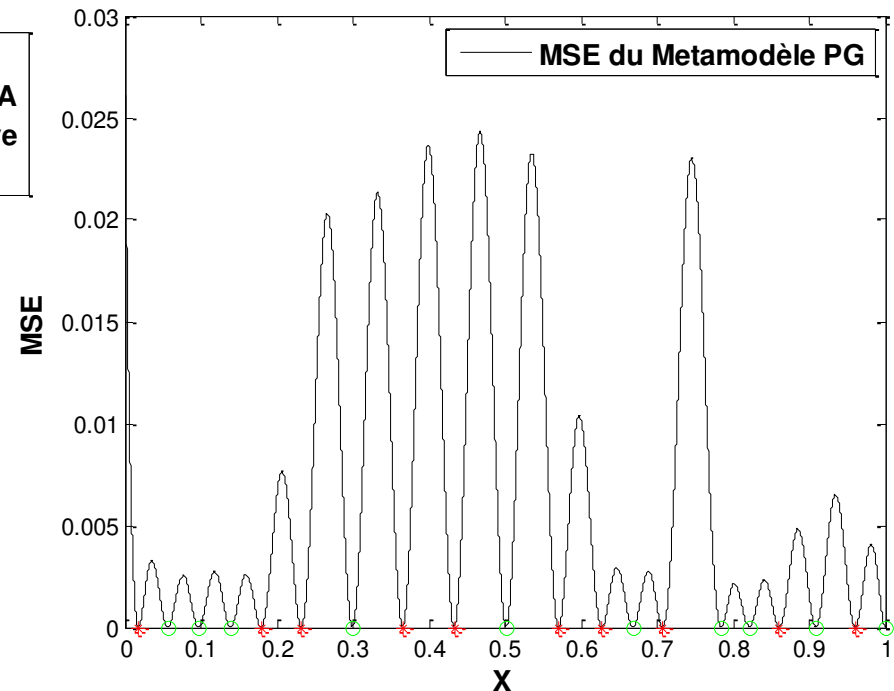
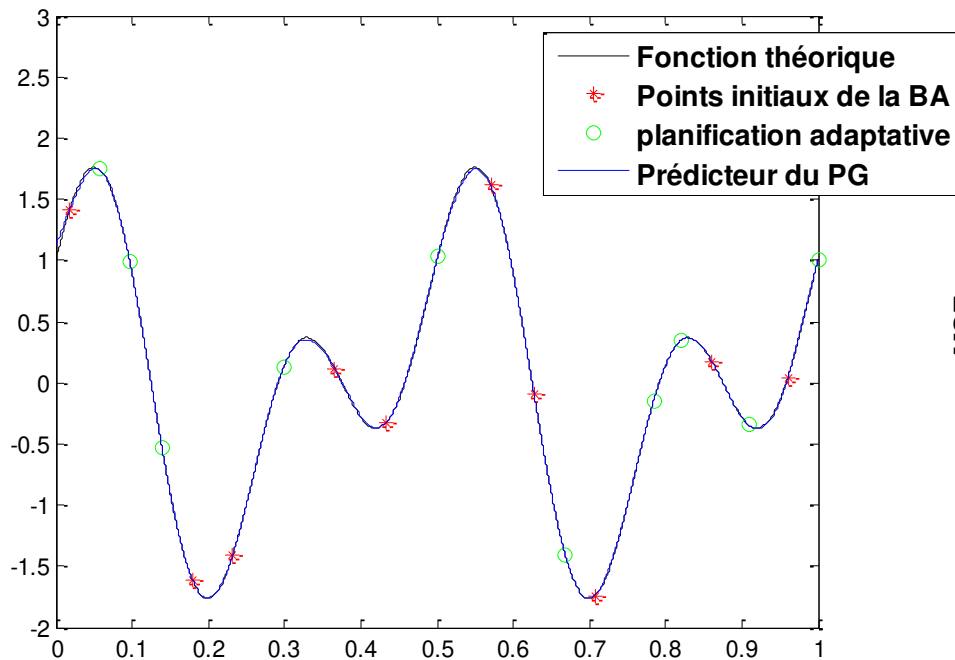
Etape 2 : Simulation de $x_{new} \Rightarrow z_{new}$

Ajout de (x_{new}, z_{new}) à la BA et mise à jour du métamodèle PG



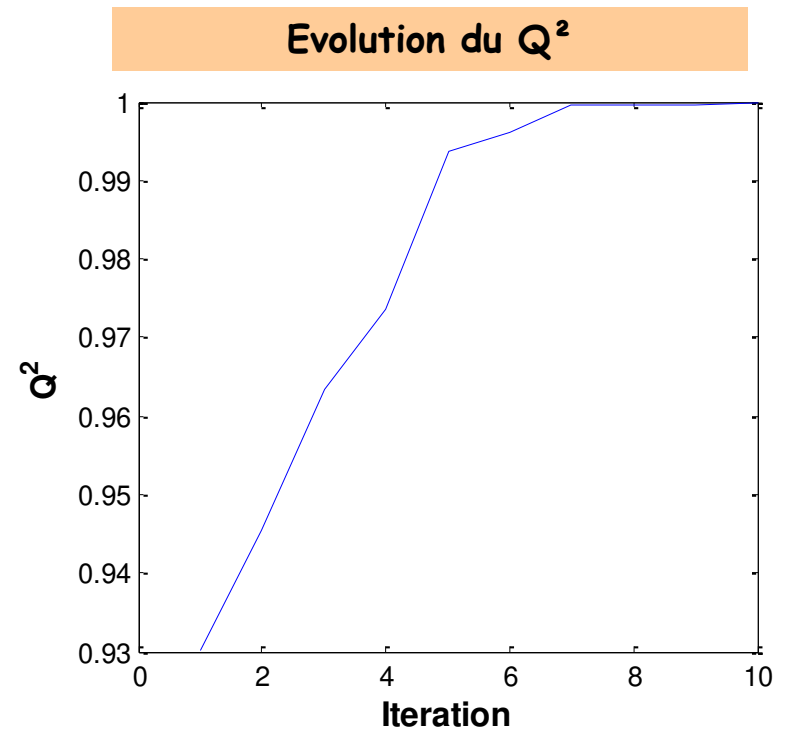
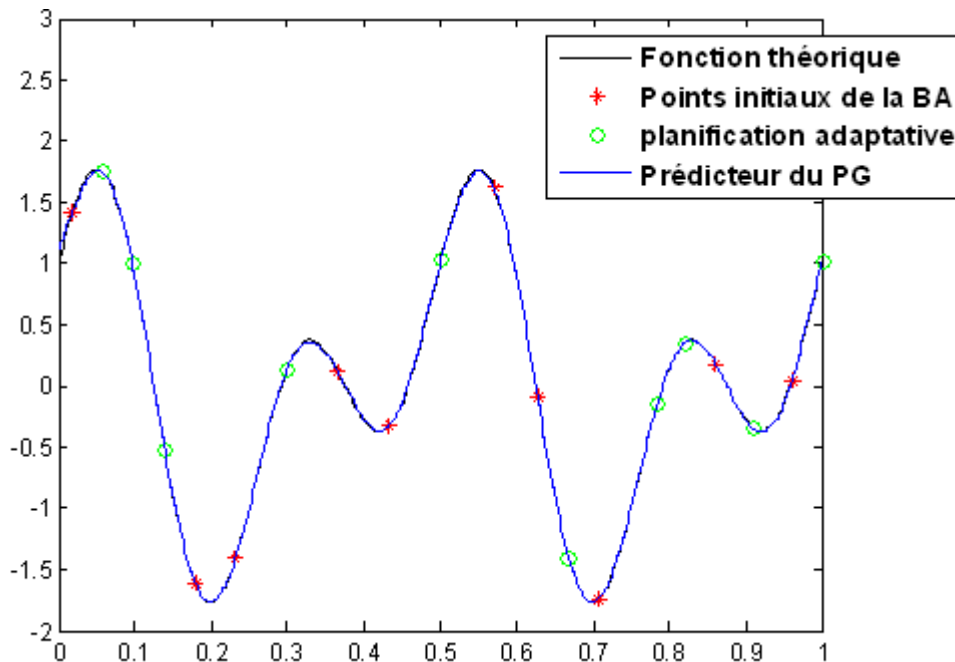
➤ Planification adaptative avec le métamodèle PG

- **Illustration : Utilisation d'un critère basé sur le MSE**
Ajout de 10 points par planification adaptative



➤ Planification adaptative avec le métamodèle PG

- **Illustration : Utilisation d'un critère basé sur le MSE**
Ajout de 10 points par planification adaptative



Rappels statistiques et théorie

- Variable, vecteur et processus gaussien

Définition du métamodèle processus gaussien (PG)

- Du krigage à la régression par processus gaussien
- Formulation analytique
- Analogie avec le krigage
- Simulation conditionnelles

Construction du métamodèle PG

- Choix de la fonction moyenne
- Choix de la fonction de covariance
- Estimation des paramètres

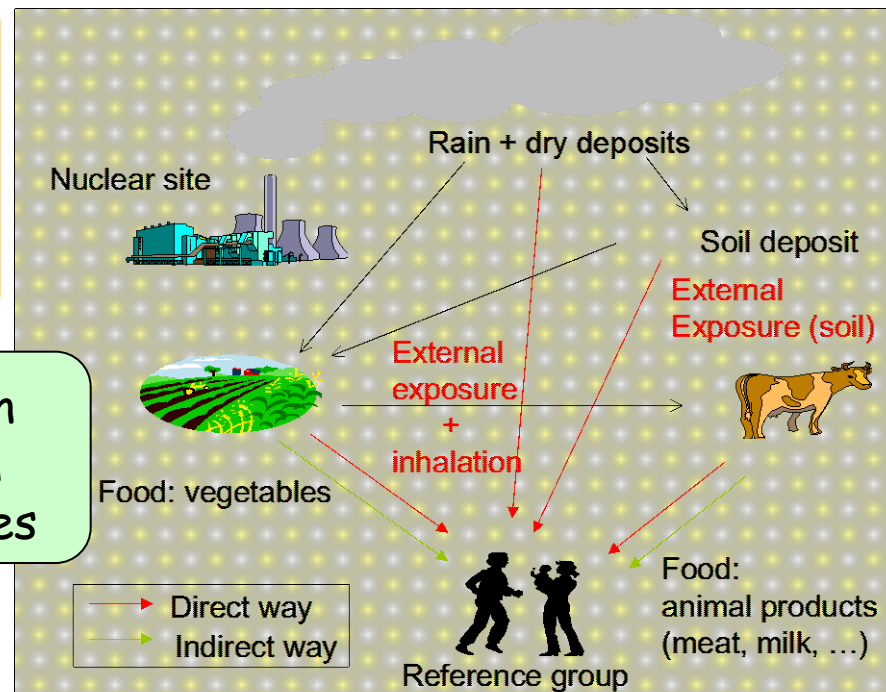
Utilisation du métamodèle PG

- Validation du métamodèle PG
- Illustration sur des exemples analytiques
- Plans d'expériences et planification adaptative

Application sur des cas industriels

Evaluation des dépôts causés par le rejet accidentel ou chronique de polluants nucléaires (rejets atmosphériques)

Modélisation de la dispersion atmosphérique basée sur un modèle à bouffées gaussiennes



Modèle gaussien de panache

Panache à modéliser

Vent u

Modèle gaussien à bouffées

Code de dispersion atmosphérique
CERES-MITHRA (C-M)
⇒ Calcul d'impact des installations nucléaires

Modélisation numérique

Entrées C-M

Chronique de rejets

- Points de rejet
- Hauteur des points de rejet
- Quantité de radionucléides rejetés
- Vitesse de dépôt des radionucléides
- Durée des rejets

Chronique météorologiques

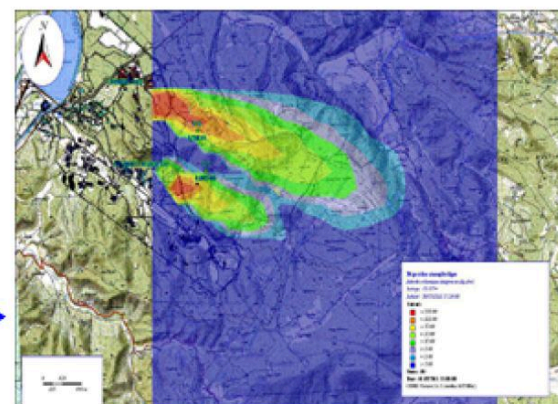
- Direction du vent
- Vitesse du vent
- Diffusion atmosphérique
- Température
- Taux d'humidité
- Durée d'observation
(≥ Durée du rejet)

Simulateur
C-M

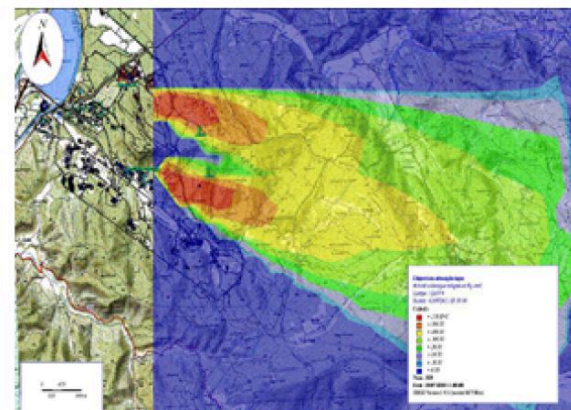
*Temps CPU de quelques
secondes à quelques
dizaines de minutes*

Sorties C-M

Cartes 2-D de la concentration de l'activité volumique intégrée (Bqsm^{-3})
à différents instants



Concentration de Cs137 après 20 minutes



Concentration de Cs137 après 2 heures

Objectifs

- Quantifier les incertitudes sur les paramètres d'entrée du modèle
- Evaluer l'influence de chaque paramètre d'entrée sur la sortie
=> analyse de sensibilité du code C-M
- Identifier l'impact sur les prédictions du modèle C-M

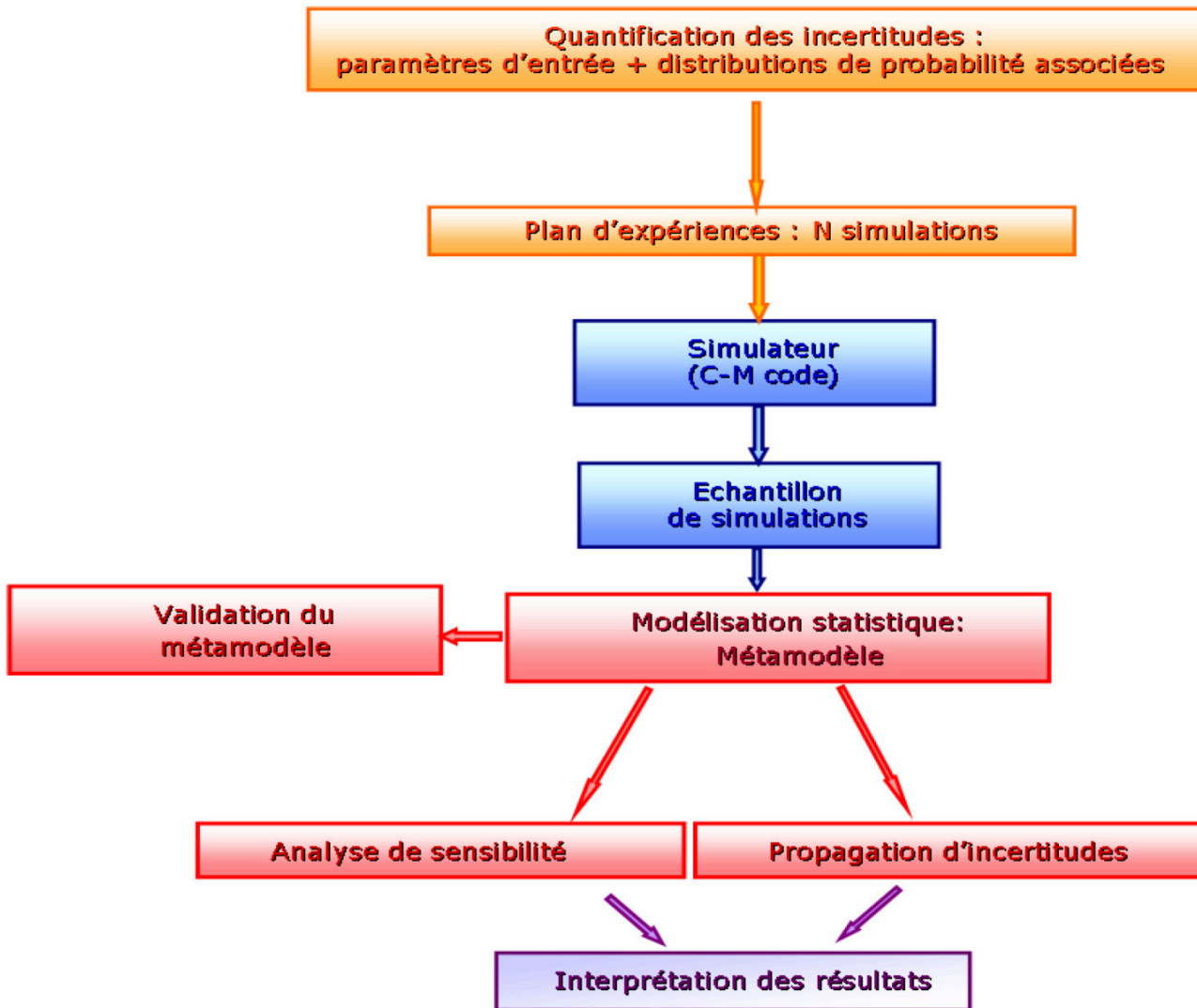
Problème

- Grand temps de calcul pour chaque simulation
- Méthodes classiques d'analyse de sensibilité globale $\Rightarrow$ des milliers de simulations
- Sorties spatiales

Solution

- Construction d'une méthodologie globale combinant différentes techniques statistiques avancées
 - ❖ Approximation du code C-M par un métamodèle
 - ❖ Analyse de sensibilité

Méthodologie



- **Quantification des incertitudes**
Création d'un simulateur réaliste des chroniques météorologiques
- **Plan d'expériences**
Hypercube latin ayant des propriétés de recouvrement optimales
- **Métamodèle**
Décomposition spatiale
Orthogonale aux Valeurs propres (POD)
+
métamodélisation des coefficients principaux de la POD par Processus Gaussiens (PG)
- **Analyse de sensibilité**
Estimation d'indices basés sur la décomposition de la variance de la sortie
↓
indices de Sobol

Quantification des sources d'incertitudes

Lois des variables incertaines du scénario

Paramètres		Intervalle de variation	Loi de probabilité associée	Lois déterminées par avis d'expert du LMTE
Hauteur des points de rejets (en m)	l_1	$[7.5; 22.5]$	Uniforme	
	l_2	$[22.5; 67.5]$	Uniforme	
Vitesse de dépôt (en $m.s^{-1}$)		$[5.10^{-4} ; 5.10^{-2}]$	Log-Uniforme	Lois modélisées après étude d'une base de données CEA
Quantité rejetée (en Bq)		$[10^8 ; 10^{10}]$	Log-Uniforme	
Direction de vent (en degré azimut)		$[249 ; 333]$	Processus temporel	
Vitesse de vent (en $m.s^{-1}$)		$[0 ; 12.5]$	Processus temporel	

Plan d'expériences et simulations

Nombre total de variables : 4 (rejet) + 12 (météo)

Choix d'un plan d'expériences numériques :

Latin hypercube Sampling maximin (LHS)

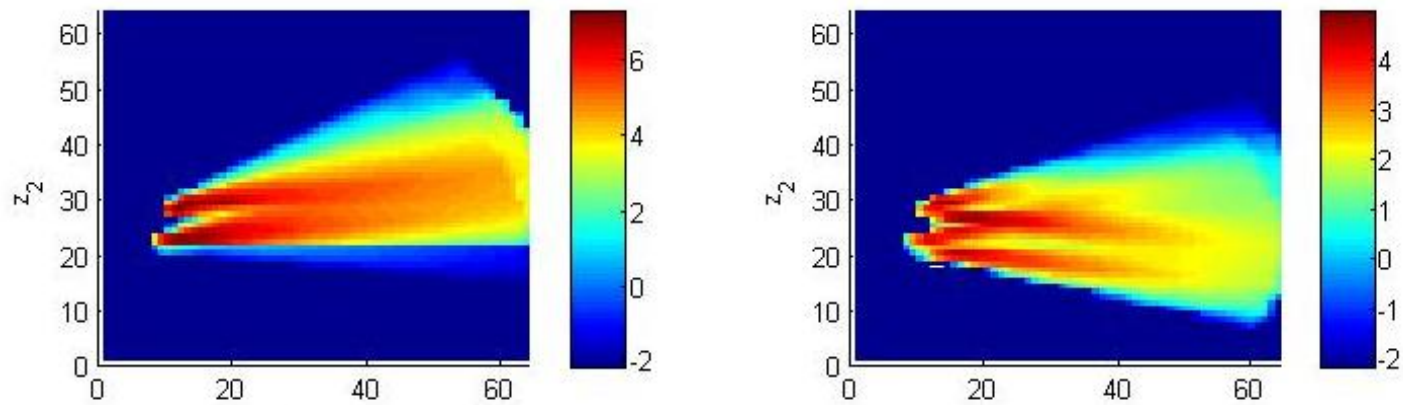
Taille du plan d'expériences :

$n=150$ simulations

bon compromis entre le nombre d'expériences et l'exploration du domaine de variation des paramètres incertains (une meilleure répartition spatiale des points)

Réalisation des simulations avec Ceres-Mithra :

Constitution de la base d'apprentissage (BA)



Exemples de cartes d'activité surfacique intégrée du Cs137 (log)

Modélisation statistique : métamodèle fonctionnel

➤ Etape 1 : ACP fonctionnelle (ou POD pour Proper Orthogonal Decomposition)

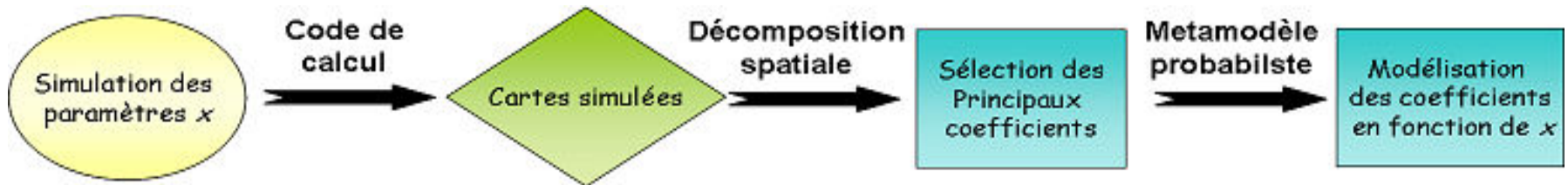
➡ Sélection des principaux coefficients de la décomposition

$$Y(X, z) \simeq \bar{Y}(z) + \sum_{l=1}^k a_l(X) \phi_l(z) \quad \text{avec} \quad \frac{|\langle Y, \phi_l \rangle|^2}{\|\phi_l\|^2} = \max_{\psi \in L^2(\Omega)} \frac{|\langle Y, \psi \rangle|^2}{\|\psi\|^2} \quad \text{où} \quad \|\phi_l\|^2 = 1$$

et $a_l(X) = \int_{\Omega} Y(X, z) \phi_l(z) dz \quad \forall l \in \{1, \dots, k\}.$

➤ Etape 2 : Modélisation des principaux coefficients en fonction de X :

➡ Métamodèle PG



➤ Prédiction :

➡ Analyse de sensibilité, propagation d'incertitude, ...



Modélisation statistique : métamodèle fonctionnel

► Validation du métamodèle

⇒ Estimation du coefficient de prédictivité Q^2

- sur base de test
- par validation croisée

$$Q^2(z) = 1 - \frac{\sum_{j=1}^{n_{test}} [\hat{Y}(x^{(j)}, z) - Y(x^{(j)}, z)]^2}{\sum_{j=1}^{n_{test}} \left[Y(x^{(j)}, z) - \frac{1}{n} \sum_{i=1}^n Y(x^{(i)}, z) \right]^2}$$

avec $\{x^{(j)}\}_{j=1 \dots n_{test}}$ les points de la base de test, z les coordonnées spatiales, Y les valeurs observées ("vraies" valeurs) et $\hat{Y}$ la prédiction du métamodèle.

Interprétation :

- pourcentage de variance expliquée par le métamodèle
- plus sa valeur est proche de 1, meilleure est l'adéquation du métamodèle au « vrai » modèle.

Rq : Dans le cas d'une sortie spatiale, on obtient une valeur de Q^2 en chaque point de l'espace.

Q^2 proche de 1 $\Rightarrow$ bon métamodèle

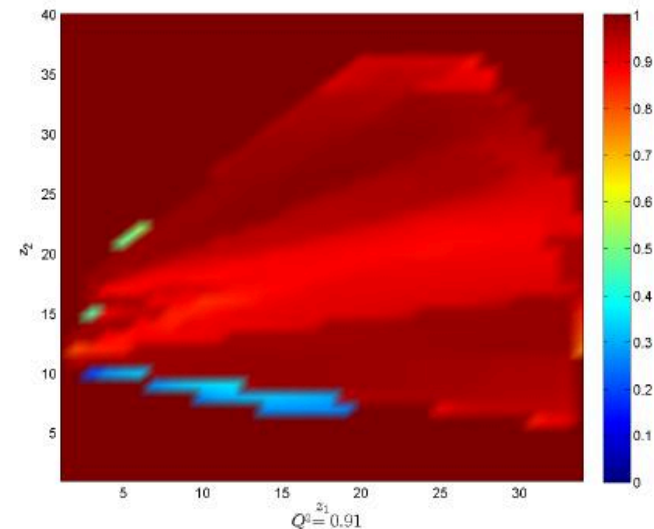


FIGURE: Q^2 du métamodèle pour $t=40\text{min}$.

Analyse de sensibilité : indices de Sobol

« Étude de l'influence des entrées d'un simulateur sur sa sortie. » Saltelli [1999]

Définition des indices de Sobol [2001]

Indices du premier ordre S_i : Mesure l'influence de l'entrée X_i indépendamment des autres.

Indices du second ordre S_{ij} : Mesure l'influence de l'interaction entre l'entrée X_i et l'entrée X_j indépendamment de leurs effets seuls.

Indices totaux S_{T_i} : Mesure l'influence de l'entrée X_i et de l'ensemble de ses interactions.

Estimation des indices de Sobol

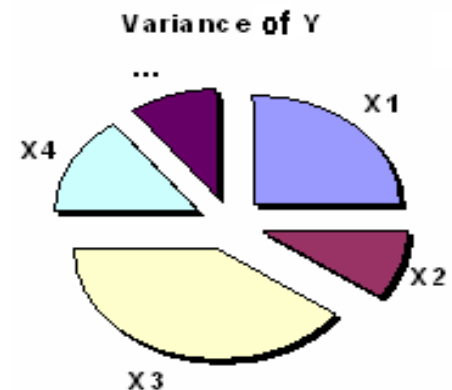
Utilisation de la méthode RBD-FAST (Mara [2009])



Nécessite des milliers de simulations

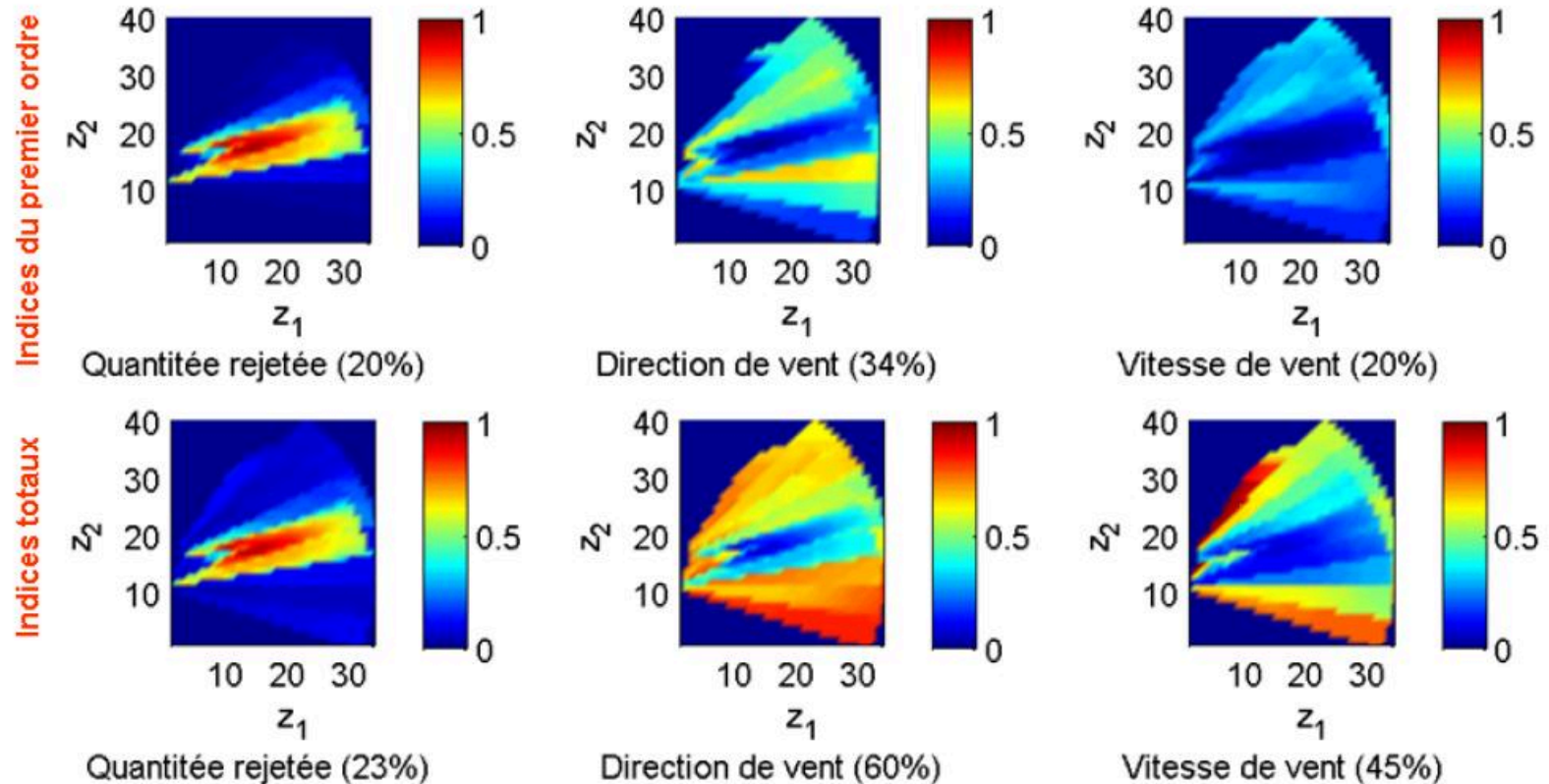


Utilisation du métamodèle

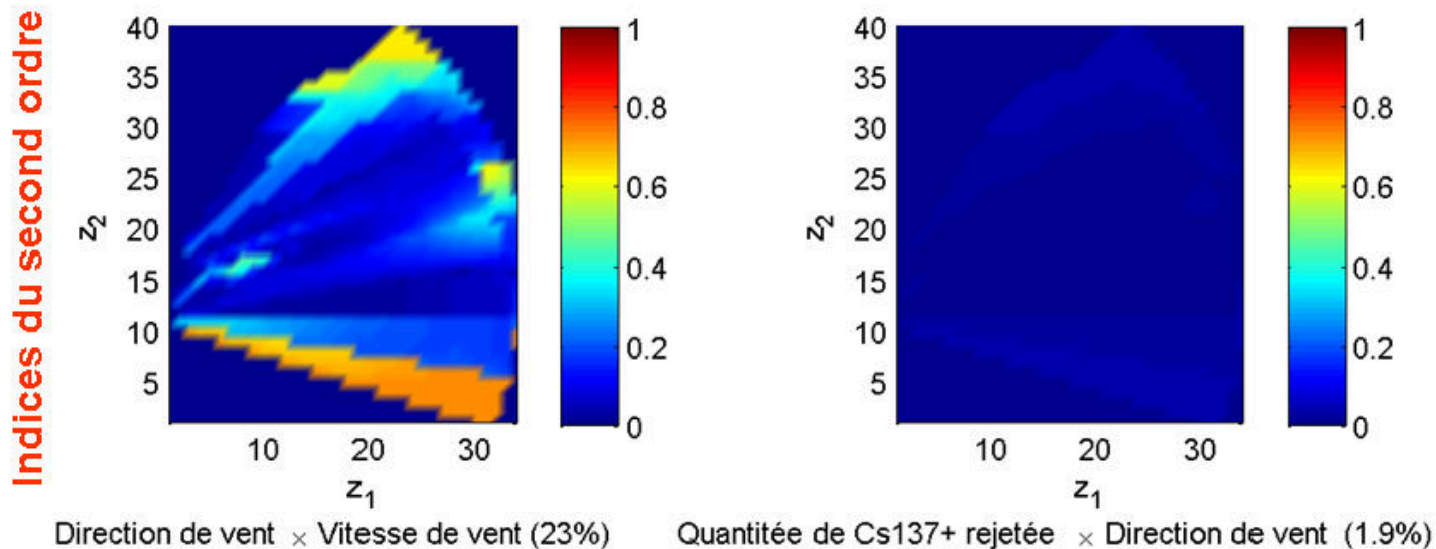


Poids de chaque entrée dans la variance de la sortie

Analyse de sensibilité : indices du 1er ordre (à $T_0+40\text{min}$)



- **Quantité rejetée** : Influe sur le coeur du panache ;
- **Direction et vitesse de vent** : Influe sur les bords du panache ;
- Somme des indices du premier ordre = 74% $\Rightarrow$ 26% d'interactions ;
- Interaction probable entre la direction et la vitesse de vent.

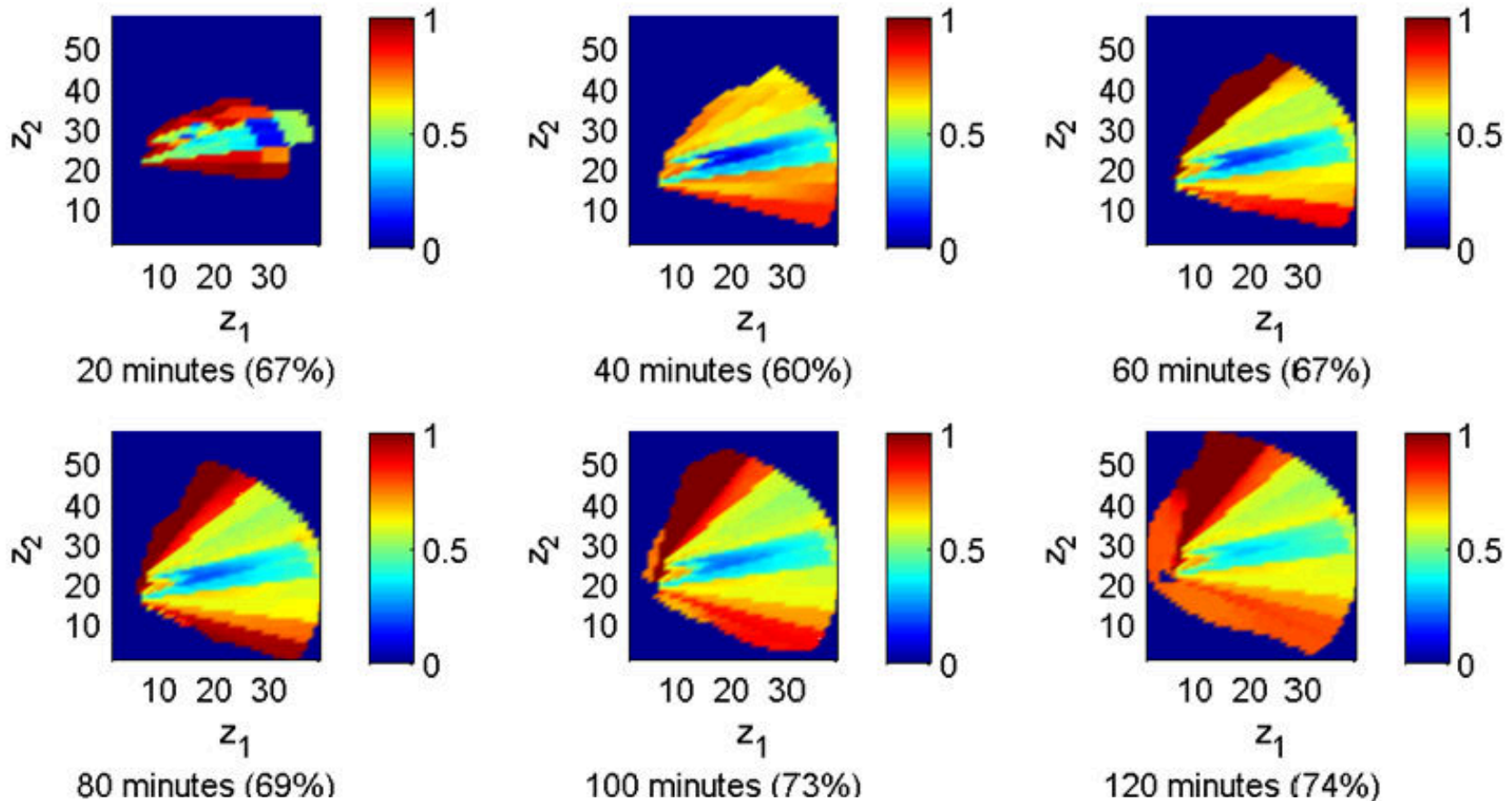
Analyse de sensibilité : interactions (à $T_0+40\text{min}$)

L'interaction *direction de vent* $\times$ *vitesse de vent* influe sur les bords du panache.

- | | | |
|---|--------|------------------------------|
| - direction de vent | (34%); | } 99% de variance expliquée! |
| - <i>direction de vent</i> $\times$ <i>vitesse de vent</i> | (23%); | |
| - quantité rejetée de Cs137+ | (20%); | |
| - vitesse de vent | (20%); | |
| - <i>direction de vent</i> $\times$ <i>quantité de Cs137+ rejetée</i> | (2%). | |

Analyse de sensibilité : évolution temporelle

Indices totaux



Forte croissance des indices totaux



Augmentation de la part d'influence des interactions de DV

➤ Processus gaussiens

- O'Hagan, A. (2006). Bayesian analysis of computer code outputs: a tutorial. *Reliability Engineering and System Safety* **91**, 1290-1300.
- Rasmussen, C. E., and Williams, C. K. I. (2006). *Gaussian Processes for Machine Learning*. Cambridge, MA: MIT Press.
- Sacks J. Welch W.J. Mitchell T.J. and Wynn H.P. (1989) Design and analysis of computer experiments. *Statistical Science*, 4, 409-435.
- Santner, T. J., Williams, B. J. and Notz, W. I. (2003). *The Design and Analysis of Computer Experiments*. New York: Springer.

➤ Géostatistiques

- Chilès & Delfiner, *Geostatistics*, Wiley, 1999
- Marcotte, *Cours de l'Ecole Polytechnique de Montreal*.

➤ Applications industrielles

- Marrel et al., Global sensitivity analysis for models with spatially dependent outputs, *Environmetrics*, 2011
- Volkova et al., *Stoch. Environ. Res. Risk Assess.*, 2008
- Marrel & Perot. Development of a surrogate model and sensitivity analysis for an atmospheric dispersion computer code. *Proceedings of PSAM 11 & ESREL 2012 Conference*, Helsinki, Finland.

➤ Packages :

- R : DiceKriging <http://cran.r-project.org/web/packages/DiceKriging/index.html>
- Matlab : DACE www2.imm.dtu.dk/~hbn/dace
- Logiciel GEM <http://www.tonyohagan.co.uk/academic/GEM/index.html>

Comme il n'y a pas que le krigeage dans la vie... => Polynômes de chaos

Base des polynômes du chaos = famille de polynômes orthogonaux $\psi_k(\mathbf{x})$ associés à une mesure $\mu_{\mathbf{x}}(\mathbf{x})$

Ex. : densité gaussienne → pol. d'Hermite ; uniforme → pol. de Legendre

Le développement de chaque v.a. d'entrée sur sa base permet de représenter $Y = f(\mathbf{X})$ par tensorisation (chaos polynomial généralisé) :

$$Y(\mathbf{X}) = \sum_{k \in \mathbb{N}} \alpha_k \psi_k(\mathbf{X}) \quad \text{avec } E[\psi_k(X) \psi_l(X)] = \delta_{k,l}$$

Limité à un ordre K de développement convenable (précision / nb de termes)

Par exemple : $p = 2$ et $K = 3$ → 10 polynômes du chaos

Estimation de α_k par intégration numérique ou par moindres carrés

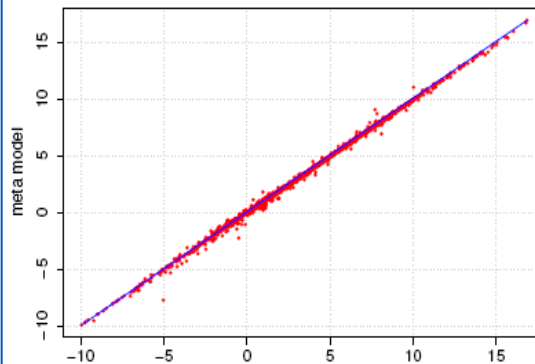
Plus : approche déterministe, taux de convergence connu, atteint un niveau arbitraire d'incertitudes, les coef. du chaos contiennent toute l'info. proba.

Minus : fléau de la dimension, adaptation des outils (en mode intrusif), problème de robustesse (non linéarités, discontinuités)

- $X = (X_1, X_2, X_3)$, $X_i \hookrightarrow \mathcal{U}(-\pi, \pi)$, independent
- $f(x_1, x_2, x_3) = \sin(x_1) + a\sin^2(x_2) + bx_3^4\sin(x_1)$
- $a = 7, b = 1/10$
- budget: 500 computations

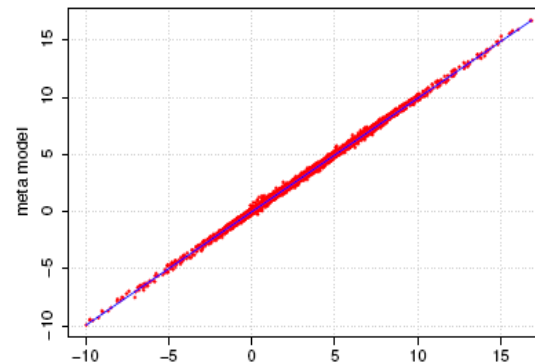
- Total degree of approx.: 9
- compute 220 coefficients

Model vs Meta-Model



- Total degree of approx.: 12
- keep 16 best coefficients

Model vs Meta-Model



Traitement des incertitudes en simulation numérique

**TP 3 : Modélisation d'expériences numériques
=> métamodèle processus gaussien**

**Compte-rendu d'évaluation :
=> CR sur les exercices 1.b), 3 et 4.**

EXERCICE 1 : SIMULATION DE TRAJECTOIRES

Objectif : Simuler des trajectoires d'un processus gaussien de moyenne et covariance données

1.a) Simulation de trajectoires 1D

Énoncé

Soit un processus gaussien centré stationnaire de covariance gaussienne

$C(t, t') = \sigma^2 e^{-\left(\frac{h}{\theta}\right)^2}$. Simuler des réalisations de $Z(t)$ de $N = 200$ points sur $[0 ; 1]$. Prendre $\sigma^2 = 0,5$, $\theta = 0,05$ puis $\theta = 0,2$.

Méthode

- simulation de N points vecteur gaussien iid => `rnorm`
- calcul matrice des distance => `dist`
- calcul matrice de covariance
- transfo en vecteur de loi => `chol`

Objectif : Simuler des trajectoires d'un processus gaussien de moyenne et covariance données

1.a) Simulation de trajectoires 1D

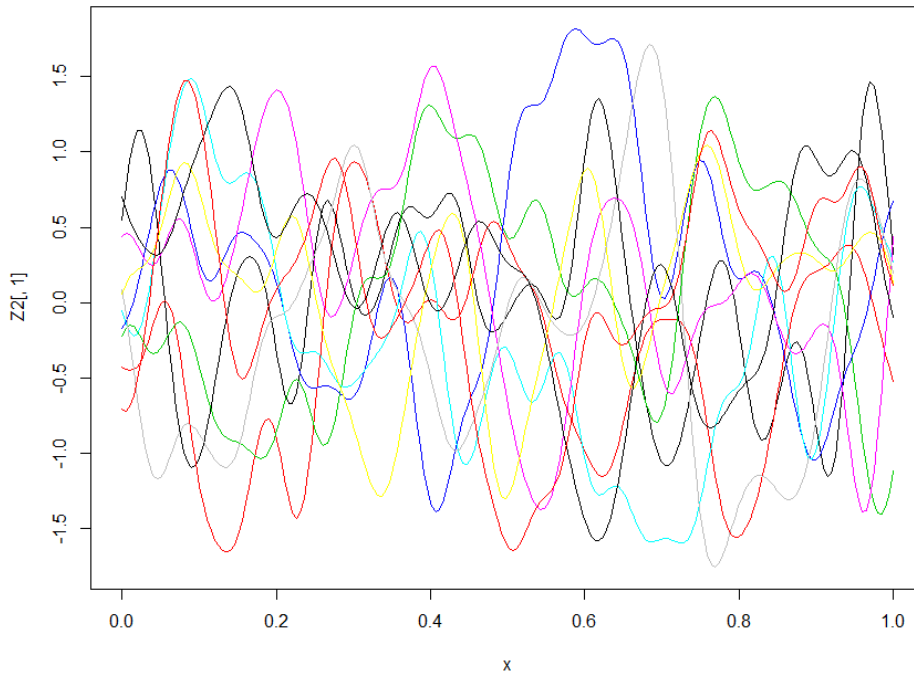
Solution

```
N <- 200 #nb de points de la trajectoire
x <- seq(0,1,length.out=N)
dist_all <- as.matrix(dist(x))
n <- 10 # simulation de n trajectoires
Z1 <- matrix(rnorm(N*n),N,n)
f0 <- rep(0,N)
sig2 <- 0.5 ; theta <- 0.05 ; lambda <- 10^(-10)
cov <- sig2*(exp(-(dist_all/theta)^2) + lambda * diag(1,N))
Z2 <- matrix(0,N,n)
for (i in 1:n){
  Z2[,i] <- f0+t(chol(cov))%*%Z1[,i]
}
plot(x,Z2[,1], main=paste('Simulation trajectoires PG de
  covariance gaussienne avec theta =',theta), type='l',
  ylim=c(min(Z2),max(Z2)))
for (i in 2:n)
  lines(x,Z2[,i],col=i)
```

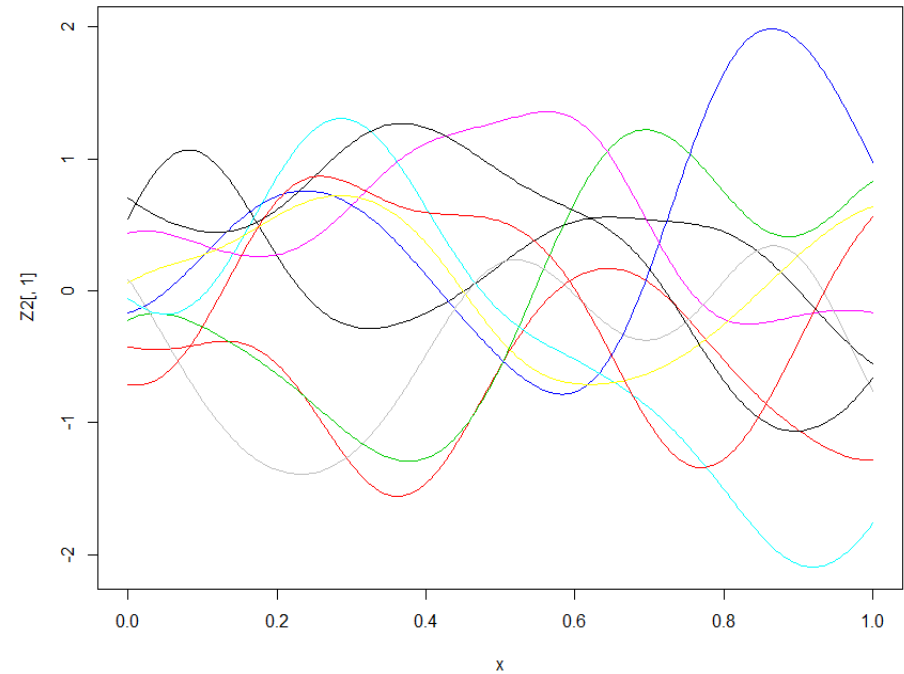
Objectif : Simuler des trajectoires d'un processus gaussien de moyenne et covariance données

1.a) Simulation de trajectoires 1D

Simulation trajectoires PG de covariance gaussienne avec $\theta = 0.05$



Simulation trajectoires PG de covariance gaussienne avec $\theta = 0.2$



Objectif : Simuler des trajectoires d'un processus gaussien de moyenne et covariance données

1.a) Simulation de trajectoires 1D

Énoncé

Soit un processus exponentiel centré stationnaire de covariance gaussienne $C(t, t') = \sigma^2 e^{-h/\theta}$. Simuler des réalisations de $Z(t)$ de $N = 200$ points sur $[0 ; 1]$. Prendre $\sigma^2 = 0,5$, $\theta = 0,05$ puis $\theta = 0,2$.

Méthode

- simulation de N points vecteur gaussien iid $\Rightarrow$ `rnorm`
- calcul matrice des distance $\Rightarrow$ `dist`
- calcul matrice de covariance
- transfo en vecteur de loi $\Rightarrow$ `chol`

Objectif : Simuler des trajectoires d'un processus gaussien de moyenne et covariance données

1.a) Simulation de trajectoires 1D

Solution

```
sig2 <- 0.5
theta <- 0.05
cov <- sig2*exp(-(dist_all/theta))
Z2 <- Z1
for (i in 1:n)
  Z2[,i] <- f0+sqrt(cov)%*%Z1[,i]

plot(x,Z2[,1], main=paste('Simulation trajectoires PG de
  covariance exponentielle avec theta =',theta), type='l',
  ylim=c(min(Z2),max(Z2)))
for (i in 2:n)
  lines(x,Z2[,i] ,col=i)
```

Objectif : Simuler des trajectoires d'un processus gaussien de moyenne et covariance données

1.a) Simulation de trajectoires 1D

Autre solution (avec DiceKriging) :

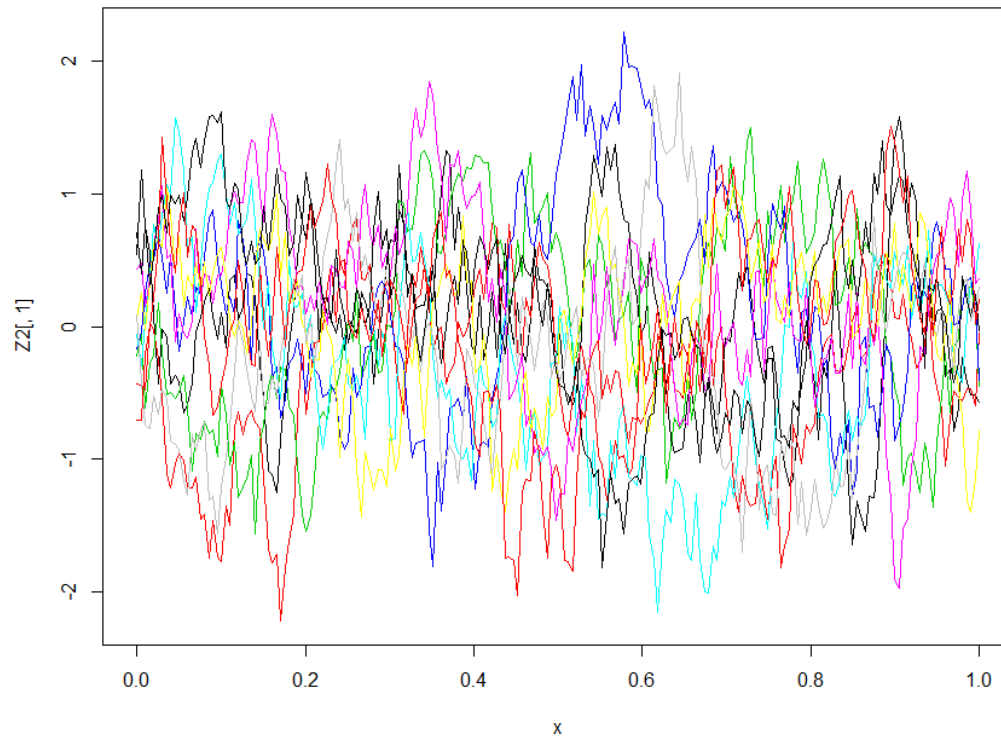
```
type <- "exp"
coef <- c(theta = 0.05)
sigma <- 0.5
model <- km(design=x, response=rep(0,N),coef.trend = 0,
            covtype=type, coef.cov=coef, coef.var=sigma)
Z2 <- simulate(model, nsim=10, newdata=NULL)

plot(x,Z2[,1], main=paste('Simulation trajectoires PG de
                           covariance exponentielle avec theta =',theta),
      type='l', ylim=c(min(Z2),max(Z2)))
for (i in 2:n)
  lines(x,Z2[i,] ,col=i)
```

Objectif : Simuler des trajectoires d'un processus gaussien de moyenne et covariance données

1.a) Simulation de trajectoires 1D

Simulation trajectoires PG de covariance exponentielle avec $\theta = 0.05$



Objectif : Simuler des trajectoires d'un processus gaussien de moyenne et covariance données

1.a) Simulation de trajectoires 1D

Énoncé

Soit un processus gaussien centré stationnaire de covariance gaussienne avec effet de pépité : $C(t, t') = \sigma^2 e^{-\left(\frac{h}{\theta}\right)^2} + \lambda \delta_0(h)$. Simuler des réalisations de $Z(t)$ de $N = 200$ points sur $[0 ; 1]$. Prendre $\sigma^2 = 0,5$, $\theta = 0,05$ puis $\theta = 0,2$.

Méthode

Ajout de l'effet de pépité => diag

Objectif : Simuler des trajectoires d'un processus gaussien de moyenne et covariance données

1.a) Simulation de trajectoires 1D

Solution

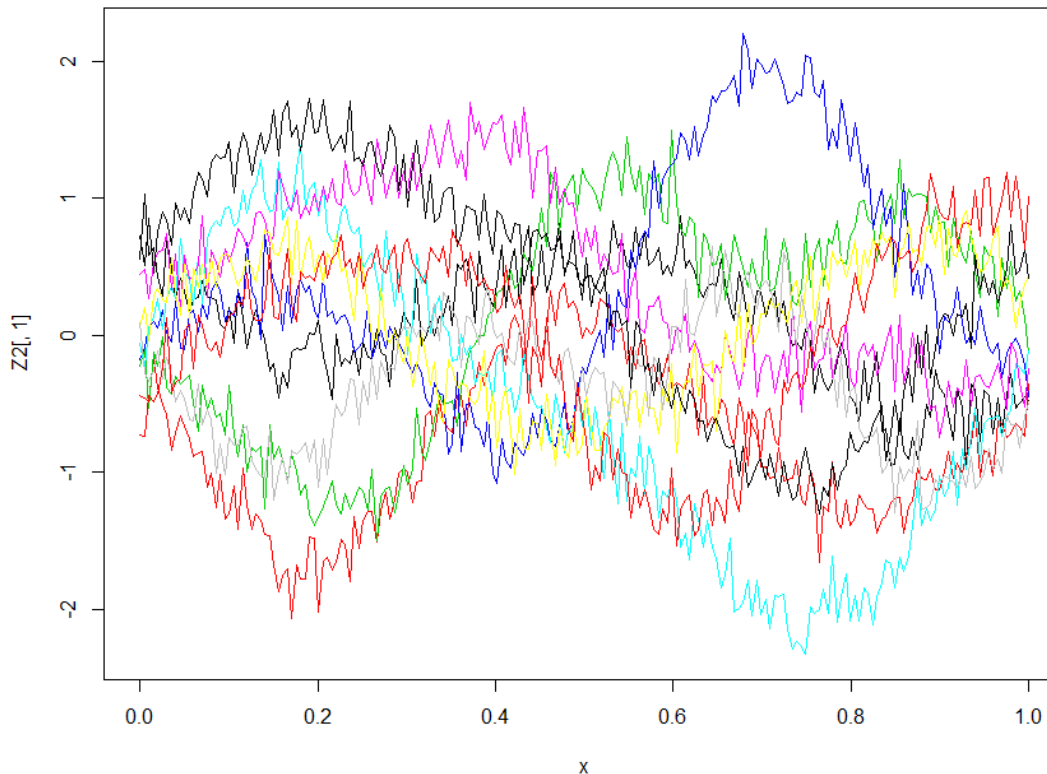
```
sig2 <- 0.5
theta <- 0.2
lambda <- 0.05
cov <- sig2*(exp(-(dist_all/theta)^2) + lambda * diag(1,N))
Z2 <- Z1
for (i in 1:n)
  Z2[,i] <- f0+t(chol(cov))%*%Z1[,i]

plot(x,Z2[,1], main=paste('Simulation trajectoires PG de covariance
  gaussienne avec theta =',theta),type='l',ylim=c(min(Z2),max(Z2)))
for (i in 2:n)
  lines(x,Z2[,i],col=i)
```

Objectif : Simuler des trajectoires d'un processus gaussien de moyenne et covariance données

1.a) Simulation de trajectoires 1D

Simulation trajectoires PG de covariance gaussienne avec $\theta = 0.2$



Objectif : Simuler des trajectoires d'un processus gaussien de moyenne et covariance données

1.b) Simulation de trajectoires 2D

Énoncé

Soit un processus gaussien centré stationnaire de covariance

gaussienne isotrope $C(t, t') = \sigma^2 e^{-\left\|\frac{h}{\theta}\right\|^2}$.

Simuler des réalisations de $Z(t)$ de $N = 50 \times 50$ points sur $[0 ; 1] \times [0 ; 1]$.

Prendre $\sigma^2 = 0,5$ et $\theta = 0,1$.

A réaliser en dehors du TP => compte-rendu

Objectif : Simuler des trajectoires d'un processus gaussien de moyenne et covariance données

1.b) Simulation de trajectoires 2D

Énoncé

Soit un processus gaussien centré stationnaire de covariance

gaussienne anisotrope $C(t, t') = \sigma^2 e^{-\left(\frac{h_1}{\theta_1}\right)^2 - \left(\frac{h_2}{\theta_2}\right)^2}$.

Simuler des réalisations de $Z(t)$ de $N = 70 \times 70$ points sur $[0 ; 1] \times [0 ; 1]$.

Prendre $\sigma^2 = 0,5$, $\theta_1 = 0,1$ et $\theta_2 = 0,03$.

A réaliser en dehors du TP => compte-rendu

Objectif : Simuler des trajectoires d'un processus gaussien de moyenne et covariance données

1.b) Simulation de trajectoires 2D

Énoncé

Soit un processus gaussien centré stationnaire de covariance

exponentielle anisotrope $C(t, t') = \sigma^2 e^{-\frac{h_1}{\theta_1} - \frac{h_2}{\theta_2}}$.

Simuler des réalisations de $Z(t)$ de $N = 70 \times 70$ points sur $[0 ; 1] \times [0 ; 1]$.

Prendre $\sigma^2 = 0,5$, $\theta_1 = 0,1$ et $\theta_2 = 0,03$.

A réaliser en dehors du TP => compte-rendu

EXERCICE 2 : KRIGEAGE 1D

► Objectif : Construire un métamodèle PG en dimension 1

Énoncé

Soit $f(x) = \sin(30(x - 0,9)^4) \cos(2(x - 0,9)) + \frac{x-0,9}{2}$

Etape 0 & 1

- Évaluer f sur *100 points équirépartis* et tracer f
- Construire un plan d'expérience et évaluer f sur ce plan
- Faire varier le nb de points de 10 à 30 et le type de plans : équirépartis et LHS.

Méthode

- Points équirépartis => seq
- Plan LHS => optimumLHS du package lhs

Objectif : Construire un métamodèle PG en dimension 1**Solution**

```

N_BA <- 15
S_BA <- optimumLHS(N_BA,1)
N_BT <- 100
S_BT <- seq(0,1,length.out=N_BT)

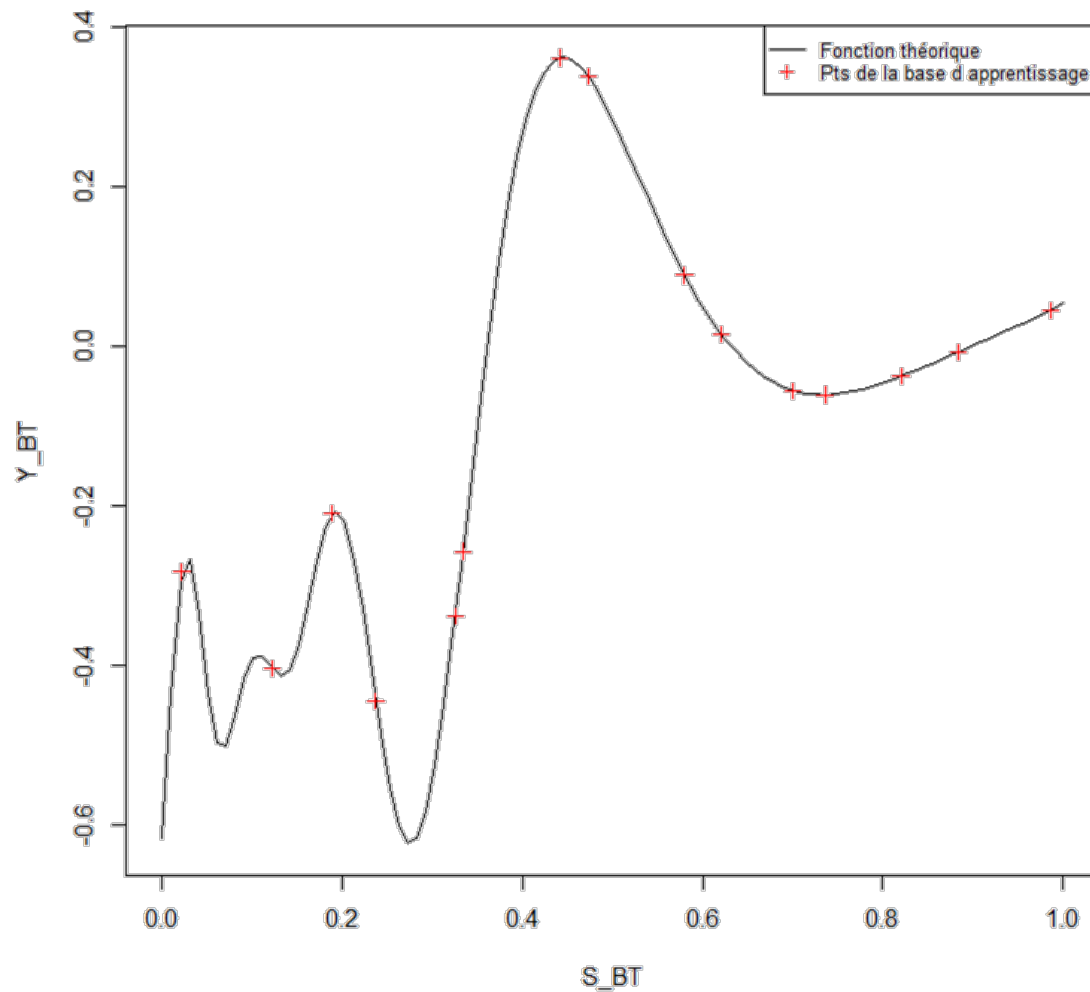
fx_exo2 <- function(X)
  sin(30*(X-0.9)^4)*cos(2*(X-0.9)) + (X-0.9)/2

Y_BA <- fx_exo2(S_BA)
Y_BT <- fx_exo2(S_BT)

plot(S_BT,Y_BT,type='l')
points(S_BA,Y_BA,pch=3,col=2)
legend("topright",y=c(0.17,0.33),c('Fonction théorique','Pts
  de la base d'apprentissage'),lty=c(1,0),pch=c(-1,3),
  col=c(1,2),cex=0.8)

```

Objectif : Construire un métamodèle PG en dimension 1



Objectif : Construire un métamodèle PG en dimension 1

Énoncé

Etape 2

Soit un processus gaussien centré stationnaire de covariance Matern3/2.

Construire le métamodèle basé sur Z conditionnellement aux points de la base d'apprentissage :

- *Estimation des paramètres du métamodèle*
- *Calcul et tracé du prédicteur sur la base de test avec intervalle de confiance*
- *Calcul et tracé du MSE sur la base de test*
- *Calcul du Q^2 sur la base de test*

Méthode

- Estimation des paramètres => `km` du package `DiceKriging`
- Calcul du prédicteur => `predict.km` du package `DiceKriging`
- Intervalle de confiance et MSE fournis par `predict.km`

Objectif : Construire un métamodèle PG en dimension 1

Solution

```

metamodel <- km(formula=~1, design=S_BA, response=Y_BA,
               covtype="matern3_2")

prediction <- predict.km(metamodel,S_BT,'UK',checkNames=FALSE)

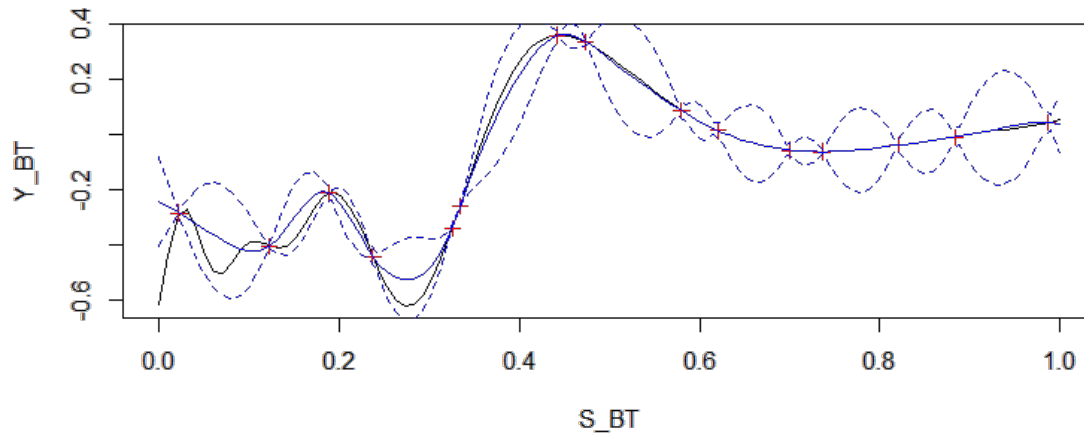
par(mfrow=c(2,1))
plot(S_BT,Y_BT,type='l')
points(S_BA,Y_BA,pch=3,col='red')
lines(S_BT,prediction$mean,col='blue')
lines(S_BT,prediction$lower95,col='blue',lty=2)
lines(S_BT,prediction$upper95,col='blue',lty=2)

plot(S_BT,prediction$sd,type='l')

Q2 <- function(y,yhat)
  1-mean((y-yhat)^2)/var(y)

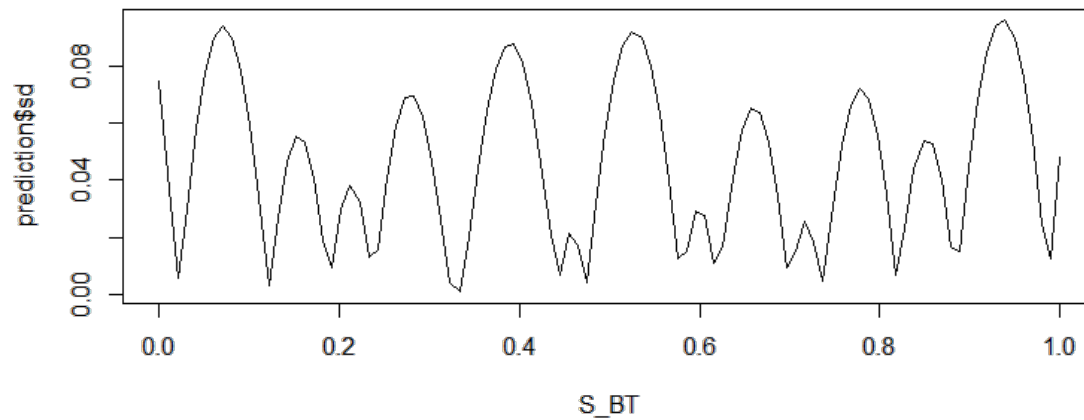
Q2_BT <- Q2(Y_BT,prediction$mean)

```

Objectif : Construire un métamodèle PG en dimension 1Estimation de f par métamodèle PG

$$Q^2 = 0,96$$

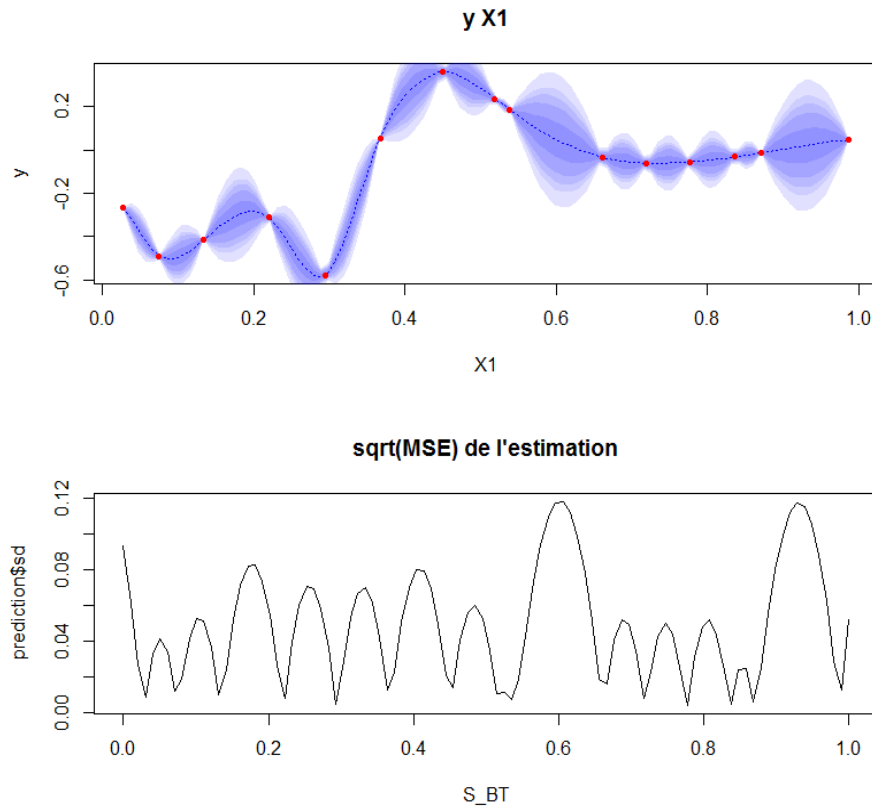
sqrt(MSE) de l'estimation



Objectif : Construire un métamodèle PG en dimension 1

Avec le package DiceView

```
par(mfrow=c(2,1))  
sectionview(metamodel)  
plot(S_BT, prediction$sd, type='l', main="sqrt(MSE) de l'estimation")
```



Objectif : Construire un métamodèle PG en dimension 1

Énoncé

Etape 3

- Calculer analytiquement les valeurs optimales des paramètres β et σ du métamodèle PG.
- Tracer la log-vraisemblance et retrouver la valeur optimale de θ .

A réaliser en dehors du TP => compte-rendu

Objectif : Construire un métamodèle PG en dimension 1

Énoncé

Étape 4

- Trouver le point x où le MSE est maximal et l'ajouter à la base d'apprentissage
- Mettre à jour le métamodèle (estimation des hyperparamètres, construction du prédicteur et du MSE)
- Tracer le prédicteur et le MSE comme à l'étape 2.

A réaliser en dehors du TP => compte-rendu

EXERCICE 3 : KRIGEAGE 2D

► Objectif : Construire un métamodèle PG en dimension 2

Énoncé

Soit $f(x_1, x_2) = -x_1 \sin \sqrt{|x_1|} - x_2 \sin \sqrt{|x_2|}$

- Évaluer f sur *70x70 points équirépartis* et tracer f
- Construire un plan d'expérience et évaluer f sur ce plan
- Faire varier le nb de points de 50 à 90 et le type de plans : équirépartis et LHS.

A réaliser en dehors du TP => compte-rendu

Objectif : Construire un métamodèle PG en dimension 2

Énoncé

Soit un processus gaussien centré stationnaire de covariance Matern3/2.

Construire le métamodèle basé sur Z conditionnellement aux points de la base d'apprentissage.

- *Calcul et tracé du prédicteur sur la base de test*
- *Tracé de l'erreur (valeur absolue) sur la base de test*
- *Calcul du Q^2 sur la base de test*

A réaliser en dehors du TP => compte-rendu

EXERCICE 4 : APPLICATION À L'ANALYSE DE SENSIBILITÉ

Exercice 4 : Application à l'analyse de sensibilité

► Objectif : Construire un métamodèle et y appliquer des méthodes d'analyse de sensibilité

Énoncé

Soit la fonction d'Ishigami :

$$f(X_1, X_2, X_3) = \sin X_1 + a(\sin X_2)^2 + bX_3^4 \sin X_1$$

On pose $a = 7$ et $b = 0,1$.

Etape 0

- Calculer les indices de Sobol du 1^{er} et du 2nd ordre théoriques de f
- Calculer les indices de Sobol totaux de f

Etape 1

- Construire la base d'apprentissage : un plan LHS de $N = 70$ à 100 points

Méthode

- Plan LHS => lhs

Exercice 4 : Application à l'analyse de sensibilité

Objectif : Construire un métamodèle et y appliquer des méthodes d'analyse de sensibilité

Solution

```
N_BA <- 80
d <- 3
S_BA <- randomLHS(N_BA,d)

ishigami <- function(X,coeff){
  X <- 2*pi*X-pi
  sin(X[,1]) + coeff[1]*sin(X[,2])^2 +
  coeff[2]*X[,3]^4*sin(X[,1])
}

coeff_fx_test <- c(7, 0.1)
Y_BA <- ishigami(S_BA,coeff_fx_test)
```

Objectif : Construire un métamodèle et y appliquer des méthodes d'analyse de sensibilité

Pour $a = 7$ et $b = 0,1$:

Indices de Sobol théoriques :

X1	X2	X3
0.314	0,442	0

Indices de Sobol théoriques totaux :

X1	X2	X3
0.558	0,442	0,244

Exercice 4 : Application à l'analyse de sensibilité

Objectif : Construire un métamodèle et y appliquer des méthodes d'analyse de sensibilité

Énoncé

Etape 2 :

- *Construire le métamodèle basé sur Z conditionnellement aux points de la base d'apprentissage.*
- *Calcul du Q^2 sur une base de test*
- *Calcul du Q^2 par validation croisée sur la base d'apprentissage*

Méthode

- métamodèle $\Rightarrow$ km

Exercice 4 : Application à l'analyse de sensibilité

Objectif : Construire un métamodèle et y appliquer des méthodes d'analyse de sensibilité

Solution

```
metamodel <- km(formula=~1, design=S_BA,response=Y_BA,
               covtype="matern5_2")

Q2 <- function(y,yhat)
  1-mean((y-yhat)^2)/var(y)

# Calcul de Q2 sur une base de test
S_BT <- randomLHS(100,3)
Y_BT <- ishigami(S_BT,coeff_fx_test)
Y_BT_hat <- predict(m, S_BT, 'UK', checkNames=FALSE)$mean
Q2(Y_BT,Y_BT_hat)

# Calcul du Q2 par validation croisée
yhat <- leaveOneOut.km(metamodel,'UK')$mean
Q2(Y_BA, yhat)
```

Exercice 4 : Application à l'analyse de sensibilité

Objectif : Construire un métamodèle et y appliquer des méthodes d'analyse de sensibilité

Q^2 sur une base de test : 0,88

Q^2 par validation croisée : 0,93

Exercice 4 : Application à l'analyse de sensibilité

Objectif : Construire un métamodèle et y appliquer des méthodes d'analyse de sensibilité

Énoncé

Etape 3 :

Calcul des indices de sensibilité

- Avec la méthode de Sobol
- Avec FAST (avec $M=4$)

Méthode

- Méthode de Sobol => `sobol`
- Méthode FAST => `fast99`

Exercice 4 : Application à l'analyse de sensibilité

Objectif : Construire un métamodèle et y appliquer des méthodes d'analyse de sensibilité

Solution

```
##Sobol
n <- 5000
X1 <- data.frame(X1=runif(n,0,1),X2=runif(n,0,1),X3=runif(n,0,1))
X2 <- data.frame(X1=runif(n,0,1),X2=runif(n,0,1),X3=runif(n,0,1))
indices.sobol <- sobol2007(f,X1,X2)

plot(indices.sobol)

## FAST
n_RBDFast <- 5000
M_RBDFast <- 4
f <- function(x)
  predict(metamodel,x,type='UK',checkNames=FALSE)$mean
indices.fast <-
fast99(f,d,n_RBDFast,M_RBDFast,q='qunif',q.arg=list(min=0,max=1))

plot(indices.fast)
```

Objectif : Construire un métamodèle et y appliquer des méthodes d'analyse de sensibilité

Indices du 1^{er} ordre

	X1	X2	X3
Théoriques	?	?	?
Sobol [2007]	?	?	?
FAST	?	?	?

Indices totaux

	X1	X2	X3
Théoriques	?	?	?
Sobol [2007]	?	?	?
FAST	?	?	?