

Ecole PECNUM

Métamodèle Processus Gaussien

Amandine Marrel
CEA

Sébastien Da Veiga
Safran

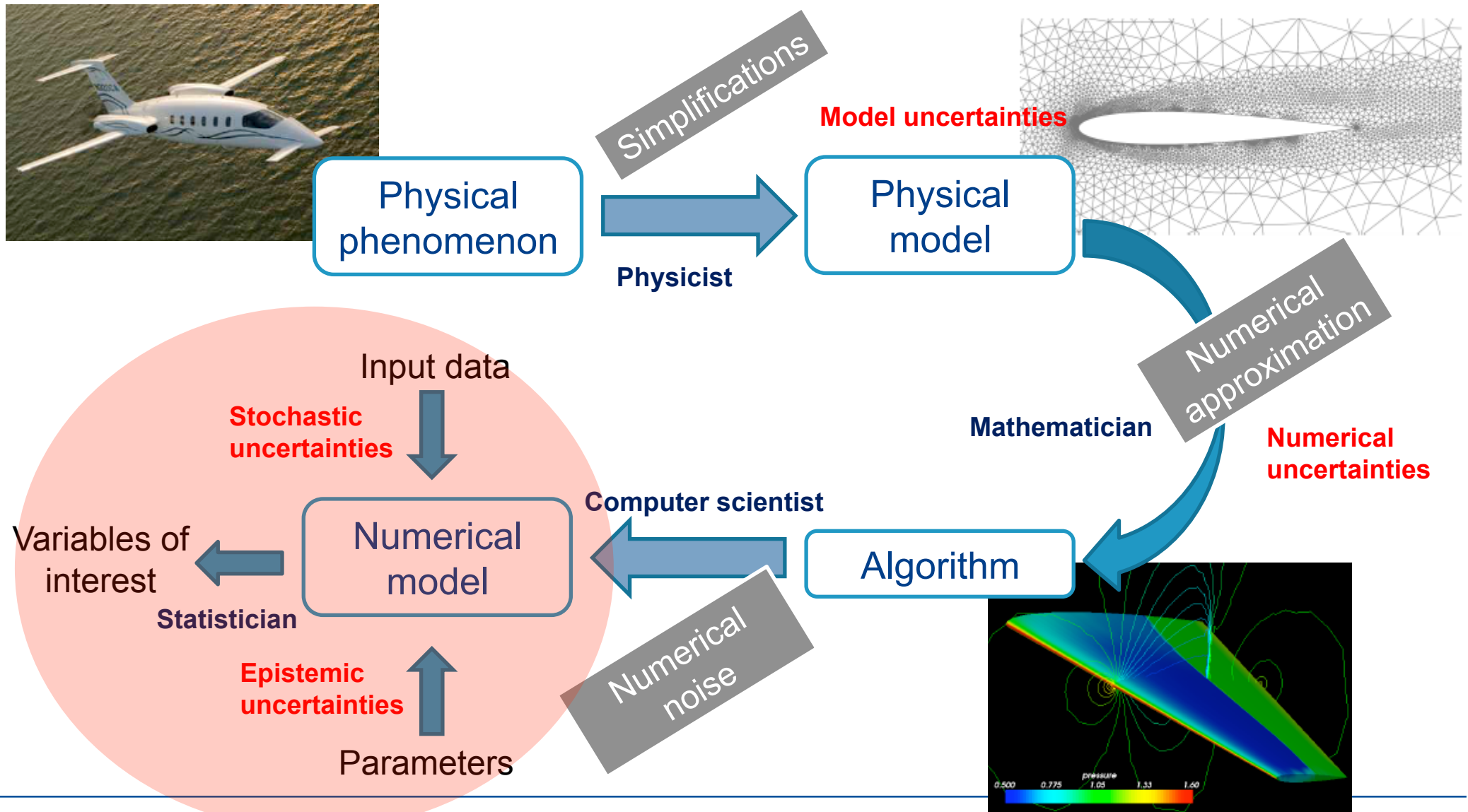
PLAN

- **Contexte et problématique**
- **Régression par processus Gaussiens**

/01/

CONTEXTE ET PROBLÉMATIQUE

CONTEXTE



CONTEXTE

➤ 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 ?

➤ Calibration de modèle :

- Paramètres à estimer pour se recalibrer sur des observations



PROPAGATION DES INCERTITUDES

➤ Problème :

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



Exploitation directe du code difficile



Utilisation d'un métamodèle

POURQUOI UN MÉTAMODÈLE ?

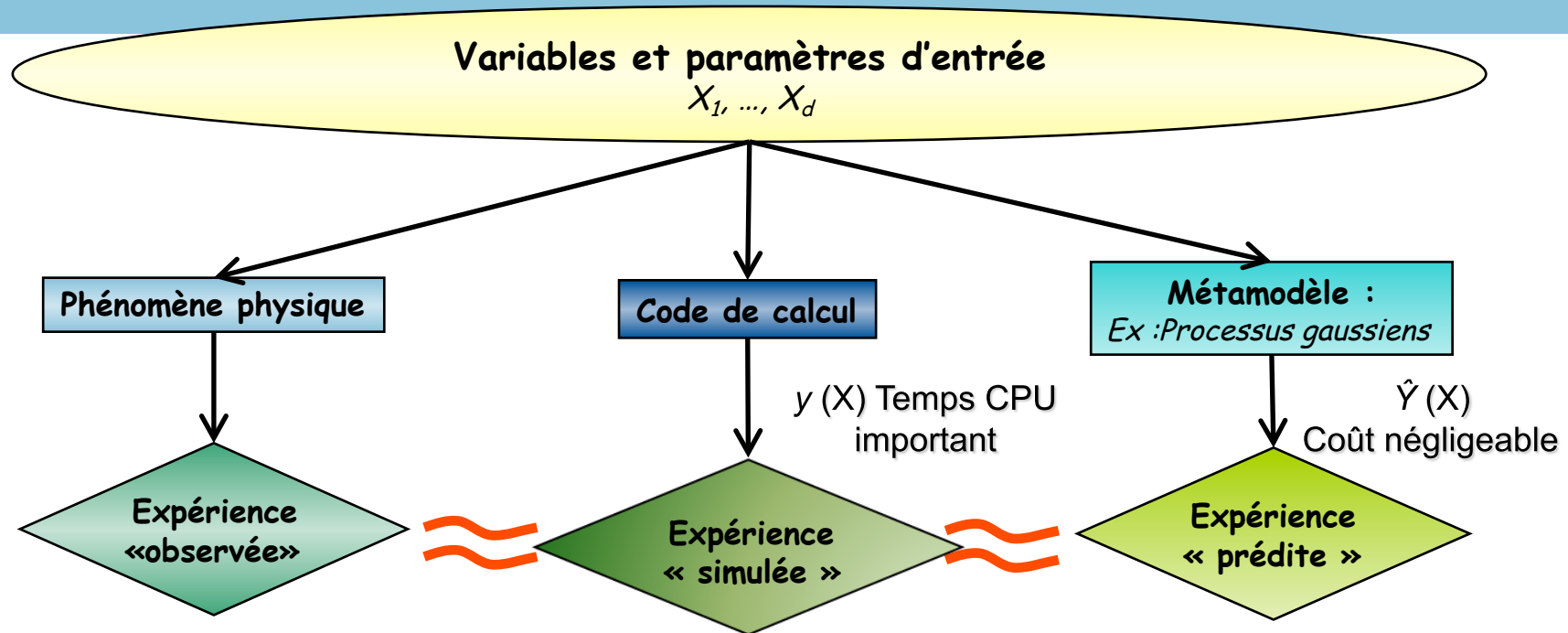
→ **Métamodèle :**

- 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

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

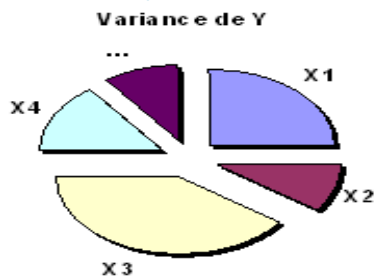
→ **Autres appellations : surface de réponse , émulateur, surrogate model, modèle simplifié,...**

POURQUOI UN MÉTAMODÈLE ?

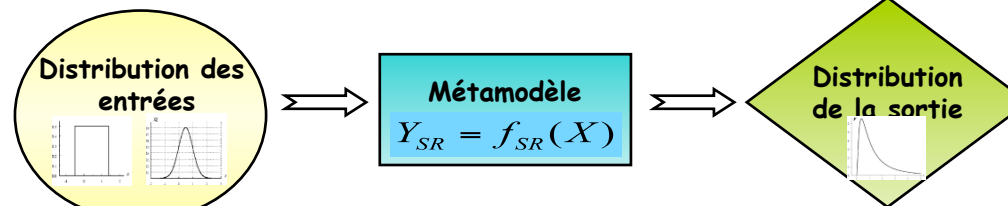


➤ Exploitation du métamodèle

▪ Analyse de sensibilité



▪ Propagation d'incertitudes



▪ Calibration

Détermination des paramètres
↓
Adéquation expériences simulées et observées

ETAPES DE LA METAMODÉLISATION

Planification d'expériences :

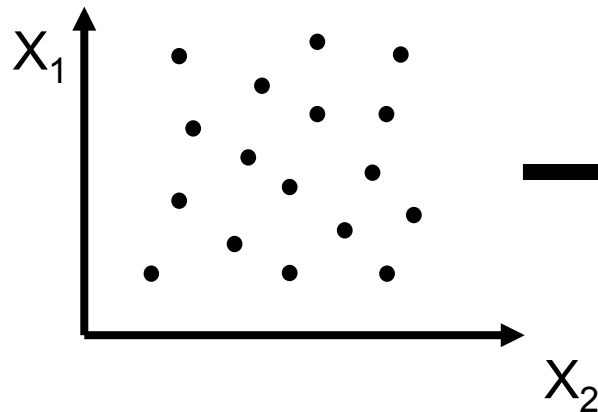
points où l'on va effectuer les simulations (Space Filling Design)

Simulation :

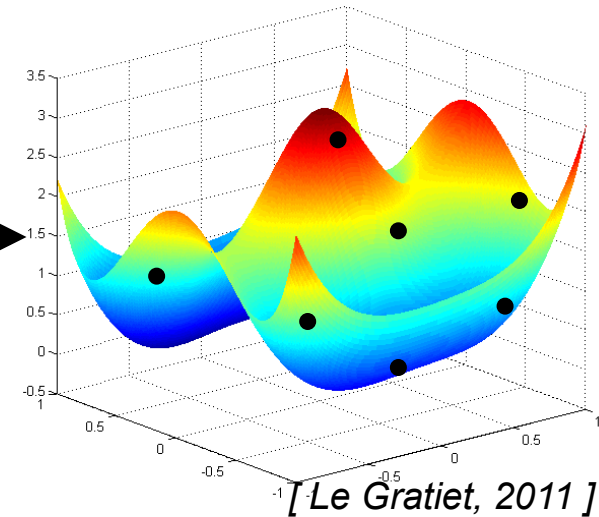
On effectue les simulations.

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

POURQUOI UN MÉTAMODÈLE ?

→ Différents types de métamodèles :

- Modèle linéaire, modèle linéaire généralisé (GLM)
- 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

POURQUOI UN MÉTAMODÈLE ?

→ Différents types de métamodèles :

- Modèle linéaire, modèle linéaire généralisé (GLM)
- 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

PLAN

→ Contexte et problématique

→ **Régression par processus Gaussiens**

/03/

RÉGRESSION PAR PROCESSUS GAUSSIENS - KRIGEAGE

RAPPELS THÉORIE

→ 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 on associe n éléments de $\mathbb{R}$
 - $X : \omega \rightarrow [X_1(\omega), X_2(\omega), \dots, X_n(\omega)]$
- 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 : $Z(x,y), Z(x,y,z), Z(x,t) \dots$

RAPPELS THÉORIE

→ Variable aléatoire vectorielle :

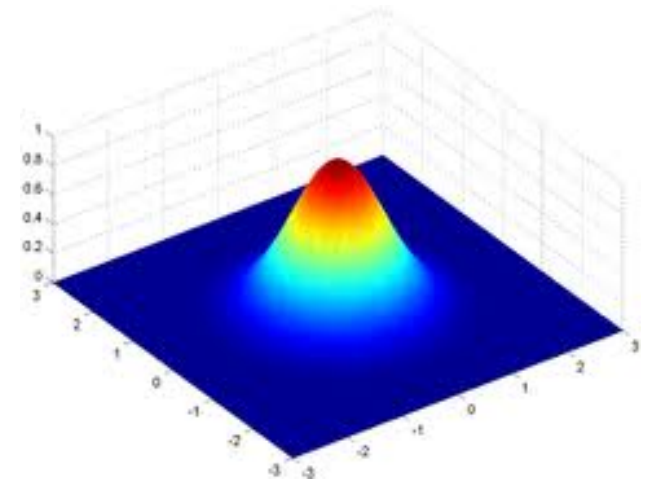
- Variable aléatoire réelle à 2 composantes $\omega \rightarrow (Z_1(\omega), Z_2(\omega)) \in \mathbb{R}^2$
- Fonction de répartition conjointe :

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

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

- 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 :

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

$$\begin{aligned} \rho_{1,2} &= \frac{\text{Cov}(Z_1, Z_2)}{\sqrt{\text{Var}(Z_1) \text{Var}(Z_2)}} = \frac{\text{Cov}(Z_1, Z_2)}{\sigma_1 \sigma_2} \\ &-1 \leq \rho_{1,2} \leq 1 \end{aligned}$$

RAPPELS THÉORIE

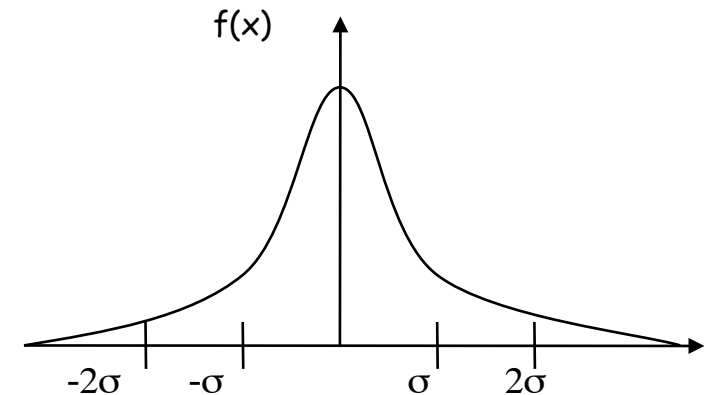
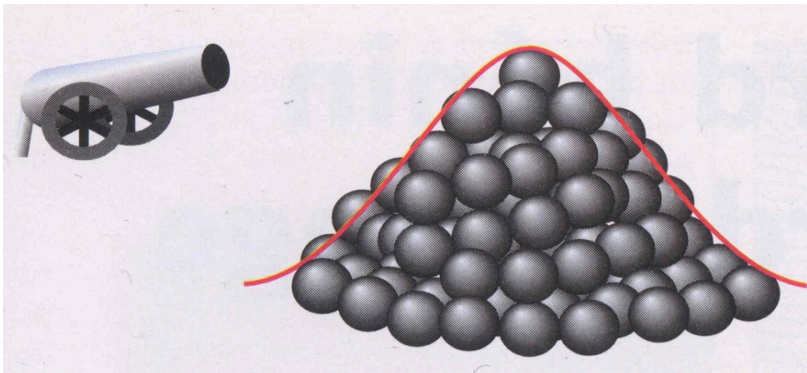
→ Loi gaussienne ou loi normale

- 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)*

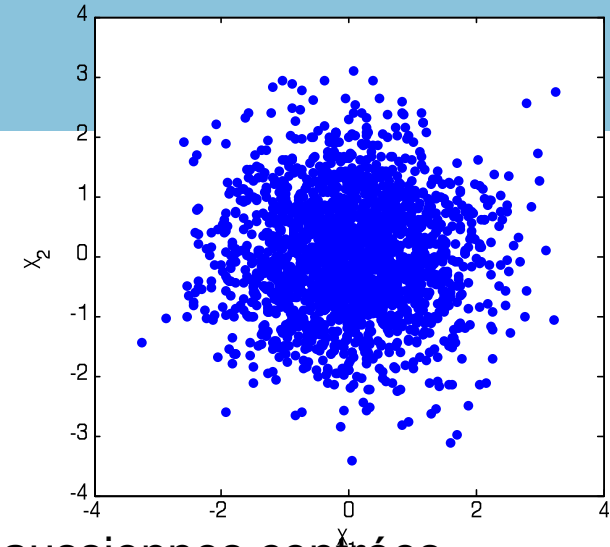
RAPPELS THÉORIE

→ Vecteur gaussien

- Loi $\mathcal{N}(0, I)$

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

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



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

- Si toute combinaison linéaire de ses composantes suit une loi gaussienne
- Si $\mathbf{Z}$ s'écrit comme une transformation affine d'une $\mathcal{N}(0, I)$

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

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

RAPPELS THÉORIE

→ Vecteur gaussien

- Propriétés : Transformation affine

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

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



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

RAPPELS THÉORIE

→ 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

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

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

RAPPELS THÉORIE

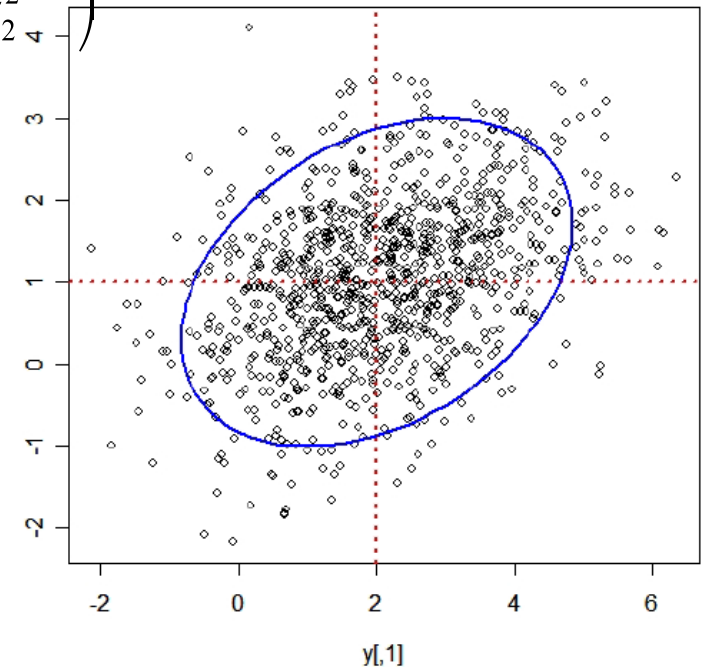
→ Vecteur gaussien

- 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%



RAPPELS THÉORIE

→ 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 = corrélation
 - Représentation de loi multi-gaussienne : Z admet une densité :

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

RAPPELS THÉORIE

→ Vecteur gaussien

- Conditionnement $\mathbf{Z} = (\mathbf{Z}_1, \mathbf{Z}_2)^\top$
- $Z_2|Z_1=z_1$ 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 : $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$

$$f_{Z_2|Z_1=z_1}(z_2) = \frac{f_{Z_2,Z_1}(z_2, z_1)}{f_{Z_1}(z_1)}$$

RAPPELS THÉORIE

→ Vecteur gaussien

- Conditionnement : extension au cas général

$$\mathbf{Z}'_1 = (\mathbf{Z}_1, \dots, \mathbf{Z}_p)^\top \text{ et } \mathbf{Z}'_2 = (\mathbf{Z}_{p+1}, \dots, \mathbf{Z}_n)^\top$$

$$(\mathbf{Z}'_1, \mathbf{Z}'_2)^\top \sim N \left(\begin{pmatrix} m'_1 \\ m'_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{1,2} \\ \Sigma_{1,2}^\top & \Sigma_{22} \end{pmatrix} \right)$$

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 $\mathbf{Z}'_2 | \mathbf{Z}'_1 = \mathbf{z}'_1$:

$$\mathbf{Z}'_2 | \mathbf{Z}'_1 = \mathbf{z}'_1 \sim N \left(m'_2 + \Sigma_{1,2}^\top \Sigma_{11}^{-1} (\mathbf{z}'_1 - m'_1) \Sigma_{22} - \Sigma_{1,2}^\top \Sigma_{11}^{-1} \Sigma_{1,2} \right)$$

RAPPELS THÉORIE

→ Processus Gaussiens (PG)

- Définition : processus aléatoire réel $Z(x)$ dont toutes les lois finies-dimensionnelles $(Z(x_1), \dots, Z(x_n))$ sont gaussiennes.
- Propriétés : entièrement caractérisé par sa moyenne et sa covariance

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

$$\text{où } m(x) = E[Z(x)]$$

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

- Lien avec le vecteur gaussien :
 - N points de simulations du processus gaussien $\{x_1, \dots, x_n\}$
 - N observations associées : $Z(x_1), \dots, Z(x_n)$
 - **$Z = [Z(x_1), \dots, Z(x_n)]$ est un vecteur gaussien**

RAPPELS THÉORIE

→ Vecteur gaussien $[Z(x_1), Z(x_2), \dots, Z(x_n)]^T \sim \mathcal{N}(m, \Sigma)$

- Vecteur de la moyenne

$$m = [m(x_1), \dots, m(x_n)]^T$$

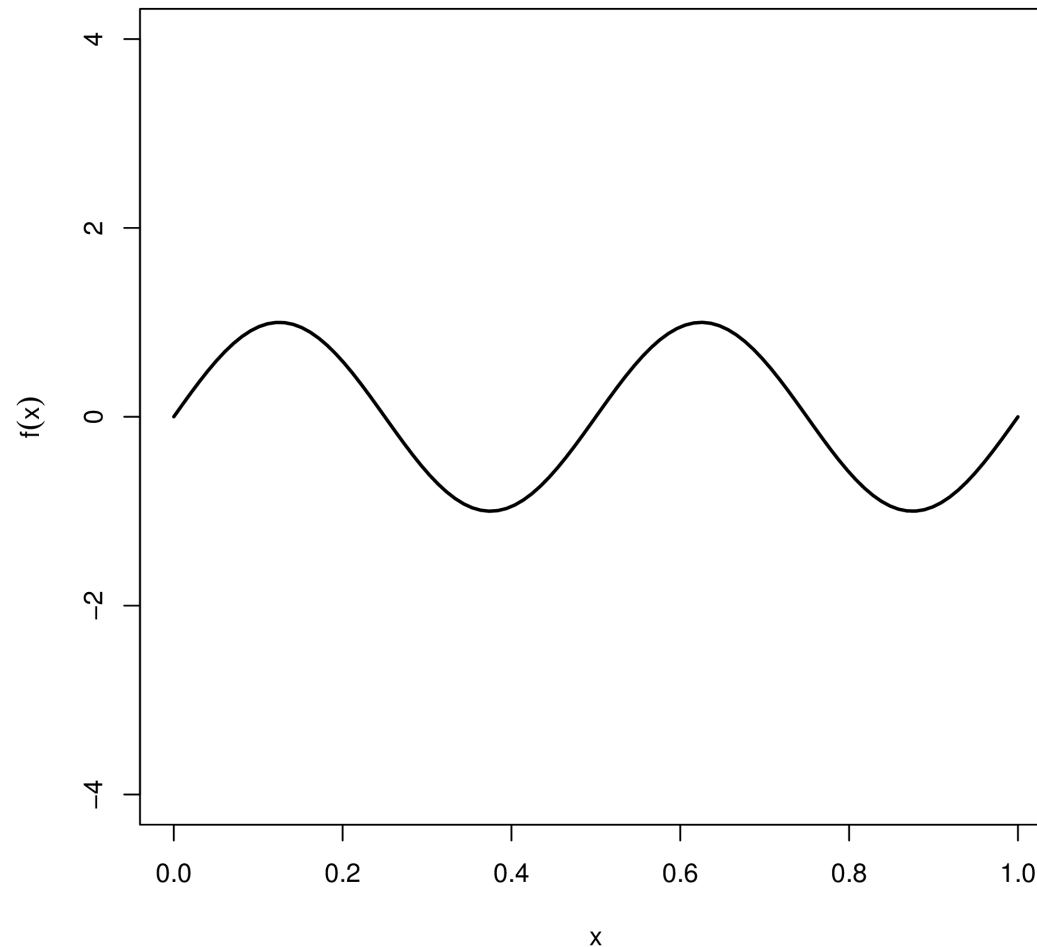
- Matrice de covariance

$$\Sigma = \begin{pmatrix} \sigma_1^2 & C(x_1, x_2) & \dots & C(x_1, x_n) \\ C(x_1, x_2) & \sigma_2^2 & \dots & C(x_2, x_n) \\ \dots & \dots & \dots & \dots \\ C(x_1, x_n) & C(x_2, x_n) & \dots & \sigma_n^2 \end{pmatrix}$$

$$\text{avec } \sigma_i^2 = \text{Var}[Z_i] = \text{Cov}(Z(x_i), Z(x_i)) = C(x_i, x_i)$$

DÉFINITION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

Fonction théorique (code de calcul e.g.), à approcher par le métamodèle



DÉFINITION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ Formulation analytique du métamodèle PG :

- Entrées du code : x dans un domaine de $\mathbb{R}^d$
- Sortie du code : $z(x)$ dans $\mathbb{R}$
- Hypothèse : la sortie déterministe $z(x)$ est considérée comme la réalisation d'un champ aléatoire $Z(x)$:

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

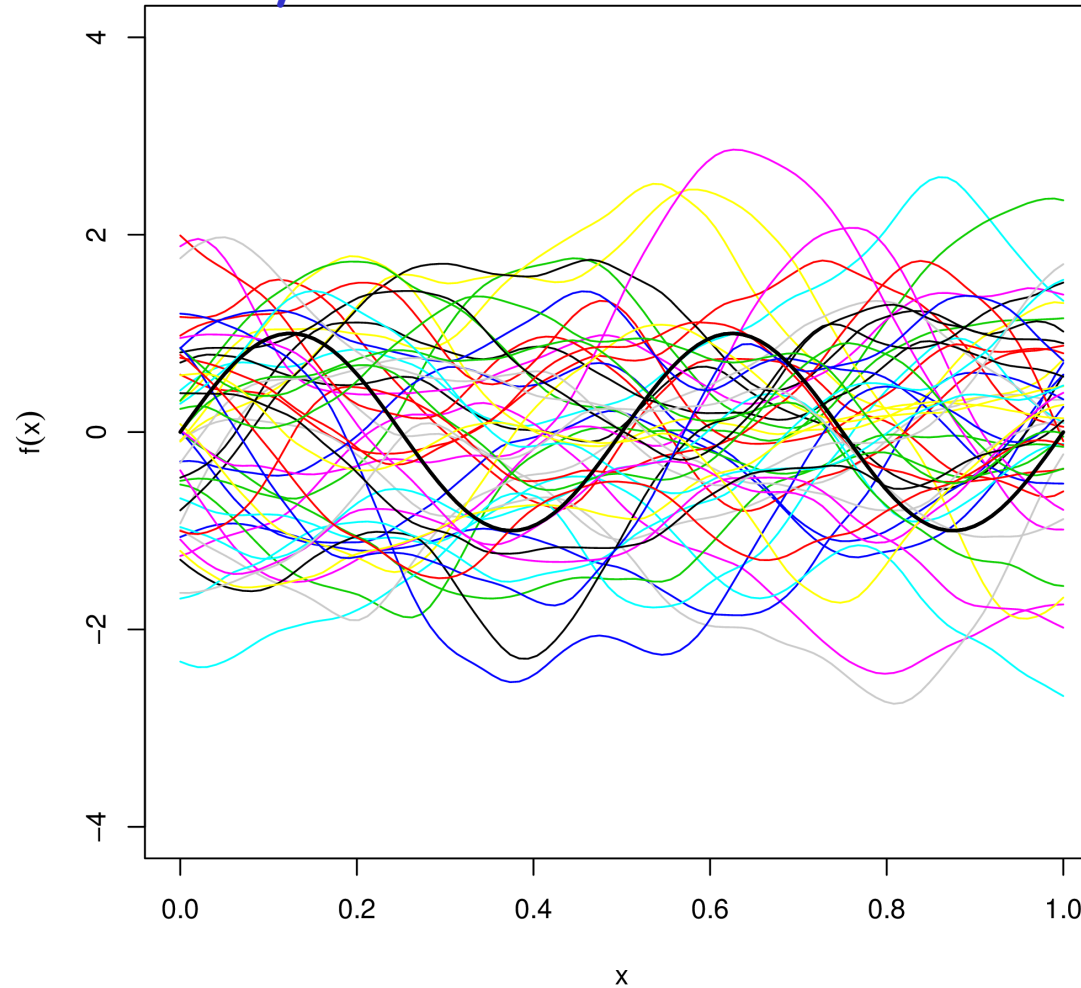
avec :

- m la fonction moyenne
- $W(x)$ un processus gaussien centré ($E[W(x)] = 0$) de variance σ^2 et de fonction de corrélation R :

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

DÉFINITION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

Exemples de réalisations du PG (trajectoires)
de moyenne m et de covariance C fixées

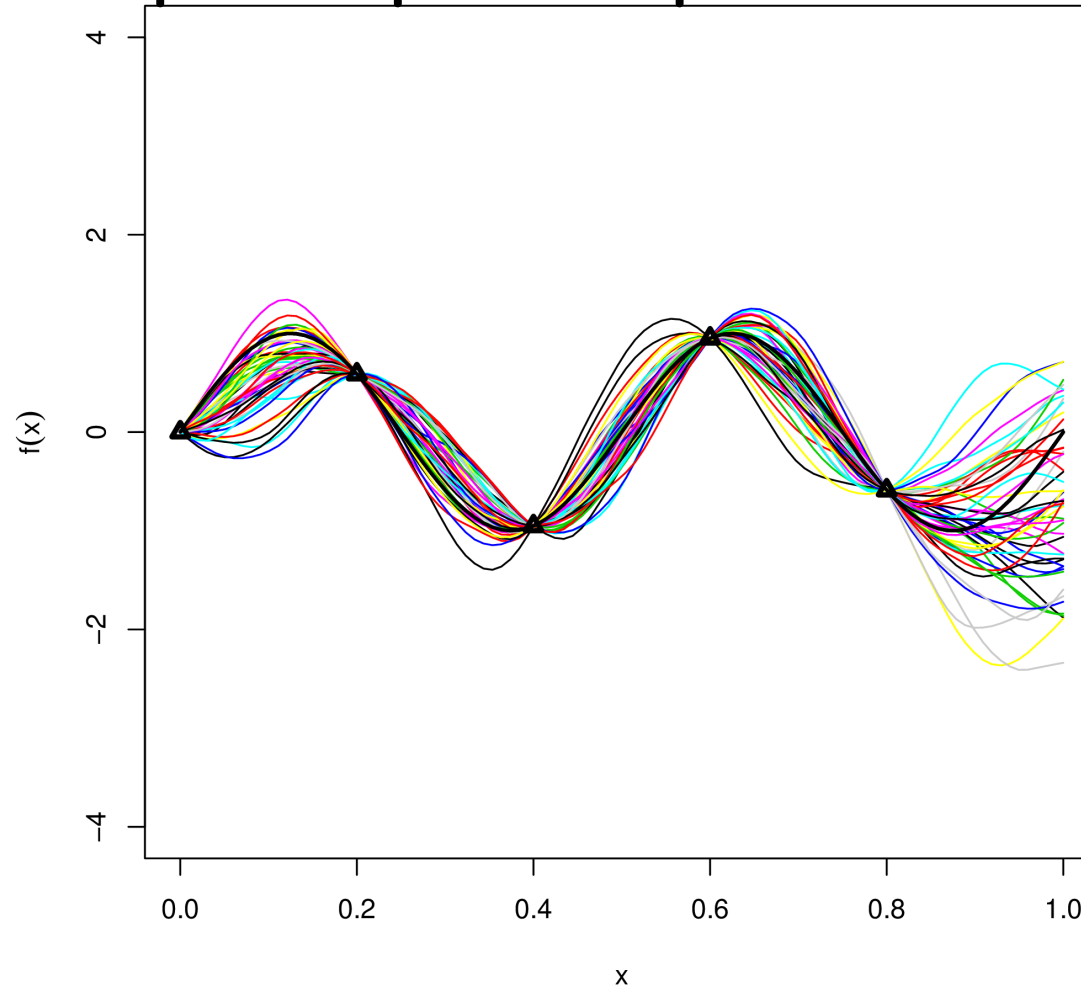


DÉFINITION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

- On conditionne ensuite le PG à passer par les n points observés (points de simulations du code de calcul)
- Le PG conditionné reste un PG (formules analytiques pour moyenne et covariance)
 - => Métamodèle PG
- Prédicteur = moyenne des réalisations du PG conditionnel (Espérance)

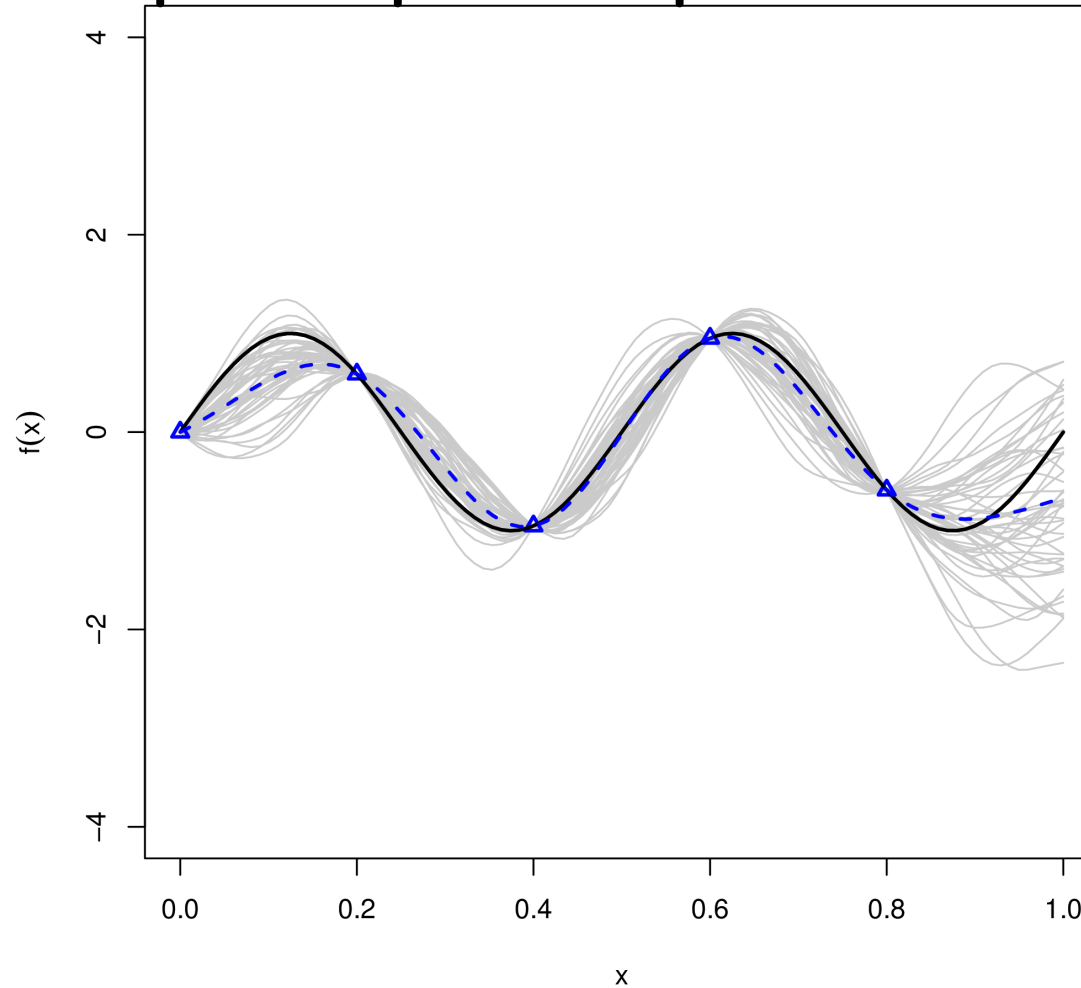
DÉFINITION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

Trajectoires du PG (de moyenne m et de covariance C fixées)
passant par les 5 points observés



DÉFINITION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

Trajectoires du PG (de moyenne m et de covariance C fixées)
passant par les 5 points observés



DÉFINITION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ 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: $\mathbf{Z}_s \sim \mathcal{N}(\mathcal{M}_s, \Sigma_s)$

avec $\mathcal{M}_s = m(\mathbf{X}_s) = [m(\mathbf{x}_1), \dots, m(\mathbf{x}_n)]^T$ 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 $\mathbf{x}^*$

DÉFINITION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ Formulation analytique du métamodèle PG :

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

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

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

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

$$\text{avec} \begin{cases} \mu(x^*) = E[Z(x^*) | Z(X_s) = Z_s] = m(x^*) + k(x^*)^T \Sigma_s^{-1} (Z_s - M_s) \\ \tilde{\sigma}^2(x^*) = Var[Z(x^*) | Z(X_s) = Z_s] = \sigma^2 - k(x^*)^T \Sigma_s^{-1} k(x^*) \end{cases}$$

DÉFINITION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ Formulation analytique du métamodèle PG :

- Prédicteur du métamodèle PG :

$$\hat{Z}(x^*) = \mu(x^*) = m(x^*) + k(x^*)^T \Sigma_S^{-1} (Z_S - M_S)$$

« 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, i.e. de variance minimale linéaire au sens combinaison linéaire des observations Z_S

DÉFINITION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ Propriétés du métamodèle PG :

- Variance du prédicteur => Erreur du prédicteur

$$MSE[\hat{Z}(x^*)] = \tilde{\sigma}^2(x^*) = \sigma^2 - k(x^*)^T \Sigma_S^{-1} k(x^*)$$

⇒ 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 géostatistique (« Weight space view »)

CONSTRUCTION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

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

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

avec :

- m la fonction moyenne
- $W(\mathbf{x})$: PG centré de variance σ^2 et de fx de corrélation R

→ Choix de la fonction « moyenne »

- Constante $m(\mathbf{x}) = c$
- Polynôme de degré 1 $m(\mathbf{x}) = \beta_0 + \sum_{i=1}^d \beta_i x_i$
- Polynôme d'ordre supérieur
- Modèle additif généralisé $m(\mathbf{x}) = \beta_0 + \sum_{i=1}^d \beta_i f_i(x_i) + \sum_{\substack{i,j=1 \\ i \neq j}}^d \beta_{i,j} f_{i,j}(x_i, x_j) + \dots$

CONSTRUCTION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

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

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

avec :

- m la fonction moyenne
- $W(x)$: PG centré de variance σ^2 et de fx de corrélation R

→ Choix de la fonction de covariance/corrélation

- Propriétés par définition:

- Symétrique $c(x, x') = c(x', x)$

- définie-positive

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

- Hypothèse : covariance stationnaire

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

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

CONSTRUCTION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ Choix de la fonction de covariance/corrélation

- Modèles paramétriques 1D: 1 seule variable d'entrée $X = X_1$

$$c(x, x') = c(x - x') = c(h) \quad \text{avec } h \in R$$

- Covariance exponentielle

$$c(h) = \sigma \exp\left(-\frac{h}{\theta}\right)$$

CONSTRUCTION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ Choix de la fonction de covariance/corrélation

- Modèles paramétriques 1D: 1 seule variable d'entrée $X = X_1$

$$c(x, x') = c(x - x') = c(h) \quad \text{avec } h \in R$$

- Covariance gaussienne

$$c(h) = \sigma \exp\left(-\left(\frac{h}{\theta}\right)^2\right)$$

CONSTRUCTION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ Choix de la fonction de covariance/corrélation

- Modèles paramétriques 1D: 1 seule variable d'entrée $X = X_1$

$$c(x, x') = c(x - x') = c(h) \quad \text{avec } h \in R$$

- Covariance de Matern

$$c(h) = \sigma \frac{1}{2^{\alpha-1} \Gamma(\alpha)} \left(\frac{h}{\theta} \right)^{\alpha} K_{\alpha} \left(\frac{h}{\theta} \right) \quad \text{avec } \alpha > 0$$

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 ($p=2$)

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

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

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

CONSTRUCTION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ Choix de la fonction de covariance/corrélation

- Généralisation en dimension supérieure: d variables d'entrée $[X_1, X_2 \dots X_d]$

- Covariance isotrope : $C(x, x') = C(\|x - x'\|) \Rightarrow c(h) = c(\|h\|)$

Hypothèse trop forte, non adaptée au cas des modèles numériques et à la nature très différente des entrées

- $\Rightarrow$ Préférable de conserver une covariance anisotrope

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

- Exponentielle

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

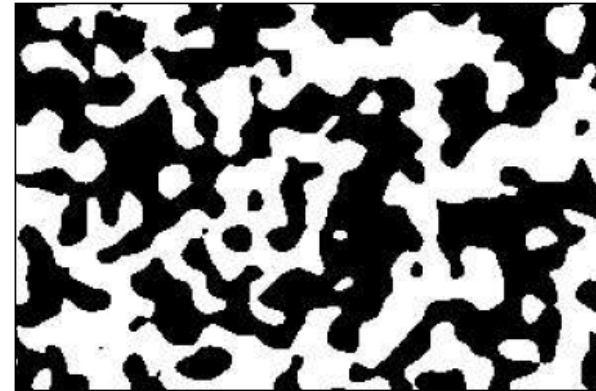
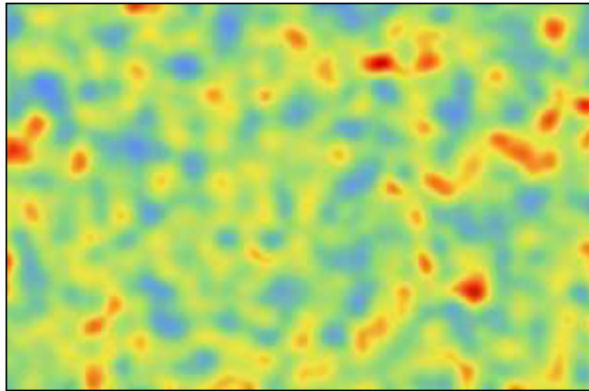
- Gaussienne

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

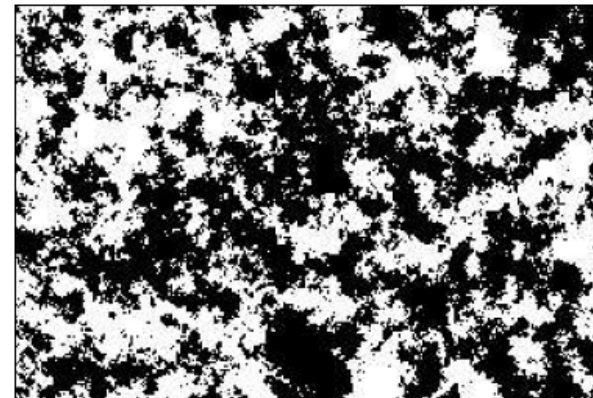
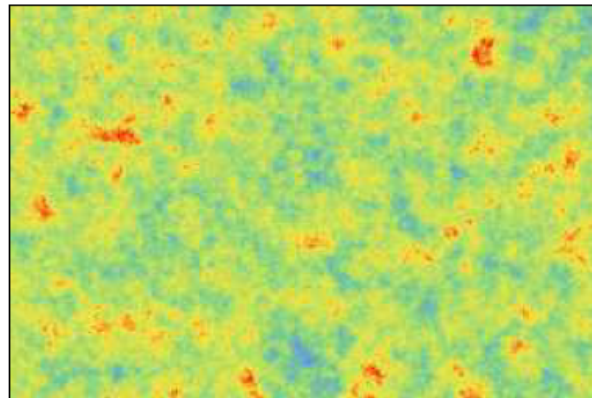
CONSTRUCTION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ Choix de la fonction de covariance/corrélation

- Covariance gaussienne 2D isotrope



- Covariance exponentielle



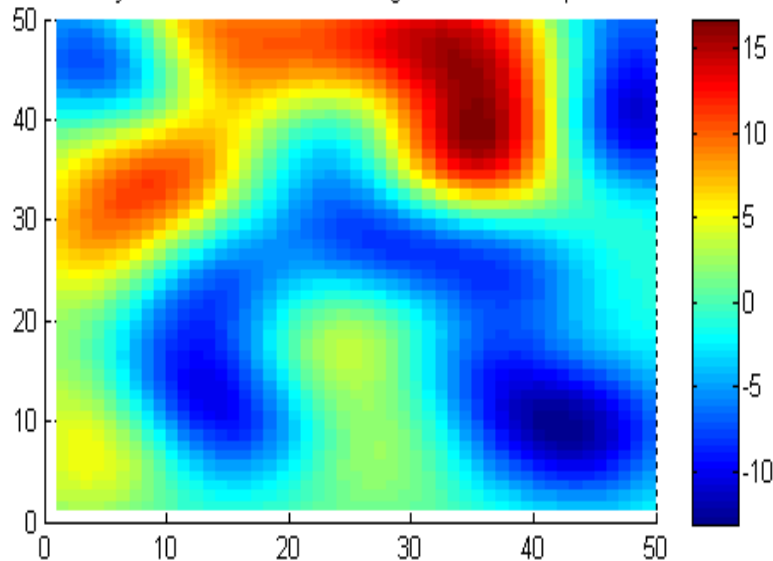
CONSTRUCTION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ Choix de la fonction de covariance/corrélation

- Covariance gaussienne 2D anisotrope

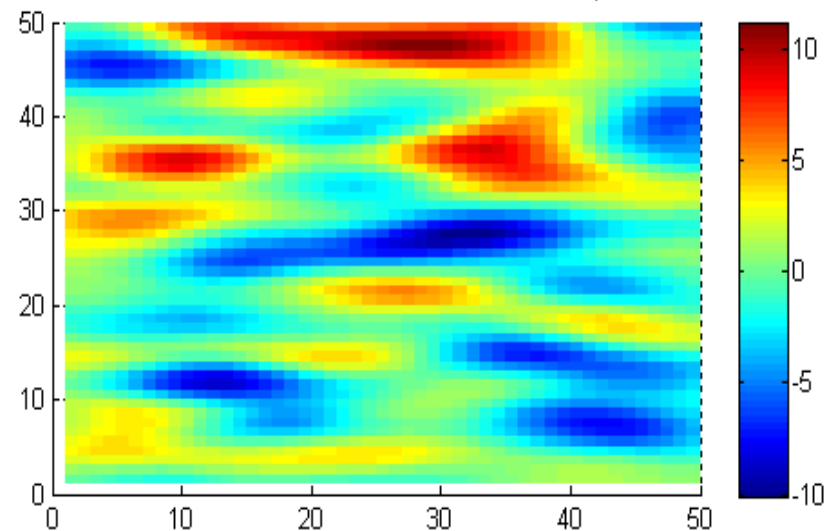
$$c(h) = \sigma^2 \exp \left(- \left(\frac{h_1}{\theta_1} \right)^2 - \left(\frac{h_2}{\theta_2} \right)^2 \right) \text{ avec } \theta_1 \neq \theta_2$$

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 $\theta_2 = 0.03$



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

CONSTRUCTION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ Choix de la fonction de covariance/corrélation

- Effet de pépite

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

- 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)

CONSTRUCTION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ 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 :

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

CONSTRUCTION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ Estimation des paramètres du métamodèle PG

- Si f_0 de forme linéaire

$$f_0(x) = \beta_0 + \sum_{i=1}^d \beta_i x_i \Rightarrow F_S = \tilde{X}_S \beta$$

$$\text{avec } \tilde{X}_S = [1, x_1^T, \dots, x_n^T]^T \text{ et } \beta = [\beta_0, \beta_1, \dots, \beta_d]^T$$

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

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

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

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

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



Optimisation numérique pour l'estimation des θ^*

CONSTRUCTION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ 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

$$\hat{Z}(x^*) = f_0(x^*) + k(x^*)^T \Sigma_S^{-1} (Z_S - F_S)$$

$$MSE[\hat{Z}(x^*)] = \hat{\sigma}^2 - k(x^*)^T \Sigma_S^{-1} k(x^*)$$

UTILISATION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ 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

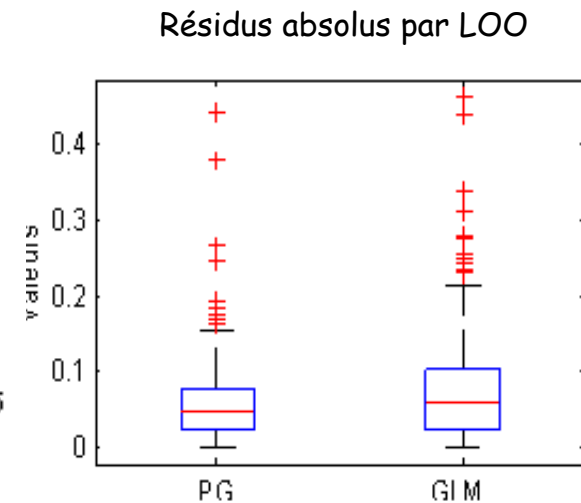
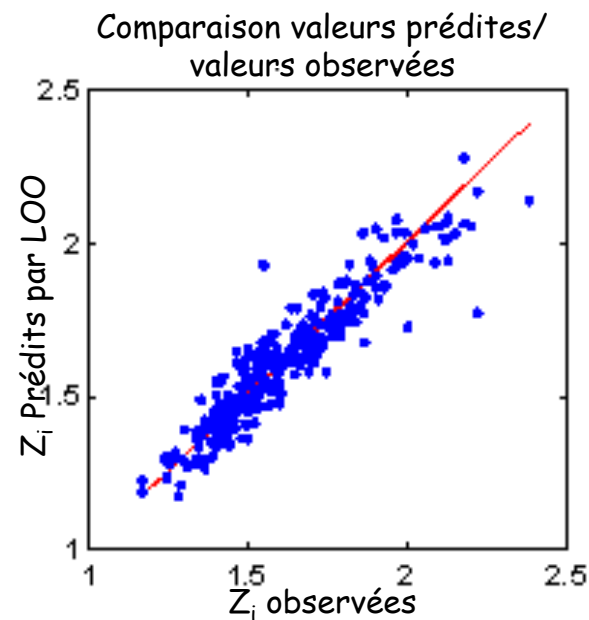
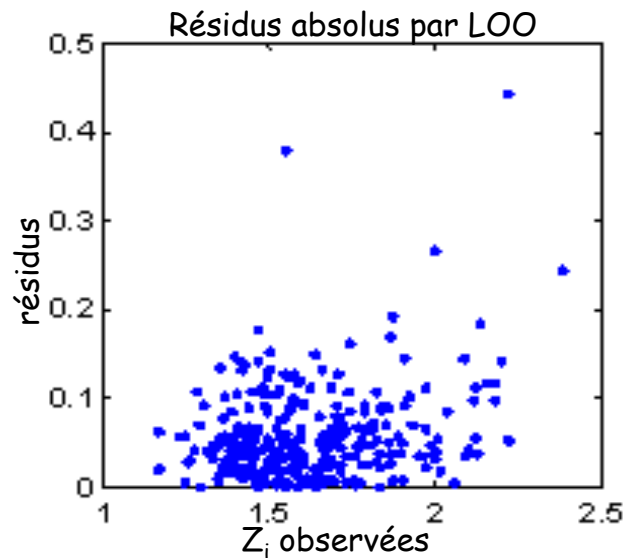


Comparaison des valeurs observées et des valeurs prédites: calcul des résidus

UTILISATION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ Validation du métamodèle PG

- Estimation des prédictions sur une base de test ou par validation croisée
- Validation graphique : Graphe des résidus, comparaison valeurs prédites/valeurs observées



UTILISATION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ Validation du métamodèle PG

- Validation quantitative :
 - Calcul du coefficient de prédictivité : Q^2
 ≈ 1 - écart quadratique normalisé entre prédictions et valeurs réelles

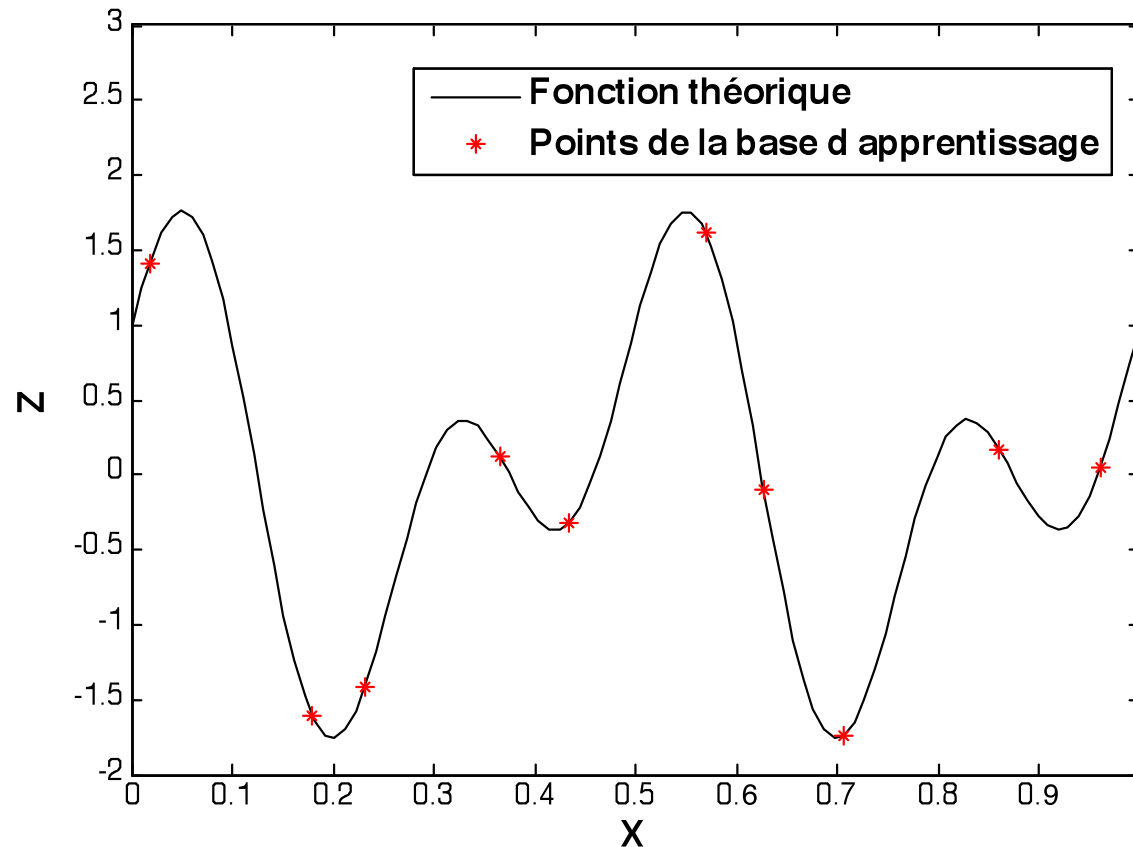
$$Q^2(Z, \hat{Z}) = 1 - \frac{\sum_{i=1}^N (Z_i - \hat{Z}_i)^2}{\sum_{i=1}^N (\bar{Z} - Z_i)^2} \quad \text{avec} \quad \begin{cases} Z_i : \text{valeurs observées} \\ \hat{Z}_i : \text{valeurs prédites par le métamodèle} \\ \bar{Z} : \text{moyenne des valeurs observées} \end{cases}$$

- 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.

UTILISATION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ Illustration sur des exemples analytiques

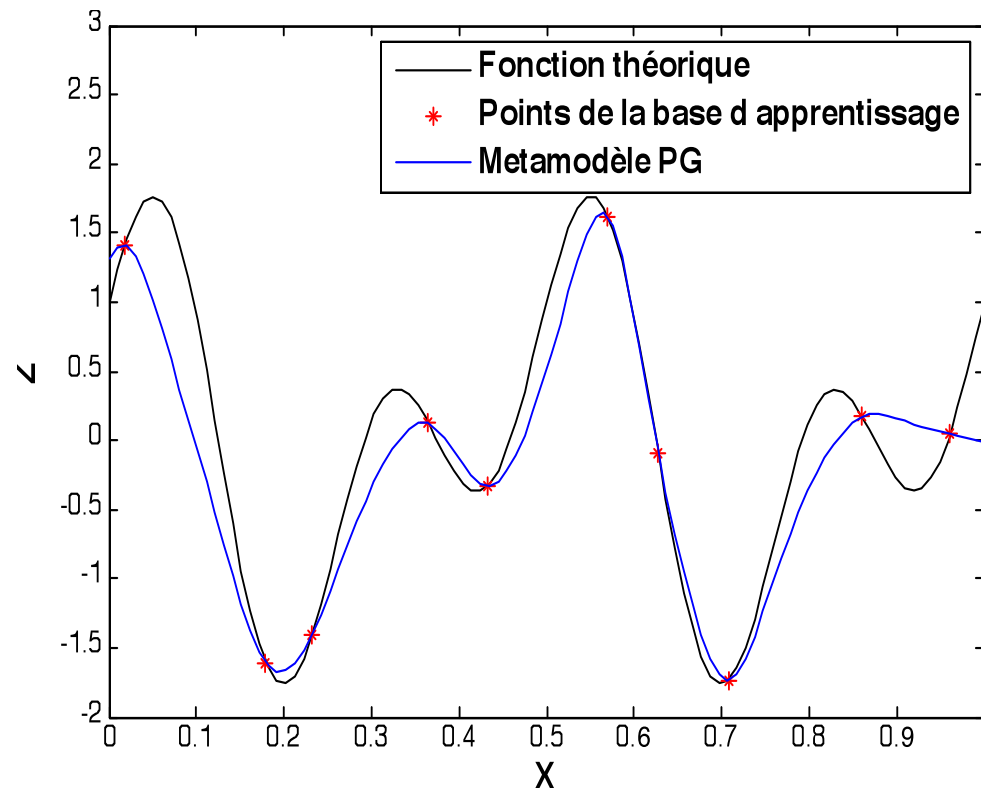
- Fonction théorique : $f(x) = \cos(4\pi X) + \sin(8\pi X)$
- Base d'apprentissage : tirage aléatoire de 10 points



UTILISATION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ Illustration sur des exemples analytiques

- Fonction théorique : $f(x) = \cos(4\pi X) + \sin(8\pi X)$
- Ajustement d'un métamodèle PG avec covariance Matern 5/2



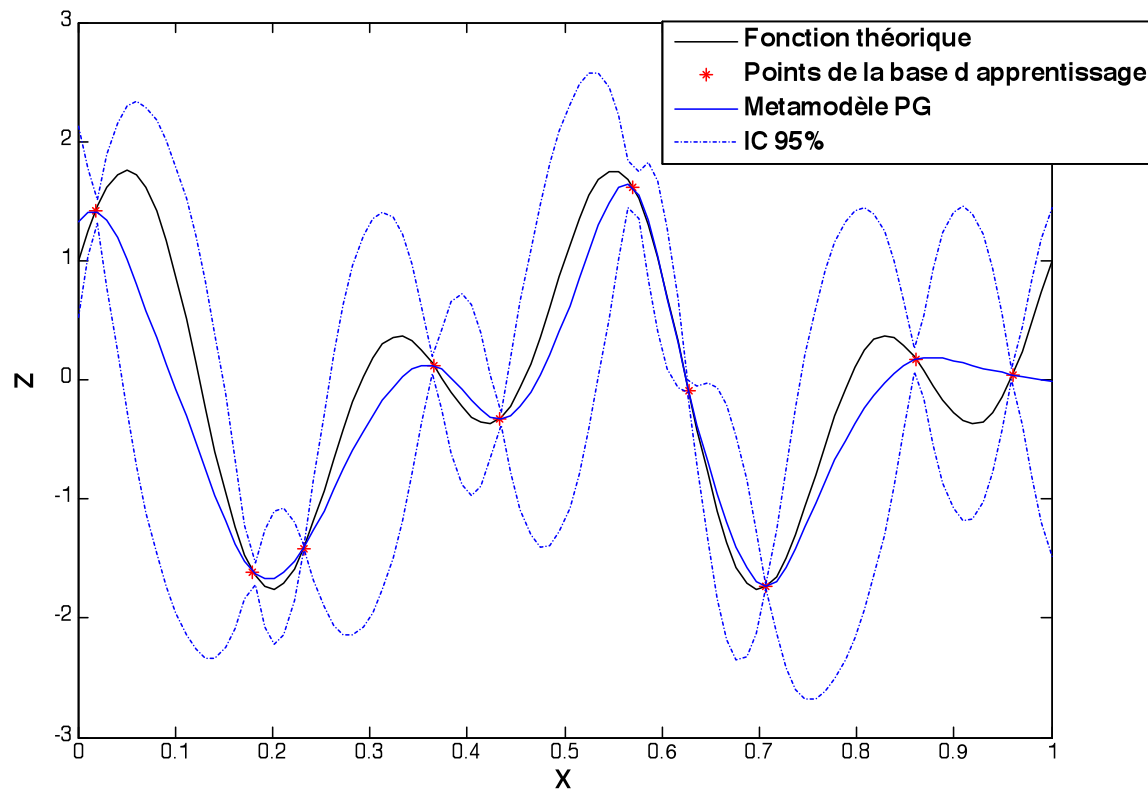
Calcul du Q^2 sur une base de test de 100 points :

$$Q^2 = 0.85$$

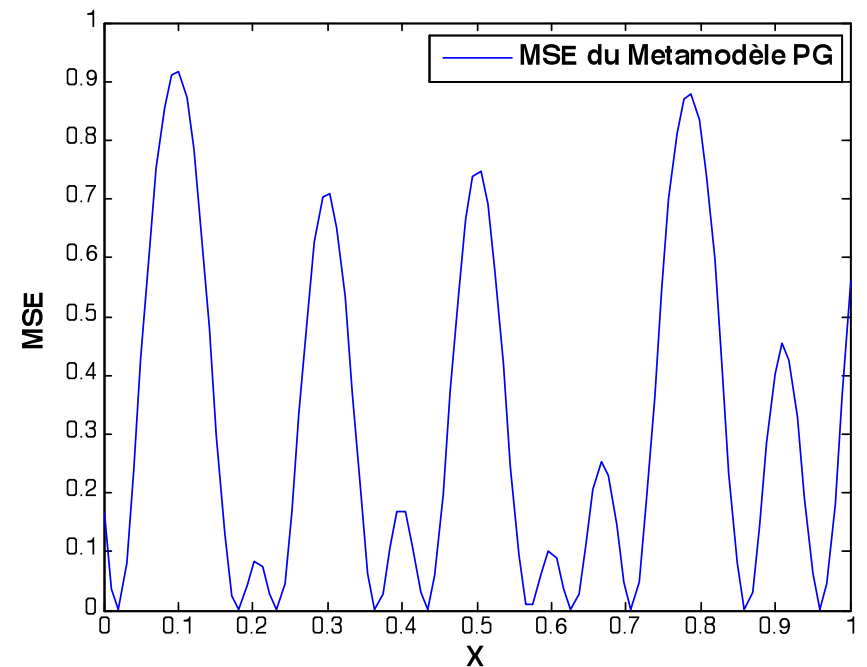
UTILISATION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ Illustration sur des exemples analytiques

- Fonction théorique : $f(x) = \cos(4\pi X) + \sin(8\pi X)$
- Utilisation du MSE du métamodèle PG



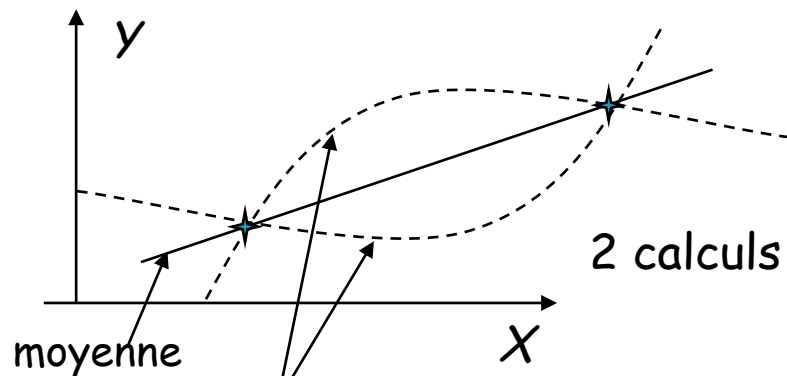
Construction d'intervalles de confiance pour les prédictions



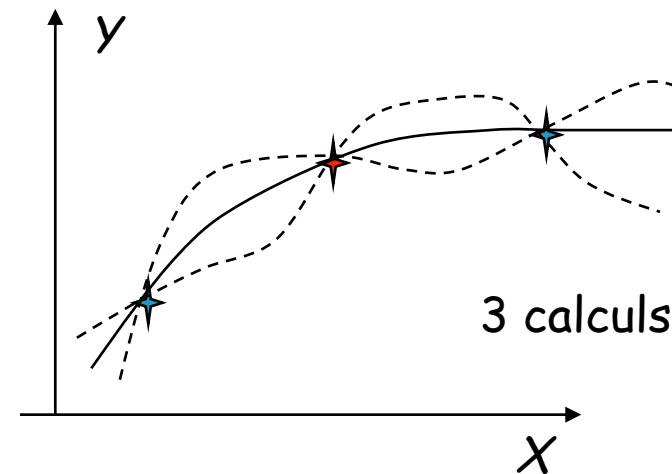
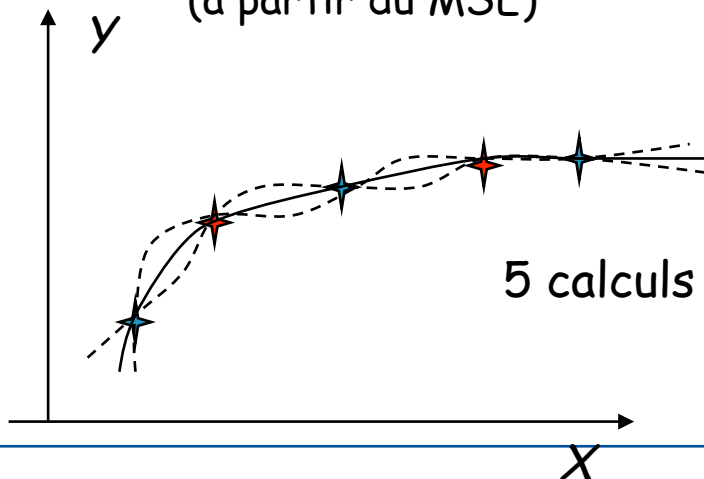
UTILISATION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ Illustration sur des exemples analytiques

- Fonction théorique : $f(x) = \cos(4\pi X) + \sin(8\pi X)$
- Utilisation du MSE du métamodèle PG



Intervalle de confiance à 95%
(à partir du MSE)



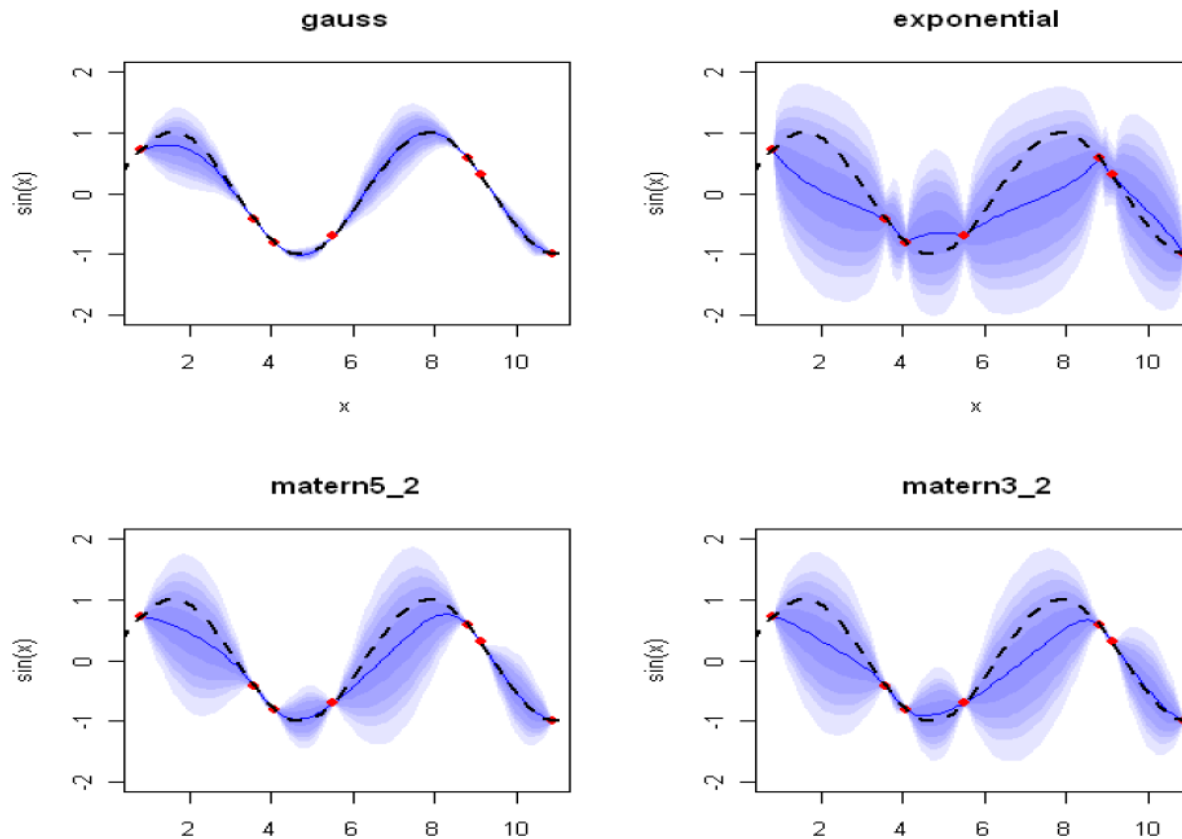
Conclusion : à partir d'un certain nb de points, le métamodèle devient précis

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

UTILISATION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ Illustration sur des exemples analytiques

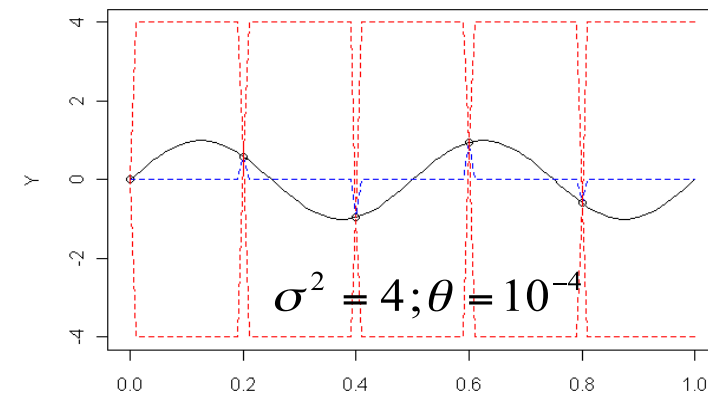
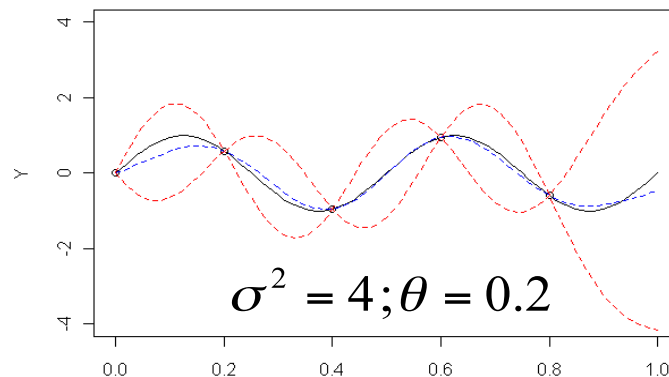
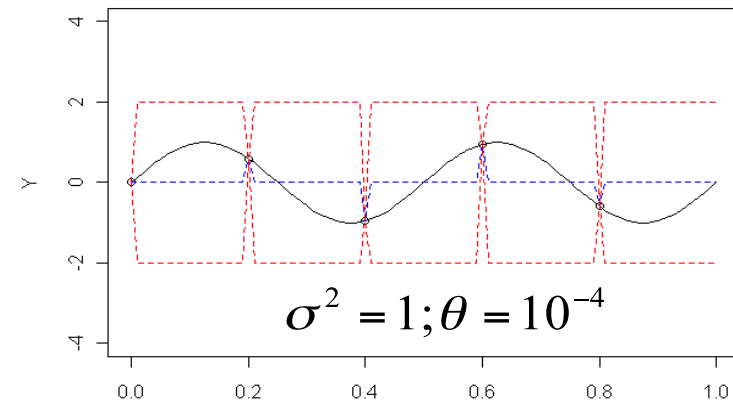
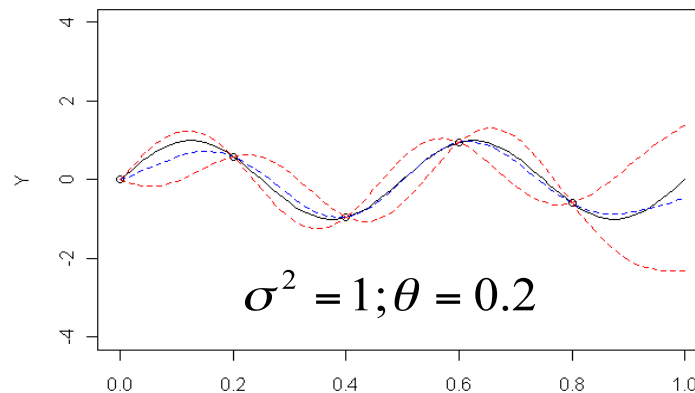
- Fonction théorique : $Y = \sin(X)$
 - Base d'apprentissage : 7 points tirés aléatoirement
- Impact du choix de la fonction de covariance



UTILISATION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ Illustration sur des exemples analytiques

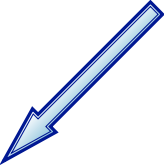
- Fonction théorique : $Y = \sin(4\pi X)$
 - Base d'apprentissage : 5 points équirépartis
- Impact des hyperparamètres



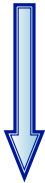
UTILISATION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ Avantages du métamodèle PG

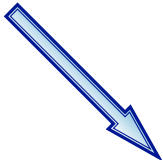
- Interpolateur exact des points de la base d'apprentissage (relaxation possible si ajout d'un effet de pépité)
- Possibilité de « jouer » sur la forme de la fonction moyenne
 - => moyenne du processus devient une fonction de x
 - => prise en compte d'avis d'expert
- Possibilité d'estimer les paramètres 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

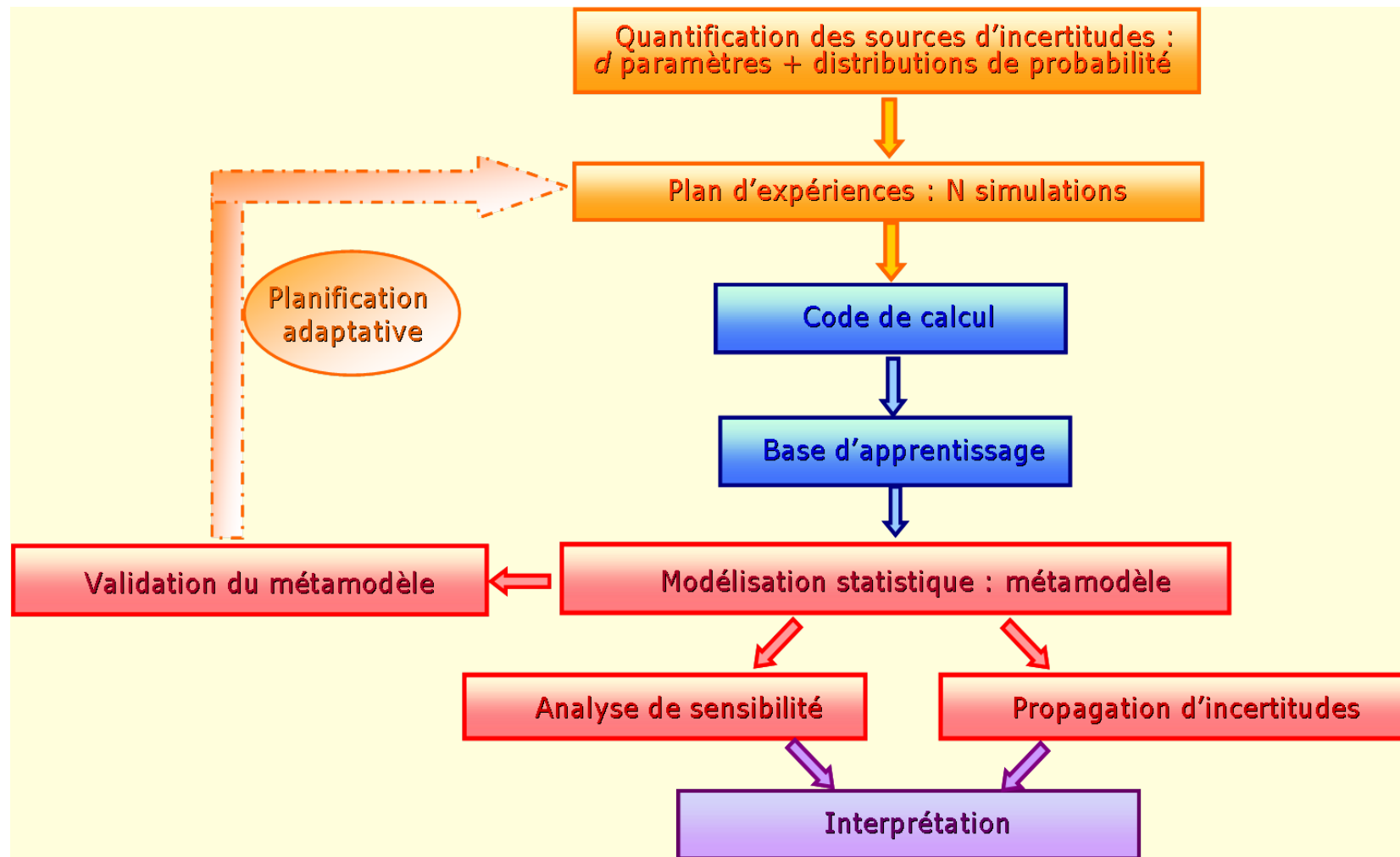
UTILISATION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ 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...

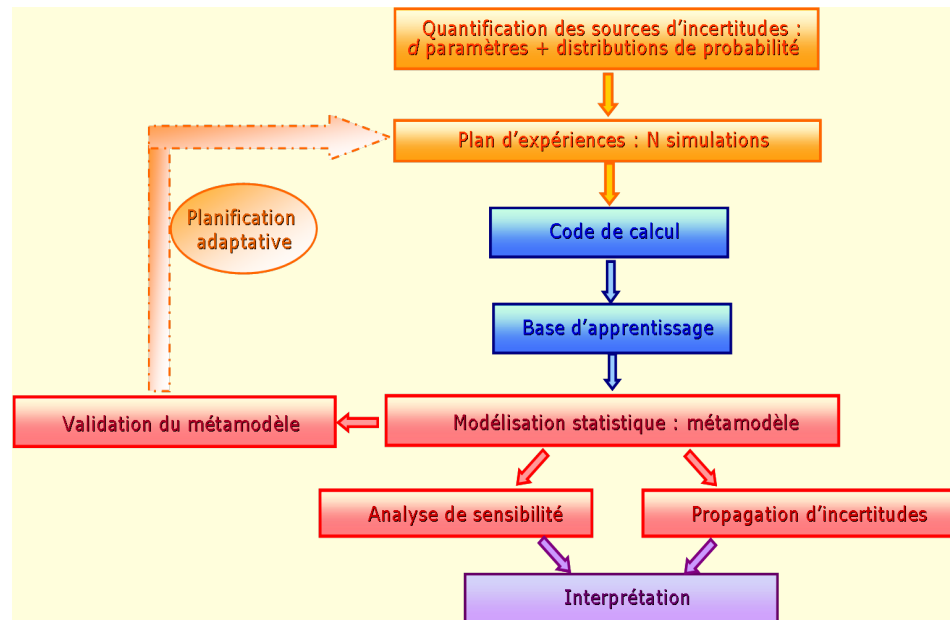
UTILISATION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ Utilisation du métamodèle PG dans une démarche «incertitudes»



UTILISATION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ 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**

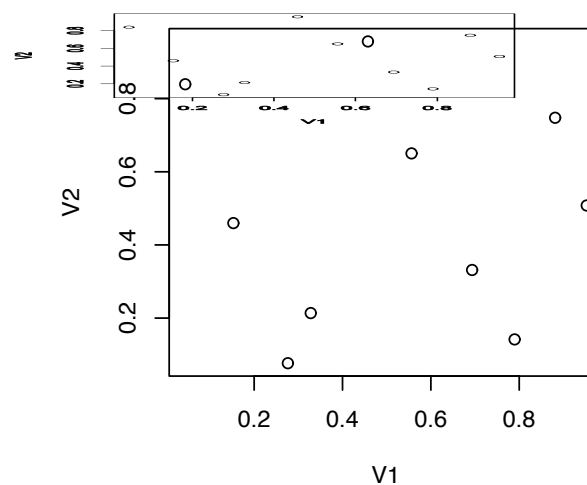
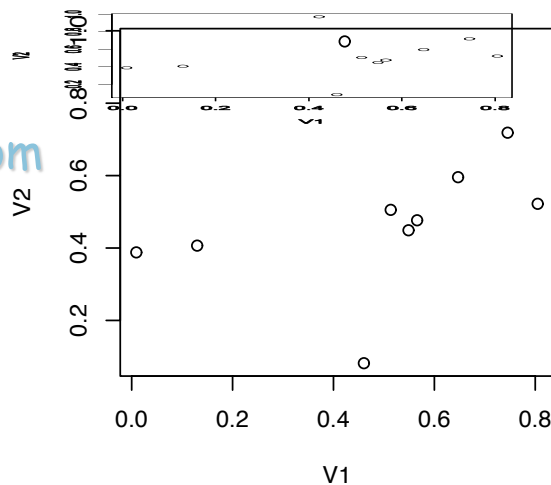
Nota Bene : l'un des intérêts importants du métamodèle est de donner la possibilité d'étudier l'impact du choix de la distribution des entrées

UTILISATION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

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

- Métamodèle PG adapté :
 - au cas des modèles numériques : expériences « déterministes »
 - quand quelques dizaines d'entrée et quelques centaines de simulations possibles = points d'expériences dispersés = « 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



Exemple : plans LHS
optimisés

ANNEXE 1 : Planification adaptative avec métamodèle PG

UTILISATION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ 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,...

UTILISATION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

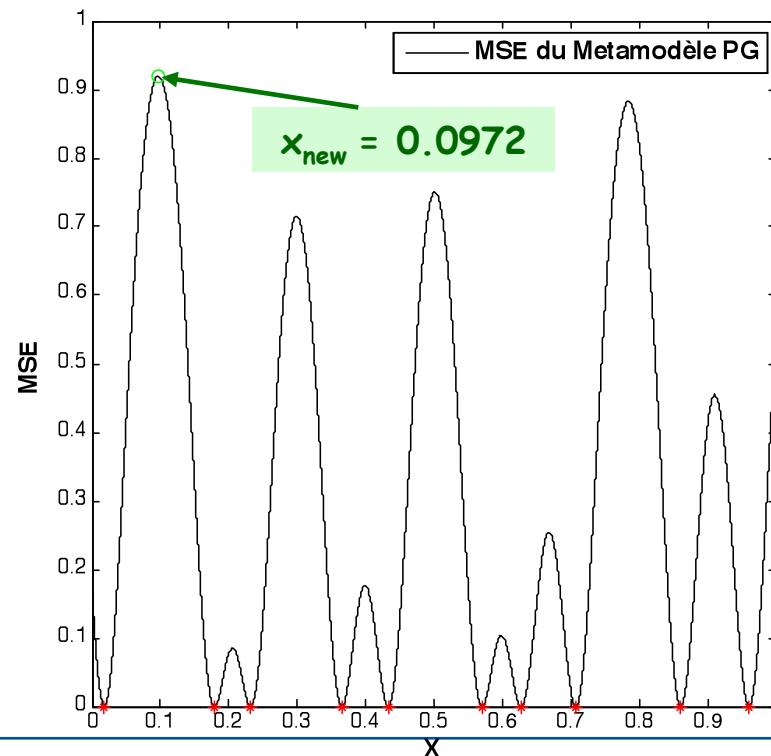
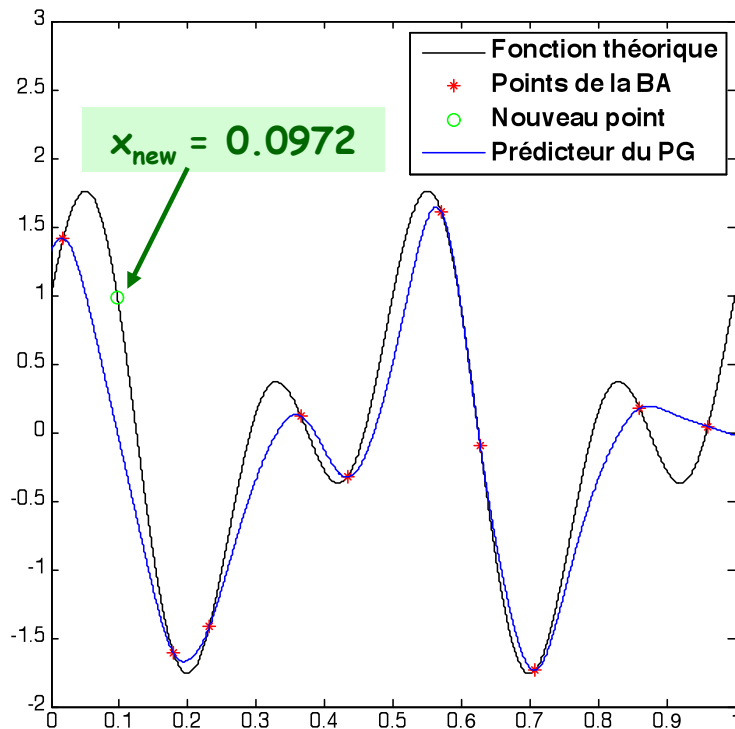
→ Planification adaptative avec le métamodèle PG

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

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

$$x_{new} = \underset{x}{\operatorname{Argmax}} MSE[\hat{Z}(x)] = \underset{x}{\operatorname{Argmax}} (\sigma^2 - k(x)^T \Sigma_S^{-1} k(x))$$

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

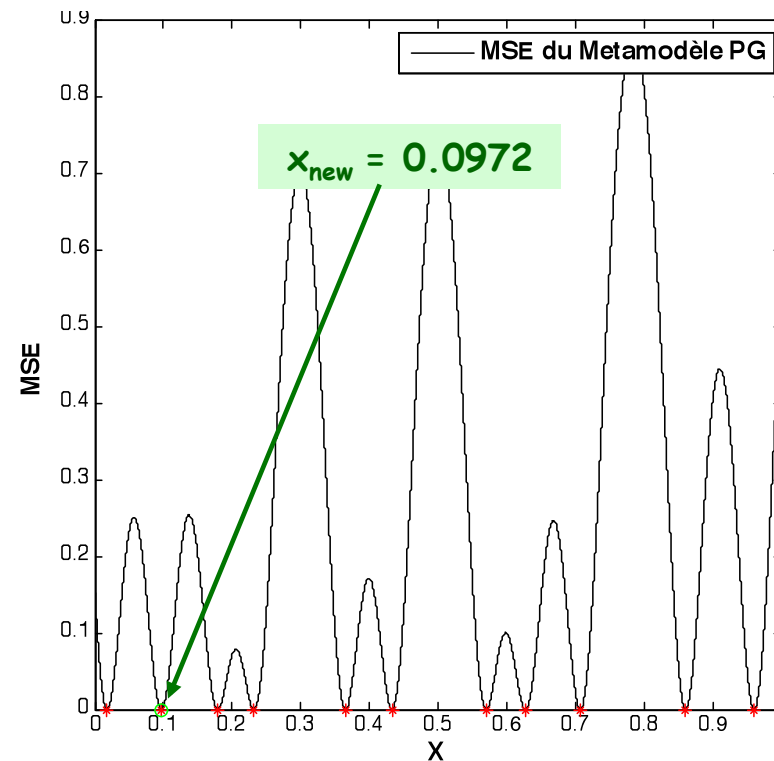
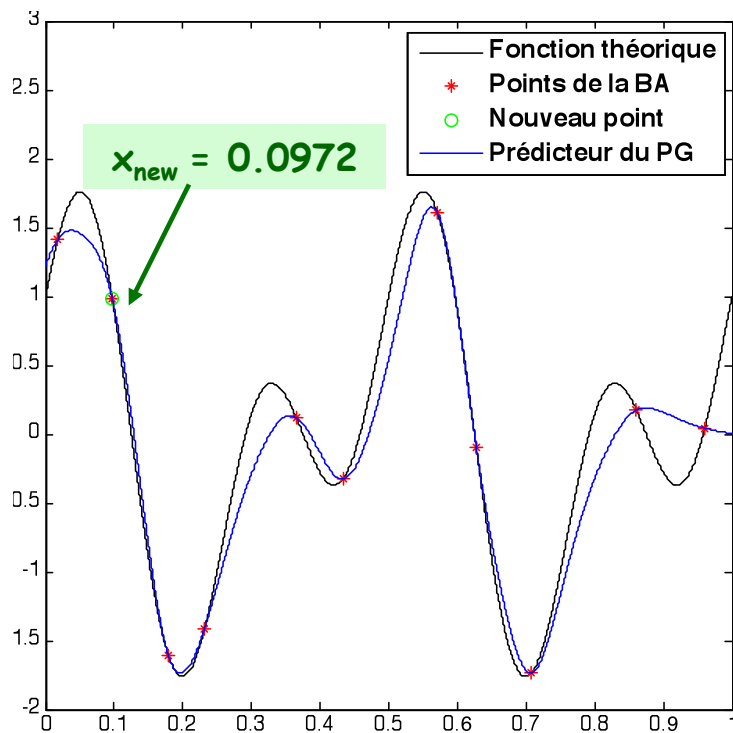


UTILISATION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ 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



UTILISATION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

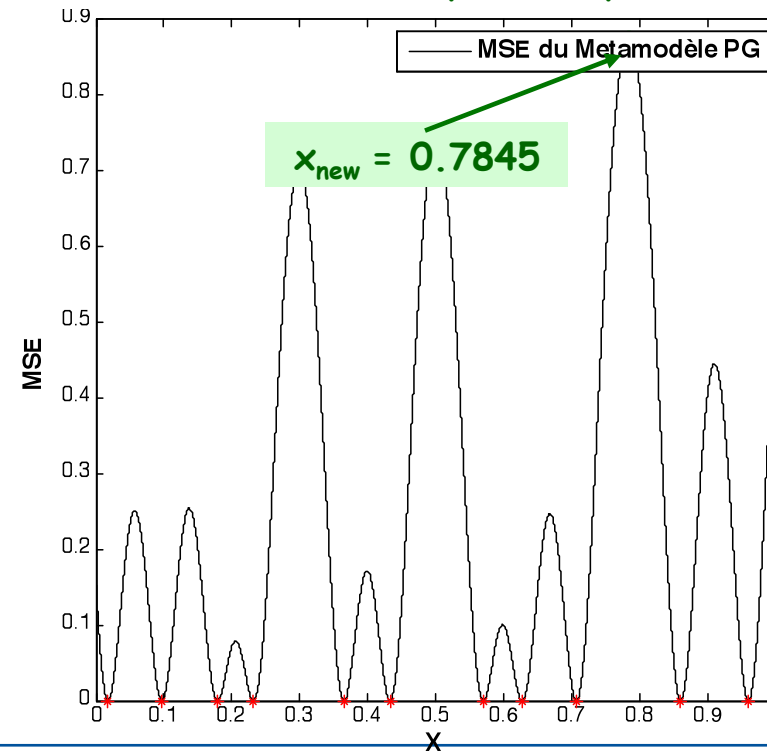
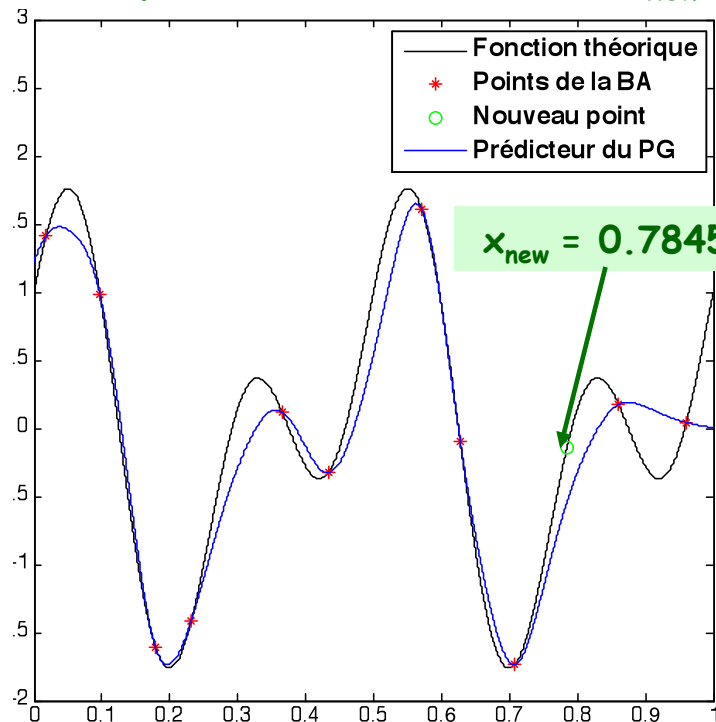
→ Planification adaptative avec le métamodèle PG

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

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

$$x_{new} = \underset{x}{\operatorname{Argmax}} MSE[\hat{Z}(x)] = \underset{x}{\operatorname{Argmax}} (\sigma^2 - k(x)^T \Sigma_S^{-1} k(x))$$

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

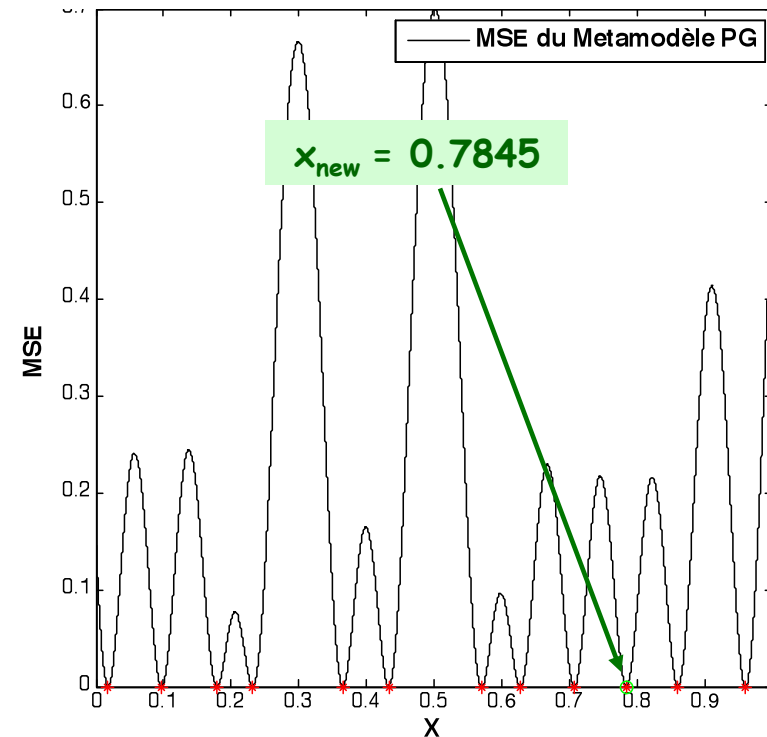
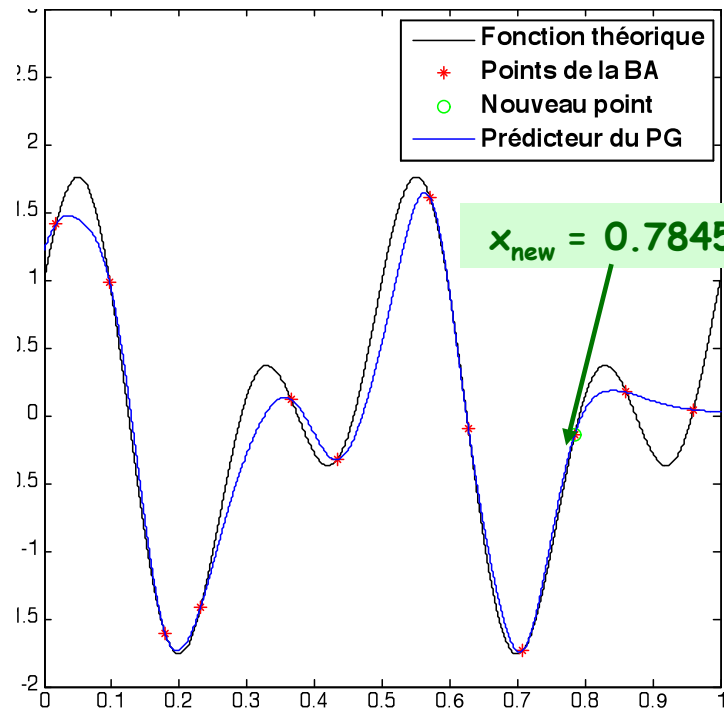


UTILISATION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ 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

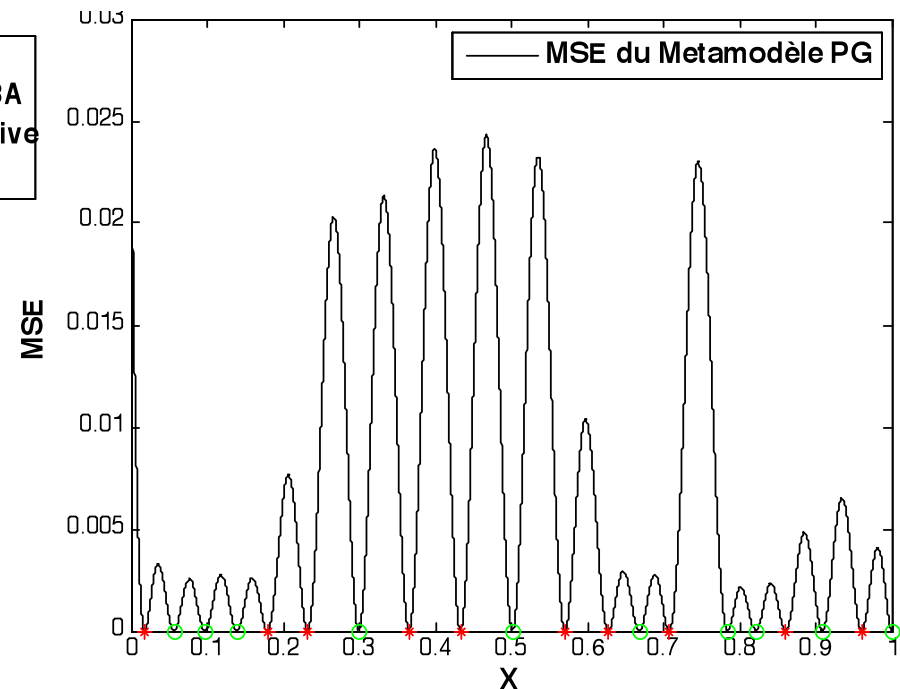
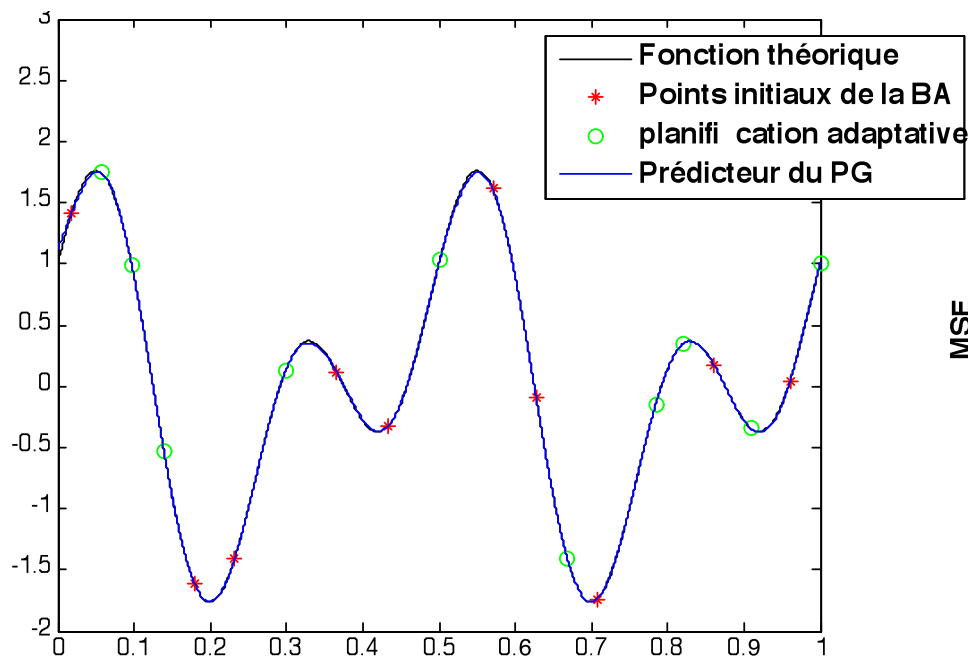


UTILISATION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ Planification adaptative avec le métamodèle PG

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

Ajout de 10 points par planification adaptative



UTILISATION DU MÉTAMODÈLE PROCESSUS GAUSSIEN

→ Planification adaptative avec le métamodèle PG

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

Ajout de 10 points par planification adaptative

