

# Module #14

## Métamodèles, modèles empiriques & processus Gaussiens

(CIV8530 - Fiabilité des structures et systèmes)

Enseignant: James-A. Goulet

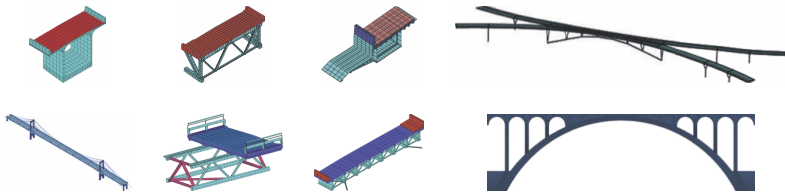
Département des génies civil, géologique et des mines  
Polytechnique Montréal



Section 8.2 – Goulet (2020)  
*Probabilistic Machine Learning for Civil Engineers*  
MIT Press

# Introduction aux métamodèles

Chaque évaluation de mon modèle me prend **>1 h... 24h...!?**



$$y = f(\mathbf{x}), \quad \mathbf{x} = [x_1, x_2, \dots, x_N]^T$$

- ▶  $f(\mathbf{x})$ : Modèle prenant comme données d'entrées  $\mathbf{x}$
- ▶  $\mathbf{x}$ : Vecteur d'attributs et de paramètres du modèle (e.g. conditions d'appuis, propriétés des matériaux, etc...)
- ▶  $y$ : Réponse du modèle (e.g. déplacement, déformation spécifique, etc...)

# Plan du module #14

---

## Introduction

## Processus Gaussien (GP)

## Métamodèles

## Modèles empiriques

## Résumé

---

### Organisation de la matière

<i>Théorie probabilité</i>	{	0	Révision algèbre & probabilités	
		1	Lois de probabilités	
		3	PDF multivariés	
<i>Fiabilité des structures &amp; systèmes</i>	{	Introduction		
		2	Formulation fiabilité composantes	$g_1(X)$
		9	Formulation fiabilité systèmes	
<i>Estimation – <math>p_f</math> échantillonnage</i>	{	4	Monte Carlo	
		12	Échantillonnage par importance	
<i>Estimation – <math>p_f</math> analytique</i>	{	5	FOSM	
		6	FORM	
		8	SORM	
<i>Utilisation des résultats</i>	{	7	Analyse de sensibilité	
		10	Incertitudes aléatoires & épistémiques	
		11	Données empiriques	
		13	Prise de décision	
		14	Métamodèles & modèles empiriques	

 Plan de la section

---

## Processus Gaussien (GP)

- 2.1 Définition
  - 2.2 Fonction de corrélation
  - 2.3 Mise à jour du GP avec des observations exactes
  - 2.4 Mise à jour du GP avec des observations inexactes
  - 2.5 Attributs multivariés:  $\mathbf{x} = [x_1, x_2, \dots, x_N]^T$ , où  $N > 1$
  - 2.6 Identification des paramètres
-



## Définition

Soit un **phénomène physique observé ou simulé** tel que

$$\underbrace{y_i}_{\text{observation/simulation}} = \underbrace{f(\mathbf{x}_i)}_{\text{modèle/réalité}}, \quad \underbrace{\mathbf{x}_i}_{\text{attributs}} = [x_1, x_2, \dots, x_N]_i^T$$

pour  $i = 1, \dots, M$

$$\mathbf{y} = [y_1, y_2, \dots, y_M]^T, \quad \mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M]_{N \times M}$$

**Processus Gaussien:**  $\mathbf{Y} = f(\mathbf{x}) \sim \mathcal{N}(\mathbf{M}_Y, \Sigma_{YY})$

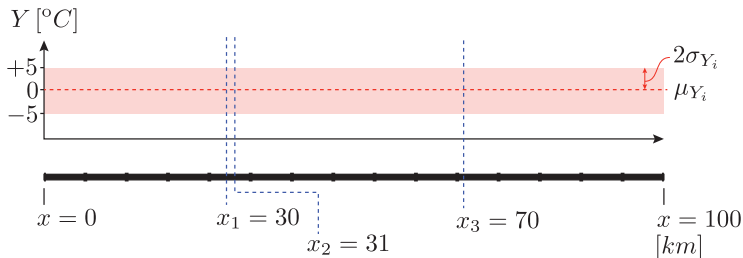
Ensemble discret de variables aléatoires Gaussienne pour lesquelles la corrélation entre deux variables  $Y_i$  et  $Y_j$  est fonction de la distance entre les attributs

$$[\Sigma_{YY}]_{ij} = \rho(x_i, x_j) \sigma_{Y_i}, \sigma_{Y_j}, \quad \rho(x_i, x_j) = \text{fct}(|x_i - x_j|)$$

## Exemple - Distribution de $T^\circ$ le long d'un oléoduc

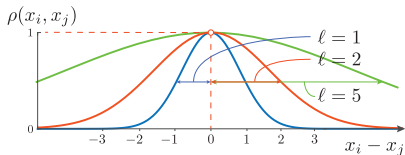


## Exemple - Distribution de T° le long d'un oléoduc



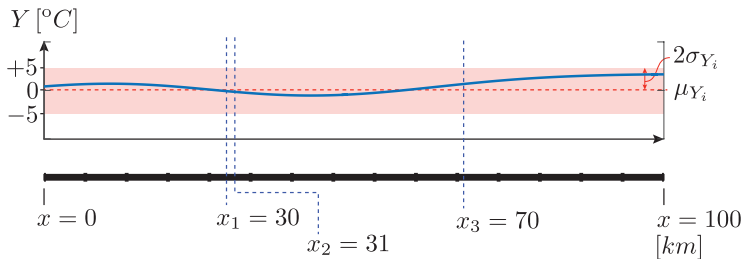
Connaissance à priori:  $\mu_{Y_i} = \mu_Y = 0^\circ\text{C}$ ,  $\sigma_{Y_i} = \sigma_Y = 2.5^\circ\text{C}$

Fonction de corrélation Gaussienne:  $\rho(x_i, x_j) = \exp\left(-\frac{1}{2} \frac{(x_i - x_j)^2}{\ell^2}\right)$



$$\text{pour } \ell = 25 \text{ km} \begin{cases} \rho(x_1, x_2) = 0.999 \\ \rho(x_1, x_3) = 0.278 \\ \rho(x_2, x_3) = 0.296 \end{cases}$$

## Exemple - Distribution de $T^\circ$ le long d'un oléoduc [ ]

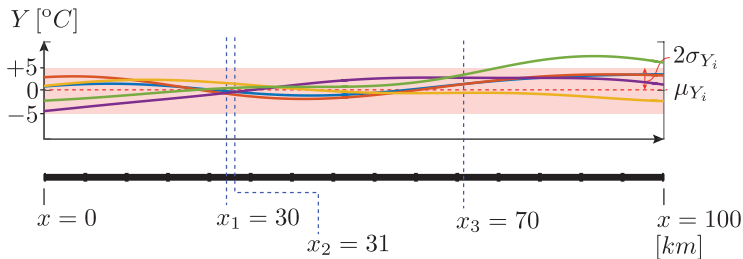


Pour  $\mathbf{x} = [0, 1, \dots, 100]^\top$

$$\overbrace{\mathbf{Y} \sim \mathcal{N}(\mathbf{0}, \Sigma_{\mathbf{Y}\mathbf{Y}}), [\Sigma_{\mathbf{Y}\mathbf{Y}}]_{ij} = \rho(x_i, x_j)\sigma_{Y_i}^2}^{\text{Connaissances à priori}}$$

$\mathbf{y}_i$ : réalisations de  $\mathbf{Y}$  représentant notre connaissance à priori

## Exemple - Distribution de $T^\circ$ le long d'un oléoduc [ ]



Pour  $\mathbf{x} = [0, 1, \dots, 100]^\top$

$$\overbrace{\mathbf{Y} \sim \mathcal{N}(\mathbf{0}, \Sigma_{\mathbf{Y}\mathbf{Y}})}^{\text{Connaissances à priori}}, \quad [\Sigma_{\mathbf{Y}\mathbf{Y}}]_{ij} = \rho(x_i, x_j)\sigma_{Y_i}^2$$

$\mathbf{y}_i$ : réalisations de  $\mathbf{Y}$  représentant notre connaissance à priori

## Mise à jour du GP avec des observations

Soit  $\mathcal{D} = \{(\mathbf{x}_i, \mathbf{y}_i), i = 1, \dots, M\}$  un ensemble de  $M$  observations et  $\mathbf{x}_*$  un ensemble de  $M_*$  attributs pour lesquels on désire prédire

$$\Pr(\mathbf{Y}_* = \mathbf{y} | \mathbf{X}_*, \mathcal{D})$$

**Rappel:** Les distribution conditionnelles sont normales [10]

$$\mathbf{Y} = \left\{ \begin{array}{c} \mathbf{Y} \\ \mathbf{Y}_* \end{array} \right\}, \quad \mathbf{M} = \left\{ \begin{array}{c} \mathbf{M}_Y \\ \mathbf{M}_{Y_*} \end{array} \right\}, \quad \Sigma = \left[ \begin{array}{c|c} \Sigma_{YY} & \Sigma_{Y_*} \\ \hline \Sigma_{Y_*}^T & \Sigma_{**} \end{array} \right]$$

Connaissances à priori

$$f_{\mathbf{Y}_* | \mathbf{x}_*, \mathcal{D}}(\mathbf{y}_* | \mathbf{x}_*, \mathcal{D}) = \mathcal{N}(\mathbf{M}_*, \Sigma_*)$$

$$[\Sigma_{YY}]_{ij} = \rho(x_i, x_j) \sigma_Y^2$$

$$[\Sigma_{Y_*}]_{ij} = \rho(x_i, x_{*j}) \sigma_Y^2$$

$$[\Sigma_{**}]_{ij} = \rho(x_{*i}, x_{*j}) \sigma_Y^2$$

$$\mathbf{M}_* = \mathbf{M}_{Y_*} + \Sigma_{Y_*}^T \Sigma_{YY}^{-1} (\mathbf{y} - \mathbf{M}_Y)$$

$$\Sigma_* = \Sigma_{**} - \Sigma_{Y_*}^T \Sigma_{YY}^{-1} \Sigma_{Y_*}$$

Connaissances à posteriori

## Exemple – distribution conditionnelle multivariée

Soit  $\underbrace{\mu_Y = 0^\circ\text{C}, \sigma_Y = 2.5^\circ\text{C}, \rho(x_1, x_2) = 0.8}_{\text{connaissance à priori}}, \underbrace{y_2 = -2^\circ\text{C}}_{\text{observation}}$

$$f_{Y_1 Y_2}(y_1, y_2) = \mathcal{N}(\mathbf{M}_{YY}, \Sigma_{YY}) \left\{ \begin{array}{l} \mathbf{M}_{YY} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \Sigma_{YY} = \begin{bmatrix} 2.5^2 & 0.8 \cdot 2.5^2 \\ 0.8 \cdot 2.5^2 & 2.5^2 \end{bmatrix} \end{array} \right.$$

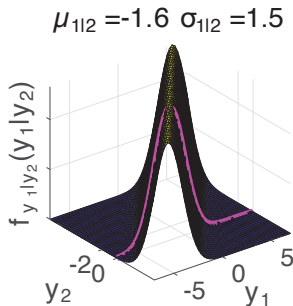
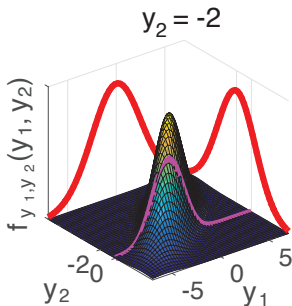
$$f_{Y_1|y_2}(y_1|y_2) = \frac{f_{Y_1 Y_2}(y_1, y_2)}{f_{Y_2}(y_2)} = \mathcal{N}(\mu_{1|2}, \sigma_{1|2}^2) \left\{ \begin{array}{l} \mu_{1|2} = \mu_1 + \rho\sigma_1 \frac{y_2 - \mu_2}{\sigma_2} \\ \sigma_{1|2} = \sigma_1 \sqrt{1 - \rho^2} \end{array} \right.$$

$$\mu_{1|2} = 0 + 0.8 \times 2.5 \frac{-2 - 0}{2.5} = -1.6^\circ\text{C}$$

$$\sigma_{1|2} = 2.5 \sqrt{1 - 0.8^2} = 1.5^\circ\text{C}$$

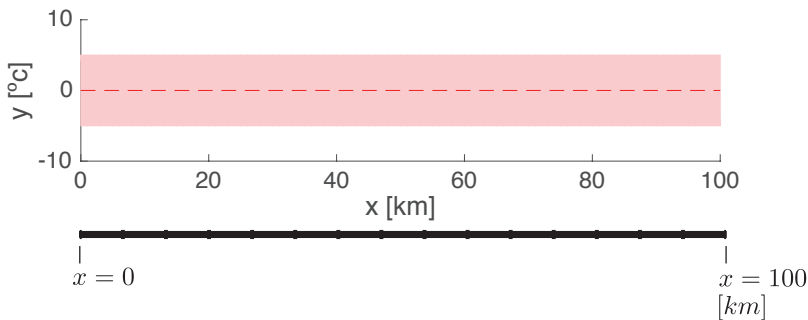
Exemple – distribution conditionnelle multivariée 

$$f_{Y_1|Y_2}(y_1|y_2) = \frac{f_{Y_1, Y_2}(y_1, y_2)}{f_{Y_2}(y_2)} = \mathcal{N}(\mu_{1|2}, \sigma_{1|2}^2) \begin{cases} \mu_{1|2} = \mu_1 + \rho\sigma_1 \frac{y_2 - \mu_2}{\sigma_2} \\ \sigma_{1|2} = \sigma_1 \sqrt{1 - \rho^2} \end{cases}$$



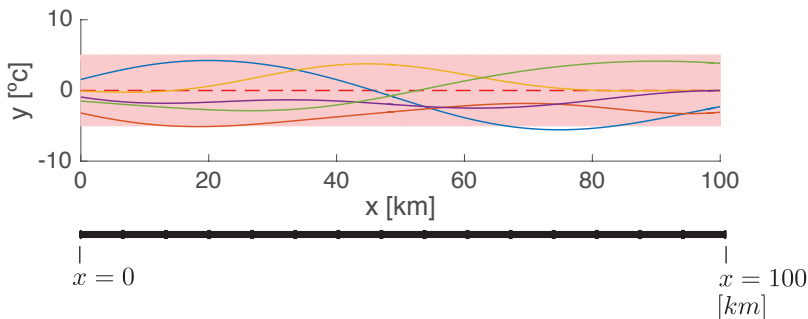


## Exemple - Distribution de $T^\circ$ le long d'un oléoduc



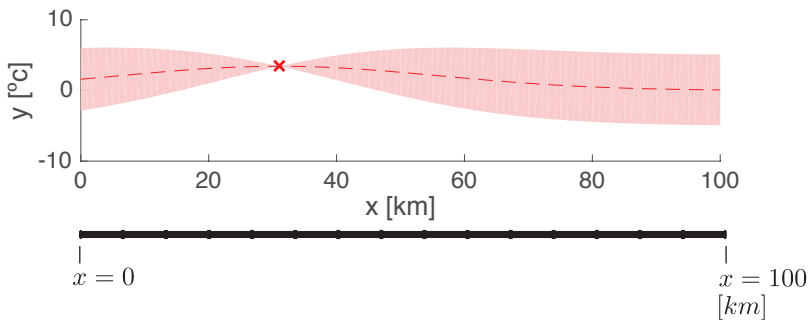
**Observations:**  $\mathcal{D} = \left\{ \left( \underbrace{[\ ]}_x, \underbrace{[\ ]}_y \right) \right\}$

## Exemple - Distribution de $T^\circ$ le long d'un oléoduc



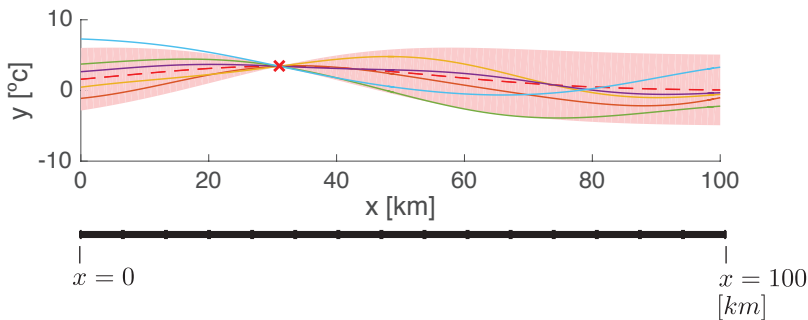
**Observations:**  $\mathcal{D} = \left\{ \left( \underbrace{[\ ]}_x, \underbrace{[\ ]}_y \right) \right\}$

## Exemple - Distribution de $T^\circ$ le long d'un oléoduc



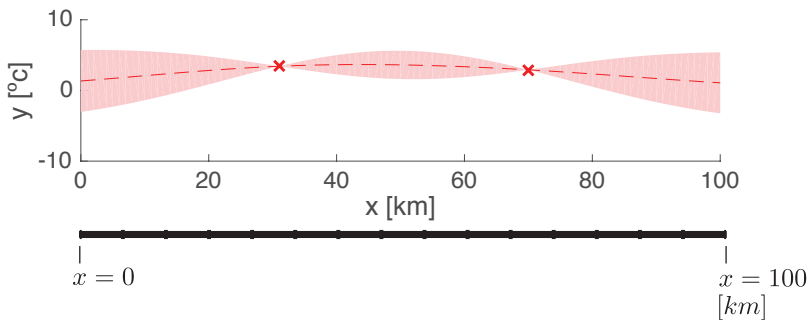
**Observations:**  $\mathcal{D} = \left\{ \left( \underbrace{[31]}_x, \underbrace{[4.13]}_y \right) \right\}$

## Exemple - Distribution de $T^\circ$ le long d'un oléoduc



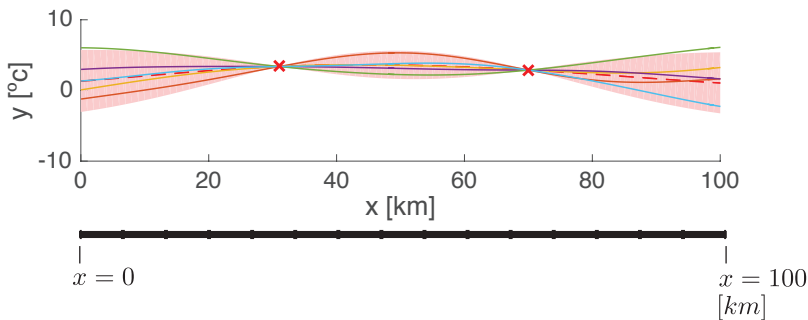
**Observations:**  $\mathcal{D} = \left\{ \left( \underbrace{[31]}_x, \underbrace{[4.13]}_y \right) \right\}$

## Exemple - Distribution de $T^\circ$ le long d'un oléoduc



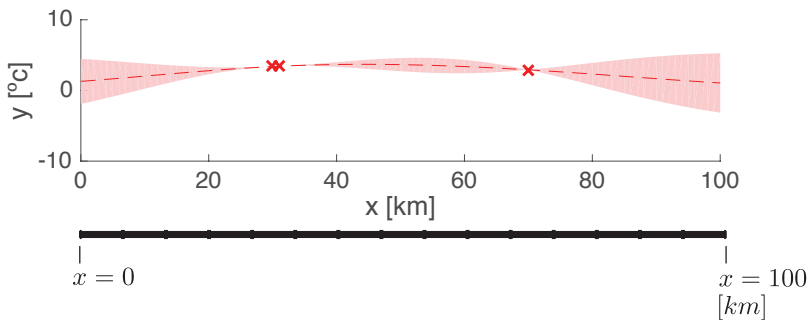
**Observations:**  $\mathcal{D} = \left\{ \left( \underbrace{[31, 70]}_x, \underbrace{[4.13, 3.62]}_y \right) \right\}$

## Exemple - Distribution de $T^\circ$ le long d'un oléoduc



**Observations:**  $\mathcal{D} = \left\{ \left( \underbrace{[31, 70]}_x, \underbrace{[4.13, 3.62]}_y \right) \right\}$

## Exemple - Distribution de $T^\circ$ le long d'un oléoduc



**Observations:**  $\mathcal{D} = \left\{ \left( \underbrace{[31, 70, 30]}_x, \underbrace{[4.13, 3.62, 3.71]}_y \right) \right\}$

## Observation inexactes

Soit un **phénomène physique** observé avec un **instrument inexact** tel que

$$Y_i = \underbrace{f(\mathbf{x}_i)}_{\text{réalité}} + \underbrace{V}_{\text{erreur de mesure}}, \quad V \sim \mathcal{N}(0, \sigma_V^2)$$

observation
erreur de mesure

$$[\Sigma_{\mathbf{Y}\mathbf{Y}}]_{ij} = \rho(x_i, x_j)\sigma_Y^2 + \sigma_V^2\delta_{ij}, \quad \delta_{ij} = 1 \text{ si } i = j$$

$$\mathbf{Y} = \left\{ \begin{array}{c} \mathbf{Y} \\ \mathbf{f}_* \end{array} \right\}, \quad \mathbf{M} = \left\{ \begin{array}{c} \mathbf{M}_{\mathbf{Y}} \\ \mathbf{M}_{\mathbf{f}_*} \end{array} \right\}, \quad \Sigma = \left[ \begin{array}{c|c} \Sigma_{\mathbf{Y}\mathbf{Y}} & \Sigma_{\mathbf{Y}\mathbf{f}_*} \\ \hline \Sigma_{\mathbf{f}_*\mathbf{Y}}^T & \Sigma_{\mathbf{f}_*\mathbf{f}_*} \end{array} \right]$$

Connaissances à priori

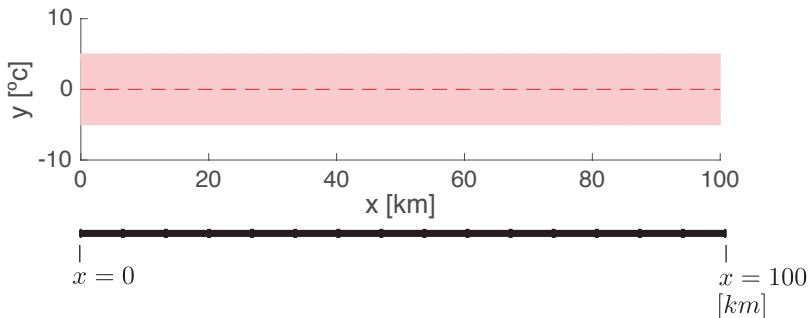
$$f_{\mathbf{Y}_*|\mathbf{x}_*, \mathcal{D}}(\mathbf{y}_*|\mathbf{x}_*, \mathcal{D}) = \mathcal{N}(\mathbf{M}_*, \Sigma_*)$$

$$\mathbf{M}_* = \mathbf{M}_{\mathbf{f}_*} + \Sigma_{\mathbf{Y}\mathbf{f}_*}^T \Sigma_{\mathbf{Y}\mathbf{Y}}^{-1}(\mathbf{y} - \mathbf{M}_{\mathbf{Y}}), \quad \Sigma_* = \Sigma_{\mathbf{f}_*\mathbf{f}_*} - \Sigma_{\mathbf{Y}\mathbf{f}_*}^T \Sigma_{\mathbf{Y}\mathbf{Y}}^{-1} \Sigma_{\mathbf{Y}\mathbf{f}_*}$$

Connaissances à posteriori



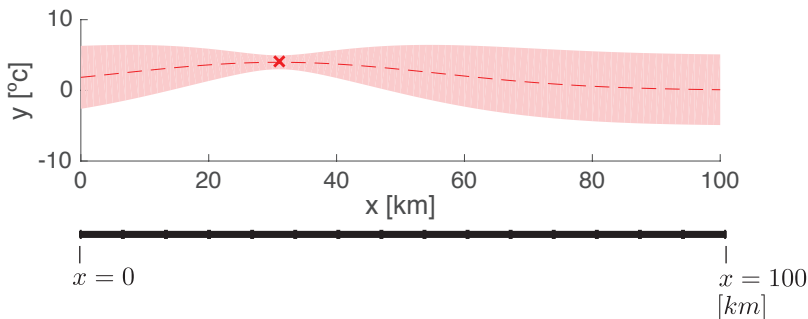
## Exemple - Distribution de $T^\circ$ ( $V \neq 0$ )



**Erreur de mesure:**  $\sigma_V = 0.5$

**Observations:**  $\mathcal{D} = \left\{ \left( \underbrace{[\ ]}_x, \underbrace{[\ ]}_y \right) \right\}$

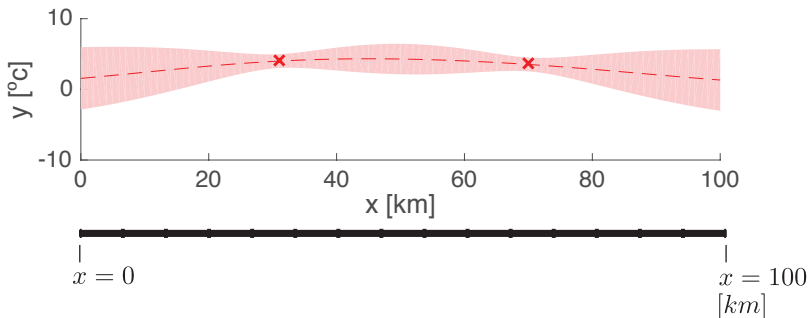
## Exemple - Distribution de $T^\circ$ ( $V \neq 0$ )



**Erreur de mesure:**  $\sigma_V = 0.5$

**Observations:**  $\mathcal{D} = \left\{ \left( \underbrace{[31]}_x, \underbrace{[3.42]}_y \right) \right\}$

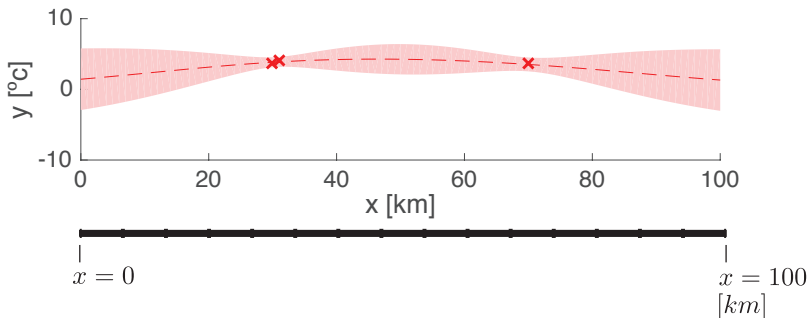
## Exemple - Distribution de $T^\circ$ ( $V \neq 0$ )



**Erreur de mesure:**  $\sigma_V = 0.5$

**Observations:**  $\mathcal{D} = \left\{ \left( \underbrace{[31, 70]}_x, \underbrace{[3.42, 2.92]}_y \right) \right\}$

## Exemple - Distribution de $T^\circ$ ( $V \neq 0$ )



**Erreur de mesure:**  $\sigma_V = 0.5$

**Observations:**  $\mathcal{D} = \left\{ \left( \underbrace{[31, 70, 30]}_x, \underbrace{[3.42, 2.92, 3.38]}_y \right) \right\}$

Attributs multivariés:  $\mathbf{x} = [x_1, x_2, \dots, x_N]^T$ , où  $N > 1$

Attributs multivariés:  $\mathbf{x} = [x_1, x_2, \dots, x_N]^T$ , où  $N > 1$

$$\underbrace{y_i}_{\text{observation/simulation}} = \underbrace{f}_{\text{modèle/réalité}}(\mathbf{x}_i), \quad \underbrace{\mathbf{x}_i}_{\text{attributs}} = [x_1, x_2, \dots, x_N]^T$$

pour  $i = 1, \dots, M$

$$\mathbf{y} = [y_1, y_2, \dots, y_M]^T, \quad \mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M]_{N \times M}, \quad \underbrace{\boldsymbol{\ell} = [\ell_1, \ell_2, \dots, \ell_N]_{N \times 1}}_{\text{un } \ell_k \text{ par attribut } \mathbf{x}_k}$$

$$\begin{aligned} \rho(\mathbf{x}_i, \mathbf{x}_j) &= \exp\left(-\frac{1}{2} \sum_{k=1}^N \frac{([x_k]_i - [x_k]_j)^2}{\ell_k^2}\right) \\ &= \exp\left(-\frac{1}{2} (\mathbf{x}_i - \mathbf{x}_j)^T \text{diag}(\ell^2)^{-1} (\mathbf{x}_i - \mathbf{x}_j)\right) \end{aligned}$$

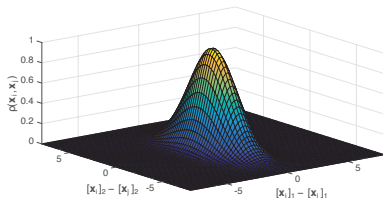
Attributs multivariés:  $\mathbf{x} = [x_1, x_2, \dots, x_N]^T$ , où  $N > 1$

## Fonction de corrélation multivariée

$$\begin{aligned} \rho(\mathbf{x}_i, \mathbf{x}_j) &= \exp\left(-\frac{1}{2} \sum_{k=1}^N \frac{([\mathbf{x}_i]_k - [\mathbf{x}_j]_k)^2}{\ell_k^2}\right) \\ &= \exp\left(-\frac{1}{2} (\mathbf{x}_i - \mathbf{x}_j)^T \text{diag}(\ell)^{-2} (\mathbf{x}_i - \mathbf{x}_j)\right) \end{aligned}$$

si  $\ell_k \gg \sigma_{X_k}$  :  $\text{Imp}(X_k \rightarrow y) \approx 0$

si  $\ell_k < \sigma_{X_k}$  :  $\text{Imp}(X_k \rightarrow y) \gg 0$



## Identification des paramètres $\sigma_Y$ & $\ell$

Jusqu'à maintenant, on a fait l'hypothèse que l'on connaît  $\sigma_Y$  et  $\ell = [\ell_1, \ell_2, \dots, \ell_N]^T$ .

En pratique, on doit apprendre les valeurs des **hyperparamètres**  $\theta = [\sigma_Y, \ell]^T$  à partir des données  $\mathcal{D} = \{(\mathbf{x}_i, \mathbf{y}_i), i = 1, \dots, M\}$ .

**Loi de Bayes:**

$$\underbrace{p(\theta | \underbrace{\mathcal{D}}_{\mathbf{x}, \mathbf{y}})}_{\text{posterior}} = \frac{\overbrace{p(\mathbf{y} | \mathbf{x}, \theta)}^{\text{likelihood}} \cdot \overbrace{p(\theta)}^{\text{prior}}}{\underbrace{p(\mathbf{y})}_{\text{cte.}}}$$

**⚠ En pratique, l'évaluation de  $p(\theta | \mathbf{x}, \mathbf{y})$  est difficile...**

# Approximation MLE

$$\underbrace{p(\theta | \mathbf{x}, \mathbf{y})}_{\text{posterior}} = \frac{\overbrace{p(\mathbf{y} | \mathbf{x}, \theta)}^{\text{likelihood}} \cdot \overbrace{p(\theta)}^{\text{prior}}}{\underbrace{p(\mathbf{y})}_{\text{cte.}}}$$

Si  $p(\theta) = \text{cte.}$ ,  $\forall \theta \rightarrow p(\theta | \mathbf{x}, \mathbf{y}) \propto p(\mathbf{y} | \mathbf{x}, \theta) \equiv p(\mathbf{y} | \mathbf{x})$

$$\theta^* = \arg \max_{\theta} p(\mathbf{y} | \mathbf{x}, \theta)$$

**Observation exactes / simulations:**

$$p(\mathbf{y} | \mathbf{x}, \theta) = p(\mathbf{y} | \mathbf{x}) = \mathcal{N}(\mathbf{0}, \Sigma_{\mathbf{Y}\mathbf{Y}})$$

**Observation inexactes:**

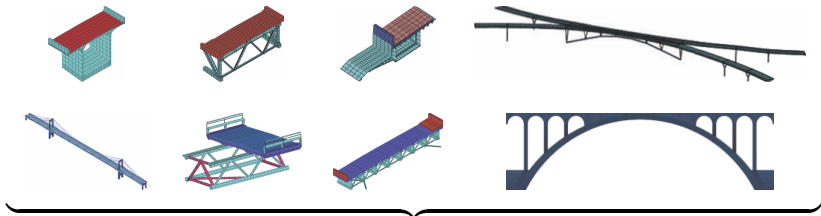
$$p(\mathbf{y} | \mathbf{x}, \theta) = p(\mathbf{y} | \mathbf{x}) = \mathcal{N}(\mathbf{0}, \Sigma_{\mathbf{Y}\mathbf{Y}} + \sigma_V^2 \cdot \mathbf{I})$$







# GP & métamodèles



$$y = f(\mathbf{x}), \quad \mathbf{x} = [x_1, x_2, \dots, x_N]^T$$

- ▶  $f(\mathbf{x})$ : Modèle prenant comme données d'entrées  $\mathbf{x}$
- ▶  $\mathbf{x}$ : Vecteur d'attributs et de paramètres du modèle (e.g. conditions d'appuis, propriétés des matériaux, etc...)
- ▶  $y$ : Réponse du modèle (e.g. déplacement, déformation spécifique, etc...)



## Considérations pratiques (cont.)

**Simulations à générer:** Générer un nombre de simulations  $M$  aussi grand que possible à partir du PDF à priori des attributs  $\mathbf{X} \sim f_{\mathbf{X}}(\mathbf{x})$

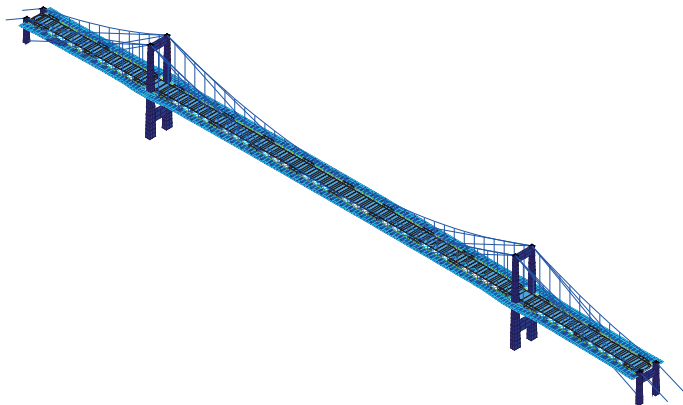
$$f_{\mathbf{Y}_*|\mathbf{x}_*\mathcal{D}}(\mathbf{y}_*|\mathbf{x}_*, \mathcal{D}) = \mathcal{N}(\mathbf{M}_*, \mathbf{\Sigma}_*) \rightarrow \{\text{MC, FORM}, \dots\}$$

## Considérations pratiques (cont.)

**Simulations à générer:** Générer un nombre de simulations  $M$  aussi grand que possible à partir du PDF à priori des attributs  $\mathbf{X} \sim f_{\mathbf{X}}(\mathbf{x})$

$$f_{\mathbf{Y}_*|\mathbf{x}_*,\mathcal{D}}(\mathbf{y}_*|\mathbf{x}_*,\mathcal{D}) = \mathcal{N}(\mathbf{M}_*, \mathbf{\Sigma}_*) \rightarrow \{\text{MC, FORM}, \dots\}$$

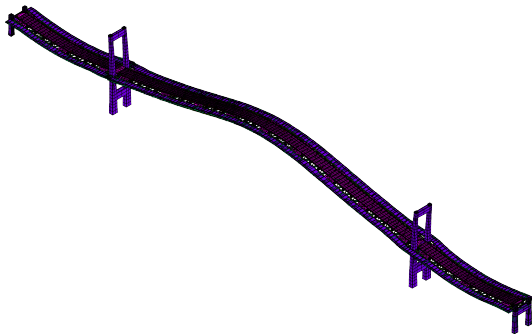
# Exemple – Metamodèle du Tamar Bridge



## Metamodèle du Tamar Bridge – Réponses $y_i$

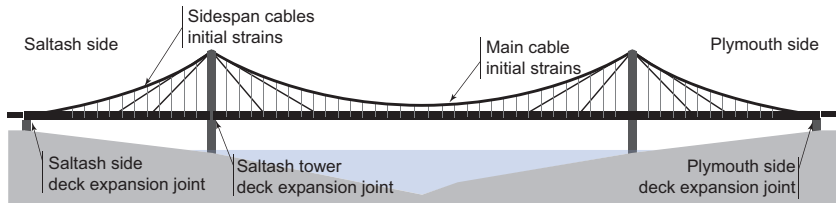
Nous désirons modéliser la **première fréquence propre** du pont

Nous avons un ensemble de **1000 simulations du modèles** pour lesquels les paramètres ont été échantillonnés aléatoirement





# Metamodelle du Tamar Bridge – Variables $\mathbf{x}_i = [x_1 : x_5]^T$



1. Main-cable initial strain:  $(5E-4, 3E-3)$  mm/mm
2. Sidespan cable initial strains  $(5E-4, 3E-3)$  mm/mm
3. Plymouth-side support long. stiff.  $10^{(4,11)}$  kN/mm
4. Saltash-side support long. stiff.  $10^{(4,11)}$  kN/mm
5. Saltash tower deck expansion joint long. stiff.  $10^{(4,11)}$  kN/mm

# Metamodèle du Tamar Bridge – Configuration du problème

## Variables et observations:

$$\boldsymbol{\theta} = \{l_1, l_2, l_3, l_4, l_5, \sigma_g\}$$

$$\mathcal{D} = \{\mathcal{D}_x, \mathcal{D}_y\} = \{(\mathbf{x}_i, y_i), i = 1 : 1000\}$$

## Modèle d'observation:

$$\underbrace{Y_i}_{\text{observation}} = \underbrace{g(\mathbf{x}_i)}_{\text{Modèle de prédiction}}$$

# Metamodèle du Tamar Bridge – Hyperparamètres

## Estimation des hyperparamètres:

$$\theta = \{\ell_1, \ell_2, \ell_3, \ell_4, \ell_5, \sigma_G\}$$

$$\theta^* = \arg \max_{\theta} \overbrace{\log p(\mathcal{D}_y | \mathcal{D}_x, \theta)}^{\text{log-likelihood}}$$

$$= \left\{ \begin{array}{l} \ell_1 = 0.0006 \text{ mm/mm} \\ \ell_2 = 0.0004 \text{ mm/mm} \\ \ell_3 = 2.21 \text{ kN/mm} \\ \ell_4 = 2.31 \text{ kN/mm} \\ \ell_5 = 1.76 \text{ kN/mm} \\ \sigma_G = 0.007 \text{ Hz} \end{array} \right\}$$



# Considérations pratiques des modèles empiriques

$$y_i = \overbrace{f(\mathbf{x}_i)}^{\text{réalité}} + \underbrace{V}_{\text{erreur de mesure}}$$

observation
erreur de mesure

$$V \sim \mathcal{N}(0, \sigma_V^2), \quad \underbrace{\mathbf{x}_i = [x_1, x_2, \dots, x_N]_i^T}_{\text{attributs}}$$

$$\begin{aligned} \mathbf{y} &= [y_1, y_2, \dots, y_M]^T \\ \mathbf{x} &= [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M]_{N \times M} \\ \ell &= \underbrace{[\ell_1, \ell_2, \dots, \ell_N]_{N \times 1}}_{\text{un } \ell_k \text{ par attribut } \mathbf{x}_k} \end{aligned}$$

pour  $i = 1, \dots, M$ , où  $M$  est le nombre d'observations

$$\mathbf{M}_Y = \mathbf{0}$$

$$\sigma_V > 0 \text{ (erreur de mesure)}$$

$$\rho(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{1}{2}(\mathbf{x}_i - \mathbf{x}_j)^T \text{diag}(\ell)^{-2}(\mathbf{x}_i - \mathbf{x}_j)\right)$$

## Exemple – Site contaminé

**Objectif:** Réhabilitation d'un site contaminé; on doit retirer toute la portion de sol où [ ] contaminants > [ ] admissible

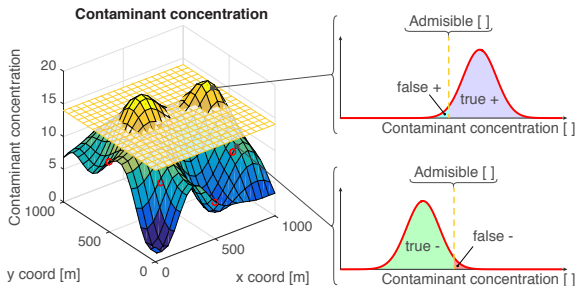
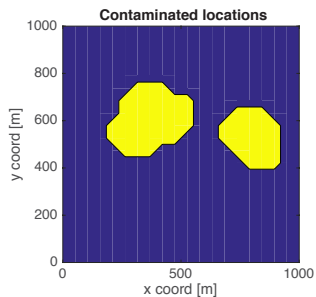
Afin de connaître où se situent les sols contaminés, on doit prendre des échantillons de sols afin de calculer la [ ]:



























- ▶ sur le terrain (200\$/essai,  $\sigma_{[ ]} = 0.5$ )
- ▶ en laboratoire (2000\$/essai,  $\sigma_{[ ]} = 0.01$ )

**Coûts:** (“+” = contaminé)

- ▶ vrai “+” = 100/ $m^2$  ( $m^2$  retiré correctement)
- ▶ vrai “-” = 0/ $m^2$  ( $m^2$  laissé en place correctement)
- ▶ faux “+” = 100/ $m^2$  ( $m^2$  retiré par erreur)
- ▶ faux “-” = 300/ $m^2$  ( $m^2$  laissé en place par erreur)

# Exemple – Site contaminé



<i>Théorie probabilité</i>	{	0	Révision algèbre & probabilités	
		1	Lois de probabilités	 
		3	PDF multivariés	 
<i>Fiabilité des structures &amp; systèmes</i>	{		Introduction	
		2	Formulation fiabilité composantes	$g_1(X)$ 
		9	Formulation fiabilité systèmes	 
<i>Estimation – <math>p_f</math> échantillonnage</i>	{	4	Monte Carlo	 
		12	Échantillonnage par importance	  
<i>Estimation – <math>p_f</math> analytique</i>	{	5	FOSM	  
		6	FORM	  
		8	SORM	  
<i>Utilisation des résultats</i> 	{	7	Analyse de sensibilité	
		10	Incertitudes aléatoires & épistémiques	
		11	Données empiriques	 
		13	Prise de décision	
		14	Métamodèles & modèles empiriques	