JU Open TURNS
07/06/2011

Méta-modèles

Thierry YALAMAS

Sommaire

- Les incertitudes
- Les surfaces de réponses
- Illustrations
- Conclusions

Méthodologie générale

Étape B

Quantification
des sources d'incertitudes

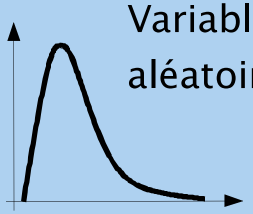
Étape A

Modèle(s)
Critère(s)

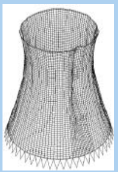
Étape C

Propagation
des incertitudes

Variables
aléatoires



Modèle physique



Densité de proba.
Moments
Prob. de défaillance

Étape C'

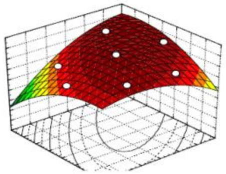
Hiérarchisation
des incertitudes

Uncertainty in industrial practice – A guide to quantitative uncertainty management,
E. de Rocquigny, N. Devictor, S. Tarantola (Eds.), John Wiley, 2008

Éléments d'une surface de réponse (1)

- La classe de modèle :
  type de représentation mathématique de la surface de réponse
  - Modèles simples : surface polynomiale (données d'entrée déterministes),
    chaos polynomial (données d'entrée aléatoires)

    Modèle le plus courant : Surface de Réponse Quadratique (SRQ), correspondant à une surface polynomiale d'ordre 2

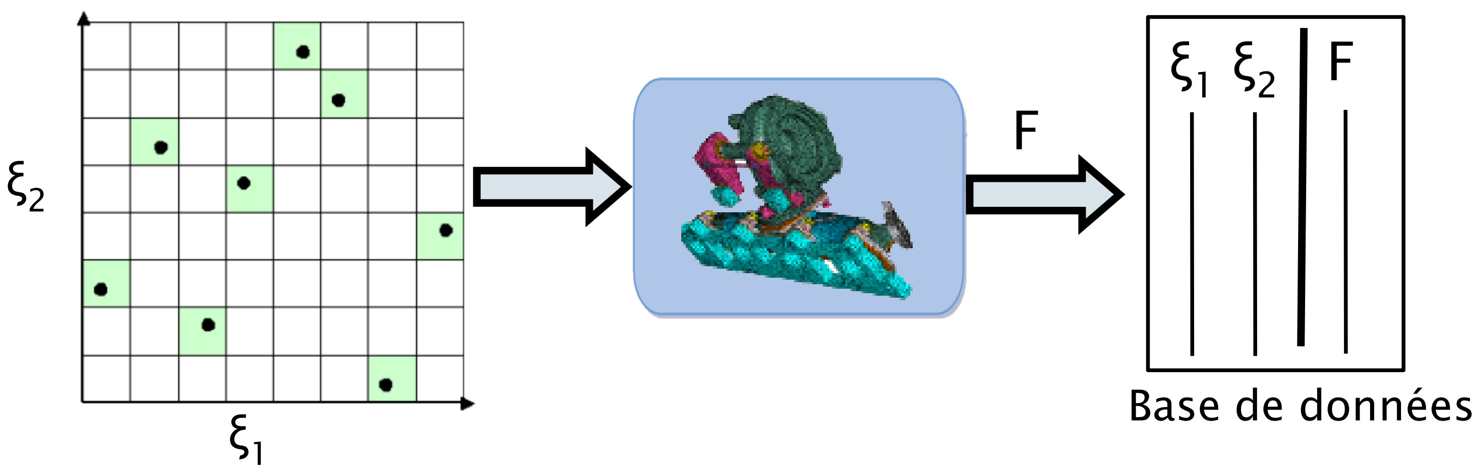
  - Modèles complexes : krigeage, réseaux de neurone, SVR
- L'algorithme de calcul :
  pour identifier les paramètres du méta-modèle à partir des points d'apprentissage de la surface de réponse
  - Régression, projection
  - Apprentissage
  - ...

Éléments d'une surface de réponse (2)

Le plan d'expérience

- Plans spécifiques au choix du modèle
- Plans génériques (utilisables avec plusieurs modèles différents)

Le plan d'expérience a pour but de sélectionner les points les plus adaptés à la construction du méta-modèle

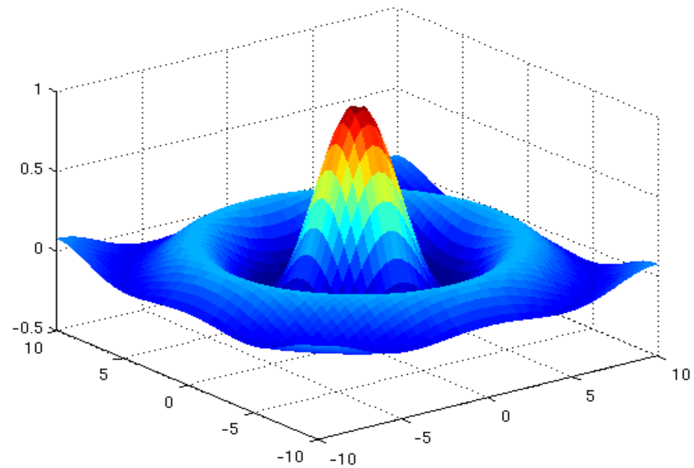


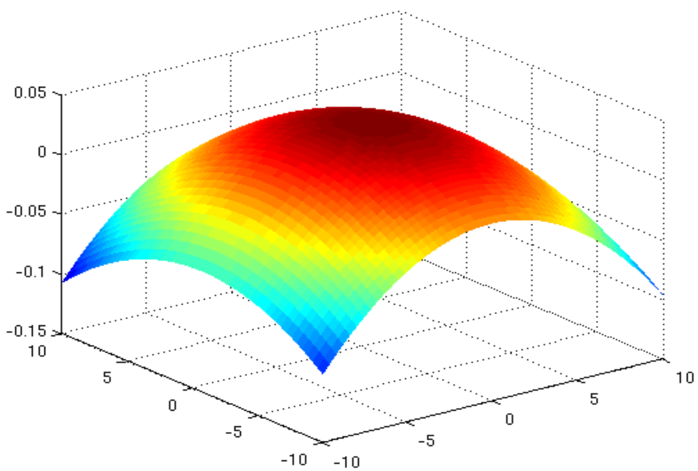
La taille du plan d'expérience est à l'origine du coût de calcul
Il est donc à choisir avec soin

Exemples : données déterministes

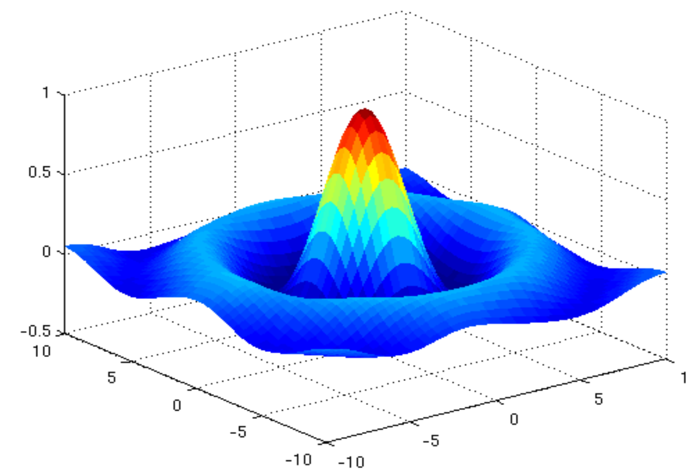
- Illustration :

Fonction sinus cardinal (sinc)





Régression polynomiale de degré 2



Régression SVM

Sommaire

Introduction aux SVM (1)

Un problème de classification linéaire :

- Un ensemble de données étiquetées X_i, i=1…n définies dans un domaine D
- On suppose 2 classes : $C_i \in \{+1;-1\}$
- Exemple en fiabilité des structures :
  - +1 : la classe sûreté,
  - -1 : la classe défaillance
- Objectif : trouver un hyperplan qui classe « le mieux possible » l'ensemble des données
- On cherche donc un séparateur des classes…

Introduction aux SVM (2)

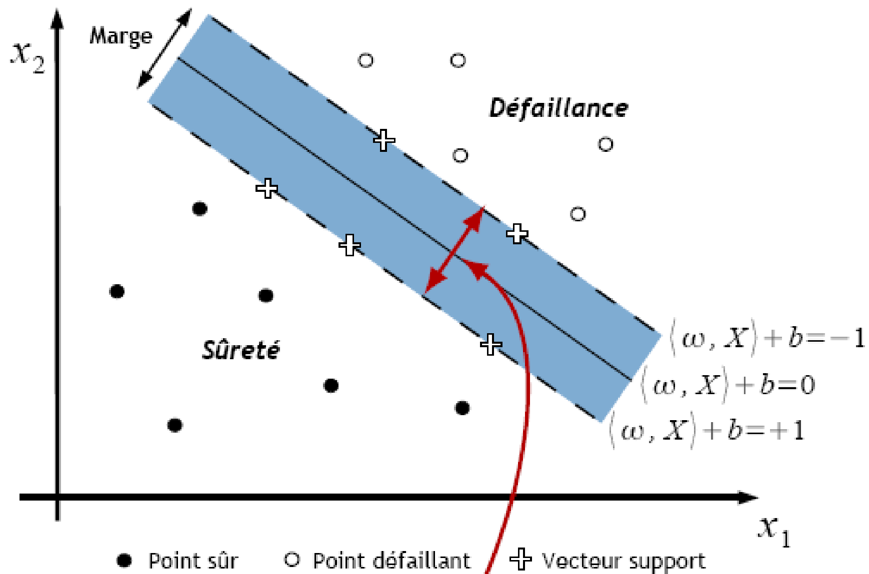
Un problème de classification linéaire :

- On cherche un hyperplan de la forme :

$$G(X) = \langle \omega, X \rangle + b$$

Vecteur de paramètres : normal à l'hyperplan

Seuil



- Comment obtenir le meilleur séparateur ?

Il faut maximiser la marge

Le séparateur est défini uniquement à partir des réalisations sur la marge : ***les points ou vecteurs supports***

Introduction aux SVM (3)

Le problème d'optimisation :

- Maximiser la marge correspond à la minimisation de la norme de ω, le problème d'optimisation s'écrivant :

➲ contraintes de bonne classification

$$\begin{cases} \min \dfrac{1}{2}\|\omega\|^2 \\ sous \; c_i\left(\langle \omega, X_i \rangle + b \geq 1\right) \quad i = 1, \ldots, n \end{cases}$$

- Formulation de Lagrange :

➲ fonction quadratique convexe

$$\boxed{L(\omega, b, \alpha) = \frac{\|\omega\|}{2} - \sum_{i=1}^{n} \alpha_i \left[c_i \left(\langle \omega, X \rangle + b \right) - 1 \right]}$$

- Conditions d'optimalités de Karush – Kuhn - Tucker :

$$\frac{\partial L(\omega, b, \alpha)}{\partial b} = 0 = \sum_{i=1}^{n} \alpha_i c_i \qquad \frac{\partial L(\omega, b, \alpha)}{\partial \omega} = 0 = \omega - \sum_{i=1}^{n} \alpha_i c_i X_i$$

Introduction aux SVM (4)

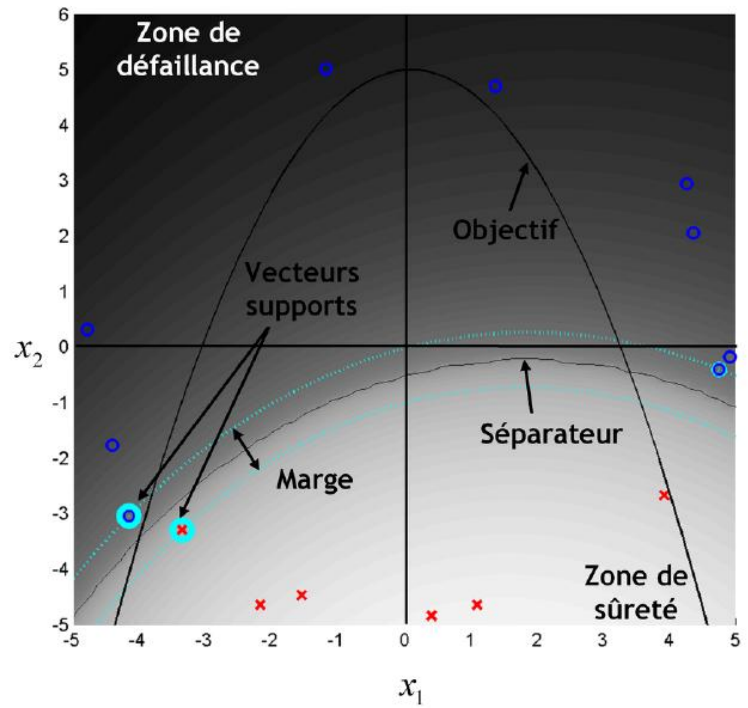
Le problème d'optimisation :

- Les αi non nuls correspondent aux réalisations sur la marge, elles sont nommées les vecteurs supports

- Le séparateur peut être uniquement défini à partir de ces réalisations, où S est le nombre de vecteurs supports

$$\omega^* = \sum_{i=1}^{S} \alpha_i c_i X_i$$

$$b^* = -\frac{1}{2}\left\langle \omega^*, X_{+1} + X_{-1} \right\rangle$$

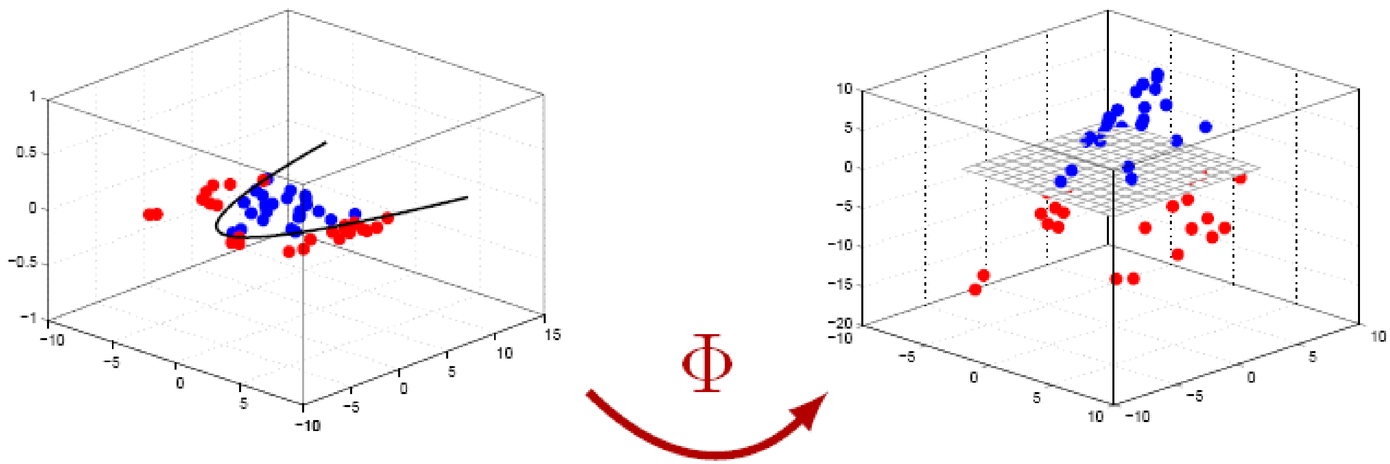


➲ Pour affiner la séparation, seules les réalisations dans la marge sont utiles, les autres peuvent être classées sans calcul à partir de la connaissance du séparateur

Introduction aux SVM (5)

Le passage au non-linéaire :

- On se ramène au cas linéairement séparable présenté précédemment par une transformation de l'espace :



- Un projecteur non linéaire *Φ* transforme l'espace de départ vers un espace de dimension supérieure : ***feature space***

→ The Kernel Trick

Le produit scalaire est remplacé par une fonction noyau *K* plus générale

$$K\left(X_1, X_2\right) = \left\langle \Phi\left(X_1\right), \Phi\left(X_2\right) \right\rangle$$

Sommaire

- Les incertitudes
- Les surfaces de réponses
  - Interpolation par Krigeage
- Illustrations
- Conclusion

La théorie de Krige & Matheron

Les hypothèses du krigeage

> $y(\mathbf{x})$ est une trajectoire *à identifier* d'un processus gaussien $Y(\mathbf{x})$ *à identifier*.

- Soit $Y(\mathbf{x})$ le *processus gaussien* non stationnaire :

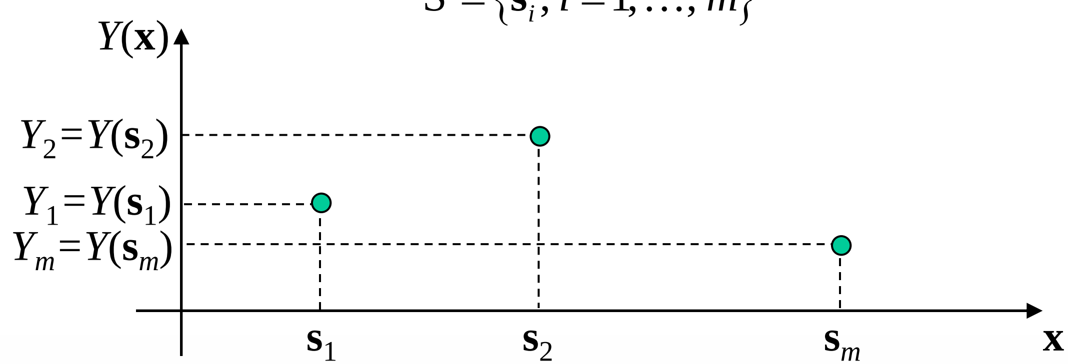
$$\underbrace{Y(\mathbf{x})}_{\text{PG non stationnaire}} = \underbrace{\mu(\mathbf{x})}_{\text{Partie déterministe (tendance)}} + \underbrace{Z(\mathbf{x})}_{\text{PG centré stationnaire}}$$

- On introduit également le *vecteur aléatoire* des observations

$$\mathbf{Y} \sim \mathrm{N}\left(\boldsymbol{\mu}_{\mathbf{Y}}, \sigma^2\left[\mathbf{R}_{\mathbf{YY}}\right]\right) \quad \text{avec } \sigma^2 R_{Y_iY_j} = Cov\left[Y_i, Y_j\right], \quad i = 1, \ldots, m \quad j = 1, \ldots, m$$

- aux sites

$$\mathrm{S} = \left\{\mathbf{s}_i, i = 1, \ldots, m\right\}$$



La théorie de Krige & Matheron

Le meilleur prédicteur linéaire sans biais de $Y(\mathbf{x})$

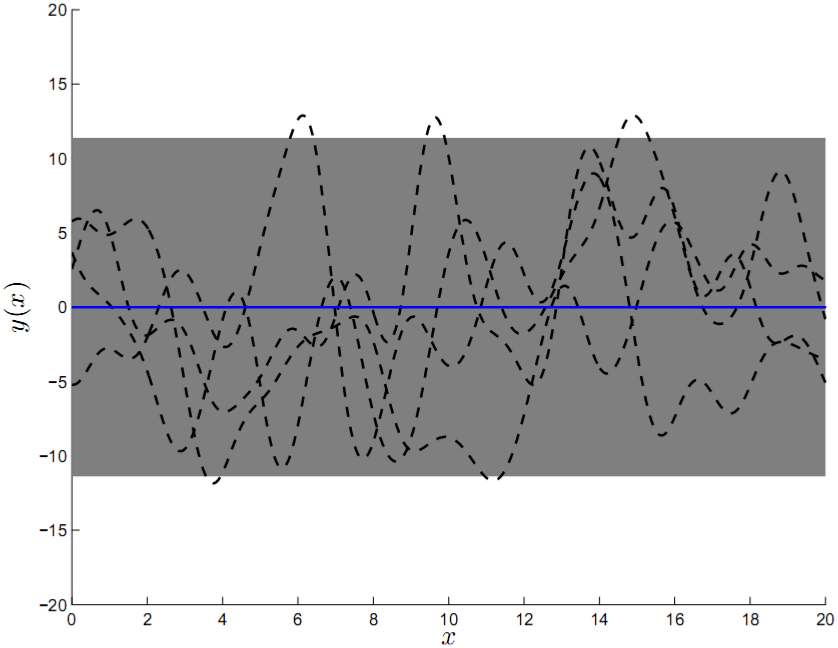
- D'où le résultat final

$$\begin{cases} \hat{Y}(\mathbf{x}) = \mu(\mathbf{x}) + \mathbf{r}(\mathbf{x})^t [\mathbf{R}_{\mathbf{YY}}]^{-1} (\mathbf{Y} - \boldsymbol{\mu}_\mathbf{Y}) \\ MSE(\mathbf{x}) = \sigma^2 \left(1 - \mathbf{r}(\mathbf{x})^t [\mathbf{R}_{\mathbf{YY}}]^{-1} \mathbf{r}(\mathbf{x})\right) \end{cases}$$

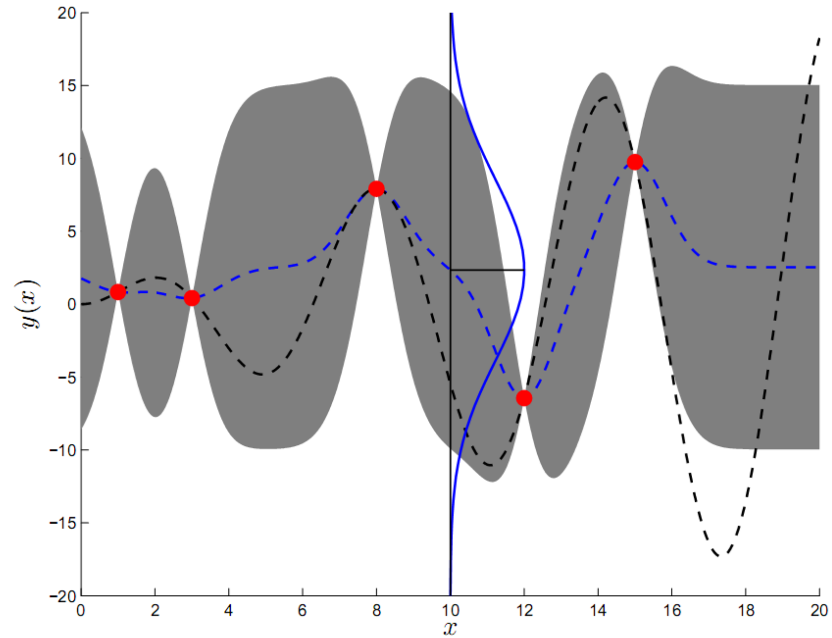
(le BLUP « moyen »)

(la « variance de krigeage »)

- Illustration (IC à 95%, fonction réelle : $y(x) = x \sin(x)$)



a priori ($\mathbf{Y}$ aléatoire)



a posteriori ($\mathbf{Y}$ = $\mathbf{y}$ observés)

La théorie de Krige & Matheron

Détermination des paramètres

- On ne connaît pas les moments du processus $Y(\mathbf{x})$, on va donc les modéliser :
  - La *moyenne* est <u>choisie</u> parmi une classe de *modèle de régression* :

$$\boxed{\mu(\mathbf{x}) = \mathbf{f}(\mathbf{x})^{\mathrm{t}} \boldsymbol{\beta}} \qquad \textit{e.g.}\ \mathbf{f}_{\mathrm{poly}}(\mathbf{x}) = \langle 1 \quad x \quad x^2 \rangle^{\mathrm{t}}$$

  - L'*autocorrélation* est également <u>choisie</u> dans une classe paramétrée connue :

$$\boxed{C(\mathbf{x}, \mathbf{x}') = \sigma^2 \mathrm{R}_{\boldsymbol{\theta}}(\mathbf{x} - \mathbf{x}')}$$

- Sous ses hypothèses additionnelles, le BLUP « moyen » s'écrit :

$$\hat{Y}(\mathbf{x}) = \mathbf{f}(\mathbf{x})^{\mathrm{t}} \boldsymbol{\beta} + \mathbf{r}_{\boldsymbol{\theta}}(\mathbf{x}) [\mathbf{R}_{\boldsymbol{\theta}}]^{-1} (\mathbf{Y} - [\mathbf{F}]\boldsymbol{\beta})$$

$$F_{ij} = f_j(\mathbf{s}_i), \quad i = 1, \ldots, m \quad j = 1, \ldots, p$$

- Et le vecteur des observations est ainsi entièrement paramétré :

$$\mathbf{Y} \sim \mathrm{N}\left([\mathbf{F}]\boldsymbol{\beta}, \sigma^2 [\mathbf{R}_{\boldsymbol{\theta}}]\right)$$

- On peut donc estimer les paramètres *optimaux* recherchés $\boldsymbol{\beta}^*$, σ^{2*}, $\boldsymbol{\theta}^*$ par les méthodes d'inférence paramétriques usuelles à partir du *vecteur des observations constatées* $\mathbf{y}$.

Pratique la plus courante : le *Maximum de Vraisemblance*

Sommaire

- Les incertitudes
- Les surfaces de réponses
- Illustration (Projet ANR Kidpocket)
- Conclusions

Objectifs

- Utilisation de méta-modèles
  - Études des moments statistiques
  - Études de quantiles particuliers
- Comparaison de méta-modèles
  - Régression par SVM (*Support Vector Machine*)
  - Interpolation par la méthode du Krigeage
  - Chaos polynomial adaptatif

Les données d'entrée

Issues d'un code de calcul électromagnétique

Les données analysées sont des valeurs de SAR (Specific Absorption Rate)

Caractéristiques des données

- A) Une Bases de données 27 points générées par un schéma de quadrature
- B) Une base de données de 25 points générées par un plan d’expérience aléatoire de type carré Latin
- Dimension de l'espace : 3 variables aléatoires dépendantes

$$\boldsymbol{R} = \begin{bmatrix} 1 & 0,33 & 0,25 \\ 0,33 & 1 & 0,69 \\ 0,25 & 0,69 & 1 \end{bmatrix}$$

- Les données sont traitées dans l'espace standard (gaussien centré et normé)

Méthodologie

Passage des données dans un espace normé

↓

Optimisation des paramètres du méta-modèle

↓

Simulations sur le méta-modèle (100 000pts)

↓

Évaluation des moments

Les outils

- Tous les résultats ont été obtenus sous Matlab/python à l’aide d’outils internes

- Les *toolbox*

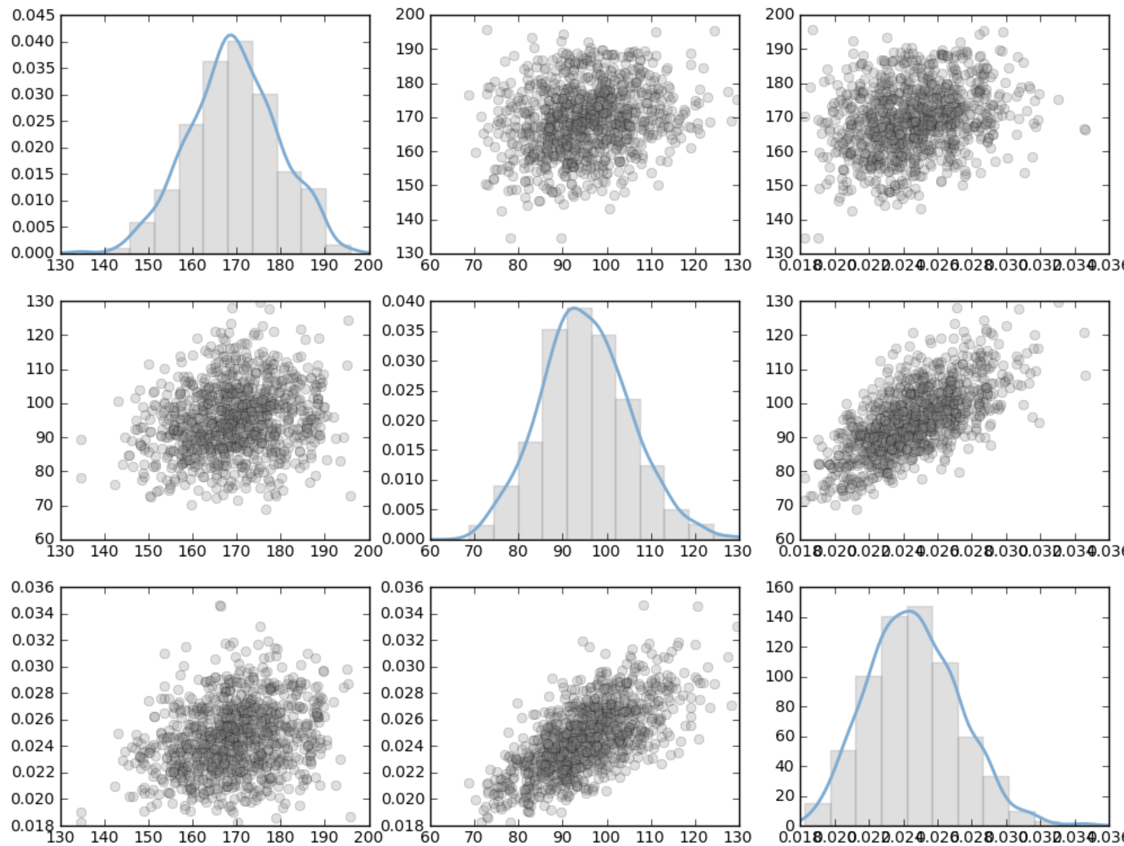
  - Support Vector regression (SVR) : SpiderSVM, interface Matlab pour LibSVM

  - Krigeage : DACE

  - Chaos adpatatif (gPC) : développements internes

Simulation des données

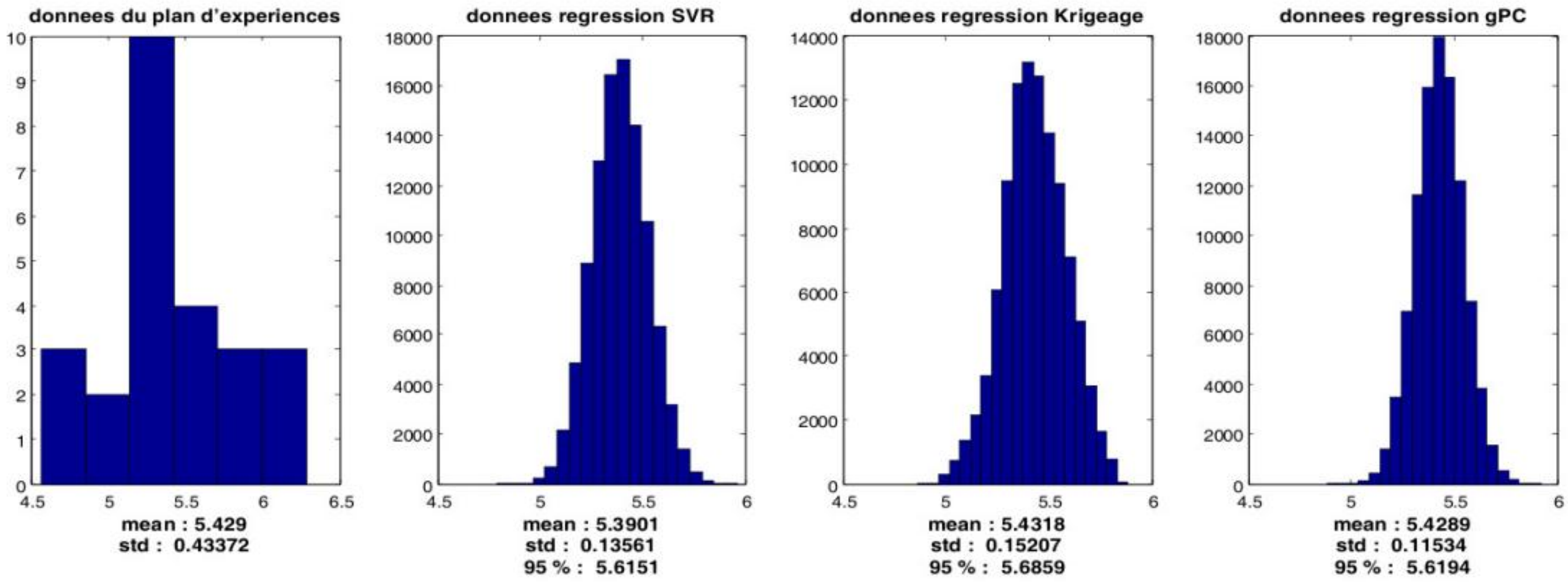
Taille Carrure devant Tr de poitrine



Population générée pour l'évaluation des moments

Résultats – base schéma de quadrature

Comparaisons des histogrammes

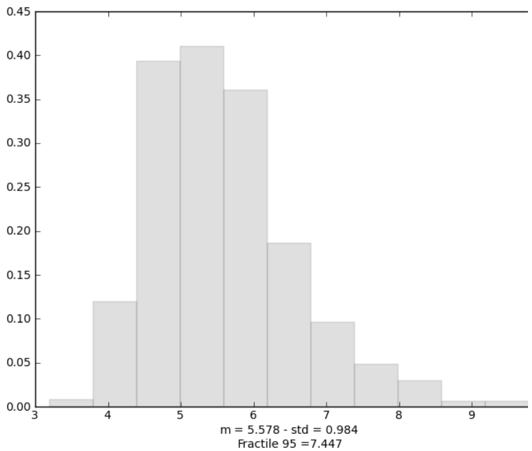


Note : les moments sont évalués sur 100 000 simulations de Monte-Carlo

Résultats – base LHS

Comparaisons des histogrammes

Chaos

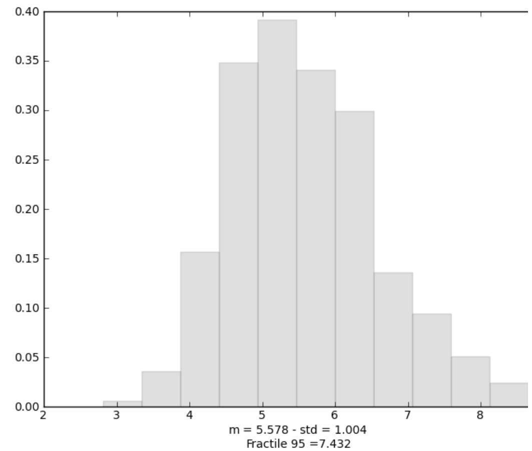


m = 5.578

σ = 0.984

95% = 7.447

Krigeage

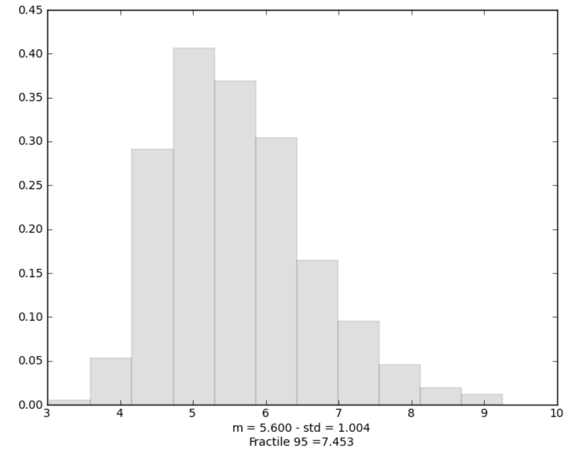


m = 5.578

σ = 1.004

95% = 7.432

SVR



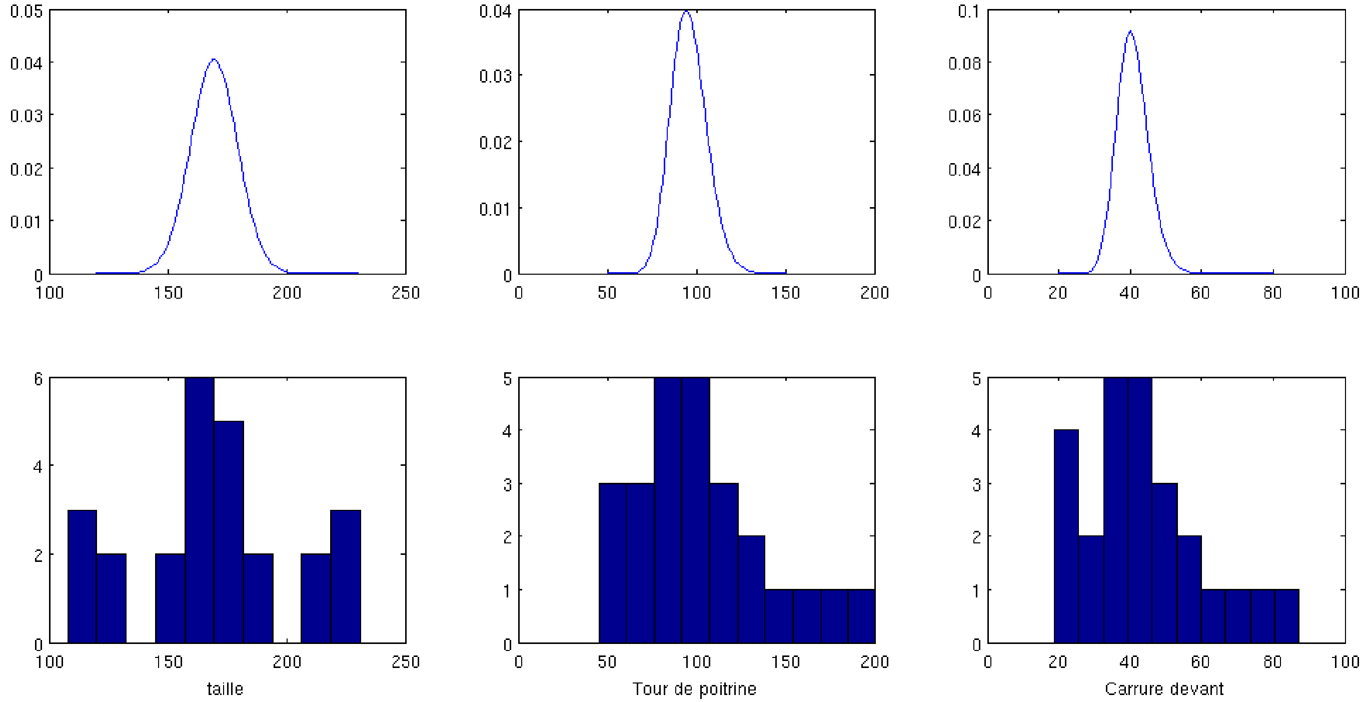
m = 5.600

σ = 1.004

95% = 7.453

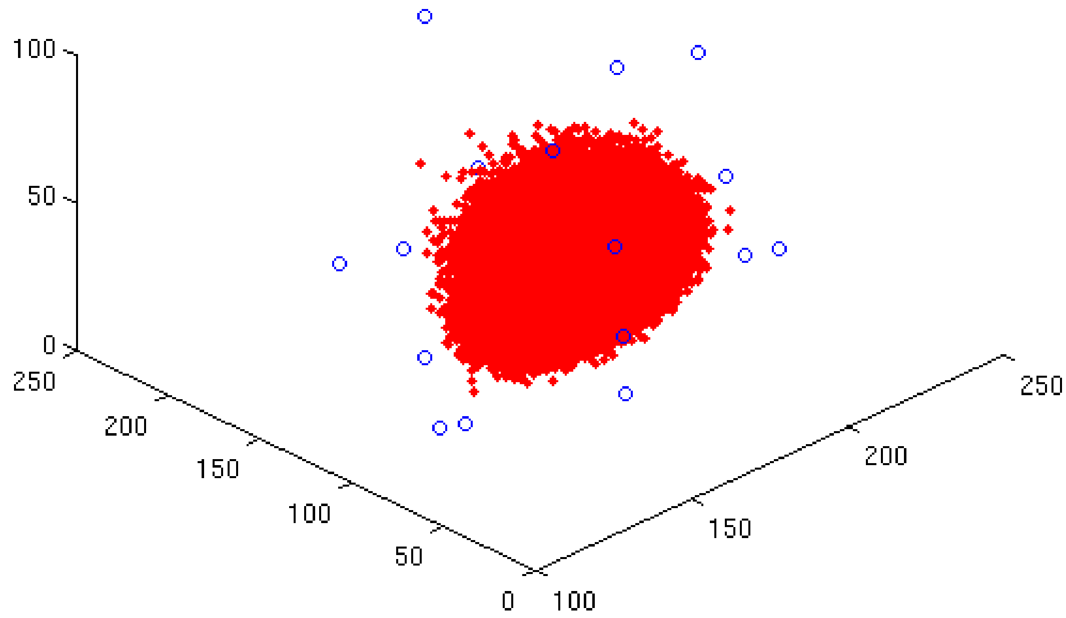
Note : les moments sont évalués sur 100 000 simulations de Monte-Carlo

Explications



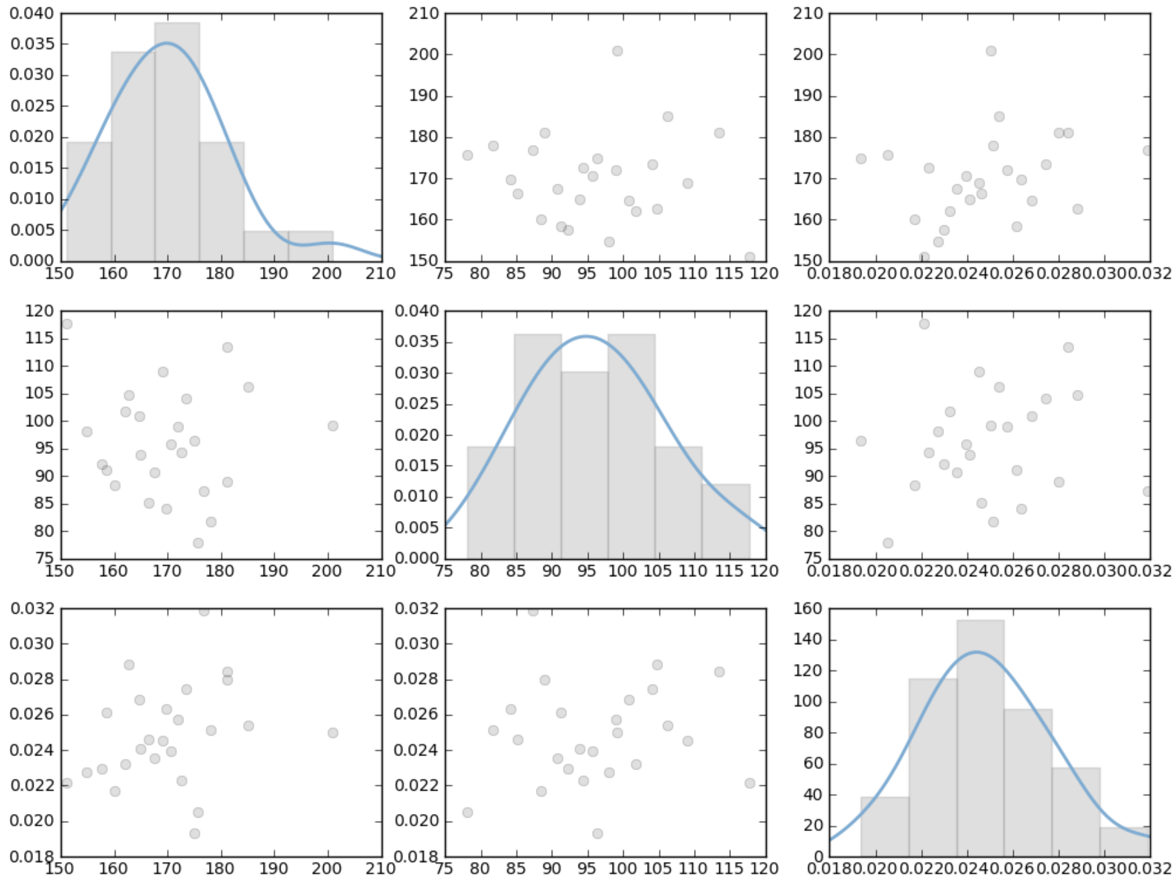
Répartition des expériences numériques du schéma de quadrature

Explications



Répartition des expériences numériques du schéma de quadrature (en bleu) et points d'évaluation pour les moments (en rouge)

Explications

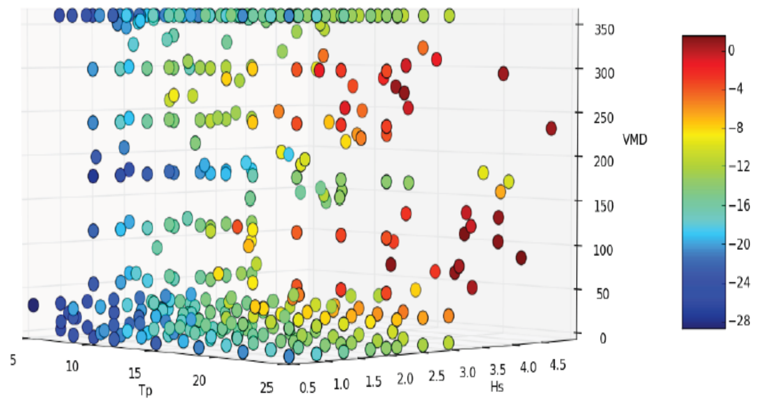


Répartition des expériences numériques du plan LHS

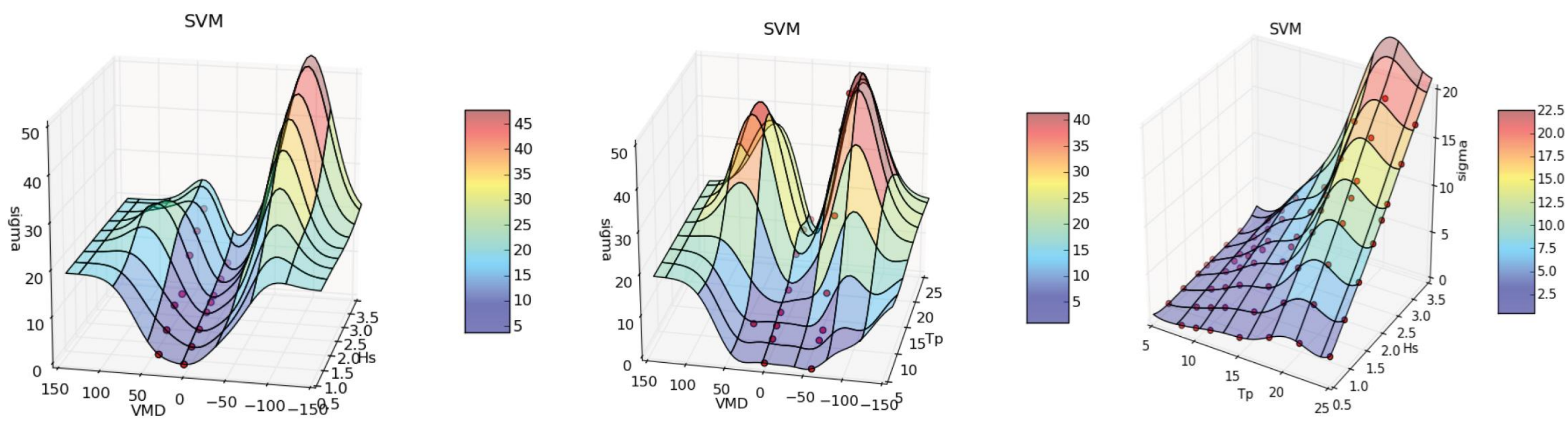
Sommaire

Elaboration du Méta-modèle

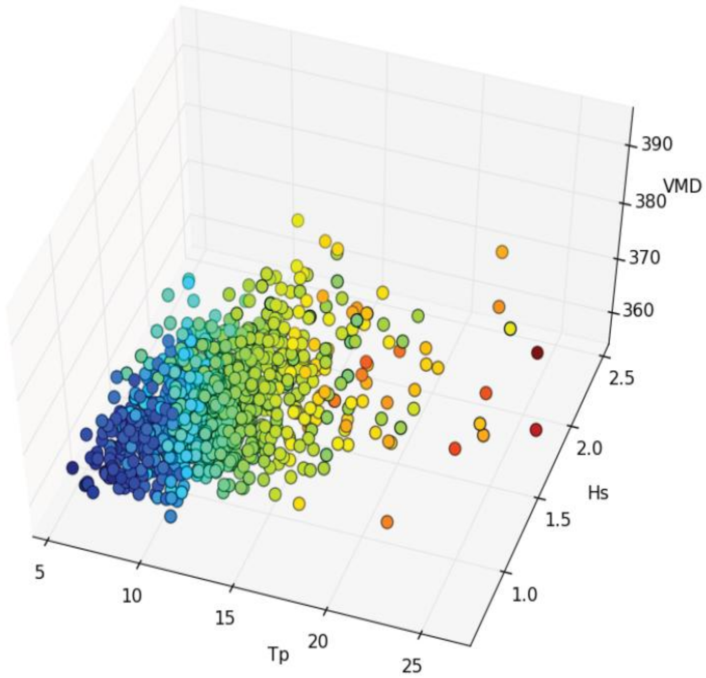
Plan d'expérience:
- LHS
- Complet sur 2 facteurs



Résultats de la régression:



Résultats de l'analyse fiabiliste



Modèle	$\log(D)$	ΔS
Krigeage	61 ans (0.85)	249 ans (0.82)
SVR	76 ans (0.95)	211 ans (0.95)

Résultats de la simulation de Monte-Carlo
- avec le méta-modèle
- suivant les distributions inférées

Conclusions

- La validité d'un méta-modèle dépend de celle du plan d'expérience (par rapport à une information recherchée)

- Le choix d'un type de méta-modèle dépend de la « physique » du phénomène représenté… dans la plage considérée