Gaussian Process Regression on Nested Spaces

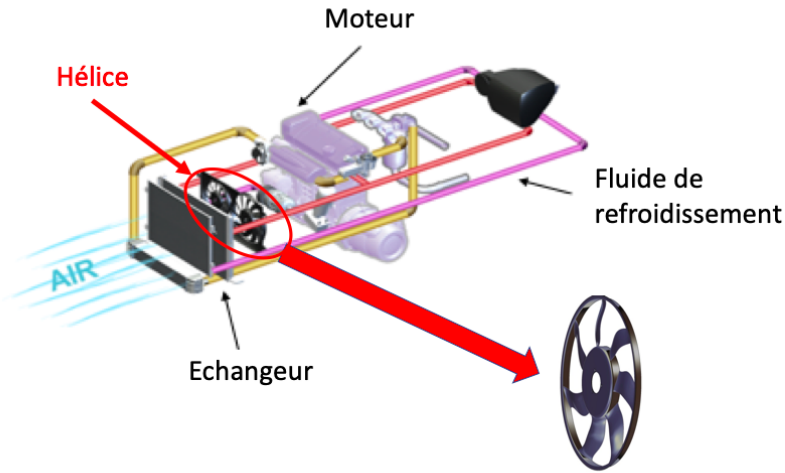
Christophette Blanchet-Scalliet

ICJ-Ecole Centrale de Lyon

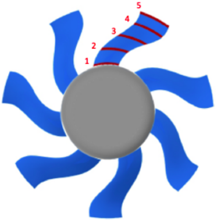
Joint work with T.Gonon (Safran), C. Helbert 5ECL) B. Demory (Valeo)

Clermont-Ferrand, Thursday 21 December 2023

Motivation- Cooling system

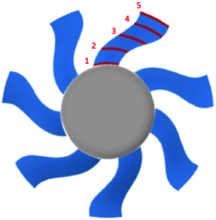


Motivation-Cooling System

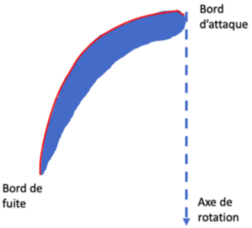


Sections de pale

Motivation-Cooling System

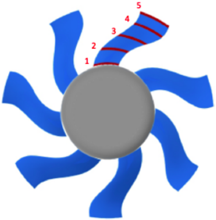


Sections de pale

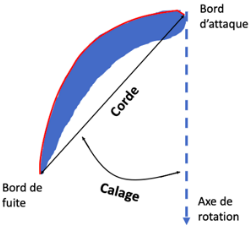


Calage et longueur de corde

Motivation-Cooling System

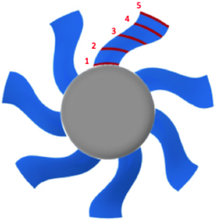


Sections de pale

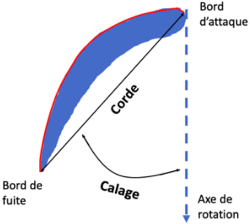


Calage et longueur de corde

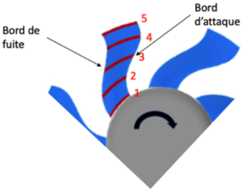
Motivation-Cooling System



Sections de pale

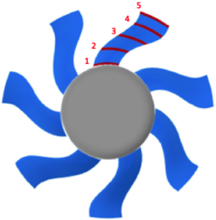


Calage et longueur de corde

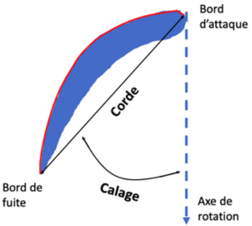


Empilement tangentiel

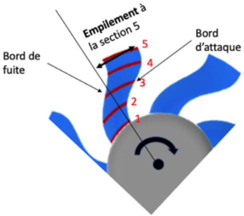
Motivation-Cooling System



Sections de pale

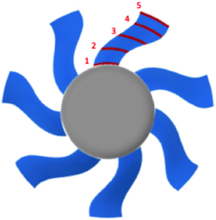


Calage et longueur de corde

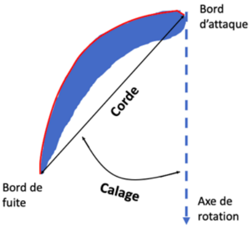


Empilement tangentiel

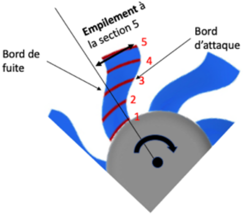
Motivation-Cooling System



Sections de pale



Calage et longueur de corde



Empilement tangentiel

Outputs

- Pressure difference : ΔP (en Pa)

Exploitation of the simulation

Computatinally expensive code $\Rightarrow$ need a metamodel.

Exploitation of the simulation

Computatinally expensive code ⇒ need a metamodel.
Problem : a lot of inputs (Stagger, chord length, sweep, ...)

Exploitation of the simulation

Computatinally expensive code ⇒ need a metamodel.
Problem : a lot of inputs (Stagger, chord length, sweep, ...)

Which study ? Untill now at each step sequentially releasing variables by group

Exploitation of the simulation

Computatinally expensive code ⇒ need a metamodel.
Problem : a lot of inputs (Stagger, chord length, sweep, ...)

Which study ? Untill now at each step sequentially releasing variables by group

- First we consider only staggers ⇒ built a DoE and a metamodel.

Exploitation of the simulation

Computatinally expensive code ⇒ need a metamodel.
Problem : a lot of inputs (Stagger, chord length, sweep, ...)

Which study ? Untill now at each step sequentially releasing variables by group

- First we consider only staggers ⇒ built a DoE and a metamodel.
- we add chord length ⇒ built a new DoE and a new metamodel.

Exploitation of the simulation

Computatinally expensive code $\Rightarrow$ need a metamodel.
Problem : a lot of inputs (Stagger, chord length, sweep, ...)

Which study ? Untill now at each step sequentially releasing variables by group

- First we consider only staggers $\Rightarrow$ built a DoE and a metamodel.
- we add chord length $\Rightarrow$ built a new DoE and a new metamodel.
- then the sweeps...

Exploitation of the simulation

Computatinally expensive code ⇒ need a metamodel.
Problem : a lot of inputs (Stagger, chord length, sweep, ...)

Which study ? Untill now at each step sequentially releasing variables by group

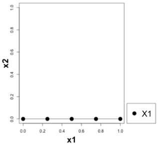
- First we consider only staggers ⇒ built a DoE and a metamodel.
- we add chord length ⇒ built a new DoE and a new metamodel.
- then the sweeps...

Which metamodel adapted to this sequential study ?
Aim : Using all the DoEs.

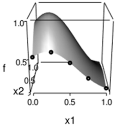
Example 2D

Let f a function of 2 variables $f(x_1, x_2) = \exp\left(\frac{(x_1-0.2)^2+(x_2-0.4)^2}{0.3}\right)$.

Inputs



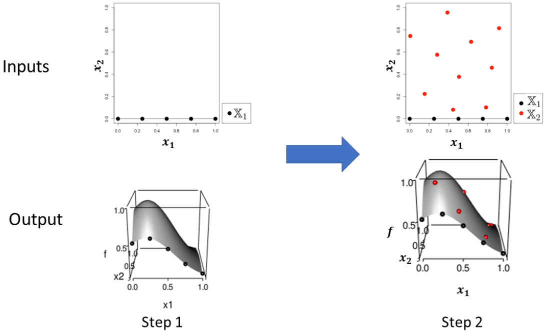
Output



Step 1

Example 2D

Let f a function of 2 variables $f(x_1, x_2) = \exp\left(\frac{(x_1-0.2)^2+(x_2-0.4)^2}{0.3}\right)$.



What kind of metamodel ?

Outline

seqGPR model
- Kriking
- New Model
- Correction process

Parameters estimation

Applications
- Analytic example in dimension 4
- Industrial Test Cases

Conclusion

Table des matières

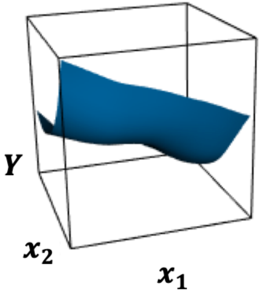
seqGPR model
- Kriking
- New Model
- Correction process

Parameters estimation

Applications
- Analytic example in dimension 4
- Industrial Test Cases

Conclusion

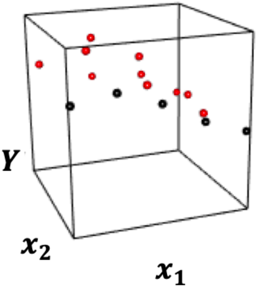
Classical Gaussian process regression [Santner et al., 2003]



The output f is the realization of a of a Gaussian process

$$Y(x_1, x_2) \sim \mathcal{PG}\left(m, k\right)$$

Classical Gaussian process regression [Santner et al., 2003]

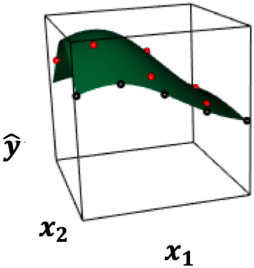


The output f is the realization of a of a Gaussian process

$$Y(x_1, x_2) \sim \mathcal{PG}\left(m, k\right)$$

The DoE $(\mathbb{X}, \boldsymbol{y})$ avec $\mathbb{X} = \mathbb{X}_1 \cup \mathbb{X}_2$.

Classical Gaussian process regression [Santner et al., 2003]



The output f is the realization of a of a Gaussian process

$$Y(x_1, x_2) \sim \mathcal{PG}\left(m, k\right)$$

The DoE $(\mathbb{X}, \boldsymbol{y})$ avec $\mathbb{X} = \mathbb{X}_1 \cup \mathbb{X}_2$.

Prediction :

$$\hat{y}(x_1, x_2) = m + k((x_1, x_2), \mathbb{X})k(\mathbb{X}, \mathbb{X})^{-1}(\boldsymbol{y} - m)$$

Parameters inference

$Y \sim \mathcal{PG}(m, k_{(\sigma^2,\theta)})$ with

$$k((x_1, x_2), (x_1', x_2')) = \sigma^2 \prod_{i=1}^{2} \left(1 + \frac{\sqrt{5}\,|x_i - x_i'|}{\theta_i} + \frac{5(x_i - x_i')^2}{3\theta_i^2}\right) \exp\left(-\frac{\sqrt{5}\,|x_i - x_i'|}{\theta_i}\right)$$

Parameters inference

$Y \sim \mathcal{PG}(m, k_{(\sigma^2,\theta)})$ with

$$k((x_1, x_2), (x'_1, x'_2)) = \sigma^2 \prod_{i=1}^{2} \left(1 + \frac{\sqrt{5}\,|x_i - x'_i|}{\theta_i} + \frac{5(x_i - x'_i)^2}{3\theta_i^2}\right) \exp\left(-\frac{\sqrt{5}\,|x_i - x'_i|}{\theta_i}\right)$$

$\eta = (m, \sigma^2, \theta_1, \theta_2)$ estimated from the DoE $(\mathbb{X}, \boldsymbol{y})$ (size $n_{\mathbb{X}}$) by maximum of likelood

Parameters inference

$Y \sim \mathcal{PG}(m, k_{(\sigma^2,\theta)})$ with

$$k((x_1, x_2), (x_1', x_2')) = \sigma^2 \prod_{i=1}^{2} \left(1 + \frac{\sqrt{5}\,|x_i - x_i'|}{\theta_i} + \frac{5(x_i - x_i')^2}{3\theta_i^2}\right) \exp\left(-\frac{\sqrt{5}\,|x_i - x_i'|}{\theta_i}\right)$$

$\eta = (m, \sigma^2, \theta_1, \theta_2)$ estimated from the DoE $(\mathbb{X}, \boldsymbol{y})$ (size $n_{\mathbb{X}}$) by maximum of likelood

$$\mathcal{L}(Y(\mathbb{X}) = \boldsymbol{y}; \eta) =$$

Parameters inference

$Y \sim \mathcal{PG}(m, k_{(\sigma^2,\theta)})$ with

$$k((x_1,x_2),(x_1',x_2')) = \sigma^2 \prod_{i=1}^{2} \left(1 + \frac{\sqrt{5}\,|x_i - x_i'|}{\theta_i} + \frac{5(x_i - x_i')^2}{3\theta_i^2}\right) \exp\left(-\frac{\sqrt{5}\,|x_i - x_i'|}{\theta_i}\right)$$

$\eta = (m, \sigma^2, \theta_1, \theta_2)$ estimated from the DoE $(\mathbb{X}, \boldsymbol{y})$ (size $n_{\mathbb{X}}$) by maximum of likelood

$$\mathcal{L}(Y(\mathbb{X}) = \boldsymbol{y}; \eta) = \frac{1}{(2\pi)^{\frac{n_{\mathbb{X}}}{2}}\,|k(\mathbb{X},\mathbb{X})|^{\frac{1}{2}}} \exp\left(-\frac{(\boldsymbol{y}-m)^T k(\mathbb{X},\mathbb{X})^{-1}(\boldsymbol{y}-m)}{2}\right)$$

seqGPR model

Can we do better ?

seqGPR model

Can we do better ? $\rightarrow$ new Model seq GPR

seqGPR model

- Can we do better ? $\rightarrow$ new Model seq GPR
- seqGPR can be view as a "multi-fidelity model" [Kennedy and O'Hagan, 2000],

seqGPR model

Can we do better ? $\rightarrow$ new Model seq GPR
seqGPR can be view as a "multi-fidelity model" [Kennedy and O'Hagan, 2000], f realization of :

$$Y_2(x_1, x_2) = Y_1(x_1) + Z_2(x_1, x_2)$$

seqGPR model

Can we do better ? $\rightarrow$ new Model seq GPR
seqGPR can be view as a "multi-fidelity model"
[Kennedy and O'Hagan, 2000], f realization of :

$$Y_2(x_1, x_2) = Y_1(x_1) + Z_2(x_1, x_2)$$

- $Y_1 \sim \mathcal{PG}(m, k_1(x_1, x_1'))$ with k_1 stationary

seqGPR model

Can we do better ? $\rightarrow$ new Model seq GPR
seqGPR can be view as a "multi-fidelity model"
[Kennedy and O'Hagan, 2000], f realization of :

$$Y_2(x_1, x_2) = Y_1(x_1) + Z_2(x_1, x_2)$$

- $Y_1 \sim \mathcal{PG}(m, k_1(x_1, x_1'))$ with k_1 stationary
- $Z_2 \sim \mathcal{PG}(0, k_2((x_1, x_2), (x_1', x_2'))$

seqGPR model

Can we do better ? $\rightarrow$ new Model seq GPR
seqGPR can be view as a "multi-fidelity model" [Kennedy and O'Hagan, 2000], f realization of :

$$Y_2(x_1, x_2) = Y_1(x_1) + Z_2(x_1, x_2)$$

- $Y_1 \sim \mathcal{PG}(m, k_1(x_1, x_1'))$ with k_1 stationary
- $Z_2 \sim \mathcal{PG}(0, k_2((x_1, x_2), (x_1', x_2'))$
- $Z_2(x_1, 0) = 0$

seqGPR model

Can we do better ? $\rightarrow$ new Model seq GPR
seqGPR can be view as a "multi-fidelity model" [Kennedy and O'Hagan, 2000], f realization of :

$$Y_2(x_1, x_2) = Y_1(x_1) + Z_2(x_1, x_2)$$

- $Y_1 \sim \mathcal{PG}(m, k_1(x_1, x_1'))$ with k_1 stationary
- $Z_2 \sim \mathcal{PG}(0, k_2((x_1, x_2), (x_1', x_2'))$
- $Z_2(x_1, 0) = 0$
- $Y_1 \perp Z_2$

seqGPR model

Can we do better ? $\rightarrow$ new Model seq GPR
seqGPR can be view as a "multi-fidelity model"
[Kennedy and O'Hagan, 2000], f realization of :

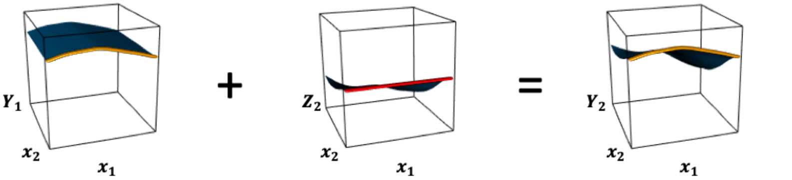
$$Y_2(x_1, x_2) = Y_1(x_1) + Z_2(x_1, x_2)$$

- $Y_1 \sim \mathcal{PG}(m, k_1(x_1, x_1'))$ with k_1 stationary
- $Z_2 \sim \mathcal{PG}(0, k_2((x_1, x_2), (x_1', x_2'))$
- $Z_2(x_1, 0) = 0$
- $Y_1 \perp Z_2$

$$f \text{ a realisation of } Y_2 \sim \mathcal{PG}(m, \boldsymbol{k_1(x_1, x_1) + k_2((x_1, x_2), (x_1', x_2'))})$$

seqGPR model

seqGPR model

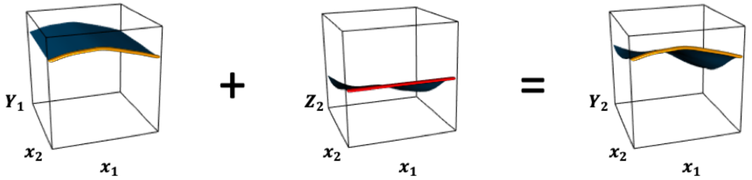


Etape 1 :
$f(x_1, 0)$ réalisation de $Y_1(x_1)$

Processus correctif :
$Z_2(x_1, x_2)$

Etape 2 :
$f(x_1, x_2)$ réalisation de
$Y_2(x_1, x_2) = Y_1(x_1) + Z_2(x_1, x_2)$

seqGPR model



Etape 1 :
$f(x_1, 0)$ réalisation de $Y_1(x_1)$

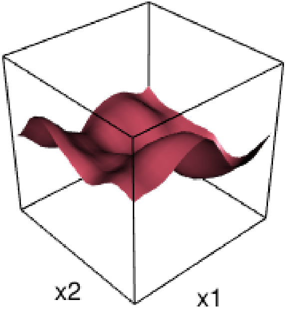
Processus correctif :
$Z_2(x_1, x_2)$

Etape 2 :
$f(x_1, x_2)$ réalisation de
$Y_2(x_1, x_2) = Y_1(x_1) + Z_2(x_1, x_2)$

Aim : Build Z_2 such that $Z_2(x_1, 0) = 0$?

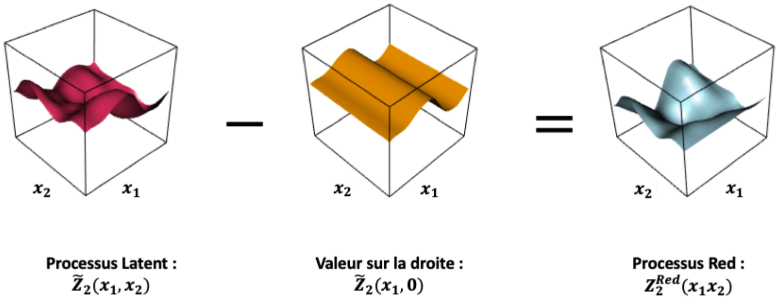
Idea

Build $Z_2(x_1, x_2)$ such that $Z_2(x_1, 0) = 0$ form a latent process $\widetilde{Z}_2 \sim \mathcal{PG}(0, \kappa)$ with a kernel κ fixed.



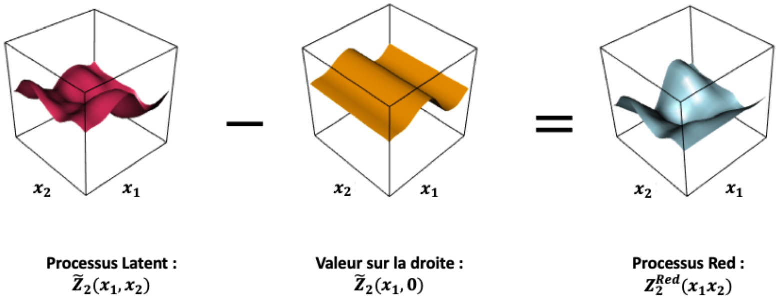
Red (reduced) process

Red process is defined by $Z_2^{Red}(x_1, x_2) = \widetilde{Z}_2(x_1, x_2) - \widetilde{Z}_2(x_1, 0)$.



Red (reduced) process

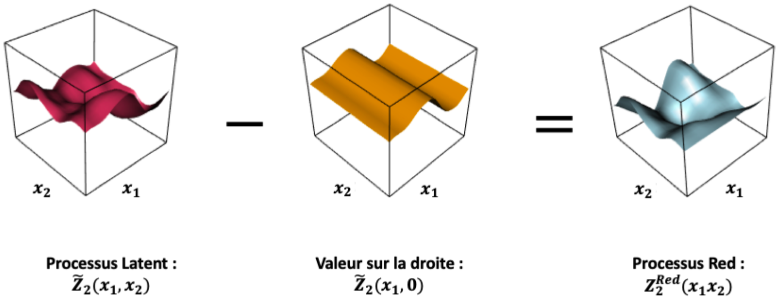
Red process is defined by $Z_2^{Red}(x_1, x_2) = \widetilde{Z}_2(x_1, x_2) - \widetilde{Z}_2(x_1, 0)$.



$$Z_2^{Red} \sim \mathcal{PG}\left(0, k_2\right)$$

Red (reduced) process)

Red process is defined by $Z_2^{Red}(x_1, x_2) = \widetilde{Z}_2(x_1, x_2) - \widetilde{Z}_2(x_1, 0)$.



$$Z_2^{Red} \sim \mathcal{PG}\left(0, k_2\right) \text{ with } k_2((x_1, x_2), (x_1', x_2')) = \kappa((x_1, x_2), (x_1', x_2')) + \kappa((x_1, 0), (x_1', 0)) - \kappa((x_1, 0), (x_1', x_2')) - \kappa((x_1, x_2), (x_1', 0)).$$

P (preconditioned) process

The process P is defined using conditional expectation introduced by [Gauthier and Bay, 2012]

$$Z_2^P(x_1, x_2) = \widetilde{Z}_2(x_1, x_2) - \mathbb{E}\left[\widetilde{Z}_2(x_1, x_2) \mid \widetilde{Z}(\mathcal{D})\right]$$

with

$$\mathcal{D} = \{(t_1, 0), t_1 \in [0, 1]\}$$

P (preconditioned) process

The process P is defined using conditional expectation introduced by [Gauthier and Bay, 2012]

$$Z_2^P(x_1, x_2) = \widetilde{Z}_2(x_1, x_2) - \mathbb{E}\left[\widetilde{Z}_2(x_1, x_2) \mid \widetilde{Z}(\mathcal{D})\right]$$

with

$$\mathcal{D} = \{(t_1, 0), t_1 \in [0, 1]\}$$

Proposition

If the process $\widetilde{Z}_2 \sim \mathcal{PG}(0, \kappa)$ has a tensor product kernel

$$\kappa((x_1, x_2), (x_1', x_2')) = \kappa_1(x_1, x_1')\kappa_2(x_2, x_2')$$

Then the conditional expectation is equal to :

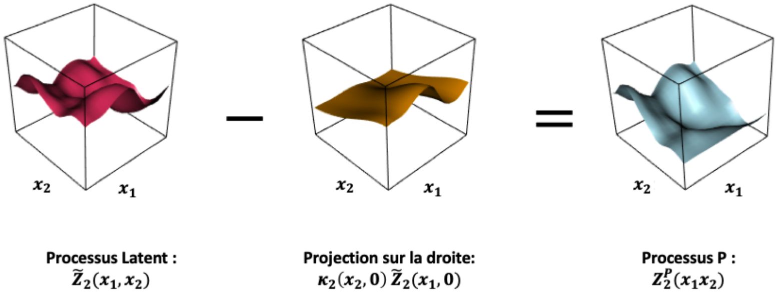
$$\mathbb{E}\left[\widetilde{Z}_2(x_1, x_2) \mid \widetilde{Z}(\mathcal{D})\right] = \kappa_2(x_2, 0)\widetilde{Z}_2(x_1, 0)$$

P (preconditioned) process

If $\widetilde{Z}_2 \sim \mathcal{PG}(0, \kappa)$ with $\kappa((x_1, x_2), (x_1', x_2')) = \kappa_1(x_1, x_1')\kappa_2(x_2, x_2')$,

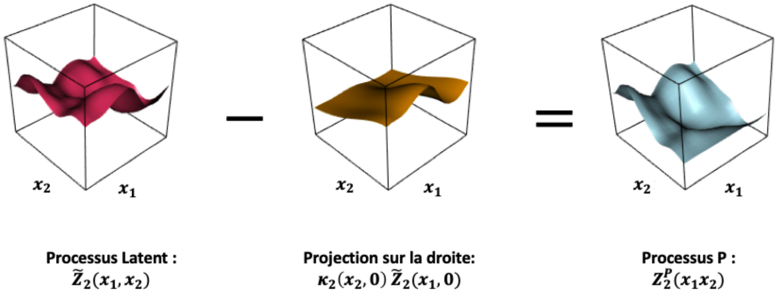
P (preconditioned) process

If $\widetilde{Z}_2 \sim \mathcal{PG}(0, \kappa)$ with $\kappa((x_1, x_2), (x_1', x_2')) = \kappa_1(x_1, x_1')\kappa_2(x_2, x_2')$, the P process is defined by $Z_2^P(x_1, x_2) = \widetilde{Z}_2(x_1, x_2) - \kappa_2(x_2, 0)\widetilde{Z}_2(x_1, 0)$.



P (preconditioned) process

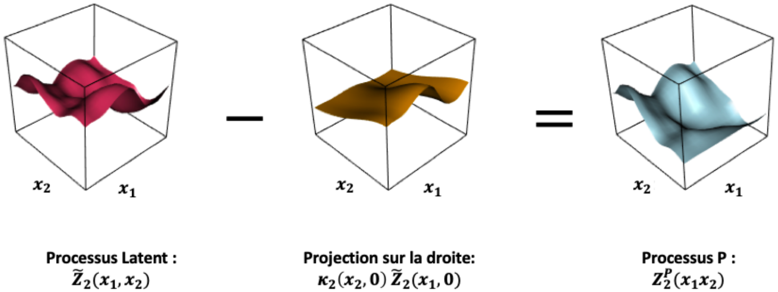
If $\widetilde{Z}_2 \sim \mathcal{PG}(0, \kappa)$ with $\kappa((x_1, x_2), (x_1', x_2')) = \kappa_1(x_1, x_1')\kappa_2(x_2, x_2')$, the P process is defined by $Z_2^P(x_1, x_2) = \widetilde{Z}_2(x_1, x_2) - \kappa_2(x_2, 0)\widetilde{Z}_2(x_1, 0)$.



$$Z_2^P \sim \mathcal{PG}(0, k_2)$$

P (preconditioned) process

If $\widetilde{Z}_2 \sim \mathcal{PG}(0, \kappa)$ with $\kappa((x_1, x_2), (x_1', x_2')) = \kappa_1(x_1, x_1')\kappa_2(x_2, x_2')$, the P process is defined by $Z_2^P(x_1, x_2) = \widetilde{Z}_2(x_1, x_2) - \kappa_2(x_2, 0)\widetilde{Z}_2(x_1, 0)$.



$$Z_2^P \sim \mathcal{PG}(0, k_2) \text{ with}$$
$$k_2((x_1, x_2), (x_1', x_2')) = \kappa_1(x_1, x_1')\,[\kappa_2(x_2, x_2') - \kappa_2(x_2, 0)\kappa_2(x_2', 0)].$$

Interpretation of the correction process

Latent process $\widetilde{Z}_2 \sim \mathcal{GP}(0, \kappa)$

Interpretation of the correction process

Latent process $\widetilde{Z}_2 \sim \mathcal{GP}(0, \kappa)$

- Red process :

$$Z_2^{Red}(x_1, x_2) = \widetilde{Z}_2(x_1, x_2) - \widetilde{Z}_2(x_1, 0)$$

$\Longrightarrow$ global disruption

Interpretation of the correction process

Latent process $\widetilde{Z}_2 \sim \mathcal{GP}(0, \kappa)$

- Red process :

$$Z_2^{Red}(x_1, x_2) = \widetilde{Z}_2(x_1, x_2) - \widetilde{Z}_2(x_1, 0)$$

$\Longrightarrow$ global disruption

- P process :

$$Z_2^{P}(x_1, x_2) = \widetilde{Z}_2(x_1, x_2) - \kappa_2(x_2, 0)\widetilde{Z}_2(x_1, 0)$$

$\Longrightarrow$ local disruption

Table des matières

seqGPR model
- Kriking
- New Model
- Correction process

Parameters estimation

Applications
- Analytic example in dimension 4
- Industrial Test Cases

Conclusion

Estimation Problem

Estimation Problem

- Model : $Y_2(x_1, x_2) = Y_1(x_1) + Z_2(x_1, x_2)$ with
  - $Y_1 \sim \mathcal{PG}(m, k_1)$, k_1 stationary
  - $Z_2 \sim \mathcal{PG}(0, k_2)$, Red or P $\Rightarrow Z_2(x_1, 0) = 0$
  - $Z_2 \perp Y_1$

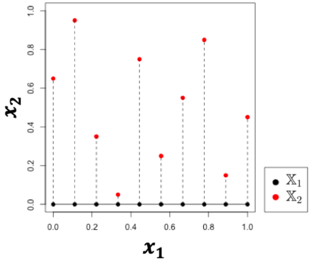
Estimation Problem

- Model : $Y_2(x_1, x_2) = Y_1(x_1) + Z_2(x_1, x_2)$ with
  - $Y_1 \sim \mathcal{PG}(m, k_1)$, k_1 stationary
  - $Z_2 \sim \mathcal{PG}(0, k_2)$, Red or P $\Rightarrow Z_2(x_1, 0) = 0$
  - $Z_2 \perp Y_1$
- We need to estimate $\eta = (\eta_1, \eta_2)$ with η_1 parameters of Y_1 and η_2 parameters of Z_2. We have $(\mathbb{X}_1, \boldsymbol{y}_1)$ and $(\mathbb{X}_2, \boldsymbol{y}_2)$.

Estimation Problem

- Model : $Y_2(x_1, x_2) = Y_1(x_1) + Z_2(x_1, x_2)$ with
  - $Y_1 \sim \mathcal{PG}(m, k_1)$, k_1 stationary
  - $Z_2 \sim \mathcal{PG}(0, k_2)$, Red or P $\Rightarrow Z_2(x_1, 0) = 0$
  - $Z_2 \perp Y_1$
- We need to estimate $\eta = (\eta_1, \eta_2)$ with η_1 parameters of Y_1 and η_2 parameters of Z_2. We have $(\mathbb{X}_1, \boldsymbol{y}_1)$ and $(\mathbb{X}_2, \boldsymbol{y}_2)$.
- Maximum of (Log-)likelihood $\mathcal{LL}(Y_1(\mathbb{X}_1) = \boldsymbol{y}_1, Y_2(\mathbb{X}_2) = \boldsymbol{y}_2; \eta)$

Nested Designs



Nested designs

Decoupled Log-likelihood:

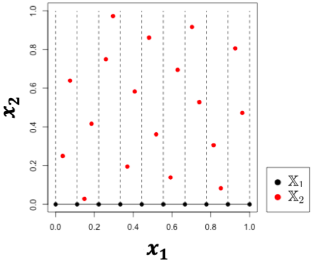
- η_1 : parameters off Y_1

$$\max_{\eta_1} \mathcal{LL}(Y_1(\mathbb{X}_1) = \boldsymbol{y}_1; \eta_1)$$

- η_2 : parameters de Z_2

$$\max_{\eta_2} \mathcal{LL}(Z_2(\mathbb{X}_2) = \boldsymbol{z}_2; \eta_2)$$

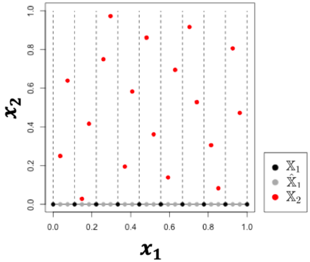
Non-Nested designs



Non-nested designs

Problem : $\mathcal{LL}(Y_1(\mathbb{X}_1) = \boldsymbol{y}_1, Y_2(\mathbb{X}_2) = \boldsymbol{y}_2; \eta)$ can't be decoupled.

Non-Nested designs

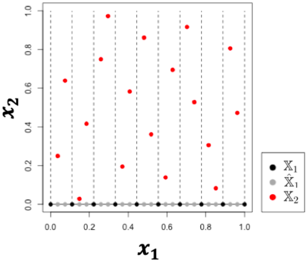


Nested designs

Problem : $\mathcal{LL}(Y_1(\mathbb{X}_1) = \boldsymbol{y}_1, Y_2(\mathbb{X}_2) = \boldsymbol{y}_2; \eta)$ can't be decoupled.

Complete data s.t $\mathbb{X}_2$ is nested in $\widetilde{\mathbb{X}}_1 = \mathbb{X}_1 \cup \hat{\mathbb{X}}_1$.

Non-Nested designs



Nested designs

Problem : $\mathcal{LL}(Y_1(\mathbb{X}_1) = \boldsymbol{y}_1, Y_2(\mathbb{X}_2) = \boldsymbol{y}_2; \eta)$ can't be decoupled.

Complete data s.t $\mathbb{X}_2$ **is nested in** $\widetilde{\mathbb{X}}_1 = \mathbb{X}_1 \cup \hat{\mathbb{X}}_1$.

$\Longrightarrow \mathcal{LL}(Y_1(\widetilde{\mathbb{X}}_1) = \widetilde{\boldsymbol{y}}_1, Y_2(\mathbb{X}_2) = \boldsymbol{y}_2; \eta)$ can be now decoupled in $\mathcal{LL}(Y_1(\widetilde{\mathbb{X}}_1) = \widetilde{\boldsymbol{y}}_1; \eta_1)\mathcal{LL}(Z_2(\mathbb{X}_2) = \boldsymbol{z}_2; \eta_2)$

Expectation-Maximization (EM) [Hastie et al., 2009]

EM algorithm is defined by sequences
$(\hat{\mathcal{L}\mathcal{L}}^{(i)}(Y_1(\widetilde{\mathbb{X}}_1) = \widetilde{\boldsymbol{y}}_1; \eta_1))_i$, $(\hat{\mathcal{L}\mathcal{L}}^{(i)}(Z_2(\mathbb{X}_2) = \boldsymbol{z}_2; \eta_2))_i$, et $(\eta^{(i)})_i$.

Expectation-Maximization (EM) [Hastie et al., 2009]

EM algorithm is defined by sequences
$(\hat{\mathcal{L}\mathcal{L}}^{(i)}(Y_1(\widetilde{\mathbb{X}}_1) = \widetilde{\boldsymbol{y}}_1; \eta_1))_i$, $(\hat{\mathcal{L}\mathcal{L}}^{(i)}(Z_2(\mathbb{X}_2) = \boldsymbol{z}_2; \eta_2))_i$, et $(\eta^{(i)})_i$.

At each step $i+1$,

Expectation-Maximization (EM) [Hastie et al., 2009]

EM algorithm is defined by sequences
$(\hat{\mathcal{L}\mathcal{L}}^{(i)}(Y_1(\tilde{\mathbb{X}}_1) = \tilde{\boldsymbol{y}}_1; \eta_1))_i$, $(\hat{\mathcal{L}\mathcal{L}}^{(i)}(Z_2(\mathbb{X}_2) = \boldsymbol{z}_2; \eta_2))_i$, et $(\eta^{(i)})_i$.

At each step $i+1$,

- Expectation :

- $\hat{\mathcal{L}\mathcal{L}}^{(i+1)}(Y_1(\tilde{\mathbb{X}}_1) = \tilde{\boldsymbol{y}}_1; \eta_1) = \mathbb{E}_{\eta^{(i)}}[\mathcal{L}\mathcal{L}(Y_1(\tilde{\mathbb{X}}_1) = \tilde{\boldsymbol{y}}_1; \eta_1) \mid Y_1(\mathbb{X}_1) = \boldsymbol{y}_1, Y_2(\mathbb{X}_2) = \boldsymbol{y}_2]$
- $\hat{\mathcal{L}\mathcal{L}}^{(i+1)}(Z_2(\mathbb{X}_2) = \boldsymbol{z}_2; \eta_2) = \mathbb{E}_{\eta^{(i)}}[\mathcal{L}\mathcal{L}(Z_2(\mathbb{X}_2) = \boldsymbol{z}_2; \eta_2) \mid Y_1(\mathbb{X}_1) = \boldsymbol{y}_1, Y_2(\mathbb{X}_2) = \boldsymbol{y}_2]$

Expectation-Maximization (EM) [Hastie et al., 2009]

EM algorithm is defined by sequences
$(\hat{\mathcal{LL}}^{(i)}(Y_1(\tilde{\mathbb{X}}_1) = \tilde{\boldsymbol{y}}_1; \eta_1))_i$, $(\hat{\mathcal{LL}}^{(i)}(Z_2(\mathbb{X}_2) = \boldsymbol{z}_2; \eta_2))_i$, et $(\eta^{(i)})_i$.

At each step $i+1$,

- Expectation :

- $\hat{\mathcal{LL}}^{(i+1)}(Y_1(\tilde{\mathbb{X}}_1) = \tilde{\boldsymbol{y}}_1; \eta_1) = \mathbb{E}_{\eta^{(i)}}[\mathcal{LL}(Y_1(\tilde{\mathbb{X}}_1) = \tilde{\boldsymbol{y}}_1; \eta_1) \mid Y_1(\mathbb{X}_1) = \boldsymbol{y}_1, Y_2(\mathbb{X}_2) = \boldsymbol{y}_2]$
- $\hat{\mathcal{LL}}^{(i+1)}(Z_2(\mathbb{X}_2) = \boldsymbol{z}_2; \eta_2) = \mathbb{E}_{\eta^{(i)}}[\mathcal{LL}(Z_2(\mathbb{X}_2) = \boldsymbol{z}_2; \eta_2) \mid Y_1(\mathbb{X}_1) = \boldsymbol{y}_1, Y_2(\mathbb{X}_2) = \boldsymbol{y}_2]$

- Maximization :

- $\eta_1^{(i+1)}$: $\max_{\eta_1} \hat{\mathcal{LL}}^{(i+1)}(Y_1(\tilde{\mathbb{X}}_1) = \tilde{\boldsymbol{y}}_1; \eta_1)$
- $\eta_2^{(i+1)}$: $\max_{\eta_2} \hat{\mathcal{LL}}^{(i+1)}(Z_2(\mathbb{X}_2) = \boldsymbol{z}_2; \eta_2)$

Table des matières

seqGPR model
- Kriking
- New Model
- Correction process

Parameters estimation

Applications
- Analytic example in dimension 4
- Industrial Test Cases

Conclusion

Function and data

Let the function

$$f(x_1, x_2, x_3, x_4) = f_1(x_1, x_2) + f_2(x_1, x_2, x_3, x_4)$$

with

$$\begin{cases} f_1(x_1, x_2) & = \left[4 - 2.1(4x_1 - 2)^2 + \frac{(4x_1-2)^4}{3}\right](4x_1 - 2)^2 \\ & + \quad (4x_1 - 2)(2x_2 - 1) + \left[-4 + 4(2x_2 - 1)^2\right](2x_2 - 1)^2 \\ f_2(x_1, x_2, x_3, x_4) & = 4\exp\left(-\left\|x - 0.3\right\|^2\right) \end{cases}$$

Function and data

Let the function

$$f(x_1, x_2, x_3, x_4) = f_1(x_1, x_2) + f_2(x_1, x_2, x_3, x_4)$$

with

$$\begin{cases} f_1(x_1, x_2) & = \left[4 - 2.1(4x_1 - 2)^2 + \frac{(4x_1-2)^4}{3}\right](4x_1 - 2)^2 \\ & + (4x_1 - 2)(2x_2 - 1) + \left[-4 + 4(2x_2 - 1)^2\right](2x_2 - 1)^2 \\ f_2(x_1, x_2, x_3, x_4) & = 4\exp\left(-\|x - 0.3\|^2\right) \end{cases}$$

- At step 1, let the restriction $f(x_1, x_2, \frac{x_1+x_2}{2}, 0.2x_1 + 0.7)$. Computer code evaluations at DoE $(\mathbb{X}_1, \mathrm{y}_1)$ of size 20.

Function and data

Let the function

$$f(x_1, x_2, x_3, x_4) = f_1(x_1, x_2) + f_2(x_1, x_2, x_3, x_4)$$

with

$$\begin{cases} f_1(x_1, x_2) & = \left[4 - 2.1(4x_1 - 2)^2 + \frac{(4x_1-2)^4}{3}\right](4x_1 - 2)^2 \\ & + \; (4x_1 - 2)(2x_2 - 1) + \left[-4 + 4(2x_2 - 1)^2\right](2x_2 - 1)^2 \\ & \\ f_2(x_1, x_2, x_3, x_4) & = 4\exp\left(-\|x - 0.3\|^2\right) \end{cases}$$

- At step 1, let the restriction $f(x_1, x_2, \frac{x_1+x_2}{2}, 0.2x_1 + 0.7)$. Computer code evaluations at DoE $(\mathbb{X}_1, y_1)$ of size 20.
- At step 2, new simulations of f at points DoE $(\mathbb{X}_2, y_2)$, a design in dimension 4 and of size 40.

Benchmark models

- **K_tot** : Kriging ($Y \sim \mathcal{PG}(m, k)$) trained on $(\mathbb{X}_1, \mathrm{y}_1)$ and $(\mathbb{X}_2, \mathrm{y}_2)$.

Benchmark models

- **K_tot** : Kriging ($Y \sim \mathcal{PG}(m, k)$) trained on $(\mathbb{X}_1, \mathrm{y}_1)$ and $(\mathbb{X}_2, \mathrm{y}_2)$.
- **K_2** : Kriging ($Y \sim \mathcal{PG}(m, k)$) trained on $(\mathbb{X}_2, \mathrm{y}_2)$.

Benchmark models

- **K_tot** : Kriging ($Y \sim \mathcal{PG}(m, k)$) trained on $(\mathbb{X}_1, \mathrm{y}_1)$ and $(\mathbb{X}_2, \mathrm{y}_2)$.
- **K_2** : Kriging ($Y \sim \mathcal{PG}(m, k)$) trained on $(\mathbb{X}_2, \mathrm{y}_2)$.
- **seqGPR** : Kriging (($Y_2 \sim \mathcal{PG}(m, k_1 + k_2)$) trained on $(\mathbb{X}_1, \mathrm{y}_1)$ and $(\mathbb{X}_2, \mathrm{y}_2)$.

$$\begin{cases} Y_2(x_1, x_2, x_3, x_4) & = & Y_1(x_1, x_2) + Z_2(x_1, x_2, x_3, x_4) \\ Z_2(x_1, x_2, \frac{x_1 + x_2}{2}, 0.2x_1 + 0.7) & = & 0, \end{cases}$$

  - $Y_1 \sim \mathcal{PG}(m, k_1)$
  - $Z_2 \sim \mathcal{PG}(0, k_2)$ a process **Red** or **P** built from $\widetilde{Z}_2$.
  - $\widetilde{Z}_2 \sim \mathcal{PG}(0, \kappa)$ independent of Y_1 with robust parametrization κ : $(\alpha, \alpha, \theta_3, \theta_4)$.

k, k_1, et κ are Matern $\frac{5}{2}$ tensor-product.

Results

RMSE on a Sobol sequence of size 10000. Median and interquartile range of RMSE on 100 samples $(\mathbb{X}_1, y_1)$ of size 20, $(\mathbb{X}_2, y_2)$ of size 40. **Red** is better.

	K_2	**K_tot**	**P**	**Red**
Médian	0.44	0.17	0.12	**0.07**
Interuqartile range	0.05	0.09	0.05	**0.02**

Results

RMSE on a Sobol sequence of size 10000. Median and interquartile range of RMSE on 100 samples $(\mathbb{X}_1, y_1)$ of size 20, $(\mathbb{X}_2, y_2)$ of size 40. **Red** is better.

	K_2	**K_tot**	**P**	**Red**
Médian	0.44	0.17	0.12	**0.07**
Interuqartile range	0.05	0.09	0.05	**0.02**

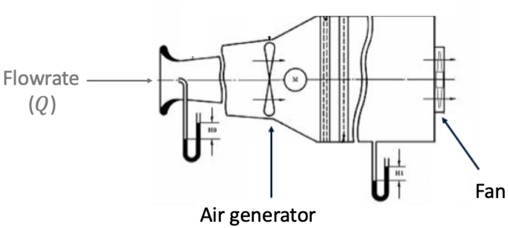
Others studies show that the size of the of the training sample has non influence.

Fan System in a cooling system

Input

Fan System in a cooling system

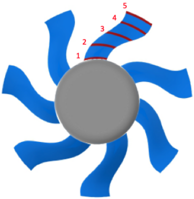
Input
Flowrate

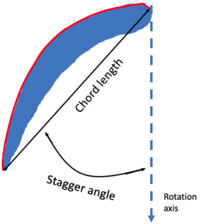


Fan System in a cooling system

Input

Flowrate
Stagger_1
Stagger_2
Stagger_3
Stagger_4
Stagger_5
LChord_1
LChord_2
LChord_3
LChord_4
LChord_5

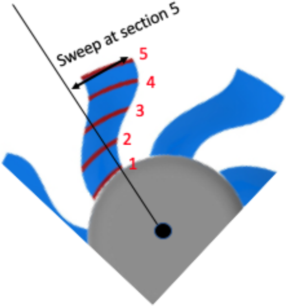




Fan System in a cooling system

Input

Flowrate
Stagger_1
Stagger_2
Stagger_3
Stagger_4
Stagger_5
LChord_1
LChord_2
LChord_3
LChord_4
LChord_5
Sweep_2
Sweep_3
Sweep_4
Sweep_5



Fan System in a cooling system

Input
Flowrate
Stagger_1
Stagger_2
Stagger_3
Stagger_4
Stagger_5
LChord_1
LChord_2
LChord_3
LChord_4
LChord_5
Sweep_2
Sweep_3
Sweep_4
Sweep_5

Output : ΔP

Data

Data

Step 1
Flowrate
Stagger_1
Stagger_2
Stagger_3
Stagger_4
Stagger_5
LChord_1
LChord_2
LChord_3
LChord_4
LChord_5
Sweep_2=1
Sweep_3=1
Sweep_4=0.82
Sweep_5=0.517645

Data

Step 1	Step 2
Flowrate	Flowrate
Stagger_1	Stagger_1
Stagger_2	Stagger_2
Stagger_3	Stagger_3
Stagger_4	Stagger_4
Stagger_5	Stagger_5
LChord_1	LChord_1
LChord_2	LChord_2
LChord_3	LChord_3
LChord_4	LChord_4
LChord_5	LChord_5
Sweep_2=1	Sweep_2
Sweep_3=1	Sweep_3
Sweep_4=0.82	Sweep_4
Sweep_5=0.517645	Sweep_5

Data

Step 1	**Step 2**
Flowrate	Flowrate
Stagger_1	Stagger_1
Stagger_2	Stagger_2
Stagger_3	Stagger_3
Stagger_4	Stagger_4
Stagger_5	Stagger_5
LChord_1	LChord_1
LChord_2	LChord_2
LChord_3	LChord_3
LChord_4	LChord_4
LChord_5	LChord_5
Sweep_2=1	Sweep_2
Sweep_3=1	Sweep_3
Sweep_4=0.82	Sweep_4
Sweep_5=0.517645	Sweep_5

- 30 models
- $\mathbb{X}_1$: 50 points in $[0,1]^{11}$
- $\mathbb{X}_2$: 50 points in $[0,1]^{15}$

Data

Step 1
Flowrate
Stagger_1
Stagger_2
Stagger_3
Stagger_4
Stagger_5
LChord_1
LChord_2
LChord_3
LChord_4
LChord_5
Sweep_2=1
Sweep_3=1
Sweep_4=0.82
Sweep_5=0.517645

Step 2
Flowrate
Stagger_1
Stagger_2
Stagger_3
Stagger_4
Stagger_5
LChord_1
LChord_2
LChord_3
LChord_4
LChord_5
Sweep_2
Sweep_3
Sweep_4
Sweep_5

- 30 models
- $\mathbb{X}_1$: 50 points in $[0,1]^{11}$
- $\mathbb{X}_2$: 50 points in $[0,1]^{15}$

	Median	**Range**
K_tot	44.6	5.9
P	**41.2**	4.7
Red	42.1	**4.6**

Table des matières

seqGPR model
- Kriking
- New Model
- Correction process

Parameters estimation

Applications
- Analytic example in dimension 4
- Industrial Test Cases

Conclusion

Conclusion

- Autoregressive model inspired by multi-fidelity models

Conclusion

- Autoregressive model inspired by multi-fidelity models
- Nullity constraints : 2 candidates (Red and P)

Conclusion

- Autoregressive model inspired by multi-fidelity models
- Nullity constraints : 2 candidates (Red and P)
- EM Algorithm to estimate the model EM pour estimer les paramètres

Conclusion

- Autoregressive model inspired by multi-fidelity models
- Nullity constraints : 2 candidates (Red and P)
- EM Algorithm to estimate the model EM pour estimer les paramètres
- Good results on the test case

Thank you for your attention

References I

Gauthier, B. and Bay, X. (2012).
Spectral approach for kernel-based interpolation.
Annales de la Faculté des sciences de Toulouse : Mathématiques, Ser. 6, 21(3):439–479.

Hastie, T., Tibshirani, R., and Friedman, J. (2009).
Model Inference and Averaging, chapter 8, pages 261–294.
Springer New York, NY.

Kennedy, M. C. and O'Hagan, A. (2000).
Predicting the output from a complex computer code when fast approximations are available.
Biometrika, 87(1):1–13.

References II

Santner, T. J., Williams, B. J., Notz, W., and Williams, B. J. (2003).
The design and analysis of computer experiments, volume 1.
Springer.