Journées Sensibilisation au Problème des Incertitudes
Gaussian Processes for code validation

François Bachoc

CEA-Saclay, DEN, DM2S, STMF, LGLS, F-91191 Gif-Sur-Yvette, France
Laboratoire de Probabilités et Modèles Aléatoires, Université Paris VII

April 2012

Introduction

- Context of code validation : Is the code in agreement with a set of reference experiments ?
- Gaussian processes validation : Modelling of the error between the code and the physical system.
- Goals :
  - Calibration of the code
  - Completion of the code by a statistical term based on a set of experiments

Least square calibration

Gaussian processes notions

Gaussian Processes Validation Model

Calibration and prediction

Model selection

Application to the thermohydraulic code Flica IV

Computation code and reality

A computation code, or parametric numerical model, is represented by a function f :

$$\begin{aligned} f \quad : \mathbb{R}^d \times \mathbb{R}^m \quad &\rightarrow \mathbb{R} \\ (x, \beta) \quad &\rightarrow f(x, \beta) \end{aligned}$$

The physical system is represented by a function Y_{real}.

$$\begin{aligned} Y_{real} \quad : \mathbb{R}^d \quad &\rightarrow \mathbb{R} \\ x \quad &\rightarrow Y_{real}(x) \end{aligned}$$

- The inputs x are the experimental conditions.
- The inputs β are the calibration parameters of the computation code.
- The outputs $f(x, \beta)$ and $Y_{real}(x)$ are the quantity of interest.

A computation code models (gives an approximation of) a physical phenomenon.

Least square calibration

We dispose of a set of experimental results : x_1, $Y_{obs}(x_1)$, ..., x_n, $Y_{obs}(x_n)$.

Least Square calibration :

- Compute :
$$\hat{\beta}_{LS} \in \arg\min_{\beta} \sum_{i=1}^{n} (f(x_i, \beta) - Y_{obs}(x_i))^2$$
- For new experimental condition x_{new} we predict the quantity of interest by : $f(x_{new}, \hat{\beta}_{LS})$.

Least square calibration as a Maximum Likelihood estimation

Assume that there exist β so that for $1 \leq i \leq n$

$$Y_{obs}(x_i) = f(x_i, \beta) + z_i$$

With $z_i \sim_{iid} \mathcal{N}(0, \sigma^2)$. z_i's are either

- A measure error.
- A measure error and a model error

Then the maximum likelihood estimator $\hat{\beta}$ of β is

$$\begin{aligned}
\hat{\beta} \quad &\in \quad \arg\max_{\beta} \frac{1}{(2\pi)^{\frac{n}{2}}(\sigma^2)^n} \exp\left(-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n}(f(x_i,\beta) - Y_{obs}(x_i))^2\right)\right) \\
&\in \quad \arg\min \beta \sum_{i=1}^{n}(f(x_i,\beta) - Y_{obs}(x_i))^2
\end{aligned}$$

Least Square calibration : Case of insufficiency

When $\boldsymbol{z}_i$'s are only measure error :

- Problem when σ^2 (or an upper-bound) is known and when the errors $f(x_i, \hat{\beta}) - Y_{obs}(x_i)$ are too large. (Statistical tests available to detect).

When $\boldsymbol{z}_i$'s are model errors.

- The physical system $x \longrightarrow f(x, \beta) + z$ would be discontinuous
- It is reasonable to assume correlation between the model errors at two neighboor points

$\longrightarrow$ A Gaussian process model is a way to answer these issues

Least square calibration

Gaussian processes notions

Gaussian Processes Validation Model

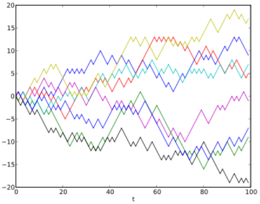
Calibration and prediction

Model selection

Application to the thermohydraulic code Flica IV

Gaussian processes (1/3)

A random function is a function $x \rightarrow F(x)$ such that $F(x)$ is a random variable. Alternalively a random function is a function that is unknown, or that depends of the hasard.



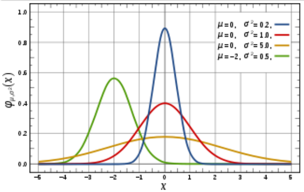
Finite dimensional distributions of a random function

Let us consider n points of $\mathbb{R}^d$: $x_1, \ldots, x_n$. By definition, the vector $(Z(x_1), \ldots, Z(x_n))$ is a random vector of $\mathbb{R}^n$. Its distribution is said to be a finite dimensional distribution of Z.

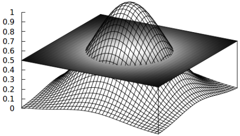
The finite dimensional distributions of Z are the set of these distributions with n et $x_1, \ldots, x_n$ varying.

Gaussian variables and vectors

A random variable is a Gaussian variable with mean μ and variance σ^2 when its probability density function is $f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right)$



A n dimensional random vector is a Gaussian vector with mean vector μ and covariance matrix R when its multidimensional probability density function is $f(x) =$
$\frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{det(R)}} \exp\left(-\frac{1}{2}(x-\mu)^t R^{-1}(x-\mu)\right)$



Gaussian processes (1/2)

A random function Z on $\mathbb{R}^d$ is a Gaussian process when its finite dimensional distributions are Gaussian.

In the sequel, we only consider Gaussian processes :

- Gaussian variables : most commonly used to represent errors.
- Gaussian properties make the treatment of the problem simpler.

Mean function $M : x \rightarrow M(x) = \mathbb{E}(Z(x))$
Covariance function $C : (x_1, x_2) \rightarrow C(x_1, x_2) = cov(Z(x_1), Z(x_2))$

- A Gaussian process is caracterized by its mean and covariance functions.

Gaussian processes (2/2)

Examples of covariance functions

Nugget covariance function $C(x,y) = \sigma^2 \mathbf{1}_{x=y}$

Gaussian covariance function $C(x,y) = \sigma^2 \exp\left(-\frac{(x-y)^2}{l_c^2}\right)$

Exponential covariance function $C(x,y) = \sigma^2 \exp\left(-\frac{|x-y|}{l_c}\right)$

Examples of realizations with Gaussian covariance function

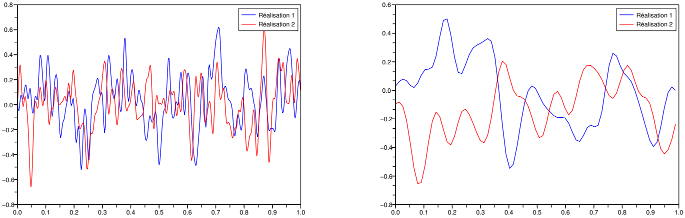


FIG.: Left : $\sigma = 0.2$, $l_c = 0.01$. Right : $\sigma = 0.2$, $l_c = 0.05$

Conditional distributions for Gaussian vectors

Theorem

Let

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} m_1 \\ m_2 \end{pmatrix}, \begin{pmatrix} R_1 & R_{1,2} \\ R_{2,1} & R_2 \end{pmatrix}\right)$$

Then, conditionaly on $X_2 = x_2$, X_1 is a Gaussian vector with

$$\mathbb{E}(X_1|X_2 = x_2) = m_1 + R_{1,2}R_{2,2}^{-1}(x_2 - m_2)$$

and

$$cov(X_1|X_2 = x_2) = R_1 - R_{1,2}R_{2,2}^{-1}R_{2,1}$$

Illustration

Let

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)$$

Then

$$\mathbb{E}(X_1|X_2 = x_2) = \rho x_2$$

and

$$var(X_1|X_2 = x_2) = 1 - \rho^2$$

Least square calibration

Gaussian processes notions

Gaussian Processes Validation Model

Calibration and prediction

Model selection

Application to the thermohydraulic code Flica IV

Model error

Statistical modelling : The physical phenomenon is one realization among a set of possible realizations. It is modeled as a realization of a random process.

Equation of the statistical model

$$Y_{real}(\omega, x) = f(x, \beta(\omega)) + Z(\omega, x)$$

- Equation that holds for a specific parameters vector β. Called "the" parameter of the numerical code.
  - No prior information case : β constant and unknown.
  - Prior information case (Bayesian case) : $\beta \sim \mathcal{N}(\beta_{prior}, Q_{prior})$
- Z is (a priori) a centered, stationary, Gaussian process. We denote by C_{mod} the covariance function of Z.

Steps of Gaussian processes code validation

- Step 1 : Estimation of the covariance function for the model error.
- Step 2 : With a given covariance function : calibration and prediction.
  - Calibration : gives a "posterior mean value" for the code parameter β and a "posterior variance".
  - Prediction : for a new experimental condition x_{new}, gives a "posterior mean value" for $Y_{real}(x_{new})$ and a "posterior variance".

Linear code and observations : notations

Linearization of the numerical model around the reference parameter :

$$\forall x : \; f(x, \beta) = \sum_{i=1}^{m} h_i(x)\beta_i$$

Observations

We observe the physical phenomenon $Y_{real}(x)$ for n inputs $x_1, \ldots, x_n$.
Define :

- $n \times m$ matrix of partial derivatives of the numerical model : H.
- Random vector of observations : y_{obs}.
- Random vector of measure error : ϵ.
- Random vector of model error : z.
- Covariance matrix of z : R_{mod}.

Matrix equation of the statistical model

The statistical model becomes, for the inputs $x_1, ..., x_n$:

$$y_{obs} = H\beta + z + \epsilon$$

Covariance matrix of $z + \epsilon$

$$R := cov(z + \epsilon) = R_{mod} + K$$

With $K := cov(\epsilon)$. K is diagonal. Most classical case : $K = \sigma^2_{mes} I$.

- No prior information case
  - When $R = \sigma^2 \mathbf{I}_n$: Classical linear regression model.
- Prior information case

$$y_{obs} \sim \mathcal{N}(H\beta_{prior}, R + HQ_{prior}H^T)$$

- Main interest of the correlation : Efficient prediction of the phenomenon when it does not have the same shape as the numerical code.

Least square calibration

Gaussian processes notions

Gaussian Processes Validation Model

Calibration and prediction

Model selection

Application to the thermohydraulic code Flica IV

Introduction

Assume we have fixed the covariance function C_{mod} of the model error.

- The statistical model is a linear regression model with a Gaussian process error.
- It is the same as the Kriging Model, well known e.g in Geostatistic and in analysis of computer experiments.
- We have closed form formulas for the calibration and the prediction.

Calibration (1/2)

Calibration problem = Statistical estimation problem

Estimation of β

- An estimator of β is a function $\hat{\beta} : \mathbb{R}^n \longrightarrow \mathbb{R}^m$.
- $\hat{\beta}(y_{obs})$ is the estimation of β according to the vector of observations y_{obs}.
- $\hat{\beta}(y_{obs})$ is a random variable because y_{obs} is a random variable.
- Quality measure of an estimator : Mean square error : $\mathbb{E}_{y_{obs},\beta}\left[||\beta - \hat{\beta}(y_{obs})||^2\right]$.

Calibration (1/2)

No prior information case

The maximum likelihood estimator of β is

$$\hat{\beta} = (H^T R^{-1} H)^{-1} H^T R^{-1} y_{obs}$$

The covariance matrix of $\hat{\beta}$ is

$$cov(\hat{\beta}) = (H^T R^{-1} H)^{-1}$$

- $\hat{\beta}$ is unbiased : $\mathbb{E}(\hat{\beta}) = \beta$
- If $y_{obs} = H\beta$, $\hat{\beta} = \beta$

Calibration (2/2)

Prior information case

Recall the a priori probability law of β is normal with mean vector β_{prior} and covariance matrix Q_{prior}. Conditionally to the observations y_{obs}, β is Gaussian with mean vector β_{post} and covariance matrix Q_{post}.

$$\beta_{post} = \beta_{prior} + (Q_{prior}^{-1} + H^T R^{-1} H)^{-1} H^T R^{-1} (y_{obs} - H\beta_{prior}).$$

$$Q_{post} = (Q_{prior}^{-1} + H^T R^{-1} H)^{-1}$$

- Best predictor according to the mean square error.
- When $Q_{prior}^{-1} \rightarrow 0$ (Uninformative prior) we find the prediction of the no prior information case, even if $\beta_{prior} \neq 0$.
- When $Q_{prior} \rightarrow 0$ $\beta_{post} \rightarrow \beta_{prior}$.

Prediction (1/4)

Goal : to complete the prediction of $f(x_0, \hat{\beta})$ at a new point x_0.

Notations

- Physical phenomenon at x_0 : $y_0 := Y_{real}(x_0)$.
- (pseudo) new observation at x_0 : $y_{obs,0}$.
- Column vector of partial derivatives of the code : h_0.
- Random variable of the model error : z_0.
- Random variable of the measure error : ϵ_0.
- Column covariance vector r_0 : $r_{0,i} := cov((z+\epsilon)_i, z_0+\epsilon_0)$.

Prediction of y_0

- A predictor of y_0 is a function $\langle y_0 \rangle : \mathbb{R}^n \to \mathbb{R}$.
- $\langle y_0 \rangle(y_{obs})$ is the prediction of y_0 according to the vector of observations y_{obs}.
- $\langle y_0 \rangle(y_{obs})$ is a random variable because y_{obs} is a random variable.
- Quality measure of a predicor : Mean square error : $\mathbb{E}_{y_{obs},y_0}\left[|y_0 - \langle y_0 \rangle(y_{obs})|^2\right]$.

Prediction (2/4) : No prior information case

Prediction

The unbiased predictor of $y_{obs,0}$ at x_0, linear with respect to the vector of observations y_{obs}, which minimizes the mean square error (the BLUP) is :

$$\langle y_{obs,0} \rangle = (h_0)^T \hat{\beta} + (r_0)^T R^{-1} (y_{obs} - H\hat{\beta})$$

with $\hat{\beta}$ the no prior information case estimator of β.

- We do not have access to the best predictor, because its expression makes use of the unknown parameter β.
- The prediction expression is decomposed into a calibration term and a Gaussian inference term of the model error.

Predictive variance

The mean square error of the BLUP is :

$$\hat{\sigma}^2_{x_0} = \mathbb{E}((z_0 + \epsilon_0)^2) - \begin{pmatrix} h_0 \\ r_0 \end{pmatrix}^t \begin{pmatrix} 0 & H^t \\ H & R \end{pmatrix}^{-1} \begin{pmatrix} h_0 \\ r_0 \end{pmatrix}$$

Confidence intervals available

Prediction (3/4) : Prior information case

Prediction

The conditional law of $y_{obs,0}$ according to the observations y_{obs} is Gaussian with mean $\langle y_{obs,0} \rangle$, with :

$$\langle y_{obs,0} \rangle = (h_0)^T \beta_{post} + (r_0)^T R^{-1} (y_{obs} - H\beta_{post})$$

- Best predictor.

Predictive variance

Conditionally to y_{obs} the variance of $y_{obs,0}$ is :

$$\hat{\sigma}^2_{x_0} = \mathbb{E}((z_0 + \epsilon_0)^2) - \begin{pmatrix} h_0 \\ r_0 \end{pmatrix}^t \begin{pmatrix} -Q^{-1}_{prior} & H^t \\ H & R \end{pmatrix}^{-1} \begin{pmatrix} h_0 \\ r_0 \end{pmatrix}$$

- When $Q^{-1}_{prior} \to 0$ (uninformative prior) we find the no prior information case.

Prediction (4/4) : from $y_{obs,0}$ to y_0

No prior information case

The BLUP of the observation equals the BLUP of the physical phenomenon :

$$\forall \lambda \in \mathbb{R}^n : \quad \mathbb{E}\left((\lambda^t y_{obs} - y_{obs,0})^2 \right) = \mathbb{E}\left((\lambda^t y_{obs} - y_0)^2 \right) + \mathbb{E}\left((\epsilon_0)^2 \right)$$

Prior information case

The conditional means are the same and the conditional variances are the same up to the measure error :

- $\mathbb{E}(y_0 | y_{obs}) = \mathbb{E}(y_{obs,0} | y_{obs})$
- $var(y_{obs,0} | y_{obs}) = var(y_0 | y_{obs}) + \mathbb{E}\left((\epsilon_0)^2 \right)$

$\boxed{\longrightarrow}$ In both cases, we keep the same prediction, and remove $\mathbb{E}\left((\epsilon_0)^2 \right)$ to the predictive variance.

Illustration of calibration (1/3)

- Observation of the physical phenomenon : $Y_{obs}(x) = x^2 + \epsilon$. $\epsilon \sim \mathcal{N}(0, \sigma^2_{mes} = 0.1^2)$
- Numerical code : $f(x, \beta) = \beta_0 + \beta_1 x$.
- Model error as a realization of a Gaussian process with covariance function : $C_{mod}(x-y) = \sigma^2 \exp\left(-\frac{|x-y|^2}{l_c^2}\right)$. $\sigma = 0.3$, $l_c = 0.5$ (known).
- Bayesian case with :

$$\beta_{prior} = \begin{pmatrix} 0.2 \\ 1 \end{pmatrix}, Q_{prior} = \begin{pmatrix} 0.09 & 0 \\ 0 & 0.09 \end{pmatrix}$$

- Observations : $x_1 = 0.2$, $x_2 = 0.4$, $x_3 = 0.6$ and $x_4 = 0.8$.

Illustration of calibration (2/3) (unnoised case)

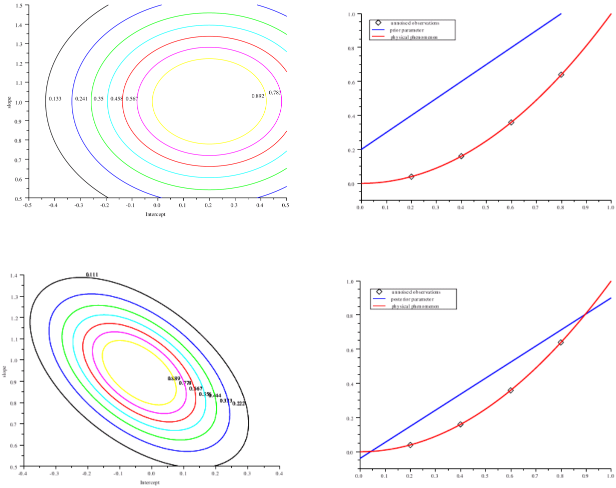


FIG.: Up-left : Prior distribution of the parameter β. Down-left : Posterior distribution of the parameter β. Right : plot of the code response corresponding to prior and posterior mean of the code parameter.

Illustration of calibration (3/3) (noised case)

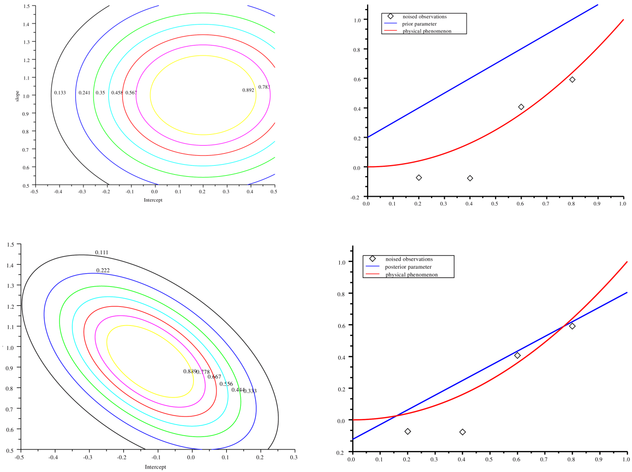
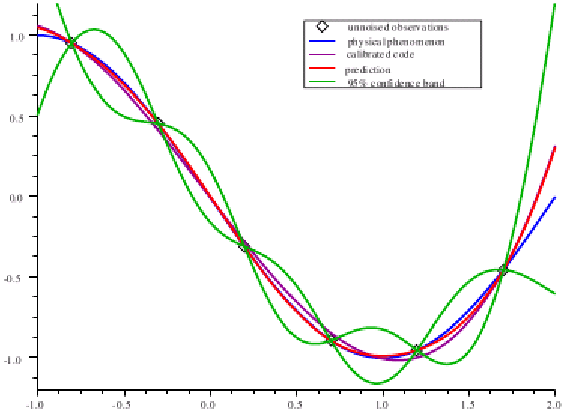


FIG.: Up-left : Prior distribution of the parameter β. Down-left : Posterior distribution of the parameter β. Right : plot of the code response corresponding to prior and posterior mean of the code parameter.

Illustration of prediction (1/3)

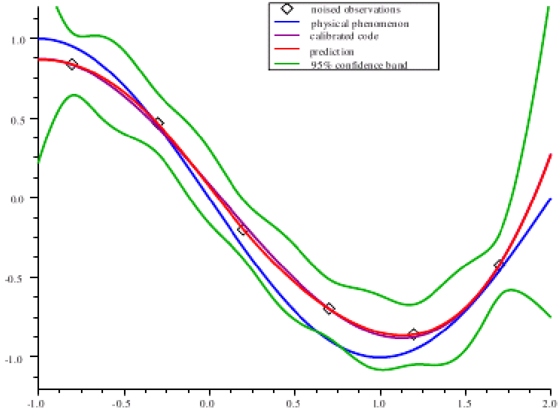
- Observation of the physical phenomenon : $Y_{obs}(x) = -sin(\frac{\pi x}{2}) + \epsilon$. $\epsilon \sim \mathcal{N}(0, \sigma_{mes}^2 = 0.1^2)$
- Numerical code : $f(x, \beta) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$.
- Model error as a realization of a Gaussian process with covariance function : $C_{mod}(x - y) = \sigma^2 \exp\left(-\frac{|x-y|^2}{l_c^2}\right)$. $\sigma = 0.3$, $l_c = 0.5$ (known).
- No prior information case.
- 6 observations regularly sampled between -0.8 and 1.7.

Illustration of prediction (2/3) (unnoised case)



- The use of the model error improves the prediction given by the numerical code.

Illustration of prediction (3/3) (noised case)



- The measure error deteriorates the quality of the predictions.
- The confidence intervals are however still reliable.

Least square calibration

Gaussian processes notions

Gaussian Processes Validation Model

Calibration and prediction

Model selection

Application to the thermohydraulic code Flica IV

Framework

- The calibration and prediction methods presented above give good results because we used a reasonable covariance function.
- The model selection is a statistical parameter estimation problem.

In our case, the covariance function of the measure error process ϵ is known for physical expertise. We want to take C_{mod} in a parametric set :

$$\left\{\sigma^2 C_{mod,\theta}\right\}$$

with $C_{mod,\theta}$ a correlation function.
Hence, with variance matrix $R_{\sigma,\theta} = \sigma^2 R_{mod,\theta} + K$, we have $(z+\epsilon) \sim \mathcal{N}(0, R_{\sigma,\theta})$ and we want to estimate σ and θ.
We present 2 methods for model selection : Restricted Maximum Likelihood and Leave One Out.

Restricted Maximum Likelihood

Principle : Estimate σ and θ independently of β (hence, same method with or without prior information).
Let C a $(n - m \times n)$ matrix of maximal rank such that $CH = 0$. Then we have :

$$w := Cy_{obs} \sim \mathcal{N}(0, CR_{\sigma,\theta}C')$$

We do maximum likelihood on the vector w.
The likelihood writes itself :

$$\ell_{\sigma,\theta}(w) \propto \frac{1}{det(CR_{\sigma,\theta}C^t)^{\frac{1}{2}}} \exp\left(-\frac{1}{2} w^t (CR_{\sigma,\theta}C^t)^{-1} w\right)$$

We maximize it :

$$\hat{\sigma}, \hat{\theta} \in \arg\max_{\sigma,\theta} \ell_{\sigma,\theta}(w).$$

Hence we estimate σ and θ to make the vector w the most probable.

Leave One Out (1/4)

We have seen that the prediction procedure (Bayesian or non-Bayesian framework) leads to a simple stochastic metamodel :
$x_0 \rightarrow \mathcal{N}\left(\langle y_{obs,0}\rangle, \hat{\sigma}^2_{x_0}\right)$. This metamodel depends on σ and θ.

- It is built according to the observations ($\approx$ learning set).

Leave One Out

- Given a vector of hyper-parameters (σ, θ).
- For i from 1 to n we learn $x_0 \rightarrow \mathcal{N}\left(\langle y_{obs,0}\rangle, \hat{\sigma}^2_{x_0}\right)$ with the reduced observations vector
$\{(x_1, y_{obs,1}), \ldots, (x_{i-1}, y_{obs,i-1}), (x_{i+1}, y_{obs,i+1}), \ldots, (x_n, y_{obs,n})\}$
- we compute the LOO errors by :

$$\epsilon_{LOO,i}(\sigma, \theta) = y_{obs,i} - \langle y_{obs,i}\rangle(y_{obs,-i}).$$

- we compute the LOO predictive variance by :

$$\hat{\sigma}^2_{LOO,i}(\sigma, \theta) = \hat{\sigma}^2_{x_i}(y_{obs,-i})$$

General utility of the Leave One Out :

- See how large the errors are.
- Check that the predictive variance are of the right size.

Leave One Out (2/4) : closed form formulas

No prior information case

With :

$$Q^{-}(\sigma,\theta) = \left(R^{-1}_{\sigma,\theta} - R^{-1}_{\sigma,\theta}H(H^T R^{-1}_{\sigma,\theta}H)^{-1}H^T R^{-1}_{\sigma,\theta}\right)$$

We have :

$$\epsilon_{LOO}(\sigma,\theta) = (diag(Q^{-}))^{-1}Q^{-}y_{obs} \quad \text{and} \quad \hat{\sigma}^2_{LOO,i}(\sigma,\theta) = \frac{1}{(Q^{-})_{i,i}}$$

Prior information case

With :

$$Q = R_{\sigma,\theta} + HQ_{prior}H^t$$

We have :

$$\epsilon_{LOO}(\sigma,\theta) = (diag(Q^{-1}))^{-1}Q^{-1}y_{obs} \quad \text{and} \quad \hat{\sigma}^2_{LOO,i}(\sigma,\theta) = \frac{1}{(Q^{-1})_{i,i}}$$

Leave One Out (3/4) : closed form formulas

- The no prior information case is the limit of the prior information case when $Q_{prior}^{-1} \rightarrow 0$.
- From a computational point of view : computing the LOO errors and predictive variance has the same order of complexity than REML and Maximum Likelihood.
- $\boxed{\longrightarrow}$ Can be use as an alternative of Maximum Likelihood techniques.

Leave One Out (4/4) : A model selection method

General principle, optimize a quality criterion based on $\epsilon_{LOO}(\sigma,\theta)$ and the $\hat{\sigma}^2_{LOO,i}$. For instance :

- Minimize norm of LOO errors.
- Set number of valid LOO p-confidence intervals close to p.
- Set $\frac{1}{n}\sum_{i=1}^{n}\frac{\epsilon^2_{LOO,i}(\sigma,\theta)}{\hat{\sigma}^2_{LOO,i}(\sigma,\theta)}$ close to 1

When the covariance matrix K of the measure error is null and no prior information case, we have $R_{\sigma,\theta}=\sigma^2 R_{mod,\theta}$, hence :

- $\epsilon_{LOO}(\sigma,\theta)$ independent of σ
- $\hat{\sigma}^2_{LOO}(\sigma,\theta)=\sigma^2\hat{\sigma}^2_{LOO}(\theta)$

Hence a classical method is :

$$\hat{\theta}\in\arg\min_{\theta}||\epsilon_{LOO}(\theta)||^2 \quad \text{and} \quad \hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}\frac{\epsilon^2_{LOO,i}(\hat{\theta})}{\hat{\sigma}^2_{LOO,i}(\hat{\theta})}$$

When $K\neq 0$ or prior information case : no classical method.

Least square calibration

Gaussian processes notions

Gaussian Processes Validation Model

Calibration and prediction

Model selection

Application to the thermohydraulic code Flica IV

The experimental results for the friction model

The experiment consists in pressurized and possibly heated water passing through a cylinder. We measure the pressure drop between the two ends of the cylinder.
Quantity of interest : The part of the pressure drop due to friction : ΔP_{fro}
Two kinds of experimental conditions :

- System parameters : Hydraulic diameter D_h, Friction height H_f, Channel width e.
- Environment variables : Output pressure P_s, Flowrate G_e, Parietal heat flux Φ_p, Liquid enthalpy h_e^l, Thermodynamic title X_{th}^e, Input temperature T_e.

We dispose of 253 experimental results. 115 are in the isothermal domain and 138 in the monophasic (non-isothermal) domain.
Important : Among the 253 experimental results, only 8 different system parameters $\rightarrow$ Not enough to use the Gaussian processes model for prediction for new system parameters $\rightarrow$ We predict for new environment variables only.

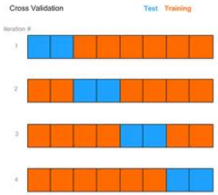
The Flica IV code for the friction model

Parameterized a_t and b_t.
Prior information (coming from previous studies) :

$$\beta_{prior} = \begin{pmatrix} 0.22 \\ 0.21 \end{pmatrix}, Q_{prior} = \begin{pmatrix} 0.11^2 & 0 \\ 0 & 0.105^2 \end{pmatrix}$$

Results

We compare predictions to observations using Cross Validation



We dispose of :

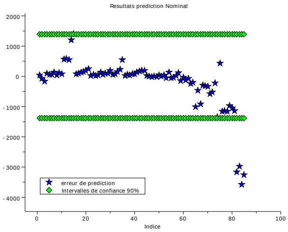
- The vector of posterior mean $\Delta\hat{P}_{fro}$ of size n.
- The vector of posterior variance σ^2_{pred} of size n.

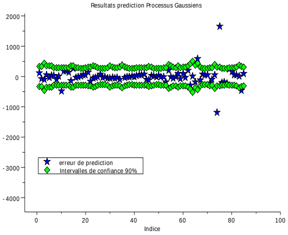
2 quantitative criteria :

- RMSE : $\sqrt{\frac{1}{n}\sum_{i=1}^{85}\left(\Delta P_{fro,i} - \Delta\hat{P}_{fro,i}\right)^2}$
- Confidence Intervals : proportion of observations that fall in the posterior 90% confidence interval.

Results : isothermal domain

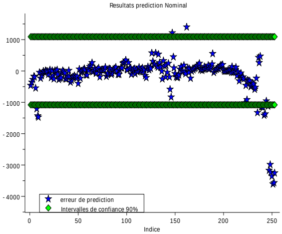
	RMSE	Confidence Intervals
Nominal code	$840Pa$	$80/85 \approx 0.94$
Gaussian Processes	$265Pa$	$79/85 \approx 0.93$

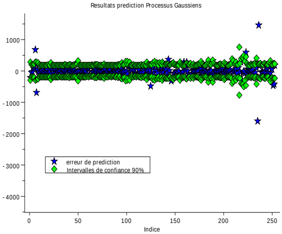




Results : isothermal and monophasic domain

	RMSE	Confidence Intervals
Nominal code	$661\,Pa$	$234/253 \approx 0.925$
Gaussian Processes	$189\,Pa$	$235/253 \approx 0.93$





Conclusion

- We can improve the prediction capability of the code by completing it with a statistical model based on the experimental results.
- Number of experimental results needs to be sufficient. No extrapolation.
- The choice of the covariance function is important.

Increasing use of probabilistic methods for numerical simulation : Kriging and Gaussian processes methods for surrogate models and code calibration and validation.

Some references

F Bachoc, G Bois, and J.M Martinez.
Contribution à la validation des codes de calcul par processus Gaussiens. Application à la calibration du modéle de frottement pariétal de flica 4.
Technical report, Commissariat à l'Energie Atomique et aux Energies Alternatives DEN/DANS/DM2S/STMF, 2012.

J Garnier.
Analyse bayésienne pour les simulations numériques.
Technical report, Commissariat à l'Energie Atomique et aux Energies Alternatives CEA/DIF/DCSA/SSEL, 2009.

M.L Stein.
Interpolation of Spatial Data Some Theory for Kriging.
Springer, 1999.

T.J Santner, B.J Williams, and W.I Notz.
The Design and Analysis of Computer Experiments.
Springer, 2003.

E. Vazquez.
Modélisation comportementale de systèmes non-linéaires multivariables par méthodes à noyaux et applications.
PhD thesis, Université Paris XI Orsay, 2005.