DE LA RECHERCHE À L'INDUSTRIE

# Calibration of nested phenomena

**G. Perrin**(1), J. Garnier(2), S. Marque-Pucheu(1,2)

(1) CEA/DAM/DIF, F-91297, Arpajon, France,
(2) Laboratoire de Probabilités et Modèles Aléatoires, Laboratoire Jacques-Louis Lions, Université Paris Diderot, 75205 Paris Cedex 13, France

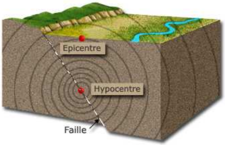
Ecole thématique ETICS - Barcelonnette |

June 2016

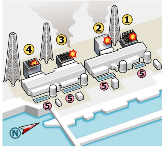
# Introduction

Two classical examples of nested codes in reliability analysis:

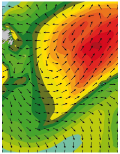
1. a nuclear power plant under seismic accelerations:



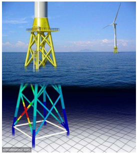

$\rightarrow$



2. an off-shore structure that weathers a storm:



$\rightarrow$



# Notations and formalism

$$x_1 \in \mathbb{R}^{d_1} \quad \overset{g_1,\beta_1,\varepsilon_1}{\longrightarrow} \quad y_1 = x_2 \in \mathbb{R}^{d_2} \quad \overset{g_2,\beta_2,\varepsilon_2}{\longrightarrow} \quad y_2$$

$$\begin{cases} y_1(x_1) = g_1(x_1;\beta_1) + \varepsilon_1(x_1), \\ y_2(x_2) = g_2(x_2;\beta_2) + \varepsilon_2(x_2). \end{cases}$$

- $g_1(\cdot;\beta_1), g_2(\cdot;\beta_2) \leftrightarrow$ available codes,
- $\varepsilon_1, \varepsilon_2 \leftrightarrow$ model errors,
- $(x_1^{(n)}, y_1^{(n)})_{1\leq n\leq N}$, $(x_2^{(m)}, y_2^{(m)})_{1\leq m\leq M} \leftrightarrow$ available observations,
- some values of $y_1^{(n)}$ can be equal to the values of $x_2^{(m)}$, such that we can have access to the results of the complete chain of codes.

## Problematic

1. Compute the posterior distribution of $\beta_c = (\beta_1, \beta_2)$,
2. Infer the distribution of $y_2$ in any non-computed point $x_1$.

## Hypotheses:

- there exists a **unique** "true" value of $\beta_c = (\beta_1, \beta_2)$,
- the experimental errors are not taken into account in this presentation,
- the model errors are supposed to be **small** compared to the responses of the calibrated codes, $\|\epsilon_1\|_{L^2} \ll \|g_1(\cdot;\beta_1)\|_{L^2}$, $\|\epsilon_2\|_{L^2} \ll \|g_2(\cdot;\beta_2)\|_{L^2}$,
- the model errors are supposed to be independent, and are modeled by (potentially vector-valued) centered Gaussian processes, which covariances are supposed to be **known** and are written $C_1$ and $C_2$ respectively,
- the second code can take additional inputs, but for the sake of clarity, these inputs do not appear in the following.

# Calibration with fixed physical system

## Bayesian framework

$$\beta_1 \sim \mathcal{U}_{\mathbb{R}^{d_1}}, \quad \beta_2 \sim \mathcal{U}_{\mathbb{R}^{d_2}}, \quad \beta_1 \perp \beta_2.$$

Gaussian priors can also be considered.

## Linearization

There exists **fixed** and **known** nominal values of the parameters, $\beta_1^{\mathrm{nom}}$ and $\beta_2^{\mathrm{nom}}$, such that:

$$g_1(x_1;\beta_1) \approx f_1(x_1)\beta_1,$$
$$g_2(x_2;\beta_2) \approx f_2(x_2)\beta_2.$$

The nested phenomenon we are interested in can therefore be written as:

$$y_2(x_1) = f_2\left(f_1(x_1)\beta_1 + \varepsilon_1(x_1)\right)\beta_2 + \varepsilon_2\left(f_1(x_1)\beta_1 + \varepsilon_1(x_1)\right).$$

# Outline

# Coupling two GPRs

From the available codes evaluations, it is possible to construct two **independent** conditioned GPR for the two nested codes:

- $\widehat{y}_1(x_1) \sim \mathcal{N}(\mu_N^1(x_1), \sigma_N^1(x_1)) \leftrightarrow \widehat{y}_1(x_1) = \mu_N^1(x_1) + u_1 \times \sigma_N^1(x_1)$,
- $\widehat{y}_2(x_2) \sim \mathcal{N}(\mu_M^2(x_2), \sigma_M^2(x_2)) \leftrightarrow \widehat{y}_2(x_2) = \mu_M^2(x_2) + u_2 \times \sigma_M^2(x_2)$,

where $u_1, u_2$ are two independent centered normalized Gaussian r.v. We denote by $\phi$ the centered normalized Gaussian PDF.

The predictor of the nested phenomenon, $Y^{\text{nest}} := \widehat{y}_2(\widehat{y}_1(x_1))$, verifies therefore:

$$\begin{aligned} Y^{\text{nest}}(x_1, u_1, u_2) = \mu_M^2 \left(\mu_N^1(x_1) + u_1 \times \sigma_N^1(x_1)\right) \\ + u_2 \times \sigma_M^2 \left(\mu_N^1(x_1) + u_1 \times \sigma_N^1(x_1)\right). \end{aligned}$$

# Coupling two GPRs

$$
\begin{aligned}
Y^{\text{nest}}(x_1, u_1, u_2) = \mu_M^2 &\left(\mu_N^1(x_1) + u_1 \times \sigma_N^1(x_1)\right) \\
&+ u_2 \times \sigma_M^2 \left(\mu_N^1(x_1) + u_1 \times \sigma_N^1(x_1)\right).
\end{aligned}
$$

The first and second statistical moments of $Y^{\text{nest}}(x_1, u_1, u_2)$, which is *a priori* not Gaussian, are given by:

$$
\begin{aligned}
\mathbb{E}_{u_1,u_2}\left[Y^{\text{nest}}(x_1, u_1, u_2)\right] &= \int_{\mathbb{R}\times\mathbb{R}} Y^{\text{nest}}(x_1, u_1, u_2)\phi(u_1)\phi(u_2)du_1du_2 \\
&= \mathbb{E}_{u_1}\left[\mu_M^2\left(\mu_N^1(x_1) + u_1 \times \sigma_N^1(x_1)\right)\right].
\end{aligned}
$$

$$
\mathbb{E}_{u_1,u_2}\left[\left(Y^{\text{nest}}(x_1, u_1, u_2)\right)^2\right] = \int_{\mathbb{R}\times\mathbb{R}} \left(Y^{\text{nest}}(x_1, u_1, u_2)\right)^2 \phi(u_1)\phi(u_2)du_1du_2
$$

$$
= \mathbb{E}_{u_1}\left[\left(\mu_M^2\left(\mu_N^1(x_1) + u_1 \times \sigma_N^1(x_1)\right)\right)^2 + \left(\sigma_M^2\left(\mu_N^1(x_1) + u_1 \times \sigma_N^1(x_1)\right)\right)^2\right]
$$

# Coupling two GPRs

- For all $x_1$, computing the two first statistical moments of $Y^{\text{nest}}(x_1, u_1, u_2)$ amounts at calculating two 1-dimensional integrals.
- Quadrature or sampling techniques can be used to compute them.

However, if

- each element of $f_2(x_2)$ can be written as a sum of functions of the form $x_2 \mapsto a_0 x_2^j \exp(a_1 x_2 + a_2 x_2)$, with $j \geq 0$, $a_1 \in \mathbb{C}$ and $a_2 \in \mathbb{R}$,
- the covariance of $\varepsilon_2$, $C_2$, is a squared exponential covariance:

$$C_2(x_2, x_2') = \sigma^2 \exp\left(-\frac{1}{2}\sum_{i=1}^{d_2}\left(\frac{x_{2,i} - x_{2,i}'}{\ell_i}\right)^2\right),$$

then **closed-form** solutions can be found for $\mathbb{E}_{u_1,u_2}\left[Y^{\text{nest}}(x_1, u_1, u_2)\right]$ and $\mathbb{E}_{u_1,u_2}\left[\left(Y^{\text{nest}}(x_1, u_1, u_2)\right)^2\right]$.

## Advantages of the method

- The conditioned GP predictors can be computed independently.
- There is no constraints on the number of available code evaluations (in particular, we are not limited by the fact that $N = M$).
- The constraints on $f_2$ and $C_2$ being not too restrictive, closed-form solutions are available for the mean and the variance of the predictor, which can be used to define iterative procedures for uncertainty quantification prospects.
- There is almost no error compensation for the calibration of $\beta_1$ and $\beta_2$.

## Coupling two GPRs

### Drawbacks of the method

- The two conditioned GP predictors being independent, the variance of the predictor can be over-estimated (in particular, the configurations $y_1^{(n)} = x_2^{(m)}$ may not be enough taken into account).
- There is almost no error compensation for the prediction of $y_2(x_1)$.
- Some evaluations of $y_2(x_2)$ can be carried out in sub-domains of $\mathbb{R}^{d_2}$ that do not belong to the image space of $y_1$.
- The closed-form expressions are valid for two nested codes only.

$\Rightarrow$ this invites us to construct a **"grouped"** predictor for $Y^{\text{nest}}$, which is based on an other linearization of the predictor.

# Outline

# Linearized approach

$$x_1 \quad \xrightarrow{f_1,\beta_1,\varepsilon_1} \quad y_1 = x_2 \quad \xrightarrow{f_2,\beta_2,\varepsilon_2} \quad y_2$$

$$y_2(x_1) = f_2(f_1(x_1)\beta_1 + \varepsilon_1(x_1))\beta_2 + \varepsilon_2(f_1(x_1)\beta_1 + \varepsilon_1(x_1)).$$

## Main objectives

- maximize the conditioning of $Y^{\text{nest}}$, to minimize its variance,
- stay in the Gaussian framework to get closed-form solutions.

## Linearization

We assume that the following linearization is valid around $\beta_1^{\text{nom}}$ and $\beta_2^{\text{nom}}$ :

$$y_2(x_1) \approx f_2(f_1(x_1)\beta_1^{\text{nom}})\beta_2 + \varepsilon_2(f_1(x_1)\beta_1^{\text{nom}}) + \frac{\partial f_2}{\partial x_2}(f_1(x_1)\beta_1^{\text{nom}})\left(f_1(x_1)(\beta_1 - \beta_1^{\text{nom}}) + \varepsilon_1(x_1)\right)\beta_2^{\text{nom}}.$$

## Vectorial representation

Hence we get:

$$Z(x_1,x_2) := \begin{pmatrix} y_1(x_1) \\ y_2(x_2) \\ y_2(x_1) \end{pmatrix} \approx \begin{pmatrix} f_1(x_1)\beta_1 + \varepsilon_1(x_1) \\ f_2(x_2)\beta_2 + \varepsilon_2(x_2) \\ \begin{pmatrix} y_2^0 + f_{31}(x_1)\beta_1 + f_{32}(x_1)\beta_2 \\ +f_{33}(x_1)\varepsilon_1(x_1) + \varepsilon_2(f_1(x_1)\beta_1^{\text{nom}}) \end{pmatrix} \end{pmatrix}$$

If $\varepsilon_1$ and $\varepsilon_2$ are two independent GP, then it comes:

$$Z(x_1,x_2) \mid \beta_1,\beta_2 \sim \mathcal{N}\left(\mu(x_1,x_2,\beta_1,\beta_2), \left[C(x_1,x_1',x_2,x_2')\right]\right).$$

By construction, $Y^{\text{nest}}$ corresponds to $Z_3$ given the available evaluations of the two codes.

$$Z(x_1, x_2) \mid \beta_1, \beta_2 \sim \mathcal{N}\left(\boldsymbol{\mu}(x_1, x_2, \beta_1, \beta_2), \left[C(x_1, x_1', x_2, x_2')\right]\right).$$

- Vector-valued GP $\boldsymbol{Z}$ can then be conditioned by all the available information gathered in $(x_1^{(n)}, y_1^{(n)})_{1 \leq n \leq N}$ and $(x_2^{(m)}, y_2^{(m)})_{1 \leq m \leq M}$.
- In particular, the fact that there exists values of $n$ and $m$ such that $y_1^{(n)} = x_2^{(m)}$ is taken into account.

$\Rightarrow$ if $(\beta_1, \beta_2)$ is a priori uniformly distributed on $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ (or Gaussian), then the posterior distributions of $(\beta_1, \beta_2)$ and $Y^{\text{nest}}$ can be analytically deduced.

## Advantages of the method

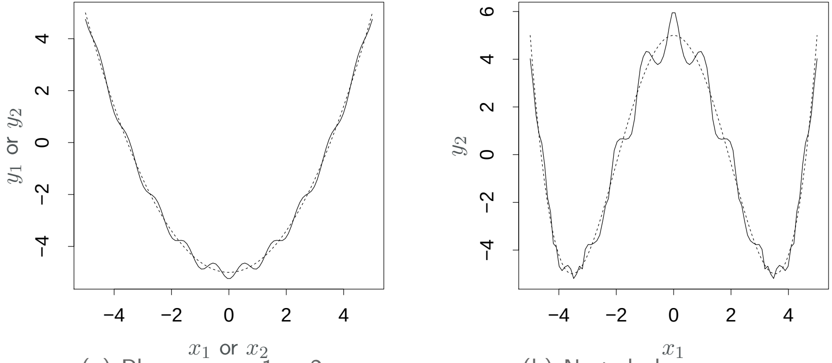
- There is no constraints on the number of available code evaluations.
- All the available information is used to constrain the predictions such that the variance of the predictor of the nested phenomenon is often reduced.
- The predictor interpolates the code evaluations when $y_1^{(n)} = x_2^{(m)}$.
- Closed-form solutions are available.
- There is *a priori* no limitation on the covariance of $\varepsilon_2$.
- There can be error compensation for the prediction of $y_2(x_1)$.
- The linearized approach can be generalized to the calibration of more than two chained codes.

## Drawbacks of the method

- There can be error compensation for the calibration of $\beta_1$ and $\beta_2$.
- The relevance of the linearization has to be controlled.

# Outline

# Numerical application




(a) Phenomenon 1 or 2

(b) Nested phenomenon

Figure: Continuous line $\leftrightarrow$ reality i.e. $y_1(x_1)$, $y_2(x_2)$ and $y_2(x_1)$, dashed line $\leftrightarrow$ computer codes $f_1(x_1)\beta_1$, $f_2(x_2)\beta_2$ and $f_2(f_1(x_1)\beta_1)\beta_2$.

$f_1(x_1) = (1, x_1^2)$, $f_2(x_2) = (1, x_2^2)$, such that the dimension of $(\beta_1, \beta_2)$ is 4.

# Numerical application

Two quantities of interest will be computed to control the relevance of the estimations :

- for the calibration :

$$\epsilon_\beta^2 = \mathbb{E}\left[\left\|\beta_c - \widehat{\beta}_c\right\|^2\right],$$

  - where $\widehat{\beta}_c$ is the estimator of the parameters given the observations,

- for the prediction capability :

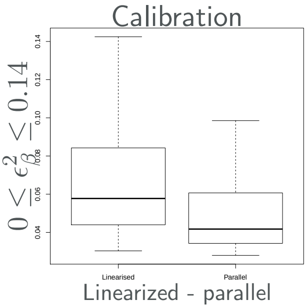
$$\epsilon_y^2 = \int_{x_1} \mathbb{E}\left[\left(y_2\left(x_1\right) - Y^{\mathsf{nest}}(x_1)\right)^2\right] dx_1,$$

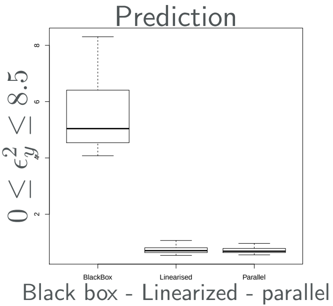
  - where $Y^{\mathsf{nest}}(x_1)$ is the predictor of $y_2(y_1(x_1))$ given the observations.

As a reference, we denoted by "black-box" the approach consisting in calibrating $y_2(y_1(x_1))$ from the only available information gathered in $(x_1^{(k)}, y_2^{(k)})_{1\leq k\leq K}$.

# Numerical application

The box-plots are computed from 50 repetitions of a calibration of $y_2(y_1(x_1))$ based on 8 observations of each phenomenon.



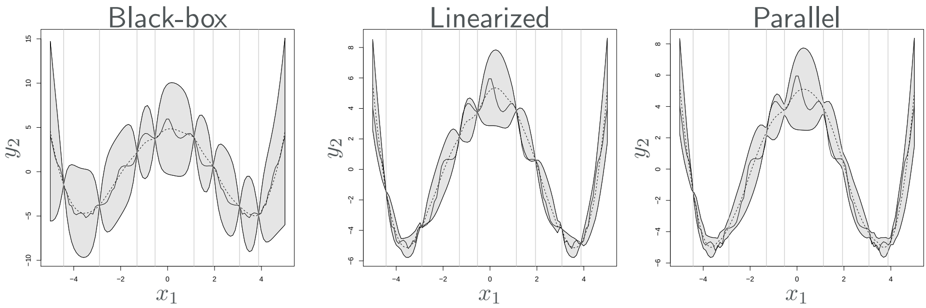




$$\epsilon_\beta^2 = \mathbb{E}\left[\left\|\beta_c - \widehat{\beta}_c\right\|^2\right] \qquad \epsilon_y^2 = \int_{x_1} \mathbb{E}\left[\left(y_2\left(x_1\right) - Y^{\text{nest}}(x_1)\right)^2\right] dx_1$$

## Calibration of the parameters

Black-box $\ll$ Linearized $\leq$ Parallel

# Numerical application




**Estimation of the prediction uncertainty**

Black-box $\ll$ Parallel $\leq$ Linearized

# Outline

# Conclusion

Two approaches have been proposed to calibrate two nested codes:

- the first one based on the coupling of two predictors,
- the second one based on a linearized approach.

Both techniques have their own advantages and drawbacks in terms of calibration and prediction.

**To go further...**

- taking into account the experimental errors,
- case when $y_1 = \{y_1(t),\ 0 \leq t \leq T\}$ is a functional output,
- sequential improvement of the predictors.

**Thank you for your attention.**