ECCOMAS Congress 2016
VII European Congress on Computational Methods in Applied Sciences and Engineering
M. Papadrakakis, V. Papadopoulos, G. Stefanou, V. Plevris (eds.)
Crete Island, Greece, 5–10 June 2016

CALIBRATION OF NESTED COMPUTER MODELS

Sophie Marque-Pucheu[1], **Guillaume Perrin**[1], **and Josselin Garnier**[2]

[1]CEA DAM DIF
F-91297, Arpajon, France
e-mail: sophie.marque-pucheu@cea.fr, guillaume.PERRIN2@cea.fr

[2] Laboratoire de Probabilités et Modèles Aléatoires, Laboratoire Jacques-Louis Lions
Université Paris Diderot, 75205 Paris Cedex 13, France
e-mail: garnier@math.univ-paris-diderot.fr

Keywords: Nested computer models, metamodeling, Bayesian calibration, Gaussian process, uncertainity quantification.

Abstract. *Thanks to computing power increase, risk quantification relies more and more on computer modelling. Methods of risk quantification based on a fixed computational budget exist, but computer models are almost always considered as a single black box.*

In this paper, we are interested in analyzing the behavior of a complex phenomenon, whose evolution can be modelled via two nested parametrized computer models. By two nested computer models, we mean that some inputs of the second model are outputs of the first model. Based on series of observations of the considered phenomenon, the idea is, first, to calibrate the parameters of each model, and then, to construct a predictor for the output of the second model, which takes into account the fact that, on the first hand, the two models are not perfect, and on the other hand, there exist uncertainties in the parameters' calibration.

Concerning the calibration of the models' parameters we distinguish between the classic "black-box" method, parallel method and grouped method. In the case of a parallel calibration, each model is calibrated separately whereas in the case of a grouped calibration, the parameters of the two models are calibrated all together. In both cases, the extent to which the proposed predictor integrates the uncertainties of the nested system is demonstrated. The proposed methods are then applied to an example.

1 INTRODUCTION

In this paper, we are interested in analyzing the behavior of a complex phenomenon, whose evolution can be modeled by two nested parametrized computer models. By two nested computer models, we mean that some inputs of the second model are outputs of the first model. The computer models are assumed linear with respect to the models' parameters $\beta_{\mathbf{1}}$ and $\beta_{\mathbf{2}}$. Both models are supposed to be affected by model errors ϵ_1 and ϵ_2. These errors are assumed to be independent. So we have :

$$\begin{aligned} y_1(x_1) &= \mathbf{h_1}(x_1)^t \beta_{\mathbf{1}} + \epsilon_1(x_1) \\ y_2(x_2) &= \mathbf{h_2}(x_2)^t \beta_{\mathbf{2}} + \epsilon_2(x_2) \\ x_2 &= y_1 \end{aligned} \tag{1}$$

where :

- $x_1 \mapsto y_1$ and $x_2 \mapsto y_2$ are the considered phenomena, $x_1 \in \mathbb{R}$, $x_2 \in \mathbb{R}$, $y_1 \in \mathbb{R}$ and $y_2 \in \mathbb{R}$
- $\mathbf{h_1}(x_1)^t \beta_{\mathbf{1}}$ and $\mathbf{h_2}(x_2)^t \beta_{\mathbf{2}}$ the deterministic computer codes, $\mathbf{h_1}(x_1) \in \mathbb{R}^{p_1} \times 1$, $\mathbf{h_2}(x_2) \in \mathbb{R}^{p_2} \times 1$, $\beta_{\mathbf{1}} \in \mathbb{R}^{p_1} \times 1$, $\beta_{\mathbf{2}} \in \mathbb{R}^{p_2} \times 1$,
- ϵ_1 and ϵ_2 the real-valued models' errors. They are modelled by centered Gaussian processes.

The outline of this paper is the following. In section 2 is presented the general method to calibrate these two nested codes. We suppose we have access to observations of the phenomena $x_1 \mapsto y_1$, $x_2 \mapsto y_2$ and $x_1 \mapsto y_2$. In sections 3 to 5 three special cases of the general method are presented. These cases, depending on the available information, are the following :

- the parallel method. In this case, there are observations only for the phenomena $x_1 \mapsto y_1$ and $x_2 \mapsto y_2$ which are denoted by $\left(\mathbf{x_{1,obs}}^{(1)}, \mathbf{y_{1,obs}}^{(1)}\right)$ and $\left(\mathbf{x_{2,obs}}^{(2)}, \mathbf{y_{2,obs}}^{(2)}\right)$,
- the "black-box" method. In this case, there are observations only for the nested phenomenon ($x_1 \mapsto y_2$), which are denoted by $\left(\mathbf{x_{1,obs}}^{(3)}, \mathbf{y_{2,obs}}^{(3)}\right)$,
- the "grouped" method. In this case, there are observations for the phenomenon $x_1 \mapsto y_1 \mapsto y_2$, which are denoted by $\left(\mathbf{x_{1,obs}}^{(1)}, \mathbf{y_{1,obs}}^{(1)}\right)$, $\left(\mathbf{x_{2,obs}}^{(2)}, \mathbf{y_{2,obs}}^{(2)}\right)$ and $\left(\mathbf{x_{1,obs}}^{(3)}, \mathbf{y_{2,obs}}^{(3)}\right)$, where $\mathbf{y_{1,obs}}^{(1)} = \mathbf{x_{2,obs}}^{(2)}$, $\mathbf{x_{1,obs}}^{(3)} = \mathbf{x_{1,obs}}^{(1)}$ and $\mathbf{y_{2,obs}}^{(2)} = \mathbf{y_{2,obs}}^{(3)}$.

Note : notations $^{(1)}$, $^{(2)}$ and $^{(3)}$ correspond to first phenomenon, second phenomenon and nested phenomenon.
Finally in section 6 the three specific cases of the general method are applied on a numerical example.

2 THE GENERAL METHOD

In this section we propose a general framework to calibrate nested codes and build posterior predictors of these codes by taking into account observations of the three phenomena ($x_1 \mapsto y_1$, $x_2 \mapsto y_2$ and $x_1 \mapsto y_2$).

2.1 Nested model linearization

According to equation (1) the relation between y_2 and x_1 is given by :

$$y_2(x_1) = \mathbf{h_2}\left(\mathbf{h_1}(x_1)^t \beta_\mathbf{1} + \epsilon_1(x_1)\right)^t \beta_\mathbf{2} + \epsilon_2\left(\mathbf{h_1}(x_1)^t \beta_\mathbf{1} + \epsilon_1(x_1)\right), \tag{2}$$

Assuming that nominal values $\bar{\beta}_\mathbf{1}$ and $\bar{\beta}_\mathbf{2}$ of $\beta_\mathbf{1}$ and $\beta_\mathbf{2}$ are available such that in the vicinity of $\bar{\beta}_\mathbf{1}$, $\mathbf{h_1}(x_1)^t\left(\beta_\mathbf{1} - \bar{\beta}_\mathbf{1}\right) + \epsilon_1(x_1)$ is small, we get a linear approximation of the nested code and its error :

$$y_3(x_1) = \mathbf{h_3}(x_1)^t \beta_\mathbf{c} + \epsilon_3(x_1), \tag{3}$$

where $\epsilon_3(x_1)$ is the proposed model error :

$$\epsilon_3(x_1) = \Delta F_{2c}\left(x_1, \bar{\beta}_\mathbf{c}\right)\epsilon_1(x_1) + \epsilon_2\left(\mathbf{h_1}(x_1)^t \bar{\beta}_\mathbf{1}\right), \tag{4}$$

and :

$$\begin{aligned}
\mathbf{h_3}(x_1) &= \begin{bmatrix} \mathbf{h_{31}}(x_1) \\ \mathbf{h_{32}}(x_1) \end{bmatrix} \\
y_3(x_1) &= y_2(x_1) + \mathbf{h_{31}}(x_1)^t \bar{\beta}_\mathbf{1} \\
\mathbf{h_{31}}(x_1) &= \Delta F_{2c}\left(x_1, \bar{\beta}_\mathbf{c}\right)\mathbf{h_1}(x_1) \\
\mathbf{h_{32}}(x_1) &= \mathbf{h_2}\left(\mathbf{h_1}(x_1)^t \bar{\beta}_\mathbf{1}\right) \\
\Delta F_{2c}\left(x_1, \bar{\beta}_\mathbf{c}\right) &= \tfrac{\partial \mathbf{h_2}}{\partial x_2}\left(\mathbf{h_1}(x_1)^t \bar{\beta}_\mathbf{1}\right)^t \bar{\beta}_\mathbf{2} \\
\bar{\beta}_\mathbf{c} &= \begin{bmatrix} \bar{\beta}_\mathbf{1} \\ \bar{\beta}_\mathbf{2} \end{bmatrix}.
\end{aligned} \tag{5}$$

Given equations (1) and (3) a joint model for the three phenomena ($x_1 \mapsto y_1$, $x_2 \mapsto y_2$ and $x_1 \mapsto y_2$) is proposed :

$$\mathbf{y_{gp}}(\mathbf{x_{gp}})\,|\beta_\mathbf{c} \sim \mathcal{N}\left(\mathbf{h_{gp}}(\mathbf{x_{gp}})^t \beta_\mathbf{c},\ \mathbf{C_{gp}}\left(\mathbf{x_{gp}}; \mathbf{x'_{gp}}\right)\right) \tag{6}$$

where :

$$\beta_\mathbf{c} = \begin{bmatrix} \beta_\mathbf{1} \\ \beta_\mathbf{2} \end{bmatrix}, \tag{7}$$

$$\mathbf{x_{gp}} = \left(x_1^{(1)}, x_2^{(2)}, x_1^{(3)}\right), \tag{8}$$

$$\mathbf{y_{gp}}\left(\mathbf{x_{gp}}\right) = \begin{bmatrix} y_1\left(x_1^{(1)}\right) \\ y_2\left(x_2^{(2)}\right) \\ y_3\left(x_1^{(3)}\right) \end{bmatrix}, \tag{9}$$

$$\mathbf{h_{gp}}\left(\mathbf{x_{gp}}\right)^t = \begin{bmatrix} \mathbf{h_1}\left(x_1^{(1)}\right)^t & 0 \\ 0 & \mathbf{h_2}\left(x_2^{(2)}\right)^t \\ \mathbf{h_{31}}\left(x_1^{(3)}\right)^t & \mathbf{h_{32}}\left(x_1^{(3)}\right)^t \end{bmatrix}, \tag{10}$$

$$\mathbf{C_{gp}}\left(x_1^{(1)}, x_2^{(2)}, x_1^{(3)}; x_1^{'(1)}, x_2^{'(2)}, x_1^{'(3)}\right) = \tag{11}$$

$$\begin{bmatrix} C_1\left(x_1^{(1)}, x_1^{'(1)}\right) & 0 & C_1\left(x_1^{(1)}, x_1^{'(3)}\right)\Delta F_{2c}\left(x_1^{'(3)}, \bar{\beta}_{\mathbf{c}}\right) \\ 0 & C_2\left(x_2^{(2)}, x_2^{'(2)}\right) & C_2\left(x_2^{(2)}, \mathbf{h_1}\left(x_1^{'(3)}\right)^t \bar{\beta}_{\mathbf{1}}\right) \\ \Delta F_{2c}\left(x_1^{(3)}, \bar{\beta}_{\mathbf{c}}\right) C_1\left(x_1^{(3)}, x_1^{'(1)}\right) & C_2\left(\mathbf{h_1}\left(x_1^{(3)}\right)^t \bar{\beta}_{\mathbf{1}}, x_2^{'(2)}\right) & C_3\left(x_1^{(3)}, x_1^{'(3)}\right) \end{bmatrix},$$

and C_1, C_2 and C_3 are the covariance functions of ϵ_1, ϵ_2 and ϵ_3 errors.
Given equation (6), in the case of uninformative prior or Gaussian prior for $\beta_{\mathbf{c}}$, the parameters' distribution and the posterior predictive distribution given the observations, are Gaussian (see [1], [6], [7], [11] [12]).
The following sections present the parameters' calibration and the prediction given series of observations according to the model presented above.

2.2 Calibration of the codes' parameters and posterior predictive distributions

Given equation (6), we have :

$$\mathbf{y_{gp,obs}}|\beta_{\mathbf{c}} \sim \mathcal{N}\left(\mathbf{H_{gp}}\beta_{\mathbf{c}}\,,\,\mathbf{R_{gp}}\right), \tag{12}$$

where :

$$\mathbf{y_{gp,obs}} = \begin{bmatrix} \mathbf{y_{1,obs}}^{(1)} \\ \mathbf{y_{2,obs}}^{(2)} \\ \mathbf{y_{3,obs}}^{(3)} \end{bmatrix}, \tag{13}$$

$$\mathbf{y_{3,obs}}^{(3)} = \mathbf{y_{2,obs}}^{(3)} + \mathbf{h_{31}}\left(\mathbf{x_{1,obs}}^{(3)}\right)^t \bar{\beta}_{\mathbf{1}}, \tag{14}$$

$$\mathbf{x_{gp,obs}} = \begin{bmatrix} \mathbf{x_{1,obs}}^{(1)} \\ \mathbf{x_{2,obs}}^{(2)} \\ \mathbf{x_{1,obs}}^{(3)} \end{bmatrix}, \tag{15}$$

$$\mathbf{H_{gp}} = \mathbf{h_{gp}}\left(\mathbf{x_{gp,obs}}\right)^t, \tag{16}$$

$$\mathbf{R_{gp}} = \mathbf{C_{gp}}\left(\mathbf{x_{gp,obs}}; \mathbf{x_{gp,obs}}\right). \tag{17}$$

Given equation (12), the posterior distribution of $\beta_{\mathbf{c}}$ given the observations $\mathbf{y_{gp,obs}}$ can be modeled in a Bayesian framework. We distinguish two cases :

- No prior information case :
$$\pi^{prior}\left[\beta_{\mathbf{c}}\right] \sim 1 \tag{18}$$

- Gaussian prior case :
$$\pi^{prior}\left[\beta_{\mathbf{c}}\right] \sim \mathcal{N}\left(\widehat{\beta}_{\mathbf{c}}^{prior}\ ,\ {\mathbf{R}_{\beta_{\mathbf{c}}}}^{prior}\right) \tag{19}$$

In a Bayesian framework, in these two cases, the parameters' posterior distribution is Gaussian :

$$\beta_{\mathbf{c}} \mid \mathbf{y_{gp,obs}} \sim \mathcal{N}\left(\ \mathbb{E}\left(\beta_{\mathbf{c}} \mid \mathbf{y_{gp,obs}}\right)\ ,\ cov\left(\beta_{\mathbf{c}} \mid \mathbf{y_{gp,obs}}\right)\ \right) \tag{20}$$

where :

- in the no prior information case :
$$\begin{aligned} cov\left(\beta_{\mathbf{c}} \mid \mathbf{y_{gp,obs}}\right) &= \mathbf{R}_{\beta_{\mathbf{c}}} = \left(\mathbf{H_{gp}}^t {\mathbf{R_{gp}}}^{-1} \mathbf{H_{gp}}\right)^{-1} \\ \mathbb{E}\left(\beta_{\mathbf{c}} \mid \mathbf{y_{gp,obs}}\right) &= \widehat{\beta}_{\mathbf{c}} = \mathbf{R}_{\beta_{\mathbf{c}}} \mathbf{H_{gp}}^t {\mathbf{R_{gp}}}^{-1} \mathbf{y_{gp,obs}} \end{aligned} \tag{21}$$

- in the Gaussian prior case :
$$\begin{aligned} \left(\mathrm{cov}\left(\beta_{\mathbf{c}} \mid \mathbf{y_{gp,obs}}\right)\right)^{-1} &= \widetilde{\mathbf{R}}_{\beta_{\mathbf{c}}}^{-1} = {\mathbf{R}_{\beta_{\mathbf{c}}}}^{-1} + \left({\mathbf{R}_{\beta_{\mathbf{c}}}}^{prior}\right)^{-1} \\ \mathbb{E}\left(\beta_{\mathbf{c}} \mid \mathbf{y_{gp,obs}}\right) &= \widetilde{\beta}_{\mathbf{c}} = \widetilde{\mathbf{R}}_{\beta_{\mathbf{c}}} \left[{\mathbf{R}_{\beta_{\mathbf{c}}}}^{-1} \widehat{\beta}_{\mathbf{c}} + \left({\mathbf{R}_{\beta_{\mathbf{c}}}}^{prior}\right)^{-1} \widehat{\beta}_{\mathbf{c}}^{prior}\right] \end{aligned} \tag{22}$$

Given equations (6), (12) and (20), the posterior predictive distribution $\mathbf{y_{gp}}\left(\mathbf{x_{gp}}\right) | \mathbf{y_{gp,obs}}$ is Gaussian. Its properties are :

$$\mathbf{y_{gp}}\left(\mathbf{x_{gp}}\right)\ |\ \mathbf{y_{gp,obs}} \sim \mathcal{N}\left(\ \widehat{y}_{gp}\left(\mathbf{x_{gp}}, \mathbb{E}\left(\beta_{\mathbf{c}} \mid \mathbf{y_{gp,obs}}\right)\right)\ ,\ \widehat{\mathbf{C}}_{\mathbf{gp}}\left(\mathbf{x_{gp}}; \mathbf{x_{gp}}\right)\ \right) \tag{23}$$

where for all $\beta_{\mathbf{c}}$:

$$\widehat{y}_{gp}\left(\mathbf{x_{gp}}, \beta_{\mathbf{c}}\right) = \mathbf{h_{gp}}\left(\mathbf{x_{gp}}\right)^t \beta_{\mathbf{c}} + \mathbf{r_{gp}}\left(\mathbf{x_{gp}}\right)^t {\mathbf{R_{gp}}}^{-1} \left(\mathbf{y_{gp,obs}} - \mathbf{H_{gp}} \beta_{\mathbf{c}}\right) \tag{24}$$

$$\mathbf{r_{gp}}\left(\mathbf{x_{gp}}\right) = \mathbf{C_{gp}}\left(\mathbf{x_{gp,obs}}; \mathbf{x_{gp}}\right) \tag{25}$$

$$\begin{aligned} \widehat{\mathbf{C}}_{\mathbf{gp}}\left(\mathbf{x_{gp}}; \mathbf{x'_{gp}}\right) = \ & \mathbf{C_{gp}}\left(\mathbf{x_{gp}}; \mathbf{x'_{gp}}\right) - \mathbf{r_{gp}}\left(\mathbf{x_{gp}}\right)^t {\mathbf{R_{gp}}}^{-1} \mathbf{r_{gp}}\left(\mathbf{x'_{gp}}\right) \\ & + \mathbf{u_{gp}}\left(\mathbf{x_{gp}}\right)^t \mathrm{cov}\left(\beta_{\mathbf{c}} \mid \mathbf{y_{gp,obs}}\right) \mathbf{u_{gp}}\left(\mathbf{x'_{gp}}\right) \end{aligned} \tag{26}$$

$$\mathbf{u_{gp}}\left(\mathbf{x_{gp}}\right) = \mathbf{h_{gp}}\left(\mathbf{x_{gp}}\right) - \mathbf{H_{gp}}{\mathbf{R_{gp}}}^{-1}\mathbf{r_{gp}}\left(\mathbf{x_{gp}}\right) \tag{27}$$

In the following sections are presented three specific cases of the general method previously introduced. The cases are distinguished according to the type of available observations.

3 THE PARALLEL APPROACH

In this section we present a particular case of the previously presented general framework. The parallel method corresponds to the case where only observations of the phenomena $x_1 \mapsto y_1$ and $x_2 \mapsto y_2$ are available.

3.1 Key features of the approach

In the parallel approach, there is no observation for the nested phenomenon, that is : $\left({\mathbf{x_{1,obs}}}^{(3)}, {\mathbf{y_{2,obs}}}^{(3)}\right) = \emptyset$.
This method includes the following steps :

1. the parameters' calibration given the observations $\left({\mathbf{x_{1,obs}}}^{(1)}, {\mathbf{y_{1,obs}}}^{(1)}\right)$ and $\left({\mathbf{x_{2,obs}}}^{(2)}, {\mathbf{y_{2,obs}}}^{(2)}\right)$. Considering that the covariance matrix given by equation (11) is a block diagonal matrix, it is like $\beta_{\mathbf{1}}$ and $\beta_{\mathbf{2}}$ were calibrated separately given the observations $\left({\mathbf{x_{1,obs}}}^{(1)}, {\mathbf{y_{1,obs}}}^{(1)}\right)$ and $\left({\mathbf{x_{2,obs}}}^{(2)}, {\mathbf{y_{2,obs}}}^{(2)}\right)$.

2. the construction of the posterior predictors given the observations :

   (a) for the phenomenon 1 : the proposed predictor is an interpolating Gaussian predictor given the observations $\left({\mathbf{x_{1,obs}}}^{(1)}, {\mathbf{y_{1,obs}}}^{(1)}\right)$

   (b) for the phenomenon 2 : the proposed predictor is an interpolating Gaussian predictor given the observations $\left({\mathbf{x_{2,obs}}}^{(2)}, {\mathbf{y_{2,obs}}}^{(2)}\right)$

   (c) for the nested phenomenon : the proposed predictor is Gaussian thanks to a linearization of the coupling of the Gaussian predictors of the phenomena 1 and 2 (see 3.2).

The following section presents how the predictor of the nested code is built in the parallel approach.

3.2 Parallel predictor for the nested code

As shown above, we have two posterior Gaussian predictors given the observations :

$$y_{pred,i}\left(x_i\right) = y_i\left(x_i\right) \mid \left({\mathbf{x_{i,obs}}}^{(i)}, {\mathbf{y_{i,obs}}}^{(i)}\right) \sim \mathcal{N}\left(\, \widehat{y}_i\left(x_i, \beta_{\mathbf{i}}\right) \,,\, \widehat{C}_i\left(x_i, x_i\right) \right) \quad i = \{1, 2\} \tag{28}$$

where :

$$\widehat{y}_i\left(x_i, \beta_{\mathbf{i}}\right) = \mathbf{h_i}\left(x_i\right)^t \beta_{\mathbf{i}} + \mathbf{r_i}\left(x_i\right)^t {\mathbf{R_i}}^{-1}\left({\mathbf{y_{i,obs}}}^{(i)} - \mathbf{h_i}\left({\mathbf{x_{i,obs}}}^{(i)}\right)^t \beta_{\mathbf{i}}\right) \tag{29}$$

$$\mathbf{r_i}\left(x_i\right) = C_i\left(\mathbf{x_{i,obs}}^{(i)}; x_i\right) \tag{30}$$

$$\mathbf{R_i} = C_i\left(\mathbf{x_{i,obs}}^{(i)}; \mathbf{x_{i,obs}}^{(i)}\right) \tag{31}$$

$$\widehat{C}_i\left(x_i; x_i'\right) = C_i\left(x_i; x_i'\right) - \mathbf{r_i}\left(x_i\right)^t \mathbf{R_i}^{-1}\mathbf{r_i}\left(x_i'\right) - \mathbf{u_i}\left(x_i\right)^t \mathbf{R_{\beta_i}}^{-1}\mathbf{u_i}\left(x_i'\right) \tag{32}$$

$$\mathbf{u_i}\left(x_i\right) = \mathbf{h_i}\left(x_i\right) - \mathbf{h_i}\left(\mathbf{x_{i,obs}}^{(i)}\right)\mathbf{R_i}^{-1}\mathbf{r_i}\left(x_i\right) \tag{33}$$

Given equation (28) the posterior predictor for the nested phenomenon can be written :

$$y_{pred,3}\left(x_1, \beta_{\mathbf{c}}\right) = \widehat{y}_2\left(\widehat{y}_1\left(x_1, \beta_{\mathbf{1}}\right) + \delta\widehat{y}_1\left(x_1\right), \beta_{\mathbf{2}}\right) + \delta\widehat{y}_2\left(\widehat{y}_1\left(x_1, \beta_{\mathbf{1}}\right) + \delta\widehat{y}_1\left(x_1\right)\right) \tag{34}$$

where :

$$\delta\widehat{y}_i\left(x_i\right) \sim \mathscr{N}\left(0\ ,\ \widehat{C}_i\left(x_i; x_i\right)\right) \quad i = \{1, 2\} \tag{35}$$

Thanks to a first order Taylor series expansion and making the following approximation according to equation (29)

$$\frac{\partial\widehat{y}_2}{\partial x_2}\left(x, \beta\right) = \left[\frac{\partial\mathbf{h_2}}{\partial x_2}\left(x\right)\right]^t \beta \tag{36}$$

the posterior predictor for the nested code is :

$$y_{pred,3}\left(x_1, \widehat{\beta}_{\mathbf{c}}\right) \simeq \widehat{y}_2\left(\widehat{y}_1\left(x_1, \widehat{\beta}_{\mathbf{1}}\right), \widehat{\beta}_{\mathbf{2}}\right) + \left[\frac{\partial\mathbf{h_2}}{\partial x_2}\left(\widehat{y}_1\left(x_1, \widehat{\beta}_{\mathbf{1}}\right)\right)\right]^t \widehat{\beta}_{\mathbf{2}}\delta\widehat{y}_1\left(x_1\right) + \delta\widehat{y}_2\left(\widehat{y}_1\left(x_1, \widehat{\beta}_{\mathbf{1}}\right)\right) \tag{37}$$

It is a linear combination of the independent Gaussian processes $\delta\widehat{y}_1$ and $\delta\widehat{y}_2$, thus it is a Gaussian predictor.
Table 1 summarizes the 'parallel' method's characteristics.

Table 2 presents the MSE of the calibrated codes and predictors for the 'parallel' method.

4 THE 'BLACK-BOX' APPROACH

In this section we present a particular case of the previously presented general framework (see 2). The 'black-box' method corresponds to the case where only observations of the phenomenon $x_1 \mapsto y_2$ are available, that is $\left(\mathbf{x_{1,obs}}^{(1)}, \mathbf{y_{1,obs}}^{(1)}\right) = \emptyset$ and $\left(\mathbf{x_{2,obs}}^{(2)}, \mathbf{y_{2,obs}}^{(2)}\right) = \emptyset$.
About the nested model's error, the following cases are distinguished :

- ϵ_1 and ϵ_2, are known, so the error is given by equation (4),
- ϵ_1 and ϵ_2 are unknown, so a specific model error for the nested code is proposed (it is assumed to be a zero-mean stationary Gaussian process).

Nested model error	The nested model error is not used because there is no observation for the nested phenomenon.
Calibration of parameters $\beta_{\mathbf{c}} = (\beta_{\mathbf{1}}, \beta_{\mathbf{2}})$	• $\beta_{\mathbf{1}}$ is calibrated given the observations for $x_1 \mapsto y_1$, • $\beta_{\mathbf{2}}$ is calibrated given the observations for $x_2 \mapsto y_2$.
Posterior predictor for phenomenon $x_1 \mapsto y_1$	Gaussian predictor given the observations of the phenomenon $x_1 \mapsto y_1$
Posterior predictor for phenomenon $x_2 \mapsto y_2$	Gaussian predictor given the observations of the phenomenon $x_2 \mapsto y_2$
Posterior predictor for phenomenon $x_1 \mapsto y_2$	Gaussian predictor thanks to a linearization of the coupling of the Gaussian predictors of the phenomena $x_1 \mapsto y_1$ and $x_2 \mapsto y_2$.

Table 1: This table summarizes the 'parallel' method's characteristics

Phenomenon	Empirical normalized integrated squared error (Empirical MSE) of the calibrated codes	Empirical normalized integrated squared error (Empirical MSE) of the posterior predictors
$x_1 \mapsto y_1$	$\frac{\sum_{x_1} \left(y_1(x_1) - \mathbf{h_1}(x_1)^t \widehat{\beta}_{\mathbf{1}} \right)^2}{\sum_{x_1} (y_1(x_1))^2}$	$\frac{1}{\sum_{x_i} (y_i(x_i))^2} \sum_{x_i} (y_i(x_i) - \mathbf{h_i}(x_i)^t \widehat{\beta}_{\mathbf{i}} - \mathbf{r_i}(x_i)^t \mathbf{R_i}^{-1} \left(\mathbf{y_{i,obs}} - \mathbf{h_i}(\mathbf{x_{i,obs}})^t \widehat{\beta}_{\mathbf{i}} \right))^2$ $i = \{1, 2\}$
$x_2 \mapsto y_2$	$\frac{\sum_{x_2} \left(y_2(x_2) - \mathbf{h_2}(x_2)^t \widehat{\beta}_{\mathbf{2}} \right)^2}{\sum_{x_2} (y_2(x_2))^2}$	
$x_1 \mapsto y_2$	$\frac{1}{\sum_{x_1} (y_2(x_1))^2} \sum_{x_1} (y_2(x_1) - \mathbf{h_2} \left(\mathbf{h_1}(x_1)^t \widehat{\beta}_{\mathbf{1}} \right)^t \widehat{\beta}_{\mathbf{2}})^2$	$\frac{\sum_{x_1} \left(y_2(x_1) - \widehat{y}_2 \left(\widehat{y}_1 \left(x_1, \widehat{\beta}_{\mathbf{1}} \right), \widehat{\beta}_{\mathbf{2}} \right) \right)^2}{\sum_{x_1} (y_2(x_1))^2}$

Table 2: This table presents the normalized empirical integrated mean squared error of the calibrated codes and predictors for the 'parallel' method

Nested model error	Two cases are distinguished : • ϵ_1 and ϵ_2, are known, so the error is given by equation (4), • ϵ_1 and ϵ_2 are unknown, so a specific model error for the nested code is proposed (it is assumed to be a zero-mean stationary Gaussian process).
Calibration of parameters $\beta_{\mathbf{c}} = (\beta_{\mathbf{1}}, \beta_{\mathbf{2}})$	$\beta_{\mathbf{c}}$ is calibrated given the observations for $x_1 \mapsto y_2$
Posterior predictor for phenomenon $x_1 \mapsto y_1$	First code using the posterior distribution of $\beta_{\mathbf{1}}$
Posterior predictor for phenomenon $x_2 \mapsto y_2$	Second code using the posterior distribution of $\beta_{\mathbf{2}}$
Posterior predictor for phenomenon $x_1 \mapsto y_2$	Gaussian predictor given the observations of the phenomenon $x_1 \mapsto y_2$

Table 3: This table summarizes the 'black-box' method's characteristics

The method includes the following steps :

1. the parameters' calibration given the observations $\left(\mathbf{x_{1,obs}}^{(3)}, \mathbf{y_{2,obs}}^{(3)}\right)$,
2. the construction of the posterior predictors given the observations :
   (a) for the phenomenon 1 : there is no observation $\left(\mathbf{x_{1,obs}}^{(1)}, \mathbf{y_{1,obs}}^{(1)}\right)$, so the proposed predictor is the calibrated code. Thanks to the code's linearity and the parameters' Gaussian distribution the calibrated code's distribution is Gaussian with :
      - mean : $\mathbf{h_1}(x_1)^t \, \mathbb{E}\left(\beta_{\mathbf{1}} \mid \left(\mathbf{x_{1,obs}}^{(3)}, \mathbf{y_{2,obs}}^{(3)}\right)\right)$
      - and variance : $\mathbf{h_1}(x_1)^t \, \mathrm{cov}\left(\beta_{\mathbf{1}} \mid \left(\mathbf{x_{1,obs}}^{(3)}, \mathbf{y_{2,obs}}^{(3)}\right)\right) \mathbf{h_1}(x_1)$

   (b) for the phenomenon 2 : there is no observation $\left(\mathbf{x_{2,obs}}^{(2)}, \mathbf{y_{2,obs}}^{(2)}\right)$, so the proposed predictor is the calibrated code. Thanks to the code's linearity and the parameters' Gaussian distribution the calibrated code's distribution is Gaussian with :
      - mean : $\mathbf{h_2}(x_2)^t \, \mathbb{E}\left(\beta_{\mathbf{2}} \mid \mathbf{y_{2,obs}}^{(2)}\left(\mathbf{x_{1,obs}}^{(3)}\right)\right)$
      - and variance : $\mathbf{h_2}(x_2)^t \, \mathrm{cov}\left(\beta_{\mathbf{2}} \mid \left(\mathbf{x_{1,obs}}^{(3)}, \mathbf{y_{2,obs}}^{(3)}\right)\right) \mathbf{h_2}(x_2)$

   (c) for the nested phenomenon : the proposed predictor is an interpolating Gaussian predictor given the observations $\left(\mathbf{x_{1,obs}}^{(3)}, \mathbf{y_{2,obs}}^{(3)}\right)$

Table 3 summarizes the 'black-box' method's characteristics.
Table 4 presents the normalized empirical integrated mean squared error of the calibrated codes and predictors of this method.

Phenomenon	Empirical normalized integrated squared error (Empirical MSE) of the calibrated codes	Empirical normalized integrated squared error (Empirical MSE) of the posterior predictors
$x_1 \mapsto y_1$	$\frac{\sum_{x_1}\left(y_1(x_1)-\mathbf{h_1}(x_1)^t\widehat{\beta}_{\mathbf{1}}\right)^2}{\sum_{x_1}(y_1(x_1))^2}$	
$x_2 \mapsto y_2$	$\frac{\sum_{x_2}\left(y_2(x_2)-\mathbf{h_2}(x_2)^t\widehat{\beta}_{\mathbf{2}}\right)^2}{\sum_{x_2}(y_2(x_2))^2}$	
$x_1 \mapsto y_2$	$\frac{1}{\sum_{x_1}(y_2(x_1))^2}\sum_{x_1}(y_2(x_1) - \mathbf{h_2}\left(\mathbf{h_1}(x_1)^t\widehat{\beta}_{\mathbf{1}}\right)^t\widehat{\beta}_{\mathbf{2}})^2$	$\frac{1}{\sum_{x_1}(y_2(x_1))^2}\sum_{x_1}(y_3(x_1) - \mathbf{h_3}(x_1)^t\widehat{\beta}_{\mathbf{c}} - \mathbf{r_3}(x_1)^t\mathbf{R_3}^{-1}\left(\mathbf{y_{3,obs}} - \mathbf{h_3}(\mathbf{x_{1,obs}})^t\widehat{\beta}_{\mathbf{c}}\right))^2$ where : $\mathbf{r_3}(x_1)^t = C_3(x_1;\mathbf{x_{1,obs}})$ $\mathbf{R_3} = C_3(\mathbf{x_{1,obs}};\mathbf{x_{1,obs}})$

Table 4: This table presents the normalized empirical integrated mean squared error of the calibrated codes and predictors for the 'black-box' method

Nested model error	Zero-mean Gaussian process, combination of the propagation of the first phenomenon's uncertainty ϵ_1 and the second phenomenon's uncertainty ϵ_2 (ϵ_1 et ϵ_2 zero-mean Gaussian process with covariance function Matérn $\frac{5}{2}$), see equation (4)
Calibration of parameters $\beta_{\mathbf{c}} = (\beta_{\mathbf{1}}, \beta_{\mathbf{2}})$	$\beta_{\mathbf{c}}$ is calibrated given the observations of the three phenomena ($x_1 \mapsto y_1$, $x_2 \mapsto y_2$ and $x_1 \mapsto y_2$)
Posterior predictor for phenomenon $x_1 \mapsto y_1$	Gaussian predictor given the observations of the three phenomena ($x_1 \mapsto y_1$, $x_2 \mapsto y_2$ and $x_1 \mapsto y_2$)
Posterior predictor for phenomenon $x_2 \mapsto y_2$	Gaussian predictor given the observations of the three phenomena ($x_1 \mapsto y_1$, $x_2 \mapsto y_2$ and $x_1 \mapsto y_2$)
Posterior predictor for phenomenon $x_1 \mapsto y_2$	Gaussian predictor given the observations of the three phenomena ($x_1 \mapsto y_1$, $x_2 \mapsto y_2$ and $x_1 \mapsto y_2$)

Table 5: This table summarizes the 'grouped' method's characteristics

5 THE 'GROUPED' APPROACH

In this section we present a particular case of the previously presented general framework (see 2). The 'grouped' method corresponds to the case where observations are chosen for the phenomenon $x_1 \mapsto y_1 \mapsto y_2$. So we have ${\mathbf{x}_{\mathbf{1},\mathbf{obs}}}^{(1)} = {\mathbf{x}_{\mathbf{1},\mathbf{obs}}}^{(3)}$, ${\mathbf{x}_{\mathbf{2},\mathbf{obs}}}^{(2)} = {\mathbf{y}_{\mathbf{1},\mathbf{obs}}}^{(1)}$ and ${\mathbf{y}_{\mathbf{2},\mathbf{obs}}}^{(2)} = {\mathbf{y}_{\mathbf{2},\mathbf{obs}}}^{(3)}$.
The method includes the following steps :

1. the parameters' calibration given the observations $\left({\mathbf{x}_{\mathbf{1},\mathbf{obs}}}^{(1)}, {\mathbf{y}_{\mathbf{1},\mathbf{obs}}}^{(1)}\right)$, $\left({\mathbf{x}_{\mathbf{2},\mathbf{obs}}}^{(2)}, {\mathbf{y}_{\mathbf{2},\mathbf{obs}}}^{(2)}\right)$ and $\left({\mathbf{x}_{\mathbf{1},\mathbf{obs}}}^{(3)}, {\mathbf{y}_{\mathbf{2},\mathbf{obs}}}^{(3)}\right)$,
2. the construction of the posterior predictors given the observations :
   (a) for the phenomenon 1 : the proposed predictor is an interpolating Gaussian predictor given the observations $\left({\mathbf{x}_{\mathbf{1},\mathbf{obs}}}^{(1)}, {\mathbf{y}_{\mathbf{1},\mathbf{obs}}}^{(1)}\right)$, $\left({\mathbf{x}_{\mathbf{2},\mathbf{obs}}}^{(2)}, {\mathbf{y}_{\mathbf{2},\mathbf{obs}}}^{(2)}\right)$ and $\left({\mathbf{x}_{\mathbf{1},\mathbf{obs}}}^{(3)}, {\mathbf{y}_{\mathbf{2},\mathbf{obs}}}^{(3)}\right)$,
   (b) for the phenomenon 2 : the proposed predictor is an interpolating Gaussian predictor given the observations $\left({\mathbf{x}_{\mathbf{1},\mathbf{obs}}}^{(1)}, {\mathbf{y}_{\mathbf{1},\mathbf{obs}}}^{(1)}\right)$, $\left({\mathbf{x}_{\mathbf{2},\mathbf{obs}}}^{(2)}, {\mathbf{y}_{\mathbf{2},\mathbf{obs}}}^{(2)}\right)$ and $\left({\mathbf{x}_{\mathbf{1},\mathbf{obs}}}^{(3)}, {\mathbf{y}_{\mathbf{2},\mathbf{obs}}}^{(3)}\right)$,
   (c) for the nested phenomenon : the proposed predictor is an interpolating Gaussian predictor given the observations $\left({\mathbf{x}_{\mathbf{1},\mathbf{obs}}}^{(1)}, {\mathbf{y}_{\mathbf{1},\mathbf{obs}}}^{(1)}\right)$, $\left({\mathbf{x}_{\mathbf{2},\mathbf{obs}}}^{(2)}, {\mathbf{y}_{\mathbf{2},\mathbf{obs}}}^{(2)}\right)$ and $\left({\mathbf{x}_{\mathbf{1},\mathbf{obs}}}^{(3)}, {\mathbf{y}_{\mathbf{2},\mathbf{obs}}}^{(3)}\right)$.

Table 5 summarizes the 'grouped' method's characteristics. Table 6 presents the normalized empirical integrated mean squared error of the calibrated codes and predictors of this method.

6 NUMERICAL EXAMPLE

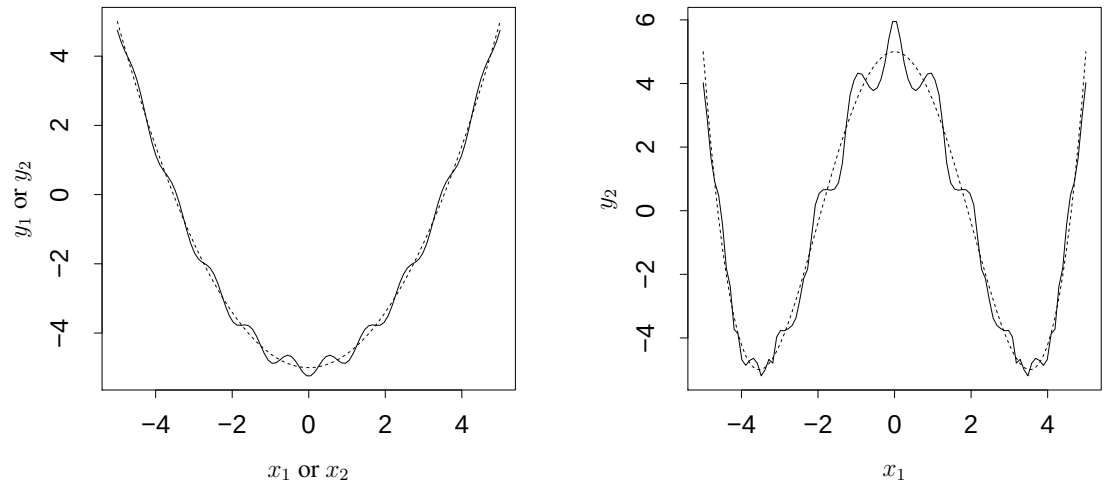
In this section we apply the previously presented particular cases of the general method on a numerical example. In this example we study a computer model coupled with itself.

Phenomenon	Empirical normalized integrated squared error (Empirical MSE) of the calibrated codes	Empirical normalized integrated squared error (Empirical MSE) of the posterior predictors
$x_1 \mapsto y_1$	$\frac{\sum_{x_1}\left(y_1(x_1) - \mathbf{h_1}(x_1)^t \widehat{\beta}_{\mathbf{1}}\right)^2}{\sum_{x_1}(y_1(x_1))^2}$	$\frac{1}{\sum_{x_1}(y_1(x_1))^2}\sum_{x_1}(y_1(x_1) - \mathbf{h_{gp}}(\mathbf{x_{gp,1}})^t \widehat{\beta}_{\mathbf{c}}$ $-\mathbf{r_{gp}}(\mathbf{x_{gp,1}})^t \mathbf{R_{gp}}^{-1}\left(\mathbf{y_{gp,obs}} - \mathbf{h_{gp}}(\mathbf{x_{gp,obs}})^t \widehat{\beta}_{\mathbf{c}}\right))^2$ where : $\mathbf{x_{gp,1}} = (x_1, \emptyset, \emptyset)$
$x_2 \mapsto y_2$	$\frac{\sum_{x_2}\left(y_2(x_2) - \mathbf{h_2}(x_2)^t \widehat{\beta}_{\mathbf{2}}\right)^2}{\sum_{x_2}(y_2(x_2))^2}$	$\frac{1}{\sum_{x_2}(y_2(x_2))^2}\sum_{x_2}(y_2(x_2) - \mathbf{h_{gp}}(\mathbf{x_{gp,2}})^t \widehat{\beta}_{\mathbf{c}}$ $-\mathbf{r_{gp}}(\mathbf{x_{gp,2}})^t \mathbf{R_{gp}}^{-1}\left(\mathbf{y_{gp,obs}} - \mathbf{h_{gp}}(\mathbf{x_{gp,obs}})^t \widehat{\beta}_{\mathbf{c}}\right))^2$ where : $\mathbf{x_{gp,2}} = (\emptyset, x_2, \emptyset)$
$x_1 \mapsto y_2$	$\frac{1}{\sum_{x_1}(y_2(x_1))^2}\sum_{x_1}(y_2(x_1)$ $-\mathbf{h_2}\left(\mathbf{h_1}(x_1)^t \widehat{\beta}_{\mathbf{1}}\right)^t \widehat{\beta}_{\mathbf{2}})^2$	$\frac{1}{\sum_{x_1}(y_2(x_1))^2}\sum_{x_1}(y_3(x_1) - \mathbf{h_{gp}}(\mathbf{x_{gp,3}})^t \widehat{\beta}_{\mathbf{c}}$ $-\mathbf{r_{gp}}(\mathbf{x_{gp,3}})^t \mathbf{R_{gp}}^{-1}\left(\mathbf{y_{gp,obs}} - \mathbf{h_{gp}}(\mathbf{x_{gp,obs}})^t \widehat{\beta}_{\mathbf{c}}\right))^2$ where : $\mathbf{x_{gp,3}} = (\emptyset, \emptyset, x_1)$

Table 6: This table presents the normalized empirical integrated mean squared error of the calibrated codes and predictors for the 'grouped' method

Method	Average variance over a test set according to the source of uncertainty
General method	$$\epsilon_{cond} = \mathbb{E}_{\mathbf{x_{gp}}^{g^{al}}=(\emptyset,\emptyset,x_1)}\left[\mathbf{C_{gp}}\left(\mathbf{x_{gp}}^{g^{al}};\mathbf{x_{gp}}^{g^{al}}\right) - \mathbf{r_{gp}}\left(\mathbf{x_{gp}}^{g^{al}}\right)^t \mathbf{R_{gp}}^{-1}\mathbf{r_{gp}}\left(\mathbf{x_{gp}}^{g^{al}}\right)\right]$$ $$\epsilon_{\beta} = \mathbb{E}_{\mathbf{x_{gp}}^{g^{al}}=(\emptyset,\emptyset,x_1)}\left[\mathbf{u_{gp}}\left(\mathbf{x_{gp}}^{g^{al}}\right)^t \text{cov}\left(\beta_{\mathbf{c}} \mid \mathbf{x_{gp,obs}}\right)\mathbf{u_{gp}}\left(\mathbf{x_{gp}}^{g^{al}}\right)\right]$$
Parallel method	$$\epsilon_{cond} = \epsilon_{cond,1} + \epsilon_{cond,2}$$ $$\epsilon_{cond,1} = \mathbb{E}_{x_1}\left[\left[\frac{\partial \mathbf{h_2}}{\partial x_2}\left(\widehat{y}_1\left(x_1,\widehat{\beta}_{\mathbf{1}}\right)\right)\right]^t \widehat{\beta}_{\mathbf{2}}\left(\sigma_1^2 - \mathbf{r_1}\left(x_1\right)^t \mathbf{R_1}^{-1}\mathbf{r_1}\left(x_1\right)\right)\right]$$ $$\epsilon_{cond,2} = \mathbb{E}_{x_1}\left[\sigma_2^2 - \mathbf{r_2}\left(\widehat{y}_1\left(x_1,\widehat{\beta}_{\mathbf{1}}\right)\right)^t \mathbf{R_2}^{-1}\mathbf{r_2}\left(\widehat{y}_1\left(x_1,\widehat{\beta}_{\mathbf{1}}\right)\right)\right]$$ $$\epsilon_{\beta} = \epsilon_{\beta,1} + \epsilon_{\beta,2}$$ $$\epsilon_{\beta,1} = \mathbb{E}_{x_1}\left[\left[\frac{\partial \mathbf{h_2}}{\partial x_2}\left(\widehat{y}_1\left(x_1,\widehat{\beta}_{\mathbf{1}}\right)\right)\right]^t \widehat{\beta}_{\mathbf{2}}\mathbf{u_1}\left(x_1\right)^t \text{cov}\left(\beta_{\mathbf{1}} \mid \mathbf{y_{1,obs}}^{(1)}\right)\mathbf{u_1}\left(x_1\right)\right]$$ $$\epsilon_{\beta,2} = \mathbb{E}_{x_1}\left[\mathbf{u_2}\left(\widehat{y}_1\left(x_1,\widehat{\beta}_{\mathbf{1}}\right)\right)^t \text{cov}\left(\beta_{\mathbf{2}} \mid \mathbf{y_{2,obs}}^{(2)}\right)\mathbf{u_2}\left(\widehat{y}_1\left(x_1,\widehat{\beta}_{\mathbf{1}}\right)\right)\right]$$
'Black-box' method	$$\epsilon_{cond} = \mathbb{E}_{x_1}\left[C_3\left(x_1;x_1\right) - \mathbf{r_3}\left(x_1\right)^t \mathbf{R_3}^{-1}\mathbf{r_3}\left(x_1\right)\right]$$ $$\epsilon_{\beta} = \mathbb{E}_{x_1}\left[\mathbf{u_3}\left(x_1\right)^t \text{cov}\left(\beta_{\mathbf{c}} \mid \mathbf{y_{2,obs}}^{(3)}\right)\mathbf{u_3}\left(x_1\right)\right]$$ $$\mathbf{u_3}\left(x_1\right) = \mathbf{h_3}\left(x_1\right) - \mathbf{h_3}\left(\mathbf{x_{1,obs}}^{(3)}\right)\mathbf{R_3}^{-1}\mathbf{r_3}\left(x_1\right)$$
Grouped method	see General method

Table 7: This table presents the sources of uncertainty of the nested phenomenon's predictors. We denote by ϵ_{cond} the part of variance coming from the conditional Gaussian process and by ϵ_{β} the part of variance coming from the parameters' posterior distribution.



(a) Model (code 1 or 2) (b) Nested model

Figure 1: The continuous line corresponds to reality i.e. $y_1(x_1)$, $y_2(x_2)$ or $y_2(x_1)$ and the dashed line correspond to the codes $\mathbf{h_1}(x_1)^t \beta_\mathbf{1}$, $\mathbf{h_2}(x_2)^t \beta_\mathbf{2}$ or $\mathbf{h_2}\left(\mathbf{h_1}(x_1)^t \beta_\mathbf{1}\right)^t \beta_\mathbf{2}$. Figure 1a presents phenomena $x_1 \mapsto y_1$, $x_2 \mapsto y_2$ and figure 1b phenomenon $x_1 \mapsto y_2$.

This section is organised as follow. Firstly, the phenomenon and its associated computer code are introduced. Secondly, the way to choose the observations is presented. Thirdly, the method to estimate the hyperparameters of the model error is exposed. Fourthly, the test sets of the posterior predictors are presented. Finally the achieved results for the 'parallel', 'black-box' and 'grouped' methods on this example are showed.

6.1 Presentation of the phenomenon

The model properties of the numerical example are :

$$\mathbf{h_1}(x) = \mathbf{h_2}(x) = \begin{bmatrix} 1 \\ x^2 \end{bmatrix} \tag{38}$$

$$\begin{gathered} \beta_\mathbf{1} = \beta_\mathbf{2} = \begin{bmatrix} -5 \\ 0.4 \end{bmatrix} \\ y_1(x) = y_2(x) = \mathbf{h_1}(x)^t \beta_\mathbf{1} - \tfrac{1}{4}\cos(2\pi x) = \mathbf{h_2}(x)^t \beta_\mathbf{2} - \tfrac{1}{4}\cos(2\pi x) \end{gathered} \tag{39}$$

Figure 1 presents the model alone and the nested model. It can be seen that the errors of the phenomena $x_1 \mapsto y_1$ and $x_2 \mapsto y_2$ are regular. The error of the nested phenomenon is more chaotic.

The three previously presented methods are applied on this example.

6.2 Observations' choice

Regarding the number of observations three cases are studied : 4, 8 or 12 observations. The observations are generated given a maximin Latin Hypercube Sampling for $\mathbf{x_{1,obs}}^{(1)}$

and $\mathbf{x_{2,obs}}^{(2)}$ in the parallel method (in this example $\mathbf{x_{1,obs}}^{(1)} = \mathbf{x_{2,obs}}^{(2)}$), for $\mathbf{x_{1,obs}}^{(3)}$ for the 'black-box' method and for $\mathbf{x_{1,obs}}^{(1)}$ and $\mathbf{x_{1,obs}}^{(3)}$ for the grouped method (in this example $\mathbf{x_{1,obs}}^{(1)} = \mathbf{x_{1,obs}}^{(3)}$). The observations' sampling, the parameters' calibration and the posterior predictors' construction are repeated 50 times for each case.

6.3 Covariance functions' hyperparameters

The covariance functions' hyperparameters are assumed to be known before the parameters' calibration. In the parallel and the grouped methods the hyperparameters of the Matérn covariance functions of ϵ_1 and ϵ_2 are estimated by a Restricted Maximum Likelihood Estimator (see [1]). In the 'black-box' method, ϵ_3 is assumed to be a zero-mean stationary Gaussian process. The hyperparameters of its Matérn covariance function ϵ_3 are estimated by a Restricted Maximum Likelihood Estimator.

6.4 Analysis of the predictors' performances

Once the predictors are built, their performances are analyzed on 100 points' grids over the input space of each phenomenon. For the three phenomena the input space is $[-5, 5]$.

6.5 The parallel approach

In the parallel approach, there are observations for the phenomena $x_1 \mapsto y_1$ and $x_2 \mapsto y_2$ which are denoted by $\left(\mathbf{x_{1,obs}}^{(1)}, \mathbf{y_{1,obs}}^{(1)}\right)$ and $\left(\mathbf{x_{2,obs}}^{(2)}, \mathbf{y_{2,obs}}^{(2)}\right)$. Figure 2 presents for the phenomena $x_1 \mapsto y_1$, $x_2 \mapsto y_2$ and $x_1 \mapsto y_2$ the calibrated codes' and the predictors' accuracy for the 'parallel' method. Figure 3 presents an example of posterior predictor for each phenomenon (1, 2 and nested). Figure 4 presents the sources of uncertainty of the nested phenomenon's predictor. The findings are :

- The more observations there are, the more accurate predictions and the better calibrated codes are.
- The predictors' performance is less sensitive to the observations' choice when the number of observations increases.
- Both mean and confidence interval of 95% of all predictors are accurate.
- The predictors for the phenomena $x_1 \mapsto y_1$ and $x_2 \mapsto y_2$ are interpolating. The predictor for the phenomenon $x_1 \mapsto y_2$ is a linearization of coupling the predictors of the first and the second phenomena, thus it is not interpolating because there is no observation of phenomenon $x_1 \mapsto y_2$.
- When there are few observations, the uncertainty of the nested phenomenon's predictor mainly comes from the parameters' posterior distribution. The uncertainty lowers with the numbers of observations. This reduction with the number of observations is particularly significant for the uncertainty coming from the parameters' distribution.

6.6 The 'black-box' approach

In the 'black-box' approach there are observations for the nested phenomenon only, which are denoted by $\left(\mathbf{x_{1,obs}}^{(3)}, \mathbf{y_{2,obs}}^{(3)}\right)$. Figure 5 presents for the phenomena $x_1 \mapsto y_1$,

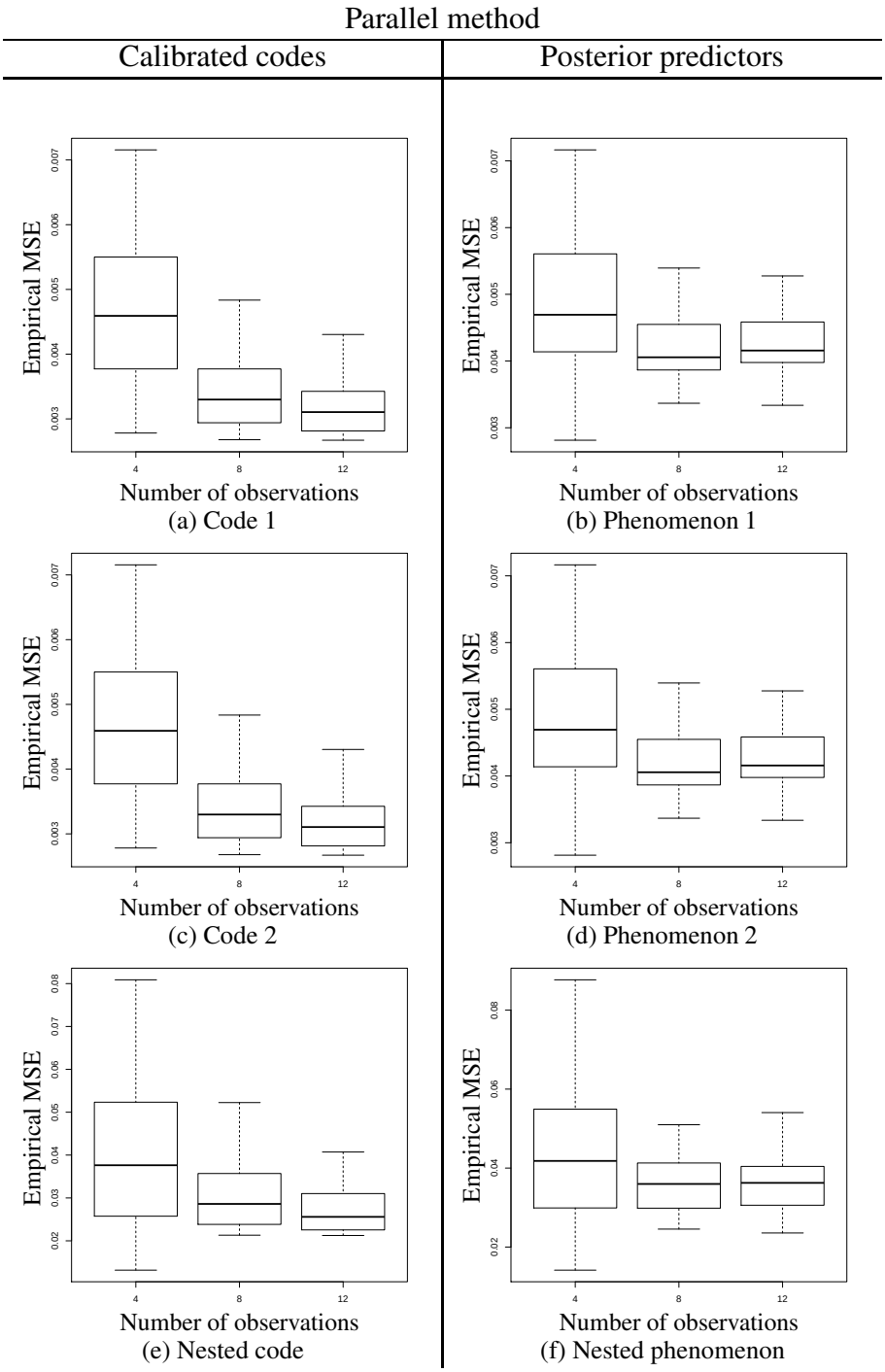
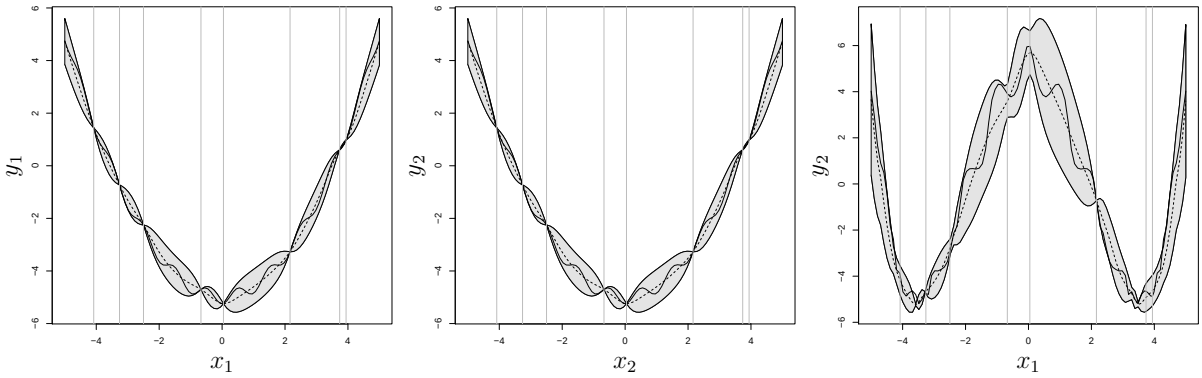


Figure 2: The figures present the empirical MSE (see table 2) of the calibrated codes (2a, 2c, 2e) and the predictors (2b, 2d, 2f) for the "parallel" method. The parameters' calibration and the posterior predictors' construction have been repeated with 50 series of observations. The same observations were used for both codes ($\mathbf{x_{1,obs}}^{(1)} = \mathbf{x_{2,obs}}^{(2)}$). The lower the empirical MSE is, the more accurate the calibrated code is. It can be seen that the accuracy of the predictors and calibrated codes increases with the number of observations and that all codes are well calibrated.

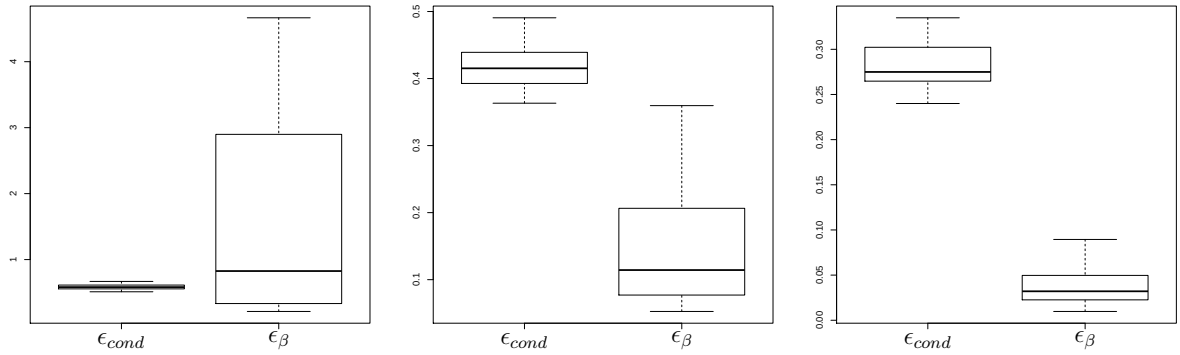
Parallel method : Posterior predictors' mean and confidence interval



(a) Phenomenon 1 (b) Phenomenon 2 (c) Nested phenomenon

Figure 3: Each figure presents an example of posterior predictor for each phenomenon. The dotted line represents the predictor's mean, the continuous line the real phenomenon and the grey area the predictor's confidence interval of 95%. The predictors are built with the same 8-point LHS maximin designs for $x_1 \mapsto y_1$ and $x_2 \mapsto y_2$ phenomena ($\mathbf{x_{1,obs}}^{(1)}$=$\mathbf{x_{2,obs}}^{(2)}$). It can be seen that the predictors' means and confidence intervals of 95% are accurate. The predictor of the nested code is not an interpolating Gaussian predictor because there is no observation for the phenomenon $x_1 \mapsto y_2$.

Parallel method : sources of uncertainty of the nested phenomenon's predictor



(a) 4 observations (b) 8 observations (c) 12 observations

Figure 4: The figures above present the distribution of the average variance of the nested phenomenon's predictor over a test set of 100 points. We denote by ϵ_{cond} the part of variance coming from the conditional Gaussian process and by ϵ_β the part of variance coming from the parameters' posterior distribution (see table 7). It can be seen that the variance lowers with the number of observations. This reduction is particularly significant for the uncertainty coming from the parameters' distribution.

$x_2 \mapsto y_2$ and $x_1 \mapsto y_2$ the calibrated codes' and predictors' accuracy for the 'black-box' method. Figure 6 presents an example of posterior predictor for each phenomenon (1, 2 and nested). Figure 7 presents the sources of uncertainty of the nested phenomenon's predictor. The findings are :

- The more observations there are, the more accurate the nested predictor is.
- When the number of observations increases, the parameters' calibration is less sensitive to the observations' choice.
- The second code is better calibrated than the first code.
- The nested code is never well calibrated.
- The mean of the nested phenomenon's predictor is accurate, but its variance is relatively high.
- The predictors of the first and the second phenomena are just calibrated codes. Their variances are relatively high.
- When there are few observations, the uncertainty of the nested phenomenon's predictor mainly comes from the parameters' posterior distribution. The uncertainty lowers with the number of observations. This reduction is particularly significant for the uncertainty coming from the parameters' distribution.

6.7 The 'grouped' approach

In this case observations of the three phenomena are available, which are denoted by $\left(\mathbf{x_{1,obs}}^{(1)}, \mathbf{y_{1,obs}}^{(1)}\right)$, $\left(\mathbf{x_{2,obs}}^{(2)}, \mathbf{y_{2,obs}}^{(2)}\right)$ and $\left(\mathbf{x_{1,obs}}^{(3)}, \mathbf{y_{2,obs}}^{(3)}\right)$. In this example we consider that $\mathbf{x_{1,obs}}^{(1)} = \mathbf{x_{1,obs}}^{(3)}$, $\mathbf{y_{1,obs}}^{(1)} = \mathbf{x_{2,obs}}^{(2)}$ and $\mathbf{y_{2,obs}}^{(2)} = \mathbf{y_{2,obs}}^{(3)}$. Figure 8 presents for the phenomena $x_1 \mapsto y_1$, $x_2 \mapsto y_2$ and $x_1 \mapsto y_2$ the calibrated codes' and the predictors' accuracy for the 'grouped' method. Figure 9 presents an example of the posterior predictors for each phenomenon (1, 2 and nested). Figure 10 presents the sources of uncertainty of the nested phenomenon's predictor. The findings are :

- Both codes are well calibrated,
- For the phenomenon $x_1 \mapsto y_1$, the accuracy of the calibrated code increases with the number of observations,
- For the phenomenon $x_2 \mapsto y_2$, the accuracy of the calibrated code decreases with the number of observations. This is probably due to an ill-conditioned problem. The space filling of the observations $\mathbf{x_{2,obs}}^{(2)}$ is not good.
- When there are few observations, the uncertainty of the nested phenomenon's predictor mainly comes from the parameters' posterior distribution. The uncertainty lowers with the number of observations. This reduction is particularly significant for the uncertainty coming from the parameters' distribution.

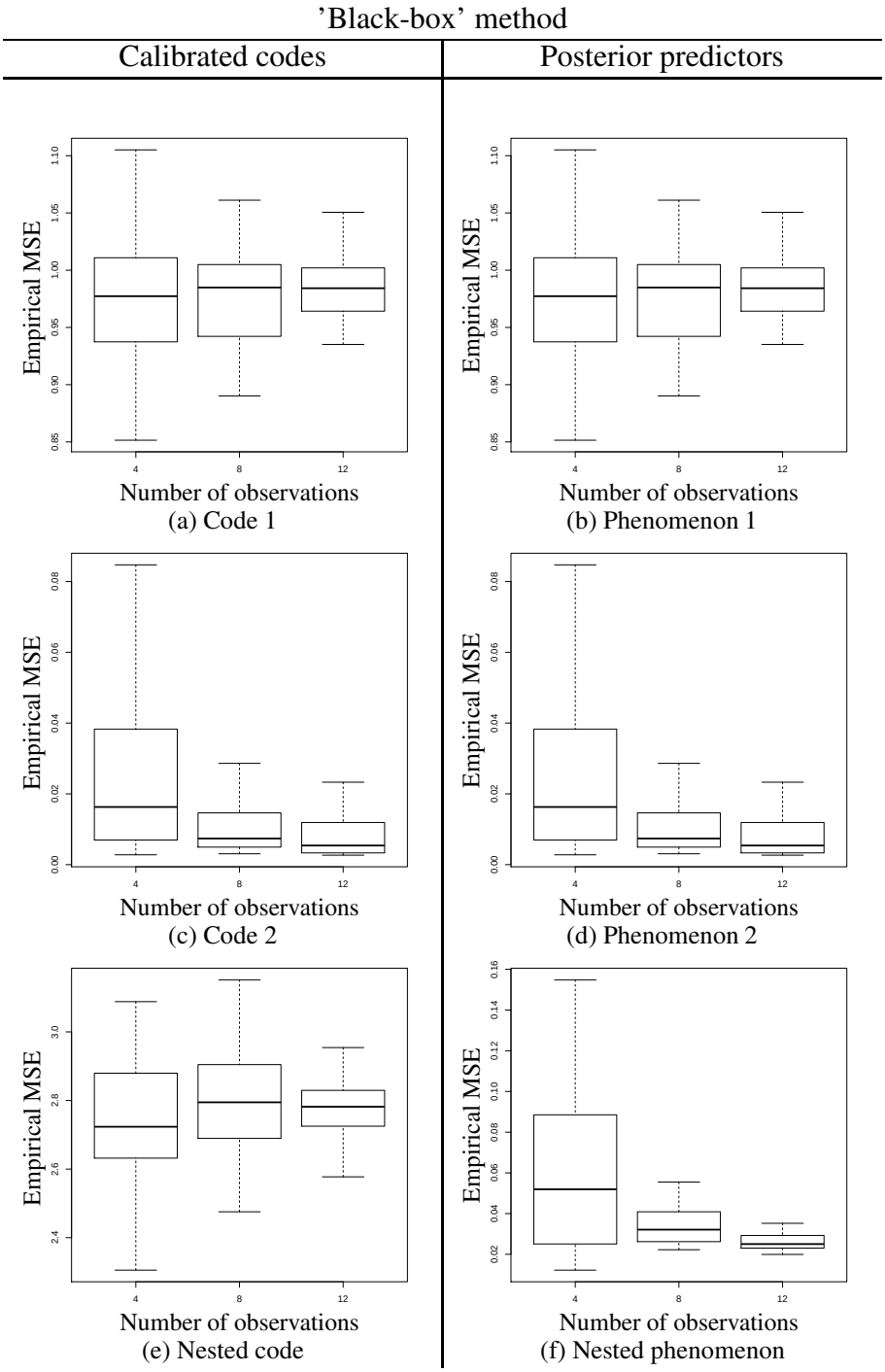
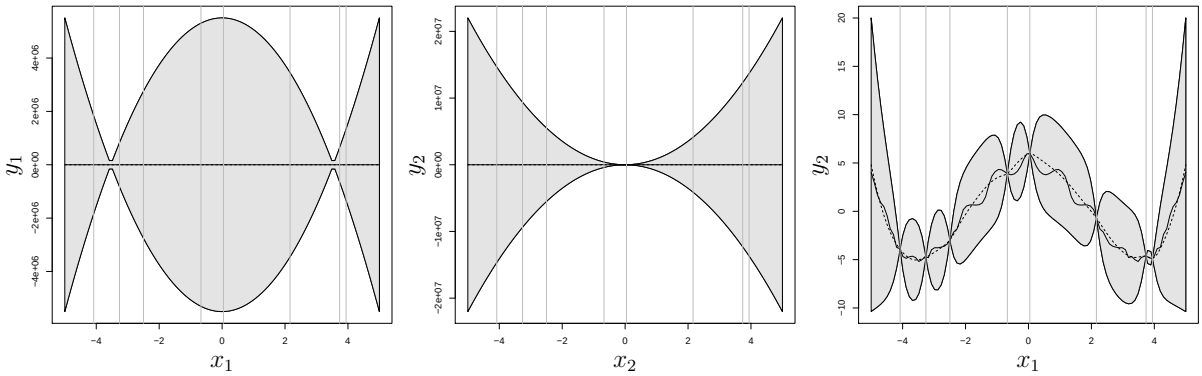


Figure 5: The figures present the empirical MSE (see table 4) of the calibrated codes (5a, 5c, 5e) and the predictors (5b, 5d, 5f) for the "black-box" method. The parameters' calibration and the posterior predictors' construction have been repeated for 50 series of observations. It can be seen that the second code is always better calibrated than the first code. When the number of observations increases, the empirical MSE lowers and is less sensitive to the observations' choice. So the calibrated codes' and predictors' robustness and accuracy increases with the number of observations.

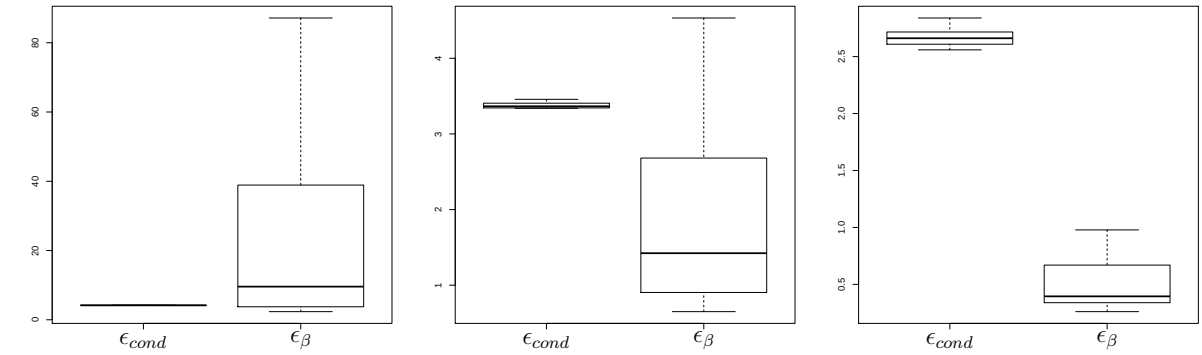
'Black-box' method : Posterior predictors' mean and confidence interval



(a) Phenomenon 1 (b) Phenomenon 2 (c) Nested phenomenon

Figure 6: Each figure presents an example of posterior predictor for each phenomenon. The dotted line represents the predictors' mean, the continuous line the real phenomenon and the grey area the confidence interval of 95%. The predictors are built with a 8-point LHS maximin design. The confidence intervals are relatively high. The predictors of the first and the second phenomena are not interpolating because there is no observation for these phenomena ($\left(\mathbf{x_{1,obs}}^{(1)}, \mathbf{y_{1,obs}}^{(1)}\right) = \emptyset$ and $\left(\mathbf{x_{2,obs}}^{(2)}, \mathbf{y_{2,obs}}^{(2)}\right) = \emptyset$). The predictors for these phenomena are just calibrated codes integrating posterior parameters distribution.

'Black-box' method : sources of uncertainty of the nested phenomenon's predictor



(a) 4 observations (b) 8 observations (c) 12 observations

Figure 7: The figures above present the distribution of the average variance of the nested phenomenon's predictor over a test set of 100 points. We denote by ϵ_{cond} the part of the variance coming from the conditional Gaussian process and by ϵ_β the part of the variance coming from the parameters' posterior distribution. The variance lowers with the number of observations, especially for the part coming from the parameters' distribution (see table 7). The variance coming from the conditional process is relatively high because the error of the nested model is chaotic (see figure 1)

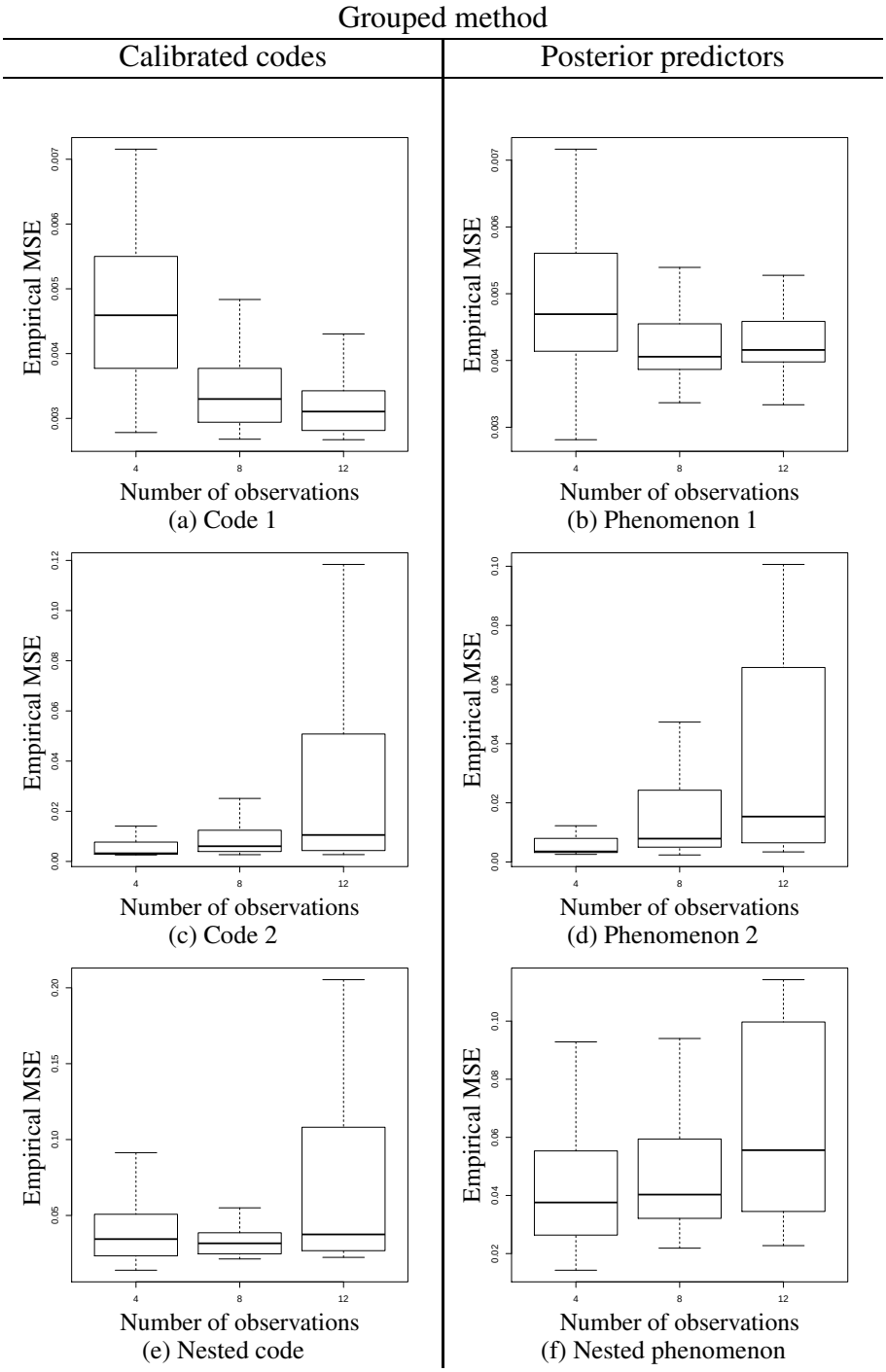
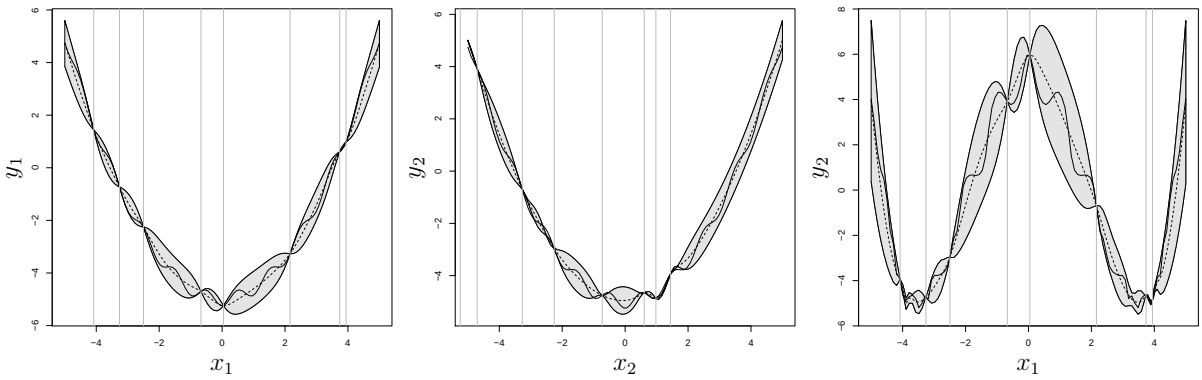


Figure 8: The figures present the empirical MSE (see table 6) of the calibrated codes (8a, 8c, 8e) and the predictors (8b, 8d, 8f) for the "grouped" method. The observations' sampling, the parameters' calibration and the posterior predictors' construction have been repeated 50 times. All codes are well calibrated. The calibrated code's and predictor's accuracy of the phenomenon 1 increases with the number of observations. The calibrated code's and predictor's accuracy of the phenomenon 2 is poorer than for the phenomenon 1 and decreases with the number of observations. This is due to the bad space filling of ${\mathbf{y}_{\mathbf{1},\mathbf{obs}}}^{(1)} = {\mathbf{x}_{\mathbf{2},\mathbf{obs}}}^{(2)}$. Indeed the observations are chosen for phenomenon $x_1 \mapsto x_2 \mapsto y_2$. So even if the observations ${\mathbf{x}_{\mathbf{1},\mathbf{obs}}}^{(1)}$ fill in well the space $[-5;5]$, the observations ${\mathbf{y}_{\mathbf{1},\mathbf{obs}}}^{(1)} = {\mathbf{x}_{\mathbf{2},\mathbf{obs}}}^{(2)}$ are not well distributed in $[-5;5]$. This leads to an ill-conditioned problem.

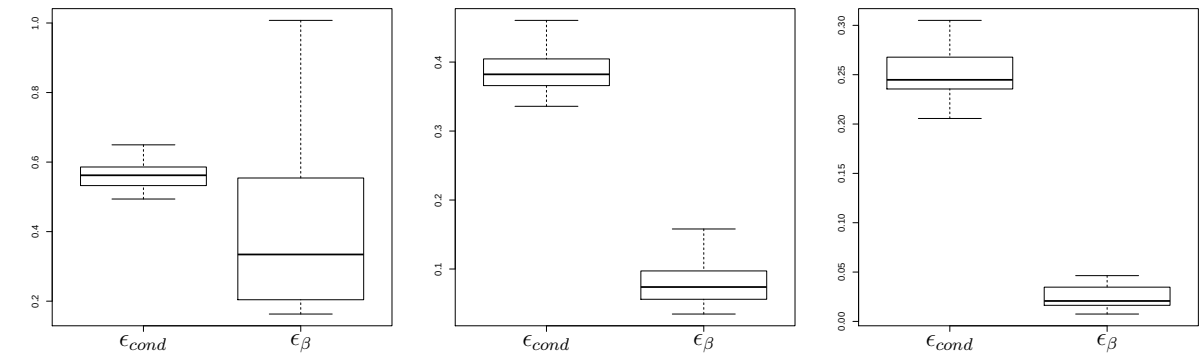
Grouped method : Posterior predictors' mean and confidence interval



(a) Phenomenon 1 (b) Phenomenon 2 (c) Nested phenomenon

Figure 9: Each figure presents an example of posterior predictor for each phenomenon. The dotted line represents the predictors' mean, the continuous line the real phenomenon and the grey area the confidence interval of 95%. The predictors are built with a 8-point LHS maximin design $x_1 \mapsto x_2 \mapsto y_2$. It can be seen that both mean and confidence interval are well estimated. The observations of the phenomenon 2 are not well distributed in $[-5; 5]$. All the predictors are Gaussian and interpolating.

Grouped method : sources of uncertainty of the nested phenomenon's predictor



(a) 4 observations (b) 8 observations (c) 12 observations

Figure 10: The figures above present the distribution of the average variance of the nested phenomenon's predictor over a test set of 100 points. We denote by ϵ_{cond} the part of variance coming from the conditional Gaussian process and by ϵ_β the part of variance coming from the parameters' posterior distribution (see table 7). It can be seen that the variance lowers with the number of observations. This reduction is particularly significant for the uncertainty coming from the parameters' distribution. It can be seen that both part of variance are relatively low even when there are few observations.

7 CONCLUSIONS AND FURTHER WORK

A general framework has been proposed to calibrate nested computer models whatever the type of available observations among the three possibilities $\left(\mathbf{x_{1,obs}}^{(1)}, \mathbf{y_{1,obs}}^{(1)}\right)$, $\left(\mathbf{x_{2,obs}}^{(2)}, \mathbf{y_{2,obs}}^{(2)}\right)$ and $\left(\mathbf{x_{1,obs}}^{(3)}, \mathbf{y_{2,obs}}^{(3)}\right)$. Predictors for all phenomena have been proposed for all possibilities of available observations. Three specific cases of this general method have been presented depending on the available observations. The performance of the codes' calibration and the posterior predictors has been analyzed on a numerical example for these three specific cases.

Questions remain about the choice of the design points. In particular our general framework could enable designs where there are not the same number of observations for the first and the second phenomena. Thus parsimonious or sequential (see [2], [3], [4], [5], [8], [9], [10]) designs could be proposed in order to improve the accuracy of the nested phenomenon's predictor.

REFERENCES

[1] F. Bachoc. Calibration and improved prediction of computer models by universal kriging. *Nuclear Science and Engineering*, 176:81–97, 2014.

[2] J. Bect, D. Ginsbourger, L. Li, V. Picheny, and E. Vazquez. Sequential design of computer experiments for the estimation of a probability of failure. *Stat. Comput.*, 22:773–793, 2012.

[3] B. J. Bichon, M. S. Eldred, L. P. Swiler, S. Mahadevan, and J. M. Mcfarland. Efficient Global Reliability Analysis for Nonlinear Implicit Performance Functions. *AIAA J.*, 46:2459–2468, 2008.

[4] C. Chevalier, J. Bect, D. Ginsbourger, and E. Vazquez. Fast parallel kriging-based stepwise uncertainty reduction with application to the identification of an excursion set. *Technometrics*, 56(4):455–465, 2014.

[5] B. Echard, N. Gayton, and M. Lemaire. AK-MCS: An active learning reliability method combining Kriging and Monte Carlo Simulation. *Struct. Saf.*, 33:145–154, 2011.

[6] M. C. Kennedy and A. O'Hagan. Predicting the output from a complex computer code when fast approximations are avalaible. *Biometrika*, 87:1–13, 2000.

[7] M. C. Kennedy and A. O'Hagan. Bayesian Calibration of Computer Models. *J. R. Stat. Soc. . Ser. B (Statistical Methodol.*, 63(3):425–464, 2001.

[8] G. Perrin. Active learning surrogate models for the conception of systems with multiple failure modes. *Reliability Engineering and System Safety*, 149:130–136, 2016.

[9] V. Picheny, D. Ginsbourger, O. Roustant, R. T. Haftka, and N.-H. Kim. Adaptive Designs of Experiments for Accurate Approximation of a Target Region. *Am. Soc. Mech. Eng.*, 132(7), 2010.

[10] P. Ranjan, D. Bingham, and G. Michailidis. Sequential Experiment Design for Contour Estimation From Complex Computer Codes. *Technometrics*, 50(4):527–541, 2008.

[11] J. Sacks, W. Welch, T. J. Mitchell, and H. P. Wynn. Design and Analysis of Computer Experiments. *Stat. Sci.*, 4:409–435, 1989.

[12] T. J. Santner, B. J. Williams, and W. Notz. *The design and analysis of computer experiments*. Springer series in statistics. Springer, New York, 2003.