

An Optimization Strategy Based on the Maximization of Matching-Targets' Probability for Unevaluated Results

M. Feraille*

IFP Energies nouvelles, 1-4 avenue de Bois-Préau, 92852 Rueil-Malmaison Cedex - France
e-mail: mathieu.feraille@ifpen.fr

* Corresponding author

Résumé — Optimisation par maximisation de la probabilité d'atteindre les cibles pour des résultats non évalués — La méthode présentée dans cet article rentre dans le cadre de l'optimisation probabiliste classique. La nouveauté consiste à construire un processus Gaussien pour chaque résultat souhaité (*i.e.* associé à chaque cible spécifiée) et de les utiliser ensuite pour estimer les densités de probabilité des résultats non évalués. Ces densités sont alors prises en compte dans le calcul de la densité *a posteriori* des paramètres à optimiser. Cette approche est adaptée lorsque chaque évaluation de la fonction est coûteuse en temps de calcul. Une description détaillée de cette méthode d'optimisation est proposée ainsi que son utilisation sur plusieurs cas tests.

Abstract — An Optimization Strategy Based on the Maximization of Matching-Targets' Probability for Unevaluated Results — *The Maximization of Matching-Targets' Probability for Unevaluated Results (MMTPUR), technique presented in this paper, is based on the classical probabilistic optimization framework. The numerical function values that have not been evaluated are considered as stochastic functions. Thus, a Gaussian process uncertainty model is built for each required numerical function result (i.e., associated with each specified target) and is used to estimate probability density functions for unevaluated results. Parameter posterior distributions, used within the optimization process, then take into account these probabilities. This approach is particularly adapted when, getting one evaluation of the numerical function is very time consuming. In this paper, we provide a detailed outline of this technique. Finally, several test cases are developed to stress its potential.*

INTRODUCTION

Optimization or inverse problems are usually associated with the minimization of an Objective Function (OF). This OF represents the mismatch between the targets to reach and the corresponding numerical function values or simulation results. Moving to a probabilistic optimization framework makes it possible to extend this OF minimization process to the maximization of the *posterior* distribution of the unknown parameters or variables. This *posterior* distribution depends on prior uncertainties on both parameters and targets (Tarantola, 2005). The bridge between this OF minimization problem and the associated classical probabilistic framework relies on assumptions about tolerances relative to targets and numerical function errors. The key points of the methodology described in this paper consist:

- in approximating these tolerances by stochastic variables; and,
- in considering Gaussian processes (Sacks *et al.*, 1989) to model numerical function errors.

Based on the classical probabilistic optimization framework, we developed a technique called Maximization of Matching-Targets' Probability for Unevaluated Results (MMTPUR). Briefly, the numerical function values that have not been evaluated are considered as stochastic functions: a Gaussian process model is computed for each simulation result associated with a target. The *posterior* distribution of the unknown parameters is then built, accounting for all of these Gaussian process models. New points are iteratively added by maximizing this distribution.

Gaussian process models are widely used either to approximate responses of costly numerical models or to perform optimization.

Jourdan (2000) and Marrel (2008) use Gaussian processes to approximate key outputs of numerical models from a fixed number of simulations. These approximations can also be used to perform uncertainty management, for instance to propagate input uncertainty towards output results and to perform sensitivity analysis.

Adaptive design strategies have also been proposed instead of fixed design to improve Gaussian process approximations for given responses (Scheidt *et al.*, 2007; Busby, 2009). Thus, new simulation points are iteratively added with various strategies, in domains either where approximations are poor, where responses are complex, or where information is required. The overall objective is to get a good approximation of the responses of interest over the entire uncertain domain. One specific feature of the approach developed by Scheidt *et al.* (2007) is to add pilot points, at each

iteration, in order to improve the Gaussian process model. These pilot points are fictitious data used to refine the final model.

As described in Schonlau (1997), Villemonteix *et al.* (2009) and Vazquez and Bect (2010), Gaussian process properties, with characterization of mean and covariance, can also be used for local or global optimization. In this case, the required OF is approximated by a Gaussian process and an improvement function corresponding to the potential reduction of the OF is defined. The classical OF minimization is then turned into the maximization of the improvement expectation, known as the expected improvement or gain. Adding a global optimization criterion to the expected improvement method yields the generalized expected improvement method. For a specific case the generalized expected improvement method corresponds to maximize the probability to be lower than the lowest available values of the OF.

Busby and Feraille (2008) and Feraille and Marrel (2012) employed Gaussian processes within a probabilistic optimization framework to handle history-matching problems in reservoir engineering. Within this context, the OF quantifies the mismatch between target production data and the corresponding numerical responses. These authors proposed to approximate the OF by a Gaussian process using an adaptive algorithm and focused on areas where the OF was low (which is the usual ultimate history-matching goal). This OF approximation provided the *posterior* distribution of the unknown parameters in a reasonable number of simulations.

The main differences between the MMTPUR technique and the methods previously introduced are listed hereafter. First, we approximate using different Gaussian process models several responses of interest instead of a single one. Then, all the models are locally improved from a sequence of adjustments within uncertain domains where the simulation results are in agreement with the targets. This is not performed over the entire uncertain domain. Finally, the optimization strategy entails the maximization of the *posterior* distribution in accounting for several Gaussian process models (one for each required numerical function result which is associated with a given target) rather than a single model.

In the following sections, we focus first on the classical probabilistic optimization framework. We then go into the details of the Gaussian process technique, which is used to approximate the output results of a simulator based upon response surface modeling. We show how to take advantage of this add-on to the classical probabilistic formulation of inverse problems: the Gaussian

process technique is used to estimate the uncertainty in numerical function values that have not been evaluated. The last section presents the MMTPUR algorithm and several application examples.

1 PROBABILISTIC INVERSE PROBLEM FRAMEWORK

Inverse problems can be solved using optimization techniques. Briefly, an iterative process is implemented to identify a set of parameters x providing a numerical response $y(x)$ as close as possible to target t . One or several values of x as well as their corresponding simulation results $y(x)$ have to be determined. The probabilistic framework described below was proposed to solve this problem (Tarantola, 2005). It entails the following data, requirements and assumptions:

- targets $T = \{T^i\}$, T^i being the i -th target, are associated with an uncertainty model characterized by probability density function $p_T(t)$, where t is a realization of T . As an example, in reservoir engineering history-matching process, the simulation model aims at simulating the production profiles for a given reservoir while the targets can be measured production profiles. Uncertainty is linked to the accuracy in measurements;
- the parameters of the simulation model, X , are associated with an *a priori* uncertainty model described by probability density function $p_X(x)$, x being a realization of X . Considering a Gaussian uncertainty model described by mean parameter values μ_X and covariance operator C_X , the *a priori* probability density function is expressed as:

$$p_X(x) \propto \exp \left\{ -\frac{1}{2} (x - \mu_X)^t C_X^{-1} (x - \mu_X) \right\} \quad (1)$$

$$= \exp (-OF_{prior}(x))$$

- simulation results $y(x) = \{y^i(x)\}$ are considered as a stochastic function, denoted $Y(x) = \{Y^i(x)\}$. It is associated with an uncertainty model described by probability density function $p_{Y(x)}(y)$, for which realization y depends on the x parameter values.

A probabilistic framework is used to define the *a posteriori* uncertainty $p_{X/Y}(x/t) = p(X = x/Y = t)$ on the X parameters that give simulation results, respecting specified targets. Using Bayes' rule this conditional density function is rewritten as:

$$p_{X/Y}(x/t) \propto p_{Y/X}(t/x) p_X(x) \quad (2)$$

$p_X(x)$ is the *a priori* probability density function (defined above) and $p_{Y/X}(t/x) = p(Y = t/X = x)$, the conditional

probability density function of obtaining simulation results y respecting target values t given parameter values x . $p_{Y/X}(t/x)$ is also known as the likelihood function $lkh(x)$. It can be derived from the independent probability density functions of T and $Y(x)$ defined above:

$$lkh(x) = p_{Y/X}(t/x) = \int_{\Gamma} p_T(t') \cdot p_{Y(x)}(t') dt' \quad (3)$$

The Γ integration domain is related to the targets' space, namely: $[-\infty; +\infty]^t$.

One often considers that the y value depends deterministically on the x value. Therefore, $p_{Y(x)}(y)$ is a Dirac probability density function $\delta(\{y(x)\})$, and the likelihood function becomes:

$$lkh(x) = p_{Y/X}(t/x) = p_T(y(x)) \quad (4)$$

Besides, assuming a Gaussian uncertainty model for T , characterized by the mean target values, μ_T , and a covariance operator, C_T , the likelihood function is expressed as:

$$lkh(x) \propto \exp \left\{ -\frac{1}{2} (y(x) - \mu_T)^t C_T^{-1} (y(x) - \mu_T) \right\} \quad (5)$$

$$= \exp (-OF_{lkh}(x))$$

Finally, we obtain the following relationship for the *posterior* probability density function:

$$p_{X/Y}(x/t) \propto lkh(x) \cdot p_X(x) \quad (6)$$

$$\propto \exp (-OF_{lkh}(x) - OF_{prior}(x)) = \exp (-OF(x))$$

This framework is used to determine x values (or realizations of X) that maximize the *a posteriori* probability of X . It consists in finding x values that maximize $p_{X/Y}(x/t)$ or that minimize Objective Function $OF(x)$.

The minimization of the Objective Function can be performed referring to many techniques, said to be global or local depending on whether they converge to global or local minima. For instance, the following types of methods can be highlighted:

- the local gradient-based methods (Tarantola, 2005), which include quasi-Newton and Gauss-Newton families;
- the global evolutionary and genetic methods (Hansen, 2006);
- the global response surface-based methods (Schonlau, 1997; Villemonteix *et al.*, 2009; Vazquez and Bect, 2010).

In the remainder of this paper, we propose to use the known simulation results to approximate $p_{Y(x)}(y)$ from

several Gaussian processes, instead of considering a Dirac probability density function. Before proceeding, the key points of the Gaussian process technique are recapped in the following section. We then show how this technique can be used to estimate probability density function $p_{Y(x)}(y)$ for unevaluated results of the numerical model. Finally, we discuss what this approach implies for the *posterior* distribution of X .

2 RESPONSE APPROXIMATION USING GAUSSIAN PROCESS

Kriging was introduced in Geostatistics (Matheron, 1963). It is nowadays widely used as an approximation to model output functions $\{y(x)\}$ (*i.e.*, the responses provided by numerical simulators) with respect to input $\{x = (x_1, \dots, x_d)\}$. This kriging approximation is a Gaussian Process (GP) characterized by a mean and a covariance structure. It can be shown that the GP or kriging mean is the best linear unbiased estimator.

In other words, we have a set of values for input parameters $x = (x_1, \dots, x_d)$, d being the number of parameters, and we derive the $y(x)$ values from the simulator. Then, we consider $y(x)$ as a realization of the following random function $Y(x)$ (Sacks *et al.*, 1989) which is a GP:

$$Y(x) = h(x) - Z(x) - U(x) \quad (7)$$

where:

- $h(x)$ is the mean of the response. Its shape is *a priori* known: $h(x) = \sum_{j=0}^k \beta_j h_j(x)$. Note that $h(x)$ can be seen as the trend of the response. $h_j(x)$ are elementary functions or terms associated with the shape of h and β are coefficients. Quite often, h_j are the terms associated with a polynomial of order 2 including all or part of the following terms: constant, linear terms (x_i) , first order interaction terms $(x_i x_j)$, quadratic terms (x_i^2) . Employing to more complex h_j functions can also be envisaged;
- $Z(x)$ corresponds to the stochastic part of model Y . It is characterized by a zero mean and covariance

$$E[Z(x_1) \cdot Z(x_2)] = \sigma^2 \exp \left(- \sum_{i=1}^d \left(\frac{|x_{1i} - x_{2i}|}{\lambda_i} \right)^{p_i} \right) = \sigma^2 \cdot R(x_1, x_2)$$

where d is the dimension of x (number of parameters) and $\{\sigma, (p_i)_{i=1, \dots, d}, (\lambda_i)_{i=1, \dots, d}\}$ are the covariance parameters. Assuming that Z is stationary (meaning that covariance does not depend on x values) makes it possible to perform the computations described hereafter;

- $U(x)$ is an independent noise that can be added to the random part of model Y . It is known as the nugget effect in geostatistics. Its usefulness appears when dealing with none-deterministic simulator y . Effectively, it allows for introducing a discontinuity at the origin into the covariance of model $Y(x)$. $U(x)$ is a Gaussian random variable with zero mean and covariance and variance equal to $\varepsilon^2 = \sigma^2 \cdot \tau$; $\tau \geq 0$. Coefficient τ represents the contribution of the nugget effect. When $U(x)$ is used, the covariance of the stochastic part of model Y becomes:

$$E[Z(x_1) \cdot Z(x_2)] = \sigma^2 [R(x_1, x_2) + \tau \cdot \delta(x_1 - x_2)]$$

with δ the Dirac's delta function.

Note that in what follows we will not consider nugget effect.

For a given sample of realizations, denoted $SA = (Y(x_i) = y_i)_{i=1, \dots, n}$ with n the number of points, it can be shown that the conditional mean and variance of Gaussian process $Y(x)$ given SA are:

$$E[Y(x)/SA] = h^t(x)\beta + r^t(x)R^{-1}(y - H\beta) \quad (8)$$

$$\text{var}[Y(x)/SA]$$

$$= \sigma^2 \left[1 - [r^t(x), h^t(x)] \times \begin{bmatrix} R & H \\ H^t & 0 \end{bmatrix}^{-1} \begin{bmatrix} r(x) \\ h(x) \end{bmatrix} \right] \quad (9)$$

$$\begin{aligned} y &= (y_1, \dots, y_n)^t \in \mathfrak{R}^n \\ h(x) &= (h_1, \dots, h_k(x))^t \in \mathfrak{R}^k \\ \text{with } H &= (h(x_i)^t)_{1 \leq i \leq n} \in \mathfrak{R}^{n \times k} \\ R &= (R(x_i, x_j))_{1 \leq i, j \leq n} \in \mathfrak{R}^{n \times n} \\ r(x) &= (R(x, x_1), \dots, R(x, x_n))^t \in \mathfrak{R}^n \end{aligned}$$

The interested reader can refer to Sacks *et al.* (1989) for more details. Covariance parameters $\{\sigma, (p_i)_{i=1, \dots, d}, (\lambda_i)_{i=1, \dots, d}\}$ are called hyper-parameters. In geostatistics, parameters $(\lambda_i)_{i=1, \dots, d}$ are also known as correlation lengths.

Functions h_j are selected on the basis of *a priori* knowledge. If such information is not available, then a constant value is taken. Coefficients β as well as hyper-parameters $\lambda = (\lambda_i)_{i=1, \dots, d}$, $p = (p_i)_{i=1, \dots, d}$ and σ are estimated by maximizing the logarithm of the following likelihood function:

$$\begin{aligned} \ell(\beta, \sigma, \lambda, p) &= -\frac{1}{2} \left[n \ln(2\pi) + n \ln \sigma^2 + \ln [\det(R)] \right. \\ &\quad \left. + \frac{1}{\sigma^2} (y - H\beta)^t R^{-1} (y - H\beta) \right] \quad (10) \end{aligned}$$

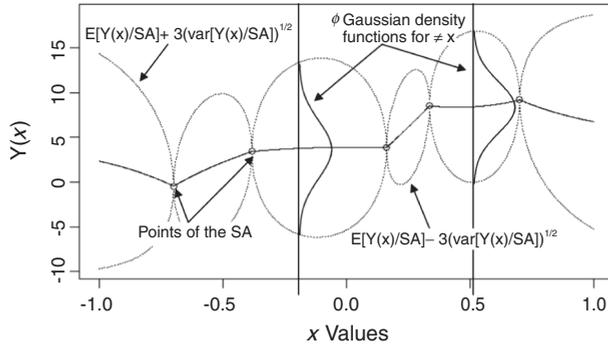


Figure 1
Example of a GP $Y(x)$.

Thus, estimate of:

- $\hat{\beta}$ is $\hat{\beta} = (H'R^{-1}H)^{-1}H'R^{-1}y$;
- $\hat{\sigma}$ is $\hat{\sigma}^2 = \frac{1}{n}(y - H\hat{\beta})'R^{-1}(y - H\hat{\beta})$;
- $(\lambda_i)_{i=1, \dots, d}$ and $(p_i)_{i=1, \dots, d}$ are obtained from a numerical optimization of $\ell(\hat{\sigma}, \hat{\beta}, \lambda, p)$.

Instead of using only this optimal value of the hyperparameters, it is possible to consider their *posterior* probability density functions that account for the previously defined likelihood $\ell(\hat{\sigma}, \hat{\beta}, \lambda, p)$ (Kennedy and O'Hagan, 2001). In such a case, the main drawback is the computational cost. For this reason, we preferred to disregard the Bayesian kriging technique.

An example of a GP $Y(x)$ is displayed in Figure 1 for a one dimensional case with no nugget effect. The circles are elements of the *SA* realization sample. A Gaussian density function ϕ is defined for each x with mean $E[Y(x)/SA]$ and standard deviation $(\text{var}[Y(x)/SA])^{1/2}$. A continuous line shows the mean of $Y(x)$ whereas dotted lines provide the 99% confidence interval (mean $\pm 3\text{std}$). The unknown realizations of $Y(x)$ should fall in this confidence interval associated with a probability of 99%. We observe that the realizations of the *SA* sample are accurately reproduced by $Y(x)$. Effectively, for these points, $x_i \in SA$: $E[Y(x_i)/SA] = y_i$ and due to the absence of nugget effect $\text{var}[Y(x_i)/SA] = 0$. Thus, the probability density function obtained for these points is a Dirac: $\delta(\{y(x_i)\})$.

3 GAUSSIAN PROCESS USED WITHIN PROBABILISTIC INVERSE PROBLEM FRAMEWORK

As stated above within the ‘‘Probabilistic inverse problem framework’’ section, simulation results

$y(x) = \{y^i(x)\}$ are considered as a stochastic function, denoted $Y(x) = \{Y^i(x)\}$. In what follows, we consider that each Y^i is a GP with no nugget effect. Thus when, the simulation is not run for x . Results $y^i(x)$ are unknown and Y^i GP yields an estimation of the probability density function of $p_{Y^i(x)}(y^i)$, as:

$$p_{Y^i(x)}(y^i) = \phi\left(\text{mean} = E[Y^i(x)], \text{std} = (\text{var}[Y^i(x)])^{1/2}\right) = \phi(\mu_{Y^i(x)}, \sigma_{Y^i(x)})$$

where ϕ is a Gaussian probability distribution function.

The likelihood probability density function for $y(x)$ results, assuming targets and simulation results to be independent, can be written as:

$$p_{Y/X}(t/x) = \int_{\Gamma} p_T(t') \cdot p_{Y(x)}(t') dt' = \prod_i \left[\int_{-\infty}^{+\infty} p_{T^i}(t') \cdot \phi(\mu_{Y^i}(t'), \sigma_{Y^i}(t')) dt' \right] \quad (11)$$

Similarly, it is shown that the *posterior* distribution of X is:

$$p_{X/Y}(x/t) \propto \left(\prod_i \left[\int_{-\infty}^{+\infty} p_{T^i}(t') \cdot \phi(\mu_{Y^i}(t'), \sigma_{Y^i}(t')) dt' \right] \right) \times p_X(x) \quad (12)$$

Let us focus on the particular cases listed below.

- each T^i respects the Gaussian law $p_{T^i} = \phi(\text{mean} = \mu_{T^i}, \text{std} = \sigma_{T^i})$. In such conditions,

$$\int_{-\infty}^{+\infty} p_{T^i}(t') \cdot p_{Y^i(x)}(t') dt' = \frac{1}{\sqrt{2\pi} \cdot \sqrt{\sigma_{T^i}^2 + \sigma_{Y^i(x)}^2}} \cdot \exp\left(-\frac{1}{2} \frac{[\mu_{T^i} - \mu_{Y^i(x)}]^2}{\sigma_{T^i}^2 + \sigma_{Y^i(x)}^2}\right).$$

This actually corresponds to the convolution at zero of the two Gaussian probabilities, p_{T^i} and $p_{Y^i(x)}$;

- each T^i respects the Uniform law defined between $[a_{T^i}, b_{T^i}]$.

$$\text{Then, } \int_{-\infty}^{+\infty} p_{T^i}(t') \cdot p_{Y^i(x)}(t') dt' = \int_{a_{T^i}}^{b_{T^i}} \frac{p_{Y^i(x)}(t')}{b_{T^i} - a_{T^i}} dt' = \frac{\Phi_{Y^i(x)}(b_{T^i}) - \Phi_{Y^i(x)}(a_{T^i})}{b_{T^i} - a_{T^i}}$$

where $\Phi_{Y^i(x)}$ is the cumulative density function of the Gaussian law $p_{Y^i(x)}$.

It is worthwhile mentioning that the independence assumption between $\{Y^i(x)\}$ leads to simplified computations and good results as shown in the examples below. However, this assumption can be removed provided we compute the GP of Y^i accounting for the covariance of

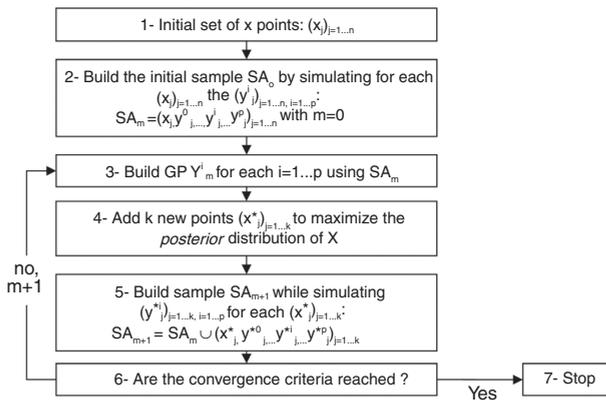


Figure 2

Successive steps of the MMTPUR technique.

several dependent $\{Y^i(x)\}$. This can be done using cokriging (Cressie, 1993) technique.

4 MAXIMIZATION OF MATCHING-TARGETS' PROBABILITY FOR UNEVALUATED RESULTS (MMTPUR) TECHNIQUE

The main feature of the MMTPUR technique consists in approximating the simulator or numerical function for unknown results associated with targets T with GP. These GP are used to compute the *posterior* distribution of X . Then, this *posterior* distribution is maximized, leading to updated x values. The successive steps of the MMTPUR technique are recapped in Figure 2 and explained hereafter:

4.1 Step 1 and 2: Initial Design

We first build the initial set of $(x_j)_{j=1, \dots, n}$ values with $(y_j^i)_{j=1, \dots, n, i=1, \dots, p}$. n stands for the number of points, p for the number of targets and m is the current iteration index. y_j^i is associated with the j -th simulation result related to the i -th target T^i . Even if it is theoretically possible to consider a single point within the initial sample, space-filling techniques like Latin hypercube sampling (McKay *et al.*, 1979; Fang *et al.*, 2006) are preferred to obtain the n point set. The objective is to end up with initial samples investigating the entire domain associated with the possible variations in parameters (x).

4.2 Step 3: Build GP Y_m^i for Each Target i Using SA_m

At this point, the GPs $Y_m^i = Y^i/SA_m$ are built using the current sample SA_m (see Eq. 8, 9).

4.3 Step 4: Add k New Points $(x^*)_{j=1, \dots, k}$ to Maximize the *Posterior* Distribution of X

As shown in Equation (12), the *posterior* distribution of X is:

$$p_{X/Y}(x/t) \quad (13)$$

$$\propto \left(\prod_i \left[\int_{-\infty}^{+\infty} p_{T^i}(t') \cdot \phi(\mu_{Y^i}(t'), \sigma_{Y^i}(t')) dt' \right] \right) \times p_X(x)$$

where: $\phi(\mu_{Y^i}, \sigma_{Y^i})$ is the Gaussian probability density function of the current Y_m^i GP (Fig. 1) with mean and standard deviation obtained from Equations (8, 9); and p_{T^i} is the probability density function for target T^i , targets being assumed to be independent.

To select the k new points $(x^*)_{i=1, \dots, k}$ to be added to the current SA_m sample, we apply the following iterative strategy:

1. select the first new point x^*_1 as the one that maximizes $p_{X/Y}(x/t)$. Several optimization techniques can be selected to solve this maximization problem, such as genetic algorithms, gradient-based optimizations, etc., or a combination of these techniques;
2. add fictitiously x^*_1 to sample SA_m and associate the mean value derived from Y_m^i : $E[Y_m^i(x^*_1)]$ to the fictitious y^i value for any i ;
3. go back to step 1 and repeat until k new points are selected.

Selecting the number k of new points is a very practical issue and depends on simulation capabilities, such as computing power (cluster, network) or the number of available simulator licenses. For example, if 10 simulation jobs can be simultaneously run, k can be set to 10. However, if the simulation jobs can be launched only sequentially, k will be set to 1.

4.4 Step 5: Build Sample SA_{m+1}

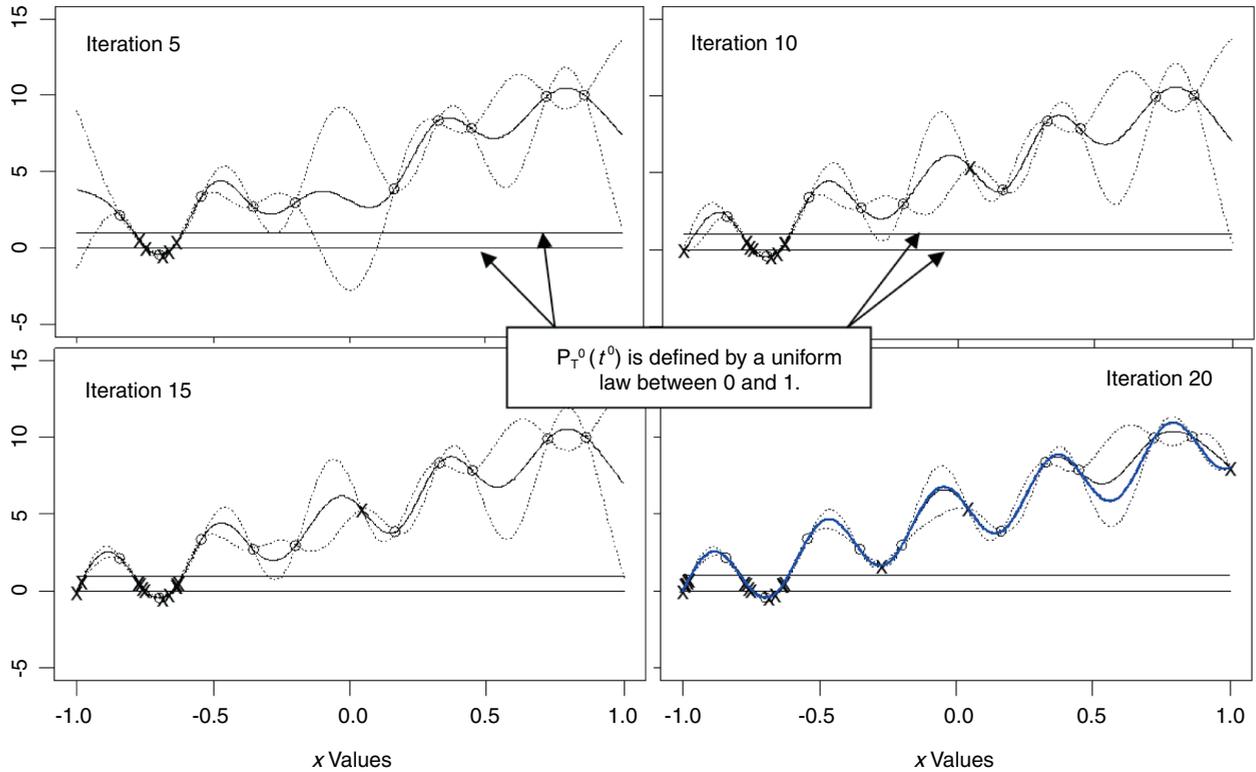
The k new $(x^*_j)_{j=1, \dots, k}$ points are simulated to obtain $(y^{*i}_j)_{j=1, \dots, k, i=1, \dots, p}$. Then, the sample is updated:

$$SA_{m+1} = SA_m \cup \left(x^*_j, y_j^{*0}, \dots, y_j^{*i}, \dots, y_j^{*p} \right)_{j=1, \dots, k}$$

4.5 Step 6 and 7: Convergence

Several convergence strategies can be envisaged based upon either:

- a maximum number of simulations affordable, or
- a threshold value for the maximum of the *posterior* distribution of X . For instance, if this maximum is lower than the threshold, the optimization is stopped.



Features for this one dimensional example

- 1 point is added at each iteration. The plots display iterations 5,10, 15 and 20
- The circled points are the 10 initial ones.
- The cross points are the additional points.

- The black curve is the mean of Y^0
- The dotted lines are the mean of $Y^0 \pm 3 \cdot \text{Standard Deviation}$
- The blue line on the final plot is the $y^0(x)$ analytical function

Figure 3

MMTPUR results for a one-dimensional case with one target.

It means that the expected probability to match the targets is very low when adding this point.

When one of these criteria is reached, the optimization process is stopped. Otherwise, index m is incremented and we go back to step 3.

5 APPLYING MMTPUR TO AN ANALYTICAL FUNCTION WITH ONE TARGET

The results, derived from the MMTPUR technique for a one-dimensional case, $x \in \mathbb{R}$, and one target, are shown in Figure 3. In this example, the simulator output $y^0(x)$ is a known analytical function. Associated target T^0 follows the probability density function $p_{T^0}(t^0)$.

More precisely, we consider that:

- y^0 is the following analytical function:

$$y^0(x) = 5(x + 1) + 2 \sin(15(x + 1))$$

- $p_{T^0}(t^0)$ is a uniform law between 0 and 1;
 - there is no *prior* probability density function for x .
- MMTPUR technique is used to determine x values that give $y^0(x)$ belonging to interval $[0, 1]$.

As explained above, an initial sample (steps 1 and 2), SA_0 , is needed. It corresponds to the 10-circled points in Figure 3. These points being known, we move to the following steps:

- a GP Y^0 is built to approximate y^0 (step 3);
- Y^0 is then used to find a new x value in order to get new $y^0(x)$ in agreement with target T^0 and associated $p_{T^0}(t^0)$ (step 4);
- the SA sample is updated (step 5) and the above steps are repeated.

The results obtained for iterations 5, 10, 15 and 20 (resp., 15, 20, 25 and 30 points) are displayed in Figure 3. Crosses indicate the added points. The evolution of the Y^0 GP is also shown: its mean is represented by a black

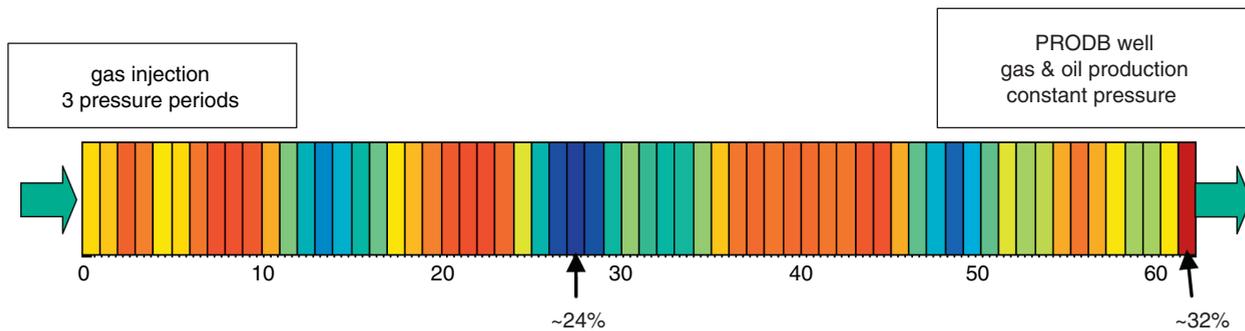


Figure 4

Description of the numerical fluid-flow model. The color represents the porosity property.

line and its mean plus or minus $3 \times$ the standard deviation, by dotted lines.

The new x points (crosses) lead to $y^0(x)$ values that mainly fall within the $[0, 1]$ interval (Fig. 3). It means that the $y^0(x)$ values respect the T^0 distribution, as required. In addition, Y^0 turns out to be accurate in areas where tolerance can be respected, thus approximating very well the analytical function y^0 at these locations.

6 APPLYING MMTPUR TO A FLUID-FLOW MODEL

In this core flood example, we apply the MMTPUR technique to a fluid-flow model depending on 4 parameters. The selected targets are synthetic cumulative oil and gas production profiles. To check the soundness of the MMTPUR technique, we compared the simulation results corresponding to a parameters' *posterior* sampling of 1 000 points with the specified targets and related tolerances.

6.1 Test Case Description

The core flood numerical model was built with the commercialized software PumaFlowTM. This one-dimensional model encompasses 62 cells. The rock sample was initially saturated with oil and irreducible water and a “pseudo-well” named PRODB was located at the outlet boundary to produce oil and gas. Porosity (Fig. 4) was measured in the laboratory: it ranges between $\sim 24\%$ to $\sim 32\%$. Permeability and capillary pressure were then derived from porosity based upon analytical models. We simulated three successive periods with different pressure drops between the outlets (Fig. 5).

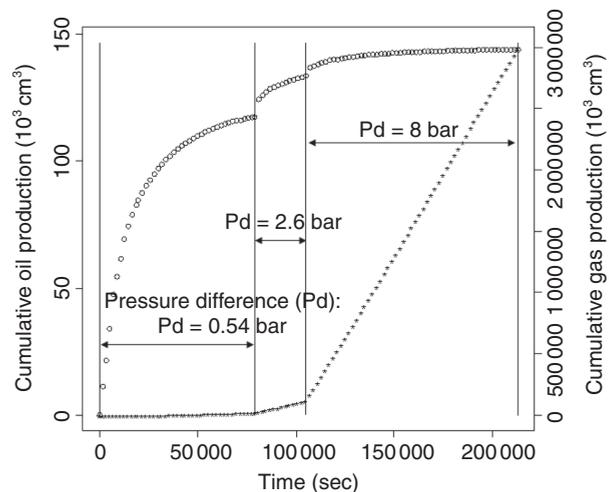


Figure 5

Synthetic cumulative oil (symbol “o”) and gas (symbol “*”) production data.

The synthetic measurements displayed in Figure 5 were simulated for a randomly-drawn forward porosity model. They consist of the cumulative oil and gas productions at the outlet.

6.2 Parameters and Targets Definition

The synthetic measurements described above are now considered as reference data. At this point, we assume that the 4 parameters listed below are unknown:

- the 2 Corey exponents associated with the gas-oil relative permeability curves: “gas Corey exp.” and “oil Corey exp.”;

TABLE 1
Parameters' definition

Name	Minimum	Maximum	Reference value
Gas Corey exp.	1.5	3.0	2.20
Oil Corey exp.	1.5	4.5	2.13
K mult.	0.1	2.0	1.96
Gas CP mult.	0.5	2.0	1.28

- a global permeability multiplier: "K mult." and
- a global gas capillary pressure multiplier: "gas CP mult."

The minimum, maximum and reference values for each of these parameters are reported in Table 1. The reference synthetic production data were generated using the reference values given in the table.

We selected 3 targets among the available synthetic data:

- the two cumulative oil and gas productions corresponding to the last time and
- the cumulative oil production at 10 hours (36 000 seconds).

The MMTPUR technique calls for the definition of tolerances for each selected targets. They were described by uniform probability laws over fixed ranges with $\pm 2\%$ variations (Fig. 6, 7). For simplicity, we disregarded prior probability $p_X(x)$.

6.3 MMTPUR Results

The initial design encompasses 5 simulations randomly selected on the basis of Latin hypercube sampling. The MMTPUR procedure was launched and stopped after the sequential addition of 6 other simulations.

Figure 6 and Figure 7 show that none of the initial simulations (black curves) provide responses respecting all targets. This is especially true for gas production, which is associated with a target much larger than the initial simulation results. The MMTPUR technique makes it possible to identify simulations (green curves) that approximately respect the required targets.

As described previously for the 4th step of the MMTPUR technique, the *posterior* parameters' distribution is built at each iteration (Eq. 13). In what follows the distribution obtained at the last iteration, are jointly used with a Markov-Chain Monte-Carlo algorithm (Geyer, 1992) to build a 1 000 points parameters' *posterior* sampling. The marginal histograms are displayed in Figure 8. These histograms are all in agreement with

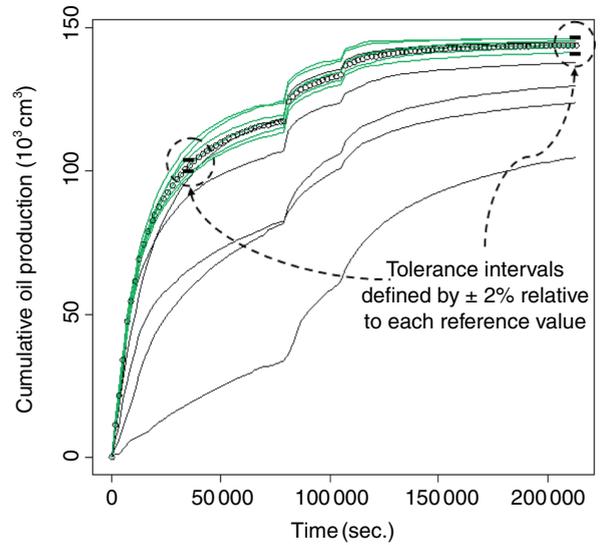


Figure 6

MMTPUR simulation results displayed with synthetic cumulative oil production data (symbol "o"). Black curves are associated with the 5 initial simulation results. Green curves correspond to the 6 simulation results iteratively added during MMTPUR procedure.

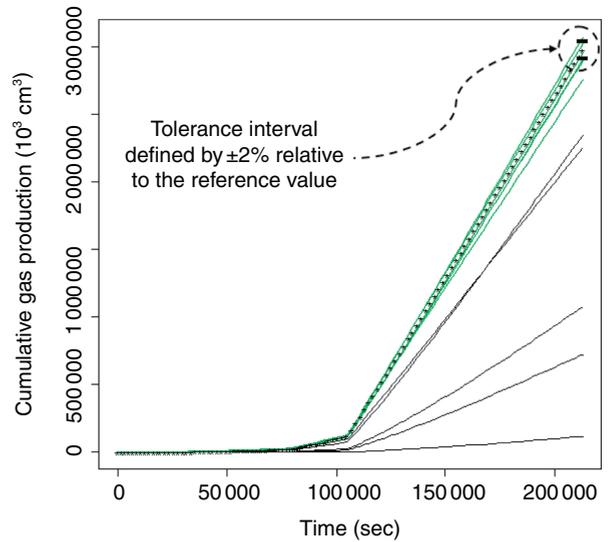


Figure 7

MMTPUR simulation results displayed with synthetic cumulative gas production data (symbol "***"). Black curves are associated with the 5 initial simulation results. Green curves correspond to the 6 simulation results iteratively added during MMTPUR procedure.

the reference values. Moreover, the wide spread of the gas capillary pressure multiplier indicates that the influence of this parameter for matching the identified targets is minor compared to the others.

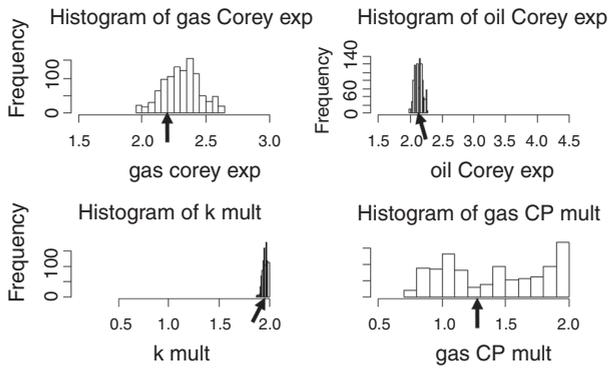


Figure 8

Marginal histogram parameters computed using the 1 000 points of the *posterior* sampling. Each plot is between parameter ranges. Arrows indicate the reference value for each parameter.

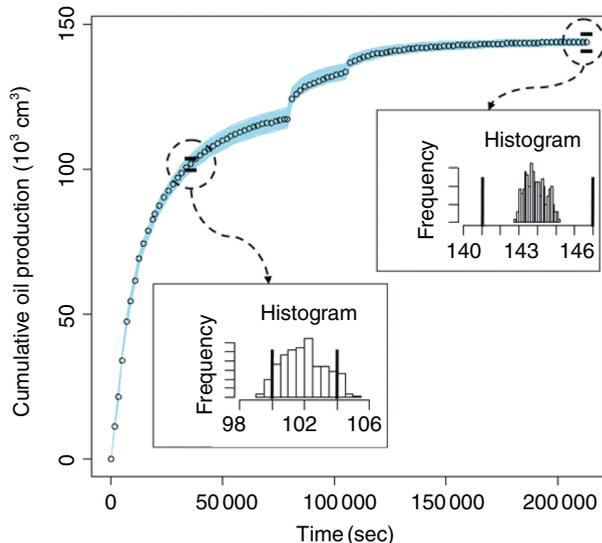


Figure 9

Cumulative oil production simulation results (light blue curves) associated with the *posterior* parameters sampling. Synthetic cumulative oil production data and tolerance intervals are also displayed.

For validation purposes, we ran the simulations for the 1 000 points of the parameters *posterior* sampling. The results are displayed in Figure 9 for oil production and in Figure 10 for gas production. We also plotted the histograms of the 1 000 simulation results for each target. The black lines indicate the chosen tolerance intervals.

The tolerance intervals for the 3 targets are almost fully respected by the 1 000 simulations. We observe that

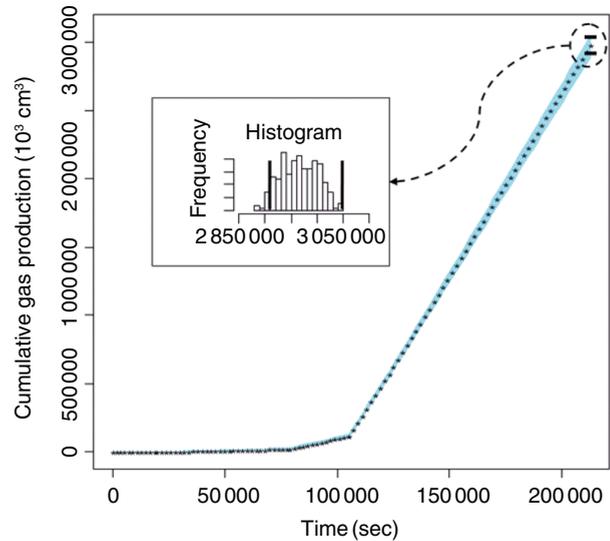


Figure 10

Cumulative gas production simulation results (light blue curves) associated with the *posterior* parameters sampling. Synthetic cumulative gas production data and tolerance interval are also displayed.

the range of the histogram obtained for the last time of oil production is narrow compared to the associated tolerance interval (Fig. 9). This can be explained by correlations between simulation results.

This 4-dimensional example highlights that the MMTPUR technique is able to define all possible simulations consistent with the tolerance intervals chosen for the 3 targets from only 11 simulations. This means that the *posterior* distributions built for parameters are good and lead to satisfactory results.

7 MAXIMIZATION OF MATCHING-TARGETS' PROBABILITY FOR UNEVALUATED RESULTS TECHNIQUE IN AN UNCERTAIN CONTEXT

The MMTPUR technique can be extended to optimization in an uncertain context. In this case, the uncertain domain is determined by two different types of parameters, controllable and uncontrollable, grouped in $x = (x_c, x_u)$. The controllable parameters, x_c , can be adjusted to determine optimal values. On the contrary, the uncontrollable parameters, x_u , are intrinsically uncertain. Thus, the goal is to find the optimal values of x_c which make it possible to fit specific targets while accounting for the uncertainty related to x_u .

As an example, a practical problem in reservoir engineering consists in identifying the location of a well with

known liquid and/or gas rates while satisfying particular targets on oil and water productions, given uncertainties about uncontrollable parameters, such as correlation lengths or the average porosity of a facies.

We associate:

- *a priori* uncertainty on the controllable parameters, through: $p_{X_c}(x_c)$;
- uncertainty on the uncontrollable parameters, through: $p_{X_u}(x_u)$.

The uncertainty associated with the uncontrollable parameters can be accounted for while calculating likelihood $p_{Y/X_c}(t/x_c)$ instead of $p_{Y/X}(t/x)$. Thus:

$$p_{Y/X}(t/x) = \int_{\Gamma} p_T(t') \cdot p_{Y(x)}(t') dt' \quad (14)$$

has to be substituted by:

$$p_{Y/X_c}(t/x_c) = \left(\int_{X_u} \left[\int_{\Gamma} p_T(t') \cdot p_{Y(x_c, x_u)}(t') dt' \right] \times p_{X_u}(x_u) dx_u \right) \quad (15)$$

where the integration domain X_u is related to the x_u space.

As a result, the *posterior* uncertainty on X_c is derived from the following equation:

$$p_{X_c/Y}(x_c/t) = \left(\int_{X_u} \left[\int_{\Gamma} p_T(t') \cdot p_{Y(x_c, x_u)}(t') dt' \right] \times p_{X_u}(x_u) dx_u \right) \times p_{X_c}(x_c) \quad (16)$$

It is then possible to add new points x_c while maximizing the *posterior* uncertainty relative to the controllable parameters X_c to get simulation results in agreement with target definition, which accounts for the uncertainty associated with X_u .

CONCLUSION

In this paper we introduce, a new optimization technique, the Maximization of Matching-Targets' Probability for Unevaluated Results, particularly adapted when getting one evaluation of the numerical function is very time consuming. MMTPUR is based upon an add-on to the classical probabilistic optimization framework considering the numerical function values that have not been evaluated as stochastic functions. The associated uncertainty model is a Gaussian process for each required numerical function result (*i.e.*, associated with

each target). It makes it possible to estimate probability density functions for each required unevaluated results. Therefore, the *posterior* distribution of the parameters as given by the MMTPUR approach, allows us to account for the error estimations related to the unevaluated function values or simulation results. As described, it is then further used to iteratively find and simulate new points that maximize this *posterior* distribution.

MMTPUR technique was also applied and presented on different simple cases with one or several targets producing good results. Moreover, an extension of the MMTPUR technique was described for a specific uncertain context. We believe that this technique could be extended to solving other problems, such as history-matching in reservoir engineering, or optimization with targets to reach.

ACKNOWLEDGMENTS

I would like to thank Mickaële Le Ravalec, Amandine Marrel and Sébastien Da Veiga for their fruitful comments and discussions.

REFERENCES

- Busby D., Feraille M. (2008) Adaptive Design of Experiments for calibration of complex simulators – An Application to Uncertainty Quantification of a Mature Oil Field, *J. Phys.: Conf. Ser.* **135**, 012026, doi: 10.1088/1742-6596/135/1/012026.
- Busby D. (2009) Hierarchical adaptive experimental design for Gaussian process emulators, *Reliab. Eng. Syst. Safe.* **94**, 1183-1193. doi: 10.1016/j.res.2008.07.007.
- Cressie N.A.C. (1993) *Statistics for Spatial Data*, Wiley, New York.
- Fang K.T., Li R., Sudjianto A. (2006) *Design and modeling for computer experiments*, Chapman and Hall/CRC.
- Feraille M., Marrel A. (2012) Prediction under uncertainty on a mature field, *Oil Gas Sci. Technol.* **67**, 2, 193-206. doi: 10.2516/ogst/2011172.
- Geyer C.J. (1992) Practical Markov chain Monte Carlo (with discussion), *Stat. Sci.* **7**, 473-511. doi:10.1214/ss/1177011137.
- Hansen N. (2006) The CMA Evolution Strategy: A Comparing Review, in *Towards a new evolutionary computation, Advances in estimation of distribution algorithms*, Lozano J.A., Larranga P., Inza I., Bengoetxea E. (eds), Springer.
- Jourdan A. (2000) Analyse statistique et échantillonnage d'expériences simulées, *Dissertation*, Université de Pau et des pays de l'Adour, France.
- Kennedy M.C., O'Hagan A. (2001) Bayesian calibration of computer models, *J. R. Stat. Soc., Ser. B Stat. Methodol.* **63**, 425-464. doi: 10.1111/1467-9868.00294.
- McKay M.D., Beckman R.J., Conover W.J. (1979) A comparison of three methods for selecting values of input variables in

the analysis of output from a computer code, *Technometrics* **21**, 239-245. doi: 10.2307/1268522.

Marrel A. (2008) Mise en œuvre et utilisation du métamodèle processus Gaussien pour l'analyse de sensibilité de modèles numériques, *Dissertation*, Institut National des Sciences Appliquées de Toulouse, France.

Matheron G. (1963) Principles of geostatistics, *Econ. Geol.* **58**, 1246-1266. doi: 10.2113/gsecongeo.58.8.1246.

Sacks J., Welch W.J., Mitchell T.J., Wynn H.P. (1989) Design and analysis of computer experiments, *Stat. Sci.* **4**, 409-423. doi:10.1214/ss/1177012413.

Scheidt C., Zabalza-Mezghani I., Feraille M., Collombier D. (2007) Toward a Reliable Quantification of Uncertainty on Production Forecasts: Adaptive Experimental Designs, *Oil Gas Sci., Technol.* **62**, 207-224. doi: 10.2516/ogst:2007018.

Schonlau M. (1997) *Computer Experiments and Global Optimization*, University of Waterloo, Canada, Dissertation.

Tarantola A. (2005) *Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM (Society for Industrial and Applied Mathematics), ISBN 978-0-89871-572-9.

Vazquez E., Bect J. (2010) Convergence properties of the expected improvement algorithm with fixed mean and covariance functions, *J. Stat. Plan. Inference.* **140**, 3088-3095. doi: 10.1016/j.jspi.2010.04.018.

Villemonteix J., Vazquez E., Sidorkiewicz M., Walter E. (2009) Global optimization of expensive-to-evaluate functions: an empirical comparison of two sampling criteria, *J. Glob. Optim.* **44**, 509-534. doi: 10.1007/s10898-008-9313-y.

Manuscript accepted in October 2012

Published online in June 2013

Copyright © 2013 IFP Energies nouvelles

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than IFP Energies nouvelles must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or a fee: Request permission from Information Mission, IFP Energies nouvelles, fax. +33 1 47 52 70 96, or revueogst@ifpen.fr.