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

Mathieu Feraille

To cite this version:


HAL Id: hal-00864214
https://hal-ifp.archives-ouvertes.fr/hal-00864214
Submitted on 20 Sep 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
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

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 $\mu_X$ and covariance operator $C_X$, the a priori probability density function is expressed as:

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

(1)

- 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 $p_{Y|X}(y|x)$, 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:

$$p_{X|Y}(x|t) \propto p_{Y|X}(t|x)p_X(x)$$

(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 $lk(x)$. It can be derived from the independent probability density functions of $T$ and $Y(x)$ defined above:

$$lk(x) = p_{Y|X}(t|x) = \int_T p_T(t') \cdot p_{Y|X}(t|x) dt'$$

(3)

The $\Gamma$ 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:

$$lk(x) = p_{Y|X}(t|x) = p_T(y(x))$$

(4)

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

$$lk(x) \propto \exp \left\{ -\frac{1}{2} (y(x) - \mu_T)^T C_T^{-1} (y(x) - \mu_T) \right\}$$

$$= \exp (-OF_{\text{lk}}(x))$$

(5)

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

$$p_{X|Y}(x|t) \propto lk(x) \cdot p_X(x) \propto \exp (-OF_{\text{lk}}(x) - OF_{\text{prior}}(x)) = \exp (-OF(x))$$

(6)

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 \( \{x = (x_1, ..., x_d)\} \). 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 \( x = (x_1, ..., x_d) \), \( d \) 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:

\[
Y(x) = h(x) - Z(x) - U(x)
\]

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 \( \beta \) 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

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

where \( d \) is the dimension of \( x \) (number of parameters) and \( \{\sigma, (\lambda_{ij})_{i,j=1,...,d}, (\beta_j)_{j=0,...,k}\} \) 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 non-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 \( \sigma^2 = \sigma^2 \cdot \tau; \tau \geq 0 \). Coefficient \( \tau \) represents the contribution of the nugget effect. When \( U(x) \) is used, the covariance of the stochastic part of model \( Y \) becomes:

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

with \( \delta \) the Dirac’s delta function.

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

For a given sample of realizations, denoted \( SA = \{Y(x_1) = y_1, ..., y_n\} \) with \( n \) the number of points, it can be shown that the conditional mean and variance of Gaussian process \( Y(x) \) given \( SA \) are:

\[
\begin{align*}
E[Y(x)/SA] & = h'(x)\beta + r'(x)R^{-1}(y - H\beta) \\
\text{var}[Y(x)/SA] & = \sigma^2 \left[ 1 - [r'(x), h'(x)] \times \begin{bmatrix} R & H \\ H^t & 0 \end{bmatrix}^{-1} r(x) \right] h(x)
\end{align*}
\]

\[
y = (y_1, \ldots, y_n)^t \in \mathbb{R}^n
\]

\[
h(x) = (h_1(x), \ldots, h_k(x))^t \in \mathbb{R}^k
\]

\[
H = (h(x))^t \in \mathbb{R}^{n \times k}
\]

\[
R = (R(x_i, x_j))^t \in \mathbb{R}^{n \times n}
\]

\[
r(x) = (R(x, x_1), \ldots, R(x, x_n))^t \in \mathbb{R}^n
\]

The interested reader can refer to Sacks et al. (1989) for more details. Covariance parameters \( \{\sigma, (\lambda_{ij})_{i,j=1,...,d}, (\beta_j)_{j=0,...,k}\} \) are called hyper-parameters. In geostatistics, parameters \( (\lambda_{ij})_{i,j=1,...,d} \) 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 \( \beta \) as well as hyper-parameters \( \lambda = (\lambda_{ij})_{i,j=1,...,d}, p = (p_j)_{j=1,...,d} \) and \( \sigma \) are estimated by maximizing the logarithm of the following likelihood function:

\[
\ell(\beta, \sigma, \lambda, p) = -\frac{1}{2} \left[ \ln(2\pi) + \ln n \sigma^2 + \ln |\det(R)| \right] + \frac{1}{2\sigma^2} [y - H\beta]^t R^{-1} [y - H\beta]
\]

(10)
Thus, estimate of:
- $\beta$ is $\hat{\beta} = (H^\dagger R^{-1}H)^{-1}H^\dagger R^{-1}y$;
- $\sigma$ is $\hat{\sigma}^2 = 1/2(\mathbf{y} - H\hat{\beta})^\dagger R^{-1}(\mathbf{y} - H\hat{\beta})$;
- $(\lambda)_i = 1, \ldots, d$ and $(p)_i = 1, \ldots, d$ 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 $\phi$ 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 $\pm$ 3std). 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, $x_i \in \text{SA}; E[Y(x_i)/SA] = y_i$ 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: $\delta(\{y_i\})$.

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 $Y$ is a GP with no nugget effect. Thus when, the simulation is not run for $x$. Results $y_i(x)$ are unknown and $Y$ GP yields an estimation of the probability density function of $p_{Y(x)}(y)$, as:

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

where $\phi$ 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:

$$p_{Y|x}(y|x) = \prod_{t} \left[ \int_{-\infty}^{+\infty} p_{T}(t) \cdot p_{Y|x}(y|x)dt \right]$$

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

$$p_{X|Y}(x|t) \propto \left( \int_{-\infty}^{+\infty} p_{T}(t) \cdot \phi(\mu_{Y(X)}(t), \sigma_{Y}(t))dt \right) \cdot p_{X}(x)$$

Let us focus on the particular cases listed below.
- each $T$ respects the Gaussian law $p_T = \phi(\text{mean} = \mu_T, \text{std} = \sigma_T)$. In such conditions,

$$\int_{-\infty}^{+\infty} p_{T}(t) \cdot p_{Y|x}(y|x)dt = \frac{1}{\sqrt{2\pi} \sigma_T} \exp \left( -\frac{(y - \mu_T)^2}{2\sigma_T^2} \right).$$

This actually corresponds to the convolution at zero of the two Gaussian probabilities, $p_T$ and $p_{Y|x}$;
- each $T$ respects the Uniform law defined between $[a_T, b_T]$.

Then, $\int_{-\infty}^{+\infty} p_{T}(t) \cdot p_{Y|x}(y|x)dt = \int_{a_T}^{b_T} \frac{y - \mu_T}{b_T-a_T} \frac{dy}{\sigma_T^2} dt' = \Phi_{\mu_{Y|x}(t'), \sigma_{Y|x}(t')}(y) - \Phi_{\mu_{Y|x}(t'), \sigma_{Y|x}(t')}(y_{\text{lower}})$

where $\Phi_{Y|x}(t')$ is the cumulative density function of the Gaussian law $p_{Y|x}$.

It is worthwhile mentioning that the independence assumption between $\{Y(x)\}$ 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 $Y'$ accounting for the covariance of...
several dependent \{ Y(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 \((x_{0})_{1}, \ldots, n\) values with \((y_{i}^{0})_{1}, \ldots, n, i = 1, \ldots, p\). \( n \) 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 \( Y^{i}_{m} \) for Each Target \( i \) Using \( SA_{m} \)

At this point, the GPs \( Y^{i}_{m} = Y^{i}/SA_{m} \) are built using the current sample \( SA_{m} \) (see Eq. 8, 9).

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

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

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

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

To select the \( k \) new points \((x^{*}_{j})_{j=1, \ldots, k}\) 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 \( Y^{i}_{m}: E[Y^{i}_{m}(x^{*}_{1})] \) 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 \((x^{*}_{j})_{j=1, \ldots, k}\) points are simulated to obtain \((y^{*}_{j})_{j=1, \ldots, k}, i=1, \ldots, p\). Then, the sample is updated:

\[
SA_{m+1} = SA_{m} \cup (x^{*}_{j}^{0}, y^{*}_{j}^{0}, \ldots, y^{*}_{j}^{p})_{j=1, \ldots, k}
\]

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.
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 \( p_r^0(\theta^0) \).

More precisely, we consider that:
- \( y^0 \) is the following analytical function:
  \[
y^0(x) = 5(x + 1) + 2 \sin(15(x + 1))
\]
- \( p_r^0(\theta^0) \) 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 \( p_r^0(\theta^0) \) (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 curve and the dotted lines are the mean of \( Y^0 \pm 3 \times \text{Standard Deviation} \).
line and its mean plus or minus \(3 \times\) the standard deviation, by dotted lines.

The new \(y_0(x)\) 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 PumaFlow\textsuperscript{TM}. 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 \(~24\%\) to \(~32\%\). 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).

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.”;
- 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 ±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 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.

<table>
<thead>
<tr>
<th>Name</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Reference value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gas Corey exp.</td>
<td>1.5</td>
<td>3.0</td>
<td>2.20</td>
</tr>
<tr>
<td>Oil Corey exp.</td>
<td>1.5</td>
<td>4.5</td>
<td>2.13</td>
</tr>
<tr>
<td>K mult.</td>
<td>0.1</td>
<td>2.0</td>
<td>1.96</td>
</tr>
<tr>
<td>Gas CP mult.</td>
<td>0.5</td>
<td>2.0</td>
<td>1.28</td>
</tr>
</tbody>
</table>
For validation purposes, we ran the simulations for the 1,000 points of the 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 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: $p_{X_i}(x_i)$;
- 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 $p_{Y/X_i}(t/x_c)$ instead of $p_{Y/X}(t/x)$. Thus:

$$p_{Y/X}(t/x) = \int_T p_T(t') \cdot p_{Y(t')}(t')dt'$$

has to be substituted by:

$$p_{Y/X_i}(t/x_c)$$

$$= \left(\int_{X_u} \left[ \int_T p_T(t') \cdot p_{Y(x_u,i)}(t')dt' \right] \times p_{X_u}(x_u)dx_u\right)$$

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

As a result, the posterior uncertainty on $X_i$ is derived from the following equation:

$$p_{X_i/Y}(x_i/t)$$

$$= \left(\int_{X_u} \left[ \int_T p_T(t') \cdot p_{Y(x_u,i)}(t')dt' \right] \times p_{X_u}(x_u)dx_u\right) \times p_{X_i}(x_i)$$

It is then possible to add new points $x_i$ while maximizing the posterior uncertainty relative to the controllable parameters $X_i$ 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 Mickaêle Le Ravalec, Amandine Marrel and Sébastien Da Veiga for their fruitful comments and discussions.

**REFERENCES**


McKay M.D., Beckman R.J., Conover W.J. (1979) A comparison of three methods for selecting values of input variables in


Manuscript accepted in October 2012
Published online in June 2013