Comput Geosci (2006) 10:389–404
DOI 10.1007/s10596-006-9031-9

Optimization of nonconventional wells under uncertainty using statistical proxies

Vincent Artus · Louis J. Durlofsky · Jérôme Onwunalu · Khalid Aziz

Received: 27 July 2005 / Accepted: 25 August 2006 / Published online: 17 November 2006

Abstract The determination of the optimal type and placement of a nonconventional well in a heterogeneous reservoir represents a challenging optimization problem. This determination is significantly more complicated if uncertainty in the reservoir geology is included in the optimization. In this study, a genetic algorithm is applied to optimize the deployment of nonconventional wells. Geological uncertainty is accounted for by optimizing over multiple reservoir models (realizations) subject to a prescribed risk attitude. To reduce the excessive computational requirements of the base method, a new statistical proxy (which provides fast estimates of the objective function) based on cluster analysis is introduced into the optimization process. This proxy provides an estimate of the cumulative distribution function (CDF) of the scenario performance, which enables the quantification of proxy uncertainty. Knowledge of the proxy-based performance estimate in conjunction with the proxy CDF enables the systematic selection of the most appropriate scenarios for full simulation. Application of the overall method for the optimization of monobore and dual-lateral well placement demonstrates the performance of the hybrid optimization procedure. Specifically, it is shown that by simulating only 10% or 20% of the scenarios (as determined by application of the proxy), optimization results very close to those achieved by simulating all cases are obtained.

Keywords advanced wells · cluster analysis · reservoir simulation · risk analysis · stochastic optimization · uncertainty management · well placement

1 Introduction

Nonconventional wells, which include horizontal, highly deviated, and multilateral wells, offer great possibilities for field development. By contacting much larger portions of the reservoir than traditional (vertical) wells, they provide a means to dramatically enhance recovery. As drilling and completion technologies improve, the proportion of nonconventional wells being drilled continues to rise [9].

The optimization of field development scenarios involving nonconventional wells, however, is particularly challenging. A key issue is the very large number of variables to consider when designing such wells. Because there are a vast number of possible well configurations (e.g., planar dual-lateral, stacked trilateral, herringbone quadrilateral) and placements, it is impossible to exhaustively evaluate (via reservoir simulation) all of the potential scenarios. Another major issue is the uncertainty in the reservoir model (referred to here as geological uncertainty), which leads to uncertainty in the performance prediction for each well configuration. The quantification of this uncertainty, which, in practical cases, is accomplished through use of some type of Monte Carlo analysis, also requires a large number of simulations. In practice, the overall uncertainty in performance does not solely arise from the reservoir model, and other stochastic parameters (such as economic

V. Artus · L. J. Durlofsky · J. Onwunalu · K. Aziz
Department of Energy Resources Engineering,
Stanford University, Stanford, CA 94305-2220, USA

V. Artus
Division Ingénierie des Réservoirs, Institut Français du Pétrole,
92500 Rueil Malmaison, France

V. Artus (✉)
KAPPA, 100 rue Albert Caquot,
06410 Sophia Antipolis, France
e-mail: artus@kappaeng.com

constraints) must be included, ideally within a decision analysis framework [5].

Our objective was to introduce and apply a new and general type of statistical proxy that has the potential to greatly accelerate optimizations of nonconventional well type and placement under geological uncertainty. Unlike previous proxies applied in this setting, this proxy provides an estimate in terms of a cumulative distribution function (CDF), which enables the quantification of proxy uncertainty. This, in turn, leads to an optimization algorithm that focuses computational effort on those scenarios to which the optimum is most sensitive. Although this treatment of uncertainty could be used with any evaluation-only optimization procedure, we incorporate it here within the context of a genetic algorithm (GA). Genetic algorithms, described in more detail below, mimic Darwinian natural selection and entail the use of a *population* of *individuals* (each individual represents a potential solution to the optimization problem) that evolve over the course of the optimization. Selection and combination rules lead to an improvement in *fitness* (value of the objective function) as the optimization proceeds.

We describe some recent studies addressing well optimization, and discuss previous uses of proxies to accelerate these calculations. Bittencourt and Horne [3] and Güyagüler *et al*. [7] used a GA to optimize the placement of multiple vertical wells for deterministic reservoir models. Güyagüler and Horne [6] investigated the problem of well placement under geological uncertainty (using a GA) by optimizing over multiple geological realizations of the reservoir model. In this case, the objective function was a utility function that depended on the performance of the particular well placement scenario over the different realizations. The utility function is a standard tool in decision analysis and accounts, in this setting, for both the variability in the performance of the scenario over the different realizations as well as for the risk attitude of the decision maker.

Yeten *et al*. [20] developed a GA-based optimization for nonconventional well deployment, both for deterministic and uncertain reservoir geologies. Their approach provided the ability to optimize the number, type, and trajectory of multilateral wells. Recently, Klie *et al*. [10] and Bangerth *et al*. [2] used simultaneous perturbation stochastic approximation (SPSA) and very fast simulated annealing (VFSA) techniques for optimizing the deployment of vertical wells in a deterministic reservoir model. They compared these algorithms with the use of a GA, showing that their approaches were less intensive computationally. However, no comparison was made in the case of optimizing nonconventional wells in the presence of geological uncertainty, for which the number of parameters and the computational cost are expected to increase dramatically for every method.

As indicated above, a considerable reduction in the computational requirements must be achieved in order to apply optimization algorithms to practical problems involving nonconventional wells. Bittencourt and Horne [3] and Güyagüler *et al*. [7] hybridized the basic GA with other search techniques in order to achieve a faster convergence. Specifically, Bittencourt and Horne [3] introduced a polytope search while Güyagüler *et al*. [7] used response surface modeling techniques. The response surface was built from a calibration pool of individuals and their associated fitness by using kriging techniques. The pool and the response surface were updated after each generation. This was shown to provide much faster convergence than the basic stand-alone GA for vertical well optimizations. When dealing with nonconventional wells, however, the dimension of the parameter space can be very large, and the response surface approach may be less appropriate in this case.

Yeten *et al*. [20] applied artificial neural networks to estimate the value of the objective function from simple well and reservoir parameters (e.g., total well length, average contacted permeability). The neural network was trained from a calibration pool of previously simulated scenarios and updated after each generation. They used this proxy as a way to identify the best individuals in the current population and only performed simulations on these individuals. This technique was shown to significantly reduce the number of simulations required for each generation. These authors also used a hill climbing procedure to achieve faster convergence and upscaling techniques to accelerate the simulations. When dealing with nonconventional wells, however, the number of parameters to take into account is very large and it can be difficult to adequately train a neural network. Moreover, the effects of the uncertainty introduced through the use of artificial neural network proxies and/or upscaling were not investigated.

This paper proceeds as follows. In Section 2, we describe the basic GA optimization procedure including the treatment of risk and uncertainty. Next, in Section 3, we present our new statistical proxies. This includes a discussion of simply computed attributes, the use of cluster analysis, prior and posterior fitness estimates, and the modified GA flowchart. Section 4 contains several examples demonstrating substantial efficiency gains from the new procedure. Conclusions are drawn in Section 5. Note that more details on the GA and statistical proxy procedures can be found in Onwunalu [15].

2 Optimization under geological uncertainty

2.1 The optimization problem

The objective is to optimize a set of parameters **I** describing a field development scenario (number, type, and trajectory of wells to be deployed in a field) in order to maximize

performance, quantified via an objective function f (e.g., cumulative oil production, net present value). Different well types can be considered during the optimization. The mainbore can be vertical, horizontal, or slanted, and several junctions can be drilled into the mainbore. A number of laterals can emanate from the junctions, with different possible orientations. The diameters of the mainbore and laterals can be either defined *a priori* or determined as part of the optimization. In this study, when one or more laterals exist, the mainbore is not perforated (i.e., not open to inflow from the reservoir). In addition, we consider only straight segments, with one segment comprising the mainbore or any lateral. Curved wells could be readily accommodated by allowing multiple segments for the mainbore and laterals.

Any well (or multiple wells) can be represented by a set of parameters specifying the heel (start) and toe (end) of the mainbore, the positions of the junctions along the mainbore (which represent the heels of the laterals), the toe of each lateral, segment diameters, etc. Well rates and other field parameters can also be included in the parameter set. In this study, we apply the parameterization described by Yeten [19].

For any development scenario (with one or several wells) defined in terms of a set of parameters $\mathbf{I}$, the value of the objective function $f(\mathbf{I})$ can be evaluated by using a reservoir simulator. In order to introduce financial considerations into the optimization process, the objective function can be prescribed as the net present value (NPV), defined as the current value of a stream of future payments:

$$\mathrm{NPV} = \sum_{i=1}^{N} \frac{\mathrm{CF}_i}{(1+r)^i}, \tag{1}$$

where N is the total number of discount periods (years), r is the discount rate, and CF_i is the cash flow for the period i, defined as:

$$\mathrm{CF}_i = R_i - E_i, \tag{2}$$

where R_i and E_i indicate the revenue and expenses. The revenue is directly proportional to the production during the considered period. The cost of a particular development scenario is highly case-specific and is influenced by many parameters [4]. We define the revenues and expenses due to production during a discount period as:

$$E_i^{\mathrm{OPEX}} = Q_i^{\mathrm{L}} C^{\mathrm{OPEX}}, \tag{3}$$

where Q_i^{L} is the total production of oil, water, and gas (in STB: Standard Barrel or SCF: Standard Cubic Feet) and C^{OPEX} represents operation revenue or costs (*$/STB or *$/SCF). The cost of the scenario at the beginning of the project (prior to any production) is given by:

$$E^{\mathrm{SCEN}} = \sum_{n=1}^{N^{\mathrm{WELL}}} \left[C^{\mathrm{CAPEX}} + L_n C^{\mathrm{DRILL}} + \sum_{j=1}^{N_n^{\mathrm{JUNC}}} C^{\mathrm{JUNC}} \right], \tag{4}$$

where N^{WELL} is the number of wells; N_n^{JUNC} is the number of junctions for the nth well; C^{CAPEX} is the fixed cost per well, including the cost of the platform and the cost of drilling to the top of the reservoir; C^{JUNC} is the cost of a junction; L_n is the length of the nth well (in feet); and C^{DRILL} is the cost per unit length of drilling the well. Although the cost of drilling varies depending on the type of the segment, the orientation and the well length, we chose to use a constant price in this study. A more general cost function could be easily implemented in this formulation.

To account for economic variability, time-dependent discount rates can be applied to production costs and revenues (dual discount NPV model). This introduces another kind of uncertainty, as it can be used to introduce optimistic or pessimistic economic scenarios in the optimization process.

2.2 Geological uncertainty and risk attitude

As discussed above, accounting for geological uncertainty is a key issue when optimizing development scenarios. Because of the probabilistic nature of the reservoir description, the performance of each scenario is also stochastic, and the optimum scenario directly depends on a strategy toward risk. For example, consider a scenario S_1 that exhibits low variability over the possible realizations of the reservoir and another scenario S_2 with a higher average (expected value) performance but with a much higher variability. For some geological realizations, therefore, S_2 will exhibit lower performance than S_1. Within an optimization framework, the relative ranking of S_1 and S_2 depends on the decision maker's attitude toward risk. It is therefore necessary to quantify this risk attitude prior to performing the optimization, as it will influence the optimization process and the optimal solution.

A simple treatment of risk was used by Yeten *et al*. [20] and Aitokhuehi *et al*. [1]. For N_{real} possible realizations of the geological model, they defined an overall objective function F that incorporates geological uncertainty as follows:

$$F(\mathbf{I}) = \langle f(\mathbf{I}) \rangle + r\sigma_{\mathbf{I}}. \tag{5}$$

In the above equation, $\langle f(\mathbf{I}) \rangle$ is the expectation of f for scenario $\mathbf{I}$ over N_{real} realizations. The $r\sigma_{\mathbf{I}}$ term accounts for both the geological uncertainty and the decision maker's strategy toward risk. Specifically, $\sigma_{\mathbf{I}}$ is the standard deviation of $f(\mathbf{I})$ over the realizations:

$$\sigma_I = \sqrt{\frac{1}{N_{real}} \sum_{n=1}^{N_{real}} (f(\mathrm{I}, n) - \langle f(\mathrm{I}) \rangle)^2}, \tag{6}$$

where $f(\mathbf{I},n)$ is the value of the performance for scenario $\mathbf{I}$ in realization n and r is the risk factor. If $r<0$, the decision

maker is risk-averse, i.e., the objective function will be higher for scenarios that minimize the standard deviation of f over the realizations. In the case of r=0, the decision maker is risk-neutral and the expected value $\langle f(\mathbf{I})\rangle$ will be optimized without any consideration to the variability over the different realizations. This procedure is close to the use of utility functions proposed by Güyagüler and Horne [6].

Here, we apply a more general approach to define the value of the objective function under geological uncertainty. From the $f(\mathbf{I},n)$ for all realizations, we can derive CDF $\{f(\mathbf{I})\}$. From this distribution, we determine $f_{10}(\mathbf{I})$, $f_{50}(\mathbf{I})$, and $f_{90}(\mathbf{I})$, where f_{10} is the value of f corresponding to a probability of 0.1 (i.e., there is a 10% probability that the actual f will be less than f_{10}), and similarly for f_{50} and f_{90}. We then define the objective function (fitness) under geological uncertainty as:

$$F(\mathbf{I}) = r_{10}f_{10}(\mathbf{I}) + r_{50}f_{50}(\mathbf{I}) + r_{90}f_{90}(\mathbf{I}), \tag{7}$$

where the values of r_{10}, r_{50}, and r_{90} depend on our strategy toward risk. A risk-neutral attitude corresponds to:

$$[r_{10} \quad r_{50} \quad r_{90}] = [0 \quad 1 \quad 0], \tag{8}$$

while a risk-averse attitude could correspond to:

$$[r_{10} \quad r_{50} \quad r_{90}] = [0.5 \quad 0.5 \quad 0], \tag{9}$$

This definition can of course be modified to include more (or different) information from the CDF of the property of interest and does not require any Gaussian assumption. As shown in Section 3, this representation will be particularly useful when we account for both geological and proxy uncertainty.

2.3 Description of the genetic algorithm

Genetic algorithms are stochastic optimization methods patterned on natural genetics and Darwinian selection. Each possible solution of the problem (a well deployment scenario in our context) is called an *individual*. The value of the objective function for each individual is referred to as *fitness*. The idea behind GAs is to perform a stochastic combination of the parameters of the most fit individuals within a given *population* in order to create new individuals. These new individuals comprise the next *generation*, and the population evolves iteratively from one generation to another. The main steps of the algorithm are described below.

The first step of a GA consists of defining a representation for any possible scenario in terms of a string of binary values, called a *chromosome* (which represents an individual). This is done by coding the values of every parameter defining the scenario into a binary format and chaining the binary codes. When this procedure has been defined, any string of binary values represents a particular development scenario, as each part of the string is a code for the value of a particular parameter. The parameterization and basic GA applied here are based on the implementation of Yeten [19].

A set of N_{ind} chromosomes can then be generated, creating an initial population of individuals. The individuals in this first generation can be randomly created or they can be explicitly specified. As each individual corresponds to a particular field development scenario, the associated fitness can be evaluated through the use of a reservoir simulator. Individuals are then ranked according to their fitness and are assigned a probability for selection as *parents* for the next generation. To increase the influence of the best individuals during the optimization process, individuals with higher ranks receive a higher selection probability. The specific correspondence between the selection probability and the rank is called the selection strategy (see Yeten [19] for more discussion).

In the next step, N_{ind} individuals are randomly selected, with probabilities based on the rank and the selection strategy, to act as parents for the next generation. The last step randomly *mates* the selected individuals via *crossover* and *mutation* operations. Crossover randomly selects a position on the chromosome string of the two parents and swaps the content of the strings after this position to produce two children. A random number is drawn before this operation, and crossover is performed only if the random number is less than the predefined crossover probability p_{c}. *Mutation* visits all of the bits of each child and flips the bit with a predefined mutation probability p_{m}. Once reproduction has been applied, a set of N_{ind} new individuals (the next generation) is obtained. During the reproduction, *elitism* can also be applied. This means that a small number of the best individuals within a population at generation i are carried directly (without any modification) to the next generation i+1. This option is the only way to ensure that the fitness of the best individual will not decrease during the optimization process.

The population evolves from one generation to another, according to the procedure described above, until a convergence criterion is met (such as no change in the best individual over a specified number of generations). The most fit individual obtained in the last generation is the solution of the overall optimization process. During the reproduction step, newly generated individuals (well configurations) can be nonphysical (e.g., extend beyond the reservoir boundaries) or violate some *a priori* constraint. In these cases, the scenario is not simulated and a very low fitness value is assigned to these invalid individuals.

3 Statistical proxies

Optimization of nonconventional wells under geological uncertainty usually requires many simulation runs. For

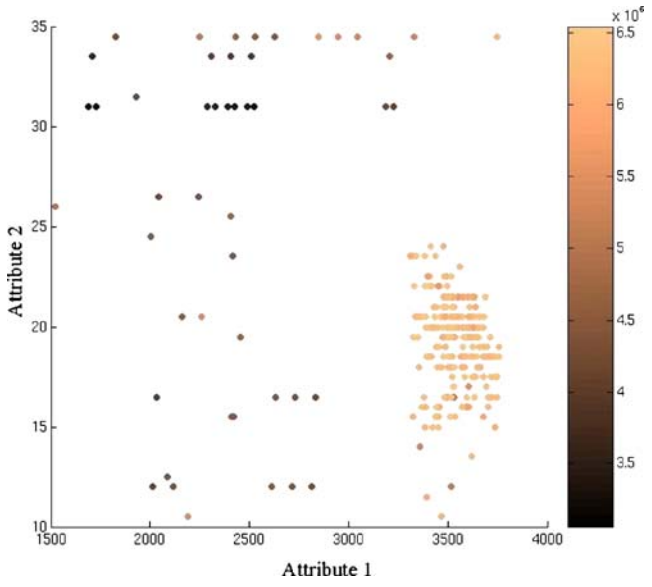

Figure 1 Partition of data from the calibration pool into the space of two attributes. Color corresponds to performance.

example, using the base-state GA without proxies, the number of simulations required is about:

$$N_{\text{sim}} = N_{\text{gen}} \times N_{\text{ind}} \times N_{\text{real}}, \tag{10}$$

where N_{gen} is the number of generations simulated, N_{ind} is the number of individuals in the population, and N_{real} is the number of geological realizations. As the number of individuals and the number of generations are often large (e.g., $\sim O(100)$), the computational requirements can quickly exceed available capabilities even if relatively few geological realizations are considered. Similar computational issues will arise with any stochastic optimization engine, i.e., the computational demands of optimization under geological uncertainty are not unique to the GA employed here.

Our objective is therefore to develop new techniques capable of the fast and (sufficiently) accurate evaluation of scenario performance in the context of field development optimization under uncertainty. These proxies should estimate fitness (i.e., reservoir performance) for general cases and in addition identify the most promising individuals in the population, which are candidates for full simulation.

3.1 Scenario attributes

We propose to generate the proxy by first calibrating the performance (fitness) to a set of *scenario attributes*. These attributes are physical properties that provide indirect information on the performance of the scenario. Examples of very simple attributes are total (perforated) well length, mean permeability along the well, wellbore diameter, and number of high permeability channels intersected. More complex and potentially more predictive attributes entail approximate numerical solutions of the performance of the scenario. These can be performed by using highly upscaled models, semianalytical approximations for single-phase flow [18], or streamline simulation with only one pressure solve (as was previously investigated for use in history matching by Idrobo *et al.* [8] and Mishra *et al.* [13]). Results from neural network estimation can also be applied [20].

A calibration step provides the link between the scenario attributes and the performance. For this calibration, we use a pool composed of all of the simulations already performed during the course of the optimization (as in Güyagüler and Horne [6] and Yeten et al. [20]). In our approach, however, the calibration is performed in the space of the attributes rather than in the space of the optimization parameters. As the dimension of the attribute space is smaller than the dimension of the optimization parameter space, and because the attributes are chosen to provide performance estimation, this approach is expected to give reliable estimates of the performance. If a relatively large number of attributes are used, principal component analysis (PCA) can be applied to remove correlations between the attributes themselves and to reduce the dimension of the calibration space. The principal components are linear combinations of the original attributes (see, e.g., Ripley [16] and Scheevel and Payrazyan [17]).

Once the scenario attributes are selected (and PCA applied if necessary) and a calibration pool is simulated, the proxy estimation can proceed. The main steps of the proxy estimation, detailed in the following sections, are as follows:

– Apply cluster analysis to provide a prior estimate of the CDF for the performance of any new scenario (given the scenario attributes). This provides a fast estimate

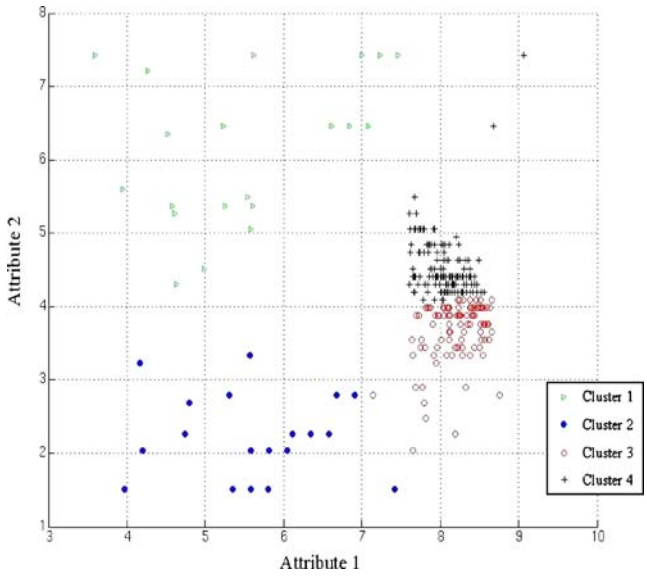

Figure 2 Partitioning of the data points from figure 1 into four clusters. Attribute values have been rescaled by their variance.

(and the associated uncertainty) of the performance of new configuration **I** for any realization n without performing any simulations.

– Average the CDFs for the performance of the new scenario over all of the realizations to give a prior estimate of the overall fitness F of scenario **I**. This also provides an estimate of the prior uncertainty due to both the geological uncertainty and the proxy error.
– Given prior estimates for all cases, a small number of scenarios and realizations are selected for simulation. This selection is based on both the prior overall fitness and the estimate of proxy error. From these simulations, the prior estimates are updated to provide posterior estimates for the selected scenarios. These estimates of the scenario fitness are then used in the GA.

We note that this framework shares some similarities with previous work on error modeling for the correction of upscaled simulations, as described by Lødøen et al. [11] and Omren and Lødøen [14]. In these studies, upscaling error (which is analogous to proxy error in the current context) was modeled by using a number of fine-scale calibration runs. The stochastic model for upscaling error was then combined with the effects of geological uncertainty to provide improved estimates for expected values and confidence interval based on coarse scale simulations.

3.2 Cluster analysis

We determine the proxy estimate for new individuals from the values of their attributes. This is accomplished by calibrating the performance to the attributes, by using a pool of available individuals for which the performance and attributes were previously computed. An initial pool is established before the optimization or over the first few generations, using random and/or predefined test scenarios. The pool is updated over the course of the optimization by using results for simulations performed for selected scenarios. We use this pool to predict the fitness of new individuals and the uncertainty associated with this prediction.

For each case {**I**,n} (this notation refers to individual **I** in realization n) in the calibration pool, the values of the attributes and the value of the performance are known. Different attributes can be analyzed in terms of their correlation with simulated performance and the most appropriate attributes selected. It is important to note that the set of attributes will in general be problem-specific, so it

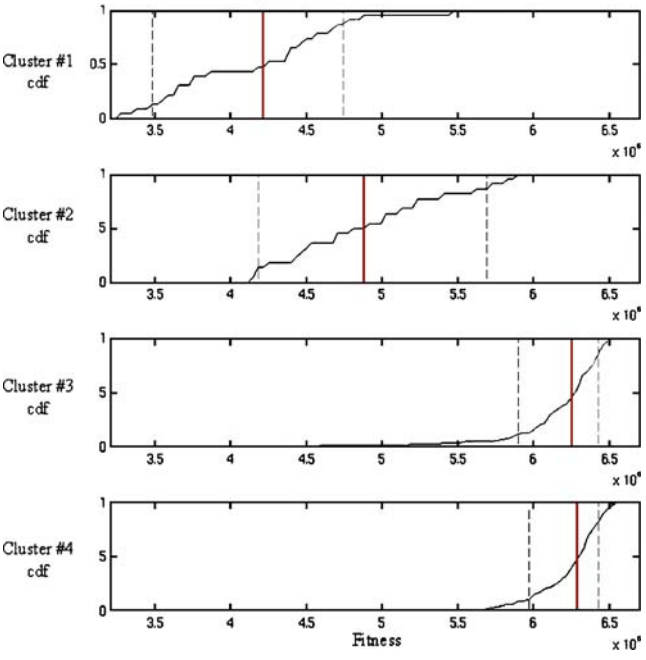

Figure 3 Cumulative distribution function of the performance for each cluster, as evaluated from figure 2.

is essential that the attributes be carefully determined for the case at hand. If the attributes are closely related to the performance, points that cluster together in attribute space should correspond to relatively similar performance. This is illustrated in figure 1, which is a simplified but realistic example based on the optimization of a producer in a channelized system, where Attribute 1 is a measure of the volume of high permeability zones contacted and Attribute 2 is the well length.

The next step is to apply clustering techniques to partition the attribute data into different classes. Different methods exist for this clustering. We use a method referred to as *c*-means or *k*-means [16]. This technique divides the calibration points into a predefined number n_c of clusters, so as to minimize the distance of each data point to the center of its cluster. The algorithm proceeds in an iterative way: first, n_c centers are drawn in the space of the attributes (rescaled by their variance). Then, every data point is related to the nearest cluster center (usually using the Euclidean norm), hence defining n_c clusters of points. The position of the center of each cluster is then updated based on the points in the cluster. Then, every data point is related again to its nearest cluster center by using the new positions of the centers. This process is repeated until convergence is reached. With this algorithm, different clusters may include different numbers of points, and some may eventually be empty. Although it always converges, this algorithm does not ensure that the optimum partition has been reached [12].

Thus this method provides a predefined number of clusters of data points in the attribute space using scenarios from the calibration pool (figure 2). It should be emphasized that this clustering is performed with respect to the attributes only, and is not based on the performance values.

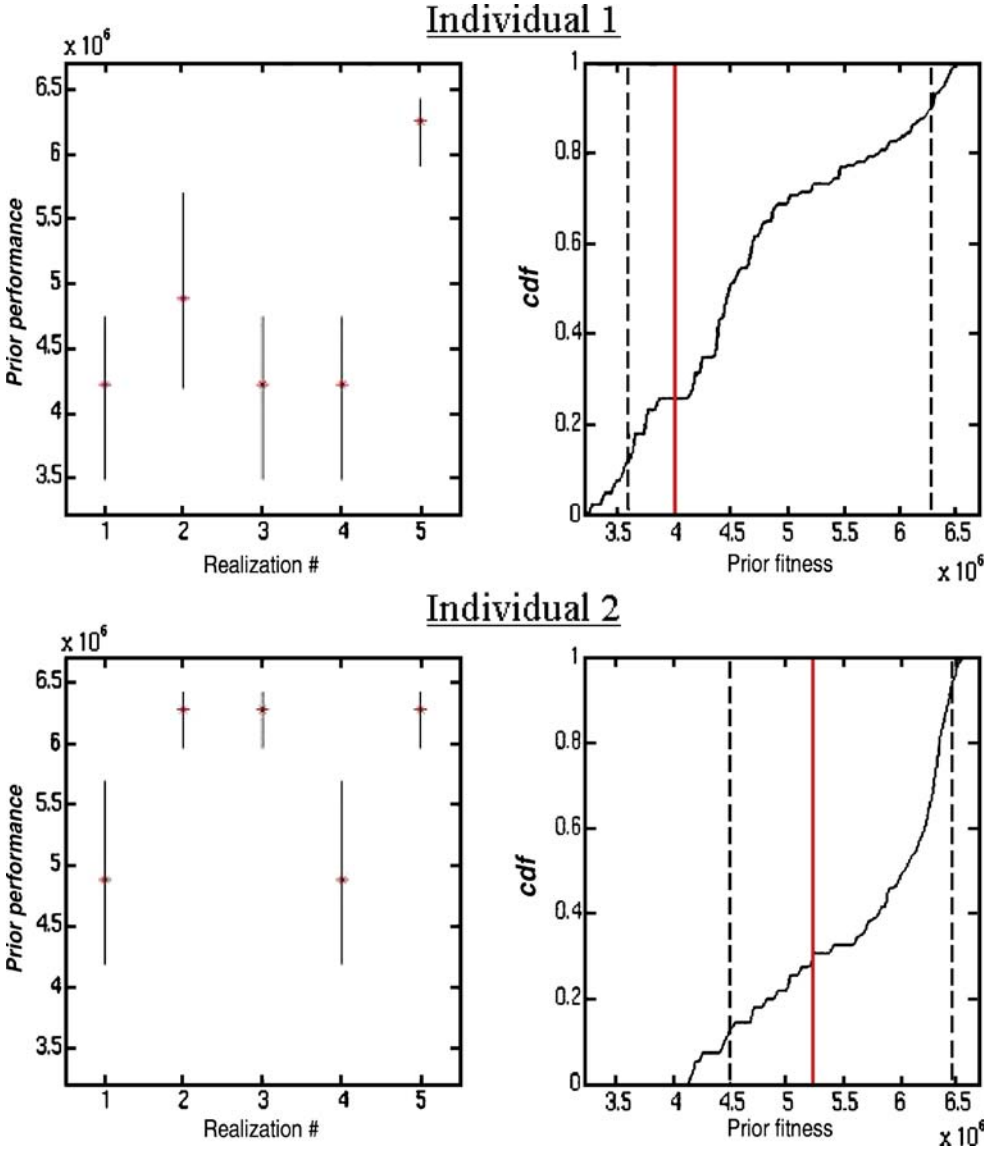

Figure 4 Prior distribution function of the performance for two individuals. Left plots show the expected performance and confidence interval for each realization. Plots on the right show the prior cumulative distribution function of the performance. On right plots, dashed lines correspond to $f_{10}^{prior}(\mathbf{I})$ and $f_{90}^{prior}(\mathbf{I})$ values, while the solid vertical lines correspond to the prior value of the fitness (using r_{10}= 0.5, r_{50}=0.5, r_{90}=0).

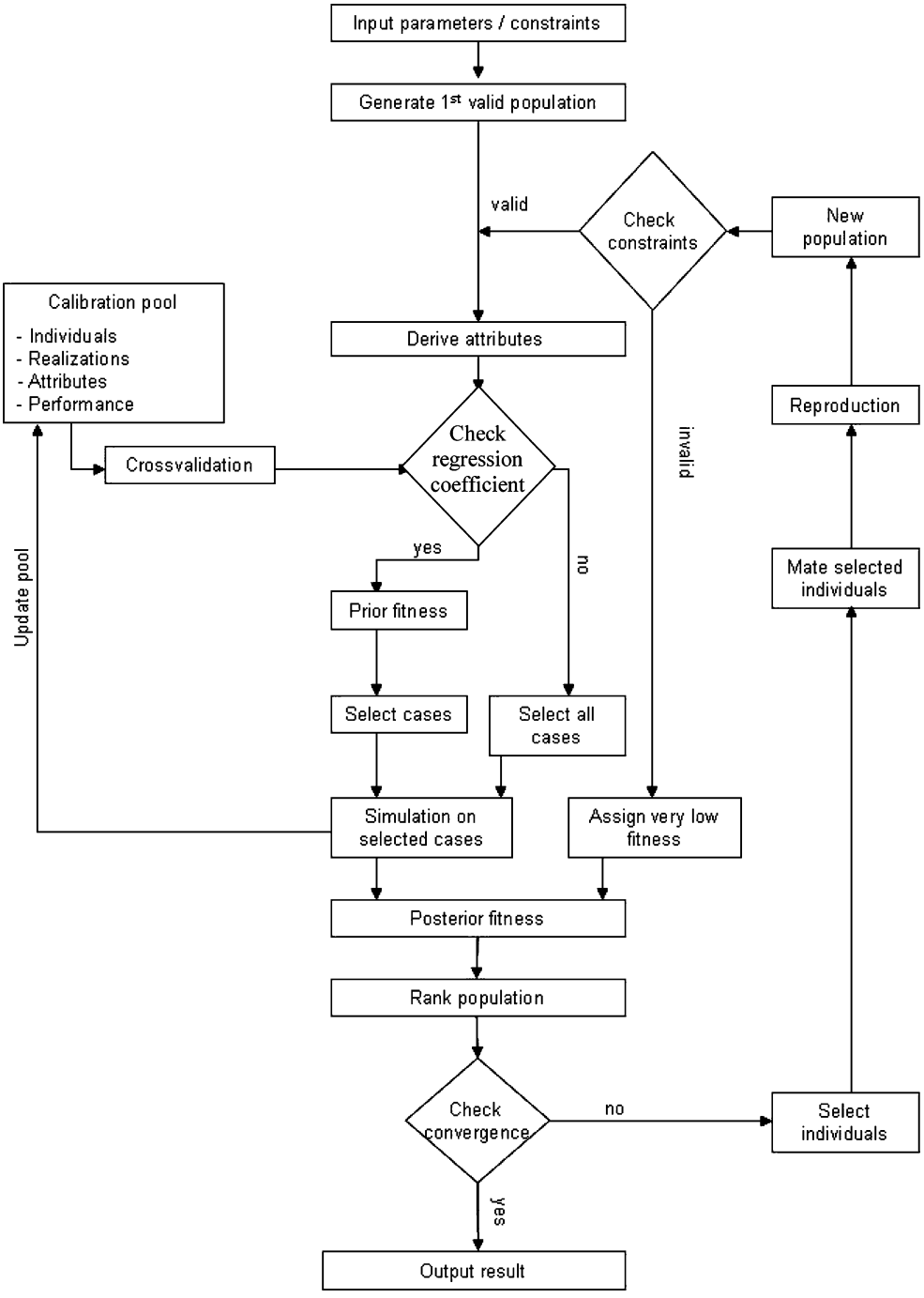

Figure 5 Flowchart of the genetic algorithm hybridized with statistical proxies.

As the performance of each individual inside each cluster is known, a simple statistical analysis provides an experimental CDF of the values of the performance for each cluster (figure 3). Then, when the prior performance of a new case $\{\mathbf{I},n\}$ is to be estimated, its attribute values are computed, indicating which cluster this case falls into. As each cluster has an associated CDF for the performance, we immediately have a prior estimate (in terms of a CDF) for the performance of the new case. This estimation is described in more detail in the next section.

3.3 Prior value of the fitness

As indicated above, depending on the value of its attributes, every new case $\{\mathbf{I},n\}$ can be related to an existing cluster and the associated CDF. This gives a direct estimate of the proxy uncertainty. As the attributes take different values depending on the geological realization, the cluster into which the scenario falls will, in general, differ for each realization. For each case $\{\mathbf{I},n\}$, therefore, there is a different CDF of the objective function, which we term

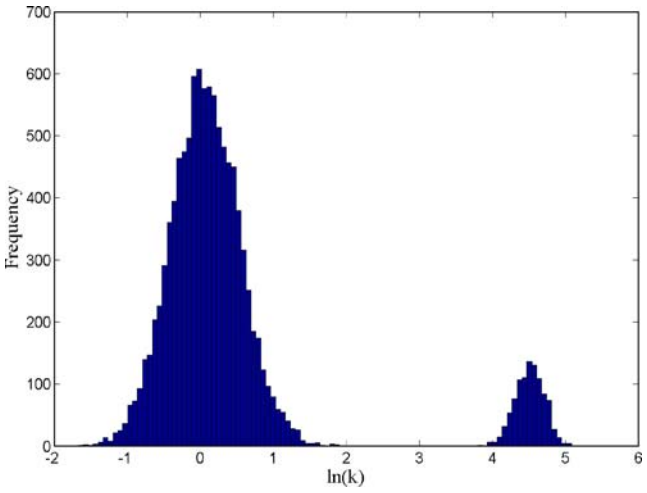

Figure 6 Histogram of the logarithm of the permeability field.

$\mathrm{cdf}_{\mathrm{prior}}\{f(\mathbf{I}|n)\}$. This function is a prior estimate because no simulations have been performed on the new individuals. From this distribution, we can determine the expected value of the fitness of this case, and the associated confidence interval in terms of $f_{10}^{\mathrm{prior}}(\mathbf{I},n)$, $f_{50}^{\mathrm{prior}}(\mathbf{I},n)$, and $f_{90}^{\mathrm{prior}}(\mathbf{I},n)$.

Two types of uncertainty therefore exist for any individual **I**. The first uncertainty is due to geological uncertainty, which results in the individual falling into different clusters for different realizations. This uncertainty would exist even if we had a perfect proxy estimate (i.e., if the CDF in each cluster was a step function). The second uncertainty is due to the uncertainty in the proxy itself and is characterized by the CDF for the particular cluster.

A simple way to combine these two uncertainties is to average the prior CDFs for individual **I** over all of the realizations. Weightings could also be introduced to reflect the probabilities associated with the various realizations. This provides a *prior cumulative distribution function* for the overall performance of the individual:

$$\mathrm{cdf}_{\mathrm{prior}}\{f(\mathbf{I})\} = \frac{1}{N_{\mathrm{real}}}\sum_{n=1}^{N_{\mathrm{real}}} \mathrm{cdf}_{\mathrm{prior}}\{f(\mathbf{I}|n)\}. \tag{11}$$

The way that this prior CDF accounts for both geological and proxy uncertainty is illustrated in figure 4 for two individuals and five realizations. In this figure, the first individual is assumed to fall in clusters 1, 2, and 4 (from

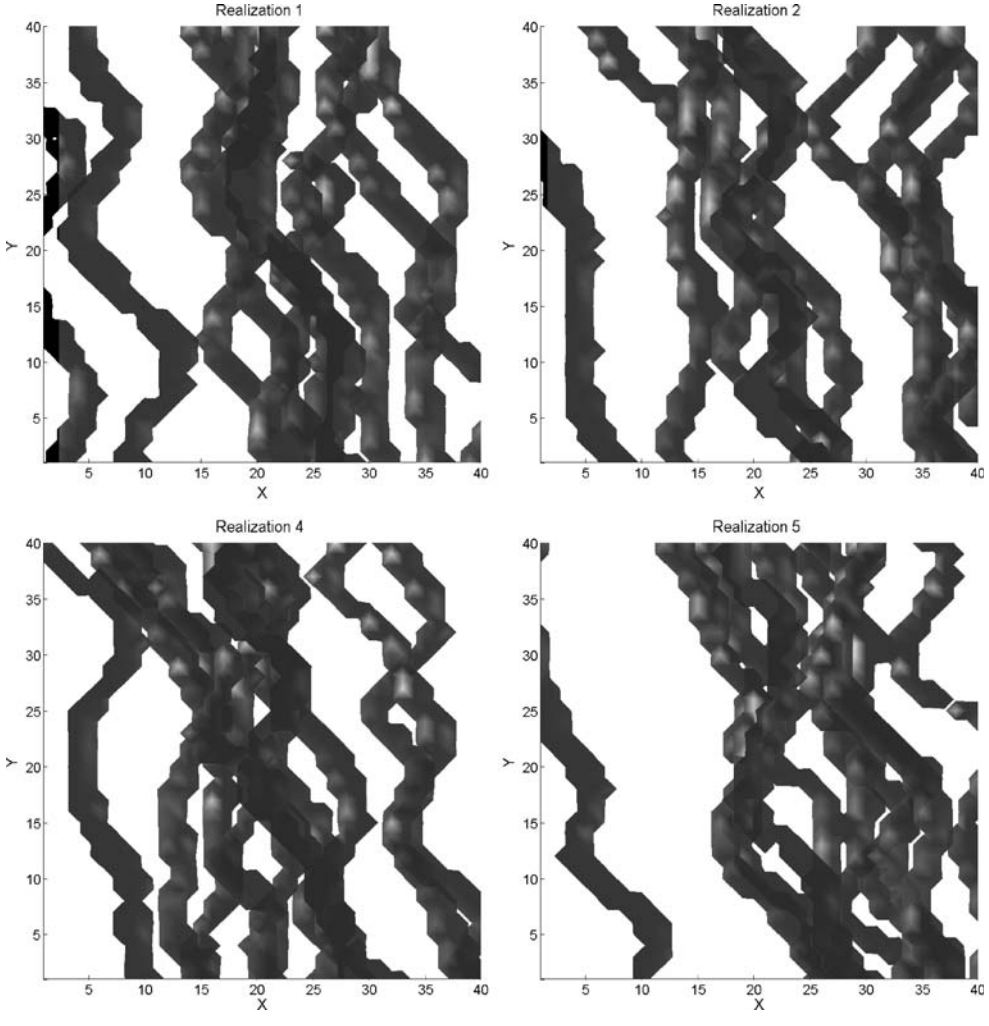

Figure 7 Four realizations of the channelized permeability field.

Table 1 Example A: reservoir and fluid properties.

Grid dimensions	40×40×7
Field dimension	6000×6000×210 ft^3
φ	0.2
$\overline{k_1}$	90 mD
$\overline{k_2}$	1 mD
C	3×10^{-5} psi^{-1}
B_o	1.3

figures 2 and 3) over the different realizations, while the second individual is assumed to fall in clusters 2 and 4. In the left figures, the stars correspond to the expected performance derived from the cluster for each realization, while the vertical bars are the proxy uncertainty associated with each cluster. Summing the CDFs associated with each cluster over the five realizations leads to the prior distributions shown on the right plots, which account for both proxy and geological uncertainty.

At this point, we can compute a prior value of the objective function for the individual based on the CDF and the strategy toward risk. For example, from the overall prior CDF we can compute $f_{50}^{\text{prior}}(\mathbf{I},n)$, $f_{10}^{\text{prior}}(\mathbf{I},n)$, $f_{90}^{\text{prior}}(\mathbf{I},n)$ and thus the prior value of the fitness as:

$$F_{\text{prior}}(\mathbf{I}) = r_{10}f_{10}^{\text{prior}}(\mathbf{I}) + r_{50}f_{50}^{\text{prior}}(\mathbf{I}) + r_{90}f_{90}^{\text{prior}}(\mathbf{I}), \tag{12}$$

where r_{10}, r_{50}, and r_{90} depend on our strategy toward risk. In the example shown, the value of the prior fitness is computed using $[r_{10}\ r_{50}\ r_{90}]=[0.5\ 0.5\ 0]$, which corresponds to a risk-averse attitude.

3.4 Posterior value of the fitness

The idea at this point is to identify a relatively small number of selected cases $\{\mathbf{I},n\}$ for simulation. This determination is based on the prior CDF as well as the estimate of proxy uncertainty, as we now describe. We focus on individuals $\mathbf{I}$ that provide high prior estimates of the objective function, $F(\mathbf{I})$, as given by equation (12). However, because this prior objective function includes proxy uncertainty, we are also interested in cases $\{\mathbf{I},n\}$ for which the uncertainty in $F(\mathbf{I})$ is reduced after the simulation.

For each individual $\mathbf{I}$ and realization n, we define a weight $W(\mathbf{I},n)$ as:

$$W(\mathbf{I},n) = W_1(\mathbf{I}) + W_2(\mathbf{I},n), \tag{13}$$

where $W_1(\mathbf{I})$ is proportional to the prior objective function for the individual and $W_2(\mathbf{I},n)$ is proportional to the uncertainty on the proxy, defined from the shape of $\text{cdf}_{\text{prior}}\{f(\mathbf{I}|n)\}$ (e.g., confidence interval). For example, we can set:

$$W_1(\mathbf{I}) = F_{\text{prior}}(\mathbf{I}), \tag{14}$$

and

$$W_2(\mathbf{I},n) = \alpha\left[f_{90}^{\text{prior}}(\mathbf{I},n) - f_{10}^{\text{prior}}(\mathbf{I},n)\right], \tag{15}$$

where α is the scaling factor that can be used to set the relative weightings of the two terms. Once all cases $\{\mathbf{I},n\}$ have received a weight $W(\mathbf{I},n)$, a predefined percentage of cases are selected for simulation. The likelihood of selection for a particular case depends on the value of $W(\mathbf{I},n)$. Assume, for example, that we have 30 individuals and 10 geological realizations (which means that we have about 300 scenarios to evaluate at each generation) and that we can afford to simulate only 10% of the cases. In the approach taken here, we simulate the 30 cases at each generation with the highest values of $W(\mathbf{I},n)$. An alternate

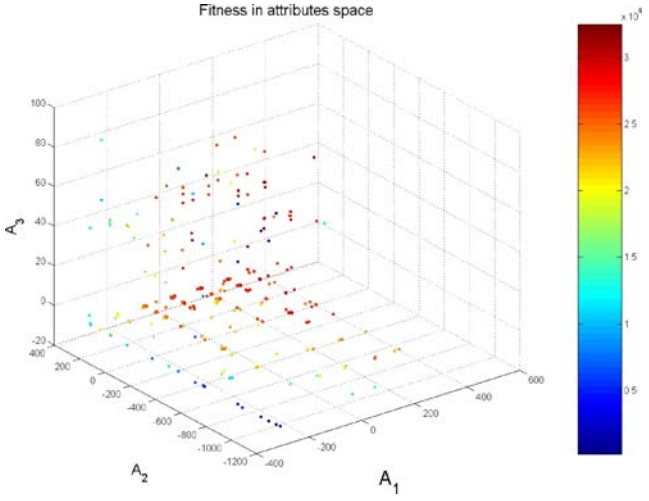

Figure 8 Example A: repartition of the calibration data in the space of the attributes after two generations for case A.2. Color corresponds to the performance. Attributes are well length, volume of channels intersected by the well, and average permeability along the well.

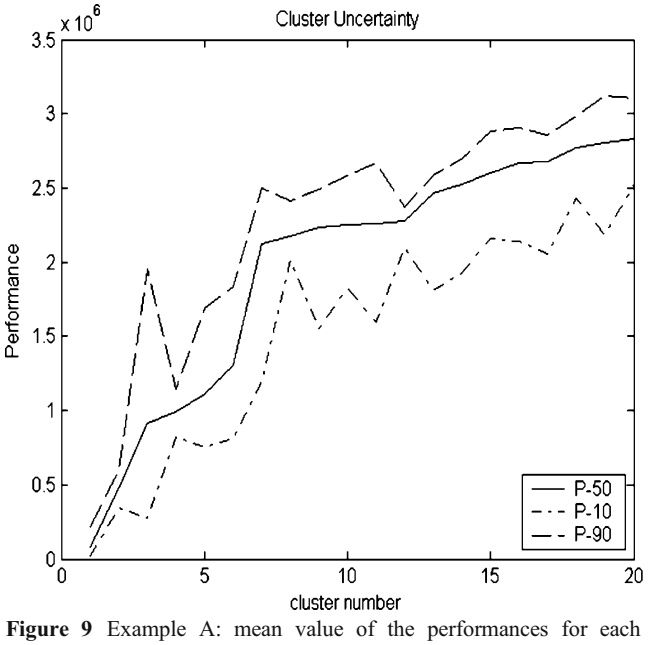

Figure 9 Example A: mean value of the performances for each cluster (solid line). Dashed lines correspond to the confidence interval.

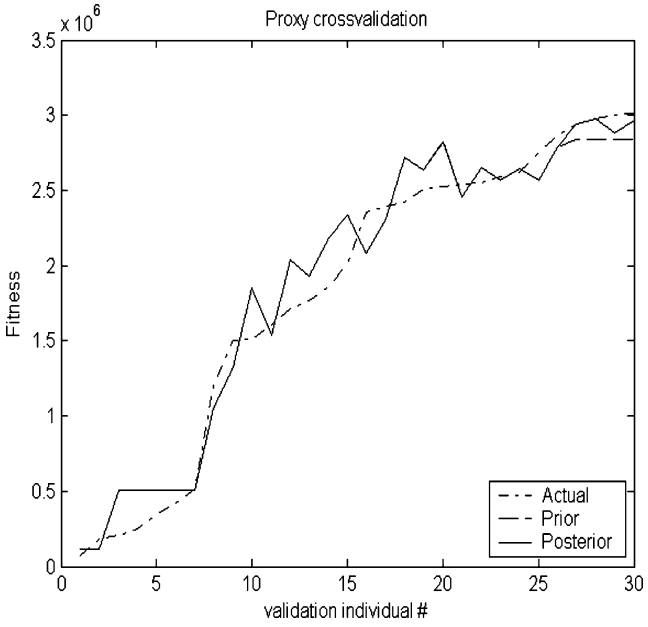

Figure 10 Example A: result of the cross-validation of the calibration pool for case A.1 after 10 generations. Only *a priori* best cases are updated for posterior estimation of the fitness.

approach would be to use a selection strategy in which cases with the highest values of $W(\mathbf{I},n)$ have the greatest probability of selection, but other cases also have nonzero selection probabilities.

The weighting process is hence a way to choose individuals with a high overall prior fitness (high $W_1(\mathbf{I})$), but also to identify, for a particular individual $\mathbf{I}$, which realizations n bring the largest uncertainty in the fitness estimate (quantified via $W_2(\mathbf{I},n)$). For example, assume that we wish to simulate only two of the 10 cases (defined by the two individuals and five realizations) from figure 4. Our weighting process leads to the selection of the second individual in realizations 1 and 4. This is because (1) this individual exhibits a higher prior fitness than the first individual and (2) individual 2 in realizations 1 and 4 corresponds to clusters with broad CDFs.

From the simulations of the selected cases, the actual value of the fitness is obtained and we can update the prior CDFs to obtain posterior CDFs. If the set $\{\mathbf{I},n\}$ was actually simulated, then we know the actual value of the fitness $f(\mathbf{I},n)$ and $\mathrm{cdf}_{\mathrm{post}}\{f(\mathbf{I}|n)\}$ is simply a step function. If $\{\mathbf{I},n\}$ was not simulated, no new information is added and $\mathrm{cdf}_{\mathrm{post}}\{f(\mathbf{I}|n)\} = \mathrm{cdf}_{\mathrm{prior}}\{f(\mathbf{I}|n)\}$.

We can now form the posterior CDF of the overall fitness:

$$\mathrm{cdf}_{\mathrm{post}}\{f(\mathbf{I})\} = \frac{1}{N_{\mathrm{real}}}\sum_{n=1}^{N_{\mathrm{real}}} \mathrm{cdf}_{\mathrm{post}}\{f(\mathbf{I}|n)\}. \tag{16}$$

This posterior distribution function for the fitness differs from the prior distribution in that the uncertainty due to the use of the proxy is reduced, particularly if this individual was simulated for several realizations. From this posterior CDF, we can derive $f_{50}^{\mathrm{post}}(\mathbf{I})$, $f_{10}^{\mathrm{post}}(\mathbf{I})$, and $f_{50}^{\mathrm{post}}(\mathbf{I})$. At this point, we compute a posterior value of the objective function for the individual as:

$$F_{\mathrm{post}}(\mathbf{I}) = r_{10}f_{10}^{\mathrm{post}}(\mathbf{I}) + r_{50}f_{50}^{\mathrm{post}}(\mathbf{I}) + r_{90}f_{90}^{\mathrm{post}}(\mathbf{I}). \tag{17}$$

The difference between $F_{\mathrm{post}}(\mathbf{I})$ and $F_{\mathrm{prior}}(\mathbf{I})$ will be illustrated in Section 4.

All the individuals in the population are now ranked according to the posterior value of the objective function, $F_{\mathrm{post}}(\mathbf{I})$. At this point, we return to the basic GA. Specifically, a selection probability, depending on the selection strategy and rank, is assigned to each individual. The next generation is then formed from the selected individuals, as described in Section 2.3.

3.5 Hybridization of the genetic algorithm: flowchart

A detailed flowchart of the new hybrid GA is illustrated in figure 5. Cross-validation can be used to check the accuracy

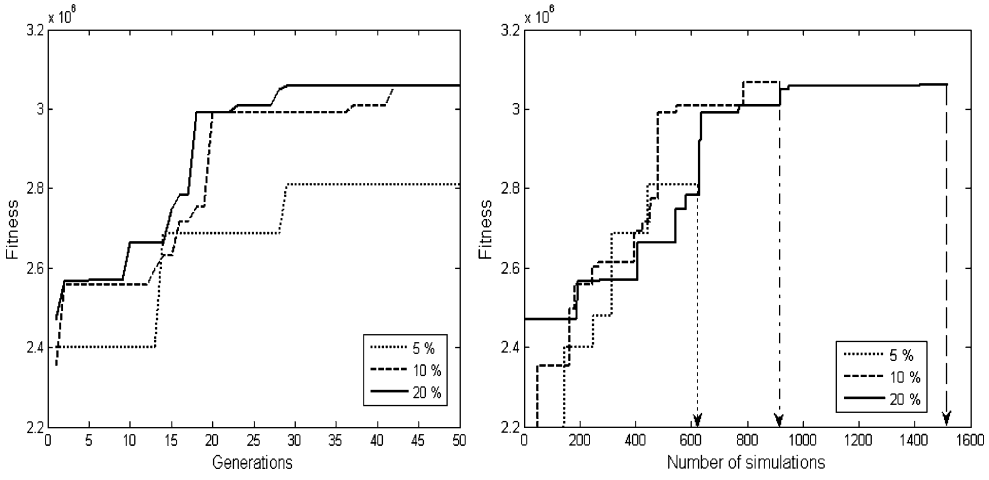

Figure 11 Example A: evolution of the best individual in the population with the number of generations and with the total number of simulations.

Table 2 Example A: results with different percentages selected for simulation.

	5%	10%	20%
Fitness of the best scenario (bbl)	2,809,000	3,066,000	3,060,000
Number of simulations	630	918	1520

of the proxy derived from the calibration pool. This is accomplished by removing some individuals {**I**,*n*} from the calibration pool and then predicting their fitness using the proxy. From this process, a regression coefficient between the actual fitness and the proxy estimate can be determined. If the value of this regression is below some predefined value, more simulations must be performed, and the proxy recalibrated, before the proxy is used. For some particular {**I**, *n*}, the values of the attributes may not correspond to any existing cluster or they may correspond to a cluster with very few points. When this occurs, the case is simulated.

The main steps of the algorithm are as follows:

1 At each generation, determine clusters in the attribute space. Associate each new case {**I**,*n*} with a cluster (unless it falls outside of all existing clusters, in which case we perform a flow simulation).
2 Compute *a priori* values for the performance for each individual **I** from the CDFs for each of the {**I**,*n*} (i.e., average the appropriate cluster CDFs).
3 Weight each case {**I**,*n*} according to equation (13). Simulate a predefined percentage of the cases (selected based on the weights).
4 Derive the posterior fitness for every individual and update the calibration pool.
5 Rank individuals based on posterior fitness and select parents for the next generation.
6 Apply crossover and mutation operations to provide the next generation of individuals.

This algorithm proceeds until some convergence criterion is met (e.g., no improvement in the most fit individual over a prescribed number of generations) or a stop criterion is reached.

4 Examples

We now present three example cases that illustrate the use of the statistical proxy described in the previous section. These examples involve the optimization of monobore or dual-lateral wells under geological uncertainty. An additional example involving the optimization of a monobore well under varying risk attitudes is presented by Onwunalu [15].

4.1 Example A: sensitivity to the proxy selection

In this example, we illustrate the sensitivity of the optimization result to the percentage of scenarios simulated. The well is constrained to be a monobore. The reservoir model is a channelized system. Five realizations constrained to data from three observation wells were randomly generated (figures 6 and 7). The key properties of the reservoir are summarized in table 1 (note that C stands for total compressibility and B_o is oil formation volume factor). Reservoir flow in this case involves only a single phase (oil) and frictional pressure losses in the well are neglected. Permeability is highly heterogeneous but locally isotropic ($k_x=k_y=k_z$).

The goal is to determine the placement of a monobore production well to maximize the cumulative oil produced over 500 days of primary depletion. Initial pressure at the top of the reservoir is 3500 psi and the bottomhole pressure

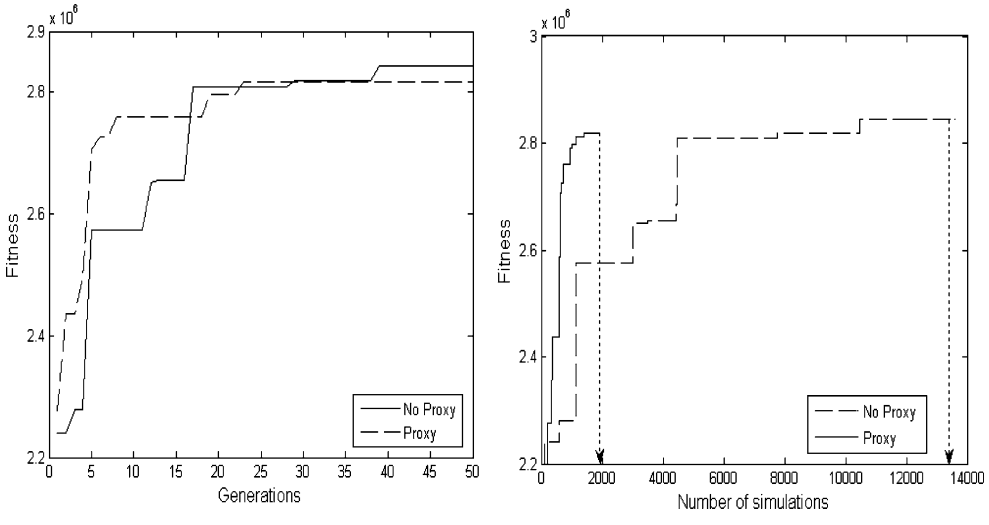

Figure 12 Example B: evolution of the best individual in the population with the number of generations and with the total number of simulations.

Table 3 Example B: comparison of the performances of the best wells found with and without proxy.

	With proxy	Without proxy
Average performance (bbl)	2,831,000	2,872,000
Standard deviation (bbl)	219,300	279,100
Number of simulations	1873	13,597

(BHP) constraint is 3200 psi. The mainbore can be oriented in any direction, but is limited in horizontal extent to a maximum of 1500 ft. We optimize under a risk-neutral attitude, which means that we seek to maximize the expected cumulative oil production over the five realizations, regardless of the variance.

We base the proxy estimate on three attributes expected to correlate with cumulative oil production – well length, volume of channels intersected by the well, and average permeability along the well. For a given individual, the determination of the proxy estimate required about 2 s processing for each realization, while each reservoir simulation consumed about 30 s. Simulations are very fast for this small problem (11,200 cells), so there is only a factor of 15 differences between the proxy calculations and the simulation runs. However, because the processing time required to determine the attribute values is largely independent of the model size, substantially greater differences between the times required for deriving attributes and for performing simulations are expected for larger problems.

In this example, we used 30 individuals and 50 generations in the GA. The calibration pool was initiated by simulating all of the individuals in each realization in the first generation (for a total of 150 simulations) and was updated at each generation using the simulated cases. Twenty clusters were used to determine the prior fitness of each individual. Prior predictions obtained from a given cluster were used only if the cluster contained more than 10 data points; otherwise, a simulation was performed.

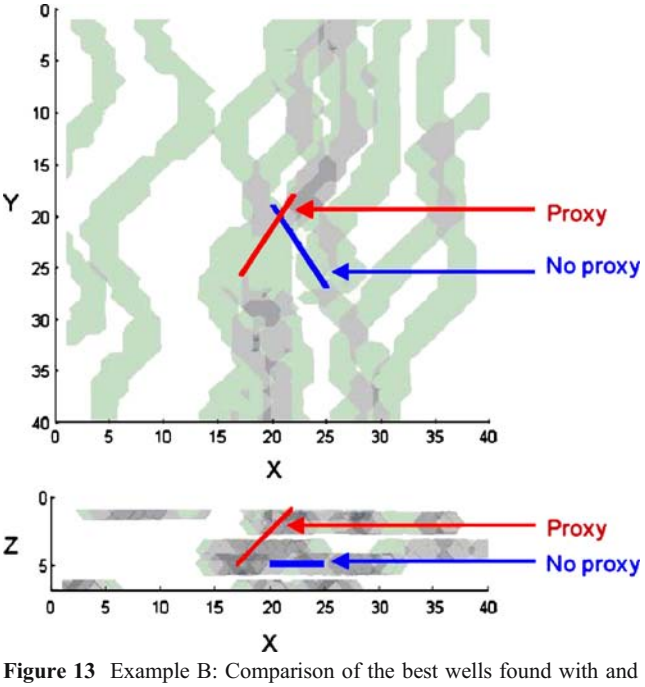

Figure 13 Example B: Comparison of the best wells found with and without proxy.

Three cases were tested in order to quantify the sensitivity of the optimization results to the percentage of (valid) scenarios simulated. These cases correspond to 5% (case A.1), 10% (case A.2), and 20% (case A.3) simulated, with the remainder estimated using the proxy. Figure 8 shows the data in the three-dimensional attribute space after two generations. It is evident from the figure that the data tend to cluster based on fitness, which suggests that the variability within a cluster will not be excessive. At the end of the second generation, there are about 300 individuals in the calibration pool (most of the cases in the second generation were also simulated because the clusters did not yet contain enough points). The fitness of all individuals in the calibration pool was then partitioned into 20 clusters using the selected attributes.

The CDF of the fitness in all clusters is now readily determined. Figure 9 depicts the $f_{10}^{\text{prior}}(\mathbf{I},n)$, $f_{50}^{\text{prior}}(\mathbf{I},n)$, and $f_{90}^{\text{prior}}(\mathbf{I},n)$ values of the performance for the 20 clusters (the clusters are now ordered in terms of increasing f_{50}^{prior}). For subsequent generations, we compute the attributes for all new cases and then determine the prior fitness. A subset of cases is then simulated as determined by the procedure described in Section 3.4. In general, there are 8, 15, and 30 actual simulations for case A.1 (5%), case A.2 (10%), and case A.3 (20%), respectively, for each generation of the optimization. A comparison of the posterior and actual fitness for case A.1 (from cross-validation of the calibration pool at generation 10) is shown in figure 10. We see from figure 10 that, although only about eight simulations are performed at each generation, there is little difference between the prior and posterior fitness. Using cross-validations of the calibration pool, updated after each generation, we compute the correlations between the prior and posterior fitness values with the actual fitness. The regression coefficient between the prior fitness and the actual fitness is seen to be quite high – greater than 0.9 at all generations of the optimization.

In figure 11 (left), we present the fitness of the best individual as a function of generation, while in figure 11

Table 4 Example C: comparison of the performances of the best wells found with and without proxy.

	With proxy	Without proxy
Average performance (bbl)	3,269,000	3,067,000
Standard deviation (bbl)	164,300	463,100
Number of simulations	1626	6677

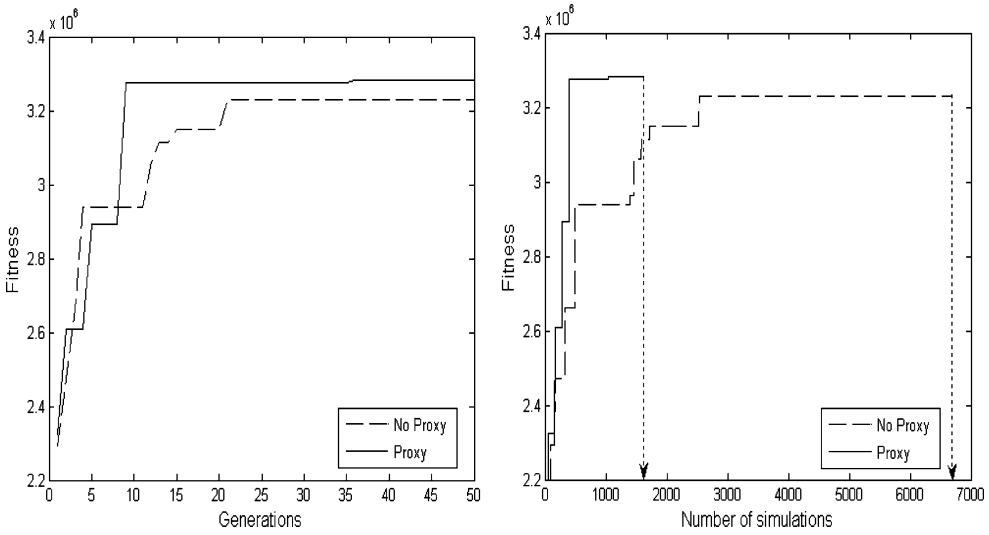

Figure 14 Example C: evolution of the best individual in the population with the number of generations and with the total number of simulations.

(right) we present the fitness of the best individual as a function of the number of simulations performed. From figure 11 (left), it is evident that cases A.2 and A.3 (10% and 20% simulated) provide very similar results in terms of best fitness, while case A.1 (5% simulated) provides a somewhat lower fitness. However, by using only 630 simulations, the best individual in case A.1 is more than 90% of the fitness in case A.2 and case A.3 (figure 11).

The key results from figure 11 are summarized in table 2. These results suggest that once a "threshold" percentage of the cases are simulated, the optimization result is relatively insensitive to higher percentages of simulated cases. Note that the number of simulations does not decrease by a factor of 2 when the percent simulated is halved. This is because considerably more than the specified percentage of cases must be simulated in early generations to build clusters with 10 points or more. In more extensive optimizations, halving the percent simulated will result in about half the number of simulations as expected.

4.2 Example B: optimization of a monobore production well

In this example, we again optimize the placement of a monobore producer in order to maximize the cumulative oil production after 500 days of primary depletion. We consider the same reservoir properties and constraints as in Example A (table 1), except that here we consider 10 realizations constrained to the three vertical observation wells. Our goal here is to assess the proxy by comparing optimization results obtained by simulating all cases with those achieved through use of the statistical proxy. Maximum well length is again 1500 ft and the optimization was again performed with 30 individuals (50 generations) under a neutral-risk mode. Following the initial generations (in which high percentages of individuals were simulated to generate 20 clusters with at least 10 data points per cluster), the proxy was applied to select 10% of the valid population for simulation at each generation (with the remainder of the individuals assessed through use of the proxy).

Figure 12 presents the evolution of the best individual over the course of the optimization. The solid line shows the result when all cases are simulated (no proxy) and the dashed line the result when the proxy is applied. From figure 12 (left), we see that the no proxy case provides a slightly better fitness (see also table 3). It is apparent from figure 12 (right), however, that the proxy result is achieved by using more than a factor of 7 fewer simulations. It is thus apparent that the use of the

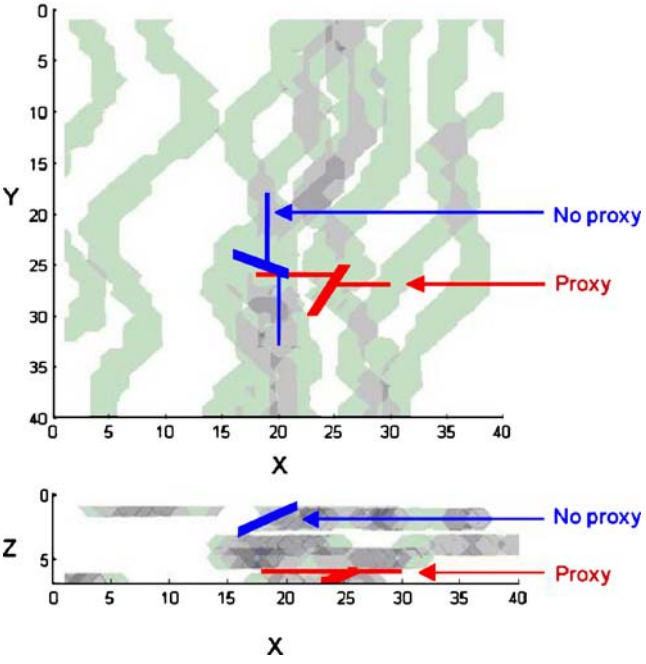

Figure 15 Example C: comparison of the best wells found with and without proxy.

proxy is able to greatly reduce the number of simulation runs while impacting only slightly the optimal fitness.

The positions of the optimal wells, as determined by the optimizations with and without the proxy, are shown in figure 13. The optimal wells clearly differ between the two optimizations, though in both cases they lie near the center of the reservoir and contact several channels. This figure suggests that a number of different wells may provide comparable optimums, although it is interesting to see that in this case the two optimizations result in wells that are similar.

4.3 Example C: optimization of a dual-lateral producer

In this example, we optimize the placement of a dual-lateral production well. We again seek to maximize the cumulative oil production over 500 days of primary depletion. We consider the same reservoir properties as in Example A (table 1), using five realizations constrained to three vertical observation wells. We again specify a BHP constraint of 3200 psi with an initial reservoir pressure of 3500 psi. The well in this case is required to have two laterals. The mainbore and the laterals can be oriented in any direction, and their horizontal lengths can take values from zero (for a vertical segment) to a maximum of 1500 feet. We again used populations of 30 individuals and ran the optimization for 50 generations. The strategy toward uncertainty was risk-neutral. In the optimizations using the proxy, we populated the clusters as described in the previous examples. When the proxy was applied, 20% of the cases were simulated.

Figure 14 compares the fitness of the best individual over the course of the optimization when all cases are simulated and when the proxy is applied. Optimization results are also summarized in table 4. It is evident that the proxy optimizations actually achieve a slightly better solution than the optimization without the proxy, in which all cases are simulated. The proxy optimizations require about a factor of 4 fewer simulations. The optimum well locations for both cases are shown in figure 15. Although the locations differ, it is clear that the wells in both cases contact several channels and are located around the middle of the reservoir.

The examples presented in this section demonstrate that the statistical proxy developed in this paper acts to significantly accelerate the GA optimization. In cases where a comparison to the full simulation (no proxy) optimization was made, the optimums achieved by the two procedures were found to be very close. As more generations and more individuals are considered in the optimization, the speedups offered by the proxy will increase, as the cost of the calibration runs will become a smaller component of the overall computations. Thus, the approach presented here clearly provides substantial speedup while maintaining the quality of the optimization procedure.

5 Summary and conclusions

In this paper, we applied a GA for optimizing the deployment of nonconventional wells under geological uncertainty. As the direct application of the base method leads to excessive computational demands, we developed and tested a new statistical proxy which acts to significantly reduce the number of cases actually simulated. This proxy incorporates ideas from cluster analysis and provides an estimate of the CDF of the fitness (or objective function). The specific conclusions from this study are as follows:

- Cluster analysis based on calibrating simply computed attributes (e.g., well length, number of channels intersected, highly simplified flow simulations) to simulation results provides a means for forming prior CDFs for the performance of a well in a particular geological model. By combining these proxy estimates for multiple geological realizations, a prior CDF for well performance, which accounts for both geological and proxy uncertainty, can be developed.
- The prior performance estimate, in conjunction with the proxy uncertainty, can be used to select a subset of cases most appropriate for full simulation. The fitness of the remainder of the cases is determined by using the proxy. It was demonstrated that the proxy is effective in terms of identifying appropriate cases for simulation and that a high degree of correlation exists between the proxy estimate and the simulated result.
- In examples involving the optimization of well placement for monobore and dual-lateral wells, the use of the proxy was shown to provide excellent results. Specifically, by simulating only 10% of the cases (as determined by application of the proxy), optimums very close to those achieved by the full procedure were attained.

Acknowledgement We are grateful to the industrial affiliates of the SUPRI-HW (Advanced Wells) research program at Stanford University for partial support of this work. V.A. also thanks the Institut Français du Pétrole (IFP) for partial funding. We thank Schlumberger for providing us with the ECLIPSE simulator, which was used for the simulations.

References

1. Aitokhuehi, I., Durlofsky, L.J., Artus, V., Yeten, B. Aziz, K.: Optimization of advanced well type and performance. In: Proc. of the 9th European Conf. on the Mathematics of Oil Recovery, Cannes, France, (2004) 30 August–2 September
2. Bangerth, W., Klie, H., Wheeler, M.F., Stoffa, P.L., Sen, M.K.: On

optimization algorithms for the reservoir oil well placement problem. Comput. Geosci., in press
3. Bittencourt, A.C., Horne, R.: Reservoir development and design optimization (paper SPE 38895). In: SPE Annual Technical Conference and Exhibition, San Antonio, Texas, (1997) 5–8 October
4. Cho, H.: Integrated optimization on a long horizontal well length. SPE Reserv. Evalu. Eng. 81–87, (2003) April
5. Couët, B., Bailey, W.J., Wilkinson, D.: Reservoir optimization tool for risk and decision analysis. In: Proc. of the 9th European Conf. on the Mathematics of Oil Recovery, Cannes, France, (2004) 30 August–2 September
6. Güyagüler, B., Horne, R.: Uncertainty assessment of well placement optimization, (paper SPE 71625). In: SPE Annual Technical Conference and Exhibition. New Orleans, Louisiana, (2001) 30 September–3 October
7. Güyagüler, B., Horne, R., Rogers, L., Rosenzweig, J.J.: Optimization of well placement in a Gulf of Mexico waterflooding. SPE Reserv. Evalu. Eng. 229–236 (2002) June
8. Idrobo, E.A., Choudhary, M.K., Datta-Gupta, A.: Swept volume calculations and ranking of geostatistical reservoir models using streamline simulation (paper SPE 62557). In: SPE/AAPG Western Regional Meeting, Long Beach, California, (2000) 19–23 June
9. Joshi, S.D.: Costs/benefits of horizontal wells (paper SPE 83621). In: SPE Western Regional/AAPG Pacific Section Joint Meeting, Long Beach, California, (2003) 19–24 May
10. Klie, H., Bangerth, W., Wheeler, M.F., Parashar, M., Matossian, V.: Parallel well location optimization using stochastic algorithms on the grid computational framework. In : Proc. of 9th European Conf. on the Mathematics of Oil Recovery, Cannes, France, (2004) 30 August–2 September
11. Lødøen, O.P., Omre, H., Durlofsky, L.J., Chen, Y.: Assessment of uncertainty in reservoir production forecasts using upscaled flow models. Proc. of 7th International Geostatistics Congress. Banff, Canada, (2004) 26 September–1 October
12. Martinez, W.L., Martinez, A.R.: Computational Statistics Handbook with Matlab. Chapman & Hall/CRC, Boca Raton, LA (2002)
13. Mishra, S., Choudhary, M.K., Datta-Gupta, A.: A novel approach for reservoir forecasting under uncertainty. SPE Reserv. Evalu. Eng. 42–48 (2002) February
14. Omre, H., Lødøen, O.P.: Improved prediction forecasts and history matching using approximate fluid flow simulators. SPE J. 339–351 (2004) September
15. Onwunalu, J.: Optimization of nonconventional well placement using genetic algorithms and statistical proxy. MSc thesis, Stanford University (2006)
16. Ripley, B.D.: Pattern Recognition and Neural Networks. Cambridge Univ. Press (1996)
17. Scheevel, J.R., Payrazyan, K.: Principal component analysis applied to 3D seismic data for reservoir property estimation. SPE Reserv. Evalu. Eng. 64–72 (2001) February
18. Wolfsteiner, C., Durlofsky, L.J., Aziz, K.: Approximate model for productivity of nonconventional wells in heterogeneous reservoirs. SPE J. 218–226 (2000) June
19. Yeten, B.: Optimum deployment of nonconventional wells. PhD thesis, Stanford University (2003)
20. Yeten, B., Durlofsky, L.J., Aziz, K.: Optimization of nonconventional well type, location and trajectory. SPE J. 200–210 (2003) September