Well rates and location optimization considering genetic algorithms and surrogate models

Eduarda de F. Andrade[1], Silvana M. B. Afonso[1], Leonardo C. de Oliveira[1], Jeferson W. O. Pinto[1]

[1]*Dept. of Civil Engineering, Federal University of Pernambuco*
Prof. Moraes Rego Avenue, Recife, 50.740-540, Pernambuco, Brazil
eduardafandrade@gmail.com, silvana.bastos@ufpe.br, leonardo.oliveira@ufpe.br, wellano@gmail.com

Abstract. In this work, we solve an optimization problem of a known reservoir of the literature through an integrated optimization by a Genetic Algorithm (GA). The location of the wells and their flow rates are the variables. The main objective of this paper is to maximize the net present value (NPV). The optimization utilized the GA from Toolbox Optimization MATLAB to define, simultaneously, the best position for the wells and best flow rates for each well in each of the three defined control cycles. During the optimization process was performed a series of function evaluations using a Reservoir Simulator. Due to the high cost of this process and aiming to avoid it, a methodology with adaptive surrogate models was employed here. As the optimization problem is restricted, using an adaptive penalty method allowed the GA to run smoothly. The Egg Model is the study reservoir in this work. It was executed 20 optimizations to verify the uniformity of the obtained results. The best solution improved the NVP by 45.32% as compared with the original case. The methodology suggested here brought consistent results with significant improvements in the NPV, the main objective of this paper.

Keywords: optimization, surrogate models, genetic algorithm, well placement and flow rate, reservoir engineering

1 Introduction

The reservoir engineering applications include considerable challenges that are of interest to the Oil & Gas Industry. Between them, the optimization of flow rates of the wells and their locations, study objects here. Both problems have their particularities and involve, generally, complex models with high computational costs. Therefore, it is necessary to develop methodologies that enable the reduction of these costs. Because of this, surrogate models were employed to reduce the CPU time cost in the optimization process. The genetic algorithm from Toolbox Optimization MATLAB [2] was used to optimize the position of the wells and their flow rates, modifying the initial configurations to both parameters. Furthermore, genetic algorithms allow working with efficiency and fast response, dealing with the optimization problem and all its complexities. The well placement problem, for example, can be quite complex due to all possible solutions, multiple scenarios that can suffer influence by a series of reservoir characteristics, e.g., the reservoir size and the number of wells arranged in the field. Thus, the main objective here is to maximize the fields' economic return through the NPV, defining the rates of producer and injector wells and the best place for the wells aiming for optimum recovery to the studied reservoir.

The principal reference for this paper is the study developed by Redouane et al. [1] about well placement optimization in fractured reservoirs utilizing a genetic algorithm (GA). Here, the methodology applied by Redouane was employed and adapted through a new approach where the location and flow rates are optimized simultaneously using the GA from Toolbox Optimization MATLAB. Previous studies have been developed about well allocation optimization or flow rates optimization with GA. Hamida et al. [3] investigated the best location for wells using a modified GA approach in oil fields. Ariadji et al. [4] studied a combinate technique with a modified GA and artificial neural network in different stages to optimize the location of horizontal wells. Onuh et al. [5] presented a study where a developed algorithm with java determines the best place for wells considering the reservoir permeability, fluid saturation, and pay zone thickness. Dada et al. [6] showed a new methodology that determines the best control settings and best well allocation through Toolbox Optimization MATLAB to all optimization performed.

Redouane et al. [1], in their study, indicated a way to work with surrogate models limiting the budget of simulations. They also explored promising regions around the best solution from the optimizer. This last method is called Minimizing an Interpolating Surface (MIS) and was presented by Jones [7]. All these techniques were applied and adapted in this work.

2 Mathematical formulation

The optimization occurs in an integrated process. Simultaneously, the location and flow rates of the wells are optimized by the Genetic Algorithm from Toolbox Optimization MATLAB to maximize the net present value (NPV) of the reservoir. The objective function values, NPV, are defined through numerical simulations using the IMEX simulator by Computer Modelling Group LTD (CMG) [8]. The lower and upper reservoir limits through their spatial coordinates, a minimum distance between wells, and the verification of the activity of the cells are the constraints of the location problem. About the last constraint, an index associated with each cell defines its activity or inactivity. It is used a centered block scheme where the coordinates i and j represent the topographic position of the wells in the reservoir grid. These variables must belong to the integers.

The non-full capacity operation (NCO) formulation by Horowitz et al. [9] was employed to optimize flow rates. The concession time of the studied reservoir was divided into three control cycles with fix times. Each cycle has limits defined by the division between the daily limit of the wells and the group's production capacity where they are connected. Also, the summation of the wells' production and injection must respect the ceiling limit of their respective platforms. These are the constraints imposed on the flow rates problem. The time of the three control cycles was determined as follows: $T_{cc1} = 0$ days (at the beginning), $T_{cc2} = 600$ days and $T_{cc3} = 2010$ days.

The formulation of the integrated optimization problem for well placement and flow rates is presented by eq. (1) as follows:

$$
\begin{aligned}
&\text{Maximize } NPV = f(\boldsymbol{x}) = \sum_{t=1}^{n_t}\left[\frac{1}{\left(1+d\right)^{\tau_t}} \cdot F(\boldsymbol{x})\right] \\
&\text{Subject to: } x_{i,w} \le n_{cI} \quad w = 1...n_w \quad x_{i,w} \in \mathbb{N}^* \\
&\qquad x_{j,w} \le n_{cJ} \quad w = 1...n_w \quad x_{j,w} \in \mathbb{N}^* \\
&\qquad \left\|\left(x_{w_a}, x_{w_b}\right)\right\| \ge dist_w \quad a,b = 1...n_w \quad \text{with } a \ne b \\
&\qquad \prod_{w=1}^{n_w}\left(\prod_{block=1}^{n_{cK}} id_{block}\right) = 1 \\
&\qquad id_{block} = \begin{cases} 0, & \text{for inactive blocks} \\ 1, & \text{otherwise} \end{cases} \qquad block = (x_{i,w}, x_{j,w}, x_{k,w}) \quad k = 1...n_{cK} \\
&\qquad \sum_{w\in P} x_{t,w} \le 1, \quad t = 1...n_t \\
&\qquad \sum_{w\in I} x_{t,w} \le 1, \quad t = 1...n_t \\
&\qquad x_{t,w}^{lb} \le x_{t,w} \le x_{t,w}^{ub} \quad w = 1...n_w \quad t = 1...n_t \\
&\qquad \text{with } x_{t,w} = \frac{q_{t,w}}{Q_{l,\max}},\ w \in P \ \text{ and } \ x_{t,w} = \frac{q_{t,w}}{Q_{inj,\max}},\ w \in I.
\end{aligned} \tag{1}
$$

Where $f(\boldsymbol{x})$ is the objective function (net present value); the design variable vector is represented by $x = \left[x_1, x_1, ..., x_w, ..., x_{n_w}, x_{t,1}, x_{t,2}, ..., x_{t,w}, ..., x_{n_t,n_w}\right]$, where $x_w = \left[x_{i,w}, x_{j,w}\right]$, i.e., i and j are design variables represented by coordinates of the wells and define the block position at the grid with x_w belonging to naturals, excluding zero ($\mathbb{N}^*$). The cash flow F at control cycle t is given by Horowitz et al. [9], the discount rate applied to capital is represented by d, the τ_t is the time at the end of the tth control cycle. n_{cI}, n_{cJ} and n_{cK} are, in this order, the maximum number of blocks in the i, j and k directions of the reservoir grid; n_w is the total number of the wells on the reservoir; $dist_w$ is a constant that establishes the minimum Euclidean distance between the well a and any well b; id_{block} is an index that defines if a block is located in an active or inactive cell; $x_{i,w}$, $x_{j,w}$ and $x_{k,w}$ are integers and real. In this study, for $x_{k,w}$ it is considered that the drilling of the well is done up to the n_{cK} limit keeping a topographic optimization as a general characteristic. $x_{t,w}$ is the vector part of the wells rates for all control cycles, n_t is the total number of control cycles. The constraints of the flow rates problem are defined by the lower and upper limits for each well described as $x_{t,w}^{lb}$ and $x_{t,w}^{ub}$, respectively. $q_{t,w}$ is the flow rate of the well w (producer P or injector I), $Q_{l,\max}$ and $Q_{inj,\max}$ are the maximum total production flow of liquids (oil and water) and the maximum total injection flow of water allowed for the platform, respectively.

3 Surrogate models

In the face of the need to reduce the high costs, risks, and time involved in computational simulations, surrogate models are an alternative to overcome all these problems inherent in high-fidelity simulations. Schmit and Farshi [10] showed that surrogate models can be very efficient in solving some complex engineering problems as an optimization process. The substitutive functions can represent the physical problem bringing a fast response by simplifying the high-fidelity functions' behavior, still enabling to getting their gradients and free of numerical noises.

3.1 Constructing the surrogate model

A surrogate model is constructed from the generation of sampling points. Here, it was used the Design of Experiments (DoE) to determine the input values to the project space. The creation of an appropriated substitutive model requires a sampling technique. Here, it was employed the Latin Hypercube Sampling (LHS), developed by Romero et al.[11], due to its uniform spread of points.

Different methods can be used to construct a surrogate model. They are organized into two groups: functional and physical. Here, the functional category was employed with the fitting data technique and the Radial Basis Function (RBF) as the adjustment model. Foroud et al. [12] compared the quadratic, radial and multiplicative model functions and concluded that the radial had the best response to the optimization process through genetic algorithms.

In problems of constrained optimization, a part of the design space can not be feasible. Oliveira [13] indicates that about 60% of the individuals from the initial population must be feasible. In this work, this recommendation is followed. The initial size of the LHS technique is 5*n*, *n* is the number of variables of the problem. Here, the total of variables is 60 (24 for location and 36 for flow rates). This results in sampling with 300 points. New groups of points are inserted during the optimization process where the model is enriched and its accuracy is increased.

4 Genetic Algorithm

An integrated approach was built to maximize the NPV. A genetic algorithm was used considering an adaptive surrogate model method to obtain the best solution for the location of the wells and their injection and production rates. The GA is characterized as a robust heuristic optimization technique, an efficient resource for dealing with constrained problems. The GA from the Toolbox Optimization MATLAB [2] has its convergence options. Here, to global search from GA, the stopping criteria adopted were the maximum number of generations, equal to 100, and the minimum changes in the objective function value of the population with each improvement. The tolerance related to this last criterion is $1e^{-6}$.

4.1 Constraints manipulation

In this paper, two constraints manipulation methods were considered. Chromosome repairing adapted from Oliveira [13] and adaptive penalty function from Lemonge and Barbosa [14]. They are used at distinct moments in the optimization process. The first one is applied due to the demand of feasible individuals in the GA's initial population in front of the problem constraints and employed at the initial search stage to correct the genes of the created individuals. The last one uses an adaptive scheme with population data, as objective function average and the level of violation of each constraint in the optimization process, instead of defining penalization parameters. It is used in the rest of the optimization process.

4.2 Adaptive surrogate reservoir model (ASRM)

In this paper, the updating of the substitutive model is made through perturbations around the best solution with the creation of 4 new points. Such perturbations affect the coordinates and flow rates based on the best result for well placement. After that, these 4 points are evaluated in the objective function by the simulator and added to the substitutive model, updating it with each new iteration. This methodology is based on Redouane et al. [1]. He established a new approach in the optimization process to create and enrich the surrogate model defining a budget of iterations (i) for the simulations. Here, the method is improved by making new points from perturbs in the optimizer's best solution using a technique that modifies some wells' location and flow rate for each point created. The ASRM procedure by GA for this paper is described in the steps below.

1. Creation of the surrogate model with $5n$ size.
2. Surrogate-based optimization by GA to find the best solution $(xstr_{ga})$.
3. Evaluation of $xstr_{ga}$ in the simulator $(MAXsurg)$. Check convergence. If convergence is ok, $BESTmax$ and $BESTpnt$ are updated.
4. Definition of a dynamic domain (D) with 4 different points generated based on the best solution found by GA. These points are created around the $xstr_{ga}$ employing random perturbations of random wells. The process is made in two steps: coordinates, first and flow rates in sequence. Initially, 5 wells from the best individual are selected randomly. Individually, it is verified the Euclidean distance (ED) of the well position to be perturbed to the center of the reservoir. The Egg model reservoir has a box shape encompassing active and inactive cells. Then, a circular influence region was considered to standardize the perturbation movements. The region diameter (d) is equivalent to reservoir half size. The size movement is equal to one or two blocks that is defined randomly. So, if the ED is smaller than the defined radius $(d / 2)$, the coordinate movement is done to the borders. Otherwise, move toward the center. About the flow rate perturbations, the same wells previously chosen have their rates modified following lower and upper limits for any changes.
5. Evaluation of the 4 new individuals at the simulator and creation of the group L_2 containing their respective objective function values.
6. The enrichment of the surrogate model with 5 points: the best solution by GA $(xstr_{ga})$ and the individuals from domain D . These points are inserted in initial population for the next search.
7. The loop (L_2) is applied *i* times until the convergence or the maximum number of iterations $(i_{\max})$ be reached. Define $i = 1$ at the beginning of the optimization process and increment every time that passes by the loop as $i = i + 1$.
8. The previous step (step 7) is repeated until the loop finishes by one of the stopping criteria. In this paper, the stopping criteria defined are the maximum number of iterations $(i_{\max} = 60)$, representing a total number of function evaluations equal to 300 ($5n$) for model enrichment, and the convergence represented by eq. (2). Then, the best individual is the optimized solution.

$$convergence = \frac{MAXsurg - fit(GA)}{MAXsurg} < 1e^{-6}. \tag{2}$$

5 Egg Model: discussion and results

The Egg Model, a benchmark reservoir proposed by Jansen et al. [15], was used in this study to validate the methodology presented here. The main features of this reservoir are the two-phase flow (oil and water) and the number of wells defined by 8 injectors and 4 producers, totalizing 12 vertical wells as showed in Figure 1. The grid is modeled with 60 by 60 by 7 representing the number of blocks in *i*, *j*, *k* directions, respectively. Thereby, in this model, there are 25,200 cells where 18,533 are active. The concession time to the exploitation of the reservoir is 3.600 days. The values of water injection cost, water produced costs, and produced oil price used by Siraj et al. [16] were employed as reference economic parameters in this paper.

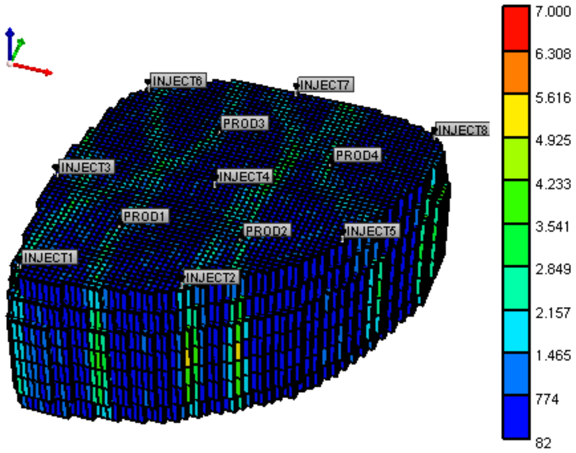

Figure 1. Egg model with permeability and initial well placement scheme (PROD: production well, INJECT: injection well).

The optimization problem formulation considered 2 variables of location for each well and 1 variable per well for each of 3 control cycles, resulting in a problem with 60 variables (24 for location and 36 for flow rates). In this paper, through the Genetic Algorithms from Toolbox Optimization MATLAB, were executed 20 optimizations runs. The size of the surrogate model, without adaptation, is 5*n*, and the size of the enrichment is 5*n*, *n* is the number of variables of the problem.

The literature previously described the initial position of the wells, representing the first part of the variables vector. One simulation with the initial layout of the wells, where the simulator controlled all the flow rates internally, defined the initial flow rate values. From this, the values corresponding to each control cycle at their respective times were extracted. The results of optimizations modified the original positions of the wells and their initial flow rates. Table 1 shows the flow data used for the base case of all runs and the optimized results of the best solution found. Figure 2 presents the best solution for well location through superposition, comparing the base and optimized cases. The well placement of producers (on the left) and injectors (on the right) are represented below.

Table 1. Flow results (m3/day) – Base case and Optimized case (best solution)

Case	Base			Optimized		
Control Cycle	1	2	3	1	2	3
P1	120.0000	120.0000	120.0000	60.7202	69.5054	81.8303
P2	84.9019	64.8176	98.0790	8.8157	54.0469	114.2730
P3	103.1475	95.1824	98.5156	110.4510	3.0747	13.2759
P4	91.9505	120.0000	83.4055	117.3860	3.2801	91.8588
I1	79.5000	79.5000	0.0000	21.6698	65.4207	51.4735
I2	79.5000	73.7037	79.5000	7.2299	6.4296	36.2244
I3	48.5690	41.0363	55.7842	64.0096	25.2547	26.7262
I4	39.3568	63.3590	62.9914	60.7281	60.7793	7.3577
I5	54.0223	39.4337	37.3381	5.7259	75.9495	2.7499
I6	30.8644	0.0000	79.5000	20.9846	27.4394	20.3574
I7	64.0231	78.1898	63.3306	60.3074	34.7712	72.2169
I8	44.1643	29.9599	21.8608	33.2188	27.9533	54.9812

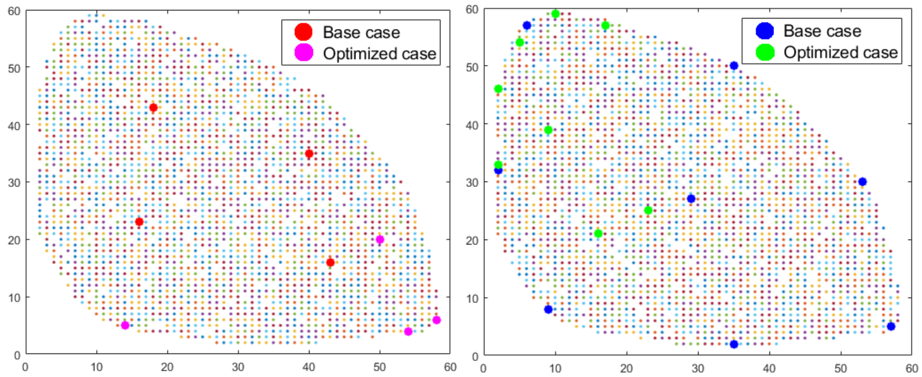

Figure 2. Well positions for producers (on the left) and injectors (on the right).

The best solution to an integrated optimization, well placement and flow rates simultaneously, by GA improved the NPV by 45.32% related to the base case. The recovery methodology employed here is the main reason for the water decrease that directly impacted the NPV increase. This reduction of produced water can be noticed through the cumulative water and water cut parameters. This reduction of produced water is noticed through the cumulative water and water cut parameters in Figure 3. The cumulative water of the optimized case presented a drop of about 80.15% over the 10 years of exploitation time compared to the base case. The water cut of the optimized case (86.62%) had a decline of about 9.03% as compared with the base case (95.65%). This decline is due to the optimized water cut rate was kept at a lower level for a significant time, especially at the beginning of the simulation. It is fundamental to mention the late water breakthrough that was crucial to best results in the optimization process.

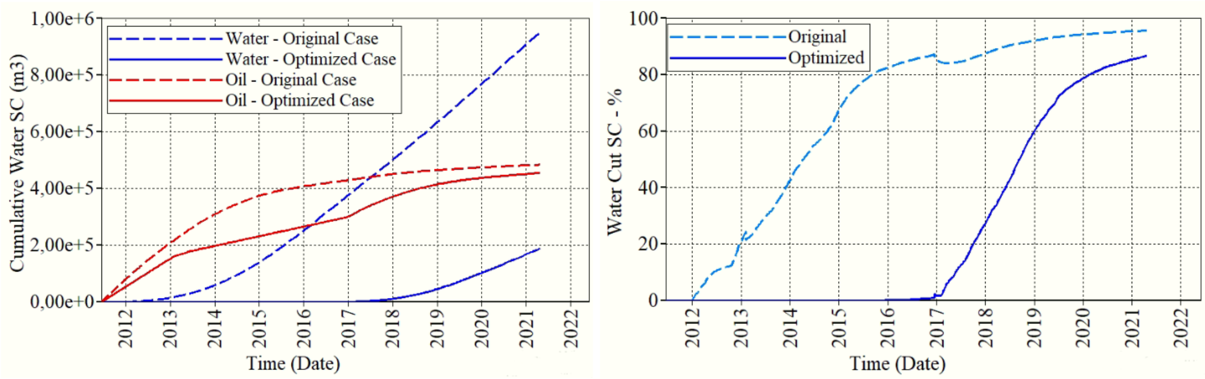

Figure 3. Water and oil cumulative production (on the left) and the water cut (on the right) for the original case (hidden lines) and the optimized case (full lines) to Egg model.

The statistical results of all optimization runs are organized in Table 2. It shows the improvement details by GA.

Table 2. Optimization results summary

Number of optimizations	Best solution (%)	Worst solution (%)	Average improvement wrt base case (%)	Standard deviation (%)	Number of objective function evaluation (best solution)	Number of objective function evaluation (average)
20	45.32%	29.88%	35.82%	4.70%	601	601

6 Conclusions

The Genetic algorithm from Toolbox Optimization MATLAB was used in this paper aiming to maximize the Net Present Value (NPV) of the Egg Model reservoir. The adopted methodology involved jointly optimizing the position and flow rates of the wells, resulting in new well placement and new flow rates for all the control cycles defined. The results of optimization demonstrated satisfactory improvements in the NPV for all of the 20 runs. The best solution improved the NPV by 45.32% as compared with the original case. The average increase is 35.82%. The standard deviation of the optimizations is 4.70%, a low value that demonstrates homogeneous and consistent results in the study developed in this paper. The excellent results of NPV improvement are due to the primary strategy adopted for oil recovery that occasioned a substantial decline of the produced water over the exploitation time of the reservoir. As previously mentioned, this decrease can be noticed through the results of water cut and cumulative water. The first parameter showed a difference of about 9.03% between the base case (95.65%) and the original case (86.62%). The cumulative produced water was reduced by approximately 80.15% in the optimized case. Concerning the accumulated oil, it was noticed a decrease in production. This is not the best scenario for the Oil & Gas Industry because one of their main interests is to maximize oil recovery and not just increase the reservoir's Net Present Value. One suggestion to improve these results is considering a strategies combination, e.g., a global strategy sequentially with a local. The methodology suggested here brought consistent solutions with significant improvements in the NPV, the main objective of this paper.

Acknowledgements. We thank the Support Research Foundation of Pernambuco (FAPECE) for the encouragement and financial assistance that provided this work's development.

Authorship statement. The authors hereby confirm that they are the sole liable persons responsible for the authorship of this work, and that all material that has been herein included as part of the present paper is either the property (and authorship) of the authors, or has the permission of the owners to be included here.

References

[1] K. Redouane, N. Zeraibi, e M. Nait Amar, “Adaptive surrogate modeling with evolutionary algorithm for well placement optimization in fractured reservoirs”, *Applied Soft Computing*, vol. 80, p. 177–191, jul. 2019.

[2] Mathworks, Inc., Natick, Massachusetts, 2018.

[3] Z. Hamida, F. Azizi, e G. Saad, “An efficient geometry-based optimization approach for well placement in oil fields”, *Journal of Petroleum Science and Engineering*, vol. 149, p. 383–392, jan. 2017.

[4] T. Ariadji, F. Haryadi, I. T. Rau, P. A. Aziz, e R. Dasilfa, “A novel tool for designing well placements by combination of modified genetic algorithm and artificial neural network”, *Journal of Petroleum Science and Engineering*, vol. 122, p. 69–82, out. 2014.

[5] C. Y. Onuh, D. Alaigba, O. J. Rotimi, e B. T. Arowolo, “Well Placement Optimization Using a Basic Genetic Search Heuristics Algorithm and a Black Oil Simulator”, *OJOGas*, vol. 02, nº 04, p. 214–225, 2017.

[6] M. A. Dada, M. Mellal, A. Makhloufi, e H. Belhouchet, “A field development strategy for the joint optimization of flow allocations, well placements and well trajectories”, *Energy Exploration & Exploitation*, vol. 39, nº 1, p. 502–527, jan. 2021.

[7] D. R. Jones, “A Taxonomy of Global Optimization Methods Based on Response Surfaces”, *J. Global Optim.*, vol. 21, p. 345–383, 2001.

[8] COMPUTER MODELLING GROUP LTD, *IMEX User Guide 2016.10: Three-phase, Blackoil Reservoir Simulator*. 2016.

[9] B. Horowitz, S. M. B. Afonso, e C. V. P. de Mendonça, “Surrogate based optimal waterflooding management”, *Journal of Petroleum Science and Engineering*, vol. 112, p. 206–219, dez. 2013.

[10] L. Schmit, Jr. e B. Farshi, “Some approximation concepts for structural synthesis”, apresentado em 14th Structures, Structural Dynamics, and Materials Conference, Williamsburg,VA,U.S.A., mar. 1973.

[11] V. J. Romero, J. V. Burkardt, M. D. Gunzburger, e J. S. Peterson, “Comparison of pure and ‘Latinized’ centroidal Voronoi tessellation against various other statistical sampling methods”, *Reliability Engineering & System Safety*, vol. 91, nº 10–11, p. 1266–1280, out. 2006.

[12] T. Foroud, A. Seifi, e H. Hassani, “Surrogate-based Optimization of Horizontal Well Placement in a Mature Oil Reservoir”, *Petroleum Science and Technology*, vol. 30, nº 11, p. 1091–1101, abr. 2012.

[13] L. C. de Oliveira, “Hybrid strategy applied to optimal reservoir management (in Portuguese)”, Doctoral Dissertation, Graduate Program in Civil Engineering, Federal University of Pernambuco, Recife, Brazil, 2013.

[14] A. C. C. Lemonge e H. J. C. Barbosa, “An adaptive penalty scheme for genetic algorithms in structural optimization”, *Int. J. Numer. Meth. Engng.*, vol. 59, nº 5, p. 703–736, fev. 2004.

[15] J. D. Jansen, R. M. Fonseca, S. Kahrobaei, M. M. Siraj, G. M. Van Essen, e P. M. J. Van den Hof, “The egg model - a geological ensemble for reservoir simulation”, *Geosci. Data J.*, vol. 1, nº 2, p. 192–195, nov. 2014.

[16] M. M. Siraj, P. M. J. Van den Hof, e J. D. Jansen, “Handling risk of uncertainty in model-based production optimization: a robust hierarchical approach”, *IFAC-PapersOnLine*, vol. 48, nº 6, p. 248–253, 2015.