Journal of Petroleum Science and Engineering 169 (2018) 560–577

Contents lists available at ScienceDirect

Journal of Petroleum Science and Engineering

journal homepage: www.elsevier.com/locate/petrol

A new methodology for history matching combining iterative discrete Latin Hypercube with multi-start simulated annealing

Célio Maschio*, Denis José Schiozer

Center for Petroleum Studies - CEPETRO, University of Campinas - Unicamp, Campinas, SP, Brazil

ARTICLE INFO

Keywords:
Reservoir simulation
History matching
Iterative Discrete Latin Hypercube
Simulated annealing

ABSTRACT

This paper introduces a new method for history matching, combining the Iterative Discrete Latin Hypercube (IDLHC) with multi-start Simulated Annealing methods. The proposed method, named IDLHCSA, combines the potential of the IDLHC in finding good matched models while preserving the diversity of solutions and the potential of the SA with local control in finding local (refined) solutions. The IDLHCSA was applied in two cases. The first is a simple reservoir model used as proof of concept. The second is a complex benchmark case (UNISIM-I-H) based on the Namorado field, located at the Campos Basin, Brazil. The robustness and effectiveness of the proposed method are demonstrated by comparison with other consolidated methods. It is demonstrated here that, when compared to other methodologies, the proposed method is more effective in finding multiple solutions for the history matching problem while maintaining solution diversity. The production forecast is analyzed and the predictive capacity of the matched models is assessed. The paper reveals that obtaining good matched models does not ensure reliable forecasts. The proposed method was able to find matched models which provided more reliable forecasts when compared to other methods.

1. Introduction

History matching involves the use of dynamic observed data to enhance the estimation of reservoir properties (attributes), such as porosity and permeability, to improve the quality of reservoir simulation models. The objective of the process is to improve the predictive capacity of reservoir models. However, finding the best matched models may not ensure reliable forecasts. Due to issues such as uncertainties in observed data (due to measurement errors for example), insufficient constraints and insufficient data, history matching is an ill-posed problem. This means that many possible combinations of attributes may result in equally good history-matched models.

The key steps of the history matching process are: (1) parameterization, which consists of defining the uncertain attributes and limits and (2) exploration of the search space, composed of the combination of the uncertain attributes, which can be carried out by an optimization or a sampling method. The effectiveness of the process depends on these two fundamental steps. There are two main aspects related to these steps: firstly, if the parameterization is wrong, no methods find correct solutions because they do not exist in the search space. Thus, the parameterization is a crucial step of the process. Secondly, if the parameterization is correct, in the sense that the search space contains good solutions (matched models according to a tolerance), the effectiveness of a given method is defined by its ability of finding these solutions. However, finding good matched models is not sufficient. A given method is really effective if it is able to find multiple solutions in disconnected local minimum regions in a search space.

The main objective of this paper is to present a new and effective methodology for history matching which is capable of finding multiple solutions by combining Iterative Discrete Latin Hypercube (IDLHC) sampling and multi-start simulated annealing with local control. The comparison of the proposed method with other well-known methods is another objective of this paper.

2. Iterative discrete Latin Hypercube (IDLHC)

Latin Hypercube (LHC), introduced by McKay et al. (1979), is a sampling technique that divides the range of input variables in n intervals of equal marginal probability $1/n$ and generates random values from each interval. The number of values from each interval is proportional to the probability corresponding to the interval. LHC ensures that the sampling will select values over the whole range of the input variables.

Discrete Latin Hypercube (DLHC) follows the same idea of the standard LHC. The main difference is that DLHC considers discrete

* Corresponding author.
E-mail addresses: celio@cepetro.unicamp.br (C. Maschio), denis@unicamp.br (D.J. Schiozer).

https://doi.org/10.1016/j.petrol.2018.06.004
Received 23 November 2017; Received in revised form 28 May 2018; Accepted 2 June 2018
Available online 05 June 2018
0920-4105/

random variables (each variable is divided into a number of uncertain levels). One advantage of the discrete version of the LHC is the possibility of working with categorical variables. In this sense, Schiozer et al. (2017) proposed a specialized version called DLHG (Discretized Latin Hypercube combined with Geostatistical realizations) applied in UNISIM-I-D benchmark case (described in Gaspar et al., 2015) for risk quantification.

In the context of history matching, Monfared et al. (2014) applied the standard Latin Hypercube method to find initial guesses which were used as starting points for the Levenberg-Marquardt algorithm. Goda and Sato (2014) also used this technique, proposing a search algorithm named Iterative Latin Hypercube Sampling (ILHS). More recently, Tanaka et al. (2017) applied the ILHS method to assisted history matching using real field data from the CO_2 storage Iwanohara project. Two cases of history matching including or excluding porosity distribution were conducted. The researchers concluded that the proposed method allowed them to optimize the reservoir model for CO_2 geological storage with limited calculation resources.

The IDLHC (Iterative Discrete Latin Hypercube) method, which is applied in this paper in conjunction with simulated annealing, was proposed by Maschio and Schiozer (2016). The method consists of applying the DLHC method in successive iterations. After each iteration, a percentage of models is selected based on the history-matching quality and a histogram is generated for each attribute. New probabilities are computed proportionally to the frequency of each uncertain level of each attribute. The probabilities computed for each attribute in a given iteration are used as input for the generation of new models in the next iteration and this iterative process continues until a stop criterion is reached (normally, the number of iterations). The selection of the best models in each iteration allows for the improvement of the probability redistribution, narrowing the search space as the sampling process is gradually intensified in the regions of high probability. To avoid discontinuities in the probability distribution, a procedure involving kernel density estimation is applied to smooth the histograms.

Maschio and Schiozer (2016) proposed a method to select the best models base on a local objective function (LOF), which is composed of influenced functions identified by means of a correlation matrix. The correlation matrix captures the influence of all attributes in all objective functions. A cut-off value (Rc) is used to select the influenced functions based on the correlation coefficient. The models are selected based on these functions. Maschio and Schiozer (2016) demonstrated that values of Rc which indicate moderate correlation can be adequate.

If a given attribute does not influence any function following the criterion based on Rc, its probability distribution is not changed in the current iteration. More details about the IDLHC method can be found in Maschio and Schiozer (2016).

3. Simulated annealing

3.1. Basic description

Simulated annealing (SA) is a stochastic optimization method inspired by the process of physical annealing in solids. The original idea of SA as an optimization method was proposed by Kirkpatrick et al. (1983). According to the authors, the simulated annealing process consists of first "melting" the system being optimized at a high temperature, then slowly reducing the temperature until the system "freezes" and no further changes occur. The term 'temperature' is used for its analogy with the physical process. The key feature of SA is that it provides a means of escaping from local minima by allowing hill-climbing moves.

A basic simulated annealing algorithm is shown in Fig. 1. Note that, if a proposed candidate has a smaller objective function (OF) compared to the current point, *i.* e, if it improves the current solution, it is always accepted. However, the algorithm also accepts uphill move (candidate with OF worse than the current point) with a certain probability. The acceptance or rejection of an uphill move is determined by the comparison of a random number between 0 and 1 (generated from a uniform distribution) with an exponential function of -ΔOF/T, as shown in Fig. 1.

Define an initial point m_0
Define an initial temperature $T_i > 0$
Set temperature change counter t = 0
Repeat
j = 0
    Repeat
        Generate $\tilde{m}_{j+1}$ neighbour of m_j
        Compute $\Delta OF = OF(\tilde{m}_{j+1}) - OF(m_j)$
        If $\Delta OF < 0$ then
            Accept $\tilde{m}_{j+1}$
        else
            Generate u = random(0,1)
            if $u < e^{-\Delta OF/T}$ accept $\tilde{m}_{j+1}$ end if
        end if
        j = j + 1
    Until j = TL
t = t + 1
Set new temperature T = T(t)
Until stopping criterion

Fig. 1. Standard simulated annealing algorithm (adapted from Eglese, 1990).

Conventional (standard) SA starts with a relatively high value of T in an attempt to avoid being prematurely trapped in local minima. The higher the value of T, the higher the acceptance probability. The value of T is gradually dropped until a very low value (close to zero). At each level of T, a certain number of iterations is done. At very low values of T, most uphill moves are rejected.

The main control parameters of a simulated annealing algorithm are: (1) the temperature, which involves the definition of an initial temperature and a cooling schedule comprising a temperature length (number of iterations at a given temperature) and a cooling ratio (rate at which the temperature is reduced); (2) the neighborhood characteristic and (3) stopping criterion, common in any optimization method. As stated by Goldstein and Waterman (1988), the performance of the simulated annealing algorithm critically depends on the choice of neighborhood characteristic. If the neighborhoods are too large, then the process essentially performs a random search. On the other hand, if the neighborhood is too small, the simulated process will not be able to move efficiently over the search space. In general, the control parameters are problem-dependent and some experimental tuning is necessary for each case. Aspects regarding problem specific choices and generic choices are discussed by Eglese (1990) and Nikolaev and Jacobson (2010).

The drawback of conventional SA is the high number of iterations necessary to guide the search toward good solutions, which can be prohibitive in problems where the evaluation of the OF requires extensive computational time, which is normally the case in history-matching applications. In this paper, we overcome this drawback by using multi-start SA (several instances of SA with different initial points) with local control and by using the benefits of parallel computing.

3.2. Application in inverse problems

Ouenes et al. (1992) applied SA to enhance gas reservoir