Year 12, Number 6
114th edition
May, 2017

UNISIM ON-LINE

Bayesian History Matching Using Parallel Interacting Markov Chain Monte Carlo

Célio Maschio

Introduction

This text is a brief compilation of a paper published in Inverse Problems in Science and Engineering (Maschio and Schiozer, 2017), which proposes a new methodology for Bayesian history matching (BHM). BHM is the manner of treating the history matching problem in a formal probabilistic point of view (using Bayesian formalism) toward a framework for quantification and reduction of uncertainty in production forecasts.

According to the Bayes' theorem, the posterior probability distribution, p(m|O), is given by:

$$p(m \mid O) = c.p(O \mid m)\, p(m) \qquad (1)$$

where m represents an instance of the parameterized reservoir model, c is a normalization constant, p(m) is the prior distribution and p(O|m) is the likelihood given by:

$$p(O \mid m) = \exp\left[-\frac{1}{2} OF(m) \right] \qquad (2)$$

where

$$OF(m) = \sum_{i=1}^{N} \left(\frac{O_i - S(m)_i}{\sigma_{d,i}} \right)^2 \qquad (3)$$

where O_i is the observed data, $S(m)_i$ is the results obtained using m in the flow simulator, N is the total number of observed data and $\sigma_{d,i}$ is the standard deviation of the observed data errors.

In the above equations, it is assumed that different vectors of observed data (observed water rate in a given well for example) are uncorrelated and that the measurement errors in observed data follow a Gaussian distribution.

Due to high nonlinearities involved in BHM problems, it is generally impossible to express the posterior distribution in a closed form. Thus, it is necessary to apply sampling techniques to solve the problem. Therefore, BHM consists of sampling the posterior distribution represented by Eq. 1.

Markov chain Monte Carlo (MCMC)

MCMC is a sequential sampling method based on Markov chain concept, which is a sequence of random variables that depends on its history only through the previous state. A Markov chain evolves by causing a perturbation around the current state. The new state is accepted or not based on a probabilistic criterion.

MCMC is a robust sampling method able to produce samples of virtually any posterior distribution and has been applied to solve a wide variety of complex practical problems. However, history matching is typically a highly non-linear inverse problem, which leads to very complicated posterior distribution, with several disconnected modes. Thus, sequential chain, the traditional method, can be trapped into local modes.

The objective of this work is to present a new methodology that uses parallel interacting MCMC to solve the BHM problem. The key idea is to start several chains in parallel, each at a different perturbation size (defined by scaled proposal variance, σ), allowing the chains to exchange information through swaps among their current states.

Methodology

The proposed methodology is composed of the following steps:

1) Set the number of parallel chains (M), the number of iteration (or states) per chain (n) and set k = 1.
2) Define M starting points (one for each chain).
3) Generate M reservoir simulation models. For k = 1, M corresponds to the starting points of the chains. For k > 1, M is the set of states (one state per chain) proposed in the Steps 6 and 7.
4) Submit the M reservoir simulation models to run in parallel in a cluster of computers.
5) After the execution of the reservoir simulations, compute p(m|O) for each model.
6) Update each chain independently using the Metropolis-Hastings criterion (within-chain moves) each chain having a specific scaled proposal.
7) At each iteration, draw two integers uniformly distributed between 1 and M (indexes of the chains) and swap the states of the chains associated to the sampled indexes (inter-chain moves).
8) While k is lesser or equal to n, increment k and return to Step 3.
9) Assess the final results.

The key idea is illustrated in Figure 1 using a simple example. The chain with greater variance are responsible for the diversification (scattered sampling) and the chain with small variance are responsible for the intensification (refined sampling). This strategy allows escaping from local modes, exploring efficiently multi-modal posterior distributions.

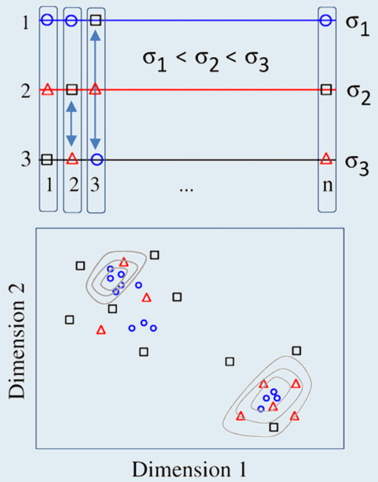
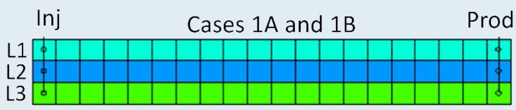


Figure 1: *Illustration of parallel MCMC method (3 parallel chains with n iterations).*

Application and Results

For validation purposes, the methodology was initially applied to a cross-section model (Figure 2). The uncertain parameter is the horizontal permeability of the layers 2 and 3 for the Case 1A and layers 1, 2 e 3 for the Case 1B. For the Case 1A, Kx1 is fixed in 1000 mD. In Maschio and Schiozer (2017) two more cases were studied: Case1C, with 5 layers and Case 2, a realistic reservoir modeled with geostatistical techniques.



K_{xL} = [1000, 600, 2000] mD (L = 1...3)

Figure 2: *Cross-section model.*

"Due to high nonlinearities involved in BHM problems, it is generally impossible to express the posterior distribution in a closed form."

Specials interests:

- UNISIM
- UNISIM Publications
- Reservoir Simulation and Management Portal
- Previous Issues

Links:

- UNICAMP
- Cepetro
- Petroleum Engineering Division
- School of Mechanical Engineering
- Petroleum Sciences and Engineering

Graduate:

Petroleum Sciences and Engineering: interested in Masters and PhD in the Simulation and Oil Reservoir Management area click here.

"We have proposed an innovative application of a new class of parallel interacting Markov chain Monte Carlo to solve the Bayesian history matching problem."

UNISIM opportunities:

If you are interested in working or developing research in the UNISIM Group, please contact us.

Immediate interest in:

- Researcher in the simulation area, management and reservoir characterization.

For further information, click here.

Research in Reservoir Simulation and Management Group

Petroleum Engineering Division - Energy Department
School Of Mechanical Engineering
Center for Petroleum Studies
University of Campinas
Campinas - SP

Phone: 55-19-3521-1220
Fax: 55-19-3289-4916

unisim@cepetro.unicamp.br

Figure 3 shows the surface of the posterior distribution for the Case 1A using discrete values of Kx2 and Kx3. To generate this figure, the ranges of Kx2 and Kx3 were divided into 30 equally spaced values and combined to form a grid with 900 combinations. After running the flow simulator for the 900 models, the value of p(m|O) was computed for each node of the grid. Clearly, we can see four disconnected modes, separated by a region with very low probability (nearly zero).

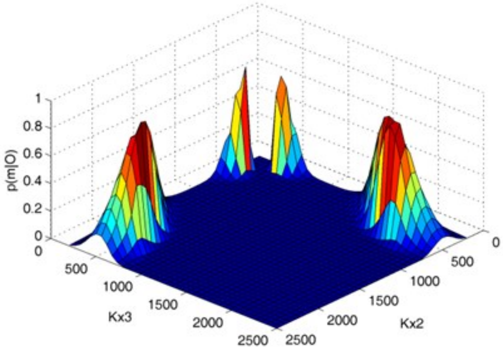


***Figure 3:** Surface of the posterior distribution for the Case 1A.*

Figure 4 shows, in green, the samples accepted during the sampling process (a sample of size 10000 was generated for both sequential and parallel algorithm). In blue, are the models with OF_N (OF_N = OF/N) smaller than 1.0 (the number of models is indicated inside the figure), which means that the model misfit is in the order of one standard deviation.

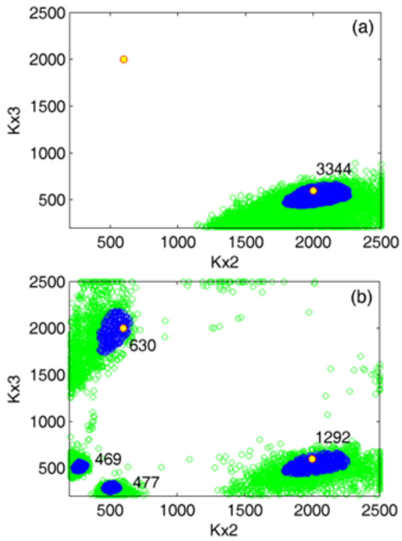


***Figure 4:** Cross plot of Kx2 and Kx3 (Case 1A): (a) sequential MCMC, (b) parallel MCMC.*

Figure 4(a) shows that the sequential algorithm got trapped in one of the local modes. On the other hand, the parallel algorithm sampled all local modes. Figure 5 shows a comparison between the sequential (a) and parallel (b) algorithms for the Case 1B. The blue points represent models with OF_N smaller than 1.0. The yellow points in these plots are permutation of the values of permeability of the three layers that were used to generate the history. As can be seen, there are a huge number of combinations, besides those yellow six points, that provide responses close to the history. It can also be seen that the parallel algorithm covered all regions formed by the connection of the yellow points.

Other details about this work can be found in Maschio and Schiozer (2017).

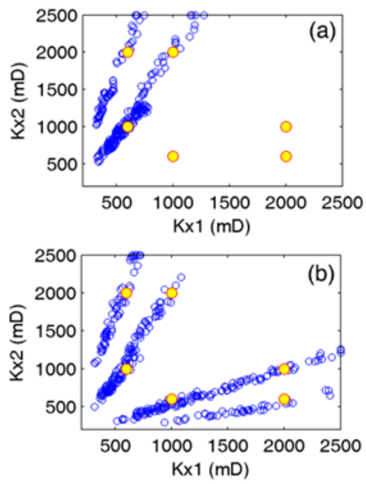


***Figure 5:** Cross plot for Kx1 and Kx2 (Case 1B): (a) sequential MCMC, (b) parallel MCMC.*

Final Remarks

1) In this work, we have proposed an innovative application of a new class of parallel interacting Markov chain Monte Carlo to solve the Bayesian history matching problem.
2) The methodology proposed in this work is well suited to overcome the drawbacks of the traditional sequential MCMC methods, because it allows escaping from local modes and explores efficiently the posterior distribution.
3) This work may contribute to disseminate the use of parallel interacting Markov chains to explore the potential of distributing computing and may also encourage the application of the described method in other cases.
4) Finally, the researches developed by the UNISIM are focused on robust methods capable of finding multiple solutions under probabilistic approaches to reduce uncertainties in reservoir attributes and production forecast. Methods that tend to deterministic solutions are not suitable to deal with this task.

Reference

Maschio, C.; Schiozer, D. J. "A New Methodology for Bayesian History Matching Using Parallel Interacting Markov Chains Monte Carlo", Inverse Problems in Science and Engineering, p. 1-32, 2017. http://dx.doi.org/10.1080/17415977.2017.1322078.

About author:

Célio Maschio is graduated in Mechanical Engineering from UNESP, holds a MS and a PhD degree in Mechanical Engineering from UNICAMP. He is a researcher at UNISIM/CEPETRO/UNICAMP.