Available online at www.sciencedirect.com
ScienceDirect
Mathematics and Computers in Simulation 118 (2015) 224–238
MATHEMATICS AND COMPUTERS IN SIMULATION
www.elsevier.com/locate/matcom

Original articles

# Multi-physics Markov chain Monte Carlo methods for subsurface flows[☆]

V. Ginting[a], F. Pereira[b,*], A. Rahunanthan[c]

[a] Department of Mathematics, University of Wyoming, Laramie, WY 82071, USA
[b] Department of Mathematical Sciences, University of Texas at Dallas, Richardson, TX 75080, USA
[c] Department of Mathematics and Statistics, University of Toledo, Toledo, OH 43606, USA

Received 30 March 2014; received in revised form 18 September 2014; accepted 20 November 2014
Available online 30 December 2014

**Abstract**

In $CO_2$ sequestration in deep saline aquifers, contaminant transport in subsurface, or oil or gas recovery, we often need to forecast flow patterns. In the flow forecasting, subsurface characterization is an important step. To characterize subsurface properties we establish a statistical description of the subsurface properties that are conditioned to existing dynamic (and static) data. We use a Markov chain Monte Carlo (MCMC) algorithm in a Bayesian statistical description to reconstruct the spatial distribution of two important subsurface properties: rock permeability and porosity. The MCMC algorithm requires repeatedly solving a set of nonlinear partial differential equations describing displacement of fluids in porous media for different values of permeability and porosity. The time needed for the generation of a reliable MCMC chain using the algorithm can be too long to be practical for flow forecasting. In this paper we develop computationally fast and effective methods of generating MCMC chains in the Bayesian framework for the subsurface characterization. Our strategy consists of constructing a family of computationally inexpensive preconditioners based on simpler physics as well as on surrogate models such that the number of fine-grid simulations is drastically reduced in the generation MCMC chains. We assess the quality of the proposed multi-physics MCMC methods by considering Monte Carlo simulations for forecasting oil production in an oil reservoir.


*Keywords:* MCMC; Porous media; Multi-physics; Multi-stage

## 1. Introduction

Scientifically correct models based on first principles, accurate numerical simulators and state-of-the-art computers are of utmost importance in producing reliable predictions of multiphase flows in the subsurface. However, uncertainty

☆ This work is partially supported by the grants from US Department of Energy (DE-FE0004832 and DE-SC0004982), the Center for Fundamentals of Subsurface Flow of the School of Energy Resources of the University of Wyoming (WYDEQ49811GNTG, WYDEQ49811PER, WYDEQ49811FRTD), 2011 Clean Coal Technologies Research Program of the School of Energy Resources of the University of Wyoming (1100 20352 2012), and from National Science Foundation (DMS-1016283).
* Corresponding author.
*E-mail addresses:* vginting@uwyo.edu (V. Ginting), luisfelipe.pereira@utdallas.edu (F. Pereira), arunasalam.rahunanthan@utoledo.edu (A. Rahunanthan).

http://dx.doi.org/10.1016/j.matcom.2014.11.023
0378-4754/

in the determination of subsurface properties, such as permeability and porosity, remains the main challenge that one has to overcome to produce accurate predictions.

The typical situation encountered in the investigation of problems such as $CO_2$ sequestration in saline aquifers, contamination of subsurface, or oil recovery is that there is very little data available to characterize subsurface properties. Only sparse data is available: at well locations permeability and porosity (static data) can be measured as well as well pressure values, production curves, and saturation or concentration values at monitoring wells (dynamic data) are usually available. Thus, the problem of determining permeability and porosity fields from available data is not well posed and a stochastic prediction methodology must be considered [23].

We are concerned with the development of a computationally fast and effective Bayesian framework for the characterization of porous media. To characterize subsurface properties we establish a statistical description of the subsurface properties that are conditioned to existing dynamic (and static) data. We use a Markov chain Monte Carlo (MCMC) algorithm in a Bayesian statistical description to reconstruct the spatial distribution of rock permeability and porosity. In this context however, acceptance or rejection criterion in the MCMC algorithm requires the calculation of likelihood function that involves repeatedly solving a set of nonlinear partial differential equations describing displacement of fluids in porous media for different values of permeability and porosity. For each proposal in the MCMC algorithm we compare measured data with its simulated counterpart [27,11,7]. The time needed for the generation of a reliable MCMC chain using the algorithm can be too long to be practical for the subsurface characterization.

The search of computationally efficient MCMC methods has been attracting the attention of several research groups. Fox and Nicholls [13] employed a perturbation technique to identify proposals that are most likely to be rejected in the posterior probability exploration. Oliver et al. [29] were perhaps among the first to apply MCMC in the petroleum reservoir simulations to characterize uncertainty in the permeability fields conditioned to the pressure data. The authors in [15] improved the algorithm in [29] by proposing a blocking MCMC approach in which each proposed member of the chain is different from the previous one over an entire, relatively large, block of cells. This blocking scheme accelerates the creation of the chain of realizations. Higdon et al. [24] presented a methodology for improving the speed and efficiency of an MCMC by combining runs on different scales. They introduced a coarse-scale to make the MCMC chain run faster and better explore the posterior, and linked the coarse chain back to the original fine-scale chain of interest. In [6] a two-stage MCMC algorithm was presented for generating samples from an unnormalized posterior distribution in which the evaluation of the posterior distribution is very difficult or computationally demanding. The algorithm was applied to recovering resistor values in a network from electrical measurements made at the boundary. Later in [11,12], the two-stage MCMC algorithm, which utilizes inexpensive coarse-scale models to screen out detailed flow and transport simulations, was used to explore the posterior distribution of the permeability field. In the first stage, a new proposal is tested at the coarse-scale model. If the proposal passes the testing at the coarse-scale model, then at the second stage the fine-scale simulation will be run and this fine-scale run is computationally very expensive when compared to the coarse-scale run. Vrugt et al. [32] developed a population-based MCMC algorithm that allows for the exchange of information among multiple chains to enhance the efficiency of MCMC sampling. Kavetski et al. [26] and Kavetski and Kuczera [25] showed that some computational complexities can be removed by smoothing the objective function in parameter optimizations of hydrological model calibration in accelerating the Bayesian posterior sampling. A multiple-try MCMC algorithm [34] was designed for a hydrological and environmental simulation model by making better use of the information generated in a costly run of the model. The authors used multiple evaluations of the posterior density in the less computationally expensive subspace of error model parameters.

Furthermore, in the same spirit of two-stage procedures [6,11,12,17], the authors introduced a multi-stage Bayesian prediction framework for subsurface flows in [19]. The authors described a predictive procedure in a Bayesian framework, which uses a single-phase flow model for characterization aiming at making prediction for a two-phase flow model. The quality of the characterization of the underlying formations was accessed through the prediction of future fluid flow production. They also considered parallelizing the generation of MCMC chains to speed-up the posterior exploration in [18,20]. In [18] the authors addressed this issue using several parallel MCMC chains for flow prediction. In [20] the authors parallelized single MCMC chains using a prefetching technique implemented on GPUs and showed that the parallelization can make the Bayesian approach computationally tractable for subsurface characterization and prediction of porous media flows.

The original contribution of the present work is the introduction of the concept of multi-physics MCMC (MP-MCMC) methods. In our strategy a family of computationally inexpensive preconditioners is constructed based