Computers & Geosciences 65 (2014) 110–117

Contents lists available at ScienceDirect

Computers & Geosciences

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

Geothermal model calibration using a global minimization algorithm based on finding saddle points and minima of the objective function

Manuel Plasencia [a,*], Andreas Pedersen [a], Andri Arnaldsson [b], Jean-Claude Berthet [b], Hannes Jónsson [a,c]

[a] Science Institute of the University of Iceland, VR-III, 107 Reykjavík, Iceland
[b] Vatnaskil Consulting Engineers, Sudurlandsbraut 50, 108 Reykjavík, Iceland
[c] Faculty of Physical Sciences, University of Iceland, VR-III, 107 Reykjavík, Iceland

ARTICLE INFO

Article history:
Received 4 March 2013
Received in revised form
8 September 2013
Accepted 11 September 2013
Available online 27 September 2013

Keywords:
Global optimization
Inverse modeling
Reservoir modeling

ABSTRACT

The objective function used when determining parameters in models for multiphase flow in porous media can have multiple local minima. The challenge is then to find the global minimum and also to determine the uniqueness of the optimized parameter values. A method for mapping out local minima to search for the global minimum by traversing regions of first order saddle points on the objective function surface is presented. This approach has been implemented with the iTOUGH2 software for estimation of models parameters. The methods applicability is illustrated here with two examples: a test problem mimicking a steady-state Darcy experiment and a simplified model of the Laugarnes geothermal area in Reykjavík, Iceland. A brief comparison with other global optimization techniques, in particular simulated annealing, differential evolution and harmony search algorithms is presented.

1. Introduction

The development of reservoir models often involves inverse modeling, i.e. an estimation of model parameters by fitting calculated values of the response of the system to measurements at discrete points in space and time. The difference between the model calculation and the measured data at the calibration points can be represented by an objective function of the model parameters. The task of estimating the best set of model parameters is thereby formulated as an optimization problem where the goal is to determine the parameter values that minimize the objective function. Even for models with only a few parameters, the resulting objective function can have more than one minimum. This is illustrated in Fig. 1, which shows a one-dimensional cut of an objective function for a geothermal reservoir model described below. Within the parameter interval shown, three local minima are present. The occurrence of multiple local minima is more likely in models with a larger number of parameters. Hence, the task becomes to find the global minimum of the objective function among several local minima. This is a challenging problem. Furthermore, it is important to know whether additional local minima, with only insignificantly higher objective function values are present since they could, for practical purposes, represent nearly as good parameter values as the global minimum.

Numerical algorithms for optimization can be broadly categorized into local optimization methods and global optimization methods. Local optimization algorithms involve an iterative process where starting from some initial guess, new parameter values are found so as to lower the value of the objective function. Such algorithms only find local minima, typically the local minimum nearest to the initial guess. Typically, local optimization methods rely on the evaluation of the gradient of the objective function. Some examples are steepest descent, conjugate gradient, Quasi-Newton and Levenberg–Marquardt methods. By carrying out multiple minimizations starting from different initial guesses, such methods can be used to find the global minimum but this becomes an inefficient procedure when many parameters are varied.

Global optimization algorithms, on the other hand, attempt to find the global minimum by also allowing the increase of the objective function during the iterative procedure. Some examples are, simulated annealing using Markov chain Monte Carlo methods and evolutionary algorithms such as differential evolutionary, harmony search, and particle swarm optimization. These methods do not make use of the gradient of the objective function and tend to converge more slowly to minima of the objective function, but have the advantage over local optimization methods that they can identify the global minimum. Three of these algorithms will be briefly described here, the simulated annealing, differential evolution and harmony search algorithms. These are implemented in the

* Corresponding author.
E-mail addresses: mpg2@hi.is, manuelplasencia@yahoo.es (M. Plasencia).

0098-3004/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.cageo.2013.09.007

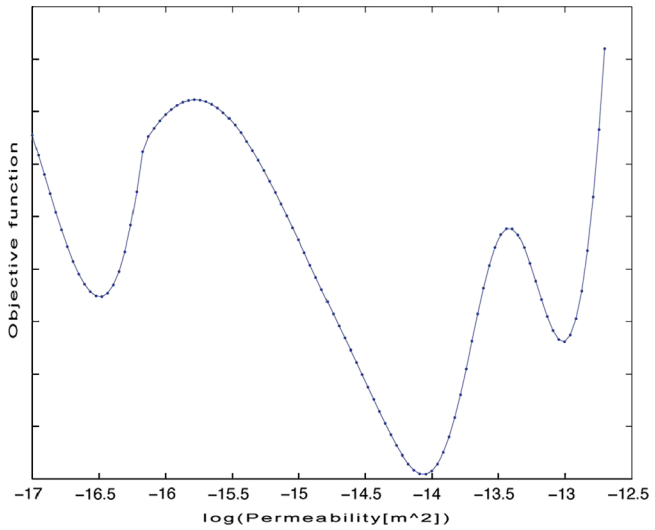

Fig. 1. A one-dimensional cut of the objective function for a geothermal reservoir model of the Laugarnes area described in Section 4. The logarithm of the permeability is varied. In addition to a global minimum (near −14.1), two local minima are present (near −16.5 and −13.0).

iTOUGH2 software, and will be compared with the global optimization method proposed here.

Simulated Annealing (Kirkpatrick et al., 1983) is an iterative procedure where an initial guess of the parameter values is iteratively updated with random increments and a selection criterion until a termination condition is reached. There, the objective function is taken to represent an 'energy' of the system, and a fictitious temperature is introduced. The temperature is introduced to control the probability of accepting increases in the objective function as an intermediate step to ultimately reach lower function values. A central issue in simulated annealing calculations is the 'time scale' of the cooling of the system from high temperature to zero temperature. The slower the cooling rate, the more likely the global minimum is found, but the computational effort becomes larger. It has been shown that in the impossible limit of infinitely long simulations with infinitesimal cooling rate, the method is guaranteed to give the global minimum (Haario and Saksman, 1991; Tsallis and Stariolo, 1996). For a given amount of computational effort, an implementation that can simulate a longer time interval is, more likely to reach the global minimum.

The Differential Evolution algorithm (Storn and Price, 1997) uses a randomly generated initial population, preferably covering the entire parameter space, which is then modified by differential mutation and crossover along with a selection criterion to find the minimum of the objective function. It has emerged as one of the simplest and most efficient techniques for solving global optimization problems. The method has been applied to diverse domains of science and engineering, such as mechanical engineering (Joshi and Sanderson, 1999), chemical engineering (Wang and Jang, 2000), machine intelligence, and pattern recognition (Das et al., 2008). Some weaknesses of the method have been identified (Lampinen and Zelinka, 2000). Furthermore, the performance of the method deteriorates as the number of parameters increases (Ali et al., 2012). Several suggestions for improving its performance have been proposed (Ali and Pant, 2011).

Harmony search (Geem et al., 2001) is also a population-based optimization algorithm using a stochastic random search (Lee and Geem, 2004). It has been applied to a wide variety of optimization problems (Geem et al., 2002, 2005; Kang and Geem, 2004; Kim et al., 2001; Lee and Geem, 2004). However, problems with the method, such as the need for parameter tuning, have been a topic of much research over the last 10 years where improvements have been proposed (Fourie et al., 2013).

The global optimization method presented here can be considered as a descendant of a method for long time scale simulations of atomic scale models of solids, known as adaptive kinetic Monte Carlo (AKMC) (Henkelman and Jónsson, 2001). The AKMC method has been successfully applied to atomic scale problems in solid-state physics and chemistry, see for example: (Henkelman and Jónsson, 2003; Karssemeijer et al., 2012; Pedersen et al., 2009a, 2009b; Pedersen and Jónsson, 2010). There, the time evolution is described by visiting local minima on the energy surface and identifying transitions by searching for first order saddle points on the objective function surface (Henkelman and Jónsson, 1999). Here, we modify the AKMC method to adapt it better to global optimization (Pedersen et al., 2012). The method, which we will refer to as global optimization using saddle traversals (GOUST) is described in detail below. It has been implemented in the EON software (Pedersen and Jónsson, 2010), which makes it possible to carry out the calculations using distributed and cloud computing

2. The GOUST method

The GOUST method relies on a fast way to identify first order saddle points on the objective function surface. We therefore first describe briefly the tool used for this purpose. A more detailed description is given in Henkelman and Jónsson (1999).

2.1. Minimum mode following

Let the number of variables of the objective function (parameters in the model to be fitted) be denoted by N. The objective function can be denoted as

$$f:\ \mathbb{R}^N \rightarrow \mathbb{R} \tag{1}$$

this defines a surface in N-dimensional space. The function is assumed to be differentiable. The extremal points where the gradient vanishes, $\nabla f = 0$, and the function value is low are of particular interest as these are local minima and low lying saddle points. To distinguish between these two kinds of extrema, the matrix of second order derivatives (the Hessian matrix, Eq. (2)) can be used. The Hessian has only positive eigenvalues at a local minimum, whereas one of the eigenvalues is negative at a first order saddle point (SPs).

$$H_{ij} = \frac{\delta^2 f}{\delta x_i \delta x_j} \tag{2}$$

To locate SPs, it is assumed that the gradient ∇f of the objective function can be evaluated readily (recent developments in automatic differentiation (see Gregory et al., 1997) could prove valuable in this context), but second derivatives are not needed. The method used to find SPs involves a minimization using a transformed gradient where the component along the minimum mode of the Hessian has been reversed

$$\nabla f^{eff} = \nabla f - 2(\nabla f \hat{\nu}_\lambda)\hat{\nu}_\lambda \tag{3}$$

here, $\hat{\nu}_\lambda$ is a normalized eigenvector corresponding to the minimum eigenvalue, λ, of the Hessian. This projection (Eq. (3)) locally transforms the gradient in the vicinity of a SP to a gradient characteristic of the vicinity of a minimum. A number of local minimization methods can then be used to converge on SPs when ∇f^{eff} is used as input, for example the conjugate gradient method. This will be referred to as the minimum mode following (MMF) method. It is important to note that only the minimum mode of the Hessian matrix is required here. The minimum mode vector can be estimated efficiently using either the dimer method (Henkelman and Jónsson, 1999) or the Lanczos method (Lanczos,