Tackling an Inverse Problem from the Petroleum Industry with a Genetic Algorithm for Sampling

Pedro J. Ballester and Jonathan N. Carter

Imperial College London, Department of Earth Science and Engineering,
RSM Building, Exhibition Road, London SW7 2AZ, UK
{p.ballester,j.n.carter}@imperial.ac.uk

Abstract. When direct measurement of model parameters is not possible, these need to be inferred indirectly from calibration data. To solve this inverse problem, an algorithm that preferentially samples all regions of the parameter space that fit data well is needed.
In this paper, we apply a real-parameter Genetic Algorithm (GA) to sample the parameter space for the inverse problem of calibrating a petroleum reservoir model. This results in several important insights into this nonlinear inverse problem.

1 Introduction

Because there is no available analytical expression between model parameters and measured data, numerical models are needed in Petroleum Engineering. The location of petroleum reservoirs (thousands of metres below ground) and their extension (typically several kilometres) make direct measurements of the model parameters mostly impossible. Only few of these parameters can be estimated from measurements at wells drilled into the reservoir. In this context, indirect measurements usually take the form of fluid production historical data and thus the inversion is called history matching. This can be posed as a search and optimisation problem by defining an objective function quantifying the mismatch between the model output and the measured production data.

In this work we consider a cross-sectional model of a layered reservoir (an extensive description of the model can be found in [1]). In order to quantify the match between the model response and the measurements, we define first an objective function for the history matching period, Δ_m, as follows

$$\Delta_m(\mathbf{m}) = \frac{1}{N_s} \sum_{j=1}^{N_s} \sum_{k=1}^{3} \frac{|S_{jk}(\mathbf{m}) - O_{jk}|}{2\sigma_{jk}} \quad . \tag{1}$$

where N_s is the number of time steps and determines the extent of the history matching period, $\mathbf{m} = (h, k_p, k_g)$ is the considered model, $S_{jk}(\mathbf{m})$ is its simulated response for the production series k at time step j, O_{jk} is the corresponding objective ('measured') value and σ_{jk}, an estimation of what would be the associated measurement error. We assume it as a 3% of the measured value, ie. $\sigma_{jk} = 0.03\ O_{jk}$. The response for the truth model is simulated to provide the measurements as $O_{jk} \equiv S_{jk}(h^0, k_p^0, k_g^0)$. The objective function for the prediction period, Δ_f, has an analogous expression.

K. Deb et al. (Eds.): GECCO 2004, LNCS 3103, pp. 1299–1300, 2004.

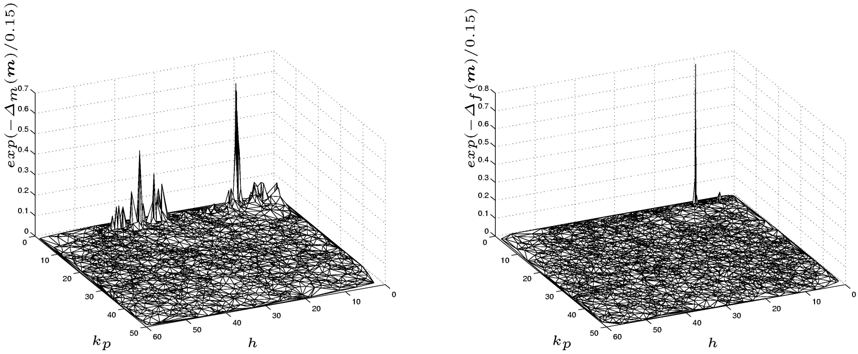

Fig. 1. GA for Sampling using $N_s = 36$ on: a) Δ_m with $\Delta_m^{best}(10.37, 1.425, 131.50) = 0.0795$ and b) Δ_f with $\Delta_f^{best}(10.40, 1.309, 131.83) = 0.0373$. The truth model was set as $h^0 = 10.4$, $k_p^0 = 1.31$ and $k_g^0 = 131.7$

2 Results and Discussion

The chosen search method is a Steady-state Real-parameter Genetic Algorithm. It combines the following features: parental selection is not fitness biased, a self-adaptative crossover operator, implicit elitism and locally scaled probabilistic replacement. Details of this GA can be found in [2]. Here we introduce an additional feature: children can only be bred within the initialisation region. The GA is used with N=50, $\lambda = 1$, NREP= 10 and $\eta = 0.01$, using a total of $7,050$ evaluations. The output of the algorithm will be the ensemble of all the individuals that entered the population during the GA run (typically a third of the total number of evaluations). In figure 1a, the very large spike, with $h \approx 10$, corresponds to the truth model. We can also see notable local optima in the regions with $0 < h < 8$ and $31 < h < 44$. Figure 1b shows the result of carrying out the inversion using the objective function for the prediction period, Δ_f. The only substantial point found corresponds to the truth model. All the other local optima that can be seen in Fig. 1a are unable to match the observations during the prediction period. We conclude that for this model you can only obtain a good prediction from the truth case, and that other good matches from the history matching phase have no predictive value. It is worth noting that the best found models in Fig. 1 are very similar to the truth model.

References

1. Ballester, P.J.: The use of Genetic Algorithms to Improve Reservoir Characterisation. PhD thesis, Department of Earth Science and Engineering, Imperial College London (2004)
2. Ballester, P.J., Carter, J.N.: An effective real-parameter genetic algorithms for multimodal optimization. In Parmee, I.C., ed.: Proceedings of the Adaptive Computing in Design and Manufacture VI. (2004) In Press.