Bayesian calibration of reservoir models using a coarse-scale reservoir simulator in the prior specification

Ole Petter Lødøen and Håkon Tjelmeland

February 28, 2007

Abstract

We consider the history matching problem in a Bayesian setting. The link between the reservoir variables and the production history is given by a reservoir simulator. To run the reservoir simulator is a computer intensive task, and this severely limits the number of runs that can be made. It is therefore natural to consider the reservoir simulator as an unknown function with a corresponding prior distribution. We define an informative prior distribution for the reservoir simulator by combining a coarser (and thereby faster) version of the reservoir simulator with parameters correcting for the bias introduced by the coarser lattice. We simulate from the resulting posterior distribution by a Metropolis–Hastings algorithm. We use the methodology on a constructed case study inspired by the Troll field in the North Sea.

1 Introduction

History matching is the process of calibrating reservoir variables to observed production data. In this paper we consider the problem in a Bayesian context. A reservoir simulator defines the link from the reservoir properties to production data. A run of the reservoir simulator may take several hours, or even days, to run. Thus, the reservoir simulator can in practice be run only for a very limited number of values for the reservoir variables. In practice the reservoir simulator is therefore largely an unknown function and, in a Bayesian setting, should be assigned a prior distribution, see also the discussion in Kennedy and O'Hagan (2001). The reservoir simulator is a very complex function going from a high dimensional input space to a high dimensional output space. To specify a prior distribution for this (unknown) function we adopt ideas from Omre and Lødøen (2004) and use a coarse scale, and thereby computationally fast, version of the reservoir simulator. The coarse scale simulator then becomes a part of the posterior distribution, so the only viable alternative for exploration of the posterior is the Metropolis–Hastings algorithm (Dellaportas and Roberts, 2003; Gamerman, 1997).

2 Model formulation

In this section we define a quite general Bayesian model. A more detailed definition of the model components is discussed in Section 3 and in Lødøen and Tjelmeland (2005).

2.1 Preliminaries

Reservoir properties that are important for fluid flow vary spatially and are typically represented on a three dimensional lattice. We let $x \in \mathbb{R}^n$ be a vector of the reservoir properties in each node. We let $\omega(x)$ denote the reservoir simulator. Of course the simulation output is also a function of the operating conditions used, but we will consider these as fixed and therefore ignore this dependency in our notation. We only consider the simulation output at a discrete set of times, so $\omega(x) \in \mathbb{R}^m$, where m is the number of production variables considered times the number of time points used. The vector $y = \omega(x)$ consists of two parts, components related to the past and components related to future production. Let D^p and D^f be matrices that picks out the components of y related to the past and future, respectively, and set $y^p = D^p\omega(x)$ and $y^f = D^f\omega(x)$. In particular, we let y_0^p denote the observed quantities from the reservoir under study and let y_0^f denote the corresponding future production that we want to predict.

We run the reservoir simulator for K values for the reservoir variables $x_1, \ldots, x_K$ and let $y_k = (y_k^p, y_k^f) = \omega(x_k), k = 1, \ldots, K$ denote the corresponding simulation results. As discussed above, we also use a coarse scale version of the reservoir simulator, and denote this by $\tilde{\omega}(x)$. Clearly, the $\tilde{\omega}(\cdot)$ really consists of two parts, an up scaling of x to a coarse grid representation and the fluid flow simulator on the coarse scale. As the same code is used for both grid representations, also $\tilde{\omega}(x) \in \mathbb{R}^m$.

2.2 Stochastic model

We use a Gaussian prior for x (after a suitable transformation) and define a prior for the unknown function $\omega(x)$ by setting,

$$\omega(x) = A\tilde{\omega}(x) + \varepsilon(x),$$

where $A \in \mathbb{R}^{m \times m}$ is a diagonal matrix, $b \in \mathbb{R}^m$ and $\varepsilon(x) \in \mathbb{R}^m$ is a multivariate process for $x \in \mathbb{R}^n$. The A and b are introduced to correct for the bias introduced by the coarse scale reservoir simulator. Independent priors are defined for A and b. Given a vector θ, we assume the residual term $\varepsilon(x)$ to be a zero mean Gaussian process with a covariance structure specified by θ. Finally, we assign a prior $\pi(\theta)$ to θ.

The observed data consists of y_0^p and $y_1, \ldots, y_K$. We consider $\omega(x)$ to be a perfect model for the fluid flow process and assume zero observation error for the production data. For the (unknown) true reservoir properties x we thereby have $y_0^p = D^p\omega(x)$. The likelihood function for the available data becomes

$$\left[\begin{array}{c} y_0^p \\ y_1 \\ \vdots \\ y_K \end{array}\right] \Bigg| \, x, A, b, \theta \sim \mathrm{N}\left(\left[\begin{array}{c} D^p(A\tilde{\omega}(x) + b) \\ A\tilde{\omega}(x_1) + b \\ \vdots \\ A\tilde{\omega}(x_K) + b \end{array}\right], \Sigma_\theta(x, x_1, \ldots, x_K) \right),$$

where $\Sigma_\theta(x, x_1, \ldots, x_K)$ follows from the covariance structure defined for the residual term $\varepsilon(x)$.

2.3 Posterior simulation

The posterior distribution of interest corresponding to the Bayesian model specified above is $\pi(x, y_0^f, A, b, \theta | y_0^p, y_1, \ldots, y_K)$. This is naturally split in two parts,

$$\pi(x, A, b, \theta | y_0^p, y_1, \ldots, y_K) \quad \text{and} \quad \pi(y^f | x, A, b, \theta, y_0^p, y_1, \ldots, y_K).$$

The first has to be explored by the Metropolis–Hastings algorithm, whereas the latter is Gaussian and can therefore be evaluated analytically. The posterior distribution $\pi(x, A, b, \theta | y_0^p, y_1, \ldots, y_K)$ includes $\tilde{\omega}(x)$, so the coarse scale reservoir simulator must be evaluated for each iteration of the Metropolis–Hastings algorithm. Even if evaluation of $\tilde{\omega}(x)$ is very fast compared to $\omega(x)$, this still puts severe restrictions on the number of Metropolis–Hastings iterations that can be run. It is therefore essential to use an algorithm with fast convergence and mixing.

3 Simulation example

In this section we consider a case study inspired by the Troll field in the North Sea. The case study was introduced in Hegstad and Omre (2001), and is also considered in Omre and Lødøen (2004). The outline of the reservoir under study is shown in Figure 1. Initially, the reservoir is fully saturated with oil and has one vertical injection well and two horizontal production wells. The reservoir is discretised on a lattice of size $50 \times 50 \times 15$. Following Hegstad and Omre (2001), we define a reference reservoir x^{true} containing true seismic reflection coefficients, porosities and permeabilities in each node of the reservoir. The coarse representation of the reservoir is on a $10 \times 10 \times 15$ lattice. Porosities and saturations are up scaled by arithmetic averages and for permeabilities harmonic averages are used. When running the reservoir simulator we assume properties as mobility ratios, rock compressibilities and relative permeability curves as known. The vertical injection well injects gas at a constant rate (65 000 mscf.day). The two production wells first produce at a constant rate (15 000 stb/day) in each well. When the pressure in either well drops below 4 100 psi, this well switches to producing at a constant pressure of 4 100 psi. We use the Eclipse reservoir simulator and evaluate the pressure in the injection well (bhp), the gas/oil ratios in the production wells ($gor1$ and $gor2$) and the oil production

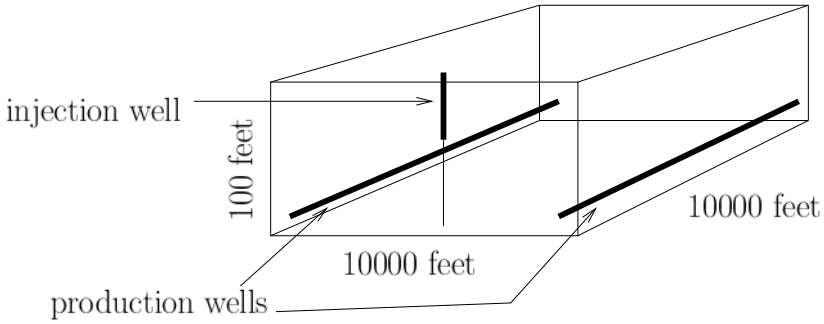

Figure 1: Outline of the reservoir. The thick lines indicate where the wells are perforated.

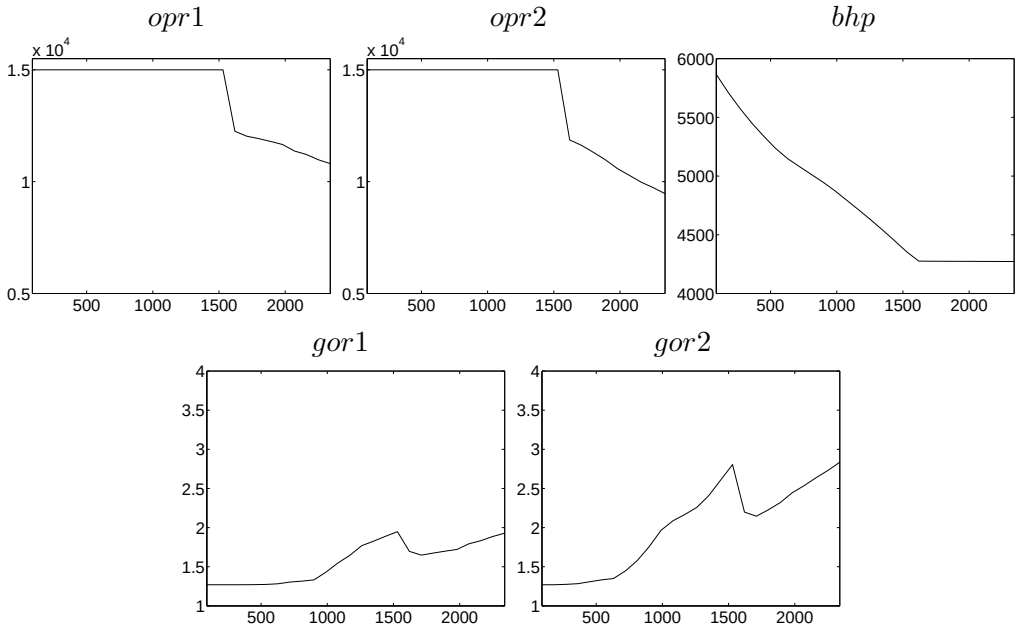

Figure 2: The production y^{true} from the reference reservoir as function of days.

rates in the production wells ($opr1$ and $opr2$), all at a quarterly basis for six and a half years. Figure 2 shows the five responses for the reference reservoir. To be able to specify a Gaussian model for the difference $\omega(x) - \tilde{\omega}(x)$ we parameterise the response variables by the length of the time period it is at the constant level and values after it has left the level. For details on this, the covariance structure used for $\varepsilon(x)$ and the hyper-parameter values used we refer to Lødøen and Tjelmeland (2005).

In our Metropolis–Hastings algorithm we use two update types, (i) updates for x, and (ii) updates of the parameters in the likelihood model, i.e. θ, A and b. When updating x we propose a small change in all nodes. The $\tilde{\omega}(x)$ must be evaluated to compute the acceptance probability for this proposal type. Random walk proposals are used for θ, A and b. No reservoir simulator run is necessary to compute the corresponding acceptance probability in this case and these updates are therefore very fast. Moreover, the updates of θ, A and b can be done in parallel with the update of x and we use this to do many updates for θ, A and b in parallel with one update of x. For more details of the simulation algorithm and convergence properties of the algorithm we again refer to Lødøen and Tjelmeland (2005).

Figure 3 shows independent realisations y_0^f from the posterior distribution generated

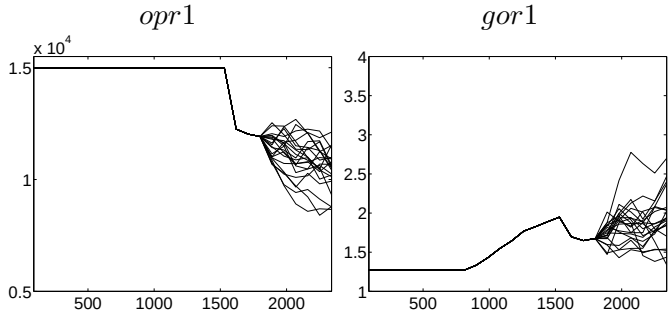

Figure 3: Independent posterior realisations of $y = (y_0^p, y^f)$.

by the Metropolis–Hastings algorithm when conditioned on production data up to 1 800 days. The posterior mean of the future production is close to the true production from the reference reservoir, but we see that the prediction uncertainty is quite large. More details of the simulation results can again be found in Lødøen and Tjelmeland (2005).

4 Closing remarks

We have formulated the history matching problem in a Bayesian setting. Using ideas from Omre and Lødøen (2004) we have used a coarse scale reservoir simulator to specify an informative prior for a fine scale reservoir simulator. We have applied the model to a constructed case study. The results discussed in Lødøen and Tjelmeland (2005) demonstrate how the use of observed production history provides information both about the bias correction parameters A and b and the reservoir properties. The result is a better characterisation of the reservoir properties and thereby better prediction of future production.

References

Dellaportas, P. and Roberts, G. O. (2003). An introduction to MCMC, *in* J. Møller (ed.), *Spatial Statistics and Computational Methods*, number 173 in *Lecture Notes in Statistics*, Springer, Berlin, pp. 1–41.

Gamerman, D. (1997). *Markov chain Monte Carlo: Stochastic simulation for Bayesian inference*, Chapman & Hall, London.

Hegstad, B. and Omre, H. (2001). Uncertianty in production forecasts based on well observations, seismic data, and production history, *SPE Journal* (June 2001): 409–424.

Kennedy, M. and O'Hagan, A. (2001). Bayesian calibration of computer models (with discussion), *J. R. Statist. Soc. B* **63**: 425–464.

Lødøen, O. and Tjelmeland, H. (2005). Bayesian calibration of hydrocarbon reservoir models using an approximate reservoir simulator in the prior specification, *Technical report, Statistics no. 5*, Department of Mathematical Sciences, Norwegian University of Science and Technology, Trondheim, Norway.

Omre, H. and Lødøen, O. (2004). Improved production forecasts and history matching using approximate fluid-flow simulators, *SPE Journal* (September 2004): 339–351.