A METHOD FOR APPROXIMATE FULLY BAYESIAN ANALYSIS OF PARAMETERS

Anne Randi Syversveen and Håvard Rue

Department of Mathematical Sciences
Norwegian University of Science and Technology
N-7034 Trondheim, Norway

September 1996

Keywords: marked point process; parameter estimation; Markov chain Monte Carlo; reservoir characterization;

ABSTRACT

Stochastic modeling of the geology in petroleum reservoirs has become an important tool in order to investigate flow properties in the reservoir. The stochastic models used contain parameters which must be estimated based on observations and geological knowledge. The amount of data available is however quite limited due to high drilling costs etc., and the lack of data prevents the use of many of the standard data driven approaches to the parameter estimation problem. Modern simulation based methods using Markov chain Monte Carlo simulation, can however be used to do fully Bayesian analysis with respect to parameters in the reservoir model, with the drawback of relatively high computational costs. In this paper, we propose a simple, relatively fast approximate method for fully Bayesian analysis of the parameters. We illustrate the method on both simulated and real data using a two–dimensional marked point model for reservoir characterization.

1 INTRODUCTION

One problem in evaluation of petroleum reservoirs is to characterize the geology in order to investigate the flow properties. An example is location of shales in a background of sand, where shales create barriers for fluid flow, and sandstone allows fluid to flow easily. Sources for information about reservoir characteristics can be well observations, seismic data, and previous production history. In the following we will consider only well observations. Usually, the number of wells is small, due to the high drilling costs. Therefore, data will be scarce. To illustrate what the geology may look like, we show in Figure 1 an outcrop taken from the

E–mail: annerand@imf.unit.no, havard.rue@imf.unit.no

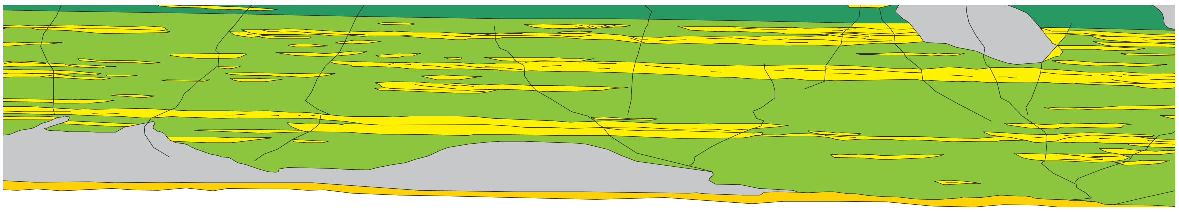

FIG. 1: Outcrop from the San Juan basin, Rocky Mountains, showing a cross section of meandering channel sandstone embedded in lower to upper delta plain mudstone, siltstone and coal.

Safari (1995) database collected by Norsk Hydro, Statoil, Saga Petroleum and The Norwegian Petroleum Directorate. The outcrop has objects of permeable sandstone in a background of non-permeable shale, and is believed to be an analog to a non-observable oil reservoir.

Haldorsen and Lake (1984) described a simple marked point model for shales in a vertical cross section of a sand matrix. The model had several severe weaknesses; it did not contain spatial interaction between the shales, and it was not possible to condition on observations from more than one well for each shale. Recently Syversveen and Omre (1997) presented an extended model with pairwise spatial interaction between objects. The model, which is described in Section 2, can handle conditioning on observations from more than one well for each object. However, the important problem of estimation of parameters in the model, was not completely treated. Some of the parameters are naturally estimated from data combined with geological knowledge, and in this paper we will consider estimation of these parameters. Prior distributions for the parameters are specified based on geological experience and combined with the observed data within the Bayesian framework. Because data are so scarce, we need the complete posterior distribution for the parameters and not only a point estimate, to judge how the data influence the prior distribution. As a consequence, more classical methods for parameter estimation in point processes (see for example the review in Diggle *et al.* (1994)), are not appropriate for our problem.

We can however do a fully Bayesian analysis of the parameters by use of modern Markov chain Monte Carlo (MCMC) methods, which are iterative simulation based approaches. In Section 3 we will give an introduction to these methods. MCMC methods are often computationally intensive, and in Section 4 we propose a relatively fast and approximate method to do a full Bayesian analysis. An important assumption is a quite low information content in the data, which is fulfilled in our problem where the data are scarce. The proposed method is presented in Section 4 and tested in Section 5 with simulated data and the outcrop in Figure 1.

2 THE MODEL

We will now present a two-dimensional model for cross sections of the reservoir, like the outcrop shown in Figure 1. Practical use of such a model involves an extension to the three–dimensional space, but we will not discuss this here. The framework is a marked point process where the object is the "mark" and its location the "point".

Each object is modeled as a rectangle with center position $\boldsymbol{C} = (C^1, C^2)$, length and height $\boldsymbol{S} = (L, H)$, and a deformation on top and bottom represented by a pair of random functions $\boldsymbol{Z}(\cdot) = (\boldsymbol{Z}_B(\cdot), \boldsymbol{Z}_T(\cdot))$. We write $\boldsymbol{U} = (\boldsymbol{C}, \boldsymbol{S}, \boldsymbol{Z}(\cdot))$, see Figure 2. The height of the rectangle

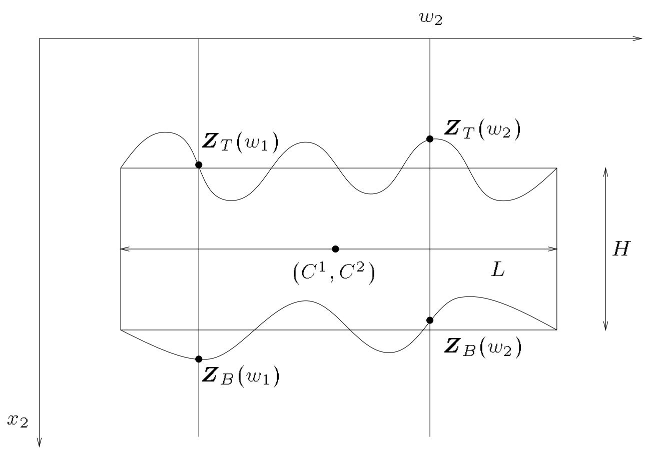

FIG. 2: Model for a single object.

is equal to the expected height of the object, while the length of the object equals the length of the rectangle. The deformation $\boldsymbol{Z}(\cdot)$ on top and bottom, are Gaussian random fields present in order to fulfill the conditioning of the data. As the deformation $\boldsymbol{Z}(\cdot)$ is not important in the parameter estimation problem, we will not discuss it further.

The prior pdf for N objects in a domain is defined in the following way: We let $\boldsymbol{\theta}_i = (c_i^1, c_i^2, l_i, h_i)$ denote the rectangle associated with object number i, and $\boldsymbol{z}_i(\cdot)$ its top and bottom positions. The joint density for the N objects can be written

$$f(\boldsymbol{\theta}, n, \boldsymbol{z}(\cdot)) = f(\boldsymbol{\theta}, n) f(\boldsymbol{z}(\cdot)|\boldsymbol{\theta}, n), \tag{1}$$

where $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n)$ and $\boldsymbol{z}(\cdot) = (\boldsymbol{z}_1(\cdot), \ldots, \boldsymbol{z}_n(\cdot))$. The last term $f(\boldsymbol{z}(\cdot)|\boldsymbol{\theta}, n)$, is Gaussian random fields (see Syversveen and Omre (1997)). The first term $f(\boldsymbol{\theta}, n)$ is a marked point process with pairwise interaction (Stoyan, Kendall, and Mecke 1995) of the form

$$\begin{aligned} f_{a,r}(\boldsymbol{\theta}, n) &= \frac{1}{c(a,r)} \exp\{\sum_{i=1}^{n} b_1(c_i^1, l_i) + b_2(c_i^2, h_i) + \sum_{j=1}^{n} \sum_{i<j} d_r(\boldsymbol{c}_i, \boldsymbol{c}_j) + na\} \\ &= h_{a,r}(\boldsymbol{\theta}, n)/c(a,r). \end{aligned} \tag{2}$$

Here $c(a, r)$ is an unknown function of the parameters a and r, the term $b_1(c_i^1, l_i) + b_2(c_i^2, h_i)$ represents the marginal density for a single object if no interaction was present. The function $d(\boldsymbol{c}_i, \boldsymbol{c}_j)$ is a pairwise interaction function between center-points of objects, defined as

$$d_r(x_i, x_j) = \begin{cases} \ln(\frac{|x_i - x_j|}{r}) & \text{if } |x_i - x_j| \leq r \\ 0 & \text{if } |x_i - x_j| > r. \end{cases} \tag{3}$$

It may have been more natural to relate the interaction between the objects to the size of the objects and not only center-points. However, we believe that interaction between center-points models the interaction sufficiently, at least to some first order. The parameter a in equation (2) is related to the intensity of objects, increasing the value of a yields an increasing (expected) number of objects.

The parameters to be estimated are the intensity parameter a and the interaction parameter r. The data are the dimension of the rectangular reservoir and exact observations of the reservoir along the transects of a number of wells. It is known which observations that belong to the same object. Observations of depth to top and bottom position of objects are contained in the vector $\boldsymbol{o}$. Note that neither $\boldsymbol{\theta}$ nor n can be observed directly. The variable $\boldsymbol{\theta}$ is not observable because of the way it is defined in the model; center-points and expected height of objects are never observed. The number of objects n is unobservable since only well observations are available, not the whole reservoir. This means that the model contains hidden variables, which makes the parameter estimation problem non–standard.

Taking observations into account, a posterior pdf for $\boldsymbol{Z}, \boldsymbol{\Theta}$, and N can be defined. See Syversveen and Omre (1997) for a definition of the likelihood function. The posterior pdf for $\boldsymbol{\Theta}, N$ is written as

$$\begin{aligned} f_{a,r}(\boldsymbol{\theta}, n|\boldsymbol{o}) =& f(\boldsymbol{\theta}, n) f(\boldsymbol{o}|\boldsymbol{\theta}, n)/c_o'(a,r) \\ =& \exp\Big\{ \sum_{i=1}^{n} (b_1(c_i^1, l_i|\boldsymbol{o}) + b_2(c_i^2, h_i|\boldsymbol{o})) \\ & + \sum_{j=1}^{n} \sum_{i<j} d_r(\boldsymbol{c}_i, \boldsymbol{c}_j) + na \Big\} I(\boldsymbol{o}|n)/c_o(a,r) \\ =& h_{a,r}(\boldsymbol{\theta}, n|\boldsymbol{o})/c_o(a,r). \end{aligned} \tag{4}$$

The function $I(\boldsymbol{o}|n)$ is an indicator function which is equal one when conditioning is fulfilled and zero otherwise, and $c_o'(a,r)$ and $c_o(a,r)$ are normalizing functions.

When a and r are given constants, we use the Metropolis-Hastings algorithm (Metropolis *et al.* 1953, Hastings 1970) to simulate from the distribution defined by the pdf (4). Proposals and acceptance rates are described in detail in Syversveen and Omre (1997).

3 PARAMETER ESTIMATION BY MONTE CARLO

In this section we review two parameter estimation methods based on Markov chain Monte Carlo (MCMC) simulations. By the first method we obtain the maximum a posteriori (MAP) estimator by doing a minor modification of the maximum likelihood estimator. The second method gives realizations from the posterior distribution for the parameters, and estimators such as the posterior mean can easily be obtained.

The basic idea behind MCMC simulation is to simulate an ergodic Markov chain, having the desired distribution $\pi(x)$ as its limiting distribution. The transition kernels can be defined in several ways, leading to different algorithms. Good review papers on MCMC are Besag *et al.*

(1995) and Tierney (1994). Green (1995) describes how to do MCMC when the state space is a union of spaces of different dimensionality, and covers MCMC for spatial point processes discussed by Geyer and Møller (1994).

3.1 The MAP estimator

The MAP estimator can be obtained by a minor modification of the maximum likelihood estimator, so we start with the latter estimator.

Monte Carlo maximum likelihood (MCML) is described by Geyer and Thompson (1992) and Geyer (1994b). The method is well suited for normalized families of densities with pdf of the form (2), where $c(a,r)$ is the normalizer of the family and $h_{a,r}(\boldsymbol{\theta}, n)$ is the unnormalized density. MCMC methods can be used to generate realizations from this kind of distributions, as mentioned in Section 2.

Suppose first that we can observe $(\boldsymbol{\theta}, n)$. Let the log-likelihood corresponding to an observation of $(\boldsymbol{\theta}, n)$ be the log-likelihood ratio against an arbitrary fixed parameter point (a_0, r_0),

$$l_{a_0,r_0}(a,r) = \ln\Big(\frac{h_{a,r}(\boldsymbol{\theta}, n)}{h_{a_0,r_0}(\boldsymbol{\theta}, n)}\Big) - \ln\Big(\frac{c(a,r)}{c(a_0,r_0)}\Big).$$

The last term is equal to $\ln(E_{a_0,r_0}[\frac{h_{a,r}(\boldsymbol{\theta},n)}{h_{a_0,r_0}(\boldsymbol{\theta},n)}])$, since

$$E_{a_0,r_0}\Big(\frac{h_{a,r}(\boldsymbol{\theta}, n)}{h_{a_0,r_0}(\boldsymbol{\theta}, n)}\Big) = \int \frac{h_{a,r}(\boldsymbol{\theta}, n)}{h_{a_0,r_0}(\boldsymbol{\theta}, n)} f_{a_0,r_0}(\boldsymbol{\theta}, n) d\mu(\boldsymbol{\theta}, n) = \frac{c(a,r)}{c(a_0,r_0)}.$$

This gives the following Monte Carlo approximation of $l(a,r)$

$$\hat{l}_{a_0,r_0}(a,r) = \ln\Big(\frac{h_{a,r}(\boldsymbol{\theta}, n)}{h_{a_0,r_0}(\boldsymbol{\theta}, n)}\Big) - \ln\Big(\frac{1}{n}\sum_{i=1}^{n} \frac{h_{a,r}(\boldsymbol{\theta}_i, n_i)}{h_{a_0,r_0}(\boldsymbol{\theta}_i, n_i)}\Big), \tag{5}$$

where $(\boldsymbol{\theta}_1, n_1)$, $(\boldsymbol{\theta}_2, n_2), \ldots$ are samples from $f_{a_0,r_0}(\boldsymbol{\theta}, n)$ generated by MCMC. The Monte Carlo approximation $\hat{l}_{a_0,r_0}(a,r)$ converges almost surely to $l(a,r)$ for any fixed (a,r) when n goes to infinity.

However, when we have hidden variables as described in Section 2, the first term in (5) can not be computed because $(\boldsymbol{\theta}, n)$ is not fully observed. Instead this term must be estimated by simulations from the posterior distribution (4). The log-likelihood is now

$$l_{a_0,r_0}(a,r) = \ln\Big(E_{a_0,r_0}\Big[\frac{h_{a,r}(\boldsymbol{\Theta}, N|\boldsymbol{O}=\boldsymbol{o})}{h_{a_0,r_0}(\boldsymbol{\Theta}, N|\boldsymbol{O}=\boldsymbol{o})}\Big]\Big) - \ln\Big(E_{a_0,r_0}\Big[\frac{h_{a,r}(\boldsymbol{\Theta}, N)}{h_{a_0,r_0}(\boldsymbol{\Theta}, N)}\Big]\Big),$$

and the Monte Carlo approximation becomes

$$\hat{l}_{a_0,r_0}(a,r) = \ln\Big(\frac{1}{n}\sum_{i=1}^{n} \frac{h_{a,r}(\boldsymbol{\theta}_i^*, n_i^*|\boldsymbol{o})}{h_{a_0,r_0}(\boldsymbol{\theta}_i^*, n_i^*|\boldsymbol{o})}\Big) - \ln\Big(\frac{1}{n}\sum_{i=1}^{n} \frac{h_{a,r}(\boldsymbol{\theta}_i, n_i)}{h_{a_0,r_0}(\boldsymbol{\theta}_i, n_i)}\Big). \tag{6}$$

The first term contains samples of $(\boldsymbol{\Theta}, N)$ from the posterior distribution (4), and the second term contains samples from the prior distribution (2), all sampled with parameter values (a_0, r_0).

When the parameters a, r are given a prior distribution $f(a, r)$, the MCML method can be extended to give the maximum a posteriori (MAP) estimator. Instead of maximizing the likelihood, we maximize the posterior density ratio
$f(a, r|\boldsymbol{o})/f(a_0, r_0|\boldsymbol{o})$ of (a, r) and (a_0, r_0). The expression to be maximized with respect to a, r is now

$$\hat{l}_{a_0,r_0}(a, r) = \ln\Big(\frac{1}{n}\sum_{i=1}^{n}\frac{h_{a,r}(\boldsymbol{\theta}_i^*, n_i^*|\boldsymbol{o})}{h_{a_0,r_0}(\boldsymbol{\theta}_i^*, n_i^*|\boldsymbol{o})}\Big) - \ln\Big(\frac{1}{n}\sum_{i=1}^{n}\frac{h_{a,r}(\boldsymbol{\theta}_i, n_i)}{h_{a_0,r_0}(\boldsymbol{\theta}_i, n_i)}\Big) + \ln\Big(\frac{f(a, r)}{f(a_0, r_0)}\Big). \quad (7)$$

Note that $f(a_0, r_0)$ can be omitted in the last term; it has been included as a normalizing term in order to get $\hat{l}_{a_0,r_0}(a_0, r_0) = 0$ as in the MCML case.

3.2 Simulated tempering

Simulated tempering is a way to use MCMC methods to sample from the whole posterior distribution for parameters, converse to the MAP–estimator which only locates the mode of the posterior pdf. The goal is to sample from the joint posterior distribution of $(\boldsymbol{\theta}, n)$ and the parameters (a, r). Then the posterior distribution for (a, r) is obtained by integrating out $(\boldsymbol{\theta}, n)$. The joint posterior probability density function for $(\boldsymbol{\theta}, n)$ and (a, r) is

$$f(\boldsymbol{\theta}, n, a, r|\boldsymbol{o}) \propto f_{a,r}(\boldsymbol{\theta}, n|\boldsymbol{o})f(a, r) = h_{a,r}(\boldsymbol{\theta}, n|\boldsymbol{o})f(a, r)/c_o(a, r)$$

using equation (4). Since $c_o(a, r)$ is unknown, it must be estimated. This can be done by using reverse logistic regression, see Geyer (1994a). The idea is to sample from several distributions with different values of (a, r), and treat these as samples from a so-called mixture distribution.

The sampling from $f(\boldsymbol{\theta}, n, a, r|\boldsymbol{o})$ is done by MCMC with two types of transitions, changing $(\boldsymbol{\theta}, n)$ or changing (a, r). In practice, $c(a, r)$ must be precomputed on a grid, since the sampling procedure will otherwise be too time-consuming.

Higdon *et al.* (1995) and Weir (1995) describe fully Bayesian analysis for images constructed from emission computed tomography data following the guidelines above.

4 APPROXIMATE FULLY BAYESIAN ESTIMATION

We have in Section 3 discussed how to use modern Markov chain Monte Carlo methods to do Bayesian inference for the two parameters of interest in the model described in Section 2, the intensity and interaction parameter. The fully Bayesian approach is preferred in our case, as it give a precise description of how the (limited) information contained in the observed data will influence the posterior pdf of the parameters. From a computational point of view however, the MAP estimator is preferable, as it is simpler and faster to compute compared to the fully Bayesian analysis through the more complicated simulated tempering approach.

To describe our approximate method for a fully Bayesian analysis, assume for a moment that the parametric family posterior pdf for the parameters of interest is known. Further assume the parameters in the pdf can be determined from the local characteristics at the mode. As an example, let the posterior distribution belong to the one-dimensional Gaussian family. The two parameters, the expectation and the variance, are then uniquely determined respectively from the mode and the second derivative of the logarithm of the density, using the fact that $\frac{d^2}{dx^2} \ln f(x) = -1/\sigma^2$.

Our method for an approximate fully Bayesian analysis of the parameters are based on the assumption about a known parametric family for the posterior pdf of the parameters of interest. With low information content in the observed data, it is a quite accurate assumption to assume that the posterior pdf belongs to same family as the prior pdf. We compute the MAP estimator and then estimate the parameters in the posterior pdf from estimated local characteristics to obtain the full posterior pdf for the parameters. The posterior pdf can of course be used later to obtain a point estimate by minimizing a suitable loss function, for example to obtain the posterior mean (which may or may not be the mode).

The big advantage with our approach is the reduction of computational complexity, both with respect to coding and computational costs. However, the assumption of a known parametric family of the posterior pdf must hold, at least approximately. In the next Section we will illustrate our approach on the presented model analyzing a simulated data-set and the outcrop in Figure 1. We will also demonstrate how we can check the validity of the assumption of a known parametric family for the posterior pdf.

5 EXAMPLES

In this section we will illustrate how to apply approximate fully Bayesian estimation in the reservoir characterization model described in Section 2, using the idea outlined in Section 4. The parameters to be estimated are the intensity parameter a and the interaction parameter r. We assign a joint Gaussian prior distribution to the parameters for two reasons. First, it has an easy interpretation for the geologists specifying the priors. Second, this choice simplifies the presentation of our main idea.

The parameters in the prior distribution can be set as follows. Based on simulations from the prior model, the geologist can evaluate the realizations and then gain insight into how the parameters affect the distribution of number of objects and how they distribute within the window of interests.

We will now see how we can determine the posterior pdf for the parameters from the local characteristics around the MAP estimate. The simulated logarithm of the posterior ratio (7) can be written as

$$\begin{aligned}\hat{l}_{a_0,r_0}(a,r) = &\ln\Big[\frac{1}{m}\sum_{i=1}^{m}\frac{\exp\left(\sum\sum d_r(\boldsymbol{c}_k^*,\boldsymbol{c}_j^*)+n_i^*a\right)}{\exp\left(\sum\sum d_{r_0}(\boldsymbol{c}_k^*,\boldsymbol{c}_j^*)+n_i^*a_0\right)}\Big]\\ &-\ln\Big[\frac{1}{m}\sum_{i=1}^{m}\frac{\exp\left(\sum\sum d_r(\boldsymbol{c}_k,\boldsymbol{c}_j)+n_ia\right)}{\exp\left(\sum\sum d_{r_0}(\boldsymbol{c}_k,\boldsymbol{c}_j)+n_ia_0\right)}\Big]+\ln\Big[\frac{f(a,r)}{f(a_0,r_0)}\Big],\end{aligned} \tag{8}$$

where the first term comes from simulations from the posterior model for $(\boldsymbol{\Theta}, N)$ and the second term from simulations from the prior model. In the Gaussian distribution, the mean equals the mode, so the posterior mean for a and r are the values that maximize (8). The posterior covariance matrix is determined as follows. Let the covariance matrix be denoted

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}.$$

Let $f(a, r|\boldsymbol{o})$ be the Gaussian posterior density for a and r. Then we have that

$$\frac{\partial^2 \log f(a, r|\boldsymbol{o})}{\partial a^2} = \frac{-1}{\sigma_1^2(1-\rho^2)}, \tag{9}$$

$$\frac{\partial^2 \log f(a, r|\boldsymbol{o})}{\partial r^2} = \frac{-1}{\sigma_2^2(1-\rho^2)}, \tag{10}$$

and

$$\frac{\partial^2 \log f(a, r|\boldsymbol{o})}{\partial a \partial r} = \frac{\rho}{\sigma_1\sigma_2(1-\rho^2)}. \tag{11}$$

Numerical values for the second derivatives are estimated from the simulations using standard finite difference approximations,

$$\frac{\partial^2 \hat{l}_{a_0,r_0}(a,r)}{\partial a^2} = \frac{1}{h_1^2}[\hat{l}_{a_0,r_0}(a-h_1,r) - 2\hat{l}_{a_0,r_0}(a,r) + \hat{l}_{a_0,r_0}(a+h_1,r)] + O(h_1^2) \tag{12}$$

$$\frac{\partial^2 \hat{l}_{a_0,r_0}(a,r)}{\partial r^2} = \frac{1}{h_2^2}[\hat{l}_{a_0,r_0}(a,r-h_2) - 2\hat{l}_{a_0,r_0}(a,r) + \hat{l}_{a_0,r_0}(a,r+h_2)] + O(h_2^2) \tag{13}$$

$$\begin{aligned}\frac{\partial^2 \hat{l}_{a_0,r_0}(a,r)}{\partial a \partial r} = &\frac{1}{4h_1h_2}[\hat{l}_{a_0,r_0}(a+h_1,r+h_2) + \hat{l}_{a_0,r_0}(a-h_1,r-h_2) - \\ &\hat{l}_{a_0,r_0}(a+h_1,r-h_2) - \hat{l}_{a_0,r_0}(a-h_1,r+h_2)] + O(h_1^2) + O(h_2^2)\end{aligned} \tag{14}$$

where the error terms are functions of fourth order derivatives. Under the assumption that the posterior is Gaussian, the error terms are all zero, since fourth order derivatives are zero.

Setting the analytical expressions (9) – (11) for the second derivatives of the posterior pdf equal to the empirical estimates based on equations (12) – (14) respectively, we obtain an estimate for Σ.

The rate of convergence for MCML (for obtaining the MAP estimate) depends on the variability in the terms in the two first sums in (8); the more variability, the slower convergence. See for example Geyer and Thompson (1992) for details on this point. We can therefore monitor the variability of each of the two terms in (8) to gain insight in how many samples from the Markov chain we need to get a reasonable estimate.

Expression (8) is only calculated for values of a and r on a grid spaced by h_1 and h_2 to avoid storage problems. Since the posterior ratio is reliable only in a small area around (a_0, r_0), we iterate until a_0 and r_0 maximize (8). If we want an estimate of the uncertainty in the estimate for Σ, we can repeat this procedure with these values for a_0 and r_0 and get an estimate of the Monte Carlo error.

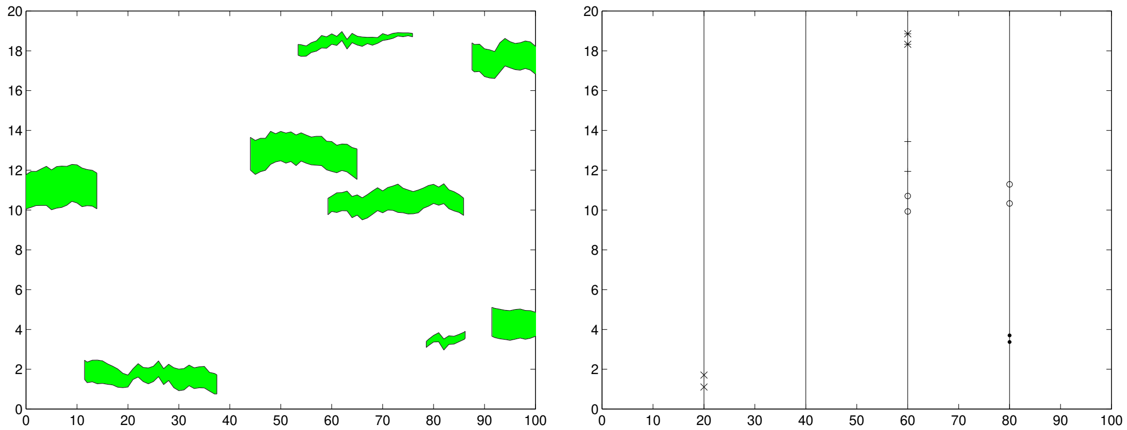

FIG. 3: To the left, realization from the prior model, regarded as the "true" reservoir. To the right, observations of top and bottom positions of objects collected from wells. Same symbol indicates same object.

The goodness of the approximation to the posterior distribution can be checked by comparing posterior density ratios $f(a,r|\boldsymbol{o})/f(a_0,r_0|\boldsymbol{o})$ for the (approximate) analytical expression with results from simulations, given by the exponential of expression (8). The comparison can be done for different parameter values a, r, a_0, and r_0.

5.1 Analyzing a simulated dataset

In our first example, we analyze a simulated dataset where the parameters a and r are drawn from their prior and then the configuration of the objects are simulated from the model to get our petroleum reservoir. The parameters in the prior mean is $[4.0 \;\; 30.0]^T$ and the covariance matrix is

$$\Sigma = \begin{pmatrix} 1.0 & 5.0 \\ 5.0 & 100.0 \end{pmatrix}.$$

It is reasonable to introduce positive correlation between a and r, because the intensity of objects is influenced by the interaction parameter. Figure 3 shows the realization at left, and the observations available at right. We drill four wells at horizontal positions 20, 40, 60, and 80 and observe top and bottom of the objects in the wells.

We use the discretization parameters $h_1 = 0.1$ and $h_2 = 1.0$ and collect 10000 samples, chosen as every 10th iteration after 1000 initial stabilizing iterations. This is repeated 10 times, all resulting in the same mode, and based on this, the posterior covariance matrix is estimated. The posterior distribution for a and r is then (approximately) Gaussian, with mean $\mu = [4.1 \;\; 30.0]^T$ and covariance

$$\Sigma = \begin{pmatrix} 0.62 \pm 0.01 & 4.23 \pm 0.09 \\ 4.23 \pm 0.09 & 66.4 \pm 2.1 \end{pmatrix}.$$

TABLE I: Posterior ratio for example with simulated data with $a_0 = 2.8$, $r_0 = 18.0$.

	Diff. estimated density - simulations				
$r \setminus a$	2.6	2.7	2.8	2.9	3.0
16.0	-0.021	-0.021	-0.009	0.018	0.061
17.0	-0.008	-0.013	-0.009	0.011	0.046
18.0	0.011	0.002	0.000	0.011	0.037
19.0	0.034	0.023	0.016	0.019	0.036
20.0	0.059	0.048	0.039	0.037	0.045

with the uncertainty given as one standard deviation. We see that the mean is almost the same as in the prior distribution, but the data have modified the covariance matrix. The variances of a and r have decreased compared to the prior distribution; that is, the observations have reduced the uncertainty about the parameter values.

We have checked the goodness of the approximation to the posterior distribution by computing the the posterior density ratio for various values of a and r both close to and more distant from a_0 and r_0 from simulations and from the estimated density. The simulated values are based on 10000 samples. Table I shows the estimated differences when $a_0 = 2.8$ and $r_0 = 18.0$. The differences are quite small and the approximation of a Gaussian posterior pdf seems quite good.

5.2 Analyzing a real dataset

In our last example we apply our method on the outcrop shown in Figure 1. The outcrop is taken from the San Juan basin, Rocky Mountains, and shows a cross section of meandering channel sandstone embedded in lower to upper delta plain mudstone, siltstone, and coal. The objects are sand and the background is nonpermeable shale. The size of the formation is rescaled and taken to be 100×16 units, where one unit is about 11 meters. We drill three wells are at position 20, 48, and 58 units, and from these wells we collect observations.

We chose the prior pdf for the two parameters as follows. We varied the parameters in the model in order to obtain a structure for realizations from the model comparable with other similar outcrops from the Safari database. This ended in the following choice for (a, r): mean $[9.0\ \ 20.0]^T$ and covariance

$$\Sigma = \begin{pmatrix} 1.0 & 2.5 \\ 2.5 & 25.0 \end{pmatrix}.$$

We used the discretization parameters $h_1 = 0.1$ and $h_2 = 1.0$. From 150000 samples from the posterior pdf, we obtained the following estimate of the parameters of the posterior pdf: the mean $[8.8\ \ 24]^T$, and covariance

$$\Sigma = \begin{pmatrix} 1.06 & 2.61 \\ 2.61 & 16.6 \end{pmatrix}.$$

We have also compared the posterior density ratio obtained analytically and with simulations. The difference is small, which support our approximation to the parametric family of the posterior pdf.

ACKNOWLEDGEMENTS

We thank the Safari project for giving us permission to use data from the Menefee formation. The first author is funded by a PhD grant from The Research Council of Norway.

BIBLIOGRAPHY

Besag, J., P. Green, D. Higdon, and K. Mengersen (1995). Bayesian computation and stochastic systems (with discussion). *Statistical Science 10*(1), 3–66.

Diggle, P., T. Fiksel, P. Grabarnik, Y. Ogata, D. Stoyan, and M. Tanemura (1994). On parameter estimation for pairwise interaction point processes. *Int. Statist. Rev. 62*(1), 99–117.

Geyer, C. (1994a). Estimating normalizing constants and reweighting mixtures in Markov chain Monte Carlo. Technical report, School of Statistics, University of Minnesota.

Geyer, C. (1994b). On the convergence of Monte Carlo maximum likelihood calculations. *J. R. Stat. Soc. B 56*(1), 261–274.

Geyer, C. and J. Møller (1994). Simulation procedures and likelihood inference for spatial point processes. *Scand. J. of Stat. 21*(4), 359–373.

Geyer, C. and E. Thompson (1992). Constrained Monte Carlo maximum likelihood for dependent data. *J. R. Stat. Soc. B 54*(3), 657–699.

Green, P. J. (1995). Reversible jump MCMC computation and Bayesian model determination. *Biometrika 82*(4), 711–732.

Haldorsen, H. and L. Lake (1984). A new approach to shale management in field-scale models. *Society of Petroleum Engineers Journal*, 447–457.

Hastings, W. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika 57*(1), 97–109.

Higdon, D., V. Johnson, T. Turkington, J. Bowsher, D. Gilland, and R. Jaszczak (1995). Fully Bayesian estimation of Gibbs hyperparameters for emission computed tomography data. Technical report, ISDS, Duke University. To appear in IEEE-TMI.

Metropolis, N., A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller (1953). Equation of state calculation by fast computing machines. *Journal of Chemical Physics 21*, 1087–1092.

Norsk Hydro, Statoil, Saga Petroleum, The Norwegian Petroleum Directorate (1995). *Safari; Sedimentary Architecture of Field Analogues for Reservoir Information.* Norsk Hydro, Statoil, Saga Petroleum, The Norwegian Petroleum Directorate.

Stoyan, D., W. Kendall, and J. Mecke (1995). *Stochastic Geometry and Its Applications.* Wiley. 2nd ed.

Syversveen, A. and H. Omre (1997). Conditioning of marked point processes within a Bayesian framework. To appear in Scandinavian Journal of Statistics.

Tierney, L. (1994). Markov chains for exploring posterior distributions (with discussion). *The Annals of Statistics 22*(4), 1701–1762.

Weir, I. S. (1995). Fully Bayesian reconstruction from single photon emission computed tomography data. To appear in Journal of American Statistical Association.