MARKED POINT MODELS FOR FACIES UNITS CONDITIONED ON WELL DATA

ANNE RANDI SYVERSVEEN AND HENNING OMRE
Department of Mathematical Sciences
Norwegian University of Science and Technology
N-7034 Trondheim, Norway

Abstract. A marked point model for modeling of facies units conditioned on observations from an arbitrary number of wells is presented. The model includes spatial interaction between units, and it does not require specification of well to well contacts. A simulation algorithm for the model is derived and demonstrated on real life data taken from an outcrop. Realizations from the model are compared to the outcrop.

1. Introduction

The paper by Haldorsen and Lake (1984) on stochastic modeling of shale units in a sand matrix is considered to be trend-setting for stochastic reservoir characterization. The model in the paper had several shortcomings, however. Since each shale was modeled by a rectangle, shale units penetrated by more than one well could not be represented. Neither did the model include spatial interaction between shales. Moreover, the simulation algorithm seems not to be exactly correct in the sense that shales penetrated by wells tend to be too short. In Syversveen and Omre (1997) we described a model that handles these shortcomings. Here we will present an extension of this model, which no longer needs information of whether observations belong to the same facies unit or not, as was assumed in Syversveen and Omre (1997). In the current paper the term shale unit according to Haldorsen and Lake (1984) will be used. The model may be used for modeling any type of facies units in a background facies. In the example, sand units in a non-permeable background is modeled.

2. Model definition

Each shale 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 $\boldsymbol{Z}(\cdot) = (\boldsymbol{Z}_B(\cdot), \boldsymbol{Z}_T(\cdot))$, and we write

$$\boldsymbol{U} = (\boldsymbol{C}, \boldsymbol{S}, \boldsymbol{Z}(\cdot)). \tag{1}$$

The height of the rectangle is the expected height of the shale, while the length of the shale equals the length of the rectangle, see Figure 1.

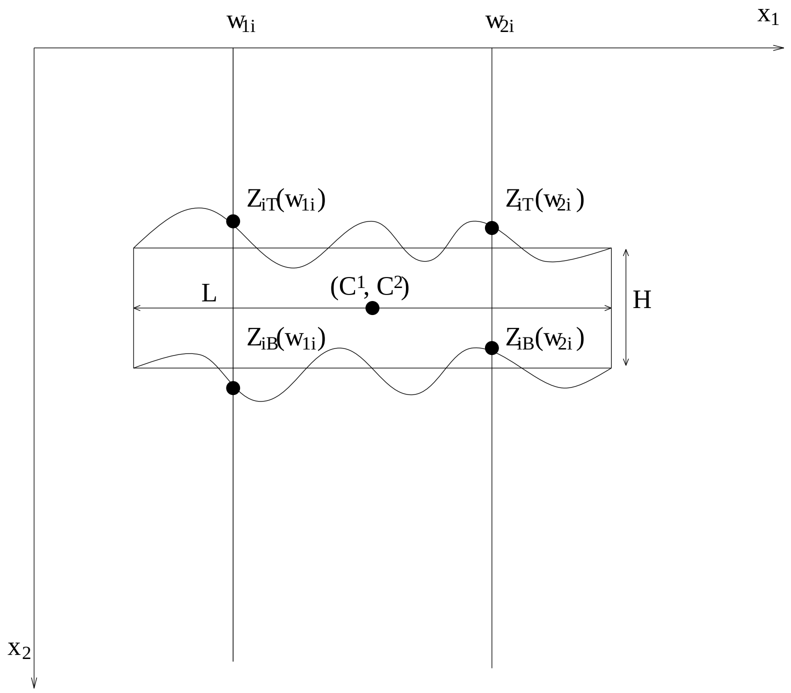

Figure 1. Model for a single shale.

The reservoir contains an unknown number of shales N, which normally interact by repulsing each other. The model for a number of shale units in the bounded reservoir $\mathcal{D}$ will be developed in a Bayesian setting.

2.1. PRIOR MODEL

We let $\boldsymbol{\theta}_i = (c_i^1, c_i^2, l_i, h_i)$ denote the rectangle associated with shale number i, and $\boldsymbol{z}_i(\cdot)$ denote its top and bottom positions for shale number i. The joint density for the N shales can be written

$$f(\boldsymbol{\theta}, n, \boldsymbol{z}(\cdot)) = f(\boldsymbol{\theta}, n) f(\boldsymbol{z}(\cdot)|\boldsymbol{\theta}, n), \tag{2}$$

where $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n)$ and $\boldsymbol{z}(\cdot) = (\boldsymbol{z}_1(\cdot), \ldots, \boldsymbol{z}_n(\cdot))$. The first term $f(\boldsymbol{\theta}, n)$ is a marked point field with pairwise interaction, see Stoyan and Stoyan (1994) of the form

$$f(\boldsymbol{\theta}, n) = \text{const} \times \exp\{\sum_{i=1}^{n} b(c_i^1, l_i) + b(c_i^2, h_i) + \sum_{j=1}^{n} \sum_{i<j} d(\boldsymbol{c}_i, \boldsymbol{c}_j) + na\}. \tag{3}$$

Here $b(c_i^1, l_i) + b(c_i^2, h_i)$ are marginal characteristics of the shales, $d(\boldsymbol{c}_i, \boldsymbol{c}_j)$ is a pairwise interaction function between center-points of shales, and the last term is related to the intensity of shales. The term $\exp(b(c_i^2, h_i))$ is defined to be a truncated bi-Gaussian pdf in order to be able to calculate parts of the posterior density analytically. The truncation is such that C^2 only takes values within the reservoir, and H is positive. The mean value of C^2 is chosen to be the center point of the reservoir in x_2 direction, and the variance will normally be chosen to be large, in order to approximate a uniform distribution over the reservoir. The mean and variance for H might be determined from the observations by an empirical Bayes approach, see Syversveen and Omre (1997). The term $\exp(b(c_i^1, l_i))$ can be chosen according to geological experience. The interaction function $d(\boldsymbol{c}_i, \boldsymbol{c}_j)$ will usually be defined such that shales repulse each other. The intensity parameter a might be estimated from the data by Monte Carlo methods, see Geyer and Thompson (1992) and Syversveen (1996).

The second term $f(\boldsymbol{z}(\cdot)|\boldsymbol{\theta}, n)$ in equation (2) is related to the top and bottom position of shale number i. The function $\boldsymbol{Z}_i(\cdot)$ will, conditioned on $\boldsymbol{\Theta}_i = \boldsymbol{\theta}_i$, be a second order stationary Gaussian random function with

$$E(\boldsymbol{Z}_i(x_1)|\boldsymbol{\Theta}_i = \boldsymbol{\theta}_i) = \begin{pmatrix} c_i^2 + h_i/2 \\ c_i^2 - h_i/2 \end{pmatrix} \tag{4}$$

and

$$Cov\ (\boldsymbol{Z}_i(x_1'), \boldsymbol{Z}_i(x_1'')|\boldsymbol{\Theta}_i = \boldsymbol{\theta}_i) = \begin{pmatrix} \sigma^2 R(x_1', x_1'') & \tau R(x_1', x_1'') \\ \tau R(x_1', x_1'') & \sigma^2 R(x_1', x_1'') \end{pmatrix} \tag{5}$$

for $x_1 \in [c_i^1 \pm l_i/2]$. The matrix $R(x_1', x_1'')$ is given by

$$R(x_1', x_1'') = \begin{pmatrix} 1 & \rho(x_1' - x_1'') \\ \rho(x_1' - x_1'') & 1 \end{pmatrix},$$

σ^2 is the variance, τ is the covariance between top and bottom, and $\rho(\cdot)$ is the spatial correlation function. In order to ensure positive definiteness of the covariance matrix, we must have $|\tau| < \sigma^2$. Note that the expected top and bottom are symmetrical around C^2. The correlation function, the variance and the covariance are chosen according to geological experience.

Assuming that the stochastic fields $\boldsymbol{Z}_i(\cdot)|\boldsymbol{\Theta}_i = \boldsymbol{\theta}_i$ in different shales are conditionally independent, the following probability density is obtained for a reservoir containing N shales:

$$f(\boldsymbol{\theta}, n, \boldsymbol{z}(\cdot)) = f(\boldsymbol{\theta}, n) \prod_{i=1}^{n} f(\boldsymbol{z}_i(\cdot)|\boldsymbol{\theta}_i). \tag{6}$$

2.2. OBSERVATIONS AND THE LIKELIHOOD FUNCTION

The observations consist of the dimension of the rectangular reservoir and exact observations of the reservoir along the transects of m wells. Observations of depths to top and bottom of shales are contained in the vector $\boldsymbol{o}$. The location of the wells are denoted w_i, $i = 1, \ldots, m$. It is assumed that no overlap between shales is observed in the wells, and that it is unknown whether observations belong to the same shale or not. Hence, the likelihood may be written as an indicator function, taking the value one when conditioning is fulfilled and zero otherwise.

With this form of the likelihood function, it is impossible to do analytical calculations on the posterior distribution, since well to well contacts are unknown. We will therefore introduce a new stochastic variable T, which is implicitely defined by $\boldsymbol{\Theta}$ and tells us which observations belong to the same shale. We call T the configuration of shales. To explain better what T is, we look at the three realizations in Figure 2, which have different configurations. For example, the three observations belonging to the same object in the "true" formation, appear as one, two, or three objects in the realizations from the posterior. Having introduced T, the likelihood function can be written

$$f(\boldsymbol{o}|\boldsymbol{\theta}, n, \boldsymbol{z}(\cdot)) = \text{const} \times \prod_{i=1}^{n} I(\boldsymbol{o}_i|\boldsymbol{\theta}_i, \boldsymbol{z}_i(\cdot), t), \tag{7}$$

where $I(\boldsymbol{o}_i|\boldsymbol{\theta}_i, \boldsymbol{z}_i(\cdot), t)$ equals one when the conditioning in shale number i is fulfilled and zero otherwise. The reason why t is present on the right hand side of equation (7), is that the knowledge of one shale unit only is not sufficient for identifying the configuration t.

The posterior model is defined by the prior, equation (2) and the likelihood, equation (7). However, we want to simulate $(\boldsymbol{\theta}, n)$ and $\boldsymbol{z}$ separately.

Therefore we will define the posterior pdf in the form of equation (9), and we look at the likelihood function for $(\boldsymbol{\Theta}, N)$,

$$\begin{aligned} f(\boldsymbol{o}|\boldsymbol{\theta}, n) &= \text{const} \times \prod_{i=1}^{n} f(\boldsymbol{o}_i|\boldsymbol{\theta}_i, t) \\ &= \text{const} \times \prod_{i=1}^{v_t} f(\boldsymbol{o}_i|c_i^2, h_i, t) \prod_{j=1}^{n} I(\boldsymbol{o}_j|c_j^1, l_j, t). \end{aligned} \tag{8}$$

Here $f(\boldsymbol{o}_i|c_i^2, h_i, t)$ is the likelihood related to the height observations in wells penetrating shale number i under configuration t, and $I(\boldsymbol{o}_j|c_j^1, l_j, t)$ is an indicator function ensuring that the lateral conditioning is honored. The number v_t is the number of penetrated shales within configuration t. Since the observations are without error, $f(\boldsymbol{o}_i|c_i^2, h_i, t)$ is directly related to $f(\boldsymbol{z}_i(\cdot)|c_i^2, h_i)$ and is therefore a multi-Gaussian pdf. The second equality follows from the fact that the horizontal and vertical dimensions in each shale are independent in the prior distribution.

In order to be able to calculate the posterior model analytically, it is important that the likelihood function can be factorized in conditionally independent parts, as we will see in the following sections.

2.3. POSTERIOR MODEL

Analogous to equation (6), the posterior model for $(\boldsymbol{\Theta}, N, \boldsymbol{Z}(\cdot))$ can be written as

$$f(\boldsymbol{\theta}, n, \boldsymbol{z}(\cdot)|\boldsymbol{o}) = f(\boldsymbol{\theta}, n|\boldsymbol{o}) \prod_{i=1}^{n} f(\boldsymbol{z}_i(\cdot)|\boldsymbol{\theta}_i, \boldsymbol{o}). \tag{9}$$

The first term is derived as follows:

$$\begin{aligned} f(\boldsymbol{\theta}, n|\boldsymbol{o}) &= \text{const} \times f(\boldsymbol{\theta}, n) f(\boldsymbol{o}|\boldsymbol{\theta}, n) \\ &= \text{const} \times \exp\{\sum_{i=1}^{n} (b(c_i^1, l_i|\boldsymbol{o}, t) + b(c_i^2, h_i|\boldsymbol{o}, t)) \\ &\quad + \sum_{j=1}^{n} \sum_{i<j} d(\boldsymbol{c}_i, \boldsymbol{c}_j) + na\} I(\boldsymbol{o}|t, n). \end{aligned} \tag{10}$$

The term $\exp(b(c_i^2, h_i|\boldsymbol{o}, t))$ is proportional to a bi-Gaussian pdf. Details can be found in Syversveen and Omre (1997). The term $\exp(b(c_i^1, l_i|\boldsymbol{o}, t))$ is equal to $\exp(b(c_i^1, l_i)) I(\boldsymbol{o}|c_i^1, l_i, t)$.

The conditional pdf for the geometry of shale number i, $f(\boldsymbol{z}_i(\cdot)|\boldsymbol{\theta}_i, \boldsymbol{o})$, consists of bi-Gaussian random functions with given parameters and conditioned on observations in a number of locations, see Journel and Huijbregts (1978).

3. Simulation from the model

The simulations are done in two steps. First the parameters $\boldsymbol{\theta}$ and the number of shales n are sampled from the density defined by equation (10) by using Metropolis-Hastings algorithm, see Metropolis *et al.* (1953) and Hastings (1970). Then the conditional Gaussian random fields $\boldsymbol{z}_i(\cdot)|\boldsymbol{\theta}_i, \boldsymbol{o}$ are sampled using conditional geostatistical simulation, see Journel and Huijbregts (1978).

The transition kernel in the Metropolis-Hastings algorithm consists of three different types of transitions; adding an unobserved shale, deleting an arbitrary unobserved shale, and changing observed shales, including a change of configuration.

When adding a new shale, the proposed new shale is drawn from the marginal prior distribution of the corresponding homogeneous marked point process. When changing observed shales, a configuration is drawn uniformly over the possible configurations. Given this configuration, each shale influenced by the change in configuration is drawn from a marginal distribution conditioned on the data as described in Syversveen and Omre (1997). To draw the shales from the conditional marginal distribution is crucial for the convergence of the algorithm.

The acceptance probabilities in the Metropolis-Hastings algorithm are defined by the posterior pdf to be sampled from and these acceptance kernels, and they are computed according to Syversveen and Omre (1997). The convergence of the algorithm do not pose critical problems since the model is defined in a relatively low dimensional space, see Syversveen and Omre (1997) for further discussion.

4. Simulation Example

In this example we model a part of the outcrop taken from the Menefee formation in the San Juan basin, Rocky Mountains, see Safari (1995), shown in Figure 2a. The outcrop is from a low to moderate sand-gross delta plain environment, and shows a cross section of meandering channel sandstones embedded in lower to upper delta plain mudstone, siltstone and coal. Hence the objects are sand and the background is nonpermeable shale. This is regarded as our "true" formation, and three artificial wells are added and shown as vertical lines in Figure 2a. Taking the observations from the wells into account, we generate realizations from the posterior model and compare them with the truth. Two laterally extensive objects are shown in a darker color than the rest. These are not modeled, because they have a geologic interpretation which is different from the objects in the model.

The size of the formation is rescaled and taken to be 100×16 units, where one unit is about 11 meters. We have chosen the following prior distributions for the parameters $\mathbf{\Theta} = (C^1, C^2, L, H)$, see equation (3). We let $\exp(b(c^1, l)) = \exp(b(c^1) + b(l))$ and $\exp(b(c^2, h)) = \exp(b(c^2) + b(h))$; that is, C^1, C^2, L and H are independent. As mentioned in section 2.1, $\exp(b(c^2, h))$ is a truncated bi-Gaussian pdf. The mean and variance of C^2 are set to $\mu_{C^2} = 8.0$ and $\sigma^2_{C^2} = 12^2$ respectively. With these parameter choices the truncated Gaussian distribution approximates a uniform distribution over $[0, 16]$. Mean and variance of H are estimated by empirical Bayes calculations, and we get $\mu_H = 0.63$ and $\sigma^2_H = 0.02$. We let $\exp(b(c^1))$ be a uniform distribution over the length of the reservoir and $\exp(b(l))$ be a mixture of two Gaussian distributions truncated at zero with parameters $\mu_{L1} = 25$, $\sigma^2_{L1} = 15^2$, $\mu_{L2} = 4$, and $\sigma^2_{L2} = 2^2$. This mixture distribution is chosen because of the large variability of the length of sand bodies. Geologists have been consulted in order to chose the parameter values in this mixture distribution.

The interaction function is defined to be

$$d(x_i, x_j) = \begin{cases} \ln(\frac{|x_i - x_j|}{15.0}) & \text{if } |x_i - x_j| \leq 15.0 \\ 0 & \text{if } |x_i - x_j| > 15.0 \end{cases}$$

which gives a weak repulsion between objects.

For the Gaussian fields on top and bottom of the objects we use correlation function $\rho(x) = \exp(-0.03x^2)$, variance $\sigma^2 = 0.38$ and the covariance $\tau = 0.05$ between top and bottom of objects.

The intensity parameter a is set to 7.0, which gives a reasonably good fit to the number of objects. In order to get a better model fit, this parameter could be estimated from the data, for example by Monte Carlo Maximum Likelihood, see Geyer and Thompson (1992) and Syversveen (1996).

In Figure 2 the “true” formation is shown together with three realizations from our model. Observe that the conditioning is fulfilled in all wells. We see that the sand bodies observed in more than one well is represented as one or more sand bodies in the realizations from the model.

5. Closing Remarks

The current paper contains an extension of the model presented in Syversveen and Omre (1997). In these papers, the traditional Haldorsen and Lake model is extended to cover facies units with spatial interaction and multi-well conditioning. The current model is demonstrated on real data from an outcrop and appears to reproduce the real geologic features reasonably well.

Figure 2. The "true" formation together with three different realizations from the model.

6. Acknowledgements

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

References

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

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

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

A.G. Journel and Ch.J. Huijbregts. *Mining Geostatistics.* Academic Press, 1978.

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

Safari – Sedimentary Architecture of Field Analogues for Reservoir Information. Norsk Hydro, Statoil, Saga Petroleum, Norwegain Petroleum Directorate, 1995.

D. Stoyan and H. Stoyan. *Fractals, random shapes and point fields.* John Wiley & Sons, 1994.

A.R. Syversveen. An approximate fully Bayesian estimation of the parameters in a reservoir characterization model. In preparation, 1996.

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