NORGES TEKNISK-NATURVITENSKAPELIGE
UNIVERSITET

Seismic reservoir prediction using Bayesian integration of rock physics relations and Markov random field models. A North Sea example.

by

Jo Eidsvik, Tapan Mukerji, Henning Omre, Per Avseth
and Gary Mavko

PREPRINT
STATISTICS NO. 2/2002

NORWEGIAN UNIVERSITY OF SCIENCE AND
TECHNOLOGY
TRONDHEIM, NORWAY

This report has URL
`http://www.math.ntnu.no/preprint/statistics/2002/S2-2002.ps`
Jo Eidsvik has homepage: `http://www.math.ntnu.no/~joeid`
E-mail: `joeid@stat.ntnu.no`
Address: Department of Mathematical Sciences, Norwegian University of Science and Technology, N-7491 Trondheim, Norway.

Seismic reservoir prediction using Bayesian integration of rock physics relations and Markov random field models. A North Sea example.

Jo Eidsvik and Henning Omre
Department of Mathematical Sciences,
Norwegian University of Science and Technology, Trondheim, Norway
Tapan Mukerji and Gary Mavko
Stanford Rock Physics Lab, Stanford, California
Per Avseth
Norsk Hydro, Bergen, Norway

Quantitative interpretation of seismic data for reservoir rock and fluid properties most commonly relies on deterministic rock physics relations and may often neglect statistical variability or spatial correlation. Geostatistical methods, on the other hand, take into account the spatial dependence, but results are often obtained without properly formalizing rock physics relations between seismic observables and rock and fluid properties. This paper presents a strategy for integrating both deterministic rock physics relations and spatial, statistical representations within a Bayesian framework.

Data from the turbidite Glitne field in the North Sea(Figure 1) provides an example for the outlined approach. The reservoir is very heterogeneous, and merging statistics with rock physics is hence particularly worthwhile in order to obtain better seismic reservoir prediction. Focus is on describing the spatial probability distribution of reservoir facies and fluid saturations along the seismic horizon representing the top Heimdal formation(Figure 1), which is capped by the Lista formation at approximately 2000m depth. Available data include mainly AVO attributes(zero offset reflectivity and AVO gradient) extracted from $3-D$ pre-stack seismic data and log based observations of facies and fluid saturation in four wells. Figure 2 shows observations of zero offset reflectivity(top, left) and AVO gradient(bottom, left) along the horizon, with the facies and fluid observations in the four wells indicated, a cross plot of zero offset reflectivity and AVO gradient(top, right) and an example of a well log from the area(bottom, right). Further, two way seismic travel times to the horizon of interest and cap rock properties(density, P- and S wave velocity) from well logs are also available(Figure 1). A grid of size 245×505 with each gridblock being $(12.5 \times 12.5)m^2$ covers the domain of interest. This corresponds to the seismic inline-xline grid in Figure 2. Facies and fluid saturations are represented on the same grid. In the example facies at each location classifies as either shale or sand, fluids as either brine or oil.

The model is presented via a Bayesian network, relating the reservoir variables to the available data by physical relations. Parameters of the probability distributions are calculated from mineralogical values and well data. The Bayesian approach consists of two steps: i) Define a prior model based on general physical knowledge and experience. This constitutes a qualified guess of what the reservoir might look like before integrating the data. ii) Define likelihood models to tie the reservoir variables(e.g. facies and fluid) to the reservoir specific observations(e.g seismic data and well logs). These steps together define the posterior probability distribution of the reservoir variables conditional on the data. The posterior probability distribution is explored by the Markov chain Monte Carlo algorithm in this paper.

Bayesian modeling of reservoir variables using rock physics. The reservoir variables of main interest are facies(Q) and fluid saturation(S) at the seismic horizon in Figure 1. Porosity(ϕ), density(ρ), P-(V_P) and S-wave velocity(V_S) are also included since they link the reservoir rock to the seismic data. All reservoir variables and available data communicate through forward models as illustrated in the network in Figure 3, with the physical relations presented in Table 1. Subscript x implies spatially varying quantities. Seismic AVO attributes, zero offset reflectivity and AVO gradient, are grouped together and denoted by $d_{s,x}$, indicating spatially varying seismic data. An arrow between two variables in Figure 3 symbolizes that the variables are directly dependent. Arrows in the network also define conditional independencies. As an example, given density, P- and S-wave velocity, the seismic AVO attributes are conditionally independent of facies, fluid saturation and porosity. This is evident from the various equations for AVO intercept and gradient such as Shuey's approximation or Aki-Richards' approximation. As another example, note that there exist arrows from S_x to $V_{P,x}$ and ρ_x, while the effect on S_x to $V_{S,x}$ goes only through ρ_x. This can be understood by low frequency fluid effects(Gassmann) which say that fluid saturation effects P-wave velocity and density, but not shear wave modulus. Hence $V_{S,x}$ is conditionally independent of S_x. Double arrows represent a deterministic relationship while single arrows represent stochastic relationships. Hence the well observations of facies and fluid saturation, denoted d_w, are assumed to be exact observations.

Arrows within the reservoir box in Figure 3 define the prior distribution of the reservoir variables. The arrows from the reservoir box to the reservoir specific observations box define the likelihood models. The Bayesian network model, and in particular the posterior probability distribution of the reservoir variables given the reservoir specific observations, is fully specified by the arrows in Figure 3 and their associated physical relations described in Table 1. Parameter values used in Table 1 are taken from mineralogical values, extracted from wells in the area and infered from prior knowledge. The relations are deterministic in the physical variables with an additive stochastic term which includes intrinsic variability for each facies as well as errors due to model simplifications. Formulating the model in a stochastic framework based on recognized physical relations seems reasonable, and offers a strategy to make the best use of deterministic and statistical techniques.

Markov random fields as prior distributions for facies and fluid saturation. In Table 1, facies and fluid saturation are defined by Markov random fields a priori. Markov random fields impose spatial continuity since the probability distribution of a variable at one location depends on the variable at neighboring locations. Markov random fields are generalizations of time domain Markov chains to the spatial domain.

As a simple $1-D$ example, consider a first order binary Markov chain. The variable X at depth z takes on one out of two possible states, and its probability distribution depends on the earlier variables only through the variable at depth $z-1$, i.e. $Pr(X(z)|X(z-1), X(z-2), \ldots) = Pr(X(z)|X(z-1))$. A Markov transition matrix defines the conditional probabilities at depth z given the variable at depth $z-1$. Suppose there are two possible states, shale(1) and sand(2). Figure 4(right) shows realizations of Markov chains with different transition matrices P. In Figure 4 the Markov transition matrices are presented on top of the realizations and the entries of the Markov transition matrix explained to the left. The leftmost realization has a probability of staying in either state equal to 0.9, and the probability of jumping to the other state is 0.1. In this Markov chain few jumps between the two states take place. This might represent for example sequences of thick, blocky sands and shales. The realization in the middle has a probability of staying in either state equal to 0.6, and the jump probabilities are 0.4. Fluctuations between the states are more common in this realization than in the leftmost one. This would be representative of thin bedded sand-shale sequences. The rightmost realization stays in shale with probability 0.9 and stays in sand with probability 0.6. This chain tends to stay in the shale state a long time before it switches to sand. The shale beds are thick, while the sand beds are thin in comparison. Simple Markov chains like this can capture a large variety of facies bed thicknesses, relative proportions, and spatial correlations.

In this study a first order binary Markov random field is used for both facies and fluid saturation as indicated in Table 1. Similar as in a Markov chain, given all other values of the Markov random field, the probability distribution of a variable at one location depends only on the values at neighboring locations as indicated in Figure 5(top). The probability of sand at any location on the grid is in this way a function of the number of sand neighbors, a number that takes one possible value out of $\{0, 1, 2, 3, 4\}$. For fluid saturation a trend depending on the seismic reflection times is added, such that the probability of oil at any location depends on the reflection time at that location in addition to the number of oil neighbors. Oil is hence modeled more likely to occur in shallow zones to take into account that oil usually overlies brine. The dependency on seismic reflection time is illustrated by an arrow in Figure 3.

Drawing realizations from the posterior probability distribution of the reservoir variables by McMC. High dimensionality of the reservoir variables represented on the 245×505 grid and non-linear relations in the network, make the posterior probability distribution of the reservoir variables analytically intractable. However, ratios of the probability distribution at different variables are computable, and this makes it possible to explore the posterior probability distribution by Markov chain Monte Carlo(McMC) sampling. McMC algorithms are iterative and they converge to realizations from the desired posterior probability distribution, regardless of the initial value. Starting with a random spatial distribution of reservoir variables, at each step of the McMC algorithm the reservoir variables are updated according to a proposal and acceptance strategy as follows:

1. Pick a gridnode at random.
2. Propose new reservoir variables to replace the current ones in this gridnode.
3. Draw a uniform random number between 0 and 1.
4. If the ratio of the posterior probability distribution(calculated using relations from Figure 3 and Table 1) at the proposed and the current reservoir variables is larger than the uniform random variable from step 3, accept the proposed variables, otherwise retain the current ones.

One iteration of the McMC algorithm is defined as $Ngrid$ such steps(where $Ngrid = 245 \times 505$ in this example), and on average each gridnode is visited once per iteration. The probability of accepting the proposed variables, given by step 4 above, depends on the seismic AVO attributes at the picked gridnode and the neighboring variables of facies and fluid saturation, in addition to the well observations. Figure 5 illustrates the idea with the neighborhood of a first order Markov random field(top), and with tentative probability density contours in the AVO plane for brine-sand, oil-sand and shale(bottom). Suppose all neighbors are oil-sand and that the observed AVO attributes fall within the oil-sand contour. If oil-sand is the proposed variable, it will surely be accepted in step 4 because the posterior probability for this is larger. If shale is proposed, it might still be accepted in step 4, but the probability of this is very small. The McMC algorithm produces a Markov chain of reservoir variables which converges to samples from the required posterior probability distribution when the number of iterations go to infinity. For practical purposes one monitors convergence of some attributes, and hence observes what number of iterations seems sufficient.

Figure 6 shows proportions of sand(top) and the facies variables(bottom) at inline 185 of the grid after each iteration of the McMC algorithm for a total of 750 iterations. In Figure 6(bottom) white denotes sand and black denotes shale. The initial value of the McMC algorithm equals shale in every gridnode, and one observes how the facies distribution changes as a function of iteration number of the McMC algorithm. After approximately 300 iterations convergence seems to have taken place, and the fluctuations hereafter represent the variability in the posterior probability distribution of facies at inline 185. At well locations the facies variables remain fixed since well observations are assumed to be exact. Fast fluctuations of the Markov chain means that the chain is mixing well, moving around in the sample space quickly. If the chain mixes well, fewer iterations are needed to describe the posterior probability density properly, and hence less computation time is required. In this study, with only one gridnode updated in each step, mixing is relatively slow, but at the same time the acceptance rate of proposed variables is large. Running the algorithm for about 1000 iterations on the 245×505 grid seems enough to represent the posterior probability distribution for most purposes, and takes about $5 - 10$ hours in C code on our modest desktop computers.

Estimation and uncertainty evaluation from multiple realizations. From the multiple realizations obtained by the McMC algorithm, probabilistic properties of facies and fluid saturation along the top Heimdal horizon are characterized. Figure 7 shows the estimated marginal maximum a posteriori value of facies and fluid saturation along the top Heimdal horizon in a combined plot with shale(black), oil-sand(gray) and brine-sand(white). This is obtained by presenting the facies and fluid saturation variables that occur most often at each location. The result is a smooth field because of the spatial dependence imposed by the Markov random field model. Figure 8 shows an estimated oil-sand probability map along the top Heimdal horizon. This is obtained by counting the proportion of oil-sand realizations at each location.

Note that oil-sand is predicted at the structural top of the turbidite system, since it is conditioned on reflection times. It is also conditioned on the seismic AVO and well data, and the shale dominated area in the south-west is identified by the AVO attributes. The lobe and the feeder channel of the turbidite system are visible from the oil-sand probability map in Figure 8.

An important advantage of this stochastic approach is that even though the original, often non-linear, rock physics relations are used, the McMC algorithm yield multiple realization from the posterior probability distribution. From these realizations functionals of interest, such as sand proportions, oil saturation, etc, can be estimated with associated uncertainty levels.

Conclusions. Inspired by seismic and well log data from the top Heimdal horizon of the Glitne field in the North Sea, this paper combines rock physics with spatial statistics in a Bayesian framework. Reservoir variables and the reservoir specific observations are integrated in a stochastic network with rock physics relations as the main ingredient, and with spatial continuity incorporated by Markov random fields.

Multiple realizations obtained from the posterior probability distribution by the Markov chain Monte Carlo algorithm probabilistically characterize the facies and fluid saturation distribution along the $2-D$ Heimdal surface.

Integrating rock physics and statistics as outlined in this paper can have a significant effect on results since it is consistent with both rock physics and spatial, statistical models. Most forward models are approximations, however, in this paper the framework exists for adding more complicated physics and for modeling other probability distributions than the Gaussian. A disadvantage of the outlined strategy is an increase in computational time, especially if the model is to be extended to $3-D$ or $4-D$ examples, or with integration of other data sources, such as production history. Effective modeling and fast McMC strategies are hence very important.

Suggested reading

"Seismic reservoir mapping from 3D AVO in a North Sea turbidite system" by Avseth et al.(GEOPHYSICS, 2001, pp 1157-1176). "Mapping lithofacies and pore fluids probabilities in a North Sea reservoir: Seismic inversion and statistical rock physics" by Mukerji et al.(GEOPHYSICS, 2001, pp 988-1001). "On the statistical analysis of dirty pictures" by Besag(Journal of Royal Statistical Society, B, 1986, pp 259-302). "Markov chain Monte Carlo in practice" by Gilks et al.(Chapman-Hall, 1996). "Computer simulation in Geology" by Harbaugh and Bonham-Carter(Wiley, 1970).

Corresponding address: joeid@math.ntnu.no

Acknowledgements

The study was financially supported by the URE initiative at NTNU, the Stanford Rock Physics Project and from the Stichting Foundation, Schlumberger. Jo Eidsvik is a PhD student funded by the Norwegian Research

Counsil. He is visiting Department of Statistics, Stanford and Stanford Rock Physics Laboratory in 2001. The authors are grateful to Norsk Hydro and Statoil for providing the data.

Figure captions

FIGURE 1: The location of the Glitne field in the North Sea and the top Heimdal horizon as illustrated by reflection times. The available data are AVO seismic attributes, well observations, two way seismic travel times and cap rock properties.

FIGURE 2: Attributes from AVO seismic data include zero offset reflectivity(top, left) and AVO gradient(bottom, left), cross plotted(top, right). Well locations and the observations(left) are assumed to be exact observations extracted from well logs. An example of a well log(bottom, right).

FIGURE 3: Network relating the reservoir variables; Facies(Q), fluid saturation(S), porosity(ϕ), density(ρ), P- (V_P) and S- wave velocity(V_S), and the reservoir variables to the data; Reflection times(t_x), well data(d_w), AVO seismic data($d_{s,x}$) and cap rock properties(cr). x denotes location, and the network holds for every location.

FIGURE 4: Three facies log realizations with various Markov transition matrices(right) illustrate the diversity of Markov chain models. The definition of the Markov transition matrix(right).

FIGURE 5: The updating step of the McMC algorithm takes into account the spatial dependence in the Markov random field(top) and the similarity to the observed AVO seismic attributes as sketched by tentative probability density contours for shale(red), brine-sand(green) and oil-sand(blue) in the AVO space(bottom).

FIGURE 6: Proportion of sand at inline 185 in the grid(top) and facies variable at inline 185(bottom) as a function of iteration number in the McMC algorithm, with sand(white) and shale(black). The McMC algorithm start out with shale at every gridnode initially.

FIGURE 7: Marginal maximum a posteriori estimate based on multiple realizations from the posterior probability distribution obtained by the McMC algorithm with shale(black), brine-sand(white) and oil-sand(gray).

FIGURE 8: Probability map of the oil-sand configuration estimated from the McMC realizations. The lobe and feeder channel of the turbidite system are illustrated.

Table captions

TABLE 1: The table gives the rules and representations used in the network. Spatial representations are used for facies(Q) and fluid saturation(S). The other variables are defined through S and Q by rock physics theory. Parameters are extracted from mineralogical values, well data and prior knowledge.

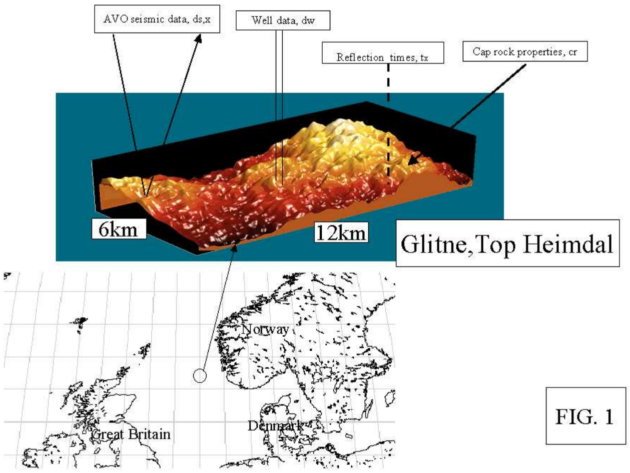

FIG. 1

Figure 1:

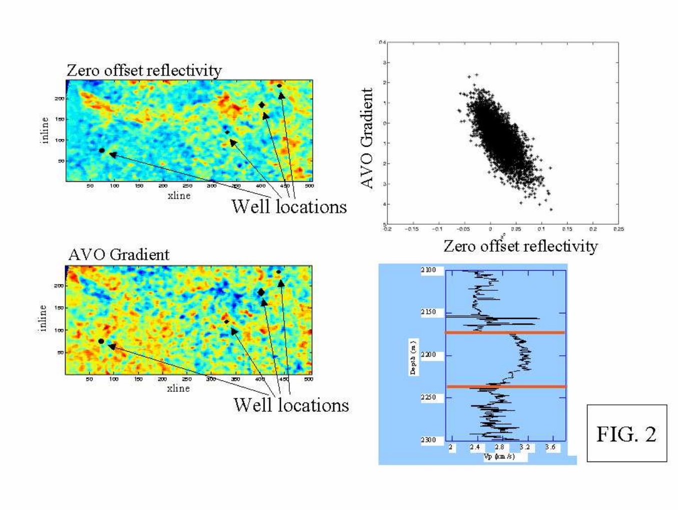

Figure 2:

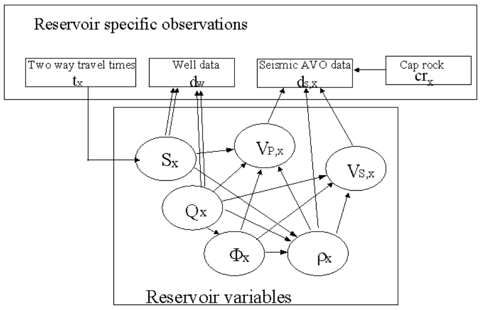

Figure 3:

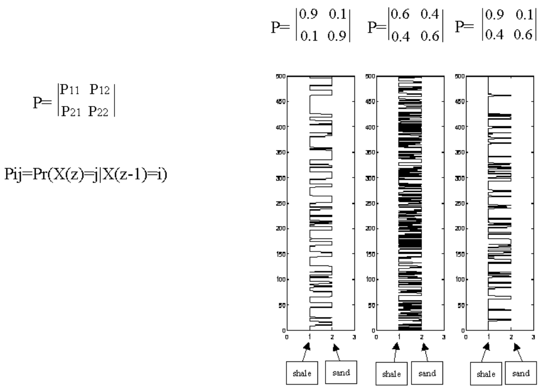

FIG. 4

Figure 4:

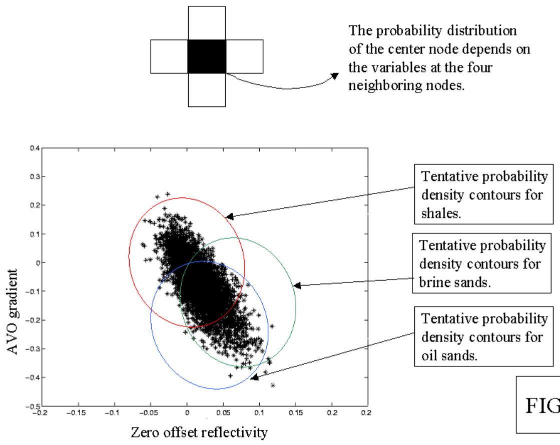

FIG. 5

Figure 5:

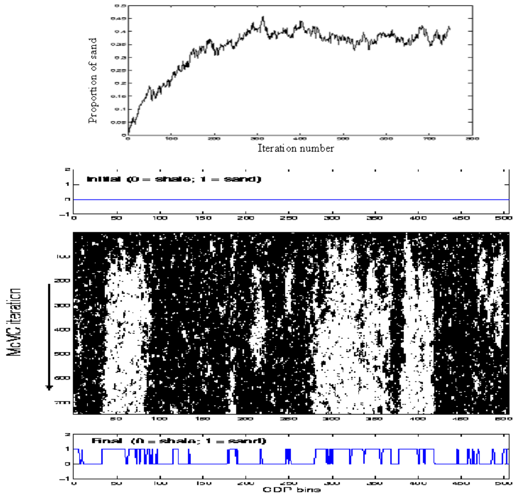

FIG. 6

Figure 6:

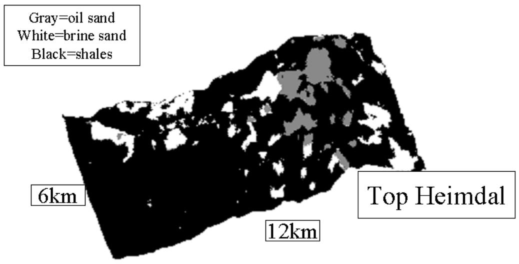

FIG. 7

Figure 7:

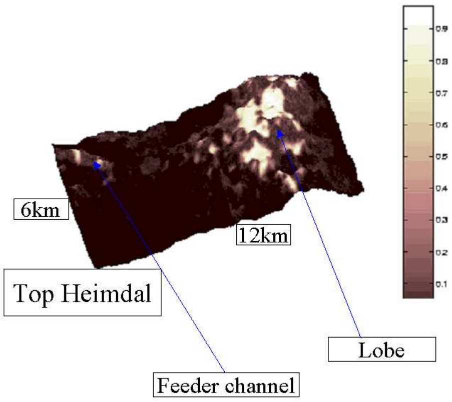

Figure 8:

<u>Facies:</u>
$[Q_x]$=Binary Markov random field; (shale, sand)
<u>Fluid saturation:</u>
$[S_x \mid t_x]$=Binary Markov random field; (brine, oil)
<u>Porosity:</u>
$[\Phi_x \mid Q_x]$=Gaussian, mean 0.3(shale) and 0.25(sand)
<u>Density:</u>
$[\rho_x \mid \Phi_x, Q_x, S_x]$=Gaussian, mean $\Phi_x \rho(S_x)+(1-\Phi_x)\rho(Q_x)$
<u>P wave velocity:</u>
$[Vp_x \mid \rho_x, \Phi_x, Q_x, S_x]$=Gaussian, mean from Vp-Vs, porosity relations from well logs and Gassmann's relation for fluid dependence.
<u>S wave velocity:</u>
$[Vs_x \mid \rho_x, Q_x]$=Gaussian, mean from Vp-Vs, porosity relations from Well logs and Gassmann's relation for fluid dependence.

Prior

<u>Seismic reflection times:</u>
t_x , exact
<u>Cap rock properties:</u>
cr , exact
<u>Well observations:</u>
d_w , exact measurements of Q and S in four wells
<u>Seismic AVO attributes (forward model):</u>
$[D_{s,x} \mid V_{P,x}, V_{s,x}, \rho_x, cr]$ = Gaussian, mean from Shuey's equations.

Likelihood

TABLE 1

Figure 9: