Seismic inversion for fluid detection, petrophysics, and stratigraphy using Bayesian models and Markov Chain Monte Carlo methods

James Gunning[1] and Michael Glinsky[2]
(1) Australian Petroleum Cooperative Research Centre
CSIRO Petroleum, PO Box 3000, Glen Waverley, Victoria, Australia 3150
E-mail: James.Gunning@csiro.au
(2) BHP Petroleum (Americas), 1360 Post Oak Boulevard
Houston, Texas, USA 77056-3020
E-mail: Michael.E.Glinsky@BHPBilliton.com

Summary

Seismic reflection data is sensitive to the presence of hydrocarbons in many kinds of reservoir rocks, as well as the spatial distribution of acoustic properties in the subsurface. Assessment of the economic worth of reservoirs in the early stages of development - when seismic data are most pertinent - should seek to exploit this fact. The maximum information content from seismic data should therefore be sought using inversion methods that model fluid distributions as well as structural and petrophysical variability. Such an inversion method will most naturally be model-based, i.e. allowing a continuum of possible reservoir structures and fluid distributions, from which a forward model of the seismic acquisition process can be generated (i.e. synthetic seismic data).

Nonetheless, inversion of seismic data is well known to be highly non-unique in the absence of constraining information. Fortunately, most early-development situations have enough well data to allow the formulation of a probabilistic prior model for the variability of acoustic properties in the pertinent rock types, and also approximate structural information derived from seismic picks.

In a Bayesian formulation, such prior expressions for the rock acoustic properties and structural uncertainties can then be combined with likelihood functions, which express the probability of various reservoir models as a function of the proximity of synthetic seismic data to observed data. The product of the prior and likelihood then forms the posterior distribution, which is the natural object from which to draw economic inferences about the reservoir.

Such inferences are invariably computed from realisations drawn from the posterior distribution, which must be generated using Markov Chain Monte Carlo techniques for problems as complex as this. The novel and challenging feature of multi-fluid prior models is that the probability space is a mixture of integer variables (which types of fluid - and in which rocks?) and continuous ones (layer surfaces, porosities, velocities etc). Drawing realisations from these high-dimensional, multiple-alternative models must be done carefully. Sampling from such models can be performed using both random-walk Metropolis algorithms and Metropolis-Hastings techniques.

The output of this stochastic inversion techniques are suites of realisations which identify the petrophysical properties (saturation, porosity, net-to-gross etc), structural properties (layer-thickness etc) and fluid types at each seismic trace location. Economic measures of reserves in place are directly calculable from these quantities.

A Bayesian vignette: treatment of information derived from seismic with multiple alternative models

If we have a probabilistic model M_k of the reservoir's structural, acoustic and petrophysical properties, whose uncertainty is specified by the prior distribution $P(M_k)$*, the posterior uncertainty in these parameters after taking into account the seismic data* D *is given by*

$$\Pi(M_k) \sim L(M_k|D)P(M_k)$$

where $L(M_k|D)$ *is the likelihood of the model* M_k*, given the seismic data. About the simplest physically reasonable form this likelihood can take is a sum over stacks using the 1D convolutional model:*

$$L(M_k|D) = \exp\left[-\sum_{\{\text{stacks } j\}} \left(\mathbf{S}_j(t)-\mathbf{w}_j(t)*\mathbf{R}_j(t)\right)^2 / 2\sigma_j^2\right]$$

where, for the jth stack, $\mathbf{S}_j(t)$ *is the seismic data,* $\mathbf{R}_j(t)$ *is the reflectivity for the appropriate incidence angle,* $\mathbf{w}_j(t)$ *is the wavelet, and* σ_j^2 *is the noise level. Here we have assumed Gaussian noise, but other forms are possible. This expression acts to force the synthetic seismic for all stacks to approximate the measured data* D*, within an acceptable error* σ_j^2*. The seismic error in the sum is acquired at approximately the Nyquist rate computed from the bandwidth.*

Some prior models k may not produce good synthetics within the scope defined by the prior (e.g. if the model postulates gas in a certain layer, resulting in a large reflected amplitude, but this is not evident in the observed data). The Bayesian scheme here will reduce the posterior probability of this model via the combined operation of the likelihood and prior distributions.

A) Construction of a prior model for the rock properties and horizons at a trace location

At each seismic trace location, the reservoir and adjacent layers are approximated as a local "layer-cake" model, comprising a set of rock layers (fig.1, part b). In general, each rock layer is a finely-laminated mixture of a reservoir rock (permeable) and non-reservoir rock (impermeable). As a simple example, pure shales or sands are endpoint instances of this mixture. Each facies used to construct the model will have available a set of regional trend curves (fig. 2), in which the variations of velocity (both p and s), porosity, and density with depth are modelled probabilistically. The net-to-gross (NG) of laminated layers is given a prior normal distribution based on regional observations or appropriately vague prior beliefs.

The permeable rocks in certain layers may have had their in-situ brine displaced by hydrocarbons. Possible hydrocarbons are oil, gas, or immobile "low-saturation gas", and appropriate prior probabilities of the occurrence of these phases in the permeable layers can be proposed, subject to various density-ordering criteria for the fluids (e.g. we may forbid brine above oil). The set of possible fluids in each layer may then be enumerated, yielding a set of prior models with associated prior probabilities. Fluid properties of the in-situ brines and possible hydrocarbons are modelled using regional data, allowing for possible uncertainties using normal approximations. Prior distributions for the layer times are developed from picked horizons on the seismic data sets (figure 3), with the priors for thin (barely resolvable) layers typically constructed using various extrapolation routines (e.g. kriging). Layer-thicknesses in depth may also be constrained by well and kriging information: such constraints act on the layer-times bounding the layer and the effective p-wave velocity (computed as per section B). The full model prior is developed from these ingredients (fig. 1a).

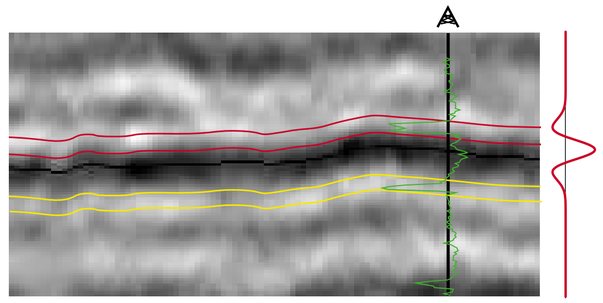


Figure 2. Suite of prior mean layer-top times, and wavelet extracted from well tie(s)

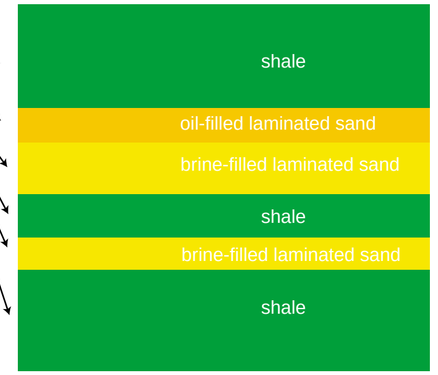
Model parameters per layer:

$\mathbf{m} = \{t_{top}, NG, \phi_R, v_{p,R}, v_{s,R}, \rho_{NR}, v_{p,NR}, v_{s,NR}, \rho_b, v_{p,b}, \rho_h, v_{p,h}, S_h\}$

R = reservoir rocks
NR = impermeable rocks
b = brine
h = hydrocarbon

Full suite of model parameters:

$\mathbf{M} = \{m_1, m_2, m_3, \ldots\}$



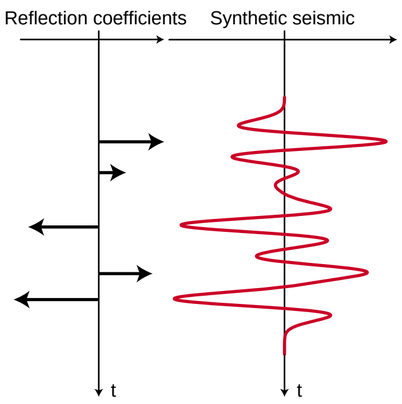


Figure 1a. Assembly of model parameters

Figure 1b. Local layer-cake model

Figure 1c. Reflectivities and synthetic seismic

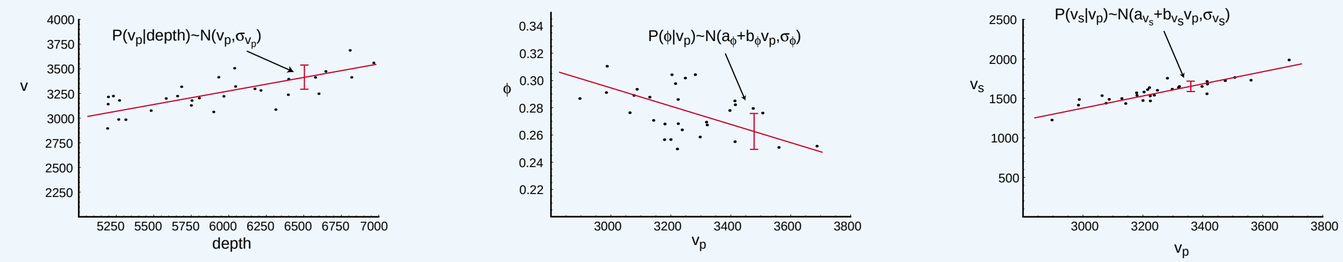


Figure 2 . Regional trends for velocities, porosity and density

References

1) Gunning, J. (2000). Constraining random field model to seismic data - getting the scale and the physics right. In *Proc. 7th European Conference on the Mathematics of Oil recovery, Baveno, Italy*
2) Aki, K., and Richards, P. (1980). *Quantitative Seismology.* Freeman.
3) Gilks, W. R. *et al*. Markov Chain Monte Carlo in Practice, Chapman and Hall

Seismic inversion for fluid detection, petrophysics and stratigraphy using Bayesian models and Markov Chain Monte Carlo methods

B) The forward seismic model

(1) **Fluid mixtures**. Permeable facies containing hydrocarbons of saturation S_h are modelled as filled with an effective fluid of bulk modulus $K_{fluid} = (S_h/K_h + (1\text{-}S_h)/K_{brine})^{-1}$. This Reuss average is appropriate for fluids well-mixed at the finest scale, which is usually appropriate for mixed phases in an exploration context.

(2) **Fluid substitutions**. The effective acoustic properties of permeable rocks filled with the effective fluid are computed from Gassman's relations. This implies low-frequency dispersion-free behaviour and isotropy of the rock matrix. Shear moduli are unchanged by fluid substitution. Densities are pore-volume weighted averages.

$$\frac{K_{sat}}{K_{grain}-K_{sat}} = \frac{K_{dry}}{K_{grain}-K_{dry}} + \frac{K_{fluid}}{\phi(K_g-K_{fluid})}$$

(3) **Laminated rock-mixtures**. Finely laminated mixtures of the fluid-substituted permeable rock with an impermeable facies (e.g. shaly sands) are modelled using Backus averages. The effective p-wave modulus ($M=\rho v_p^2$) is given by $M_{eff} = (NG/M_R + (1\text{-}NG)/M_{NR})^{-1}$. A similar expression for the s-wave modulus applies.

(4) **Reflection coefficients**. From the effective densities and velocities of the two layers forming an interface, the p-p reflection coefficient for incident waves of angle θ to the normal is given by the first-order expression (first-order contrasts in properties being consistent with the convolutional model):

$$R_{pp} = \frac{1}{2}\left(\frac{\Delta\rho}{\langle\rho\rangle} + \frac{\Delta v_p}{\langle v_p\rangle}\right) + \theta^2\left(\frac{\Delta v_p}{2\langle v_p\rangle} - \frac{2\langle v_s\rangle^2\left(\frac{\Delta\rho}{\langle\rho\rangle} + \frac{2\Delta v_s}{\langle v_s\rangle}\right)}{\langle v_p\rangle^2}\right)$$

Here the notation follows the pattern $\langle\rho\rangle = (\rho_1+\rho_2)/2$, $\Delta\rho=(\rho_2-\rho_1)$ etc. The influence of the shear velocity is clearly apparent only in the larger offset angle stacks (θ appreciable).

(5) **Convolutions**. Synthetic seismic data are formed by convolving the reflectivities with the appropriate wavelet for the stack (fig. 1c). The seismic error ("true - synthetic") is accumulated over a fixed time range which excludes the influence of reflections that may lie outside the scope of the model. This time range must be fixed for each trace to prevent biases stemming from variable degrees-of-freedom in the likelihood function.

C) Getting approximate posterior modes

For each possible fluid combination k, the posterior probability surface is a surface which may contain several local peaks, or modes (fig. 4). The most plausible of these are located by using standard (e.g. BFGS) minimisation algorithms starting from suitable initial points (e.g. the prior means). The complete set of modes is the union of all modes for each fluid combination.

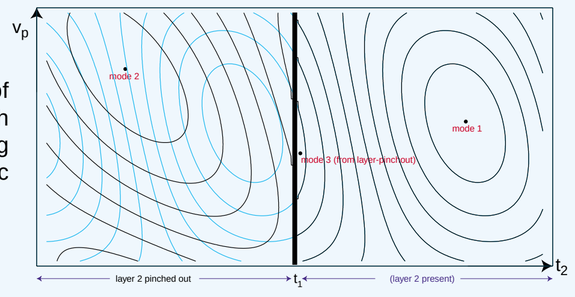


Figure 4. Typical multimodality of posterior distribution associated with pinchouts. Mode location using Newton methods and quadratic programming may be required.

Each mode has a local curvature, which is estimated by applying linearisation approximations to the likelihood functions at the mode, and updating the prior Gaussian distribution using standard expressions from inverse theory. If the prior distribution is $P(M) = N(\overline{M},C)$ and the likelihood $L=\exp(-(f(M)-S)^2/\sigma^2)$ is linearised at the point M_0 with $X = \nabla f(M_0)$, the approximate posterior is $P(M) = N(\tilde{M},\tilde{C})$ with

$$\tilde{M} = \tilde{C}(X^T(S + XM_0 - f(M_0)/\sigma^2) + C^{-1}\overline{M})$$
$$\tilde{C} = (X^T X/\sigma^2 + C^{-1})^{-1}$$

The approximate posterior covariance $\tilde{C}$ is crucial in the random walk Metropolis algorithm used to draw samples from the posterior distribution.

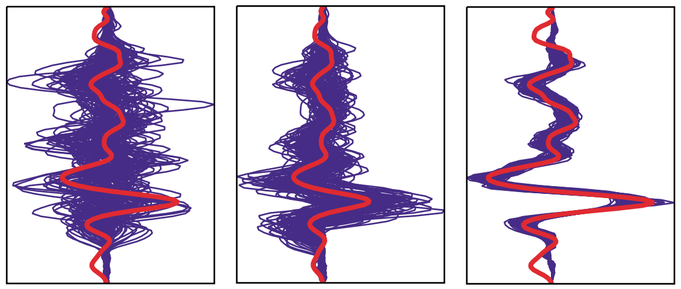


Figure 5. Synthetic seismic traces overlaid by the observed seismic trace at a particular location in the survey. **Left:** sample models drawn from the prior with brine filling a reservoir layer lower down in the model. **Middle:** the same, but with oil in the reservoir layer. **Right:** traces computed from models drawn from the posterior distribution, conditional on oil present in the reservoir layer.

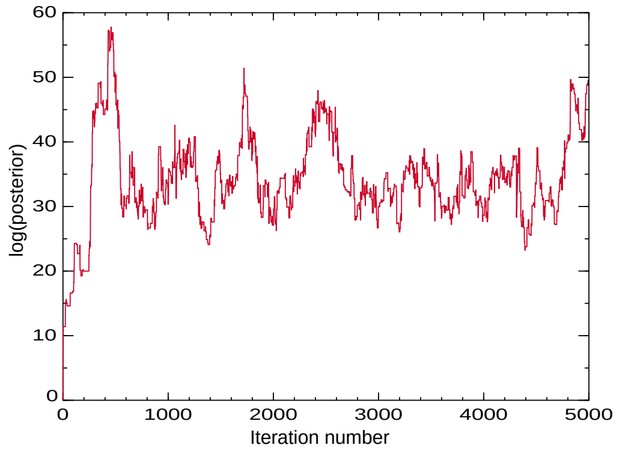


Figure 6. Typical convergence of a within-model MCMC random-walk Metropolis sampler for the entire model posterior probability. All models are initialised at the approximate maximum-likelihood point, hence the initial decay to equilibrium. Correlation lengths in these chains typically scale as 1/d (for a model with d dimensions) and optimal sampling efficiency for fully Gaussian models is known to be about 0.3/d.

D) Sampling from the posterior distribution using a MCMC scheme

Drawing samples from the posterior distribution of the various models M_k can be performed using a Markov Chain Monte Carlo (MCMC) algorithm (see ref. 3). These algorithms are the only known methods available for high dimensional problems where the posterior distribution cannot be analytically derived. For problems comprising a suite of candidate models of varying dimension, the chain requires a proposal kernel which can produce both transitions *between* models and draw new candidates *within* the same model.

For the problem of drawing proposal candidates within the same model, the posterior surface will be near-quadratic within the region of interest. Consequently, some established results for drawing from Gaussian distributions using random-walk Metropolis-Hastings algorithms can be used. New models M' are drawn by adding a jump ΔM from the current point, where the jump is drawn from a symmetric "hopping" distribution $p_H(\Delta M)$ centered at the current point. The new model, M'=M+ΔM, is accepted with probability

$$\alpha = \min(1, \frac{\Pi(M')}{\Pi(M)})$$

Optimally efficient sampling from the posterior occurs when the "hopping" distribution $p_H(\Delta M)$ has a Gaussian form with covariance $C_H = (5.8/d)\tilde{C}$, (where d is the dimension of the model), hence the importance of establishing the shape of the local covariance via the linearisation assumption described in section C. Poor estimates of the local covariance result in slow mixing of the MCMC chain and inadequate sampling of the posterior. A typical sampling sequence is shown in fig. 6.

Jumps between models can be performed in a variety of ways, but must be performed with care so as to preserve the reversibility of the chain. The simplest algorithm takes advantage of the fact that the models are *nested,* and preserves as many parameters at their current values as possible when model jumping. For example, when jumping from model M to M', with d' > d, the non-nested parameters *m* in M' are drawn from a suitable distribution *q(m)* (say the prior), which will then enter the acceptance criterion as

$$\alpha = \min(1, \frac{\Pi(M')}{\Pi(M)q(m)})$$

E) Open-source software details.

The software used to generate these inversions is part of a suite of open-source software projects being sponsored by BHP-Billiton, which includes some significant extensions to the well known *Seismic Unix* (SU) toolkit from the Colorado School of Mines. The inversion codes used here are developed in java, using high-performance numerical libraries from CERN and other sources. Source code will duly be available on the CSIRO website (contact the author for details).

CSIRO

bhpbilliton

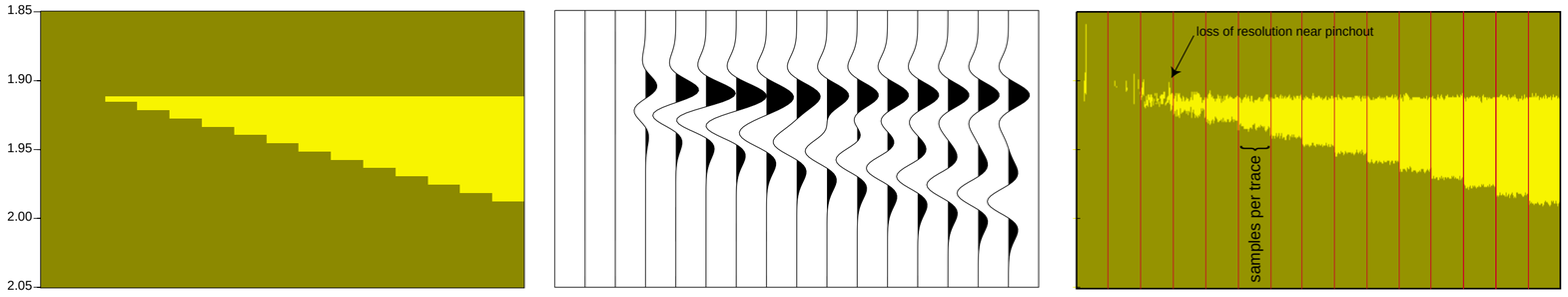


Figure 7. Inversion of a sand wedge in shales, illustrating loss of horizon-time certainty near pinchouts. **Left**: truth-case sand body in shale background. **Middle**: seismic formed from truth-case model. **Right**: sequence of samples (about 40 per trace) of layer times, grouped so as to illustrate uncertainty in time horizons. The uncertainty is constant when the horizons are resolvable, but degrades within tuning thickness. The prior horizon distributions are quite broad. The signal to noise ratio is about 4:1.