



SPE 110771

Imposing Multiple Seismic Inversion Constraints on Reservoir Simulation Models

Subhash Kalla, SPE, Louisiana State University, Christopher D. White, SPE, Louisiana State University, James Gunning, SPE, CSIRO, and Michael E. Glinsky, SPE, BHP-Billiton Petroleum, Inc.

Copyright 2007, Society of Petroleum Engineers.

This paper was prepared for presentation at the 2007 SPE Annual Technical Conference and Exhibition held in Anaheim, California, U.S.A., 11-14 November 2007.

This paper was selected for presentation by an SPE Program Committee following review of information contained in an abstract submitted by the author(s). Contents of the paper, as presented, have not been reviewed by the Society of Petroleum Engineers and are subject to correction by the author(s). The material, as presented, does not necessarily reflect any position of the Society of Petroleum Engineers, its officers, or members. Papers presented at SPE meetings are subject to publication review by Editorial Committees of the Society of Petroleum Engineers. Electronic reproduction, distribution, or storage of any part of this paper for commercial purposes without the written consent of the Society of Petroleum Engineers is prohibited. Permission to reproduce in print is restricted to an abstract of not more than 300 words; illustrations may not be copied. The abstract must contain conspicuous acknowledgment of where and by whom the paper was presented. Write Librarian, SPE, P.O. Box 833836, Richardson, TX 75083-3836, U.S.A., fax 01-972-952-9435.

Abstract

Well data reveal reservoir layering with relatively high vertical resolution but are areally sparse, whereas seismic data have low vertical resolution but are areally dense. Improved reservoir models can be constructed by integrating these data. The proposed method combines stochastic seismic inversion, finer-scale well data, and geologic continuity models to build ensembles of geomodels.

Stochastic seismic inversions operating at the mesoscale (≈ 10 m) generate rock property estimates that are consistent with regional rock physics and true-amplitude imaged seismic data. These can be used in a cascading workflow to generate ensembles of fine-scale reservoir models, wherein each realization from the Bayesian seismic inversion is treated as an exact constraint for an associated finer scale stochastic model. We use two-point statistical models for the fine-scale model, modeling thickness and porosity of multiple facies directly. The update of these fine-scale models by the seismic constraints yields highly correlated truncated Gaussian distributions. These generate potentially rich pinchout behavior and flexible spatial connectivities in the fine scale model.

The seismic constraints confine the fine-scale models to a posterior subspace corresponding to the constraint hypersurface. A Markov Chain Monte Carlo samples the posterior distribution in this subspace using projection methods that exploit the reduced dimensionality that comes with the exact constraints.

These methods are demonstrated in three-dimensional flow simulations on a cornerpoint grid, illustrating the effects of stratigraphic variability on flow behavior.

Introduction

Reservoirs are sparsely sampled by well penetrations, but seismic survey results provide controls for reservoir stratigraphy and properties such as average porosity. However, beds thinner than about $1/8$ to $1/4$ the dominant seismic wavelength cannot be resolved in these surveys.^{1,2} At a depth of 3000 m, the maximum frequency in the signal is typically about 40 Hz and for average velocities are *circa* 2,000 m/s; this translates to best resolutions of about 10 m. The resolution limits and errors inherent in seismic-derived estimates complicate use of seismic inversion data.³

Mesoscale (≈ 10 m) reservoir models obtained by seismic inversion using rock-physics concepts and effective-media ideas are a manageable basis for Bayesian seismic integration because seismic is usefully informative at this scale as explained above. An attractive route to typical geocellular scale (≈ 1 m) models is downscaling mesoscale models to meter-scale models using constraint equations embodying the effective media laws. In particular, downscaling specific realizations drawn from the posterior of a stochastic mesoscale inversion produces sum or average constraint equations for fine scale models.

We use probabilistic depth and thickness information originating from the layer-based seismic inversion code DELIVERY⁴ as input to a downscaling algorithm operating on a cornerpoint grid. Seismic constraints and priors are modeled on the quasivertical block edges, analogous to seismic traces. Simulation at the edges preserves geometric detail required for cornerpoint reservoir models used in many commercial reservoir simulators (*e.g.*, ECLIPSE⁵). Block-center properties such as porosity are obtained by averaging the edge properties.

Realization ensembles from the inversion code DELIVERY⁴ carry rich information about interproperty and vertical interzone correlations induced by the seismic information (**Fig. 1**). These ensembles are generated assuming no trace-to-trace correlation, and the traces generally do not coincide with cornerpoint edges in the flow grid. The DELIVERYMESSAGE^{6,7} code augments the interproperty and interzone correlations with the mesoscale lateral correlation structures required for geological continuity, and constructs models or model-samples at the quasivertical cornerpoint edges of the flow grid. Each realiza-

tion from DELIVERYMESSAGER thus captures vertical, horizontal, and interproperty correlations at the mesoscale (**Fig. 2**).

These realizations are natural inputs to the downscaling problem we describe. They contain the requisite coupling between geometry and rock properties that seismic inversion induces, plus the necessary spatial correlations required for geological smoothness. These mesoscale models provide explicit sum and average constraints on the corresponding subseismic layers. Such constraints are nontrivial to respect using conventional geostatistical algorithms for fine-scale heterogeneity.

Specifically, we consider a fine-scale model of $k \in \{1 \dots K\}$ layers, each layer k with thickness h_k and porosity ϕ_k . We use t as an untruncated surrogate for layer thickness, so $h_k = \max(0, t_k)$: the proxy t may take on negative values, whereas h is truncated at zero.

If one wishes to ensure consistency of both thickness and average porosity in a downscaling problem, the following constraints must be imposed at each column of gridblock corners:

$$\sum_{k=1}^K F_k \max(0, t_k) = H_s \quad (1)$$

$$\sum_{k=1}^K F_k \max(0, t_k \phi_k) = \Phi H_s \quad (2)$$

$$\sum_{k=1}^K (1 - F_k) \max(0, t_k) = H_{sh} \quad (3)$$

The right-hand sides of these equations are obtained from a mesoscale joint realization of net thickness H_s , non-net thickness H_{sh} , and the mesoscale average porosity Φ . Mesoscale porosity is the net thickness weighted average. Here $F_k \in \{0, 1\}$ is a predetermined facies indicator for layer k , where there are $K_s \leq K$ “sand” layers with $F = 1$, and $K_{sh} = K - K_s$ “shale” layers with $F = 0$.

The fine-scale model prior to these constraints is a joint multi-Gaussian distribution of t, ϕ (over all layers and traces) with means and covariances constructed in the usual ways, such as from log data or variogram analysis of outcrop analogs.⁸ The principal challenge of this approach is that the downscaling constraints potentially force the posterior of the fine-scale model to have truncated regions (*e.g.*, the mode of layer thickness or porosity may be zero). The nonlinearity embedded in the $\max(0, t_k)$ terms makes linear estimation inadequate in many cases.

Use of Terms. The following conventions are used (**Fig. 3**).

Zones (≈ 10 m) are mesoscale reservoir sequences which seismic inversion data resolves. Constraints are applied to zones.

Layers (≈ 1 m) are fine-scale features not generally resolved by inversion data, but layers can be identified in wells. In this paper, well data is used only at the layer scale — log and core data must be upscaled to layers before this data integration method can be applied.

Traces are near-vertical lines at which constraints are applied and layer thicknesses will be estimated. They may correspond to true seismic traces, but more generally they will be data that

have been simulated from seismic inversions using DELIVERYMESSAGER.⁷ Traces define the xy edges of the cornerpoint gridblocks⁹ (*viz.*, COORD records).⁵ Conditioning data at wells are a type of trace. A trace record comprises vectors of thickness and average properties for all flow grid layers.

Paths are sequences in which traces (or blocks of traces) are visited. We use quasirandom multigrid paths.¹⁰

Problem Formulation

We combine diverse data elements in prior and likelihood/constraint expressions to obtain a posterior probability. The overall posterior distribution is approximated by a multigrid sequential simulation passing over all traces of the grid (Sequential Methods and Marginalization, later). Properties at each trace are simulated by sampling from a Bayesian posterior distribution whose prior is assembled from hard data and previously visited traces, and whose likelihood is comprised by the local coarse scale constraints (sums and averages from the massaged seismic inversions).

The prior distribution for each layer at a given trace is obtained by solving an ordinary kriging system¹¹ with well observations and previously simulated values as the informing data. At a trace, kriging for layer k uses only layer k data: vertical correlation is assumed to be zero.

This is the usual sequential simulation approximation for multi-Gaussian models. An efficient approximation to the local-prior covariance matrix is crucial to the success of this Bayesian approach. This matrix must be factored at each step of the sequential simulation algorithm, and some approximations make the sampling process more efficient (Assumptions on Prior Covariance, later).

Algorithm Outline. The algorithm is similar to many sequential algorithms, except that Markov Chain Monte Carlo (MCMC) is used to generate conditional samples at each spatial location (**Fig. 4**). A multigrid random path specifies simulation sequence. At each point on the path, the local conditional prior is constructed using kriging, and sampling of the local posterior is performed using MCMC in a reduced-dimensionality subspace with projections back to the constraint surface in the original space (Sampling with Exact Constraints, later). A sample from the converged MCMC chain is chosen, and the algorithm proceeds to the next trace in the multigrid path.

Prior and Posterior Formulation. The posterior probability π for the thicknesses and porosities is, from Bayes’ rule,

$$\pi(\mathbf{t}_s, \phi_s, \mathbf{t}_{sh} | H_s, \Phi, H_{sh}, \mathbf{d}_{\ell k}) = \frac{p(H_s, \Phi, H_{sh} | \mathbf{t}_s, \phi_s, \mathbf{t}_{sh}, \mathbf{d}_{\ell k}) p(\mathbf{t}_s, \phi_s, \mathbf{t}_{sh} | \mathbf{d}_{\ell k})}{p(H_s, \Phi, H_{sh} | \mathbf{d}_{\ell k})} \quad (4)$$

where $\mathbf{d}_{\ell k}$ is a vector of the all neighboring conditioning or previously simulated traces in layer k in the neighborhood of trace ℓ . The size of the vectors $\mathbf{t}_s, \phi_s,$ and \mathbf{t}_{sh} are $K_s, K_s,$ and K_{sh} respectively. The second term on the numerator is the prior for the local fine-scale variables, constructed using typical kriging apparatus. The first term in numerator (the likelihood) is

an exact constraint in the proposed cascading workflow, so that the posterior probability of the model is either (1) identically zero if the constraint is not satisfied, or (2) proportional to the prior if the constraint is satisfied. Assuming multi-Gaussian and independent distributions (Assumptions on Prior Covariance, later), the variables $\mathbf{t}_s, \phi_s, \mathbf{t}_{sh}$ are normally distributed with standard deviations $\sigma_{t_s}, \sigma_{\phi_s}$ and $\sigma_{t_{sh}}$, respectively. The multivariate prior distribution of \mathbf{t}_s for example is

$$p(\mathbf{t}_s | \mathbf{d}_{\ell k}) = \frac{1}{(2\pi)^{\frac{K_s}{2}} |\mathbf{C}_{ps}|^{\frac{1}{2}}} \exp \left[-\frac{1}{2} (\mathbf{t}_s - \bar{\mathbf{t}}_s)^T \mathbf{C}_{ps}^{-1} (\mathbf{t}_s - \bar{\mathbf{t}}_s) \right] \quad (5)$$

where \mathbf{C}_{ps} is the prior or kriging covariance matrix for thicknesses of sand, which has rank K_s . We solve the kriging system using a Cholesky factorization, obtaining prior means and variances for all layers on a given trace.

The constraints [Eqns. (1-3)] will yield posterior distributions that may be far from normal and (unlike the priors) are correlated (Simulations of Two-Layer Systems, later).

Assumptions on Prior Covariance. A number of assumptions make calculation of the local conditional prior distributions more efficient.

1. Traces can be approximated as vertical when computing the kriging covariance matrix: the areal separation between the current trace and each of its neighbors is assumed constant for all layers for each trace-neighbor pair. This assumption is particularly useful when kriging thicknesses in a cornerpoint grid: the separation distance actually depends on the estimated variable (thickness) for these cases. The assumption is acceptable if the traces (or COORDs) are nearly parallel, or if their lateral displacement within the reservoir interval is small compared to cornerpoint/trace spacing.
2. Vertical correlation of layer properties is neglected, so that all kriged estimates are two-dimensional. This leads to a simpler, smaller kriging system. This assumption also makes the prior covariance matrices diagonal for the exact constraint problem. This assumption seems reasonable for many episodic depositional systems, such as turbidites. Trends could be introduced in the priors if required.
3. *Optionally*, the covariance models are the same for all the shales and sands. This leads to a simple Toeplitz structure that is computationally efficient. If this assumption is not made, the prior covariance matrix will be anisotropic and must be transformed in the MCMC calculations. This assumption is important for inexact constraints,¹⁰ but is less essential for exact constraints.

Recapitulating, the benefits of these assumptions are: (1) the kriging matrices are identical for all layers, and therefore only three kriging systems (for t_s, t_{sh} , and ϕ_s) are solved at each trace, (2) the priors for each layer can be computed separately, and (3) the prior variances in each constraints are constant, but the prior means vary layer-by-layer. The problem is still tractable without these assumptions. However, the additional

kriging work will make the approach more expensive, especially if assumptions 1 and 2 are not valid.

Sampling with Exact Constraints

Exact constraints lower the dimensionality of the sampling problem. The maximum a posteriori (MAP) estimate is commonly obtained using Lagrange multipliers.^{12,13} For sampling, however, we reformulate the problem in a lower-dimensional space and perform appropriate projections to ensure the constraint is honored.

Orthogonal Subspace Projection. At each trace, three constraints (Eqns. 1-3) are applied in this downscaling method. The system has $2K_s + K_{sh}$ random variables (sand thicknesses, sand porosities, and shale thicknesses), and thus $d = 2K_s + K_{sh} - 3$ degrees of freedom. The constrained posterior is sampled by a MCMC random walk in a lower dimensional subspace, with projection to the constraint surface to reconstruct the full set of variables. Because of the assumptions (Assumptions on Prior Covariance, earlier), the priors of \mathbf{t}_s, ϕ_s , and \mathbf{t}_{sh} are uncorrelated. The problem can thus be solved sequentially for the three blocks $\mathbf{t}_s, \mathbf{t}_{sh}, \phi_s$.

We reparametrize, projecting all points in \mathbf{T} , the original space of \mathbf{t}_s , onto the plane normal to $\mathbf{t}_s = (1, 1, \dots, 1)$. The basis for this new space, \mathbf{R} , is obtained by singular value decomposition (SVD).¹³ The space \mathbf{R} is a $K_s - 1$ dimensional basis orthogonal to $\mathbf{u} = (1, 1, \dots, 1)/\sqrt{K_s}$ (**Fig. 5**). In this $K_s - 1$ subspace a point is denoted by \mathbf{r} , and a point in K_s dimensions is $\boldsymbol{\tau} = (\delta, \mathbf{r})$, where the first element is parallel to \mathbf{u} . Although potentially confusing, the use of these three entities (\mathbf{r} in *rotated* $K_s - 1$ \mathbf{R} space, $\boldsymbol{\tau}$ in *rotated* K_s space, and \mathbf{t}_s in the original K_s \mathbf{T} space) is essential to explaining and implementing the algorithm. The transformation matrix \mathbf{U} rotates the K_s -long augmented vector $\boldsymbol{\tau} = (\delta, \mathbf{r})$ back to original coordinates in \mathbf{T} , *viz.*, $\mathbf{t}_s = \mathbf{U}\boldsymbol{\tau}$.

The random walk is performed in \mathbf{R} space. This rotated ($K_s - 1$)-dimensional subspace is useful because each basis direction in the original space \mathbf{T} is aligned with some constraint facets (**Fig. 5**), whereas \mathbf{u} has a component normal to all facets. The alignment of facets with the bases of \mathbf{T} precludes a unique mapping from the constraint surface to a coordinate-aligned facet of the constraint hypersurface in the original space. Indeed, some points (*e.g.*, the point P in **Fig. 5**) cannot be projected back to any facet of the constraint hypersurface along any single original basis vector of \mathbf{T} . On the other hand, any point in \mathbf{R} can be uniquely projected back to the constraint surface by moving a distance δ along \mathbf{u} .

For any point in \mathbf{R} space, the point \mathbf{t}_s is obtained by first transforming $\mathbf{t}_s = \mathbf{U}\boldsymbol{\tau}$ (δ is unknown) and then solving for δ :

$$\sum_{j=1}^{K_s} \max(0, t_{s(j)}(\delta, \mathbf{r})) = H_s \quad (6)$$

Eqn. (6) is piecewise linear and monotonic in \mathbf{t}_s , and therefore has a unique solution. The solution can be found in $O(K_s \log K_s)$ time by sorting \mathbf{t}_s (without unknown δ) and searching between its elements for the δ that satisfies Eqn. (6).

Once δ is estimated, $\boldsymbol{\tau} = (\delta, \mathbf{r})$ and $\mathbf{t}_s = \mathbf{U}\boldsymbol{\tau}$. The K_{sh} shale thicknesses are obtained similarly. For porosity the constraint is

$$\sum_{k=1}^{K_s} \max(0, t_{s(k)}) \max(0, \phi_{s(k)}) = \Phi H_s \quad (7)$$

which can be solved by sorting the $\phi_{s(k)}$ for which $t_{s(k)}$ is positive, and finding ϕ_s . Eqn. (7) uses left truncation to ensure all porosities are positive.

Effects of Projection on MCMC Sampling. The random walk discussed above is a Markov chain sampling on the constraint surface. Because the constrained problem is nonlinear (although piecewise linear), a Jacobian is used to compute the correct acceptance probability α for a jump from $\boldsymbol{\tau}$ to $\boldsymbol{\tau}'$ in the Metropolis-Hastings formula

$$\alpha = \min \left(1, \frac{\pi(\boldsymbol{\tau}') P_J(\mathbf{t}|\mathbf{t}')}{\pi(\boldsymbol{\tau}) P_J(\mathbf{t}'|\mathbf{t})} \right) \quad (8)$$

These Jacobian terms P_J correct for changes in the angle of projection (Fig. 5) if the walk moves between facets of the constraint hypersurface. To compute the Jacobian, we first form the normal to the active constraint facet. The facet used to form \mathbf{n} differs for the two P_J in Eqn. (8): for $P_J(\mathbf{t}|\mathbf{t}')$ it is the normal to the facet that \mathbf{t} lies on, and for $P_J(\mathbf{t}'|\mathbf{t})$ the facet containing \mathbf{t}' is used. The Jacobian in the numerator of Eqn. (8) is

$$P_J(\mathbf{t}|\mathbf{t}') = \frac{1}{\mathbf{u} \cdot \mathbf{n}} \quad (9)$$

where \mathbf{n} is based on the location of \mathbf{t} , as discussed above. The angle θ is measured between \mathbf{u} and \mathbf{n} (Fig 5). The ratio of the Jacobians is 1 if \mathbf{t} and \mathbf{t}' lie on the same facet, causing the Metropolis-Hastings α to correctly revert to its usual form. Including the Jacobians P_J preserves reversibility of the MCMC steps if \mathbf{t} and \mathbf{t}' are not on the same facet, a necessary property for correct MCMC calculations.¹⁴ Because of their prior lack of correlation, the Jacobians for the three variables (sand and shale thickness, plus porosity) can be combined, *e.g.*, $P_J(\mathbf{t}_s|\mathbf{t}'_s)P_J(\mathbf{t}_{sh}|\mathbf{t}'_{sh})P_J(\phi_s|\phi'_s)$.

Generating and Accepting Proposals

In this section, we focus on simulation at a given trace ℓ . The overall simulation proceeds by visiting all ℓ that are not in the conditioning data set by a specific, random, multigrad path. At a particular trace, while generating samples, transitions between points are proposed in the rotated space \mathbf{R} . If the prior covariance \mathbf{C}_p (in original t - or ϕ -coordinates) is not isotropic (Assumptions on Prior Covariance, earlier), then the covariance corresponding to the rotated subspace must be transformed when computing proposals, *via* $\tilde{\mathbf{C}} = \mathbf{U}\mathbf{C}_p\mathbf{U}^T$. The Schur complement yields the covariance of the \mathbf{R} -space vector \mathbf{r} (using $\delta = H/\sqrt{K}$), $\tilde{\mathbf{C}}_r = \tilde{\mathbf{C}}_{22} - \tilde{\mathbf{C}}_{21}\tilde{\mathbf{C}}_{11}^{-1}\tilde{\mathbf{C}}_{12}$, where $\tilde{\mathbf{C}}$ is split up into the blocks (1,2) corresponding to the (δ, \mathbf{r}) pieces. Thus proposals should be sampled from the distribution $\Delta\mathbf{r} \sim N(0, s^2\tilde{\mathbf{C}}_r)$. The Cholesky factor $\tilde{\mathbf{L}}_r$ of the covariance

matrix $\tilde{\mathbf{C}}_r$ is multiplied into a scaled d -long vector of standard normal deviates $\mathbf{a} \sim [N(0, 1)]$

$$\Delta\mathbf{r} = s\tilde{\mathbf{L}}_r\mathbf{a}. \quad (10)$$

Here s is a scalar chosen for sampling efficiency,¹⁴ typically $s = 5.76/d$ for large d . The constraint is imposed by projecting \mathbf{r} and transforming to new coordinates (Orthogonal Subspace Projection, earlier). The proposal prior probability is computed at the new point \mathbf{t}'_s , using Eqn. (5). The Metropolis-Hastings transition probability is then¹⁴ computed using Equation (8). The proposed transition is accepted with probability α , and the algorithm iterates until equilibrium of the MCMC sample. One of the realizations from the equilibrium distribution is randomly chosen and this process is continued until all the traces are visited.

Prior Distributions and Constraints in 2D and 3D

Important features of higher-dimensional cases are easier to visualize with two layers. Simple parameter choices are used to clarify the explanation.

In two dimensions the constraint surface comprises two orthogonal line segments joined by a 45 degree ‘‘chamfer’’ in the 1st quadrant (Fig. 5). In three dimensions, the 45 degree line in Fig. 5 corresponds to a triangle with vertices on each t -axis at H (Fig. 6); increasing H shifts the high-likelihood region away from the origin, but with no change in slope. There are six other facets comprising the constraint in three dimensions; for K variables there are $2^K - 1$ facets. On one facet, all K of the t are positive. Between 1 and $K - 1$ layers have $t \leq 0$ on the remaining $2^K - 2$ facets. Depending on the prior, few or many of these facets, corresponding to layer configurations, may occur in the posterior.

If no layer kriging data were used and the seismic data were considered exact, any \mathbf{t} on any facet of the constraint hypersurface could be used. Conversely, in a sequential simulation not conditioned to seismic, the layer thicknesses would simply be drawn from the prior.

Simulations of Two-Layer Systems

We consider a case with prior mean thicknesses of 3 m and 1 m, 1 m standard deviations, and constraints of total thickness $H_s = 4$ m and $\Phi H = 1.0$; the layer porosity prior means are 0.2 and 0.3, with a standard deviation of 0.05 (Table 1). We consider only one facies, $K = K_s = 2$. In this case the sum of the prior thickness means is equal to the thickness constraint, but the individual layer thicknesses are poorly resolved (Fig. 7). Because the means are consistent with the constraint, the posterior means of \mathbf{t} are near their prior means, and the two layer distributions look rather like mirror images, with their high (layer 1, $h_1 \approx 4$) and low (layer 2, $h_2 \approx 0$) complementing one another. The thickness constraint is fulfilled for every realization (Fig. 8). The univariate histograms of t are non-Gaussian because of the constraints. The thickness constraint induces a nearly perfect negative correlation in the layer thicknesses (Table 1; $\rho_{12} = C_{12}/\sqrt{C_{11}C_{22}} = -0.98$); the departure from -1 occurs because a small fraction of the realizations are not on the

45 degree portion of the constraint surface. There will be a substantial negative correlation in thickness if the prior means are larger than the prior standard deviations, such that most points lie on the hypersurface facet in the first quadrant.

Histograms for the porosity (Fig. 9) show that the posterior means are greater than the prior means. This is because the thickness averaged mean porosity for the priors is less than the thickness averaged porosity constraint. The distribution of ϕ_1 and ϕ_2 (Fig. 10) also reveals the constraint information: the porosity of both layers cannot be higher than or less than 0.25 at the same time, which gives a striking inclined hourglass shape to the posterior (Fig. 10). The distribution of $\phi_1 t_1$ and $\phi_2 t_2$ satisfies the ΦH constraint (Fig. 11).

Cross distributions of t_1 and ϕ_1 and t_2 and ϕ_2 (Figs. 12 and 13) yield rich information about how the constraints interact with the prior. All the samples in these two pictures follow the condition $\phi_i t_i < 1$ because $\Phi H_s = 0.25 \times 4 = 1$. As the thickness of a layer increases, the porosity of the layer converges to $\phi = \frac{\Phi H_s}{H_s}$, which is 0.25 for this case.

Synthetic Three-Dimensional Cases

Synthetic 3D cases test and illustrate the MCMC simulation method. The prior distribution is varied by changing the range (b) and sill (σ^2) of the Gaussian semivariograms,

$$\gamma(\Delta x) = \sigma^2 \left[1 - \exp\left(-\frac{(3\Delta x)^2}{b^2}\right) \right] \quad (11)$$

where Δx is the lag.

Trends in H -related parameters are varied to illustrate behavior. The simulations are conditioned on four traces, all of which have average thicknesses imposed (H_s/K_s for sands and H_{sh}/K_{sh} for shales; Table 2). Porosity at these traces increases downward linearly from 0.2 to 0.3 in the sands. For all cases, the grid size is $I \times J \times K$ is $100 \times 100 \times 10$ and the $x - y$ extent is 1000×1000 m; $L = 1000$ m. The total sand thickness, $H_s = 14$ m, shale thickness, $H_{sh} = 6$ m, and porosity thickness, $\Phi H_s = 3.5$ m, are uniform. The porosity prior is held constant (with range of 500 m and standard deviation of 0.025). We specify alternating shales and sands, $K_s = K_{sh} = 5$. The system comprises $100 \times 100 \times 5 = 5 \times 10^4$ values for each of the three variables, t_s , t_{sh} , and ϕ_s (150,000 total).

Illustrative Cases. The correlation range and sill are varied for the sandstone and shale layer thicknesses (Table 2). Increasing the range gives the geobodies larger average dimensions, whereas increasing the variance (at constant mean thickness) makes them truncate (*i.e.*, have zero thickness) more frequently. This difference is apparent when comparing realizations for the base case prior and for priors with a greater sand sill (cases 0000 and 00+0, Table 3; Fig. 14a,b). The high-sill case has greater variability in sand thickness; there will be more truncated beds, and to fulfill the thickness constraints there must also be more thick beds. The high-sill case bedding geometry is thus more “come-and-go” and less “layer-cake.”

The realizations displayed in Fig. 14 are chosen to illustrate differences; they are not chosen to be “typical” realizations for their respective cases.

This cascading data integration method has several features that distinguish it from conventional geostatistical methods. There is little surface rugosity; the layer thicknesses vary smoothly and plausibly. Near-zero thicknesses do not appear in isolated areas because of the truncation rules and the smooth Gaussian variogram. All three constraints are fulfilled at every trace.

Performance. A large step size in the MCMC proposals helps explore the posterior rapidly. On the other hand, large steps are more likely to be rejected, “wasting” computations on a sample that is not accepted. A good choice appears to be to scale the covariance to yield acceptance rates of about 30 percent.¹⁴

The computational cost of a single simulation is examined component-by-component (Table 4). Several features are striking. First, 98 percent of the work is done in the deepest part of the sampling loop, which requires random number draws, sorting vectors, and multiplication of random normal vectors by the Cholesky factor. Because the kriging system is solved only three times per trace — and is two-dimensional, with an efficient k -d neighbor search — the associated work is small, about 0.005 percent. The overall cost of about 8 minutes for 150,000 variables on a laptop computer does not seem prohibitive. Because the algorithm is trace-wise sequential, the overall cost scales linearly in the number of traces; this extrapolates to 10^6 variables in about 55 minutes.

Three-dimensional Cornerpoint Flow Simulations

Flow Model Descriptions. The effects of varying stratigraphy through the proposed downscaling method are investigated using flow modeling. All models use the same seismic constraint data; here, constant *mesoscale* values of thickness and porosity over the flow domain (Synthetic 3D Cases, earlier). The permeability is assigned using an exponential transform of porosity,

$$k = 20 \exp(10\phi) \quad (12)$$

The differences in flow behavior are due only to differences in continuity (as expressed in the prior) and stochastic fluctuations. A range of models with distinct ranges and variances for both sand and shale layers are created using the proposed methods for downscaling.

The four factors are considered using a 2-level full factorial experimental design, with 6 replicates of the base case (0000, Table 3) to investigate stochastic fluctuation and 6 replicates high sand sill case to reconcile fluctuations and effects (00+0, Table 3; total of 28 runs). The models are single-phase tracer simulations for simplicity and efficiency. The geometry is one-quarter of a five-spot (a symmetry element if rates were balanced, properties were homogeneous, and the pattern were infinite). The only response analyzed is the recovery when the produced normalized tracer concentration exceeds 10 percent, referred to as N_{pD} .

Analysis of Flow Model Results. Although honoring the same mesoscale constraints and conditioning data, the flow models have notably different flow responses, with recovery at

10 percent tracer fraction ranging from $N_{pD} = 0.39$ to 0.65. The priors allow quite diverse models, so this variability is unsurprising (Table 3). The flow behavior for distinct prior models appears quite different, as expected (Fig. 14c,d). The greater variability in thickness and more frequent termination in the high sand sill case increases tortuosity and cause lower recovery.

The variance over the set of 16 factorial points is 3.2 times greater than the variance of the replicates at the centerpoint. While seeming to suggest that variation in the prior may dominate fluctuations, $F = 3.2$ with 15 and 5 degrees of freedom is not significant for a two-sided test. That is, stochastic fluctuations appear to be roughly comparable to the effects of varying the prior, even with the large variability in the priors used in the experimental design.

We examine the effects of prior variability *versus* stochastic fluctuation more closely using a pair of points, sets of six replicates for the base case (0000; design center) and for a higher sand sill value (00+0; design face center). For these two sets, a Welch two-sample t -test¹⁵ indicates that the mean responses for the different priors are not the same ($t = 1.875$; critical t for 95 percent confidence is 0.12). That is, the means are different, and the specification of the prior has a significant effect compared with stochastic fluctuations.

In summary, (1) prior specification has a statistically significant effect on response, and (2) prior variability and stochastic fluctuations may both make substantial contributions to overall response variability.

Discussion

Related Methods when Constraint is not Exact. It is possible to partially account for the uncertainty in the seismic inversion without the cascading sampling workflow advocated here (Abstract and Introduction, earlier). At particular traces, full uncertainties of all inverted quantities are available from DELIVERY,⁴ and these can be used in the form of soft constraints on the downscaling problem. In this framework, the likelihood does not induce a dimensionality reduction, but the posterior surface still exhibits a similar “kinked” behavior around the constraint lines. Efficient sampling from these distributions is challenging, but methods such as use of auxiliary variables are promising.¹⁰ A remaining challenge in this approach to incorporate the transverse spatial correlation in the inversion uncertainty; this difficulty motivated the cascading approach advocated in this paper.

Sequential Methods and Marginalization. A difficult question in nonlinear downscaling problems is whether the system posterior distribution can be adequately factored into the product of conditional distributions, as implied by the sequential pass over the columns of gridblocks.¹⁶ Generating conditional distributions at the current trace in sequential simulation requires integrating information over “unvisited” sites (computing analytical marginal distributions) and conditioning only on “visited” sites. The integrability requirement is usually met only by exponential family distribution functions, and the posteriors in our problem do not always follow exponential forms

(*e.g.*, Fig. 9). Nonetheless, the approximations we make can be improved by blockwise sequential schemes, though a block approach increases the dimensionality of the MCMC sampling subproblem and exacerbates the configurational complexity of handling more pinchout transitions. If we compare methods with rigorous marginals to local approximations for several test problems, the local method errors increase as transverse correlation lengths increase.¹⁷ The intuitively clear reason for this is that the information from “unvisited” sites should increasingly shape the local marginal as the correlations strengthen. Although global schemes handle this correctly, these approaches increase sampling dimensionality and the configurational complexity of handling more pinchout transitions simultaneously.

Interestingly, within the modeling framework of treating imaged seismic data as dense independent information, rigorous global methods that correctly develop the local marginal distribution may overstate the statistical certainty at any particular location, especially when the transverse correlation lengths are long. Loosely speaking, this is because all n seismic traces within a correlation length contribute to the reduction in uncertainty at the current point, and the resulting $1/\sqrt{n}$ reduction which occurs if all the traces are treated as independent measurements is over-optimistic. In practice many systematic effects in acquisition and processing make the imaged seismic data less informative than this.

Conclusions

Stochastic seismic inversion models can be integrated with a truncated Gaussian geostatistical model for fine-scale layer thicknesses and porosity using a Markov chain Monte Carlo algorithm. Mesoscale seismic inversion realizations (which act as exact constraints) of net-sand, gross sand, and porosity are “stochastically downscaled,” using a Metropolis Hastings sampler exploiting dimensionality reduction and projection to the constraint surface. Reproduction of limiting results and model truncation behavior is demonstrated in two-dimensional test cases. Synthetic three-dimensional cases demonstrate that the proposed data integration procedure is acceptably efficient and is capable of producing models consistent with seismic data and exhibiting diverse flow behavior. New methods are required to create, screen, and analyze model in this powerful but complex cascading workflow.

Acknowledgments

BHP-Billiton funded this research by research agreements and an unrestricted gift to Louisiana State University, and research agreements with CSIRO. The authors are grateful for this support, and for permission to publish this paper.

Nomenclature

Roman Symbols

a	random number
b	variogram range
C_p	prior covariance matrix based on kriging, m^2
\tilde{C}	covariance matrix in new coordinates, m^2
F	facies index
h	nonnegative layer thickness, m
H	total thickness at trace, m
K, k	total number of layers and layer indices respectively
ℓ	indices over traces
L	extent of simulation model, m
\mathbf{n}	normal vector to a surface
p	probability density
P	probability
P_J	Jacobian term in Metropolis-Hastings transition
\mathbf{r}	vector for the subspace \mathbf{R}
\mathbf{R}	subspace in transformed coordinates
s	scaling factor
t	Gaussian proxy for h , may be negative, m
\mathbf{T}	space in original coordinates
\mathbf{u}	vector orthogonal to the first-quadrant hyper-surface facet
\mathbf{U}	rotation matrix
N_{pD}	recovery factor at breakthrough, pore volumes
x, y, z	coordinates, m

Greek Symbols

α	Metropolis-Hastings transition probability
δ	magnitude of the direction vector \mathbf{u}
κ	number of layers at a trace with $t_k > 0$
λ	dimensionless range, b/L
ν	coefficient of variation
ϕ	layer porosity
Φ	trace average porosity
π	posterior
σ^2	variance
$\boldsymbol{\tau}$	vector of properties in the transformed axes

Diacritical Marks

$\bar{\quad}$	mean
\cdot	proposed point, may become new point
$\tilde{\quad}$	rotated

References

- Dobrin, M.B. and Savit, C.H.: *Introduction to Geophysical Prospecting*, McGraw-Hill Companies, New York, fourth edition (1988).
- Widess, M.B.: "How thin is a thin bed?" *Geophysics* (December 1973) **38**, 1176.
- Deutsch, C.V., Srinivasan, S. and Mo, Y.: "Geostatistical reservoir modeling accounting for precision and scale of seismic data," *SPE Annual Technical Conference and Exhibition*, Dallas, Texas (1996) .
- Gunning, J.G. and Glinsky, M.E.: "Delivery: an open-sourced Bayesian seismic inversion program," *Computers and Geosciences* (2004) **30**.
- Schlumberger Technology Co., Oxfordshire: *Eclipse 100 97A Reference Manual* (2004).
- Glinsky, M. *et al.*: "Integration of uncertain subsurface information into multiple reservoir simulation models," *The Leading Edge* (2005) **24**, 990.
- Gunning, J., Glinsky, M.E. and White, C.: "DeliveryMessenger: a tool for propagating seismic inversion information into reservoir models," *Computers and Geosciences* (May 2007) **33**, 630.
- Willis, B.J. and White, C.D.: "Quantitative outcrop data for flow simulation," *Journal of Sedimentary Research* (2000) **70**, 788.
- Ponting, D.K.: "Corner Point Grid Geometry in Reservoir Simulation," *Proc., First European Conference Math. Oil Recovery*, Cambridge (1989) .
- Kalla, S., White, C.D., Gunning, J.S. and Glinsky, M.E.: "Consistent Downscaling of Seismic Inversions to Cornerpoint Flow Models," *SPE Annual Technical Conference and Exhibition*, San Antonio, Texas (September 24-27, 2006) SPE paper no. 103268.
- Goovaerts, P.: *Geostatistics for Natural Resources Evaluation*, Applied Geostatistics Series, Oxford University Press, Oxford (1997).
- Michalak, A.M. and Kitanidis, P.: "A method for enforcing parameter nonnegativity in Bayesian inverse problems with an application to contaminant source identification," *Water Resources Research* (2003) **39**.
- Golub, G.H. and van Loen, C.F.: *Matrix Computations*, The John Hopkins University Press, Baltimore, third edition (1996).
- Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B.: *Bayesian Data Analysis*, CRC Press, Boca Raton, second edition (2003).
- R Development Core Team: *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria (2007), ISBN 3-900051-07-0.
- Liu, J.S.: *Monte Carlo Strategies in Scientific Computing*, Springer, New York, first edition (2001).
- Kalla, S., White, C.D. and Gunning, J.: "Downscaling Seismic Data to the Meter Scale: Sampling and Marginalization," *Petroleum Geostatistics*, EAGE, Cascais, Portugal (10-14 September 2007) .

Table 1: Parameters and results for 2-layer simulation

Layer	Prior				Posterior					
	\bar{t}	σ_t	$\bar{\phi}$	σ_ϕ	\bar{h}	C_t		$\bar{\phi}$	C_ϕ	
1	3	1.	0.20	0.05	2.89	(0.44 -0.48)		0.23	(0.0008 0.0)	
2	1	1.	0.30	0.05	1.11	(-0.48 0.55)		0.31	(0.0 0.0021)	

Table 2: Design of 3D flow simulations

Factor	Sands			Shales		
	Low (-)	Base (0)	High (+)	Low (-)	Base (0)	High (+)
Range, $\lambda = a/L$	0.25	0.5	1.0	0.25	0.5	1.0
Sill, $\nu = \sigma_t/\mu_t$	0.50	1.0	2.0	0.5	1.0	2.0

Table 3: Results of 3D flow simulations

Run	ν_{sh}	λ_{sh}	ν_s	λ_s	N_{pD}
0	0	0	0	0	0.564 ^a
17	0	0	+	0	0.505 ^b
1	-	-	-	-	0.649
2	-	-	-	+	0.658
3	-	-	+	-	0.399
4	-	-	+	+	0.578
5	-	+	-	-	0.624
6	-	+	-	+	0.553
7	-	+	+	-	0.579
8	-	+	+	+	0.516
9	+	-	-	-	0.603
10	+	-	-	+	0.599
11	+	-	+	-	0.481
12	+	-	+	+	0.645
13	+	+	-	-	0.649
14	+	+	-	+	0.608
15	+	+	+	-	0.392
16	+	+	+	+	0.595

^a value is mean of 6 replicates, $\sigma = 0.046$

^b value is mean of 6 replicates, $\sigma = 0.061$

Table 4: Performance summary for the 3D example (one complete simulation)^a

Process	Work in seconds ^b
Kriging work	2.32
Total overhead all traces	4.45
Samples, 5000 per trace, all traces	485.46
Cost of example simulation, excluding io	492.23

^a Model size, $100 \times 100 \times 10$; 5000 samples per trace

^b Using a 2 GHz Pentium-M (laptop) processor with 1 GB of RAM.

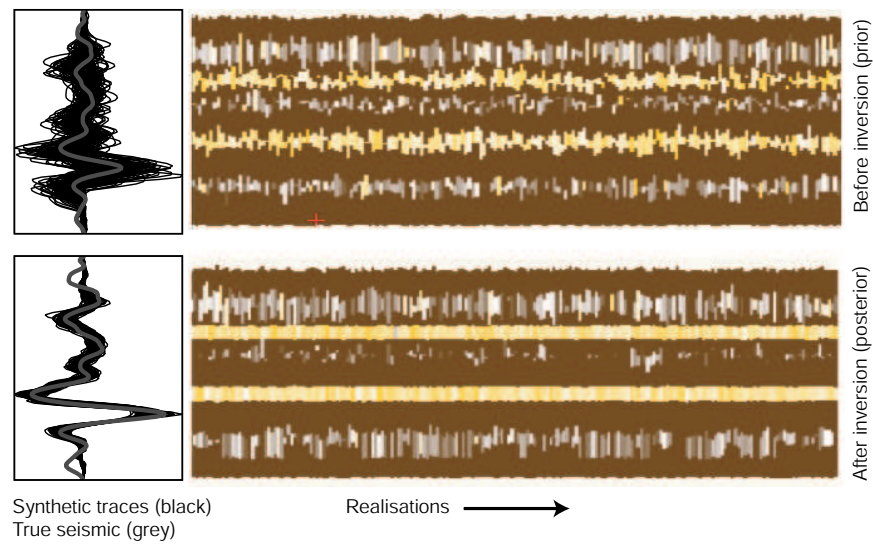


Figure 1: Layer-based realizations before (upper right) and after (lower right) seismic inversion, as produced by the DELIVERY code at a particular seismic trace. Synthetic traces corresponding to particular realizations are shown in black with the actual seismic data (grey) in the left insets.

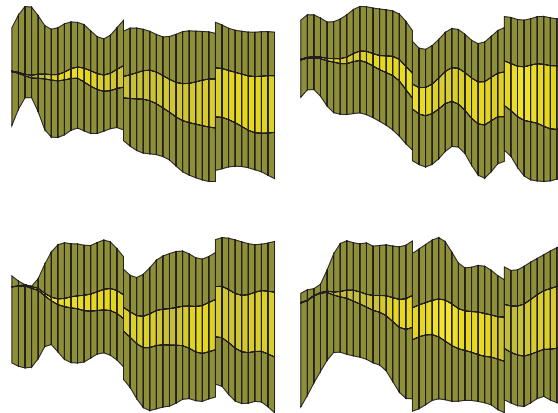


Figure 2: Four mesoscale layer-based realization cross-sections *in depth* of a simple 3-layer shale/sand/shale wedge test problem with Graben-like fault, as output from the DELIVERYMESSAGER code.⁷ Reflection surfaces that are well defined in time may still have appreciable depth variability, due to transverse variations and uncertainties in the velocity structure.

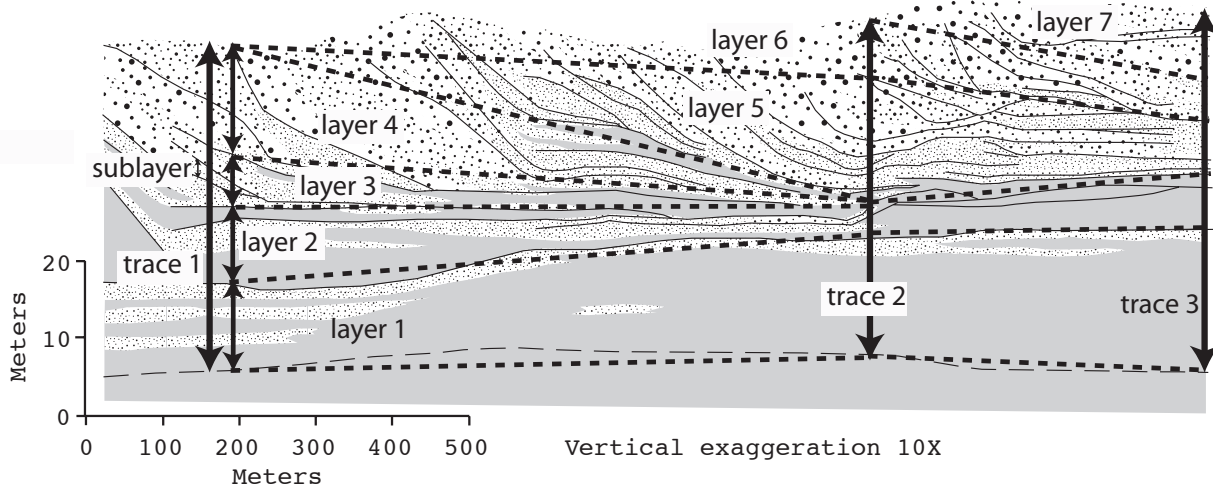


Figure 3: A trace is a line with composite properties informed by seismic data. It may comprise many layers. Sublayers are not modeled in this paper. This image is an interpreted outcrop data set.⁸

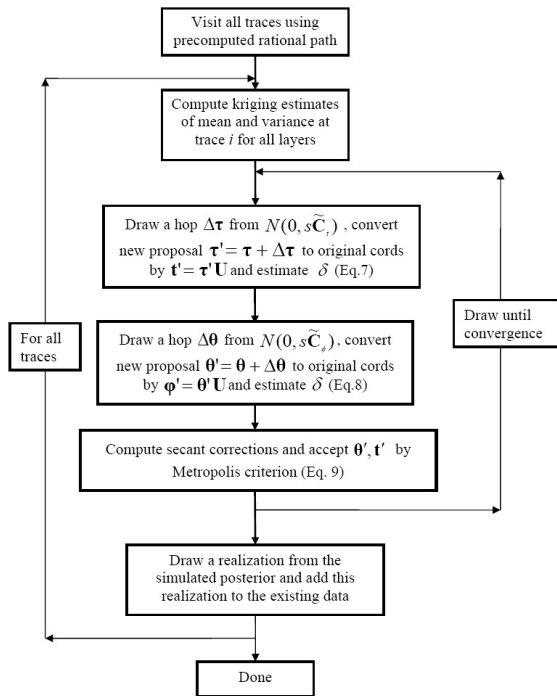


Figure 4: Flow chart for sequential simulation using Markov Chain Monte Carlo.

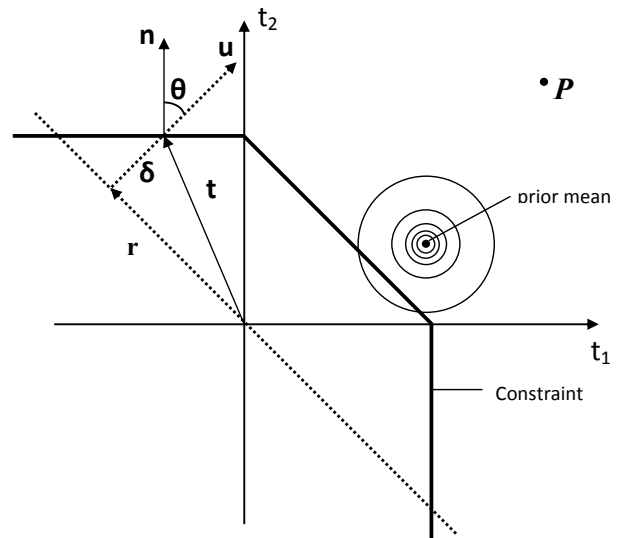


Figure 5: Exact constraint geometry with the prior sum more than the sum constraint. The point r is projected distance δ along u until it lies on the constraint surface.

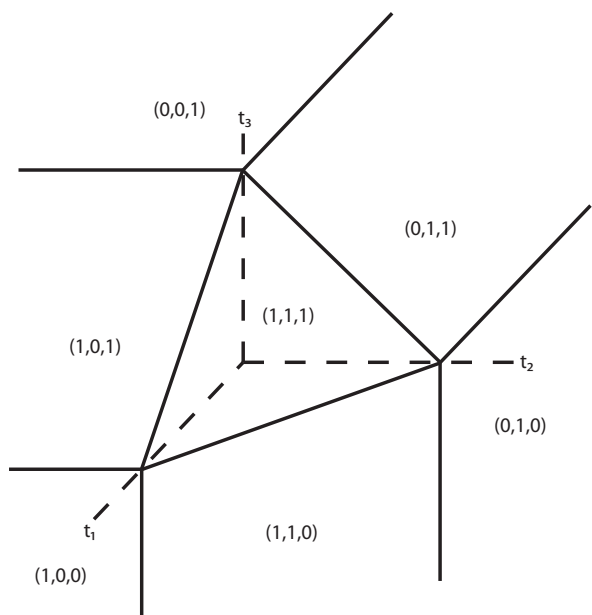


Figure 6: The three-dimensional constraint surface for a three layer case has $2^3 - 1 = 7$ facets. Hyperplanes intersect all t -axes at the total thickness constraint, H .

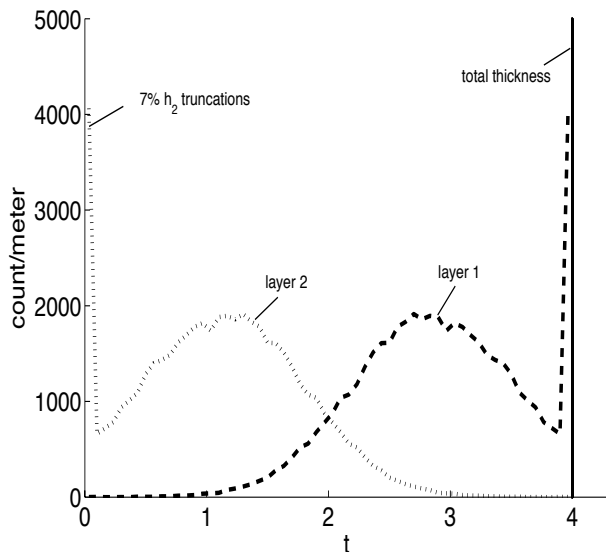


Figure 7: Thickness distribution for two layers. The prior and constraint are such that 7 percent of the realizations have $h_2 = 0$ ($t_2 < 0$) which creates the spike for layer 1 at $h_1 = 4$ m to yield $H = 4$ m.

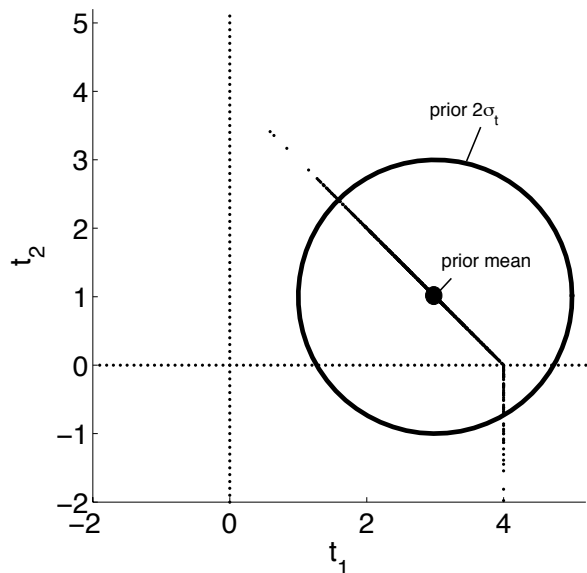


Figure 8: Simulations results for a two layer case with thickness priors $\bar{t} = (3 \text{ m}, 1 \text{ m})$, $\sigma_t = 1 \text{ m}$, and $H = 4 \text{ m}$. All the realizations are exactly on the constraint surface.

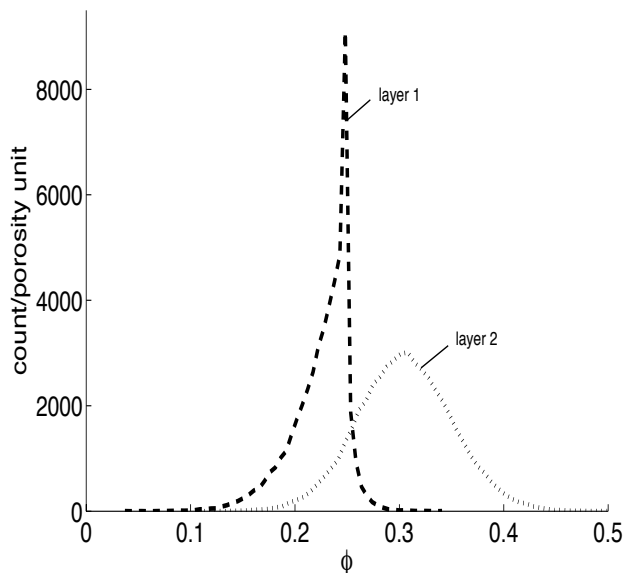


Figure 9: Porosity distributions for two layers. The constraints Φ and H are such that realizations having $\phi_2 > 0.25$ are not very probable and this skews the curve steeper on the right side.

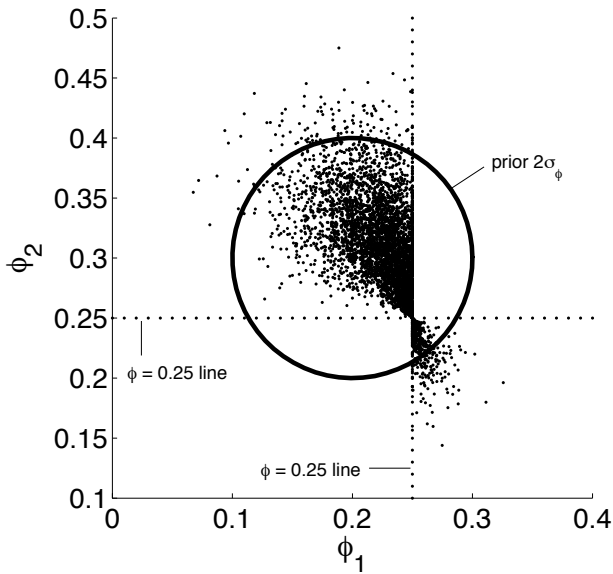


Figure 10: Simulation results of porosity for two layers. Because of the constraints the porosity of both the layers cannot be greater or less than 0.25 simultaneously. The prior for porosity is $\bar{\phi} = (0.2, 0.3)$, $\sigma_\phi = 0.05$.

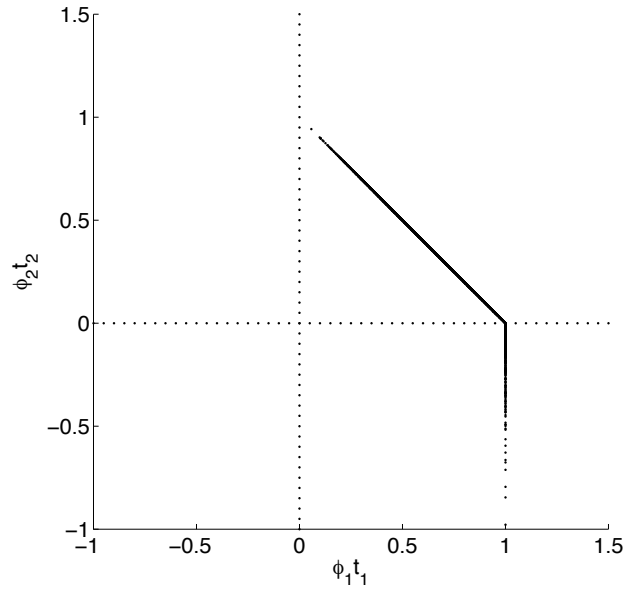


Figure 11: Simulations results for a two layer case with constraint $\Phi H = 1$ ($\Phi = 0.25$) m. All the realizations are exactly on the constraint surface.

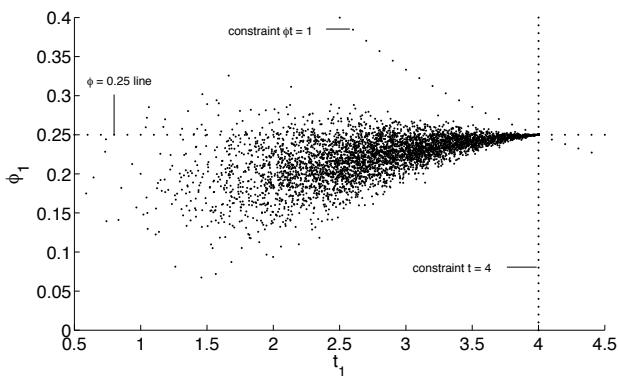


Figure 12: Cross plot for the distribution of porosity and thickness for layer one. As layer thickness reaches the total thickness constraint $h_1 = H = 4$ m, the porosity converges to average porosity and $\phi_1 = \Phi = 0.25$.

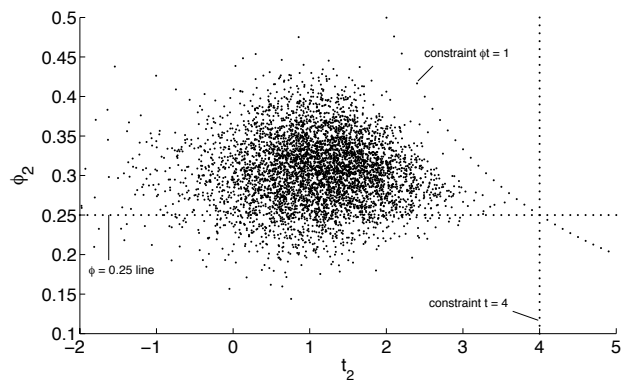
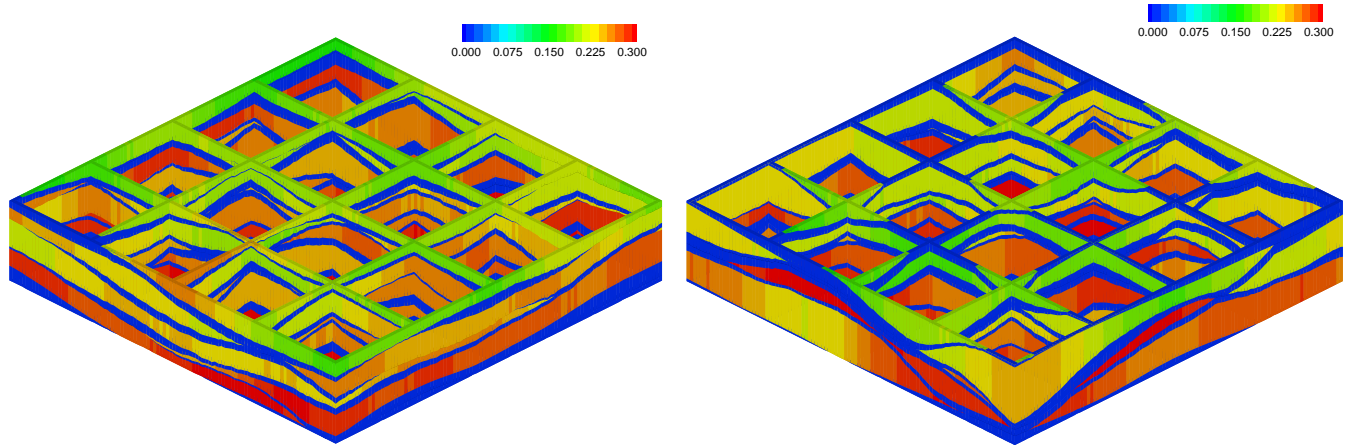
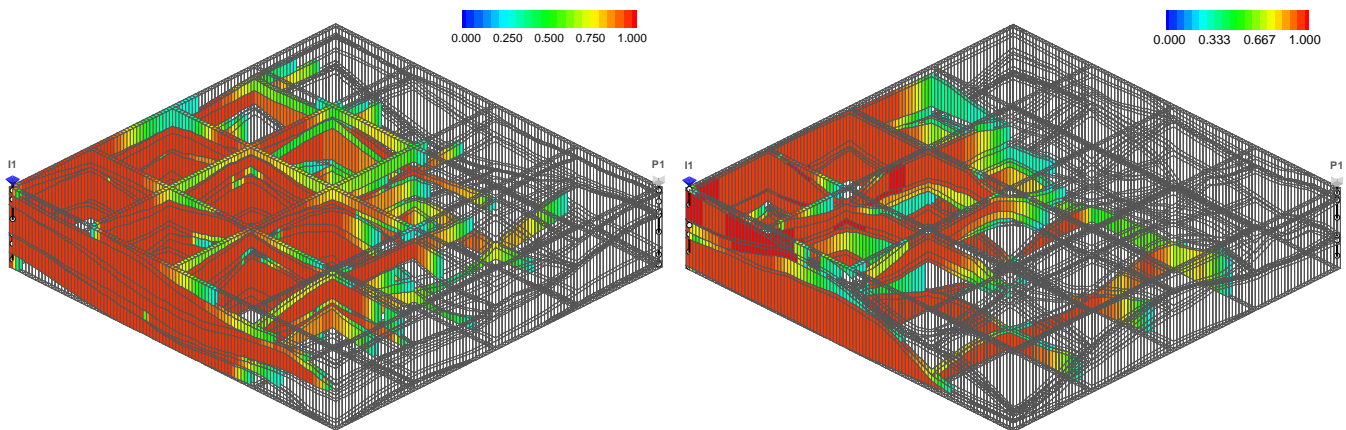


Figure 13: Cross plot for the distribution of porosity and thickness for layer two. As layer thickness reaches the total thickness constraint $h_2 = H = 4$ m, the layer porosity must converge to average porosity and $\phi_2 = \Phi = 0.25$.



(a) Porosity realization, base case (0000)

(b) Porosity realization, higher sand sill (00+0)



(c) Tracer concentration at breakthrough, base case (0000).

(d) Tracer concentration at breakthrough, higher sand sill (00+0).

Figure 14: Simulation on $100 \times 100 \times 10$ cornerpoint grids, areal extent is $X = Y = L = 1000$ m. Constraints used are $H = 20$ m, $H_s = 14$ m, and $\Phi H_s = 3.5$ m; $\Phi = 0.25$. Vertical exaggeration is tenfold for all figures. Dark blue layers are zero-porosity shales in (a) and (b). Compared with the base case, the high sand sill case has layers that are of less uniform thickness and have more truncation. Thus, the high-sill case has lower recovery efficiency.