

Consistent Downscaling of Seismic Inversion Thicknesses to Cornerpoint Flow Models

Subhash Kalla, SPE, Louisiana State University; Christopher D. White, SPE, Louisiana State University; James Gunning, SPE, Commonwealth Scientific and Industrial Research Organization; and Michael E. Glinsky, SPE, BHP-Billiton Petroleum Inc.

Summary

Accurate reservoir simulation requires data-rich geomodels. In this paper, geomodels integrate stochastic seismic inversion results (for means and variances of packages of meter-scale beds), geologic modeling (for a framework and priors), rock physics (to relate seismic to flow properties), and geostatistics (for spatially correlated variability). These elements are combined in a Bayesian framework. The proposed workflow produces models with plausible bedding geometries, where each geomodel agrees with seismic data to the level consistent with the signal-to-noise ratio of the inversion. An ensemble of subseismic models estimates the means and variances of properties throughout the flow simulation grid.

Grid geometries with possible pinchouts can be simulated using auxiliary variables in a Markov chain Monte Carlo (MCMC) method. Efficient implementations of this method require a posterior covariance matrix for layer thicknesses. Under assumptions that are not too restrictive, the inverse of the posterior covariance matrix can be approximated as a Toeplitz matrix, which makes the MCMC calculations efficient. The proposed method is examined using two-layer examples. Then, convergence is demonstrated for a synthetic 3D, 10,000 trace, 10 layer cornerpoint model. Performance is acceptable.

The Bayesian framework introduces plausible subseismic features into flow models, whilst avoiding overconstraining to seismic data, well data, or the conceptual geologic model. The methods outlined in this paper for honoring probabilistic constraints on total thickness are general, and need not be confined to thickness data obtained from seismic inversion: Any spatially dense estimates of total thickness and its variance can be used, or the truncated geostatistical model could be used without any dense constraints.

Introduction

Problem Statement. Reservoir simulation models are constructed from sparse well data and dense seismic data, using geologic concepts to constrain stratigraphy and property variations. Reservoir models should integrate sparse, precise well data and dense, imprecise seismic data.

Because of the sparseness of well data, stochastically inverted seismic data can improve estimates of reservoir geometry and average properties. Although seismic data are densely distributed compared to well data, they are uninformative about meter-scale features. Beds thinner than about $1/8$ to $1/4$ the dominant seismic wavelength cannot be resolved in seismic surveys (Dobrin and Savit 1988; Widess 1973). For depths of ≈ 3000 m, the maximum frequency in the signal is typically about 40 Hz, and for average velocities of $\approx 2,000$ m/s, this translates to best resolutions of about 10 m. Besides the limited resolution, seismic-derived depths and thicknesses are uncertain because of noise in the seismic data and uncertainty in the rock physics models (Gunning and Glinsky 2004, 2006). This resolution limit and uncertainties associated with seismic depth and thickness estimates have commonly limited the use of seismic data to either inferring the external geometry

or guiding modeling of plausible stratigraphic architectures of reservoirs (Deutsch et al. 1996).

In contrast, well data reveal fine-scale features but cannot specify interwell geometry. To build a consistent model, conceptual stacking and facies models must be constrained by well and seismic data. The resulting geomodels must be gridded for flow simulation using methods that describe stratal architecture flexibly and efficiently.

Objective. Our objective is to use probabilistic depth and thickness information from the layer-based seismic inversion code Delivery (Gunning and Glinsky 2004) to inform a downscaling algorithm operating on a cornerpoint grid. Delivery provides ensembles of coarse-scale geomodels that contain thickness and other property constraint information. These coarse-scale models must be downscaled to the flow model scale, honoring well data such as layer thicknesses, porosity and permeability (Doyen et al. 1997; Behrens et al. 1998). The downscaling must embrace conceptual geologic models for stratigraphic frameworks, especially layer correlation models between sparse conditioning points. This problem fits inside a larger workflow, where this integration of the geomodel, well data, and seismic data is referred to as "enforcement," and the associated algorithms comprise the software package known as Enforcer.

Gridding Considerations. Seismic constraints and priors are modeled on the quasivertical block edges, analogous to seismic traces. Simulation at the edges preserves geometric detail in cornerpoint models. The stochastic inversion assumes no trace-to-trace correlation, and the traces are not necessarily coincident with cornerpoint edges in the flow model. Geologically plausible lateral correlations are introduced, and seismic data are kriged to the (possibly nonvertical) cornerpoint edges using methods implemented in DeliveryMessenger; greater integration of the geomodel and a flow simulation is a subject of ongoing work (Glinsky et al. 2005; Gunning et al. 2007; Kalla et al. 2007b). Analogous seismic-scale frameworks are used in Delivery (Gunning and Glinsky 2004) for constructing prior estimates of layer locations, and are typically constructed using geomodeling software (Pet 2005), although quasimechanistic depositional modeling (Merriam and Davis 2001) or surface-oriented geostatistics algorithms (Pyrzc 2004) are possible alternatives.

Nature of the Seismic Constraints. The data used by the downscaling problem are typically realizations of the seismic inversion coarse-scale model "massaged" to the edges of columns of the cornerpoint grid. These inverted models contain the requisite coupling between geometry and rock properties which seismic inversion induces, plus the necessary spatial correlation behavior forced by the massaging algorithm. These coarse-scale models provide explicit constraints on the corresponding subgridded models, which are nontrivial to respect using conventional geostatistical algorithms for fine-scale heterogeneity.

A characteristic difficulty is that parameters of the fine-scale model such as thickness may have one-sided or mixture distributions (e.g., the mode of layer thickness may be zero in a cornerpoint model). Because of constraints to be imposed, linear estimation may prove inadequate. For example, if one wishes to ensure consistency both in thickness and in average porosity

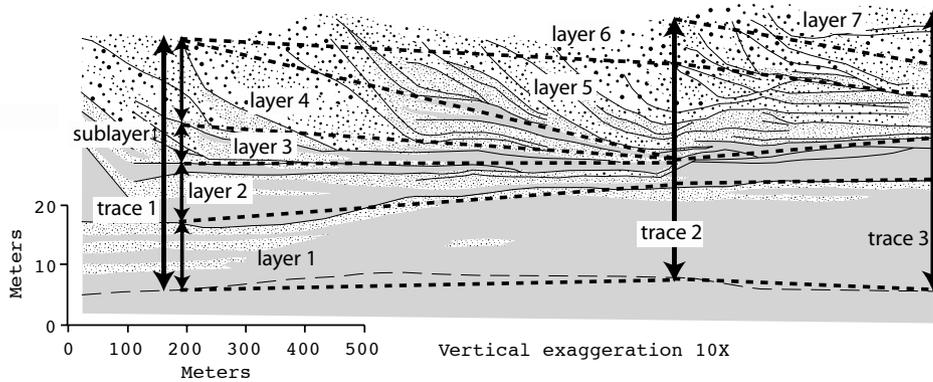


Fig. 1—A trace is a line with composite properties informed by seismic data. It may be composed of many layers. Sublayers are not modeled in this paper. This image is an interpreted outcrop data set (Willis and White 2000).

in a downscaling problem consisting only of vertical gridding refinement, the following equations must be considered at column of gridblock corners: $\sum_{k=1}^K h_k = H$, and $\sum_{k=1}^K h_k \phi_k = \bar{\Phi} H$, where K is the number of layers, k indicates a particular layer, ϕ is the porosity, h is a layer thickness, H is the total thickness predicted by seismic, and $\bar{\Phi}$ is the estimated average porosity at the trace scale. If layer porosity and thickness must be jointly estimated, the problem is nonlinear.

In summary, seismic downscaling to well and stratigraphic data on an arbitrary cornerpoint grid is a difficult problem, chiefly on account of the constraints, but also because of nonlinearities.

Use of Terms. The following conventions are used:

- Layers are generally not resolved by seismic data, but can be identified in wells. This terminology is illustrated in **Fig. 1** (Willis and White 2000). Sublayers might exist if some geomodel layers are not resolved in the cornerpoint grid layers. In this paper, well data is used only at the layer scale—sublayer log and core data must be upscaled.
- Traces are a segment of reservoir whose average properties are constrained by seismic, and will generally contain many layers. Traces correspond to the edges of the cornerpoint gridblocks [viz., COORD records, (Ecl 2004); (Ponting 1989)]. Conditioning data are a type of trace; order, properties, and thickness are specified at conditioning traces.
- A path is a sequence in which traces (or layers, or blocks) are visited. We use a quasirandom multigrid path.
- Multigrid paths are paths that preferentially visit widely spaced points early.
- The resolution matrix is the inverse of the covariance matrix, and closely related to the Hessian in an optimization problem.

Problem Formulation

Our approach is to combine diverse data elements in prior and likelihood expressions to obtain a posterior probability. The overall posterior distribution is approximated by the posterior obtained by a multigrid sequential simulation passing over all columns or column-blocks of the cornerpoint grid. Each column of blocks is simulated by sampling from a Bayesian posterior distribution conditional on hard data and previously visited columns by means of the priors, and collocated coarse-scale constraints by means of the likelihood. The prior distribution for each column is determined by solving an ordinary kriging system (Goovaerts 1997) using observations and previously simulated values. The seismic data are incorporated by means of a constraint on the sum of the layer thicknesses, which comes from a stochastic seismic inversion. In the proposed approach, layer thicknesses are modeled as truncated

Gaussian processes to allow for pinchouts; this model complicates imposition of the seismic sum constraint (Sampling Approach, later). The prior data and thickness constraints are combined in a Bayesian posterior form. Finally, the posterior is sampled using MCMC methods with auxiliary variables (Gelman et al. 2003).

An efficient approximation to the posterior covariance matrix is crucial to the success of this Bayesian approach. In this study, efficiencies are gained by assumptions regarding particular form of the covariance, which yield a computationally tractable matrix (see the Estimating the Prior subsection). This posterior covariance matrix is required by the sequential simulation algorithm, and encapsulates the compromise between prior information from kriging and total thickness constraints derived from seismic information.

For simplicity, we consider systems with a single thickness constraint. More general constraints are addressed in the Discussion section and other studies (Kalla et al. 2007b). Numerical methods and sampling methods are also discussed in later sections.

The Truncated Proxy for Thickness. A proxy t for thickness h is used. The untruncated proxy t is kriged to obtain prior distributions because kriging assumes variables are continuous but actual thickness h is non-negative. The proxy t may take on negative values, whereas h is truncated at zero. The probability of $t_k \leq 0$ corresponds to the probability that layer k is absent, locally:

$$P(h_k = 0) = \int_{-\infty}^0 dP(t_k). \dots \dots \dots (1)$$

Algorithm Outline. Before discussing details, the algorithm framework is presented (**Fig. 2**). First, the untruncated Gaussian surrogate for all conditioning data with $h = 0$ must be simulated. Then, a multigrid random path for a sequential simulation is generated. At each point on the path, the prior is estimated by kriging and the likelihood is used to update thicknesses at the trace by seismic data. To treat the possibility of zero thicknesses (or pinchouts), auxiliary variables are used, followed by a Metropolis-Hastings step to propose a new thickness vector. The chain is iterated to convergence, a sample vector \mathbf{t} is drawn, and the simulation then moves to the next trace in the path. Multiple paths can be used to generate multiple chains, in the same way sequential Gaussian simulations generate multiple realizations (Deutsch and Journel 1998).

Estimating the Prior. This step in the algorithm supplies prior means \bar{t} and variances $\sigma_{t_k}^2$ for all layers on a given trace. A few assumptions can simplify the kriging solution, and greatly improve efficiency (see the Numerical Considerations section).

- For many block shapes and grid spacings, traces can be

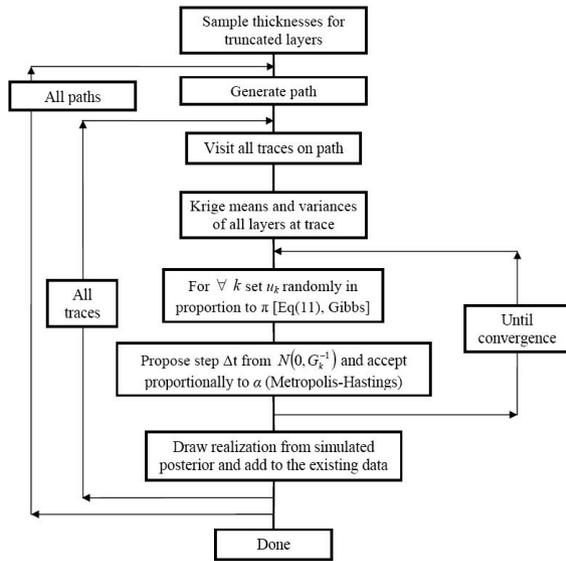


Fig. 2—Flow chart for sequential simulation using MCMC.

approximated as vertical when computing the kriging covariance matrix (i.e., small lateral trace displacement compared to trace spacing). Then the areal separation between the visited trace and each of its neighbors is constant for all layers and all trace-neighbor pairs.

- If in addition the covariance models are the same for all layers, then the covariance matrices will be the same on a layer-by-layer basis as well.

- Layer thicknesses may be *a priori* uncorrelated vertically at each trace. This may be reasonable, as the lateral thickness variations are likely more informative than the thicknesses of the layers above and below. This assumption seems particularly appropriate for turbidite systems, in which meter-scale beds may correspond to individual depositional events. Bed thicknesses then correlate strongly only within beds, with between-bed correlations being weak or even negative if compensatory deposition or scouring were occurring.

If all of these assumptions are reasonable, then the priors for each layer can be computed separately; the kriging matrices are identical for all layers, and therefore only one kriging system needs to be solved at each trace; and the prior variances in each column are then uniform. The prior means vary layer by layer. The tracewise-constant prior variance allows more efficient solution methods (see the Numerical Considerations section). These assumptions need not be imposed: this would make the kriging system(s) more expensive to solve, and the approximation to the posterior covariance will be more expensive to compute.

The neighbor list is extracted from the list of conditioning data and previously simulated points using a k-d tree (Bentley 1975) with specifications of desired points per quadrant. This search strategy is more efficient than most alternatives, especially on irregular grids. Also, assuming only two-dimensional layer thickness correlation implies that a two-dimensional search suffices, further improving search efficiency.

Cokriging or collocated kriging could be used to get prior covariances (Goovaerts 1997). Such a result could be integrated well with the seismic data, which provide local correlated estimates of trace-scale properties (Gunning and Glinsky 2004). Alternatively, these essential rock physics correlations can be preserved using a cascading workflow originating from seismic inversions (Kalla et al. 2007b).

If vertical correlations are included, separate neighbor lists may be required for each of the K_ℓ layers at the trace, or a single list could be used for all layers. While the single list might require

solving a larger kriging system, it would only require solving one kriging system for all K layers.

The Posterior Resolution Matrix. The seismic data are combined with the prior to obtain posterior probability. The seismic data are incorporated as a constraint on the total thickness, \bar{H} , with resolution $\frac{1}{\sigma_H^2}$ obtained from a stochastic inversion using Delivery (Gunning and Glinsky 2004).

The posterior probability for any thickness vector \mathbf{t} is, from Bayes' rule,

$$\pi(\mathbf{t}|H, \mathbf{d}_{\ell k}) = \frac{p(H|\mathbf{t}, \mathbf{d}_{\ell k}) p(\mathbf{t}|\mathbf{d}_{\ell k})}{p(H|\mathbf{d}_{\ell k})}, \dots \dots \dots (2)$$

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 product of the likelihood and prior are proportional to the posterior, without normalizing term in the denominator, which does not depend on \mathbf{t} . That is,

$$\pi(\mathbf{t}|H, \mathbf{d}_{\ell k}) \propto p(H|\mathbf{t}, \mathbf{d}_{\ell k}) p(\mathbf{t}|\mathbf{d}_{\ell k}). \dots \dots \dots (3)$$

We assume that departures from the prior (\bar{t}_k) and updating (\bar{H}) data means are normally distributed with standard deviations σ_{t_k} and σ_H , respectively. The assumptions apply to departures, not values, and so the resulting posterior probabilities are not assumed to be normal, as will be demonstrated in later examples. The multivariate prior distribution of \mathbf{t} is

$$p(\mathbf{t}|\mathbf{d}_{\ell k}) = \frac{1}{(2\pi)^{\frac{\kappa}{2}} |\mathbf{C}_p|^{\frac{1}{2}}} \exp \left[-\frac{1}{2} (\mathbf{t} - \bar{\mathbf{t}})^T \mathbf{C}_p^{-1} (\mathbf{t} - \bar{\mathbf{t}}) \right], \dots (4)$$

where \mathbf{C}_p is the prior or kriging covariance matrix, which is of rank K with the kriging variances $\sigma_{t_k}^2$ along the diagonal. The number of active layers (with $t_k > 0$) is κ .

Similarly, we can express the updating constraint on H as a Gaussian likelihood,

$$p(H|\mathbf{t}, \mathbf{d}_{\ell k}) = \frac{1}{\sqrt{2\pi}\sigma_H} \exp \left[-\frac{(H - \bar{H})^2}{2\sigma_H^2} \right], \dots \dots \dots (5)$$

where

$$H = \mathbf{t}^T \mathbf{T}, \dots \dots \dots (6)$$

and

$$T_k = \begin{cases} 0 & \text{if } t_k < 0 \\ 1 & \text{otherwise} \end{cases} \dots \dots \dots (7)$$

The conditioning on $\mathbf{d}_{\ell k}$ in Eq. 5 is indirect, due to the conditioning of \mathbf{t} on $\mathbf{d}_{\ell k}$. The product of Eqs. 4 and 5 is the proportional to the posterior, Eq. 3. This product can be converted to a quadratic form by taking the logarithm, giving

$$-2 \ln [\pi(\mathbf{t}|H, \mathbf{d}_{\ell k})] = \ln \left[(2\pi)^K |\mathbf{C}_p| \right] + \ln \left(2\pi\sigma_H^2 \right) + (\mathbf{t} - \bar{\mathbf{t}})^T \mathbf{C}_p^{-1} (\mathbf{t} - \bar{\mathbf{t}}) + \frac{(\mathbf{t}^T \mathbf{T} - \bar{H})^2}{\sigma_H^2}. \dots (8)$$

We seek a stationary point in the posterior probability by setting the gradient with respect to \mathbf{t} of Eq. 8 to zero, viz.,

$$\mathbf{C}_p^{-1} (\mathbf{t} - \bar{\mathbf{t}}) + \frac{(\mathbf{T}\mathbf{T}^T \mathbf{t} - \bar{H})}{\sigma_H^2} = 0.$$

The Hessian, \mathbf{G} of Eq. 8 is the desired resolution matrix (which is the inverse of the posterior covariance):

$$\mathbf{G} = \mathbf{C}_p^{-1} + \mathbf{T}\mathbf{T}^T / \sigma_H^2. \dots \dots \dots (9)$$

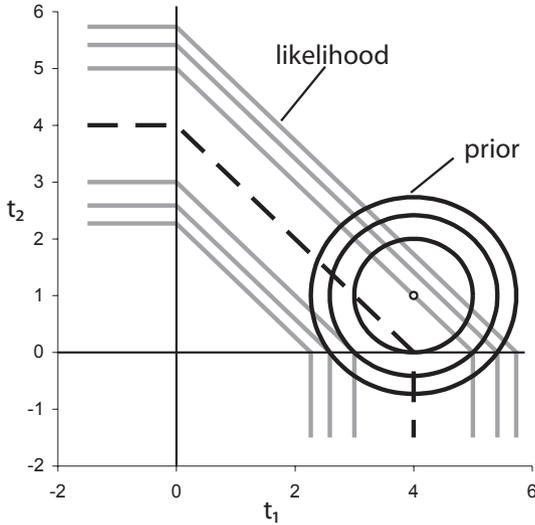


Fig. 3—Contours of minus log likelihood and minus log prior distributions for a two-layer system, with $\bar{H} = 4, \sigma_H = 1, \bar{\mathbf{t}} = (4, 1)$, and $\sigma_t = 1$. Contours are in increments of one in $\Delta\left(\frac{t-\bar{t}}{\sigma_t}\right)$ or $\Delta\left(\frac{\Sigma t-H}{\sigma_H}\right)$, with values of zero along exactly honoring the thickness sum (dashed line) and where $\mathbf{t} = \bar{\mathbf{t}}$ (small circle). Consistent units.

If the prior covariance matrix is diagonal, \mathbf{C}_p^{-1} and \mathbf{G} are easy to compute. For $T_k = 1, \forall k$, the Hessian has the form

$$\mathbf{G} = \begin{pmatrix} \frac{1}{\sigma_{t_1}^2} + \frac{1}{\sigma_H^2} & \frac{1}{\sigma_H^2} & \cdots & \frac{1}{\sigma_H^2} \\ \frac{1}{\sigma_H^2} & \frac{1}{\sigma_{t_2}^2} + \frac{1}{\sigma_H^2} & \cdots & \frac{1}{\sigma_H^2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{\sigma_H^2} & \frac{1}{\sigma_H^2} & \cdots & \frac{1}{\sigma_{t_K}^2} + \frac{1}{\sigma_H^2} \end{pmatrix} \dots \quad (10)$$

If the prior variances $\sigma_{t_k}^2$ are all equal (see the Problem Formulation section), \mathbf{G} is Toeplitz (Golub and van Loan 1996), and in fact a particularly simple form, with all super- and subdiagonals equal. Note that the Hessian is constant except for the dependence of \mathbf{T} on \mathbf{t} ; this is a lurking nonlinearity.

Prior and Likelihood Distributions in 2D

Important features of higher-dimensional cases are easily visualized for a system with two layers (Fig. 3). The dashed line in Fig. 3 is the thickness sum constraint, and lines parallel to it are isoprobability contours. In three dimensions, the dashed line in Fig. 3 corresponds to a triangle with vertices on each t -axis at \bar{H} ; increasing \bar{H} shifts the high-likelihood region away from the origin, but with no change in slope. Tighter seismic constraints will narrow the width of the high-likelihood region.

The assumption of equal prior variances implies the prior has the circular shape shown in Fig. 3; it would be ellipsoidal if prior variance differed by layer, and it would be an inclined ellipsoid if the layer thicknesses were correlated. Such priors could be sampled using methods discussed in this paper, but the resolution matrices would be non-Toeplitz and the algorithms would be slower.

In this example, the prior mean thicknesses (4m and 1m for the two layers) sum to greater than the mean trace thicknesses (4m), so the prior center of mass [circles in Fig. 3; Eq. 4] lies above the maximum likelihood line [dashed line in Fig. 3; Eq. 5, for $t_k > 0, \forall k \in \{1, 2\}$]. Because \bar{t}_2 is small compared to \bar{H} , there is substantial prior (and posterior) probability that t_2 is negative, yielding many realizations with $h_2 = 0$.

If no layer kriging data were used and the seismic data were considered exact, any layer thickness pair (t_1, t_2) along the dashed line with 45 degree slope could be used. Conversely, in a sequential simulation not conditioned to seismic, the layer thicknesses would simply be drawn from the prior (Fig. 3).

Sampling problems are caused by the nonlinearity [Eqs. (6, 7)] apparent as slope discontinuities in the likelihood where the axes intersect the contours of the likelihood surface (Fig. 3). This nonlinearity may dominate sampling where the prior admits significant probability of one or more thicknesses being zero (as is the case for layer 2 in Fig. 3). In higher dimensions, many layers may be pinched out at any given trace, and a method to move around these corners while sampling is needed (see the Auxiliary Variables to Treat Pinchouts subsection).

Sampling Approach

Because the log-posterior surface is quadratic with constraints (Eq. 9), the most likely *a posteriori* thickness vector could be found by constrained quadratic programming (Nocedal and Wright 1999). However, our goal is simulation, not maximum *a posteriori* estimation, so we sample from the posterior. We use an MCMC method (Fig. 2).

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, multigrid path.

Observed Thicknesses of Zero. Some layers may be absent at conditioning points, $h_k = 0$. For these points, we only know that $t_k \leq 0$ at these points, but require a particular value of t_k to use in estimating means at the traces to be simulated. One could simply draw random numbers in the range $[0, P(h_k = 0)]$ and apply an inverse normal transformation, but this decorrelates the variables. Instead, we precondition these data using a Gibbs sampler to preserve the correlation (see Appendix).

Auxiliary Variables to Treat Pinchouts. The posterior distribution has marked slope discontinuities at the interfaces in parameter space where layers pinch out (i.e., the hyperplanes $t_k = 0$; Fig. 3). Standard MCMC methods based on small jumping proposals will diffuse around such distributions very slowly. It has been shown that introducing auxiliary variables \mathbf{u} can promote mixing, or alteration between states, in difficult MCMC problems with related “configurational stiffness” characteristics (Higdon 1998). Auxiliary variable methods use an augmented posterior probability space:

$$\pi(\mathbf{u}, \mathbf{t}) = \pi(\mathbf{t})\pi(\mathbf{u}|\mathbf{t}), \dots \dots \dots \quad (11)$$

where the augmented binary variables \mathbf{u} ($u_k \in \{0, 1\} \forall k \in \{1 \dots K\}$) are chosen to align samples in the directions of maximum posterior considering the bends in the likelihood. When the sampling kernel in the MCMC algorithm is near the slope discontinuities, these auxiliary variables can change from zero and one (or *vice versa*), and allow the sampling direction to change.

The term $\pi(\mathbf{u}|\mathbf{t}) [= \prod_{k=1}^K \pi(u_k|t_k)]$ is a conditional probability for the auxiliary variables, which may be constructed in any helpful way. In our case, we construct the conditional to help detect the kinks in the posterior that occur when layers pinch out. One possible choice of a symmetric form is

$$\pi(u_k = 1|t_k) = \begin{cases} 1 - \frac{1}{2 + t_k/\sigma_{\pi k}} & \text{if } t_k \geq 0 \\ \frac{1}{2 - t_k/\sigma_{\pi k}} & \text{otherwise} \end{cases}, \dots \dots \quad (12)$$

where $\sigma_{\pi k}$ is a univariate approximation to the multivariate posterior covariance,

$$\frac{1}{\sigma_{\pi k}^2} = \frac{1}{\sigma_{t_k}^2} + \frac{\kappa}{\sigma_H^2} \dots \dots \dots \quad (13)$$

That is, $\sigma_{\pi k} \approx \sum_{j=1}^K G_{kj}$, (Eq. 10). κ is the current number of active layers; $\kappa = \sum_{k=1}^K T_k \leq K$.

Sampling from the augmented posterior distribution is performed by alternating Gibbs samples for the auxiliary variables with the Metropolis-Hastings samples for the thicknesses t_k . The Gibbs sampling scans over the layers. At each layer, a uniform $[0, 1]$ random number is drawn. If the random number is less than $\pi(u_k = 1|t_k)$, u_k is assigned 0. When the u_k for all K layers have been simulated, we construct a resolution matrix (for step size and direction dependent on \mathbf{u}) from which jumping proposals are formed, which are well tuned for the current configuration of the system. The auxiliary variables create an adaptively varying proposal kernel that does not break reversibility.

The Gibbs sample gives a list of likely active layers at the current iterate in \mathbf{u} .

Metropolis-Hastings Step. The new kernel obtained from the Gibbs step (previous section) is used to sample a new thickness vector \mathbf{t} using a Metropolis-Hastings step. Let the number of active layers be κ , $\kappa \leq K$. At each trace, a resolution matrix of rank K is constructed and its Cholesky factors are computed. The resolution matrix $\mathbf{G}_\kappa = \mathbf{C}_\pi^{-1} + \mathbf{u}\mathbf{u}^T/\sigma_H^2$ is used to make the MCMC jumping proposal (Eq. 14). The appropriate resolution and inverse matrices are computationally inexpensive for the simple Toeplitz resolution matrix used in the proposed approach (see the Numerical Considerations section). The Hessian \mathbf{G} and the posterior covariance $\mathbf{C}_\pi = \mathbf{G}^{-1}$ are of rank K , but matrix inverse used in sampling is of lower rank κ (Numerical Considerations, later). The Cholesky factor $\mathbf{L}_{C\pi}$ of the covariance matrix (the Cholesky factorization is $\mathbf{C}_\pi = \mathbf{L}_{C\pi}\mathbf{L}_{C\pi}^T$) is multiplied into a κ -long vector of random normal variables $\mathbf{r} \sim [N(0, 1)]$ to produce a vector $\Delta\mathbf{t}$ of proposed changes in \mathbf{t} ,

$$\Delta\mathbf{t} = s\mathbf{L}_{C\pi}\mathbf{r}, \dots\dots\dots (14)$$

so that $\Delta\mathbf{t} \sim N(0, s^2\mathbf{G}_\kappa^{-1})$, where s is a scalar chosen for sampling efficiency. Typically $s^2 = 5.76/\kappa$ for large κ (Gelman et al. 2003). This vector is rank κ , and the changes must be sorted back into \mathbf{t} by referencing \mathbf{u} . We can compute the likelihood at the new point $\mathbf{t}' = \mathbf{t} + \Delta\mathbf{t}$, using Eq. 5. The Metropolis-Hastings transition probability is then (Gelman et al. 2003)

$$\alpha = \min \left(1, \frac{\pi(\mathbf{t}'|H, \mathbf{d}_{\ell k}) \prod_{k=1}^K \pi(u_k|t'_k)}{\pi(\mathbf{t}|H, \mathbf{d}_{\ell k}) \prod_{k=1}^K \pi(u_k|t_k)} \right) \dots\dots\dots (15)$$

Eq. 15 is similar to the standard Metropolis-Hastings ratio, but has been modified to include the auxiliary variables so that reversibility is maintained. The proposed transition $\Delta\mathbf{t}$ is then accepted with probability α , and the algorithm proceeds to the next Gibbs sample for the auxiliary variables.

Numerical Considerations

The Toeplitz form of the posterior resolution matrix and subsidiary assumptions simplify computations (see the Estimating the Prior subsection). Because of these simplifications, only two matrix solutions are required per trace: a Cholesky factorization of the kriging matrix (which is dense and not Toeplitz, with rank equal to the number of neighbors used, N_ℓ), and the factorization of the inverse of the Toeplitz resolution matrix (rank K_ℓ and very inexpensive). If the Toeplitz-yielding assumptions were not made, K_ℓ rank- $\sum_{k=1}^{K_\ell} N_{\ell k}$ kriging systems are required at each trace ℓ . Even more prohibitive, the posterior resolution matrix \mathbf{G} would have to be refactored every time any t_k flips from a positive to nonpositive state. Because this occurs deep within the sampling method (see the Sampling Approach section), this would result in a remarkable loss in efficiency.

To carry out the simulation, we need the Cholesky factor $\mathbf{L}_{C\pi}$ of the posterior covariance matrix, $\mathbf{C}_\pi = \mathbf{G}^{-1}$. With $\mathbf{L}_{C\pi}$, we can generate correlated normal deviates, $\Delta\mathbf{t}$, from uncorrelated random normal input vectors, $\mathbf{r} \sim N(0, 1)$, $\Delta\mathbf{t} = \mathbf{L}_{C\pi}\mathbf{r}$ (see the Metropolis-Hastings Step subsection) (Goovaerts 1997). For the special Toeplitz matrices, the factor $\mathbf{L}_{C\pi}$ can be computed from the Cholesky factor of the resolution matrix \mathbf{G} . That is, factor \mathbf{G} to get

\mathbf{L}_G , invert \mathbf{L}_G by backsubstitution to get \mathbf{L}_G^{-1} (inexpensive because the matrix is triangular), and take the persymmetric transpose (Golub and van Loan 1996) of \mathbf{L}_G^{-1} . This is the Cholesky factor of \mathbf{C}_π , $\mathbf{L}_{C\pi}$.

The rank “downdate” from K to $\kappa < K$ is the lower rank- κ triangle of $\mathbf{L}_{C\pi}$. The matrix rank changes whenever the auxiliary variable transitions between zero and nonzero. Because of the Toeplitz form, the required factored correlation matrices $\mathbf{L}_{C\pi\kappa}$, regardless of the number of active layers κ (or rank), can be computed from a single factoring of the rank- K covariance and inverse to get $\mathbf{L}_{C\pi}$ and taking the appropriate rank- κ submatrix.

In combination, the efficient factorization method for the posterior rank- K covariance matrix and determination of $\mathbf{L}_{C\pi\kappa}$ for all possible pinchout combinations makes this algorithm efficient. Precise work estimates for these matrix calculations have not been done, but an upper bound is the work done for a general Toeplitz matrix (Golub and van Loan 1996), inverting the resolution matrix and factoring that inverse to get $\mathbf{L}_{C\pi}$. For that less efficient approach, the inverse of the Toeplitz resolution matrix requires $W \propto K^3$ floating operations (flops), and further work $W \propto K^4$ flops is required for the factoring. In comparison, the proposed method is at worst $W \propto K^3$ for the inverse and all factors, a full order of improvement (see the Performance subsection).

Simulations of Two-Layer Systems

Several two-layer simulations illustrate the behavior of the data integration algorithm. Different combinations of prior and updating data variance are considered, along with perfectly consistent vs. slightly contradictory prior means and constraints. Results are summarized in **Table 1**.

Tight Sum Constraint. This case assumes the sum of the layer prior means is equal to the trace mean, but the layer thicknesses are poorly resolved (**Fig. 4**). Because the means are consistent and the constraint variance is relatively small, the simulations tightly cluster around the constraint line, and the posterior means of \mathbf{t} are near their prior means, although the correlation induced by the constraint is marked (covariance column, Table 1). Moreover, many realizations have \mathbf{t} near $(4, 0)^T$ (which is very unlikely in the prior) because of the relatively tight seismic constraint ($\sigma_t/\sigma_H = 10$). The bend in the posterior caused by the pinchout is clearly seen below $t_2 = 0$ (**Fig. 4a**). The posterior layer variances are reduced because of the added data in the constraint (eigenvalues, Table 1). The axial (maximum) standard deviation is the same for the posterior as for the (isotropic) prior, but the transverse standard deviation is significantly reduced. The univariate histograms of t are slightly non-Gaussian, and truncation makes the histograms of h depart even more. The strict seismic constraint has transformed the uncorrelated prior into a posterior in which the thicknesses are strongly negatively correlated, a natural outcome of a sum constraint.

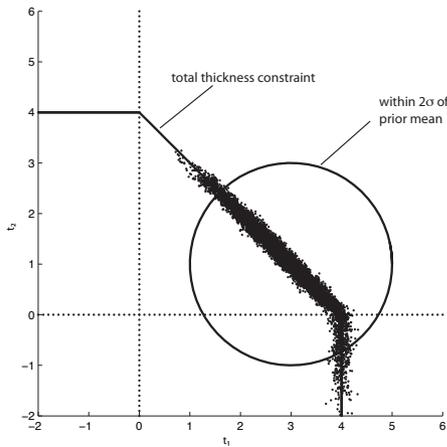
Loose Constraint and Prior. As for the previous case, the prior means are taken to be consistent with the seismic constraint. However, the variances of both prior and constraint are higher for this case. The data are therefore more dispersed, and it is more likely that layer 2 is assigned a zero thickness (**Fig. 5**). As before, although t appears nearly Gaussian in the univariate histograms, h will be truncated to nonnegative values and is thus non-Gaussian, and the bend in the posterior at $t_2 = 0$ is observed.

Sum of Prior Means less than Constraint. A mismatch between the prior layer means and the thickness constraint shifts the axis of the cloud of simulations points above or below the constraint line (**Fig. 6**). In this case, both layer thicknesses are increased from their priors to better match the seismic constraint. For the moderate standard deviation and prior means much greater than zero, few truncations occur and the posteriors are nearly Gaussian. For this nearly multi-Gaussian case, the constraint has transformed the isotropic, uncorrelated prior thicknesses (**Fig. 3**)

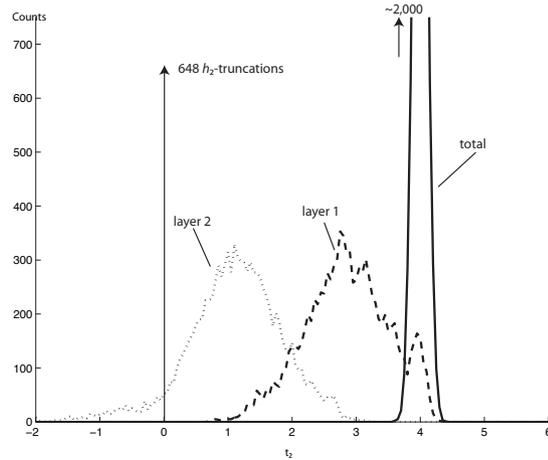
TABLE 1—PARAMETERS AND RESULTS FOR 2-LAYER SIMULATION

Case	Prior			Constraint		Posterior						
	\bar{t}_1	\bar{t}_2	σ_t	\bar{H}	σ_H	\bar{t}_1	\bar{t}_2	Covariance of t		σ_t^*	\bar{H}	σ_H
Tight	3.0	1.0	1.0	4.0	0.1	2.86	1.11	0.46	-0.50	1.01	4.00	0.10
Loose	3.0	1.0	1.0	4.0	0.5	2.97	0.97	-0.50	0.59	0.14	4.00	0.49
$\mathbf{T}^T \mathbf{t} < \bar{H}$	3.0	1.0	0.5	6.0	0.5	3.65	1.66	0.53	-0.46	1.03	4.00	0.49
								0.16	-0.08	0.49	5.31	0.41
								-0.08	0.16	0.28		

*These are the square roots of the largest and smallest eigenvalues, respectively, of the posterior covariance matrix



(a) Scattergram, $N = 8000$



(b) Histogram, bin size $\Delta t = 0.05$

Fig. 4—Simulation results for a two-layer case with inaccurate layer thickness but total thickness ($h_1 + h_2$) tightly constrained. $\bar{H} = 4$, $\bar{\mathbf{t}} = (3, 1)^T$, $\sigma_H = 0.1$, and $\sigma_t = 1$; consistent units.

to a strongly correlated, more compact posterior. Because the prior and constraint variances are equal, the mean of the scatter cloud is shifted roughly one-half the distance from the prior toward the constraint, as would be expected (Table 1) (Gelman et al. 2003).

Convergence. MCMC methods may converge too slowly to be practical, or may have multiple modes such that multiple chains or methods to switch between modes are needed. In numerical experiments undertaken so far, these potential problems do not appear to be too severe in this algorithm.

Convergence is critiqued by examining posterior distribution statistics over many iterations (Gelman et al. 2003). For a variety of cases examined, the means converge in no more than $\approx 1,000$ iterations, and the variances stabilize in no more than $\approx 2,500$ iterations. That is, some 2,500 iterations are needed for the chain to begin sampling the posterior reliably; this is referred to as the “burn-in;” samples prior to burn in are discarded before the chain is used to simulate the posterior. This number of iterations, while large, is not prohibitive if the proposal method is computationally inexpensive (see the Numerical Considerations section) and the acceptance rate is not too small. For a realistic 3D synthetic problem, the proposed method attains a sampling rate of almost 200,000 iterations per second and an acceptance rate averaging ≈ 0.4 , which makes such long burn-in requirements manageable (see the Synthetic 3D Cases section).

Chains started in widely dispersed parts of t -space converge to the same posterior (Fig. 7). This was expected, based on the relatively simple form of the posterior resolution matrix, \mathbf{G} . The early behavior depends on the starting point (Fig. 7a): chains that move in from the flanks of the constraint (transverse paths) take large, efficient steps; those moving along the axis zig-zag and advance more slowly. The latter is the classic behavior of movement

along a trough in a minimization problem where the eigenvalues of the Hessian differ markedly (Table 1). After many iterations, all chains are sampling the same region (Fig. 7b), and the post-burn-in chains are statistically indistinguishable.

The simple 2D examples indicate the algorithm is reproducing expected results in limiting cases.

Synthetic 3D Cases

A synthetic 3D data set is used to test and illustrate the MCMC simulation method. Prior (range and sill of semivariogram, R) and updating data (trends in \bar{H} and σ_H) parameters are varied to illustrate behavior, and algorithm performance is discussed.

For all cases, x to y extent is 1000×1000 m, the number of grids in those directions are 100×100 respectively, and the number of layers is 10. The framework for the reference model was created by randomly placing objects with scaled bi-Gaussian thickness variations in x and y ; for the 1 km areal grid, an isotropic standard deviation, $\sigma = 500$ m, was used to compute layer thickness with $h(x, y) = h_{\max} \exp \left[\frac{(x-\bar{x})^2 + (y-\bar{y})^2}{\sigma^2} \right]$. This object-based method with Gaussian thickness variations is not the same as a Gaussian covariance process. The object models are used only to create conditioning data. Twenty-five traces were used in cases discussed in this section; the algorithm has also been used with no conditioning traces and with up to 200 conditioning traces.

Illustrative Cases. Four different cases show features of the data integration method (Fig. 8). With short ranges, termination is more common, although the average layer thickness is similar to the longer range (Figs. 8a and 8b). There is little noise, unlike what is commonly observed in Gaussian processes; the layer thicknesses vary smoothly and plausibly, and near-zero thicknesses do not

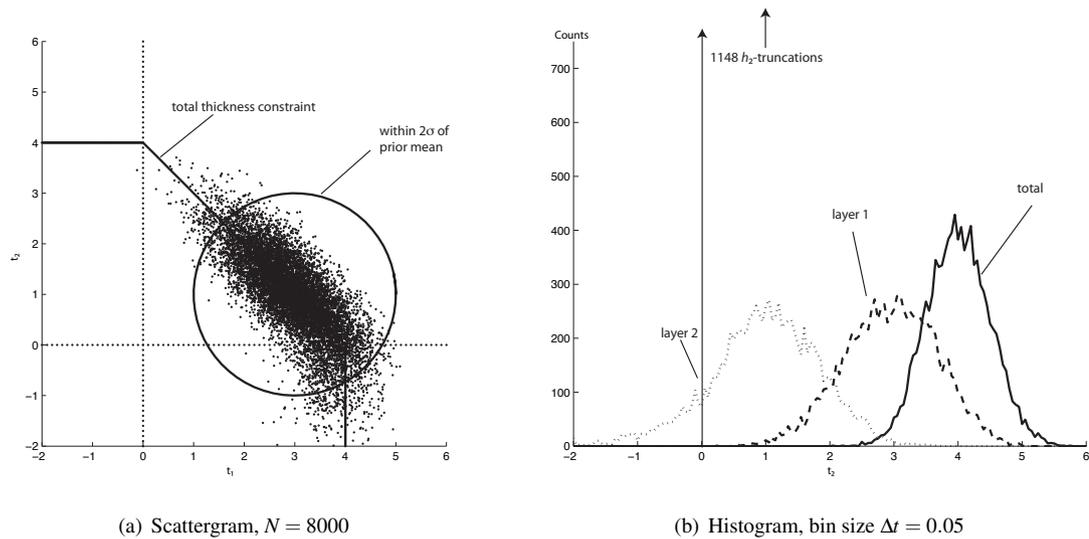


Fig. 5—Simulation results for a two-layer case with inaccurate layer and total thicknesses ($h_1 + h_2$). $\bar{H} = 4$, $\bar{\mathbf{t}} = (3, 1)^T$, $\sigma_H = 0.5$, and $\sigma_t = 1$; consistent units.

appear in isolated areas; this results from the truncation rules and the smooth Gaussian variogram. The pinchout pattern is clearer in the longer-range case (Fig. 8b). In particular on the first cross-section in the left, the light layer near the base and the dark layer in the middle appear to taper and pinch out smoothly; this behavior is more characteristic of object models than most covariance-based simulations.

Seismic data may imply a thickness trend (Fig. 8c). The seismic trend will be reproduced in the simulation, with a precision conditioned on the inferred seismic thickness variance, σ_H . If the seismic variance is higher for smaller mean thickness, low thicknesses fluctuate more, as can be seen by comparing the left front edges of Figs. 8c and 8d. For the low variance case (Fig. 8c), the edge panel is of nearly uniform thickness; the nonuniform variance case (Fig. 8d) has much greater fluctuation on the left edge.

Although based on a synthetic case, these results indicate that the proposed method can reproduce complex pinchout layering and plausible seismic trends. The number of pinchouts can be quite large in complex cornerpoint grids; 30,608 of 100,000 trace segment are zero-thickness in one of the example cases (Fig. 8c). The complex pinchout structure is obtained even though the conditioning data are not especially dense (Fig. 8d).

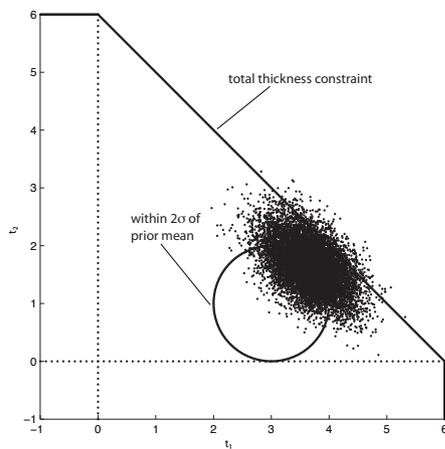
Performance. For adequate performance, an MCMC simulation should converge to its target distribution in as few steps as possible. A large step size helps explore the posterior in few steps. On the other hand, large steps are more likely to be rejected, wasting computations on a sample that is not retained. The step size is usually adjusted indirectly, by scaling the posterior covariance (which is used to generate steps; see the Metropolis-Hastings Step subsection). For the system examined, the covariance is not scaled; this gives a step size of the order of the square root of the smallest diagonal element in the posterior covariance matrix. In high-dimensional problems, it may be more appropriate to use $\hat{\mathbf{C}}_\pi = \frac{5.76}{K} \mathbf{C}_\pi$ to ensure adequate acceptance rates (Gelman et al. 2003). Although the unscaled covariance yields larger steps for $K = 10$, the test cases had acceptance rates of 30 to 40 percent. This step size and acceptance rate appears to yield good convergence, thorough exploration of the posterior, and smooth posterior samples (where they should be smooth: e.g., if the prior implies truncations are very unlikely or almost certain). The best choice of scaling is problem-dependent.

The computational cost of a single simulation (for the case of

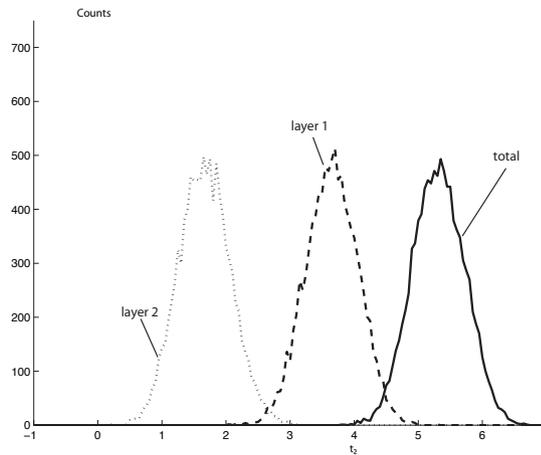
TABLE 2—PERFORMANCE SUMMARY FOR THE 3D EXAMPLE (ONE COMPLETE SIMULATION)*	
Process	Work in Seconds**
Kriging work	5.95
Toeplitz solver work	0.22
Total overhead all traces	6.17
Samples, 5000 per trace, all traces	299.20
Cost of example simulation, excluding io	305.37
* model size, $100 \times 100 \times 10$; 5,000 samples per trace	
** using a 2 GHz Pentium-M processor with 1 GB of RAM	

Fig. 8a) is examined component-by-component in **Table 2**. Several features are striking. First, 97.98 percent of the work is done in the deepest part of the sampling loop, which requires random number draws, extractions of submatrices, and multiplication of random normal vectors by lower triangular matrices (the Cholesky factor of the posterior covariance matrix, $\mathbf{L}_{C\pi\kappa}$). None of these operations is particularly expensive, but a total of 5×10^7 iterations were performed for this case ($\approx 164,000$ samples accepted per second). Because the kriging system is solved only once per trace—and is 2D, with an efficient k-d neighbor search—the associated work is small, about 1.95 percent. The Toeplitz manipulations are practically cost-free, only about 0.07 percent of the total work. Finally, the overall cost of about 5 minutes on a laptop computer (for 10^5 unknowns) does not seem prohibitive.

Because it is a tracewise sequential algorithm, this MCMC method scales linearly in the number of block edges, or traces. Thus, a model with 10^6 traces and 10 layers should require approximately 8.5 hrs if attempted on a single Pentium-M processor with adequate memory: not too alarming, for a model with 10^7 unknowns. The Toeplitz covariance and inversion work scales approximately with the third power of layer count (see the Numerical Considerations section), and linearly for generating samples at traces. However, Toeplitz solver work takes less than 1 percent of the computing time (Table 2). That is, although the cubic scaling is unfavorable for large K , the multiplier for the Toeplitz work is small and this component does not control the total work required. This is because proposing samples consumes most of the work, and each trace has thousands of proposals and requires only one K^3 Toeplitz solve. The total, sampling-dominated work scales with K rather than K^3 . Therefore, a model with 20 layers takes approximately twice as long as the 10-layer model used in the illustrations.



(a) Scattergram, $N = 8000$



(b) Histogram, bin size $\Delta h = 0.05$

Fig. 6—Simulation results for a two-layer case with prior sum less than the sum constraint. $\bar{H} = 6$, $\bar{\mathbf{t}} = (3, 1)^T$, $\sigma_H = 0.5$, and $\sigma_t = 0.5$; consistent units.

Discussion

Sequential Methods. A difficult aspect of these nonlinear downscaling problems is discerning whether the overall system posterior distribution can be safely factored into the product of conditional distributions implied by the sequential pass over the columns of gridblocks. This factorization requires computing both analytical marginal distributions (integrating over unvisited sites), and conditional distributions dependent only on visited sites. This requirement is usually met only by exponential family distribution functions. The posterior in our problem does not strictly satisfy these requirements. Nonetheless, the approximations we make can doubtless be improved by blockwise sequential schemes, though a block approach increases both the dimensionality of the MCMC sampling subproblem and the configurational complexity of handling more pinchout transitions.

Notwithstanding these concerns, we have demonstrated that using auxiliary variables greatly facilitates effective sampling of a complicated high-dimensional posterior distribution that arises in the downscaling problem we address. Similar difficulties will arise in any more or less rigorous recasting of the problem, so the technique we demonstrate should be widely applicable. Possible extensions are use of mixture-independence samplers (Gilks et al. 1996) that take advantage of the piecewise quadratic form of the log-posterior function, and generalization to multiple correlated variables in the model and associated likelihood.

Related Methods. As discussed in Simulation of Two-Layer Systems, if no layers are likely to be absent, the posterior distribution remains multi-Gaussian, and simulation and estimation methods are linear. In this case, the proposed method is a variant of collocated cokriging, where the collocated data are a sum rather than a constraint on a single thickness (Goovaerts 1997). The proposed methods are needed only when there is substantial likelihood of layers terminating laterally, in which case untruncated Gaussian models will fail.

Previous work on reservoir characterization with truncated Gaussian fields has focused on categorical simulations (Xu and Journel 1993; Matheron et al. 1987). In contrast, the proposed method combines aspects of categorical and continuous simulations. The condition $t_k \leq 0$ on the thickness proxy is equivalent to setting an indicator for layer occurrence to zero. However, in the categorical case all $t_k > 0$ would be identical (for a binary case), whereas we use values $t_k > 0$ to model the continuous variable h_k . This hybrid

approach could be applied without constraints, yielding sequential truncated Gaussian simulations of thickness; this corresponds closely to the cases with high σ_H presented above, and the resulting images would be similar.

Cornerpoint Grids. The MCMC simulation is over the block edges, or traces. This is different from many geostatistical modeling approaches, which are commonly block-centered. However, geometry—especially pinchouts or discontinuities at faults—can be modeled more accurately using cornerpoints. The porosity and other rock properties should be simulated or estimated at the same point, because these properties are generally correlated through the rock physics model and seismic response. Even for cornerpoint grids, reservoir simulators use block centered values for rock properties such as porosity. The trace properties must be averaged appropriately to the block center. A simple mean is probably adequate for thickness and porosity-thickness. However, the permeability must be upscaled more carefully, especially for nonrectangular blocks; a good method might be to integrate the Jacobian over the half-block domains (Peaceman 1996). Even for uniform permeability, the Jacobian integration correctly provides face- and direction-dependent transmissibilities for a nonrectangular grid. The method could also be used to perform approximate upscaling for sublayer heterogeneities and compute more accurate pore and bulk volumes.

Extensions. Three extensions are discussed in other, related work. First, several distinct facies are subjected to separate seismic thickness constraints (Kalla et al. 2007b). This permits, for example, conditioning on net and gross thickness separately. Second, product constraints, of the form $\sum_{k=1}^K h_k \phi_k = \bar{\Phi}H$, can be imposed; these constraints are nonlinear (Kalla et al. 2007b). More general scale linkages have been implemented using Markov random fields (Lee et al. 2002). Third, block methods or other approaches were considered by Kalla et al. (2007a) to address difficulties with the computation of marginal distributions in non-Gaussian sequential simulation (see the Sequential Methods subsection).

Conclusions

Stochastic seismic inversion computations can be integrated with a truncated Gaussian geostatistical model for layer thicknesses using an MCMC method. Truncation makes the problem nonlinear, which is ameliorated by the introduction of auxiliary variables and

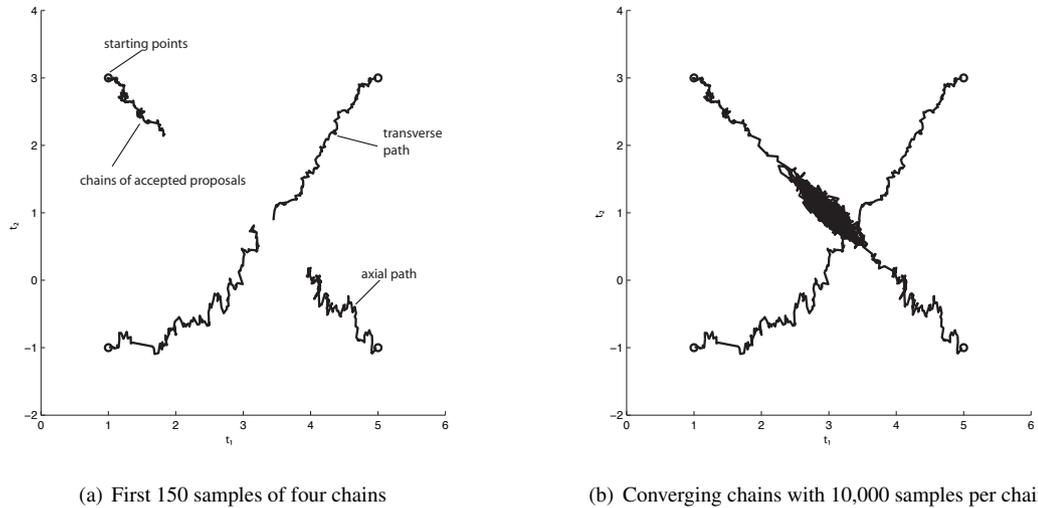


Fig. 7—Four Markov chains starting from diverse points tend to migrate toward the most likely region. (a) Convergence is slower for points that must move along the axis to reach the area of the mode. (b) Results are practically identical for long chains, because the posterior is unimodal. The prior and constraint data are the same as in Fig. 4.

a mixed Gibbs-Metropolis-Hastings sampling procedure. Under reasonable assumptions, the posterior resolution matrix is a special form of Toeplitz matrix; the special form can be exploited to make MCMC sample proposals more efficient to evaluate. Proposal efficiency is critical to the usefulness of the method, because many thousands of proposals must be evaluated at each trace for a single cornerpoint grid realization. The ability of the method to reproduce limiting case results and correctly model truncations is verified by examining algorithm behavior in two dimensions. A synthetic 3D case demonstrates that the procedure is acceptably fast. Although many issues remain—especially implementation of more complex constraints and integration with fine-scale geomodels—the proposed method appears to offer a foundation for further development.

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. Schlumberger provided access to Petrel reservoir modeling software, used for visualization in this research. Two anonymous reviewers provided insightful, detailed corrections and suggestions, which have improved the paper significantly.

Nomenclature

C_p	= prior covariance matrix based on kriging, m^2
C_π	= posterior covariance matrix, m^2
G	= posterior resolution matrix or Hessian, m^{-2}
d	= neighboring conditioning
h	= nonnegative layer thickness, m
H	= total thickness at trace, m
L	= Cholesky factor of covariance matrix, m
$N(\mu, \sigma^2)$	= normal distribution function with mean μ and variance σ^2
$N^{-1}(\mu, \sigma^2; r)$	= inverse normal distribution function with mean μ and variance σ^2 , at a cumulative probability of r
p	= probability density
P	= probability
r	= random number
R_x	= covariance range parameter in direction x , m

s	= scaling factor
t	= Gaussian proxy for h , may be negative, m
u	= auxiliary variable correlated to layer state
U	= uniform distribution function
T	= $T_k = \frac{1}{2}(\text{sgn}(t_k) + 1)$
W	= computational work, flops
x, y, z	= coordinates, m
X, Y, Z	= grid extents, m
α	= Metropolis-Hastings transition probability
γ	= semivariogram model
Δ	= separation vector for variogram models, m
ϕ	= layer porosity
$\bar{\Phi}$	= trace average porosity
κ	= number of layers at a trace with $t_k > 0$
π	= posterior
σ^2	= variance

Indices and Special Subscripts

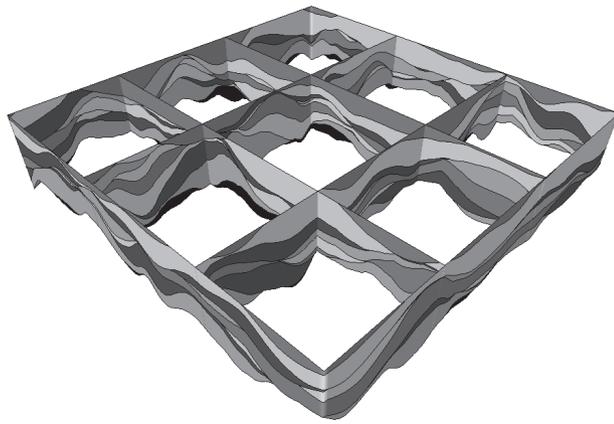
D	= number of nonzero conditioning data
k	= indices over layers
K	= total number of layers
ℓ	= indices over traces
L	= total number of traces
p	= prior
λ, Λ	= zero thickness data index and count

Diacritical Marks

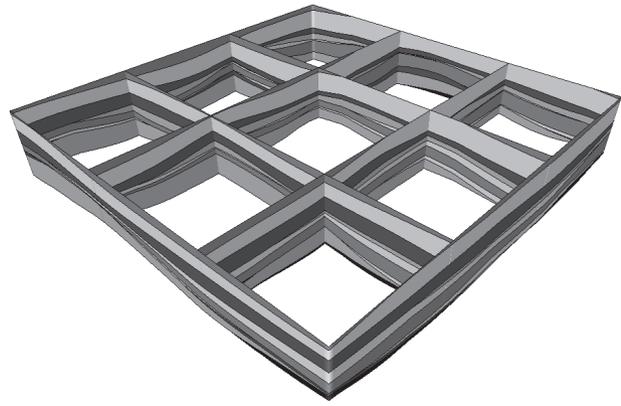
$\bar{\cdot}$	= mean
\cdot'	= proposed point, may become new point

References

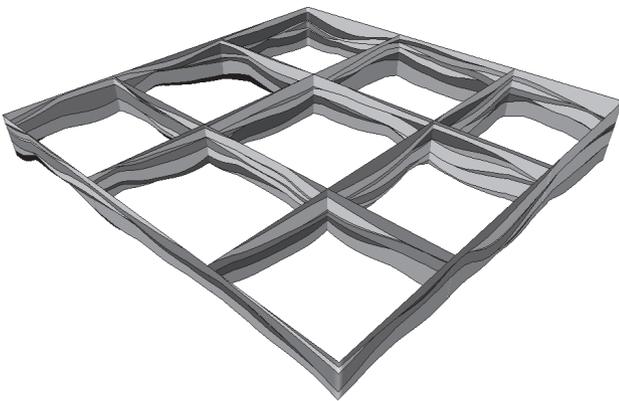
- Behrens, R.A., MacLeod, M.K., and Tran, T.T., 1998. Incorporating seismic attribute maps in 3d reservoir models. *SPEEE*, **1** (2): 122–126. SPE-36499-PA. DOI: 10.2118/36499-PA.
- Bentley, J.L., 1975. Multidimensional binary search trees used for associative searching. *Communications of the ACM*, **18** (9): 509–517.
- Deutsch, C.V. and Journel, A.G., 1998. *GLSIB: Geostatistical Software Library and User's Guide*. New York: Applied Geostatistics Series. Oxford University Press, Oxford, New



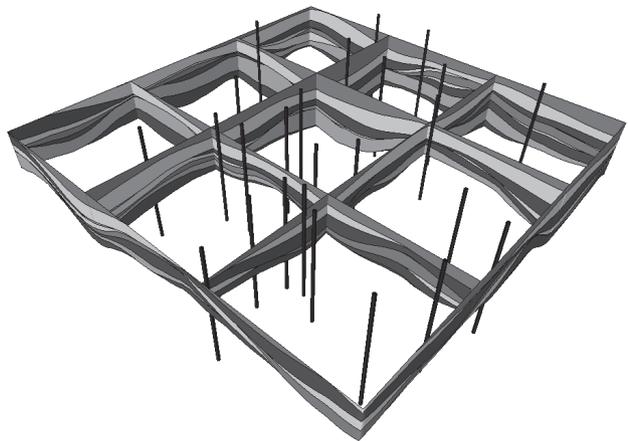
(a) Short range, $R = 200$



(b) Long range, $R = 750$



(c) Seismic thickness trend, $\bar{H} = 7 + \frac{13x}{x}$, $R = 350$; $x = 0$ is on the left front front



(d) Noise varies, $\sigma_H = 5 - \frac{3x}{x}$; R and \bar{H} as in (c); $x = 0$ is on the left front front

Fig. 8—Simulations on $100 \times 100 \times 10$ cornerpoint grids, areal extent is $X = Y = 1000$ m, and 25 conditioning traces are used. Unless otherwise noted, $\bar{H} = 20$ and $\sigma_H = 2$. All realizations use a Gaussian semivariogram with $R_x = R_y = R$, $\gamma(\Delta) = 1 - \exp[-(|\Delta|/R)^2]$, m^2 . All models flattened on the topmost surface. Range, thickness, and standard deviation are in m. $7.5 \times$ vertical exaggeration for all figures. Vertical black lines in (d) are conditioning traces.

York, second edition.

Deutsch, C.V., Srinivasan, S., and Mo, Y., 1996. Geostatistical reservoir modeling accounting for precision and scale of seismic data. Paper SPE 36497 presented at SPE Annual Technical Conference and Exhibition, Dallas, Texas, 6-9 October. DOI: 10.2118/36497-MS.

Dobrin, M.B. and Savit, C.H., 1988. *Introduction to Geophysical Prospecting*. McGraw-Hill, New York, fourth edition.

Doyen, P.M., Psaila, D.E., den Boer, L.D., and Jans, D., 1997. Reconciling data at seismic and well log scales in 3-d earth modeling. Paper SPE 38698 presented at SPE Annual Technical Conference and Exhibition, San Antonio, Texas, 5-8 October. DOI: 10.2118/38698-MS.

Eclipse 100 Reference Manual. 2004. Schlumberger Technology Co., Oxfordshire, UK.

Gelman, A., Carlin, J.B., Stern, H.S., and Rubin, D.B., 2003. *Bayesian Data Analysis*. CRC Press, Boca Raton, Florida, second edition.

Gilks, W.R., Richardson, S., and Spiegelhalter, D.J., 1996. *Markov Chain Monte Carlo in Practice*. Chapman and Hall, London.

Glinsky, M.E., Asher, B., Hill, R., Flynn, M., Stanley, M., Gunning, J., Thompson, T., et al., 2005. Integration of uncertain subsurface information into multiple reservoir simulation models. *The Leading Edge*, **24** (10): 990-999. DOI: 10.1190/1.2112372.

Golub, G.H. and van Loan, C.F., 1996. *Matrix Computations*. Johns Hopkins University Press, Baltimore, Maryland, third edition.

Goovaerts, P., 1997. *Geostatistics for Natural Resources Evaluation*. Applied Geostatistics Series. Oxford University Press, Oxford, New York.

Gunning, J.G. and Glinsky, M.E., 2004. Delivery: An open-source model-based Bayesian seismic inversion program. *Computers & Geosciences*, **30** (6): 619-636. DOI: 10.1016/j.cageo.2003.10.013.

Gunning, J.G. and Glinsky, M.E., 2006. Wavelet extractor: A Bayesian well-tie wavelet derivation program.

Computers & Geosciences, **32** (5): 681–695. DOI: 10.1016/j.cageo.2005.10.001.

Gunning, J.G., Glinsky, M.E., and White, C.D., 2007. Delivery messenger: A tool for propagating seismic inversion information into reservoir models. *Computers & Geosciences*, **33** (5): 630–648. DOI: 10.1016/j.cageo.2006.09.004.

Higdon, D.M., 1998. Auxiliary variable methods for Markov chain Monte Carlo with applications. *Journal of the American Statistical Association*, **93** (442): 585–595. DOI: 10.2307/2670110.

Kalla, S., White, C.D., and Gunning, J., 2007a. Downscaling seismic data to the meter scale: Sampling and marginalization. EAGE, Presented at Petroleum Geostatistics, Cascais, Portugal, 10-14 September.

Kalla, S., White, C.D., Gunning, J., and Glinsky, M.E., 2007b. Imposing multiple seismic inversion constraints on reservoir simulation models using block and sequential methods. Paper SPE 110771 presented at SPE Annual Technical Conference and Exhibition, Anaheim, California, 11-14 November. DOI: 10.2118/110771-MS.

Lee, S.H., Malallah, A., Datta-Gupta, A., and Higdon, D.M., 2002. Multiscale data integration using markov random fields. *SPEEE*, **5** (1): 68–78. SPE-76905-PA. DOI: 10.2118/76905-PA.

Matheron, G., Beucher, H., de Fouquet, C., and Galli, A., 1987. Conditional simulation of the geometry of fluvio-deltaic reservoirs. Paper SPE 16753 presented at SPE Annual Technical Conference and Exhibition, Dallas, Texas, 27-30 September. DOI: 10.2118/16753-MS.

Merriam, D.F. and Davis, J.C. (eds.), 2001. *SedSim in Hydrocarbon Exploration*, 71–97. Kluwer Academic+Plenum Publishers, New York.

Nocedal, J. and Wright, S.J., 1999. *Numerical Optimization*. Springer Science+Business Media, New York.

Peaceman, D.W., 1996. Calculation of transmissibilities of grid-blocks defined by arbitrary cornerpoint geometry. Richardson, Texas. Paper SPE 37306 available from SPE.

Petrel Workflow Tools, Introduction Course. 2005. Schlumberger Technology Co., Oslo, Norway.

Ponting, D.K., 1989. Corner point grid geometry in reservoir simulation. *Proc.*, Joint IMA/SPE European Conference on the Mathematics of Oil Recovery, Cambridge, UK.

Pyrz, M.J., 2004. *The Integration of Geologic Information into Geostatistical Models*. Ph.D. thesis, University of Alberta, Edmonton, Alberta, Canada.

Widess, M.B., 1973. How thin is a thin bed? *Geophysics*, **38** (6): 1176–1180. DOI: 10.1190/1.1440403.

Willis, B.J. and White, C.D., 2000. Quantitative outcrop data for flow simulation. *Journal of Sedimentary Research*, **70** (4): 788–802. DOI: 10.1306/2DC40938-0E47-11D7-8643000102C1865D.

Xu, W. and Journel, A.G., 1993. GTSIM: Gaussian truncated simulations of reservoir units in a W. Texas carbonate field. Richardson, Texas. Paper SPE 27412 available from SPE.

Appendix—Zero Thickness Conditioning Data

In this paper, the untruncated Gaussian proxy \mathbf{t} is kriged, not the actual thickness \mathbf{h} . At simulated traces, \mathbf{t} is computed and stored, and only converted to \mathbf{h} for output. Conditioning data present more of a challenge. If we observe some layer k on trace ℓ has $h_{\ell k} = 0$, the value of $t_{\ell k}$ is indeterminate; we only know $t_{\ell k} \leq 0$. The conditioning data might be decorrelated if we used a simple but reasonable draw such as

$$t_k = N^{-1} \left(\bar{t}_k, \sigma_{t_k}^2; r \right), r \sim U[0, P(h_k = 0)], \dots \quad (\text{A1})$$

where $P(h_k = 0)$ is given by Eq. 1, N is normal distribution function, and U is the uniform distribution function. Instead, we model the correlation as follows, with a loop over all layers.

- Find all zero conditioning data in this layer, k ; the list of the locations of zero data is indexed over $\lambda_k \in \{0 \dots \Lambda_k\}$. The positive conditioning data in layer k are indexed by $d \in \{0 \dots D_k\}$.
- Initialize all Λ_k zero thickness observations in layer k with random draws, using Eq. A1.
- Visit each point λ , forming a kriging system of size $D_k + \Lambda_k - 1$, composed of all points in this layer except the current point. Compute the mean and variance, and draw $r \sim U[0, P(h_k = 0)]$; in the first iteration, the kriging weights and variances are stored for reuse. $P(h_k = 0)$ is computed using the new mean and standard deviation of t_k . The new simulated value t_k is the inverse of $N(\bar{t}_k, \sigma_{t_k}^2)$ at cumulative probability r .
- Generate a chain and store.
- Repeat $\forall k \in \{1 \dots K\}$

The stored chains can be used at the beginning in later simulations of layer thickness. Before simulating any new points, sets of the zero-thickness conditioning data are drawn from the stored chain.

Subhash Kalla is a PhD student in petroleum engineering at Louisiana State University. email: skalla2@lsu.edu. He holds an MS degree in petroleum engineering from Louisiana State University. and a BTech degree in chemical engineering from the Regional Engineering College (now National Institute of Technology), Warangal, India. His research interests include reservoir simulation, reservoir characterization, and uncertainty analysis. **Christopher D. White** is an associate professor of petroleum engineering at Louisiana State University. He holds a BS degree from the University of Oklahoma and MS and PhD degrees from Stanford University, all in petroleum engineering. email: cdwhite@lsu.edu. Formerly, White was a research engineer for Shell Development Company and a research scientist for the Bureau of Economic Geology, the University of Texas at Austin. His research interests include general reservoir engineering, reservoir simulation, and statistics. **James Gunning** is senior research scientist in the CSIRO petroleum reservoir characterisation area. He holds a BS degree in physics, a BEE degree in electrical engineering, and a PhD degree in applied mathematics, all from Melbourne University. His current research interests are in spatial statistics, Bayesian methodologies, uncertainty quantification, data integration, and inverse problems from remote sensing tools like seismic and electromagnetic methods. **Michael E. Glinsky** is the Section Leader of Quantitative Interpretation for BHP Billiton Petroleum. He holds a BS degree in physics from Case Western Reserve University and a PhD degree in physics from the University of California, San Diego. He worked as a Department of Energy Distinguished Postdoctoral Research Fellow at Lawrence Livermore National Laboratory for five years. He worked for three years at the Shell E&P Technology Co. researching on Bayesian AVO and 4D inversion. He has published more than 25 papers in the scientific literature on subjects as varied as plasma physics, signal processing for oil exploration, x-ray diagnostics, application of exterior calculus to theoretical mechanics, and laser biological tissue interactions.