2008

Reservoir characterization using seismic inversion data

Subhash Kalla
Louisiana State University and Agricultural and Mechanical College, skalla2@lsu.edu

RESERVOIR CHARACTERIZATION USING SEISMIC INVERSION DATA

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Craft & Hawkins Department of Petroleum Engineering

by
Subhash Kalla
B.Tech. in Chemical Engineering, REC Warangal, 2002
M.S. in Petroleum Engineering, Louisiana State University, 2005
December 2008

Acknowledgments

This dissertation would not be possible without several contributions. It is a pleasure to thank my adviser Dr. Christopher D. White. During last six years he encouraged me to be an independent researcher, gave me freedom in choosing research direction, and offered stimulating suggestions. It is my pleasure to work with Dr. James Gunning of CSIRO. I want to thank him for his thoughtful insights and help in my understanding of Bayesian problems. I am also grateful to work with Dr. Mike Glinsky of BHP-Billiton Petroleum Quantitative Interpretation and want to thank him for supporting and guiding my research. Every discussion with these three is valuable and educational. This work was motivated by a perceived need to integrate Geoscience and Geostatistics and funded by an unrestricted gift from BHP-Billiton.

I want to thank my committee members Dr. Stephen Sears, Dr. Juan Lorenzo, and Dr. Seung Kam for providing me with valuable suggestions. Special thanks to all the Petroleum Engineering faculty for valuable discussions.

I extend many thanks to my friends especially Hong Tang, Xin Li, Shengkai Duan, and Charisma Lattao and colleagues especially Mike Shook of Chevron, Ricard Sech of Exxon-Mobil and Bill Cleary of BHP-Billiton for their valuable guidance.

Finally, I'd like to thank my family. My father and mother have been constant sources of support and encouragement. My brother, even though younger, switched roles and advised me critically. My wife enchanted me with her care and support.

This dissertation is dedicated to my parents.

Table of Contents

List of Tables

List of Figures

Abstract

Reservoir architecture may be inferred from analogs and geologic concepts, seismic surveys, and well data. Stochastically inverted seismic data are uninformative about meter-scale features, but aid downscaling by constraining coarse-scale interval properties such as total thickness and average porosity. Well data reveal detailed facies and vertical trends (and may indicate lateral trends), but cannot specify intrawell stratal geometry. Consistent geomodels can be generated for flow simulation by systematically considering the precision and density of different data. Because seismic inversion, conceptual stacking, and lateral variability of the facies are uncertain, stochastic ensembles of geomodels are needed to capture variability.

In this research, geomodels integrate stochastic seismic inversions. At each trace, constraints represent means and variances for the inexact constraint algorithms, or can be posed as exact constraints. These models also include stratigraphy (a stacking framework from prior geomodels), well data (core and wireline logs to constrain meter-scale structure at the wells), and geostatistics (for correlated variability). These elements are combined in a Bayesian framework.

This geomodeling process creates prior models with plausible bedding geometries and facies successions. These prior models of stacking are updated, using well and seismic data to generate the posterior model. Markov Chain Monte Carlo methods sample the posteriors. Plausible subseismic features are introduced into flow models, whilst avoiding overtuning to seismic data or conceptual geologic models. Fully integrated cornerpoint flow models are created, and methods for screening and simulation studies are discussed. The updating constraints on total thickness and average porosity need not be from a seismic survey: any spatially dense estimates of these properties may be used.

Chapter 1
Introduction

A review of geomodeling, sedimentology, seismic data, and seismic inversion is given in this chapter. These are the fundamental concepts, ideas, and the context on which this research is built. The details in the upcoming chapters discuss some frequently used geomodeling concepts. Discussion about geomodels and their role in petroleum engineering is given in section 1.1. Geologic and stratigraphic data central to the algorithms in this dissertation, as well as their value and limitations are discussed in section 1.2. Sections 1.3 and 1.4 discuss scale and resolution issues of seismic data and seismic inversion, respectively. Seismic inversion data is used in this research. Therefore, it is fundamental in understanding the resolution of seismic interpretation and the role of seismic in modeling structural and rock property variations in geologic models. Problems in using seismic derived properties in building geomodels and the current approach in solving some of the issues are discussed in section 1.5. Objectives and significance of the research are discussed in section 1.6 and section 1.7.The approach taken to solve the problems and the outline of the thesis are discussed in 1.8 and 1.9.

1.1 Geomodeling

Geomodels are geometric and petrophysical representations of oil and gas fields, used to predict the flow of fluids through a porous rock. Realistic geomodels are needed to predict the effect of heterogeneities in the reservoir under various recovery scenarios. They are the link between geologic concepts that define structure of deposition and the properties used in mathematical flow simulations. They integrate diverse information from different sources and ultimately represent 3D reservoirs in 3D. Since flow modeling has become more important for asset management, the importance of geomodeling has increased.

1.1.1 Challenges in Geomodeling

Geomodels should be geologically realistic, but must balance geologic detail versus computational efficiency in reservoir simulation. The more detail in a geologic model, the more grid cells are required to represent the complex geology, and consequently, time to simulate fluid flow in the geologic model increases. It is known that not all the details in the deposition influence the fluid flow between injectors and producers often separated by 2 km (Willis and White 2000). Simple geomodels with incorporated key heterogeneities are the best compromise models to understand the fluid flow in the reservoir. As a result, one of the main challenges in building geomodels is to understand what scale and how significant the various heterogeneities which are present in the depositional environment will influence the recovery mechanism. In this study, models are built at the scale of ~ 1 m vertical resolution, which may be close to the "right" scale to simulate and predict. That is, we construct the model at the flow scale, rather than the more common practice of constructing over-resolved geomodels and upscaling. The "right scale" for downscaling is where fluid flow is not sensitive to heterogeneity below that resolution. Even if significantly influencing heterogeneities are well understood, 3D modeling of geomodels is complex, because it needs to incorporate such diverse (and challenging) geologic features as stratigraphic stacking (section 1.2, chapter 5) and diverse data, such as seismic inversion (section 1.4). The phase of model building that ensures the models honor the available data is called data conditioning.

1.1.2 Geomodel Construction

Methods to construct geologic models depend upon geologic variability, and this in turn depends on the sedimentary process that creates the reservoir. In reservoirs with low variability like sheet sands in distal turbidities, layers and their properties laterally may vary little. In these reservoirs, deterministic methods based on interpolation can be used to build reservoir architecture and model rock properties. Common deterministic methods are based on either kriging or weighted inverse distance. In inverse distance methods, weights of the hard data

points for interpolation are assigned based on the inverse of distance criterion; the larger the distance from estimated point, the less weight the datum will have. Kriging uses the concepts of random variables, weak stationarity (higher order moments do not change when shifted in time or space) and autocorrelation (correlation of variable with itself against a time or space shift) to predict values of a variable distributed in space. Performance of these methods depends on the availability of conditioning data like well control, and may represent geology well when the properties are not varying over distances less than the data spacing. Uses of these methods are limited as deterministic methods generate artificially smooth property variations. Such variations are uncharacteristic of actual variation in many geosystems. For instance, permeability may vary at a much smaller scale than well spacing.

The counterpart of deterministic processes is the stochastic process. Rather than estimating a single possible outcome of the reservoir model, there is some variability in a stochastic process. Even if all the conditioning data are the same, there are many possibilities on how the model will be realized (although some of the models are more probable than others). The most common stochastic processes are as follows:

- **Gaussian Simulation:** This method is used to simulate a random field extended by using multi-Gaussian (or multinormal) distributions. The models generated are conditioned to well constraints (hard data) and use the variogram model(soft data), which is a autocorrelation function that characterizes the spatial variability of the Gaussian variable. Several algorithms like simulated annealing, sequential Gaussian simulation (SGS), and LU decomposition are based on Gaussian simulation. In this research, a variation of SGS is used as explained briefly below. The SGS proceeds sequentially and predicts a random variable at each grid cell; the SGS follows a preset path, eventually visiting all the grid nodes. The steps are shown below:

    1. Define a random path through all of the nodes to be simulated.

2. At a node in the random sequence, use simple or ordinary kriging to estimate the mean and variance of the random variable, which is assumed to be Gaussian. Both simulated and original data are used to condition the estimate.

3. Draw a random value $P \in [0, 1]$, then use the mean μ and variance σ^2 of kriging estimate and assign that value to the node being simulated, assuming the attribute is distributed as $N(\mu, \sigma)$; the simulated value is the deviate that has a cumulative probability P.

4. Visiting nodes sequentially generate the kriging estimates and assign values at all simulation nodes.

- **Multipoint Geostatistics:** This method also assumes some form of weak-stationarity and simulates the random variable being modeled sequentially. Multipoint geostatistics (MPG) also integrate hard and soft data, but the soft data is a training image rather than a variogram in kriging-based methods. A training image represents a conceptual image of the sedimentary heterogeneity to be reproduced in the simulated model. The training image is used as a template to find conditional probability distributions of a given pixel pattern. In the sequential simulation, probability of the pixels belonging to each pattern is calculated from conditional functions, and the pattern is assigned to the pixel proportional to the probability (Strebelle 2002).

- **Object-Based Method:** This method was first used in geologic modeling to simulate shales which are not correlated between wells (Haldorsen and Lake 1984). Later, this method was used to model reservoirs with distinct geometries, such as fluvial channel systems. This is because geometry of the channel complexes, channels, or overbank deposits can be easily parameterized as templates characterized by simple shapes (Chessa 1995). A type of object-based model called a surface-based model is useful in simulating stratigraphic surfaces. The surfaces are based on predetermined templates for a geologic setting. The surface-based or object-based models place these

templates stochastically in the reservoir, while matching the observations and models. When there are many conditioning data, it is hard to condition surface or object based models (Pyrcz, Catuneanu, and Deutsch 2005).

1.2 Geology Basics

1.2.1 Data Types and Scales

Reservoir characterization is performed by analyzing different geologic data. Understanding the source of data and its limits is essential in any data integration techniques. In this research, geologic data coming from various sources are to be used depending on both scale and resolution. Integrating all these data in building geomodels is crucial as incorporating all this information increases confidence in the resulting geomodels and prediction performance. However, current methods are not able to include these diverse and multiscale data. Geologic data for modeling comes from the sources shown below.

- **Core:** Core samples are useful for calibrating well log estimates. They are also useful for examining small scale heterogeneities, thereby inferring depositional environment. Even though the resolution of core is high, the coverage of the reservoir is small. Typical core diameters are $\sim$ 10 cm and lengths are $\sim$ 10 m.

- **Well Logs:** Well logs measure rock properties Such as resistivity, sonic velocity and radioactivity; models and correlations allow estimation of reservoir properties like porosity and water saturation. Well logs are also useful to understand stratigraphy and layering of reservoir if enough wells are available. Vertical resolution is $\sim$ 10 cm, a resolution commonly available throughout the well (or at least the productive section). Still, the areal coverage is low and the volume of reservoir sampled by wells is low. For the algorithms described in this study, core and well log data must be upscaled from decimeter scale to meter scale. Because the main properties modeled in this dis-

sertation are porosity and thickness, upscaling from core- or log-scale to the flow- or meter-scale is both simple and accurate.

- **Seismic:** Seismic gives information about the larger scale heterogeneity and may help estimate reservoir properties through inversion, but the resolution is low $\sim$ 10 m (Liner 2004).

- **Outcrop:** Outcrops are exposures of sedimentary bedrock at the earth's surface. Outcrops allow direct observation of deposition environment features such as bedding orientation, paleo-current directions, and lithofacies architecture. An outcrop will provide analog information with excellent vertical resolution, but showing only 2D cross-sections. Model data used in this work for example variogram and stacking patterns, could come from outcrop data.

1.2.2 Clastics Sedimentology Basics

Distribution, continuity, and internal characteristics of sandstones depend on the depositional environment. Sedimentary structures of sand bodies are similar if the depositional environments match closely. In this section, sedimentary deposits in major environments are described, mostly based on the research of R. J. Le Blanc (1976).

1. **Alluvial(Fluvial):** This environment is characterized by deposits of sediment, transferred by fluvial systems across continental regions towards delta mouths or subcanyon regions. This system is mainly divided into two main sub-environments

    - **Alluvial Fan and Braided Stream:** Sands of this origin are derived from mountainous erosion and are transported to alluvila fans through canyons, finally deposited on braided systems.

    - **Meandering Stream:** Sedimentation occurs in this system because of channel migration and abandonment. When a channel migrates along a caving-bank area, it deposits a point bar on the other side of the channel.

2. **Aeolian:** These are sand dunes deposited by wind in arid areas with low vegetation in coastal and desert areas.

3. **Deltaic:** These are shore sediments lying between the flood-tide and ebb-tide and are generally divided into two main sub-environments.

    - **Distributary Channel:** Distributary channels of the deltaic plain bring sand deposits to the delta front.
    - **Delta Fringe Sands:** These are the sands deposited in front of the river mouth as bars. Once sands are brought in by channel, they are dispersed by the marine processes of waves, tides, and currents.

4. **Coastal Plain Sands:** These are present on coastal plains, such as barrier islands and tidal channel sandstones. Barrier-island sands are long and narrow belts, found parallel to the shoreline. The thickness depends on the depth of inner continental shelf waters.

5. **Shallow Marine:** In transgressive conditions, shallow marine sediments are formed as the sea advances toward land. Transgressions can occur with a rising sea level, a subsiding coastal plain, or an abandoned delta.

6. **Deep Marine:** Sediments from the continental slope (where deltas and carbonate reefs are present) reach the ocean floor by passing through canyons as debris flows and turbidity currents. Ultimately, sediments are deposited in the deep marine.

The shape and size of river mouth bars in deltaic reservoir depend on the rate of sediment deposition, subsidence, flood cycles, and reworking processes by the sea. Variability in geometry and properties between sand bodies also occur under similar depositional circumstances. Geomodeling must incorporate the different shapes and sizes of geobodies (variogram range and sill) to determine how they assemble together. More discussion on modeling clastics is given in 1.2.4.

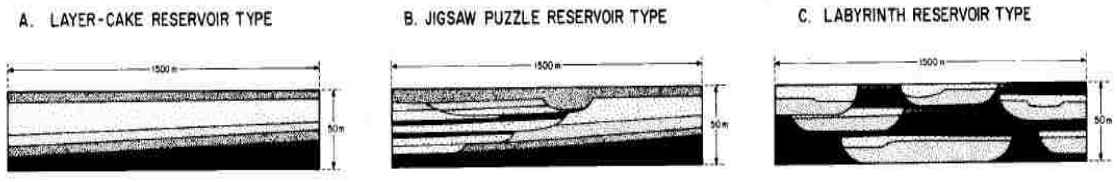


Figure 1.1: Three major reservoir types for geologic modeling from Weber and van Geuns (1990).

1.2.3 Stratigraphy Basics

Stratigraphy is the study of layers. Sequence stratigraphy relates the layering of rocks to sea-level changes by studying, the interaction of sedimentation with respect to sea level changes, sediment supply, and subsidence. Stratigraphy uses the principles of original horizontality, lateral continuity, superposition, cross-cutting, and faunal successions to determine the relation of a layer to the other layers above and below within a geologic system (Embry 2002). In this study, stacking patterns, an arrangement of vertical facies successions of sedimentary body surfaces, is used as one of the inputs for data integration. This is a major element in stratigraphic interpretation of the depositional settings. For example, stacking patterns for unconfined-sheet sands can prograde (move seaward), retrograde (move landward), or aggrade (move vertically).

1.2.4 Modeling Clastic Reservoirs

Depending on variability of properties like permeability (vertically and laterally) and discontinuous sands of variable thickness, Weber and van Geuns (1990) divided most of the reservoirs with different depositional settings into three major divisions for modeling purposes (Fig. 1.1).

- **Layer cake reservoirs** are reservoirs where properties change little, either laterally or vertically. Few wells are required to correlate the layers between the wells, as sandbodies

	Continental	Transitional	Marine
Layer cake	Sheet flood deposits Lacutrine sheet sand Aeolian dunes	Barrier bars Chenier deposits Transgressive Sands	Shallow marine sheet sands Offshore bars Outer-fan turbidites
Jigsaw Puzzle	Braided river deposits Point bars Mixed lacustrine/fluvial Mixed aeolian/wadi	Combined facies complexes, e.g., barrier bar plus tidal channel-fill channel-fill/mouth bar combinations with high NGR	Storm sand lenses Mid-fan turbidites
Labyrinth	Fluvioglacial deposits with low NGR Low-sinuosity channel-fills	Low-sinuosity distributary channel-fills	Upper fan turbidites Slumps Storm deposits with low NGR

Figure 1.2: Classification of clastics depositional environments into three reservoir types from Weber and van Geuns (1990).

laterally are very extensive. These reservoirs are modeled as a package of superimposed extensive layers with no sudden thickness variations.

- **Jigsaw-puzzle reservoirs** are sand body objects that stack together without much background low permeable facies. Layer properties significantly change vertically and laterally due to layer truncations and low permeability bodies which are embedded between the sand bodies. Several wells in a square kilometer are needed to characterize these types of reservoirs. Usually, stochastic methods are implemented to model such reservoirs.

- **Labyrinth reservoirs** are rarely correlated between wells with realistic well spacing. These reservoirs are characterized by sand bodies or lenses in a low permeable rock background. The sand bodies are connected by low permeable thin sand stone sheets.

Even though there are a number of depositional settings with different continuity and connectivity characteristics (Clastics Sedimentology Basics, earlier), understanding the closest reservoir type for the depositional setting provides an indication for a geomodel construction method. For layer-cake models, deterministic methods may be adequate, as there are very minor lateral variations. For labyrinth reservoirs, object-based models are more appropriate.

The objects with particular shape and size are placed under a low-permeability background. However a number of realizations are required to understand the connectivity between the wells.

The proposed surface-based approaches (chapters 2 and 3) are especially appropriate for jigsaw-puzzle reservoirs where the layers pinchout and geobodies are stacked upon one another. To mimick geologic pinchouts in reservoir modeling using kriging based methods creates a challenge. Current kriging-based models that generate models with pinchout behavior and conditioning to seismic and well data prove even more difficult. However, stacking patterns provide rich, geologic information. Although integrating this information is challenging using kriging-based methods, yet this becomes feasible, using a surface-based approach with stratigraphic information. Surface-based models can be used to generate realistic-looking stacking patterns, but they are hard to condition to seismic and well data. Despite inherent difficulties, an ensemble of surface–based models can be used to infer probabilistic descriptions of stacking patterns, and thereby incorporate this stratigraphic information into stochastic geomodels (chapter 5).

1.3 Seismic Data

Even in the earliest forms, seismic surveys reduced risk in exploring for oil and gas. The advent of 3D seismic and better seismic processing techniques has further reduced risk in exploration phase and improved characterization of the reservoirs. But generally, reservoir characterization using seismic data alone cannot resolve spatial heterogeneities and stratigraphy at the scale desired for flow simulation. Using seismic data, it is difficult to identify beds thinner than 10 m from 3-D seismic signals when the reservoir depth is greater than 3,000 m (Dobrin and Savit 1988; Widess 1973); $\lambda/4$ could be resolved where λ is the wavelength of seismic. The limit and errors associated with seismic estimates, therefore, allocate seismic data either to inference of the external architecture or guiding the probable internal stratigraphic architecture of a reservoir.

Nonetheless, seismic data acquisition and processing provide excellent lateral coverage of the reservoir, with spacing often as dense as 12.5 m. The lack of vertical resolution of the seismic data can be partly offset by using model parameters (for example, layer thickness, porosity and permeability)derived from core and well-logs; well data have high vertical resolutions.

Principles of seismology are used to estimate the properties of the subsurface by sending controlled seismic energy into earth and observing reflected seismic waves at receivers. The seismic energy source can be dynamite, air gun, or vibrators. Seismic (acoustic) waves are a form of elastic wave that propagates through the subsurface. The wave propagation depends on the property of the medium called seismic impedance I which is the product of wave velocity and density of the rock ($I = \rho v$). If a seismic wave passes through the boundary of two different rocks with varying impedances, a portion of the seismic wave energy reflects, while some transmits through the boundary. The amplitude of the reflected wave called reflection coefficient (R) depends on the impedance contrast between the two layers that created the wave.

$$R = (I_2 - I_1)/(I_2 + I_1) \tag{1.1}$$

where $I_i = \rho v$. From seismic study we have travel time, which is the time it takes for a reflection to arrive at the receiver from a boundary. If the seismic velocity in the rock is known, the travel time may be used to estimate the depth to the geologic boundary that generated the reflection. For a wave traveling vertically, the travel time (t) from the source to the reflector and back is provided by the formula $t = 2d/v$, where d is the depth of the reflector and v is the wave velocity in the rock. (This discussion greatly simplifies the processing step, omitting important issues such as migration and stacking. That level of detail is not considered essential in this context.)

Travel time and amplitude of the reflected waves are used to interpret the subsurface. A brief explanation of interpreting seismic data utilizing the reflectors is given below.

1.3.1 Synthetic Seismic at Wells

Both a $v(z)$ velocity model and a $\rho(z)$ density model, which are a function of depth are needed to create a synthetic seismic at wells. The product of these two models offers an acoustic impedance model of the subsurface. Because a vertical or nearly vertical well is similar to a seismic trace, an impedance model may be estimated for nearby seismic traces. We can compare this with the amplitude data from the seismic survey for quality control, attribute correlation, and wavelet extraction. The impedance model may be obtained from the sonic logs and density (or neutron) logs.

Sonic logs record interval transit time, that is, estimates of the reflection time between layers which is shown as estimated as thickness of layers (from logs) divided by velocity.

$$\Delta t = \frac{z_2 - z_1}{\bar{v}} \tag{1.2}$$

Velocity of a layer also is obtained from this equation because sonic gives the transit time (Δt) and most logs give thickness of a layer ($z_2 - z_1$). Using the density and velocity models, reflection coefficients are estimated at the rock boundaries.

R at each interface as a function of travel time (from sonic logs) is called the reflection coefficient (RC) series. A synthetic is created by convolving the RC series with a wavelet as shown in Fig. 1.3. A wavelet is a wave pulse approximation for a seismic source (generated by an air gun, dynamite, or other sources) which contains many frequencies and is time-limited. If the wavelet is a good approximation for the wave produced by the actual source, comparing the reflectors between the synthetic and actual seismic data gives the geological horizons that cause seismic reflections at the well (or trace). A series of related reflections on several vertical traces is referred to as a reflection event. By correlating reflection events across the wells, an estimated cross-section of the geologic structure can be interpreted (Liner 2004).

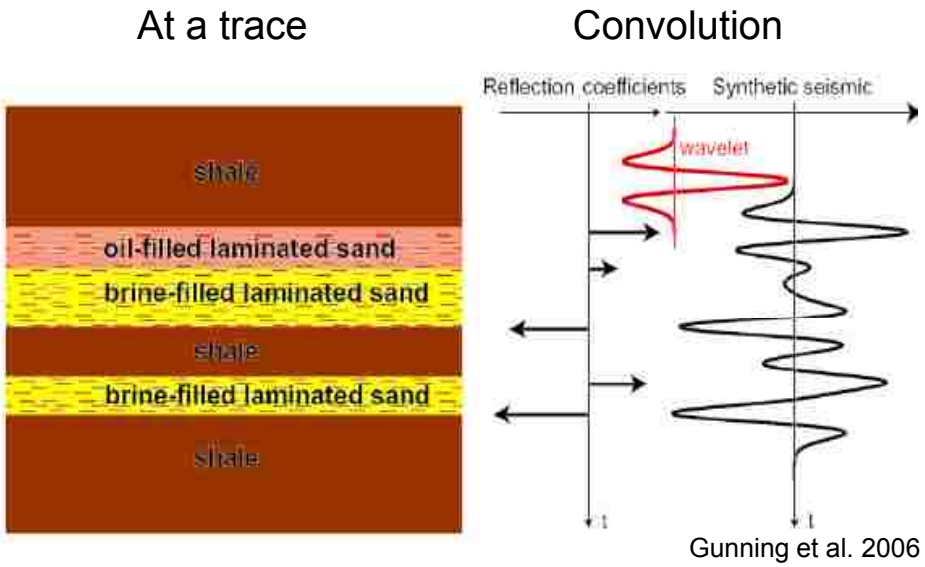


Figure 1.3: Generating synthetic at a trace by starting with a earth model and convolving with wavelet.

1.3.2 Structure Interpretation

Once seismic reflection events have been correlated with geological horizons using synthetics at wells, structural interpretation proceeds in three steps as follows:

1. **Fault detection**: Near-vertical faulting can be seen on seismic sections (inline and crossline sections) and low-dip faults are best seen in horizontal sections (time slices).

2. **Horizon tracking**: Tracking of amplitude–related quantities (also known as attributes) such as peak amplitude or RMS amplitude can be used to pick horizons.

3. **Time to depth conversion**: Depth conversion is carried out by combining the seismic, well control, and velocity model to create a structure map in depth rather than time. For linear depth-velocity models (Liner 2004), the cross plot between the average velocity $\tilde{v}$ and depth z (from logs) at wells gives the velocity gradient k and surface velocity v_0 at those wells (related as $\tilde{v} = v_0 + kz$), z being vertical depth. Average velocity is computed by using log picks and seismic time picks from sonic logs (Eq. 1.2). A map is created by contouring the v_0 values at each well. These surface velocities at

every location, velocity gradients, and observed seismic time picks, gives the depth of a horizon throughout the reservoir.

1.3.3 Stratigraphic Interpretation

Stratigraphic understanding has improved since the advent of 3D seismic. Stratigraphic features tend to be subtle and are usually indicated by variations in amplitude, phases, and terminations (Liner 2004). These features are best seen in horizontal time slices. Such studies improve understanding of the depositional environment and possible recognition of features such as channel complexes, which may constrain the geologic model further.

1.3.4 Reservoir Properties Estimation

Seismic attributes are used to estimate reservoir properties. There are many ways to extract seismic attributes which are secondary quantities from seismic amplitude, *e.g.*, normalized amplitude. If these attributes are correlated to reservoir properties at wells they are used to predict reservoir properties between wells.

1.4 Seismic Inversion

Given a set of reflectors and the physical laws of wave propagation, we can develop a subsurface model structure and its physical properties. These types of problems are generally called inverse problems; when applied to reflection seismology, it is called seismic inversion. Seismic inversion is used to transform a noisy, seismic trace into density and sonic logs, the inverse of transforming these two logs into a synthetic (Synthetic Seismic at Wells, earlier) (Latimer 2006).

Seismic inversion helps remove peculiarities of wavelets and then estimates reservoir properties with a better resolution. Inversion may be considered in several ways.

- Acoustic seismic amplitude is inverted to generate physical rock properties like impedance. Using rock physics, it is possible to generate reservoir parameters that are directly used

in flow simulation like porosity, layer depths, and fluid saturations (Gunning and Glinsky 2004).

- Inversion provides higher resolution images, because it removes wavelet effects such as tuning *via* deconvolution.

- Acoustic impedance inversion requires incorporation of well log data. So, inversion is a data integration step, and the resulting data tie to the wells and also approximately match seismic data.

Compared to seismic amplitudes, inversion gives higher resolution stratigraphic images and better estimates of reservoir properties, because it removes tuning and wavelet effects.

1.4.1 Rock Physics Analysis

In a seismic survey, a compressional wave (P-wave) is generated using a source. When a seismic wave propagates through elastic boundaries within the earth, both P-waves and S-waves are generated. Reflected waves are recorded at different offsets. Then the common mid-point (CMP) gather uses a group of traces whose source-receiver midpoint lies in a small region of the Earth's surface (Fig. 1.4). Basically, inversion creates impedance values at all CMP bins. So, CMP gathers contain information about P-wave velocity, S-wave velocity, and density (Mavko, Mukerji, and Dvorkin 2003).

Rock properties like the bulk modulus (k), shear modulus (μ), porosity, fluid type, lithology, clay content, and gas saturation affect seismic wave propagation through rocks. This is because any elastic medium is characterized by velocity and density (Seismic Data, earlier), and they are dependent on the rock properties through elastic constants. Rocks are elastic and elastic theory provides the relation between velocity (P and S waves) and density to elastic constants (shear and bulk modulus that are functions of rock properties) under

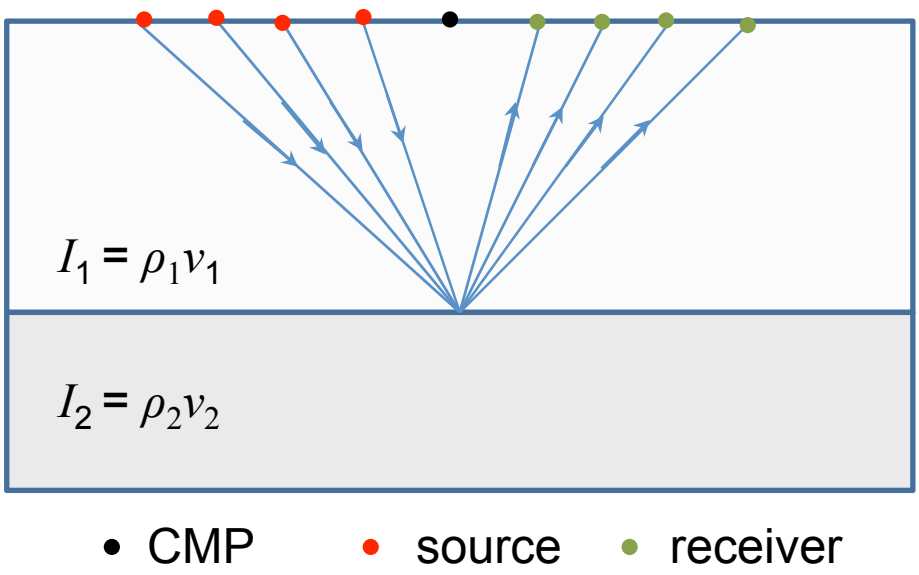


Figure 1.4: Location of CMP for a set of receivers and sources.

certain assumptions. For isotropic media (Liner 2004),

$$\text{Compressional wave velocity}(V_p) = \left(\frac{\lambda + 2\mu}{\rho}\right)^{1/2}$$

$$\text{Shear wave velocity}(V_s) = \left(\frac{\mu}{\rho}\right)^{1/2}$$

$$\text{where Lame's parameter}(\lambda) = k - 2\mu/3$$

Further, theory gives the relation between porosity, fluid, and other properties of reservoir rock to seismic propagation in reservoirs. For example, the Gassmann (1951) equation gives the effect of pore fluid on k. It calculates the bulk modulus of the rock with fluid using the bulk moduli of the solid frame, the matrix, and the pore fluid (Wang 2001):

$$k^* = k_d + \frac{\left(1 - \frac{k_d}{k_m}\right)^2}{\frac{\phi}{k_f} + \frac{1-\phi}{k_m} - \frac{k_d}{k_m^2}}$$

where k^* is the bulk modulus of rock which depends on k_f, k_d, and k_m, called the bulk moduli of fluid, frame, and matrix respectively. There are many assumptions in deriving the Gassmann equation; in particular, there is an effective medium of macroscopic isotropy. In short, the wavelength is assumed long, compared to the grain and pore sizes. Similarly, other

reservoir properties affect elastic properties such as k and μ, and they further affect, velocity and density of the rock.

To predict lithology, fluid, and porosity, which are of interest to reservoir engineers, we must estimate P-wave velocity, S-wave velocity, and density of layers from seismic amplitude data; this is called full elastic inversion. Both P-wave velocity and S-wave velocity are needed as they propagate through rock differently, depending on lithology and fluid content. What is observed from seismic study are the amplitudes of the reflected waves; that relationship between amplitudes and velocity is needed for seismic inversion. The amplitudes of reflected and transmitted P- and S-waves for any angle of incidence are given by Zoeppritz equation (Aki and Richards 2002).

1.4.2 Seismic Parameter Inversion

Zoeppritz equations relating to wave reflection coefficient and impedance simplify significantly when the angle of incidence is zero and the layers are flat.

$$R_{Pi} = \frac{I_{Pi+1} - I_{Pi}}{I_{Pi+1} + I_{Pi}} \tag{1.3}$$

where R_{Pi} is the zero offset P-wave reflection coefficient at the interface i and $I_{Pi} = \rho_i v_{Pi}$ is the P-impedance of the layer i. To use this approximation, CMP gathers can be stacked (added) and approximated as zero offset reflections. This is because the average of all the angles of reflection implies a mean angle of incidence of zero (Russell and Hampson 1991).

A seismic trace is not an RC series, but it is generated by the convolution of the wavelet and the RC series

$$\mathbf{S}(t) = \mathbf{W}(t) * \mathbf{R}(t) + \mathbf{N}(t) \tag{1.4}$$

where $\mathbf{S}(t)$ is the seismic trace, $\mathbf{W}(t)$ is the seismic wavelet, $\mathbf{R}(t)$ is the reflectivity, $*$ denotes convolution, and $\mathbf{N}(t)$ is the noise.

There are different techniques to invert seismic data and many common parametric inversion methods involve the following steps:

- Start with an initial guess of model parameters $\mathbf{m}$, like P- and S- wave velocities.
- Define an objective function, $f(\mathbf{m}) = \| \mathbf{S} - \mathbf{D} \|$ that characterizes the misfit between observed seismic data $\mathbf{D}$ and corresponding prediction $\mathbf{S}$ (1.4).
- Minimize the objective function $f(\mathbf{m})$ to get a most-likely estimate of $\mathbf{m}$.

Inversion using the steps explained above along with rock physics gives the reservoir parameters (Tarantola 2004).

1.5 Motivation

Uncertainties of reservoir properties are minimized by integrating available multiple types of data. Densely distributed seismic data reveal only decameter-scale features, and so these subseismic geological layers and their properties should be integrated with hard well data; but well data are sparse. Also, conceptual geological models provide continuity information between two lateral points and stacking patterns; this information is not provided by trace-based stochastically inverted seismic data or the well data. All this information is mainly at two different scales: one is seismic scale information, generally with less precision, but denser; the other is subseismic scale information estimated be means of well data, with high precision at the well locations but wells are sparse. The former constrains the latter for thickness and porosity in a form like

$$\sum h_i = H$$
$$\sum \phi_i h_i = \phi H$$

where H is the total thickness predicted by seismic, ϕ is the porosity and h is the layer thickness. Incorporating additional information such as constraints on the sum of the values increases the precision of estimates (Deutsch, Srinivasan, and Mo 1996; Behrens et al. 1998; Doyen et al. 1997). Another issue, apart from constraining the individual thickness sum to seismic thickness, is that all estimates must be nonnegative (Michalak and Kitanidis 2003).

Even though geologic parameters like thickness and porosity are nonnegative, the way these parameters are approximated in modeling by Gaussian functions may make the transformed parameters negative; for thickness, the negative value indicates a pinchout layer. Transformed parameters can be made nonnegative by data transformation or by using Lagrange multipliers (Fletcher 2000). The commonly used power transformation (a special case is the log transformation, with a power of zero) cannot be used for truncated distributions in which there is a significant probability that the variable could be zero. Also, transformations distort additive properties so that sum constraints (as above) cannot be imposed. For example, if we were to use $\ln(h)$ to model a log-normal variable and the constraint is $\sum h = H$, then because $\sum \ln h \neq lnH$, an additive constraint is hard to impose. Using Lagrangian multipliers with nonnegative constraints could solve the problem. This would require expanding the parameter space $\mathbf{x}$ with Lagrangian multipliers ($\boldsymbol{\lambda}$), expanding the original function $f(\mathbf{x})$ into the Lagrange function, and solving the function (1.5) with k active constraints where k is the number of layers in each column.

$$L(\mathbf{x}, \boldsymbol{\alpha}) = f(\mathbf{x}) + \sum_{i=1}^{k} \lambda_i g_i(\mathbf{x}) \tag{1.5}$$

The solution is obtained by differentiating the Lagrangian with respect to $\mathbf{x}$ and $\boldsymbol{\lambda}$. This gives equations (1.6) that can be solved to obtain the best estimate of f with constraints $g_i(\mathbf{x}) \geq 0$ in our case $g_i(x) = x$, the layer thickness.

$$\begin{aligned} \nabla_x L(\mathbf{x}, \boldsymbol{\lambda}) &= 0 \\ g_i(\mathbf{x}) &\geq 0, i \in K \\ \lambda_i &\geq 0, i \in K \\ \lambda_i g_i(\mathbf{x}) &= 0, \forall i \end{aligned} \tag{1.6}$$

This a quadratic programming (QP) problem where $f(x) = \mathbf{x}^T C\mathbf{x} + \mathbf{g}\mathbf{x}$, with inequality constraints. If the constraints $g_i(\mathbf{x}) \geq 0$ for the problem are affine functions and f is quadratic and positive definite, then convexity can be exploited to obtain the solution. Addressing the

problem as QP gives results which can be interpreted as a maximum likelihood estimate (MLE), but individual realizations to capture uncertainty are either difficult or impossible. If a single most likely estimate of the properties is to be obtained, then using the Lagrange maximum likelihood method suffices. But when conditional realizations from the posterior are desired, then a single MLE is not adequate.

1.6 Objectives

The methods developed to generate geologic models must integrate data from seismic, well, and geologic information consistently. This integration is challenging because the resolution and accuracy of these data types differ. Consistency requires downscaling methods not biased toward any particular data. Seismic inversion results have the lowest resolution but are spread throughout the reservoir; away from wells, seismic data provide constraints which should be honored. The well data has the resolution needed for geomodels, but they are areally sparse. Well data should be honored precisely when geomodels are built. Geologic correlation lengths quantify how continuous a geologic body can be, and this informs the probability that a layer is present or not at a given location. Another form of geologic knowledge is the conceptual stacking pattern that describes how layers are related with each other; stacking models should be included when geomodels are built.

1.7 Significance of the Research

Integrating diverse data improves the accuracy of reservoir models. Integrated geologic models with seismic, well, and conceptual information better describe the range of geomodel properties and geometry. The reason for preparing geologic models is that they are required for reservoir simulation, which is essential to assessing reservoir performance under different development scenarios. Models built with all the data will be more accurate and so give more relevant production performance. In this dissertation, the methods used to integrated different scale information are based on Bayesian methods; because of their generality, they are

applicable for other types of information also. Any information that has low or high resolution and present everywhere or few locations can be integrated using the proposed methods. The formulation in terms of a Gaussian proxy for thickness or porosity allows modeling layers that pinch out or have nonpay (zero porosity) regions. The proposed Bayesian formulation includes seismic constraints, pinch-out modeling, and integration of geologic data in the form of correlation lengths, seismic inversion and well data. This comprehensive formulation is an original contribution of this dissertation to geomodeling research.

In this current work geologic information like correlation lengths and stacking patterns are also integrated. Incorporating this information is especially important when fewer well data are available. This is because stacking pattern information helps to extrapolate the geologic models whenever there is little well data to guide the subseismic layers. Extrapolation is a challenge for traditional kriging based algorithms. Stratal architecture is integrated using surface-based models, and the resulting geomodels generated are consistent with the well and seismic information. The problem of integrating stratigraphic data in building geologic models using surface-based model is equivalent to surface-based model conditioning. The method to condition these models is shown in this study using the ensemble of surface-based models. This procedure circumvents the need to condition individual surface-based realizations, and so avoids the problems of creating artifacts around wells. Surface-based models are generally difficult to condition, so the proposed preprior approach is an important contribution.

The proposed methods create ensembles of realizations to characterize uncertainty. A multivariate probabilistic screening method is proposed to select realizations that span the flow response space; the approach is also flexible in the responses considered, rigorous in its use of high-discrepancy sampling methods, and original in its conception and implementation. It can be used to make complex geomodeling workflows more computationally tractable.

1.8 Current Approach

Previous research has introduced an open source tools that support this workflow, including Delivery(Gunning and Glinsky 2004) and Massager (Gunning and Glinsky 2006; Glinsky et al. 2005). Delivery performs a fully probabilistic seismic inversion using a layer-based model of the reservoir (where the layering and rock physics comprise the prior model), and Massager introduces lateral correlation into the point estimates and then maps the seismic grid data to corner point grid data. The results are gridded arrays of relevant reservoir parameters such as layer thickness, hydrocarbon probability, and auto- and cross-covariances. An ensemble of realizations of the reservoir properties can be generated both for volumetric calculations, as well as statistical reservoir modeling.

In this research stochastic seismic inversion models generated by tools like Delivery and Massager are downscaled. Two possible approaches to downscale the stochastic inversion models impose the seismic data as a constraint *via* Bayesian likelihood, but differ in the formulation of that likelihood. Both the methods integrate seismic data by visiting traces sequentially, which is similar to sequential Gaussian simulation. One approach downscales seismic constraints with precision scaled on the seismic noise. Although each seismic inversion model coming from stochastic inversion gives particular values, *e.g.*, sum and average constraints for total thickness and average porosity, an ensemble of these realizations can be used to characterize the variances and an averages of constraints. If these results are assumed to follow Gaussian distributions, downscaling seismic properties using inexact constraints should be weighted toward the mean seismic constraint, with weighting related to the inverse of the seismic noise. The second downscaling method uses seismic information as exact constraints, downscaling one particular seismic inversion model. The exact constraint method is preferred because each seismic realization has thickness, porosity, saturation and other properties is consistent with the amplitude, velocity model, density model, and wavelet information (and their covariances). Thus, the exact constraint approach yields models with

the proper correlations between the flow properties. For example, if the travel time of the wavelet is high, then the layer thickness may be high or porosity of that layer is high. Again, these correlations are honored when one particular layer is downscaled but downscaling one realization with exact constraints is a challenging problem. Seismic uncertainty is addressed by considering many different inversions as constraints; thus the workflow has a cascading, hierarchical form (Kalla et al. 2007).

Even with all this data in a geomodel still there is uncertainty as there are many unknown geomodel properties to estimate from relatively few measurements. Characterization of this uncertainty is feasible when seismic inversion data are downscaled. One way of modeling this uncertainty is by generating ensembles of models; each ensemble realization is called a realization. When seismic are used as exact constraint, each seismic inversion models has many possible downscaled models corresponding to it, and the need to consider multiple inversion models combinatorically generates more realizations. In the inexact case, multiple models could be generated from downscaling ensemble seismic statistics. All these models should be consistent with all available data.

A selection of $O(10)$ of the $O(10^2) - O(10^3)$ realizations will be used in multiphase flow simulations; full-physics flow simulations are expensive. Because some of the realizations are similar, they will respond to fluid flow in a similar way. The realizations selected should be different from each other when responding to fluid flow, and must be selected in a way that maintains a rigorous probabilistic framework. One of the aims of this study is to sample realizations that are identified as diverse so that they capture the uncertainty in flow responses.

1.9 Outline of Thesis

The next chapter of this thesis chapter 2 incorporates seismic constraints with noise. The algorithm proposed augments additional variables which are indicators of a layer being zero in the SGSIM algorithm, and uses MCMC techniques to sample the posteriors (Kalla et al.

2006). This sample from unnormalized truncated Gaussian priors of the layer thickness serves to integrate data at two different scales to create an ensemble of reservoir property maps. The final product of this study is a set of horizon depths filled with layers and the property maps for each layer. In chapter 3, formulation for the exact constraint problem is discussed. When the seismic constraint is exact, an algorithm rotates the basis and samples on a hypersurface. This formulation has the attractive property of preserving correlations in the rock physics and inversion results.

A block solver technique for avoiding sequential simulation pitfalls is discussed in chapter 4. In this chapter, a sequential simulation similar to SGSIM is compared with a global method similar to LU decomposition and a sequential method with approximate marginals. Integrating stratigraphic information to the downscaling algorithm by using surface based models is the topic of chapter 5. Adding stratigraphic information to seismic and well control constraints the geologic model and generates more realistic-looking as well as more data-rich reservoir models. Important aspects of building stratigraphic models using surface based models are discussed. Choosing representative realizations is discussed in chapter 6. Importance of sampling rather than ranking in choosing realizations is discussed in section 6, before providing overall summary and conclusions in chapter 8.

Chapter 2
Downscaling of Seismic Inversion Inexact Thickness*

2.1 Integrating Geologic, Seismic and Well Data

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 spare, 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. Besides the limited resolution, seismic-derived depths and thicknesses are uncertain, due to noise in the seismic data and uncertainty in the rock physics models (Gunning and Glinsky 2004; Gunning and Glinsky 2006). Thus, the 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, Srinivasan, and Mo 1996).

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

* Portions of this chapter appeared in 2006 SPE conference paper no. 103268.

as "enforcement," and the associated algorithms comprise the software package known as ENFORCER.

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 DELIVERYMASSAGER. 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 (Schlumberger Technology Co. 2005), although quasimechanistic depositional modeling (Griffiths et al. 2001) or surface-oriented geostatistics algorithms (Pyrcz 2004) are possible alternatives.

2.1.1 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, one may wishe to ensure consistency both in thickness and in average porosity in a downscaling problem consisting only of vertical gridding refinement.

Therefore, the following equations must be considered at column of gridblock corners:

$$\sum_{k=1}^{K} h_k = H$$

$$\sum_{k=1}^{K} h_k \phi_k = \bar{\Phi} H$$

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 corner-point grid is a difficult problem, chiefly on account of the constraints, but also because of nonlinearities.

2.1.2 Summary of Related Work

Several researchers worked on integrating seismic information in building geologic models using 2D and 3D seismic data. Early work was done by Araktingi and Bashore (1992), Bashore et al. (1993), Araktingi, Hewett, and Tran (1993), and others. These researchers integrated seismic and well data using kriging with an external drift and cokriging using a Markov-Bayes approximation (Goovaerts 1997). Seismic data are used to impact the large-scale features in the approach, considering well-log as primary data and the seismic data as secondary. The external drift method changes the kriging weights to satisfy the seismic trends. In the Markov-Bayes approach, variograms for seismic and cross-correlation are obtained from well data variogram. An inherent assumption in the study is that the variogram is scale dependent, and may be used for the same variogram with both well and seismic data.

A collocated cokriging technique using Bayesian updating for a kriging solution is used to integrate dense, but low resolution seismic data by Doyen, den Boer, and Pillet (1996). Doyen, den Boer, and Jans (1997), and Behrens and Tran (1999) considered a seismic constraints introduction into Bayesian formulation through the likelihood functions. The approach in

this thesis is also to use similar framework to make a similar assumption, by ignoring the vertical correlation with the vertical cells in order to implement the algorithm faster. The seismic likelihood function is used to update point kriging estimates at each simulated point in the 3-D model. Behrens and Tran (1998) used sequential simulation with a block kriging that treats seismic data as soft constraints to integrate seismic data. Covariance function for seismic data is derived by integrating subseismic covariance functions. Property estimates are obtained by weighted linear combinations of neighboring cells, as well as a seismic average of the current trace. Lee et al. (2002) talked about sampling posterior using Metropolis-Hasting and Gibbs algorithms. They used mismatch between seismic and sum of the layer thickness as the likelihood, and estimated prior from interpolation. This research also applies a similar framework, but for a more general problem, such as when layers pinchout.

If there is a significant possibility of thickness or porosity being zero then methods in this chapter using auxiliary variables are more appropriate then methods explained before. However, none of the approaches explained before can ensure preservation of the correlations inherent in the rock physics, seismic data, and seismic inversions. The cascading workflow, which preserves these correlations are discussed using exact-constraint approach in chapter 3 (Kalla et al. 2007).

2.1.3 Use of Terms

Layers are generally not resolved by seismic data, but can be identified in wells. 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 (Ponting 1989; *viz.*, `COORD` records, Schlumberger Technology Co. 2004). 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.

2.2 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 *via* the priors, and collocated coarse-scale constraints *via* 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 *via* 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 Markov chain Monte Carlo 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 a particular form of the covariance, which yields a computationally tractable matrix (Estimating the Prior, later). 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.

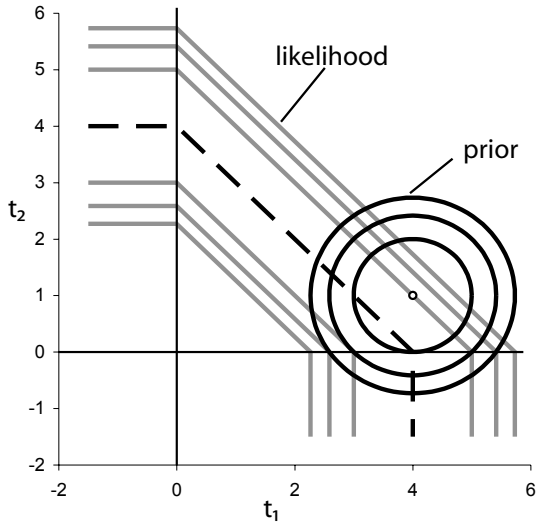


Figure 2.1: Contours of minus log likelihood and and minus log prior distributions for a two-layer system, with $\bar{H} = 4, \sigma_H = 1, \bar{\mathbf{t}} = (4, 1),$ and $\sigma_t = 1$. Contours are in increments of one in $\Delta\left(\frac{t-\bar{t}}{\sigma_t}\right)$ or $\Delta\left(\frac{\Sigma t - H}{\sigma_H}\right)$, with values of zero along exactly honoring the thickness sum (dashed line) and where $\mathbf{t} = \bar{\mathbf{t}}$ (small circle). Consistent units.

For simplicity, we will consider systems with a single thickness constraint. More general constraints are addressed in the later discussion and in the next chapter. Numerical methods and sampling methods are also discussed in later sections.

2.2.1 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 continous, yet the 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:

$$P(h_k = 0) = \int_{-\infty}^{0} dP\left(t_k\right) \tag{2.1}$$

2.2.2 Algorithm Outline

Before discussing details, the algorithm framework is presented (**Fig. 2.2**). First, the untruncated Gaussian surrogate for all conditioning data with $h = 0$ must be simulated. Then,

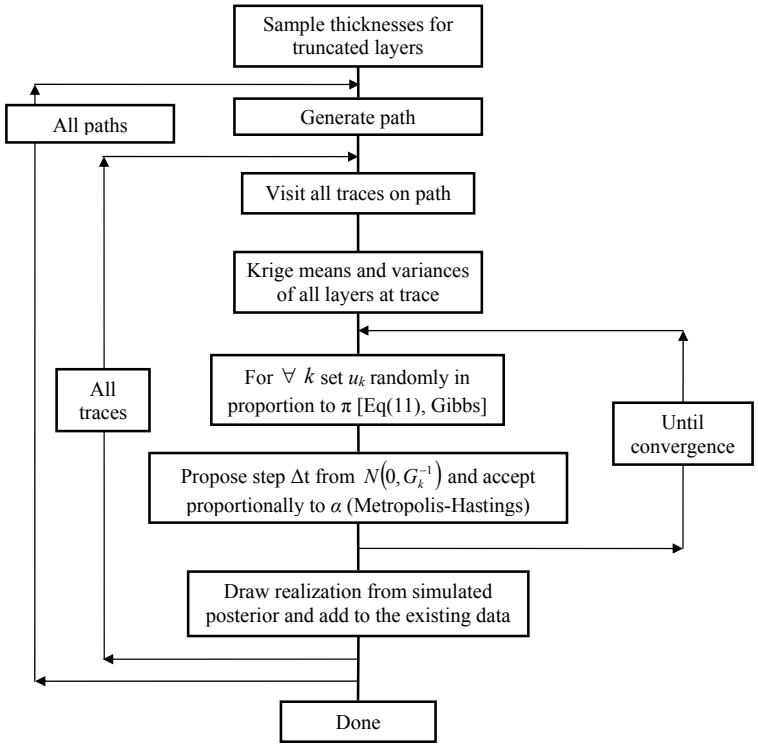


Figure 2.2: Flow chart for sequential simulation using Markov Chain Monte Carlo.

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), auxillary 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 may be used to generate multiple chains, in the same way that sequential Gaussian simulations generate multiple realizations (Deutsch and Journel 1997).

2.2.3 Estimating the Prior

This step in the algorithm supplies prior means $\bar{t}$ and variances σ^2_{tk} for all layers on a given trace.

A few assumptions can simplify the kriging solution, and greatly improve efficiency (Numerical Considerations, later).

1. For many block shapes and grid spacings, traces can be 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.
2. 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.
3. 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 (1) the priors for each layer can be computed separately; (2) the kriging matrices are identical for all layers, and therefore only one kriging system needs to be solved at each trace; and (3) 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 (Numerical Considerations, later). 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.

This would assume that a two-dimensional layer thickness correlation implies that a two-dimensional search would suffice, thus improving search efficiency.

Cokriging or collocated kriging could be used to get prior covariances (Goovaerts 1997). Such a result would integrate well with the seismic data, which would 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. 2007).

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. Although the single list might require solving a larger kriging system, it would only require solving one kriging system for all K layers.

2.2.4 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 use a simple but reasonable draw such as

$$t_k = N^{-1}\left(\bar{t}_k, \sigma^2_{tk}; r\right), r \sim U[0, P(h_k = 0)] \tag{2.2}$$

where $P(h_k = 0)$ is given by Eqn. (2.1), N is the normal distribution function, and U is the uniform distribution function. However, we model the correlation with a loop over all layers, as follows:

1. Find all zero conditioning data in this layer, k; the list of the locations of zero data is indexed over $\lambda_k \in \{0 \ldots \Lambda_k\}$. The positive conditioning data in layer k are indexed by $d \in \{0 \ldots D_k\}$.

2. Initialize all Λ_k zero thickness observations in layer k with random draws, using Eqn. (2.2).
3. 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 $N(\bar{t}_k, \sigma^2_{tk})$ at cumulative probability r.
4. Generate a chain and store.
5. 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.

2.2.5 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 the resolution $\frac{1}{\sigma^2_H}$obtained from a stochastic inversion, using DELIVERY (Gunning and Glinsky 2004).

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

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

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 a normalizing term in the denominator, which does not depend on $\mathbf{t}$. That is,

$$\pi\left(\mathbf{t}|H, \mathbf{d}_{\ell k}\right) \propto p\left(H|\mathbf{t}, \mathbf{d}_{\ell k}\right) p\left(\mathbf{t}|\mathbf{d}_{\ell k}\right) \tag{2.3}$$

We assume that departures from the prior ($\bar{t}_k$) and updating ($\bar{H}$) data means are normally distributed with standard deviations σ_{tk} 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

$$p(\mathbf{t}|\mathbf{d}_{\ell k}) = \frac{1}{(2\pi)^{\frac{K}{2}}|\mathbf{C}_p|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(\mathbf{t}-\bar{\mathbf{t}})^T\mathbf{C}_p^{-1}(\mathbf{t}-\bar{\mathbf{t}})\right] \tag{2.4}$$

where $\mathbf{C}_p$ is the prior or kriging covariance matrix, which is of rank K with the kriging variances σ_{tk}^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,

$$p(H|\mathbf{t}, \mathbf{d}_{\ell k}) = \frac{1}{\sqrt{2\pi}\sigma_H} \exp\left[-\frac{(H-\bar{H})^2}{2\sigma_H^2}\right] \tag{2.5}$$

where

$$H = \mathbf{t}^T\mathbf{T}$$

and

$$T_k = \begin{cases} 0 & \text{if } t_k < 0 \\ 1 & \text{otherwise} \end{cases} \tag{2.6}$$

The conditioning on $\mathbf{d}_{\ell k}$ in Eqn. (2.5) is indirect, due to the conditioning of $\mathbf{t}$ on $\mathbf{d}_{\ell k}$. The product of Eqns. (2.4) and (2.5) is the proportional to the posterior, Eqn. (2.3). This product can be converted to a quadratic form by taking the logarithm, giving

$$\begin{aligned} -2\ln\left[\pi\left(\mathbf{t}|H, \mathbf{d}_{\ell k}\right)\right] = &\ln\left[(2\pi)^K|\mathbf{C}_p|\right] + \ln\left(2\pi\sigma_H^2\right) + \\ &(\mathbf{t}-\bar{\mathbf{t}})^T\mathbf{C}_p^{-1}(\mathbf{t}-\bar{\mathbf{t}}) + \frac{(\mathbf{t}^T\mathbf{T}-\bar{H})^2}{\sigma_H^2} \end{aligned} \tag{2.7}$$

We seek a stationary point in the posterior probability by setting the gradient with respect to $\mathbf{t}$ of Eqn. 2.7 to zero, *viz.*,

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

The Hessian, $\mathbf{G}$ of Eqn. (2.7) is the desired resolution matrix (which is the inverse of the posterior covariance):

$$\mathbf{G} = \mathbf{C}_p^{-1} + \mathbf{T}\mathbf{T}^T/\sigma_H^2 \tag{2.8}$$

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

$$\mathbf{G} = \begin{pmatrix} \frac{1}{\sigma_{t1}^2} + \frac{1}{\sigma_H^2} & \frac{1}{\sigma_H^2} & \cdots & \frac{1}{\sigma_H^2} \\ \frac{1}{\sigma_H^2} & \frac{1}{\sigma_{t2}^2} + \frac{1}{\sigma_H^2} & \cdots & \frac{1}{\sigma_H^2} \\ \vdots & \ddots & \ddots & \vdots \\ \frac{1}{\sigma_H^2} & \frac{1}{\sigma_H^2} & \cdots & \frac{1}{\sigma_{tK}^2} + \frac{1}{\sigma_H^2} \end{pmatrix} \tag{2.9}$$

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

2.3 Prior and Likelihood Distributions in 2D

Important features of higher-dimensional cases are easily visualized for a system with two layers (**Fig. 2.1**). The dashed line in Fig. 2.1 is the thickness sum constraint, and lines parallel to it are isoprobability contours. In three dimensions, the dashed line in Fig. 2.1 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. 2.1; it would be ellipsoidal if prior variance differed by layer, and an inclined ellipsoid should the layer thicknesses be correlated. Such priors could be sampled by 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 (4 m and 1 m for the two layers) sum to greater than the mean trace thicknesses (4 m), so the prior center of mass [circles in Fig. 2.1; Eqn. (2.4)] lies above the maximum likelihood line [dashed line in Fig. 2.1; Eqn. (2.5), 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) having a 45 degree slope along the dashed line could be used. Conversely, in a sequential simulation not conditioned to seismic, the layer thicknesses would simply be drawn from the prior (Fig. 2.1).

Sampling problems are caused by the nonlinearity [Eqns. (2.2.5, 2.6)] apparent as slope discontinuities in the likelihood where the axes intersect the contours of the likelihood surface (Fig. 2.1). 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. 2.1). 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 (Auxiliary Variables to Treat Pinchouts, later discussed).

2.4 Sampling Approach

Because the log–posterior surface is quadratic with constraints (Eqn. (2.8)), 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. Samples are generated using the Markov chain Monte Carlo (MCMC) method (**Fig. 2.2**).

In this section, the focus is 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.

2.4.1 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 $[0, P(h_k = 0)]$ to apply an inverse normal transformation, but this would decorrelate the variables. Instead, we precondition these data using a Gibbs sampler to preserve the correlation (Zero Thickness Conditioning, earlier discussion).

2.4.2 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. 2.1). 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, *i.e.*, alteration between states, in difficult MCMC problems with related "configurational stiffness" characteristics (Higdon 1998). Auxiliary variable methods use an augmented posterior probability space:

$$\pi(\mathbf{u}, \mathbf{t}) = \pi(\mathbf{t})\pi(\mathbf{u}|\mathbf{t}) \tag{2.10}$$

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 $\pi(\mathbf{u}|\mathbf{t})$ $\left[= \prod_{k=1}^{K} \pi(u_k|t_k)\right]$ 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

$$\pi\left(u_k = 1 | t_k\right) = \begin{cases} 1 - \frac{1}{2 + t_k/\sigma_{\pi k}} & \text{if } t_k \geq 0 \\ \\ \frac{1}{2 - t_k/\sigma_{\pi k}} & \text{otherwise} \end{cases} \tag{2.11}$$

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

$$\frac{1}{\sigma_{\pi k}^2} = \frac{1}{\sigma_{tk}^2} + \frac{\kappa}{\sigma_H^2} \tag{2.12}$$

That is, $\sigma_{\pi k} \approx \sum_{j=1}^{K} G_{kj}$, [Eqn. (2.9)]. 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 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 (step size and direction are 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 $\mathbf{u}$.

2.4.3 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 $\mathbf{G}_\kappa = \mathbf{C}_p^{-1} + \mathbf{u}\mathbf{u}^T/\sigma_H^2$ is used to make the MCMC jumping proposal [Eqn. (2.13), later]. The appropriate resolution and inverse matrices are computationally inexpensive for the simple Toeplitz resolution matrix used in the proposed approach (Numerical Considerations, discussed later). The Hessian $\mathbf{G}$ and the posterior covariance $\mathbf{C}_\pi = \mathbf{G}^{-1}$ are of rank K, but the matrix inverse used in this sampling is of lower

rank κ (Numerical Considerations, later). The Cholesky factor $\mathbf{L}_{C\pi}$ of the covariance matrix (the Cholesky factorization is $C_\pi = \mathbf{L}_{C\pi}\mathbf{L}_{C\pi}^{\mathrm{T}}$) 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 $\mathbf{t}$,

$$\Delta\mathbf{t} = s\mathbf{L}_{C\pi}\mathbf{r} \tag{2.13}$$

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 that likelihood at the new point $\mathbf{t}' = \mathbf{t}+\Delta\mathbf{t}$, using Eqn. (2.5). The Metropolis-Hastings transition probability is then (Gelman et al. 2003)

$$\alpha = \min\left(1, \frac{\pi\left(\mathbf{t}'|H, \mathbf{d}_{\ell k}\right)\prod_{k=1}^{K}\pi(u_k|t'_k)}{\pi\left(\mathbf{t}|H, \mathbf{d}_{\ell k}\right)\prod_{k=1}^{K}\pi(u_k|t_k)}\right) \tag{2.14}$$

Eqn (2.14) 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.

2.5 Numerical Considerations

The Toeplitz form of the posterior resolution matrix and subsidiary assumptions simplify computations (Estimating the Prior, earlier). Because of these simplifications, only two matrix solutions are required per trace: (1) 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 (2) 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 will be 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 (Sampling Approach, earlier), 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, $\mathbf{\Delta t}$, from uncorrelated random normal input vectors, $\mathbf{r} \sim N(0,1)$, $\Delta\mathbf{t} = \mathbf{L}_{C\pi}\mathbf{r}$ (Metropolis-Hastings Step, earlier; 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, (1) Factor $\mathbf{G}$ to get $\mathbf{L}_G$, (2) invert $\mathbf{L}_G$ by back-substitution to get $\mathbf{L}_G^{-1}$ (inexpensive because the matrix is triangular), and (3) take the persymmetric transpose (Golub and van Loan 1996) of $\mathbf{L}_G^{-1}$. This is the Cholesky factor of $\mathbf{C}_\pi$, $\mathbf{L}_{C\pi}$.

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 the 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), by 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 in a full order of improvement (further discussion in Performance, later).

2.6 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

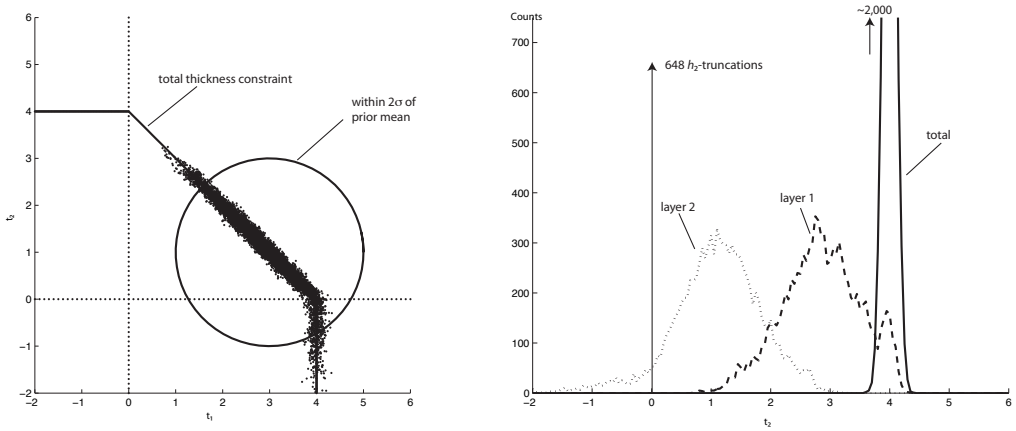
Table 2.1: Parameters and results for 2-layer simulation

Case	Prior $\bar{t}_1$	Prior $\bar{t}_2$	Prior σ_t	Constraint $\bar{H}$	Constraint σ_H	Posterior $\bar{t}_1$	Posterior $\bar{t}_2$	Posterior Covariance of t	Posterior $\bar{H}$	Posterior σ_H
Tight	3.0	1.0	1.0	4.0	0.1	2.86	1.11	$\begin{pmatrix} 0.46 & -0.50 \\ -0.50 & 0.59 \end{pmatrix}$	4.00	0.10
Loose	3.0	1.0	1.0	4.0	0.5	2.97	0.97	$\begin{pmatrix} 0.53 & -0.46 \\ -0.46 & 0.72 \end{pmatrix}$	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	$\begin{pmatrix} 0.16 & -0.08 \\ -0.08 & 0.16 \end{pmatrix}$	5.31	0.41

consistent versus slightly contradictory prior means and constraints. Results are summarized in **Table 2.1**.

2.6.1 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. 2.3**). 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 2.1). 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. 2.3(a)]. The posterior layer variances are reduced, because of the added data in the constraint (eigenvalues, Table 2.1). 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.



(a) Scattergram, $N = 8000$

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

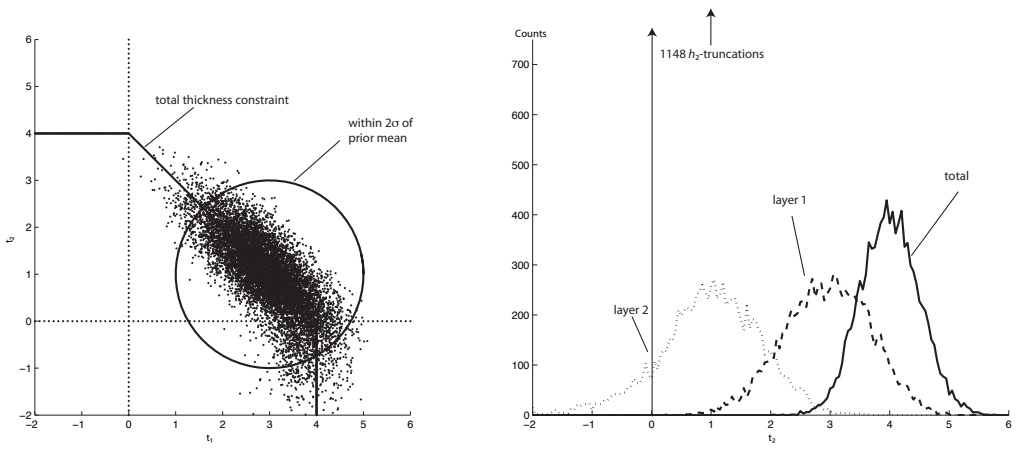
Figure 2.3: Simulation results for a two-layer case with inaccurate layer thickness but total thickness $(h_1 + h_2)$ tightly constrained. $\bar{H} = 4$, $\bar{\mathbf{t}} = (3, 1)^T$, $\sigma_H = 0.1$, and $\sigma_t = 1$; consistent units.

2.6.2 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. 2.4**). 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.

2.6.3 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. 2.5**). 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; the posteriors are nearly Gaussian. For this nearly multi-Gaussian case, the constraint has transformed the isotropic, uncorrelated prior thicknesses (Fig. 2.1) to a strongly correlated, more compact posterior. Since the prior and constraint variances



(a) Scattergram, $N = 8000$

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

Figure 2.4: 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.

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 2.1; Gelman et al. 2003).

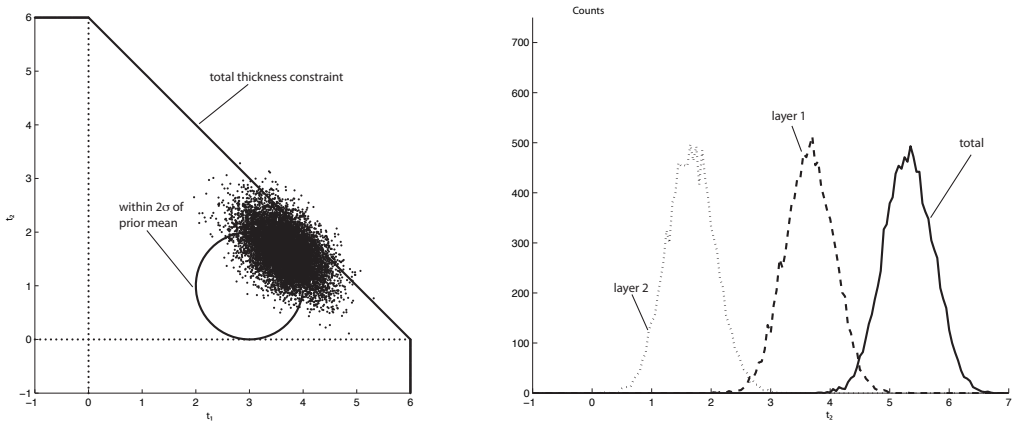
2.7 Synthetic 3D Cases

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, $x - y$ 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



(a) Scattergram, $N = 8000$

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

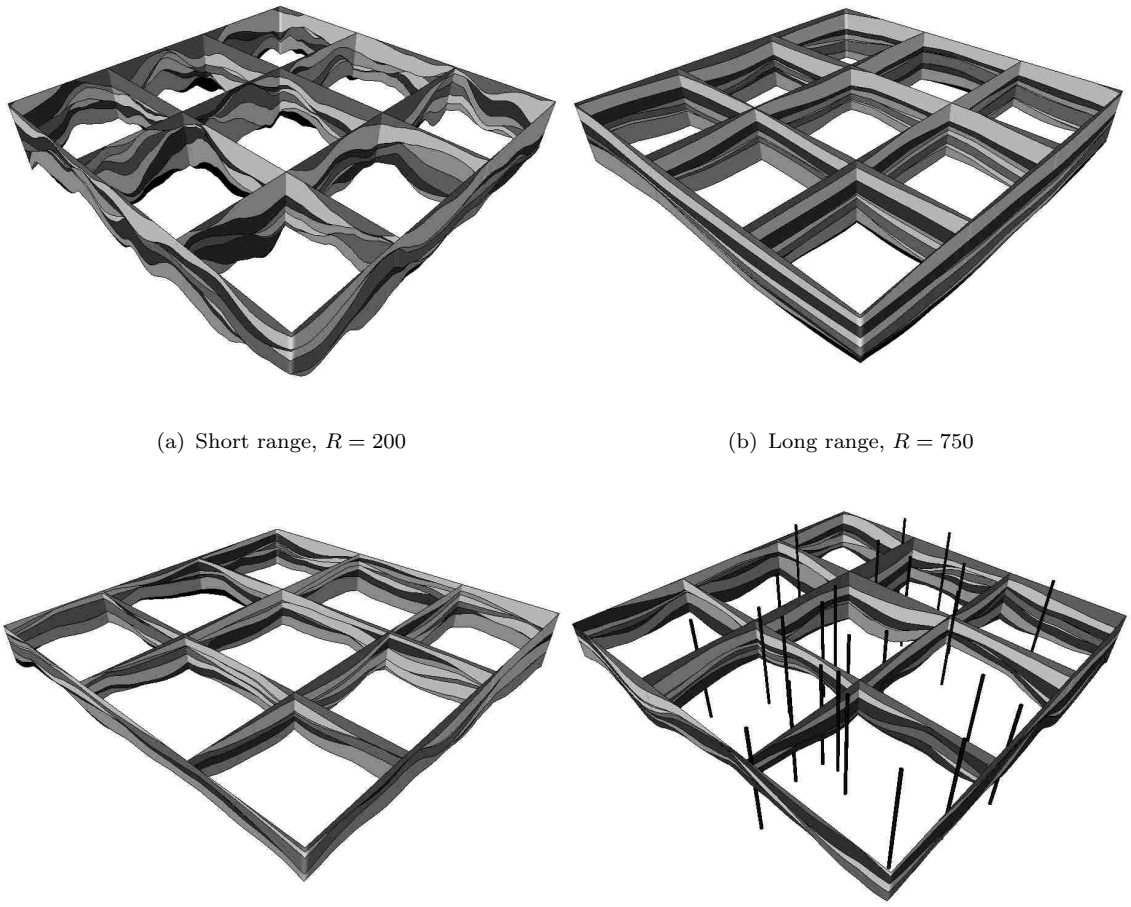
Figure 2.5: Simulation results for a two layer case with prior sum less than the sum constraint. $\bar{H} = 6$, $\bar{\mathbf{t}} = (3, 1)^T$, $\sigma_H = 0.5$, and $\sigma_t = 0.5$; consistent units.

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.

2.7.1 Illustrative Cases

Four different cases show features of the data integration method (**Fig. 2.6**). With short ranges, termination is more common, although the average layer thickness is similar to the longer range [Figs. 2.6(a,b)]. 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 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 (b). 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. 2.6(c)]. 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 thicknesses, low thicknesses



(a) Short range, $R = 200$

(b) Long range, $R = 750$

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

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

Figure 2.6: 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 $R_x = R_y = R$, $\gamma(\Delta) = 1 - \exp\left[-\left(||\Delta||/R\right)^2\right]$, m^2. All models flattened on the topmost surface. Range, thickness, and standard deviation are in m. 7.5× vertical exaggeration for all figures. Vertical black lines in (d) are conditioning traces.

fluctuate more, as may be seen by comparing the left front edges of Figs. 2.6(c) and (d). For the low variance case (c), the edge panel is of nearly uniform thickness while the nonuniform variance case (d) 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 segments are zero-thickness in one of the example cases [Fig. 2.6(c)]. The complex pinchout structure is obtained, even though the conditioning data are not especially dense [Fig. 2.6(d)].

2.8 Discussion

2.8.1 Cornerpoint Grids

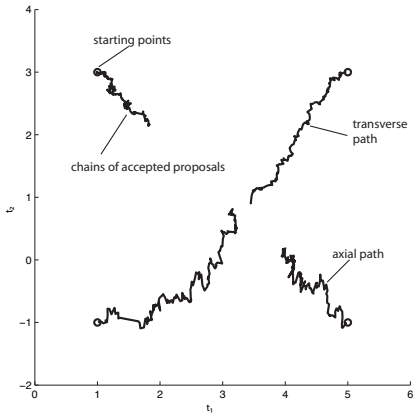
The MCMC simulation is over the block edges, or traces. This is different from many geostatistical modeling approaches, which are commonly block-centered. However, geometry – especially pinchouts or discontinuities at faults – can be modeled more accurately using cornerpoints. The porosity and other 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 1993). 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.

2.8.2 Convergence of Inexact MCMC Simulations

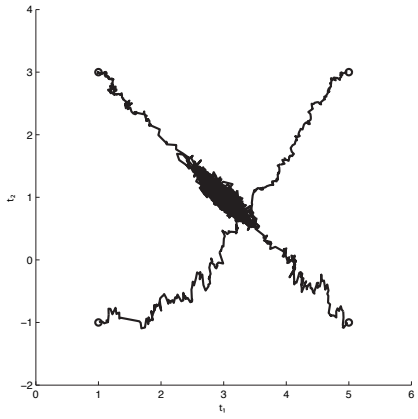
MCMC methods may converge too slowly to be practical, or may have multiple modes, such that multiple chains and/or methods to switch between modes are needed. In numerical experiments undertaken so far in this algorithm, these potential problems do not appear to be too severe.

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 $\approx$ 1000 iterations, and the variances stabilize in no more than $\approx$ 2500 iterations. That is, some 2500 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 (Numerical Considerations, previously), 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 $\approx$ 0.4, which makes such long, burn-in requirements manageable (Synthetic 3D Cases, later). Chains started in widely dispersed parts of **t**-space converge to the same posterior (**Fig. 2.7**). This was expected, based on the relatively simple form of the posterior resolution matrix, **G**. The early behavior depends on the starting point [Fig. 2.7(a)]: 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 2.1). After many iterations, all chains are sampling the same region [Fig. 2.7(b)], and the post-burn-in chains are statistically indistinguishable.

The simple, two-dimensional examples indicate the algorithm is reproducing expected results in limiting cases.



(a) First 150 samples of four chains



(b) Converging chains with 10,000 samples per chain (excluding burn-in)

Figure 2.7: 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. 2.3.

2.8.3 Performance of Inexact MCMC Simulations

For adequate performance, an MCMC simulation should converge to its target distribution in as few steps as possible. A larger step size helps explore the posterior in fewer steps. On the other hand, large steps are more likely to be rejected, "wasting" computations on a sample that would not be retained. The step size is usually adjusted indirectly, by scaling the posterior covariance (which is used to generate steps; Metropolis-Hasting step, earlier). For the system examined, the covariance is not scaled; this gives a step size applying the square root of the smallest diagonal element in the posterior covariance matrix. In high-dimensional problems, it may be more appropriate to use $\tilde{\mathbf{C}}_\pi = \frac{5.76}{K}\mathbf{C}_\pi$ 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, as well as smooth posterior samples (where they should be smooth: for example, if the prior

Table 2.2: Performance summary for the 3D example (one complete simulation)[a]

Process	Work in seconds[b]
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

[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.

implies truncations are either very unlikely or almost certain). Therefore the best choice of scaling is problem-dependent.

The computational cost of a single simulation [for the case of Fig. 2.6(a)] is examined component-by-component in **Table 2.2**. 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 two-dimensional, 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 five 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 it is not too alarming, for a model with 10^7 unknowns. The Toeplitz covariance and inversion work, then, scales approximately with the third power of layer count (Numerical Considerations, previously) and linearly for generating samples at traces. However, Toeplitz solver work takes less than 1

percent of the computing time (Table 2.2). That is, although the cubic scaling is unfavorable for large K, the multiplier for the Toeplitz work is small; as a result, this component does not control the total work required. This is because proposing samples consume most of the work; each trace has thousands of proposals and therefore 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.

2.8.4 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 a 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 with similar resulting images.

Many issues remain – especially implementation of more complex constraints, as well as integration with fine-scale geomodels in an auxiliary variable framework; the proposed method appears to offer a foundation for further development.

Chapter 3

Downscaling Multiple Seismic Inversion Constraints to Fine–Scale Flow Models*

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 attain typical geocellular scale (≈1 m) models is to downscale mesoscale models to meter-scale models by using constraint equations embodying the effective media laws. In particular, downscaling specific realizations similar to exact constraints drawn from the posterior of a stochastic mesoscale inversion, produces sum constraints for fine scale models.

We use probabilistic depth and thickness information originating from the layer–based seismic inversion code DELIVERY (Gunning and Glinsky 2004) 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.,* Schlumberger Technology Co. 2004). Block-center properties such as porosity are obtained by averaging the edge properties.

Realization ensembles from seismic inversions (*e.g.*, DELIVERY; Gunning and Glinsky 2004) carry rich information about interproperty and vertical interzone correlations, induced by the seismic information (**Fig. 3.1**). These ensembles are generated, assuming there is no trace-to-trace correlation, and the traces generally do not coincide with cornerpoint edges in the flow grid. This must be corrected by augmenting the interproperty and interzone correlations with the mesoscale lateral correlation structures required for geological continuity, and constructing models or model–samples at the quasivertical cornerpoint edges of the flow grid (*e.g.*, DELIVERYMASSAGER; Gunning, Glinsky, and White 2007). Each realization from

* Portions of this chapter appeared in 2007 SPE conference paper no. 110771.

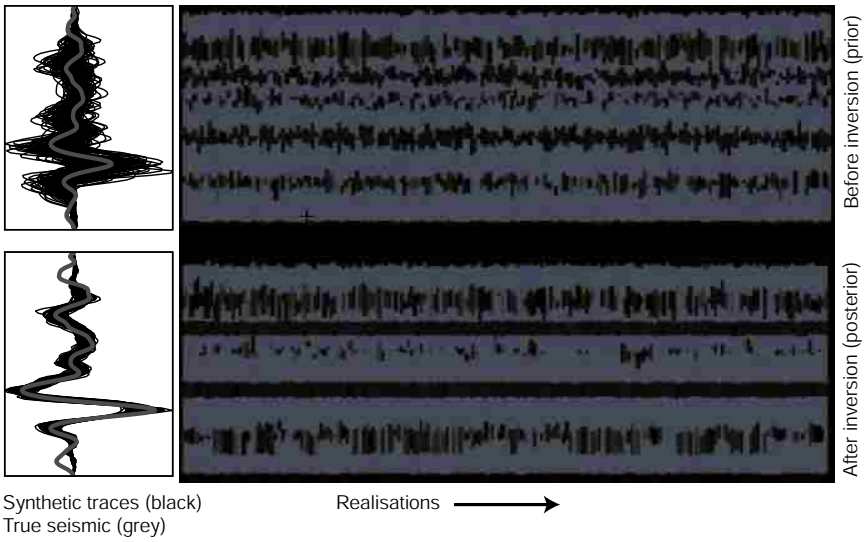


Figure 3.1: Layer–based realizations before (upper right) and after (lower right) seismic inversion, as produced by the DELIVERY code at a particular seismic trace. Synthetic traces corresponding to particular realizations are shown in black with the actual seismic data (grey) in the left insets.

deliveryMassager thus captures vertical, horizontal, and interproperty correlations at the mesoscale (**Fig. 3.2**).

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 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 ($k \in \{1 \dots K\}$) layers, each layer k with thickness h_k and porosity ϕ_k. We use t as an untruncated surrogate for layer thickness, $h_k = \max(0, t_k)$: 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:

$$\sum_{k=1}^{K} I_k \max(0, t_k) = H_s \tag{3.1}$$

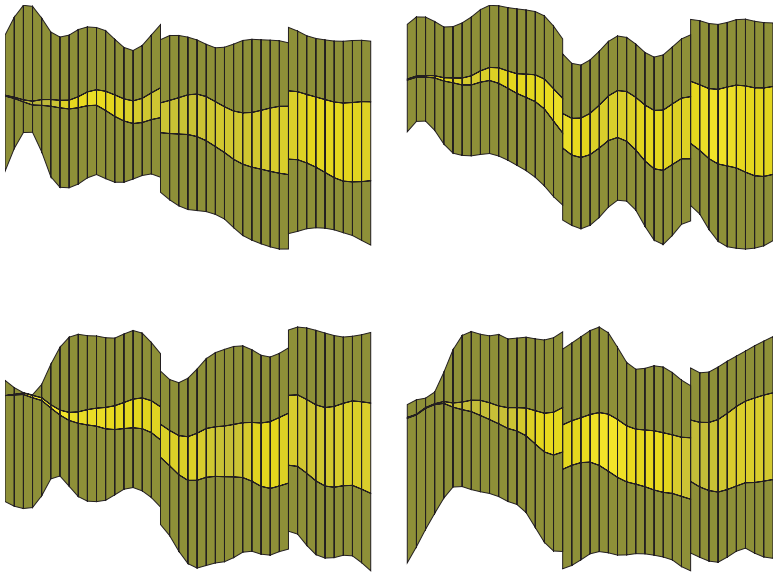


Figure 3.2: Four mesoscale layer–based realization cross–sections *in depth* of a simple 3–layer shale/sand/shale wedge test problem with Graben–like fault, as output from the DeliveryMassager code.(Gunning, Glinsky, and White 2007) Reflection surfaces that are well defined in time may still have appreciable depth variability, due to transverse variations and uncertainties in the velocity structure.

$$\sum_{k=1}^{K} I_k \max(0, t_k \phi_k) = \Phi H_s \tag{3.2}$$

$$\sum_{k=1}^{K} (1 - I_k) \max(0, t_k) = H_{sh} \tag{3.3}$$

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 net porosity-thickness ΦH_s. Mesoscale porosity Φ is the net thickness weighted average. Here $I_k \in \{0, 1\}$ is a predetermined facies indicator for layer k, where there are $K_s \leq K$ "sand" layers with $I_k = 1$, and $K_{sh} = K - K_s$ "shale" layers with $I_k = 0$.

The fine–scale model prior to these constraints is a joint multi–Gaussian distribution of t, ϕ (over all layers at each trace) with means and covariances constructed in the usual ways, such as from log data or variogram analysis of outcrop analogs (Willis and White 2000). 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.

If all thicknesses and porosities have prior means much higher than the corresponding prior standard deviations (*i.e.*, pinchout and nonpay intervals are unlikely), methods introduced by Doyen et al. (1997), Behrens et al. (1998), and Lee et al. (2002) could be used to integrate seismic constraints with noise (inexact constraints). On the other hand, if there is a significant possibility of thickness or porosity being zero, then methods similar to Kalla et al. (2006) using auxiliary variables are more appropriate. However, neither of these approaches can ensure preservation of the correlations inherent in the rock physics, seismic data, and seismic inversions. The cascading workflow, which preserves these correlations, is a novel feature of the proposed exact-constraint approach. This method could also be used for downscaling conceptual models and possibly to condition surface– (Pyrcz, Catuneanu, and Deutsch 2005) or process–based (Wellner et al. 2007) models.

Terms used are consistent with the last chapter. Additionally, *Truncated Gaussian* is used in a nonstandard way in this chapter. Generally, it implies a rescaled probability density, *viz.*,

$$f^*(h(t)) = \begin{cases} 0 & t < 0 \\ \frac{f(t)}{1-F(0)} & t \geq 0 \end{cases}$$

here, instead, we intend

$$f^*(h(t)) = \begin{cases} 0 & t < 0 \\ F(t)\delta(t) & t = 0 \\ f(t) & t > 0 \end{cases}$$

where (for this equation only) $\delta(t)$ indicates the Dirac delta function. Although this usage is nonstandard, it at least has the virtue of being descriptive.

3.1 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. 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 from the massaged seismic inversions). The prior distribution for each layer at a given trace is obtained by solving a kriging system (Goovaerts 1997) with well observations and previously simulated values as the informing data. This procedure is similar to 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 Monte Carlo 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).

3.1.1 Algorithm Outline

The algorithm is similar to many sequential algorithms, except that Markov Chain Monte Carlo (MCMC) is used to generate conditional samples that match seismic constraints at each spatial location (**Fig. 3.3**). A multigrid random path specifies the 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.

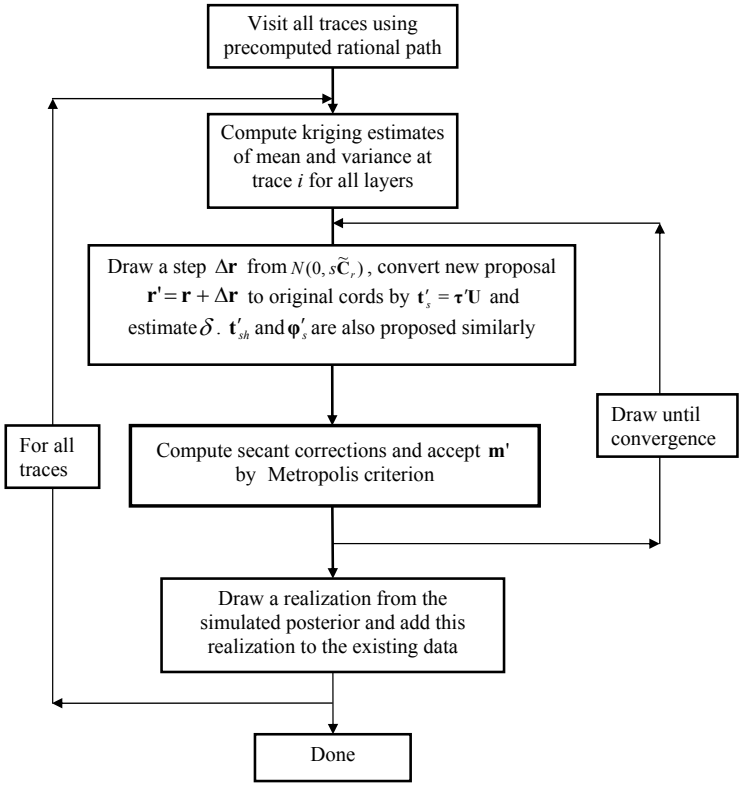


Figure 3.3: Flow chart for sequential simulation using Markov Chain Monte Carlo.

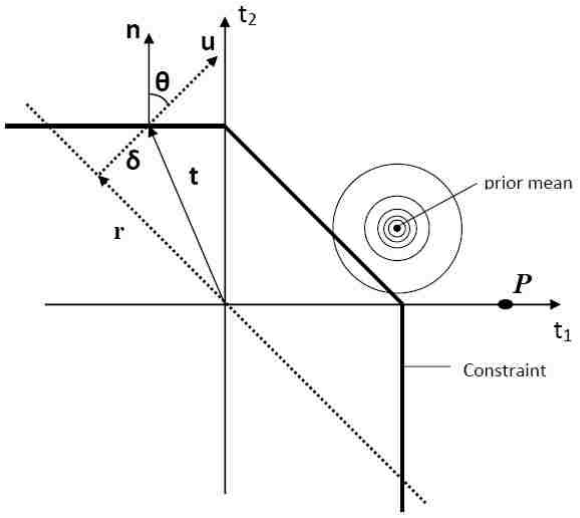


Figure 3.4: Exact constraint geometry with the prior sum more than the sum constraint. The point **r** is projected distance δ along **u** until it lies on the constraint surface.

3.1.2 Prior and Posterior Formulation

The posterior probability π for the thicknesses and porosities is, from Bayes' rule,

$$\pi\left(\mathbf{t}_s, \boldsymbol{\phi}_s, \mathbf{t}_{sh} | H_s, \Phi, H_{sh}, \mathbf{d}_\ell\right) = \frac{p\left(H_s, \Phi, H_{sh} | \mathbf{t}_s, \boldsymbol{\phi}_s, \mathbf{t}_{sh}\right) p\left(\mathbf{t}_s, \boldsymbol{\phi}_s, \mathbf{t}_{sh} | \mathbf{d}_\ell\right)}{p\left(H_s, \Phi, H_{sh}\right)} \tag{3.4}$$

where $\mathbf{d}_\ell$ is a vector of the all neighboring conditioning or previously simulated traces in the neighborhood of trace ℓ. The size of the vectors $\mathbf{t}_s, \boldsymbol{\phi}_s$ and $\mathbf{t}_{sh}$ are K_s, K_s and K_{sh}, respectively. The likelihood (first term in numerator) 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. The second term on the numerator is the prior for the local fine–scale variables, constructed using typical kriging apparatus. The denominator is a scaling term, which need not be computed for MCMC methods.

Assuming multi–Gaussian and independent distributions (Assumptions on Prior Covariance, later), the variables $\mathbf{t}_s, \boldsymbol{\phi}_s$ and $\mathbf{t}_{sh}$ are normally distributed with standard deviations σ_{ts},$\sigma_{\phi s}$ and σ_{tsh}, respectively. The multivariate prior distribution of $\mathbf{t}_s$ for example is

$$p(\mathbf{t}_s|\mathbf{d}_\ell) = \frac{1}{(2\pi)^{\frac{K_s}{2}} |\mathbf{C}_{ps}|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(\mathbf{t}_s - \overline{\mathbf{t}}_s)^T \mathbf{C}_{ps}^{-1} (\mathbf{t}_s - \overline{\mathbf{t}}_s)\right] \tag{3.5}$$

where $\mathbf{C}_{ps}$ is the prior or simple kriging covariance matrix for thicknesses of sand, which has rank K_s. Ordinary kriging is used to estimate the means. Cokriging (*e.g.*, of t and ϕ) could be used for the means and variances. We solve the kriging system, using a Cholesky factorization.

The constraints [Eqns. (3.1-3.3)] will yield posterior distributions that may be far from normal and (unlike the priors) are correlated (Simulations of Two-Layer Systems, later).

3.1.3 Assumptions on Prior Covariance

A number of assumptions make calculation more efficient for the local conditional prior distributions.

1. Traces may 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 not strictly valid for nonparallel cornerpoint block edges, or traces (the lines along which the k-index varies in a grid). However, the assumption is acceptable (1) if the traces are nearly parallel, (2) if their lateral displacement within the reservoir interval is small compared to trace spacing, or (3) if the effect on the computed covariances is generally small. If the numbers of sand and shale facies are N_s and N_{sh}, respectively, only $2N_s + N_{sh}$ kriging systems must be solved. This number would typically be much less than the number of parameters to be estimated, $2K_s + K_{sh}$.
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; for example, in deltaic sediments one could stipulate an upward-increasing mean thickness to simulate the residuals of the trend. The form of the trend could vary by area. This rescaling of mean or variance has been widely used in geostatistics; Goovaerts (1997) recommends this approach for systems with strong, known trends. Such trends could be used for both thickness and porosity.
3. Optionally, the covariance models are the same for all the shales and sands. 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, because this assumption leads to a simple Toeplitz structure that is computationally efficient (Kalla et al. 2006), yet less important for exact constraints.

Recapitulating, the benefits of these assumptions are: (1) the kriging matrices are identical for all layers, and therefore only $2N_s + N_{sh}$ kriging systems need be solved at each trace (in this paper, 3 for t_s, t_{sh}, and ϕ_s); (2) the priors for each layer can be computed separately; and (3) the prior variances in each property are constant, but the prior means vary layer-by-layer. The problem is still tractable without these assumptions (Performance of the MCMC Simulations, later).

3.2 Sampling with Exact Constraints

Exact constraints lower the dimensionality of the sampling problem. The maximum a posteriori (MAP) estimate is commonly obtained by using Lagrange multipliers (Michalak and Kitanidis 2003; Golub and van Loan 1996). For sampling, however, we reformulate the problem in a lower-dimensional space and perform appropriate projections to ensure the constraint is honored.

3.2.1 Orthogonal Subspace Projection

At each trace, three constraints (Eqs. 3.1-3.3) are applied in this downscaling method. The system has $2K_s + K_{sh}$ random variables (sand thicknesses, sand porosities, and shale thicknesses), and thus $2K_s + K_{sh} - 3$ 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 it is assumed that there is no vertical correlation, the variables $\boldsymbol{t}_s$, $\boldsymbol{t}_{sh}$, and $\boldsymbol{\phi}_s$ are *a priori* uncorrelated between layers (*i.e.*, $\mathrm{Cov}(\boldsymbol{t}_s, \boldsymbol{t}_s)$, $\mathrm{Cov}(\boldsymbol{t}_{sh}, \boldsymbol{t}_{sh})$ and $\mathrm{Cov}(\boldsymbol{\phi}_s, \boldsymbol{\phi}_s)$ are diagonal and $\mathrm{Cov}(\boldsymbol{t}_s, \boldsymbol{t}_{sh}) \equiv \mathbf{0}$). Moreover, no correlation is assumed between ϕ_s and t_s; $\mathrm{Cov}(\boldsymbol{t}_s, \boldsymbol{\phi}_s) \equiv \mathbf{0}$ (this assumption could be modified using appropriate trend models; Assumptions on Prior Covariance, earlier). Thus, the three variables can be simulated in three separate blocks, with the only restriction being that t_s must be simulated before ϕ_s, because the simulated t_s is used in imposing the total porosity thickness constraint ΦH_s.

The following description describes the procedure for the sand thickness vector $\mathbf{t}_s$ at a single trace ℓ. The constraints for shale thickness $\mathbf{t}_{sh}$ and sand porosity $\boldsymbol{\phi}$ are honored in the same way.

3.2.2 The Projection

We reparametrize, projecting all points in $\mathbf{T}$, the original space of $\mathbf{t}_s$, onto the plane normal to $\mathbf{u} = (1, 1, \ldots, 1)/\sqrt{K_s}$. The basis for this new space of $\mathbf{R}$ is obtained by singular value decomposition or SVD (Golub and van Loan 1996). The space $\mathbf{R}$ is a $(K_s - 1)$-dimensional basis orthogonal to $\mathbf{u} = (1, 1, \ldots, 1)/\sqrt{K_s}$ (Fig. 3.4). 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* $(K_s - 1)$-dimensional $\mathbf{R}$ 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 the $(K_s - 1)$-dimensional $\mathbf{R}$ subspace because the seismic constraint reduces the dimensionality by one. Rotation is useful because some of the directional vectors in the original $\mathbf{T}$ space are aligned with some constraint facets (Fig. 3.4). In Fig. 3.4, t_1 and t_2 are the directional vectors for a 2D problem; they are parallel to the constraint surface when $t_1 < 0$ or $t_2 < 0$, whereas $\mathbf{u} = (1, 1)/\sqrt{2}$ is not parallel to any constraint facet. For this 2D example, sampling is done on the 1D $\mathbf{r}$ ($\perp$ to $\mathbf{u}$) and projected back to the constraint. In comparison, if sampling were done on the reduced basis t_1 (without rotation), a sampled t_1 greater than H cannot be projected back to the constraint surface by using the other basis vector t_2.

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

$$\sum_{k=1}^{K_s} \max(0, t_{s(k)}) = H_s \tag{3.6}$$

Eqn. (3.6) 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 $\mathbf{t}_s$ and searching between its elements for the δ that satisfies Eqn. (3.6). Once δ is known, $\boldsymbol{\tau} = (\delta, \mathbf{r})$ and $\mathbf{t}_s = \mathbf{U}\boldsymbol{\tau}$. The K_{sh} shale thicknesses are obtained similarly. For porosity, the constraint is

$$\sum_{k=1}^{K_s} \max(0, t_{s(k)}) \max(0, \phi_{s(k)}) = \Phi H_s \tag{3.7}$$

which can be solved by sorting the $\phi_{s(k)}$ for which $t_{s(k)}$ is positive, and finding $\boldsymbol{\phi}_s$. Eqns. (3.7) and (3.6) uses truncation to ensure only positive porosities and thicknesses are used when matching constraints.

3.2.3 Effects of Projection on MCMC Sampling

The constrained posterior is sampled by a random walk in $2K_s + K_{sh} - 3$ dimensions, followed by a projection to the constraints to simulate the other three variables. Therefore, for one MCMC step, we sample a vector $\mathbf{m}$ defined as

$$\mathbf{m} = \{t_{s(1)} \ldots t_{s(Ks)}, \phi_{s(1)} \ldots \phi_{s(Ks)}, t_{sh(1)} \ldots t_{sh(Ksh)}\}$$

Since the earlier assumptions (Assumptions on Prior Variance, earlier) deem $\mathbf{t}_s, \boldsymbol{\phi}_s$, and $\mathbf{t}_{sh}$ to be uncorrelated, we solve this problem with a three-fold blocking. The random walk is a Markov chain on the constraint surface. Since original sampling is done in linear $\mathbf{R}$ (one dimensional lower) for each constraint, with that sample being projected onto the nonlinear constraint, a Jacobian is used to compute the correct acceptance probability α for a jump from $\mathbf{m}$ to $\mathbf{m}'$ in the Metropolis-Hastings formula

$$\alpha = \min\left(1, \frac{\pi(\mathbf{m}')\, P_J(\mathbf{m}|\mathbf{m}')}{\pi(\mathbf{m})\, P_J(\mathbf{m}'|\mathbf{m})}\right) \tag{3.8}$$

These Jacobian terms P_J correct for changes in the angle of projection (Fig. 3.4) should the walk move between facets of the constraint hypersurface. Due to a prior lack of correlation, the Jacobians for the three variables (sand and shale thickness, plus porosity) are $P_J(\mathbf{m}|\mathbf{m}') = P_J(\mathbf{t}_s|\mathbf{t}'_s)P_J(\mathbf{t}_{sh}|\mathbf{t}'_{sh})P_J(\boldsymbol{\phi}_s|\boldsymbol{\phi}'_s)$. Similarly, the prior for the properties is

$\pi(\mathbf{m}) = \pi(\mathbf{t}_s)\pi(\mathbf{t}_{sh})\pi(\boldsymbol{\phi}_s)$. The facet used to form $\mathbf{n}$ differs for the two P_J in Eqn. (3.8): for $P_J\left(\mathbf{t}_s|\mathbf{t}'_s\right)$ it is normal to the facet that $\mathbf{t}_s$ lies on, and for $P_J\left(\mathbf{t}'_s|\mathbf{t}_s\right)$ the facet containing $\mathbf{t}'_s$ is used. The Jacobian for the jump from $\mathbf{t}_s$ to $\mathbf{t}'_s$ is

$$P_J\left(\mathbf{t}_s|\mathbf{t}'_s\right) = \frac{1}{\mathbf{u}\cdot\mathbf{n}} = \frac{1}{|\mathbf{u}||\mathbf{n}|\cos\theta} \tag{3.9}$$

where $\mathbf{n}$ is based on the location of $\mathbf{t}_s$, as discussed above. The angle θ is measured between $\mathbf{u}$ and $\mathbf{n}$ (Fig. 3.4). The ratio of the Jacobians is 1, if $\mathbf{t}_s$ and $\mathbf{t}'_s$ lie on the same facet, correctly reducing the Metropolis-Hastings α to its usual form. Including the Jacobians, P_J preserves reversibility of the MCMC steps if $\mathbf{t}_s$ and $\mathbf{t}'_s$ are not on the same facet, as required for correct MCMC calculations (Gelman et al. 2003).

3.2.4 Generating and Accepting Proposals

This section focuses 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 quasirandom, multigrid 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 $\phi-$coordinates) is not isotropic (Assumptions on Prior Covariance, earlier), then the covariance corresponding to the rotated subspace must be transformed when computing proposals, *via* $\tilde{\mathbf{C}} = \mathbf{U}\mathbf{C}_p\mathbf{U}^{\mathrm{T}}$. The Schur complement yields the covariance of the $\mathbf{R}$-space vector $\mathbf{r}$ (using $\delta = H/\sqrt{K}$, the distance from the origin to the central facet), $\tilde{\mathbf{C}}_r = \tilde{\mathbf{C}}_{22} - \tilde{\mathbf{C}}_{21}\tilde{\mathbf{C}}_{11}^{-1}\tilde{\mathbf{C}}_{12}$, where $\tilde{\mathbf{C}}$ is split up into the blocks (1,2) corresponding to the $(\delta, \mathbf{r})$ 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 $(K-1)$-long vector of standard normal deviates $\mathbf{a} \sim [N(0,1)]$ to get a properly correlated random step:

$$\Delta\mathbf{r} = s\tilde{\mathbf{L}}_r\mathbf{a}. \tag{3.10}$$

Here s is a scalar chosen for sampling efficiency (Gelman et al. 2003), typically $s = 5.76/(K-1)$ for large K. The constraint is imposed by projecting $\mathbf{r}'(= \mathbf{r} + \Delta\mathbf{r})$ and transforming to

new coordinates (Orthogonal Subspace Projection, earlier). The proposal prior probability is computed at the new point $\mathbf{t}'$, using Eqn. (2.4). The Metropolis-Hastings transition probability is then computed (Gelman et al. 2003) using Equation (3.8). 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 added to the "conditioning" data for later traces. This process is continued until all the traces are visited.

3.3 Prior Distributions and Constraints for Two or Three Properties

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

For two layer thicknesses, the constraint surface comprises two orthogonal line segments, joined by a 45 degree "chamfer" in the 1st quadrant (**Fig. 3.4**). In three dimensions, the chamfer in Fig. 3.4 is a triangle with vertices on each t-axis at H (**Fig. 3.5**); increasing H shifts the high-likelihood region away from the origin, but with no change in slope. There are six additional facets comprising the constraint in three dimensions; for K variables, there are $2^K - 1$ facets. On one facet (analogous to the 45 degree facet in 2D), all K of the t are positive. Between 1 and $K - 1$, layers have $t \leq 0$ on the remaining $2^K - 2$ facets; each facet corresponds to a distinct pinchout configuration. Depending on the prior and the constraint, the posterior of $\mathbf{t}$ may be distributed on few or many of these facets.

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

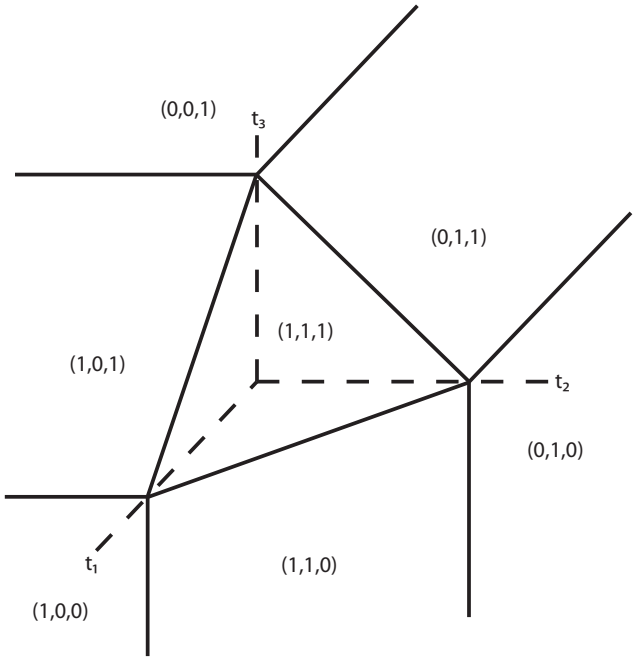


Figure 3.5: The three-dimensional constraint surface for a three layer case has $2^3 - 1 = 7$ facets. Hyperplanes intersect all t-axes at the total thickness constraint, H.

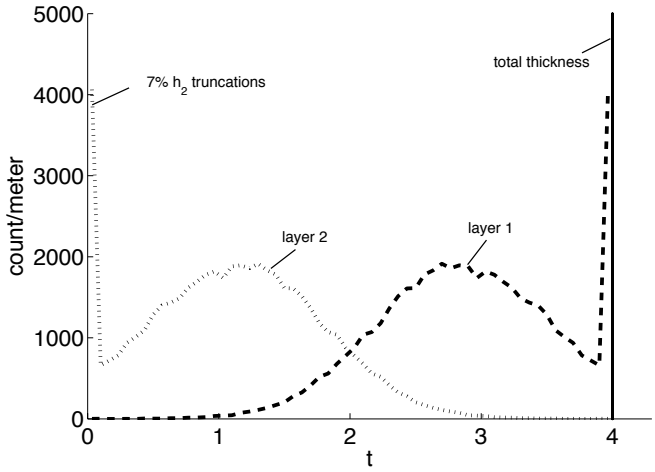


Figure 3.6: Thickness distribution for two layers. The prior and constraint are such that 7 percent of the realizations have $h_2 = 0$ ($t_2 < 0$) which creates the spike for layer 1 at $h_1 = 4$ m to yield $H = 4$ m.

Table 3.1: Parameters and results for 2-layer simulation

Layer	Prior				Posterior			
	$\bar{t}$	σ_t	$\bar{\phi}$	σ_ϕ	$\bar{h}$	$\boldsymbol{C}_t$	$\bar{\phi}$	$\boldsymbol{C}_\phi$
1	3	1.	0.20	0.05	2.89	$\begin{pmatrix} 0.44 & -0.48 \\ -0.48 & 0.55 \end{pmatrix}$	0.23	$\begin{pmatrix} 0.0008 & 0.0 \\ 0.0 & 0.0021 \end{pmatrix}$
2	1		0.30		1.11		0.31	

3.4 Simulations of Two-Layer Systems

We now 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 $\Phi H_s = 1.0$ m; the layer porosity prior means are 0.2 and 0.3, with a standard deviation of 0.05 (**Table 3.1**). We will 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. 3.6**). 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 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 exactly for every realization (**Fig. 3.7**). The univariate histograms of t are nonGaussian, due to the constraints. The thickness constraint induces a nearly perfect negative correlation in the layer thicknesses (Table 3.1; $\rho_{12} = C_{12}/\sqrt{C_{11}C_{22}} = -0.98$); the departure from -1 occurs because a small fraction of the realizations are not found 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.

In contrast, the posterior means of the two layer porosities $(0.23, 0.31)$ are greater than the prior means $(0.20, 0.30)$; **Fig. 3.8**, Table 3.1. The posterior must shift because the sum of the prior porosity thickness, $\phi_1 t_1 + \phi_2 t_2 = 0.2 \times 3 + 0.3 \times 1 = 0.9$ m, is less than the seismic porosity-thickness constraint, $\Phi H_s = 1$ m. This shows that priors need not match

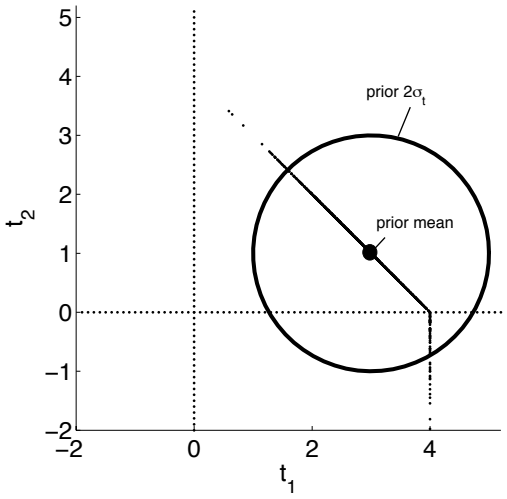


Figure 3.7: Simulations results for a two layer case with thickness priors $\bar{\mathbf{t}} = (3\text{ m}, 1\text{ m}), \sigma_t = 1$ m, and $H = 4$ m. All the realizations are exactly on the constraint surface.

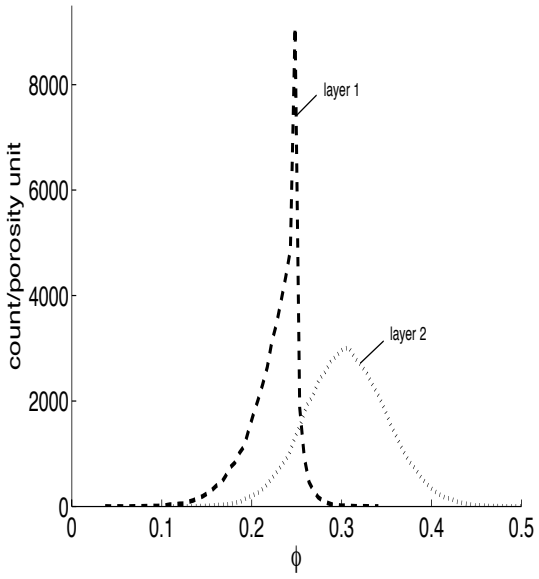


Figure 3.8: Porosity distributions for two layers. The constraints Φ and H are such that realizations having $\phi_2 > 0.25$ are not very probable and this skews the curve steeper on the right side.

the constraint exactly, and posterior reconciles the prior and likelihood. The distribution of ϕ_1 and ϕ_2 (**Fig. 3.9**) also reveals the constraint information: the porosity of both layers cannot simultaneously be higher than or less than 0.25; this gives a striking inclined hourglass shape to the posterior (Fig. 3.9). All simulated $\phi_1 t_1$ and $\phi_2 t_2$ exactly satisfy the ΦH constraint (**Fig. 3.10**).

The interactions of constraints and priors can be quite complex (**Figs. 3.11**). All realizations in this crossplot of t_1 and ϕ_1 are constrained by the condition $\phi_1 t_1 < 1$ (or $\phi_1 = 1/t_1$, which is the hyperbola at the upper right of the figure)since $\Phi H_s = 1$. As the thickness of a layer increases, the porosity of the layer converges to $\phi = \Phi H_s / t_1$, which is 0.25 for $t_1 = 4$.

These simple two layer cases demonstrate several features of the procedure. First, constraints are indeed honored exactly and the truncation behavior is captured – that is, pinchout configuration varies between realizations. The nonlinearity caused by truncation and product constraints (*viz.*, ϕh) give rise to particular correlations and clearly non–Gaussian

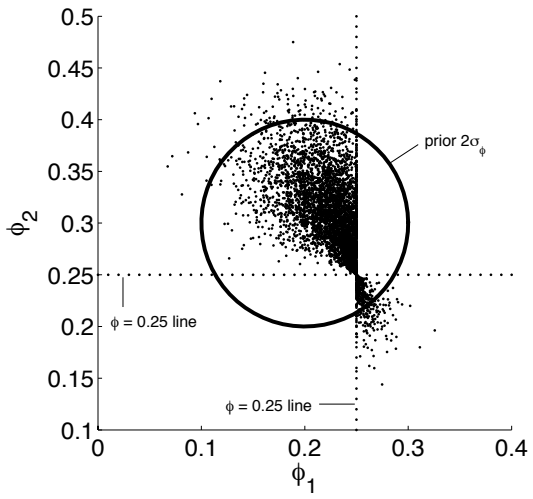


Figure 3.9: Simulation results of porosity for two layers. Because of the constraints the porosity of both the layers cannot be greater or less than 0.25 simultaneously. The prior for porosity is $\bar{\boldsymbol{\phi}} = (0.2, 0.3), \sigma_\phi = 0.05$.

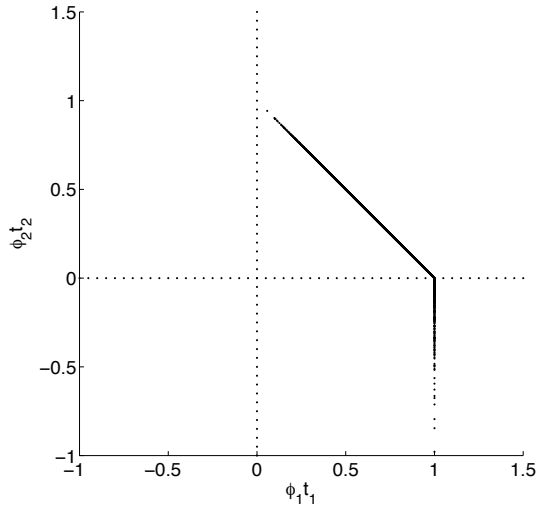


Figure 3.10: Simulations results for a two layer case with constraint $\Phi H = 1(\Phi = 0.25)$ m. All the realizations are exactly on the constraint surface.

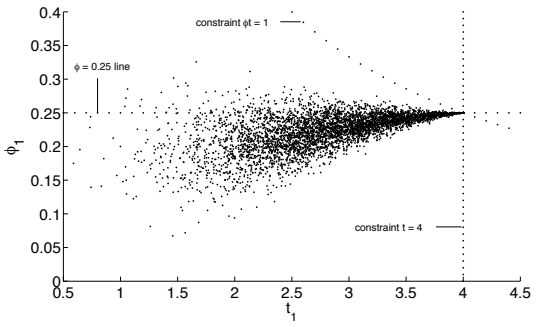


Figure 3.11: Cross plot for the distribution of porosity and thickness for layer one. As layer thickness reaches the total thickness constraint $h_1 = H = 4$ m, the porosity converges to average porosity and $\phi_1 = \Phi = 0.25$.

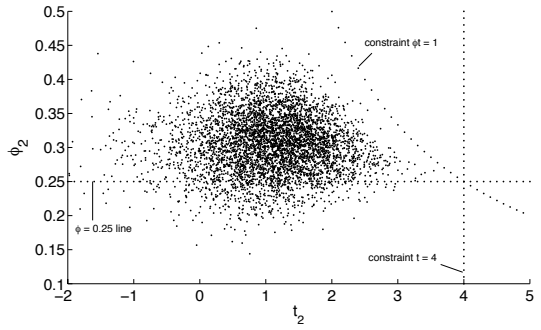


Figure 3.12: Cross plot for the distribution of porosity and thickness for layer two. As layer thickness reaches the total thickness constraint $h_2 = H = 4$ m, the layer porosity must converge to average porosity and $\phi_2 = \Phi = 0.25$.

Table 3.2: Design of 3D flow simulations

	Sands			Shales		
Factor	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

posteriors. These insights and verifications allow us to move on to a three-dimensional flow model.

3.5 Synthetic Three-Dimensional Examples

Synthetic 3D cases test and illustrate the MCMC simulation method.

3.5.1 Geomodel Construction

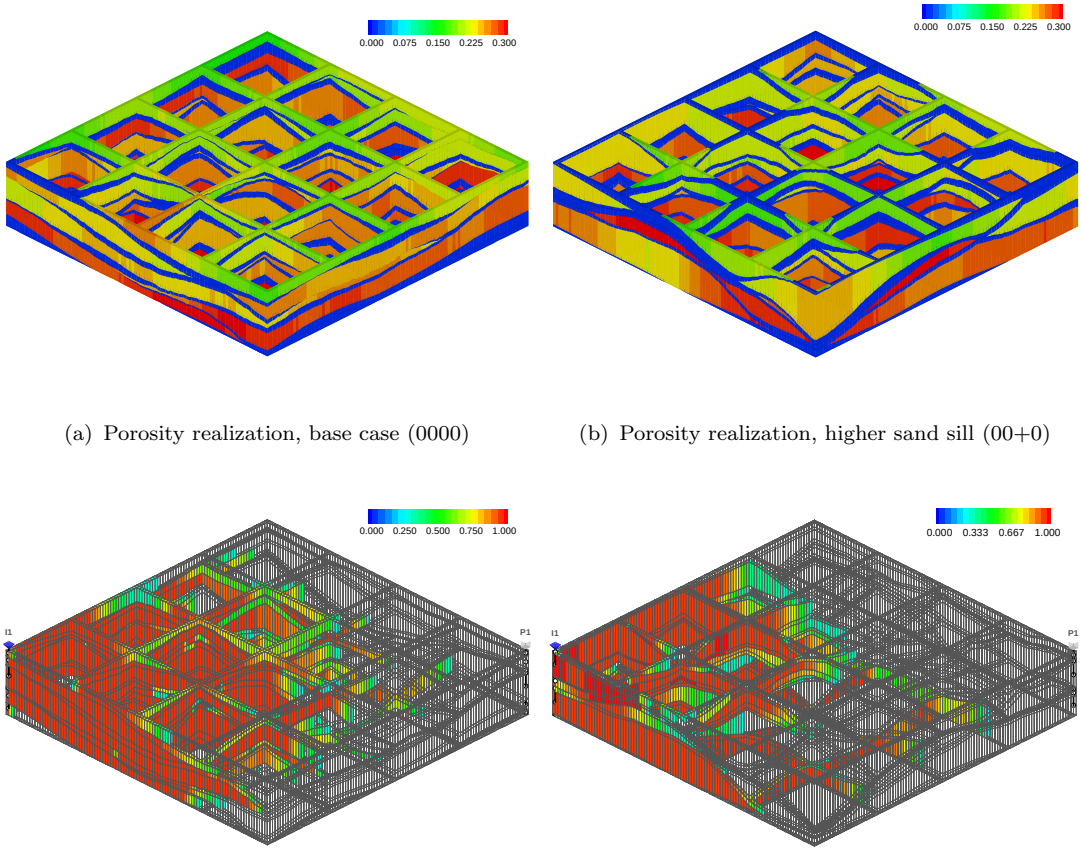
The prior distribution is varied by changing the range (b) and sill (σ^2) of the Gaussian semivariograms,

$$\gamma(\Delta x) = \sigma^2 \left[1 - \exp\left(-\frac{(3\Delta x)^2}{b^2}\right)\right] \tag{3.11}$$

where Δx is the lag.

Trends in H-related parameters are varied to illustrate behavior. The simulations are conditioned on four traces which have $h_{s(k)} = H_s/K_s \forall k \in \{1 \dots K_s\}$ for sands and $h_{sh(k)} = H_{sh}/K_{sh} \forall k \in \{1 \dots K_{sh}\}$ for shales (**Table 3.2**). 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. Seismic is informative at the scale of H_s and H_{sh}, but not for the layer thicknesses h_s or h_{sh}. The porosity prior is a variogram range of 500 m and variogram standard deviation of 0.025 for all cases. The permeability is assigned using an exponential transform of porosity,

$$k = 20 \exp(10\phi) \tag{3.12}$$



(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 3.13: Simulation on $100\times100\times10$ 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.

We specify alternating shales and sands, $K_s = K_{sh} = 5$. The system comprises $100\times100\times5 = 5 \times 10^4$ mesoscale constraints for each of the three variable:, t_s, t_{sh}, and ϕ_s (150,000 total).

The correlation range and sill are varied for the sandstone and shale layer thicknesses (Table 3.2). Increasing the range gives larger average dimensions to the geobodies, whereas increasing the variance (at constant mean thickness) makes the geobodies vary more in thickness and be more likely to truncate (*i.e.,* have zero thickness or "pinch out") 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 3.3**; **Fig. 3.13** a,b).

Table 3.3: Results of 3D flow simulations

Run	ν_{sh}	λ_{sh}	ν_s	λ_s	N_{pD}
0 (base)	0	0	0	0	0.564^a
17 (high sill)	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, $\sigma = 0.061$

The high-sill case has greater variability in sand thickness; there will be additional truncated beds; to fulfill the thickness constraints, there must also be an increase in thick beds. The high-sill case bedding geometry is thus more "come-and-go" and less "layer-cake."

The realizations displayed in Fig. 3.13 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. Further, near-zero thicknesses do not appear in isolated areas because of the truncation rules and the smooth Gaussian variogram. Finally, all three constraints are fulfilled at every trace.

3.5.2 Three-Dimensional Flow Modeling and Analysis

The effects of varying stratigraphy through the proposed downscaling method are investigated by means of flow modeling. All models use the same seismic constraint data; here, constant *mesoscale* values of thickness and porosity over the flow domain are found as discussed above. 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.

Four factors are considered using a two-level full factorial experimental design, with six replicates of the base case (0000, Table 3.3) to investigate stochastic fluctuation and six replicates of the high sand sill case to reconcile fluctuations and effects (00+0, Table 3.3; 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 patterns were infinite). The only response analyzed is recovery, *i.e.*, when the produced normalized tracer concentration exceeds 10 percent, referred to as N_{pD}.

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 various priors allow quite diverse models, so this variability is unsurprising (Table 3.3). The flow behavior for distinct prior models appears quite different, as expected (Fig. 3.13c,d). The greater variability in thickness and more frequent termination in the high sand sill case increase tortuosity and cause lower recovery.

Changing the prior can have a significant effect on responses such as recovery. We examine the effects of prior variability *versus* stochastic fluctuation when using a pair of points. Sets of six replicates were examined for the base case (0000; design center) and for a higher sand sill value (00+0; design face center). For these sets, a Welch two-sample t-test (R

Development Core Team 2007) indicates that the mean responses for the different priors are not the same ($t = 1.88$; for 95 percent confidence, the critical value $t_c = 0.12$). Therefore, the means are different, and the specification of the prior has a significant effect compared with stochastic fluctuations. However, if the factorial is analyzed using a linear model, the variance captured by the model does not dominate the residuals (variance ratio $F = 3.2$, with 15 and 5 degrees of freedom, not significant at 95 percent confidence). Thus, stochastic fluctuations or nonlinear effects are not negligible when compared to the linear trend.

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. Thus, it is important to use prior models — here, variogram ranges and sills — that reflect actual reservoir variability. These might be inferred from modern systems, outcrop exposures, or mechanistic depositional models.

3.6 Discussion

3.6.1 Challenges in Field Applications

One assumption in the proposed approach is that total number of layers (K) are known and correlations of these layers among the wells are also known. Deterministic prior knowledge on layer stratification may not be available, requiring a stochastic approach to layer identification and correlation. Such an approach would be very useful for field applications.

In addition, the wireline log and data at wells are commonly of a higher resolution than the flow-model layer scale. Thus, these data must be upscaled to the layer scale (*circa* 1 m) before downscaling the seismic data. Fortunately, the data used for seismic constraints – porosity and thickness – are straightforward to upscale. After the downscaling step, the flow model can be infilled with permeability and other properties (which may be upscaled through use of existing methods; *e.g.*, Li, Beckner, and Kumar (1999))

If the inversion neglects lateral correlations (*e.g.*, DELIVERY), the trace data should be processed to approximately reconstruct the lateral correlation expected in the reservoir.

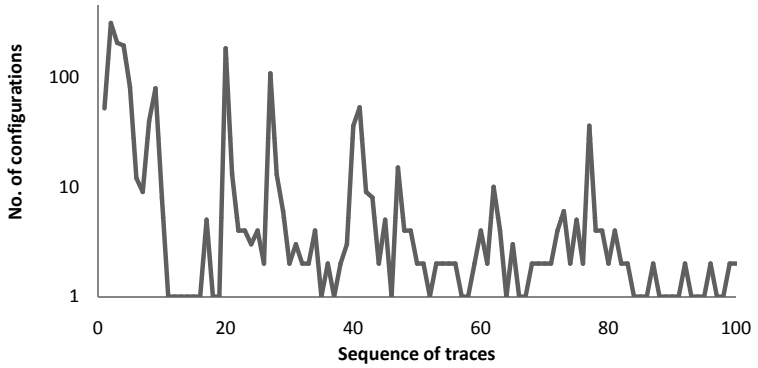


Figure 3.14: log(configurations) occurring at first 100 traces in the sequence, while simulating thickness for the base case (0000). Maximum number of configurations for 10 layers is $2^{10} - 1 = 1023$

The smoothing process must preserve interproperty correlations at each trace, as well (*e.g.*, DeliveryMassager).

If faults deform the reservoir, the modeling team must specify whether thickness and porosity are correlated across faults. If they are not correlated downscaling has to be done separately in fault block; this would increase bookkeeping, but is computationally trivial.

The presented results did not consider fluid saturation and its effects of seismic data. A seismic scale constraint on saturation could be added using a form like

$$\Phi H_s \bar{S}_o = \sum_{k=1}^{K_s} h_{sk} \phi_k S_{ok}$$

The sequential method used to compute ϕ after h_s could be used to compute S_{ok}.

3.6.2 Flow Model Diversity

Stochastic flow models may fail to capture the full range of uncertainty. This can lead to an underestimation of risk, or may cause ensemble inversion methods to fail (Gu 2006). In both cases, a more diverse set of flow models will mitigate the problem.

One way of assessing model diversity is to examine how many distinct layering configurations occur in a model. In other words, at each trace, the number of permutations of truncated and present layers may be counted. For the ten–layer case in the examples, there

are $2^{10} - 1 = 1023$ possible configurations at each of the 10^4 traces (over 10^7 possible stacking patterns, in principle). The count for the base case (0000) (**Fig. 3.14**) shows that no trace is nearly that diverse. This is expected, because the prior information from wells and previously simulated data restrict the range of *a priori* probable configurations. This effect is clear, because traces that are simulated later in the sequential process generally have fewer configurations.

However in a single realization, 13,909 configurations were simulated, with the most fecund trace totaling 311 (occurred at 2^{nd} trace) alternative configurations. This wide range of stratigraphic behavior was attained, even though there were four conditioning wells and 10,000 mesoscale constraints.

Flow responses also provide evidence of model diversity. The flow models (Synthetic Three-Dimensional Examples, earlier) are highly constrained (10,000 mesoscale constraints on each of H_s, H_{sh}, and ΦH_s; plus four wells with fine–scale constraints on all layer h and ϕ). Nonetheless, the flow responses are very diverse, with N_{pD} ranging from 0.39 to 0.65 – with all meso– and fine–scale data honored *exactly* and remain identical for all models. It appears that allowing thickness variations and layer termination tends to increase model diversity, compared with methods that only simulate intensive properties like k and ϕ.

Model diversity might be further increased by treating the layer labeling or correlation stochastically (Challenges in Field Applications, earlier).

3.6.3 Convergence of the Exact MCMC Simulation

MCMC methods may converge too slowly to be practical, or may have multiple modes such that multiple chains and/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 the proposed algorithm.

Convergence is critiqued by examining posterior distribution statistics over many iterations (Gelman et al. 2003). For the 2D problem, the means converge in no more than $\approx$ 100

Table 3.4: 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.

iterations, and the variances stabilize in no more than ≈ 200 iterations. That is, some 200 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, and the stabilized portion of the chain is used to simulate the posterior. This study used 5000 iterations to ensure adequate sampling; this is not prohibitive, if the proposal method is computationally inexpensive and the acceptance rate is not too small. For a realistic 3D synthetic problem, the proposed method attains a sampling rate of almost 100,000 iterations per second and an acceptance rate averaging ≈ 0.4, which makes such long chains manageable. Convergence is improved by starting MCMC sampling from the posterior mean, which can be estimated in these problems.

These 3D tests also show good convergence (typically, ≈ 1000 iterations).

3.6.4 Performance of the Exact MCMC Simulations

A large step size in the MCMC proposals allows rapid exploration of the posterior. On the other hand, large steps are more likely to rejected, causing wasted computations on a sample that is not accepted. A good compromise appears to be to scale the covariance to yield acceptance rates of about 30 percent (Gelman et al. 2003).

The computational cost of a single simulation is examined component–by–component (**Table 3.4**). 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. The kriging system is solved only three times per trace, and is two–dimensional with an efficient k-d neighbor search (Bentley 1975), causing the associated work to be small, about 0.5 percent. Yet the overall cost of eight minutes for 150,000 variables on a laptop computer does not seem prohibitive. Since the algorithm is trace–wise sequential, the overall-cost scales linearly in the number of traces; this extrapolates to 10^6 variables in about fifty five minutes on a laptop.

Even if the prior variances in each constraint are not constant (Assumptions on Prior Covariance, earlier), the problem is still manageable, especially with little change in efficiency related to kriging work. For the 10-layer case discussed if the prior variances are not constant, the extra work will be approximately five percent of the total work and in computation time, around 20 seconds. Most of the extra work in this case involves estimating covariance corresponding to the rotated subspace and then multiplying Cholesky factor $\tilde{\mathbf{L}}_r$ into $\mathbf{a}$ for each proposal (Generating and Accepting Proposals, earlier.). This work required scales according to $O(K^2)$ (where K is the layer count, not the block count). We estimate this would add about 40 percent to computation time, and three minutes for the 3D problem discussed above.

Chapter 4

Sequential Sampling and Marginalization in Seismic Downscaling

In the last two chapters, the idea is to incorporate seismic data in mesoscale Bayesian seismic inversions that treat subseismic heterogeneity *via* effective–media theory; subsequently, we downscale these inversions to meterscale models by using constraint equations to embody the effective media laws (Gunning, Glinsky, and White 2007). The proposed approach models layer thicknesses as "marked–surfaces", with truncations of negative thicknesses to allow pinchouts. For example, a set of K sublayers (thicknesses $t_k, k \in \{1 \dots K\}$), drawn from a meso-scale layer of total interval thickness H, implies the downscaling constraint for each column

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

The constraint is embedded in a likelihood expression to yield a Bayesian posterior

$$\pi\left(\mathbf{t}|H, \mathbf{d}\right) \propto L\left(H|\mathbf{t}, \mathbf{d}\right) p\left(\mathbf{t}|\mathbf{d}\right) \tag{4.1}$$

where $\mathbf{t}$ is a vector of thicknesses, H is the "target thickness", and $\mathbf{d}$ any additional hard data. The posterior distribution for the fine scale parameters is generally high dimensional, so we seek a sequential simulation algorithm by passing over all columns of the grid. Each column is simulated by sampling from a Bayesian posterior distribution, conditional on hard data and previously visited columns *via* the priors, and collocated coarse scale constraints *via* the likelihood. A suitable likelihood, with "accuracy" σ_H for K layers at a column; with expected total net-sand thickness, H is

$$L(H|\mathbf{t}, \mathbf{d}) \propto \exp\left[-\left[\left(\sum_{k=1}^{K} \max(0, t_k)\right) - H\right]^2 / 2\sigma_H^2\right] \tag{4.2}$$

Equation 4.2 corresponds to the inexact constraints discussed in Chapter 2. Prior distribution for the K layers is determined by kriging surrounding layer thickness (using data and previous

simulations); the distributions are $\mathbf{t} \sim N(\bar{\mathbf{t}}, \mathbf{C}_p)$, where $\bar{\mathbf{t}}$ and $\mathbf{C}_p$ are the kriged estimates and errors, respectively. A local, linearised posterior covariance derived from (4.1) is

$$\tilde{\mathbf{C}} = (\mathbf{C}_p^{-1} + \mathbf{X}\mathbf{X}^T/\sigma_H^2)^{-1} \tag{4.3}$$

where $\mathbf{X}$ is a design matrix comprising 1's if a layer k is present ($t_k > 0$) and zero otherwise; $\mathbf{X}$ depends on $\mathbf{t}$. This nonlinearity makes the posterior a piece-wise Gaussian, which is difficult to sample.

4.1 Linear Theory

If we partition the model vector $\mathbf{t}$ into I parts ($\mathbf{t}_1\ \mathbf{t}_2\ \ldots\ \mathbf{t}_I$), distribution $\pi(\mathbf{t})$ may be simulated by decomposition

$$\pi(\mathbf{t}) = \pi(\mathbf{t}_1)\pi(\mathbf{t_2}|\mathbf{t_1}) \ldots \pi(\mathbf{t}_I|\mathbf{t}_1 \ldots \mathbf{t}_{I-1})$$

which is the basis for sequential simulation. To simulate $\pi(\mathbf{t})$ sample from $\pi(\mathbf{t}_1), \pi(\mathbf{t}_2|\mathbf{t}_2), \ldots$ and finally from $\pi(\mathbf{t}_I|\mathbf{t}_1 \ldots \mathbf{t}_{I-1})$, $\pi(\mathbf{t}_i|\mathbf{t}_1 \ldots \mathbf{t}_{i-1})$ is the marginal distribution at location i, conditional on already simulated points from 1 to $i-1$ (Fig. 4.1). This marginal distribution may be obtained by integrating $\pi(\mathbf{t})$ over unvisited sites from $i+1$ to I

$$\pi(\mathbf{t}_i|\mathbf{t}_1 \ldots \mathbf{t}_{i-1}) = \int_{-\infty}^{\infty} \pi\left(\mathbf{t}\right) \mathrm{d}\mathbf{t}_{i+1} \ldots \mathrm{d}\mathbf{t}_I$$

In our case, the posterior is linked to prior, and in Eq. 4.1, likelihood is a simple product over all traces; therefore, posterior conditional on $\mathbf{d}$ and $\mathbf{H}$ is

$$\pi(\mathbf{t}_i|\mathbf{t}_1 \ldots \mathbf{t}_{i-1}, \mathbf{d}, \mathbf{H}) \propto \int_{-\infty}^{\infty} \prod_{j=1}^{I} L\left(\mathbf{H}_j|\mathbf{t}_j, \mathbf{d}\right) p\left(\mathbf{t}|\mathbf{d}\right) \mathrm{d}\mathbf{t}_{i+1} \ldots \mathrm{d}\mathbf{t}_I \tag{4.4}$$

Eqn. (4.4) is integrated over all $\mathbf{t}_j, j \in \{(i+1) \ldots I\}$.

Assume the prior $\mathbf{t} \sim N(\boldsymbol{\mu}, \mathbf{C})$

$$p(\mathbf{t}|\mathbf{d}) = \frac{1}{(2\pi)^{\frac{K}{2}}|\mathbf{C}|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(\mathbf{t}-\boldsymbol{\mu})^T\mathbf{C}^{-1}(\mathbf{t}-\boldsymbol{\mu})\right]$$

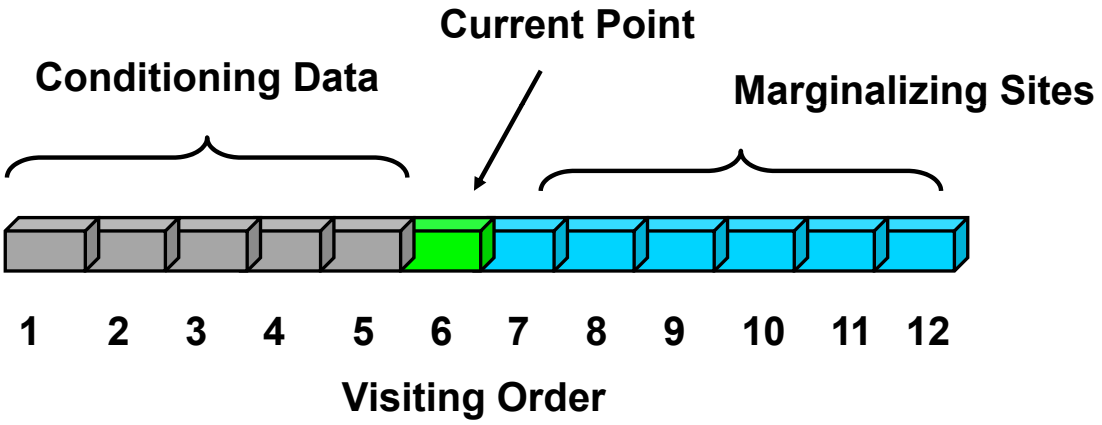


Figure 4.1: Conditioning and unsimulated traces when simulating at a trace $i = 6$ and for $I = 12$.

Consider the simple partitioning of $\mathbf{t}$ into current ($\mathbf{t}_1$) and "unvisited" ($\mathbf{t}_2$) sites, $\mathbf{t} = (\mathbf{t}_1\ \mathbf{t}_2)$. This partitioning makes the prior look like

$$p(\mathbf{t}|\mathbf{d}) \propto \exp\left[-\frac{1}{2}\begin{pmatrix}\mathbf{t}_1-\boldsymbol{\mu}_1\\ \mathbf{t}_2-\boldsymbol{\mu}_2\end{pmatrix}^T\begin{pmatrix}\mathbf{C}_{11} & \mathbf{C}_{12}\\ \mathbf{C}_{21} & \mathbf{C}_{22}\end{pmatrix}^{-1}\begin{pmatrix}\mathbf{t}_1-\boldsymbol{\mu}_1\\ \mathbf{t}_2-\boldsymbol{\mu}_2\end{pmatrix}\right]$$

Using the notation $\boldsymbol{\Sigma} = \mathbf{C}^{-1}$ gives

$$\begin{pmatrix}\boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12}\\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22}\end{pmatrix}\begin{pmatrix}\mathbf{C}_{11} & \mathbf{C}_{12}\\ \mathbf{C}_{21} & \mathbf{C}_{22}\end{pmatrix} = \begin{pmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{pmatrix} \tag{4.5}$$

Assume we have observations $\mathbf{H} = (\mathbf{H}_1, \mathbf{H}_2)$ from a linear observation process and error model $\mathbf{H} - \mathbf{X}\mathbf{t} \sim N(0, \mathbf{C}_H)$. Likelihood is proportional to

$$L(\mathbf{H}_1, \mathbf{H}_2|\mathbf{t}) \propto \exp\left[-\frac{1}{2}\begin{pmatrix}\mathbf{X}_1\mathbf{t}_1-\mathbf{H}_1\\ \mathbf{X}_2\mathbf{t}_2-\mathbf{H}_2\end{pmatrix}^T\begin{pmatrix}\mathbf{C}_{H_1}^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{C}_{H_2}^{-1}\end{pmatrix}\begin{pmatrix}\mathbf{X}_1\mathbf{t}_1-\mathbf{H}_1\\ \mathbf{X}_2\mathbf{t}_2-\mathbf{H}_2\end{pmatrix}\right]$$

The covariance $\mathbf{C}_H^{-1}$ is usually diagonal; without loss of generality, it may be factored into $\mathbf{X}$ and $\mathbf{H}$ (Kalla, White, and Gunning 2007). This form is apposite for linear observations or constraints that apply to distinct blocks of $\mathbf{t}$.

$$L(\mathbf{H}_1, \mathbf{H}_2|\mathbf{t}) \propto \exp\left[-\frac{1}{2}\begin{pmatrix}\mathbf{X}_1'\mathbf{t}_1-\mathbf{H}_1'\\ \mathbf{X}_2'\mathbf{t}_2-\mathbf{H}_2'\end{pmatrix}^T\begin{pmatrix}\mathbf{X}_1'\mathbf{t}_1-\mathbf{H}_1'\\ \mathbf{X}_2'\mathbf{t}_2-\mathbf{H}_2'\end{pmatrix}\right]$$

where $\mathbf{X}' = \mathbf{C}^{-1/2}\mathbf{X}$ and $\mathbf{H}' = \mathbf{C}^{-1/2}\mathbf{H}$.

The product of prior and likelihood is proportional to posterior, and the log of bayesian posterior is equal to

$$-2\log\pi(\mathbf{t}|\mathbf{d},\mathbf{H}) = \begin{pmatrix} \mathbf{t}_1 - \boldsymbol{\mu}_1 \\ \mathbf{t}_2 - \boldsymbol{\mu}_2 \end{pmatrix}^T \begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{t}_1 - \boldsymbol{\mu}_1 \\ \mathbf{t}_2 - \boldsymbol{\mu}_2 \end{pmatrix} + \begin{pmatrix} \mathbf{X}_1'\mathbf{t}_1 - \mathbf{H}_1' \\ \mathbf{X}_2'\mathbf{t}_2 - \mathbf{H}_2' \end{pmatrix}^T \begin{pmatrix} \mathbf{X}_1'\mathbf{t}_1 - \mathbf{H}_1' \\ \mathbf{X}_2'\mathbf{t}_2 - \mathbf{H}_2' \end{pmatrix} + \text{const} \tag{4.6}$$

If the log posterior can be written as a quadratic function with mean $\tilde{\boldsymbol{\mu}}$ and covariance $\tilde{\mathbf{C}}$

$$-2\log\pi(\mathbf{t}|\mathbf{d},\mathbf{H}) = \begin{pmatrix} \mathbf{t}_1 - \tilde{\boldsymbol{\mu}}_1 \\ \mathbf{t}_2 - \tilde{\boldsymbol{\mu}}_2 \end{pmatrix}^T \begin{pmatrix} \tilde{\mathbf{C}}_{11} & \tilde{\mathbf{C}}_{12} \\ \tilde{\mathbf{C}}_{21} & \tilde{\mathbf{C}}_{22} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{t}_1 - \tilde{\boldsymbol{\mu}}_1 \\ \mathbf{t}_2 - \tilde{\boldsymbol{\mu}}_2 \end{pmatrix} + \text{const} \tag{4.7}$$

then comparing the coefficients of $\mathbf{t}^T\mathbf{t}$ in equations 4.6 and 4.7 gives the inverse of covariance of the posterior $\tilde{\mathbf{C}}^{-1}$

$$\begin{pmatrix} \tilde{\mathbf{C}}_{11} & \tilde{\mathbf{C}}_{12} \\ \tilde{\mathbf{C}}_{21} & \tilde{\mathbf{C}}_{22} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{pmatrix}^{-1} + \begin{pmatrix} \mathbf{X}_1' & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_2' \end{pmatrix}^T \begin{pmatrix} \mathbf{X}_1' & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_2' \end{pmatrix}$$

Using equation 4.5 we can write the above equation as

$$\begin{pmatrix} \boldsymbol{\Sigma}_{11} + \mathbf{X}_1'^T\mathbf{X}_1' & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} + \mathbf{X}_2'^T\mathbf{X}_2' \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{C}}_{11} & \tilde{\mathbf{C}}_{12} \\ \tilde{\mathbf{C}_{21}} & \tilde{\mathbf{C}}_{22} \end{pmatrix} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \tag{4.8}$$

Then the marginal of $\mathbf{t}_1$, given $\mathbf{H}$ has covariance obtained by expanding above matrix and solving for $\tilde{\mathbf{C}}_{11}$, is

$$\tilde{\mathbf{C}}_{11} = (\boldsymbol{\Sigma}_{11} + \mathbf{X}_1'^{\mathrm{T}}\mathbf{X}_1' - \boldsymbol{\Sigma}_{12}(\boldsymbol{\Sigma}_{22} + \mathbf{X}_2'^{\mathrm{T}}\mathbf{X}_2')^{-1}\boldsymbol{\Sigma}_{21})^{-1} \tag{4.9}$$

By comparing coefficients of $\mathbf{t}$ in equations 4.6 and 4.7, we obtain the mean of the posterior, while expanding and solving the equation similar to Eq.(4.8) gives, the marginal mean as

$$\tilde{\mathbf{t}}_1 = \tilde{\mathbf{C}}_{11}\mathbf{X}_1'^{\mathrm{T}}(\mathbf{H}_1' - \mathbf{X}_1'\boldsymbol{\mu}_1) + \tilde{\mathbf{C}}_{11}\mathbf{X}_2'^{\mathrm{T}}(\mathbf{H}_2' - \mathbf{X}_2'\boldsymbol{\mu}_2) + \boldsymbol{\mu}_1 \tag{4.10}$$

In sequential simulation, the dimensionality of $\mathbf{t}_2$ (and rank of $\mathbf{C}_{22}$) can be very large. Important $\tilde{\mathbf{C}}_{11}$ (Eq.4.9), $\mathbf{\Sigma}_{22}$ and other matrices are obtained by inverting $\mathbf{C}_{22}$. Therefore we must approximate equations 4.9 and 4.10. One plausible approximation, based on a "weak correlations" (small $\mathbf{C}_{12}$) Taylor series expansion yields

$$\tilde{\mathbf{C}}_{11} = (\mathbf{C}_{11}^{-1} + \mathbf{X}_1^{'\mathrm{T}}\mathbf{X}_1^{'} + \mathbf{C}_{11}^{-1}\mathbf{C}_{21}^{\mathrm{T}}\mathbf{X}_2^{'\mathrm{T}}\underbrace{\mathbf{X}_2^{'}\mathbf{C}_{21}\mathbf{C}_{11}^{-1}}_{\mathbf{X}_{2,\mathrm{eff}}^{'}})^{-1} \tag{4.11}$$

and

$$\tilde{\mathbf{t}}_1 = \tilde{\mathbf{C}}_{11}(\mathbf{X}_1^{'\mathrm{T}}(\mathbf{H}_1^{'} - \mathbf{X}_1^{'}\boldsymbol{\mu}_1) + \mathbf{X}_{2,\mathrm{eff}}^{'\mathrm{T}}(\mathbf{H}_2^{'} - \mathbf{X}_2^{'}\boldsymbol{\mu}_2)) + \boldsymbol{\mu}_1. \tag{4.12}$$

This removes the need to invert a (potentially very large) $\mathbf{C}_{22}$ matrix block. Equation 4.12 is a standard Bayesian formula to update $\mathbf{t}_1$ given $\mathbf{H}_1^{'}$, with the contribution of secondary data $\mathbf{H}_2^{'}$, is attenuated by the modified sensitivity matrix $\mathbf{X}_{2,\mathrm{eff}}^{'}$. This is a manageable approximation for a marginal that includes the effect of information at unvisited sites.

4.2 Sampling within Nonlinear Constraints

The above equations are for linear constraints $f(\mathbf{t}) = \mathbf{X}\mathbf{t}$. For the nonlinear constraints $f(\mathbf{t}) = \mathbf{X}(\mathbf{t})\mathbf{t}$, as in the downscaling problem, additional approximations are needed to make the marginal for $\mathbf{t}_1$ tractable. The marginal of $\mathbf{t}_1$ is conditional on seismic information $\mathbf{H}$

$$\pi\left(\mathbf{t}_1|\mathbf{H}\right) \propto \int L\left(\mathbf{H}_1|\mathbf{t}_1\right) L\left(\mathbf{H}_2|\mathbf{t}_2\right) p\left(\mathbf{t}_1, \mathbf{t}_2\right) \mathrm{d}\mathbf{t}_2$$

If we neglect the nonlinearities in $\mathbf{t}_2$, then the marginal is analytically integrable, because it fits the linear theory as explained in the last section

$$\pi\left(\mathbf{t}_1|\mathbf{H}\right) \propto e^{-\frac{(f(\mathbf{t}_1)-\mathbf{H}_1)^2}{2\sigma_{\mathbf{H}_1}^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(\mathbf{X}_2\mathbf{t}_2-\mathbf{H}_2)^T\mathbf{C}_{\mathbf{H}_2}^{-1}(\mathbf{X}\mathbf{t}_2-\mathbf{H}_2)} p\left(\mathbf{t}_1, \mathbf{t}_2\right) \mathrm{d}\mathbf{t}_2 \tag{4.13}$$

which we call sequential simulation with marginalization (SM). A heavier approximation neglects lateral correlations between the current and unsimulated columns, $\boldsymbol{t}_1$ and $\boldsymbol{t}_2$, thus yielding sequential simulation without marginalization (SS),

$$\pi\left(\mathbf{t}_1|\mathbf{H}\right) \propto e^{-\frac{(f(\mathbf{t}_1)-\mathbf{H}_1)^2}{2\sigma_{\mathbf{H}_1}^2}} p\left(\mathbf{t}_1\right) \tag{4.14}$$

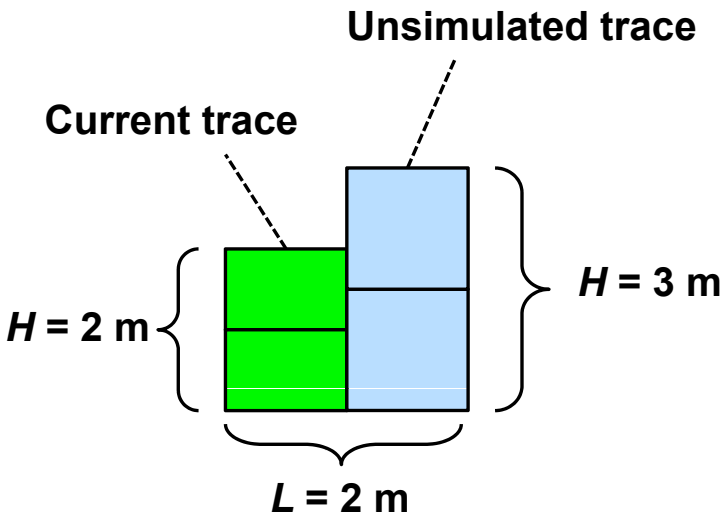
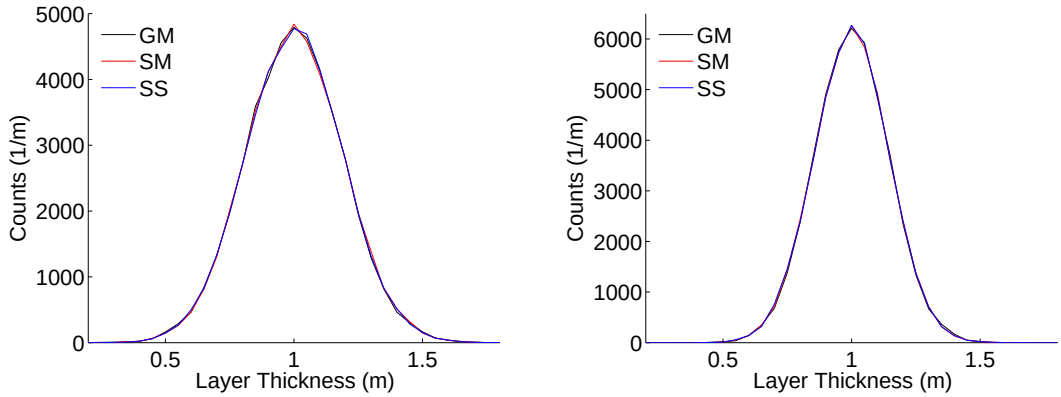


Figure 4.2: 2D example for comparing GM,SM, and SS methods.

4.3 2D Examples

A small 2 layer × 2 column 2D example compares SM and SS algorithms with a global MCMC method (GM). Columns have 1 m separation (Fig. 4.2). Two constrasting cases of constraint uncertainty (σ_H) and lateral correlation (range, λ_x) are considered. The seismic thickness constraint ($H = 2$ m) at trace one, $H = 3$ m at the second trace, and σ_H, are stationary. The prior means (for μ_1 in Eqn. (4.12)) are $\bar{t}_k = 1$ m $\forall k$; autocovariances are Gaussian with sill related to a stationary prior standard deviation of $\sigma_t = \bar{t}_k$. These parameters cause a low probability for layers to pinch out, in which case the SM method works well. The assumption in the derivation of SM method is that the system is linear. For these examples, auxiliary variables are used for sampling (Kalla et al. 2006).

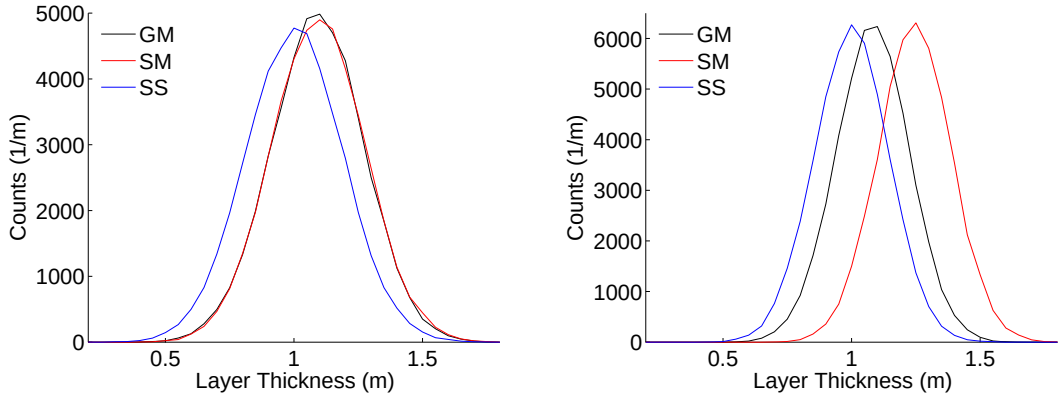
Weak geologic correlation ($\lambda_x = 0.25$ m). The marginals for the first column visited for the global, marginalized, and standard sequential methods [Fig.(4.3)] have only small differences. This shows that marginalization is not required, if lateral correlation is weak.



(a) Strong seismic data, weak correlation

(b) Weak seismic data, weak correlation

Figure 4.3: Global (GM), sequential marginalized (SM), and standard sequential (SS) simulation results for weak correlation cases. Results are for layer 1.



(a) Weak seismic data, strong correlation

(b) Strong seismic data, strong correlation

Figure 4.4: Global (GM), sequential marginalized (SM), and standard sequential (SS) simulation results for strong correlation cases. Results are for layer 1.

Weak seismic constraint and strong geologic correlation ($\sigma_H = 0.25$ m and $\lambda_x = 4$ m). The marginals for the global method differs [Fig.(4.4a)] from the standard sequential method. The approximate marginal method closely follows the rigorous MCMC result.

Strong seismic constraint and strong geologic correlation ($\sigma_H = 0.025$ m and $\lambda_x = 4$ m). The marginals for the global method are different than the SM and SS method [Fig.(4.4b)]. The SS method is found to be closer than the rigorous SM result. This poor result for the SM method is due to a seismic constraint in the second trace which is inconsistent with the

current trace. If there is a strong correlation, then both the constraints should be very close, *i.e.*, data is not consistent in this case.

Sampling the uncertainty in these nonlinear downscaling problems is difficult. Global MCMC methods are accurate but expensive, which motivates consideration of sequential methods. Cheaper sequential methods are reasonably accurate, provided the lateral correlation is not high, and the constraints are weak. If the correlation is high and constraints are strong, naïve sequential simulation poorly approximates the marginals. For such cases, the proposed approximate marginals offer improved sampling at a moderate cost.

4.4 Discussion

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 (Liu 2001).

Generating conditional distributions at the current trace in sequential simulation requires integrating information over "unvisited" sites in computing analytical marginal distributions and conditioning only on "visited" sites. The integrability requirement is usually met only through exponential family distribution Functions; yet the posteriors in our problem do not tend to follow exponential forms. Nonetheless, the approximations we make can be improved by blockwise sequential schemes, although a block approach not only increases the dimensionality of the MCMC sampling subproblem, but also exacerbates the configurational complexity of handling more pinchout transitions. Comparing methods with rigorous marginals to local approximations for several test problems, the local method errors increase as transverse correlation lengths increase. 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 simultaneously increase sampling dimensionality and the configurational complexity of handling more pinchout transitions.

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 result, a $1/\sqrt{n}$ reduction which occurs if all the traces are treated as independent measurements, may be over–optimistic. In practice, many systematic effects found in acquisition and processing tend to make the imaged seismic data less informative.

Chapter 5

Stratigraphic Stacking Patterns in Downscaling to Fine–Scale Flow Models

Subsurface models are poorly constrained, due to sparse sampling of the depositional heterogeneity by wells and low-resolution seismic data. In geostatistical modeling, geologic continuity information, such as correlation lengths, are included in kriging, sequential Gaussian simulation and their variants. This information comes from geologic analogues like outcrops, conceptual models, and process-based models. In the downscaling approach discussed, the integrating of stacking patterns as observed in analogues, should result in more realistic and more constrained models.

In clastics, analogues aid in modeling sand body distribution and continuity, which are controlled by depositional environment (Section 1.2). Clastic reservoirs may have continuous and/or discontinuous shales; their locations are governed by the geometry of the bedding surfaces along which they were deposited. Since these important bedding characteristics are understood from analogues, they must be included when building geologic models and are usually included as models in the form of priors.

5.1 Integrating Stacking Patterns

Generating the models and parameters for stacking patterns and integrating at the current trace-wise algorithm is challenging. For integration to be viable at each trace a "preprior" mean and variance for $t_k, \forall k \in [1, K]$ must be specified. This "preprior" gives the probability of a layer being present at a particular trace and provides a statistical, stratigraphic description of the entire reservoir. For example, in a prograding geologic system from left to right, the lowest layers on the left should have relatively low mean thicknesses, yet have a significant probability of pinching out. The mean thickness of layers then should increase in the middle and decrease in the upper layers as deposition moves rightward. As the proba-

bility of truncation is inversely related to the mean thickness, so probability of truncation decreases in the middle and increases in the upper layers. For Gaussian functions, any of the two parameters (mean, variance or probability of truncation) define a unique function. If each layer mean and variance are available at all points on a corner point grid, then this information may be integrated while estimating the prior by kriging. Preprior can be used as

- The mean and variance of preprior is used to estimate the simple kriging mean and variance while estimating prior (Section 5.3)
- The harmonic mean of variances and a variance-weighted mean of preprior and kriging can be used as prior variance and mean respectively

The new prior is used in exactly the same way as before in the downscaling algorithm. The only difference is that the preprior has stratigraphic detail, which is also included in the downscaling approach.

5.2 Preprior Using Surface Based Modeling

Surface based models (SBM) have been used to mimic stratigraphic architecture; ensembles of these are used to estimate the preprior in this case. Other methods can be used to generate the preprior but integrating the preprior information into the downscaling algorithm will be similar, as discussed in section 5.3. Surface-based models are a variant of object based models and are applied to different depositional settings. In this chapter, surface based models are used to generate compensational stacking of the distal part of a turbidite deposit (Pyrcz 2004). This is a sample problem to show how the method works, but this procedure would be similar for other depositional environments.

5.2.1 Generating Lobes Using Surfaces

For compensational stacking of turbidite lobes, the general procedure is to generate an initial surface with certain geometry and place it on the initial base topography. The body between the surface and base topography is a geobody defined as a lobe. The next surface is generated and stacked on the initial surface to obtain the second lobe, and the procedure continues. The bathymetry after inserting a surface depends on where the surface is inserted. If we have K lobes present in the reservoir, K surfaces are generated and stacked one after the other. The K surfaces and the bottom known bathymetry, creates the K lobes required.

This simple stacking can be utilized to generate realistic and complex geologic patterns but artifacts are easy to interfere. This is because physics does not driv the forward model or stacking but it is driven by a set of rules. Geobodies are created based on simple shapes and inserted in the model, depending on prior rules. For example, one rule in stacking could be not to stack a layer exactly on the top of a previous layer. A lot of work has gone forward generating realistic surface-based models; Pyrcz (2004) provides a good discussion. In this research, simple prograding stratigraphic models are built by using surface-based modeling; the method to integrate the model statistics using an ensemble is also discussed. Again, this very simple stacking algorithm is not advocated for geomodeling, but is only used (as other, more complex models could be used) to generate the preprior.

5.2.2 Lobe Geometry and Surface Positioning

Surface models generated here can be thought of as filling the 4^{th} order Bouma sequence detected by the seismic with 3^{rd} order lobes (Fig. 5.1). Initial bathymetry is needed, as explained earlier, for surface-based modeling; here, we fill the seismic volume with lobes, so that the bottom of a seismic zone becomes the base surface. In our synthetic case a simple flat surface is assumed as base surface and lobes are stacked to mimic a prograding system from left to right. Lobe geometry is a simple quadratic function, and it is used to estimate thickness of lobe as a function of location

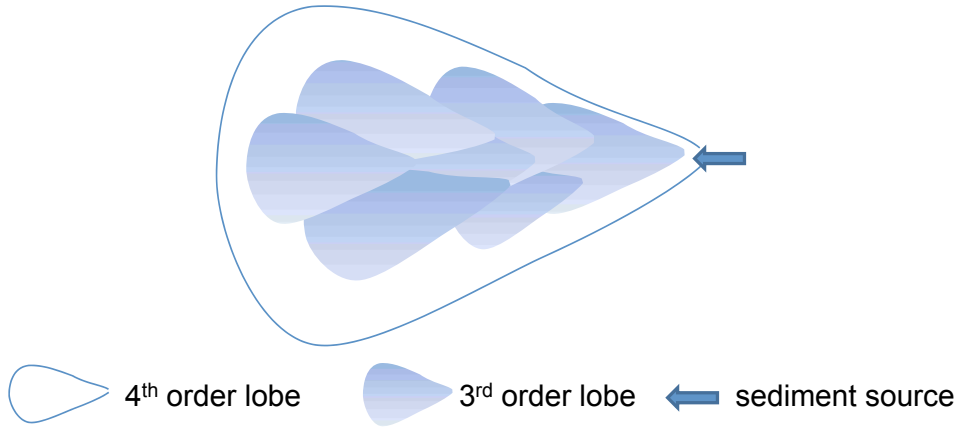


Figure 5.1: Filling of 4^{th} order lobes by the 3^{rd} order prograding lobes.

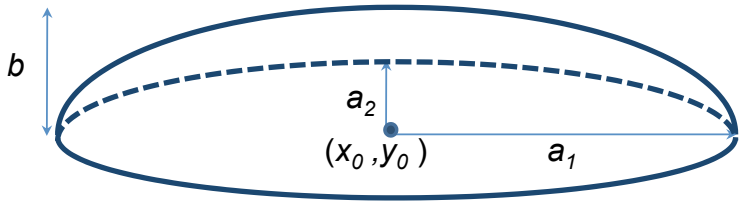


Figure 5.2: Surface template and its geometric parameters.

$$t(x,y) = b\left[1 - \left(\frac{x-x_0}{a_1}\right)^2 - \left(\frac{y-y_0}{a_2}\right)^2\right] \tag{5.1}$$

Constants a_1 and a_2 are radii of elliptical lobe from center x_0 and y_0. The constants are simulated from a normal distribution function, with a given mean and variance. Parameter b is the maximum height of the lobe, again generated by a random normal deviate with mean height and variance given as input. These input statistics usually come from analogue and well data. The surface then can be placed, depending on deposition, by maintaining inter-relationships between the lobes. This relation is controlled by placement of the center of the lobe x_0 and y_0. In this study the geologic system is prograding x_0 of a lobe, obtained by adding a scaled random deviate to the center of the last lobe x_0. y_0 is assumed to migrate from left to right as a normal deviate from centerline. Surface models are built in a larger area than required in order to remove the boundary effects.

5.2.3 Conditioning

Surface-based models are uncertain; conditioning one of these models to well and seismic without creating artifacts is difficult. In this study, an ensemble of these models are used to characterize surface-based model uncertainty. Ensemble means and variance of variables, such as thickness or porosity of all layers and at all traces, is used to condition SBM. These ensemble properties as estimated then could be used to build and to condition the geomodels as the geological uncertainty is estimated; more importantly, this removes the need to condition one specific surface-based realization. Conditioning the SBM to seismic and well data is equivalent to downscaling the seismic constraints while incorporating stratigraphic detail, which is the current problem.

5.3 Integrating Surface Based Models

One way to integrate the surface-based ensemble statistics is to use the mean and variance data available at all points on a corner point grid, while doing simple kriging. Some advantages on taking this path are given here. The terminology used in deriving the simple kriging apparatus is similar to Goovaerts (1997) work.

5.3.1 Using Ensemble Statistics in Kriging Apparatus

Kriging estimates the value of a continuous attribute Z at location $\mathbf{u}_i$ conditioning on n data at $\mathbf{u}_\alpha$ denoted as $Z(\mathbf{u}_\alpha), \alpha = 1, \ldots, n$. Kriging estimate $Z^*(\mathbf{u}_i)$, a linear regression estimate

$$Z^*(\mathbf{u}_i) - m(\mathbf{u}_i) = \sum_{\alpha=1}^{n} \lambda(\mathbf{u}_\alpha) \left[Z(\mathbf{u}_\alpha) - m(\mathbf{u}_\alpha)\right] = \boldsymbol{\lambda}^{\mathrm{T}}(\mathbf{u}_\alpha) \left[\mathbf{Z}(\mathbf{u}_\alpha) - \mathbf{m}(\mathbf{u}_\alpha)\right] \tag{5.2}$$

Where $\lambda(\mathbf{u}_\alpha)$ is the weight assigned to data point, $Z(\mathbf{u}_\alpha)$. $m(\mathbf{u}_i)$ and $m(\mathbf{u}_\alpha)$ are the expected values of random variables $Z(\mathbf{u}_i)$ and $Z(\mathbf{u}_\alpha)$. Residual, $R(\mathbf{u}) = Z(\mathbf{u}) - m(\mathbf{u})$ is modeled as a stationary random function with zero mean and covariance $C(h)$, where h is the distance between two locations. A auto-covariance function $C(h)$ is built, using the residuals. $\boldsymbol{\lambda}(\mathbf{u}_\alpha), \mathbf{Z}(\mathbf{u}_\alpha)$, and $\mathbf{m}(\mathbf{u}_\alpha)$ are column vectors with size n holding information about neigh-

boring data and their weights. Similarly, a column vector $\mathbf{m}$ and matrix $\mathbf{C}$ are defined with rank $n+1$ and partitioned as in Eq. 5.3. For simplicity, $\mathbf{u}$ is dropped from the equations. For example, $Z(\mathbf{u}_\alpha)$ is written as Z_α

$$\mathbf{Z} = \begin{pmatrix} Z_i \\ \mathbf{Z}_\alpha \end{pmatrix} \mathbf{m} = \begin{pmatrix} m_i \\ \mathbf{m}_\alpha \end{pmatrix} \mathbf{C} = \begin{pmatrix} C_{ii} & \mathbf{C}_{i\alpha} \\ \mathbf{C}_{\alpha i} & \mathbf{C}_{\alpha\alpha} \end{pmatrix} \tag{5.3}$$

$\mathbf{C}$ is the covariance matrix between the $n+1$ conditioning traces and the simulated trace. $\mathbf{C}_{\alpha\alpha}$ is the $n \times n$ covariance matrix between conditioning data; C_{ii} is the variance at the simulated trace, and $\mathbf{C}_{i\alpha} = \mathbf{C}_{\alpha i}^{\mathrm{T}}$ is the conditioning data to simulated data covariance matrix. $\mathbf{m}$ holds the mean values for current trace and conditioning traces, and $\mathbf{Z}$ holds the estimated values for current trace and conditioning traces.

From the theory of multivariate normal distributions,

$$p\left(\mathbf{Z}\right) = \frac{1}{(2\pi)^{(n+1)/2}|\mathbf{C}|^{1/2}} \exp\left\{-\frac{1}{2}\left(\mathbf{Z}-\mathbf{m}\right)^{\mathrm{T}} \mathbf{C}^{-1}\left(\mathbf{Z}-\mathbf{m}\right)\right\}$$

the distribution of Z_i conditional on $\mathbf{Z}_\alpha$ is multivariate normal function

$$p\left(Z_i|\mathbf{Z}_\alpha\right) \propto$$
$$\exp\left\{-\tfrac{1}{2}\left(Z_i - m_i - \mathbf{C}_{i\alpha}\mathbf{C}_{\alpha\alpha}^{-1}(\mathbf{Z}_\alpha - \mathbf{m}_\alpha)\right)^{\mathrm{T}} \mathbf{C}_{|\mathbf{C}_{\alpha\alpha}}^{-1}\left(Z_i - m_i - \mathbf{C}_{i\alpha}\mathbf{C}_{\alpha\alpha}^{-1}(\mathbf{Z}_\alpha - \mathbf{m}_\alpha)\right)\right\}$$

with mean

$$Z_i^* = m_i + \boldsymbol{\lambda}_\alpha\left[\mathbf{Z}_\alpha - \mathbf{m}_\alpha\right] \tag{5.4}$$

where $\boldsymbol{\lambda}_\alpha = \mathbf{C}_{i\alpha}\mathbf{C}_{\alpha\alpha}^{-1}$ and variance

$$\mathbf{C}_{|\mathbf{C}_{\alpha\alpha}} = \sigma_i^2 = C_{ii} - \boldsymbol{\lambda}_\alpha \mathbf{C}_{\alpha i} \tag{5.5}$$

The same results can also be obtained by linear regression (Eqn. 5.2). This derivation, rather than linear regression, gives a better understanding of what kriging is doing and what it means to the assumptions of the stationarity of mean and variance. In simple kriging estimates (Goovaerts 1997), stationarity of mean is assumed $[m_i = m_\alpha = m]$ or a trend in mean when kriging is applied with a trend. In our case, we do not hold stationarity or trend in

a local mean. This formulation is possible, because the means and variances data may be obtained from ensemble surface-based models at all traces, including the well locations. The variogram for the residuals is computed by using the well data. One of the issues with this modeling is labeling the layer numbers and identifying the well picks, as discussed in section 5.3.3.

5.3.2 Non-Stationarity and Negative Values

The covariance matrix **C** is built by using a residual variogram with the assumption of second order stationarity. That is, the residual variance will be constant everywhere in the domain. If residual variance at a trace is not equal everywhere, one could rescale the residual by using the standard deviation, then use SK or OK and correlograms to krige the scaled deviate and concept of a purely local variance with a global normalized deviate correlation. Covariance models for normalized residuals should be used.

However, if the layers pinch-out, the usual truncated Gaussian distributions of the layer properties, such as thickness, would arise at a given trace. Surface-based modeling gives only positive values. Therefore in our downscaling problem, negative values are required to get Gaussian function for all variables. The surfaced-based models can be built to obtain negative values by either

- Extrapolating Eq. 5.1 for the whole domain, which gives negative values wherever layers pinch out. If ensemble properties after extrapolation are approximately Gaussian, the mean and variance are used to estimate local distribution by using Eqs. 5.4 and 5.5.

- The ensemble of surface based models gives thickness and other property distributions which are truncated Gaussian. An estimate of Gaussian mean and variance are needed in Eqs. 5.4 and 5.5; these can be estimated from left truncated Gaussian parameters, since the ensemble provides the mean, variance, and number of samples truncated to the total number of samples, using methods explained in Cohen (1949).

Assumptions in this work are similar to SGS, in that surface based models are consistent with well data. The discrete well data must be a plausible member of the distributions obtained from surface models.

5.3.3 Choosing the Well Picks

The K layers present in our surface-based model are to be related with the layers that are observed at the wells. This step is required before the conditioning of layers to well and seismic data. At wells, elevations, rather than thicknesses, are used to choose the layers. This is because elevation of layer boundaries (or depth if measured if the reference is a top surface) are better approximated as Gaussian functions (except at the lower boundaries) when using surface-based models. Elevation of a boundary is simply the sum of all the thicknesses below the specific boundary, if traces are vertical. At a well trace using ensemble realizations, the mean $q_k, \forall k \in [1, K]$ and variance $\sigma^2_{qk}, \forall k \in [1, K]$ of all the boundary elevations are estimated. Let us assume that at the same well, the elevations of all the layers $e_z, \forall z \in [1, Z]$ are observed, where $0 \leq Z \leq K$ (Fig.5.3). These Z layer observations are to be mapped as K total layers in the reservoir. At a well identifying the well observation with surface-model ensemble elevations is like finding one out of many multivariate normal populations to which a new observation belongs.

Estimating

Assume there are K layers in the reservoir. Elevation for each layer k at a well trace is assumed Gaussian with probability density function $f_k(\cdot)$, where $k \in 1...K$. Then for each observation e_z at the well, where $z \in 1...Z$ (discrete picks of layer boundaries at wells), the likelihood of e_z being a member of each population is calculated and then the population with the largest likelihood is taken.

For example, let us assume we have a two-layer reservoir (K=2) but only one layer is observed at a well (Z=1). An ensemble of surface models will give the distribution of elevations

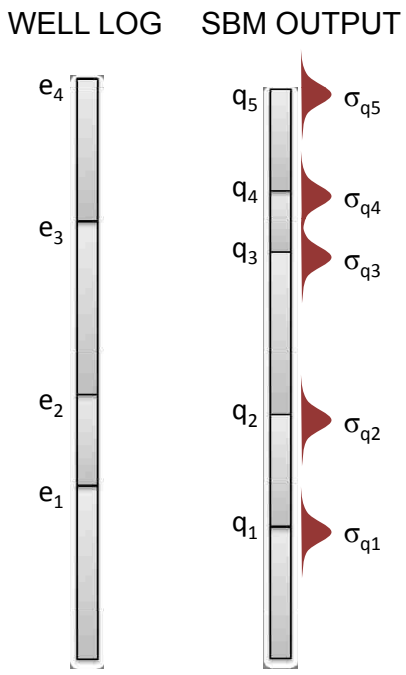


Figure 5.3: Mapping of well layers observed ($Z = 4$) to SBM layers ($K = 5$).

for both the layers $N(q_1, \sigma_{q1})$ and $N(q_2, \sigma_{q2})$. Let the discrete well elevation be e_1 (Fig. 5.4). If the likelihood of e_1 belonging to the first boundary ($L_1(e_1)$) is more plausible than the second boundary ($L_2(e_1)$), then e_1 is associated with the first layer boundary.

$$L_1(e_1) > L_2(e_1) \iff -(e_1 - q_1)^2/(2\sigma_{q1}^2) - \log(\sigma_{q1}) > -(e_1 - q_2)^2/(2\sigma_{q2}^2) - \log(\sigma_{q2})$$

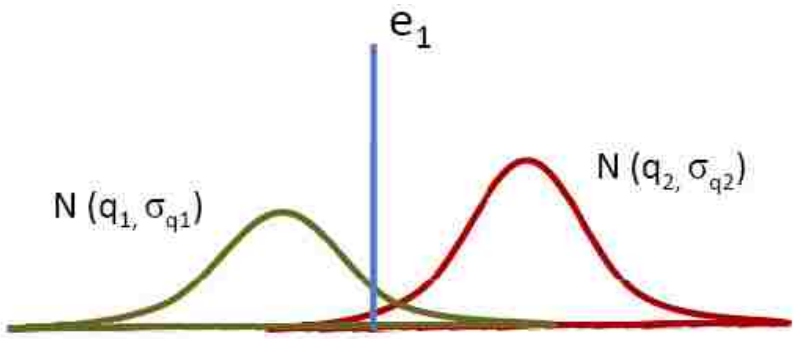


Figure 5.4: Assaigning the layer observed to either of 2 layers by likelihood.

Sampling

In the last section, estimating the most probable mapping of well observations to the K layers present in the reservoir is deterministic. Probabilistic mapping of the Z layers observed at a well to the K layers present in the model is discussed here. Let Y_k be a variable which takes value one, if the well observation e_z is associated with k^{th} boundary, and zero otherwise. The observation e_z is allocated to each group according to the probability of its belonging to the k^{th} group.

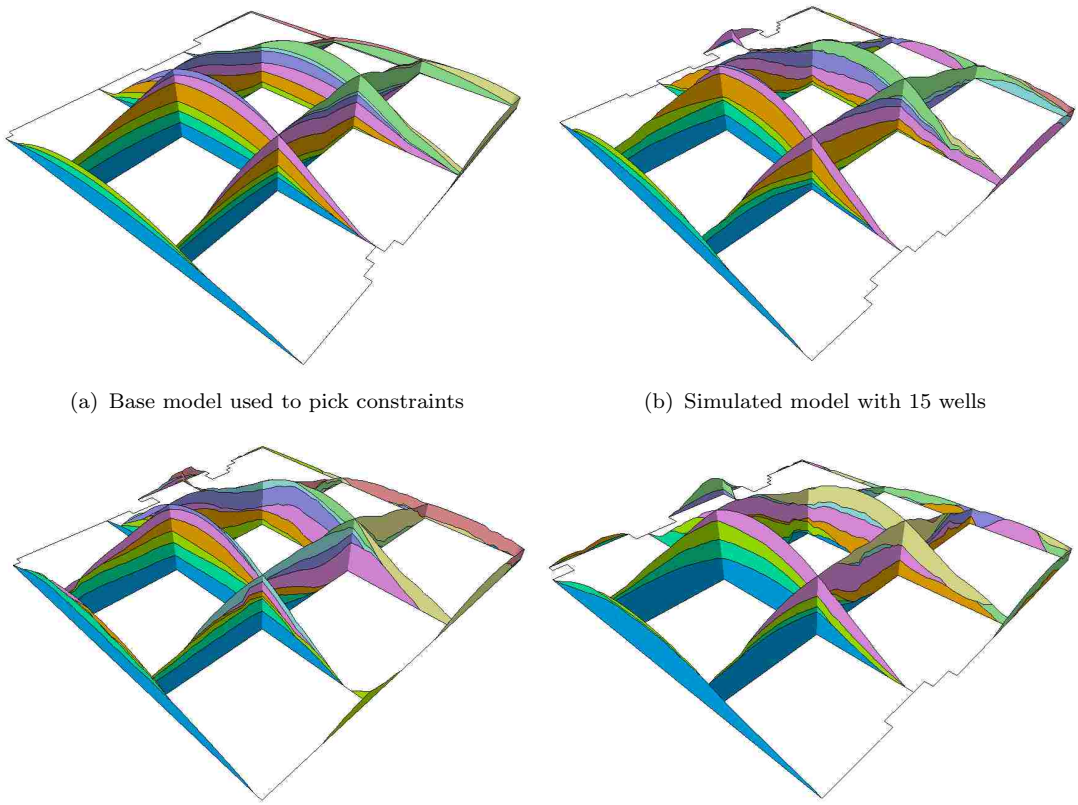
$$P(Y_k = 1|e_z) = \frac{f_k(e_z)}{\sum_{k=1}^{K} f_k(e_z)} \tag{5.6}$$

At a well trace, the first layer e_1 from the bottom could be associated with any layer $k, \forall k \in [1, K]$, with probability given by Eq. 5.6; let a sampled model layer mapped with e_1 be the i^{th} layer. Then the second layer at the well e_2 is assigned to model the layer between layers $i + 1$ and K by a sampling proportional to the probability of e_2 belonging to any of those layers. The procedure repeats until all the model layers are tagged with all the well layers.

If any of the layers at the well is known to correlate with the model layer, layers above and below that layer are tagged with the above procedure as if the known layer is the reference layer. It can be extended to any known number of layers.

5.4 Results

A simple code that generates stacking layers using distributions for length, height, and the center location of a lobe is used. A ten-layer model is generated, assuming a prograding system. One instance of 100 realizations from a stacker is taken as the real case, shown in Fig. 5.5(a). Fifteen wells and seismic constraints are taken from this model. Also, the average and variance of thicknesses for each layer are founded by running the stacker 100 times. The ensemble averages are used as means in the simple kriging. Scaled residuals are kriged and the variogram for scaled residuals is obtained at the wells, where well observations for all



(a) Base model used to pick constraints

(b) Simulated model with 15 wells

(c) Simulated model with 5 wells

(d) Simulated model with no wells.

Figure 5.5: Simulation on $50 \times 50 \times 10$ cornerpoint grids, areal extent is $X = Y = L = 1000$ m. Constraints used are well and seismic (exact). Also ensemble statistics of surface based models are used in simple kriging. Vertical exaggeration is twentyfold for all figures. The more the number of wells, the closer the model is to the base case, but all the constraints are satisfied in every model.

layers and ensemble statistics from 100 realizations are available. For all cases, the grid size is $I \times J \times K$ is $50 \times 50 \times 10$ and the $x - y$ extent is 1000×1000 m; $L = 1000$ m. The total thickness (H) is given by the seismic constraints.

Fig. 5.5(b) is generated using ensemble statistics, well data, and seismic data. Seismic data is integrated at each and every trace as exact constraints in using the algorithm found in Chapter 3. As a result, the total thicknesses of both the base and the simulated model are exactly the same at all the traces. The integrated model is very similar to the reference case (Fig. 5.5(a)). All the layers terminating in the reference case are very similar to the base

case. There is also very little noise in the simulated model, unlike the truncated Gaussian or indicator simulations.

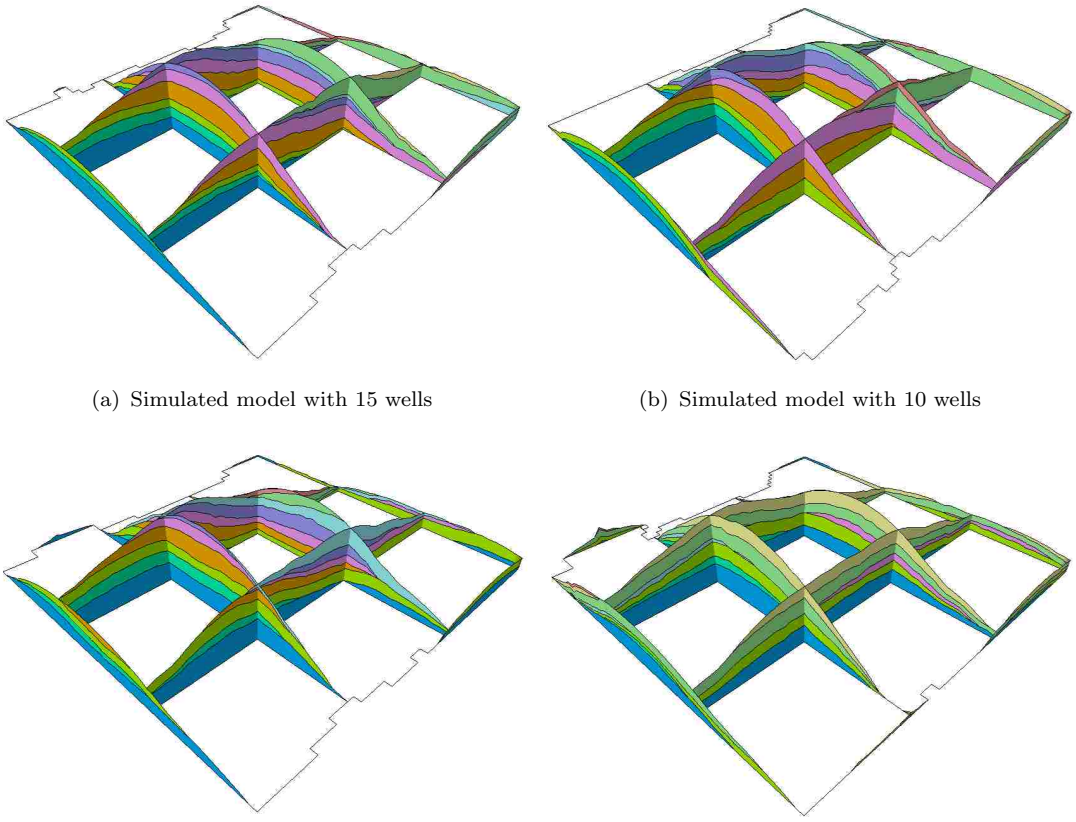
5.5 Discussion

5.5.1 Sensitivity to Well Data

Although based on a synthetic case, the data integration method demonstrates how to reproduce complex pinchout layering and to match seismic constraints exactly, which is required to condition surface-based models. In 15 well examples, an abundance of well control is used to show that this method can integrate many wells, which constitutes a major problem with most methods used to condition the surface-based models. The above procedure uses kriging and as such, the interpolation of added conditioning data is better. As a result, the simulated case matches the reference case very well. If there are less number of wells (Fig. 5.5(c,d)), due to less constraining data, various realizations would present differently yet all the seismic, well, and ensemble statistics would be honored. For instance, in Fig. 5.5(d), no well data is used yet the system is prograding and seismic constraints are honored. This is because we are integrating stacking patterns through preprior. If there is no well data, preprior dictates the layering, an essential process in reservoir modeling.

5.5.2 Sensitivity to Preprior Information

The reservoir models in Fig 5.6 are built similar to previous cases but without preprior. In these cases, ordinary kriging is used which applies well data to estimate the kriging mean and variance, rather than simple kriging, which uses ensemble statistics as in previous cases. When models are simulated without well conditioning (Fig. 5.6(d)), layers are not prograding, but stacking on top of one another. This is because no well data or preprior tells the simulation which layers are prefered in which side, as a prograding system. However, as the number of well data increases, this additional information informs the stacking pattern and then the models looks more like the base case (Fig. 5.6(a,b,c)). When there is little



(a) Simulated model with 15 wells

(b) Simulated model with 10 wells

(c) Simulated model with 5 wells

(d) Simulated model with no wells.

Figure 5.6: In this simulation on $50 \times 50 \times 10$ cornerpoint grids, the areal extent is $X = Y = L = 1000$ m. Constraints used are well and seismic (exact). No ensemble statistics of surface-based models are used. Vertical exaggeration is twentyfold for all figures. Less well models are not only prograding, but in fact are randomly filling the seismic volume. As the number of wells increase, the models are closer to the base case.

well data for the reservoir or parts of the reservoir, the use of preprior information helps to mimic the stratigraphic patterns. The stationarity issues with kriging can also be decreased by using preprior.

Chapter 6

Selecting Geostatistical Realizations Using Fast Flow Simulations

Heterogeneities at different scales in the reservoir influence flow. The true effect of all the heterogeneity on fluid flow could only be estimated by exhaustive sampling of the entire reservoir. Due to lack of such detailed (and practically speaking, impossible) sampling, knowledge of the spatial variations in properties such as porosity and permeability is not possible. However, geostatistical models (under certain assumptions) are conditioned to geologic, seismic, well control and other data, and are able to present an array of plausible subsurface geologic models. Even with very good subsurface data, capturing the true image of all the geology is not possible because of the complex and hierarchical nature of the heterogeneities present in the subsurface. In addition, there is the question of flow model adequacy, which is not addressed in this dissertation. The uncertainty is caused by undersampling spatially, which can be addressed (albeit indirectly, and weakly) by generating multiple geostatistical models. However, although these models honor all the available data and their spatial structure, each model could be quite different from other models in response to the fluid flow or the spatial distribution of static average properties like net-to-gross. This arises from a lack of constraining data. The fewer the constraining data, the more uncertainty and the more different the realizations become in response to production.

In any case, the realizations are intended to characterize the uncertainty. Some of these realizations are not plausible because of an inability to integrate production data; building geologic models honoring production data remains an ongoing research (Li et al. 2007). Even when not considering the rejected models, the number of feasible models available could run to the hundreds.

Significantly, these realizations are generated by using a particular algorithm based on certain assumptions like linearity or stationarity (Goovaerts 1997). Under these assumptions, all

the realizations are "reasonable." If these assumptions are not met, then generating models with the algorithm might give unrealistic geologic models, and then more appropriate modeling techniques would be used. This chapter addresses uncertainties in lack of information, and the resulting stochastic fluctuations in flow responses, rather than geostatistical model *per se*; an alternative geostatistical formulation is discussed in chapter 5.

6.1 The Need to Screen

Flow simulation is widely used for reservoir forecasting and management. Flow models are built from geomodels, which are uncertain because of sparse data and complex geologic origins as discussed above. Increasingly, engineers quantify uncertainty by considering a range of plausible geomodels. However, it is expensive and difficult to simulate many geomodels using full-physics flow models. A small set of geomodels are to be selected from hundreds of realizations. Many of these initial set of realizations are redundant. The redundancy or similarity of the models should be judged upon response of these models to the fluid flow rather than visual inspection. Some models may look different visually but could respond to fluid flow very similarly. From the initial set of realizations, similar geomodels are removed and so the remaining are different and they would capture the uncertainty. We need efficient and robust methods to select relatively small sets of geomodels that are nonetheless diverse and representative.

6.2 Sampling Not Ranking Realizations

Realizations have commonly been selected by ranking (Deutsch and Srinivasan 1996). Ranking uses an easy-to-estimate *secondary* response like effective permeability; the chosen response should correlate with full-physics (or *primary*) responses like cumulative oil recovery. Secondary responses are useful if the rank correlations with the primary responses are high. However, the ranking may change if factors like well pattern change; generally, the rank differs for various responses. Primary responses, like breakthrough time, depend on connectivity

and continuity of good quality rock, also many secondary responses or their combinations measure the continuity. When there are many secondary responses that are of interest, a single ranking is hard to obtain; therefore selection of low, medium, and high cases using a single ranking does not represent uncertainty. Conversely, we must consider the joint, multivariate distribution of many secondary responses in order to choose realizations that are diverse in all responses. A quasirandom sampling method called Hammersley sequences ensures the sample is representative in the multivariate distribution of secondary responses.

6.3 Simple Static and Dynamic Responses

Simple summary statistics, computed by estimating volume average properties, can be used as secondary responses to differentiate various realizations. Average properties for a realization such as net to gross or net pore volume are estimated by first defining the cut-off for net sand, using a porosity or permeability threshold. Below that threshold, the porosity and permeability of the rock is considered zero and an indicator of net sand $I = 0$. $I = 1$ means that the cell is a net sand, and can contribute to fluid flow. The volume average net-to-gross for N cell block or reservoir is

$$NTG = \frac{1}{NV_p} \sum_{i=1}^{N} I_i v_i \tag{6.1}$$

Volume average porosity (Φ) is estimated as

$$\Phi = \frac{1}{V_p} \sum_{i=1}^{N} \phi_i v_i \tag{6.2}$$

where v_i and ϕ_i are the volume and porosity of cell i, N is the total number of cells in the reservoir, and V_p is the total pore volume of the reservoir.

Upscaled permeability can also be used as a secondary response, although upscaling permeability is usually harder than other parameters. Upscaling properties for the coarser scale grid are obtained from fine scale grids; as a result, flow responses for both the grids are the same for a chosen range of boundary conditions. The most accurate method to upscale permeability is to solve flow equations with constant pressure and no flow boundary conditions

(called the pressure solver technique). There are other techniques which are less accurate, but much faster. The method introduced by Cardwell and Parsons (1945) and improved by Li, Cullick, and Lake (1995) and others is fast and close to the pressure solver technique. Upscaled effective permeability is between the upper and lower bounds, which presents harmonic and arithmetic static averages. The horizontal effective permeability is similar to the horizontal directional upper bounds, and the vertical effective permeability is likewise near the vertical directional lower bounds. Upscaling a section of a reservoir would involve both arithmetic and harmonic average as explained in Li, Beckner, and Kumar (1999). Whether upscaled dynamically or statically, significant computation, or at least workflow management, is needed to manage many upscaled models for ranking realizations (Deutsch and Srinivasan 1996; Kupfersberger and Deutsch 1999; Ates et al. 2005). In this current research, upscaled permeability of a subgrid can be one of the secondary responses for sampling realizations.

There are other dynamic measures used other than flowing upscaled models. One such measure is to use a shortest path algorithm that calculates time of flight between two locations under single-phase constant rate flow conditions. These random path algorithms are similar to streamline simulations, which are also used for ranking realizations (Ates et al. 2005). Other ranking methods use a tracer test analysis, which is also used in this dissertation. However in this work, tracer simulations provide some of the secondary responses, because other secondary responses (*e.g.*, upscaled injectivities) are also considered for sampling the realizations.

6.4 Screening Using Fast Flow Simulations

In this work, to make uncertainty assessment feasible, we propose a method to select relevant models by using simple and fast simulations, assuming a single phase tracer flow. Although the screening simulation does not include all of the physics or operational constraints of the full-field model, it does incorporate many important effects of heterogeneity; and it is used to select models, rather than approximate them.

Here, single-phase tracer simulations are used for the secondary responses. Many responses can be computed from tracer simulations like injectivity, Lorenz coefficients, and several residence time statistics. Although the screening simulation must simplify the physics and operational constraints, tracer simulations include geomodel heterogeneity and geometry. Various injector-producer pairs sample reservoir anisotropy. The number of injectors and producers depends on the size and scale of heterogeneities needed to capture. This screening method samples flow model variability with far less computation than full-physics simulations in high-dimensional geomodel spaces.

At the same time, the sampling preserves variability in flow responses. This method includes other statistics like average porosity, together with tracer test statistics to allow sample realizations to be easily done. The method is also easy to adapt to other approximate flow models and alternative sampling methods. Sampling multivariate secondary responses leads to realizations that differ for a variety of production scenarios that are not otherwise possible with ranking based on a single secondary response.

6.4.1 Tracer Test Analysis

If a tracer is injected as a slug (or pulse injected) through injectors, it mixes and disperses with the reservoir fluid before reaching a producer, where tracer concentration C is measured as a function of time t (volume of tracer is measured if rate is not constant). This tracer response curve (C *vs.* t) is influenced by the reservoir heterogeneity. For homogeneous reservoirs, the tracer breakthrough time is longer and the spread of C *vs.* t curve remains small. On the other hand, for more heterogeneous reservoirs, the breakthrough time is shorter, and the spread of the tracer distribution curve is larger. Tracer analysis and the method of moments based on ideal tracer assumption as shown here is mainly from Shook and Forsmann (2005). An ideal tracer is a tracer that is fully miscible with an injected phase, and the tracer does not partition, decay, or adsorb from the injected phase. The tracers are assumed not to change the density or viscosity of the injected phase; therefore they are "ideal."

Under constant rate injection, the volume swept by the tracer is proportional to the cumulative tracer produced ($\int_0^\infty C dt$), and the mean residence time of the tracer ($\bar{t}$) is (Saad, Maroongroge, and Kalkomey 1996)

$$\bar{t} = \frac{\int_0^\infty tCdt}{\int_0^\infty Cdt}$$

The above equation is for pulse tracer injection; if the tracer is injected as a slug with the duration of injection t_s, the mean residence time is modified as

$$\bar{t} = \frac{\int_0^\infty tCdt}{\int_0^\infty Cdt} - \frac{t_s}{2}$$

This solution is for multiple wells and for any well pattern. Variance of the tracer response curve gives the spread of the curve, and it is a good indicator of heterogeneity in the reservoir. It is estimated as

$$\sigma^2 = \frac{\int_0^\infty t^2Cdt}{\int_0^\infty Cdt} - \bar{t}^2$$

In many situations, tracer simulations are not run until the concentration of the tracer goes back to zero. Usually the tracer response curve has a large tail. Not running the simulation without observing this tail leads to an underestimation of the mean response time. If simulation is stopped at time t_b, an exponential decline of the tracer can approximate the concentration at a late time, $C(t) \sim b\exp^{(-at)}, \forall\, t > t_b$. Coefficients a and b are estimated by fitting a late time trend just before t_b with the exponential decline curve. After considering the extrapolation $\bar{t}$ is

$$\bar{t} = \frac{\int_0^{t_b} tCdt + \int_{t_b}^\infty tCdt}{\int_0^{t_b} Cdt + \int_{t_b}^\infty Cdt} = \frac{\int_0^{t_b} tCdt + \frac{b}{a^2}\exp^{(-at)}(1+at_b)}{\int_0^{t_b} Cdt + \frac{b}{a}\exp^{(-at)}}$$

F-Φ Curves

Shook and Forsmann (2005) found that tracer test analysis can estimate flow geometry parameters, such as flow and storage capacities. If tracer flow paths may be imagined as streamlines going through a reservoir, each streamline path has a volumetric capacity and a velocity, depending on time of flight (Ates et al. 2005). By arranging streamlines in decreasing

volumetric capacity, cumulative flow capacity (F_i) of streamline i is defined as the sum of all flow capacities (k_iA_i/L_i for i, from Darcy's law) greater than streamline i divided by the sum of all flow capacities. Cumulative storage capacity (Φ_i) also is similar, except that it replaces flow capacity with storage capacity ($\phi_iA_iL_i$ which is pore volume) of streamline i. Flow and storage capacities are useful for quantifying the heterogeneities as the ratio of permeability to porosity that is proportional to interstitial velocity in a single phase flow. Flow and storage capacities can be estimated from tracer tests. The cumulative storage capacity, Φ_i, is

$$\Phi_i = \frac{\int_0^t tCdt}{\int_0^\infty tCdt}$$

The cumulative flow capacity, F_i, is the cumulative tracer recovery at time t divided by complete recovery

$$F_i = \frac{\int_0^t Cdt}{\int_0^\infty Cdt}$$

The shape of $F - \Phi$ curve gives the fraction of pore volume contribution to a fraction of recovery.

Lorenz and Dykstra-Parsons Coefficients

From $F - \Phi$ curve other estimates of reservoir heterogeneity like Lorenz coefficient (L_C) and Dykstra-Parsons coefficient (V_{DP}) can be estimated (Lake and Jensen 1989). Lorenz coefficient is estimated as

$$L_C = 2\left[\int_0^1 Fd\Phi - \frac{1}{2}\right]$$

When L_C is close to zero, the reservoir is homogeneous and when it is close to one Reservoir, is heterogeneous. L_C cannot be greater than 1 as it is normalized.

Dykstra-Parsons Coefficient is estimated from a tracer test by taking the derivative of $F - \Phi$ plot. The derivative (F') represents instantaneous fluid velocity. The coefficient is

$$V_{DP} = \frac{F'|_{\Phi=0.5} - F'|_{\Phi=0.841}}{F'|_{\Phi=0.5}}$$

V_{DP} uses the F' at the mean ($\Phi = 0.5$) and one standard deviation above the mean ($\Phi = 0.841$). The statistics come from tracer test analysis, much like mean residence time,

Lorenz and Dykstra-Parsons coefficients, and tracer breakthrough time are used as secondary responses. Sampling realizations in this research with multiple secondary responses are done by Principal Component Analysis.

6.4.2 Principal Component Analysis (PCA)

The central idea in this chapter is to use many secondary responses, both static and dynamic, to sample realizations. Various injector-producer pairs sample anisotropy by using several dynamic responses, together with many upscaled properties to assess the variability of different sections of reservoir.

The dimensionality of all the secondary responses generated can be in the hundreds. Many of these are intercorrelated and redundant and therefore can be eliminated while retaining the variation in the data set by using Principle Component Analysis (PCA). This is achieved by rotating the secondary responses data cloud axis and generating a new set of variables called principle components (PCs), aligned with the data cloud and ordered such that the first few PCs retain most of the variability present in the original variables.

Extraction of principal components amounts to a variance maximizing rotation of the original variable space (Hair et al. 2005). After finding the PC (a vector) on which the variance is maximal, there remains some variability around this PC. In PCA, after the first PC has been extracted (that is, after the first vector has been drawn through the data), we continue to define another PC that maximizes the remaining variability, and so on. In this manner, consecutive factors are extracted. Since each consecutive factor is defined to maximize the variability that is not captured by the preceding factor, consecutive factors are orthogonal to each other (under linearity assumption). This new orthogonal basis gives a set of factors (PCs) that accounts for most of the variability in the system.

Let the data set be $\mathbf{X}$, an $m \times n$ matrix, where m is the number of samples or realizations and n is the number of variables. Also assume the covariance of the $\mathbf{X}$ be $\mathbf{\Sigma}$ $(= \mathrm{E}\left[(\mathbf{X} - \mathrm{E}[\mathbf{X}])(\mathbf{X} - \mathrm{E}[\mathbf{X}])^{\mathrm{T}}\right])$, a matrix whose $(i,j)^{\mathrm{th}}$ element is the covariance between

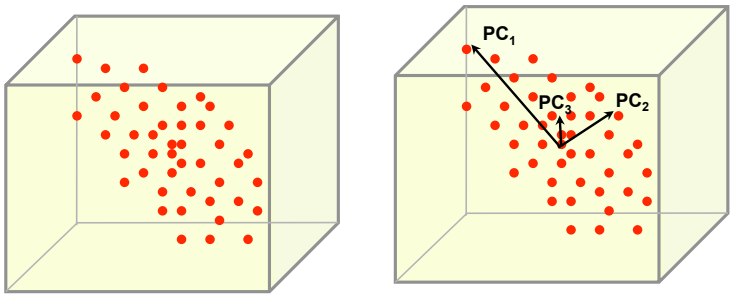


Figure 6.1: Example PCA for 3D problem

i^{th} and j^{th} variables of $\mathbf{X}$. Principle components are the *eigenvectors* of $\mathbf{\Sigma}$ and *eigenvalues* of this matrix, giving the variance along each principle component. The higher the eigenvalue of a PC, the higher the contribution of that PC towards the variability of the data set. Generally a small subset of PC's can capture most of the variability.

6.5 Synthetic Example

6.5.1 Geomodel Construction

A downscaled model with high sand sill ($\sigma_s^2 = 36$) and medium range ($b/L = 0.5$) is taken as a reference case (Three-dimensional Cornerpoint Flow Simulations, Chapter 3, earlier). Downscaling is done with four conditioning wells, all of which have an average thickness imposed and a porosity trend that increases downward linearly from 0.2 to 0.3. Fifty realizations are generated with a $100 \times 100 \times 10$ grid with 1000×1000 m areal extent. The total sand thickness $H_s = 14$ m, total shale thickness, $H_{sh} = 6$ m, and average porosity $\Phi = 0.25$ constraints are imposed at each trace. Alternative sand and shale architecture is used. Permeability is assigned, using an exponential transform of porosity,

$$k = 20e^{10\phi}$$

All the 50 realizations have the same well and seismic constraints but very few wells are used to condition the model. As a result, stochastic fluctuations are very dominant and 50 realiza-

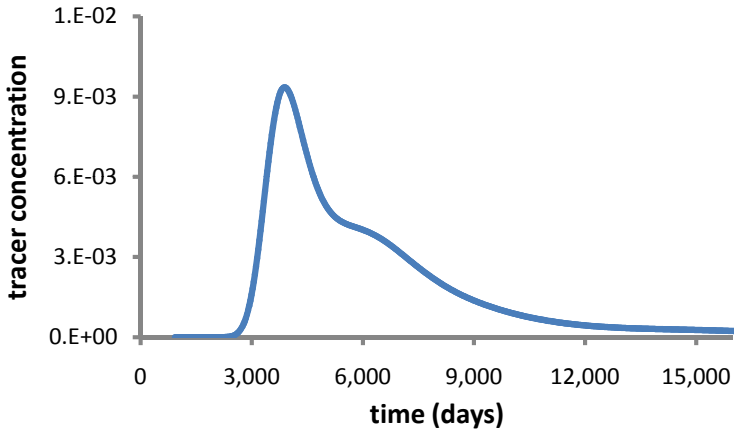


Figure 6.2: Residence time distribution of tracer at producer 1 for a realization

tions are very different. This is a practical situation, reflecting the early field development, where little well conditioning data is available, and the uncertainty is huge.

6.5.2 Secondary Responses Considered

Single phase tracer simulation uses a geometry of one-quarter of a five-spot. Several responses were considered for each producer: tracer breakthrough time, mean residence time, volume swept, mean residence time without exponential correction, standard deviation, and skewness of tracer distribution, Lorenz coefficient, and dynamic effective permeability. Tracer statistics are obtained from residence time distribution of the tracer at each producer (Fig. 6.2). Dynamic effective permeability is obtained by the ratio of flowrate to pressure difference between an injector and a producer; this variable is proportional to effective permeability. Mean residence time, with or without exponential correction, is estimated to understand the effect of later time tracer distribution on sampling (Tracer Test Analysis, earlier).

6.5.3 PCA Analysis

Twenty-four responses are obtained, since there are three producers and eight responses at each producer. Lower dimensional parameterization of 24-dimensional data is done by estimating principle components (PCs) of the data, which will be 24-dimensional vectors

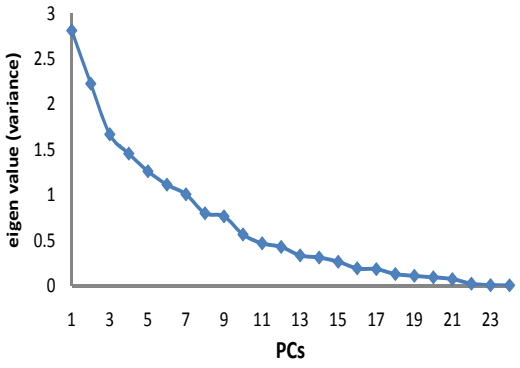


Figure 6.3: Variance captured by each PC.

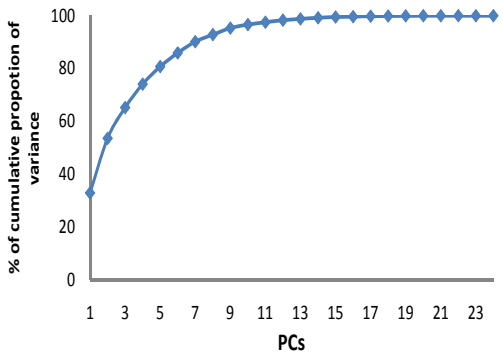


Figure 6.4: Cumulative proportion of variance.

aligned with the directions of maximum variance of the data. The PCs are ordered by data variance as discussed in earlier section. In Fig. 6.3, variance captured by each PC is shown. The first PC is aligned with the direction of maximum variance, and the second PC in the orthogonal direction contributes the most variance, and so on. In this case, seven PCs are capturing 90 percent of the total variance (Fig. 6.4); as a result, the other PCs are discarded.

All 50 realizations can be transformed and sampled in the new PC space with no correlations. If the correlations are zero, sampling methods like Monte-Carlo, Latin hypercubes, and Hammersley sequences (a quasi-Monte Carlo method) can be directly used to sample the principal component space. Low dimensionality in a PC space makes sampling methods more accurate and efficient. Sampling in this lower dimensional multivariate space can also be done by using experimental design methods such as orthogonal arrays (Kalla and White 2005).

6.5.4 Hammersley Sampling on PCA Space

Five samples are selected using Hammersley sequence on the PC space; the nearest neighbors (using Euclidean distance) among 50 are chosen. The samples generated by Hammersley are space-filling, with a relatively small but still representative sample (Kalagnanam and Diwekar 1997). These selected samples are distinct from one another and thus capture most of the

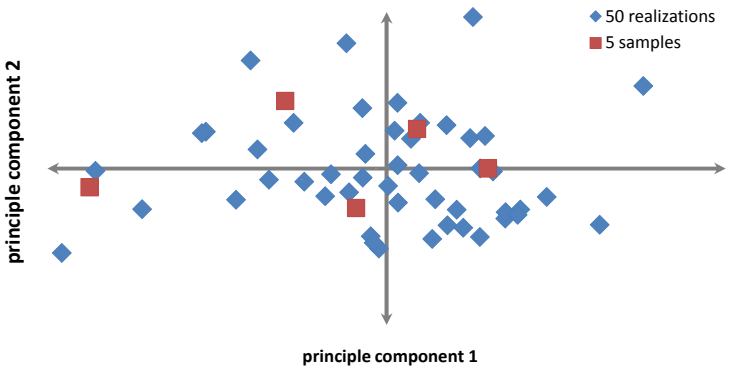


Figure 6.5: Loading parameters for first two significant PCs.

variability. Fig. 6.5 shows PC_1 and PC_2 and the 50 realizations. The five samples chosen by Hammersley are also shown. As expected, they fill the PC space very well. Fig. 6.6 shows the mean residence time response of producer 1 and 2 for all 50 samples on the original axis. They are negatively correlated. Also shown in the figure are the five samples chosen by Hammersley, these samples cover the data cloud and thereby sampl the original space as well. If better coverage were required, more samples should be used; using five samples in a 7-dimensional PC space (24 dimensional in the original responses) is rather aggressive.

The correlation coefficient between mean residence time, with or without exponential correction at a late time, is almost equal to one. This shows that for sampling, there is no need to do late time tracer correction. Mean residence time, however, is different with and without correction; if the residence time *per se* were sought, rather than a secondary sampling variable, then the correction should be applied.

6.5.5 Validation Using Full Physics Simulations

The five samples selected from 50 realizations using the tracer distribution and other secondary responses are to be diverse in primary response space. Sampled points primary response distribution should be representative of the distribution of the original realizations. In other words, the statistics from the five samples and 50 realizations should be closer.

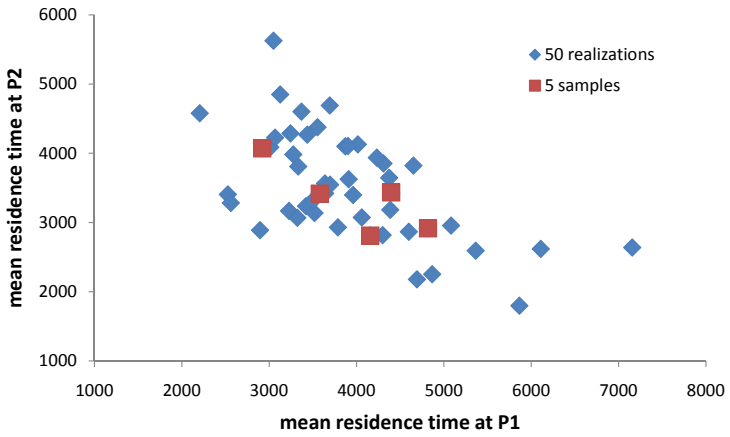


Figure 6.6: Two correlated responses from tracer flooding.

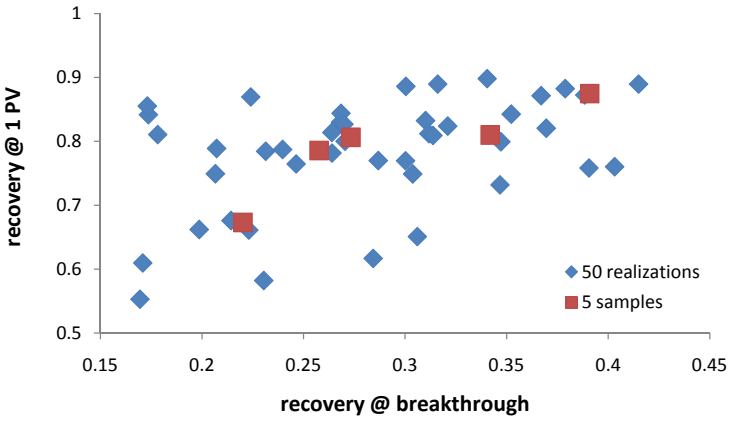


Figure 6.7: Two full-physics responses for the 50 realizations.

To validate the sampling, primary responses like recovery at breakthrough time (BT) and recovery at one pore volume (1 PV) is estimated by running two phase flow simulations. For all 50 realizations, two phase flow simulations are done with geometry of one-quarter of a five-spot, like the tracer flooding. Breakthrough is considered when 100 STB/day of water is produced from the field and recovery is estimated as the amount of oil produced compared to the initial mobile oil in place. Typical fluid properties and relative permeability curves are used for this study (Appendix A). Injectors are rate-controlled and producers are controlled by bottom hole pressure.

Table 6.1: Recovery summary for the 50 realizations and the 5 samples

	Recovery for 50 at BT (%)	Recovery for 5 at BT (%)	Recovery for 50 after 1PV (%)	Recovery for 5 after 1PV (%)
Mean	28.1	29.7	77.8	78.9
Stand Dev	7.2	6.9	9.1	7.4

Fig. 6.7 shows the recovery at breakthrough and after 1 PV injected for all 50 realizations. Also shown in the figure are the five samples chosen by Hammersley using secondary responses, which cover the 50 realizations well. The recovery at breakthrough for 50 realizations varies from 12.7 percent to 41.5 percent, and the recovery after 1 PV for the 50 realizations varies from 55.2 percent to 89.8 percent. In Table 6.1, mean and standard deviation for the 50 realizations and five samples are shown.

6.6 Discussion

A Welch two-sample t-test indicates that the means of recovery (means are 28.1 and 29.7, $t = -0.5$; 95 percent confidence interval: -0.1 0.07) at breakthrough time and recovery after 1PV (means are 77.8 and 78.9, $t = -0.3$; 95 percent confidence interval: -0.1 0.08) for the five samples and 50 realizations are not the same. In other words, the five sample mean does not specify the 50 realization mean for both primary responses. However, the F-test presents a variance ratio of $F = 1.1$ and 1.5 for breakthrough and recovery, which is within the 95 percent confidence interval. Therefore, we see that, the variance captured by the five samples does not differ from the 50 original realizations for both primary responses.

Even though the means are different at the 95 percent confidence interval, p-value for the two cases (64% for BT and 76% for 1 PV) is around 70%. So for both the primary responses, it is likely that both samples belong to the same population. More work is needed with a greater number of realizations (more than 200), with more samples (around 20), in order to provide a more concrete understanding. Influences of gravity, capillary, and viscous forces on

sampling efficiency, as well as the number of samples required for capturing the distribution of primary responses, should be addressed in the future studies.

This work, by incorporating fast flow simulations, PCA, and Hammersley, in sampling rather than ranking, is novel in its approach. Further, any fast flow simulation method such as shortest path algorithms or other sampling techniques like simple Monte Carlo sampling may be used with this current approach.

Chapter 7
Concluding Discussion

Focused discussions have been presented in previous chapters. This chapter briefly addresses issues that span many topics and future work that may be fruitful in integrated reservoir modeling.

7.1 Future Research

There are outstanding issues that should be addressed to improve the data integration techniques proposed in this work.

7.1.1 Surface Modeling Data

The surface-based models used in this research require data such as geologic correlation lengths and possibly layer architecture, which is difficult to estimate. However, these data may be obtained from a variety of historical, experimental, or numerical approaches.

- **Historical (Outcrop Data):** Outcrop data supplies a wealth of small-scale facies and spatial information unattainable in subsurface investigations. Some of the several types of information that may be obtained from outcrops are depositional feature orientations like paleocurrent directions and grading and stacking patterns. White et al. (2004) investigated tide-influenced deltaic sandstones in an outcrop of the Frontier Formation in central Wyoming. Stratigraphy in the upper and lower portions of this analogue is influenced by the episodal tidal deposition and ebb-directed currents respectively, and deposited geobodies show those characteristics. These studies allow understanding for the extent of the sand bodies in similar depositional settings. Outcrops give detailed, small-scale heterogeneity; therefore their effect on fluid flow may be studied extensively. However these results are for analogue, rather than the reservoir under

study. They are also 2D cross-sections; as such, the outcrops will involve intensive work to characterize the heterogeneity. Finally, each outcrop is but one realization of the sedimentary process. Although extensive exposures would mitigate this problem, that process may cause difficulty in developing statistical models of the form proposed in this dissertation.

- **Experimental (Flume or Tank Models):** Experimental stratigraphy is generated by using flume or tank models to fill up "basins" under a controlled supply of sediment, subsidence, base level variations, and transport mechanisms. For doing such a controlled experiment, state of the art, an experimental setup is required, much like the Jurassic Tank at the University of Minnesota (Paola et al. 2001). The experiments imitate those physical processes involving the sediment transport with shoreline and stratigraphic records, evolving over time. Only the boundary conditions may be controlled in these experiments, but as in many actual reservoir cases, the paleocurrent directions, the rate of sediment influx, and other boundary conditions are unknown. Therefore, it is hard to mimic stratigraphy of a particular reservoir, but these experiments in general give a plethora of information on how sediment records evolve over time, as well as the interaction of different boundary conditions on a sedimentary environment. However, each experiment is time-consuming, and given the experiment's size and complexity, is expensive. This makes it difficult to formulate statistical models solely from experimental data.

- **Numerical (Mechanistic Sedimentation Models):** Mechanistic sedimentary models of a depositional setting are generated by solving sediment transport equations. These models are process-imitating, forward models that incorporate the physics behind the deposition. Numerical models are very similar to tank models, but numerical models allow easy interpretation, simply because it is easier to scan through these models. Yet these models also have the same limitations of tank models in boundary

conditions. It follows that this uncertainty about the sediment rate in the paleoconditions, as well as the timing of the big sediment influx, also makes mimicking the reservoir architecture a difficult procedure. However, an ensemble of these models may provide a useful statistical model for use as a preprior as formulated in this dissertation. The computational effort of running a large suite of such models may limit this application in the near future.

Many realizations will be needed for these models, if we are to develop robust statistical models. One approach would be to derive properties from outcrop, tank experiments, and numerical experiments, in order to build surface-based models, as discussed earlier in chapter 5. Typically, the rules for the surface-based models are controlled by physical and numerical models. The inherent uncertainty (even though this uncertainty is difficult to estimate) found in the parameters for surface-based models are easy to adjust; as a result, an ensemble of these models can provide the robust statistics required. Again, any other approach may be used to build the necessary statistics, as explained in the last paragraph; ensembles of numerical models, providing they are relatively inexpensive, could be useful.

7.1.2 Stratigraphic Model Inference

In this dissertation, the stratigraphic model (or preprior) has been formulated as a spatially varying probability field, $p(t|l,k,K)$, where l is the current trace (areal location), k is the layer, and K is the total number of layers. As noted in the previous section, this requires many realizations to compute a stable estimate. One goal of future research should be to formulate guidelines for the number of realizations required, based on *a priori* parameters or convergence monitoring.

An additional complication is heteroscedacity, the upward-increase of stratigraphic model variance. Briefly, this is caused by the sequential nature of deposition. This is noticeable in simple surface-based stacking models and also in numerical models. This chaotic behavior is evinced in the unpredictability of natural phenomena such as deltaic lobe autocyclicity. As

deposition (or simulation of deposition) proceeds, the placement of each succeeding layer is conditional on an increasingly large number of prior choices (which determine the accessibility of sediment supply and areal variation of accommodation, Reading 1996). Fundamentally, sedimentation is *not* a Markov process: the new locus of sedimentation depends on many previous sedimentation episodes, not just the most current episode. In terms of impact on the stratigraphic model, in a compensational stacking geologic environment, the layers on the top will vary widely for different realizations, and the preprior (which is the average of all the realizations) will be less informative about how layers are arranged with respect to one another at the top, as compared to the bottom of the depositional succession.

Although this chaotic nature may be a true reflection of natural sedimentation, it tends to be expressed unrealistically, as local noise rather than global uncertainty in the current implementation of the stratigraphic model. At this point, it is not clear how to formulate the model to partition the noise correctly amongst various scales. Nonetheless, the Bayesian formulation does ensure that the realizations will be less dependent on the stacking pattern information at the top of the succession where preprior variance is high, thereby allowing seismic and well data to exert a comparatively greater influence. Clearly, a more robust and consistent method for estimating the preprior is needed.

7.1.3 Prior Model Inference

The proposed approaches use variograms of t to regularize the grossly underdetermined, downscaling problem. The untruncated proxy t is kriged rather than truncated h to obtain prior distributions, because kriging assumes variables are Gaussian distributions (continuous distributions with a range of $-\infty$ to $+\infty$), but actual thickness h is non-negative. Unfortunately, while the variogram of h might be observable in the data sources discussed above, the variogram of the untruncated Gaussian t is utterly unobservable and must be inferred from h. This inference problem was not addressed in this dissertation.

In principle, a Gibbs sampling procedure, similar to the approach used to convert $h = 0$ well data to t (chapter 2.2.4), could be used to address this. If we observe some layer k on trace ℓ has $h_{\ell k} = 0$, the value of $t_{\ell k}$ is determinate by this procedure; we can use $t_{\ell k} \leq 0$ at these traces and other traces where $t = h$ to determine the variogram for t. However, the estimation is nontrivial and will require careful formulation and validation versus a variety of historical, experimental, and numerical data sets.

7.1.4 Selecting the Right Ensemble

The transfer of uncertainty from the geologic model to production forecasting is optimized by selecting realizations that are different from one another (chapter 6). The proposed method of running fast flow simulations and using statistics like average permeability, and then sampling by using Hammersley sampling as well as a principal component analysis, allows a small, selected sample to remain diverse to span the multidimensional response. Five samples chosen from 50 realizations span multiple responses (Fig. 6.7) but these few samples (5) cannot mimic all the primary responses distribution statistics. However, even this small sample correctly reproduced the variance in a blind prediction of a waterflood process (which was not used to select the small sample).

A more comprehensive study, involving around 200 initial realizations and a sample of *circa* 20 from those, together with a comparison of the primary response statistics, might prove that the responses distributions show similarities. Also, other fast flow simulation techniques may prove useful, such as flowing through the upscaled models, shortest path algorithms. The influence of well spacing (optimal well spacing) for fast flow simulations and the influence of gravity, viscous, and capillary forces effects on sampling should be further studied.

7.2 Model Testing

7.2.1 Verification Remains Elusive for Earth Models

Most 3D models to this point are not well tested. It can be argued that geomodels and other models in earth sciences like weather forecasting models can never be verified or validated, because the modeled system is never closed and models require input parameters that are incompletely known (Oreskes, Shrader-Frechette, and Belitz 1994). Verification can be done by comparing the geomodel generated to the truth case, but architecture and properties of the actual reservoir that is being modeled are always unknown, and so cannot be verified.

Despite this difficulty, it is worthwhile to compare different algorithms against one another, and to compare model output against observed data. Geoscientists could generate well-constrained, quantitative, field studies of specific environments that represent the prime sources to be modeled. Then a comparison may be made between the reference model and the model generated by algorithm. Even then, it will be a challenge to generate goodness-of-fit criteria and to design testing protocols. Model selection techniques such as Bayesian information criterion (BIC) may prove to be unuseful, as the number of free parameters to be estimated are huge. When used to compare different realizations on an average, all the realizations may have same BIC (Liddle 2007). More work is needed to show the applicability of these techniques in comparing models.

7.2.2 Validation

Validation does not require the establishment of truth. The internal consistency required for validation is more appropriate for the computer code, and relates only weakly to comparisons of 3D model predictions to physical reality. Nevertheless, simplified 2D models provide insight and validate the algorithm. When seismic is used as an inexact constraint, 2D case studies (Figs. 2.3, 2.4, and 2.5) have shown that the marginal distributions of the two layers and distribution of the total thickness are consistent with the prior and the likelihood. All the

samples generated by MCMC sampling are consistent with the well data (through the prior) and the seismic data (through likelihood). Similarly, when seismic is used as exact constraints, the 2D cases show that the marginals are consistent with the well, and the total thickness is exactly equal to the seismic constraint (Fig. 3.6). The 2D cases validate the algorithm and the mathematical procedure.

7.2.3 Consistency

Validation of the 3D results is weak, because various realizations will give different results. Yet consistency of the realizations can be established; the validation is only probabilistic. Further, a 3D realization is valid, depending on the input parameters and the various approximations that are used to build that model. For the example 3D cases provided in this research, geologic frameworks provided by many realizations involving different prior and likelihood and in inexact and exact constraint problems, static models built are consistent with all of the geological and geophysical information. In inexact constraint 3D problems, as the noise of the seismic increases (Fig.2.6c and Fig.2.6d) the sum of layers are more influenced by well information. Therefore, the total thickness deviates from the seismic mean. On the other hand, if the seismic noise is less, the total thickness closely matches the seismic mean. When the variability increases (in exact constraint problem; Fig.3.13a and Fig.3.13b) and the geologic correlation length decreases (in an inexact constraint problem; Fig.2.6a and Fig.2.6b) the layers pinch out more frequently. These model parameters and their behavior are honored in the realizations, while being consistent with all other constraints.

The consistency of results is also evident when stratigraphic data (preprior) is incorporated. When more well data and stratigraphic data are included, the realizations become close to the reference case, as expected (Fig.5.5). When the preprior and the well data are not used, the realizations generated are aggrading rather than prograding (Fig.5.6). When there is no well or preprior data, the relative location of the layers is unknown to the algorithm; as a result, stratigraphy is not reproduced. Nevertheless, the prior information (correlation

lengths) and the seismic data are honored. Realizations with either the wells or preprior are able to replicate the reference model. All these results show the consistency of realizations as generated by use of the downscaling algorithm.

7.2.4 Calibration

Comparing the flow response of the actual production of real case and geomodel, called history matching in the petroleum industry, may be of great help to show consistency of the flow results (Oliver 1994). Yet it must be understood that even if the flow responses are consistent, we neither validate nor verify the model. The actual truth case could be very different from the realization. This is because the flow average, or integrates the heterogeneities between the injector and the producer. History matching can be seen as a (nonlinear) deconvolution or downscaling process, functionally similar to the seismic-based methods discussed in this dissertation. Similarly to the results of this dissertation, history matching is non-unique. The results are best viewed probabilistically (Li 2008).

One could use history matching to calibrate a surface-based model to numerical, experimental, or historical stratigraphic models. That is, the prior and preprior could be inferred using yet a third inversion process, in addition to the flow- and seismic-based modeling. The process is similar to the variogram inference used for surface-based models in a somewhat simpler context, applying only well data and shape templates (Pyrcz 2004). This comprehensive geomodeling workflow has yet to be realized.

Chapter 8
Summary and Conclusions

8.1 Summary

In this thesis, new methods are proposed to integrate multiscale data. The methods involve downscaling seismic inversion models. A seismic model is uncertain because of its low resolution and noise; further, multiple realizations are used to characterize this uncertainty. When the downscaling is done by picking up a realization it is called an exact method. The advantages are that downscaling a single realization incorporates the correlation between the layer thickness, porosity, net-to-gross and other parameters that are honored in seismic inversion. Downscaling inexact seismic constraints incorporates the seismic noise which is obtained by an ensemble of realizations. Both exact and inexact methods are useful and apply sequential algorithm which decomposes the domain, based on multi-Gaussian assumptions. It is shown that when the correlations are not tight, the decomposition is a good approximation, but if they are tight, sequential methods with marginals are more appropriate.

Including stratigraphy into reservoir models by using kriging-based algorithms is challenging. In this work, surface-based models are used to build stacking patterns, used in the sequential algorithm, together with seismic and well data. An ensemble of surface-based models is generated, and the distributions of thicknesses and properties obtained are assumed Gaussian and sunsequently used in simple kriging. This seismic downscaling approach, using an ensemble of surface-based models, could be regarded as conditioning surface-based models. Any other approach that can characterize the stratigraphic detail can also be integrated in a similar fashion; therefore more appropriate methods for generating such details should be investigated.

A method to select realization by using fast flow simulations and secondary responses, such as upscaled permeability, is shown. This procedure uses both principal component analysis

and sampling techniques to sample a few realizations from an original set of realizations. A large number of an original set of ealizations is generated to capture uncertainty in the primary responses. Selecting realizations using fast simulations removes the need to run expensive two- and three-phase simulations for all original realizations.

8.2 Conclusions

Stochastic seismic inversion computations can be integrated by means of a truncated Gaussian geostatistical model for layer thickness, using a Markov chain Monte Carlo method. Truncation makes the problem nonlinear, which is ameliorated by the introduction of auxiliary variables and a mixed Gibbs-Metropolis-Hastings sampling procedure in an inexact constraint problem. 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. Mesoscale seismic inversion realizations (which act as exact constraints) of net-sand, gross sand, and porosity are "stochastically downscaled", using a Metropolis Hastings sampler by projection to the exact constraint surface. Use of exact constraints from stochastic seismic inversion realizations preserves correlations implied by rock physics and seismic data.

Proposal efficiency is critical to the usefulness of both these methods, 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 modeled truncations, are verified by examining algorithm behavior in two dimensions. Synthetic three–dimensional cases demonstrate that the proposed data integration procedures are acceptably efficient and are capable of producing models consistent with seismic data while exhibiting diverse flow behavior.

Sampling the uncertainty in these nonlinear downscaling problems is difficult. Global MCMC methods are accurate but expensive, which motivates consideration of sequential methods. Cheaper sequential methods are reasonably accurate, if the lateral correlation is

not high and if the constraints are weak. If the correlation is high and constraints are strong, naïve sequential simulation will poorly approximate the marginals. For such cases, the proposed approximate marginals offer improved sampling at a moderate cost.

A new approach to integrate stratigraphic data using surface-based models is also shown. This approach involves generating an ensemble of surface based models; this removes the need to condition one specific realization and also integrates the uncertainty of the surface-based models. Since this approach is Bayesian, the importance of seismic, well, and stratigraphic data are honored without overtuning to one particular information. All these methods, once explained, can be used in general when other information is available; yet the methods are not limited to surface-based models to generate the preprior.

Sampling rather than ranking, is proposed for selecting realizations from uncertain geologic models. In this research, tracer simulations a type of fast flow simulations are used to screen the realizations. The PCA and Hammerlsey sampling techniques are then used to select small but representative samples from the original set of realizations. This general procedure can be extended to other fast flow simulations, such as streamline simulations, and other sampling techniques like Latin hypercubes. New methods are proposed to create, screen, and analyze models in this powerful, but complex workflow.

Bibliography

Aki, Keiiti, and Paul G. Richards. 2002. *Quantitative Seismology: Theory and Methods.* New York: University Science Books.

Araktingi, U. G., and W. M. Bashore. 1992, October. "Effects of Properties in Seismic Data on Reservoir Characterization and Consequent Fluid-Flow Predictions When Integrated With Well Logs." *SPE Annual Technical Conference and Exhibition.* Washington, DC. SPE paper no.24752.

Araktingi, U. G., T. A. Hewett, and T. T. B. Tran. 1993. "GEOLITH: An Interactive Geostatistical Modeling Application." *SPE Computer Applications* 4 (April): 17–23.

Ates, H., A. Bahar, S. El-Abd, M. Charfeddine, M. Kelkar, and A. Datta-Gupta. 2005. "Ranking and Upscaling of Geostatistical Reservoir Models Using Streamline Simulation: A Field Case Study." *SPE Reservoir Engineering and Evaluation* 8, no. 1.

Bashore, W. M., U. G. Araktingi, Marjorie Levy, and W. J. Schweller. 1993, October. "The Importance of the Geological Model for Reservoir Characterization Using Geostatistical Techniques and the Impact on Subsequent Fluid Flow." *SPE Annual Technical Conference and Exhibition.* Houston, Texas. SPE paper no.26474.

Behrens, R. A., M. K. MacLeod, T. T. Tran, and A. C. Alimi. 1998. "Incorporating Seismic Attribute Maps in 3D Reservoir Models." *SPEREE* 1, no. 2 (April).

Behrens, R. A., and T. T. B. Tran. 1998. "Incorporating Seismic Attribute Maps in 3D Reservoir Models." *SPE Reservoir Evaluation and Engineering* 4:122–126.

———. 1999. "Incorporating Seismic Data of Intermediate Vertical Resolution Into Three-Dimensional Reservoir Models: A New Method." *SPE Reservoir Evaluation and Engineering* 2:325–333.

Bentley, Jon Louis. 1975. "Multidimensional Binary Search Trees Used for Associative Searching." *Commun. ACM* 18, no. 9 (September).

Cardwell, W. T., and R. L. Parsons. 1945. "Average Permeability of Hetreogeneous Porous Media." *Trans. AIME* 160:34–42.

Chessa, A. G. 1995. "Conditional Simulation of Spatial Stocastic Models for Reservoir Heterogeneity." Ph.D. diss., Delft University.

Cohen, A. C. 1949. "On Estimating the Mean and Standard Deviation of Truncated Normal Distributions." *Journal of the American Statistical Association* 44 (248): 518–525 (Dec).

Deutsch, C. V., and S. Srinivasan. 1996, April 21-24. "Improved Reservoir Management Through Ranking Stochastic Reservoir Models." *SPE/DOE Tenth Symposium on Improved Oil Recovery.* Tulsa, Oklahoma. SPE paper no. 35411.

Deutsch, C. V., S. Srinivasan, and Y. Mo. 1996, October 6-9. "Geostatistical Reservoir Modeling Accounting for Precision and Scale of Seismic Data." *SPE Annual Technical Conference and Exhibition*. Denver, Colorado. SPE paper no. 36497.

Deutsch, Clayton V., and André G. Journel. 1997. *GSLIB: Geostatistical Software Library and User's Guide*. Second. Oxford: Oxford University Press.

Dobrin, Milton B., and Carl H. Savit. 1988. *Introduction to Geophysical Prospecting*. Fourth. New York: McGraw-Hill Companies.

Doyen, P. M., L. D. den Boer, and D. Jans. 1997, October. "Reconciling Data at Seismic and Well Log Scales in 3-D Earth Modelling." *SPE Annual Technical Conference and Exhibition*. San Antonio, Texas. SPE paper no.38698.

Doyen, P. M., L. D. den Boer, and W. R Pillet. 1996, October. "Seismic Porosity Mapping in the Ekofisk Field Using a New Form of Collocated Cokriging." *SPE Annual Technical Conference and Exhibition*. Denver, Colorado. SPE paper no.36498.

Doyen, P. M., D. E. Psaila, L. D. den Boer, and D. Jans. 1997, October 5-8. "Reconciling Data at Seismic and Well Log Scales in 3-D Earth Modeling." *SPE Annual Technical Conference and Exhibition*. San Antonio, Texas.

Embry, Ashton F. 2002. "Transgressive - Regressive (T-R) Sequence Stratigraphy." *Gulf Coast Association of Geological Societies Transactions* 52:151–172.

Fletcher, R. 2000. *Practical Methods of Optimization*. Second. New York: John Wiley.

Gassmann, F. 1951. "Uber die Elastizitat Poroser Medien." *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zurich* 96:1–23.

Gelman, Andrew, John B. Carlin, Hal S. Stern, and Donald B. Rubin. 2003. *Bayesian Data Analysis*. Second. Boca Raton: Chapman and Hall.

Glinsky, M. E., B. Asher, R. Hill, M. Flynn, M. Stanley J. Gunning, T. Thompson, J. Kalifa, S. Mallat, C̃. D. White, and D. Renard. 2005. "Integration of Uncertain Subsurface Information into Multiple Reservoir Simulation Models." *The Leading Edge* 24 (10): 990–999.

Golub, Gene H., and Charles F. van Loan. 1996. *Matrix Computations*. Third. Baltimore: The Johns Hopkins University Press.

Goovaerts, Pierre. 1997. *Geostatistics for Natural Resources Evaluation*. Applied Geostatistics Series. Oxford: Oxford University Press.

Griffiths, C. M., C. P. Dyt, E. Paraschivoiu, and K. Liu. 2001. "Sedsim in Hydrocarbon Exploration." Edited by D. F. Merriam and J. C. Davis, *Geologic Modeling and Simulation: Sedimentary Systems*. New York: Kluwer Academic, Plenum Publishers, 71–97.

Gu, Yaqing. 2006. "Ensemble Kalman Filter for Reservoir Model Updating." Ph.D. diss., Mewbourne School of Petroleum and Geological Engineering, The University of Oklahoma.

Gunning, James, Michael E. Glinsky, and Christopher D. White. 2007. "*DeliveryMassager*: a Tool for Propagating Seismic Inversion Information into Reservoir Models." *Computers and Geosciences* 33 (5): 630–648 (May).

Gunning, James G., and Michael E. Glinsky. 2004. "Delivery: an Open-Sourced Bayesian Seismic Inversion Program." *Computers and Geosciences*, vol. 30.

———. 2006. "Wavelet Extractor: a Bayesian Well-Tie Wavelet Derivation Program." *Computers and Geosciences*, vol. 32.

Hair, Joseph F., Bill Black, Barry Babin, Rolph E. Anderson, and Ronald L. Tatham. 2005. *Multivariate Data Analysis*. New Jersey: Prentice Hall.

Haldorsen, H. H., and L. W. Lake. 1984. "A New Approach to Shale Management in Field Scale Simulation Models." *SPEJ Journal* 24 (August): 447–457.

Higdon, David M. 1998. "Auxiliary Variable Methods for Markov Chain Monte Carlo with Applications." *Journal of the American Statistical Association* 93 (442): 585–595.

Kalagnanam, J. R., and U. M. Diwekar. 1997. "An Efficient Sampling Technique for Off-Line Quality Control." *Technometrics* 39 (3): 308–319.

Kalla, S., C. D. White, and J. Gunning. 2007, 10-14 September. "Downscaling Seismic Data to the Meter Scale: Sampling and Marginalization." *EAGE Conference on Petroleum Geostatistics*. Cascais, Portugal.

Kalla, S., C. D. White, J. Gunning, and M. E. Glinsky. 2007, 11-14 November. "Imposing Multiple Seismic Inversion Constraints on Reservoir Simulation Models." *SPE Annual Technical Conference and Exhibition*. Anaheim, Ca. SPE paper no. 110771.

Kalla, Subhash, and Christopher D. White. 2005, October 9-12. "Efficient Design of Reservoir Simulation Studies for Development and Optimization." *SPE Annual Technical Conference and Exhibition*. Dallas, Texas. SPE paper no. 95456.

Kalla, Subhash, Christopher D. White, James Gunning, and Michael E. Glinsky. 2006, September 24-27. "Consistent Downscaling of Seismic Inversions to Cornerpoint Flow Models." *SPE Annual Technical Conference and Exhibition*. San Antonio, Texas. SPE paper no. 103268.

Kupfersberger, H., and C. V. Deutsch. 1999. "Ranking Stochastic Realizations for Improved Aquifer Response Uncertainty Assessment." *Journal of Hydrology* 223:54–65.

Lake, L. W., and J. L. Jensen. 1989. "A Review of Heterogeneity Measures Used in Reservoir Characterization." SPE USMS paper no.020156.

Latimer, Rebecca. 2006. "Uses, Abuses, and Examplees of Seismic-Derived Acoustic Impedance Data: What Does the Interpreter Need to Know?" *AAPG Bulletin*, vol. 90.

Lee, Sang Heon, Adel Malallah, Akhil Datta-Gupta, and David M. Higdon. 2002. "Multiscale Data Integration Using Markov Random Fields." *SPE Reservoir Evaluation and Engineering*, April, 68–78.

Li, D., B. Beckner, and A. Kumar. 1999, October. "A New Efficient Averaging Technique for Scaleup of Multimillion-Cell Geologic Models." *SPE Annual Technical Conference and Exhibition*. Houston, Texas. SPE paper no.56554.

Li, D., A. S. Cullick, and L. W. Lake. 1995. "Global Scale-Up of Reservoir Model Permeability with Local Grid Refinement." *Petroleum Science and Engineering* 14:1–13.

Li, Xin. 2008. "Continuous Reservoir Model Updating by Ensemble Kalman Filter on Grid Computing Architectures." Ph.D. diss., Louisiana State University. 199 pages.

Li, Xin, Christopher D. White, Zhou Lei, and Gabrielle Allen. 2007, 10-14 September. "Reservoir Model Updating by Ensemble Kalman Filter- Practical Approaches Using Grid Computing Technology." *EAGE Conference on Petroleum Geostatistics*. Cascais, Portugal.

Liddle, Andrew R. 2007. "Information Criteria for Astrophysical Model Selection." *Monthly Notices of the Royal Astronomical Society* 377 (1): 74–78.

Liner, Christopher L. 2004. *Elements of 3D Seismology*. Tulsa: PennWell.

Liu, Jun S. 2001. *Monte Carlo Strategies in Scientific Computing*. First. New York: Springer.

Matheron, G., H. Beucher, C. de Fouquet, and A. Galli. 1987, September 27-30,. "Conditional Simulation of the Geometry of Fluvio-Deltaic Reservoirs." *SPE Annual Technical Conference and Exhibition*. Dallas, Texas. SPE paper no. 16753.

Mavko, Gary, Tapan Mukerji, and Jack Dvorkin. 2003. *The Rock Physics Handbook: Tools for Seismic Analysis of Porous Media*. Cambridge University Press.

Michalak, A. M., and P.K. Kitanidis. 2003. "A Method for Enforcing Parameter Nonnegativity in Bayesian Inverse Problems with an Application to Contaminant Source Identification." *Water Resources Research* 39, no. 2.

Nocedal, Jorge, and Stephen J. Wright. 1999. *Numerical Optimization*. New York: Springer Science+Business Media, Inc.

Oliver, Dean S. 1994. "Incorporation of Transient Pressure Data into Reservoir Characterization." *In Situ* 18(3):243–275.

Oreskes, Naomi, Kristin Shrader-Frechette, and Kenneth Belitz. 1994. "Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences." *Science* 263

(5147): 641–646.

Paola, Chris, Jim Mullin, Chris Ellis, David C. Mohrig, John B. Swenson, Gary Parker, Tom Hickson, Paul L. Heller, Lincoln Pratson, James Syvitski, Ben Sheets, and Nikki Strong. 2001. "Experimental Stratigraphy." *Geological Society of America* 11 (7): 4–9.

Peaceman, Donald W. 1993. "Calculation of Transmissibilities of Gridblocks Defined by Arbitrary Cornerpoint Geometry." unsolicited manuscript, SPE paper no. 37306.

Ponting, David K. 1989. "Corner Point Grid Geometry in Reservoir Simulation." *Proc., First European Conference Math. Oil Recovery*. Cambridge.

Pyrcz, Michael J. 2004. "The Integration of Geologic Information into Geostatistical Models." Ph.D. diss., University of Alberta. 250 pages.

Pyrcz, Michael J., Octavian Catuneanu, and Clayton V. Deutsch. 2005. "Stochastic Surface-Based Modeling of Turbidite Lobes." *AAPG Bulletin* 89 (2): 177–191 (February).

R Development Core Team. 2007. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.

Reading, Harold G. 1996. *Sedimentary Environments: Processes, Facies and Stratigraphy*. Third. Bognor Regis: Wiley-Blackwell.

R. J. Le Blanc, Sr. 1976, October 3-6. "Distributions and Continuity of Sandstone Reservoirs – Part 1." *SPE Annual Fall Technical Conference and Exhibition*. New Orleans, Louisiana. SPE paper no. 6137.

Russell, B., and D. Hampson. 1991. "A Comparision of Post-Stack Seismic Inversion Methods." *Society of Exploration Geophysists*. 876–878. Ann. Mtg. Abstracts.

Saad, N., V. Maroongroge, and C. T. Kalkomey. 1996, April. "Ranking Geostatistical Models Using Tracer Production Data." *SPE European 3-D Reservoir Modeling Conference*. Stavanger, Norway. SPE paper no.35494.

Schlumberger Technology Co. 2004. *Eclipse 100 97A Reference Manual*. Oxfordshire: Schlumberger Technology Co.

———. 2005, March. *Petrel Workflow Tools, Introduction Course V.2004*. Schlumberger Technology Co.

Shook, G. M., and J. H. Forsmann. 2005. "Tracer Interpretation Using Temporal Moments on a Spreadsheet." *Idaho National Laboratory*. INL/EXT-05-00400.

Strebelle, Sebastien. 2002. "Conditional Simulation of Complex Geological Structures Using Multi-Point Statistics." *Mathematical Geology* 34 (1): 1–21.

Tarantola, Albert. 2004. *Inverse Problem Theory and Methods for Model Parameter Estimation*. Philadelphia: SIAM.

Wang, Zhijing. 2001. "Fundamentals of Seismic Rock Physics." *Geophysics* 66 (2): 398–412 (March).

Weber, K. J., and L. C. van Geuns. 1990. "Framework for Constructing Clastic Reservoir Simulation Models." *Journal of Petreoleum Technology* 42 (October): 1248–1297.

Wellner, Robert William, David Hoyal, Ben Sheets, Tao Sun, John Van Wagoner, and Anthony Sprague. 2007. "A New Model for Development of River-Dominated Deltas." *AAPG Bulletin*, vol. 91.

White, C. D., B. J. Willis, J. P. Bhattacharya, S. P. Dutton, and K. Naryanan. 2004. "Sedimentologic, Statistical, and Flow Characteristics of a Tide-Influenced Delta." Edited by G. M. Grammer, P. M. Harris, and G. P. Erbeli, *Integration of outcrop and modern analogs in reservoir modeling*, Volume 80 of *AAPG Memoir Series*. Tulsa: AAPG, 134–152.

Widess, M. B. 1973. "How Thin is a Thin Bed?" *Geophysics* 38 (6): 1176–1180 (December).

Willis, Brian J., and Christopher D. White. 2000. "Quantitative Outcrop Data for Flow Simulation." *Journal of Sedimentary Research* 70:788–802.

Xu, Wenlong, and A. G. Journel. 1993. "GTSIM: Gaussian Truncated Simulations of Reservoir Units in a West Texas Carbonate Field." unsolicited manuscript, SPE paper no. 27412.

Appendix A: Eclipse Data Input File

```
RUNSPEC

TITLE
3d Waterflood Simulation for Screening

DIMENS
- - NDIVIX NDIVIY NDIVIZ
100 100 10 /

OIL

WATER

FIELD

WELLDIMS
- - NWMAXZ NCWMAX NGMAXZ NWGMAX
10 10 10 10 /

UNIFOUT

NSTACK
100 /

START
1 'JAN' 2000 /

GRID
INIT

INCLUDE
GRID_DATA /

COPY
PERMX PERMZ /
PERMX PERMY /
/

MULTIPLY
PERMZ .2 /
/
```

```
MINPV
20 /

PINCH
0.1 'NOGAP' /

PROPS

GRAVITY
- -Oil Water Gas
50.0 1.05 0.60 /

Rock
- -RefP Cf
7500 8.3E-6 /

PVDO
- -PRES FVF VIS
400 1.0120 1.160
1200 1.0040 1.164
2000 0.9960 1.167
2800 0.9880 1.172
3600 0.9802 1.177
4400 0.9724 1.181
5200 0.9646 1.185
5600 0.9607 1.190
6000 0.9560 1.195
/

- - RELATIVE PERMEABILITY AND CAPPILARY PRESSURE CURVES
SWOF
0.27 0.000 0.900 0
0.35 0.012 0.596 0
0.40 0.031 0.438 0
0.45 0.060 0.304 0
0.50 0.099 0.194 0
0.55 0.147 0.109 0
0.60 0.204 0.048 0
0.65 0.270 0.012 0
0.70 0.346 0.000 0
0.75 0.432 0.000 0
0.80 0.527 0.000 0
0.85 0.631 0.000 0
0.90 0.744 0.000 0
0.95 0.867 0.000 0
1.00 1.000 0.000 0 /
```

```
PVTW
- - RefP Bw Cw Visc Viscosibility
7500 1.0211 2.68E-06 0.3959 0 /

RSCONST
- - RS BUBPRES
0.8 50 /

SOLUTION

EQUIL
9875 6000 20480 0 100 0 1 /

SUMMARY

RUNSUM
SEPARATE
TCPU
RPTONLY

FOPT
FWPT
FWPR

WWIR
'I1' /

WOPR
'P2'
'P3'
'P4' /

WBHP
'I1'
'P2'
'P3'
'P4' /

WWPR
'P2'
'P3'
'P4' /

SCHEDULE
```

```
RPTSCHED
'RESTART=1' 'FIP=1' /

WELSPECS
'I1' 'P' 1 1 1* 'water' /
'P2' 'P' 1 100 1* 'OIL' /
'P3' 'P' 100 1 1* 'OIL' /
'P4' 'P' 100 100 1* 'OIL' /
/

COMPDAT
'I1' 1 1 1 10 'OPEN' 1* 1* 0.75 1* 0 /
'P2' 1 100 1 10 'OPEN' 1* 1* 0.75 1* 0 /
'P3' 100 1 1 10 'OPEN' 1* 1* 0.75 1* 0 /
'P4' 100 100 1 10 'OPEN' 1* 1* 0.75 1* 0 /
/

WCONINJE - - Name Phase Status Mode Qsc, Lsc, BHP
'I1' 'WATER' 'OPEN' 'RATE' 5000 1* 8000 /
/

WCONPROD - - NAME Status Mode Qo,Qw,Qg,Ql,Qr,BHP
'P2' 'OPEN' 'BHP' 5* 4000 /
'P3' 'OPEN' 'BHP' 5* 4000 /
'P4' 'OPEN' 'BHP' 5* 4000 /
/

TUNING
/
/ - - LITMIN LITMIN MXWSIT
20 1 100 1 8 /

TSTEP
200*5/
TSTEP
48*30.4 /

END
```

Appendix B: Nomenclature

Roman Symbols

a	random number
b	variogram range
$\tilde{\mathbf{C}}$	covariance matrix in new coordinates, m^2
$\mathbf{C}_p$	prior covariance matrix based on kriging, m^2
$\mathbf{C}_\pi$	posterior covariance matrix, m^2
$\mathbf{d}$	neighboring conditioning
$\mathbf{G}$	posterior resolution matrix or Hessian, m^{-2}
h	nonnegative layer thickness, m
H	total thickness at trace, m
I	facies indicator (1 for pay, 0 otherwise)
$\mathbf{L}$	Cholesky factor of covariance matrix, m
$\mathbf{m}$	all variables simulated at trace
$\mathbf{n}$	normal vector to a surface
$N(\mu, \sigma^2)$	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
N_{pD}	recovery factor at breakthrough, pore volumes
p	probability density
P	probability
P_J	Jacobian term in Metropolis-Hastings transition
$\mathbf{R}$	subspace in transformed coordinates; vector for the subspace is $\mathbf{r}$
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

$\mathbf{u}$	vector orthogonal to the first-quadrant hypersurface facet
U	uniform distribution function
$\mathbf{U}$	rotation matrix
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

Greek Symbols

α	Metropolis-Hastings transition probability
δ	magnitude of the direction vector $\mathbf{u}$
$\delta(t)$	Dirac delta function
Δ	separation vector for variogram models, m
γ	semivariogram model
κ	number of layers at a trace with $t_k > 0$
λ	dimensionless range, b/L
ν	coefficient of variation
ϕ	layer porosity
Φ or $\bar{\Phi}$	trace average porosity
π	posterior
σ^2	variance
$\boldsymbol{\tau}$	vector of properties in the transformed axes

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
s	sand
sh	shale
λ, Λ	zero thickness data index and count

Diacritical Marks

$\bar{\cdot}$	mean
$\cdot'$	proposed point, may become new point
$\tilde{\cdot}$	rotated
$\cdot^*$	truncated

Appendix C: SPE Permissions

SPE Copyright Requirements

Author-Retained Rights and Terms and Conditions

The term "employers" in the following means the companies, universities, or organizations for which the authors worked at the time the paper was written.

1. Authors/employers retain all intellectual property rights, including any idea, process, procedure, or article of manufacture described in the paper.

2. Authors/employers may reproduce and distribute copies of the paper internally to employees of the company or organization for which the author worked at the time the paper was written. Such distribution includes posting of the paper on the employer's intranet accessible only to company employees.

3. Authors/employers may reproduce, or authorize reproduction of, and distribute up to 50 paper copies of the paper outside the company or organization for which the author worked at the time the paper was written for personal, business, or educational purposes provided that the SPE copyright notice is included, the copies are not used in any way that implies SPE endorsement of a product or service, and the copies themselves are not offered for sale.

4. Authors/employers may make an oral presentation of the same material provided proper acknowledgement of SPE copyright ownership is made.

5. Authors/employers may incorporate all or part of the paper in future writings or presentations. If the entire paper or a portion thereof is used in substantially unchanged form, proper acknowledgement of SPE copyright must be made. If the paper is substantially changed or altered in substance so that it is a new work of authorship, reference should be made to the SPE publication.

6. Authors/employers may request return of one-time journal publication rights to enable publication of the paper in a journal or magazine if the paper is not being considered for publication in an SPE journal. Such requests should be made in writing to SPE Customer Service. Requests for return of one-time journal publication rights will not be granted for papers submitted to SPE for peer review unless the paper is declined for publication or it is at least 6 months after the submission date.

7. In the case of work performed under a U.S. government contract or grant, SPE recognizes that, if the contract/grant so requires, the U.S. government has royalty-free permission to reproduce all or portions of the paper and to authorize others to do so for official U.S. government purposes only.

Company or companies as used herein means the company the author worked for at the time the paper was written, such companys parent (if any), and any company for which the company (or such ultimate parent) owns or controls, directly or indirectly, fifty percent or more of the stock or voting rights.

8. For all uses not covered by Items 2 through 6, authors/employers must request permission from SPE to reproduce or authorize reproduction of the paper.

9. Although authors are permitted to re-use all or portions of the paper in other works, this does not include granting third-party requests for reprinting, republishing, or other types of re-use. SPE must handle all such third-party requests

Vita

Subhash Kalla was born in August, 1981, in Madhavaram, Andhra Pradesh, India. He finished his undergraduate studies at Regional Engineering College, Warangal, in May 2002. He earned a master of science degree in petroleum engineering from Louisiana State University in August 2005. He is currently a candidate for the degree of Doctor of Philosophy in petroleum engineering, to be awarded in December 2008.