3-D GEOSTATISTICAL SEISMIC INVERSION WITH WELL LOG CONSTRAINTS

Jonathan Kane, William Rodi, and M. Nafi Toksöz

Earth Resources Laboratory
Department of Earth, Atmospheric, and Planetary Sciences
Massachusetts Institute of Technology
Cambridge, MA 02139

ABSTRACT

Information about reservoir properties usually comes from two sources: seismic data and well logs. The former provide an indirect, low resolution image of rock velocity and density. The latter provide direct, high resolution (but laterally sparse) sampling of these and other rock parameters. An important problem in reservoir characterization is how best to combine these data sets, allowing the well information to constrain the seismic inversion and, conversely, using the seismic data to spatially interpolate and extrapolate the well logs.

We develop a seismic/well log inversion method that combines geostatistical techniques for well log interpolation (i.e., kriging) with a Monte Carlo search method for seismic inversion. We cast our inversion procedure in the form of a Bayesian maximum *a posteriori* (MAP) estimation in which the prior is iteratively modified so that the algorithm converges to the model that maximizes the likelihood function.

We follow the approach used by Haas and Dubrule (1994) in their sequential inversion algorithm. Kriging is applied to the well data to obtain velocity estimates and their covariances for use as *a priori* constraints in the seismic inversion. Inversion of a complete 3-D seismic section is performed one trace at a time. The velocity profiles derived from previous seismic traces are incorporated as "pseudo well logs" in subsequent applications of kriging. Our version of this algorithm employs a more efficient Monte Carlo search method in the seismic inversion, and moves sequentially away from the wells so as to minimize the kriging variance at each step away from the inverted wells.

Numerical experiments with synthetic data demonstrate the viability of our seismic/well data inversion scheme. Inversion is then performed on a real 3-D data set provided by Texaco.

INTRODUCTION

The most accurate method for obtaining information about the subsurface properties of the Earth is to drill a hole, extract the rocks, and/or use well logging tools to sample petrophysical properties therein. This, unfortunately, is too expensive and difficult to do in more than a few sparse, isolated locations. To obtain the values of a petrophysical property over a larger three-dimensional domain requires the accurate extrapolation of known values as well as the use of indirect information supplied by remote sensing techniques.

This situation is germane to the petroleum industry, which seeks to infer the location of hydrocarbons given sparse well log data and indirect seismic data. The density and compressional wave velocity of seismic waves can be sampled at a small scale in the well logs and at a large scale, indirectly, with seismic waves. Results of decades of research have been applied to inferring the velocity and density of the Earth from seismic data. Since the 1980's this research has been applied to three dimensional exploration. Standard techniques for performing this inference are illustrated in several sources, including Yilmaz (1987). Typically, these methods seek only to locate singularities in the subsurface parameters, i.e., locations at which the petrophysical parameter fields are nondifferentiable (such as jump discontinuities or thin beds). Migration is a technique which takes seismic data and "migrates" the singularities back to their point of origin, either in time (time migration) or depth (depth migration). A summary of migration techniques can be found in Gardner (1985). Therefore, it is standard practice not to infer the actual values of the parameter of interest, but rather to generate images of the geometry of subsurface structures through migration. We will refer to standard methods of obtaining images of subsurface structures as "imaging," and call the actual estimation of the petrophysical parameter values "inversion." In this paper, we attempt to provide a methodology for taking seismic data that have been time migrated along with a few well logs and perform inversion for acoustic velocity. Although density is also a relevant parameter in seismic wave propagation, we assume it constant for computational reasons.

We first give a brief overview of random field theory. This provides a framework for describing a field of petrophysical parameters in probabilistic fashion. We then use this framework to pose the inverse problem of inferring a wave velocity field given sets of data related to it by an operator. Inversion is performed on a synthetic 2-D data set and then on a synthetic 3-D data set. This allows for an assessment of the performance of the inversion method. Finally, we demonstrate the inversion method on a real 3-D data set provided by Texaco.

STOCHASTIC DESCRIPTION OF GEOLOGY

Let us denote a parameter field of interest as $v(\mathbf{x})$, where v is defined over a spatial field $\mathcal{D}$ which can contain either a finite, countably infinite, or uncountably infinite set of locations $\mathbf{x}$. A description of a random field requires a multivariate probability density $p(\cdot)$ to be defined over $v(\mathbf{x})$. The parameter value at each spatial location is then a random variable (RV) and has its own marginal probability density function as well as a (possibly infinite) number of joint moments with the RVs at other spatial locations.

We will make certain assumptions which simplify analysis and the usage of random fields. First, we assume that the RVs are jointly Gaussian. Then, the joint probability distribution has the following form:

$$p(v(\mathbf{x})) = \frac{\exp[-\frac{1}{2}(v(\mathbf{x}_i) - \mu_v(\mathbf{x}_i))^T \mathbf{C}_v(\mathbf{x}_i, \mathbf{x}_j)^{-1}(v(\mathbf{x}_i) - \mu_v(\mathbf{x}_i))]}{(2\pi)^{\frac{N}{2}} |\mathbf{C}_v|^{\frac{1}{2}}}$$

$$\text{where} \quad N := \text{number of elements in} \quad \mathbf{v}.$$

From this we see that a Gaussian field is fully described by its bivariate moments (covariances) and its mean. The covariance of such a field is defined as:

$$\mathbf{C}_v(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{C}_v(v(\mathbf{x}_i), v(\mathbf{x}_j)) = E[v(\mathbf{x}_i)v(\mathbf{x}_j)] - \mu_{\mathbf{x}_i}\mu_{\mathbf{x}_j} \tag{1}$$

where

$$\mu_{\mathbf{x}_i} = E[v(\mathbf{x}_i)].$$

The function $\mathbf{C}(\cdot)$ is known as the *covariance function* and describes the covariance between every two points in the region $\mathcal{D}$. It is also known as the (auto-)covariogram, or, if normalized to have a maximum of 1, the (auto-)correlogram. We assume that $\mu_{\mathbf{x}_i} = 0$ for simplification.

A second assumption that simplifies the description of a random field is to assume that the random field of interest is *stationary*. Roughly, this implies that the statistics of the random field do not change with spatial location. The covariance function then takes a simple form that depends only on spatial separation, not spatial location:

$$\mathbf{C}(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{C}(\mathbf{s})$$

$$\mathbf{s} = \mathbf{x}_i - \mathbf{x}_j.$$

The distance it takes for this function to go from its maximum to almost 0 is known as the *range* or *correlation length*. This covariance function can take a number of forms. Two common ones are Gaussian and exponential. These functions have the following forms, respectively,

$$\mathbf{C}(\mathbf{s}) = \sigma^2 e^{-\mathbf{s}^2/a^2} \quad , \quad \mathbf{C}(\mathbf{s}) = \sigma^2 e^{-|\mathbf{s}|/a}$$

where a is the correlation length and σ^2 is the variance at 0 separation. We use the Gaussian covariance function exclusively.

An alternate, but equivalent, representation of the covariance function exists called the *(semi-)variogram*, or *structure function*, $2\gamma(\mathbf{s})$. The relationship between the variogram and the covariance function in the case of a zero-mean, stationary random field is:

$$2\gamma(\mathbf{s}) = 2(\mathbf{C}(0) - \mathbf{C}(\mathbf{s})).$$

We show that the variogram exists for certain random fields for which the covariance function does not (nonstationary processes that have stationary increments). An example of such a process is Brownian motion, which has the following power law variogram:

$$2\gamma(\mathbf{s}) = \alpha \mathbf{s}^{\beta}$$

$$\alpha \geq 0 \qquad 0 \leq \beta < 2.$$

A third, and necessary, simplifying assumption is to discretize the parameter field over the region $\mathcal{D}$. This allows numerical manipulation and estimation to be performed on a computer. The covariance structure of a discrete, countably finite Gaussian field can be put into a table that lists the covariance between each two RVs in the field. This table is known as the covariance matrix. If $\mathbf{v}$ is real then $\mathbf{C}$ is automatically symmetric. The assumption of stationarity causes the covariance matrix to have a *Toepplitz* form, i.e., it has constant diagonals.

Thus, we are limiting ourselves to a countably finite number of parameters in $\mathcal{D}$. This final simplification inherently assumes that our discretization of the continuous field is sufficient to capture all relevant details.

Generating Realizations From Random Fields

Covariance functions lend themselves to generating realizations of random fields, which is required by the Monte Carlo inversion method used in this paper. This requires the generation of a random vector $v(\mathbf{x})$ such that it has the prescribed $\mathbf{C}(\mathbf{s})$. Assuming a zero mean field, the following line of reasoning demonstrates how to generate the desired $v(\mathbf{x})$. We henceforth write $v(\mathbf{x}_i)$ as $\mathbf{v}$ for brevity and use $\mathbf{v}^T$ to denote its transpose.

A computationally efficient method for generating realizations makes use of the Fourier transform. Denote the Fourier transform of a finite dimensional vector as $\mathbf{F}$. We find that for a symmetric, positive definite matrix $\mathbf{C}$ we have the following relationship: $\mathbf{C} = \mathbf{F}\Lambda\mathbf{F}^T$ where Λ is the *spectrum* of the covariance function. Λ has the useful property that it is a diagonal matrix with $\sqrt{\Lambda} = \sqrt{\Lambda}^T$. This allows us to do the following decomposition:

$$\begin{aligned} E[\mathbf{v}\mathbf{v}^T] &= \mathbf{F}\Lambda\mathbf{F}^T = \mathbf{F}\sqrt{\Lambda}\sqrt{\Lambda}\mathbf{F}^T \\ &= \mathbf{F}\sqrt{\Lambda}\mathbf{I}\sqrt{\Lambda}\mathbf{F}^T = \mathbf{F}\sqrt{\Lambda}E[\mathbf{w}\mathbf{w}^T]\sqrt{\Lambda}\mathbf{F}^T \\ &= E[\mathbf{F}\sqrt{\Lambda}\mathbf{w}\mathbf{w}^T\sqrt{\Lambda}\mathbf{F}^T] = E[\mathbf{F}\sqrt{\Lambda}\mathbf{w}(\mathbf{F}\sqrt{\Lambda}\mathbf{w})^T]. \end{aligned}$$

This tells us that $\mathbf{v} = \mathbf{F}\sqrt{\Lambda}\mathbf{w}$. Thus we are filtering white noise to obtain a realization of our random field. We use this method throughout the paper.

BAYESIAN INVERSION

We present standard results of Bayesian inversion. For derivation and rigor see Tarantola (1987). Inversion theory seeks to infer the value of a vector, $\mathbf{v} \in \mathcal{V}$, given only a data vector $\mathbf{y} \in \mathcal{Y}$, a noise vector $\mathbf{n} \in \mathcal{N}$, and a mapping, $H(\mathcal{V}) \mapsto \mathcal{Y}$. We call the set of elements $\mathbf{v} \in \mathcal{V}$ the *model space* and the elements $\mathbf{y} \in \mathcal{Y}$ the *data space*. Mapping $H(\mathcal{V}) \mapsto \mathcal{Y}$ can be noninvertible (no existence of an inverse) or not uniquely invertible (the existence of many inverses for a given data vector $\mathbf{y}$). Also, the inverse mapping $H^{-1}(\mathcal{Y})$ may not be a continuous mapping even if $H(\mathcal{V})$ is. This amounts to the problem that slight perturbations in $\mathbf{y}$ lead to large perturbations in $H^{-1}(\mathbf{y})$.

Each element of a vector $\mathbf{v}$, $\mathbf{n}$, or $\mathbf{y}$ is itself an element in another vector space. This is the vector space of finite variance random variables. The addition of this structure upon $\mathcal{V}$, $\mathcal{N}$, and $\mathcal{Y}$ makes $\mathbf{v}$, $\mathbf{n}$, and $\mathbf{y}$ random vectors and allows us to perform Bayesian inversion.

To illustrate Bayesian inversion, we use a simple example common to many fields of science. The problem is that of extrapolation. Extrapolation can be posed in the following way: Let vector $\mathbf{v}$ represent an element of the model space and mapping H_K be a subset of the rows of the identity matrix, $\mathbf{I}$. The operator selects which model parameters are observed and (possibly) adds some noise to them to give us the observed data $\mathbf{y}$. $\mathbf{v}$ and $\mathbf{n}$ are random Gaussian vectors with associated covariance matrices, $\mathbf{C_v}$ and $\mathbf{C_n}$.

We write the forward problem as:

$$\mathbf{y} = H_K\mathbf{v} + \mathbf{n}.$$

Matrix H_K will be an $P \times Q$ matrix, where P is the number of observed elements $\mathbf{y}$, and Q is the dimension of $\mathbf{v}$. Thus, we have $P < Q$ and are faced with the problem of nonuniqueness. There are many models that fit the data, i.e., there are many interpolations that go through the data points but are different elsewhere.

Rearranging, we have:

$$\mathbf{n}(\mathbf{v}, \mathbf{y}) = \mathbf{y} - H_K\mathbf{v}.$$

We call this the *noise* function over the joint model and data space. Sometimes this is called the *error* function, but we will reserve the word *error* for the norm of the misfit function, that is,

$$e(\mathbf{v}, \mathbf{y}) = \|\mathbf{n}(\mathbf{v}, \mathbf{y})\|^p$$

for some $1 \leq p$. For a given $\mathbf{y} = \mathbf{y}_{obs}$, we seek as our estimate:

$$\hat{\mathbf{v}} = \min_{\mathbf{v}} E(e(\mathbf{v}, \mathbf{y} = \mathbf{y}_{obs})).$$

Standard results give us the following expressions for the estimate $\hat{\mathbf{v}}$ and its covariance matrix:

$$\hat{\mathbf{v}} = \mathbf{C_{vy}}\mathbf{C_y^{-1}}\mathbf{y}$$

$$\mathbf{C}_{\hat{\mathbf{v}}} = \mathbf{C_v} - \mathbf{C_{vy}}\mathbf{C_y^{-1}}\mathbf{C_{vy}^T},$$

where $\mathbf{C_{vy}}$ is the *cross*-covariance between $\mathbf{v}$ and $\mathbf{y}$. In Earth science these equations are known as the *kriging* system of equations and are used to extrapolate sparse well data to unknown locations. $\mathbf{C_v}$ and $\mathbf{C}_{\hat{\mathbf{v}}}$ are often called the *prior* and *posterior* covariances, respectively.

For cases in which H is a nonlinear operator, inversion becomes more difficult. The optimal answer depends on which norm is used to define the error. The minimization procedure is more difficult because the error function is no longer quadratic and may have more than one inflection point. The goal remains to find the minimum of $E(e(\mathbf{v}, \mathbf{y} = \mathbf{y}_{obs}))$ but $\hat{\mathbf{v}}$ is no longer a linear function of $\mathbf{y}$. Thus, a procedure is required that iterates over either the inverse or forward operator in order to minimize $E(e(\mathbf{v}, \mathbf{y}))$. In the following sections, we use a Monte Carlo method for nonlinear estimation.

Nonlinear Seismic Operator

The seismic forward modeling operator used in this paper, denoted $H_S(\cdot)$, is the most simple operator possible. It is known as the convolutional model. It assumes vertical incidence of plane seismic waves, no conversion to S-waves, no multiple scattering, and no loss of amplitude in the seismic signal as it travels through the medium. In order for this model to hold, the geology must be horizontally stratified and the raw seismic data time migrated. This processing produces a seismic trace at each vertical location of the model that is a function only of the acoustic velocity and density at that vertical location. We make a further simplification by assuming that the density is constant, which greatly reduces the computational complexity and allows us to invert for only one parameter.

$H_S(\cdot)$ consists of three operations. First, an operation converts the vertical velocity field as a function of depth into velocity as a function of two-way traveltime:

$$v(z) \mapsto v'(t).$$

v' is then resampled onto a regular time series. The next operation produces the *reflectivity series*:

$$r(t) = \frac{1}{2}\frac{d}{dt}\ln v'(t) \cong \frac{v'(t+\delta t) - v'(t)}{v'(t+\delta t) + v'(t)}.$$

Finally, $r(t)$ is convolved with known seismic wavelet, $w(t)$, to produce the seismogram:

$$s(t) = r(t) * w(t).$$

For a derivation of this operator see Sengbush *et al.* (1960).

We write the relation between seismograms and velocity as

$$\mathbf{s} = H_S(\mathbf{v}) + \mathbf{n}.$$

Although the dimension of $\mathbf{s}$ may be greater than $\mathbf{v}$, there is no unique inverse to this operator. There are an infinite number of velocity vectors that all result in the same seismogram. To verify this, simply multiply $v'(t)$ by a constant and the resulting seismogram does not change.

In the next section, we address how to find an inverse to the seismic operator and how to constrain it to choose one velocity vector from the infinite that fit the data.

Monte Carlo Inversion

In recent years, adaptive search methods have gained popularity as a tool for finding optimal solutions to geophysical inverse problems. Their popularity has been motivated by vast increases in computing power and the fact that these methods are more effective than traditional gradient-based methods in finding globally, rather than just locally, optimal solutions to nonlinear problems. Recent applications of adaptive search methods to seismic inversion include genetic algorithms (Sen *et al.*, 1991), simulated annealing (Rothman, 1985; Stoffa and Sen, 1991; Vestergaard and Mosegaard, 1991), and iterative improvement (Haas and Dubrule, 1994).

The Monte Carlo method for estimating a solution to an inverse problem differs greatly from the linearized inverse methods presented above. The Monte Carlo method poses the inverse problem as an integration:

$$\hat{\mathbf{v}} = \frac{1}{\mathcal{V}} \int_{\mathcal{V}} e(\mathbf{s} = \mathbf{s}_{obs}, \mathbf{v}) d\mathbf{v}$$

where

$$e(\mathbf{s} = \mathbf{s}_{obs}, \mathbf{v}) = \|\mathbf{s}_{obs} - H_S(\mathbf{v})\|^p$$

$$p \geq 1 \quad , \quad \mathcal{V} := \text{region of integration.}$$

$H_S(\cdot)$ is the convolutional seismic operator from the previous section and $\mathbf{s}_{obs}$ is the observed seismic data. Depending on the norm used to define the error function, a different solution will be obtained in the integration. The $p = 1$ norm gives the median of posterior distribution, the $p = 2$ norm gives the mean, and the $p = \infty$ norm leads to a MAP solution. Furthermore, if the integrand can be split into two functions, we can obtain the same solution by sampling from one function and evaluating the other. This situation applies to Bayesian inversion where the posterior distribution can be split into

a prior and a likelihood function. It is usually easier to sample from the prior than the posterior[1].

The Monte Carlo method approximates the above integral by a summation over the number of samples. It can be shown that, as the number of samples goes to ∞, the sample estimate approaches the solution $\hat{\mathbf{v}}$.

The search algorithm employed by Haas and Dubrule (1994) is a simple Monte Carlo method, whereby the velocity estimate[2], $\hat{\mathbf{v}}$, obtained from a given seismic trace, $\mathbf{s}$, is a realization of a specified random process. Many realizations of the process are generated and the one yielding the maximum correlation is chosen as the solution. This is an easier criterion to fit than minimizing error. We will call Haas and Dubrule's method *iterative improvement*. The prior distribution that they use to generate velocity realizations is the posterior mean and covariance resulting from kriging well data.

Our Monte Carlo search algorithm differs in two respects. First, we are minimizing error, not maximizing correlation. Second, we borrow a useful concept from simulated annealing: We allow the process from which realizations are generated to change adaptively during the search. In simulated annealing, this is accomplished by specifying a temperature schedule, whose choice is one of the most difficult aspects of the algorithm. Our approach is to gradually shift the mean and reduce the variance of the prior such that, like simulated annealing, the search becomes more and more local as better solutions are found. The "annealing" in our algorithm is not controlled by a temperature schedule, but is determined only by the total number of trials. Our algorithm is basically a MAP estimation with a prior that progressively moves toward the minimum error model. This makes it a maximum likelihood estimator.

Proposed Monte Carlo Method

Our new Monte Carlo inversion method is easy to apply and provides good results for the convolutional seismic problem. The crux of this method is based on some properties of the following vector

$$\mathbf{v} = \alpha\mathbf{r} + \beta\mathbf{z},$$

where $\mathbf{r}$ and $\mathbf{z}$ are independent Gaussian vectors. The covariance matrix, $\mathbf{C_v}$, of $\mathbf{v}$ can therefore be expressed as

$$\mathbf{C_v} = \alpha^2\mathbf{C_r} + \beta^2\mathbf{C_z}.$$

If we further impose the constraint that

$$\mathbf{C_v} = \mathbf{C_r} = \mathbf{C_z} = \mathbf{C},$$

[1]There are methods to sample directly from the posterior. Foremost among them is the Metropolis sample rejection method, which, in the case of posterior distribution estimation, leads to Gibbs sampling and in the case of maximum likelihood estimation leads to simulated annealing.

[2]Actually, Haas and Dubrule (1994) use acoustic impedance, but this equals velocity in our case of constant density.

the constants α and β must satisfy

$$\alpha^2 + \beta^2 = 1.$$

We therefore have three zero mean Gaussian random vectors, $\mathbf{v}, \mathbf{r}, \mathbf{z}$, with the same covariance matrix. Below, we clarify the purpose of constructing $\mathbf{v}$ in such a way.

The algorithm proceeds as follows:

1. Perform kriging at a vertical location adjacent to a well log using well data within a specified distance of the kriged location. This gives a kriged mean vector, $\mu_{\mathbf{v}}$ for the pseudo well log and a kriged covariance matrix, $\mathbf{C_v}$, which is approximately stationary as a function of depth. This can be written as:

$$\mathbf{v} = \mu_{\mathbf{v}} + \mathbf{v}^*.$$

2. Specify the total number of iterations N_{tot} to be performed.

3. For $i = 1 \ldots N_{tot}$, let

$$\mathbf{v}_i^* = \alpha_i \mathbf{r}_i + \beta_i \mathbf{z}_i$$

where

$$\mathbf{z}_i = \mathbf{F}\sqrt{\Lambda}\mathbf{w}_i$$

and

$$\mathbf{r}_i = \mathbf{v}_{i-1}^* \quad \text{if} \quad e_{i-1} < e_{i-2}$$

or

$$\mathbf{r}_i = \mathbf{r}_{i-1} \quad \text{if} \quad e_{i-1} \geq e_{i-2}.$$

4. As the algorithm proceeds, let

$$\alpha^2 = \frac{N_{tot} - i + 1}{N_{tot}} \quad \text{and} \quad \beta^2 = \frac{i-1}{N_{tot}}.$$

Thus, more and more weight will be applied to $\mathbf{r}$ while still ensuring that all realizations of $\mathbf{v}^*$ have covariance $\mathbf{C_v}$.

KRIGING AND SEISMIC INVERSION ON SYNTHETIC DATA

In this section, we apply the above inverse methods to synthetic data sets, first in 2-D and then in 3-D. A 2-D realization of a Gaussian random field is generated by the Fourier method. This field will be used to represent acoustic wave velocity. The input stochastic parameters needed to generate this field are the covariance function and the mean. C(s) in this case is an anisotropic Gaussian function with a horizontal correlation length $a_x = 200$ ft and a vertical correlation length, $a_z = 10$ ft. The variance of this field is $\sigma^2 = 250{,}000$ ft^2/s^2 and the mean is $\mu = 5000$ ft/s. The realization is shown in Figure 1. In performing all synthetic inversion, we assume we know the stochastic parameters *a priori*.

A 2-D seismic data set is generated from the velocity field using the $H_S(\cdot)$ operator discussed above (see Figure 2). For inversion purposes we have, *a priori*, the seismic wavelet used to create the synthetic seismic data. Estimating this wavelet in reality is a difficult task (see below).

Two wells are also extracted from this data set and, along with the seismic data, are used to do the inversion. The synthetic wells are shown in Figure 3. In this case, we can test the efficiency of the inversion process because we have the true model to compare it to after inverting the data.

We want to make optimal use of all the data in performing the inversion. Since each seismogram is dependent only on the velocity field at the same vertical location, we can sequentially invert one seismogram at a time. The first question is: "In what order should it be done?" The intuitively logical answer is to start inverting seismograms near the location of existing wells and proceed outward. If there is any horizontal similarity in the velocity field (which we know *a priori* there is) then this procedure will make inversion easier by constraining the inversion method to look for solutions similar to adjacent well data. This brings us to the second question: "How do we make use of the well data that we have?" The answer is to extrapolate the data with kriging. Kriging provides both an estimated mean and covariance at a vertical location. From these post-kriging stochastic parameters we can generate "pseudo" well logs. We use the Fourier method to generate these simulations. Normally, the post-kriging random field is quite nonstationary but in the special case of extrapolating vertical well logs to other vertical locations, the post-kriging field is approximately stationary. By way of generating these simulations, we employ our Monte Carlo method to search for the maximum likelihood model. When the best model is found it is included as more well data for the purpose of kriging further locations.

Some important differences should be pointed out between the method in Haas and Dubrule (1994) and the one presented here. They limit their parameter space to 30 samples in the vertical direction corresponding to a total of 120 ms. This corresponds to a few wiggles of their seismic trace, thus making it easier to fit the data. In contrast, the models used in our synthetic inversion are sampled in depth at 10 ft intervals for 1000 ft of depth and our seismic data consist of 256 samples at 0.002 ms discretization. Thus,

we have 100 parameters to invert for. Also, their algorithm performs sequential inversion by leap frogging horizontally through field locations. Haas and Dubrule perform the leap frogging because it helps give horizontal continuity to the inverted field. But, by doing the kriging at locations far away from the well data, the kriging variance is much higher. Thus, since there are an infinite number of models that fit the seismic data, there is a greater chance of converging to a suboptimal one. We, in contrast, perform the kriging at locations adjacent to well data. This minimizes the kriging variance and helps better constrain the inversion algorithm.

The 2-D field is inverted twice to show the variability of the inversion results. These two inversions are shown in Figures 4 and 5. Note how similar they are As a matter of fact, they look more like each other than the true field. It is unclear why this is so, however, they all reproduce the main features of the true field, and can be compared with the kriging results shown in Figure 6. We can see that performing kriging without conditioning it to the seismic data fails to reproduce the interwell variability.

We now turn to 3-D inversion. Again, a Gaussian synthetic velocity field is created (Figure 7). It has identical stochastic parameters to the 2-D field with the inclusion of another horizontal correlation length $a_y = 200$ ft. A 3-D seismic data cube is also produced from this model (not shown). We again use an *a priori* seismic wavelet in the inversion. Five well logs are extracted from this field and are shown in Figure 8. Using these wells we first perform simple kriging in Figure 9. Some of the main features of the field are reproduced close to the wells but not beyond about a correlation length. We next perform sequential inversion on the same wells and seismic data. The results are shown in Figure 10, where we see great improvement over the kriging results.

There are 1836 seismic traces in the synthetic 3-D model and we perform 1000 Monte Carlo simulations for each trace. One entire 3-D inversion takes approximately four hours to complete. There are only 51 traces in the 2-D inversion, and again 1000 simulations are performed per trace. This takes only approximately seven minutes to run. All runs were performed on a 400 MHz Pentium 2 PC.

We should point out how the sequential algorithm moves from one vertical location to another. For 3-D inversion, the marching order is shown in Figure 11. This figure is a map view of the first 65 locations visited. Thus the inversion algorithm "propagates" out from each true well location. The same method is used for the 2-D model. Note also which data are included each time a location is kriged. It would seem logical to include all true well logs *and* previously inverted "pseudo" logs every time kriging is performed. This, however, gave poor results and increased computational time by an order of magnitude. Much better results were obtained when only the closest pseudo logs and a few far pseudo logs were included in the kriging. The farther ones are included just to help ensure continuity of the inverted field. At present, we cannot explain why this method is superior.

APPLICATION TO TEXACO DATA SET

The Texaco data set consists of five velocity well logs in a 3-D field (Figure 12). These logs were sampled at 0.5 ft intervals in depth. The seismic data consist of 1836 traces each sampled at 0.002 ms for a total of 121 samples. Thus, 242 ms of data are recorded at each trace. The well logs are truncated so that only the portion corresponding to 242 ms of time is included in the inversion. The average velocity in the well logs is 15,075 ft/s. This leads to approximately 3000 parameters to krig and invert at each pseudo log location. This is far too computationally expensive for kriging and inversion, and may not be warranted if the well logs are significantly smooth at small scales. Unfortunately, the Texaco logs, like most logs, exhibit great variability even at the smallest scales. Despite this fact, we chose to smooth the logs in order to reduce the number of parameters. The velocity logs are smoothed enough to allow three decimations, that is, the sampling interval becomes 4 ft instead of 0.5 ft; we are left with ≈ 400 parameters to deal with in each log.

For the synthetic model used in the first inversion experiment, we knew *a priori* the covariance function describing the random field. This, however, is not the case for the Texaco data set. We attempt to estimate a variogram function for the velocity field in the vertical direction from available well logs. The variogram fit is shown in Figure 13. It is fit with a Gaussian model with a vertical correlation length $a_z = 12$ ft.

There is no satisfactory way to estimate the horizontal correlation lengths. Haas and Dubrule (1994) estimate them from the seismic data. From our experience, we find that the horizontal correlation lengths do not affect the inversion as long as they are chosen large enough to include immediately neighboring locations. Therefore, seeing that the seismic data are significantly correlated horizontally, we simply choose the horizontal correlation lengths to be 20 times the vertical length: $a_x = 240ft$, $a_y = 240$ ft. The variance of the velocity field is also obtained from the variogram and is 1,750,000 ft^2/s^2. We use all these parameters in the kriging and inversion of the Texaco data.

The main difficulty in working with real seismic data is that we need to extract a seismic wavelet. This is needed for the H_S operator in the Monte Carlo search method. To extract a wavelet we need a seismogram and reflectivity series at the same place. We perform deconvolution on these two series to obtain the seismic wavelet. This was done at each of the well log locations of the Texaco data with unsatisfactory results. Five different wavelets were extracted that looked significantly different from each other (see Figure 14). They also were not temporally limited the way one would expect a seismic wave to be. This problem illustrates the limitation of the convolutional model and questions its applicability. To perform sequential inversion, we can have only one wavelet, thus for lack of a better method, the wavelets were averaged and the resulting wavelet was used in the inversion.

We follow the same extrapolation pattern illustrated in Figure 11 with the Texaco data. Also, we allow 1000 Monte Carlo iterations at each pseudo log location that is inverted, just as was done for synthetic inversion. This number of iterations is again

adequate to fit the seismic data even though the number of parameters being inverted for has increased four times. The results of the Texaco data set inversion are shown in Figure 16, and can be compared with the kriged results in Figure 15. The inversion brings out somewhat more detail but both have trends that persist horizontally through the field. Note that the figures show only the first 400 ft of the inverted model out of 3000 ft. Matlab could not plot the entire model, hence only a sub volume is shown.

We cannot check the accuracy of this inversion because we do not have the true field for comparison as we did with the synthetic models. However, we can perform a *cross-validation* by performing the entire inversion again without one of the true well logs and comparing the results of the inversion at the location of the true well log. The results of the cross-validation at the true log location are shown in Figures 17 and 18 (again, only the first 400 ft). We show the reflectivity because this is what the inversion is sensitive to. Remember the nonuniqueness—an infinite number of velocity models all have the same reflectivity series. While the velocities fit each other reasonably well, the reflectivity fits even better.

We can explain our good results by the fact that the seismograms only change slightly with horizontal distance. This implies that the underlying field is quite laterally continuous. Therefore, it seems easier for the inversion method to estimate the Texaco model than the synthetic model.

Note, finally, that the entire 3-D inversion of the Texaco data required six hours of computing time.

CONCLUSIONS

We presented a method for inverting post-stack seismic data given limited well log data. Numerous methods exist to estimate petrophysical properties away from wells. Many methods also exist to invert seismic data. However, few researchers have attempted to combine these methodologies in a coherent way. The results obtained show that sequentially extrapolating wells and inverting seismograms result in an accurate inversion of an unknown velocity field. The sequential extrapolation constrains the inversion with existing well data and the Monte Carlo algorithm handles the nonlinearity of the inversion.

Two main questions still need to be answered:

1. We do not yet know the theoretically optimal way to incorporate inverted psuedo logs into the kriging process when sequentially extrapolating. The current method is somewhat ad-hoc.

2. The greatest problem with the method is extracting a seismic wavelet from real data. As was shown, the obtained wavelets do not match each other and are not temporally limited. This illustrates the limitations of the convolutional model and motivates us to consider applying the method to prestack instead of post-stack seismic data.

Despite the second problem, the method still produced satisfactory results. This may mean that the exact shape of the wavelet is not important, and all that might be required is something that sufficiently smooths the reflectivity series.

ACKNOWLEDGMENTS

We thank Christie Callender of Texaco and Texaco for financial support as well as the use of their 3-D post-stack data set. This work was supported by the Borehole Acoustics and Logging/Reservoir Delineation Consortia at the Massachusetts Institute of Technology.

REFERENCES

Gardner, G.H.F. (editor), 1985, Migration of seismic data, *Geophysics Reprint Series, No. 4,* Soc. Expl. Geophys.

Haas, A. and Dubrule, O., 1994, Geostatistical inversion—a sequential method of stochastic reservoir modeling constrained by seismic data, *First Break, 12.*

Rothman, D.H., 1985, Nonlinear inversion, statistical mechanics, and residual statics estimation, *Geophysics, 50,* 2784–2796.

Sen, M.K. and Stoffa, P.L., 1991, Nonlinear one-dimensional seismic waveform inversion using simulated annealing, *Geophysics, 56,* 1624–1638.

Sengbush, R.L., Lawrence, P.L., and McDonal, F.J., 1960, Interpretation of synthetic seismograms, *Geophysics, 26,* 138–157.

Stoffa, P.L. and Sen, M.K., 1991, Nonlinear multiparameter optimization using genetic algorithms: Inversion of plane-wave seismograms, *Geophysics, 56,* 1794–1810.

Tarantola, A., 1987, *Inverse Problem Theory—Methods for Data Fitting and Model Parameter Estimation,* Elsevier Pub.

Vestergaard, P.D. and Mosegaard, K., 1991, Inversion of post-stack seismic data using simulated annealing, *Geophys. Prosp., 39,* 613–624.

Yilmaz, O., 1987, Seismic data processing, *Investigations in Geophysics, 2,* Soc. Expl. Geophys.

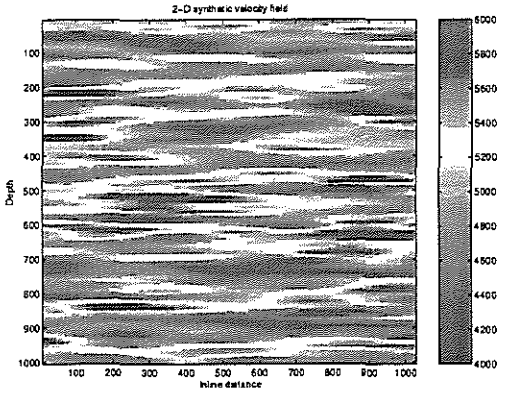

Figure 1: Synthetic 2-D velocity field

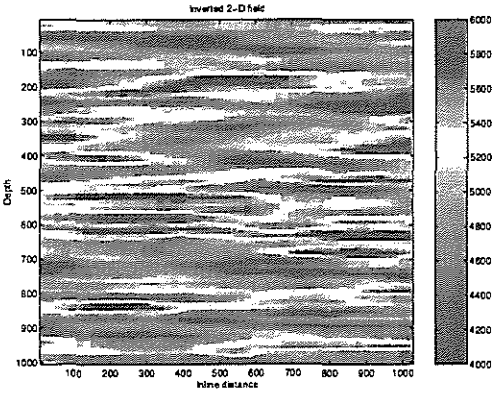

Figure 4: Inverted velocity field #1

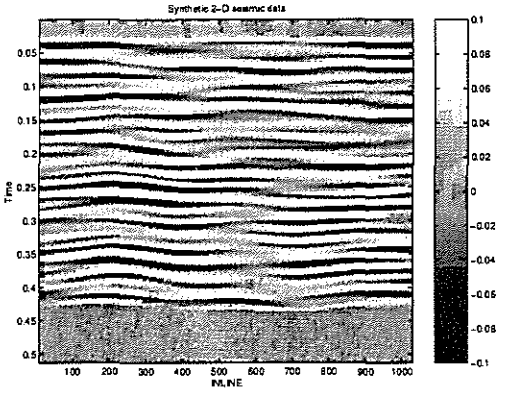

Figure 2: Seismic data

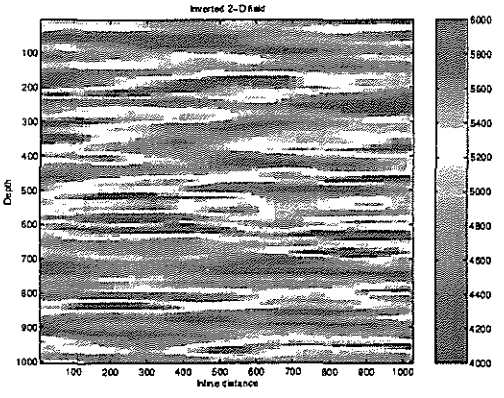

Figure 5: Inverted velocity field #2

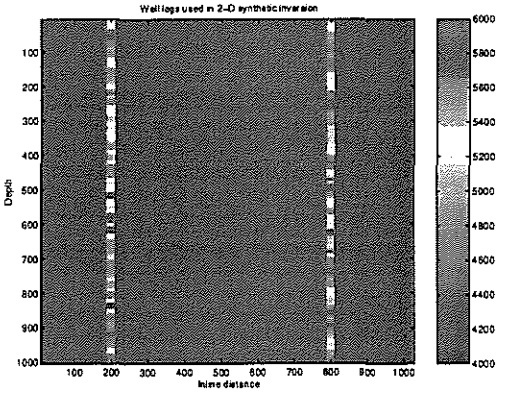

Figure 3: Well data used in inversion

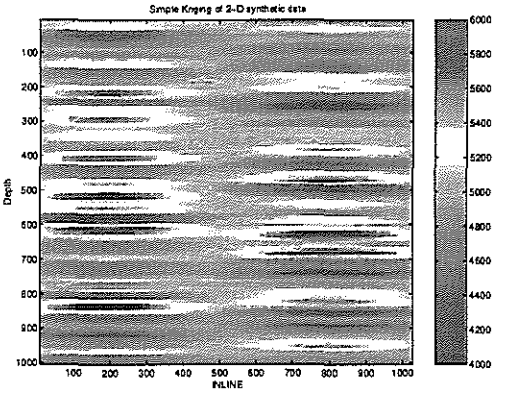

Figure 6: Simple kriging on well data

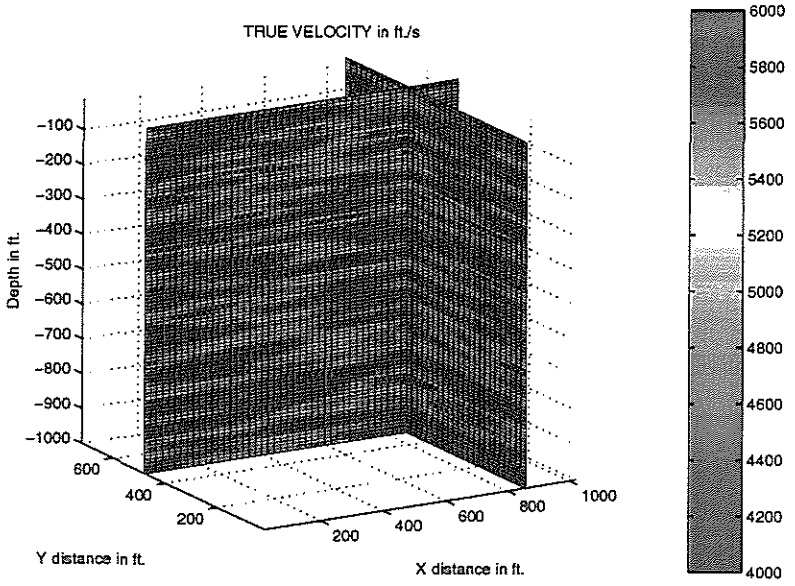

Figure 7: Synthetic 3-D velocity field

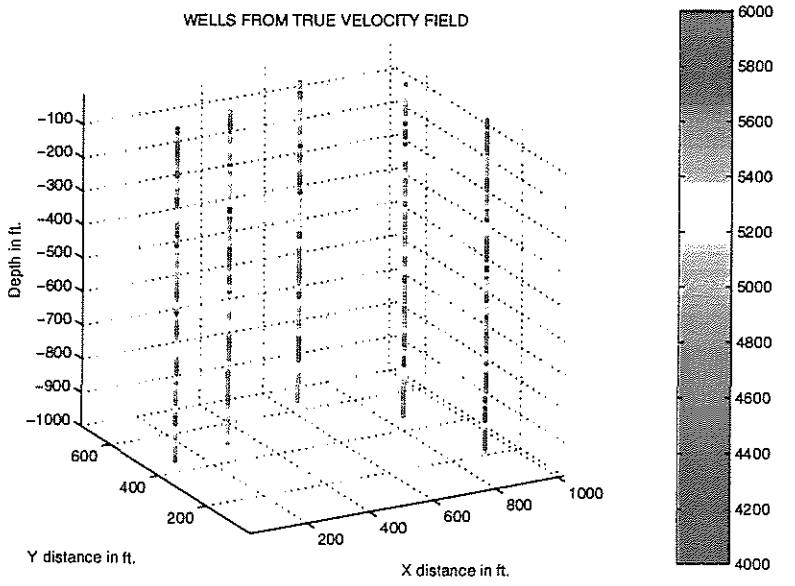

Figure 8: Wells from synthetic model

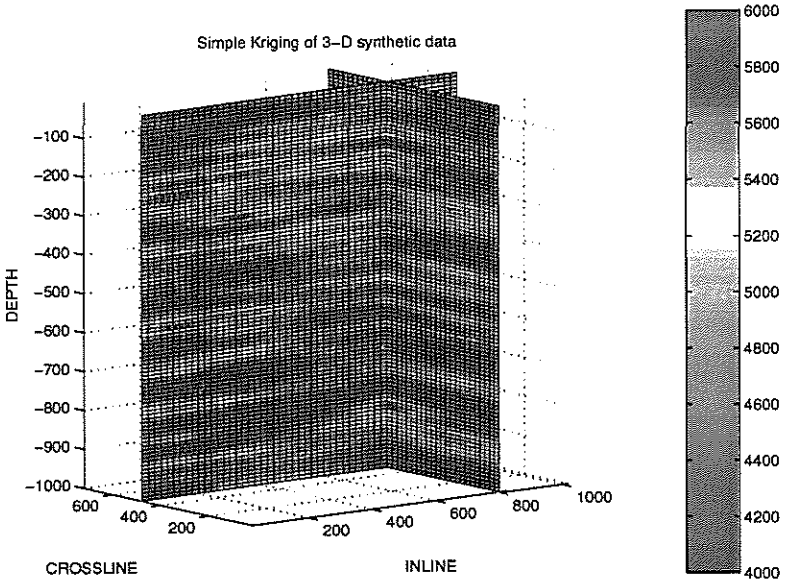

Figure 9: Simple kriging of wells

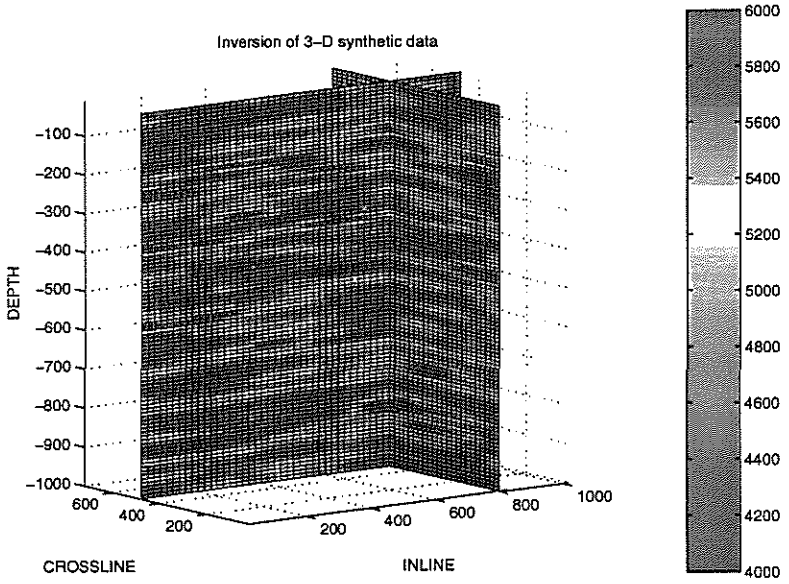

Figure 10: Inverted 3-D velocity field

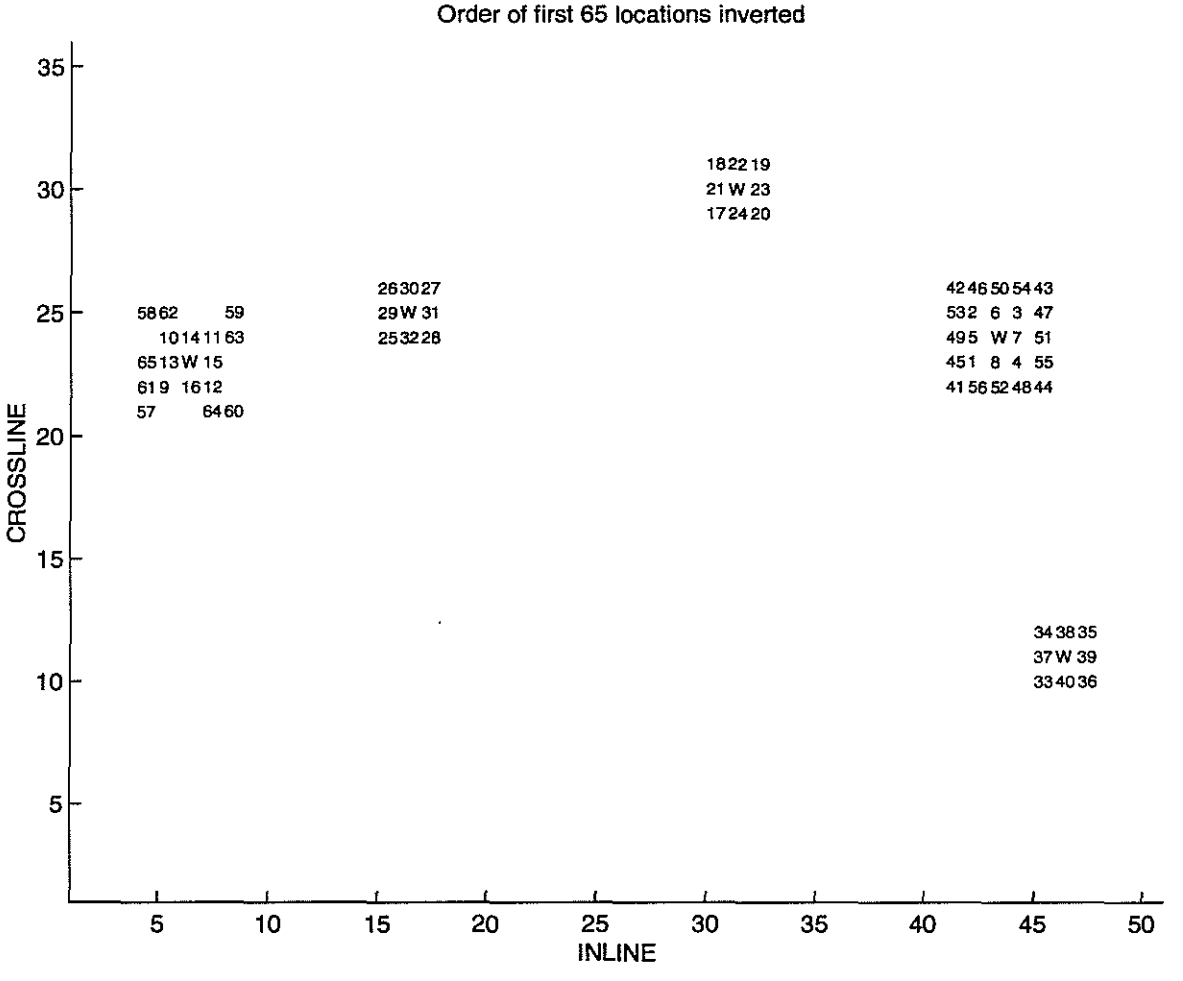

Figure 11: Order of locations inverted in 3-D (map view).

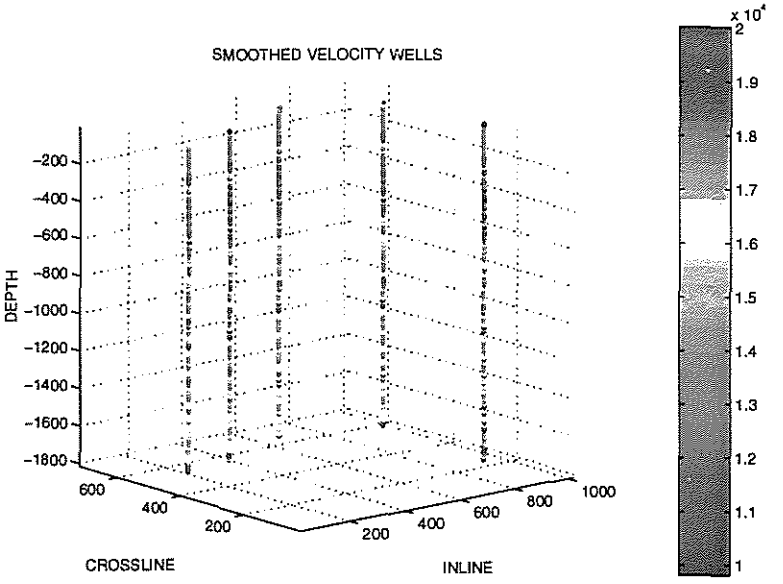

Figure 12: Smoothed Texaco wells

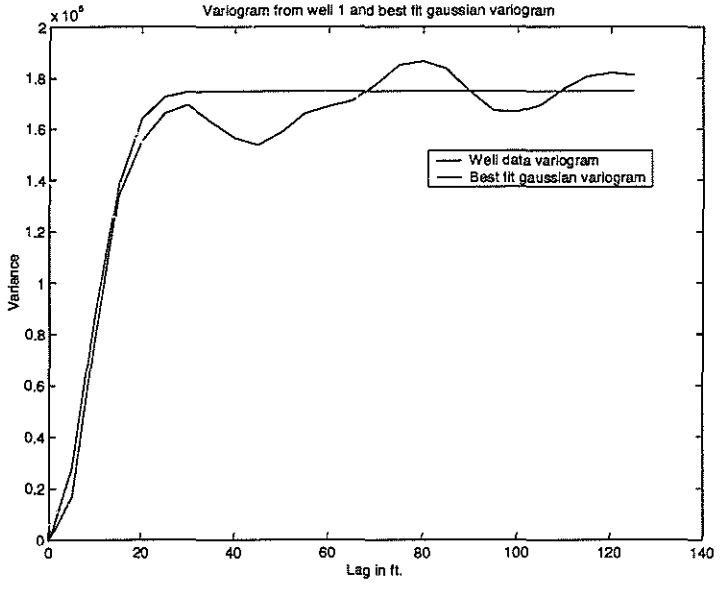

Figure 13: Vertical variogram from Texaco wells

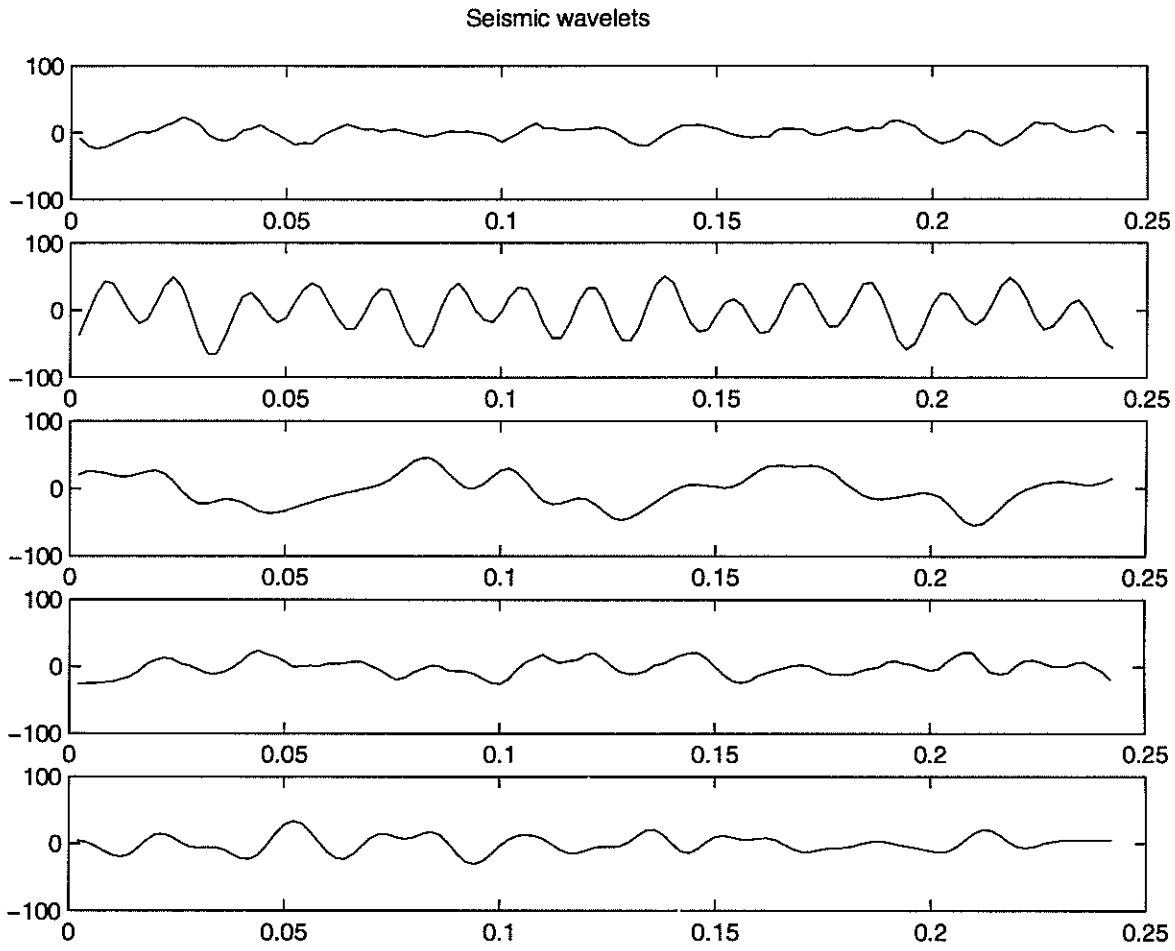

Figure 14: Seismic wavelets extracted from Texaco wells.

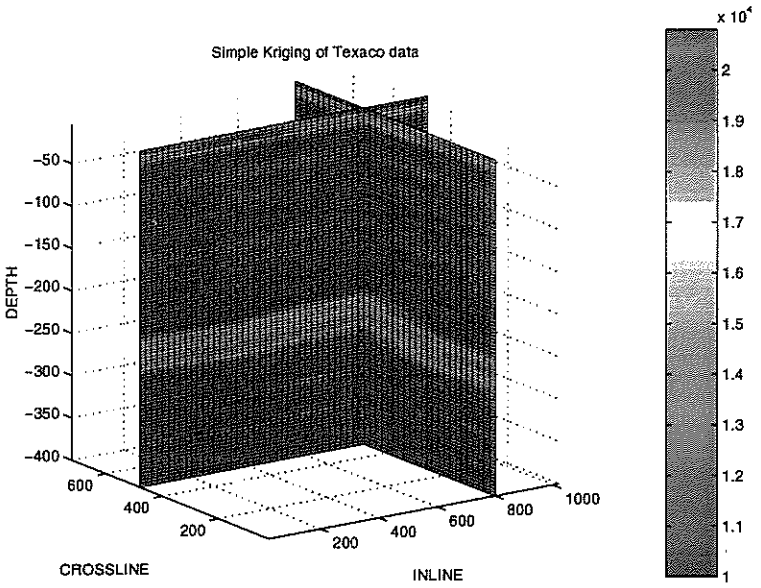

Figure 15: Simple kriging applied to Texaco well data(first 400 ft.)

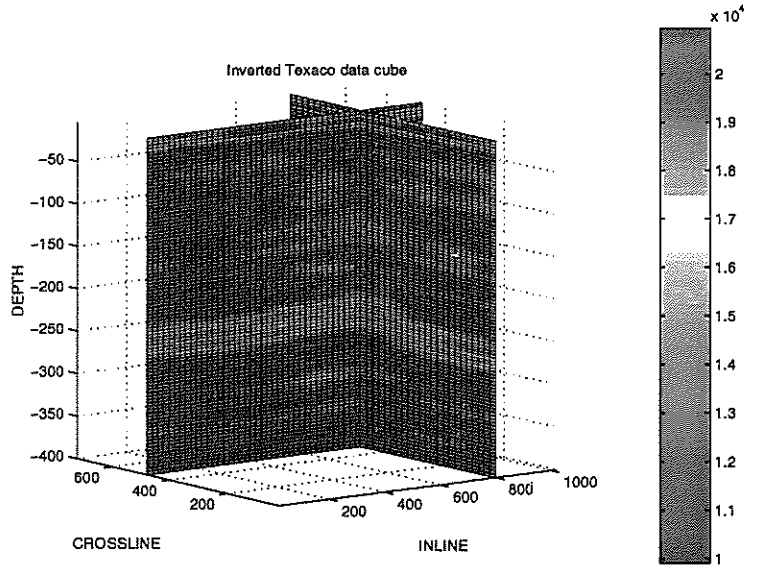

Figure 16: Inverted Texaco velocity field (first 400 ft.)

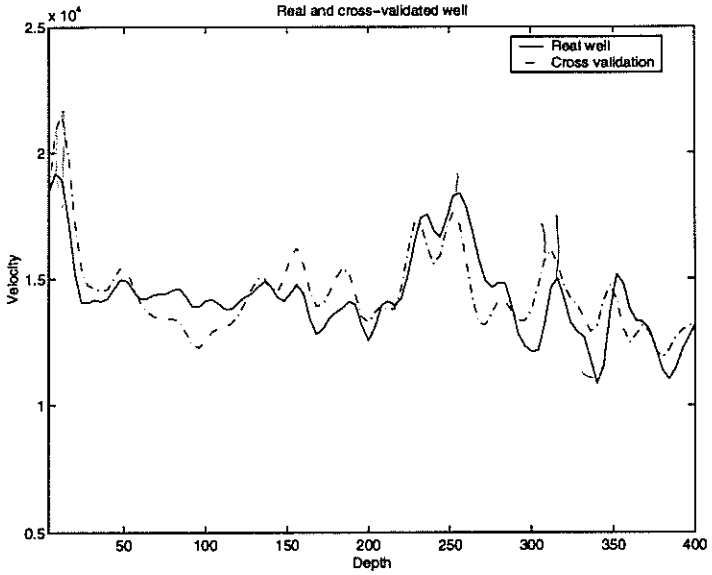

Figure 17: Cross validation of velocity

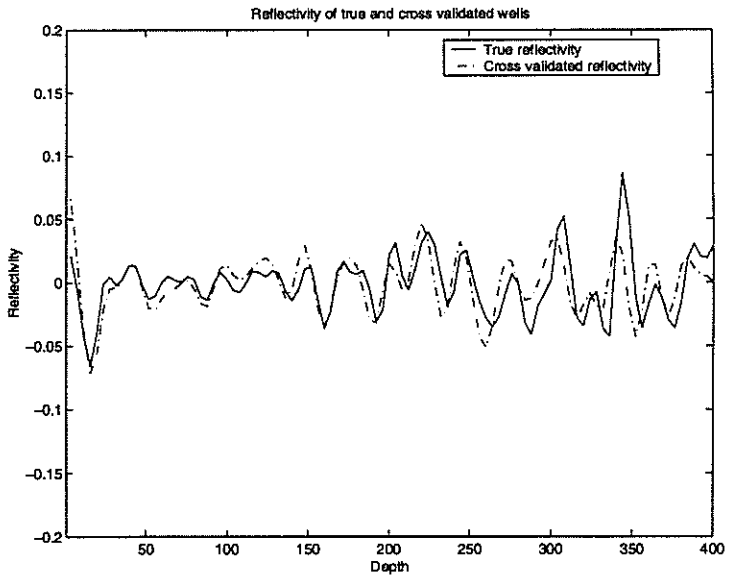

Figure 18: Cross validation of reflectivity