Enforcing smoothness and assessing uncertainty in non-linear one-dimensional prestack seismic inversion

Omar J. Varela,1,3 Carlos Torres-Verdín1* and Mrinal K. Sen2
1Department of Petroleum and Geosystems Engineering, One University Station, Mail Stop C0300, The University of Texas at Austin, Austin TX 78712, USA, 2Institute for Geophysics, Campus Mail Code R2200, The University of Texas at Austin, Austin TX 78712, USA
3Now with: ExxonMobil Upstream Research Company, USA

Received July 2003, revision accepted November 2005

ABSTRACT
Estimation of elastic properties of rock formations from surface seismic amplitude measurements remains a subject of interest for the exploration and development of hydrocarbon reservoirs. This paper develops a global inversion technique to estimate and appraise 1D distributions of compressional-wave velocity, shear-wave velocity and bulk density, from normal-moveout-corrected PP prestack surface seismic amplitude measurements. Specific objectives are: (a) to evaluate the efficiency of the minimization algorithm (b) to appraise the impact of various data misfit functions, and (c) to assess the effect of the degree and type of smoothness criterion enforced by the inversion. Numerical experiments show that very fast simulated annealing is the most efficient minimization technique among alternative approaches considered for global inversion. It is also found that an adequate choice of data misfit function is necessary for a reliable and efficient match of noisy and sparse seismic amplitude measurements. Several procedures are considered to enforce smoothness of the estimated 1D distributions of elastic parameters, including predefined quadratic measures of length, flatness and roughness.

Based on the general analysis of global inversion techniques, we introduce a new stochastic inversion algorithm that initializes the search for the minimum with constrained random distributions of elastic parameters and enforces predefined autocorrelation functions (semivariograms). This strategy readily lends itself to the assessment of model uncertainty. The new global inversion algorithm is successfully tested on noisy synthetic amplitude data. Moreover, we present a feasibility analysis of the resolution and uncertainty of prestack seismic amplitude data to infer 1D distributions of elastic parameters measured with wireline logs in the deepwater Gulf of Mexico. The new global inversion algorithm is computationally more efficient than the alternative global inversion procedures considered here.

INTRODUCTION
Prestack seismic amplitude data are often used to estimate subsurface petrophysical properties. Independent elastic parameters derived from prestack surface seismic amplitude data include compressional-wave velocity $V_P$, shear-wave velocity $V_S$ and bulk...
density $\rho_b$ as functions of depth or seismic traveltime. This paper develops a global inversion technique to estimate and appraise 1D distributions of the three elastic parameters from normal-moveout (NMO)-corrected PP prestack surface seismic amplitude measurements.

The problem of estimating 1D distributions of elastic parameters from prestack seismic amplitude data can be approached using non-linear inversion. Such a procedure is equivalent to the minimization of an objective function written as the metric of the difference between the measured and numerically simulated prestack surface seismic amplitude data. In most applications, inversion requires an efficient and accurate forward operator to simulate the measured seismic amplitude data. Estimation methods based on local minimization often fail to produce a global minimum when the starting solution is far from the optimal point and the objective function exhibits multiple local minima (Tarantola 1987). Applications of local minimization strategies can be found in the open technical literature (e.g. Tarantola 1986; Pan, Phinney and Odom 1988). On the other hand, exhaustive trial-and-error search methods in model space are difficult to implement in an efficient manner because the model space is often extremely large (Sen and Stoffa 1991). The efficiency of global minimization methods remains largely controlled by both the expediency of the search algorithm in model space and the computer power available to simulate the measurements. Examples of global minimization techniques can be found in Kirkpatrick, Elatt and Vecchi (1983), Sen and Stoffa (1991) and Stoffa and Sen (1991).

Various factors control the efficiency and reliability of inversion algorithms, including type of minimization technique, selection of objective function, selection of the initial model, sampling strategy in model space, and spatial smoothness, among others. As a prelude to understanding the new global inversion algorithm developed here, we present a numerical study of the effect of these factors on the efficiency of the inversion algorithm and on the reliability of the estimated 1D distributions of elastic parameters. To evaluate the computer efficiency of the algorithm, we implement various random-walk search methods associated with a global minimization procedure. These search methods include Metropolis (Metropolis et al. 1953), heatbath (Geman and Geman 1984; Rothman 1986) and very fast simulated annealing (VFSA; Ingber 1989). In addition, a detailed study is performed of the sensitivity of the inverted 1D distributions of elastic parameters to the type of data misfit function implemented by the inversion (e.g. $L_1$, $L_2$-norm metrics) and to the degree and type of smoothness criterion (e.g. flatness, roughness) enforced in model space.

Guided by the above analysis of global inversion, we introduce a new efficient stochastic algorithm to invert 1D prestack seismic amplitude data using initialization, sampling and smoothing strategies borrowed from the field of geostatistical estimation. We describe examples of the performance of this algorithm when applied to the inversion of noisy NMO-corrected synthetic amplitude data. An additional example is described to assess the resolution and uncertainty of prestack seismic amplitude data for vertical variations of elastic parameters measured with wireline logs in the deepwater Gulf of Mexico. The central technical contribution of this study is the quantitative evaluation of the uncertainty of 1D prestack seismic amplitude data via inversion.

**PRESTACK STOCHASTIC INVERSION**

**Background**

The physical process of reflection, transmission and mode conversion of plane waves at a horizontal boundary as a function of angle was described by Aki and Richards (2002), among others. Inversion of prestack seismic amplitude data yields a 1D distribution of elastic parameters (i.e. $V_p$, $V_s$ and $\rho_b$) from the information content available in both time and source–receiver space. Approaches to this problem include the use of different approximations of either Zoeppritz equations (Wang 1999) or angle stacks (Simmons and Backus 1996). The inversion algorithms considered here make use of the reflectivity method (Fuchs and Muller 1971; Kennett 1983) to compute the full-wave response of a stack of horizontal layers. More specifically, we use the reflectivity method for efficient computation of synthetic seismograms, devoid of transmission losses for P-wave propagation (primaries only) in the offset–time domain.

**Formulation**

Inversion of prestack seismic amplitude data into a 1D distribution of elastic parameters remains a highly non-linear, non-unique process that requires significant forward modelling capabilities. Prestack seismic waveforms are modelled assuming a locally 1D
distribution of elastic parameters. The problem of estimating the corresponding 1D distribution of elastic parameters is cast as the global minimization of an appropriately constructed objective function. The global minimization methods described here implement a Monte-Carlo guided search method and a search schedule based on simulated annealing (SA). These methods are designed to consider the presence of multiple minima in the objective function and to pursue the global minimum, regardless of the starting point in parameter space. The estimation of model parameters is also performed while enforcing physical constraints (e.g. trends) that eliminate the need to search for inconsistent models that may also honour the measurements.

Global minimization technique

Simulated annealing is a global minimization method that replicates the thermodynamic cooling of a multiparticle physical system. The range of possible global energy values adopted by such a system corresponds to the range of values considered by the objective function. The Metropolis procedure simulates the natural process whereby crystal lattices of glass or metal relax to a lower energy state of thermal equilibrium. This process is usually referred to as annealing (Metropolis et al. 1953). The concept behind SA is that the objective function relates global system energy to a given state of the system and various ways to bring the multiparticle system to a lower energy state of thermodynamic equilibrium.

The minimization algorithm based on SA starts with an initial model $m_0$, which has an associated energy $E(m_0)$. In this context, $m$ refers to a vector of size $N$ that contains the unknown model parameters indexed in an appropriate manner and $E$ designates the objective function. A random walk process (i.e. a random process consisting of a sequence of discrete steps of fixed length) is then used to select a new location $m_i$ in model space. The global energy $E(m_i)$, associated with the new location in model space, is subsequently calculated and compared with the prescribed acceptance test. If the objective function decreases, then the new point $m_i$ in model space is accepted unconditionally ($m_0 = m_i$); otherwise, the change is accepted but only with probability equal to $\exp[-(E(m_i) - E(m_0))/T]$, where $T$ is a control parameter called ‘temperature’. Iteratively, many locations $m_i$ in model space are explored in sequence. Each new location may be either accepted or rejected in the search of the global minimum according to this criterion. Such a procedure eventually reaches the global minimum. It can be shown that if the rate of temperature decrease is sufficiently low, the global minimum can be statistically reached for the global energy function (Ingber 1989, 1993). Various SA algorithms are considered here, namely, Metropolis, heatbath and VFSA. These algorithms make use of the same acceptance/rejection criterion introduced by Metropolis et al. (1953). The difference between them is that Metropolis SA formulates a new sample in model space from a given probability density function (PDF), while heatbath SA produces weighted samples in model space that are always accepted, and VFSA formulates a new sample in model space from a 1D Cauchy PDF that is a function of temperature. Metropolis SA and VFSA algorithms are two-step processes whereas heatbath SA is a one-step process.

The search algorithm allows a cooling schedule in which temperature decreases exponentially with iteration number or annealing-time $k$, and is given by

$$T = T_0 \exp(-ck^{1/N}),$$

where $T_0$ is a specified initial temperature, $c$ is a specified temperature decay rate and $N$ is the dimensionality of the model parameter space (Ingber 1989). These specified parameters are adjusted in a heuristic manner and are often refined to improve the efficiency of the algorithm. The reader is referred to Ingber (1993) for a general summary of practical applications of SA.

Sampling techniques

The approach to the formulation of an initial model needs to take into account both the inclusion of a priori model information and the optimization of the computational efficiency of the inversion algorithm. Many sampling strategies can be implemented, depending upon the available a priori information or knowledge about the model. One of the simplest assumptions could be the lack of specific knowledge about the model parameters. In that case, only the minimum and maximum values of the model parameters would be known a priori. The initial model would then be formulated from uniform PDFs spanning the corresponding range of variation. With more specific a priori information about the model, global PDFs such as Gaussian distributions, for instance, can be used to constrain the search for model parameters, including the selection of an initial model. If available, the
spatial distribution (i.e. vertical trends) can be used in conjunction with the PDFs of model parameters to formulate the initial model and to constrain the subsequent selection of model parameters. In the latter case, each elastic property in the 1D subsurface model will exhibit a vertical property trend and each layer will be associated with a specific PDF. If more specific information were available about the 1D distribution of elastic parameters, then this a priori knowledge could be used to further constrain the inversion and hence reduce non-uniqueness.

Data misfit function

In order to assess the similarities or differences between synthetic and measured prestack seismic amplitude data, several types of misfit function were considered in the inversion algorithm. By definition, a solution to the inverse problem, m, entails the smallest misfit or prediction error above the threshold assumed for measurement noise. There are different metrics, or norms, available to quantify the length or size of the misfit vector. The first type of data misfit function was constructed in the time domain. Equations 2 and 3 below define the L_1- and L_2-norms of the data misfit vector, computed for all the discrete measurement times and source–receiver offsets, namely,

\[ \| e_i \|_1 = \sum_{i=1}^{N_{off}} \sum_{j=1}^{N_f} |e_{ij}| = \sum_{i=1}^{N_{off}} \sum_{j=1}^{N_f} \left| S(x_i, t_j)^{obs} - S(x_i, t_j)^{est} \right| \] (2)

and

\[ \| e_i \|_2 = \left( \sum_{i=1}^{N_{off}} \sum_{j=1}^{N_f} |e_{ij}|^2 \right)^{1/2} = \left( \sum_{i=1}^{N_{off}} \sum_{j=1}^{N_f} \left| S(x_i, t_j)^{obs} - S(x_i, t_j)^{est} \right|^2 \right)^{1/2} \] (3)

where \( S^{est} \) and \( S^{obs} \) are the synthetic and measured prestack seismic amplitude data, respectively, given as a function of source–receiver distance \( x_i \) and seismic traveltime \( t_j \). In the above equations, \( N_t \) is the number of discrete time samples per trace for a given source–receiver offset and \( N_{off} \) is the number of source–receiver offsets included in the prestack gather.

The second type of data misfit function was constructed in the frequency domain using the real parts of the geometric (L_g) and harmonic (L_h) metrics of the correlation between synthetic and measured prestack seismic amplitude data introduced by Sen and Stoffa (1991), i.e.

\[ \| e_i \|_g = \sum_{i=1}^{N_{off}} \left[ \frac{\sum_{j=1}^{N_f} S(x_i, f_j)^{obs} S^*(x_i, f_j)^{est}}{\left( \sum_{j=1}^{N_f} |S(x_i, f_j)^{obs} S^*(x_i, f_j)^{obs}| \right)^{1/2}} \right]^{1/2} \] (4)

and

\[ \| e_i \|_h = \sum_{i=1}^{N_{off}} \left[ \frac{\sum_{j=1}^{N_f} S(x_i, f_j)^{obs} S^*(x_i, f_j)^{est}}{\left( \sum_{j=1}^{N_f} |S(x_i, f_j)^{obs} S^*(x_i, f_j)^{obs}| \right)^{1/2}} + \left( \sum_{j=1}^{N_f} |S(x_i, f_j)^{est} S^*(x_i, f_j)^{est}| \right)^{1/2} \right] \] (5)

respectively.

In the above equations, \( f_j \) is frequency, \( N_f \) is the number of frequencies, and the superscript * indicates the complex conjugate operator. Weighted norms may also be considered in which time–frequency data for a given source–receiver offset may be assigned a specific weight. These weights could be adjusted to put all of the source–receiver measurements on equal footing and hence to increase the sensitivity of all the source–receiver traces to a perturbation of the model parameters. Exploring the use and optimal selection of source–receiver weights, however, goes beyond the scope of this paper.
Measures of model smoothness

Additional terms in the objective function are used to bias the solution constructively toward a preconceived notion of model properties. A general objective function that includes different metrics, constraints and a priori information is given by

$$E(m) = E_{\text{data}} (S^{\text{src}}, S^{\text{obs}}) + \alpha E_{\text{model}}(m),$$

where $\alpha$ is a user-defined parameter that controls the relative importance of the two additive components, $E_{\text{data}}$ represents the data misfit function (in the present study this term could be any of the misfit functions defined in (2)–(5)), and $E_{\text{model}}$ represents any measure of a priori information and constraints enforced on the model parameters. Normally, the relative weight applied to the data misfit and smoothness terms of the objective function is determined by trial and error as there is no universal method to calculate it in a deterministic fashion. The continuous version of the term $E_{\text{model}}$ in (6) is defined as operating on a seismic traveltime-domain function $m = m(t)$ that contains the continuous model parameters, i.e. $V_p(t)$, $V_S(t)$ and $\rho_0(t)$, and is given by

$$E_{\text{model}}(m) = \left[ \omega \int_{t_0}^{t_1} |m|^2 \, dt + \beta \int_{t_0}^{t_1} \frac{dm}{dt} \, dt + \gamma \int_{t_0}^{t_1} \left| \frac{d^2m}{dt^2} \right|^2 \, dt \right] + \left[ \chi \int_{t_0}^{t_1} \left| \frac{dm}{dt} \right| \, dt \right].$$

where $\omega, \beta, \gamma$ and $\chi$ are user-defined positive numbers that control the smoothness of the unknown model $m(t)$, and $t_0$ and $t_1$ are lower and upper time limits of the estimation domain, respectively. The first three additive terms of (7) are measures of the length of the solution, based on an $L_2$-norm metric. These terms are usually called size, flatness ($F$) and roughness ($R$), respectively. Menke (1989) emphasized that such a measure of length is the simplest type of a priori assumption about a particular model. The last term of (7) is a measure of flatness ($F_1$) based on an $L_1$-norm metric. We emphasize that the use of a mixed $L_2$- and $L_1$-norm in model space is desirable in some cases, in order to reduce Gibb’s phenomena in the estimation of the continuous function $m(t)$ (Oldenburg, Scheuer and Levy 1983). Often, the size and smoothness properties of the model function $m(t)$ are measured with respect to an a priori reference model (i.e. $m \sim m_{\text{ref}}$). Such a strategy was not used as an explicit measure of simplicity in this paper. By excluding the first additive term of the general objective function, we arrive at the discrete version of (7), namely,

$$E_{\text{model}}(m) = \left[ m^T (\beta W_F^T W_F + \gamma W_R^T W_R) m \right] + \left[ \chi \left\| W_F m \right\|_1 \right],$$

where the superscript $T$ indicates the transpose operator, the vector $m$ describes the model parameters, i.e. $m = (\rho_{b1}, \ldots, \rho_{bN}, V_p1, \ldots, V_pN, V_S1, \ldots, V_SN)$, and $W$ is a weighting block matrix. For the cases considered here, model parameters change with vertical location (i.e. seismic traveltime), whereupon model flatness can be calculated using the discrete version of the model parameters, and is given by

$$F(m) = m^T [W_F^T W_F] m,$$

where the block diagonal matrix $W_F$ can be written as

$$W_F = \begin{bmatrix} W_{\rho b} & 0 \\ 0 & W_{V_p} \\ 0 & W_{V_S} \end{bmatrix}. \quad (10)$$

Each of the block matrices included in (10) is determined by the type of approximation used to calculate the derivative (e.g. backward, forward, three-point or five-point central difference approximation). For the case of a backward first-difference approximation used for model parameters that are not regularly spaced, $W_X$ is given by

$$W_X = \begin{bmatrix} -1 & 1 & \frac{1}{\sqrt{\Delta t_1}} \\ \frac{1}{\sqrt{\Delta t_1}} & -1 & \frac{1}{\sqrt{\Delta t_1}} \\ \frac{1}{\sqrt{\Delta t_2}} & \frac{1}{\sqrt{\Delta t_2}} & -1 \\ \frac{1}{\sqrt{\Delta t_{N-1}}} & \frac{1}{\sqrt{\Delta t_{N-1}}} & \frac{1}{\sqrt{\Delta t_{N-1}}} \end{bmatrix}. \quad (11)$$

where \( X \) is a generic designation for any of the three elastic parameters and \( \Delta t \) is the interval between consecutive time samples. The roughness \( R \) of the model parameters can also be calculated in a similar manner to the flatness, by discretizing the second-derivative operator in place of matrix \( W_X \). In the most general case, the inversion algorithms considered here are formulated as the minimization of the objective function

\[
E(m) = E_{data}(S^{st}, S^{obs}) + \beta |F(m)| + \gamma |R(m)| + \chi |F_1(m)|.
\]  

**NUMERICAL EXPERIMENTS AND DISCUSSION**

Numerical experiments of global inversion were performed using different noisy synthetic data sets that included only PP seismic reflection amplitudes. All the examples assume input data in the form of NMO-corrected prestack seismic gathers. These numerical experiments are described in three separate sections. The first section (examples 1–4) considers a sensitivity analysis to evaluate the influence of the most important inversion parameters. Results from this analysis are implemented in subsequent exercises and serve as reference for inversion results obtained with the new global inversion algorithm in the second section (examples 5–7). In the final section (example 8), the new global inversion algorithm is used to assess the sensitivity of prestack seismic data to the estimation of 1D distributions of elastic parameters constructed from well-log data acquired in the deepwater Gulf of Mexico.

**ANALYSIS OF INVERSION PARAMETERS**

**Description of the synthetic data set**

Figure 1 shows a 1D synthetic subsurface model consisting of two sandbodies embedded in a background shale layer. The upper sandbody is water-filled, whereas the lower one is saturated with oil. Panels (a)–(d) in this figure show the corresponding well logs of bulk density, compressional-wave velocity, shear-wave velocity and lithology. These well logs were calculated using a rock-physics model that includes the effect of compaction. Panel (e) shows the input NMO-corrected prestack seismic amplitude data simulated for a single source–receiver gather. The simulated prestack seismic amplitude data were also contaminated with 5%, zero-mean additive Gaussian random noise. A zero-phase Ricker wavelet centred at 35 Hz was assumed in the simulation of the prestack seismic amplitude data. Simulation of seismic amplitude data was performed assuming 10 source–receiver offsets with a uniform receiver spacing of 300 m and a constant time sampling rate of 2 ms in the interval from 0 to 1.4 s. The number and spacing of receivers were selected to ensure sufficient variability in the seismic amplitudes across the prestack seismic gather. In this particular case, the maximum offset-to-depth ratio is approximately equal to 2 and the maximum angle of incidence allowed by the simulations is 60°. An appropriate maximum offset-to-depth ratio is critical for ensuring measurable sensitivity of the seismic amplitude data to the elastic parameters, especially in the estimation of shear-wave velocity and bulk density (Castagna and Backus 1993). The discretization of the elastic parameters was performed directly from the well logs shown in Fig. 1. This discretization consisted of 50 layers of an average traveltime ‘thickness’ of 28 ms for each layer, spanning the same time interval as the prestack seismic amplitude data (i.e. from 0 to 1.4 s).

**Example no. 1: Assessment of the efficiency of the annealing technique**

In this example, inversion was performed using various annealing techniques (Metropolis, heatbath and VFSA) to yield estimates of compressional- and shear-wave velocities and bulk density. The inversions were implemented under the same operating assumptions and the corresponding computer time was clocked to assess relative computer efficiency. We made no attempt to compare the relative efficiency of different seismic forward modelling techniques. Figure 2 shows the CPU time associated with each SA technique in relation to the CPU time entailed using VFSA. The exercises considered here consistently indicated that VFSA remained the most efficient SA algorithm for global inversion. Similar results have been reported in the open technical
Example no. 2: Choice of objective function

A similar inversion exercise was performed to appraise the sensitivity of the estimated elastic parameters to the choice of data misfit function. Selecting an appropriate data misfit function is crucial to the inversion because some of these functions are not sensitive to absolute differences in seismic amplitudes. Figure 3 shows the original prestack seismic data set (panel a) and the seismic residuals derived from the inversion with the use of the $L_1$ (panel b), $L_2$ (panel c), geometric $L_6$ (panel d) and harmonic $L_h$ (panel e) data misfit functions described by (2) (3) (4) and (5), respectively. In this example, all inversions included the same number of iterations in the minimization process. The harmonic misfit function is sensitive to the absolute seismic amplitude differences and hence produces superior results in terms of data residuals, total misfit and computational performance. Such a data misfit function is used as the default option in the remaining examples considered here.

Example no. 3: Sampling techniques in model space

Various approaches were considered here to formulate an initial model and to constrain subsequent model changes in the search for the minimum of the objective function. The first approach assumes that there is no well-log data available and hence general distributions are used as constraints (see Fig. 4). In the second approach, we use both well-log data and trends (see Fig. 5). These
approaches are of significance in enforcing a priori model information, in optimizing computer efficiency, and in assessing the influence of a priori information on the inversion. For simplicity, but without sacrifice of generality, the example described in this section makes use of only the first term of (12).

Figure 4 shows inversion results obtained when assuming various types of a priori information about the 1D distribution of elastic parameters (e.g. uniform, normal), using various types of sampling strategies. All inversion exercises achieved the same similarity in data space (i.e. small prestack seismic data misfit). Panels (a) and (b) in Fig. 4 show the uniform and Gaussian PDFs of compressional-wave velocity, shear-wave velocity, and bulk density used to formulate stochastically the initial model entered in the inversion. This example assumes no specific trend with respect to seismic time (or depth). Therefore, the same PDF is used for all layers. Panels (c) and (d) show the actual (line) and inverted (open circles) elastic parameters estimated using uniform and Gaussian PDFs. Correlation coefficients \( r^2 \) is used as a scale-independent measure of similarity of two discrete functions) between actual and inverted parameters are also annotated in these panels. For the case of scant prior information about the unknown model, no reasonable assumption can be made about the 1D variations of the elastic properties and hence a uniform PDF remains the most appropriate choice (panel a) to sample model parameters. In this case, the two sets of actual and inverted elastic parameters do not correlate very well as indicated by the relatively low correlation coefficients shown in panel (c) \( r^2_{V_p} = -0.54, \ r^2_{V_s} = 0.42, \ r^2_{\rho_b} = 0.20 \) and the 1D distribution of elastic parameters shown in the same panel. Use of a priori model information in the form of a Gaussian PDF (panel b) yields 1D distributions of elastic parameters closer to the actual ones, as indicated by the corresponding correlation coefficients shown in panel (d) \( r^2_{V_p} = -0.11, \ r^2_{V_s} = 0.51, \ r^2_{\rho_b} = 0.67 \) and the comparison between actual and inverted 1D distributions of elastic parameters shown in the same panel. In this particular case, the enforced Gaussian PDF was constructed from histograms sampled from the well logs shown in Fig. 1.

Figure 5 shows inversion results obtained when a priori model information includes (1) knowledge that the model exhibits a specific vertical trend (2) knowledge of lithology (i.e. sand or shale) and (3) specific ranges of variation about the vertical trend for each of the three elastic parameters. In all cases, the inversions entailed the same similarity in data space enforced in previous examples. Panel (a) in Fig. 5 shows the vertical trends and the lower, mean and upper bounds for the 1D distribution of unknown elastic parameters. Upper and lower bounds for the three elastic parameters were defined as \( \pm 10\% \) of their corresponding means, respectively. Gaussian PDFs may also be enforced for each layer to constrain the variability of the unknown elastic parameters. Panel (b) shows the corresponding 1D distribution of actual and inverted elastic parameters. Open circles indicate the inverted elastic parameters in this panel. Panel (c) shows the evolution of the data misfit function as a function of iteration number; panel (b) shows the correlation coefficients between the actual and inverted elastic parameters. These results exhibit better correlation.
Enforcing smoothness and assessing in 1D prestack seismic inversion

Figure 3 Comparison in data space of the performance of the inversion algorithm for different types of data misfit function after the same number of iterations of the VFSA algorithm. Measured prestack seismic amplitude data (a), and data residuals yielded by the inversion using several data misfit functions, namely: (b) $L_1$-norm (c) $L_2$-norm (d) geometric norm $L_g$, and (e) harmonic norm $L_h$. The various misfit functions are described by equations 2, 3, 4, 5, respectively.

coefficients ($r_{V_p}^2 = 0.35$, $r_{V_S}^2 = 0.85$, $r_{p_b}^2 = 0.72$) than those obtained for the examples described in Fig. 4. Subsequent inversion exercises described here employ the same model sampling strategy assumed to obtain the results shown in Fig. 5.

Example no. 4: Smoothness criterion

The inverted elastic parameters can sometimes exhibit low correlation coefficients with the actual parameters, even though these parameters entail a high degree of similarity in data space. Such a situation can also occur in cases where the unknown models are constrained by a priori information such as vertical trends, lithology and PDFs for each layer (see Fig. 5). This common situation arises because of the non-uniqueness implicit in the non-linear relationship between the data and the model, especially in the presence of noisy and sparse measurements. Common practice shows that a more general objective function, such as that shown in (12), can be used to address non-uniqueness of the inversion by biasing the search for the unknown model solution towards a specific subset in model space.

© 2006 European Association of Geoscientists & Engineers, Geophysical Prospecting, 54, 239–259
Figure 4 Effect of the model sampling technique on the inverted 1D distribution of elastic parameters for the case of no vertical trend imposed a priori on the elastic parameters. Panels (a) and (b) show the uniform and Gaussian probability density functions enforced by the inversion to estimate the elastic parameters shown in panels (c) and (d), respectively. In panels (c) and (d), the inverted distributions of elastic parameters are indicated by open circles and $r^2$ is the correlation coefficient between the actual and inverted elastic parameters.

Figure 6 shows results obtained for different types and degrees of smoothness enforced by the inversion algorithm. All the inversion results shown in this figure entailed the same similarity in data space. The second (model flatness) and third (model roughness) additive terms of (12) were the two types of smoothness criterion used to obtain the results shown in Fig. 6. The degree of smoothness was controlled by different values of the user-defined parameters $\beta$ and $\gamma$ in (12). Panels (a)–(d) in Fig. 6 show cross-plots of actual and inverted elastic properties when 1%, 5%, 10% and 20% of model flatness was enforced by the inversion. Correlation coefficients between actual and inverted elastic parameters in general increase as the enforced degree of model flatness increases. For instance, for a model flatness of 1%, the correlation coefficients are $r^2_{Vp} = 0.39$, $r^2_{Vs} = 0.85$ and $r^2_{\rho_b} = 0.76$, and for a model flatness of 10%, the correlation coefficients are $r^2_{Vp} = 0.51$, $r^2_{Vs} = 0.90$ and $r^2_{\rho_b} = 0.76$. Panels (e) and (f) in Fig. 6 also show cross-plots of actual and inverted elastic parameters when measures of both model flatness and model roughness are included in the inversion. To obtain the results shown in Fig. 6, the model flatness was kept constant at 20% and the model roughness was assumed equal to 10% and 20% in the inversion algorithm. The correlation coefficients between actual and inverted elastic parameters increase as the enforced degree of model flatness increases for a given value of model flatness. For the case of 20% model flatness and 10% model roughness, the correlation coefficients are $r^2_{Vp} = 0.69$, $r^2_{Vs} = 0.91$ and $r^2_{\rho_b} = 0.84$, whereas for the case of 20% model flatness and 20% model roughness, the correlation coefficients are $r^2_{Vp} = 0.76$, $r^2_{Vs} = 0.96$.
Enforcing smoothness and assessing in 1D prestack seismic inversion

Figure 5 Effect of the model sampling technique on the inverted 1D distribution of elastic parameters for the case of a vertical trend imposed a priori on the elastic parameters. Panel (a) shows the vertical trend, minimum, mean and maximum a priori values of the elastic parameters and the corresponding Gaussian probability density functions enforced by the inversion to estimate the elastic parameters shown in panel (b). The evolution of the data misfit function as a function of iteration number is shown in panel (c). In panel (b), the inverted distributions of elastic parameters are indicated by open circles and $r^2$ is the correlation coefficient between the actual and inverted elastic parameters.

and $r^2_{\rho_b} = 0.87$. Correlations substantially improve when both model flatness and model roughness are included in the objective function described by (12).

Next, the $L_1$-norm was used to calculate the fourth additive term of the objective function shown in (12). This was the type of smoothness criterion used to obtain the results shown in Fig. 7. The degree of smoothness was controlled by different values of the user-defined parameter $\chi$. Panels (a) and (b) in Fig. 7 show the actual and inverted 1D distributions of elastic parameters when 5% and 10% of $L_1$-norm model smoothness was enforced by the inversion algorithm, respectively. For the case of 5% of model smoothness (panel a), the correlation coefficients are $r^2_{V_p} = 0.44$, $r^2_{V_S} = 0.86$ and $r^2_{\rho_b} = 0.73$, whereas for the case of 10%
Figure 6 Cross-plots of actual and inverted elastic parameters obtained using various types of smoothness criteria and the same measure of similarity in data space. Panels (a)–(d) show results obtained when enforcing a model flatness criterion of 1%, 5%, 10% and 20%, respectively. Panels (e) and (f) show results obtained when enforcing simultaneously model flatness and model roughness criteria of 20% and 10%, and 20% and 20%, respectively. In all the panels, $r^2$ is the correlation coefficient.
Enforcing smoothness and assessing in 1D prestack seismic inversion

Figure 7 One-dimensional distributions and cross-plots of actual and inverted elastic parameters obtained using an $L_1$-norm metric to enforce model flatness, i.e. by making use of the fourth additive term of the objective function (equation 12). Panels (a) and (b) show results obtained when enforcing a model flatness criterion of 5% and 10%, respectively. In panels (a) and (b), the inverted distributions of elastic parameters are indicated by open circles and $r^2$ is the correlation coefficient between the actual and inverted elastic parameters.

of model smoothness (panel b), the correlation coefficients are $r_{Vp}^2 = 0.42$, $r_{Vs}^2 = 0.89$ and $r_{\rho_b}^2 = 0.82$. Compared to the results shown in Fig. 5, results improve when the $L_1$-norm smoothness criterion is enforced. However, as shown in Fig. 6, more accurate results were obtained when using an $L_2$-norm measure of smoothness.

The above examples of the enforcement of smoothness in the inversion are presented here for reference purposes, given that such measures of smoothness are commonly used in seismic inversion. As shown below, the new global inversion algorithm described here provides an altogether different way of enforcing smoothness of the unknown 1D distributions of elastic parameters.

DESCRIPTION AND EVALUATION OF THE NEW GLOBAL INVERSION ALGORITHM

New global inversion algorithm

Based on the above numerical experiments, the proposed global inversion algorithm makes use of the most efficient annealing technique (i.e. VFSA), the most efficient data misfit function (i.e. harmonic), and it introduces a sampling strategy based on geostatistical concepts. This inversion algorithm makes explicit use of only the data misfit term of the general objective function described in (12). Figure 8 shows a generalized flow diagram of the proposed global inversion algorithm. The main steps are described below:

(1) At the outset, an initial model is randomly drawn from PDFs of model parameters (i.e. $V_p$, $V_s$ and $\rho_b$). The type of PDF and the associated value-range constraints depend on the a priori information available about the model.

(2) Next, a specified number of hard points in model space (i.e. points that are initially fixed) are randomly chosen.

(3) A random walk designates the next time sample to be considered for analysis. Estimation of the corresponding elastic parameters (i.e. $V_p$, $V_s$ and $\rho_b$) is performed via kriging (i.e. a geostatistical modelling technique that is widely used for data interpolation and extrapolation) provided that the time sample remains within the assumed time range of the specified correlation function (semivariogram); otherwise, the elastic parameters for this time sample are determined using a VFSA solution rule.

Kriging of elastic parameters is performed using a fixed semivariogram (autocorrelation) model (e.g. spherical, Gaussian, etc.) and fixed semivariogram parameters (e.g. range, sill, etc.). Different semivariogram models can be used for a specific lithology and for a specific elastic parameter provided that a priori information about the corresponding parameters is available.

4) A Metropolis acceptance/rejection criterion is enforced by the inversion algorithm. Once all the discrete time samples are visited, the current iteration is completed and the target temperature (see equation 1) is decreased according to a prescribed cooling schedule.

5) The process described above is performed iteratively until either the data misfit function decreases to a previously specified value or the number of iterations reaches a previously specified upper bound. According to this iterative procedure, the final solution remains conditioned by the initial realization of model parameters. Therefore, various seeds are necessary to evaluate the variability (uncertainty) of the final solution. Such a strategy provides an efficient way of assessing non-uniqueness and hence of appraising quantitatively the reliability and stability of the inversion results in the presence of noisy and sparsely sampled prestack seismic amplitude data.

As emphasized above, the new global inversion algorithm makes use of a prescribed semivariogram (or autocorrelation) function to enforce a measure of smoothness on the estimated model parameters. In parenthesis, we mention that Tarantola (1987) showed that when an exponential covariance operator (i.e. an exponential semivariogram) is used as an integral kernel \(K\) to estimate a smooth version \(m_{\text{smooth}}\) of the original continuous model \(m\), using the equation, \(m_{\text{smooth}}(y) = \int K(x, y)m(x)dx\), this operation is equivalent to enforcing a model norm written as the weighted sum of the \(L_2\)-norm of the function and the \(L_2\)-norm of its first derivative. In a similar fashion, the kriging operation included in the proposed global inversion algorithm to estimate elastic parameters at time samples between hard points implicitly enforces a mixed quadratic model norm in the inversion. The exact representation of such a mixed quadratic norm depends on the specific choice of semivariogram model and the associated parameters.
Enforcing smoothness and assessing in 1D prestack seismic inversion

Figure 9 Summary of the results obtained with the new global inversion algorithm. Panels (a) and (b) show inversion results in data space and model space, respectively. In panel (b), the inverted distributions of elastic parameters are indicated by open circles and \( r^2 \) is the correlation coefficient between the actual and inverted elastic parameters.

Example no. 5: Single-model realization

Figure 9 shows results in data and model space obtained with the new global inversion algorithm. Panel (a) shows the measured and inverted prestack seismic amplitude data as well as the seismic amplitude residuals. Panels (b) shows the actual and inverted 1D distributions of elastic parameters and their associated correlation coefficients (\( r_{VP}^2 = 0.86 \), \( r_{VS}^2 = 0.98 \), \( r_{\rho_b}^2 = 0.95 \)). These coefficients are higher than those in previous inversion exercises. The number of hard points used by the inversion algorithm is equal to 10\% of the total number of time samples. A zero-nugget, spherical semivariogram of range and normalized sill equal to 0.60 s and 1, respectively, was used for kriging (i.e. to interpolate) elastic parameters between hard points. All inversion results described in Fig. 9 entailed the same similarity in data space and made use of the harmonic data misfit function (equation 5). Smoothness was implicitly enforced through the time-sampling strategy used by the new global inversion algorithm.

Example no. 6: Effect of semivariogram range and number of hard points

Tables 1 and 2 summarize the results obtained when different ranges of vertical correlation and numbers of hard points were used by the inversion algorithm, respectively. All inversion results described in these tables entailed the same similarity in data space and made use of the harmonic data misfit function. The correlation coefficients shown in Table 1 between the actual and inverted bulk density and shear-wave velocity are approximately the same for all vertical correlation ranges under consideration. For the case of compressional-wave velocity, however, the correlation coefficient reaches a peak at a semivariogram range of approximately 0.70 s. As emphasized earlier, the kriging of elastic parameters is, by default, performed using a spherical semivariogram model with a correlation range equal to 0.60 s. Such a range was calculated experimentally from the compressional-wave-velocity log. In the most general case, each elastic parameter and the lithology could be associated with a specific semivariogram model. Table 1 also shows the normalized computer CPU time required to estimate each inversion result. This CPU time increases slightly as the correlation range increases. Table 2 summarizes similar experimental results obtained when using different numbers of hard
Table 1 Summary of the relationship between spatial correlation length, correlation coefficient and CPU time when performing the new global inversion algorithm on the synthetic prestack data set shown in Fig. 1

<table>
<thead>
<tr>
<th>Range [s] (two-way time)</th>
<th>$r_{Vp}^2$</th>
<th>$r_{Vs}^2$</th>
<th>$r_{\rho}^2$</th>
<th>Relative CPU time</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>0.057</td>
<td>0.974</td>
<td>0.976</td>
<td>0.878</td>
</tr>
<tr>
<td>0.20</td>
<td>0.078</td>
<td>0.974</td>
<td>0.956</td>
<td>0.899</td>
</tr>
<tr>
<td>0.50</td>
<td>0.763</td>
<td>0.964</td>
<td>0.951</td>
<td>0.878</td>
</tr>
<tr>
<td>0.70</td>
<td>0.814</td>
<td>0.9464</td>
<td>0.944</td>
<td>0.906</td>
</tr>
<tr>
<td>0.90</td>
<td>0.594</td>
<td>0.971</td>
<td>0.962</td>
<td>0.846</td>
</tr>
<tr>
<td>1.10</td>
<td>0.556</td>
<td>0.957</td>
<td>0.949</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Table 2 Summary of the relationship between number of hard points, correlation coefficient and CPU time when performing the new global inversion algorithm on the synthetic prestack data set shown in Fig. 1

<table>
<thead>
<tr>
<th>No. of hard points</th>
<th>$r_{Vp}^2$</th>
<th>$r_{Vs}^2$</th>
<th>$r_{\rho}^2$</th>
<th>Relative CPU time</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.846</td>
<td>0.995</td>
<td>0.973</td>
<td>1.000</td>
</tr>
<tr>
<td>5</td>
<td>0.890</td>
<td>0.981</td>
<td>0.976</td>
<td>0.912</td>
</tr>
<tr>
<td>9</td>
<td>0.846</td>
<td>0.938</td>
<td>0.953</td>
<td>0.906</td>
</tr>
<tr>
<td>11</td>
<td>0.795</td>
<td>0.996</td>
<td>0.982</td>
<td>0.820</td>
</tr>
<tr>
<td>15</td>
<td>0.894</td>
<td>0.995</td>
<td>0.940</td>
<td>0.832</td>
</tr>
<tr>
<td>17</td>
<td>0.824</td>
<td>0.914</td>
<td>0.945</td>
<td>0.952</td>
</tr>
</tbody>
</table>

points. Neither the correlation coefficients between the actual and inverted elastic parameters nor the computer CPU times exhibit significant differences for the various numbers of hard points considered in the experiments.

Example no. 7: Assessment of uncertainty

Figure 10 shows results from the evaluation of uncertainty of the inverted 1D distributions of elastic parameters. To construct the results shown in this figure, 22 independent estimations were performed with the new global inversion algorithm. Each of the inversions was initialized with an independent starting model but with the same number of fixed points. The inversion results shown in this figure entailed the same similarity in data space. Subsequently, histograms for each set of parameters were sampled at a particular time and used as indicators of variability. These histograms are shown in colour-coded format in Fig. 10. Clearly, the compressional-wave velocity exhibits the greatest variability among the three elastic parameters. Such behaviour is due to the relatively large range of constraints enforced by the inversion algorithm for this elastic parameter (see Fig. 5a). In general, the inverted distributions of elastic parameters remain consistent with the actual distributions (identified by red lines in Fig. 10). The overall computer efficiency of the proposed algorithm surpasses the efficiency of the alternative global inversions techniques considered here.

FEASIBILITY ANALYSIS FOR THE INVERSION OF FIELD DATA

Description of the Gulf of Mexico data set

The 1D model under consideration was constructed from actual well-log data acquired in the deepwater Gulf of Mexico. Figure 11 shows the corresponding well logs along the depth interval from 8000 ft (2438.4 m) to 13 200 ft (4023.4 m). A
Enforcing smoothness and assessing in 1D prestack seismic inversion

Figure 10 Evaluation of the uncertainty of the 1D distributions of elastic parameters yielded by the new global inversion algorithm. Panels (a), (b) and (c) show colour-coded normalized histograms for bulk density, compressional- and shear-wave velocities, respectively, calculated from 22 independent realizations of inverted elastic parameters. All the realizations entail the same similarity in data space. The red line indicates the actual distribution of elastic parameters.

A close-up view of the same logs is also shown in the interval of interest from 11 600 ft (3535.7 m) to 13 200 ft (4023.4 m). The well-log data were converted from depth to normal seismic traveltime using the compressional-wave-velocity log. Subsequently, the logs were sampled at a constant rate of 2 ms to construct the discretized version of elastic parameters. A zero-phase Ricker wavelet centred at 35 Hz was used for the simulation of the 1D prestack seismic data in the interval from 2.4 to 3.7 s. The assumed source–receiver prestack gather consists of 10 traces, corresponding to the same number of receivers spaced at 800 m intervals. The maximum offset-to-depth ratio for the simulations is approximately equal to 2 and the maximum angle of incidence is 60°.

Figure 12 shows the simulated prestack seismic traces contaminated with 5% of zero-mean additive Gaussian random noise. Noisy prestack and NMO-corrected seismic amplitude data input to the new global inversion algorithm comprised only the zone of interest, from 3.3 to 3.7 s. This same time interval comprised the discretized elastic parameters, hence amounting to a total of 200 equal traveltime layers considered for inversion. The experimental spherical semivariogram used in the inversion algorithm was inferred from the well-log data and exhibited a correlation range of 0.04 s and a normalized sill of 1.

Example no. 8: Assessment of uncertainty

Figure 13 shows graphically the results obtained from the evaluation of uncertainty for the Gulf of Mexico 1D subsurface model using the new global inversion algorithm. 22 independent inversions were performed to estimate the 1D distribution of elastic parameters. All inversion results shown in this figure entailed the same similarity in data space. Inversion results are displayed in the form of colour-coded normalized histograms as a function of seismic traveltime. Each histogram was constructed at a particular time from samples of elastic parameters obtained from the 22 independent estimations. As indicated in Fig. 13, shear-wave velocity and bulk density exhibit the most variability with respect to the actual well-log parameters (identified by
solid red lines in each panel). In general, the inverted 1D distributions of elastic parameters are in qualitative agreement with the well-log distributions of the same parameters. The variability (uncertainty) of the inverted results increases considerably within thin sand layers.

CONCLUSIONS

Various simulated annealing algorithms, objective functions, sampling strategies and smoothing criteria were explored to perform 1D global inversion of prestack seismic amplitude data. Both the selection of a specific random-walk search method and the construction of the objective function significantly constrained the efficiency of the inversion and the accuracy and reliability of the results. Sampling strategies allowed us to enforce a priori information in the inversion, including an additive smoothness term that helps to constrain the range of possible solutions in model space. This paper introduced a new global inversion algorithm based on simulated annealing that enforced sampling and smoothness strategies widely used in the field of geostatistics. The
Figure 12 Prestack seismic traces simulated numerically for the 1D distributions of elastic parameters constructed with the well logs shown in Figure 11. Offset receivers are assumed uniformly spaced at 800 m intervals. Simulation of the prestack seismic amplitude data was performed with a 35 Hz Ricker wavelet. The simulated prestack seismic amplitude data shown in this figure were further corrected for normal moveout and contaminated with 5% zero-mean Gaussian additive noise (maximum offset-to-depth ratio approximately equal to 2). Inversion was performed using data within the time interval from 3.3 to 3.7 s.

The overall computer efficiency of the proposed algorithm is superior to alternative inversion techniques evaluated in this study. Moreover, the new global inversion algorithm provides estimates of uncertainty in an efficient manner. The flexibility of the proposed algorithm also makes it possible to estimate petrophysical properties, such as porosity and water saturation, which are either deterministically or stochastically related to the elastic parameters.
Figure 13 Appraisal of resolution and uncertainty of prestack seismic data for the subsurface model in the deepwater Gulf of Mexico. Elastic parameters estimated with the new global inversion algorithm: panels (a) (b) and (c) show colour-coded normalized histograms for bulk density, compressional- and shear-wave velocities, respectively, calculated from 22 independent realizations of inverted elastic parameters. All of the realizations entailed the same similarity in data space. The red line indicates the actual well-log data.

ACKNOWLEDGEMENTS

We are grateful to Dr Roy White and two anonymous reviewers for their constructive technical and editorial feedback. The work reported in this paper was supported by the US Department of Energy under contract No. DE-FC26–00BC15305.

REFERENCES