GEOPHYSICS, VOL. 68, NO. 1 (JANUARY-FEBRUARY 2003); P. 185–198, 16 FIGS., 3 TABLES.
10.1190/1.1543206

Bayesian linearized AVO inversion

Arild Buland* and Henning Omre‡

ABSTRACT

A new linearized AVO inversion technique is developed in a Bayesian framework. The objective is to obtain posterior distributions for P-wave velocity, S-wave velocity, and density. Distributions for other elastic parameters can also be assessed—for example, acoustic impedance, shear impedance, and P-wave to S-wave velocity ratio. The inversion algorithm is based on the convolutional model and a linearized weak contrast approximation of the Zoeppritz equation. The solution is represented by a Gaussian posterior distribution with explicit expressions for the posterior expectation and covariance; hence, exact prediction intervals for the inverted parameters can be computed under the specified model. The explicit analytical form of the posterior distribution provides a computationally fast inversion method. Tests on synthetic data show that all inverted parameters were almost perfectly retrieved when the noise approached zero. With realistic noise levels, acoustic impedance was the best determined parameter, while the inversion provided practically no information about the density. The inversion algorithm has also been tested on a real 3-D data set from the Sleipner field. The results show good agreement with well logs, but the uncertainty is high.

INTRODUCTION

The objectives of geophysical inverse problems are to make inferences about model parameters based on general knowledge and a set of geophysical measurements. In general, these inverse problems are multidimensional and ill posed, and they are often strongly affected by noise and measurement uncertainty.

In a statistical perspective, the solution of an inverse problem is not limited to a single set of predicted parameters but is represented by a probability density function (PDF) on the model space. The aim of inversion is not only to find a best-fitting set of model parameters but also to characterize the uncertainty in the inversion results. A Bayesian setting is a natural choice for many geophysical inverse problems, where it is possible to combine available prior knowledge with the information contained in the measured data (Tarantola and Valette, 1982; Tarantola, 1987; Duijndam, 1988a,b; Scales and Tenorio, 2001; Ulrych et al., 2001). The solution of a Bayesian inverse problem is represented by the posterior distribution. Analysis of nonuniqueness and uncertainty in the solutions of inverse problems is often called resolution analysis. In a Bayesian setting, all questions about resolution are addressed by the posterior distribution. The Bayesian concept is general and is recognized both in statistics and in geophysics, but the specific model definition and the associated solution of an actual problem may be complicated.

A general geophysical inverse problem has a nonlinear relationship between the model parameters and the measured data, and the posterior distribution can usually be found only by use of stochastic simulation techniques (Monte Carlo) (Mosegaard and Tarantola, 1995; Sen and Stoffa, 1996; Mosegaard, 1998; Eide et al., 2002). However, stochastic simulation of inverse problems involving seismic data is in most cases impractical because of high computational costs. For linear problems, or problems that can be linearized, analytical solutions may in some cases be found. A solution with an explicit analytical form is usually computationally superior to iterative search- and simulation-based solutions.

Amplitude versus offset (AVO) inversion is a seismic prestack inversion technique for estimating elastic subsurface parameters. AVO inversion utilizes the fact that the reflection strength from subsurface interfaces depends on the reflection angle and on the material properties where the reflections take place. The AVO inversion methods can be separated into nonlinear methods (Dahl and Ursin, 1991; Bulaand et al., 1996; Gouveia and Scales, 1997; 1998) and linearized inversion (Smith and Gidlow, 1987; Lörtzer and Berkhout, 1993). Gouveia and Scales (1998) define a Bayesian nonlinear model

Manuscript received by the Editor October 30, 2000; revised manuscript received March 11, 2002.
*Statoil Research Centre, Postuttak, N-7005 Trondheim, Norway, and Norwegian University of Science and Technology, N-7491 Trondheim, Norway. E-mail: abu@statoil.com.
‡Norwegian University of Science and Technology, Department of Mathematical Sciences, N-7491 Trondheim, Norway. E-mail: omre@stat.ntnu.no.

and optimize the posterior distribution via a nonlinear conjugated gradient procedure to find the maximum a posteriori model (MAP). They performed the uncertainty analysis by making a Gaussian approximation of the posterior distribution centered at the MAP solution. This approach is much faster than using Monte Carlo simulation; but since the problem is complex and strongly nonlinear, the method is still computationally intensive compared to linearized AVO inversion. In the linearized AVO inversion methods, the Zoeppritz equation for the reflection coefficient is linearized assuming weak contrasts. The seismic data must be processed prior to linear AVO inversion to remove nonlinear relations between the model parameters and the seismic response, such as offset-dependent traveltime and geometrical spreading.

In this paper, a Bayesian linearized AVO inversion method is presented, where the objective is to obtain posterior distributions for the P-wave velocity, S-wave velocity, and density. Closest to our approach is the method presented in Lörtzer and Berkhout (1993). Their inversion algorithm, however, was formulated for relative contrasts of the elastic parameters and not for the elastic parameters themselves. Transformation to the elastic parameters may not be simple in a stochastic setting. Lörtzer and Berkhout use standard AVO techniques based on single interface theory, which are then applied independently on a sample-by-sample basis. We base our inversion algorithm on a weak contrast reflectivity function defined for continuous seismic traveltime. The inversion problem is solved simultaneously for all times in a given time window, which makes it possible to include the wavelet by convolution and temporal correlation between model parameters close in time. The solution is represented by a Gaussian posterior distribution with explicit expressions for the posterior expectation and covariance. The explicit analytical form of the posterior distribution provides a computationally fast inversion method. A continuous definition of the elastic parameters makes it possible to integrate various types of observations with varying sampling density and sampling support. Moreover, generalization to a full 3D model with spatial dependencies is in principle straightforward, but an efficient numerical technique is required.

The inversion result is obtained from true-amplitude processed seismic prestack data. Important steps in the processing are the removal of multiples and corrections for the effects of geometrical spreading and absorption. We assume that wave mode conversions, interbed multiples, and anisotropy effects can be neglected after processing. The data should be prestack migrated, such that dip-related effects and the moveouts are removed. After prestack migration, we assume that each single bin gather can be regarded as the response of a local 1-D earth model. Finally, the gathers must be transformed from offsets to angles, either by common-angle migration or by a transform after the migration.

In the following sections, the methodology, synthetic tests, and application of the inversion method on a real 3-D data set are presented.

METHODOLOGY

Seismic reflection coefficients depend on the material properties of the subsurface. An isotropic, elastic medium is completely described by three material parameters. We have chosen $\{\alpha(t), \beta(t), \rho(t)\}$, where α is P-wave velocity, β is S-wave velocity, ρ is density, and t is the two-way vertical seismic traveltime. Several other parameterizations can also be used, for example, $\{Z_P(t), Z_S(t), \rho(t)\}$, where $Z_P = \alpha\rho$ is the acoustic impedance and $Z_S = \beta\rho$ is the shear impedance. The shear impedance is often substituted by the P-wave to S-wave velocity ratio α/β. When inversion results are obtained for a chosen parameter set, the corresponding results can be generated for another parameter set, including uncertainty.

The inversion method is based on a weak contrast approximation to the PP reflection coefficient (Aki and Richards, 1980):

$$c_{PP}(\theta) = a_\alpha(\theta)\frac{\Delta\alpha}{\bar{\alpha}} + a_\beta(\theta)\frac{\Delta\beta}{\bar{\beta}} + a_\rho(\theta)\frac{\Delta\rho}{\bar{\rho}}, \quad (1)$$

where

$$a_\alpha(\theta) = \frac{1}{2}(1 + \tan^2\theta), \quad (2)$$

$$a_\beta(\theta) = -4\frac{\bar{\beta}^2}{\bar{\alpha}^2}\sin^2\theta, \quad (3)$$

$$a_\rho(\theta) = \frac{1}{2}\left(1 - 4\frac{\bar{\beta}^2}{\bar{\alpha}^2}\sin^2\theta\right). \quad (4)$$

In addition, $\bar{\alpha}$, $\bar{\beta}$, and $\bar{\rho}$ are averages over the reflecting interface; $\Delta\alpha$, $\Delta\beta$, and $\Delta\rho$ are the corresponding contrasts; and θ is the reflection angle. We discuss only PP reflectivity in this paper, but the extension to include converted waves is straightforward.

The single-interface reflection coefficient in equation (1) can be extended to a time-continuous reflectivity function (Stolt and Weglein, 1985):

$$c_{PP}(t,\theta) = a_\alpha(t,\theta)\frac{\partial}{\partial t}\ln\alpha(t) + a_\beta(t,\theta)\frac{\partial}{\partial t}\ln\beta(t) + a_\rho(t,\theta)\frac{\partial}{\partial t}\ln\rho(t), \quad (5)$$

where $a_\alpha(t,\theta)$, $a_\beta(t,\theta)$, and $a_\rho(t,\theta)$ are generalizations of the coefficients in expressions (2)–(4) with time-dependent velocities $\bar{\alpha}(t)$ and $\bar{\beta}(t)$. We assume that $\bar{\alpha}(t)$ and $\bar{\beta}(t)$ can be represented by a constant or slowly varying known background model, such that $\bar{\alpha}(t)$ and $\bar{\beta}(t)$ are the average or moving average of $\alpha(t)$ and $\beta(t)$ in a time window. The inversion algorithm requires that this background model is known prior to the inversion so the coefficients can be precalculated.

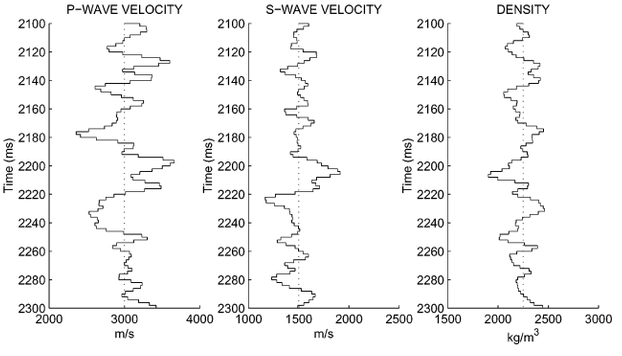

FIG. 1. The P-wave velocity, S-wave velocity, and density in well A (solid lines). The constant background model is dotted.

In the formulas above, the reflection angle θ is used as an independent variable. However, the seismic data are recorded as a function of source–receiver distance h (offset). The transform of the data from the (t, h) domain to the (t, θ) domain depends on the velocity function. This transform can be performed by common-angle migration, ray tracing, or standard approximate offset-angle relations based on a smooth background model.

The prior model

In a Bayesian setting, the prior model defines a statistical model for the prior information of the material parameters, expressed mathematically by PDFs. The prior information must be independent of the seismic data and is established from other available information and knowledge. It is in a good statistical tradition to be vague when specifying the prior model.

The material parameters $\{\alpha(t), \beta(t), \rho(t)\}$ are assumed to be log-Gaussian, which implies that the parameters are restricted to take positive values. This assumption is required for later analytical treatment because of expression (5). The logarithm of these material parameters defines a continuous Gaussian vector field

$$\mathbf{m}(t) = [\ln \alpha(t), \ln \beta(t), \ln \rho(t)]^T \tag{6}$$

with expectation

$$\mathrm{E}\{\mathbf{m}(t)\} = \boldsymbol{\mu}(t) = [\mu_\alpha(t), \mu_\beta(t), \mu_\rho(t)]^T, \tag{7}$$

where the elements $\mu_\alpha(t), \mu_\beta(t)$, and $\mu_\rho(t)$ are the expectations of $\ln \alpha(t)$, $\ln \beta(t)$, and $\ln \rho(t)$, respectively. We assume that the expectation functions are smooth. The covariances between $\ln \alpha$, $\ln \beta$, and $\ln \rho$ at times t and s are denoted

$$\mathrm{Cov}\{\mathbf{m}(t), \mathbf{m}(s)\} = \boldsymbol{\Sigma}(t, s). \tag{8}$$

A simple example is the stationary covariance function

$$\boldsymbol{\Sigma}(t, s) = \boldsymbol{\Sigma}_0\ \nu_t(\tau), \tag{9}$$

where $\nu_t(\tau)$ is a temporal correlation function and τ is the time lag $\tau = |t - s|$. The time-invariant covariance matrix $\boldsymbol{\Sigma}_0$ is

$$\boldsymbol{\Sigma}_0 = \begin{bmatrix} \sigma_\alpha^2 & \sigma_\alpha\sigma_\beta\nu_{\alpha\beta} & \sigma_\alpha\sigma_\rho\nu_{\alpha\rho} \\ \sigma_\alpha\sigma_\beta\nu_{\alpha\beta} & \sigma_\beta^2 & \sigma_\beta\sigma_\rho\nu_{\beta\rho} \\ \sigma_\alpha\sigma_\rho\nu_{\alpha\rho} & \sigma_\beta\sigma_\rho\nu_{\beta\rho} & \sigma_\rho^2 \end{bmatrix}, \tag{10}$$

where the diagonal elements are the variances and where $\nu_{\alpha\beta}$, $\nu_{\alpha\rho}$, and $\nu_{\beta\rho}$ are the correlations between $\ln \alpha(t)$, $\ln \beta(t)$, and $\ln \rho(t)$, respectively. The temporal correlation function $\nu_t(\tau)$ must be a positive definite function, take values in the interval $[-1, 1]$, and have the property that $\nu_t(0) = 1$. One such correlation function is the second-order exponential function

$$\nu_t(\tau) = \exp\left[-\left(\frac{\tau}{d}\right)^2\right], \tag{11}$$

where d is a range parameter characterizing the temporal dependency.

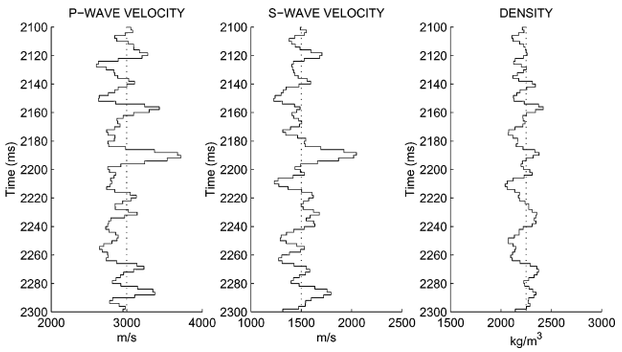

FIG. 2. The P-wave velocity, S-wave velocity, and density in well B (solid lines). The constant background model is dotted.

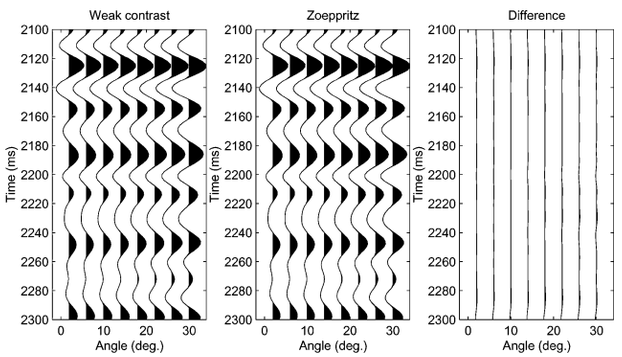

FIG. 3. Weak contrast modeling in well A with constant background (left), ray tracing with full Zoeppritz equation (middle), and residual (right).

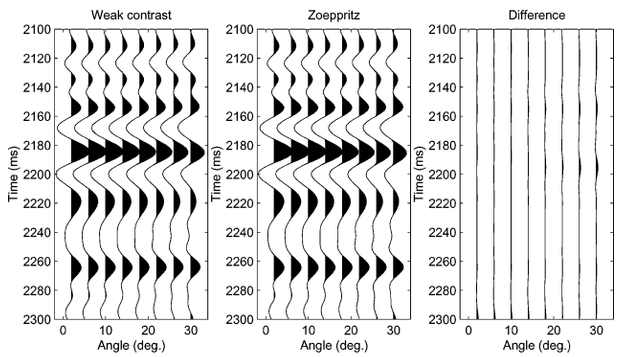

FIG. 4. Weak contrast modeling in well B with constant background (left), ray tracing with full Zoeppritz equation (middle), and residual (right).

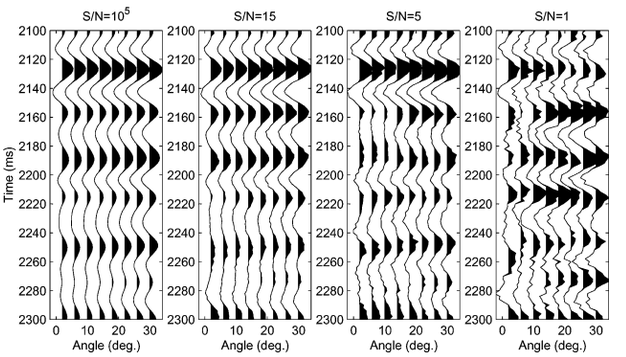

FIG. 5. Synthetic CDP gathers for model A with different noise levels.

The time derivative of $\mathbf{m}(t)$ defines a new vector field

$$\mathbf{m}'(t) = \left[\frac{\partial}{\partial t}\ln\alpha(t), \frac{\partial}{\partial t}\ln\beta(t), \frac{\partial}{\partial t}\ln\rho(t)\right]^T, \qquad (12)$$

where the elements can be recognized in the reflectivity function, expression (5). This is also a Gaussian vector field due to the linearity of the differentiation process (Christakos, 1992). The expectation of $\mathbf{m}'(t)$,

$$\mathrm{E}\{\mathbf{m}'(t)\} = \frac{\partial}{\partial t}\boldsymbol{\mu}(t) = \boldsymbol{\mu}'(t), \qquad (13)$$

and the covariance

$$\mathrm{Cov}\{\mathbf{m}'(t), \mathbf{m}'(s)\} = \frac{\partial^2}{\partial t\,\partial s}\boldsymbol{\Sigma}(t,s) = \boldsymbol{\Sigma}''(t,s), \qquad (14)$$

are both derived from the specified prior model for $\mathbf{m}(t)$. The stochastic model for the differentiated field $\mathbf{m}'(t)$ is needed later but is discussed here because of the close relationship to $\mathbf{m}(t)$. The cross-covariance between $\mathbf{m}'(t)$ and $\mathbf{m}(s)$ is

$$\mathrm{Cov}\{\mathbf{m}'(t), \mathbf{m}(s)\} = \frac{\partial}{\partial t}\boldsymbol{\Sigma}(t,s) = \boldsymbol{\Sigma}'(t,s). \qquad (15)$$

For the stationary covariance function in expression (9), the covariance in expression (14) and the cross-covariance in

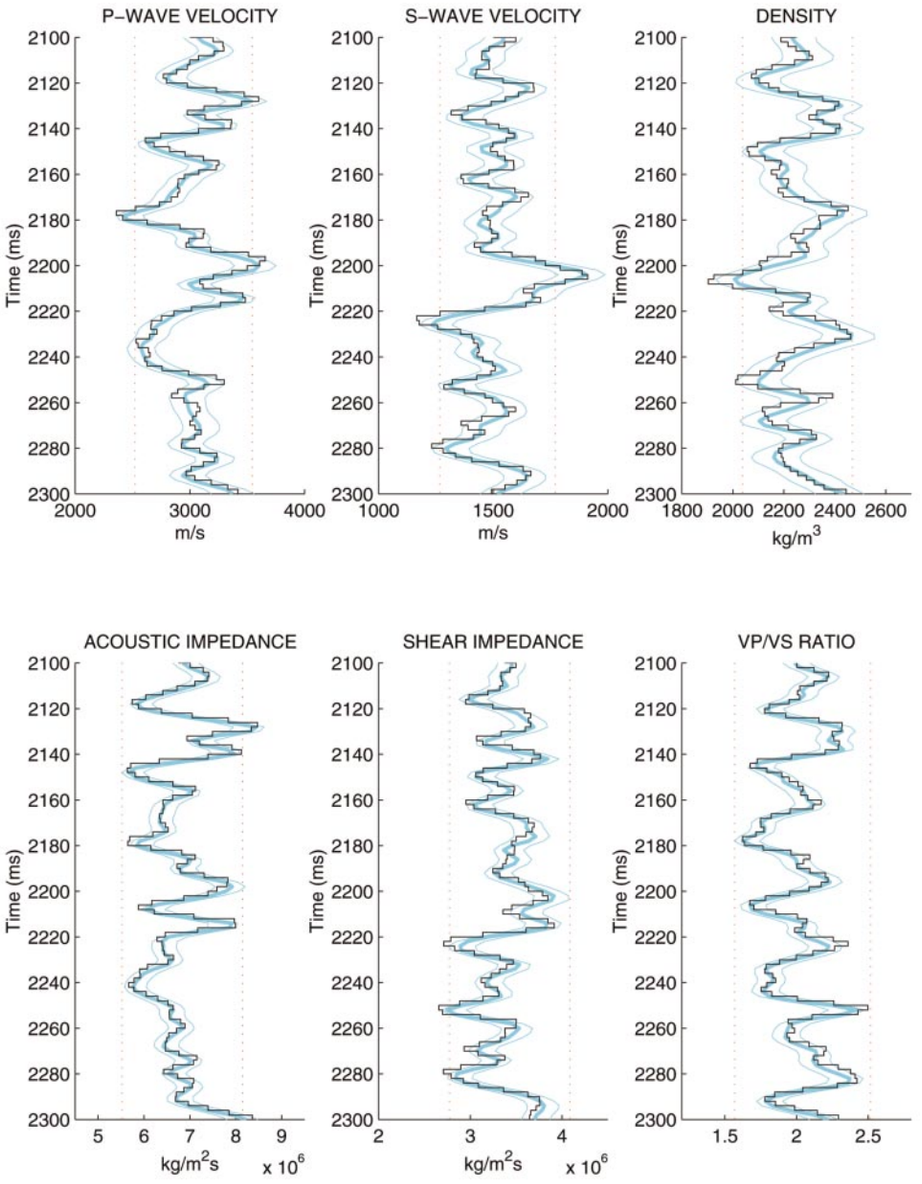

FIG. 6. The MAP solution (thick blue line) of model A with S/N ratio 10^5 with 0.95 prediction interval (thin blue lines), the true earth profile (black line), and 0.95 prior model interval (red dotted lines).

expression (15) are

$$\mathrm{Cov}\{\mathbf{m}'(t), \mathbf{m}'(s)\} = -\boldsymbol{\Sigma}_0 \frac{\partial^2}{\partial \tau^2} \nu_t(\tau) \tag{16}$$

and

$$\mathrm{Cov}\{\mathbf{m}'(t), \mathbf{m}(s)\} = \mathrm{sign}(t-s)\boldsymbol{\Sigma}_0 \frac{\partial}{\partial \tau} \nu_t(\tau), \tag{17}$$

respectively, where sign$(\cdot)$ returns the sign of the argument.

The continuous form of the Gaussian field $\mathbf{m}(t)$ makes it possible to give a proper definition of the time differentiated Gaussian field $\mathbf{m}'(t)$. In a computer program, the continuous fields are represented on a grid. The grid density should be determined by the temporal variability of the elastic parameters, and not by the sampling density of the seismic data. A discrete representation of $\mathbf{m}(t)$ in a time interval is Gaussian:

$$\mathbf{m} = [\ln \boldsymbol{\alpha}^T, \ln \boldsymbol{\beta}^T, \ln \boldsymbol{\rho}^T]^T \sim \mathcal{N}_{nm}(\boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m), \tag{18}$$

where the vectors $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\rho}$ are discrete representations of $\alpha(t)$, $\beta(t)$, and $\rho(t)$, respectively. The logarithm operates elementwise on the vector elements, $\sim$ means distributed as, and $\mathcal{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes an n-dimensional Gaussian distribution with expectation $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$ (see appendix A). The dimension of $\mathbf{m}$ is n_m, and $\boldsymbol{\mu}_m$ and $\boldsymbol{\Sigma}_m$ are defined from the continuous analogs $\boldsymbol{\mu}(t)$ and $\boldsymbol{\Sigma}(t, s)$, respectively. A discrete representation of the differentiated field $\mathbf{m}'(t)$ in a time interval is Gaussian:

$$\mathbf{m}' \sim \mathcal{N}_{nm}(\boldsymbol{\mu}'_m, \boldsymbol{\Sigma}''_m), \tag{19}$$

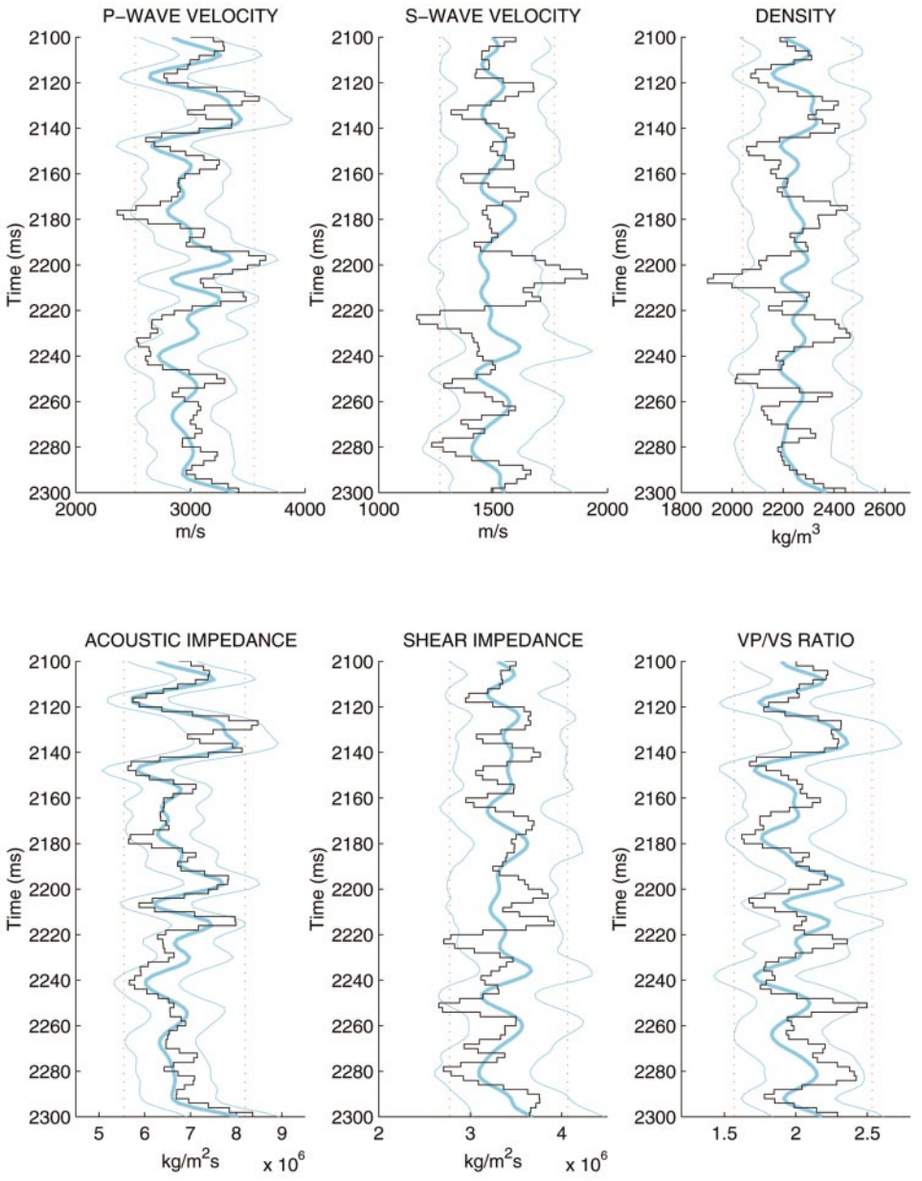

FIG. 7. The MAP solution (thick blue line) of model A with S/N ratio = 5. A 0.95 prediction interval (thin blue lines), the true earth profile (black line), and 0.95 prior model interval (red dotted lines).

with $\boldsymbol{\mu}'_m$ and $\boldsymbol{\Sigma}''_m$ being defined from the continuous analogs in expressions (13) and (14), respectively. A discrete version of the cross-covariance in expression (15) is denoted $\boldsymbol{\Sigma}'_m$.

Seismic forward modeling

In Bayesian AVO inversion, the likelihood function can be defined through a PDF for the observed seismic data $\mathbf{d}_{obs}$ given a specific parameter vector $\mathbf{m}$, denoted $p(\mathbf{d}_{obs}|\mathbf{m})$. The likelihood function describes how likely the observed data are for these specific parameters. The unconditional PDF for the seismic data, denoted $p(\mathbf{d}_{obs})$, is related to the likelihood function by

$$p(\mathbf{d}_{obs}) = \int p(\mathbf{d}_{obs}|\mathbf{m})\, p(\mathbf{m})\, d\mathbf{m}, \tag{20}$$

where $p(\mathbf{m})$ is the prior model for the parameter vector $\mathbf{m}$. Below, the PDF $p(\mathbf{d}_{obs})$ is defined and is later used to derive the posterior distribution $p(\mathbf{m}|\mathbf{d}_{obs})$.

A discrete version of the reflectivity function $c_{PP}(t,\theta)$ [expression (5)] for a given time interval and a set of reflection angles can be written

$$\mathbf{c} = \mathbf{A}\mathbf{m}', \tag{21}$$

where the matrix $\mathbf{A}$ is defined by the coefficients $a_\alpha(t,\theta)$, $a_\beta(t,\theta)$, and $a_\rho(t,\theta)$ (see appendix B). The corresponding seismic angle gather $\mathbf{d}_{obs}$ is represented by the convolutional model,

$$\mathbf{d}_{obs} = \mathbf{S}\mathbf{c} + \mathbf{e} = \mathbf{S}\,\mathbf{A}\mathbf{m}' + \mathbf{e}, \tag{22}$$

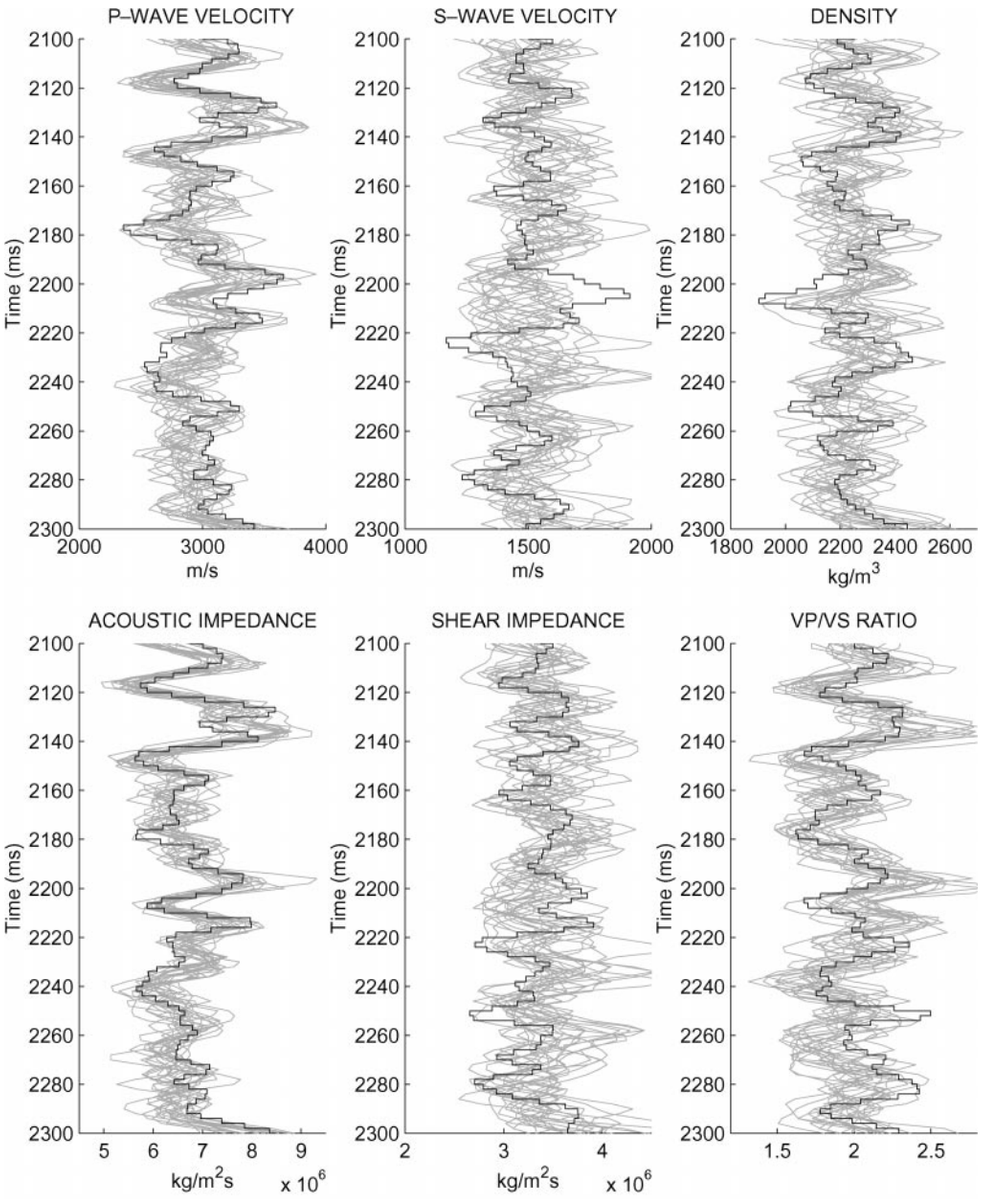

FIG. 8. Realizations from the posterior distribution (gray lines) for model B with S/N ratio = 5. The true earth profile is shown in black.

where the convolution is formulated by the matrix-vector multiplication **Sc**. The matrix **S** represents one wavelet for each reflection angle (see appendix B). If we assume the error term **e** is zero-mean Gaussian

$$\mathbf{e} \sim \mathcal{N}_{n_d}(\mathbf{0}, \boldsymbol{\Sigma}_e) \tag{23}$$

and independent of **m**, then $\mathbf{d}_{obs}$ is Gaussian

$$\mathbf{d}_{obs} \sim \mathcal{N}_{n_d}(\boldsymbol{\mu}_{d_{obs}}, \boldsymbol{\Sigma}_{d_{obs}}), \tag{24}$$

where

$$\boldsymbol{\mu}_{d_{obs}} = \mathbf{S}\mathbf{A}\boldsymbol{\mu}_m', \tag{25}$$

$$\boldsymbol{\Sigma}_{d_{obs}} = \mathbf{S}\mathbf{A}\boldsymbol{\Sigma}_m''\mathbf{A}^T\mathbf{S}^T + \boldsymbol{\Sigma}_e \tag{26}$$

using result 1 in appendix A.

Table 1. Percentwise decrease of the width of the 0.95 interval from the prior model to the posterior model for model A.

S/N ratio	α	β	ρ	Z_P	Z_S	α/β
10^5	71	66	58	79	75	78
15	37	20	12	46	26	39
5	33	11	8	40	11	30
1	21	3	4	25	3	17

Table 2. Percentwise decrease of the width of the 0.95 interval from the prior model to the posterior model for model B.

S/N ratio	α	β	ρ	Z_P	Z_S	α/β
10^5	77	73	69	81	78	75
15	45	29	33	49	36	22
5	39	21	28	41	26	16
1	25	12	20	27	17	7

Table 3. Processing sequence.

1.	SEGD read and data editing
2.	Navigation merge
3.	Gun, cable, and filter delay correction (−38.6 ms)
4.	Deterministic zero phasing
5.	Lowcut filter 6 Hz
6.	Gain correction t^2
7.	Tidal correction
8.	Swell noise and interference attenuation
9.	Forward $\tau - p$ transformation
10.	Predicted deconvolution 64/300-ms gap/operator
11.	Inverse $\tau - p$ transformation
12.	Old velocity input
13.	NMO
14.	k-filter spatial resampling
15.	CMPs to 75 m trace spacing
16.	Sort to 39 offset volumes, 50% cross-line binning
17.	$f - x$ trace interpolation
18.	Static correction
19.	DMO, reduction to 19 offsets
20.	Prestack time migration with minimum velocity
21.	Inverse NMO with old velocities
22.	New velocity picking
23.	NMO with new velocities
24.	Transform to time angle by ray tracing

The posterior model

A combination of equations (18) and (24) gives the joint distribution

$$\begin{bmatrix} \mathbf{m} \\ \mathbf{d}_{obs} \end{bmatrix} \sim \mathcal{N}_{n_m+n_d}\left(\begin{bmatrix} \boldsymbol{\mu}_m \\ \boldsymbol{\mu}_{d_{obs}} \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_m & \boldsymbol{\Sigma}_{m,d_{obs}} \\ \boldsymbol{\Sigma}_{d_{obs},m} & \boldsymbol{\Sigma}_{d_{obs}} \end{bmatrix}\right), \tag{27}$$

where $\boldsymbol{\Sigma}_{d_{obs},m}$ is the cross-correlation between $\mathbf{d}_{obs}$ and **m**,

$$\boldsymbol{\Sigma}_{d_{obs},m} = \mathrm{Cov}\{\mathbf{d}_{obs}, \mathbf{m}\} = \mathbf{S}\mathbf{A}\boldsymbol{\Sigma}_m', \tag{28}$$

and $\boldsymbol{\Sigma}_{m,d_{obs}}$ is the transpose of $\boldsymbol{\Sigma}_{d_{obs},m}$.

The posterior distribution for **m** given $\mathbf{d}_{obs}$ is Gaussian

$$\mathbf{m}|\mathbf{d}_{obs} \sim \mathcal{N}_{n_m}\left(\boldsymbol{\mu}_{m|d_{obs}}, \boldsymbol{\Sigma}_{m|d_{obs}}\right), \tag{29}$$

where the posterior expectation and covariance are

$$\boldsymbol{\mu}_{m|d_{obs}} = \boldsymbol{\mu}_m + (\mathbf{S}\mathbf{A}\boldsymbol{\Sigma}_m')^T\boldsymbol{\Sigma}_{d_{obs}}^{-1}\left(\mathbf{d}_{obs} - \boldsymbol{\mu}_{d_{obs}}\right), \tag{30}$$

$$\boldsymbol{\Sigma}_{m|d_{obs}} = \boldsymbol{\Sigma}_m - (\mathbf{S}\mathbf{A}\boldsymbol{\Sigma}_m')^T\boldsymbol{\Sigma}_{d_{obs}}^{-1}\mathbf{S}\mathbf{A}\boldsymbol{\Sigma}_m', \tag{31}$$

using the general formula for a conditional Gaussian distribution (see appendix A). The posterior distribution, which is found on an explicit analytical form, contains the complete solution of the inverse problem, including uncertainty. A set of possible solutions can be generated by stochastic simulation from the posterior distribution: $\mathbf{m}^{(1)}, \mathbf{m}^{(2)}, \ldots, \mathbf{m}^{(n)}$. Since **m** represents the logarithm of the elastic material parameters, the corresponding set of simulated solutions of the P-wave velocity, S-wave velocity, and density are obtained by the inverse transform $\exp[\mathbf{m}^{(i)}]$, $i = 1, \ldots, n$.

The maximum a posteriori (MAP) solution for **m** is equal to the posterior expectation $\mathrm{MAP}\{\mathbf{m}\} = \boldsymbol{\mu}_{m|d_{obs}}$ since the posterior distribution is Gaussian. The MAP solution is generally smoother than a single realization $\mathbf{m}^{(i)}$. When the noise level in expression (22) increases, the posterior distribution converges to the prior distribution. In the case where the noise level is high, the MAP solution is defined mainly from the maximum of the prior distribution, which we assume to be smooth. A

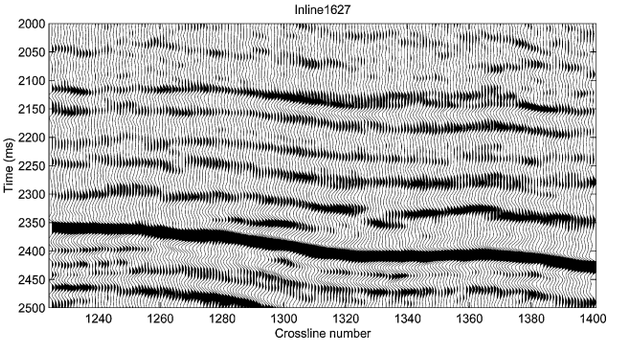

FIG. 9. Stack section of in-line 1627. The data are stacked after the processing listed in Table 3. The strong continuous reflector appearing between 2350 and 2450 ms is the Top Shetland horizon (base reservoir)

single realization, however, is a possible solution with the full variability defined mainly from the prior distribution. In the opposite case, when the noise approaches zero, the solution is determined mainly from the seismic data and the influence of the prior model decreases.

The posterior distributions for the P-wave velocity, S-wave velocity, and density are log-Gaussian. The MAP solution for the elastic parameters $\exp[\mathbf{m}]$ can be determined componentwise by

$$\mathrm{MAP}\{\exp[m]\} = \exp\left[\mu_{m|d_{obs}} - \sigma^2_{m|d_{obs}}\right], \tag{32}$$

while the posterior expectation is

$$\mathrm{E}\{\exp[m]\} = \exp\left[\mu_{m|d_{obs}} + \sigma^2_{m|d_{obs}}/2\right], \tag{33}$$

where m is a component of $\mathbf{m}$, $\mu_{m|d_{obs}}$ is the posterior expectation for m, and $\sigma^2_{m|d_{obs}}$ is the posterior variance. Typically, $\sigma^2_{m|d_{obs}} \ll \mu_{m|d_{obs}}$, such that the log-Gaussian distribution is close to symmetric and $\exp[\mu_{m|d_{obs}}]$ is a good approximation for both $\mathrm{MAP}\{\exp[m]\}$ and $\mathrm{E}\{\exp[m]\}$.

A $(1-\epsilon)$ prediction interval for a specific parameter m—for example, $\ln \alpha(t)$—is given by

$$\mu_{m|d_{obs}} \pm z_{\frac{\epsilon}{2}} \sigma_{m|d_{obs}}, \tag{34}$$

where z_ϵ is the ϵ-quantile in the standard Gaussian distribution. A 0.95 interval for $\alpha(t)$ is $\exp[\mu_{m|d_{obs}} \pm 1.96\sigma_{m|d_{obs}}]$. Prediction intervals for Z_P, Z_S, α/β, and other combinations of α, β, and ρ can be obtained by variable transforms or by stochastic simulation from the posterior distribution.

A SYNTHETIC EXAMPLE

The inversion method is tested on two synthetic earth profiles, wells A and B, shown in Figures 1 and 2. The two profiles

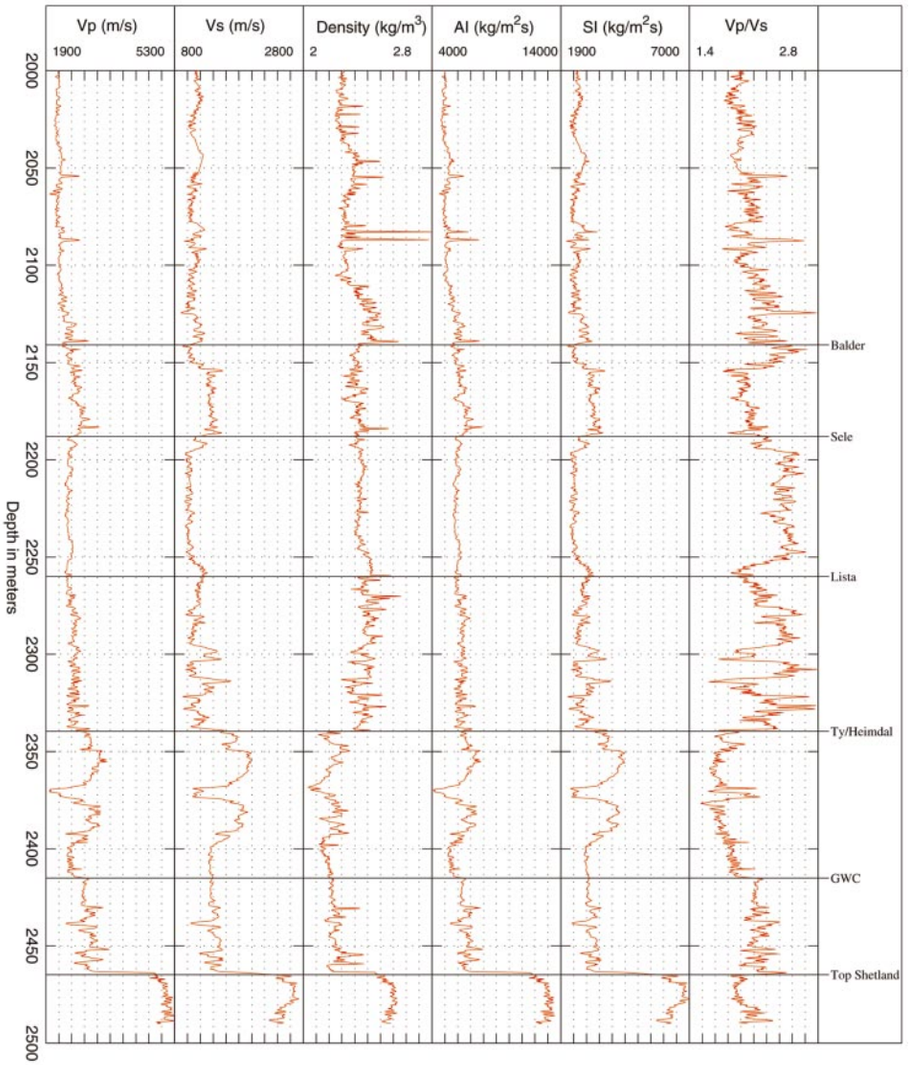

FIG. 10. The well logs plotted in the depth interval 2000–2500 m. The top reservoir is at 2340 m, the gas–water contact is at 2415 m, and the base reservoir is at 2465 m.

are simulated from a stationary prior distribution with constant expectation

$$\boldsymbol{\mu}(t) = \boldsymbol{\mu} = [\mu_\alpha, \mu_\beta, \mu_\rho]^T = [8.006, 7.313, 7.719]^T, \tag{35}$$

corresponding to the logarithm of 3000 m/s P-wave velocity, 1500 m/s S-wave velocity, and 2250 kg/m^3 density and a stationary covariance function with $\sigma_\alpha^2(t) = 0.0074, \sigma_\beta^2(t) = 0.0074$, and $\sigma_\rho^2(t) = 0.0024$. A second-order exponential correlation function [equation (11)] with range $d = 5$ ms is used for both wells. The two wells have different correlations between $\ln\alpha$, $\ln\beta$, and $\ln\rho$: The logs in well A are uncorrelated, while the logs in well B have a relatively strong correlation, $\nu_{\alpha\beta} = \nu_{\alpha\rho} = \nu_{\beta\rho} = 0.7$.

Modeling test

The weak contrast reflection coefficient approximation in equation (1) is adequate for moderate reflection angles. A further approximation is utilized in this inversion algorithm, where the coefficients $a_\alpha(t,\theta)$, $a_\beta(t,\theta)$, and $a_\rho(t,\theta)$ [equations (2)–(4)] are calculated from a constant or slowly varying prior known background model. The seismic forward modeling is performed by convolution directly in the time-angle domain. Real seismic data are recorded in the time-offset domain and are transformed to the time-angle domain prior to inversion.

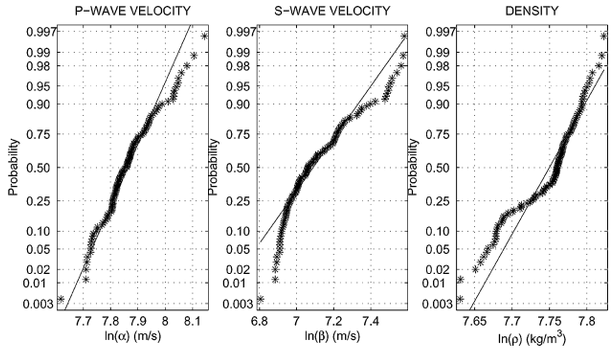

FIG. 11. Gaussian probability plot of the logarithm of the P-wave velocity (left), S-wave velocity (middle), and density (right).

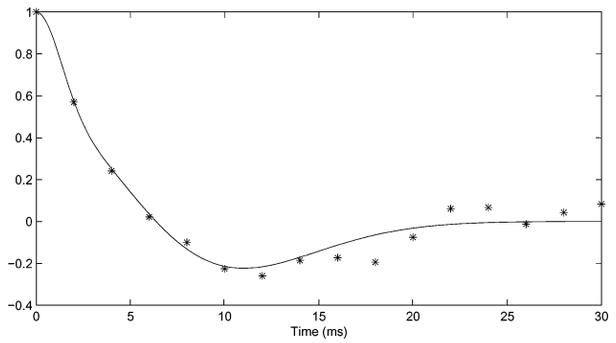

FIG. 12. Correlation function estimated from well logs (stars), and an analytical correlation function (black line) derived from two exponential second-order correlation functions.

Modeling tests are shown in Figures 3 and 4 for wells A and B, respectively. The wavelet is a Ricker wavelet with 25 Hz center frequency and normalized amplitude. The CDP gathers to the left are calculated using weak contrast modeling directly in the time-angle domain. The coefficients a_α, a_β, and a_ρ are calculated from the constant background model defined in equation (35). The CDP gathers in the middle of the figures are produced by ray tracing in the time-offset domain using the full Zoeppritz equation, and the gathers are then transformed to the time-angle domain. The gathers to the right show the differences between these two modeling methods. The relative squared data errors are only 0.0019 and 0.0027 for the two examples, which justifies the use of the weak-contrast reflectivity function and convolution directly in the time-angle domain. Compared to typical noise levels in real seismic data, the modeling errors in these examples are negligible.

Inversion test

The AVO inversion is tested on synthetic seismic data sets with different noise levels. A mixture of white and colored noise of the form

$$\mathbf{e} = \mathbf{e}_1 + \mathbf{S}\mathbf{e}_2 \tag{36}$$

was added to the synthetic data. The first term represents the white Gaussian noise with variance σ_1^2, $\mathbf{e}_1 \sim \mathcal{N}_{n_d}(\mathbf{0}, \sigma_1^2\mathbf{I})$. The error term $\mathbf{S}\mathbf{e}_2$ represents source-generated noise—for example, remaining multiples—and is assumed to be correlated between different angle traces by an exponential correlation function $\nu_\theta = \exp[-|\theta_i - \theta_j|/d_\theta]$, where $d_\theta = 20^\circ$ is the correlation range. For each angle θ_i, the elements in the error term $\mathbf{e}_2$ corresponding to θ_i are white Gaussian with variance σ_2^2. Noise gathers with four different levels were simulated with variances $\sigma_1^2 = \sigma_2^2 = 0.00005^2, 0.008^2, 0.015^2$, and 0.03^2, corresponding to S/N ratios 10^5, 15, 5, and 1, respectively. The CDP gathers for model A with noise are shown in Figure 5.

The inversion results of data set A with S/N ratio 10^5 are displayed in Figure 6, showing the MAP solutions with 0.95 prediction intervals for P- and S-wave velocities, density, acoustic impedance, shear impedance, and P- to S-wave velocity ratio. The solutions for the P- and S-wave velocities and the density are analytically obtainable, while stochastic simulation from the posterior distribution is used for the acoustic impedance, shear impedance, and P- to S-wave velocity ratio. The prior model applied is optimal in the sense that it is the

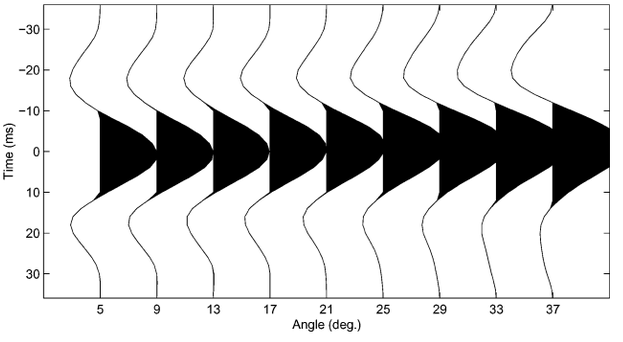

FIG. 13. Estimated wavelets at reservoir level.

same statistical model used to make the synthetic well logs. The 0.95 intervals of the prior model are shown in Figure 6 with dotted lines. Also, the correct noise covariance and the correct wavelet have been used. With this low noise level, all inverted parameters are retrieved almost perfectly with low uncertainty.

The inversion results of data set A with S/N ratio 5 are displayed in Figure 7. The confidence regions are much wider than in Figure 6, and the MAP solutions are smoother. Realizations from the posterior distribution for the data set with S/N ratio 5 are shown in Figure 8. These realizations have a higher but more realistic variability than the smooth MAP solution. Each of the realizations is a possible solution with this noise level. The large variations among these realizations explain the wide confidence regions.

A comparison of the 0.95 confidence regions for the prior and the posterior models makes a picture of the information content in the seismic prestack data. For model A, the seismic data provide mostly information about the acoustic impedance, the P- to S-wave velocity ratio, and the shear impedance for the case with S/N ratio 10^5, where the width of the confidence regions is reduced by 75% to 79% (see Table 1). In the case with an S/N ratio of 1, the seismic data provide mostly information about the acoustic impedance, the P-wave velocity and the P- to S-wave velocity ratio, but the intervals are only reduced by about 20%. In this case with strong noise, the seismic data provide practically no information about the S-wave velocity, the density, or the shear impedance (3–4% reduction of the interval width).

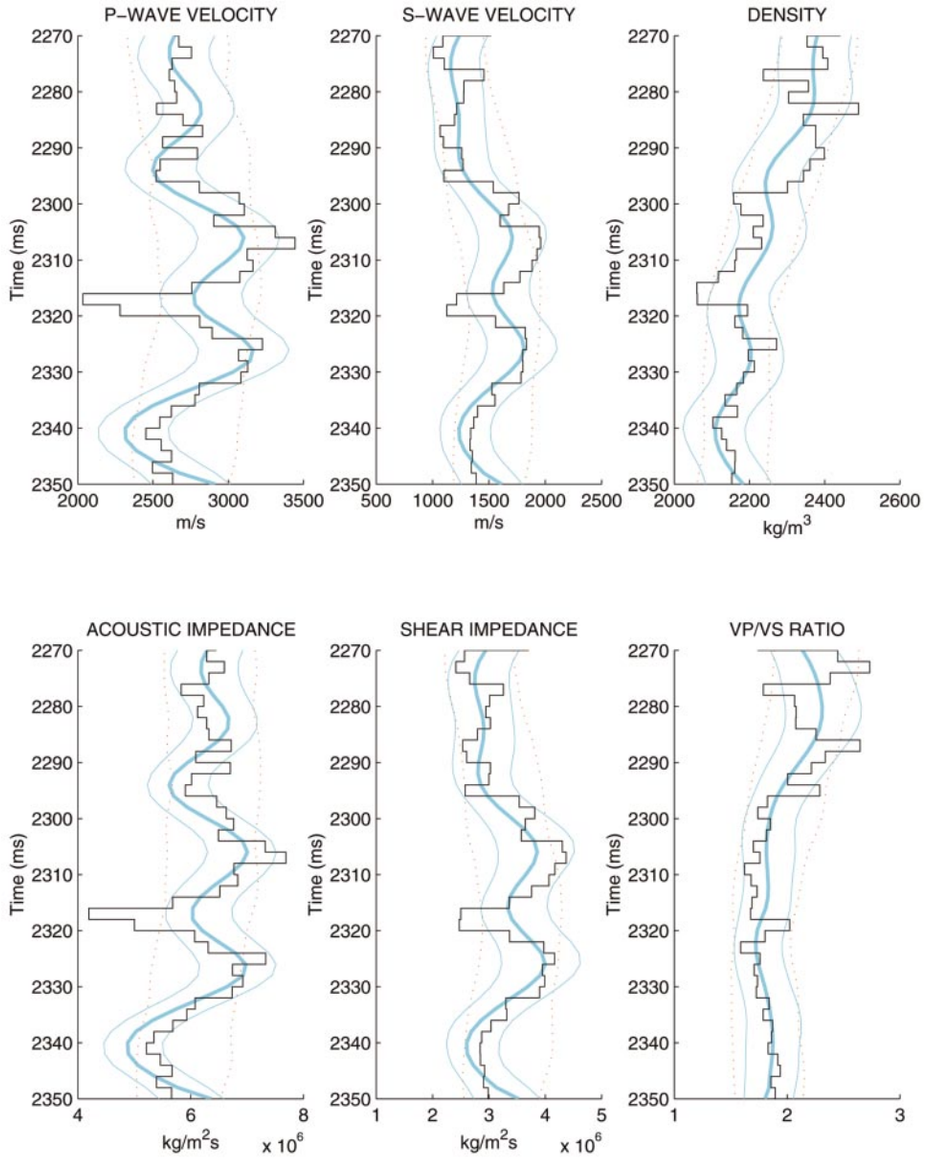

FIG. 14. The MAP solution (thick blue line) in the well position with 0.95 prediction interval (thin blue lines), the well log (black line), and 0.95 prior model interval (red dotted lines).

In well A, the P-wave velocity, S-wave velocity, and density are uncorrelated, but this is unrealistic. Well B is simulated with a strong correlation (0.7) between these three logs. The results in Table 2 shows that the confidence intervals for most parameters have been reduced more than in case A.

INVERSION OF SLEIPNER DATA

The Sleipner Øst field is located on the eastern side of the south Viking Graben in Norwegian block 15/9. The gas/condensate is trapped in a submarine-fan sandstone complex of the early Tertiary Ty Formation. The depth of the reservoir sand is in the range of 2270 to 2500 m subsea with the gas–water contact at 2415 m sub-sea. The reservoir underlies a thick shale package. The Ty Formation is composed of mainly clean sandstone with some thin shale and siltstone layers. The base reservoir is a sand/chalk interface (Top Shetland Formation).

A rectangular portion of the seismic survey, defined from in-lines 1411 to 1751 and from cross-lines 1225 to 1400, is used in this inversion. The inversion area covers 9.3 km^2, representing 12% of the total survey. The seismic data were processed by a contractor applying the processing sequence shown in Table 3. The processing sequence was defined such that the final prestack amplitudes should image the reflection strength of the subsurface interfaces as correctly as possible. The stack section of in-line 1627 is shown in Figure 9, where the strong, continuous reflector appearing between 2350 and 2450 ms is the Top Shetland horizon (base reservoir).

A well is located at in-line 1627 and cross-line 1291. The well logs are displayed as a function of depth in Figure 10, where the impedance logs and the P- to S-wave velocity ratio log are calculated from the P-wave velocity, S-wave velocity, and density logs.

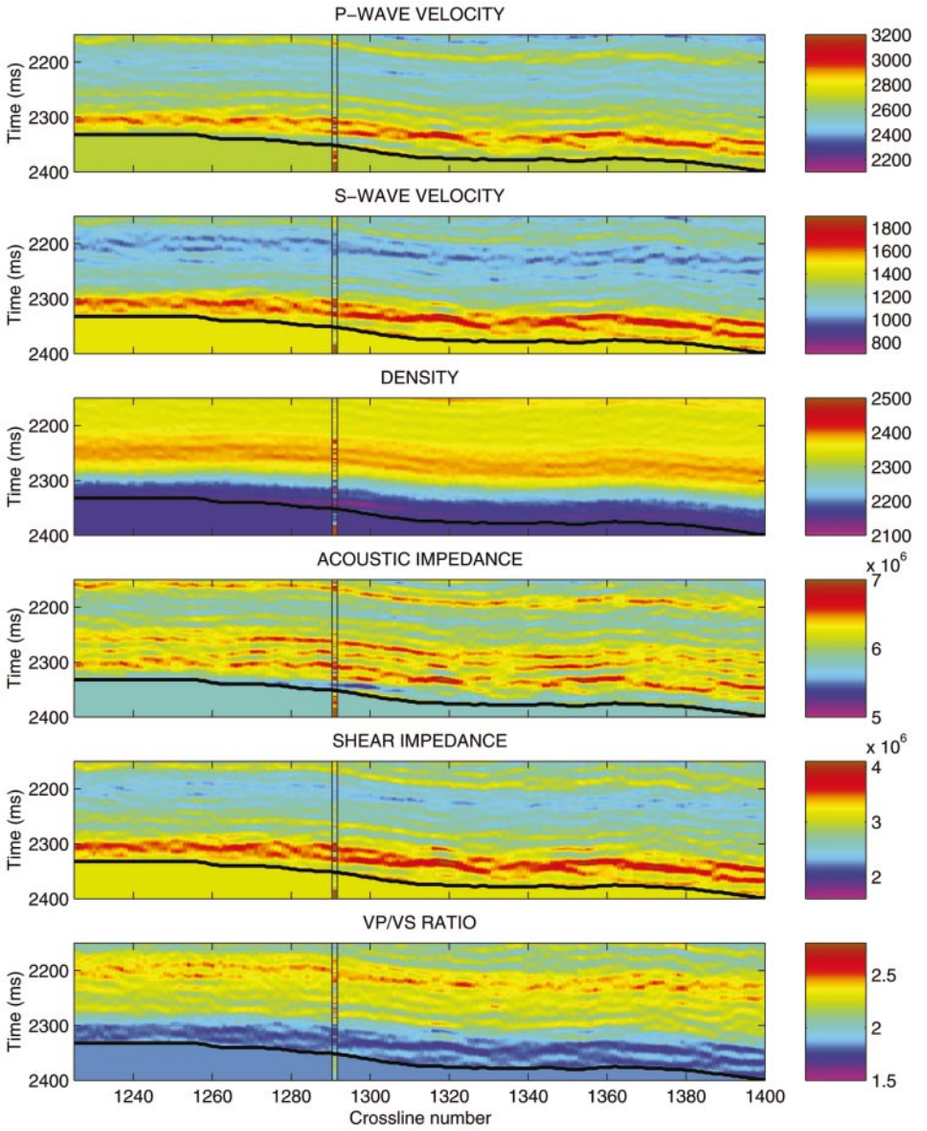

FIG. 15. MAP solution of in-line 1627.

The top reservoir is characterized by an increase in P-wave velocity (20%), a large increase in S-wave velocity (60%), and a small decrease in density (−8%). The acoustic impedance has only a minor increase from the shale to the gas sand (12%). The increase in the S-wave velocity may therefore be a better indicator for the top reservoir.

The base reservoir is well delineated on the seismic stack by the strong Top Shetland reflector. In the well, the P-wave velocity increases from about 3000 m/s in the sand reservoir to about 5000 m/s in the Top Shetland chalk formation, the S-wave velocity increases from about 1400 to 2600 m/s, and the density increases from 2200 to 2550 kg/m^3, which gives a reflection coefficient of about 0.3. Since the inversion algorithm is based on a weak contrast assumption, the inversion window is defined as a 250-ms window ending 32 ms above the Top Shetland (a half wavelet). At the well, the inversion window is from 2100 to 2350 ms.

Prior model estimation

The specification of the prior model is a controversal part of the Bayesian analysis. The extensive use of mathematically convenient prior distributions and the precision level of the prior distributions are debatable. Often, the available prior information is insufficient to define a unique parametric prior distribution. This problem can partly be handled by uninformative prior models or by including the uncertainty in the prior model by a hierarchical prior model. When some prior observations are available, a pragmatic approach is to select a parametric distribution and then estimate the neccessary parameters from the available prior observations. The most obvious pitfall with this approach is that a prior distribution with too low variance may be specified.

In the current study, the most important source of prior information is the available well logs. Ideally, the prior model should be estimated from several wells in the area to include lateral heterogeneity. Unfortunately, only one well is available in the portion of the survey selected for this inversion test.

The parameter vector $\mathbf{m}$ is assumed to be Gaussian. This assumption can be evaluated graphically in the well position by a Gaussian probability plot. If the parameters at this well are exactly Gaussian, the plot will be linear. Gaussian probability plots of $\ln\alpha$, $\ln\beta$, and $\ln\rho$ calculated from the well logs in the time interval 2100–2350 ms are shown in Figure 11. Despite some curvature in these plots, we consider the Gaussian a priori assumption to be acceptable. The computational cost of rejecting the Gaussian assumption is dramatic, since we lose the analytical form of the solution.

From the well logs, estimates of the elements in $\mathbf{\Sigma}_0$ are obtained by standard estimators. The estimated variances for $\ln\alpha$, $\ln\beta$, and $\ln\rho$ are $\widehat{\sigma_\alpha^2} = 0.0037$, $\widehat{\sigma_\beta^2} = 0.0116$, and $\widehat{\sigma_\rho^2} = 0.0004$. The estimated correlation coefficients are $\widehat{\nu_{\alpha\beta}} = 0.65$, $\widehat{\nu_{\alpha\rho}} = 0.11$, and $\widehat{\nu_{\beta\rho}} = -0.07$.

The temporal correlation functions are estimated for certain time lags from the well logs (see Figure 12). The temporal correlation function is modeled by an analytic correlation function,

$$\nu(\tau; d_1, d_2) = \frac{1}{2}\exp\left[-\left(\frac{\tau}{d_1}\right)^2\right] + \frac{1}{2}\left(1 - \frac{2\tau^2}{d_2^2}\right)\exp\left[-\left(\frac{\tau}{d_2}\right)^2\right], \quad (37)$$

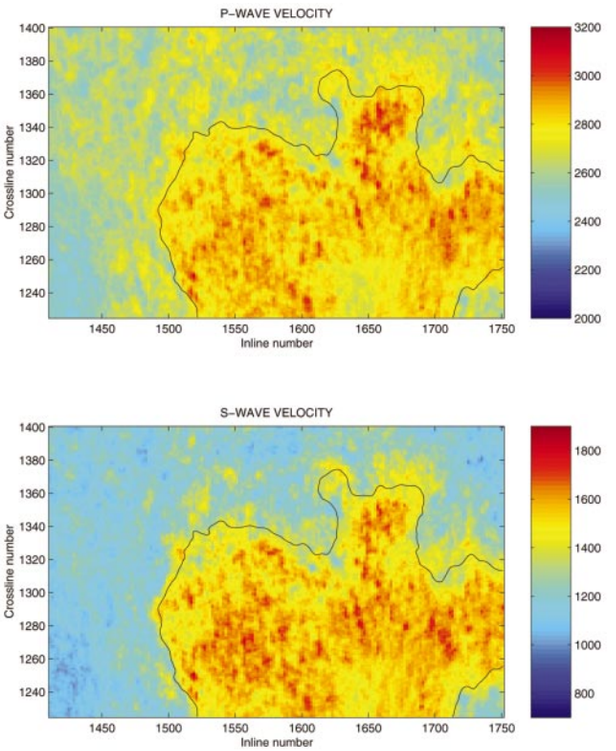

FIG. 16. Time slice of MAP solution. The interpreted boundary of the top reservoir is shown by the black line.

defined by the sum of an exponential second-order correlation function with range $d_1 = 1.8$ ms and a normalized second derivative of an exponential second-order correlation function with range $d_2 = 9$ ms. The fit to the estimated correlation function is considered to be good (see Figure 12).

Estimation of wavelets and noise covariance

The last step before inversion is to estimate wavelets and the noise covariance matrix from the well logs and the CDP gather at the well position. A Bayesian wavelet estimation method has been used, where seismic noise, possible mis-tie between the time axis of the well logs and the seismic traveltimes, and errors in the log measurements are included (see Buland and Omre, 2001). The possible mis-tie between the time axis of the seismic data and the well logs is handled by allowing for shift, stretch, and squeeze of the time axis of the well logs. The noise term is assumed to be a mixture of white and colored noise as in expression (36). The solution of the estimation problem is obtained by Markov chain Monte Carlo (MCMC) simulation, using 2000 iterations. The estimated wavelets are shown in Figure 13. The two variance terms, $\mathbf{e}_1$ and $\mathbf{e}_2$, are estimated to $\widehat{\sigma_1^2} = 0.0114$ and $\widehat{\sigma_2^2} = 0.0001$, and the estimated correlation range for $\mathbf{e}_2$ is $\widehat{d_\theta} = 10^\circ$.

Inversion results

The inversion results in the well position are displayed in Figure 14, showing the MAP solutions and 0.95 prediction intervals. With the estimated noise level in this data set, the prediction intervals are only marginally reduced compared to the prior model. The acoustic impedance and the P-wave velocity are the best determined parameters, but the prediction intervals are only reduced by 31% and 27%, respectively. The prediction intervals for the P- to S-wave velocity ratio, the shear impedance, and the S-wave velocity are reduced by about 15%, while the interval for the density is reduced by only 5%. Top reservoir is located at 2300 ms and is characterized by an increase in both P- and S-wave velocity. This increase in the two velocities is seen in the inversion results, but the thin, low-velocity layer at 2317 ms is not retrieved by the inversion.

The MAP solutions of in-line 1627 are shown in Figure 15. The bottom of the inversion window is shown by a thick black line 32 ms above the Top Shetland horizon. The well logs are plotted for comparison and show good agreement with the inversion results. Time slices (2320 ms) of the P- and S-wave velocities are shown in Figure 16. The boundary of the top reservoir, interpreted from stacked data, is shown by a black line.

The inversion algorithm is fast. The inversion of the 3-D test cube was finished in less than 3 minutes on a single 400-MHz Mips R12000 CPU. The algorithm is suitable for parallelization. On an 8-CPU parallel machine, the whole survey (77.5 km^2) could be inverted in approximately 3 minutes.

DISCUSSION AND CONCLUSIONS

We have developed a Bayesian AVO inversion method where the objective is to obtain posterior distributions for P-wave velocity, S-wave velocity, and density. The solution of the AVO inversion is given by a Gaussian posterior distribution. The explicit analytical form of the posterior distribution provides a fast inversion method. The uncertainty in the inverted parameters is an integral part of the solution.

Inversion tests on synthetic data with S/N $= 10^5$ show high agreement between the estimated and the correct model, but weak noise (S/N $= 15$) has a dramatic effect on the uncertainty of the predicted parameters. The inversion of the field data from Sleipner shows good agreement with well logs, but the uncertainty is high. Acoustic impedance is the best determined parameter, while the AVO inversion provides practically no information about the density. Generally, the resolution of the different parameters depends on the prior model and the noise covariance.

ACKNOWLEDGMENTS

We thank Statoil and the Sleipner licence (Statoil, Exxon/Mobil, Norsk Hydro, and TotalFinaElf) for permission to publish this paper. We also thank the reviewers for a thorough reading and useful comments.

REFERENCES

Aki, K., and Richards, P. G., 1980, Quantitative seismology: W. H. Freeman & Co.

Anderson, T. W., 1984, An introduction to multivariate statistical analysis: John Wiley & Sons, Inc.

Buland, A., and Omre, H., 2001, Bayesian wavelet estimation from seismic and well data: http://www.math.ntnu.no/preprint/statistic.s/2001.

Buland, A., Landrø, M., Andersen, M., and Dahl, T., 1996, AVO inversion of Troll field data: Geophysics, **61**, 1589–1602.

Christakos, G., 1992, Random field models in earth sciences: Academic Press Inc.

Dahl, T., and Ursin, B., 1991, Parameter estimation in a one-dimensional anelastic medium: J. Geophys. Res., **96**, 20217–20233.

Duijndam, A. J. W., 1988a, Bayesian estimation in seismic—inversion—part I: Principles: Geophys. Prosp., **36**, 878–898.

——— 1988b, Bayesian estimation in seismic inversio—art II: Uncertainty analysis: P Geophys. Prosp., **36**, 899–918.

Eide, A. L., Omre, H., and Ursin, B., 2002, Prediction of reservoir variables based on seismic and well observations: J. Am. Statistical Assn., **97**, no. 457, 18–28.

Gouveia, W. P., and Scales, J. A., 1997, Resolution of seismic waveform inversion: Bayes versus Occam: Inverse Problems, **13**, 323–349.

——— 1998, Bayesian seismic waveform inversion: Parameter estimation and uncertainty analysis: J. Geophys. Res., **103**, 2759–2779.

Lörtzer, G. J. M., and Berkhout, A. J., 1993, Linearized AVO inversion of multi-component seismic data, *in* Castagna, J., and Backus, M., Eds., Offset-dependent reflectivity—Theory and practice of AVO analysis: Sot. Expl. Geophys., 317–332.

Mosegaard, K., 1998, Resolution analysis of general inverse problems through inverse Monte Carlo sampling: Inverse Problems, **14**, ,405–426.

Mosegaard, K., and Tarantola, A., 1995, Monte Carlo sampling of solutions to inverse problems: J. Geophys. Res., **100**, 12431–12447.

Scales, J. A., and Tenorio, L., 2001, Prior information and uncertainty in inverse problems: Geophysics, **66**, 389–397.

Sen, M. K., and Stoffa, P. L., 1996, Bayesian inference, Gibbs' sampler and uncertainty estimation in geophysical inversion: Geophys. Prosp., **44**, 313–350.

Smith, G. C., and Gidlow, P. M., 1987, Weighted stacking for rock property estimation and detection of gas: Geophys. Prosp., **35**, 993–1014.

Stolt, R. H., and Weglein, A. B., 1985, Migration and inversion of seismic data: Geophysics, **50**, 2458–2472.

Tarantola, A., 1987, Inverse problem theory: Elsevier Science Publ. Co., Inc.

Tarantola, A., and Valette, B., 1982, Inverse problems = quest for information: J. Geophys., **50**, 159–170.

Ulrych, T. J., Sacchi, M. D., and Woodbury, A., 2001, A Bayes tour of inversion: A tutorial: Geophysics, **66**, 55–69.

APPENDIX A
GAUSSIAN DISTRIBUTION

A random field $r(t)$ is a spatially correlated function where the function value is random for any t. A Gaussian random field $r(t)$ is a random field where all multidimensional distribution for vectors $\mathbf{r} = [r(t_1), \ldots, r(t_n)]^T$ are Gaussian for any n and any configuration $\{t_1, \ldots, t_n\}$. The Gaussian random field is completely specified by the expectation function $\mu(t)$ and the covariance function $\Sigma(t, s) = \mathrm{Cov}\{r(t), r(s)\}$. The expectation can be arbitrary, but the covariance must be symmetric, $\Sigma(t, s) = \Sigma(s, t)$, and positive definite. The Gaussian probability density function for $\mathbf{r}$ is

$$p(\mathbf{r}) = \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{r}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{r}-\boldsymbol{\mu}) \right], \tag{A-1}$$

where n is the dimension of the vector $\mathbf{r}$ and where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are the expectation vector and the covariance matrix defined from $\mu(t)$ and $\Sigma(t, s)$, respectively. A compact notation is $\mathbf{r} \sim \mathcal{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. For further reading on multivariate statistical analysis, see e.g., Anderson (1984), where the two basic results below are presented.

Result 1

If the vector $\mathbf{r}$ is Gaussian, $\mathbf{r} \sim \mathcal{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, and $\mathbf{M}$ is an $m \times n$ matrix, then

$$\mathbf{Mr} \sim \mathcal{N}_m(\mathbf{M}\boldsymbol{\mu}, \mathbf{M}\boldsymbol{\Sigma}\mathbf{M}^T). \tag{A-2}$$

Result 2

Consider two multivariate Gaussian variables $\mathbf{r}_1 \sim \mathcal{N}_{n_1}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$ and $\mathbf{r}_2 \sim \mathcal{N}_{n_2}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_{22})$ with joint distribution

$$\begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{bmatrix} \sim \mathcal{N}_{n_1+n_2}\left(\begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix} \right). \tag{A-3}$$

Then the conditional distribution of $\mathbf{r}_1$ given $\mathbf{r}_2$ is Gaussian:

$$\mathbf{r}_1|\mathbf{r}_2 \sim \mathcal{N}_{n_1}(\boldsymbol{\mu}_{1|2}, \boldsymbol{\Sigma}_{1|2}), \tag{A-4}$$

where the conditional expectation and covariance are

$$\boldsymbol{\mu}_{1|2} = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{r}_2 - \boldsymbol{\mu}_2), \tag{A-5}$$

$$\boldsymbol{\Sigma}_{1|2} = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}. \tag{A-6}$$

APPENDIX B
THE FORWARD MODELING

A discrete version of the continuous reflectivity function

$$c_{PP}(t, \theta) = a_\alpha(t, \theta)\frac{\partial}{\partial t}\ln\alpha(t) + a_\beta(t, \theta)\frac{\partial}{\partial t}\ln\beta(t) + a_\rho(t, \theta)\frac{\partial}{\partial t}\ln\rho(t) \tag{B-1}$$

in a time interval and for a set of reflection angles can be written as

$$\mathbf{c} = \mathbf{A}\mathbf{m}'. \tag{B-2}$$

The sparse matrix $\mathbf{A}$ is defined by

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_\alpha(\theta_1) & \mathbf{A}_\beta(\theta_1) & \mathbf{A}_\rho(\theta_1) \\ \vdots & \vdots & \vdots \\ \mathbf{A}_\alpha(\theta_{n_\theta}) & \mathbf{A}_\beta(\theta_{n_\theta}) & \mathbf{A}_\rho(\theta_{n_\theta}) \end{bmatrix}, \tag{B-3}$$

where $\mathbf{A}_\alpha(\theta_i)$, $\mathbf{A}_\beta(\theta_i)$, and $\mathbf{A}_\rho(\theta_i)$ are $(n_m/3) \times (n_m/3)$ diagonal matrices containing discrete time samples of $a_\alpha(t, \theta_i)$, $a_\beta(t, \theta_i)$, and $a_\rho(t, \theta_i)$, respectively; n_θ is the number of reflection angles; and n_m is the dimension of $\mathbf{m}$ and $\mathbf{m}'$.

The convolution of the reflection coefficients $\mathbf{c}$ with the wavelets can be formulated as a matrix–vector multiplication

$$\mathbf{d}_{obs} = \mathbf{S}\mathbf{c} + \mathbf{e}, \tag{B-4}$$

where $\mathbf{S}$ is a block-diagonal matrix containing one wavelet for each reflection angle. In an expanded form, expression (B-4) can be written

$$\begin{bmatrix} \mathbf{d}_{obs}(\theta_1) \\ \vdots \\ \mathbf{d}_{obs}(\theta_{n_\theta}) \end{bmatrix} = \begin{bmatrix} \mathbf{S}(\theta_1) & & \\ & \ddots & \\ & & \mathbf{S}(\theta_{n_\theta}) \end{bmatrix} \times \begin{bmatrix} \mathbf{c}(\theta_1) \\ \vdots \\ \mathbf{c}(\theta_{n_\theta}) \end{bmatrix} + \begin{bmatrix} \mathbf{e}(\theta_1) \\ \vdots \\ \mathbf{e}(\theta_{n_\theta}) \end{bmatrix}, \tag{B-5}$$

where $\mathbf{d}_{obs}(\theta_i)$ is the seismic time trace for angle θ_i and where $\mathbf{c}(\theta_i)$ and $\mathbf{e}(\theta_i)$ are the corresponding reflection coefficients and error samples, respectively. The block matrix $\mathbf{S}(\theta_i)$ represents the wavelet for angle θ_i

$$\mathbf{S}(\theta_i) = \begin{bmatrix} s_1(\theta_i) & & & & & & \\ s_2(\theta_i) & s_1(\theta_i) & & & & & \\ \vdots & & \ddots & & & & \\ s_{n_s}(\theta_i) & \cdots & \cdots & s_1(\theta_i) & & & \\ & \ddots & & & \ddots & & \\ & & s_{n_s}(\theta_i) & \cdots & \cdots & s_1(\theta_i) \\ & & & \ddots & & \vdots \\ & & & & \ddots & \vdots \\ & & & & & s_{n_s}(\theta_i) \end{bmatrix}, \tag{B-6}$$

where $(s_1(\theta_i), \ldots, s_{n_s}(\theta_i))$ are the samples of the wavelet for angle θ_i. In this example, the sampling of the wavelet is equal to the sampling of the seismic data. If the sampling of $\mathbf{c}$ and $\mathbf{d}_{obs}$ are different, the rows of $\mathbf{S}$ contains wavelets corresponding to the sampling of $\mathbf{c}$, and the rows are shifted relatively according to the sampling of $\mathbf{d}_{obs}$.