Stochastic inversion of seismic PP and PS data for reservoir parameter estimation

Shortened Title: **Stochastic inversion of PP and PS data**

Jinsong Chen (jchen@lbl.gov)

Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

Michael E. Glinsky (Michael.glinsky@iongeo.com),

ION Geophysical

The initial paper was presented on MS 1: Vp/Vs, Tomography, and AVO/AZ Inversion Session of SEG Annual Meeting 2013

(SEG Expanded Abstracts 32, 305, 2013)

Submission to Geophysics

ABSTRACT

We investigate the value of isotropic seismic converted-wave (i.e., PS) data for reservoir parameter estimation using stochastic approaches based on the floating-grain rock-physics model. We first perform statistical analysis on a simple two-layer model built on actual borehole logs and compare the relative value of PS data versus AVO gradient data for estimating floating-grain fraction. We find that PS data are significantly more informative than AVO gradient data for estimating floating-grain fraction in terms of likelihood functions, and the combination of PS and AVO gradient data together with full PP stack data provides the maximal value for the reservoir parameter estimation. To evaluate the value of PS data under complex situations, we develop a hierarchical Bayesian model to combine seismic PP and PS data and their associated time registration. We extend a model-based Bayesian method developed previously for inverting seismic PP data by including PS responses and time registration as additional data and two-way PS travel time and PS reflectivity as additional variables. We apply the method to a synthetic six-layer model that closely mimics real field scenarios. The case study results show that PS data provide more information than AVO gradient data for estimating floating-grain fraction, porosity, net-to-gross, and layer thicknesses when their corresponding priors are weak.

INTRODUCTION

Multicomponent seismic surveying has been used for hydrocarbon exploration for decades because it can capture the seismic wave-field more completely than conventional single-element techniques (Stewart et al., 2002). Although several types of energy conversion may occur when seismic waves pass through the underlying earth, transmitted or multiple conversions generally have much lower amplitudes than the P-down and S-up reflection (Rodriguez-Saurez, 2000). Consequently, among many applications of multi-component seismic data, the use of converted-wave or PS images receives much more attention (Stewart et al., 2002; Mahmoudian and Margrave, 2004; Veire and Landro, 2006). However, the high acquisition cost of collecting multicomponent seismic data compared to conventional seismic surveys and the challenge in processing multicomponent data, makes the use of converted-wave data as a routine practice difficult.

The interest in using multicomponent seismic data again for hydrocarbon applications is inspired by recent advances in seismic data acquisition technologies, such as ocean-bottom seismometer techniques (e.g., ocean-based cables and ocean-based nodes) (Hardage et al., 2011; Pacal, 2012). With the use of new techniques, multicomponent seismic data can be collected more reliably compared to conventional seismic survey techniques. Another major reason for using multicomponent seismic data is the need to estimate spatially-distributed ductile fraction (Glinsky et al., 2013), and to characterize fractures for unconventional resources because shear-wave splitting provides an effective approach to image fracture orientation and density (Bale et al., 2013). There are many other successful applications of converted-wave data, such as time-lapse monitoring of geomechanical changes (Davis et al., 2013), and reservoir characterization (Brettwood et al., 2013).

In this study, we use stochastic approaches to investigate the value of converted-wave data for reservoir parameter estimation based on a floating-grain rockphysics model developed by DeMartini and Glinsky (2006). The model is well-documented (Gunning and Glinsky, 2007) and appropriate for porous sedimentary rocks in which some solid materials are “floating” or not involved in loading support because it can explain the observed variation in P-wave velocity versus density trends, and lack of variation in the P-wave velocity versus S-wave velocity trends. The rockphysics relationship can be modified and applied to unconventional shale resource exploration as done by Glinsky et al. (2013), where the media is considered as a binary mixing of brittle and ductile materials and ductile fraction plays the same role as floating-grain fraction.

We employ stochastic methods in the study because they have many advantages over traditional deterministic approaches in reservoir parameter estimation using multiple geophysical data sets when dealing with complex issues involving uncertainty (Chen et al., 2008). We start from analyzing a simple two-layer model by comparing the relative value of full PS versus AVO gradient data for estimating floating-grain fraction according to their likelihoods when both rockphysics models and seismic data are subject to uncertainty. We then focus on more complicated cases involving multiple layers and develop a hierarchical Bayesian model to combine seismic PP and PS data and their associated time registration.

We extend the model-based Bayesian method developed by Gunning and Glinsky (2004) for inverting seismic AVO data, by revising their open-source Java codes (i.e., 'Delivery') to allow isotropic converted-wave responses and PS event time registration as additional data. We use the same rock physics models and Markov chain Monte Carlo (MCMC) (Gilks et al., 1996) sampling strategies as Delivery. Since this study is built on the previous work, the subsequent descriptions will be focused on the new development and applications, and the details of other parts can be found in Gunning and Glinsky (2004).

ROCKPHYSICS MODEL AND ANALYSIS OF TWO-LAYER MODELS

Floating-grain rockphysics model

We use the floating-grain rockphysics models developed by Demartini and Glinsky (2006) and Gunning and Glinsky (2007) to link reservoir parameters to seismic attributes. In the model, the subsurface is considered as a binary mixture of reservoir members (say sand) and non-reservoir members (say shale). For sand, we assume that some solid materials are "floating" in pore space and seismic properties (i.e., seismic P- and S-wave velocity and density) can be characterized by two fundamental parameters. One of them is loading depth (z), which is measure of effective pressure; the other is floating-grain fraction (x). The general model is given below

$$v_p = a_{vp} + b_{vp} z + c_{vp} x + \varepsilon_{vp}, \tag{1}$$

$$v_s = a_{vs} + b_{vs} v_p + \varepsilon_{vs}, \tag{2}$$

$$\rho = a_\rho + b_\rho v_p + c_\rho x + \varepsilon_\rho. \tag{3}$$

In equations 1-3, symbols v_p, v_s, and ρ represent seismic P- and S-wave velocity and density, and symbols ε_{vp}, ε_{vs}, and ε_ρ represent uncertainty associated with their corresponding regression equations. We assume that ε_{vp}, ε_{vs}, and ε_{ρ} have Gaussian distributions with zero mean and variance of σ_{vp}^2, σ_{vs}^2, and σ_{ρ}^2, respectively.

We rewrite equations 2-3 in terms of loading depth z and floating-grain fraction x as follows

$$v_s = (a_{vs} + a_{vp}b_{vs}) + b_{vs}b_{vp}z + b_{vs}c_{vp}x + (b_{vs}\varepsilon_{vp} + \varepsilon_{vs}), \tag{4}$$

$$\rho = (a_{\rho} + a_{vp}b_{\rho}) + b_{\rho}b_{vp}z + (b_{\rho}c_{vp} + c_{\rho})x + (b_{\rho}\varepsilon_{vp} + \varepsilon_{\rho}). \tag{5}$$

We can see that in the rockphysics model, seismic properties linearly depend on the reservoir parameters with uncertainty.

We can use different relationships for shale because seismic properties in shale do not depend on floating-fraction. As it has been done in Gunning and Glinsky (2007), we drop floating-grain fraction from equations 1 and 4 and use the power-law form of the Gardner relationship (Gardner et al., 1974) for density, i.e., $\rho = av_p^b + \varepsilon_{\rho}$, where both a and b are coefficients. By fitting actual borehole logs from suitable field sites, we obtain all those coefficients and their associated standard errors for sand and shale members. Table 1 is a summary of all those values.

Reflectivity coefficients

We use the linearized Zoeppritz approximations (Aki and Richard, 1980) for small contrasts and angles to obtain PP and PS reflectivity coefficients at an interface, which are given below:

$$R_{pp}(\theta) = \frac{1}{2}\left(\frac{\Delta v_p}{v_p} + \frac{\Delta\rho}{\rho}\right) + \frac{1}{2}\left[\frac{\Delta v_p}{v_p} - 4r_{sp}^2\left(\frac{\Delta\rho}{\rho} + 2\frac{\Delta v_s}{v_s}\right)\right]\theta^2, \tag{6}$$

$$R_{ps}(\theta) = -\frac{1}{2}\left[\frac{\Delta\rho}{\rho} + 2r_{sp}\left(\frac{\Delta\rho}{\rho} + 2\frac{\Delta v_s}{v_s}\right)\right]\theta. \tag{7}$$

In equations 6 and 7, $v_p = (v_{p1} + v_{p2})/2$, $v_s = (v_{s1} + v_{s2})/2$, $\rho = (\rho_1 + \rho_2)/2$, $r_{sp} = v_s / v_p$, $\Delta v_p = v_{p2} - v_{p1}$, $\Delta v_s = v_{s2} - v_{s1}$, and $\Delta\rho = \rho_2 - \rho_1$, where ($v_{p1}$, v_{s1}, ρ_1) and (v_{p2}, v_{s2}, ρ_2) are P- and S-wave velocity and density in the layers above and below the interface. Symbol θ is the PP incident angle in the unit of radius. The PP and PS reflectivities have the fourth and third order of accuracy in terms of the incident angle.

For ease of description, we let A_0 be the first term in equation 6, A_1 be the first term in equation 7 excluding θ, and A_2 be the second term in equation 6 excluding θ^2. We thus have the following relationship:

$$\begin{pmatrix} A_0 \\ A_1 \\ A_2 \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 & 0 \\ -1/2 - r_{sp} & 0 & -2r_{sp} \\ -2r_{sp}^2 & 1/2 & -4r_{sp}^2 \end{pmatrix} \begin{pmatrix} \Delta\rho/\rho \\ \Delta v_p / v_p \\ \Delta v_s / v_s \end{pmatrix}. \tag{8}$$

We use the capital letter $\mathbf{A}$ to represent the vector on the left side of equation 8 and use $\mathbf{M}_a$ and $\Delta\mathbf{C}$ represent the matrix and the vector on the right side of the equation. Thus equation 8 becomes $\mathbf{A} = \mathbf{M}_a \Delta\mathbf{C}$. These notations will be used in the subsequent text.

Synthetic two-layer model

To demonstrate the value of PS data, we start from a simple two-layer model based on actual borehole logs from Gunning and Glinsky (2007), with the first layer being shale and the second layer being sand whose rockphysics models are given in Table 1. Since we focus on estimation of floating-grain fraction in the sand layer, we fix the loading depth as $z_1 = 17{,}060$ ft and $z_2 = 17{,}457$ ft for the first and second layers. By using the shale regression equations with coefficients given in Table 1, we have $v_{p1} = 10,756$ ft/s, $v_{s1} = 5,245$ ft/s, and $\rho_1 = 2.49$ g/cc. By using equations 1, 4, and 5, we can get

$$\Delta\rho = (a_\rho + a_{vp}b_\rho + b_{vp}b_\rho z_2 - \rho_1) + (b_\rho c_{vp} + c_\rho)x + (\varepsilon_\rho + b_\rho \varepsilon_{vp}), \tag{9}$$

$$\Delta v_p = (a_{vp} + b_{vp}z_2 - v_{p1}) + c_{vp}x + \varepsilon_{vp}, \tag{10}$$

$$\Delta v_s = (a_{vs} + a_{vp}b_{vs} + b_{vp}b_{vs}z_2 - v_{s1}) + b_{vs}c_{vp}x + (\varepsilon_{vs} + b_{vs}\varepsilon_{vp}). \tag{11}$$

Let

$$\begin{cases} w_\rho = a_\rho + a_{vp}b_\rho + b_{vp}b_\rho z_2 - \rho_1 \\ w_{vp} = a_{vp} + b_{vp}z_2 - v_{p1} \\ w_{vs} = a_{vs} + a_{vp}b_{vs} + b_{vp}b_{vs}z_2 - v_{s1} \end{cases} \tag{12}$$

We have

$$\begin{pmatrix} \Delta\rho/\rho \\ \Delta v_p / v_p \\ \Delta v_s / v_s \end{pmatrix} = \begin{pmatrix} w_\rho/\rho \\ w_{vp}/v_p \\ w_{vs}/v_s \end{pmatrix} + \begin{pmatrix} (c_\rho + b_\rho c_{vp})/\rho \\ c_{vp}/v_p \\ b_{vs}c_{vp}/v_s \end{pmatrix} x + \begin{pmatrix} (\varepsilon_\rho + b_\rho\varepsilon_{vp})/\rho \\ \varepsilon_{vp}/v_p \\ (\varepsilon_{vs} + b_{vs}\varepsilon_{vp})/v_{vs} \end{pmatrix}. \tag{13}$$

Let $\mathbf{W}_0$, $\mathbf{W}_1$, and $\boldsymbol{\varepsilon}_w$ represent the first, second, and third vectors on the right side of equation 13. We get $\Delta\mathbf{C} = \mathbf{W}_0 + \mathbf{W}_1 x + \boldsymbol{\varepsilon}_w$. By assuming that the errors in equations 1-3 are independent, we can obtain the following covariance matrix Σ_w

$$\Sigma_w = \begin{pmatrix} (\sigma_\rho^2 + b_\rho^2\sigma_{vp}^2)/\rho^2 & b_\rho\sigma_{vp}^2/(\rho v_p) & b_\rho b_{vs}\sigma_{vp}^2/(\rho v_s) \\ b_\rho\sigma_{vp}^2/(\rho v_p) & \sigma_{vp}^2/v_p^2 & b_{vs}\sigma_{vp}^2/(v_p v_s) \\ b_\rho b_{vs}\sigma_{vp}^2/(\rho v_s) & b_{vs}\sigma_{vp}^2/(v_p v_s) & (\sigma_{vs}^2 + b_{vs}^2\sigma_{vp}^2)/v_{vs}^2 \end{pmatrix}. \tag{14}$$

Synthetic seismic data and likelihood function

For the purpose of this analysis, we consider the PP and PS reflectivities at the interface as data. Specifically, we use a full PP stack with an incident angle of zero and a full PS stack with an incident angle of θ, and an AVO gradient stack with an incident angle of θ. Let vector $\mathbf{R}_m$ be the data with additive Gaussian random noise $\boldsymbol{\varepsilon}_m$. We thus have (see Appendix A for the detailed derivation)

$$\mathbf{R}_m = \mathbf{M}_\theta \mathbf{M}_a \Delta\mathbf{C} + \boldsymbol{\varepsilon}_m = \mathbf{M}_\theta \mathbf{M}_a (\mathbf{W}_0 + \mathbf{W}_1 x) + (\boldsymbol{\varepsilon}_m + \mathbf{M}_\theta \mathbf{M}_a \boldsymbol{\varepsilon}_w), \tag{15}$$

where the angle dependent matrix $\mathbf{M}_\theta$ is defined as follows:

$$\mathbf{M}_\theta = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \theta & 0 \\ 0 & 0 & \theta^2 \end{bmatrix}. \tag{16}$$

The second term on the right side of equation 15 is residuals, which include two parts, i.e., the measurement errors and the uncertainty caused by the rockphysics model. Since we assume both parts have multivariate Gaussian distribution, their summation also has a multivariate Gaussian distribution (Stone, 1995). Let $\boldsymbol{\Sigma}_m$ be the covariance matrix of measurement errors and $\boldsymbol{\Sigma}_w$ be the covariance matrix of uncertainty in rockphysics models. The combined covariance thus is given by $\Sigma_c = \mathbf{M}_\theta (\Sigma_m + \mathbf{M}_a \Sigma_w \mathbf{M}_a^T) \mathbf{M}_\theta^T$, where Σ_w is given by equation 14. Consequently, the likelihood function of x given data $\mathbf{R}_m$ is a multivariate Gaussian distribution as follows:

$$\begin{aligned} f(\mathbf{R}_m \mid x) &\propto \left|\Sigma_c\right|^{-1/2} \\ &\times \exp\left\{-(\mathbf{R}_m - \mathbf{M}_\theta \mathbf{M}_a \mathbf{W}_0 - \mathbf{M}_\theta \mathbf{M}_a \mathbf{W}_1 x)^T \Sigma_c^{-1} (\mathbf{R}_m - \mathbf{M}_\theta \mathbf{M}_a \mathbf{W}_0 - \mathbf{M}_\theta \mathbf{M}_a \mathbf{W}_1 x)\right\}. \end{aligned} \tag{17}$$

Model comparison

We compare the estimation results by using four different combinations of seismic data by specifying their angle-dependent matrices: (1) using only the full PP stack, (2) using the full PP and PS stacks, (3) using the full PP and AVO gradient stacks, and (4) using all the seismic data (see Appendix B to find angle-dependent matrix for each case).

Their corresponding data can be represented by $\mathbf{R}_m^{(1)}$, $\mathbf{R}_m^{(2)}$, $\mathbf{R}_m^{(3)}$, and $\mathbf{R}_m^{(4)}$. To avoid the effects of prior distribution on floating-grain fraction, we focus on the likelihood functions $f(\mathbf{R}_m^{(k)} \mid x)$ ($k = 1,2,3,4$) for those models.

Figure 1 compares the likelihood functions for the true floating-grain fraction being 0.0 (Figure 1a) and 0.035 (Figure 1b). The noise levels for all the data are equal to 0.01 in the unit of reflection coefficients (RFC). The black, green, red, and blue curves represent the likelihoods obtained using (1) the full PP stack only, (2) the full PP and AVO gradient stacks, (3) the full PP and PS stacks, and (4) all the seismic data. As we expect, the true values have the maximum likelihood in both cases. It is clear that the likelihoods of using the full PP and AVO gradient data (green curves) are considerably larger than those of using the full PP stack only (black curves). The likelihoods of using the full PP and full PS stacks (red curves) are significantly larger than those of using the full PP and AVO gradient stacks (green curves). This suggests that the combination of the full PP and PS stacks are more informative for estimating floating-grain fraction than that of full PP and AVO gradient stacks. When we use all the data, we can get the largest likelihoods (blue curves). This implies that the full PS and AVO gradient stacks may complement to each other to some degree. In addition, we can see that the clean sand (Figure 1a) overall has larger likelihoods with the maximum of 24 than the sand with floating grain (Figure 1b) with the maximum of 14.

The above comparison depends on the noise levels in the seismic data. In practice, full PS and AVO gradient stacks are typically have larger errors than full PP stacks. To investigate the effects of noise on the likelihood analysis, we vary noise levels in both full PS and AVO gradient stacks from 0.01 RFC to 0.1 RFC while fixing the noise level of the full PP stack as 0.01 RFC. We first calculate the maximum likelihoods for each combination of seismic data and then normalize the results by the values of using the full PP stack only to get the following likelihood ratios

$$r_k = \frac{\max\{f(\mathbf{R}_m^{(k)} \mid x)\}}{\max\{f(\mathbf{R}_m^{(1)} \mid x)\}}. \tag{18}$$

Figure 2 shows the likelihood ratios for the true floating-grain fraction of 0.0 and 0.035. The green, red, and blue curves are the likelihood ratios of using the full PP and AVO gradient stacks, the full PP and PS stacks, and all the seismic data. Generally, as the noise levels in the full PS and AVO gradient data increase, the likelihood ratios decrease and approach to 1, the result of using the full PP stack only. Additionally, we can see the likelihood ratios of using the full PP and PS stacks always have larger likelihood ratios than those of using full PP and gradient stacks and the likelihood ratios of using all the data always have the largest values. This means that the combination of full PP and full PS stacks is more informative than that of full PP and AVO gradient stacks even under larger noise levels.

BAYESIAN MODEL FOR MULTIPLE LAYERS

Hierarchical Bayesian models

Although analysis of two-layer models allows us to understand the value of PS data for floating-grain fraction estimation, it is just marginal analysis of relative changes of compaction and floating-grain fraction across an interface under simple conditions. In the case of multiple layers, we need to develop a hierarchical Bayesian model to combine seismic PP and PS data and their time registration. This model is an extension of the model-based Bayesian method by Gunning and Glinsky (2004) with converted wave responses and PS time registration as additional data and two-way PS travel time and PS reflectivity as additional unknowns.

We consider effective seismic P-wave and S-wave velocity ($\mathbf{v}_p$ and $\mathbf{v}_s$) and density ($\boldsymbol{\rho}$), and seismic PP and PS reflectivity ($\mathbf{R}_{pp}$ and $\mathbf{R}_{ps}$) as unknowns. They are functions of rock physics parameters through suitable rock physics models. We consider PP traveltime ($\mathbf{t}_{pp}$) as a primary unknown, and both layer thickness ($\mathbf{d}$) and PS traveltime ($\mathbf{t}_{ps}$) can be derived from the PP traveltime and associated effective seismic attributes. The data used for inversion include seismic PP and PS full-waveforms ($\mathbf{S}_{pp}$, and $\mathbf{S}_{ps}$) and PP and PS event registration time ($\mathbf{T}_{pp}$ and $\mathbf{T}_{ps}$). If available, we can also include other types of information from nearby boreholes, such as depth constraints ($\mathbf{D}_b$).

Figure 3 shows all the unknowns, available data, and their relationships; the dashed rectangle highlights our extension to Delivery. Specifically, we add two unknowns related to converted wave (i.e., $\mathbf{t}_{ps}$, and $\mathbf{R}_{ps}$) and two types of new data sets (i.e., $\mathbf{T}_{ps}$ and $\mathbf{S}_{ps}$). Following the direct graphical model, we have the following hierarchical Bayesian model:

$$\begin{aligned}
&f(\boldsymbol{\alpha},\mathbf{t}_{pp},\mathbf{t}_{ps},\mathbf{d},\mathbf{v}_p,\mathbf{v}_s,\boldsymbol{\rho},\mathbf{R}_{pp},\mathbf{R}_{ps} \mid \mathbf{S}_{pp},\mathbf{S}_{ps},\mathbf{T}_{pp},\mathbf{T}_{ps},\mathbf{D}_b)\\
&\propto f(\mathbf{S}_{pp}\mid\mathbf{t}_{pp},\mathbf{R}_{pp})f(\mathbf{S}_{ps}\mid\mathbf{t}_{ps},\mathbf{R}_{ps})f(\mathbf{T}_{pp}\mid\mathbf{t}_{pp})\\
&\times f(\mathbf{T}_{ps}\mid\mathbf{t}_{ps})f(\mathbf{D}_b\mid\mathbf{d})f(\mathbf{R}_{pp}\mid\mathbf{v}_p,\mathbf{v}_s,\boldsymbol{\rho})\\
&\times f(\mathbf{R}_{ps}\mid\mathbf{v}_p,\mathbf{v}_s,\boldsymbol{\rho})f(\mathbf{d}\mid\mathbf{t}_{pp},\mathbf{v}_p)f(\mathbf{t}_{ps}\mid\mathbf{t}_{pp},\mathbf{v}_p,\mathbf{v}_s)\\
&\times f(\mathbf{v}_p,\mathbf{v}_s,\boldsymbol{\rho}\mid\boldsymbol{\alpha})f(\boldsymbol{\alpha})f(\mathbf{t}_{pp}).
\end{aligned} \tag{19}$$

Equation 19 defines a joint posterior probability distribution function of all unknown parameters up to a normalizing constant. The first five terms on the right side of the equation are the likelihood functions of available data, which link data to the associated unknowns; other terms on the right side are the prior probability distributions, which are derived from other sources of information, such as rockphysics models. We define all the likelihood functions and prior distributions in a similar way to Delivery (see Gunning and Glinsky, 2004). In the following, we only describe the new development.

Equation 19 is a general Bayesian model for combining seismic PP and converted-wave data, and we can simplify or vary the equation in different ways depending on specific applications. For example, we can consider PP and PS reflectivities as functions of $\mathbf{v}_p$, $\mathbf{v}_s$, and $\boldsymbol{\rho}$, but ignore their associated uncertainties and consider depth $\mathbf{d}$ as a function of P-wave velocity and two-layer PP travel time. Since in Bayesian statistics (Bernardo and Smith, 2000), data affect unknowns only through likelihood functions, we can use statistics $\mathbf{Q}(\mathbf{S}_{pp},\mathbf{S}_{ps})$ of seismic data $\mathbf{S}_{pp}$ and $\mathbf{S}_{ps}$ in the Bayesian model, for example, the rotation and truncation of original seismic data through principal component analysis (Venables and Ripley, 1999) or other methods. Consequently, we can have the following Bayesian model:

$$\begin{aligned}&f(\boldsymbol{\alpha},\mathbf{t}_{pp},\mathbf{t}_{ps},\mathbf{v}_p,\mathbf{v}_s,\boldsymbol{\rho}\,|\,\mathbf{S}_{pp},\mathbf{S}_{ps},\mathbf{T}_{pp},\mathbf{T}_{ps})\\ &\propto f(\mathbf{Q}\,|\,\mathbf{t}_{pp},\mathbf{t}_{ps},\mathbf{v}_p,\mathbf{v}_s,\boldsymbol{\rho})f(\mathbf{T}_{pp}\,|\,\mathbf{t}_{pp})f(\mathbf{T}_{ps}\,|\,\mathbf{t}_{ps})\\ &\times f(\mathbf{t}_{ps}\,|\,\mathbf{t}_{pp},\mathbf{v}_p,\mathbf{v}_s)f(\mathbf{v}_p,\mathbf{v}_s,\boldsymbol{\rho}\,|\,\boldsymbol{\alpha})f(\boldsymbol{\alpha})f(\mathbf{t}_{pp}).\end{aligned} \tag{20}$$

Likelihood function of seismic data

We describe a general form of the likelihood function in terms of statistics of seismic data, with the likelihood function of original seismic data as a special case of the form. Let $\mathbf{G}(\mathbf{t}_{pp},\mathbf{t}_{ps},\mathbf{v}_p,\mathbf{v}_s,\boldsymbol{\rho})$ be the response vector of a suitable forward model that links seismic statistics $\mathbf{Q}$ to unknowns. Let vector $\boldsymbol{\varepsilon}_m$ represent the residuals. We assume that the residuals have the multivariate Gaussian distribution with zero mean and the covariance matrix of Σ_m. We thus have

$$f(\mathbf{Q} \mid \mathbf{t}_{pp}, \mathbf{t}_{ps}, \mathbf{v}_p, \mathbf{v}_s, \boldsymbol{\rho}) = (2\pi)^{-k/2} \left|\Sigma_m\right|^{1/2}$$
$$\times \exp\left(-\frac{1}{2}\left(\mathbf{Q} - \mathbf{G}(\mathbf{t}_{pp}, \mathbf{t}_{ps}, \mathbf{v}_p, \mathbf{v}_s, \boldsymbol{\rho}\right)^T \Sigma_m^{-1} \left(\mathbf{Q} - \mathbf{G}(\mathbf{t}_{pp}, \mathbf{t}_{ps}, \mathbf{v}_p, \mathbf{v}_s, \boldsymbol{\rho}\right)\right). \quad (21)$$

In equation 21, k is the dimension of the multivariate Gaussian distribution and $\left|\Sigma_m\right|$ is the determinant of the covariance matrix Σ_m. One of the main advantages of using statistics in equation 21 is that we can have more options in defining likelihood functions so that we make their residuals uncorrelated.

Likelihood functions of PP and PS event time registration

The use of event time registration as data is one of main advantages of Delivery, as well as the current extension, because PP event time is directly related to P-wave velocity and PS event time directly related to P-wave and S-wave velocity. They provide additional information to constrain the estimates of P-wave and S-wave velocity beyond the reflectivity-based PP and PS full-waveforms.

Traditional methods for joint inversion of PP and PS data are primarily based on mapping of PS data to PP time (or domain conversion), in which PS data are considered as additional seismic stacks. Although this approach is simple to implement, it suffers from difficulties, such as wavelet distortion (Bansal and Matheney, 2010), because the conversion of PS time to PP time needs interval seismic P-to-S velocity ratios, which are not known *a priori*.

In this study, we avoid the PP-to-PS domain conversion and use PS data directly in the PS time domain. We pick a PS event from PS seismograms that has a good correspondence with a PP event in the PP seismograms along the same profile, and we refer it to as the master PS horizon. In the PS forward simulation, we calculate all the PS times relative to the master horizon. The relative PP and PS time for a given layer is calculated by:

$$\Delta t_{ps} = \frac{1}{2}\left(1 + \frac{v_p}{v_s}\right)\Delta t_{pp}. \tag{22}$$

In equation 22, both PP and PS velocity are interval velocity and are unknown; they will be estimated in inversion procedures.

The likelihood functions of PP and PS event registration time are determined by assuming the errors have multivariate Gaussian distribution. Let Σ_{pp} and Σ_{ps} be the covariance matrices of PP and PS event time, respectively. We have the following likelihood functions after ignoring the constant:

$$\begin{aligned} f(\mathbf{T}_{pp} \mid \mathbf{t}_{pp}) &\propto (2\pi)^{-k_1/2} \left|\Sigma_{pp}\right|^{-1/2} \exp\left(-\frac{1}{2}\left(\mathbf{T}_{pp} - \mathbf{t}_{pp}\right)^T \Sigma_{pp}^{-1} \left(\mathbf{T}_{pp} - \mathbf{t}_{pp}\right)\right), \text{ and} \\ f(\mathbf{T}_{ps} \mid \mathbf{t}_{ps}) &\propto (2\pi)^{-k_2/2} \left|\Sigma_{ps}\right|^{-1/2} \exp\left(-\frac{1}{2}\left(\mathbf{T}_{ps} - \mathbf{t}_{ps}\right)^T \Sigma_{ps}^{-1} \left(\mathbf{T}_{ps} - \mathbf{t}_{ps}\right)\right). \end{aligned} \tag{23}$$

In equation 23, k_1 and k_2 are the dimensions of $\mathbf{T}_{pp}$ and $\mathbf{T}_{ps}$; $\left|\Sigma_{pp}\right|$ and $\left|\Sigma_{ps}\right|$ are the determinants of the covariance matrix Σ_{pp} and Σ_{ps}.

Conditionals of unknowns and Markov chain Monte Carlo sampling methods

We use Markov chain Monte Carlo methods to draw many samples from the joint distribution given in equation 20. To do this, we need first derive conditional distributions of each type of unknowns given all other variables and data. The normalizing constants of each conditional are irrelevant when we use MCMC methods to draw samples. Therefore, we only need to keep the term on the right of equation 20 to get its conditional, which are given below:

$$f(\mathbf{t}_{pp}\,|\,\cdot) \propto f(\mathbf{Q}\,|\,\mathbf{t}_{pp},\mathbf{t}_{ps},\mathbf{v}_p,\mathbf{v}_s,\boldsymbol{\rho}) f(\mathbf{T}_{pp}\,|\,\mathbf{t}_{pp}) f(\mathbf{t}_{ps}\,|\,\mathbf{t}_{pp},\mathbf{v}_p,\mathbf{v}_s) f(\mathbf{t}_{pp}), \tag{24}$$

$$f(\mathbf{t}_{ps}\,|\,\cdot) \propto f(\mathbf{Q}\,|\,\mathbf{t}_{pp},\mathbf{t}_{ps},\mathbf{v}_p,\mathbf{v}_s,\boldsymbol{\rho}) f(\mathbf{T}_{ps}\,|\,\mathbf{t}_{ps}) f(\mathbf{t}_{ps}\,|\,\mathbf{t}_{pp},\mathbf{v}_p,\mathbf{v}_s), \tag{25}$$

$$f(\mathbf{v}_p,\mathbf{v}_s,\boldsymbol{\rho}\,|\,\cdot) \propto f(\mathbf{Q}\,|\,\mathbf{t}_{pp},\mathbf{t}_{ps},\mathbf{v}_p,\mathbf{v}_s,\boldsymbol{\rho}) f(\mathbf{t}_{ps}\,|\,\mathbf{t}_{pp},\mathbf{v}_p,\mathbf{v}_s) f(\mathbf{v}_p,\mathbf{v}_s,\boldsymbol{\rho}\,|\,\boldsymbol{\alpha}), \tag{26}$$

$$f(\boldsymbol{\alpha}\,|\,\cdot) \propto f(\mathbf{v}_p,\mathbf{v}_s,\boldsymbol{\rho}\,|\,\boldsymbol{\alpha}) f(\boldsymbol{\alpha}). \tag{27}$$

For equations 24-26, we cannot obtain analytical forms of those conditionals because PP and PS registration time and seismic attributes $\mathbf{v}_p$, $\mathbf{v}_s$, and $\boldsymbol{\rho}$ are nonlinear functions of other variables. We have to use MCMC methods (Gilks et al., 1996) to draw many samples from the joint posterior distribution.

In equation 27, we use the floating-grain rockphysics model given in equations 1-3 to link layered seismic attributes to their corresponding reservoir parameters, which is a linear function in this case. Let vector $\mathbf{r}$ be the combined vector of $\mathbf{v}_p$, $\mathbf{v}_s$, and $\boldsymbol{\rho}$ arranged by the layer indices and vector $\boldsymbol{\alpha}$ be the corresponding reservoir parameters. We thus have $\mathbf{r} = \boldsymbol{\mu}_r + \mathbf{H}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}_r$, where vector $\boldsymbol{\varepsilon}_r$ represents uncertainty associated with the linear relationship. We assume that it has a multivariate Gaussian distribution with zero mean and the covariance matrix of Σ_r. The detailed derivation and specific forms are given in Appendix C.

If we use a multivariate Gaussian prior for $\boldsymbol{\alpha}$, i.e., $f(\boldsymbol{\alpha}) \sim N(\boldsymbol{\mu}_p, \Sigma_p)$, we can obtain the analytical formula of posterior distribution, $f(\boldsymbol{\alpha} | \cdot) \sim N(\boldsymbol{\mu}_u, \Sigma_u)$, which is given below

$$\begin{cases} \Sigma_u^{-1} = \mathbf{H}^T \Sigma_r^{-1} \mathbf{H} + \Sigma_p^{-1}, \\ \Sigma_u^{-1} \mathbf{u} = \mathbf{H}^T \Sigma_r^{-1} (\mathbf{r} - \boldsymbol{\mu}_r) + \Sigma_p^{-1} \boldsymbol{\mu}_p. \end{cases} \tag{28}$$

We can obtain many samples of the joint posterior distribution given in equation 20 by using MCMC sampling methods. In this study, we draw many samples using revised Delivery developed by Gunning and Glinsky (2004).

CASE STUDY OF MULTIPLE LAYERS

We use the second example of Gunning and Glinsky (2007) to demonstrate the benefits of including converted-wave data into estimation of floating-grain fraction. Figure 4 shows various logs from an actual borehole, including P- and S-wave velocity, density, P- and S-wave velocity ratios, and P-wave impedance. According to the logs, we can build a synthetic model with six layers, which are (1) hard marl, (2) soft marl, (3) shale, (4) upper sand, (5) shale, and (6) lower sand from shallow to deep (see Figure 5). Both upper and lower sands are oil reservoirs with an oil saturation of 0.62 and thicknesses of 700 ft and 360 ft, respectively. Table 2 summarizes the main reservoir parameters and effective seismic properties. As shown in the table, layers 4 and 6 have relatively low Vp/Vs ratios and include the floating-grain fraction of 0.035.

PP and PS reflectivities and synthetic seismic data

We use the PP and PS wavelets typical of those derived from field borehole logs for sparse-spike inversion using the method by Sassen and Glinsky (2013). The PP and PS wavelets have the peak frequencies of 23 Hz and 13 Hz, respectively (see Figure 6).

We generate synthetic PP and PS data by first using equations 6 and 7 to calculate PP and PS reflectivities and then convolve the reflectivities with their corresponding wavelets. Table 3 shows the calculated full PP and the angle-weighted full PS and AVO gradient reflectivities at the five interfaces. As we can see that at the top interface, the reflectivities of the full PP stack (incident angle = 0 degrees) is very strong, which dominates the reflection from other deeper layers. Except at the interface between Layers 4 and 5, the magnitudes of PS reflectivities at all the interfaces are much larger than those of their corresponding AVO gradient reflectivities. We convolve those reflectivities with the given wavelets to get seismic full waveforms. Figure 7 shows the synthetic seismic data without noise added, where full PP and AVO gradient stacks are in the PP time domain whereas the full PS stacks are in the PS time domain. For inversion, we assume those data have uncorrelated Gaussian random noise with the standard deviation of 0.01 RFC.

Priors for the inversion

Since our main focus is on the demonstration of the value of PS data for reservoir parameter estimation, we mainly focus on the estimation of floating-grain fraction and net-to-gross in the upper and lower pay layers. Similar to Gunning and Glinsky (2007), we first consider prior $X \sim N(0.02, 0.03^2)$, which is a strong prior for the true floating-grain fraction of 0.035. Secondly, we consider a weak prior $X \sim N(0.0, 0.05^2)$, which give significant prior probability to the zero floating-grain fraction or clean sand. For net-to-gross (NG), we also consider two types of priors: (1) $NG \sim N(0.6, 0.1^2)$, and (2) $NG \sim N(0.5, 0.3^2)$.

Since we use model-based inversion methods, we can set a wide range of priors and consider many parameters as unknowns. For example, we assume that PP travel time to each interface has the normal distribution with the true values as mean and 10 milliseconds as the standard deviation. We assume the uncertainty in the thickness of Layer 4 is 70 ft (i.e., 10% of the thickness) and 20 ft for other layers. We increase the uncertainty of Layer 4 later to study its sensitivity to PS data.

Inversion cases

To demonstrate the usefulness of PS data for improving parameter estimation, we invert synthetic seismic data under the following four scenarios: (1) using only the full PP stack, (2) using full PP and AVO gradient stacks, (3) using full PP and PS stacks, and (4) using all the seismic data. We compare the posterior estimates of each case with their corresponding prior distributions to evaluate the benefit of using PS data.

Since the above comparisons usually depend on inversion situation, we consider the following three factors: (1) prior on floating-grain fraction (i.e., $X \sim N(0.02,\ 0.03^2)$ or $X \sim N(0.0,\ 0.05^2)$), (2) prior on net-to-gross (i.e., $NG \sim N(0.6,\ 0.1^2)$ or $NG \sim N(0.5,\ 0.3^2)$). We consider two sets of noise levels. The first one is that all seismic data have a noise level of 0.01 RFC, and the other is that PP full stacks have a noise level of 0.01 RFC, but AVO gradient stacks and PS stacks have a noise level of 0.02 RFC.

By changing priors and noise levels, we obtain many sets of posterior distributions. We use MCMC methods to draw 20,000 samples and keep the later half for analysis. With the use of those samples, we can obtain wide ranges of statistics, such as means, medians, modes, density functions, and predictive intervals. In the following several subsections, we selectively report our results.

Estimation of floating-grain fraction, porosity, and net-to-gross

We compare the estimates of reservoir parameters (i.e., floating-grain fraction, net-to-gross, and porosity) under different prior distributions. To investigate the effects of priors about floating-grain fraction, we use a strong prior about net-to-gross, i.e., $NG \sim N(0.6, 0.1^2)$, and noise levels for all the data types of 0.01 RFC. This implies we have the same quality for all the seismic data. We will explore the effects of noise levels later on.

Figure 8 compares the posterior probability densities (PDFs) of floating-grain fraction, porosity, and net-to-gross with their corresponding prior PDFs (black curves) for Layer 4 (i.e., upper-pay layer). The red, green, and blue curves represent the posterior PDFs after conditioning on full PP stacks, full PP and AVO gradient stacks, and full PP and PS stacks, respectively. For floating-grain fraction, even under the good prior (i.e, $X \sim N(0.02, 0.03^2)$), the mode of the prior probability corresponds to the zero floating-grain fraction or clean sand (see black curves in Figure 8a). After conditioning to seismic data (i.e., full PP, full PP plus AVO stacks, or full PP plus full PS stacks), the modes of the posterior PDFs corresponds to the true values 0.035 (see the red, green, and blue curves), with the results of using full PP and PS stacks slightly better than other two. As shown in Figure 9a, if we use a biased prior to clean sand, say $X \sim N(0.0, 0.05^2)$, the posterior estimates of floating-grain fraction using PP data only and using both full PP and AVO gradient stacks provide biased results (i.e., clean sand). However, the combination of full PP and full PS data provides correct estimates of the true value.

We can get similar results for comparison of porosity PDFs (see Figure 8b vs. Figure 9b). Under the good prior of floating-grain fraction, the modes of the posterior estimates for all the combinations of seismic data correspond to the true value quite well. But under the biased prior of floating-grain fraction, only the posterior estimates obtained using full PP and PS stacks provide good estimates of porosity. Since we use a very strong prior about net-to-gross (i.e., $NG \sim N(0.6, 0.1^2)$) for the true value of 0.65, we expect the updating of the prior is minimal for all the posterior estimates (see Figure 8c and Figure 9c).

We have similar comparisons of posterior PDFs for the lower-pay layer (i.e., Layer 6). Although overall the posterior estimates of floating-grain fraction and porosity are worse than those in the upper-pay layer, the combination of full PP and PS stack provides more information than full PP stacks only or the combination of full PP and AVO gradient stacks for updating the priors of floating-grain fraction and porosity.

Effects of the prior about net-to-gross and noise levels in seismic data

To explore the effects of prior about net-to-gross, we use less informative prior (i.e., $NG \sim N(0.5, 0.3^2)$) for net-to-gross and good prior about floating-grain fraction ($X \sim N(0.02, 0.03^2)$). Since the properties in the lower-pay layer are much less sensitive to seismic data, we only do the comparison for the upper-pay layer. Similar to what we found earlier, the combination of full PP and PS stacks significantly improve the estimates of floating-grain fraction and porosity (see the blue curves over the red and green curves in Figure 10a and 10b). Unlike previous comparison in Figures 8c and 9c, we found the combined use of full PP and PS stacks in this case significantly improve the estimates of net-to-gross when it has significant uncertainty (see Figure 10c).

In reality, it is more difficult to collect and process full PS and AVO gradient stacks compared to full PP data. Therefore, they are likely subject to larger noise. To explore the effects of noise levels on reservoir parameter estimation, we let the prior of floating-grain fraction be $X \sim N(0.02, 0.03^2)$ and let net-to-gross prior be $NG \sim N(0.6, 0.1^2)$. We set the noise level in full PP stacks as 0.01 RFC but noise levels in full PS and AVO gradient stacks as 0.02 RFC. Figure 11 shows the posterior PDFs of floating-grain fraction, porosity and net-to-gross. Although the estimated results slightly worse than those obtained using noise levels of 0.01 RFC (see Figure 8), the conclusions remain the same.

Comparison of discrepancies between the estimated and the true values

Since we use sampling-based methods for inversion, we can obtain many samples of other variables as given in equation 20, such as effective P-wave and S-wave velocity, density, layer-thickness, etc. With the use of those samples, we can not only visually compare prior and posterior PDFs but also calculate a wide range of statistics. In the previous comparisons, we qualitatively compare the posterior estimates with their corresponding priors. To demonstrate the value of PS data, in this section, we quantitatively compare the estimated results with their true values.

We first compare the difference between the estimated median and the true value, which measures how accurate a chosen point estimator (in this case, median) to the true value of a given parameter. Figure 12a compares the differences between the estimated floating-grain fraction, porosity, and net-to-gross values with their true values. The priors for floating-grain fraction is $X \sim N(0.0, 0.05^2)$ and for net-to-gross is $NG \sim N(0.6, 0.1^2)$, and the noise levels are 0.01 RFC for full PP stacks and 0.02 for other data sets. We normalize the results by the difference obtained from prior distributions. For net-to-gross, as we demonstrated early, under the good prior, the estimated medians do not improve the prior medians. The value slightly over 1.0 may reflect the effects of noise in seismic data or sampling variations during the inversion procedure. For floating-grain fraction and porosity, when conditioning to full PP data, the differences are significantly reduced (see the triangles and circles corresponding to 'R0' in Figure 12a). When adding AVO gradient stacks, the improvement is minimal (see the points corresponding to 'R0R2' in Figure 12a), but adding PS data leads to significant reduction (see the points corresponding to 'R0R1').

Figure 12b compares the differences for effective P-wave and S-wave velocity, effective density, and layer thickness. For effective P-wave velocity and density, conditioning to seismic full PP data significantly improves the accuracy, and further adding AVO gradients or PS data does not lead significantly reduction. However, for effective S-wave velocity and layer thickness, either adding AVO gradient stacks or PS stacks lead to further reduction of the discrepancies, but adding of PS data causes more reduction. For density (see the crosses in Figure 12b), adding full PS data does not lead to significant reduction in uncertainty. This is because for the current case study, after conditioning to full PP data, the uncertainty is already very small, leaving less room for further improvement.

Comparison of widths of uncertainty bounds

The MCMC-based methods also allow us to quantitatively compare the uncertainty associated with all the estimation. In this study, we calculate the widths of 95% predictive intervals. Similar to the comparison of the discrepancies, we normalize the results by those obtained from the prior PDFs.

Figure 13 shows the results for reservoir parameters and for effective parameters. For reservoir parameters (i.e., floating-grain fraction, porosity, and net-to-gross), the reduction of uncertainty is small and the maximum values are around 20%. The use of various combinations of seismic data seems not make significant difference. For P-wave velocity and density, after conditioning to full PP data, adding AVO gradient data or full data does not lead further significant reduction. However, for S-wave velocity and layer thickness, adding PS data causes significantly more reduction in the uncertainty than adding AVO gradient data.

Comparison of predictive probabilities

In the previous sections, we compare the discrepancy between the estimated and true value and the widths of uncertainty bounds, both of which just compare one aspect of posterior PDFs. A better evaluation is to compare the predictive probabilities of a small interval around the true value, which is given by

$$\mathrm{Prob}(\theta \in [(1-\varepsilon)\theta^{True}, (1+\varepsilon)\theta^{True}] \mid \mathrm{Data}) . \quad (29)$$

In equation 29, we set $\varepsilon = 2.5\%$ for effective density and 5% for other parameters because the posterior density has much smaller uncertainty compared to other effective properties. The large predictive probability means that the data provide stronger evidence to support the occurrence of the true values. Again, we normalize the probabilities by the prior predictive probability.

Figure 14a compares the predictive probability ratios of floating-grain fraction, porosity, and net-to-gross. These results are more consistent than those shown in Figures 12a and 13a as the ratios of net-to-gross are very close to 1.0. This means that for the tight prior of net-to-gross ($NG \sim N(0.5, 0.1^2)$), the updating is ignorable. For floating-grain fraction and porosity, the use of full PP stacks significantly increases the predictive probabilities. Adding AVO gradient stacks does not cause significant improvement. However, adding of full PS stacks leads to significant improvement again. Figure 14b shows the similar comparison for effective P-wave and S-wave velocity, effective density, and layer thickness. Similar statements to Figures 12b and 13b hold, but their results are more consistent.

CONCLUSIONS

We started from likelihood analysis of a simple two-layer model based on the floating-grain rock-physics model and found that seismic PS data are significantly more informative than AVO gradient data for reservoir parameter estimation. This motivated us to develop a hierarchical Bayesian model to combine PP and PS data under more complicated cases (e.g., multiple layers, a large number of unknowns, etc). We inverted PS data directly in the PS time domain unlike many previous methods that first convert PS time to PP time and then invert PS data in the PP time domain. The alignment of PP and PS time was carried out by identifying a common reflection interface and using the PP and PS time to the interface as references. This avoids many difficulties caused by the conversion of PS time to PP time, such as the distortion of wavelets, and the requirement of knowing internal P-wave to S-wave velocity ratios a-priori. Since we considered PS time registration as data in the model, we can obtain more information during the joint inversion compared to previous methods for inverting PS data.

We revised the open-source Java codes, 'Delivery', to implement the developed Bayesian model by adding converted wave responses and PS time registration as additional data and two-way PS travel time and PS reflectivity as additional unknowns. We took advantage of general structures for model setup, convenient ways to specify priors, efficient MCMC sampling methods, and various post analysis tools. We demonstrated the use of the revised codes by applying them to a six-layer synthetic model built on actual borehole logs; the codes are effective and convenient for joint inversion of seismic data at multiple CDPs.

We performed comparison studies based on the synthetic six-layer model to demonstrate the value of PS data for inversion of reservoir parameters. We compared the inversion results obtained from using full PP stacks only, full PP stacks plus AVO gradient stacks, and full PP stacks plus PS full stacks. We found that PS data are very helpful for improving the estimates of porosity and floating-grain fraction and for improving the estimates of effective S-wave velocity and layer-thickness under a range of priors and noise levels in seismic data. Net-to-gross is relatively less sensitive to PS data. Compared to the posterior results obtained from full PP plus AVO gradient stacks, we found that full PP stacks are most informative for parameter estimation, then full PS stacks, and finally AVO gradient stacks. This suggests that to improve the estimates of reservoir parameters, full PS stacks are more valuable because PS data can provide complementary information to PP full data, and give similar but better information than

AVO gradient data. Consequently, they have the potential of significantly improving parameter estimation results.

ACKNOWLEDGMENTS

We thank ION Geophysical for funding and for permission to publish this work. We thank James Gunning from CSIRO for providing help in understanding the Delivery codes.

APPENDIX A

DERIVATION OF MEASUREMENT EQUATION FOR TWO-LAYER MODELS

Let $R_{fullPP}(0)$ be the full PP stack with the incident angle of zero and let $R_{fullPS}(\theta)$ and $R_{AVO-gradient}(\theta)$ be the full PS and AVO gradient stacks with the incident angle of θ. From equations 6-8, we have $R_{fullPP}(0) = A_0$, $R_{fullPS}(\theta) = A_1\theta$, and $R_{AVO-gradient}(\theta) = A_2\theta^2$. Let $\mathbf{R} = (R_{fullPP}(0), R_{fullPS}(\theta), R_{AVO-gradient}(\theta))^T$ be the reflectivities at the interface, where T represents the transpose of a vector or matrix. Let $\mathbf{R}_m$ be the measurements with noise added, i.e., $\mathbf{R}_m = \mathbf{R} + \boldsymbol{\varepsilon}_m$. Thus, we have

$$\mathbf{R}_m = \begin{pmatrix} R_{fullPP}(0) \\ R_{fullPS}(\theta) \\ R_{AVO-gradient}(\theta) \end{pmatrix} + \boldsymbol{\varepsilon}_m = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \theta & 0 \\ 0 & 0 & \theta^2 \end{pmatrix} \begin{pmatrix} A_0 \\ A_1 \\ A_2 \end{pmatrix} + \boldsymbol{\varepsilon}_m = \mathbf{M}_\theta \Delta\mathbf{C} + \boldsymbol{\varepsilon}_m . \qquad \text{(A-1)}$$

APPENDIX B

ANGLE-DEPENDENT MATRICES FOR SYNTHETIC TWO-LAYER MODELS

For the case of using only full PP stacks, we set $\mathbf{M}_\theta = (1, 0, 0)$. For the case of using full PP and full PS stacks, we set

$$\mathbf{M}_\theta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \theta & 0 \end{pmatrix}. \qquad \text{(B-1)}$$

Similarly, for the case of using full PP and AVO gradient, we set

$$\mathbf{M}_\theta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & \theta^2 \end{pmatrix}. \qquad \text{(B-2)}$$

For the case of using all seismic data, we set

$$\mathbf{M}_\theta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \theta & 0 \\ 0 & 0 & \theta^2 \end{pmatrix}. \qquad \text{(B-3)}$$

APPENDIX C

DERIVATION OF MEAN VECTOR AND COVARIANCE MATRICES

In the current study, we assume the reservoir parameters under estimation are loading-depth and floating-grain fraction. Let v_{pi}, v_{si}, ρ_i, z_i, and x_i be seismic P- and S-wave velocity, density, loading depth, and floating-grain fraction at the i-th layer, respectively. From the rockphysics model in given equation 1-5, we have

$$\begin{aligned}\mathbf{r}_i = \begin{pmatrix} v_{pi} \\ v_{si} \\ \rho_i \end{pmatrix} &= \begin{pmatrix} a_{vp} \\ a_{vs} + a_{vp} a_{vs} \\ a_\rho + a_{vp} b_\rho \end{pmatrix} + \begin{pmatrix} b_{vp} & c_{vp} \\ b_{vp} b_{vs} & c_{vp} b_{vs} \\ b_{vp} b_\rho & c_{vp} b_\rho + c_\rho \end{pmatrix} \begin{pmatrix} z_i \\ x_i \end{pmatrix} + \begin{pmatrix} \varepsilon_{vp} \\ b_{vs}\varepsilon_{vp} + \varepsilon_{vs} \\ b_\rho \varepsilon_{vp} + \varepsilon_\rho \end{pmatrix} \\ &= \boldsymbol{\mu}_{ri} + \mathbf{H}_i \boldsymbol{\alpha}_i + \boldsymbol{\varepsilon}_{ri}. \end{aligned} \qquad \text{(C-1)}$$

We can form vectors and matrices for all the layers by stacking those layer-based vectors and matrices, i.e., $\mathbf{r} = (\mathbf{r}_1^T, \mathbf{r}_2^T, \cdots, \mathbf{r}_n^T)^T$, $\boldsymbol{\mu}_r = (\boldsymbol{\mu}_1^T, \boldsymbol{\mu}_2^T, \cdots, \boldsymbol{\mu}_n^T)^T$, $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_1^T, \boldsymbol{\alpha}_2^T, \cdots, \boldsymbol{\alpha}_n^T)^T$, $\boldsymbol{\varepsilon}_r = (\boldsymbol{\varepsilon}_1^T, \boldsymbol{\varepsilon}_2^T, \cdots, \boldsymbol{\varepsilon}_n^T)^T$, and $\mathbf{H} = (\mathbf{H}_1^T, \mathbf{H}_2^T, \cdots, \mathbf{H}_n^T)^T$.

It is straightforward to derive covariance matrix from equation B-1 by assuming residuals ε_{vp}, ε_{vs}, and ε_{ρ} in equations 1-3 have Gaussian distributions with zero mean and variances of σ_{vp}^2, σ_{vs}^2, and σ_{ρ}^2, respectively. Specifically, the matrix is

$$\Sigma_{ri} = \sigma_{vp}^2 \begin{pmatrix} 1 & b_{vs} & b_{\rho} \\ b_{vs} & b_{vs}^2 + \sigma_{vs}^2 / \sigma_{vp}^2 & b_{vs} b_{\rho} \\ b_{\rho} & b_{vs} b_{\rho} & b_{\rho}^2 + \sigma_{\rho}^2 / \sigma_{vp}^2 \end{pmatrix}. \quad \text{(C-2)}$$

The covariance matrix $\Sigma_r = diag(\Sigma_1, \Sigma_2, \cdots, \Sigma_n)$.

REFERENCES

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

Bale, R., T. Marchand, K. Wilkinson, K. Wikel, and R. Kendall, 2013, The signature of shear-wave splitting: Theory and observations on heavy oil data: The Leading Edge, **32**, 14-24.

Bansal, R., and M. Mathency, 2010, Wavelet distortion correction due to domain conversion: Geophysics, **75**, V77-V87.

Bernardo, J. M., and F. M. Smith, 2000, Bayesian theory: John Wiley & Sons, LTD.

Brettwood, P., J. P. Leveille, and S. Singleton, 2013, C-wave data improve seismic imaging: The American Oil & Gas Reporter, **1**.

Chen, J., A. Kemna, and S. Hubbard, 2008, A comparison between Gauss-Newton and Markov chain Monte Carlo based methods for inverting spectral induced polarization data for Cole-Cole parameters: Geophysics, **73**, F247-F259.

Davis, T. L., A. Bibolova, S. O'Brien, D. Klepacki, and H. Robinson, 2013, Prediction of residual oil saturation and cap-rock integrity from time-lapse, multicomponent seismic data, Delhi field, Louisiana: The Leading Edge, **32**, 26-31.

DeMartini, D. C., and M. E. Glinsky, 2006, A model for variation of velocity versus density trends in porous sedimentary rocks: Journal of Applied Physics, **100**, 014910.

Gardner, G. H. F., L. W. Gardner, and A. R. Gregory, 1974, Formation velocity and density – The diagnostic basics for stratigraphic traps: Geophysics, 39, 770-780.

Gelman, A., and D. B. Rubin, 1992, Inference from iterative simulation using multiple sequences: Statistical Science, **7**, 457-472.

Glinsky, M. E., A. Cortis, D. Sassen, H. Rael, and J. Chen, 2013, Rockphysics and geophysics for unconventional resources, multicomponent seismic, quantitative interpretation: The 2nd International Workshop on Rock-Physics (2IWRP), 2013, South Hampton, United Kingdom, August 4-9, http://arxiv.org/abs/1304.6048.

Gilks, W., S. Richardson, and D. Spiegelhalter, 1996, Markov Chain Monte Carlo in Practice: Chapman & Hall/CRC, New York.

Gunning, J., and M. E. Glinsky, 2004, Delivery: an open-source model-based Bayesian seismic inversion program: Computers and Geosciences, **30**.

Gunning, J., and M. E. Glinsky, 2007, Detection of reservoir quality using Bayesian seismic inversion: Geophysics, **72**, R37-R49.

Hardage, B. A., M. V. DeAngelo, P. E. Murray, and D. Sava, 2011, Multicomponent seismic technology: Society of Exploration Geophysicists.

Mahmoudian, F., and G. F. Margrave, 2004, Three-parameter AVO inversion with PP and PS data using offset-binning: CREWES Report, **16**.

Pacal, E. E., 2012, Seismic imaging with ocean-bottom nodes (OBNs): new acquisition designs and the Atlantis 4C OBN: Master Thesis, University of Houston.

Rodriguez-Saurez, C., 2000, Advanced marine methods: ocean- bottom and vertical cable analyses: PhD Thesis, University of Calgary.

Sassen, D., M. E. Glinsky, 2013, Noise-thresholding sparse-spike inversion with global convergence: calibration and applications: 83th Annual International Meeting, SEG, Expanded Abstracts, 3073-3076.

Stewart, R. R., J. E. Gaiser, R. J. Brown, and D. C. Lawton, 2002, Tutorial: Converted-wave seismic exploration: methods: Geophysics, **67**, 1348-1363.

Stone, C. J., 1995, A course in probability and statistics: Duxbury Press.

Tarantola, A., 2005, Inverse problem theory and methods for model parameter estimation: The Society for Industrial and applied mathematics (SIAM), Philadelphia, PA.

Veire, H. H., and Landrø, M., 2006, Simultaneous inversion of PP and PS seismic data: Geophysics, **71**, R1–R10.

Venables, W. N., and B. D. Ripley, 1999, Modern applied statistics with S-Plus, Third Edition, *Springer*, New York.

FIGURE CAPTIONS

Figure 1: Likelihoods of floating-grain fraction given various data combinations for the true value of (a) 0.0% and (b) 3.5%. The black, green, red, and blue curves show the likelihoods of using (1) full PP stacks only, (2) full PP and AVO gradient stacks, (3) full PP and PS stacks, and (4) all three types of seismic data.

Figure 2: Likelihood ratios of using various combinations of seismic data to that of using full PP stacks only as a function of measurement errors in full PS and AVO gradient stacks. The green, red, and blue curves show the results of using full PP and AVO gradient stacks, full PP and PS stacks, and all three data sets, respectively.

Figure 3: Dependent relationships among unknown parameters and data.

Figure 4: Various logs from an actual borehole as a function of depth: (a) P-wave velocity (ft/s), (b) S-wave velocity (ft/s), (c) density (g/cc), (d) Vp/Vs, and (e) P-Impedance (MPa). The solid red line segments in (a) are the approximate layer interfaces.

Figure 5: Six-layer model, where Layers 4 and 6 are oil reservoir with oil saturation of 0.62, net-to-gross of 0.65, and floating-grain fraction of 0.035.

Figure 6: Normalized PP and PS wavelets extracted from an actual field site.

Figure 7: Seismic data without noise added.

Figure 8: Posterior probability distribution of floating-grain fraction when priors about floating-grain fraction and net-to-gross are strong (i.e., $X \sim N(0.02,\ 0.03^2)$ and $NG \sim N(0.6,\ 0.1^2)$, the reference case).

Figure 9: Posterior probability distribution of floating-grain fraction when the prior about floating-grain fraction is weak (i.e., $X \sim N(0.0,\ 0.05^2)$).

Figure 10: Posterior probability distribution of floating-grain fraction when the prior about net-to-gross is weak (i.e., $NG \sim N(0.5,\ 0.3^2)$).

Figure 11: Posterior probability distribution of floating-grain fraction when errors in AVO gradient and full PS stacks are doubled.

Figure 12: Comparison of differences between the true values and estimated medians for priors $X \sim N(0.02,\ 0.03^2)$, $NG \sim N(0.5,\ 0.1^2)$, and noise of 0.01 for full PP and 0.02 for others.

Figure 13: Comparison of half-widths of 95% predictive intervals for priors $X \sim N(0.02,\ 0.03^2)$, $NG \sim N(0.5,\ 0.1^2)$, and noise of 0.01 for full PP and 0.02 for others.

Figure 14: Comparison of predictive probability of the true values for priors $X \sim N(0.02,\ 0.03^2)$, $NG \sim N(0.5,\ 0.1^2)$, and noise of 0.01 for full PP and 0.02 for others.

Table 1: Sand and shale rock-physics model coefficients from actual borehole logs

	Regression Equations	Standard Errors	Units
Sand	$v_p = 2.12\times10^3 + 5.08\times10^{-1} Z + 1.80\times10^4 X$	344	ft/s
	$v_s = -4.00\times10^3 + 8.94\times10^{-1} v_p$	226	ft/s
	$\rho = 1.70 + 5.04\times10^{-5} v_p + 1.56X$	0.0149	g/cc
Shale	$v_p = (-5.38\times10^3) + 9.46\times10^{-1} Z$	476	ft/s
	$v_s = -3.3730\times10^3 + 8.0120\times10^{-1} v_p$	226	ft/s
	$\rho = 0.534 v_p^{0.166}$	0.030	g/cc

Table 2: Synthetic six-layer model parameters

	Layer-1	Layer-2	Layer-3	Layer-4	Layer-5	Layer-6
Geology	Hard Marl	Soft Marl	Shale	Upper Sand	Shale	Lower Sand
Vp (km/s)	3.67	2.85	3.32	3.49	3.48	3.58
Vs (km/s)	1.75	1.17	1.63	1.89	1.76	1.97
Density (g/cc)	2.54	2.38	2.50	2.39	2.52	2.41
Vp/Vs	2.10	2.44	2.04	1.84	1.98	1.81
Net-to-gross (NG)	0.0	0.0	0.0	0.65	0.0	0.65
Porosity				0.187		0.181
Floating-grain fraction				0.035		0.035

Table 3: PP and PS reflectivities of six-layer models

Interfaces	Full PP Stack (0 degrees)	Angle-weighted AVO gradient Stack (45 degrees)	Angle-weighted Full PS Stack (45 degrees)
Layers 1 and 2	-0.159	0.135	0.328
Layers 2 and 3	0.100	-0.133	-0.272
Layers 3 and 4	0.004	-0.070	-0.088
Layers 4 and 5	0.025	0.031	0.019
Layers 5 and 6	-0.010	-0.054	-0.057

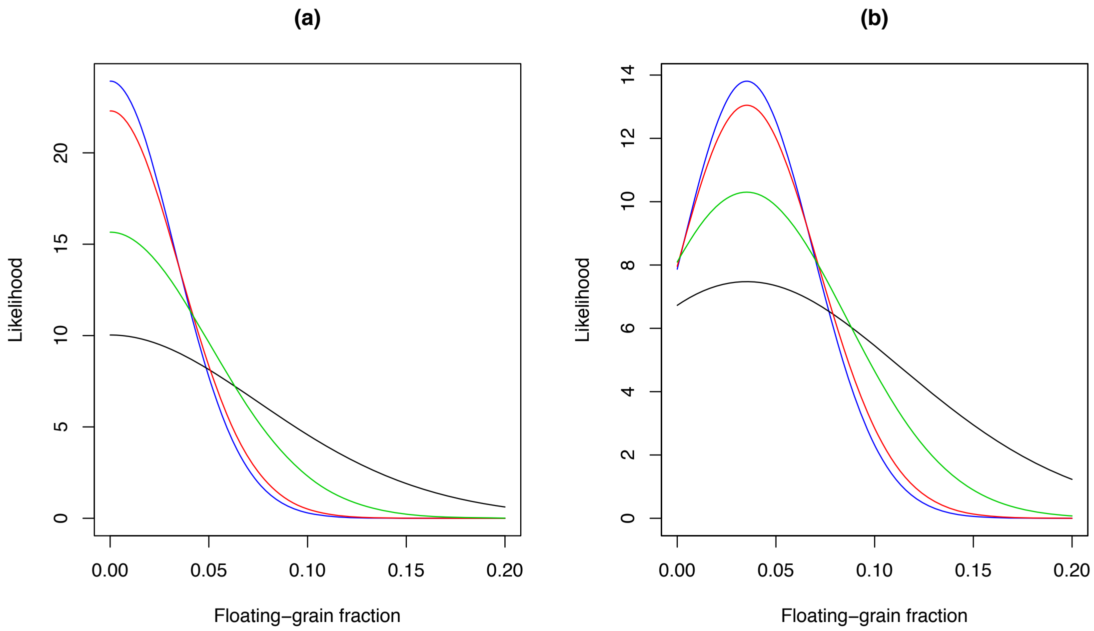

Figure 1. Likelihoods of floating-grain fraction given various data combinations for the true value of (a) 0.0% and (b) 3.5%. The black, green, red, and blue curves show the likelihoods of using (1) full PP stacks only, (2) full PP and AVO gradient stacks, (3) full PP and PS stacks, and (4) all three types of seismic data.

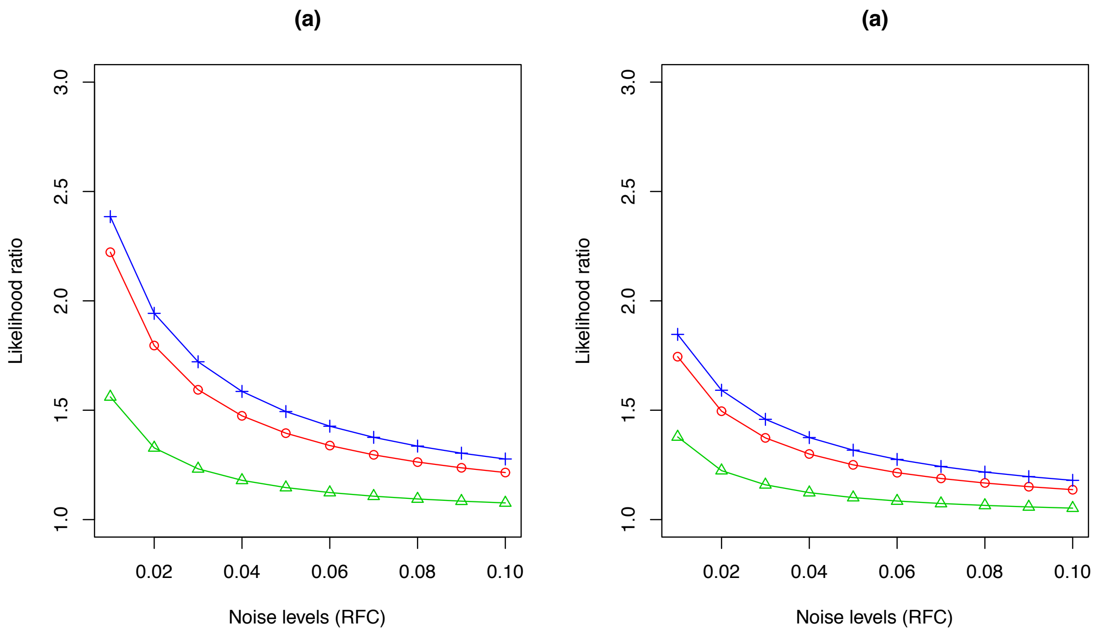

Figure 2. Likelihood ratios of using various combinations of seismic data to that of using full PP stacks only as a function of measurement errors in full PS and AVO gradient stacks. The green, red, and blue curves show the results of using full PP and AVO gradient stacks, full PP and PS stacks, and all three data, respectively.

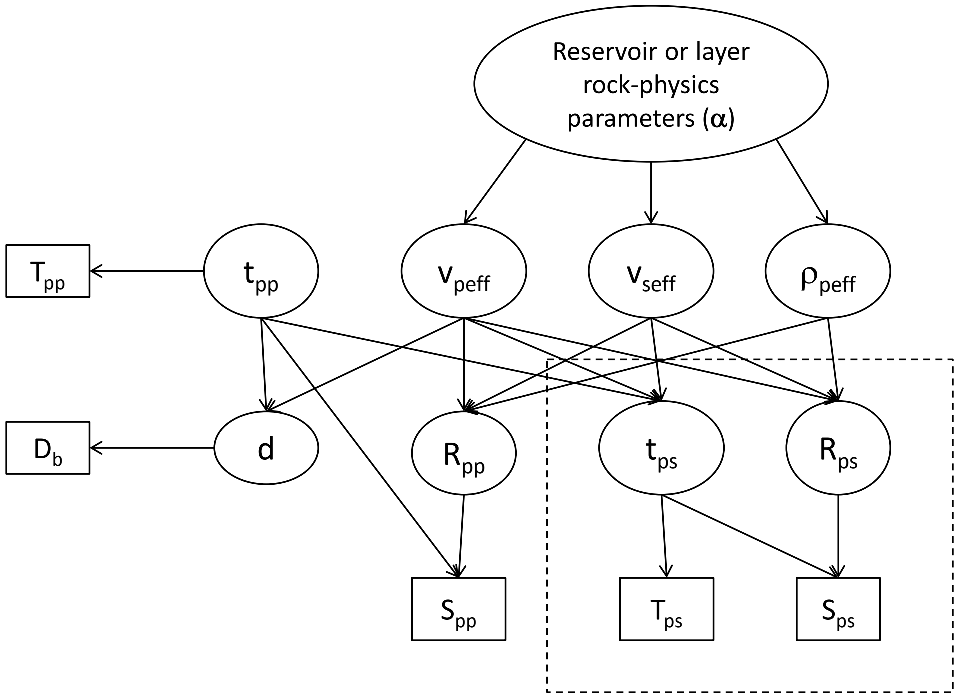

Figure 3. Dependent relationships among unknown parameters and data.

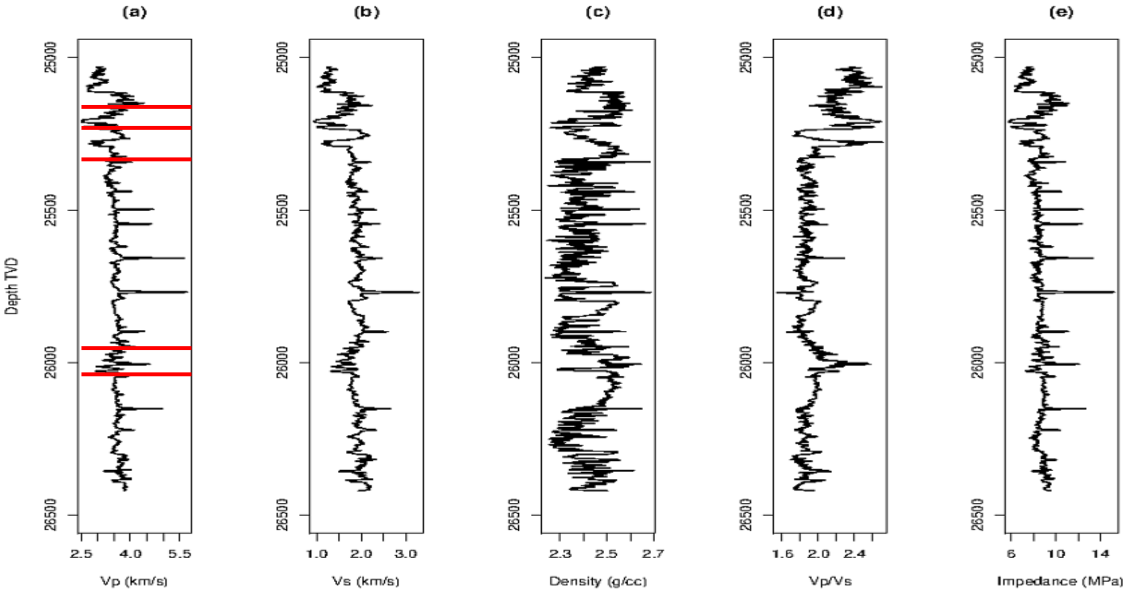

Figure 4. Various logs from an actual borehole as a function of depth: (a) P-wave velocity (ft/s), (b) S-wave velocity (ft/s), (c) density (g/cc), (d) Vp/Vs, and (e) P-Impedance (MPa). The solid red line segments in (a) are the approximate layer interfaces.

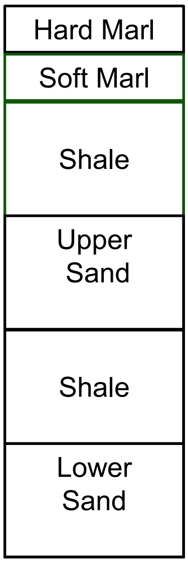

Figure 5. Six-layer model, where layers 4 and 6 are oil reservoir with oil saturation of 0.62, net-to-gross of 0.65, and floating-grain fraction of 0.035.

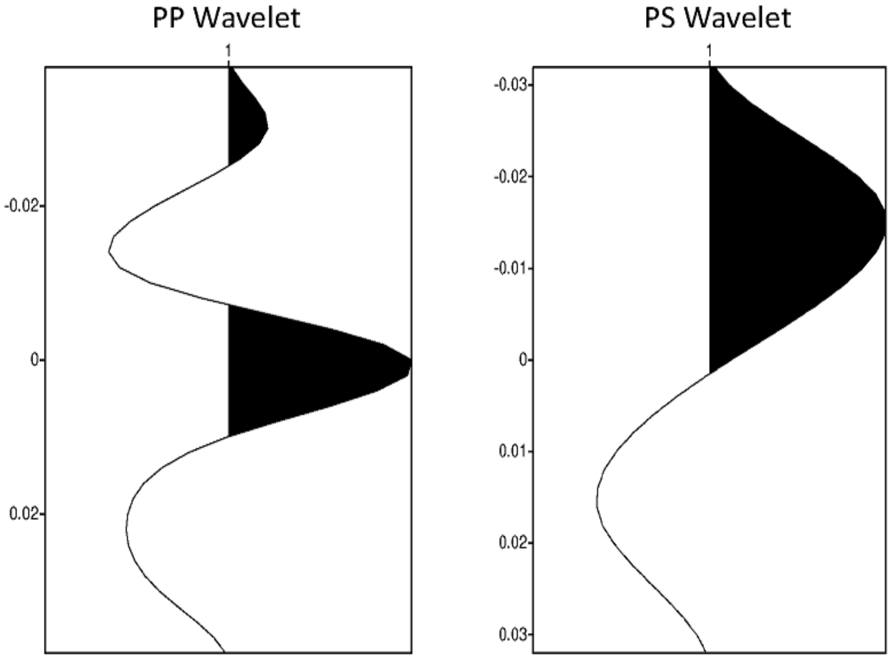

Figure 6. Normalized PP and PS wavelets extracted from an actual field site.

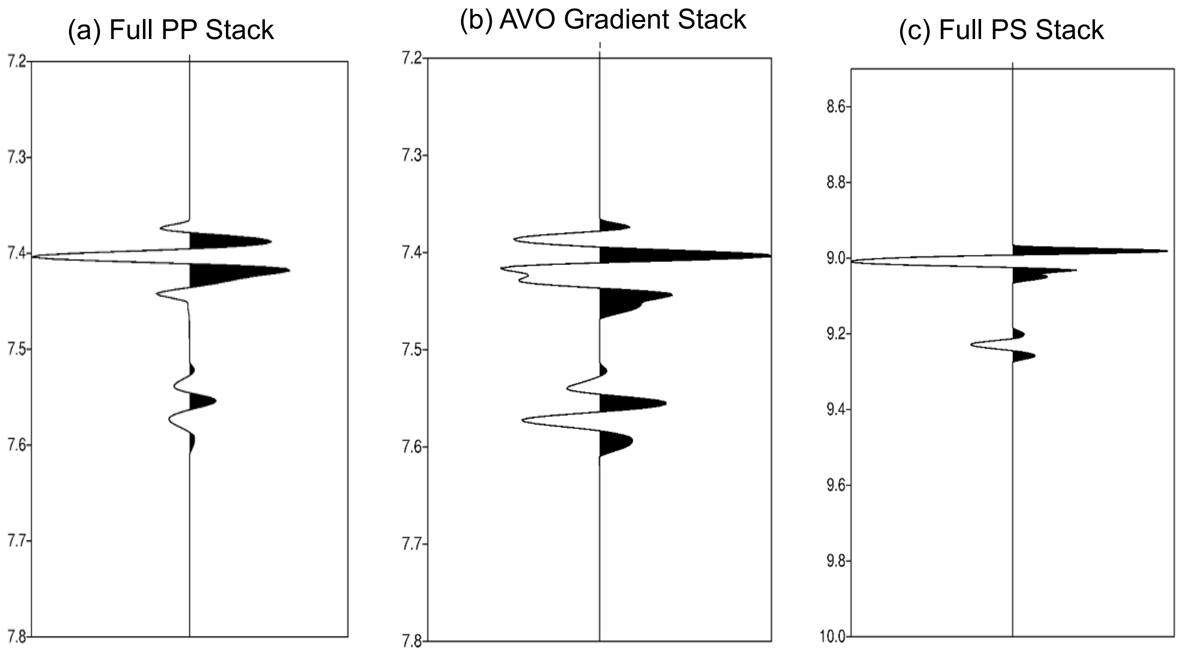

Figure 7. Seismic data without noise added.

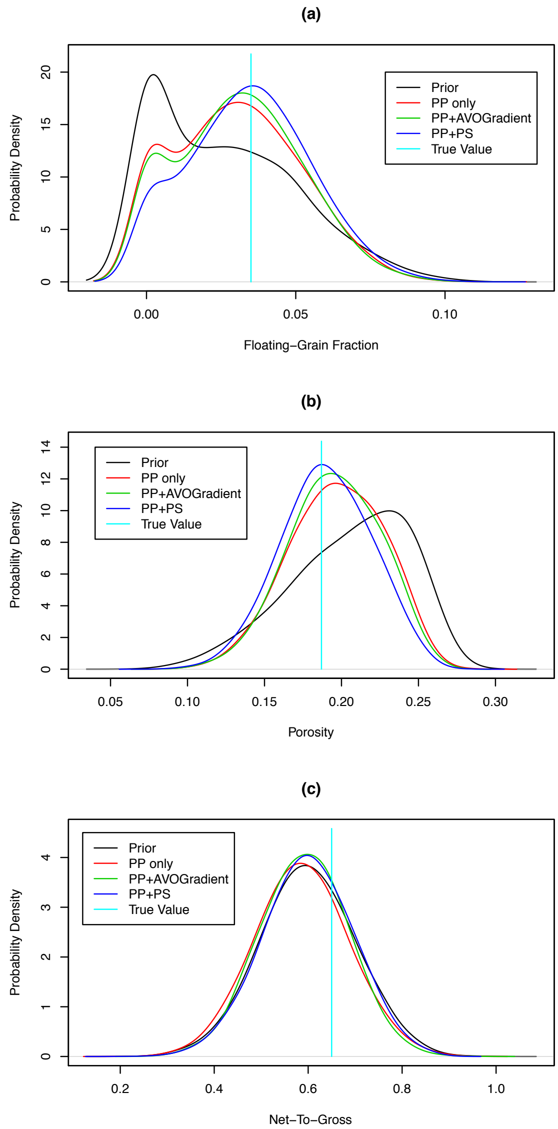

Figure 8. Posterior probability distribution of floating-grain fraction when priors about floating-grain fraction and net-to-gross are strong (i.e., $X \sim N(0.02, 0.03^2)$ and $NG \sim N(0.6, 0.1^2)$, the reference case).

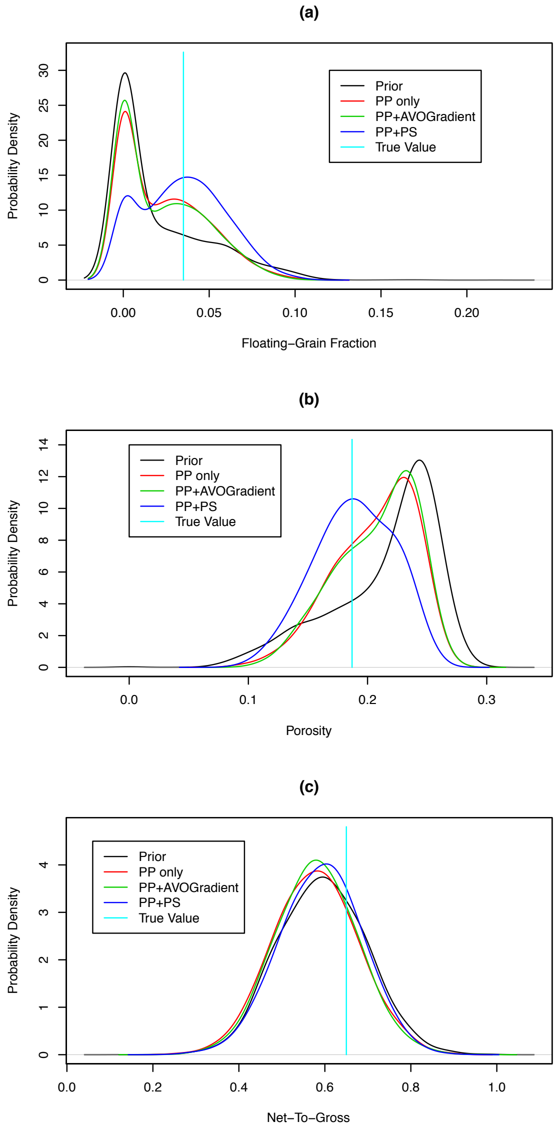

Figure 9. Posterior probability distribution of floating-grain fraction when the prior about floating-grain fraction is weak (i.e., $X \sim N(0.0, 0.05^2)$).

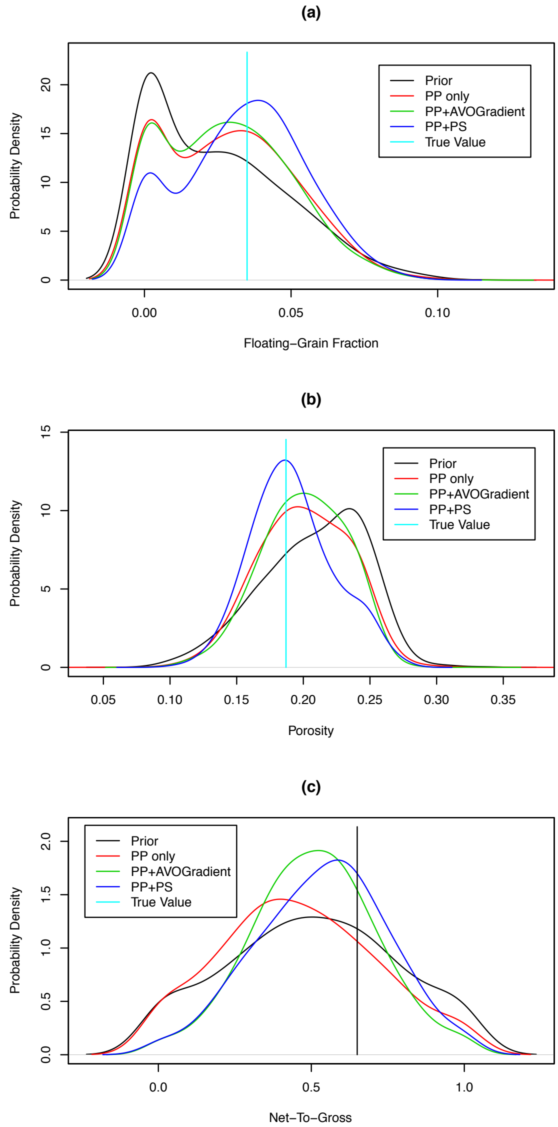

Figure 10. Posterior probability distribution of floating-grain fraction when the prior about net-to-gross is weak (i.e., $NG \sim N(0.5, 0.3^2)$).

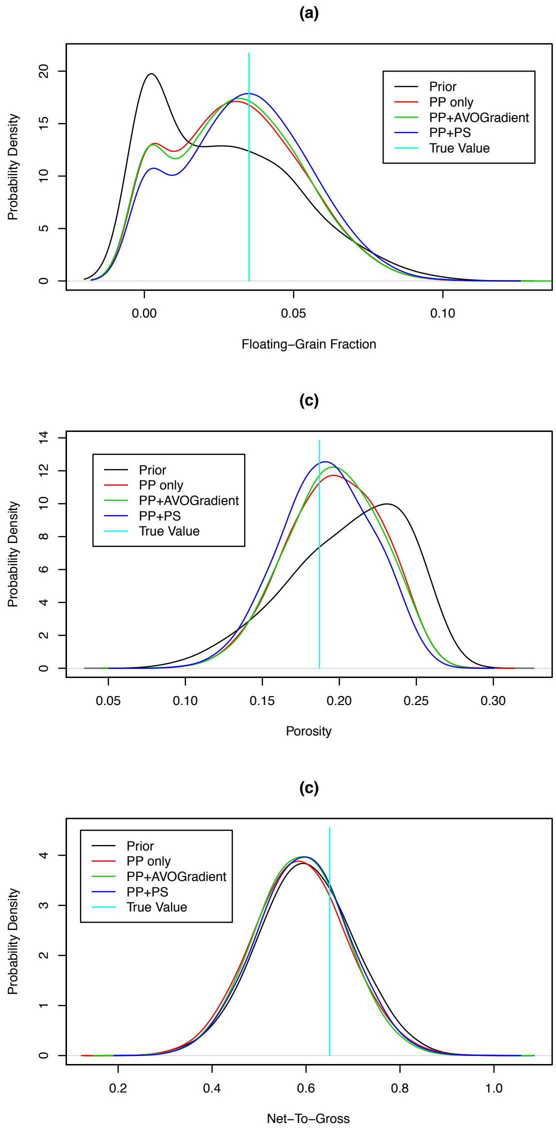

Figure 11. Posterior probability distribution of floating-grain fraction when errors in AVO gradient and full PS stacks are doubled.

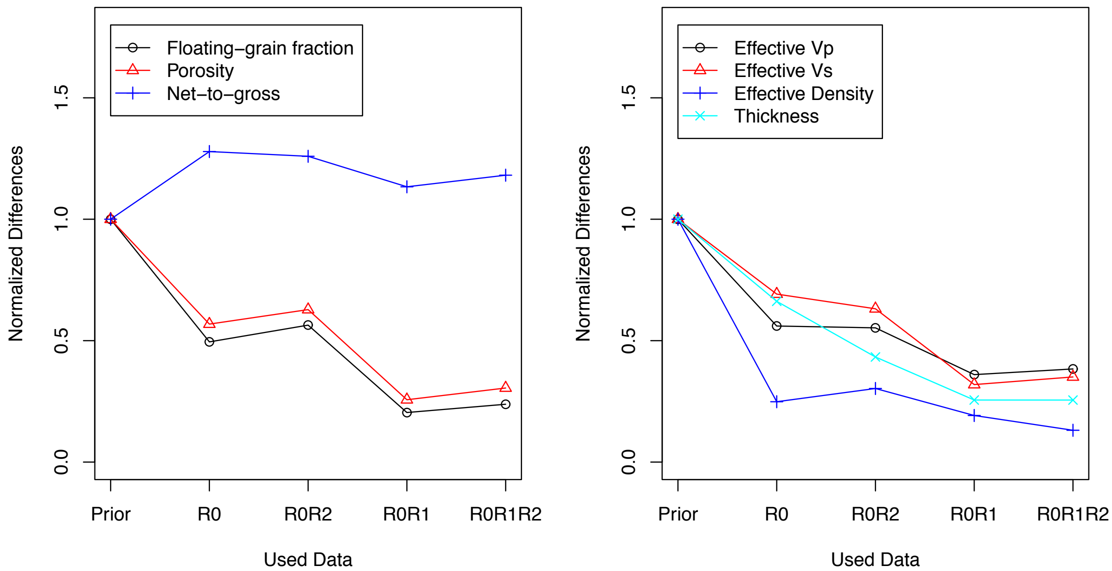

Figure 12. Comparison of differences between the true values and estimated medians for priors $X \sim N(0.02, 0.03^2$, $NG \sim N(0.5, 0.1^2)$, and noise of 0.01 for full PP and 0.02 for others.

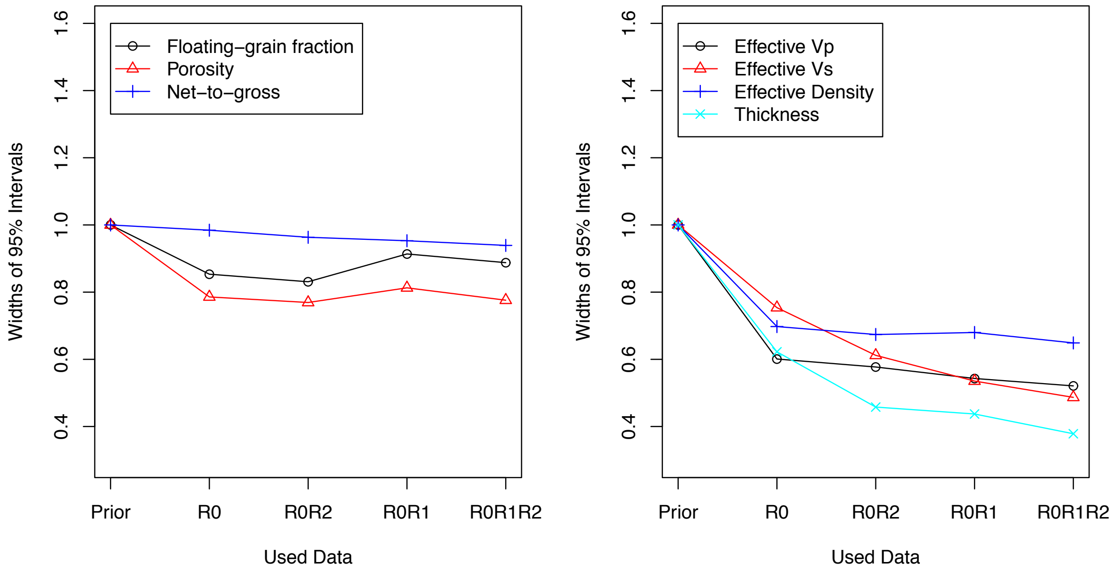

Figure 13. Comparison of half-widths of %95 predictive intervals for priors $X \sim N(0.02, 0.03^2$, $NG \sim N(0.5, 0.1^2)$, and noise of 0.01 for full PP and 0.02 for others.

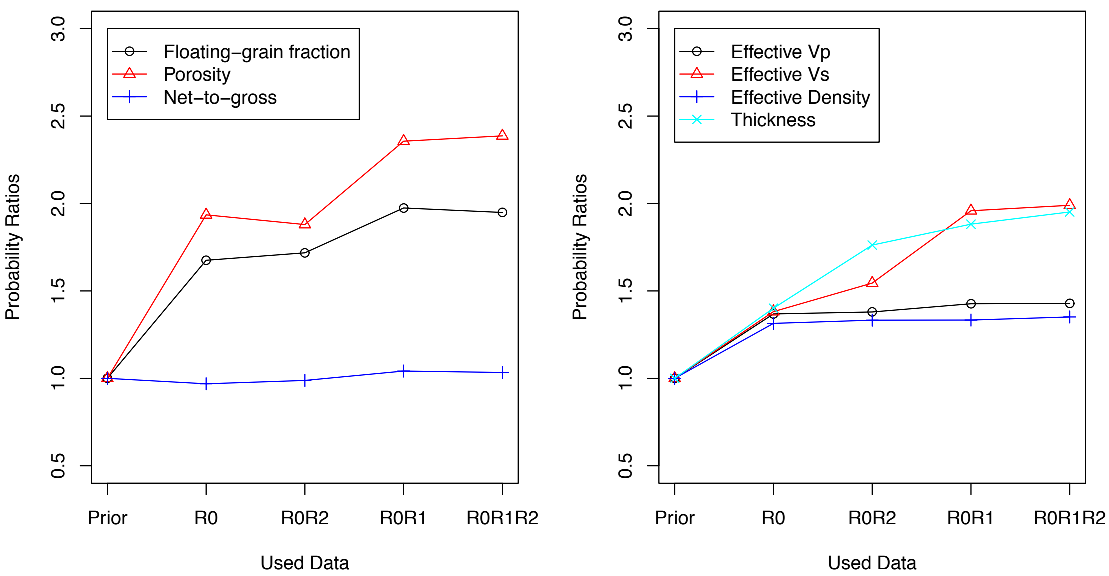

Figure 14. Comparison of predictive probability of the true values for priors $X \sim N(0.02, 0.03^2$, $NG \sim N(0.5, 0.1^2)$, and noise of 0.01 for full PP and 0.02 for others.