Elastic wavefield tomography with probabilistic petrophysical clustering

Odette Aragao & Paul Sava
Center for Wave Phenomena, Colorado School of Mines

ABSTRACT
Full waveform inversion (FWI) aims to use all information provided by seismic data to deliver high resolution models of subsurface parameters. However, multiparameter FWI suffers from an inherent trade-off between parameters and from ill-posedness due to the highly nonlinear nature of FWI. Also, the models recovered using elastic FWI are subject to local minima if the initial models are far from the optimal solution. In addition, an objective function purely based on the misfit between recorded and modeled data may honor the seismic data, but disregard the geological context. Hence, the inverted models may be geologically inconsistent, and not represent feasible lithological units. We propose that all the aforementioned difficulties can be alleviated by explicitly incorporating petrophysical information into the inversion through a penalty function based on multiple probability density functions (PDFs), where each PDF represents a different lithology with distinct properties. We treat lithological units as clusters and use unsupervised K-means clustering to separate the petrophysical information into different units of distinct lithologies are not easily distinguishable. Through several synthetic examples, we demonstrate that the proposed framework leads FWI to elastic models that are superior to models obtained either without incorporating petrophysical information, or with a probabilistic penalty function based on a single PDF.

Key words: wavefield tomography, elastic, petrophysics, clustering

1 INTRODUCTION

The capability to deliver high resolution subsurface models has made full waveform inversion (FWI) a powerful process for building models of seismic properties underlying wave propagation (Tarantola, 1984; Pratt, 1999; Sirgue and Pratt, 2004; Virieux and Operto, 2009). Conventionally, FWI functions as a data-fitting iterative procedure, i.e., the minimization of an objective function based on the misfit between recorded and simulated data, where the latter is modeled using a two-way wave equation in order to consider all wave propagation mechanisms. At each iteration, the considered wave equation employs the current subsurface model, which is updated by perturbations derived from the gradient of an objective function, usually computed using the adjoint state method (Plessix, 2006).

However, the inverse problem constrained by wave-equation presents many difficulties. First, the inverse problem is ill-posed and highly non-linear in relation to the model parameters, essentially because seismic waves are oscillatory in nature (Symes, 2008). Second, if the initial subsurface models are not kinematically accurate enough to predict data with an error lower than the dominant wavelenght, the data residual suffers from cycle skipping and the inversion leads to models corresponding to local minima.

Compared to implementations using the isotropic acoustic wave equation (Tarantola, 1984; Pratt, 1990; Pratt et al., 1996), elastic FWI is more accurate in describing the physics of wave propagation in the earth (Tarantola, 1988; Pratt, 1990). An isotropic elastic wave equation can be written in terms of P- and S-wave velocities or Lamé parameters, λ and μ. However, the inverse problem is more ill-posed when estimating multiple parameters, as we introduce more degrees of freedom in the model space (Operto et al., 2013). Moreover, different elastic properties may be coupled and updates of one parameter leak into others, an effect known as crosstalk. In addition, the combinations of different components in the recovered model parameters can be lithologically implausible or impossible as model parameters are updated simultaneously, but independently.

One can reduce the number of possible solutions for the ill-posed FWI problem by adding a model regularization term to the objective function (Tikhonov and Arsenin, 1977; Tarantola, 2005). First- and second-order Tikhonov regularization can be used to generate smooth updates of the parameters, while total variation (TV) regularization preserves sharp fea-

tures in the models (Guitton, 2012). Additionally, one can also include a prior model term if additional information (well logs or geological information, for example) is available prior to the inversion. However, these techniques do not solve the problem of model updates that represent implausible lithologies. Moreover, as the model regularization terms operate in the spatial domain, they can guide inversion towards models that do not truly represent the subsurface, when accurate geological information about the entire studied region is not taken into account.

Petrophysical information provided by well logs describes the model only at sparse locations and therefore well log information cannot easily form the basis of model regularization in the space domain. As an alternative, one can apply petrophysical constraints in the model parameter space, in order to constrain the relationship between the model parameters during the inversion. This is the approach we take in this paper. Baumstein (2013) and Peters et al. (2015) use the projection onto convex sets (POCS method) to incorporate constraints into acoustic FWI, which leads to more stable simulations and geologically plausible structure. Manukyan et al. (2018) constrain the inversion by assuming that the elastic parameters have structural similarity and thus mitigate parameter tradeoff problems and the different spatial resolution of elastic parameters. Duan and Sava (2016) and Rocha and Sava (2018) use a logarithmic penalty function to constrain, respectively, elastic wavefield tomography and elastic reflection waveform inversion, assuming a general linear relationship between the model parameters.

Aragao and Sava (2018) apply probabilistic model constraints, by explicitly using petrophysical information during the inversion in order to recover models that honor both geophysical and petrophysical data. They achieve this goal by incorporating a penalty function, based on probability density functions (PDFs), into the FWI objective function. They demonstrate through elastic models that incorporating a probabilistic petrophysical penalty into the inversion objective function leads to reliable models that are superior to models obtained either without constraints or with approximate analytic constraints. The petrophysical data are treated in the model parameter space, while the geophysical data are treated in the physical space. The updates at different locations are driven by the data misfit term, but the petrophysical penalty prevents incorrect model updates in regions where data are inneficient and do not constrain sufficiently the subsurface properties.

However, Aragao and Sava (2018) use a single PDF to constrain the entire subsurface model, whereas the available petrophysical information may indicate the occurrence of many lithological units. On the one hand, this indication can be evident, such that different lithologies are well defined and isolated in the petrophysical space. In this case, the PDF represents a multimodal distribution. On the other hand, when building the petrophysical PDF from well log data, one may not be able to differentiate distinct lithological units because the range of elastic parameters for these lithologies overlap. Still, one can use classification methods, such as K-means clustering (MacQueen, 1967), to categorize the petrophysical information into different units. We refer to a lithological unit as a *cluster*, representing a group of objects that have similar or correlated properties (Hartigan, 1975) and to the techniques that perform the classification task as *clustering*.

In order to recover subsurface models that represent different lithologies with distinct properties through constrained FWI, we propose to apply modified petrophysical constraints that use different PDFs for each lithological unit consistent with the available petrophysical information. Before this probabilistic penalty term can be applied, we categorize each point of the updated model to the most representative cluster. For this purpose, we determine the distance from each point in the model to the center of the considered clusters, and then we use the closest cluster to build the petrophysical penalty.

We illustrate the performance of the proposed framework with three synthetic examples. We show the critical effect of the probabilistic penalty term when we use multiple PDFs compared to the cases of unconstrained FWI or of a single PDF used to describe the entire model. We also show that, even when we can not differentiate classes of distinct lithologies in crossplots of model parameters, we can group the petrophysical data into different clusters, representing specific lithological units. We also highlight how cycle skipping may be mitigated when we explicitly incorporate petrophysical constraints into the inversion.

2 ELASTIC WAVEFIELD TOMOGRAPHY

We use the isotropic elastic wave equation that assumes slowly varying Lamé parameters, such that their spatial gradients can be neglected,

$$\rho\ddot{\mathbf{u}} - \lambda[\nabla(\nabla\cdot\mathbf{u})] - \mu\left[\nabla\cdot\left(\nabla\mathbf{u}+\nabla\mathbf{u}^T\right)\right] = \mathbf{f}, \quad (1)$$

where $\mathbf{u}(e,\mathbf{x},t)$ is the elastic wavefield, $\mathbf{f}(e,\mathbf{x},t)$ is the source function, $\lambda(\mathbf{x})$ and $\mu(\mathbf{x})$ are the Lamé parameters, $\rho(\mathbf{x})$ is the density and e, $\mathbf{x}$ and t are, respectively, the experiment index, space coordinates and time.

The full waveform inverse problem is solved by minimizing an objective function $\mathcal{J}(\mathbf{u}_s,\lambda,\mu)$ consisting of three parts:

- a term that estimates the misfit between the predicted and observed data $\mathcal{J}_D(\mathbf{u}_s,\lambda,\mu)$,
- a model regularization term $\mathcal{J}_M(\lambda,\mu)$ that operates in the spatial domain and emphasize spatial correlation among the model parameters,
- a term that applies probabilistic model constraints using a petrophysical penalty term, $\mathcal{J}_P(\lambda,\mu)$.

Then,

$$\mathcal{J}(\mathbf{u}_s,\lambda,\mu) = \mathcal{J}_D(\mathbf{u}_s,\lambda,\mu) + \mathcal{J}_M(\lambda,\mu) + \mathcal{J}_P(\lambda,\mu). \quad (2)$$

We can define the misfit term $\mathcal{J}_D(\mathbf{u}_s,\lambda,\mu)$ as

$$\mathcal{J}_D(\mathbf{u}_s,\lambda,\mu)=\sum_e \frac{1}{2}\left\|\mathbf{W}_u(e,\mathbf{x},t)\mathbf{u}_s(e,\mathbf{x},t)-\mathbf{d}_{obs}(e,\mathbf{x},t)\right\|^2, \tag{3}$$

where $\mathbf{d}_{obs}(e,\mathbf{x},t)$ are the observed data and $\mathbf{W}_u(e,\mathbf{x},t)$ are weights that restrict the source wavefield $\mathbf{u}_s(e,\mathbf{x},t)$ to the known receiver locations.

We use a gradient-based method (Tarantola, 1988) to update the model, and we compute the gradient of $\mathcal{J}_D$ with respect to λ and μ using the adjoint-state method (Plessix, 2006), which consists of four steps:

(i) Calculate the seismic wavefield $\mathbf{u}_s(e,\mathbf{x},t)$, such that $\mathbf{L}\mathbf{u}_s-\mathbf{f}_s=\mathbf{0}$, where $\mathbf{f}_s$ and $\mathbf{L}$ represent, respectively, the source function and the linear elastic wave operator computed from equation 1.

(ii) Calculate the adjoint source $\mathbf{g}_s=\partial\mathcal{J}_D/\partial\mathbf{u}_s$, which uses the misfit between the observed and simulated data, according to equation 3.

(iii) Calculate the adjoint wavefield $\mathbf{a}_s(e,\mathbf{x},t)=\mathbf{L}^T\mathbf{g}_s$, where $\mathbf{L}^T$ is the adjoint elastic wave operator.

(iv) Calculate the gradient of $\mathcal{J}_D$ with respect to the parameters λ and μ by

$$\begin{bmatrix}\dfrac{\partial\mathcal{J}_D}{\partial\lambda}\\ \dfrac{\partial\mathcal{J}_D}{\partial\mu}\end{bmatrix}=\sum_e\begin{bmatrix}-\nabla(\nabla\cdot\mathbf{u})\star\mathbf{a}_s\\ -\nabla\left(\nabla\mathbf{u}+\nabla\mathbf{u}^T\right)\star\mathbf{a}_s\end{bmatrix}, \tag{4}$$

where the symbol $\star$ represents zero-lag time crosscorrelation.

Incorporating a model regularization term into the objective function reduces the number of possible solutions for the ill-posed FWI problem (Tikhonov and Arsenin, 1977; Tarantola, 2005; Asnaashari et al., 2013; Jun et al., 2017; Zhang et al., 2018). One can perform regularization by adding the first-order Tikhonov regularization term $\mathcal{J}_M^{Tik}(\lambda,\mu)$ into the objective function in order to provide smoothness to the estimated models, such that

$$\mathcal{J}_M^{Tik}(\lambda,\mu)=\frac{1}{2}\|\nabla\lambda(\mathbf{x})\|^2+\frac{1}{2}\|\nabla\mu(\mathbf{x})\|^2, \tag{5}$$

where ∇ represents the first-order spatial derivative operator. For a second-order Tikhonov regularization, one can replace ∇ by the second-order Laplacian operator. An alternative regularization approach uses total variation (TV) by adding another term $\mathcal{J}_M^{TV}(\lambda,\mu)$ to the objective function in order to preserve sharp interfaces in the recovered models (Guitton, 2012; Zhang et al., 2019),

$$\mathcal{J}_M^{TV}(\lambda,\mu)=\int_\Omega\sqrt{\epsilon_\lambda^2+\|\nabla\lambda(\mathbf{x})\|^2}\,d\mathbf{x}+\int_\Omega\sqrt{\epsilon_\mu^2+\|\nabla\mu(\mathbf{x})\|^2}\,d\mathbf{x}, \tag{6}$$

where Ω represents the subsurface spatial domain, and ϵ_λ and ϵ_μ are predefined parameters that make equation 6 differentiable.

One can also include a prior model regularization term $\mathcal{J}_M^{prior}(\lambda,\mu)$ into the objective function (Tarantola, 2005; Asnaashari et al., 2013; Zhang et al., 2018), such that

$$\mathcal{J}_M^{prior}(\lambda,\mu)=\frac{1}{2}\left\|\,\mathbf{W}_\lambda(\mathbf{x})(\lambda(\mathbf{x})-\bar{\lambda}(\mathbf{x}))\right\|^2+\frac{1}{2}\left\|\mathbf{W}_\mu(\mathbf{x})(\mu(\mathbf{x})-\bar{\mu}(\mathbf{x}))\right\|^2, \tag{7}$$

where $\bar{\lambda}(\mathbf{x})$ and $\bar{\mu}(\mathbf{x})$ are prior models. $\mathbf{W}_\lambda(\mathbf{x})$ and $\mathbf{W}_\mu(\mathbf{x})$ are weighting operators in the model space, and represent the inverse square-root of the model covariance matrices. Prior information can be based on knowledge about the model geometry, e.g., smoothness or prior geological information.

When one considers the objective function for the elastic FWI problem as $\mathcal{J}_D+\mathcal{J}_M$, the components of the model parameters λ and μ are updated independently. Therefore, they can become physically inconsistent with each other, leading to an infeasible geologic representation of the subsurface. Additionally, the terms in equations 5, 6 and 7 operate in the spatial domain, therefore, they can cause the elastic inversion to deliver unphysical and artifact-contaminated models.

Alternatively, if we have access to petrophysical information obtained, for example, from well logs, we can impose a penalty term that operate in the model parameter space. Aragao and Sava (2018) include a petrophysical penalty term $\mathcal{J}_P(\lambda,\mu)$, built using a probabilistic approach, into the objective function. From available well log data, they create a probability density function (PDF), whose cells of indices (i,j) represent the probability of occurance of the model components $(\lambda(i),\mu(j))$. For any given model with components $(\lambda_\mathbf{x},\mu_\mathbf{x})$ they evaluate in the model parameter space the distance to a point in the model space by

$$d_x(i,j)=\sqrt{\left[\lambda(i)-\lambda_\mathbf{x}\right]^2+\left[\mu(j)-\mu_\mathbf{x}\right]^2}, \tag{8}$$

for $i=1\cdots N_\lambda$ and $j=1\cdots N_\mu$ where N_λ and N_μ are the numbers of discrete intervals defined for parameters λ and μ, respectively. They calculate a weighted distance from $(\lambda_\mathbf{x},\mu_\mathbf{x})$ to the whole PDF by using the probability of each cell $P(i,j)$:

$$D_\mathbf{x}=\sum_{i,j}\frac{P(i,j)}{d_x(i,j)}, \tag{9}$$

and then define the petrophysical penalty term $\mathcal{J}_P$ as

$$\mathcal{J}_P(\lambda,\mu)=\eta\sum_\mathbf{x}\frac{1}{D_\mathbf{x}}, \tag{10}$$

where η is a scalar parameter that determines the strength of $\mathcal{J}_P(\lambda,\mu)$. The gradient of $\mathcal{J}_P$ with respect to the model parameters λ and μ is

$$\begin{bmatrix}\dfrac{\partial\mathcal{J}_P}{\partial\lambda}\\ \dfrac{\partial\mathcal{J}_P}{\partial\mu}\end{bmatrix}=-\eta\begin{bmatrix}\left(\sum_{i,j}\dfrac{P(i,j)\left[\lambda(i)-\lambda_x\right]}{\left[d_x(i,j)\right]^3}\right)\sum_{i,j}\dfrac{1}{D_\mathbf{x}^2}\\ \left(\sum_{i,j}\dfrac{P(i,j)\left[\mu(j)-\mu_x\right]}{\left[d_x(i,j)\right]^3}\right)\sum_{i,j}\dfrac{1}{D_\mathbf{x}^2}\end{bmatrix}. \tag{11}$$

By incorporating $\mathcal{J}_P(\lambda,\mu)$ into the FWI object function, the

elastic parameters are linked during the inversion. Aragao and Sava (2018) also show that, by using the term $\mathcal{J}_P$, one can:

- reduce the non-linearity and ill-posed nature of FWI,
- avoid geologically implausible earth models,
- mitigate artifacts produced by interparameter crosstalk,
- maintain robust and realistic petrophysical relationships between the model parameters.

However, the method proposed by Aragao and Sava (2018) uses a petrophysical penalty term based on a single PDF to represent the entire model, i.e., a single PDF to describe many lithologies. However, the occurance of many lithological units may be indicated by the available petrophysical information, with different lithologies well defined and isolated in the petrophysical space. In this case, we propose to use a petrophysical penalty term based on multiple PDFs, each representing a different lithological unit. In order to do so, we categorize each given model with components $(\lambda_{\mathbf{x}}, \mu_{\mathbf{x}})$ to a specific cluster, by calculating the distance between $(\lambda_{\mathbf{x}}, \mu_{\mathbf{x}})$ and the center of each PDF. Then, we use the closest PDF to build the petrophysical penalty related to $(\lambda_{\mathbf{x}}, \mu_{\mathbf{x}})$.

Nevertheless, the occurance of different lithologies may not be evident in the available petrophysical data. In this case, one may not be able to differentiate distinct lithological units in the PDF. One solution is to apply K-means clustering to separate the petrophysical information into different PDFs, where each PDF represents a lithological unit. K-means clustering consists of five steps (MacQueen, 1967; Hartigan, 1975):

(i) define the number of clusters k,
(ii) initiate the algorithm with randomly selected center for each cluster,
(iii) assign each point of the available petrophysical data to the closest center,
(iv) update the center of each cluster as the mean of the points that were assigned to that cluster,
(v) repeat steps 3 and 4 until no points are re-assigned, i.e., there is no change in center values.

After defining the clusters, we build a different PDF for each cluster and then follow the same approach as discussed above.

3 EXAMPLES

We illustrate our EFWI method with three examples. For all examples, we compare inversions using only the data misfit term $\mathcal{J}_D$, the data misfit with the petrophysical penalty term based on a single PDF $\mathcal{J}_D + \mathcal{J}_{P1}$, and the data misfit with the petrophysical penalty term based on multiple PDFs $\mathcal{J}_D + \mathcal{J}_{P2}$.

For the first example, different lithologies are well defined and isolated in the petrophysical space. We aim to show how the probabilistic penalty term based on multiple PDFs guides the inversion toward more accurate models compared to the cases of unconstrained FWI or FWI using a probabilistic penalty term based on a single PDF.

The second example uses the same setup as for the first example, but a different initial model. In particular, we show how the petrophysical penalty based on multiple PDFs alleviates the lack of low wavenumbers in the initial model and, as a consequence, helps mitigating cycle skipping.

In the third example, we use a different correct model, where the available petrophysical information do not indicate the occurrence of separate lithological units. We use the K-means clustering technique to categorize the petrophysical information into clusters and to create different PDFs for the petrophysical penalty term. We demonstrate that K-means clustering is an effective classification method to separate the petrophysical data into several lithological units.

3.1 Clustered Marmousi 2 model

The first and second synthetic examples use a portion of the Marmousi 2 model (Martin et al., 2002). We modify the model to represent well-defined and distinct lithologies. We consider that prior petrophysical data from three well logs are available and we use this information to build the petrophysical model penalty. Figure 1 shows the correct λ and μ models, the well locations and the model parameters crossplot. We simulate a multicomponent seismic survey with a line of receivers and 20 evenly spaced sources located at the surface ($z = 0$ km). Figure 2 shows the crossplot of λ and μ models extracted at the wells located at $x = 1.1$, 2.0 and 3.0 km (Figure 1), which indicates the presence of three separate lithological units (Figure 2). We use a 7.5 Hz peak Ricker source wavelet and perform two different experiments using the same correct λ and μ models (Figure 1) and the same petrophysical information (Figure 2), but with different starting models.

3.1.1 Experiment A: clustered Marmousi 2 model with a 2D smooth initial model

For the first experiment, we use the initial λ and μ models shown in Figure 3. Figures 4(a) and 4(b) show, respectively, the vertical and horizontal components of the data residual before inversion. Notice that the data residual computed using the initial model (Figure 3) shows reflections, diving and head waves, thus highlighting the kinematic inaccuracy of the initial λ and μ models.

Figure 5 shows the recovered λ and μ models using the objective function $\mathcal{J}_D$. As we do not include petrophysical information during inversion, the μ model shows spurious artifacts due to the trade-off between the elastic parameters λ and μ. The recovered μ model also has more structural detail than the recovered λ model, as the S-wave wavelenght is shorter than that of P waves. Therefore, the different models recovered from unconstrained inversion have distinct spatial resolution. Additionally, the amplitude of the low velocity anomaly presented in the true λ model (Figure 1) at $x = 3$ km and $z = 0.5$ km is not well recovered during the inversion. Moreover, the crossplot shown in Figure 5 does not resemble that of the true model (Figure 1). Figures 6(a) and 6(b) show, respectively, the vertical and horizontal components of the residual

constructed using the recorded and the predicted data computed using the recovered models shown in Figure 5. The data residuals contain almost as many reflections, diving and head waves as in the case when we predict the data using the initial model (Figures 4(a) and 4(b)). Therefore, the recovered λ and μ models using the objective function $\mathcal{J}_D$ are kinematically inaccurate.

In order to perform the inversion using the petrophysical penalty term, we carry out two tests. For the first test, we use a single PDF to constrain the entire model, even though the crossplot of λ and μ provided by the well logs (Figure 2) indicates well defined separate lithologies. Therefore, the first test uses the objective function $\mathcal{J}_D + \mathcal{J}_{P1}$. Figure 7(a) shows the single PDF used in $\mathcal{J}_{P1}$, which corresponds to three clusters. For the second test, we use different PDFs to constrain the model, and the objective function uses $\mathcal{J}_D$ and $\mathcal{J}_{P2}$. Figure 7(b) shows the different PDFs used in $\mathcal{J}_{P2}$. The black dot inside each PDF in Figure 7(b) represents the center of the PDF and we use these points to decide which PDF is used to constrain each point in the model, i.e., decide to which cluster each point of the updated model is more likely to represent.

Figure 8 shows the recovered λ and μ models using the objective function $\mathcal{J}_D + \mathcal{J}_{P1}$. Different from the unconstrained inversion (Figure 5), the recovered λ and μ models have similar spatial resolution. Additionally, the inversion is able to recover the amplitude of the low velocity anomaly present in the true λ model (Figure 1) at $x = 3\,\text{km}$ and $z = 0.5\,\text{km}$ better than the inversion using the objective function $\mathcal{J}_D$ (Figure 5). However, the inverted models depicted in Figure 8 are noisy and do not show blocky structures as in the true models (Figure 1). In addition, the crossplot shown in Figure 8 shows that the inversion could not recover model with $\lambda > 11.5\,\text{GPa}$ and $4.0 < \mu < 5.0\,\text{GPa}$. Also, the majority of the points in the crossplot of Figure 8 are inside the distribution with $\lambda < 11.0\,\text{GPa}$ and $\mu < 4.5\,\text{GPa}$. Therefore, when petrophysical information indicates the presence of distinct lithologies, but we choose to use a penalty term based on a single PDF for the entire model, the cluster with larger probability attracts most model points during the inversion. Figures 9(a) and 9(b) show, respectively, the vertical and horizontal components of the data residual computed using the recorded and the predicted data computed using the recovered models shown in Figure 8. Notice that the amplitudes of the data residuals are smaller compared to the case where the data are predicted using the models recovered with the unconstrained inversion (Figures 6(a) and 6(b)), which indicates that the recovered λ and μ models using the objective function $\mathcal{J}_D + \mathcal{J}_{P1}$ are more accurate kinematically compared to the models recovered using the objective function $\mathcal{J}_D$.

Figure 10 shows the recovered λ and μ models using the objective function $\mathcal{J}_D + \mathcal{J}_{P2}$. The contours correspond to three PDFs, each representing a cluster. The inverted λ and μ models shown in Figure 10 have comparable resolution and are closer to the true models compared with the cases when we use the objective functions $\mathcal{J}_D$ (Figure 5) or $\mathcal{J}_D + \mathcal{J}_{P1}$ (Figure 8). Additionally, the artifacts in the μ model caused by interparameter crosstalk are reduced. Notice that the crossplot built with the inverted models (Figure 10) shows points inside all three PDFs. Figures 11(a) and 11(b) show, respectively, the vertical and horizontal components of the data residual obtained using the recorded and the predicted data computed using the recovered models depicted in Figure 10. Note that the amplitudes of the data residuals are much smaller compared to the case where the data are predicted using the models obtained for the unconstrained objective function $\mathcal{J}_D$ (Figure 5) or for the objective function $\mathcal{J}_D + \mathcal{J}_{P1}$ (Figure 8).

3.1.2 Experiment B: clustered Marmousi 2 model with a 1D smooth initial model

For the second experiment, we use the initial λ and μ models shown in Figure 12. Note that the initial models are laterally homogeneous (1D), therefore FWI is subject to local minina in the absence of low frequencies. Figure 13 shows the comparison between the vertical component of the simulated data computed using the homogeneous model (Figure 12), depicted in the background as variable density, and the vertical component of observed data, shown as an overlay in black and white wiggles. The red and blue colors in the variable density plot represent, respectively, positive and negative amplitudes. Note that the time delay is larger than half a period, as the positive amplitudes of the simulated data occur after the peak of each wiggle trace. Figure 13 proves that cycle skipping occurs using the basic 1D models.

Figure 14 shows the recovered λ and μ models using the objective function $\mathcal{J}_D$. As for the experiment A (Figure 5), the recovered μ model has more structural detail than the recovered λ model (Figure 14). Additionally, the inversion is not able to recover the amplitude of the high velocity anomalies present at the bottom part of the true models (Figure 1), as the initial model is too far from the truth (Figure 12). Moreover, the crossplot shown in Figure 14 does not resemble the true model (Figure 1).

Figure 15 shows the recovered λ and μ models using the objective function $\mathcal{J}_D + \mathcal{J}_{P1}$. As for the experiment A (Figure 10), the contours represent a single PDF with three clusters, which is used in the penalty term. The inversion recovers better the amplitude of the high velocity anomalies in the models compared with the cases when we use the objective function $\mathcal{J}_D$ (Figure 14).

Figure 16 shows the recovered λ and μ models using the objective function $\mathcal{J}_D + \mathcal{J}_{P2}$. As for the experiment A (Figure 10), the contours correspond to three PDFs. The inversion recovers the high velocity model anomalies, so λ and μ Figure 16 are better recovered compared with the cases when we use the objective functions $\mathcal{J}_D$ (Figure 14) and $\mathcal{J}_D + \mathcal{J}_{P1}$ (Figure 15). The models shown in Figure 16 prove that, in the absent of low frequencies, the petrophysical penalty term based on multiple PDFs helps to avoid cycle skipping and leads inversion to converge towards models (Figure 16) close to the true models (Figure 1).

3.2 Marmousi 2 model with clustering

The third synthetic example uses a portion of the Marmousi 2 model (Martin et al., 2002), but we do not modify the models to represent well-defined and distinct lithologies. Figure 17 shows the correct λ and μ models, the well locations and the crossplot of the model parameters. For this example, we simulate a multicomponent ocean bottom seismic survey (OBS) with a line of receiver at the water bottom ($z = 0.17$ km) and 20 evenly spaced pressure sources located at $z = 0.05$ km. We use a 7.5 Hz peak Ricker source wavelet. Figure 18 shows the initial λ and μ models and the corresponding crossplot in the model parameter space.

Figure 19 shows the recovered λ and μ models using the objective function $\mathcal{J}_D$. Once again, one can notice that the inverted μ model has higher spatial resolution. Also, the μ model has spurious artifacts due to parameter crosstalk. The main structures in the initial λ model (Figure 18) conform to the true λ model (Figure 17) and are not preserved during the unconstrained inversion (Figure 19). Additionally, the crossplot shown in Figure 19 does not resemble that of the true model (Figure 17) because the updates are not restricted to a feasible petrophysical region.

For the inversions using petrophysical information, we use the crossplot shown in Figure 20, which corresponds to the λ and μ values extracted at the wells located at $x = 1.1$, 2.0 and 3.0 km (Figure 17). Note that this crossplot does not describe distinct lithologies. Therefore, one can choose to use a single PDF (Figure 21(a)) to constrain the entire model using the objective function $\mathcal{J}_D + \mathcal{J}_{P1}$. However, we use K-means clustering to separate the petrophysical information depicted in Figure 20 into multiple PDFs and then employ the objective function $\mathcal{J}_D + \mathcal{J}_{P2}$. A crucial step for K-means clustering technique is defining the number of clusters. One can do it by analyzing and interpreting the well logs and identifying different classes of lithology, such as, shale, sandstone, limestone and salt. For this example, we define four clusters (Figure 21(b)), consistent with this portion of the Marmousi 2 model that we use has four lithologies.

Figure 22 shows the recovered λ and μ models using the objective function $\mathcal{J}_D + \mathcal{J}_{P1}$. Note that both λ and μ models have resolution and are better recovered compared with the cases when we use the objective functions $\mathcal{J}_D$ (Figure 19). Additionally, the μ model artifacts caused by interparameter crosstalk are attenuated. However, the inverted λ model depicted in Figure 22 is noisy and does not show blocky structures above $x = 1.0$ km.

Figure 23 shows the recovered λ and μ models using the objective function $\mathcal{J}_D + \mathcal{J}_{P2}$. The contours correspond to the four PDFs used to define the petrophysical probabilistic penalty term. Both λ and μ inverted models depicted in Figure 23 show high resolution and are closer to the true models compared with the cases when we use the objective functions $\mathcal{J}_D$ (Figure 19) or $\mathcal{J}_D + \mathcal{J}_{P1}$ (Figure 22). The μ model shown in Figure 23 shows reduced crosstalk, and the λ model has blocky structure that conform to the true models (Figure 17).

4 CONCLUSIONS

We develop a strategy for incorporating petrophysical information into elastic full waveform inversion through a penalty term based on multiple probability density functions. We demonstrate that imposing this kind of penalty improves the quality of the recovered subsurface models and better represents feasible lithologies. Through several synthetic examples, we verify the effectiveness of the proposed method in leading inversion towards the true models. We also show that the quality of the recovered models can be compromised when the penalty function is based on a single probability density function, while petrophysical information indicates the occurrence of different lithologies in the subsurface. We demonstrate that K-means clustering can be used to separate the petrophysical data into different lithological units in order to compute multiple corresponding probability density functions. In addition, we prove that explicitly incorporating petrophysical information into the FWI framework alleviates cycle skipping and mitigates the crosstalk between elastic parameters.

5 ACKNOWLEDGMENTS

We acknowledge CAPES (Brazilian Federal Agency for Support and Evaluation of Graduate Education) for financial support. We thank the sponsors of the Center for Wave Phenomena (CWP), whose support made this research possible. The synthetic examples in this paper uses the Madagascar open-source software package (Fomel et al., 2013) freely available from http://www.ahay.org.

REFERENCES

Aragao, O., and P. Sava, 2018, Elastic wavefield tomographywith probabilistic petrophysical model constraints: Manuscript submitted for publication.

Asnaashari, A., R. Brossier, S. Garambois, F. Audebert, P. Thore, and J. Virieux, 2013, Regularized seismic full waveform inversion with prior model information: Geophysics, **78**, R25–R36.

Baumstein, A., 2013, POCS-based geophysical constraints in multiparameter full wavefield inversion: Presented at the 75th European Association of Geoscientists and Engineers Conference and Exhibition.

Duan, Y., and P. Sava, 2016, Elastic wavefield tomography with physical model constraints: Geophysics, **81**, no. 6, R447–R456.

Fomel, S., P. Sava, I. Vlad, Y. Liu, and V. Bashkardin, 2013, Madagascar: open-source software project for multidimensional data analysis and reproducible computational experiments: Journal of Open Research Software, 8th ed., DOI: http://dx.doi.org/10.5334/jors.ag.

Guitton, A., 2012, Blocky regularization schemes for full-waveform inversion: Geophysical Prospecting, **60**, 870–884.

Hartigan, J. A., 1975, Clustering algorithms: John Wiley & Sons.

Jun, H., J. Kwon, C. Shin, H. Zhou, and M. Cogan, 2017, Regularized Laplace-Fourier-domain full waveform inversion using a weighted L-2 objective function: Pure and Applied Geophysics, **174**, no. 3, 955–980.

MacQueen, J. B., 1967, Some methods for classification and analysis of multivariate observations: Presented at the 5th Berkeley Symposium on Mathematical Statistics and Probability.

Manukyan, E., H. Maurer, and A. Nuber, 2018, Improvements to elastic full-waveform inversion using cross-gradient constraints: Geophysics, **83**, no. 2, R105–R115.

Martin, G., K. Marfut, and S. Larsen, 2002, Marmousi-2, an updated model for the investigation of avo in structurally complex areas: Presented at the SEG Technical Program Expanded Abstracts.

Operto, S., Y. Gholami, V. Prieux, A. Ribodetti, R. Brossier, L. Metivier, and J. Virieux, 2013, A guided tour of multiparameter full-waveform inversion with multicomponent data: from theory to practice: The Leading Edge, **32**, no. 9, 1040–1054.

Peters, B., Z. Fang, B. R. Smithyman, and F. J. Herrmann, 2015, Regularizing waveform inversion by projections onto intersections of convex sets: Presented at the 77th European Association of Geoscientists and Engineers Conference and Exhibition.

Plessix, R. E., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: Geophysical Journal International, **167**, no. 2, 495–503.

Pratt, R. G., 1990, Inverse theory applied to multi-source cross-hole tomography: Geophysical Prospecting, **38**, 311–329.

———, 1999, Seismic waveform inversion in the frequency domain. Part 1: Theory and verification in a physical scale model: Geophysics, **64**, 888–901.

Pratt, R. G., Z. M. Williamson, and M. Warner, 1996, Two-dimensional velocity model from wide-angle seismic data by wavefield inversion: Geophysical Journal International, **124**, 323–340.

Rocha, D., and P. Sava, 2018, Elastic reflection waveform inversion with petrophysical model constrains: Presented at the SEG Technical Program Expanded Abstracts.

Sirgue, L., and R. Pratt, 2004, Efficient waveform inversion and imaging: a strategy for selecting temporal frequencies: Geophysics, **69**, 231–248.

Symes, W. W., 2008, Migration velocity analysis and waveform inversion: Geophysical Prospecting, **56**, no. 6, 765–790.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.

———, 1988, Theoretical background for the inversion of seismic waveforms including elasticity and attenuation: Pure and Applied Geophysics, **128**, 365–399.

———, 2005, Inverse Problem Theory and Methods for Model Parameter Estimation: Society for Industrial and Applied Mathematics.

Tikhonov, A., and V. Arsenin, 1977, Solution of ill-posed problems: Winston and Sons.

Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, no. 6, WCC1–WCC26.

Zhang, Z., T. Alkhalifah, E. Z. Naeini, and B. Sun, 2018, multiparameter elastic full waveform inversion with facies-based constraints: Geophysical Journal International, **213**, 2112–2127.

Zhang, Z., T. Alkhalifah, Z. Wu, Y. Liu, B. He, and J. Oh, 2019, Normalized nonzero-lag crosscorrelation elastic full-waveform inversion: Geophysics, **84**, no. 1, R1–R9.

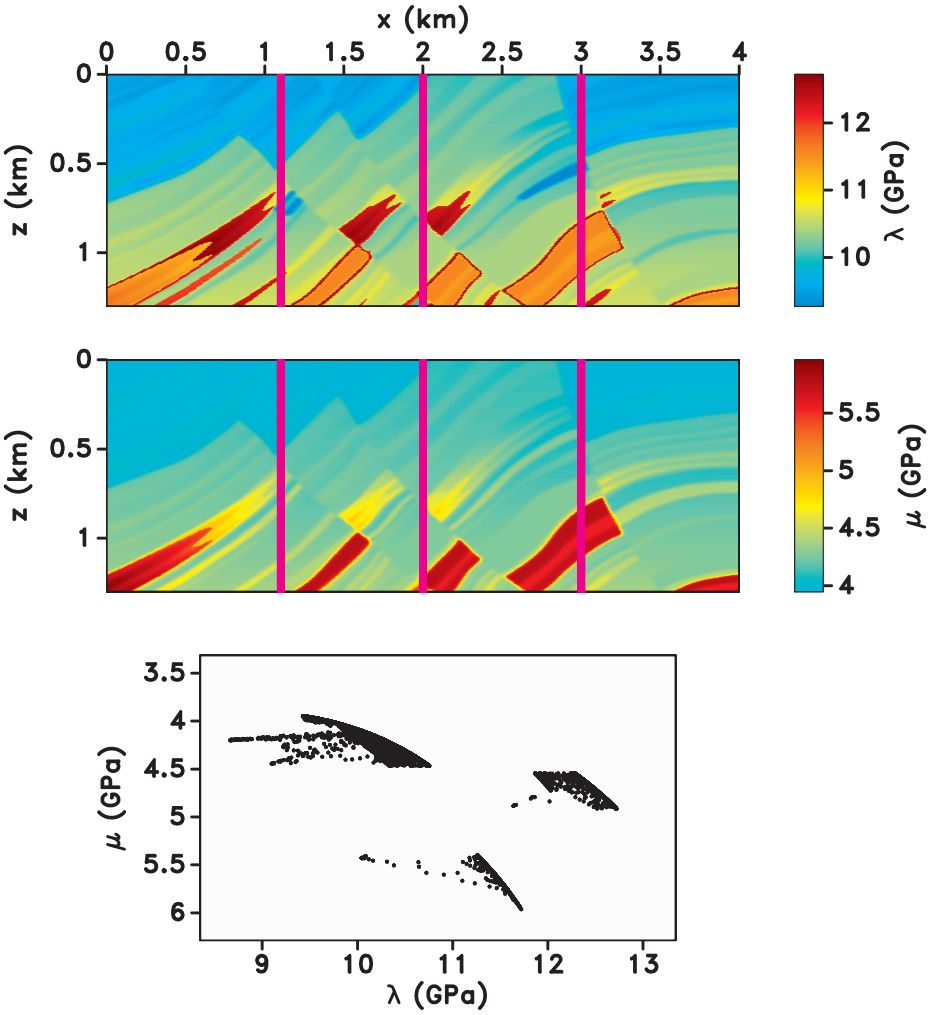


Figure 1. Clustered Marmousi 2 model: true λ and μ models and their crossplot for experiments A and B. The vertical lines correspond to well locations that provide information for the petrophysical penalty term.

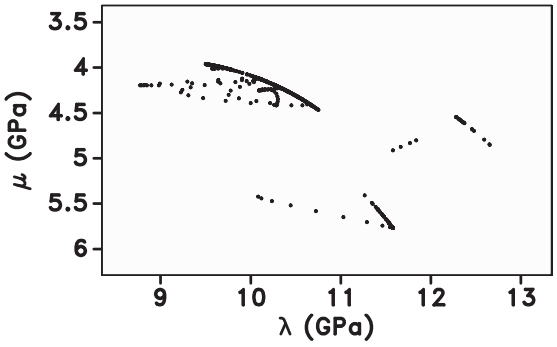


Figure 2. Crossplot of λ and μ provided by the well logs indicated in Figure 1. Note that different lithologies are well defined and isolated in the petrophysical space.

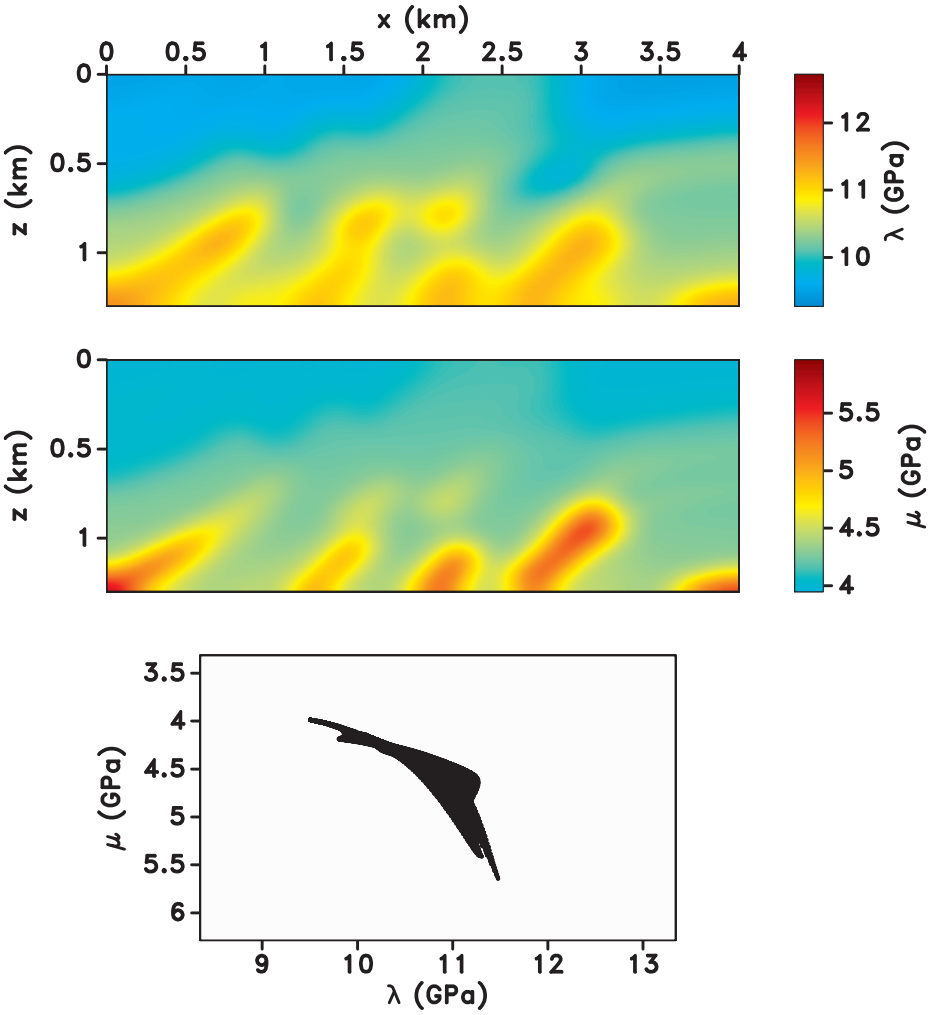


Figure 3. Experiment A: Initial λ and μ models and their crossplot in the $\lambda - \mu$ space. We obtain these initial models by smoothing the true model, which helps the inversion to converge to a global minimum.

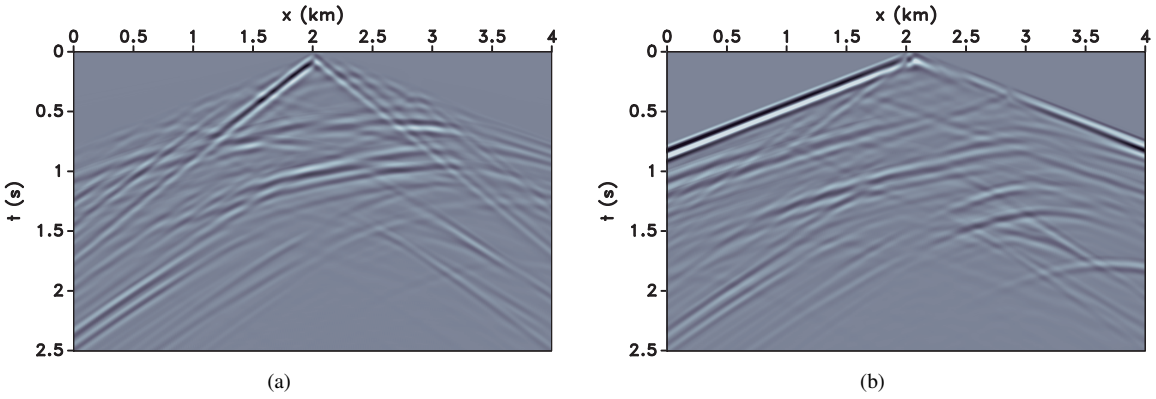


Figure 4. Experiment A: (a) vertical and (b) horizontal components of the data residual. The predicted data are computed using the initial models shown in Figure 3.

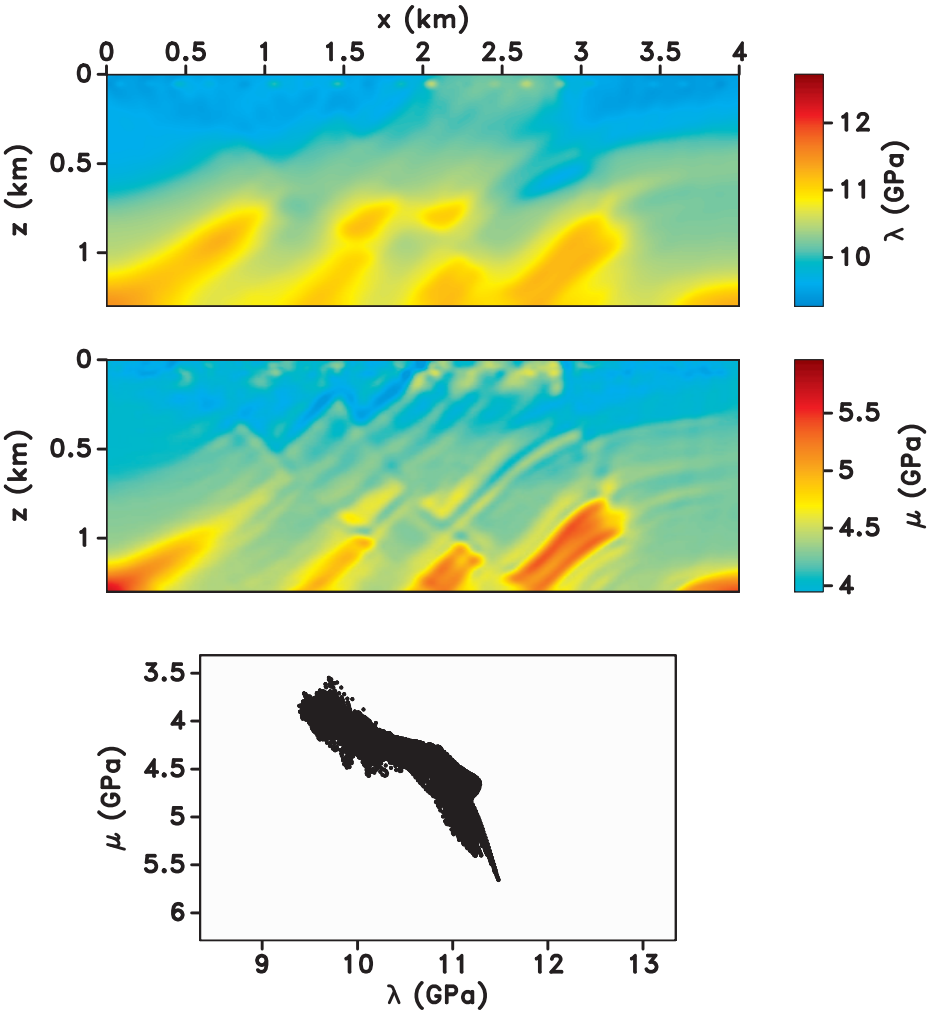


Figure 5. Experiment A: Recovered λ and μ models obtained using the objective function $\mathcal{J}_D$. Note that the μ model has more structural details than the λ model.

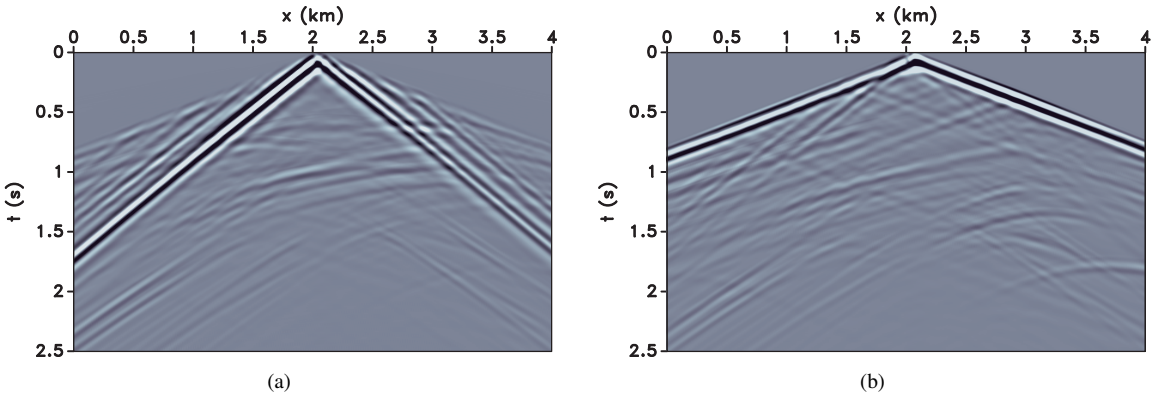
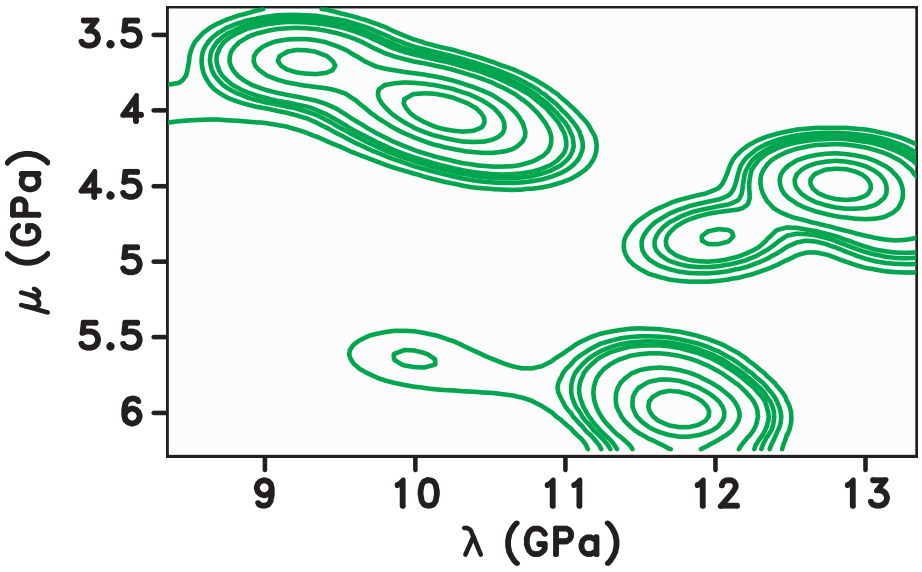
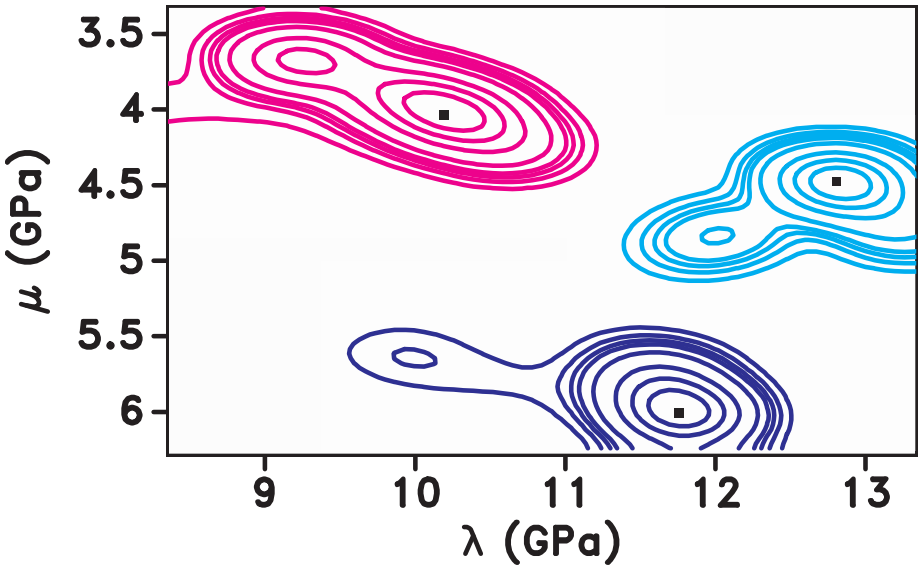


Figure 6. Experiment A: (a) vertical and (b) horizontal components of the data residual for the unconstrained inversion. The predicted data are computed using the recovered models shown in Figure 5.



(a)



(b)

Figure 7. (a) PDF associated with the term $\mathcal{J}_{P1}$ and (b) PDF associated with the term $\mathcal{J}_{P2}$. We estimate the PDFs in (a) and (b) from the well logs shown in Figure 2.

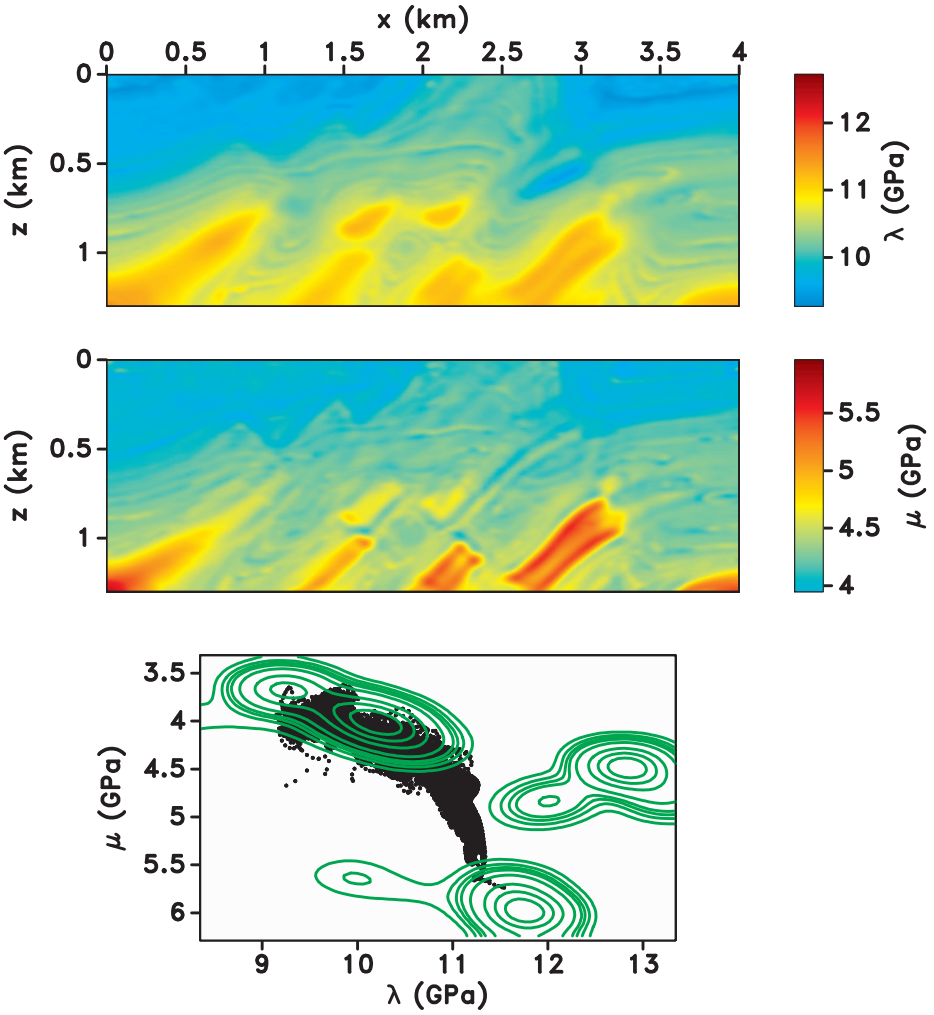


Figure 8. Experiment A: Recovered λ and μ models using the objective function $\mathcal{J}_D + \mathcal{J}_{P1}$. The green contours overlain on the crossplot illustrate the probabilistic penalty term based on a single PDF $\mathcal{J}_{P1}$.

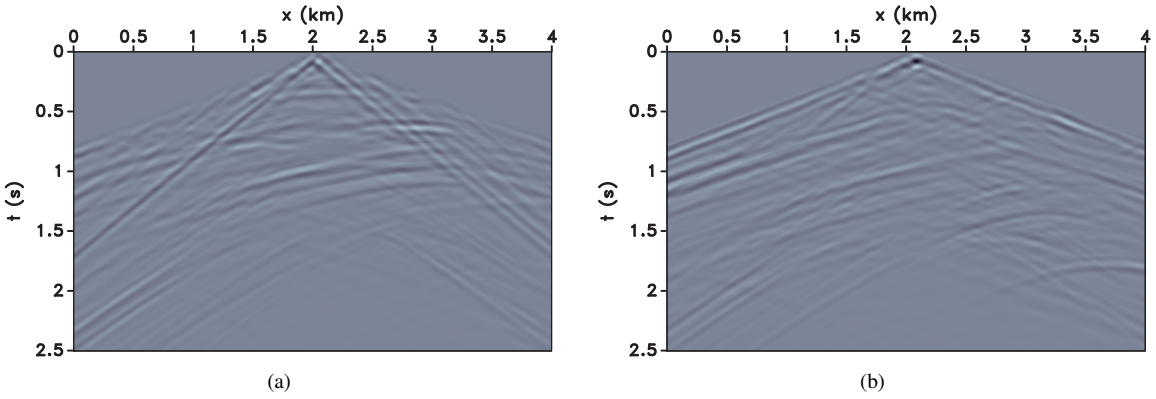


Figure 9. Experiment A: (a) vertical and (b) horizontal components of the data residual for the unconstrained inversion. The predicted data are computed using the recovered models shown in Figure 8.

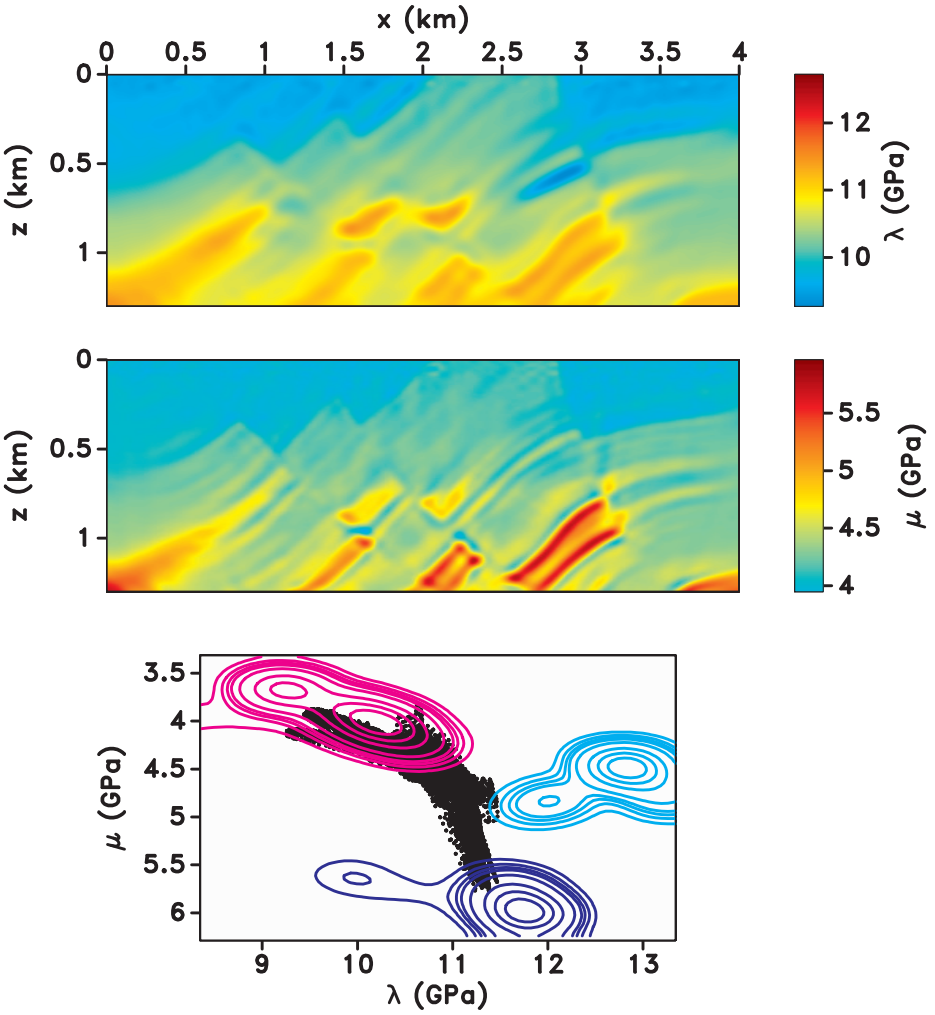


Figure 10. Experiment A: Recovered λ and μ models using the objective function $\mathcal{J}_D + \mathcal{J}_{P2}$. The contours with different colors overlain on the crossplot illustrate multiple PDFs used in the probabilistic penalty $\mathcal{J}_{P2}$. Note that three separate PDFs are used to constrain the model, although we display them together for simplicity.

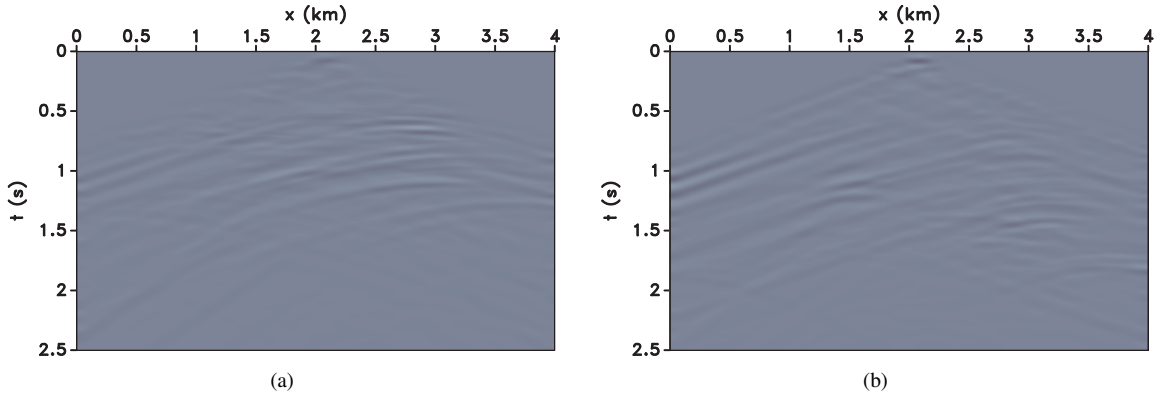


Figure 11. Experiment A: (a) vertical and (b) horizontal components of the data residual for the unconstrained inversion. The predicted data are computed using the recovered models shown in Figure 10.

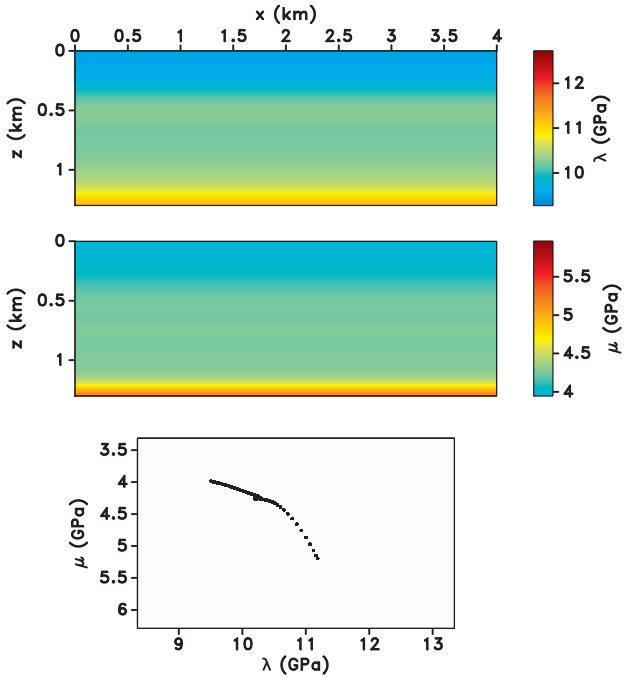


Figure 12. Experiment B: Initial λ and μ models and their crossplot in the $\lambda - \mu$ space. As these models are far from the true models (Figure 1), the inversion does not converge to a global minimum.

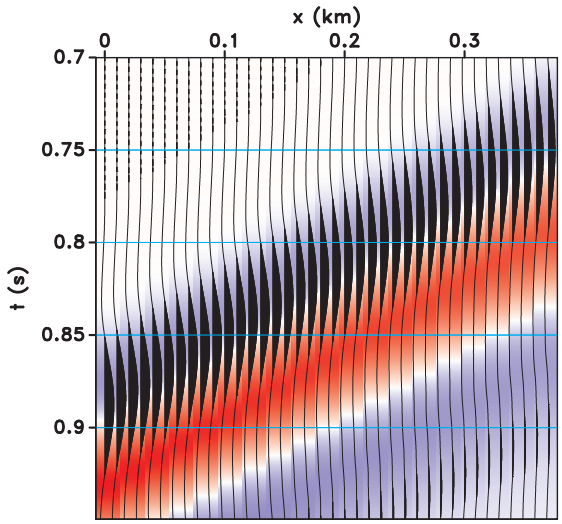


Figure 13. Experiment B: illustration of cycle skipping with the 1D starting model. The variable density plot in the background represents the vertical component of the simulated data computed using the initial 1D model (Figure 12) and the overlaying black and white wiggles are the vertical component of observed data.

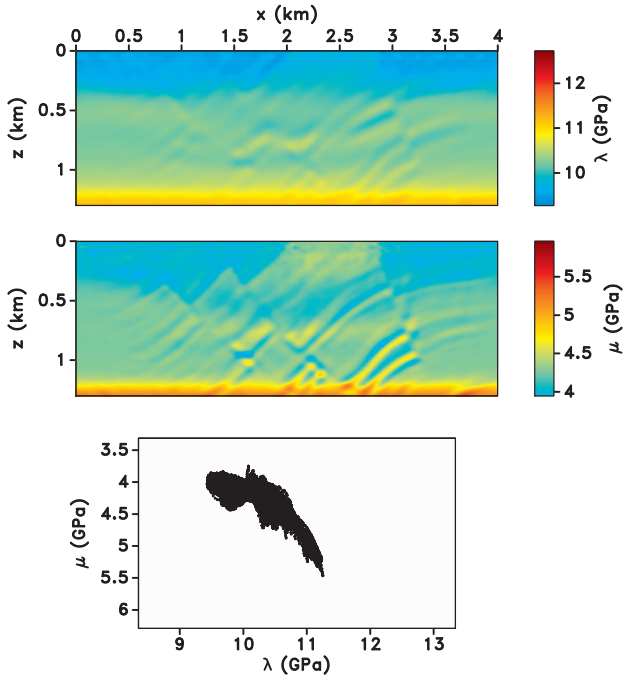


Figure 14. Experiment B: Recovered λ and μ models using the objective function $\mathcal{J}_D$. The inaccuracy of the initial models (Figure 12) leads FWI a local minimum.

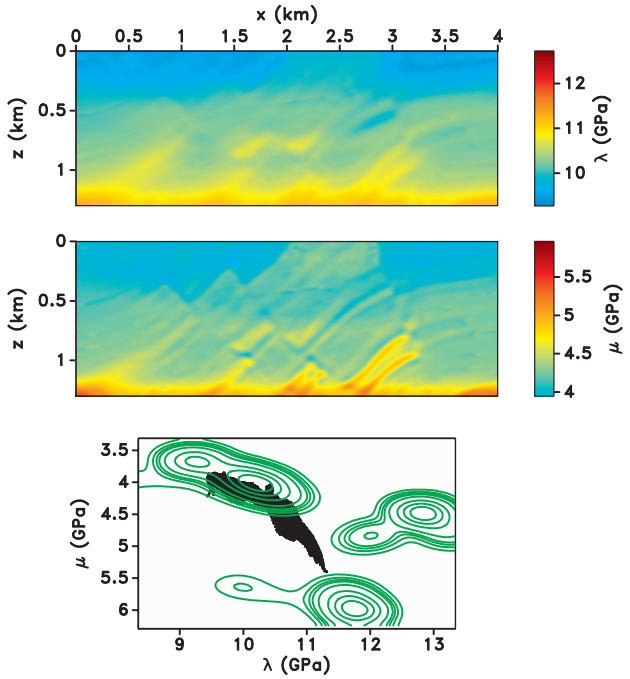


Figure 15. Experiment B: Recovered λ and μ models using the objective function $\mathcal{J}_D + \mathcal{J}_{P1}$. The green contours overlain on the crossplot illustrate the probabilistic penalty term based on a single PDF.

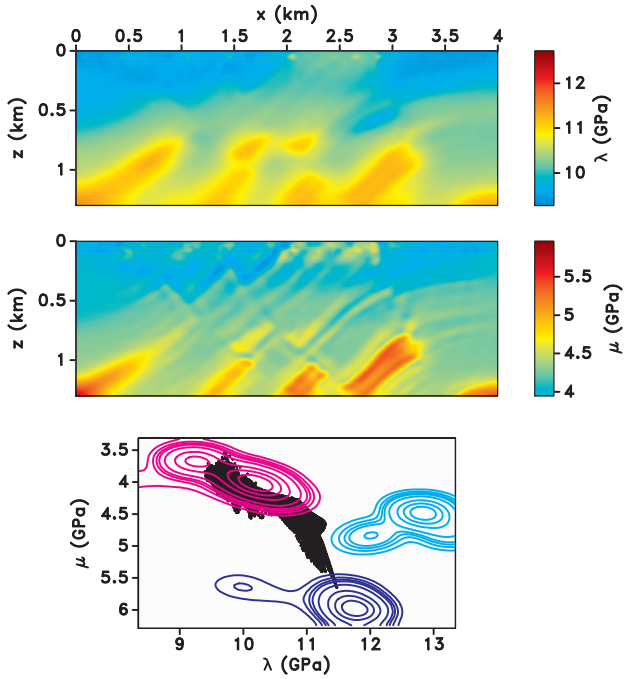


Figure 16. Experiment B: Recovered λ and μ models using the objective function $\mathcal{J}_D + \mathcal{J}_{P2}$. The contours with different colors overlain on the crossplot illustrate the multiple PDFs used in the probabilistic penalty.

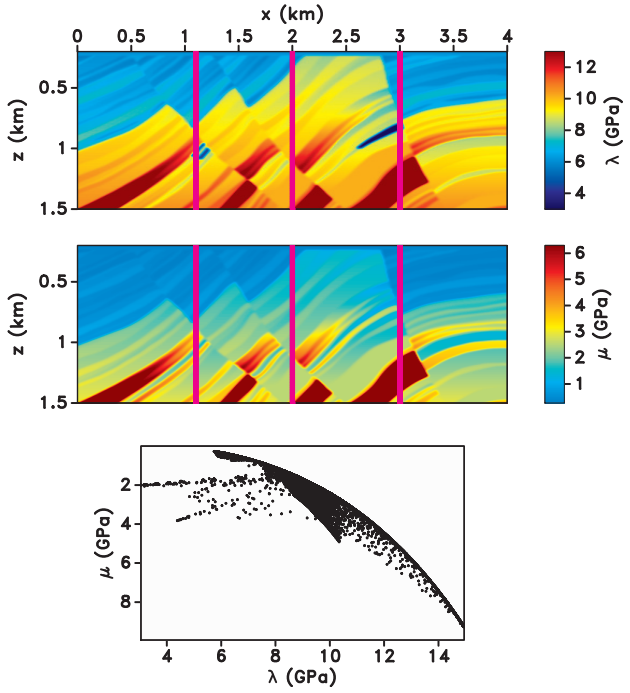


Figure 17. True λ and μ models and their crossplot for the inversion with clustering. The vertical lines correspond to well locations that provide information for the petrophysical penalty term.

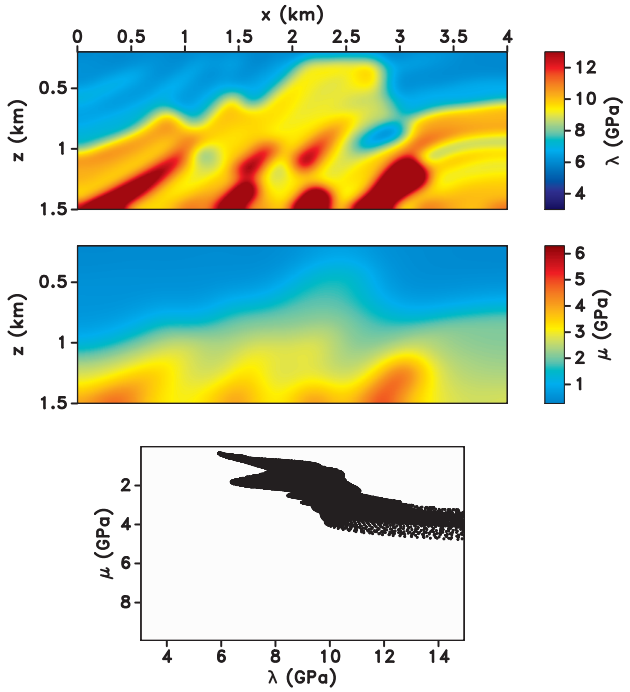


Figure 18. Initial λ and μ models and their crossplot in the $\lambda - \mu$ space for the inversion with clustering.

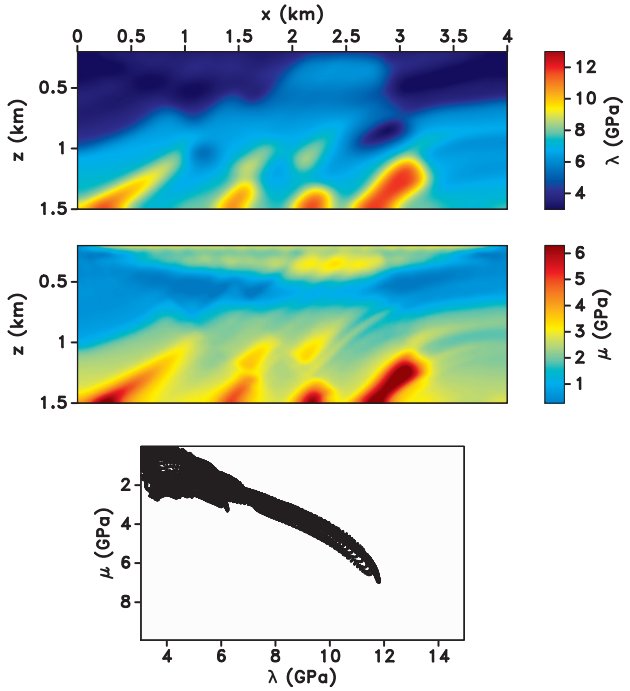


Figure 19. Recovered λ and μ models obtained using the objective function $\mathcal{J}_D$ for the inversion with clustering.

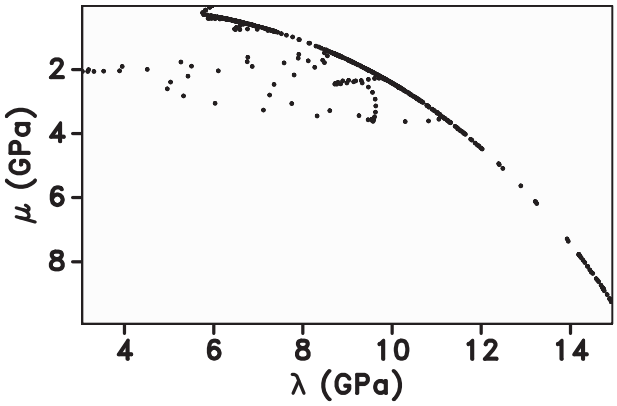


Figure 20. Crossplot of λ and μ provided by the well logs indicated in Figure 17. Note that different lithologies are not distinct in the petrophysical space.

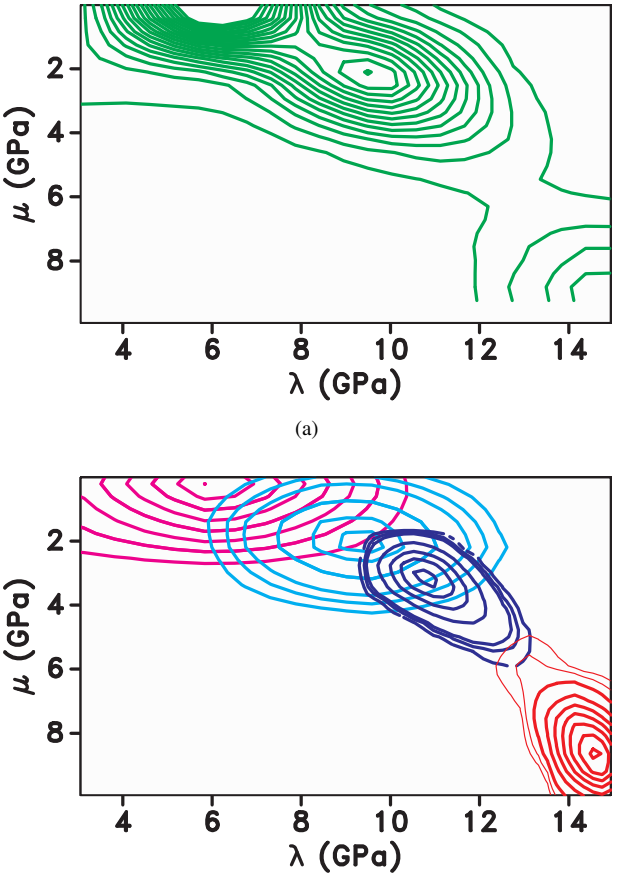


Figure 21. (a) PDF associated with the term $\mathcal{J}_{P1}$ and (b) PDF associated with the term $\mathcal{J}_{P2}$ for the inversion with clustering. We estimate the PDFs in (a) and (b) from the well logs shown in Figure 20.

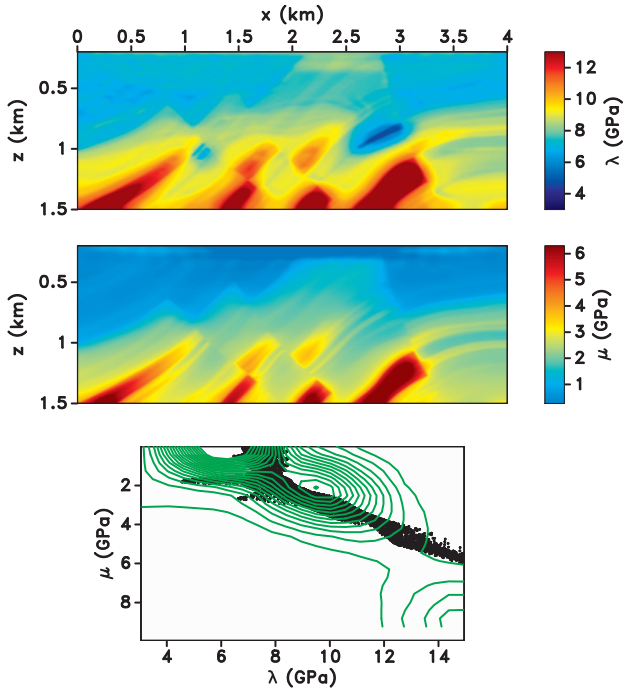


Figure 22. Recovered λ and μ models using the objective function $\mathcal{J}_D + \mathcal{J}_{P1}$ for the inversion with clustering. The green contours overlain on the crossplot illustrate the probabilistic penalty term based on a single PDF $\mathcal{J}_{P1}$.

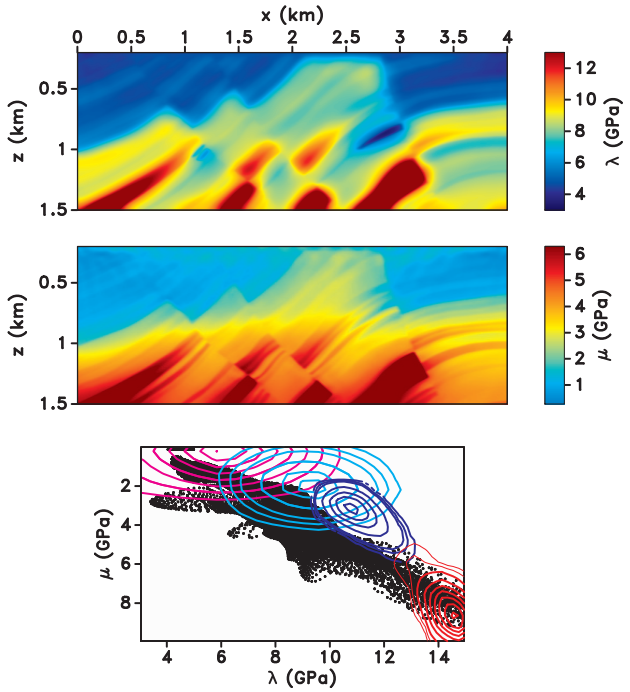


Figure 23. Experiment A: Recovered λ and μ models using the objective function $\mathcal{J}_D + \mathcal{J}_{P2}$ for the inversion with clustering. The contours with different colors overlain on the crossplot illustrate the multiple PDFs used in the probabilistic penalty $\mathcal{J}_{P2}$.