Probabilistic moveout analysis by time warping[a]

Yanadet Sripanich[1], Sergey Fomel[2], Jeannot Trampert[3], William Burnett[4], and Thomas Hess[2]
[1]Formerly Utrecht University; presently PTT Exploration and Production Public Company Limited
[2]The University of Texas at Austin
[3]Utrecht University
[4]ExxonMobil Upstream Research Company

ABSTRACT

Parameter estimation from reflection moveout analysis represents one of the most fundamental problems in subsurface model building. We propose an efficient moveout inversion method that is based on the process of automatic flattening of common-midpoint (CMP) gathers using local slopes. We show that as a byproduct of this flattening process, we can also estimate reflection traveltimes corresponding to the flattened CMP gathers. This traveltime information allows us to construct a highly overdetermined system and subsequently invert for moveout parameters including normal-moveout (NMO) velocities and quartic coefficients related to anisotropy. We utilize the 3D generalized moveout approximation (GMA) that can accurately capture the effects of complex anisotropy on reflection traveltimes as the basis for our moveout inversion. Due to the cheap forward traveltime computations by GMA, we employ a Monte Carlo inversion scheme for an improved handling of the non-linearity between reflection traveltimes and moveout parameters. This choice also allows us to set up a probabilistic inversion workflow within a Bayesian framework, where we can obtain the posterior probability distributions that contain valuable statistical information on estimated parameters such as uncertainty and correlations. We use synthetic and real-data examples including the data from the SEAM Phase II unconventional reservoir model to demonstrate the performance of our proposed method and discuss insights into the problem of moveout inversion gained from analyzing the posterior probability distributions. Our results suggest that the solutions to the problem of traveltime-only moveout inversion from 2D CMP gathers are relatively well-constrained by the data. However, parameter estimation from 3D CMP gathers associated with more moveout parameters and complex anisotropic models are generally non-unique and that there are trade-offs among inverted parameters, especially the quartic coefficients.

INTRODUCTION

Moveout analysis is arguably one of the most important steps in seismic processing, which leads to an expeditious construction of subsurface models. The classical work of Taner and Koehler (1969) represents the foundation for modern moveout analysis, where the reflection traveltimes of pure P waves are approximated by a Taylor expansion around zero offset. The degree of matching between the data and the computed traveltime can be judged based upon the semblance criterion (Neidell and Taner, 1971). Velocities that correspond to a higher degree of matching (semblance) are picked, and an accurate subsurface model can be inferred. In a 2D isotropic subsurface, only one parameter — the normal-moveout (NMO) velocity — needs to be estimated but more parameters are required when 3D data and effects from anisotropy are under consideration (Alkhalifah and Tsvankin, 1995; Grechka and Tsvankin, 1998a; Yilmaz, 2001; Tsvankin, 2012; Thomsen, 2014; Masmoudi et al., 2017; Oh and Alkhalifah, 2016). Aiming to improve the process, many researchers have proposed strategies such as refining the semblance criterion for higher resolution semblance peaks (Kirlin, 1992; Kirlin and Done, 2001; Abbad et al., 2009; Fomel, 2009; Luo and Hale, 2012; Chen et al., 2015; Wilson and Gross, 2017), taking into account amplitude variations with offset (AVO) (Sarkar et al., 2001, 2002; Yan and Tsvankin, 2008), and implementing inversions by maximizing semblance-based objective function to avoid too many costly parameter scans (Grechka and Tsvankin, 1999; Vasconcelos and Tsvankin, 2006). However, it can generally be said that the higher the number of parameters to scan, the lower the computational efficiency of semblance-based techniques.

Fomel (2007) proposes an alternative approach to this problem and formulated an automatic NMO correction of CMP gathers based on local slopes, which can directly and automatically be estimated from the gathers themselves (Fomel, 2002). This concept evolves from the original velocity-independent imaging strategies of Ottolini (1983) and early results of Wolf et al. (2004), which rely on the connection between the slopes of reflection traveltime surfaces and the time-domain processing parameters. As a result, the process of subsurface model building simply amounts to mapping from the automatically measured slopes to the corresponding NMO velocity. In the simple case of 2D hyperbolic reflection traveltimes, the NMO velocity v_{nmo} can be found from the slope p as follows (Fomel, 2007):

$$\frac{1}{v_{nmo}^2(t_0, x)} = \bar{W}(t_0, x) = p(t_0, x)\frac{t}{x} \ , \tag{1}$$

where t is the hyperbolic reflection traveltime, t_0 denotes the zero-offset traveltime, x is offset, and $p = dt/dx$. The slowness squared $\bar{W}(t_0, x)$ is obtained as an attribute volume in terms of t_0 and x. In conventional time processing workflow (Yilmaz, 2001), a profile of slowness squared $W(t_0)$ at each CMP location as opposed to a volume of $\bar{W}(t_0, x)$ is generally preferred, which leads to a research question on how one can best obtain $W(t_0)$ from $\bar{W}(t_0, x)$. Moreover, reflection traveltime approximations are often nonhyperbolic with more independent parameters needed to honor the

effects from anisotropy and heterogeneity. Stovas and Fomel (2016) present several relationships between local slopes/curvatures (d^2t/dx^2) and moveout parameters for different nonhyperbolic traveltime approximations. Nonetheless, the problem of using local slopes to achieve both moveout correction (event flattening) and parameter estimation (velocity analysis) automatically in this regime remains a challenging task.

Alternatively, Burnett and Fomel (2009) and Burnett (2011) propose to accomplish both processes by a combination of non-physical flattening and moveout inversion. The former is achieved by using predictive painting (Fomel, 2010), which automatically traces the CMP events according to the local structural information from slopes without relying on any prior assumption regarding their shape. This is notably different from directly mapping local slopes to moveout parameter as in equation 1. The traced events are subsequently *warped* (numerical nonstationary shifts) until they are flat and a seismic image can be obtained from stacking these flattened gathers. The term 'non-physical' is used here to emphasize the difference between the conventional physical flattening based on an assumption of some particular shape of traveltime surfaces (e.g., hyperbolic in equation 1 or any other nonhyperbolic forms extracted from the physics of wave propagation) and the warping approach, which flattens the gathers numerically. In view of subsurface model building, the byproduct of this process, the nonstationary shifts, can be stored and applied to a time attribute volume — thereby the name *time-warping* — which yields reflection traveltimes of the corresponding flattened CMP events. We can then invert for the best-fit parameters of any preferred moveout approximation from this knowledge on reflection traveltimes of all CMP events and obtain properties of the subsurface.

Many moveout approximations suitable for various settings exist in the literature and choosing which one to use depends largely on the data coverage, the types of subsurface media, the desired traveltime accuracy, and the computational efficiency — more parameters to scan is often more computationally inefficient. Conventionally, moveout approximations have the well-known hyperbolic expression, which is exact for homogeneous isotropic or elliptically anisotropic media, but is only approximately applicable to small-offset data in other cases. For long-offset data, the moveout become nonhyperbolic and require more parameters to capture the effects from possible anisotropy (Hake et al., 1984; Castle, 1994; Tsvankin and Thomsen, 1994; Alkhalifah and Tsvankin, 1995; Al-Dajani et al., 1998; Ursin and Stovas, 2006; Blias, 2009; Aleixo and Schleicher, 2010; Alkhalifah, 2011; Golikov and Stovas, 2012; Farra and Pšenčík, 2013; Farra et al., 2016). Fomel and Stovas (2010) propose the 2D highly accurate generalized moveout approximation (GMA), which requires five independent parameters. This approximation is named 'generalized' due to the fact that its functional form can reduce to several other known moveout approximations with different choices of parameters, and therefore provides a systematic view on the effects of various parameter choices on the approximation accuracy. Its 3D extension (3D GMA) is proposed by Sripanich et al. (2017) and requires a total of seventeen parameters to accurately describe reflection traveltimes in 3D layered anisotropic media.

In this study, we first show that using the time-warping workflow, we can con-

struct a highly overdetermined inverse problem for moveout parameters. We improve upon the work of Burnett and Fomel (2009) and Burnett (2011) that rely on second-order moveout approximations (NMO ellipse), and utilize in our approach instead the more accurate 3D GMA appropriate for 3D layered anisotropic models. Furthermore, we acknowledge the cheap forward traveltime computations with 3D GMA and propose to employ a global inversion scheme within a transdimensional (hierarchical Bayesian) framework (Malinverno and Briggs, 2004; Sambridge et al., 2006, 2013), instead of the previous linearized least-squares inversion by Burnett (2011). This allows us to obtain valuable statistical posterior probability distributions of all estimated moveout parameters and data uncertainty as opposed to only one 'best-fit' solution according to the least-squares minimization criterion. We provide synthetic and real-data examples including one from the SEAM Phase II unconventional reservoir model to demonstrate the performance and discuss important implications from using the proposed traveltime-only inversion approach for subsurface model building.

THEORY

Moveout inversion by time warping

In general, we can formulate the problem of moveout inversion in the context of time-warping as follows (Burnett and Fomel, 2009):

$$t^2 - t_0^2 = F(\mathbf{m}, \mathbf{x}) \,, \tag{2}$$

where $t^2 - t_0^2$ denotes the measured traveltime squared shifts from non-physical flattening, F is a linear or non-linear operator describing the geometry of the reflection traveltime surface, i.e, the moveout approximation of choice, $\mathbf{m}$ denotes the vector of pertaining moveout parameters associated with that choice of F, and $\mathbf{x}$ denotes the vector of offsets (x and y). Here, t denotes the recorded reflection traveltime and t_0 is its value at zero offset. For each event at t_0 in a CMP gather of N traces, we have N equations with only several unknowns ($\mathbf{m}$) depending on the choice of F. Therefore, this generally leads to a highly overdetermined system for estimating the associated moveout parameters $\mathbf{m}$. For example, in the case of 2D hyperbolic reflection traveltime, we have $F(\mathbf{m}, \mathbf{x}) = x^2 \, W(t_0)$ leading to

$$t^2 - t_0^2 = x^2 \, W(t_0) \,. \tag{3}$$

Therefore, this highly overdetermined system can be solved using, for instance, a least-squares inversion (Burnett and Fomel, 2009; Burnett, 2011). The number of parameters in $\mathbf{m}$ varies depending on the choice of F and the kind of considered data (2D or 3D). Nevertheless, in general, the number of traces in a CMP gather far exceeds the number of unknown parameters and the resulting system is almost always overdetermined. We note that in practice, the automatic traveltime picks from recorded data may include complications from various causes such as noise and errors from slope estimation, which may lead to a system of equations that is mathematically

inconsistent — no set of values for **m** satisfies the system exactly. In this study, we propose to use the Bayesian inversion framework to find the posterior distributions of possible solutions instead of seeking only one solution associated with least-squares misfit minimization criterion.

Time-warping workflow

Steps to implement the entire time-warping workflow can be summarized (Figures 1 and 2) as follows (Burnett and Fomel, 2009):

1. The inputs of the workflow include a CMP gather $d(t, x, y)$ and a time attribute volume as shown in Figures 1a and 1b. The latter is generated simply by having a similarly-sized volume as the CMP gather with each element storing the corresponding value of traveltime along the axis. After warping, this volume will store the reflection traveltimes from different CMP events.

2. Apply slope estimation on the CMP gather (Figure 1a) using tools such as plane-wave destruction filter (Fomel, 2002). The resulting slopes indicate the local structures and are used to automatically trace the CMP events in the predictive painting method (Fomel, 2010).

3. The picked events are then numerically warped until flattened by inverse interpolation. The warping process can be viewed as applying a nonstationary shifting filter. An application of this filter to $d(t, x, y)$ produces $d(t_0, x, y)$ in Figure 2a. Applying the same filter to the time attribute volume in Figure 1b, we obtain $t(t_0, x, y)$ in Figure 2b. Here, we restrict our application of warping to only the parts with CMP events between time 1 and 2 s.

4. In the considered time range (1–2 s), we can compute the traveltime shift volume by the element-wise subtraction of $t^2(t_0, x, y) - t_0^2$.

5. The resulting traveltime shifts will be used in the left-hand side of equation 2 for the moveout parameter inversion.

The above five steps are carried out in the same way, regardless of the choice of moveout approximation. These steps could also be applied to migrated common-image-point gathers, or many other gather or domain types. In this study, we focus on its application to the problem of moveout inversion, which is dependent upon the choice of moveout approximation F and the inversion method. Different moveout approximations represent different shapes of reflection traveltime surfaces and involve a different number of moveout parameters to be fitted. An appropriate inversion method must also be chosen in accordance with the choice of moveout approximation. Whereas Burnett and Fomel (2009) solved for the three parameters of NMO ellipse moveout model using least-squares inversion, here we use a Monte Carlo inversion to infer the many parameters of the more comprehensive 3D GMA model.
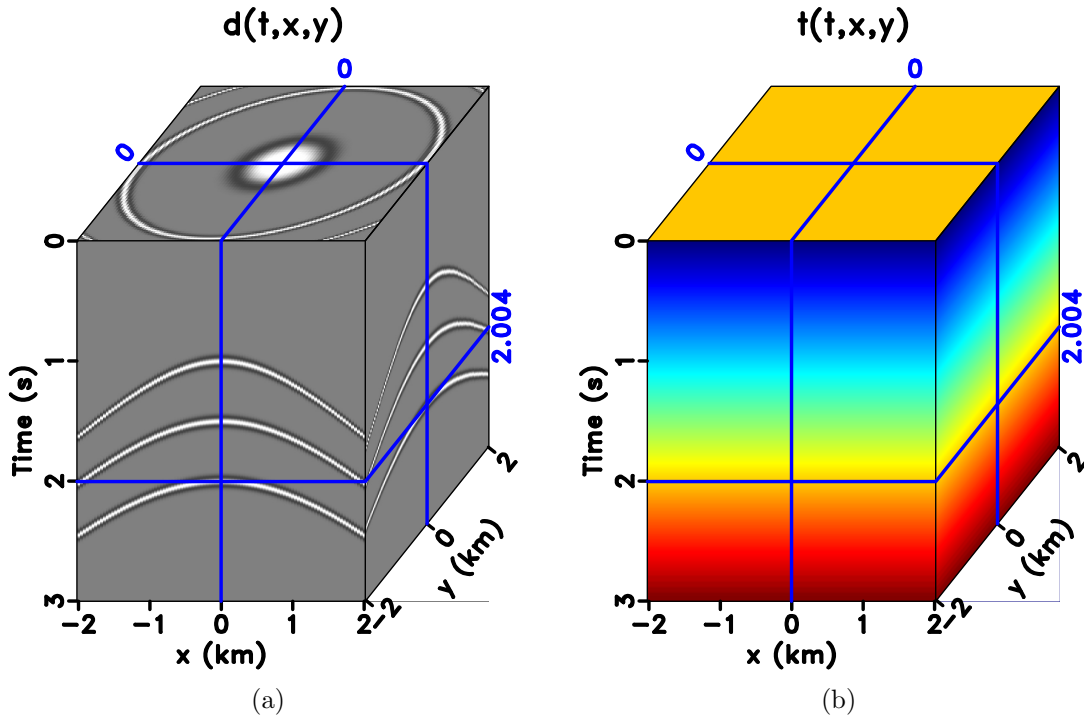
Figure 1: Original volumes of (a) CMP data $d(t, x, y)$ and (b) time attribute .
mcmc/ cmpcube,timecube

3D generalized nonhyperboloidal moveout approximation

Using similar notations as before, we express the reflection traveltime of each CMP event in terms of offset x and y in a given acquisition coordinate frame as $t(x, y)$. The 3D generalized nonhyperboloidal moveout approximation (3D GMA) can be expressed as (Sripanich et al., 2017):

$$t^2(x, y) \approx t_0^2 + W(x, y) + \frac{A(x, y)}{t_0^2 + B(x, y) + \sqrt{t_0^4 + 2t_0^2 B(x, y) + C(x, y)}} \ , \qquad (4)$$

where

$$\begin{aligned}
W(x, y) &= W_1 x^2 + W_2 xy + W_3 y^2 \ , \\
A(x, y) &= A_1 x^4 + A_2 x^3 y + A_3 x^2 y^2 + A_4 xy^3 + A_5 y^4 \ , \\
B(x, y) &= B_1 x^2 + B_2 xy + B_3 y^2 \ , \\
C(x, y) &= C_1 x^4 + C_2 x^3 y + C_3 x^2 y^2 + C_4 xy^3 + C_5 y^4 \ , \qquad (5)
\end{aligned}$$

and t_0 again denotes the reflection traveltime at zero offset. The total number of independent parameters in equation 4 is seventeen including t_0, W_i, A_i, B_i, and C_i. In the case of $x = 0$ or $y = 0$, equation 4 reduces to the generalized nonhyperbolic moveout approximation of Fomel and Stovas (2010). Sripanich et al. (2017) shows that
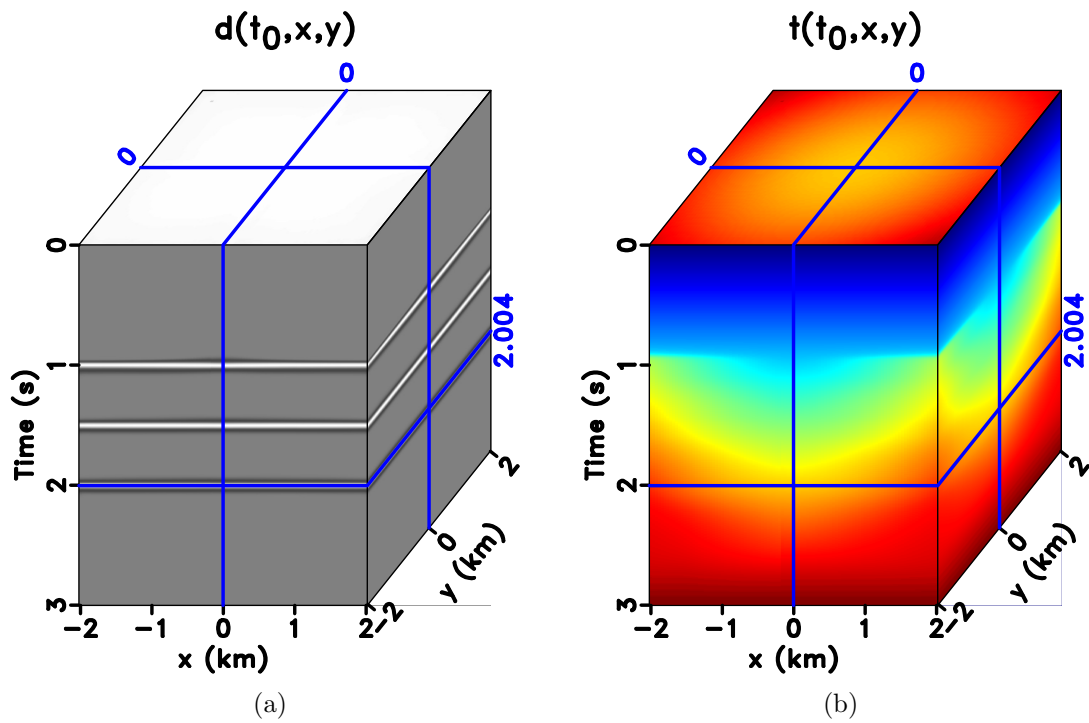
Figure 2: Warped volumes of (a) flattened CMP $d(t_0, x, y)$ and (b) time $t(t_0, x, y)$. We only apply warping to the time range of 1–2 s because CMP events are present only in this t_0 range. mcmc/ flat3cube,warpedtimecube

the 3D GMA (equation 4) can predict reflection traveltimes in 3D homogeneous or complex horizontally layered anisotropic models with a nearly exact level of accuracy for all practical purposes.

In equation 4, the parameters t_0, W_i, and A_i are defined with respect to the zero-offset ray, whereas B_i and C_i are obtained from finite-offset rays (Sripanich et al., 2017). With these definitions, the parameters t_0, W_i, and A_i are related to previously known time-processing parameters: W_i simply denote the reciprocals of NMO velocities and govern the so-called NMO ellipse (Grechka and Tsvankin, 1998a). A_i represent the quartic coefficients that govern the nonhyperbolicity of reflection traveltimes. A common approach to simplify and relate the quartic coefficients (A_i) to anisotropic parameters of the subsurface medium is done using the pseudoacoustic approximation, where A_i can then be expressed in terms of η_i (Alkhalifah and Tsvankin, 1995; Alkhalifah, 2003; Stovas, 2015; Sripanich and Fomel, 2016; Sripanich et al., 2017). With regard to this study, we keep the A_i notation to maintain the generality of the proposed moveout inversion approach.

On the other hand, the B_i and C_i are designed to be dynamic and can vary with respect to different choices of finite-offset rays, which lead to better flexibility and accuracy of traveltime fitting. As a result, they cannot be directly related to medium parameters and one may expect that there are many possibilities of B_i and C_i that can lead to equally good approximation accuracy. Therefore, we expect that the non-uniqueness of B_i and C_i can lead also to the non-uniqueness in the final inverted models.

Monte Carlo inversion

Equipped with the general concepts on time-warping and the choice of 3D GMA described in the previous sections, we propose to solve this moveout inversion problem with a global Monte Carlo inversion method. This choice is encouraged by the non-uniqueness of the solution of this inverse problem that only relies on traveltime (kinematic) data and the fact that the 3D GMA depends non-linearly on associated moveout parameters. By employing a global optimization scheme within a Bayesian framework, we can map out the non-uniqueness of possible solutions and present them as posterior probability distributions, while also honoring the non-linearity of the problem. Here, we largely follow the notations of Mosegaard and Tarantola (1995) and Tarantola (2005), where we express the posterior probability density function $\sigma(\mathbf{m})$ as

$$\sigma(\mathbf{m}) = k\rho(\mathbf{m})L(\mathbf{m}) . \tag{6}$$

$\rho(\mathbf{m})$ denotes the prior probability density function of model parameters $\mathbf{m}$ and $L(\mathbf{m})$ represents the likelihood function (with an appropriate normalization constant k) that measures the degree of fit between predicted data and observed data. Assuming a prior $\rho(\mathbf{m})$ that we can draw samples from, the goal is to design a selection process based on $L(\mathbf{m})$ that will result in samples whose density directly represents $\sigma(\mathbf{m})$.

Mosegaard and Tarantola (1995) showed that given samples of $\rho(\mathbf{m})$ and a choice of $L(\mathbf{m})$, the selection process can be done by means of the Metropolis-Hastings algorithm. In this study, we define $L(\mathbf{m})$ as:

$$L(\mathbf{m}) = \frac{1}{(2\pi S)^{N/2}} \exp\left(-\frac{1}{2S^2} \sum_{n=1}^{N} (F_n^{\text{obs}} - F_n(\mathbf{m}))^2\right) , \qquad (7)$$

where for each t_0, the misfit between the observed F_n^{obs} and modeled $F_n(\mathbf{m})$ for all N traces in the considered CMP gather is evaluated and summed according to equation 7. Here, S represents a 'hyperparameter' related to data uncertainty, which we choose to express as

$$S = \frac{S_\%}{100} \sqrt{\frac{\sum_{n=1}^{N} (F_n^{\text{obs}})^2}{N}} , \qquad (8)$$

with $S_\%$ denoting the magnitude of uncertainty expressed in percent with respect to the root mean square (RMS) of the data F_n^{obs}. Because t_0 is constant for each CMP event, S can also be considered as a relative uncertainty in the estimated reflection traveltime squared (equation 2).

We emphasize that data uncertainty is generally not easily quantified in practice due to the workflow of estimating F. To account for this complication, we propose to utilize the transdimensional (hierarchical Bayesian) inversion framework (Malinverno and Briggs, 2004; Sambridge et al., 2006, 2013), which permits us to treat $S_\%$ as yet another parameter to be estimated during the inversion. In this way, we obviate the need to estimate the data uncertainty explicitly, while correctly honoring its effects in our inversion. Our implementation of the Monte Carlo inversion with Metropolis-Hastings rule can, therefore, be summarized as follows:

1. Assume that we start at a point in the model space j with a guess set of $\mathbf{m}_j$ drawn from the prior distributions of each model parameter. In the case of the 3D GMA, $\mathbf{m}$ consists of W_i, A_i, B_i, C_i, and $S_\%$. For uniform prior distributions, a random number generator can be used for drawing samples; whereas, for Gaussian prior distributions, a combination of a random number generator and the Box-Muller transform can be used.

2. Evaluate the current likelihood function $L_j = L(\mathbf{m}_j)$ according to equation 7 with a choice of cutoff offset that depends on the example at hand.

3. Attempt to move to another point i with a new guess of $\mathbf{m}_i$ drawn from the same distribution. This attempt is evaluated according to the following Metropolis-Hastings rule.

   - If $L_i/L_j \geq 1$, then accept the move to i.
   - If $L_i/L_j < 1$, then make a random decision of staying at j or accepting the move to i with a probability of L_i/L_j.

4. Record the values of **m** after the selection.

5. Repeat steps 2–4 as many times as needed.

In our workflow, we propose to perform moveout parameter estimation with two inversion runs:

1. The first run is aimed at estimating W_i using only small-offset data. At this stage, we assume that the prior distribution is uniform within the chosen bounds (min and max values) for each model parameter in **m** and specify a small cutoff offset (e.g., where the offset-to-depth ratio is assumed to be unity).

2. Subsequently, the resulting distributions of W_i from the first run are assumed to be Gaussian and will be used as new prior distributions for W_i in the second inversion run with all available (small- and large-offsets) data to estimate A_i, B_i, and C_i. Better prior distributions of W_i from the first run are expected to help with the estimation of A_i, B_i, and C_i. The prior distributions for other pertaining parameters are again assumed to be uniform within some chosen bounds (min and max values).

After the two runs, the final recorded **m** can be visualized as histograms that represent 1D marginal posterior probability distributions of the estimated moveout parameters **m**. The corresponding posterior probability density function $\sigma(\mathbf{m})$ can subsequently be obtained with appropriate normalization of the histograms. In order to quantify the 'gain of information' from the moveout inversion process, we use the Kullback-Leibler divergence $D_{KL} \in [0, \infty)$ given by,

$$D_{KL} = \int \sigma(\mathbf{m}) \ln \left(\frac{\sigma(\mathbf{m})}{\rho(\mathbf{m})} \right) d\mathbf{m} \;, \tag{9}$$

which estimates the relative difference between the prior $\rho(\mathbf{m})$ and the posterior $\sigma(\mathbf{m})$. A high D_{KL} value indicates that the posterior distribution is very different from the prior, and therefore, has gained new information on the moveout parameters through the inversion process using the provided data (i.e, moveout traveltime shifts F_n^{obs} in our problem). On the other hand, $D_{KL} = 0$ indicates that the two distributions are identical and no gain has been achieved upon seeing the data. In all of our examples, we report D_{KL} with respect to the input uniform prior distributions to the first inversion run.

Even though D_{KL} is generally unbounded, it is possible to compute a benchmark value for some simple case. For example, if we consider Gaussian prior and posterior distributions, the Kullback-Leibler divergence (D_{KL}) can be expressed as

$$D_{KL} = \ln \left(\sqrt{\frac{v_\rho}{v_\sigma}} \right) + \frac{v_\sigma + (\mu_\sigma - \mu_\rho)^2}{2v_\rho} - \frac{1}{2} \;, \tag{10}$$

where the v_i and μ_i are the variance and mean of the distributions, respectively. It follows that if no information is gained during the inversion process and $\sigma(v_\sigma, \mu_\sigma) = \rho(v_\rho, \mu_\rho)$, then $D_{KL} = 0$. If the standard deviation of the posterior reduces to half that of the prior with the same mean, the Kullback-Leibler divergence is equal to

$$D_{KL} = \ln(2) + \frac{1}{8} - \frac{1}{2} = 0.318 \ . \tag{11}$$

EXAMPLES

2D synthetic model with inverse moveout

We first consider a simple 2D example where the input CMP gather is generated from the inverse moveout process based on the 2D version (Fomel and Stovas, 2010) of equation 4 (2D GMA) given by

$$t^2(x, y) \approx t_0^2 + Wx^2 + \frac{Ax^4}{t_0^2 + Bx^2 + \sqrt{t_0^4 + 2Bt_0^2 x^2 + Cx^4}} \ . \tag{12}$$

This case can be associated with a horizontal reflector beneath a homogeneous layer of transversely isotropic medium with vertical symmetry axis (VTI). Relying on the pseudoacoustic approximation and the choice of zero-offset and horizontal rays to fit moveout parameters, Fomel and Stovas (2010) show that for a homogeneous VTI medium, it is possible to specify A, B, and C in terms of η (Alkhalifah and Tsvankin, 1995) — a commonly used parameter to characterize the nonhyperbolicity of moveout traveltimes. This process leads to

$$A = -4W^2\eta \ , \ \ B = \frac{W(1 + 8\eta + 8\eta^2)}{1 + 2\eta} \ , \ \text{and} \ C = \frac{W^2}{(1 + 2\eta)^2} \ , \tag{13}$$

where

$$\eta = \frac{\epsilon - \delta}{1 + 2\delta} \ . \tag{14}$$

ϵ and δ are Thomsen's parameters that characterize the property of the overlying VTI medium.

In this example, we consider a CMP event generated by the inverse moveout process (Figure 3) based on the 2D GMA (equation 12) with A, B, and C specified as in equation 13. We use the VTI parameters of Dry Green River shale (Thomsen, 1986) given by $V_p = 3.292$ km/s, $V_s = 1.768$ km/s, $\epsilon = 0.195$, and $\delta = -0.22$. These numbers can be translated to $V_{nmo} = V_p\sqrt{1 + 2\delta} = 1/\sqrt{W} = 2.46$ km/s and $\eta = 0.741$, which by equation 13, correspond to $W = 0.165$, $A = -0.0805$, $B = 0.7516$, and $C = 0.00441$.

Using this configuration, we first test our algorithm and its capability to estimate both wanted model parameters and the hyperparameter $S_\%$ that governs the data uncertainty. We consider the setup for the first inversion run in our workflow, where we restrict our data range to small offsets. The benefit of this test is two-fold:
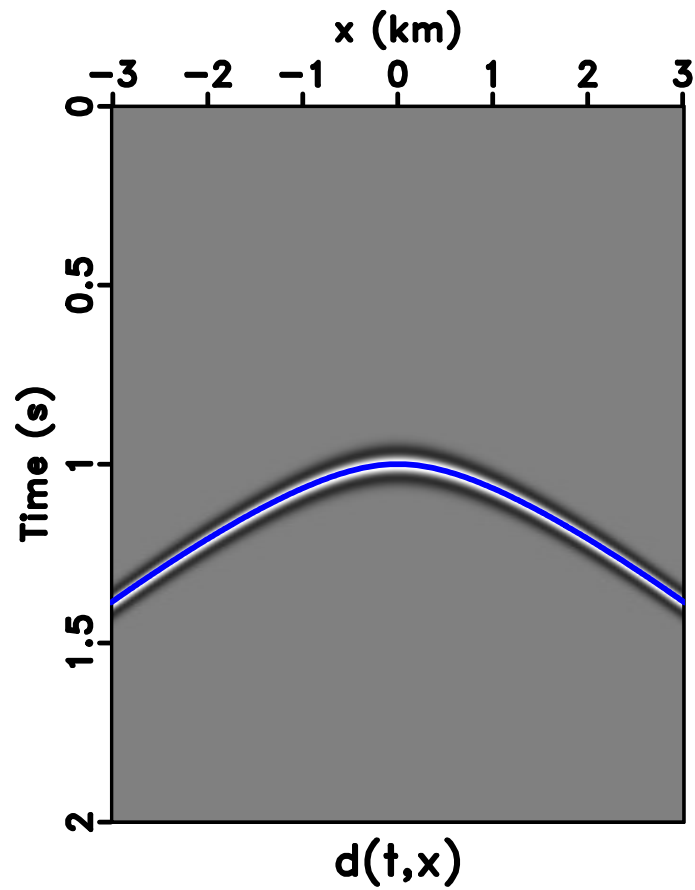
Figure 3: The CMP gather 2D Dry Green River shale VTI model overlain by an automatically picked event (blue) for non-physical flattening. mcmc2D/ pick2d

1. Given that our algorithm functions correctly, we should be able to recover both W and $S_\%$, while seeing how the posterior distributions of W changes with varying data uncertainty $S_\%$.

2. If the estimated maximum likelihood points (peaks of histograms) are not centered around the true values, the results then suggest some trade-offs among inverted parameters intrinsic to our traveltime-only moveout inversion setup.

In this test, we assume that the picked moveout curve for non-physical flattening is approximately error-free and add artificial Gaussian noise with a variance equal to S^2 (equation 8) with $S_\%$=0.1, 0.5, 2, 5, and 25 to the input F_n^{obs} for the inversion process. We record model parameters at every 100 iterations to increase the independence between recorded models and stop the iterations when the number of recorded model reaches 20,000. We specify the minimum and maximum values for the uniform prior distributions as shown in Table 1 except for the range of $S_\%$, which is set to 0–60 in this case. Here, we use the relationships between GMA parameters (A, B, and C) and η (equation 13) to help guide our selection of the min and max values and we also specify relatively large ranges to reflect the reality of this situation in practice. The resulting posterior distributions for W and $S_\%$ from this experiment with the cutoff offset specified at 1.25 km are shown in Figure 4. We can observe that our algorithm performs well and manages to correctly estimate posterior distributions for both W and $S_\%$ centered around the true values. Moreover, at lower $S_\%$ (small data uncertainty/noise), the estimated posterior distribution of W has smaller variance and has its peak closer to the true value. On the other hand, the estimated posterior distribution of W becomes wider (higher variance) with increasing $S_\%$ as would be anticipated. Looking closer at Figure 4, we also see that at higher levels of $S_\%$, the peak of the posterior distribution of W shifts towards lower value suggesting a trade-off between inverted parameters despite the use of only small-offset data. This observation is supported by the 2D marginal posterior distributions of W and the other three parameters: A, B, and C as shown in Figure 5 for the case of $S_\%$=2. We can observe a clear correlation between the $W - A$ pair. A decrease in estimated W correlates with an increase in estimated A suggesting a trade-off between the two parameters in our inversion setup. Correlations among the other pairs of $W - B$ and $W - C$ are not observed in this experiment.

Parameters	W	A	B	C	$S_\%$
Min	0.1	-0.1	0.5	0.0	0.0
Max	0.3	0.0	1.0	0.006	10

Table 1: Minimum and maximum values of moveout parameters that define the uniform prior distributions for our Monte Carlo inversion for Dry Green River shale example.

After successful testing, we use the same example and implement the proposed two-run inversion without any additional noise to the input data. In this case, we
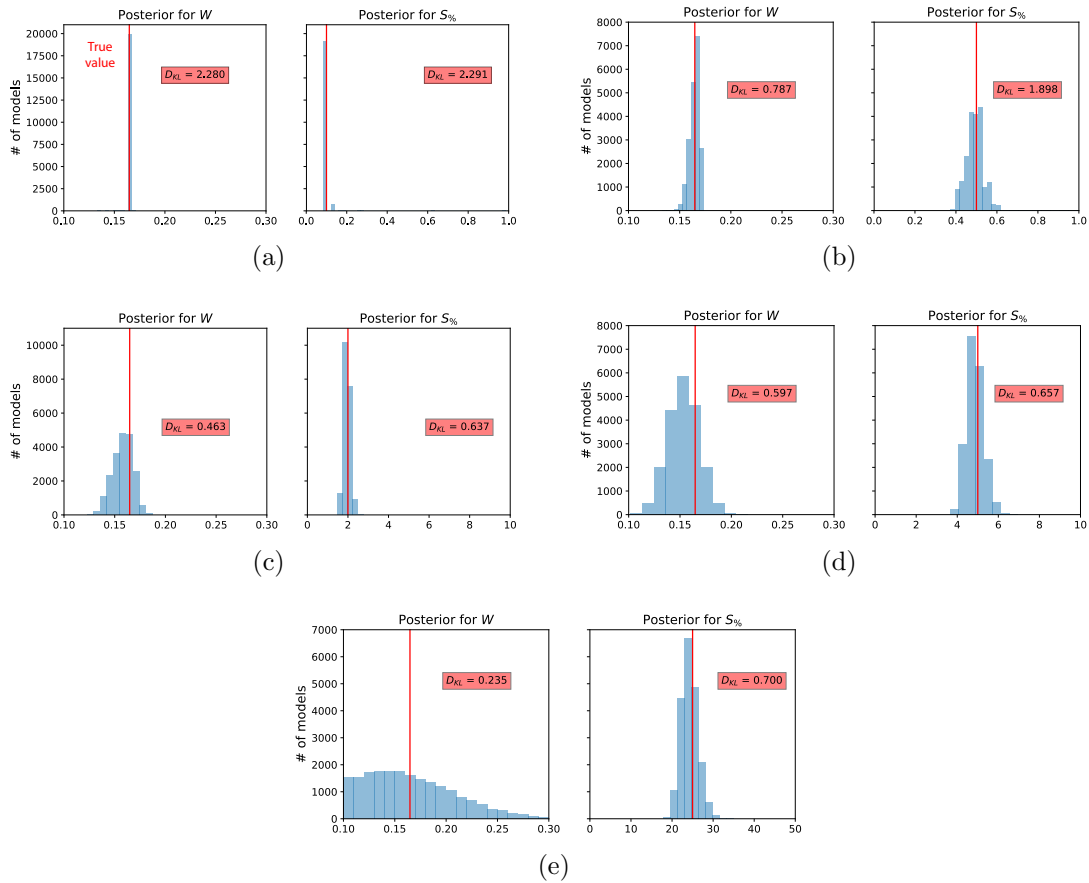
Figure 4: A comparison of estimated posterior probability distributions for W and $S_\%$ of the Dry Green River shale example with cutoff offset at 1.25 km for different levels of added Gaussian noise with $S_\% =$ (a) 0.1, (b) 0.5, (c) 2, (d) 5, and (e) 25. The solid vertical lines denote the true values.
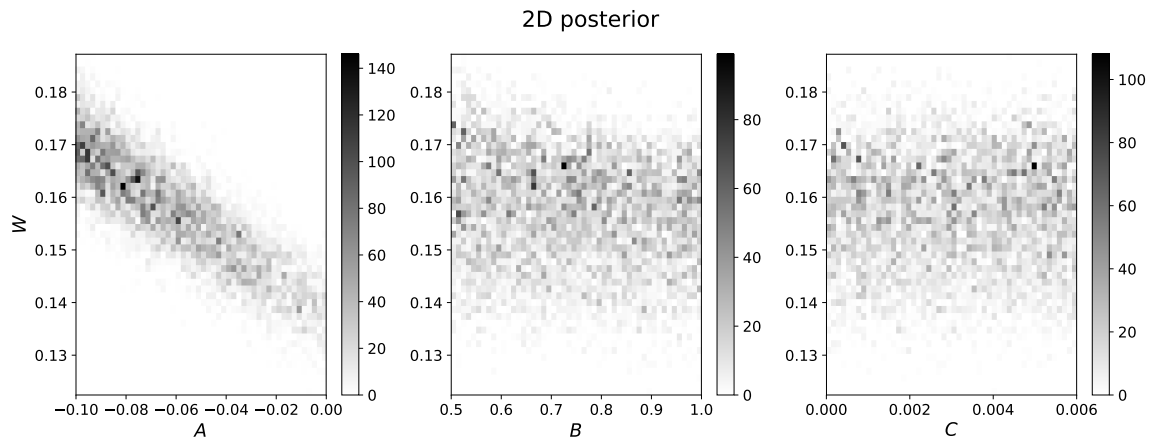

Figure 5: 2D marginal posterior distributions of W and A, B, and C. Prominent correlations between the $W - A$ pair can be observed suggesting trade-offs in estimating the two parameters.

record the model at every 500 iterations and stop when the total number of recorded model reaches 20,000. Similar min and max values for the uniform prior distributions in Table 1 are used in this experiment. The resulting posterior distributions of all estimated parameters are shown in Figure 6. We can observe that the posterior distributions of W and A have their peaks close to the true values with high D_{KL} suggesting that information on both parameters has been gained from the inversion and they appear to be recoverable relatively high fidelity. On the other hand, the distributions for B and C are more diffuse with no clear peak and lower D_{KL}. We emphasize that the results on B and C agree with the properties of the GMA, where both parameters are designed to be non-unique for dynamic traveltime fitting with high accuracy. Finally, the posterior distribution for $S_\%$ has only one peak close to zero verifying our assumption that the picked traveltime curve in this example is approximately noise-free.
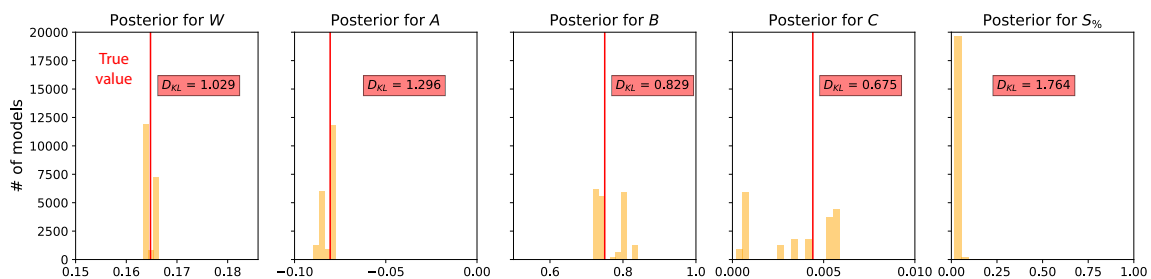
Figure 6: Estimated posterior probability distributions in the form of histograms from the proposed two-run inversion for the Dry Green River shale example. The solid vertical lines denote the true values. The maximum likelihood points are located close to the true values suggesting that moveout parameters from 2D gathers are well recoverable from the proposed approach.

2D Mobil Viking Graben data set

Next, we consider an example with real 2D dataset from the Viking Graben (Figure 7) from Keys and Foster (1998). In this example, we utilize the flexibility of the proposed time-warping workflow that can be implemented at any selected locations related to different reflectors to help assessing velocity estimated from conventional semblance-based technique. We can achieve this selection by choosing an appropriate t_0 (a slice of the $t(t_0, x, y)$ cube) that corresponds to the desired reflector and implement a Monte Carlo inversion only at such location.

Figure 8 shows the CMP gather number 1000 of the Viking Graben data overlain by two picked moveout curves associated with prominent reflectors to be considered in our inversion. Due to the availability of only small-offset data, we are obliged to restrict our inversion to W and $S_\%$. Their ranges for the input uniform prior distribution are 0.08–0.5 for W and 0–10 for $S_\%$. The cutoff offsets used for the two events are at 1.8 (blue) and 2.2 km (red) approximately corresponding to trace number
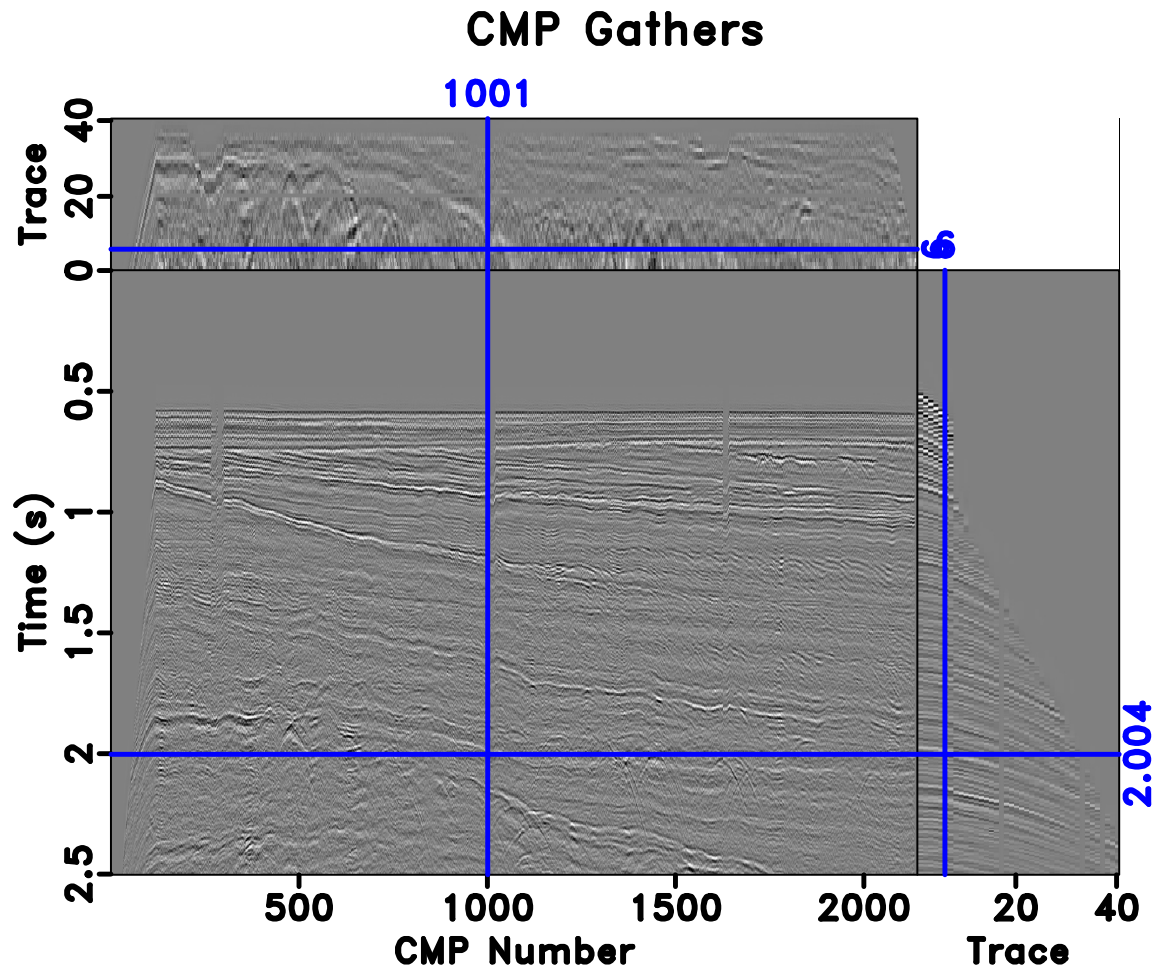
Figure 7: The Mobil Viking Graben dataset. viking/ cmpsvk

27 and 35, respectively. A comparison of the semblance-based velocity and those from the proposed workflow is shown in Figure 9 with the solid black line denoting semblance-based velocity picks and white '+' signs denoting the maximum likelihood points from the proposed inversion workflow. The estimated posterior distributions of both events are shown in Figure 10. We can see that in this example the results from both approaches generally agree well with each other despite relying on different theoretical bases (signal coherency vs. traveltime-only inversion) to estimate W. This result gives further assurance to the estimated parameters. Nevertheless, we note that using the proposed workflow, we obtain posterior distributions that contain additional statistical information for more assessment and quality control of both the estimated moveout parameters (only W in this case) and the input data uncertainty ($S_\%$).
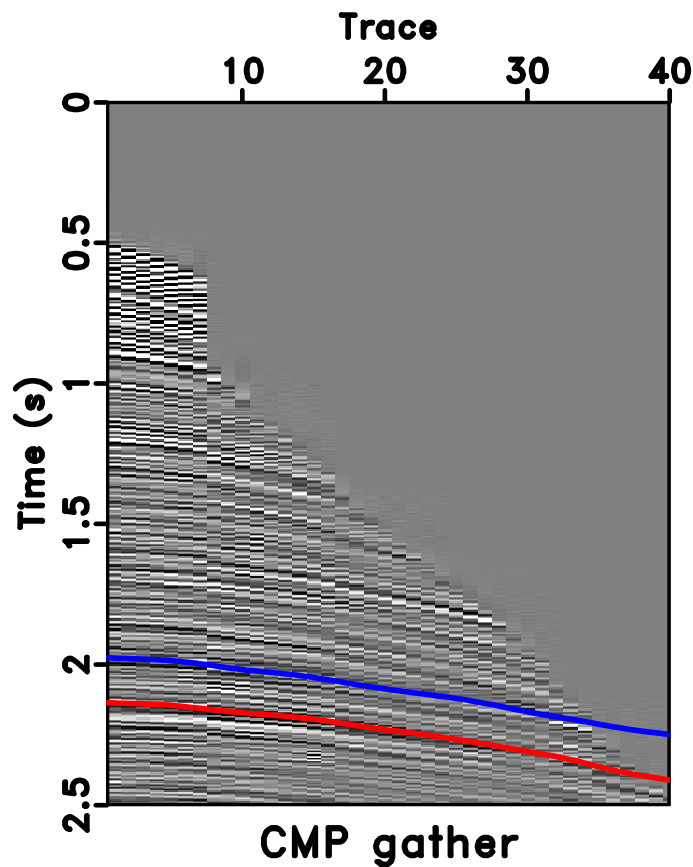
Figure 8: CMP number 1000 from the Mobil Viking Graben dataset overlaid by picked moveout curve used in the time-warping workflow. The cutoff offsets at 1.8 (for blue) and 2.2 (for red) km are approximately at trace number 27 and 35. viking/ pickvk

3D synthetic model with inverse moveout

Turning to a 3D case, we consider an example where the input CMP gather is generated from the inverse moveout process similar to the previous 2D case but now
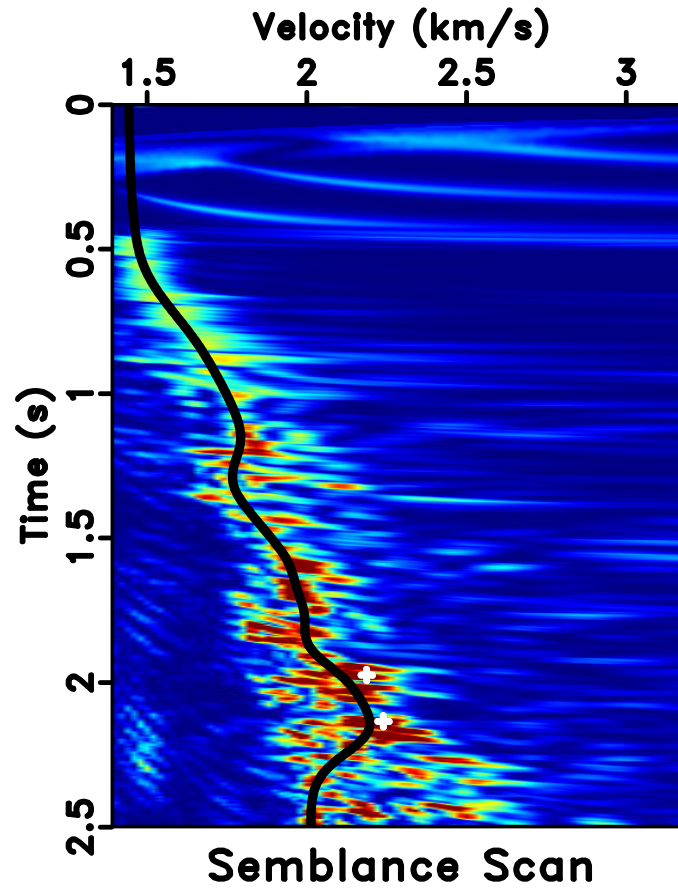
Figure 9: Semblance scan and automatically picked velocity (black line) associated with CMP number 1000 from the Mobil Viking Graben dataset. The white '+' signs are the peaks of estimated posterior distribution of W from the proposed inversion workflow. viking/ vscanvk
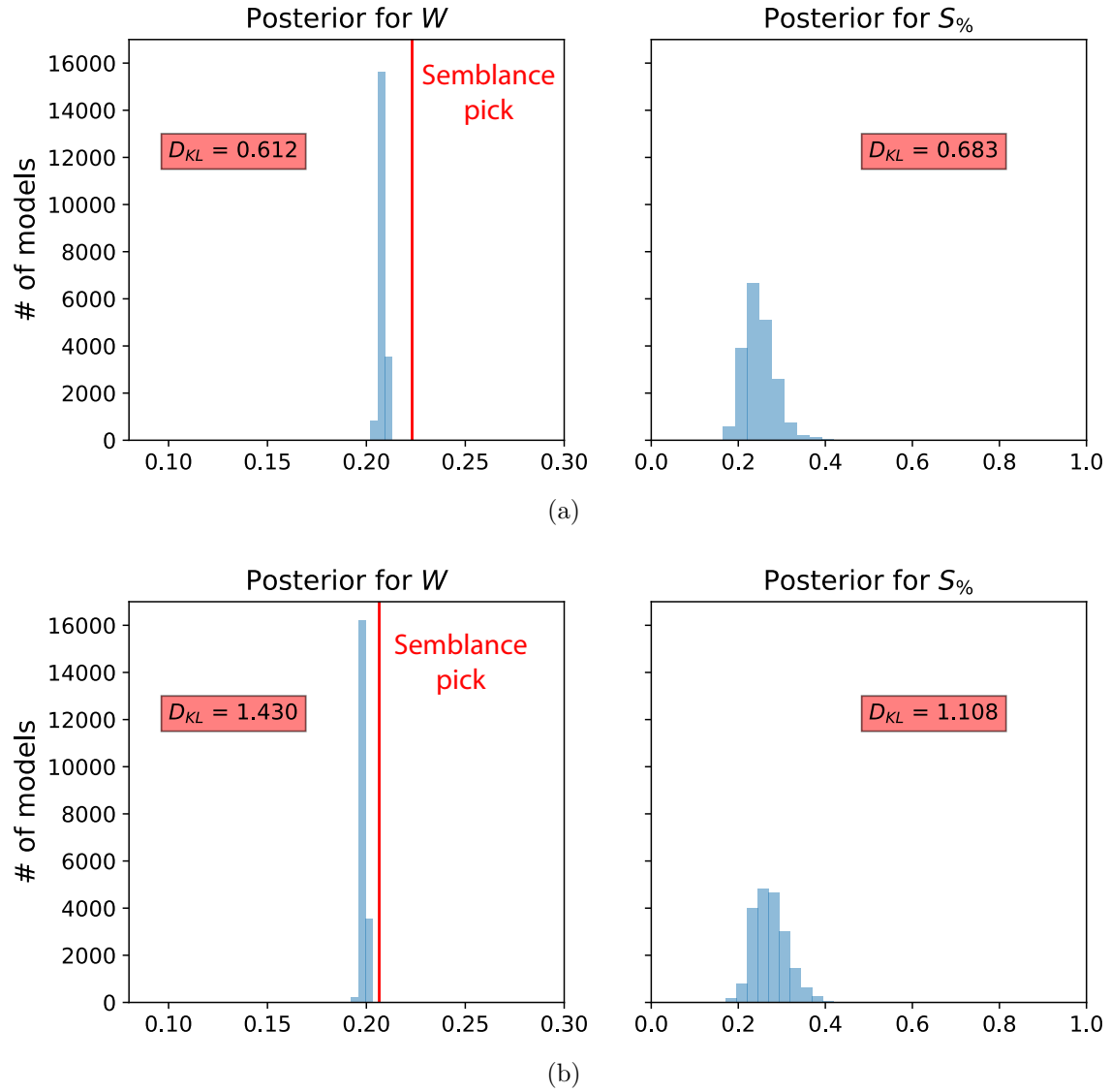
Figure 10: Estimated posterior probability distributions in the form of histograms for the (a) blue and (b) red events (Figure 8) associated with CMP number 1000 from the Mobil Viking Graben dataset. The solid lines denote picked values from semblance-based analysis.

with the 3D GMA in equation 4. All pertaining parameters W_i, A_i, B_i, and C_i are non-zero and can represent the case of an orthorhombic subsurface with azimuthally rotated principal axes with respect to the local Cartesian coordinates. The inputs of the time-warping workflow for this experiment have previously been used in our explanation on the method and are shown in Figure 1. We purposefully specify three CMP events at $t_0 = 1$, 1.5, and 2 s, which correspond to the maximum offset-to-depth ratios of approximately 2.326, 1.789, and 1.5 in the x direction and of 2.474, 1.876, and 1.5625 in the y direction reflecting our choice of W_1 and W_3, respectively. According to the general rule of thumb of offset-to-depth ratio > 1 for nonhyperbolic moveout analysis (Alkhalifah, 1997; Tsvankin, 2012; Thomsen, 2014), we anticipate that our inverted posterior probability distributions for A_i from the topmost event should demonstrate the best performance. We also add Gaussian noise with $S_\% = 2.5$ (top event), 2.75 (middle event), and 3.0 (bottom event) to the input traveltime data from automatically picked moveout curves. Because we are dealing with a more challenging 3D gather, we anticipate that the automatically picked moveout curve will no longer be error-free and therefore, the final estimates of $S_\%$ would reflect total data uncertainty from both the artificially added noise and the automatic picking process.

In this example, we specify the minimum and maximum values for the prior distributions as shown in Table 2. Here, we use the computational experiments with 3D GMA from Sripanich et al. (2017), as well as equation 13, to provide a rough estimate for the magnitude of each parameter and guide our choices of their min and max values, while keeping a relatively large range. The model parameters are recorded every 500 iterations until reaching a total of 20,000 models. The cutoff offsets for the first inversion in our workflow are at 1, 1.25, and 2 km for the top, middle, and bottom events, respectively. The estimated posterior probability distributions of the three events from our proposed workflow are shown in the form of histograms in Figures 11 – 16. It is apparent from Figure 11 that the W_i are estimated well with relatively high D_{KL} values.

Parameters	W_1	W_2	W_3	A_1	A_2	A_3	A_4	A_5	B_1	B_2	B_3	C_1	C_2	C_3	C_4	C_5	$S_\%$
Min	0.1	-0.15	0.12	-0.09	-0.045	-0.05	-0.035	-0.09	0.25	0.01	0.25	0.0015	0.001	0.0045	0.001	0.0015	0.0
Max	0.3	0.	0.32	-0.001	-0.02	-0.03	-0.015	-0.001	0.8	0.1	0.75	0.005	0.0045	0.007	0.0035	0.0045	5

Table 2: Minimum and maximum values of moveout parameters that define the uniform prior distributions for our Monte Carlo inversion for 3D inverse moveout example.

We proceed with the second inversion run with all available data and the resulting W_i from the first run. The inverted A_i from this experiment require more detailed analyses. First, we emphasize that the posterior distributions of A_i appear more diffuse and multi-modal (Figure 12). This observation suggests that there are many equivalently good values of A_i that can lead to an accurate prediction of the moveout traveltime, which indicates the inherent non-uniqueness of moveout inversion problem based on traveltime data alone. Moreover, it is not obvious that having a larger offset-to-depth ratio (Figure 12a associated with the topmost event) helps to mitigate this
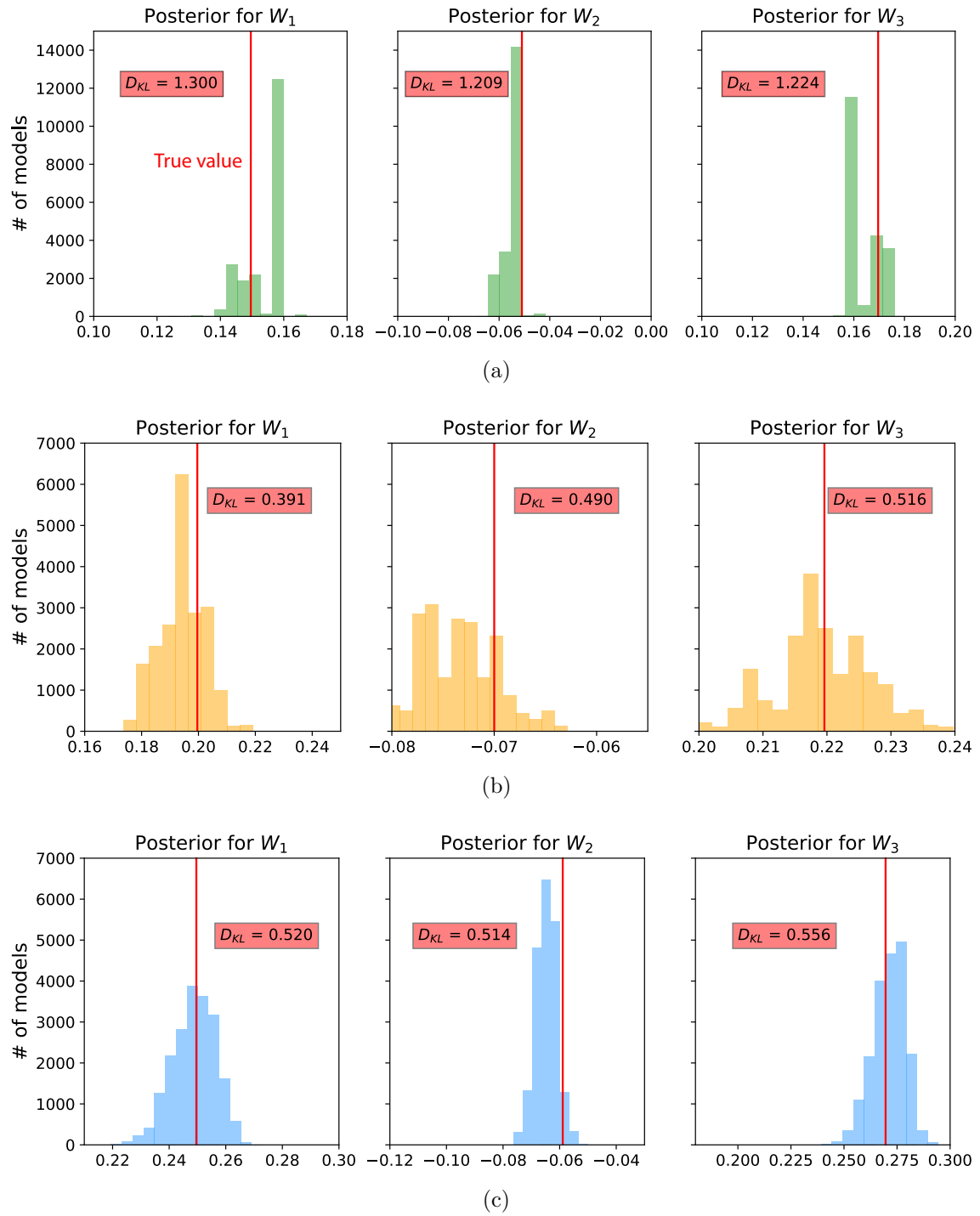
Figure 11: A comparison of estimated posterior probability distributions for W_i in the form of histograms for events at (a) 1 s, (b) 1.5 s, and (c) 2 s of the 3D inverse moveout model. The solid vertical lines denote the true values.

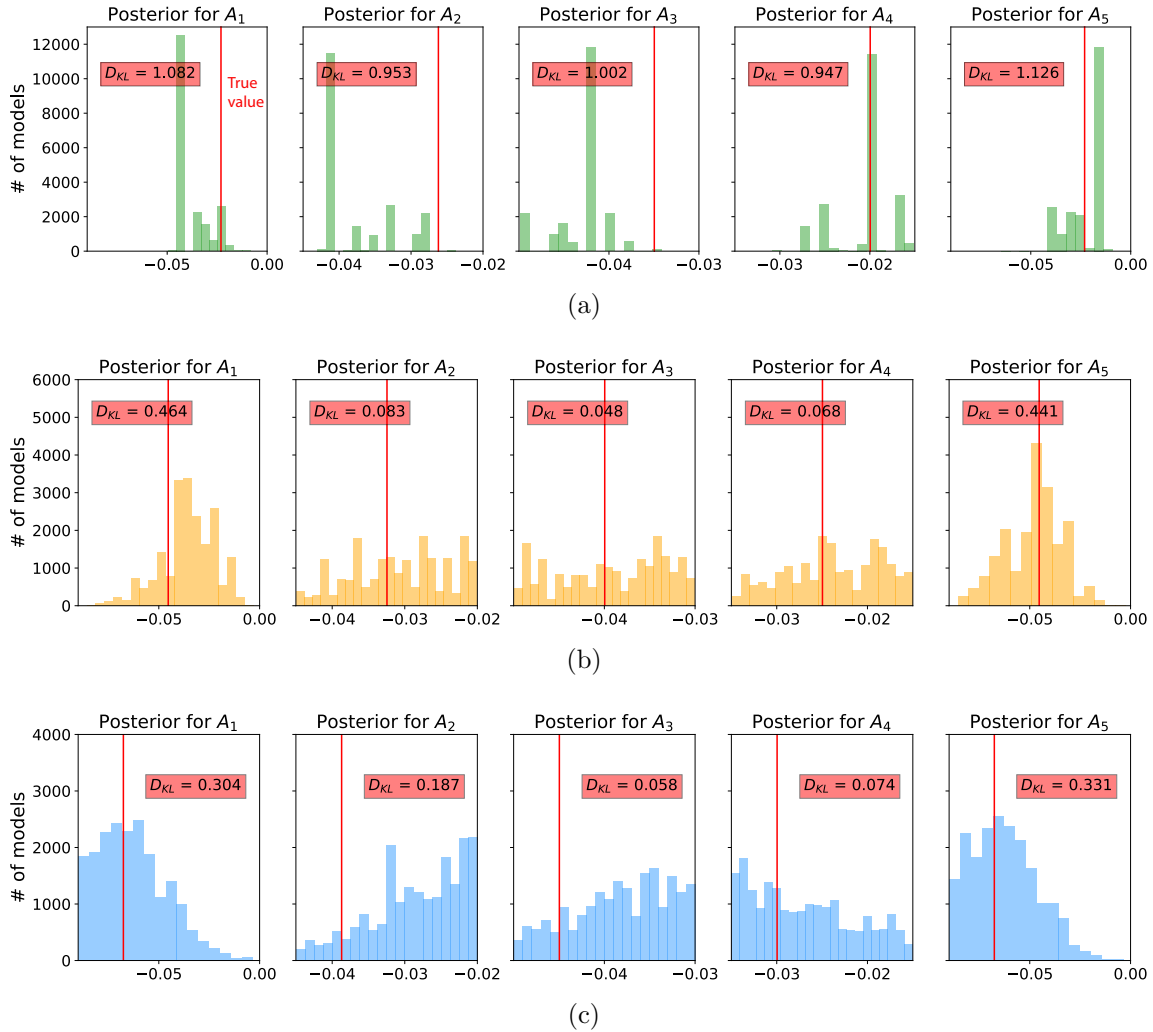non-uniqueness of possible solutions.

(a)

(b)

(c)

Figure 12: A comparison of estimated posterior probability distributions for A_i in the form of histograms for events at (a) 1 s, (b) 1.5 s, and (c) 2 s of the 3D inverse moveout model. Note that they are more diffuse than the inverted W_i in Figure 11. The solid vertical lines denote the true values.

A closer look at Figures 11a, reveals that the peak of the posterior distribution of W_1 are slightly higher than the true value, whereas that of W_3 is slightly lower. This becomes less obvious in Figures 11b and 11c, which correspond to smaller offset-to-depth ratios. One explanation of this behavior can be given by considering Figures 11a and 12a together. In Figures 12a, the maximum likelihood values of the posterior A_i also slightly deviate from the true values. Therefore, similar to the previous to 2D case, there appears to be a trade-off between inverted W_i and A_i on top of the non-uniqueness of possible solutions, which seems to be noticeable with data from larger offset-to-depth ratio. This observation is supported by the 2D marginal posterior distributions of W_i and A_i shown in Figure 13, where we can observe clear correlations between the $W_1 - A_1$ pair and the $W_3 - A_5$ pair. A decrease in estimated W_1 (W_3)

correlates with an increase in estimated A_1 (A_5). A lesser degree of correlations can also be seen from $W_2 - A_2$ and $W_2 - A_4$ pairs, where a decrease in W_2 correlated with an increase in A_2 and A_4. Correlations among other pairs of W_i and A_i are not observed from this experiment. Note that we display the 2D marginal posterior distributions from the bottom event at 2 s in Figure 13. This is done to facilitate the visualization of possible correlations because the associated posterior distributions are more diffuse. However, our observation remains valid and applicable to other events because this trade-off behavior is a consequence of our inversion configuration with the 3D GMA.
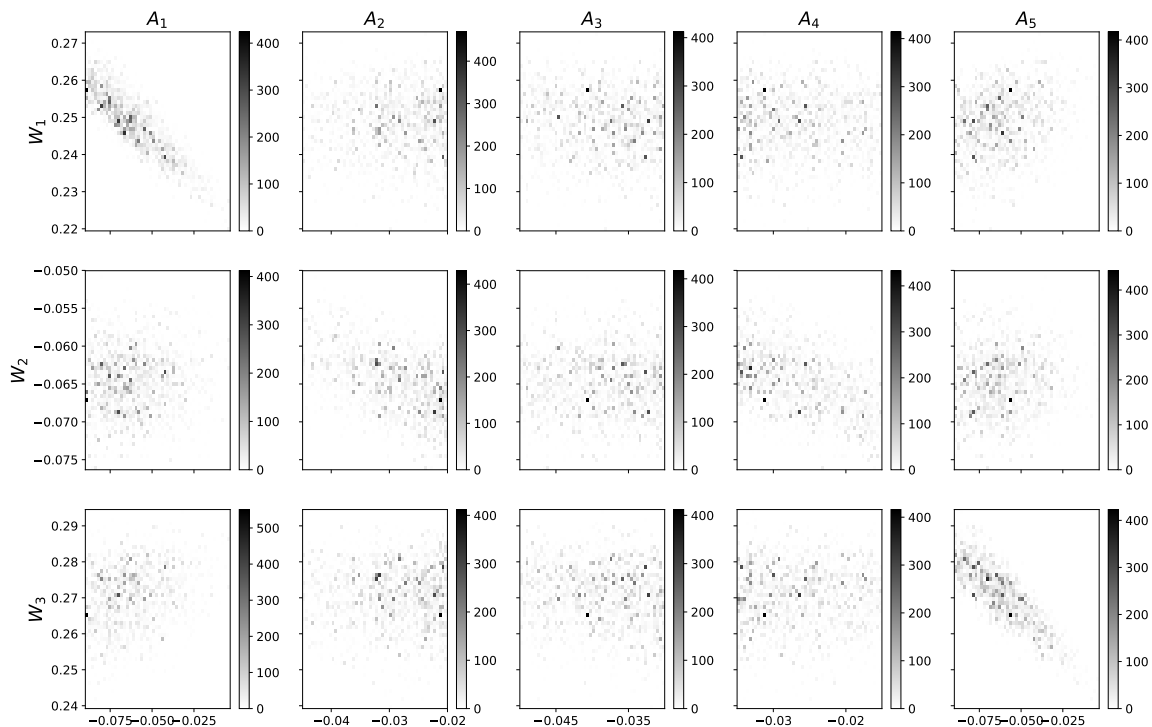
Figure 13: 2D marginal posterior distributions of W_i and A_i from the 3D inverse moveout example for the event at 2 s that suggest correlations and trade-offs among inverted parameters. The rows denote W_i and the columns denote A_i. Prominent correlations between the $W_1 - A_1$ pair and $W_3 - A_5$ pair can be observed.

Analogous to the estimated A_i, the resulting distributions of B_i (Figure 14) and C_i (Figure 15) are also fairly diffuse and multi-modal. Therefore, similar conclusions can be drawn regarding the non-uniqueness of possible solutions. However, similar to the 2D case, we emphasize that both parameters, whose values depend on the choice of finite-offset rays, are designed to be dynamic in the 3D GMA and this behavior is to be expected. Figure 16 shows the resulting posterior distributions of $S_\%$ for the three events. We can observe that the peak estimates of $S_\%$ are slightly higher than the level of added Gaussian noise agreeing with our expectation that the automatically picked moveout curve for this 3D gather contains a small degree of uncertainty, which can be captured and quantified in our workflow.
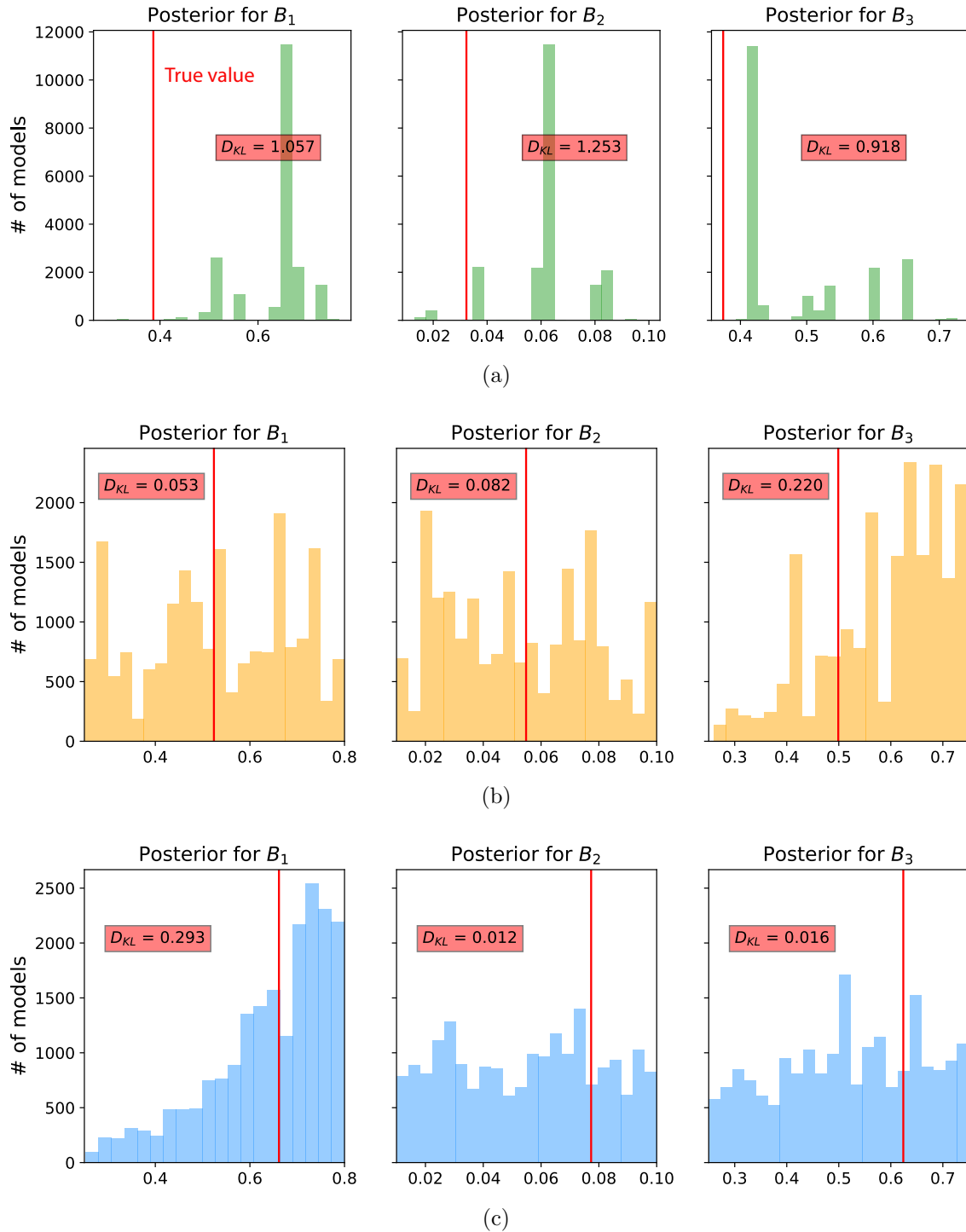
Figure 14: A comparison of estimated posterior probability distributions for B_i in the form of histograms for events at (a) 1 s, (b) 1.5 s, and (c) 2 s of the 3D inverse moveout model. Note that they are moderately diffuse and sometimes multi-modal similar to Figure 12. The solid vertical lines denote the true values.
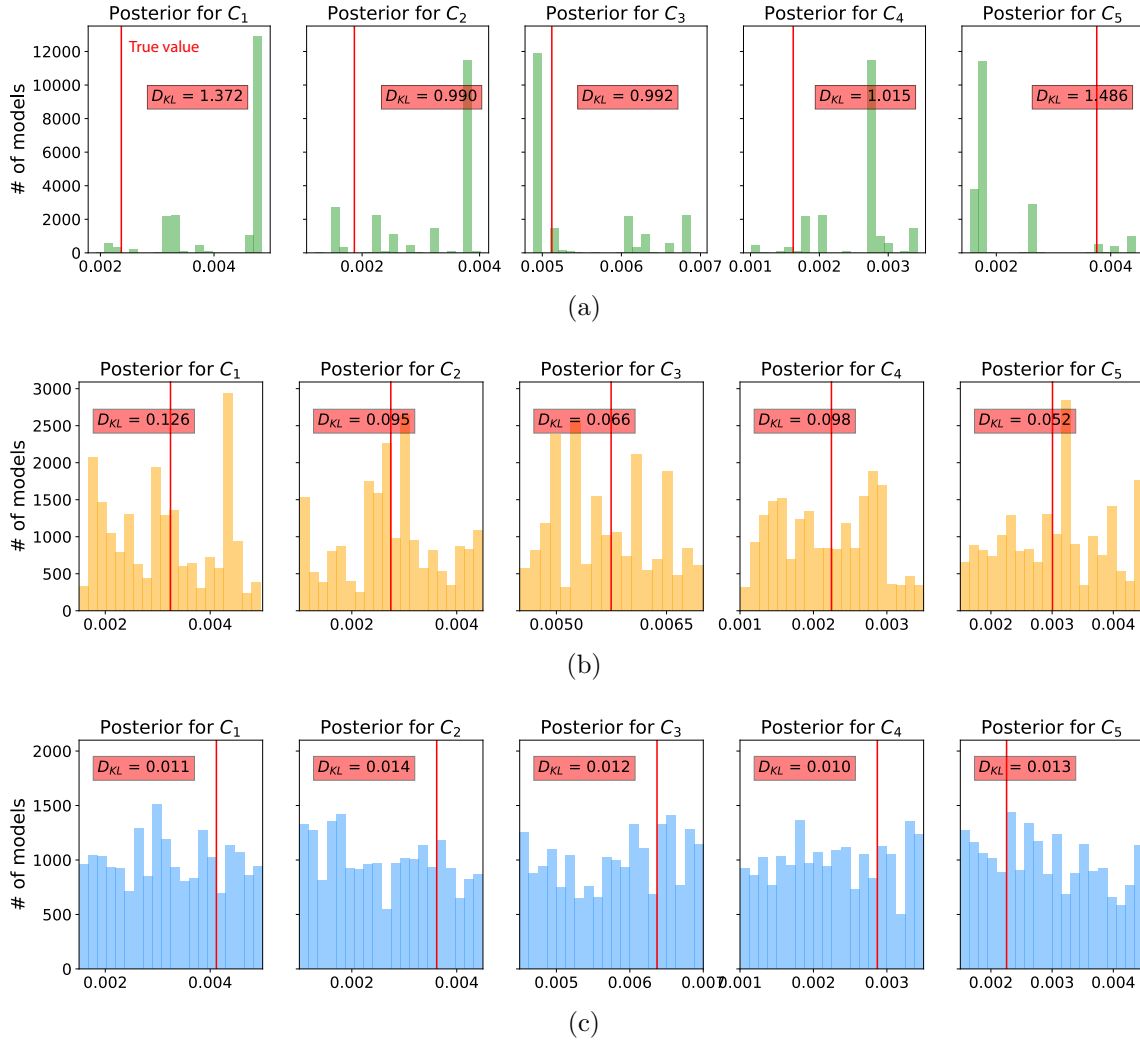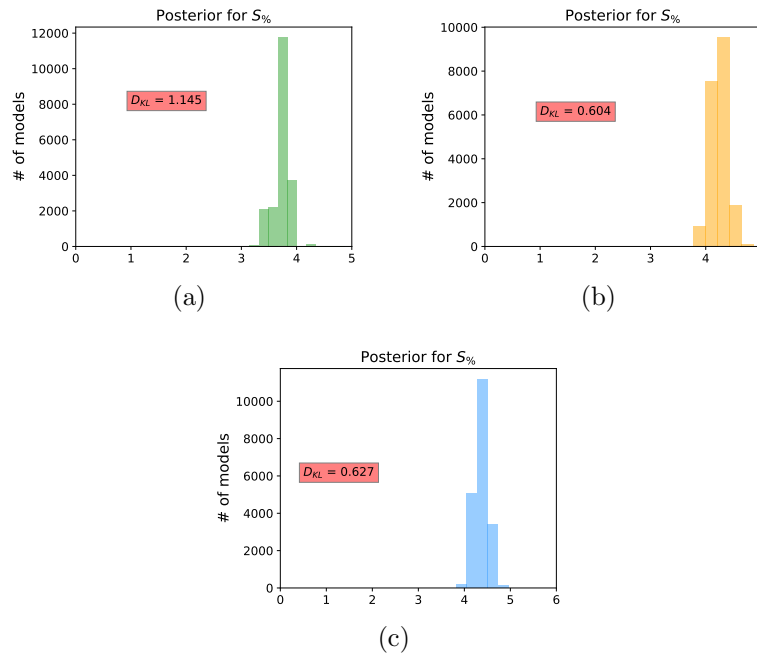
Figure 15: A comparison of estimated posterior probability distributions for C_i in the form of histograms for events at (a) 1 s, (b) 1.5 s, and (c) 2 s of the 3D inverse moveout model. Note that they are moderately diffuse and sometimes multi-modal similar to Figure 12. The solid vertical lines denote the true values.

Figure 16: A comparison of estimated posterior probability distributions for $S_\%$ in the form of histograms for events at (a) 1 s, (b) 1.5 s, and (c) 2 s of the 3D inverse moveout model reflecting the total uncertainty from both the automatically picked 3D moveout curve and the artificially added Gaussian noise. The peaks in the distributions are slightly higher than the levels of added Gaussian noise for the three events at $S_\% = $ 2.5, 2.75, and 3 suggesting the small degree of uncertainty from the automatic picking process in a more challenging case of 3D gather.

We emphasize that the above analysis is permissible due to the use of a Bayesian framework for the inversion that gives posterior probability distribution as a result. Together with our analysis from the previous 2D inverse moveout example, our results suggest a limitation of using moveout traveltime inversion to estimate quartic coefficients, which becomes more difficult in 3D with a higher number of wanted parameters associated with a more complex subsurface model (e.g., orthorhombic media).

SEAM Phase II unconventional reservoir model

For a more complex example, we consider a CMP gather from the 3D SEAM II unconventional reservoir model shown in Figure 17 (Oristaglio, 2015). This model has a largely flat structure suitable for moveout analysis and includes unaligned orthorhombic sublayers that require 3D multi-azimuth processing. We choose one CMP gather located at $x = 4.735$ km and $y = 2.881$ km, eliminate multiples, and bin to 39×39 offset samples from -3.8 km to 3.8 km in both x and y directions. We note that the maximum depth of this model is 3.75 km, therefore the offset-to-depth ratio of intermediate layers are moderately larger than unity.
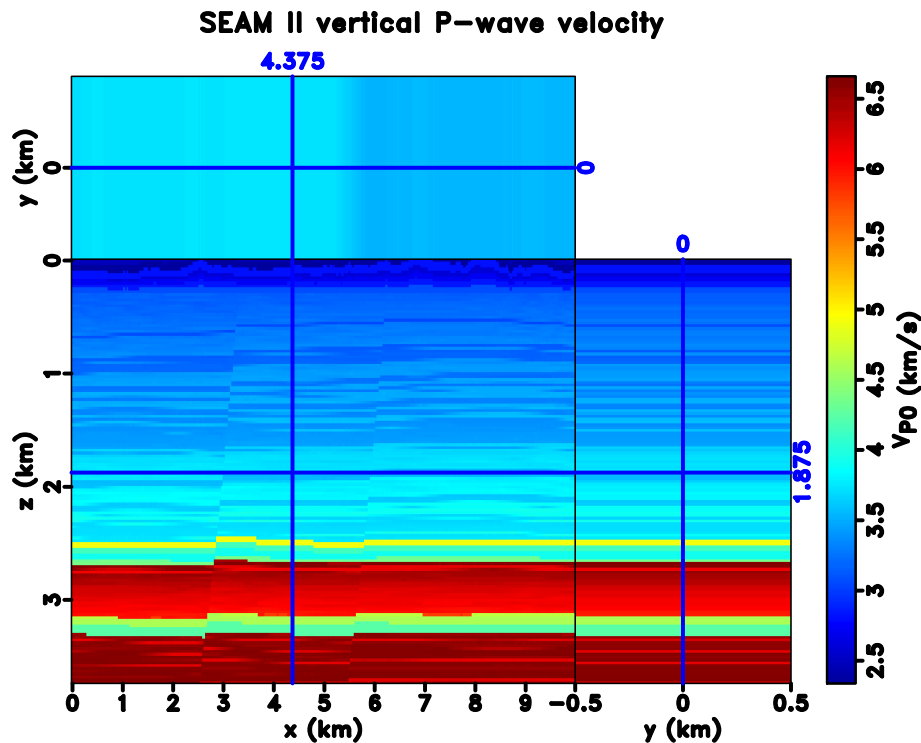

Figure 17: Vertical P-wave velocity of SEAM II unconventional reservoir model. seam2mcmc/ vpseam3d

Figure 18 shows the considered CMP gather overlain by automatically traced events for non-physical flattening. We specifically focus on two events at t_0=1.315 s and 1.815 s. The parameter ranges for prior distributions are given in Table 3, where

we use the same notion as in the previous 3D inverse moveout case to choose the min and max values. We also assume that B_i and C_i are largely unknown and specify relative large ranges for their values. The model parameters are recorded every 500 iterations until reaching a total of 10,000 models. The cutoff offsets for the first inversion are at 0.89 km and 1.5 km for the two events, respectively. In this example, we choose to use 1D average values of moveout parameters to benchmark our results. These values are computed by taking the true model parameters at the considered CMP location and use them to compute effective moveout parameters appropriate for a stack of horizontally layered unaligned orthorhombic media (Sripanich and Fomel, 2016; Koren and Ravve, 2017). They are exact in the case of truly 1D media but represent good approximate values of moveout parameters in this case of mildly laterally heterogeneous SEAM II model. We note also that in this example, the contribution from the quartic terms (A_i) as shown in their 1D average values are relatively small in comparison with that from the second-order terms (W_i). Therefore, this condition is expected to further hamper the parameter estimation process in addition to the inherent non-uniqueness of possible solutions.
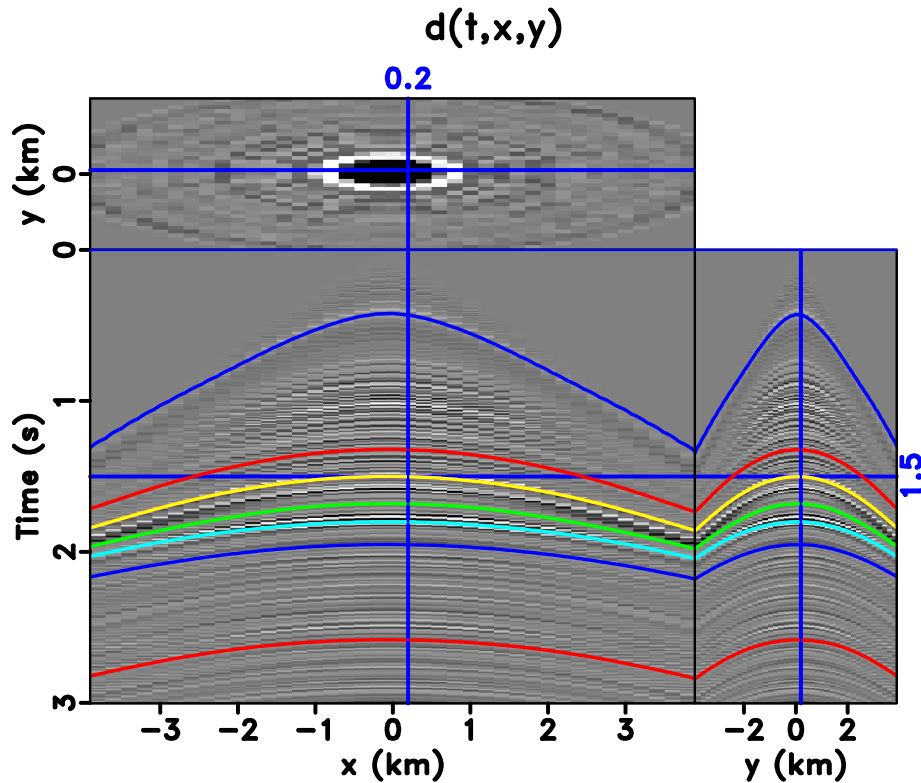

Figure 18: The CMP gather of SEAM II model at x=4.375 km and y=2.881 km overlain by automatically picked events for non-physical flattening. seam2mcmc/ pickongather0

Example posterior histograms for events at $t_0 = 1.315$ s and 1.815 s are shown in Figures 19–21. It follows from the results that in this example of SEAM II, the proposed approach can capture W_1 and W_3 with clearly defined maximum likelihood

Parameters	W_1	W_2	W_3	A_1	A_2	A_3	A_4	A_5	B_1	B_2	B_3	C_1	C_2	C_3	C_4	C_5	$S_\%$
Min	0.05	-0.05	0.05	-0.01	0.0	-0.005	0.0	-0.01	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Max	0.15	0.05	0.15	0.0	0.001	-0.001	0.001	0.0	0.5	0.1	0.5	0.005	0.005	0.005	0.005	0.005	5.0

Table 3: Minimum and maximum values of moveout parameters that define the uniform prior distributions for our Monte Carlo inversion for SEAM II example.

values close to the 1D average values and a high value of D_{KL}. However, results on W_2 parameter that define the orientation of the NMO ellipse suggest that its value is centered around zero different from the 1D average value. The estimated posterior distributions of A_i related to subsurface anisotropy exhibits the same characteristics as before — diffuse and multi-modal, although several maximum likelihood values and the 1D average values fall within the same ballpark. Agreeing with our previous analysis on 3D synthetic example, these results suggest that it is possible to estimate the quartic moveout coefficients from a 3D gather associated with complex anisotropic media, although there is a limitation on the inherent non-uniqueness of solutions due to more diffuse and and multi-modal posterior distributions in comparison with W_i, which are generally well-constrained.

Spiral-sorted (snail) gather

Instead of looking at 3D CMP gathers, Burnett (2011) shows that the time-warping workflow can accommodate gathers with other choices of sorting scheme, which may provide greater flexibility. This is due to the fact that in the time-warping workflow, one utilizes non-physical flattening that is not restricted to any particular event shape. The algorithm merely functions based on aligning events between neighboring traces and will be able to produce meaningful results as long as the automatic picker (predictive painting) can follow the events from trace to trace.

The first and most intuitive candidate is offset-sorted gather, where traces corresponding to the same CMP position is not binned but instead sorted and shown in 2D with increasing absolute offset. This leads directly to a practical advantage of requiring only 2D slope estimation as opposed to a 3D version. Additionally, neighboring CMP gathers can be offset-sorted and juxtaposed along the third axis for simultaneous processing with 3D slope estimator and additional regularization across CMP positions. Nonetheless, in this offset-sorting scheme, neighboring traces may be from completely random azimuths (Figure 22), which leads to a nonoptimal setup for capturing azimuthal variations of moveout curves.

Alternatively, Burnett (2011) proposes that the choice of spiral sorting is particularly attractive because it holds a similar advantage of enabling the use of 2D slope estimator, yet captures azimuthal information. An illustration of a map-view spiral-sorted gather in comparison with offset-sorted gather is shown in Figure 22. We can see that spiral sorting the gather will allow the events to vary smoothly in both offset and azimuth, while maintaining a minimum physical proximity between neighboring
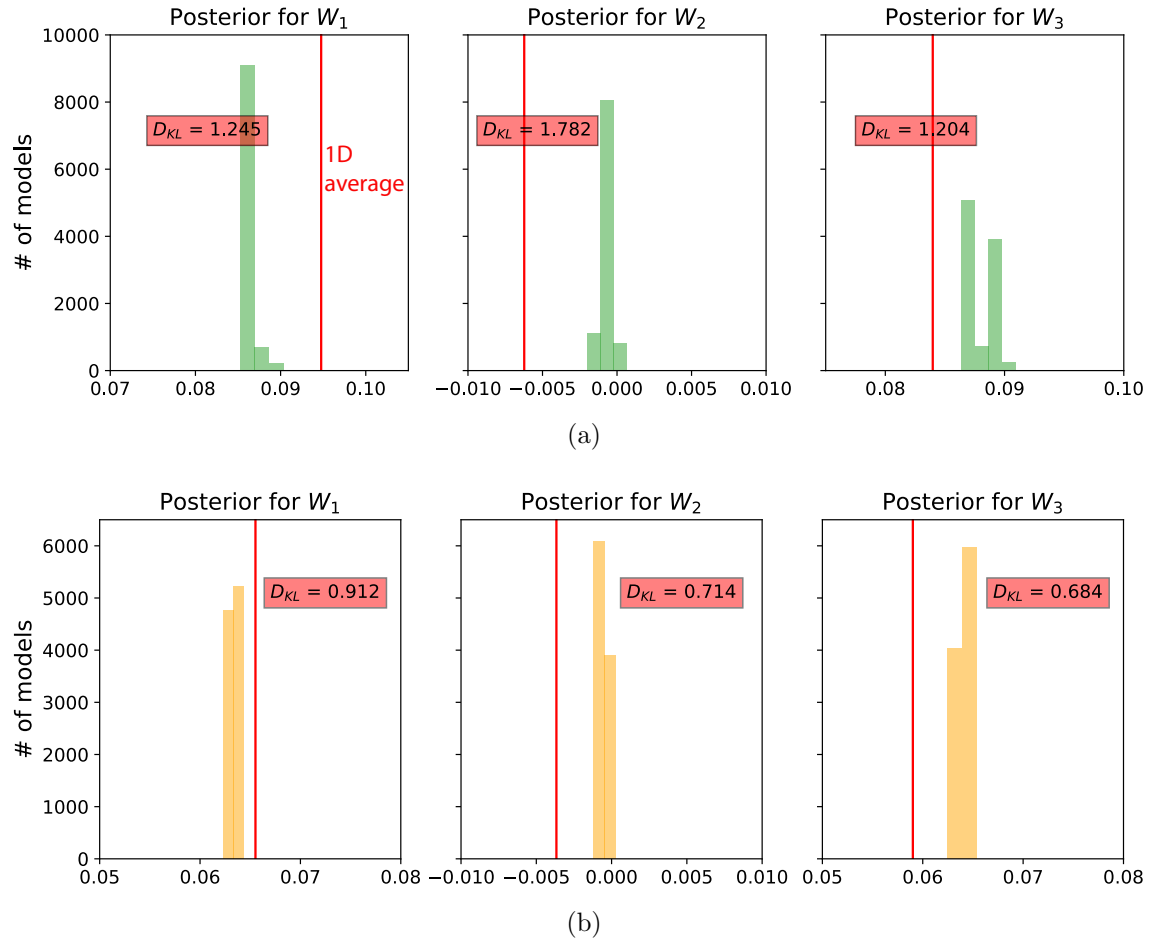
(a)

(b)

Figure 19: A comparison of estimated posterior probability distributions for W_i in the form of histograms for events at (a) 1.315 s and (b) 1.815 s of the SEAM II model. The solid vertical lines denote the 1D average values.
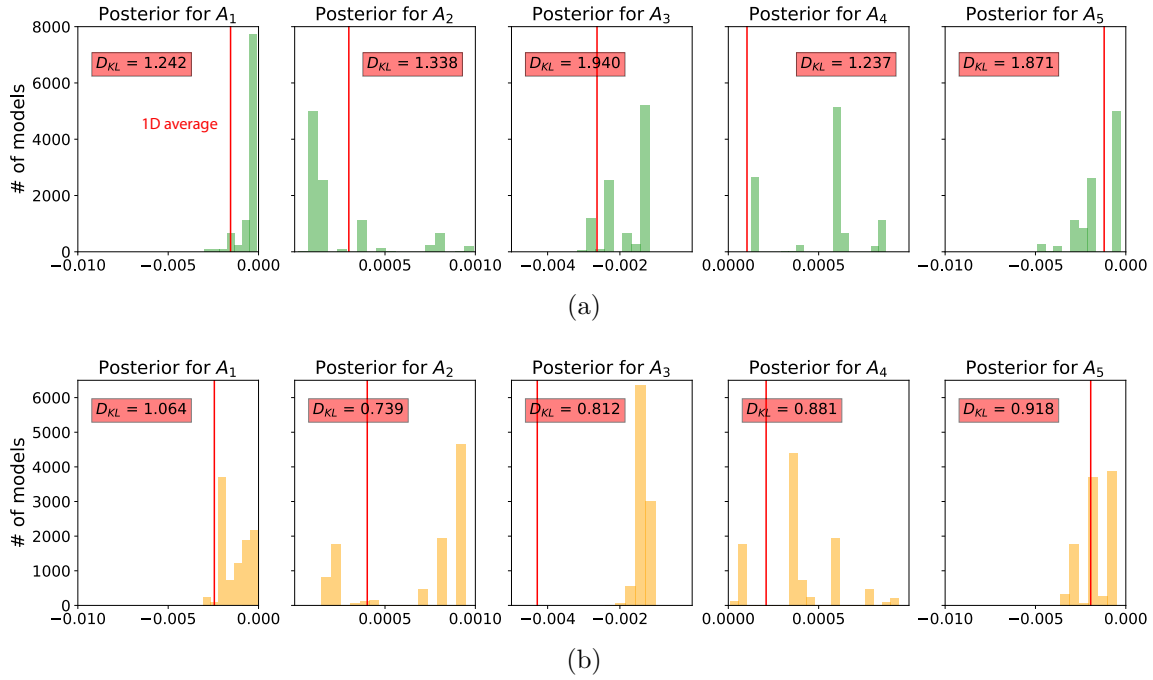
Figure 20: A comparison of estimated posterior probability distributions for A_i in the form of histograms for events at (a) 1.315 s and (b) 1.815 of the SEAM II model. They are not as well-constrained as the inverted W_1 and W_3, which indicate large uncertainties. The solid vertical lines denote the 1D average values.
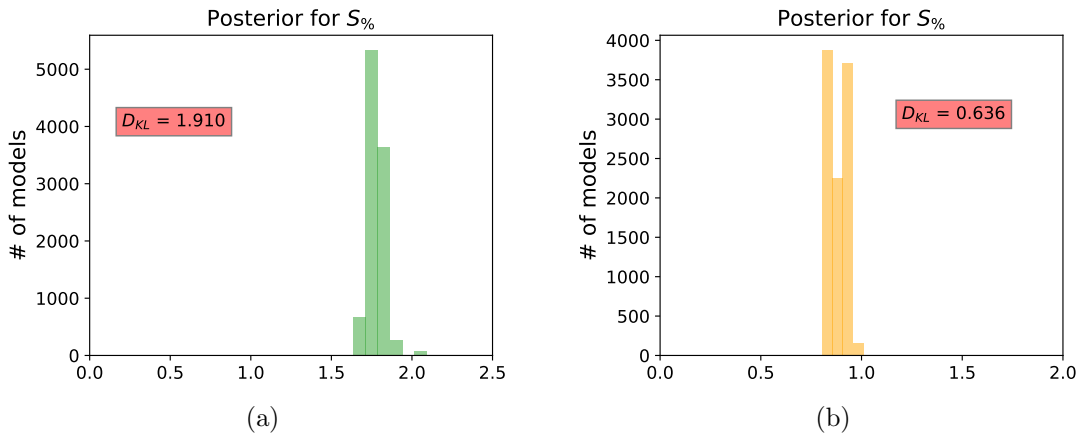

Figure 21: A comparison of estimated posterior probability distributions for $S_\%$ in the form of histograms for events at (a) 1.315 s and (b) 1.815 of the SEAM II model.
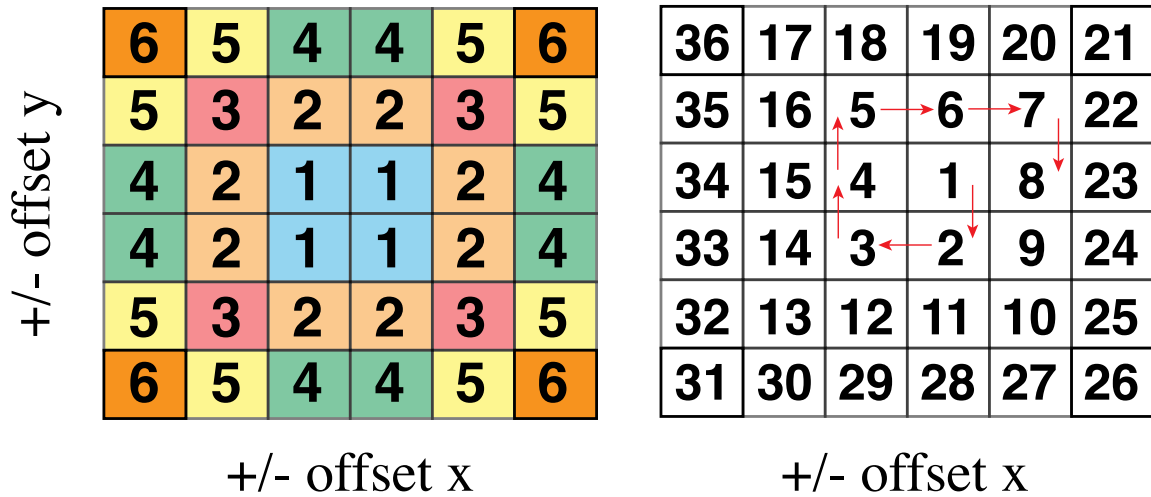
Figure 22: Map-view setups for offset-sorted gather (left) and spiral-sorted gather (right). Different numbers indicate different traces that have to be sorted from the smallest to the largest. Azimuthal information in the former is arranged inconsistently because the absolute offsets are equal for all four traces and the order in which these four traces will be arranged is random. However, in the latter, the azimuthal information can be traced smoothly from small to large radial offsets, which enhances continuity and stability for the slope estimation process.

traces.

Using the same CMP gather of SEAM II data from the previous example, we spiral sort the gather instead of binning, which enables us to look at the total 3200 traces available at this particular CMP position. Subsequent 2D slope estimation, as opposed to 3D, is carried out and the automatically traced events are shown overlaying the spiral-sorted gather in Figure 23. We use the same prior distributions as in the previous case (Table 3) and the same general setup by recording the model every 500 samples with a total of 10,000 models. Similar histograms of estimated posterior distributions for the same two events are shown in Figures 24– 26. In this particular example of SEAM II data, apart from saving computational cost by allowing 2D slope estimation, spiral sorting the gather appears to slightly improve estimated posterior probability distributions in comparison with those from the previous section (Figures 19 and 20) — some of the maximum likelihood A_i are closer to the 1D average values. Similar general observations and conclusions regarding the non-uniqueness of possible solutions can still be drawn.

DISCUSSION

We first emphasize that one of the main advantages of time-domain processing is efficiency. Despite its well-known limitations in complex geologic environments (e.g.,
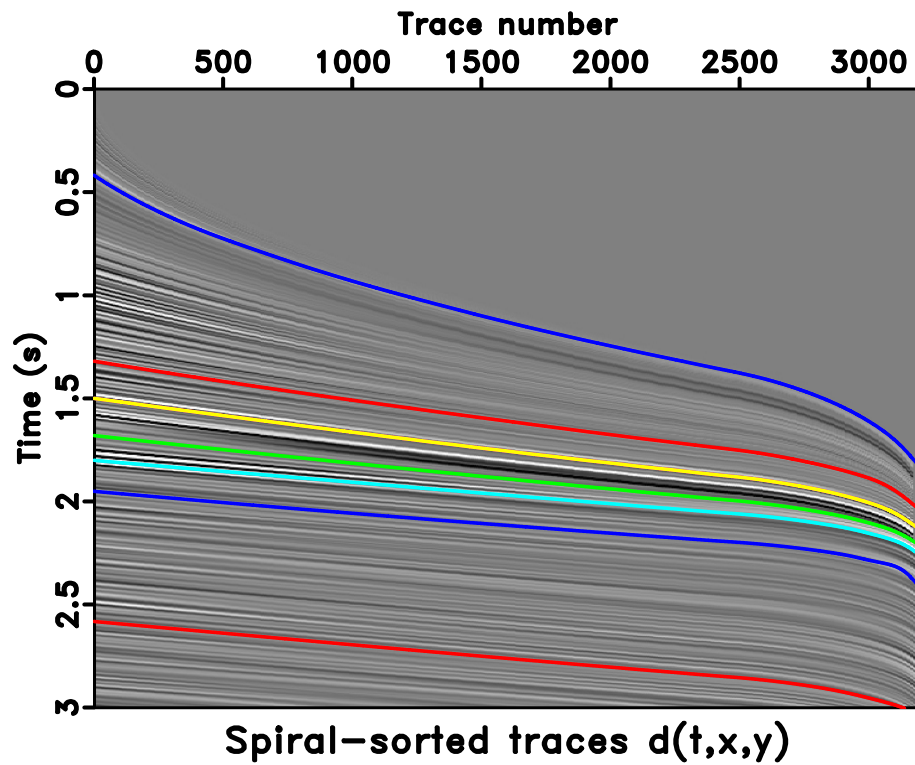
Figure 23: The spiral-sorted CMP gather of SEAM II model at x=4.375 km and y=2.881 km overlain by automatically picked events for non-physical flattening. seam2mcmcspiral/ pickongather
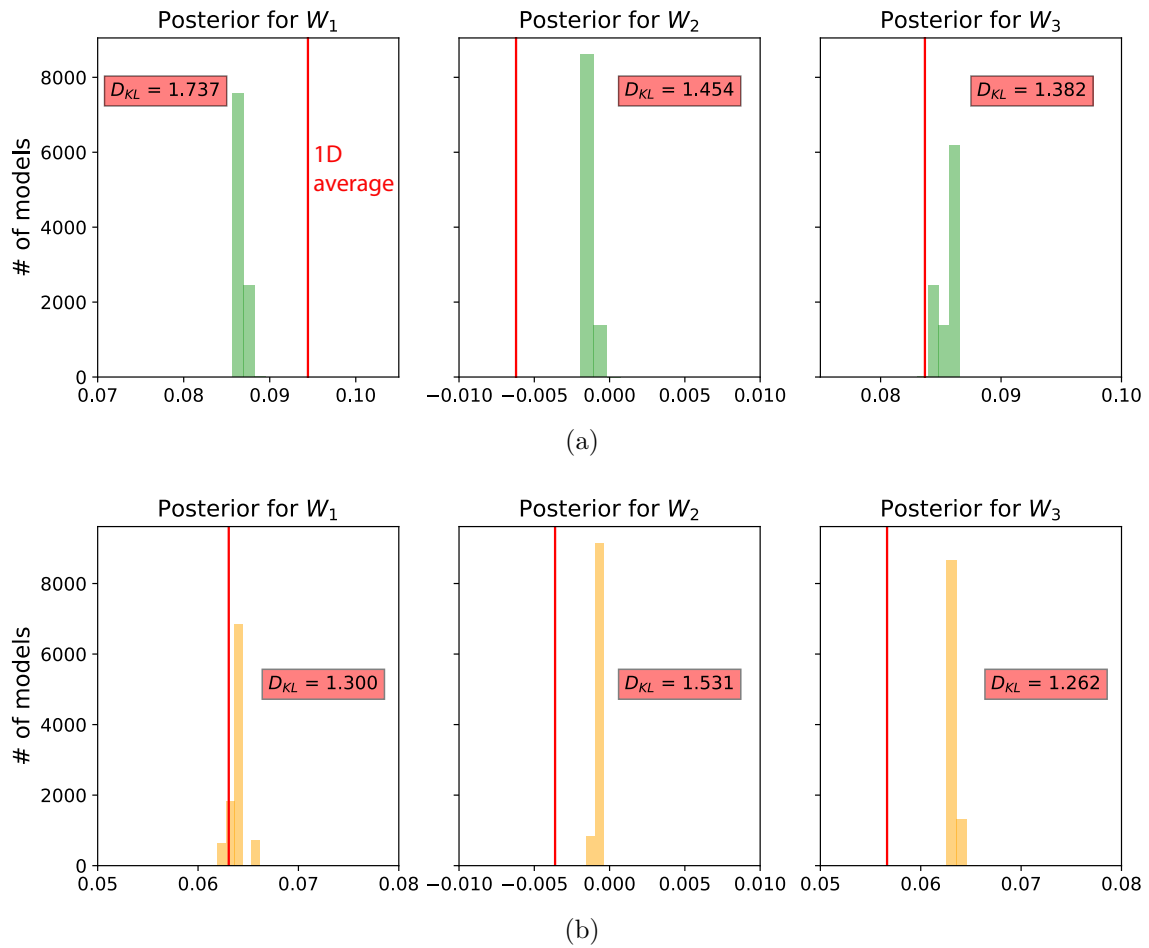
Figure 24: A comparison of estimated posterior probability distributions for W_i in the form of histograms for events at (a) $1.315\ s$ and (b) $1.815\ s$ of the SEAM II model from the inversion with spiral-sorted gather. The solid vertical lines denote the 1D average values.
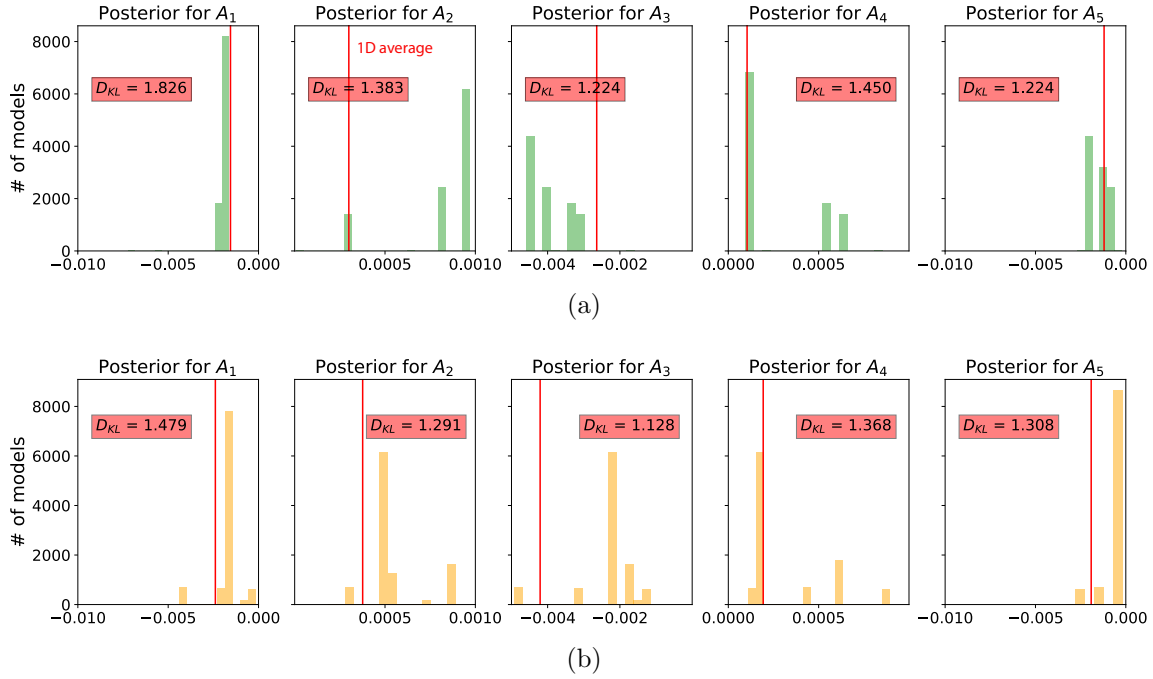
(a)

(b)

Figure 25: A comparison of estimated posterior probability distributions for A_i in the form of histograms for events at (a) 1.315 s and (b) 1.815 of the SEAM II model from the inversion with spiral-sorted gather. They are not as well-constrained as the inverted W_1 and W_3, which indicate large uncertainties. The solid vertical lines denote the 1D average values. However, some of the maximum likelihood A_i are closer to the 1D average values than the results from regularly sorted gather (Figure 20).
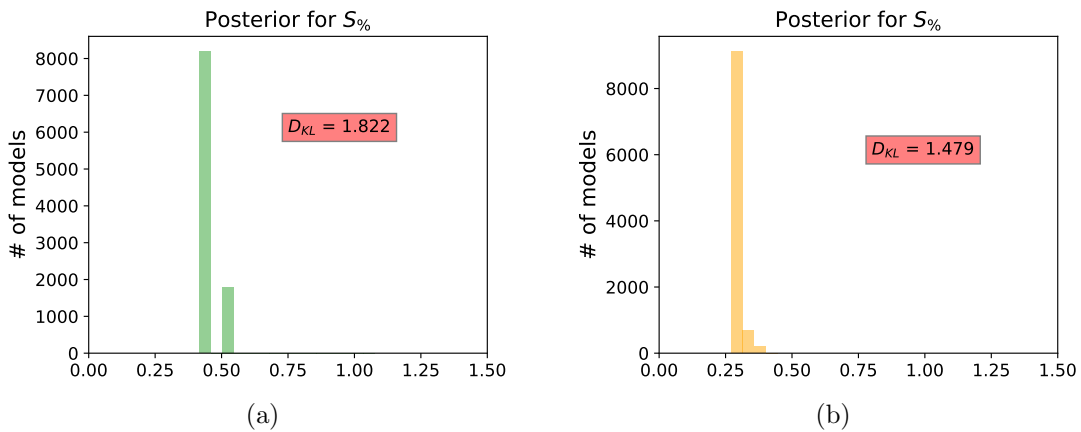

(a)   (b)

Figure 26: A comparison of estimated posterior probability distributions for $S_\%$ in the form of histograms for events at (a) 1.315 s and (b) 1.815 of the SEAM II model from the inversion with spiral-sorted gather.

subsalt), it remains an attractive choice in some other important areas such as unconventional reservoirs, where the geology is largely layered (e.g, SEAM II model in Figure 17) and an expensive treatment of depth-imaging may not be necessary (Fomel, 2014). Our time-warping approach presented here further improves the current automatic time-domain processing workflow by proposing the use of highly accurate 3D traveltime approximations (equation 4) and adding a probabilistic component for a better analysis of the inversion results. We show that these additions can easily be implemented, while maintaining the original goal of efficiency and automation. We note, however, that the degree of efficiency of our time-warping approach primarily depends on the number of sampled models, which must be tailored to any specific problem on a case-by-case basis.

In addition, even though we propose a workflow with two inversion runs to help obtain better prior distributions for W_i by taking advantage of one of the characteristics of our problem (i.e., W_i can be estimated with only small-offset data), our resulting posterior probability distributions of A_i are still estimated relative to the choice of uniform priori probability distributions. Therefore, it remains possible to improve the proposed approach if a better knowledge on the prior distributions $\rho(\mathbf{m})$ is available. For example, one may choose to use, instead of a uniform prior, a Gaussian distribution over some known mean value and standard deviation. Such information may be obtained on the basis of a previous velocity model in the region or any nearby well measurement. A better characterization of the prior probability distributions will surely enhance the efficiency of Monte Carlo sampling for solutions to the moveout inversion problem.

We note that in practice, it is sensible to consider a large range for uniform prior distributions of different parameters. However, this may not be the best choice in view of computational efficiency. We emphasize that this decision has to be made on a case-by-case basis and one may choose to follow our method by utilizing information such as relationships between GMA parameters and η in equation 13, as well as, results from forward computational experiments with 3D GMA in various media (Sripanich et al., 2017) to help coming up with a rough numeric range for each parameter.

Moreover, one could be tempted to use a different inversion sequence (e.g., estimate W_i, A_i, and $S_\%$ before B_i and C_i) with more runs to better constrain the results or to use a different parameterization as that in equation 4 to reduce the trade-offs between parameters. However, we would like to emphasize that A_i, B_i and C_i are inherently connected. This follows from the fact that to maintain the accuracy of moveout traveltime approximations, the quartic (x^4) term needs to have both the numerator (related to A_i) and the denominator (related to B_i and C_i). Without a proper denominator, the series would lead to inferior accuracy (higher error) at large offsets (Tsvankin, 2012; Sripanich et al., 2017). Therefore, it is unclear how we can separately invert for A_i, B_i and C_i in different runs without introducing any potential bias to the results. One way to approximately handle this complexity is to assume a relationship among A_i, B_i, and C_i such as that in equation 13, which effectively reduces the number of inverted parameters at the expense of the moveout approxima-

tion accuracy. The result from this experiment with our 2D inverse moveout example (Figure 3) is shown in Figure 27, where we can observe that the maximum likelihood points of estimated W and η lie close to the true values. Nonetheless, the 3D counterpart of equation 13 is not yet established for complex 3D layered orthorhombic media with azimuthally rotated principal axes such as the SEAM II model.
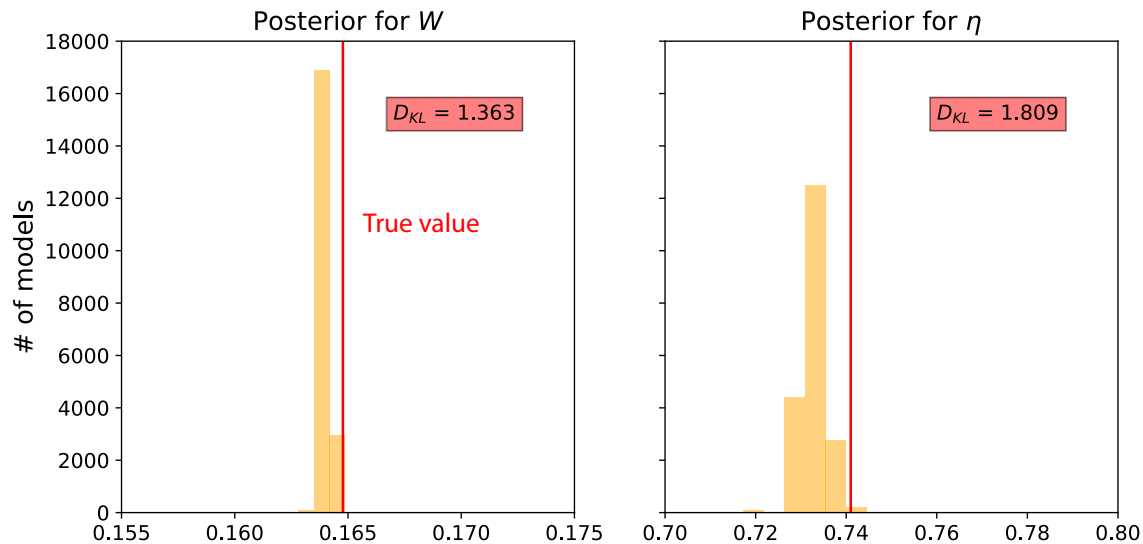

Figure 27: Estimated posterior probability distributions in the form of histograms from the proposed two-run inversion for the Dry Green River shale example with the η constraint in equation 13. The maximum likelihood points are located close to the true values suggesting that moveout parameters from 2D gathers are well recoverable from the proposed approach.

In our inversion workflow, we propose to use the Metropolis-Hastings algorithm for its simplicity. There are more advanced guided random walk methods that can lead to better performance and several of them are reviewed by Sambridge and Mosegaard (2002). Other choices of likelihood functions $L(\mathbf{m})$ are also possible (Mosegaard and Tarantola, 1995; Tarantola, 2005). Nonetheless, it is expected that our observations on the characteristics of the resulting posterior probability distributions (i.e, multimodal and diffuse) should remain valid.

With regard to the resulting posterior probability distributions, it is clear that there is non-uniqueness in the moveout inversion based on traveltime fitting and trade-offs between inverted parameters. Grechka and Tsvankin (1998b) show that the inverted nonhyperbolic moveout parameter— η in their case— is highly sensitive to the errors in reflection traveltimes and the spreadlength of available data. This observation agrees with our posterior probability distributions that can be both multi-modal and diffuse, which suggest that there are many equivalently good sets of possible moveout parameters that satisfy the same level of accuracy in predicting reflection traveltimes. Moreover, the B_i and C_i in 3D GMA also remain undetermined and contribute to the non-uniqueness in the inverted models. Therefore, several fac-

tors contribute to the non-uniqueness of possible solutions and choosing which specific choice of representative model parameters to commit to is not straightforward. The results obtained from our approach could be used as starting models for more sophisticated tomographic or full-waveform inversion techniques for more accuracy and better convergence.

In the time-warping workflow, it is important to reiterate that we do not map local slopes to moveout parameters directly as that requires a prior assumption of the traveltime approximation form (Stovas and Fomel, 2016). Rather, we rely on measured slope fields and predictive painting to automatically trace CMP events and capture the moveout curves regardless of their shape. Consequently, conditioning of local slopes field is critical to the quality of the input traveltime data for the inversion. In light of this topic, it is possible to further improve the time-warping method by addressing potential problems that may arise from interfering dips (Schleicher et al., 2009) and predictive painting across discontinuties (Xue et al., 2018). Nonetheless, given that the information on reflection traveltimes can be obtained, our proposed Bayesian inversion will produce posterior distributions of estimated parameters that reflect the level of data uncertainty.

We stress that there are many possible choices of moveout approximations that can be used. Each of them can achieve a different level of approximation accuracy and involves a different set of moveout parameters. The 3D GMA is chosen here primarily due to its convenient applicability to complex 3D anisotropic models (e.g, layered orthorhombic with unaligned vertical symmetry planes in SEAM II) and its high accuracy in traveltime prediction, which technically allows us to accurately invert for moveout parameters and relate them to subsurface properties. This comes at a price of working with a large number of moveout parameters (seventeen). Other less accurate moveout approximations with fewer parameters can be used but a poorer performance in traveltime prediction may lead also to more uncertainty in the inverted moveout parameters. Nevertheless, an analysis has to be conducted on a case-by-case basis to draw meaningful conclusions regarding the trade-offs between the accuracy of traveltime approximation and the number of resolvable moveout parameters. Therefore, in general, it appears to be practicable to perform moveout inversion in a hierarchical manner starting from 3D GMA, which is applicable to a complex 3D model before reducing the number of inverted parameters based on resolvability of parameters (e.g., with an analysis on the posterior probability distributions and D_{KL}) or different assumptions such as a simpler anisotropic model of subsurface medium.

The estimated moveout parameters (W_i and A_i) are well-known to also be affected by lateral heterogeneity (Lynn and Claerbout, 1982; Grechka and Tsvankin, 1999; Blias, 2009; Sripanich et al., 2019). Only when the subsurface is laterally homogeneous (1D) can the moveout parameters directly be related to medium parameters. In our approach, we do not specify their expressions explicitly in terms of medium parameters and treat them as 'effective' moveout parameters that are estimated based on the 3D GMA. Interval parameter estimation can subsequently be accomplished by layer-stripping given that the medium is assumed to be 1D (Sripanich and Fomel,

2016; Koren and Ravve, 2017).

Most importantly, regardless of which moveout approximation one uses, the time-warping workflow, nonetheless, remains the same. The final estimated models are expressed in the time (t_0) domain, which can be related to the depth model via time-to-depth conversion (Cameron et al., 2007; Li and Fomel, 2015; Sripanich and Fomel, 2018).

Finally, we note that alternative traveltime approximations such as various forms of common-reflection-surface (CRS) approximations exist, which consider traveltime variations in both offset and midpoint directions (e.g., Jäger et al., 2001). In general, CRS methods aim to achieve better stacked images and objective functions for global CRS parameter estimation are normally considered based on coherency measurements such as semblance (e.g., Barros et al., 2015; Garabito, 2018; Bloot et al., 2018). An alternative estimation of CRS parameters can also be done by mapping directly from local slopes and curvatures (Santos et al., 2011; Waldeland et al., 2018). In view of the proposed approach, inversion-based CRS parameter estimation methods can potentially benefit from a Bayesian formulation, which will lead to statistical posterior distributions of estimated parameters. Nonetheless, in this study, we restrict ourselves to the CMP domain and the corresponding problem of parameter estimation from reflection moveout only in the offset direction. We focus on investigating the problem of 3D anisotropic model building at the earliest stage of processing from reflection moveouts. As demonstrated in our examples, the proposed framework is relatively flexible and can be used to bring additional statistical information on inverted parameters at any selected locations (reflectors). We anticipate that this method could represent an auxiliary tool that can help assess the quality of inverted moveout parameters from other automatic estimation methods.

CONCLUSION

We propose a general workflow for moveout processing and inversion by means of time-warping. The method achieves event flattening and moveout parameter estimation in an automatic and efficient manner with minimal user inputs. We show that our method doesn't rely on multi-parameter scans but instead on data-driven slope measurements. Automatic event flattening can be achieved with the measured slopes and produces the information on moveout traveltimes for all corrected CMP events. Such information can then be used in a moveout inversion process for estimating effective parameters associated with any preferred choice of moveout approximation. Due to the cheap traveltime modeling with moveout approximation, we choose to solve this inverse problem using a global Monte Carlo optimization method. Using both synthetic and real-data examples, we demonstrate that the resulting moveout parameters estimated using our approach are represented by posterior probability distributions that contain valuable statistical information. Unlike the results from other estimation methods such as semblance-based picking and least-squares inversion, the posterior probability distributions are unbiased towards any particular solution but

represent the entire space of possible solutions and their associated probability of occurrence. This information can be used in conjunction with results from other estimation methods for better assessment and quality control of estimated parameters associated with any selected reflectors. Our results also suggest that the solutions to the problem of moveout inversion from traveltime fitting for 2D CMP gathers are generally well-constrained with only a few parameters to estimate despite a trade-off between the second-order W and quartic moveout coefficient A. On the other hand, for 3D CMP gathers associated with complex anisotropic subsurface, the solutions appear non-unique and that there are trade-offs among inverted parameters, especially the quartic coefficients (A_i). Only W_i parameters that govern the NMO ellipse seem to be generally well-constrained by the traveltime data. Further limitations on the data such as noise and sparse measurements can hamper the inversion and lead to even inferior results as shown by more diffuse estimated posterior distributions at a high level of data uncertainty. Therefore, one has to be mindful to draw conclusions on representative model parameters to be used in subsequent stages of processing and imaging, especially for 3D anisotropic models.

ACKNOWLEDGMENTS

DATA AND MATERIALS AVAILABILITY

Data associated with this research are available and can be obtained by contacting the corresponding author except the SEAM II data.

REFERENCES

Abbad, B., B. Ursin, and D. Rappin, 2009, Automatic nonhyperbolic velocity analysis: Geophysics, **74**, no. 2, U1–U12.

Al-Dajani, A., I. Tsvankin, and M. N. Toksöz, 1998, Non-hyperbolic reflection moveout for azimuthally anisotropic media: 68th Annual International Meeting Expanded Abstracts, Society of Exploration Geophysicists, 1479–1482.

Aleixo, R., and J. Schleicher, 2010, Traveltime approximations for q-P waves in vertical transversely isotropy media: Geophysical Prospecting, **58**, 191–201.

Alkhalifah, T., 1997, Velocity analysis using nonhyperbolic moveout in transversely isotropic media: Geophysics, **62**, 1839–1854.

———, 2003, An acoustic wave equation for orthorhombic anisotropy: Geophysics, **68**, 1169–1172.

———, 2011, Scanning anisotropy parameters in complex media: Geophysics, **76**, no. 2, U13–U22.

Alkhalifah, T., and I. Tsvankin, 1995, Velocity analysis for transversely isotropic media: Geophysics, **60**, 1550–1566.

Barros, T., R. Ferrari, R. Krummenauer, and R. Lopes, 2015, Differential evolution-based optimization procedure for automatic estimation of the common-reflection surface traveltime parameters: Geophysics, **80**, no. 6, WD189–W200.

Blias, E., 2009, Long-offset NMO approximation for a layered VTI model: Model study: 79th Annual International Meeting Expanded Abstracts, Society of Exploration Geophysicists, 3745–3748.

Bloot, R., T. A. Coimbra, J. H. Faccipieri, and M. Tygel, 2018, Common-reflection-surface method in weakly anisotropic vertical transverse isotropic media: Geophysics, **83**, no. 3, C99–C113.

Burnett, W., and S. Fomel, 2009, Moveout analysis by time-warping: 79th Annual International Meeting Expanded Abstracts, Society of Exploration Geophysicists, 3710–3714.

Burnett, W. A., 2011, Multiazimuth Velocity Analysis Using Velocity-Independent Seismic Imaging: PhD thesis, The University of Texas at Austin.

Cameron, M. K., S. Fomel, and J. A. Sethian, 2007, Seismic velocity estimation from time migration: Inverse Problems, **23**, 1329–1369.

Castle, R. J., 1994, Theory of normal moveout: Geophysics, **59**, 983–999.

Chen, Y., T. Liu, and X. Chen, 2015, Velocity analysis using similarity-weighted semblance: Geophysics, **80**, no. 4, A75–A82.

Farra, V., and I. Pšenčík, 2013, Moveout approximations for P- and SV-waves in VTI media: Geophysics, **78**, no. 5, WC81–WC92.

Farra, V., I. Pšenčík, and P. Jílek, 2016, Weak-anisotropy moveout approximations for P waves in homogeneous layers of monoclinic or higher anisotropy symmetries: Geophysics, **81**, no. 2, C17–C37.

Fomel, S., 2002, Applications of plane-wave destruction filters: Geophysics, **67**, 1946–1960.

———, 2007, Velocity-independent time-domain seismic imaging using local slopes: Geophysics, **72**, no. 3, S139–S147.

——, 2009, Velocity analysis using AB semblance: Geophysical Prospecting, **57**, 311–321.

——, 2010, Predictive painting of 3-D seismic volumes: Geophysics, **75**, no. 4, A25–A30.

——, 2014, Recent advances in time-domain seismic imaging: 84th Annual International Meeting Expanded Abstracts, Society of Exploration Geophysicists, 4400–4404.

Fomel, S., and A. Stovas, 2010, Generalized nonhyperbolic moveout approximation: Geophysics, **75**, no. 2, U9–U18.

Garabito, G., 2018, Global optimization strategies for implementing 3D common-reflection-surface stack using the very fast simulated annealing algorithm: Application to real land data: Geophysics, **83**, no. 4, V253–V261.

Golikov, P., and A. Stovas, 2012, Accuracy comparison of nonhyperbolic moveout approximations for qP-waves in VTI media: Journal of Geophysics and Engineering, **9**, 428–432.

Grechka, V., and I. Tsvankin, 1998a, 3-D description of normal moveout in anisotropic inhomogeneous media: Geophysics, **63**, no. 3, 1079–1092.

——, 1998b, Feasibility of nonhyperbolic moveout inversion in transversely isotropic media: Geophysics, **63**, no. 3, 957–969.

——, 1999, 3-D moveout inversion in azimuthally anisotropic media with lateral velocity variation: Theory and a case study: Geophysics, **64**, 1202–1218.

Hake, H., K. Helbig, and C. S. Mesdag, 1984, Three-term Taylor series for $t^2 - x^2$ curves over layered transversely isotropic ground: Geophysical Prospecting, **32**, 828–850.

Jäger, R., J. Mann, G. Höcht, and P. Hubral, 2001, Common-reflection-surface stack: Image and attributes: Geophysics, **66**, 97–109.

Keys, R. G., and D. J. Foster, 1998, Comparison of seismic inversion methods on a single real data set: Society of Exploration Geophysicists Press.

Kirlin, R. L., 1992, The relationship between semblance and eigenstructure velocity estimators: Geophysics, **57**, no. 8, 1027–1033.

Kirlin, R. L., and W. J. Done, 2001, Covariance Analysis for Seismic Signal Processing: Society of Exploration Geophysicists.

Koren, Z., and I. Ravve, 2017, Normal moveout coefficients for horizontally layered triclinic media: Geophysics, **82**, no. 4, WA119–WA145.

Li, S., and S. Fomel, 2015, A robust approach to tome-to-depth conversion and interval velocity estimation from time migration in the presence of lateral velocity variation: Geophysical Prospecting, **63**, 315–337.

Luo, S., and D. Hale, 2012, Velocity analysis using weighted semblance: Geophysics, **77**, no. 2, U15–U22.

Lynn, W. S., and J. F. Claerbout, 1982, Velocity estimation in laterally varying media: Geophysics, **47**, no. 6, 884–897.

Malinverno, A., and V. A. Briggs, 2004, Expanded uncertainty quantification in inverse problems: Hierarchical Bayes and empirical Bayes: Geophysics, **69**, no. 4, 1005–1016.

Masmoudi, N., A. Stovas, and T. Alkhalifah, 2017, Scanning anisotropy parameters

in horizontal transversely isotropic media: Geophysical Prospecting, **65**, no. 4, C127–C137.

Mosegaard, K., and A. Tarantola, 1995, Monte Carlo sampling of solutions to inverse problems: Journal of Geophysical Research, **100**, 12431–12447.

Neidell, N. S., and M. T. Taner, 1971, Semblance and other cohenrency measures for multichannel data: Geophysics, **36**, no. 3, 482–497.

Oh, J., and T. Alkhalifah, 2016, The scattering potential of partial derivative wavefields in 3D elastic orthorhombic media: An inversion prospective: Geophysical Journal International, **206**, 1740–1760.

Oristaglio, M., 2015, Elastic anistropy in SEAM Phase II models: The Leading Edge, **34**, no. 8, 964–970.

Ottolini, R., 1983, Velocity independent seismic imaging: SEP, **37**, 59–68.

Sambridge, M., T. Bodin, K. Gallagher, and H. Tkalčić, 2013, Transdimensional inference in the geosciences: Philosophical Transactions of The Royal Society A Mathematical Physical and Engineering Sciences, **371**, 20110547.

Sambridge, M., K. Gallagher, A. Jackson, and P. Rickwood, 2006, Trans-dimensional inverse problems, model comparison and the evidence: Geophysical Journal International, **167**, 528–542.

Sambridge, M., and K. Mosegaard, 2002, Monte Carlo methods in geophysical inverse problems: Reviews of Geophysics, **40**, 1–29.

Santos, L., J. Schleicher, J. C. Costa, and A. Novais, 2011, Fast estimation of common-reflection-surface parameters using local slopes: Geophysics, **76**, no. 2, U23–U34.

Sarkar, D., R. T. Baumel, and K. L. Larner, 2002, Velocity analysis in the presence of amplitude variation: Geophysics, **67**, 1664–1672.

Sarkar, D., J. P. Castagna, and W. J. Lamb, 2001, AVO and velocity analysis: Geophysics, **66**, 1284–1293.

Schleicher, J., J. C. Costa, L. T. Santos, A. Novais, and M. Tygel, 2009, On the estimation of local slopes: Geophysics, **74**, no. 4, P25–P33.

Sripanich, Y., and S. Fomel, 2016, Theory of interval parameter estimation in layered anisotropic media: Geophysics, **81**, no. 5, C253–C263.

———, 2018, Fast time-to-depth conversion and interval velocity estimation in the case of weak lateral variations: Geophysics, **83**, no. 3, S227–S235.

Sripanich, Y., S. Fomel, and A. Stovas, 2019, Effects of lateral heterogeneity on time-domain processing parameters: Geophysical Journal International, **219**, no. 2, 1181–1201.

Sripanich, Y., S. Fomel, A. Stovas, and Q. Hao, 2017, 3D generalized nonhyperboloidal moveout approximation: Geophysics, **82**, no. 2, C49–C59.

Stovas, A., 2015, Azimuthally dependent kinematic properties of orthorhombic media: Geophysics, **80**, no. 6, C107–C122.

Stovas, A., and S. Fomel, 2016, Mapping of moveout attributes using local slopes: Geophysical Prospecting, **64**, 31–37.

Taner, T., and F. Koehler, 1969, Velocity spectra – Digital computer derivation and applications of velocity functions: Geophysics, **34**, 859–881.

Tarantola, A., 2005, Inverse Problem Theory: Society for Industrial and Applied Mathematics.

Thomsen, L., 1986, Weak elastic anisotropy: Geophysics, **51**, 1954–1966.

——, 2014, Understanding Seismic Anisotropy in Exploration and Exploitation, 2 ed.: Society of Exploration Geophysicists and European Association of Geoscientists and Engineers.

Tsvankin, I., 2012, Seismic Signatures and Analysis of Reflection Data in Anisotropic Media: Society of Exploration Geophysicists.

Tsvankin, I., and L. Thomsen, 1994, Non-hyperbolic reflection moveout in anisotropic media: Geophysics, **59**, 1290–1304.

Ursin, B., and A. Stovas, 2006, Traveltime approximations for a layered transversely isotropic medium: Geophysics, **71**, no. 2, D23–D33.

Vasconcelos, I., and I. Tsvankin, 2006, Non-hyperbolic moveout inversion of wide-azimuth P-wave data for orthorhombic media: Geophysical Prospecting, **54**, 535–552.

Waldeland, A. U., H. Zhao, J. H. Faccipieri, A. H. S. Solberg, and L. J. Gelius, 2018, Fast and robust common-reflection-surface parameter estimation: Geophysics, **83**, no. 1, O1–O13.

Wilson, H., and L. Gross, 2017, Amplitude variation with offset-friendly bootstrapped differential semblance: Geophysics, **82**, no. 5, V297–V309.

Wolf, K., D. Rosales, A. Guitton, and J. Claerbout, 2004, Robust moveout without velocity picking: 74st Annual International Meeting Expanded Abstracts, Society of Exploration Geophysicists, 2423–2426.

Xue, Z., X. Wu, and S. Fomel, 2018, Predictive painting across faults: Interpretation, **6**, no. 2, T449–T455.

Yan, J., and I. Tsvankin, 2008, AVO-sensitive semblance analysis for wide-azimuth data: Geophysics, **73**, no. 2, U1–U11.

Yilmaz, O., 2001, Seismic data analysis, 2 ed.: Society of Exploration Geophysicists.