Wave-equation migration velocity analysis using partial-stack power maximization

Yang Zhang and Guojian Shan

ABSTRACT

We proposed to use a partial-stack power maximization objective function in wave-equation migration velocity analysis. Instead of stacking the angle-domain common-image gathers all at once, the partial-stack power maximization objective function stacks them in smaller groups. It improves the robustness against the cycle-skipping problem and can achieve better global convergence. We also added a normalization term to the partial-stack power maximization objective function to balance the different reflector amplitudes. We tested our objective function using the Marmousi model. The results demonstrate that using the partial-stack power maximization criterion can achieve better global convergence. We also observed that the normalization of reflector amplitudes is very important in order to better constrain the tomography problem.

INTRODUCTION

Introduced by Gardner (1974) and Sattlegger (1975), migration velocity analysis (MVA) belongs to a family of methods for estimation of migration velocity. Instead of looking at travel-times in the seismic data, MVA extracts the velocity information from the migrated images. Etgen (1990) and van Trier (1990) proposed the first formulations of tomographic MVA using surface offset-domain common-image-gathers (ODCIGs) obtained by Kirchhoff migration. Recently, the tomographic MVA method has been extended to wave-equation migration velocity analysis (WEMVA), which uses the wave-equation rather than the ray-based model to resolve the velocity information (Chavent and Jacewitz, 1995; Biondi and Sava, 1999). The wave-equation-based methods are more accurate than ray-based methods, because the wave-equation better describes wave-propagation physics and provides physically more realistic sensitivity kernels for the velocity update. The wave-equation model can behave quite differently than the ray-based model when applied to complex velocity models.

To estimate velocity, WEMVA solves an optimization problem. Evaluating the flatness of the subsurface angle-domain common-image gathers (ADCIGs) is currently a popular choice when forming WEMVA objective functions (Biondi and Sava, 1999; Clapp and Biondi, 2000; Biondi and Symes, 2004). The objective function is usually optimized by applying gradient-based algorithms. The computation of the gradient is performed in two steps: 1) computation of a perturbation in the migrated image

domain, and 2) back-projection of the image perturbation into the velocity model using the image-space wave-equation tomographic (ISWET) operator (Sava and Biondi, 2004).

Several ADCIGs-based WEMVA objective functions have been proposed in the literature. The stack power maximization (SPM) method directly maximizes the stack of the ADCIGs of all detectable angles, but similar to full-waveform inversion (FWI) (Tarantola, 1984), it is prone to cycle-skipping when the starting model is not sufficiently accurate and the data do not contain very low frequencies (Symes, 2008). Differential-semblance optimization (DSO) (Symes and Carazzone, 1991; Shen et al., 2005; Shen and Symes, 2008) penalizes the difference between the ADCIGs of neighboring angles (or the unfocused energy on the *subsurface* ODCIGs). This objective function can achieve much better global convergence. However, the differencing operator amplifies the high-frequency (with respect to the angle axis) image variations in the ADCIGs and can generate unwanted artifacts in the gradient (Fei and Williamson, 2010), which slows down convergence.

If we compare the two objective functions that previously mentioned, notice that "stacking all angles" is a special case of the smoothing operations, and it extracts nothing but the zero-frequency component (with regard to the angle axis), while "differencing neighboring angles" extracts all frequency components but boosts higher-frequency components. The partial-stack power maximization objective function is a compromise between the two. It utilizes many non-zero frequency components as DSO does, while still using the stacking operator (in contrast to the differencing operator) with smaller windows. Having higher-frequency components in the objective functions ensures better global convergence, and not using the differencing operator avoids amplifying the high-frequency noise in the ADCIGs. Therefore the partial-stack power maximization objective function combines the merits of the SPM and DSO objective functions.

The rest of this paper is divided into two parts: first we present the mathematical formulation; then we demonstrate the effectiveness of our method with the Marmousi examples.

THEORY

Partial-Stack power maximization

For the sake of simplicity, we assume two-dimensions in our derivation; however, extending the theory to 3-D is straightforward for this method. We denote the prestack image as $I(z, \gamma, x)$, (x, z are the depth and horizontal axis, and γ is the reflection-aperture angle).

To enforce the goal of ADCIG flatness, we have multiple options in choosing the objective functions. The stack power maximization (Soubaras and Gratacos, 2007)

maximizes the full angle stack of the ADCIG:

$$\max_v J_{SPM}(v) = \frac{1}{2}\|\sum_\gamma I(z,x,\gamma;v)\|_2^2, \tag{1}$$

in which $I(z,x,\gamma;v)$ is the ADCIGs migrated using the current velocity model v. This approach can yield a high-resolution model, however when the velocity error is large, the angle gathers will become strongly curved and demonstrate significant residual moveout (RMO) (different amounts of event shifts at different angles). Because we stack all angles at once, as the difference of event shifts between angles becomes bigger than half wavelength of the dominant frequency, the stacking becomes incoherent and would result in two separate events instead of one. The cycle-skipping phenomenon arises from such situations.

As another option to enforce angle-gather flatness, the differential semblance optimization (DSO) objective function proposed by Shen et al. (2005) overcomes this cycle-skipping issue by using a local operator that operates on each individual angle and its immediate neighbors:

$$\min_v J_{DSO}(v) = \frac{1}{2}\|\frac{\partial I(z,x,\gamma;v)}{\partial \gamma}\|_2^2. \tag{2}$$

Cycle-skipping is very unlikely to happen using the DSO objective function, because the relative shifts in the gathers between one angle and its neighbors is generally very small.

However, the differencing operator is poorly conditioned (i.e., the operator is very short, spanning over only two angles and therefore requires many iterations to make all angles the same), and it magnifies the high-frequency variations of the ADCIGs along the axis of reflection angle. In several cases these high-frequency variations are not desired (for example, when high-frequency noise is present in ADCIGs, or when the variations are caused mainly by non-uniform subsurface illumination at each reflection angle).

In order to combine the advantages of both approaches, Shen and Symes (2008) introduce a bi-objective function that includes both terms using the weighted sum:

$$\min_v J_{CMB}(v) = J_{DSO}(v) - \beta J_{SPM}(v). \tag{3}$$

As expected, this approach can achieve both global and local convergence, but the disadvantage brought by the differencing operator still remains, and adjusting the parameter β ($\beta \geq 0$) might not be trivial.

The partial-stack power maximization (partial SPM) objective function is an alternative way to combine the SPM and DSO objective functions. We use a partially stacking operator that has a span between those of the full stacking operator and the differencing operator:

$$\max_v J_{PSPM}(v) = \frac{1}{2}\|\sum_\gamma \{g(\gamma) * I(z,x,\gamma;v)\}\|_2^2, \tag{4}$$

in which $*$ means convolution and $g(\gamma)$ is the windowing function whose support corresponds to the span of the partial-stacking operator. The partial-stacking objective function serves as a transition between SPM and DSO. SPM uses only the zero-frequency component (with respect to the angle axis), and DSO uses all components but zero-frequency; while the partial SPM uses both zero and non-zero frequency components. Nonetheless, partial-stack is still a low-pass stacking operator, thus it does not amplify high frequencies. Additionally, partial SPM objective function has a single term, and the user does not have to choose proper values for the relative weight parameter β as in objective function (3).

In practice, subsurface offset CIGs are more convenient for implementation than subsurface angle-domain CIGs. To find out the subsurface offset-domain counterparts of the objective functions we just discussed, Sava and Fomel (2003) showed that the transform between subsurface offset gathers and angle gathers is analogous to a Fourier transform with respect to γ and h; that is, convolution in the angle-domain corresponds to multiplication in the offset domain. Therefore we can see that for the SPM objective function, the offset-domain counterpart is to maximize the zero subsurface-offset image (because the Fourier transform of a constant function is a spike at the origin):

$$\max_v J_{SPM\_O}(v) = \frac{1}{2}||I(z,x,h=0;v)||_2^2; \tag{5}$$

for the DSO objective function in the offset-domain, the differencing operator will map to a weighting function $f(h) = h$:

$$\min_v J_{DSO\_O}(v) = \frac{1}{2}||I(z,x,h;v)h||_2^2. \tag{6}$$

Following the same logic, the partial SPM objective function will be:

$$\max_v J_{PSPM\_O}(v) = \frac{1}{2}||I(z,x,h;v)G(h)||_2^2, \tag{7}$$

in which $G(h)$ corresponds to the Fourier transform of $g(\gamma)$. Figure 1 illustrates the comparisons of the three objective functions in the angle-domain (a) and offset-domain (b).

Balancing the impact of strong and weak reflectors

Notice that all of the objective functions we discussed previously implicitly put more weight on the strong-amplitude events, which causes the inversion to spend very little effort on the unflatness of the ADCIGs at weak reflectors. Not only will the inverse problem become less constrained, but also the inverted model might depart far from the true model, as we can see in the examples later.

There are multiple ways to normalize these objective functions so that they are more independent of reflector strength; here, for simplicity, we choose to normalize

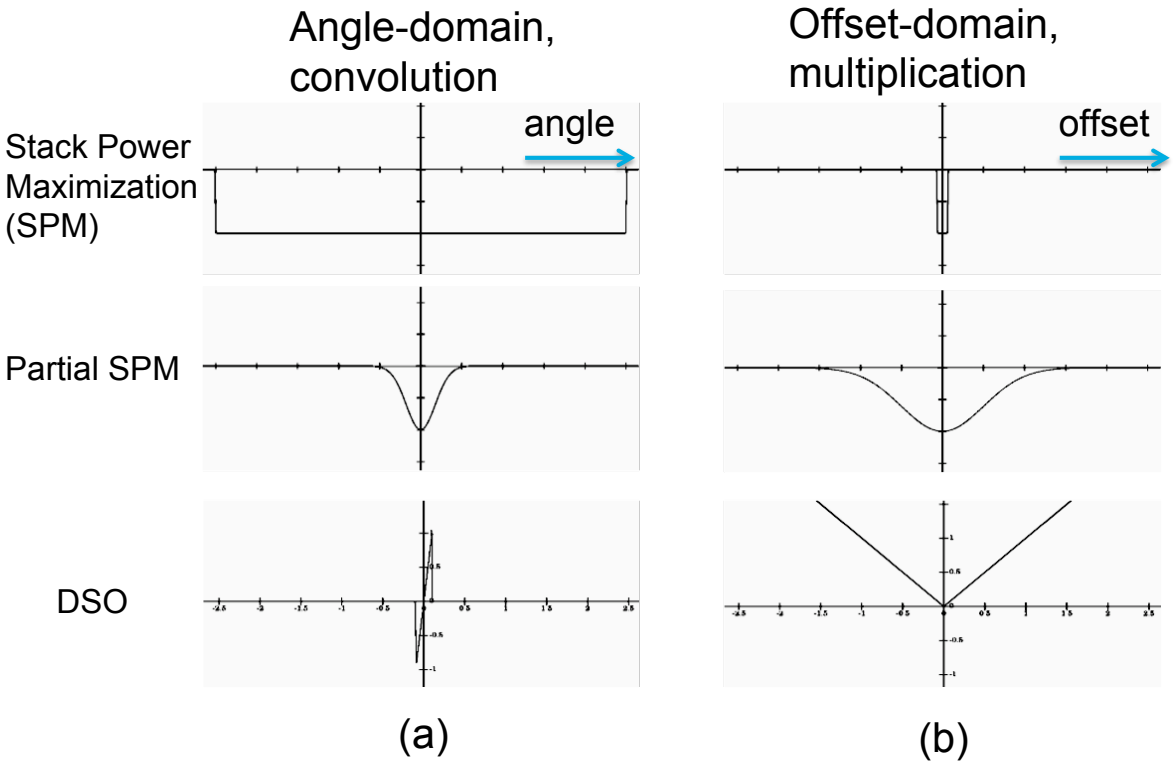

Figure 1: The comparison of the three objective functions in angle-domain(a) and offset-domain(b). [**NR**]

them for each inline location x (Tang, 2011). The normalized version of objective function (7) is:

$$\max_v J_{PSPM\_OB}(v) = \frac{1}{2}\sum_x \frac{\sum_{z,h}[I(z,x,h;v)G(h)]^2}{\sum_{z,h} I^2(z,x,h)}. \tag{8}$$

The next step is to find the model gradient from the objective function, as we will use gradient-based methods to solve this optimization problem. For conciseness and without bringing confusion, we omit the variables in $I(z,x,h;v)$ and simply denote it as I, and define

$$\bar{I} = \sum_{z,h} I^2 \qquad \text{and} \qquad \bar{I_G} = \sum_{z,h} I^2 G^2(h). \tag{9}$$

Then the gradient of the objective function (8) is

$$\frac{\partial J_{PSPM\_OB}}{\partial v} = \frac{\partial I}{\partial v}\frac{(\bar{I}G^2(h) - \bar{I_G})I}{\bar{I}^2}. \tag{10}$$

Eq. (10) indicates that first we compute the image perturbation $\Delta I = \frac{(\bar{I}G^2(h)-\bar{I_G})I}{\bar{I}^2}$; then we back-project ΔI using the image-space wave-equation tomographic operator.

NUMERICAL EXAMPLES

We use the synthetic Marmousi model to test the effectiveness of the partial SPM objective function. In our implementation, we use a two-way acoustic wave-equation propagator, and for our optimization algorithm, we use non-linear conjugate-gradient with Polak-Ribiere formula for search direction.

We implement the offset-domain representation of the partial SPM objective function (eq. (7) and eq. (8)). The selection of weighting function $G(h)$ is not unique, as long as it can be considered as the frequency spectrum of a certain low-pass filter. In our examples, we use Gaussian functions for $G(h)$. We started with a wide $G(h)$ that would include most unfocused energy of the ODCIGs at non-zero subsurface offsets, as the inversion proceeds, the gathers will become more focused; we then reduce the width of $G(h)$ so that our objective function gradually approaches the traditional SPM objective function. To control the resolution of the inversion, we precondition the model gradient with a triangular smoothing operator and reduce the extent of smoothing gradually within each iteration.

Marmousi Example

The model size is 9 km in x and 3.2 km in z. The spatial sampling is 20 m. The survey geometry is of split-spread type, with sources and receivers located on the top of the

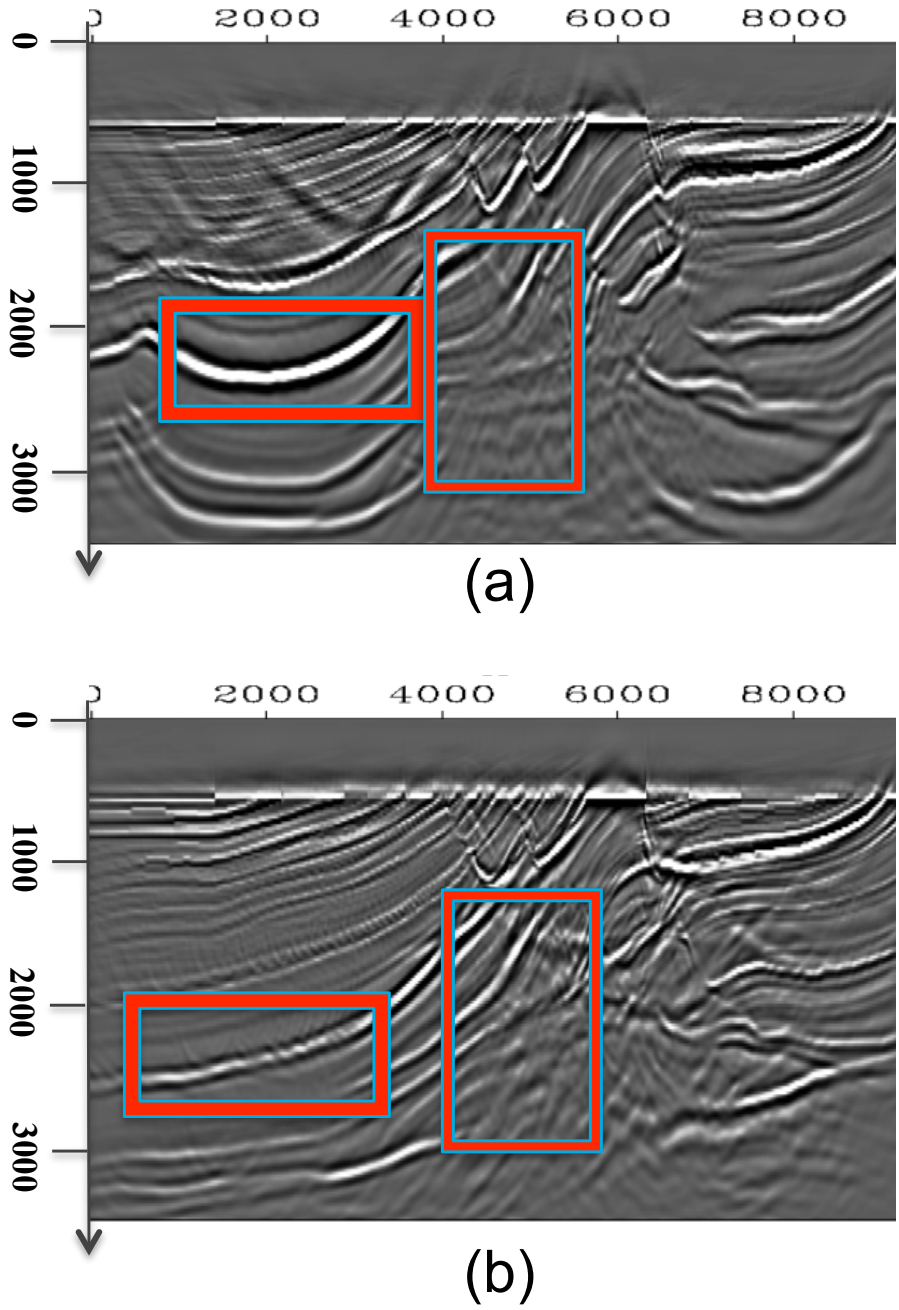

Figure 2: The migrated image (a) using the velocity model inverted from the objective function (7) (without normalization) and (b) using the initial model. The display clip we use for (a) is 4 times of the clip used in (b), and we can see that in (a) the single reflector marked in the left box becomes very strong and coherent. In contrast, many weaker reflectors (for e.g., marked in the right box) that are initially present in the initial image (a) are imaged much poorly in (b). Therefore it is important to give the weak reflectors more weight in the objective function. [**NR**]

model. There are 451 receivers with 20 m spacing that fully covers the model surface. We simulate 226 shots with shot spacing of 40 m. The central frequency of the Ricker wavelet we use is 10 Hz. For the inversion, we invert all frequencies simultaneously rather than from low to high frequencies. We run 40 nonlinear iterations for the inversion.

Figure 4(a) shows the true velocity model and 4(b) shows the starting model $v(z)$. Figures 4(c) and (d) show the inverted velocity model using the un-normalized partial SPM objective function and the normalized partial SPM objective function respectively. As we can see, the inverted result in 4(c) is stuck in local minima, because without normalization, the inversion will attempt to increase the focusness of a few large, strong reflectors, while ignoring and even sacrificing the coherency of smaller, weaker events. The observation of the corresponding migrated images in figure 2 further confirms our conclusion. The result in figure 4(d) shows global convergence towards the correct model. The long-wavelength part of the velocity model is well captured up to depth 2.2 km, as can be seen from the migrated images in figure 3.

CONCLUSION

We propose to use the partial stack power maximization objective function in wave-equation migration velocity analysis. This objective function merges the advantages of both conventional stack power maximization and differential-semblance-optimization objective functions, and it can achieve good global convergence, while retaining the relatively high resolution of the stack power maximization objective function. We have successfully applied our approach to the Marmousi model. We have also verified that the normalization of the reflector amplitude in the objective function is not only preferred but necessary in the Marmousi case.

ACKNOWLEDGEMENT

The authors thank Chevron Energy Technology Company for the permission to publish, and thank Yue Wang and Lin Zhang for suggestions and comments. The first author thanks Prof. Claerbout for the initial discussion on WEMVA objective functions.

REFERENCES

Biondi, B. and P. Sava, 1999, Wave-equation migration velocity analysis: SEG Technical Program Expanded Abstracts, **18**, 1723–1726.

Biondi, B. and W. W. Symes, 2004, Angle-domain common-image gathers for migration velocity analysis by wavefield-continuation imaging: Geophysics, **69**, 1283–1298.

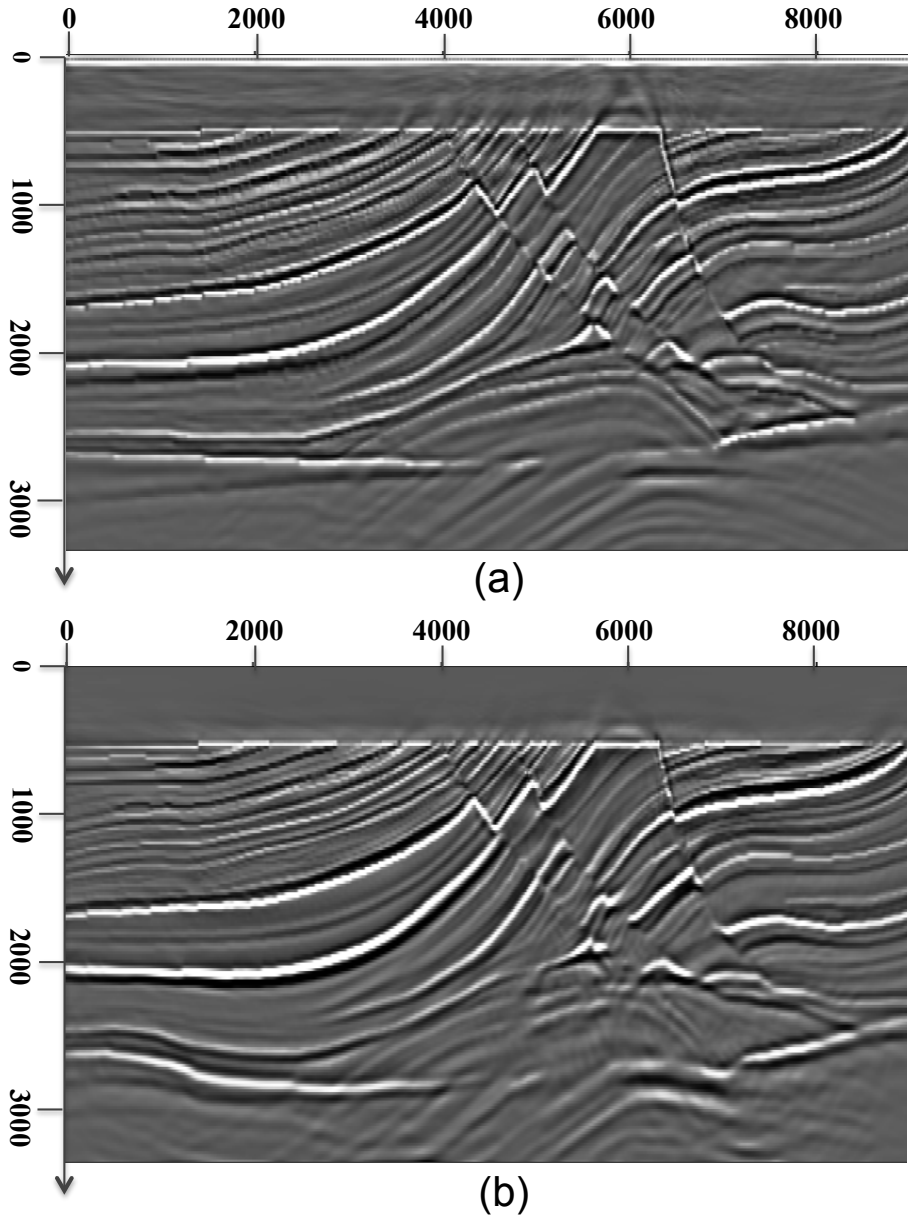

Figure 3: The comparison of the migrated images using the true velocity model (a) and using the inverted model with normalized partial SPM objective function (b). The two images match well up to 2.2 km depth. [**NR**]

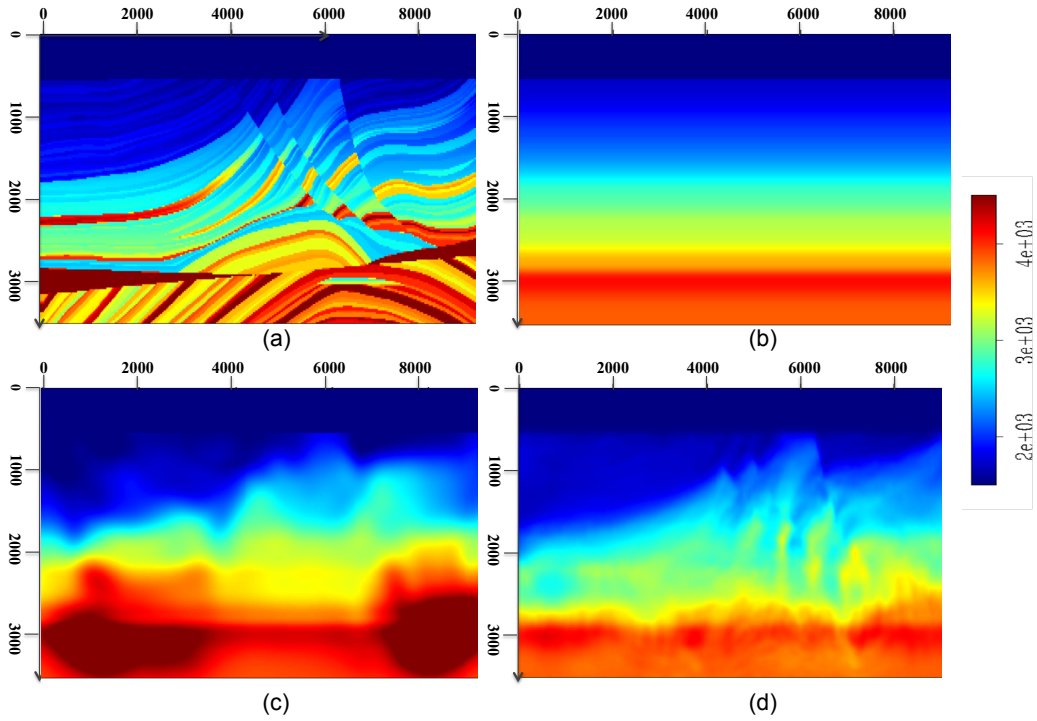

Figure 4: The Marmousi true velocity model (a), the starting velocity model (b), the inverted velocity models using partial SPM objective function without normalization (c) (eq. (7)) and with normalization (d) (eq. (8)). We can see that without normalization, the inversion yield a model that is very different from the true one. [**NR**]

Chavent, G. and C. A. Jacewitz, 1995, Determination of background velocities by multiple migration fitting: Geophysics, **60**, 476–490.

Clapp, R. G. and B. Biondi, 2000, Tau domain migration velocity analysis using angle crp gathers and geologic constraints: SEG Technical Program Expanded Abstracts, **19**, 926–929.

Etgen, J., 1990, Residual prestack migration and interval velocity estimation: PhD thesis, Stanford University.

Fei, W. and P. Williamson, 2010, On the gradient artifacts in migration velocity analysis based on differential semblance optimization: SEG Technical Program Expanded Abstracts, **29**, 4071–4076.

Gardner, G. H. F., 1974, Elements of migration and velocity analysis: Geophysics, **39**, 811.

Sattlegger, J. W., 1975, Migration velocity determination: Part i. philosophy: Geophysics, **40**, 1–5.

Sava, P. and B. Biondi, 2004, Wave-equation migration velocity analysis. I. Theory: Geophysical Prospecting, **52**, 593–606.

Sava, P. C. and S. Fomel, 2003, Angle-domain common-image gathers by wavefield continuation methods: Geophysics, **68**, 1065–1074.

Shen, P. and W. W. Symes, 2008, Automatic velocity analysis via shot profile migration: Geophysics, **73**, VE49–VE59.

Shen, P., W. W. Symes, S. Morton, A. Hess, and H. Calandra, 2005, Differential semblance velocity analysis via shot profile migration: SEG Technical Program Expanded Abstracts, **24**, 2249–2252.

Soubaras, R. and B. Gratacos, 2007, Velocity model building by semblance maximization of modulated-shot gathers: Geophysics, **72**, U67–U73.

Symes, W., 2008, Migration velocity analysis and waveform inversion: Geophysical Prospecting, **56**, 765–790.

Symes, W. W. and J. J. Carazzone, 1991, Velocity inversion by differential semblance optimization: Geophysics, **56**, 654–663.

Tang, Y., 2011, Imaging and velocity analysis by target-oriented wavefield inversion: PhD thesis, Stanford University.

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

van Trier, J., 1990, Tomographic determination of structural velocities from depth migrated seismic data: **66**.