Stanford Exploration Project, Report 94, November 11, 1997, pages 223–239

Stanford Exploration Project, Report 94, November 11, 1997, pages 223–239

Attenuation of long period multiples

T. Curtis Holden[1]

keywords: *multiple suppression, beam stack, beam stack inversion*

ABSTRACT

I present a multiple suppression method that utilizes inverse beam stacking to model and remove long period multiples from CMP gathers. The method is applied to synthetic data and is compared with the Parabolic Radon Transform (PRT) method. The results are very good except at the nearest offsets.

INTRODUCTION

Inversion of shot gathers or CMP gathers to a model space that distinguishes multiple and primary energy is a well recognized robust method of filtering multiple energy. Thorson and Claerbout (1985) introduced the method of inverse velocity stacking. Inverse velocity stacking is the representation of data as the superposition of constant amplitude hyperbolas such that the original data can be reconstructed from a model. Hampson (1986) used inverse velocity stacking as a tool for filtering multiple energy, though Hampson improved on the speed of inverse velocity stacking by reformulating the inversion as the modeling of data with constant amplitude parabolas instead of constant amplitude hyperbolas.

In order for a transform to be valuable for multiple suppression the representation in model space of the multiple and primary energy must be such that the multiple and primary energy map into distinct regions of model space. Inverse velocity transforms fit this requirement best when the amplitudes and trajectories of events in the data match the amplitudes and trajectories of the transforming curve. Amplitude variations can be incorporated into inversion schemes to improve the compactness of events in model space. This assumes that the amplitudes of interest are actually known in advance, which is not a realistic assumption in many cases. The PRT method performs well in cases in which events in NMO-corrected gathers approximate constant amplitude parabolas. This situation is not satisfied in real data in all situations. Examples of such situations include cases of: 1) strong AVO on multiples and/or primary events 2) situations in which the primary events are strongly non-hyperbolic, and 3) cases in which the moveout of multiples relative to primaries is so

[1]**email:** curt@sep.stanford.edu

large that the multiples can not be approximated by a parabola in an NMO-corrected gather.

I present a method that utilizes inverse beam stacking for the purpose of filtering multiple energy. The method presented in this paper performs well in cases in which the PRT method produces suboptimal results. The results of the beam stack method are much cleaner and more complete than the results of the PRT method except at the near offsets. The speed of the beam stack method is on the same order as the PRT method.

BEAM STACK

The beam stack operator performs the transform of data parameterized by time and offset (t, x) to a model parameterized by time, offset and slowness (τ, x_m, p). I formulated the beam stack operator with the forward operation mapping from the model space to the data space. This is represented by:

$$\mathbf{Hm} = \mathbf{d} \tag{1}$$

Where $\mathbf{H}$ is the forward beam stack operator, $\mathbf{m}$ is the model and $\mathbf{d}$ is the data. The forward operation maps energy from each point in model space to local trajectories in data space. The shape of these trajectories is a function of the location in model space the energy is being mapped from. The adjoint operator maps energy to each point in model space via a summation of energy over a local trajectory in data space. Again, the shape of the trajectory is a function of the location in model space energy is being mapped to.

The adjoint of the generalized beam stack operator samples each (t, x) location in the data and stacks over a localized window of offsets, $(x - l, x + l)$. The trajectory of this stack is a function designed to evaluate the local component of energy with dip equal to the slowness parameter p. This energy is mapped to the model space location $(\tau = t, x_m = x, p)$. In order to evaluate the dip, the trajectory function must have a slope at the point (t, x) in the data equal to p. One function that can be used to evaluate the local dip is a dipping line. This is often referred to as a local slant stack (Biondi, 1990). The resolution of the local slant stack is limited by the Fresnel zone of the the linear trajectory across the curved event. In order to increase resolution in model space, I chose to use a parabolic trajectory for the local stack. The parabolic trajectory estimates the local curvature of hyperbolic events from the offset and stepout being evaluated. This estimation of the wavefront curvature results in a larger Fresnel zone. A hyperbolic function is also an option as a stacking trajectory but is prohibited in the frequency domain because of time dependent curvature. Nevertheless, the parabolic approximation of the hyperbolic curvature is quite good for local segments of hyperbolic trajectories.

Beam Stack Model Space and Data Space: An Illustration

Each of the three panels in figure 1 represents a slice of an example beam stack model space with constant slowness. Together these three slices constitute an entire beam stack model space with slowness parameters: $p_1 = 0.0s/km, p_2 = 0.025s/km, p_3 = 0.5s/km$. The energy in this model space is zero at all locations except three points. The forward operation applied to the model space will map the nonzero energy in model space to the parabolic trajectories in data space illustrated in figure 2. The left panel of figure 1 corresponds to the zero slowness slice, $p = 0s/km$, and as such the point with nonzero energy maps to the zero dip trajectory that is displayed on the left of figure 2. The middle panel corresponds to the slowness slice $p = 0.025s/km$ and maps the point of nonzero energy to the trajectory displayed in the middle row of figure 2. The right panel of figure 1 corresponds the slowness slice $p = 0.05s/km$ and maps the point of nonzero energy to the trajectory at the right in figure 2. Figure 2 illustrates the forward mapping from individual points in model space. These three trajectories in themselves do not constitute the complete mapping to model space; to obtain the complete forward modeling, each forward mapping from individual points in model space must be summed together. This is displayed in figure 3. This data has multi-valued dips at the t-x location $(0.12, 1.6)$. Unraveling multi-valued dips is one of the problems that needs to be addressed in modeling multiples and primaries.

The adjoint operation maps the data that lies on the trajectories displayed in figure 2 back to the model space. The complete adjoint operation sums over trajectories for each location in data space. The beam stack operator is not unitary and as such, the mapping of the result of the forward operation, mentioned above and displayed in figure 3, will not map exactly back to the original model displayed in figure 1. The result of the adjoint operation applied to the data displayed in figure 3 is displayed in figure 4. The great disparity between the original model displayed in figure 1 and the adjoint applied to the forward modeled data displayed in figure 3 is a consequence of the original model chosen. The original model includes points of energy, which are not a natural representation of events in beam stack model space. The beam stack transform models dips which necessarily have a greater extent than a point. I chose this model to illustrate what a beam stack trajectory looks like and to illustrate the relation between the model and data space.

Frequency Domain Formulation

The beam stacking operator is represented in the time domain by:

$$m(\tau, x_m, p) \quad = \quad \sum_{x=-l+x_m}^{l+x_m} w(x, x_m) d(t = \tau + \frac{p}{2}(\frac{x^2}{x_m} - x_m), x) \tag{2}$$

where (t, x) is the time and offset of the input gather and (τ, x_m, p) is the time, offset and slowness in model space. Each location in model space, (τ, x_m, p), corresponds to the component of energy with the dip p at the location $(t = \tau, x = x_m)$ in the data.

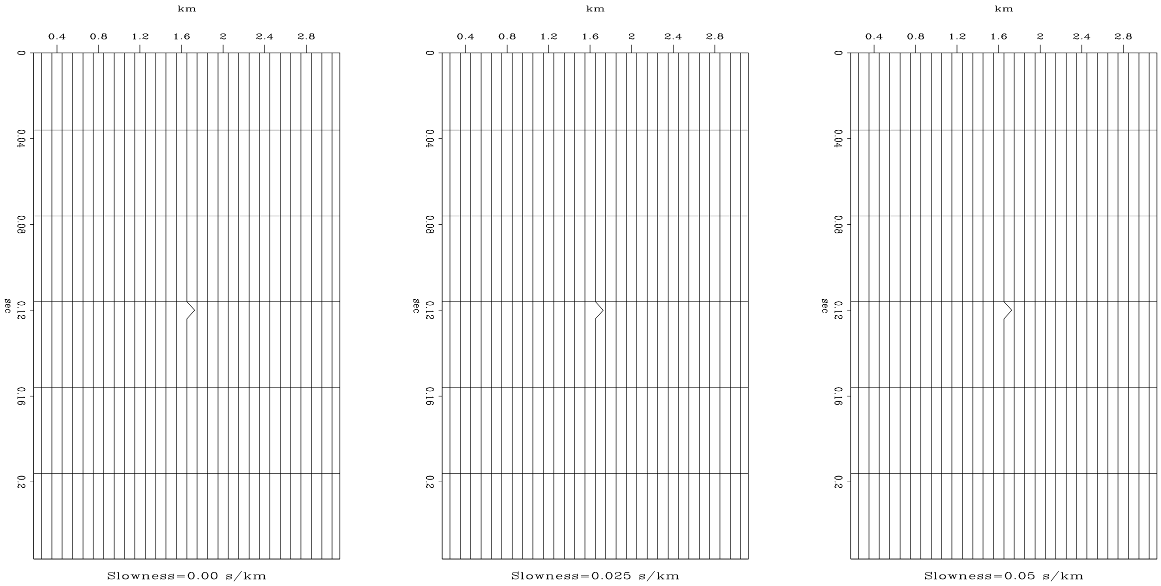

Figure 1: Example model space of the beam stack transform. Slowness values are $p_1 = 0.0s/km, p_2 = 0.025s/km, p_3 = 0.05s/km$. curt-model [ER]

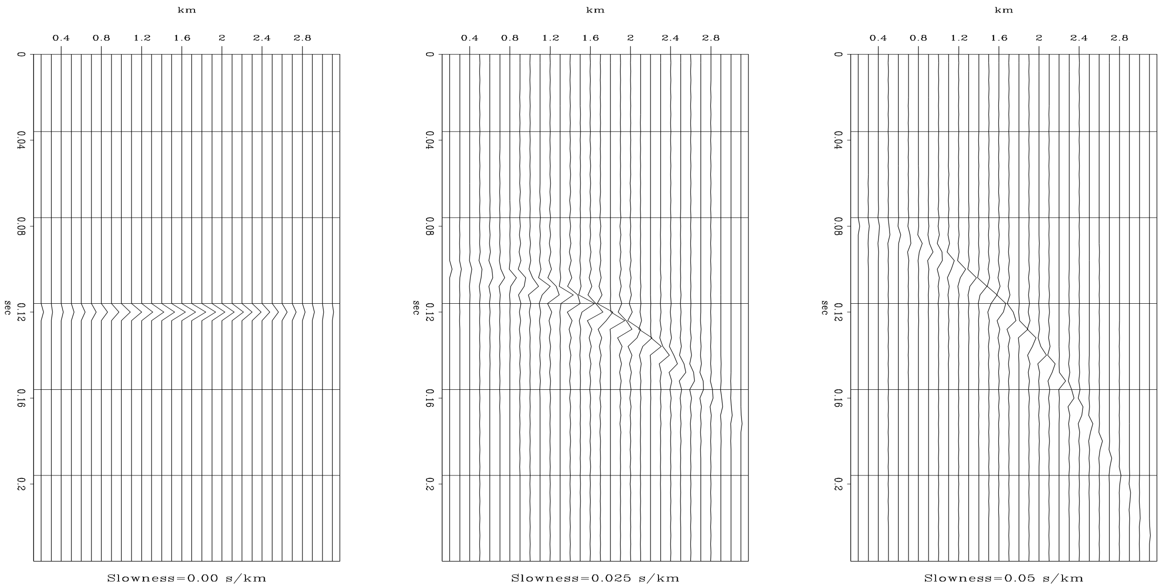

Figure 2: Trajectories resulting from the forward modeling of each of the non-zero points individually in the example model space. curt-stakTraj [ER]

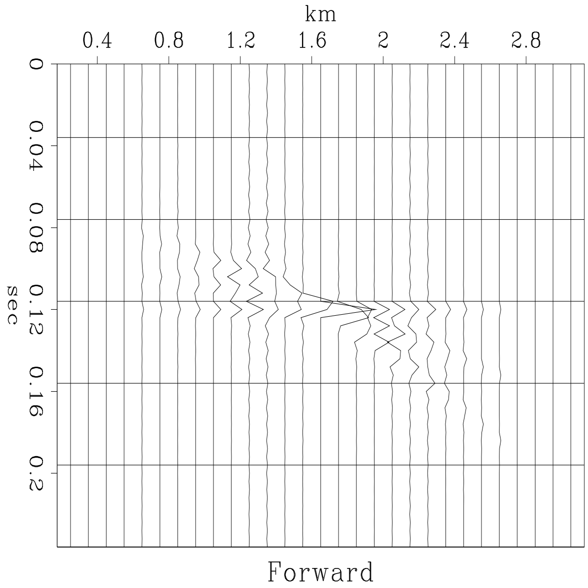

Figure 3: Data resulting from the forward modeling of the example model space. curt-data [ER]

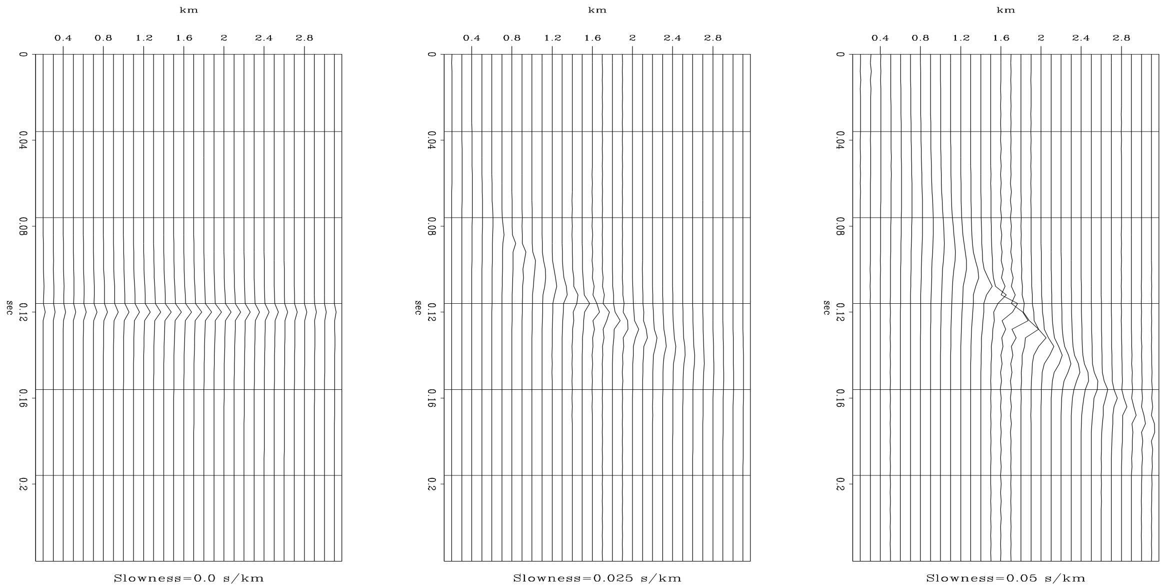

Figure 4: Adjoint operation applied to the forward modeled example model space. curt-adj [ER]

The factor $w(x, x_m)$ is a weighting function that is generally used for the purpose of reducing truncation artifacts. The forward operator is as follows:

$$d(t,x) \;=\; \sum_{p=p_1}^{p_{n_p}} \sum_{x_m=-l+x}^{l+x} w(x,x_m) m(\tau = t - \frac{p}{2}(\frac{x^2}{x_m} - x_m), x_m, p) \tag{3}$$

Fourier transforming both sides of the above equation in time results in a nice formulation in the frequency domain:

- Adjoint:

$$m(\omega, x_m, p) \;=\; \sum_{x=-l+x_m}^{l+x_m} w(x,x_m) d(\omega,x) e^{j\omega\frac{p}{2}(\frac{x^2}{x_m} - x_m)} \tag{4}$$

- Forward:

$$d(\omega,x) \;=\; \sum_{p=p_1}^{p_{n_p}} \sum_{x_m=-l+x}^{l+x} w(x,x_m) m(\omega,x_m,p) e^{-j\omega\frac{p}{2}(\frac{x^2}{x_m} - x_m)} \tag{5}$$

Examination of the above frequency domain representation reveals that the action of the operator appears in a multiplicative exponential term. This representation allows the operator to be split into separate operators for each frequency. This frequency domain representation allows the operator to be applied selectively to those frequencies of interest.

Synthetic data

To test the inverse beam stack multiple suppression method I have prepared three models of synthetic data composed of primaries and multiples. These models are displayed in figure 5. Each model consists of two primaries, one at $2.0s$ having constant amplitude and the other at $1.5s$ having a polarity reversal. Both primaries have a zero offset RMS velocity of $2.1km/s$. The primaries in model2 have non-hyperbolic moveout which is due to a velocity function that is proportional to the cube of the offset. The far offset velocity of the primaries in model2 are 1.15 times higher than the near offset RMS velocity. The primaries generated in model2 might be expected to occur in the presence of a salt body. Superimposed on these primaries are multiple events. Two multiple events are superimposed on the primary events in model1 and model2 and occur at $1.4s$ and $1.92s$. The multiples in model1 and model2 have RMS velocities of $1.5km/s$ and have amplitudes that are six times the zero offset amplitudes of the primary events. The multiples in model3 are made up of two periodic trains of events, one train with RMS velocity of $1.5km/s$ and the other train with RMS velocity of $1.7km/s$. The period of the first train is $0.12s$ while the period of the second is $0.1s$.

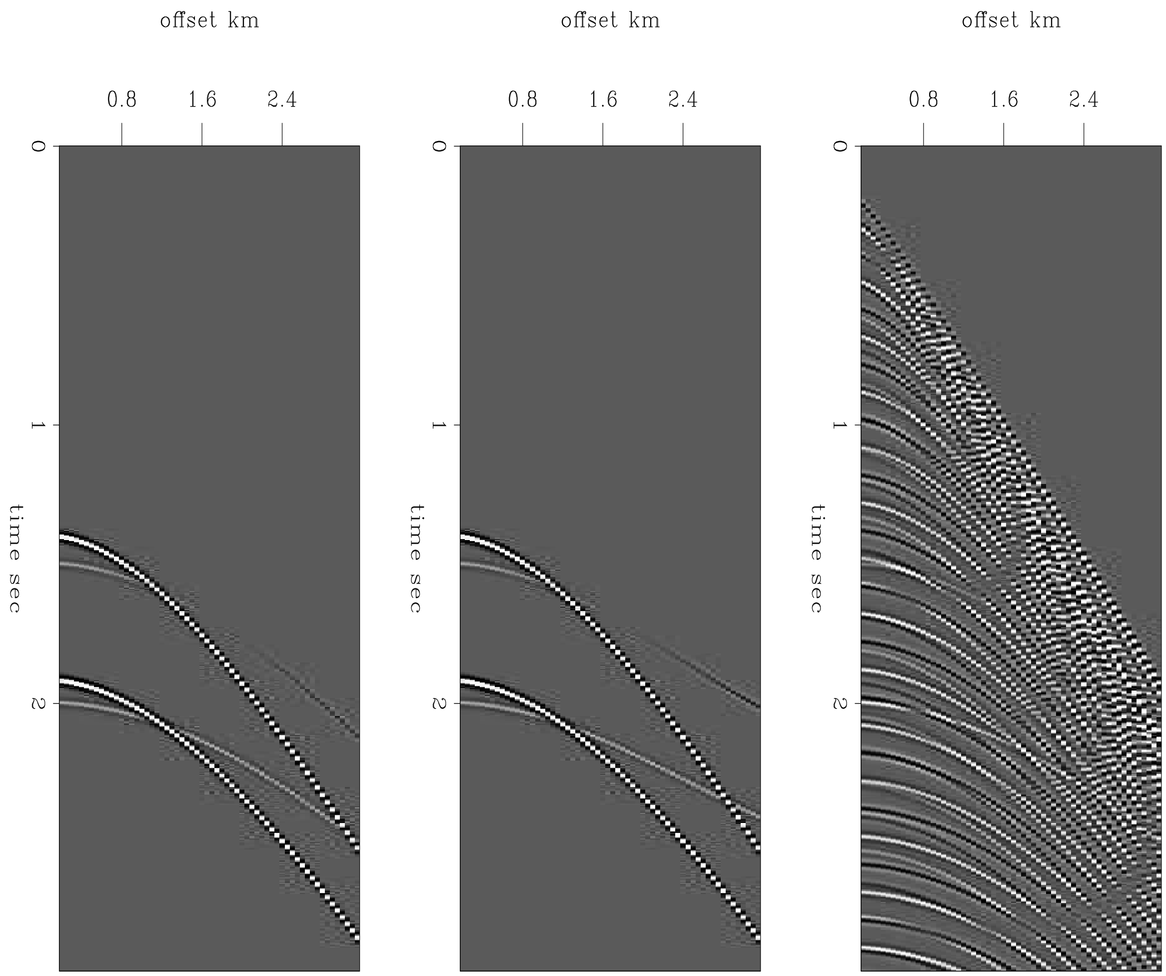

Figure 5: Synthetic data. Left: model1, Center: model2, Right: model3 curt-modmult [ER]

Beam Stack Inversion

In this paper I formulate the beam stack inversion as an iterative least squares problem in the frequency domain. I use a conjugate gradient solver to find the model, $\mathbf{m}$, that minimizes the residual, $\mathbf{r}$, in the following system:

$$\mathbf{r} \quad = \quad ||(\mathbf{d} - \mathbf{Hm})||^{\mathbf{2}} \tag{6}$$

In my previous work with Biondi, we formulated the iterative problem in the time domain. As opposed to the iterative time domain representation, the frequency domain formulation is appreciably faster.

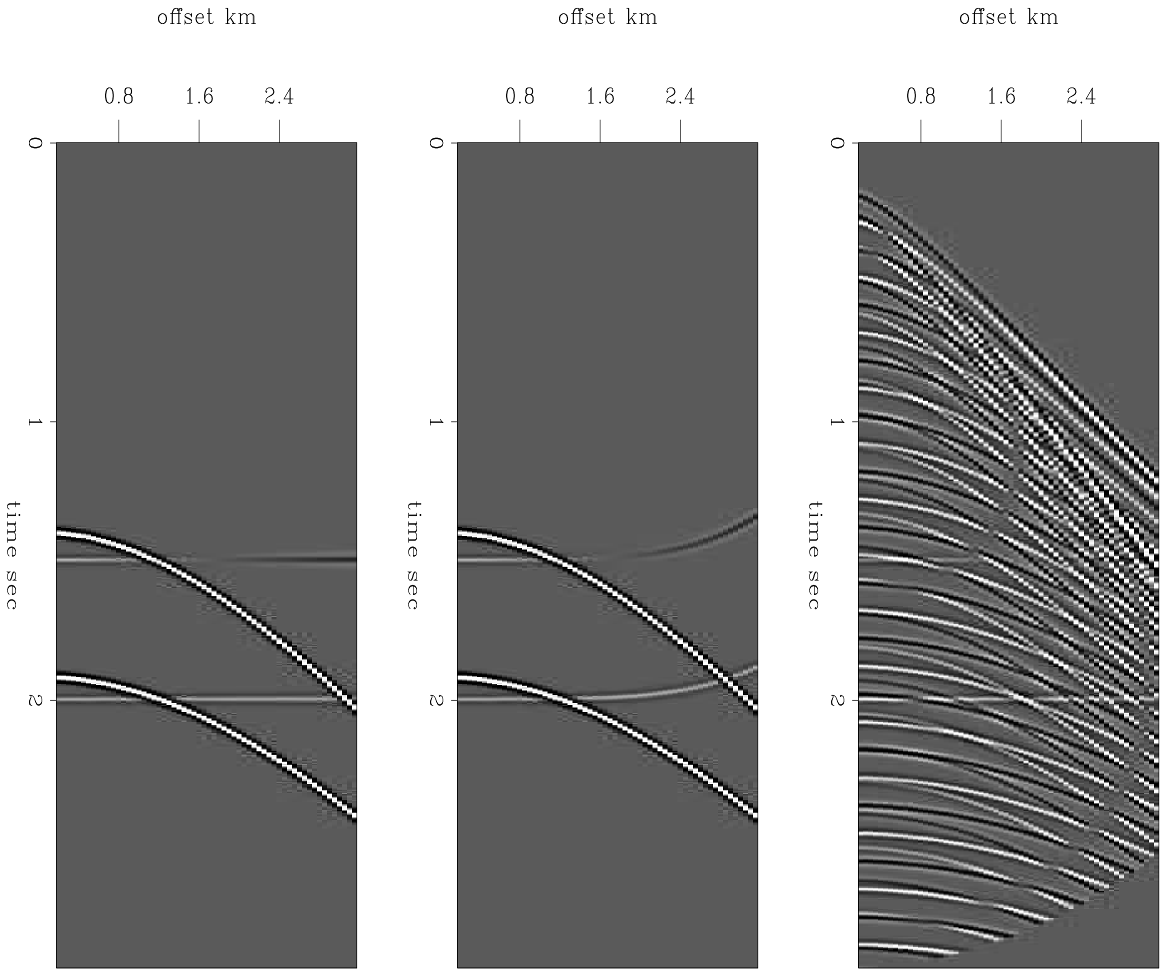

Figure 6: NMO-corrected synthetic data. curt-nmo [ER]

I applied the inversion to the synthetic data displayed in figure 5 after applying NMO. The NMO-corrected synthetic data is displayed in figure 6. The application of NMO has the effect of limiting the dip range in the data. This reduces the size of the model space necessary to represent the input data.

The inversion results for the NMO-corrected data appear in figures 7,8 + 9. Each figure displays all the slowness slices from the model space of the beam stack inversion. It is apparent from figures 7+9 that the primary energy is confined to the zero slowness slice of the model space, while figure 8 shows that the primary energy resides in zero and negative slowness slices.

I forward modeled each inversion result to produce a reconstruction of the NMO-corrected synthetic data. The difference in total energy between the reconstructed synthetic data and original synthetic data is less than 2% in each case. This difference was achieved in five iterations and required three minutes of computing time on one processor of an SGI power challenge.

An alternative to the iterative least-squares solution is a direct inversion. To implement the direct inversion, the data covariance matrix, $\mathbf{HH}^{\dagger}$, must be inverted. This matrix, and the beam stack operator itself, can be partitioned into a separate matrix for each frequency of interest, thus reducing the size of the data covariance matrix significantly. The partitioned beam stack operator in matrix form is of size $(N_x)(N_xN_p)$ while the partitioned data covariance matrix is of size $(N_x)(N_x)$. In order to apply a direct least squares inversion, the data covariance matrix would have to be inverted for each frequency of interest and applied as a pre-conditioner to the application of the adjoint to the data.

Hampson (1986) used this approach to formulate the direct least squares inverse of the parabolic radon transform. In the case of the radon transform, the model covariance matrix must be inverted. The model covariance matrix of the PRT is of size $(N_p)(N_p)$ where p is the slowness parameter of the PRT. A typical range of the PRT parameter p is 50. Kostov (1990) has shown that the family of radon transforms results in a model covariance matrix that is Toeplitz in structure and as such can be inverted with the Levinson method (and other methods) at a cost proportional to n^2 instead of n^3, where n is the size of the matrix to be inverted (Claerbout, 1976). The Toeplitz property of the PRT model covariance matrix coupled with its reasonable size, around, 50×50, elements, make a direct inversion of the PRT model covariance matrix reasonable. The parabolic beam stack data covariance matrix is not Toeplitz and it's size is equal to the square of the number of offsets in the data. Because of these features, the parabolic beam stack operator is not a great candidate for a direct inversion.

MULTIPLE SEPARATION

NMO correction makes separation of the multiple energy from primary energy straightforward. Primary events in NMO-corrected gathers are flat and map to the zero slowness, $p = 0s/km$, slice of model space. The energy in the rest of the model is multiple energy and can easily be isolated by windowing out all slices with slowness greater than zero. This isolated multiple energy can be forward modeled and subtracted from the original NMO-corrected data or the primary energy can be forward modeled

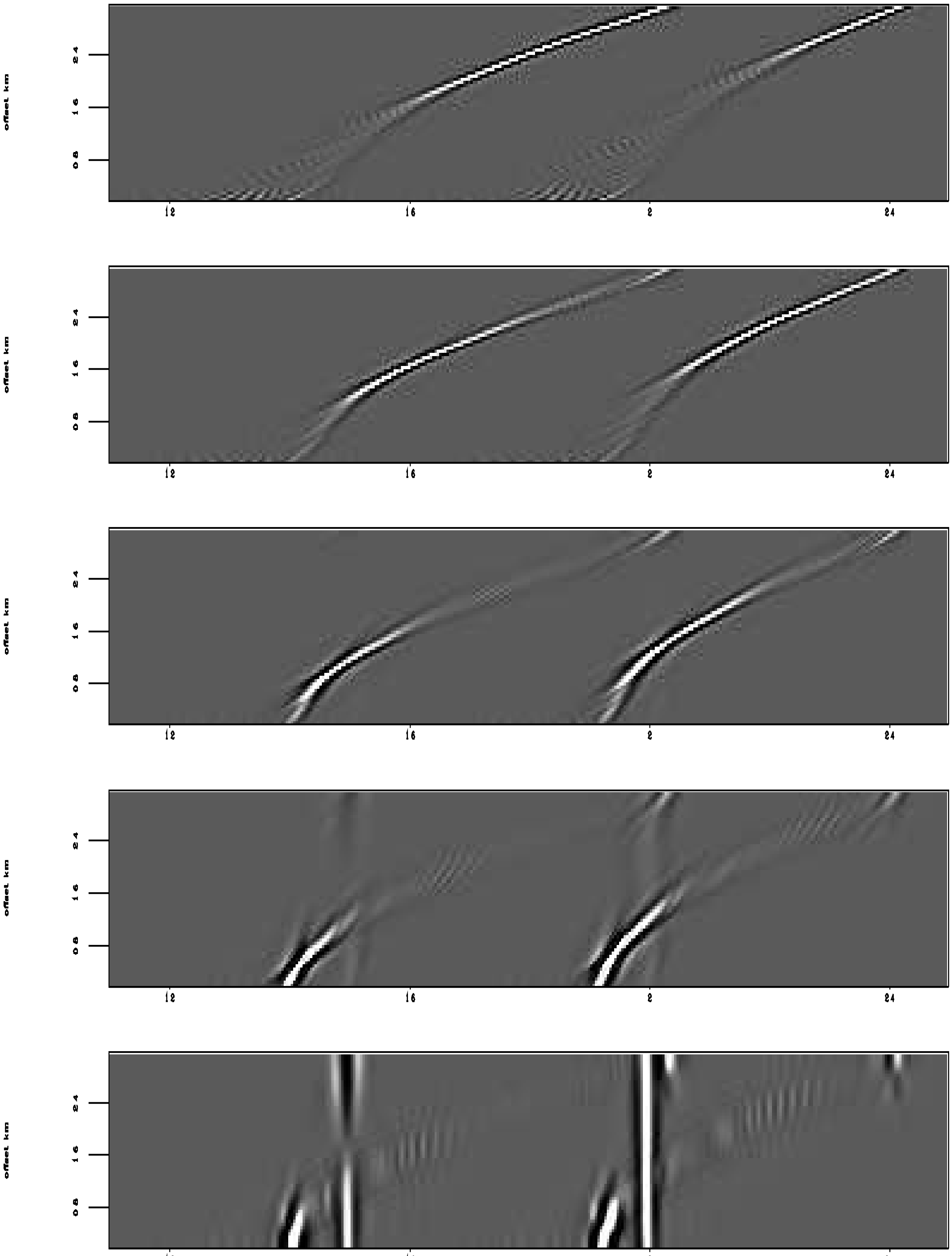

Figure 7: Inversion of NMO-corrected model1. The slices displayed from bottom to top have the following constant slowness values(s/km): $p_1 = 0.0, p_2 = 0.075, p_3 = 0.15, p_4 = 0.225, p_5 = 0.3$ $\boxed{\text{curt-invnmo1}}$ [ER]

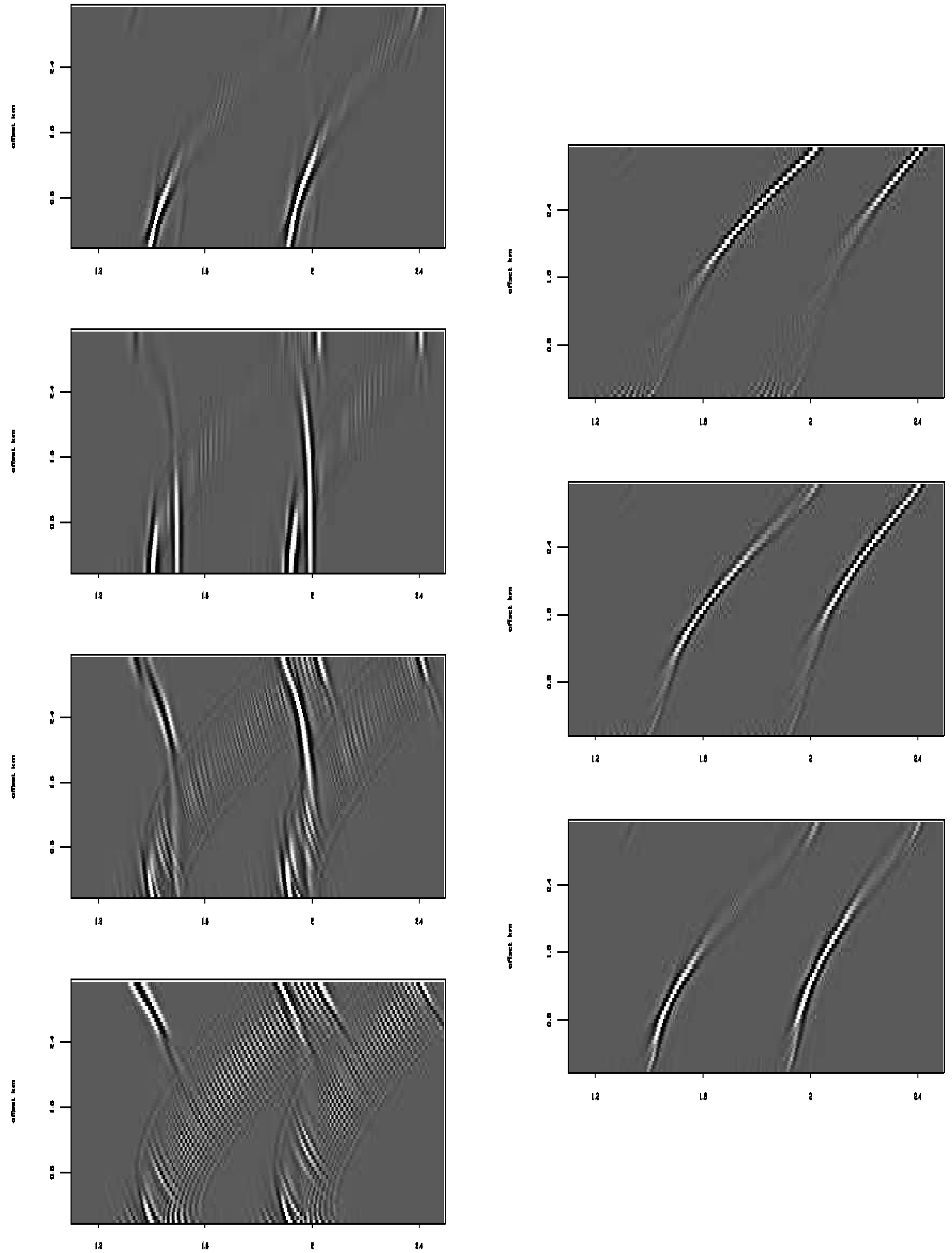

Figure 8: Inversion of NMO-corrected model2. The slices displayed from bottom-left to top-right have the following constant slowness values(s/km): $p_1 = -0.15, p_2 = -0.075, p_3 = -0.0, p_4 = 0.075, p_5 = 0.15, p_6 = 0.225, p_7 = 0.3$ curt-invnmo2 [ER]

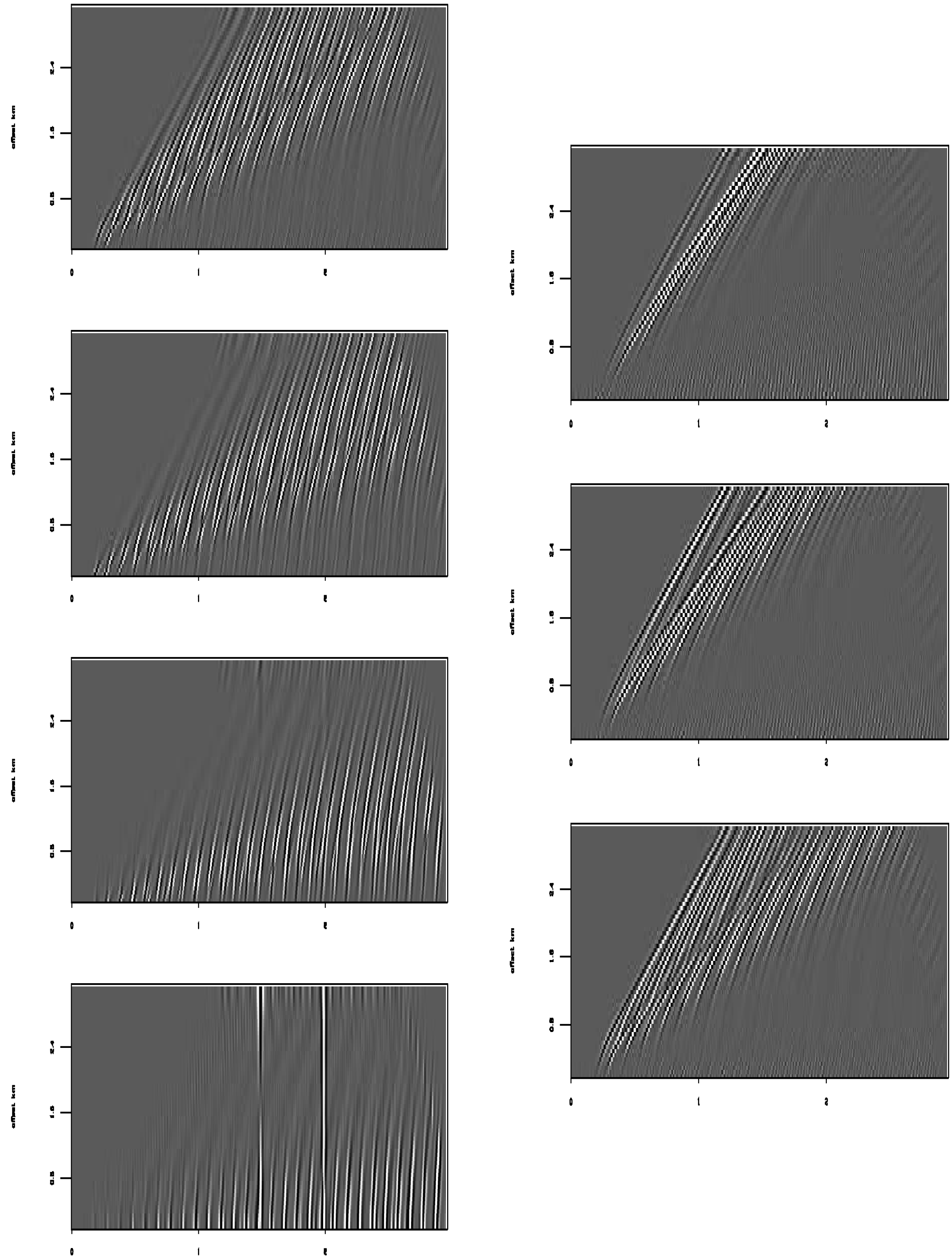

Figure 9: Inversion of NMO-corrected model3. The slices displayed from bottom-left to top-right have the following constant slowness values(s/km): $p_1 = 0.0, p_2 = 0.075, p_3 = 0.15, p_4 = 0.225, p_5 = 0.3, p_6 = 0.375, p_7 = 0.45$ $\boxed{\text{curt-invnmo3}}$ [ER]

directly from the zero slowness slice. The NMO approach avoids the somewhat difficult approach taken by Holden and Biondi (1996) of applying a time-offset dependent mute to inverted data without NMO correction. A potential drawback to the NMO approach is that primary events in NMO-corrected gathers may not always be flat. This may be compensated for if there is prior knowledge of the primary event shape. In this respect the NMO method may not be as robust as the method of applying a time-offset dependent mute to inverted data without NMO correction

The multiple filtering results are displayed in figure 10. These results were obtained by simply windowing the zero slowness slice of each inversion result displayed in figures 7,8 + 9, and then forward modeling. The windowing of the primary events in the model2 case involved three slowness slices due to the negative dips present on the primaries. Like the primaries, the multiples will have a component of very nearly zero slowness at the nearest offset. This energy will map into the zero slowness panel with the primaries. This energy appears as stubs at the near offsets of the results in figure 10. The reflecting boundary conditions contribute to the amplitude of these stubs.

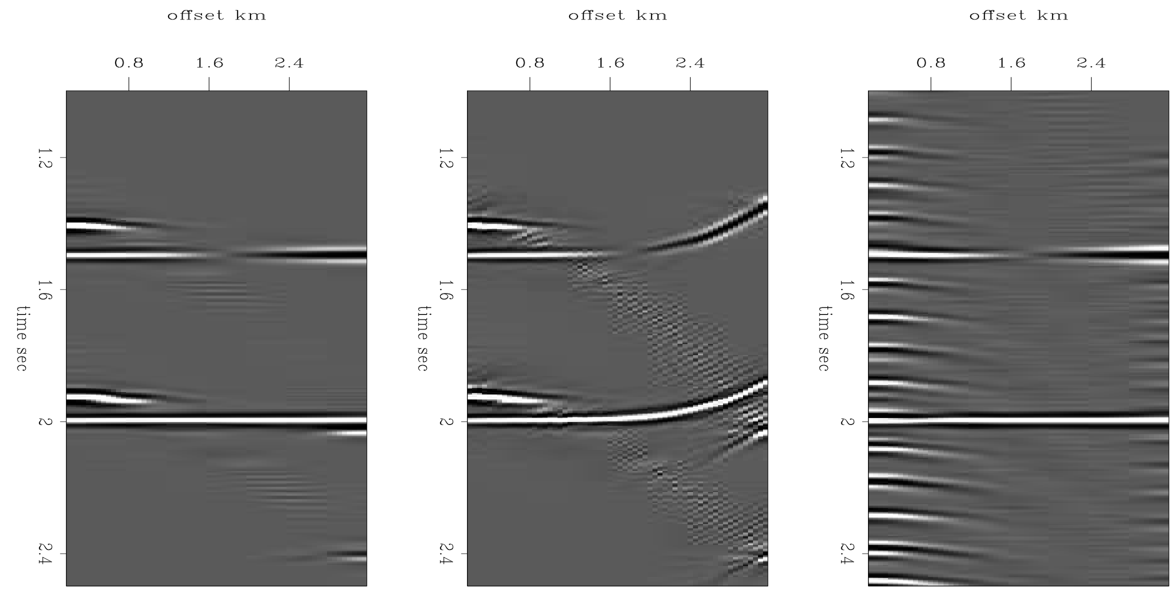

Figure 10: Result of beam stack multiple suppression. Left: model1 estimated primaries, Center: model2 estimated primaries, Right: model3 estimated primaries. curt-fwdprime [ER]

I have prepared results of the PRT multiple suppression method as a comparison. The estimated primaries of the NMO-corrected data using the PRT multiple suppression method are displayed in figure 12. The PRT transform of the NMO-corrected data are displayed in figure 14. Multiple energy is obviously present in all three results in figure 12, the most notable cases are the distinct strong multiple ghosts present with the estimated primaries of model1 and model2. These ghosts are caused by the inability of the PRT to compactly represent the significantly non-parabolic multiples in PRT transform space such that the masking procedure can not separate

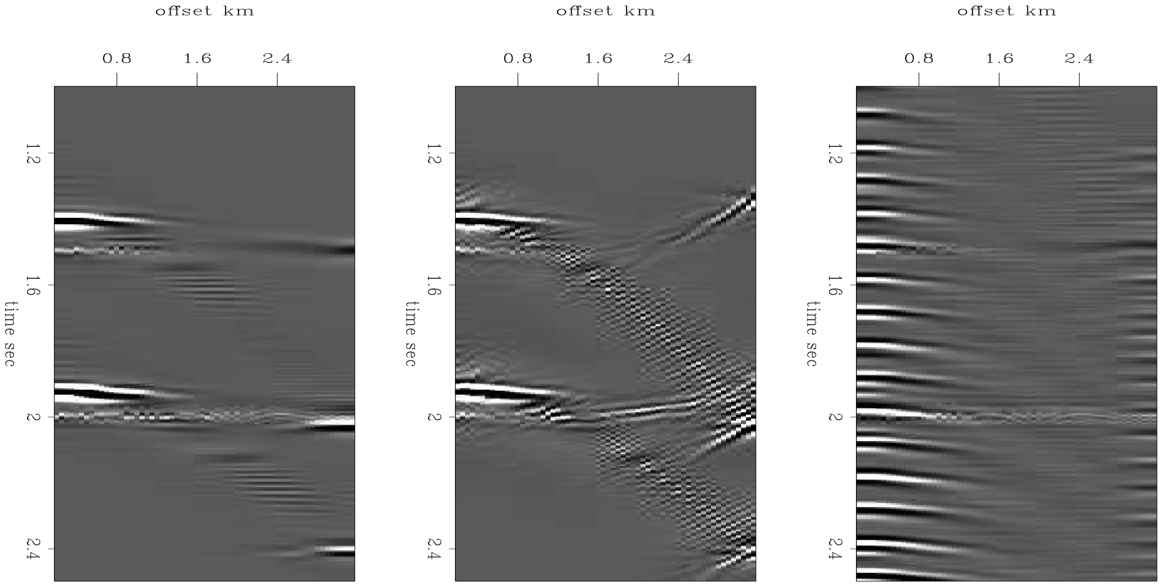

Figure 11: Energy difference of original primaries and result of beam stack multiple suppression. curt-diffbs [ER]

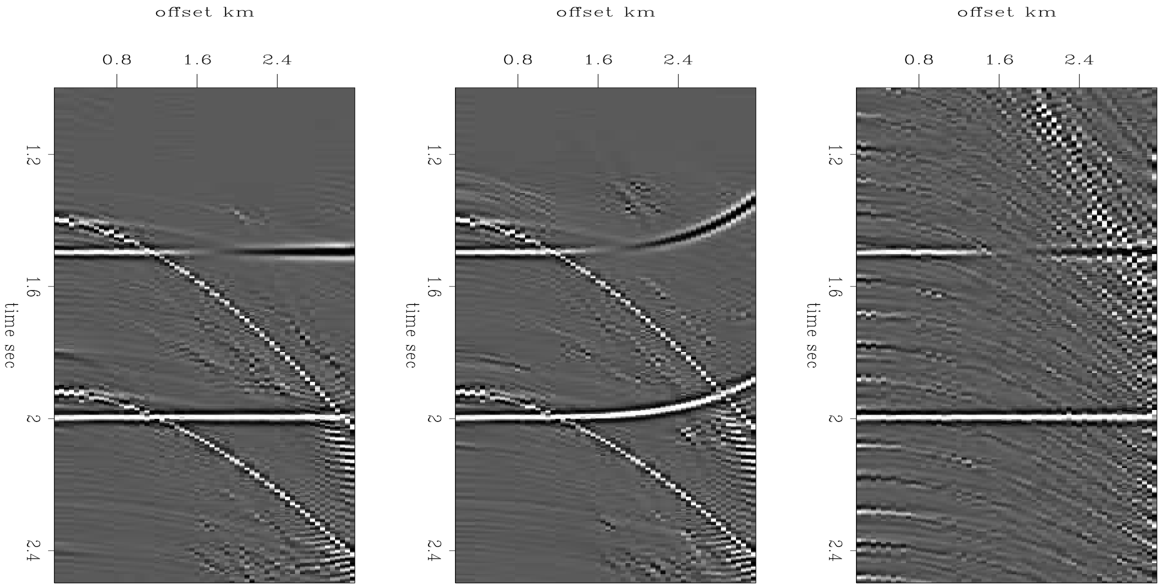

Figure 12: Results of PRT multiple suppression method. Left: model1 estimated primaries, Center: model2 estimated primaries, Right: model3 estimated primaries. curt-prtprime [CR]

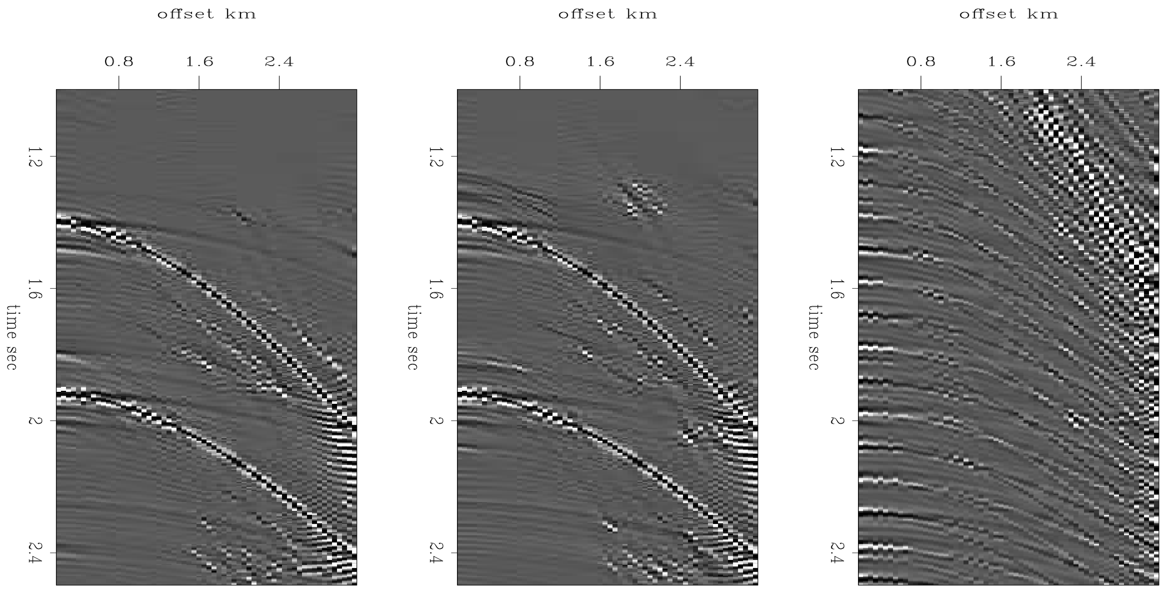

Figure 13: Energy difference of original primaries and result of PRT multiple suppression. curt-diffprt [CR]

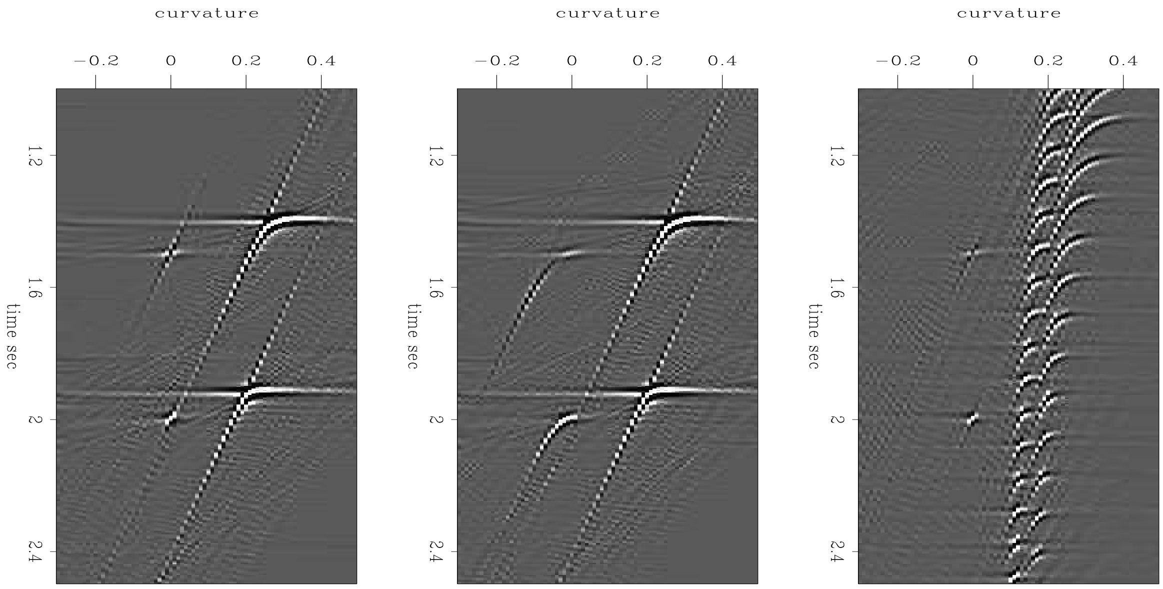

Figure 14: PRT transform space. Left: model1, Center: model2, Right: model3 curt-prttrans [CR]

the primary from the multiple energy. Examination of figure 14 reveals that multiple energy spreads over a large region of the PRT transform space and is probably not properly isolated by the masking procedure associated with the PRT method.

The difference between the estimated primaries from each method and the original multiple-free primaries appear in figure 11 for the beam stack method and in figure 13 for the PRT method. The PRT method leaves a significant amount of multiple energy in the multiple suppressed results for all three cases at all times and offsets, where as the beam stack method contains multiple energy only at the near offsets. Increasing the sampling of the slowness parameter or range of the slowness parameters of the PRT transform does not change the PRT result significantly.

CONCLUSION

The of inverse beam stack multiple suppression method has shown cleaner results than the results of the PRT method on the same synthetic data examples. As such, the beam stack method can be attractive method when the PRT method is not optimal. The beam stack method is especially attractive in situations in which prestack amplitudes are important. The beam stack method performs well in the presence of very strong multiples, in the presence of non-hyperbolic moveout and in cases of significant AVO on primaries or multiples. The speed is comparable to the PRT method. Unfortunately the near offset remnant multiple energy is a problem for the beam stack method. The near offset energy may be handled using predictive deconvolution. Another possible shortcoming is that the beam stack method may be sensitive to primary events in NMO-corrected gathers that are not completely flat.

REFERENCES

Biondi, B., 1990, Seismic Velocity Estimation By Beam Stack: SEP–**64**, 123.

Claerbout, J. F., 1976, Fundamentals of geophysical data processing: Blackwell.

Hampson, D., 1986, Inverse velocity stacking for multiple elimination: J. Can. Soc. Expl. Geophys., **22**, no. 1, 44–55.

Holden, T. C., and Biondi, B., 1996, Suppression of water bottom multiple energy in beam stacked data: SEP–**93**, 151–160.

Kostov, C., 1990, Seismic signals from a drillbit source: SEP–**63**.

Thorson, J. R., and Claerbout, J. F., 1985, Velocity-stack and slant-stack stochastic inversion: Geophysics, **50**, no. 12, 2727–2741.