

Migration velocity analysis of incomplete/irregular data using least squares Kirchhoff migration

Abdolnaser Yousefzadeh, John C. Bancroft, and Gary F. Margrave

ABSTRACT

In order to attenuate the migration artifacts and increase the spatial resolution of the subsurface reflectivity, conventional migration can be replaced by the least squares prestack migration (LSPSM). However, this is a costly procedure and also requires a good knowledge of background velocity.

This paper focuses on migration velocity analysis by measuring semblances on common image gather (CIG)s and extends the method to LSPSM CIGs. We show that when the data are sampled incompletely or irregularly, the LSPSM offset domain and shot domain common image gathers give higher resolution in the semblance or unnormalized crosscorrelation panel for choosing the correct velocity than conventional migration common image gathers.

The velocity information extracted by these methods is accurate enough to give a good convergence in the least squares conjugate gradients and also a good reconstruction of missing data.

INTRODUCTION

In exploration seismology a source releases energy in the form of elastic waves into the subsurface. Designed receiver systems digitize and record the energy (wavefield) that arrives directly or reflects back to the earth surface.

The sampled wavefield will be processed to be interpretable for geophysicists, geologists, or mining engineers. In the seismic data processing procedures, noises and unwanted signals are attenuated. After normal move out (NMO) correction, all traces corresponding to the same common midpoint (CMP) position are stacked to increase the signal to noise ratio. There are many other steps in the seismic data processing to give the best possible picture of the subsurface. Proper deconvolution, filtering, and multiple attenuations are some examples of these steps.

However, there may be some pre-requisites that are violated or some noises that has not been considered. In the simple theory of seismic reflection method, reflectors are assumed to be horizontal and also earth is horizontally homogenous. This is not true in the real earth. In the simplest case, ignoring this fact leads to an image in which the dipping reflectors are not in the correct positions and may also have wrong dip angles.

Migration is an effective tool in the seismic data processing in order to overcome some of these mentioned problems. Ideally, migration not only moves the dipping events

to the correct spatial position, but also collapses diffracted energies to the scatterpoints. Compared to the other steps in seismic processing, migration is a very complex and expensive procedure. It is mostly implemented on prestack data, a usually a very large dataset. The result of migration is an image which ideally shows the earth's reflectivity model as a function of time or depth.

Hagedoorn (1954) introduced the first computational method of migration, diffraction summation. In the 70's Claerbout and Doherty (1972) showed how migration is an approximate solution to the wave equation. The integral formulation of the Kirchhoff migration was introduced by Schneider (1978). Gazdag (1978) and Stolt (1978) showed how to perform migration using properties of the Fourier transform (Gary et. al., 2000). Reverse time migration, one of the most expensive wavefield continuation methods of migration was introduced later (Baysal et al., 1983; Whitmore, 1983). Hill (1990) introduced the Gaussian beam migration based on the decomposition of data and source function in Gaussian beams. Researchers still are looking for new ways of migration or improvement of the older methods to extend them to be able to image the area with a complex geology.

Each point in the geological domain may be regarded as a scatterpoint. Kirchhoff wave equation migration collapses all diffracted energies back to the original scatterpoint. In the resulted migration image, each dipping reflector is moved to the correct location with true dip and length. Because of its many advantages, when compare to the other methods of migration, Kirchhoff remained one of the main migration methods in the seismic exploration. Kirchhoff migration algorithms are computationally fast and can handle incomplete and/or irregularly sampled data. It can be implemented on both poststack and prestack data, and can also be used for the migration of the converted waves. Handling irregular topography is easier in with the Kirchhoff migration. Its ability to migrate a part of the image is another advantage of Kirchhoff migration.

With a simple geology structure and in case of regularly and densely sampled data, Kirchhoff migration produces an acceptable subsurface reflectivity image. In reality, it is difficult to have a dense, regularly sampled seismic data, especially in the 3D surveys. Human-made obstacles, mountains, rivers, land topography, acquisition layout, and economical aspects, usually leave some gaps or cause an un-even distribution of azimuth (in 3D) and offsets in the acquired seismic data. Marine data also lack a wide distribution of azimuths, and may have gaps due to the marine construction or streamer feathering.

The ability of easily handling incomplete and irregular seismic data is probably the main advantages of the Kirchhoff to the other methods of migration. However, incomplete data produce migration artifacts and give a blurred image of the earth reflectivity. Least squares migration (LSM) helps to avoid this problem.

In order to overcome this shortcoming, Kirchhoff migration can be improved by a generalized inverse as an approximation to the exact inverse (Tarantola, 1984). This approach is called least squares migration or inversion (Nemeth et. al., 1999, Duquet et. al., 2000, Kuehl and Sacchi, 2001). This study focuses on the Kirchhoff least squares prestack time migration (LSPSM). Kirchhoff type of LSPSM proposed by Nemeth et. al. (1999) and Duquet et. al. (2000). Nemeth et. al. (1999) showed that LSM can be used to

reduce the acquisition footprints in the seismic data. LSPSM is not limited to the Kirchhoff method. Wave propagation migration and migration methods that utilize Fourier domain can also be used in the least squares scheme. However, wave propagation migration methods (Kuehl and Sacchi, 2001) have higher cost compare to the Kirchhoff method.

Higher resolution images of the LSM can be used for data reconstruction. However, this method require an accurate estimate of the velocity. Migration velocity analysis on the Common Image Gather (CIG) is a useful tool to find the required migration velocity. the sensitivity of the LSM to the accuracy of the velocity information is a key to use LSM for migration velocity analysis, especially with incomplete and irregular data. The high sensitivity of LSM to the accuracy of velocity and its low sensitivity to the incompleteness and irregularities of data make it a practical choice for the velocity analysis.

KIRCHHOFF LEAST SQUARES MIGRATION

Many seismic processing and imaging problems can be written in the general form of:

$$\mathbf{d} = \mathbf{G}\mathbf{m}, \quad (1)$$

where \mathbf{d} is the observed seismic data, \mathbf{m} is the earth reflectivity model and \mathbf{G} is an operator acting on \mathbf{m} in order to produce \mathbf{d} . Radon transform and deconvolution are the two processes which can be formulated as linear problems. Kirchhoff seismic modelling or de-migration is another example of an operator that can be written in terms of a linear problem.

The inverse process,

$$\mathbf{m} = \mathbf{G}^{-1}\mathbf{d}, \quad (2)$$

recovers the earth model or reflectivity from the seismic data. Since matrix \mathbf{G} may not be square or non-invertible or it may be extremely large, calculating the inverse of \mathbf{G} may be difficult or impossible. Thus, approximations to the inversion are used. The first approximation uses the transpose (adjoint) of \mathbf{G} :

$$\hat{\mathbf{m}} = \mathbf{G}^T\mathbf{d}, \quad (3)$$

where $\hat{\mathbf{m}}$ is the migrated image and \mathbf{G}^T is the migration operator.

By defining Kirchhoff modelling (equation 1) as a forward operator and Kirchhoff migration (equation 3) as its adjoint (conjugate transpose) operator, seismic imaging can be considered as an inversion problem. Substitution of \mathbf{d} from equation 1 into equation 3 results in:

$$\hat{\mathbf{m}} = \mathbf{G}^T\mathbf{G}\mathbf{m}. \quad (4)$$

If $\mathbf{G}^T \mathbf{G} = \mathbf{I}$, then Kirchhoff migration would reconstruct the true earth reflectivity model. However, due to the geometrical spreading loss during modelling or irregularity and/or incompleteness of sampled seismic data, $\mathbf{G}^T \mathbf{G}$ is different from an identity matrix. In such cases Kirchhoff pre-stack migration produces artifacts in the migrated image.

In the LSM, migration artifacts are attenuated by minimizing the difference between the observed data, \mathbf{d} , and the modeled data, $\mathbf{G}\tilde{\mathbf{m}}$, expressed by $|\mathbf{G}\tilde{\mathbf{m}} - \mathbf{d}|$, where $\tilde{\mathbf{m}}$ is an approximation to \mathbf{m} . Since data include some errors, trying to find a model to fit the data perfectly will be replaced by

$$\mathbf{e} = \mathbf{G}\tilde{\mathbf{m}} - \mathbf{d}, \quad (5)$$

where \mathbf{e} is an error vector (Sacchi, M. D., 2005, course note in MITACS 2005, University of Alberta). Minimum norm or Euclidian norm solutions includes finding a model, $\tilde{\mathbf{m}}$, that minimizes the following cost function:

$$J(\tilde{\mathbf{m}}) = \|\tilde{\mathbf{m}}'\tilde{\mathbf{m}}\|^2, \quad (6)$$

subject to the data constraint:

$$\|\mathbf{G}\tilde{\mathbf{m}} - \mathbf{d}\|^2 = \epsilon. \quad (7)$$

These two equations, together implies the minimization of a cost function in the form of:

$$J(\tilde{\mathbf{m}}) = \|\mathbf{G}\tilde{\mathbf{m}} - \mathbf{d}\|^2 + \mu^2 \|\tilde{\mathbf{m}}\|^2. \quad (8)$$

where μ is the regularization weight.

In a more general form, an objective function to reduce the migration artifacts can be written as (Nemeth, 1999):

$$J(\tilde{\mathbf{m}}) = \|\mathbf{G}\tilde{\mathbf{m}} - \mathbf{d}\|^2 + \mu^2 \mathcal{R}(\tilde{\mathbf{m}}). \quad (9)$$

The first term in the right hand side of equation 8 is data misfit. In the minimization of $J(\tilde{\mathbf{m}})$, this term recovers a model to fit the data. Second term on the right hand side of equations 8 and 9 is a regularization term.

By ignoring the regularization term, $\mu = 0$, after taking the first derivative of the remaining cost function and setting it equal to zero, the following normal equation is achieved:

$$\mathbf{G}^T \mathbf{G} \mathbf{m}_{LS} - \mathbf{G}^T \mathbf{d} = \mathbf{0}. \quad (10)$$

This gives the least squares solution, \mathbf{m}_{LS} ,

$$\mathbf{m}_{LS} = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{d}. \quad (11)$$

Minimum norm is the simplest form of the regularization function, ($\mathcal{R}(\mathbf{m}) = \|\mathbf{m}\|_2^2$), which leads to the damped least squares solution, \mathbf{m}_{DLS} , to the problem:

$$\mathbf{m}_{DLS} = (\mathbf{G}^T \mathbf{G} + \mu^2 \mathbf{I})^{-1} \mathbf{G}^T \mathbf{d}. \quad (12)$$

Smoothing in the offset direction (smoothing in the direction of the hyperbola trajectories in a shot gather) is another constraint which can be performed on the LSPSM by assuming $\mathcal{R}(\mathbf{m}) = \|\mathbf{D}_h \mathbf{m}\|_2^2$, where \mathbf{D}_h is the first derivative in the offset direction and mathematically is expressed by multiplication of the following matrix to the model vector:

$$\mathbf{D}_h = \begin{bmatrix} -1 & +1 & 0 & \cdots \\ 0 & -1 & +1 & \cdots \\ 0 & 0 & -1 & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix}. \quad (13)$$

First derivative operator acts as a high pass filter. Therefore, minimization of a cost function with this regularization term is equal to penalizing high frequency contents. Solution to a LSPSM problem with smoothing as the regularization, \mathbf{m}_{SLS} , is:

$$\mathbf{m}_{SLS} = (\mathbf{G}^T \mathbf{G} + \mu^2 \mathbf{D}_h^T \mathbf{D}_h)^{-1} \mathbf{G}^T \mathbf{d}. \quad (14)$$

Practiced LSPSTM

Since the images resulted from the least squares migration have higher resolution than the images from migration, they can then be used in a forward problem to reproduce or interpolate the missing data (Nemeth, 1999).

There are some issues associated with the replacement of the migration with the least squares migration. The main problem is that the convergence of the method to a solution depends on the accuracy of the background velocity model. Without a reasonably accurate background velocity, least squares migration does not converge to the desired solution. Least squares migration is more sensitive to the accuracy of the velocity information than the migration itself (Yousefzadeh, 2008).

The second problem is that least squares migration is more computer time and memory consuming procedure than the migration. LSPSM ends up with solving a large system of linear equations. A modified version of Conjugate Gradient (CG) (Hestenes and Steifel, 1952) method, Least Squares Conjugate Gradient (LSCG) (Scales, 1987)

method has been widely used as a solver for the LSM equation (Nemeth et. al., 1999, Duquet et. al., 2000, Kuehl and Sacchi, 2001, and Yousefzadeh, 2008). However, this is still an expensive method. In solving the equation with LSCG method, each iteration requires more than two migration running time.

To show the advantage of using LSPSM instead of migration consider the acquisition geometry in Figure 1 which shows 16 sources and 96 receivers per sources. Source interval is 187.5 m and receiver interval is 15.625 m and sampling rate is 2 ms. This geometry is considered on an earth with horizontal and faulted dipping and folded layers with the velocity shown in Figure 2. The background velocity starts at 2300 m/s at the top and increases to 6300m/s at the bottom.

Using forward Kirchhoff prestack time modelling operator synthetic data are generated. Then, %0.2 random noise is added to data. Synthetic data is prestack migrated (Figure 3) and least squares prestack Kirchhoff time migrated with 20 iterations using LSCG (Figure 4). The LSPSM produced higher resolution image than the migration itself. The fault, shallower reflectors, and folds show themselves better on the image resulted from LSPSM. The migration artifacts are more attenuated if we use regularized LSM with smoothing in the offset direction as the regularization. Figure 5 shows the effect of implying this regularization with $\mu = 2 \times 10^{-7}$. Comparing Figure 5 to Figure 4, regularized LSM shows slightly higher resolution image.

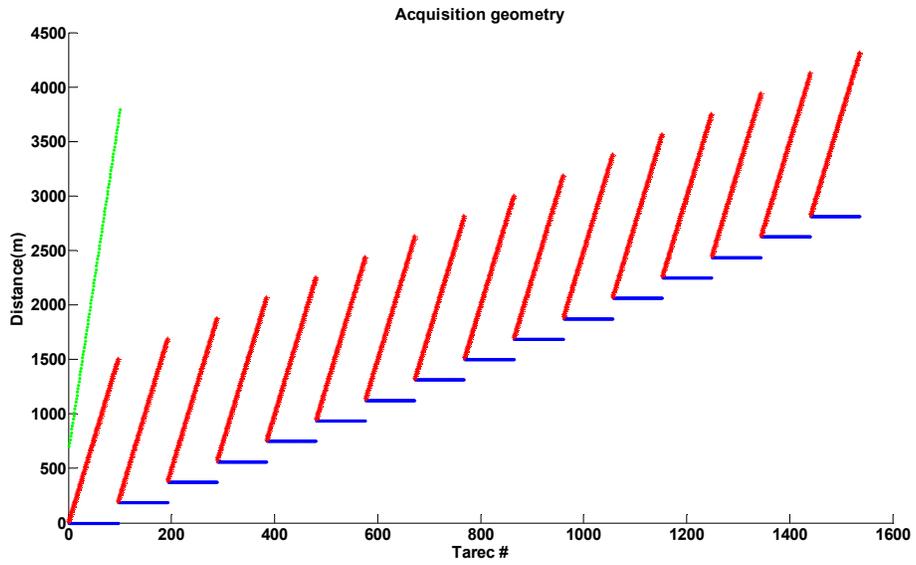


FIG. 1. Acquisition geometry used to generate synthetic data. blue: sources, red: receivers, green: CMP or image points.

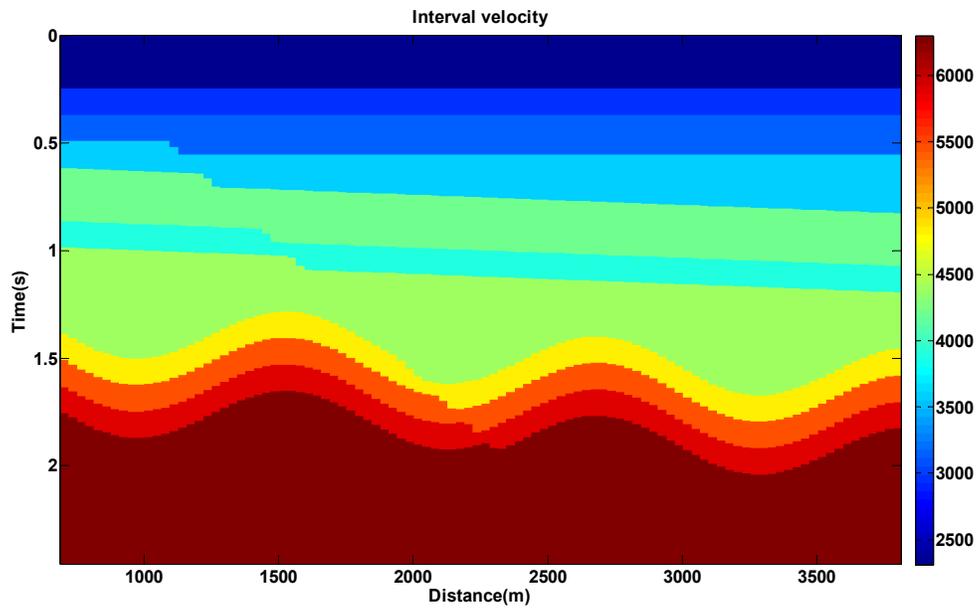


FIG. 2. Interval velocity model used for generation of synthetic data.

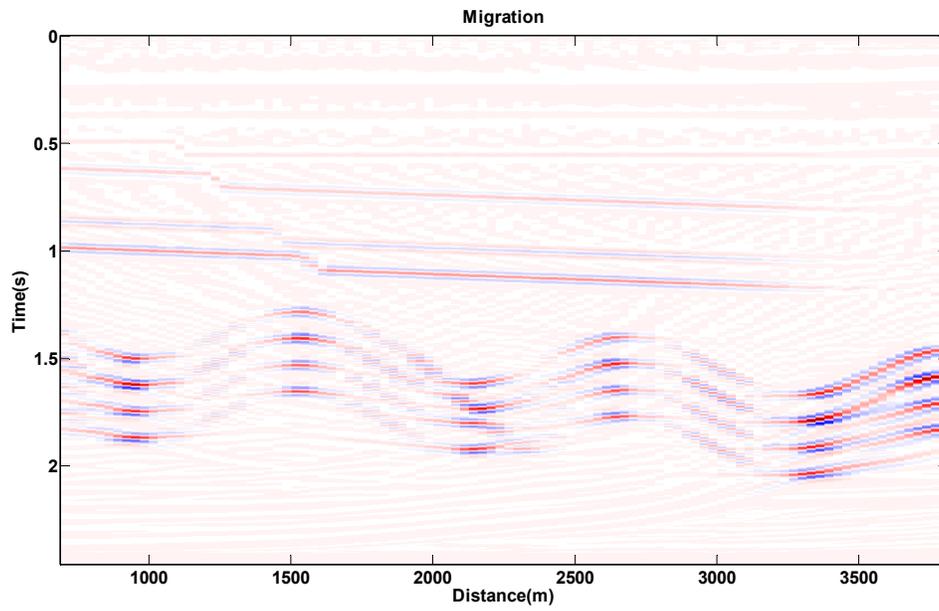


FIG. 3. Migration of the synthetic data.

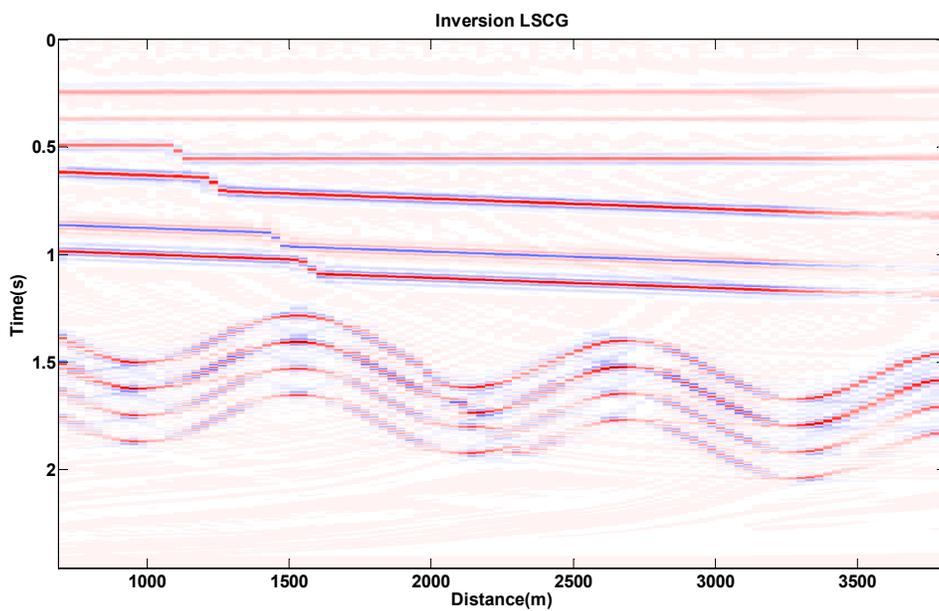


FIG. 4. LSPSM of the synthetic data.

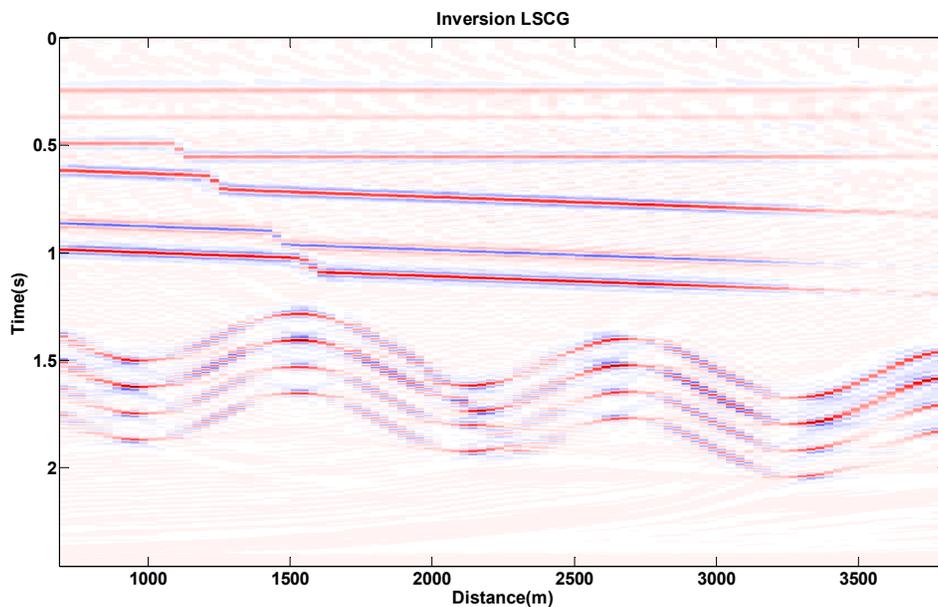


FIG. 5. Regularized (smoothing in the offset direction) LSPSM of synthetic data, $\mu = 2 \times 10^{-7}$.

Decimating some traces increases the migration artifacts. Figure 6 shows the geometry when we decimate data by regularly removing %75 of the traces. Migration of these dataset is shown in Figure 7. Applying the LSPSM to these data gives the image in Figure 8 in which most migration artifacts are attenuated.

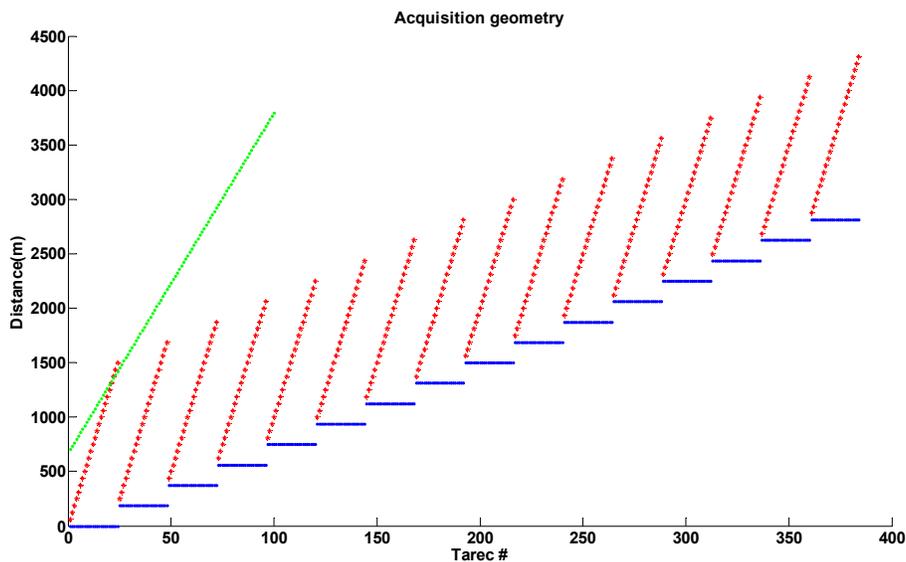


FIG. 6. Acquisition geometry for the synthetic data, when %75 of traces is removed regularly. blue: sources, red: receivers, green: image points.

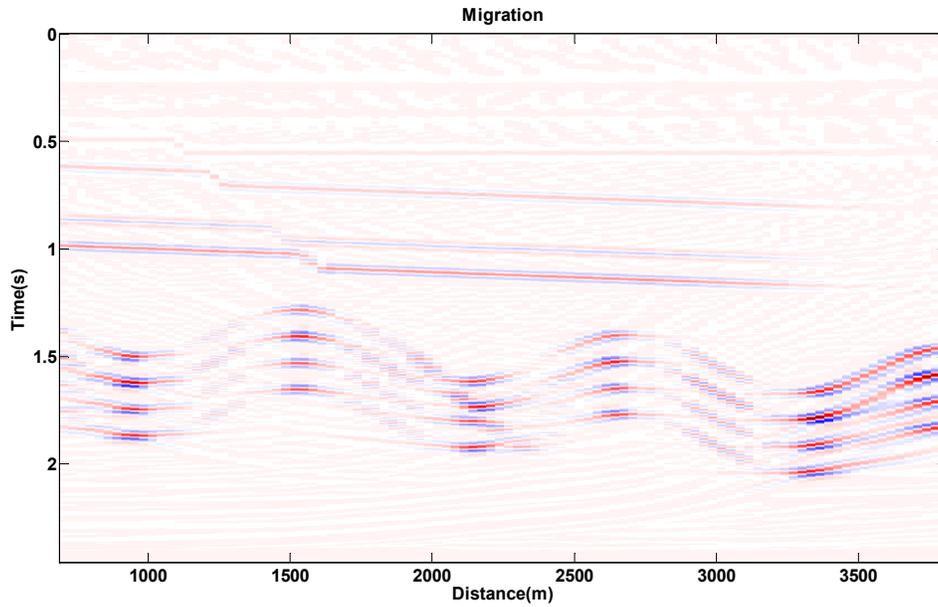


FIG. 7. Migration of synthetic data, when %75 of traces is removed regularly.

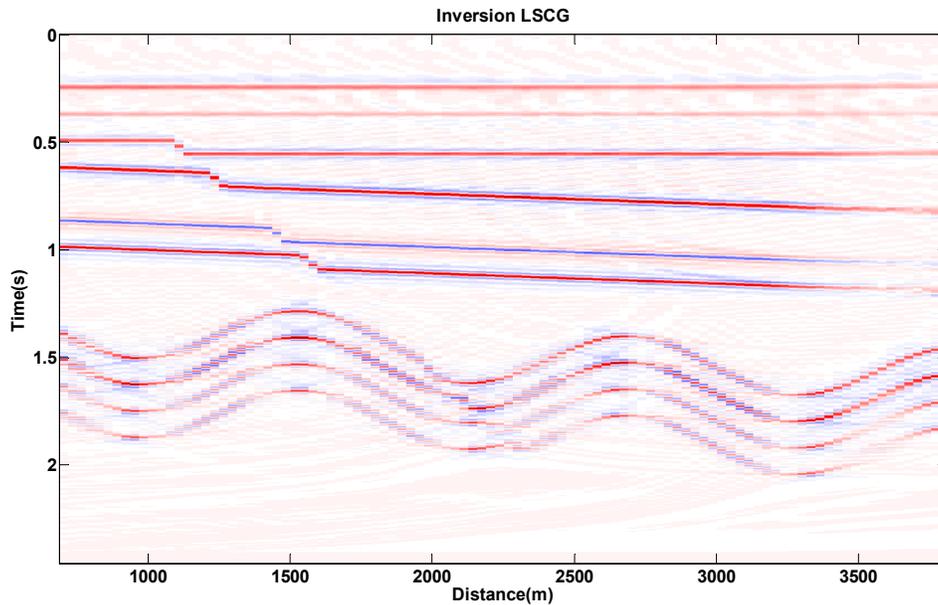


FIG. 8. LSPSM of synthetic data, when %75 of traces is removed regularly.

The same improvement with LSPSM is achieved when the data are decimated randomly and remaining portions random located as shown in Figure 9. Figures 10 and 11 show the result of migration and LSPSM of this irregularly sampled data, respectively.

As seen in these synthetic examples, LSPSM is able to attenuate migration artifacts which can be due to data incompleteness or irregularities or both.

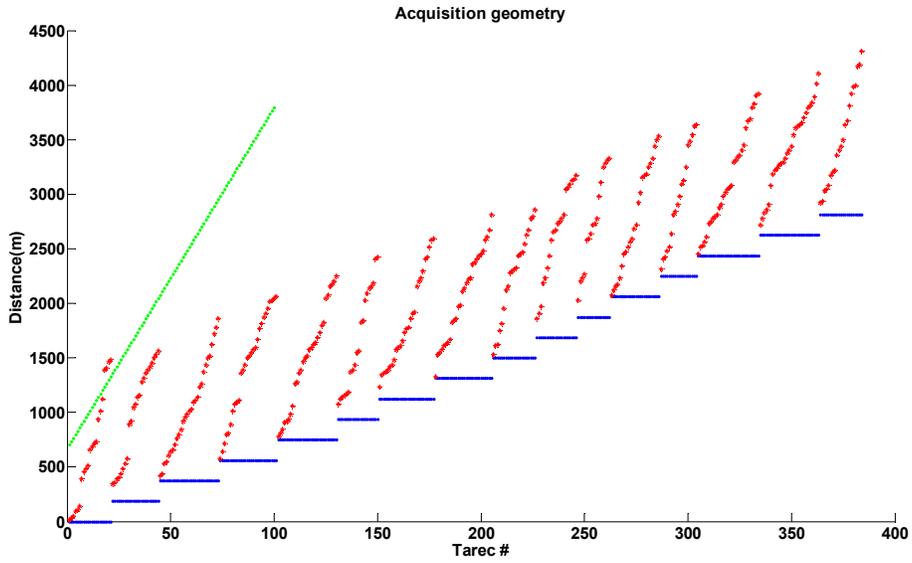


FIG. 9. Acquisition geometry for the synthetic data, when %75 of traces is removed randomly. blue: sources, red: receivers, green: image points.

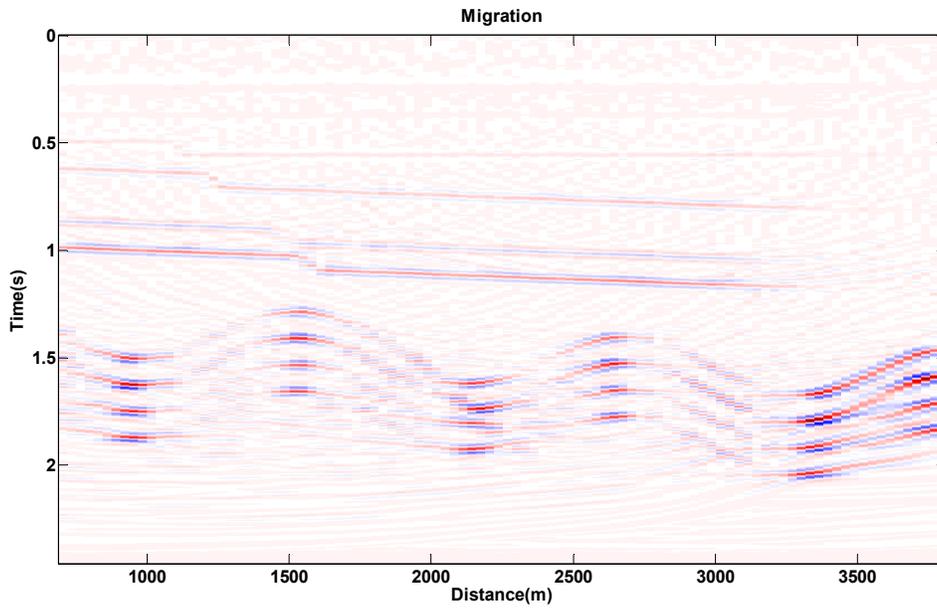


FIG. 10. Migration of synthetic data, when %75 of traces is removed randomly.

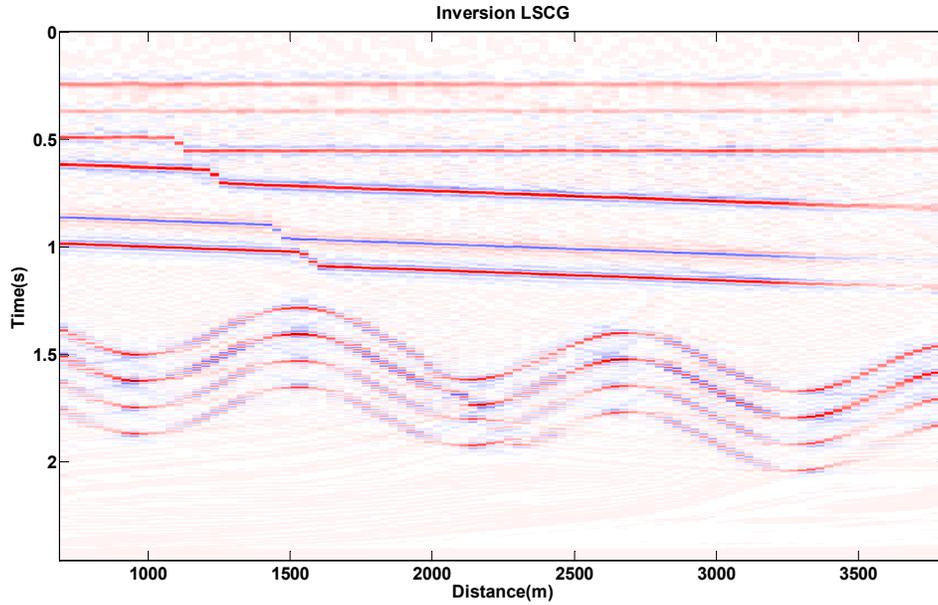


FIG. 11. LSPSM of synthetic data, when %75 of traces is removed randomly.

Data reconstruction with LSM

Data reconstruction is one of the least squares migration capabilities. There are many processes in the seismic data analysis which ask for regularly sampled data. Some methods of multiple attenuation and some powerful methods of migration that are mostly being done in the frequency domain require regularly sampled data. Replacing missing traces with traces having zero amplitude for all samples introduces some artifacts in the resulted images.

Nemeth et al. (1999) suggested data reconstruction or interpolation using LSM. Since the resulting image of LSM is a high resolution image, it can be used for the data reconstruction. Suppose \mathbf{G}_i and \mathbf{d}_i show the forward modelling operator for incomplete data, and incomplete data, respectively. Then the least squares image may be calculated by,

$$\mathbf{m}_{LS} = (\mathbf{G}_i^T \mathbf{G}_i + \mu^2 \mathbf{I})^{-1} \mathbf{G}_i^T \mathbf{d}_i. \quad (15)$$

Then complete data, \mathbf{d} , are achievable by forward modelling of \mathbf{m}_{LS} :

$$\mathbf{d} = \mathbf{G} \mathbf{m}_{LS} = \mathbf{G} (\mathbf{G}_i^T \mathbf{G}_i + \mu^2 \mathbf{I})^{-1} \mathbf{G}_i^T \mathbf{d}_i. \quad (16)$$

This method of data reconstruction is robust. However this method is relatively more expensive than the other methods of data reconstruction which usually use properties of the Fourier transform and works with a small part of data (one shotgather for instance) at

each time (for methods see: Spitz, 1991, Porsani, 1999, Gulunay, 2003). Using LSPSM, the observed data must tolerate several migration and modelling performances in the LSCG scheme until the algorithm converges to a reasonable model, and then there is an additional forward modelling with the new geometry of the sources and receivers.

To show the ability of data reconstruction with LSPSM, we used the image of LSPSM in Figure 8 for reconstruction when %75 of data which were decimated regularly. Results in the first 50 traces from shotgather # 10 are shown in Figure 12. Figures 12a, 12b, 12c, and 12d show the original, decimated, reconstructed, and residual data, respectively. With having only %25 of data, LSPSM is a powerful method to reconstruct the missing data. The same results are achieved with irregularly decimated data as shown in Figure 13.

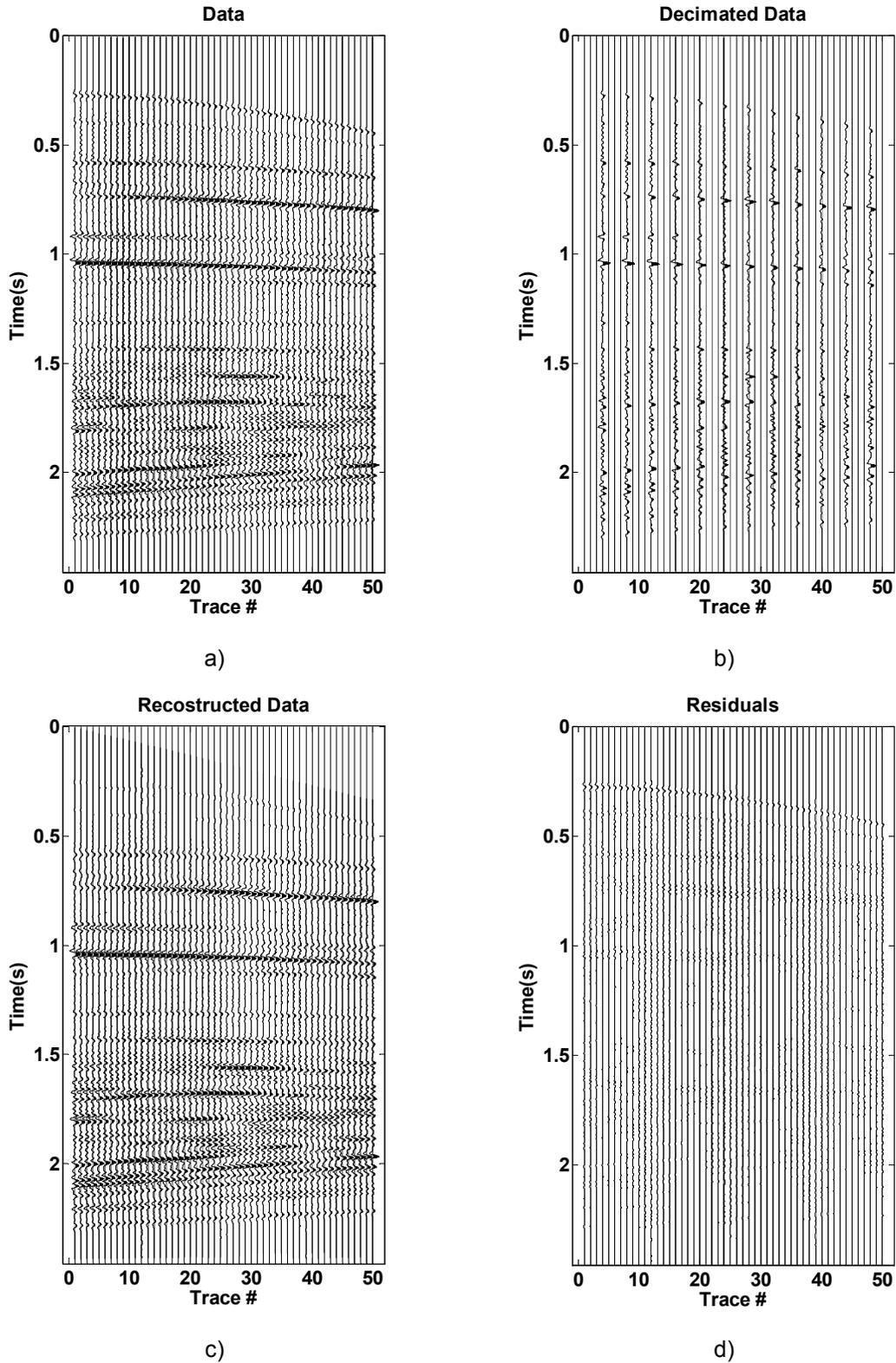


FIG. 12. Reconstruction of data by LSM. %75 of data are removed regularly. a) data, b) decimated data, c) reconstructed data, d) residuals. First 50 traces from shotgather # 10 are shown.

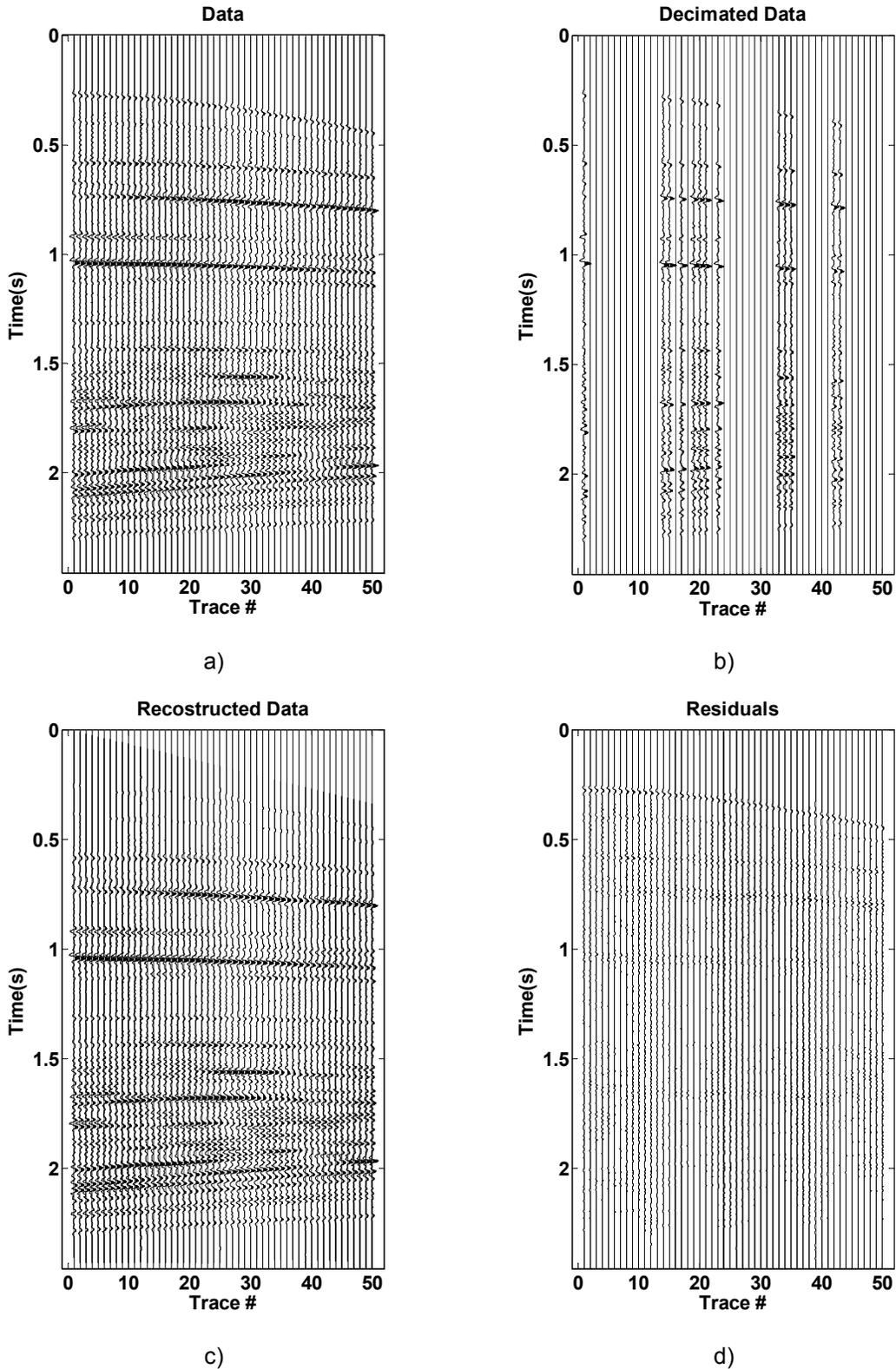


FIG. 13. Reconstruction of data by LSM. %75 of data are removed randomly. a) data, b) decimated data, c) reconstructed data, d) residuals. First 50 traces from shotgather # 10 are shown.

Solving the least squares migration equation

The performance of least squares seismic inversion requires solving an usually large system of linear equations. When $\mathbf{G}^T \mathbf{G}$ is not a large matrix, inverse problem can be solved using direct methods. However, in the real seismic imaging problems, \mathbf{G} is a large matrix and its size equals to the number of grids in the model domain multiplied by the number of observations (data). In such cases the direct methods are not efficient.

When $\mathbf{G}^T \mathbf{G}$ is a positive definite matrix, Steepest Descent is a powerful method to solve equation (9). In this method, the cost function has a paraboloid shape and its minimum is the solution to equation (9) (Shewchuk, 1994). The Steepest Descent algorithm starts at an arbitrary point and follows the opposite direction of maximum gradient until gradient changes, and then it changes the direction to find a new minimum. The procedure continues until it arrives at a global minimum or exact solution.

In the method of CG which is an improvement to the Steepest Descent, the new direction is orthogonal to the all previous directions. Therefore CG finds the solution in less iterations compare to the method of Steepest Descent (Hestenes and Steifel, 1952).

CG requires that $\mathbf{G}^T \mathbf{G}$ be a symmetric and positive definite matrix. LSCG, a modified version of CG method, does not require this condition and directly works with \mathbf{G} and \mathbf{G}^T matrices. If equation $\mathbf{Gm} = \mathbf{d}$ is an overdetermined problem, then $\mathbf{G}^T \mathbf{G}$ is nonsingular and LSCG converges to solve equation $\mathbf{G}^T \mathbf{Gm}_{LS} = \mathbf{G}^T \mathbf{d}$. However, since the condition number of the matrix $\mathbf{G}^T \mathbf{G}$ is square of the condition number of the matrix \mathbf{G} , convergence of LSCG method is slower than the convergence of the CG method (Shewchuk, 1994).

A LSCG's algorithm to solve $\mathbf{G}^T \mathbf{Gm}_{LS} - \mathbf{G}^T \mathbf{d} = \mathbf{0}$, is in the follow:

$$\mathbf{m}_0 = \mathbf{0}$$

$$\mathbf{s}_0 = \mathbf{d} - \mathbf{Gm}_0$$

$$\mathbf{r}_0 = \mathbf{G}^T \mathbf{s}_0$$

$$\mathbf{p}_0 = \mathbf{r}_0$$

$$\mathbf{q}_0 = \mathbf{Gp}_0$$

$$i = 0$$

For $i = 1$: # of iterations

$$\alpha_{i+1} = \frac{\mathbf{r}_i \cdot \mathbf{r}_i}{\mathbf{q}_i \cdot \mathbf{q}_i}$$

$$\mathbf{m}_{i+1} = \mathbf{m}_i + \alpha_{i+1} \mathbf{p}_i$$

$$\mathbf{s}_{i+1} = \mathbf{s}_i - \alpha_{i+1} \mathbf{q}_i$$

$$\mathbf{r}_{i+1} = \mathbf{G}^T \mathbf{s}_{i+1} \quad *$$

$$\beta_{i+1} = \frac{\mathbf{r}_{i+1} \cdot \mathbf{r}_{i+1}}{\mathbf{r} \cdot \mathbf{r}}$$

$$\mathbf{p}_{i+1} = \mathbf{r} + \beta \mathbf{p}_i$$

$$\mathbf{q}_{i+1} = \mathbf{G} \mathbf{p}_{i+1} \quad **$$

$$i = i + 1$$

endfor

There are two multiplications of \mathbf{G} or \mathbf{G}^T with a vector (in the lines marked by “*” and “**”, respectively) in each iteration. Since there is no decomposition of \mathbf{G} or \mathbf{G}^T in the algorithm, it is possible to use operators instead of explicit forms of \mathbf{G} or \mathbf{G}^T matrixes. This reduces the required memory. However, this means that each iteration in the LSCG has twice computational cost of one migration (or modelling). For example 10 iterations in the LSCG are equal to 20 migrations in the time cost.

Each Kirchhoff prestack time migration includes calculation and/or application of the Double Square Root equation, proper weight function, rho filter, antialiasing filter, cross-correlation, and interpolation. In order to have a modelling operator which is exactly the adjoint of the migration operator, the adjoint of all these steps must be considered in the modelling as well. This procedure causes the modelling operator to be even more expensive than the migration operator. For instance, adding triangle anti aliasing filter to the migration operator makes the modelling operator two times more expensive than the migration.

For linear problems with any starting point, the Steepest Descent and CG methods always converge to the solution after n iterations, where n is the number of unknowns (Hestenes and Steifel, 1952). However, in the case of nonlinear problems, the best result is achievable when the starting point is considered to be close to the global minimum of the quadratic function. In a linear problem, any rough estimate of \mathbf{m} is suitable as the initial guess (Shewchuk, 1994). If there is not any other information about \mathbf{m} , $\mathbf{m} = \mathbf{0}$ is best choice.

LSPSM images in Figures 4, 8, and 11 are calculated with 20 iterations in LSCG. Convergence rate in the CG method is an appropriate way of measuring the ability of the LSPSM in finding the best reflectivity image. In this paper, convergence rate is the relative Euclidean norm of the difference between the original data and the data achieved by forward modelling of LSPSM image after each iteration. Figure 14 shows the convergence rate for these images. As seen in this figure, LSCG is an efficient method to solve LSM problems. With known background velocity, residuals reduced to %10 in less than 10 iterations.

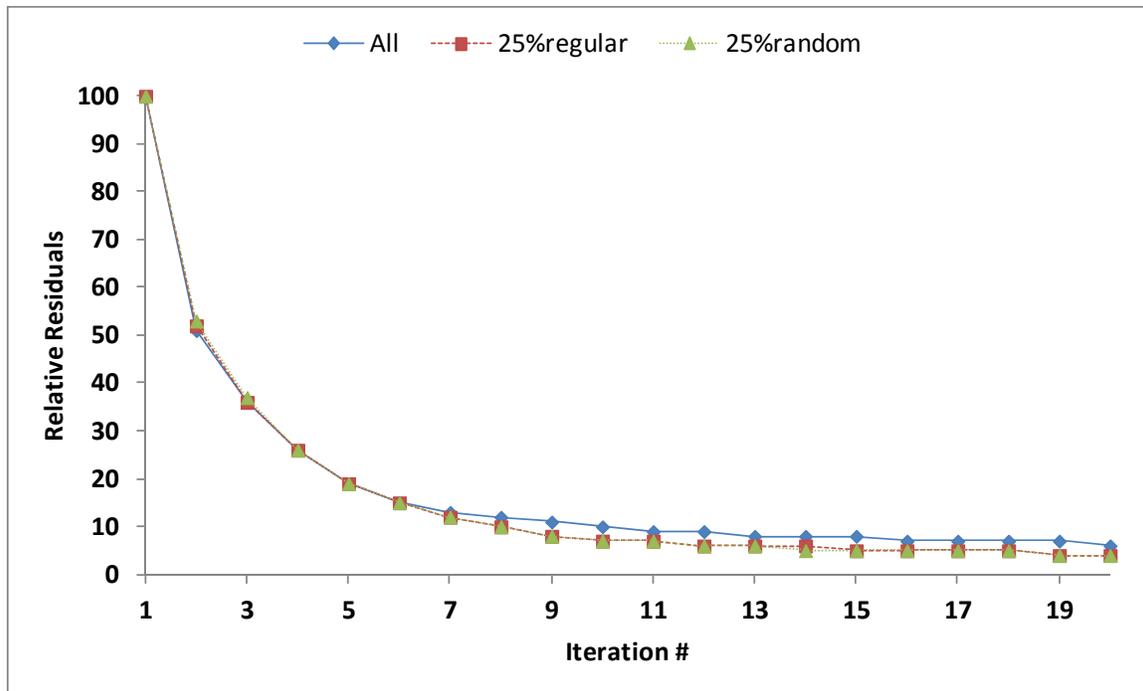


FIG. 14. Convergence of LPSTM with all data, and %75 regularly and randomly decimated data.

MIGRATION AND LSPSM VELOCITY ANALYSIS

As the industry moves to exploit smaller hydrocarbon resources in the areas with sophisticated geology, more advanced imaging techniques are required to provide a higher resolution of the subsurface image. LSM is one of these methods. Any advanced imaging method requires relatively accurate velocity information. Extracting proper velocity information is a main step for subsurface imaging and in seismic data processing in general. LSPSM recovers a high resolution image of subsurface reflectivity. In order to recover the unknown reflectivity, \mathbf{m} , in LSPSM we have to define Kirchhoff modelling and migration operators, \mathbf{G} and \mathbf{G}^T , as accurately as possible. These operators need a good estimation of the rms velocity to do migration and modelling correctly.

A proper subsurface velocity information leads to a well-focused migration image. An inaccurate velocity model distorts the migration image. However, these distortions have some useful information about the velocity and can be used to modify the velocity model used in imaging.

In the simplest method of velocity analysis in the data domain, one can use a CMP gather which has been NMO corrected several times using a plausible range of constant velocities. Correct velocity at each time results in flat events at that time. With a lower or higher velocity implemented, events will be over- or under-corrected, respectively. Flatness of the events can be used as a measurement tool for finding better velocity estimation at each time. Proper velocity at each time gives the maximum coherency at that time. Semblance is a commonly used type of the coherency measurements which can be calculated by,

$$S = \frac{1}{M} \frac{\sum_t (\sum_{i=1}^M f_{i,t(i)})^2}{\sum_t \sum_{i=1}^M f_{i,t(i)}^2}, \quad (17)$$

where $f_{i,t(i)}$ is the amplitude of i th trace at time $t(i)$ and M is the number of traces that contributes to the measurements. Semblance is the energy of the stack normalized by the energy of the components of stack (Sheriff, 2002).

By calculating semblance at each time and for all constant velocities in the range, a panel of semblance is obtained. In this panel the horizontal axis is the velocity and the vertical axis is two way traveltime. By choosing the maximum semblance at each time, the best velocity for that time in the corresponding CMP position is achieved. Yilmaz (2008) motioned a complete list of the other methods of coherency measurements which are useful for velocity analysis.

NMO correction and semblance analysis are based on horizontal reflectors but will focus dipping events with higher velocity referred to as stacking velocities. With the rms velocity, NMO correction is not able to flatten dipping events. Velocity analysis on image domain instead of data domain can be used.

Migration velocity analysis can be implemented by performing migration with a range of constant velocities and creating many migration images. Then comparing best image focusing for each time, we achieve a migration velocity function. In a better way this can be done by doing velocity analysis on each CIG, instead of the whole image. Here, we extend this method to the CIGs resulted from LSPSM.

Prestack Kirchhoff migration of 2D data produces a migrated section with two spatial coordinates, x and τ . It is not difficult to divide the integration of Kirchhoff migration to a few sets of different recording offsets (and azimuths in 3D) (Biondi, 2007). The result will be a cube (or hypercube, 4D) and each section is a “prestack partial image” as called by Biondi (2007). CIG or Common Reflection Point (CRP) referred as a part of the image that corresponds to the specific surface point.

Offset domain CIG are equivalent to CMP gathers after NMO correction. However, in the image gathers, the focused energy is migrated and comes from the area below the CMP position. Hence, they are better than the data domain gathers for velocity estimation of dipping layers. Velocity analysis based the on semblance method can be extended to the migration velocity analysis on offset domain CIGs.

There are three common types of CIG. Offset domain CIG is probably easiest one to produce and work with. In the offset domain CIG, horizon axis is the absolute source-receiver distance. In the angle domain CIG, horizon axis shows the incident angle parameter. In shot domain CIG, the distance between image point and source is the horizontal axis. Both, offset domain and shot domain CIGs can be can be sassily obtained by prestack Kirchhoff migration. Our focus in this paper is on the offset domain and shot domain CIGs. We show the effect of error in velocity information on the LSPSM images and then we show relative advantages of using CIGs from LSPSM instead of CIGs from migration.

Effect of the velocity accuracy on LSPSM

To see the effect of the velocity accuracy on LSPSM, let study two cases with error in the velocity information. First, the velocity is %10 more than the true velocity, and second, the velocity is %10 less than the true velocity.

Consider a data acquisition geometry as in Figure 1 and the velocity model shown in Figure 15. Data are produced using these geometry and model, then after adding one percent random noise, the data are decimated by removing %75 of the traces, regularly. This data are migrated using the true velocity. Figure 16 shows the result of migration. Result of LSPSM is shown in Figure 17. LSPSM image has higher resolution and the reflectors show better resolution and continuity as we expected from the previous discussions.

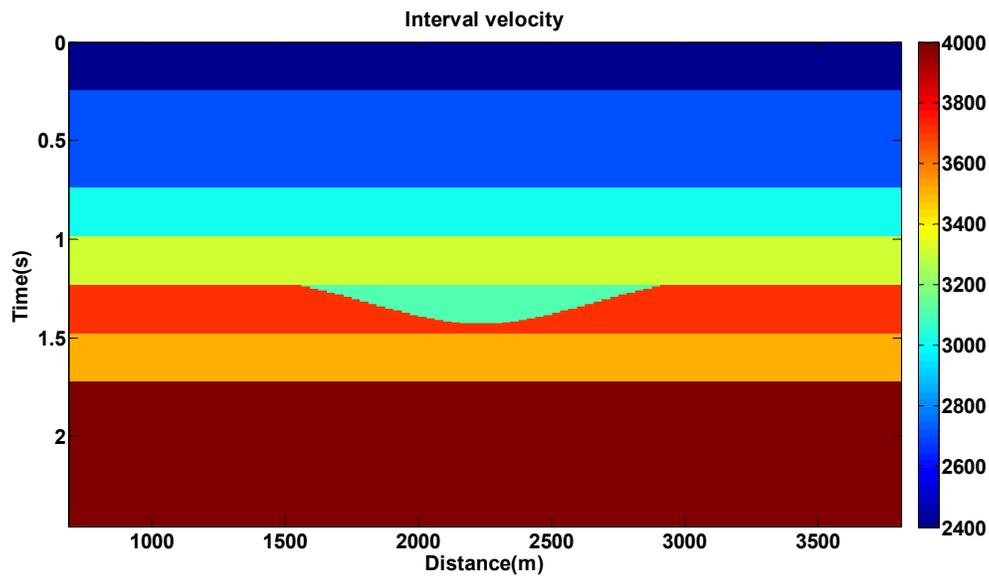


FIG. 15. Velocity model used for studying the effect of velocity errors on LSPSM.

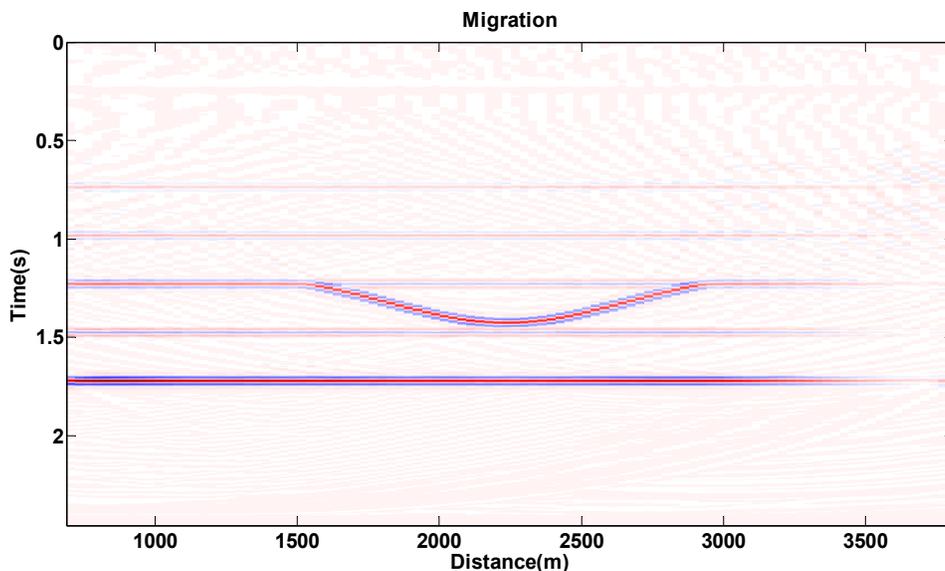


FIG. 16. Migration image of %75 decimated data with the true velocity.

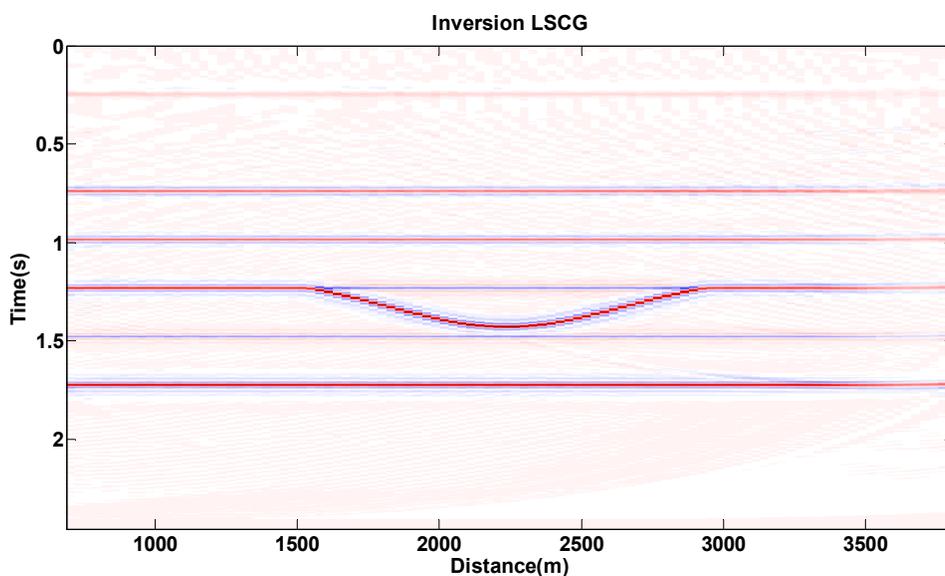


FIG. 17. LSPSM image of %75 decimated data with the true velocity.

Now, the complete data, (not the decimated data), are migrated with a velocity that is %10 lower than the true velocity (Figure 18). LSPSM of same data with %10 lower than the true velocity is shown in Figure 19. This time, LSPSM not only does not improve the migration resolution, but also it seems to introduce more noise to the resulted image. This image has been used for data reconstruction. Figure 20 shows the result of the data. Figures 20a, 20b, 20c, and 20d are the first 50 traces from shot number 10 in the original data, reconstructed data, the difference between original and reconstructed data, and the difference between the original and reconstructed data when true velocity is used in

LSPSM, respectively. Where Figure 20d shows a very good reconstruction of data, Figure 20c shows the data reconstruction failed because of using only %10 lower than the true velocity. Same result is achieved with using a velocity that is %10 higher than the true velocity. Figures 21, 22, and 23 show the migration image, LSPSM image, and data reconstruction, respectively, when velocity is %10 higher than the true velocity and complete data are being used.

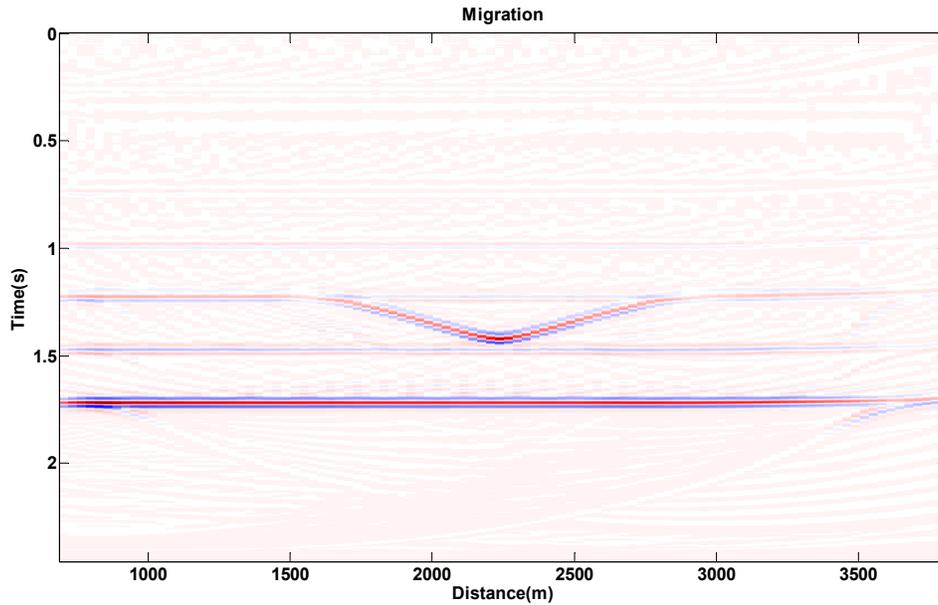


FIG. 18. Migration of complete data with the velocity %10 lower than the true velocity.

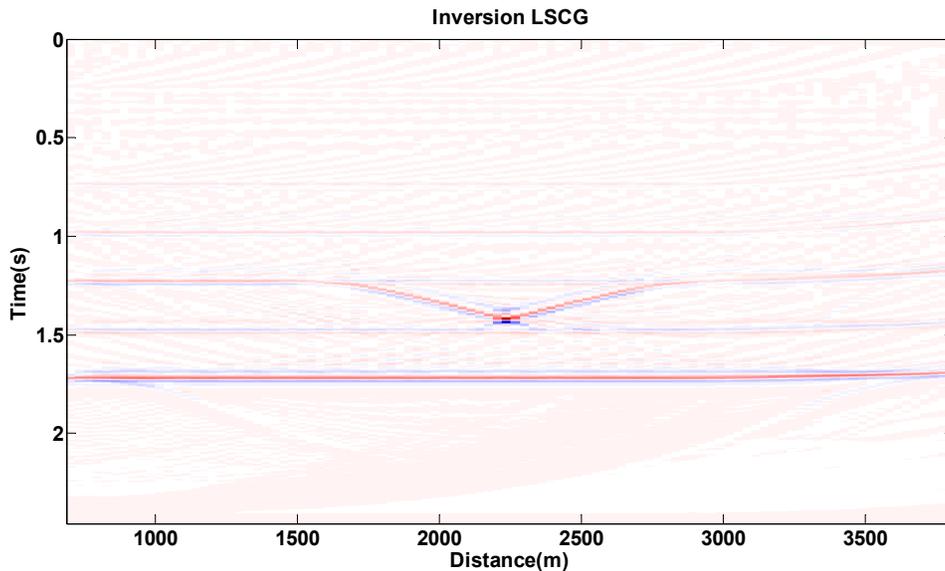


FIG. 19. LSPSM of complete data with the velocity %10 lower than the true velocity.

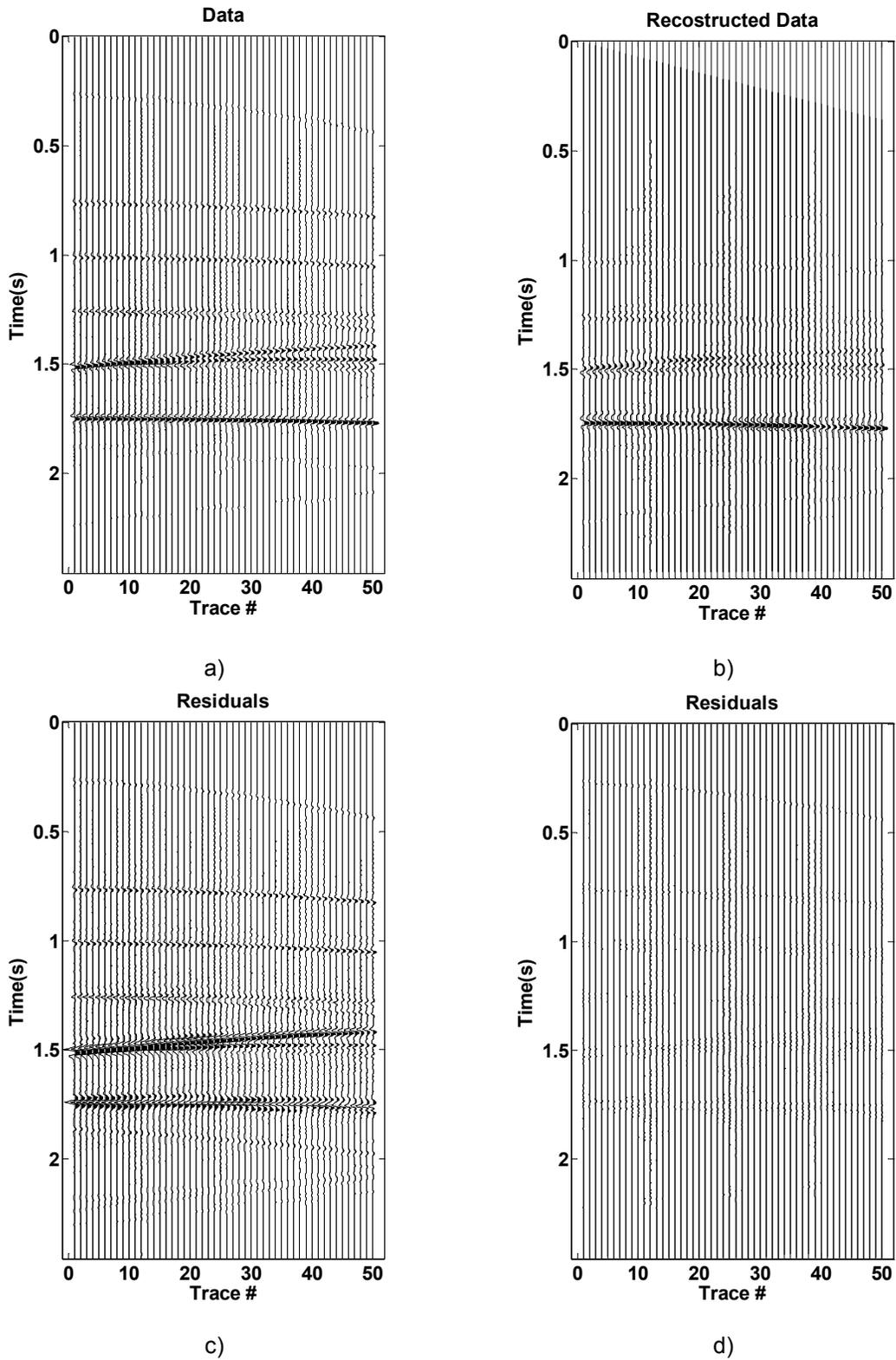


FIG. 20. Reconstruction of data by LSPSM. Complete data are used but velocity is 10% lower than the true velocity. a) data, b) reconstructed data, c) residual with wrong velocity, d) residual with true velocity. First 50 traces from shotgather # 10 are shown.

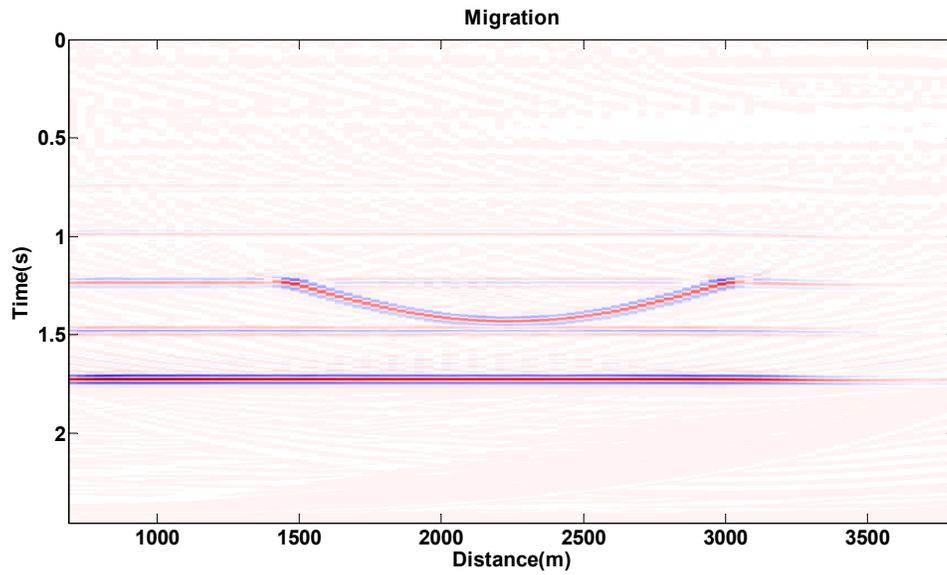


FIG. 21. Migration of complete data with the velocity %10 higher than the true velocity.

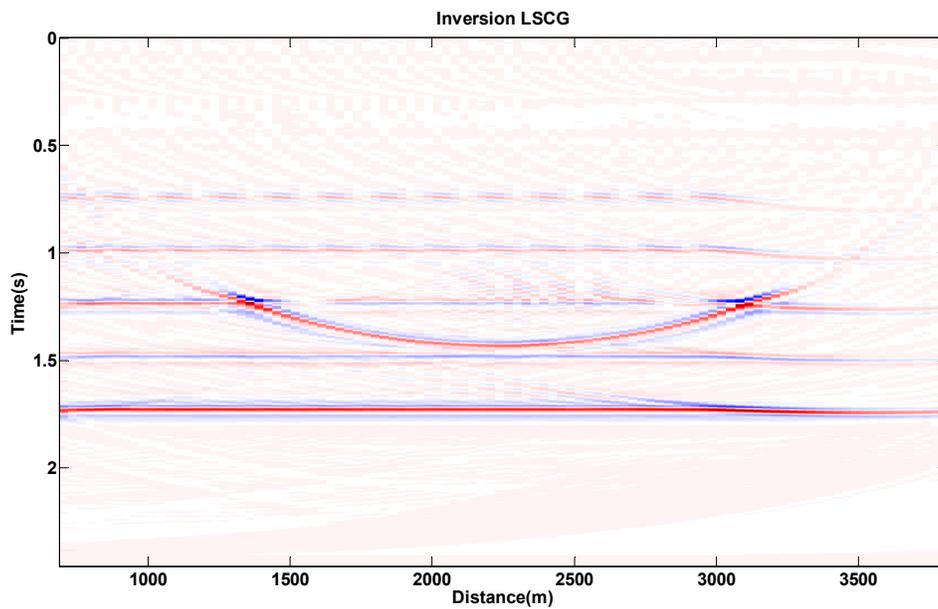


FIG. 22. LSPSM of complete data with the velocity %10 higher than the true velocity.

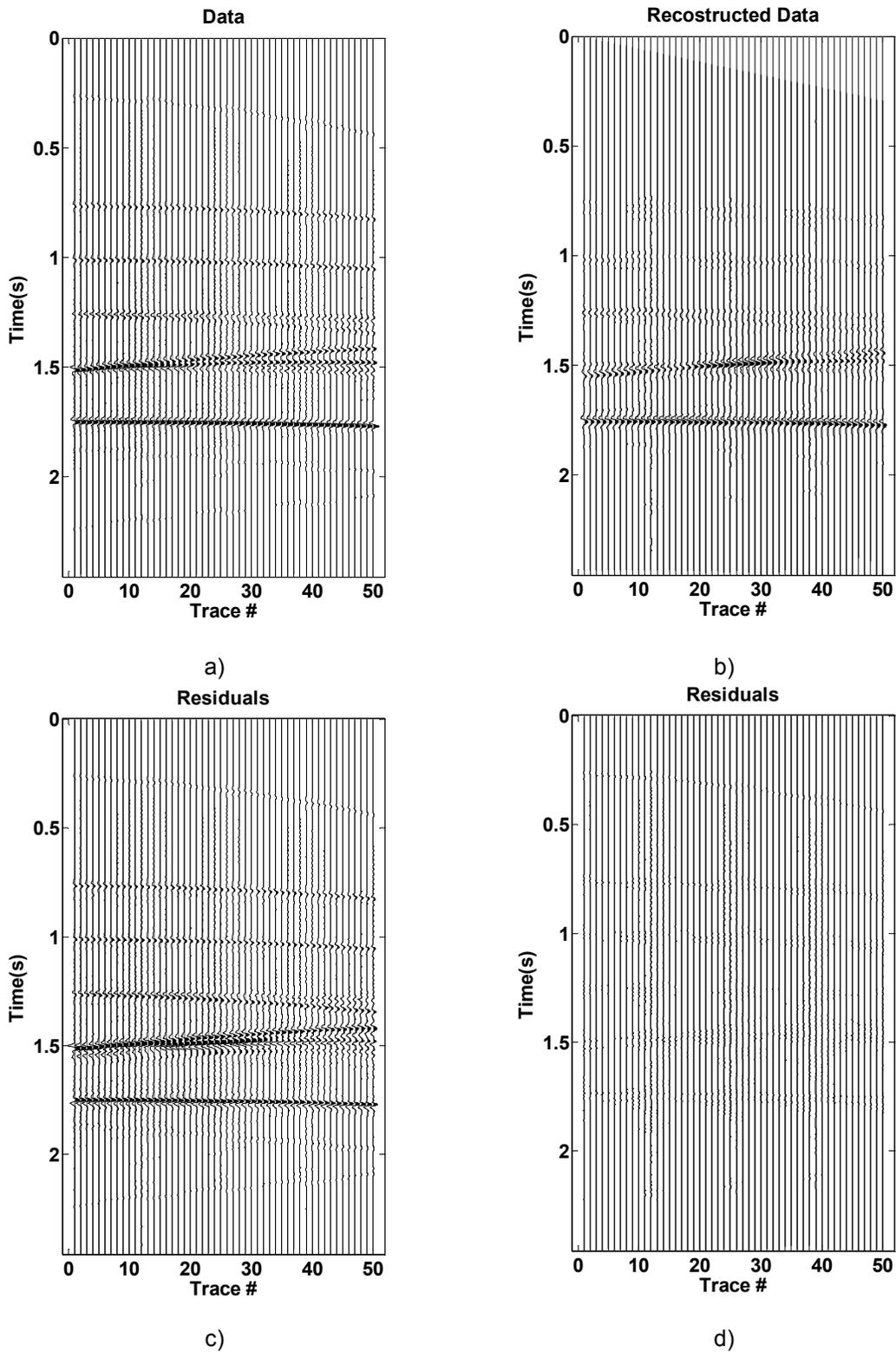


FIG. 23. Reconstruction of data by LSPSM. Complete data are used but velocity is 10% higher than the true velocity. a) data, b) reconstructed data, c) residual with wrong velocity, d) residual with true velocity. First 50 traces from shotgather # 10 are shown.

Figure 24 shows the convergence rate for these three examples of LSPSM. As seen when the residuals converges to %10 in three iterations in case of using true velocity, it will not go to less than %75 with %10 higher velocity and %85 with %10 lower velocity is used in the LSPSM.

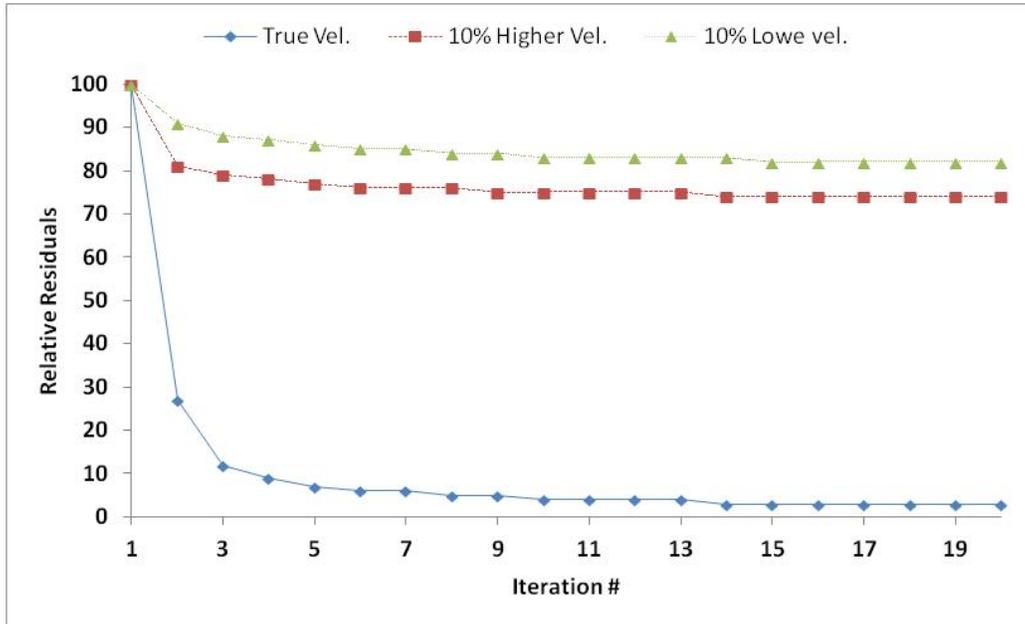


FIG. 24. Convergence of LSPSM in 20 iteration with true, %10 more, and %10 less velocity

Velocity analysis on LSPSM offset domain CIGs

Now we construct the offset domain CIGs by dividing data into ten offset bins and migrating and performing LSPSM on the each offset bin separately using true velocity. The final image can be obtained by stacking images from all bins together. Figure 25 shows the migration of first offset bin and Figure 26 shows the stack of images from all offset bins. Figures 27a and 27b show the corresponding offset domain CIG at $x = 2250\text{ m}$ with using %25 of regularly sampled data, after migration and LSPSM, respectively. Figures 28 and 29 show same CIG when the implemented velocity is %10 lower or higher than the true velocity, respectively. Offset domain CIGs from LSPSM have slightly higher resolution especially in the shallower parts of the CIGs. The higher resolution image of offset domain CIGs resulted from LSPSM compare to migration image make it attractive to use them for velocity estimation by the semblance method.

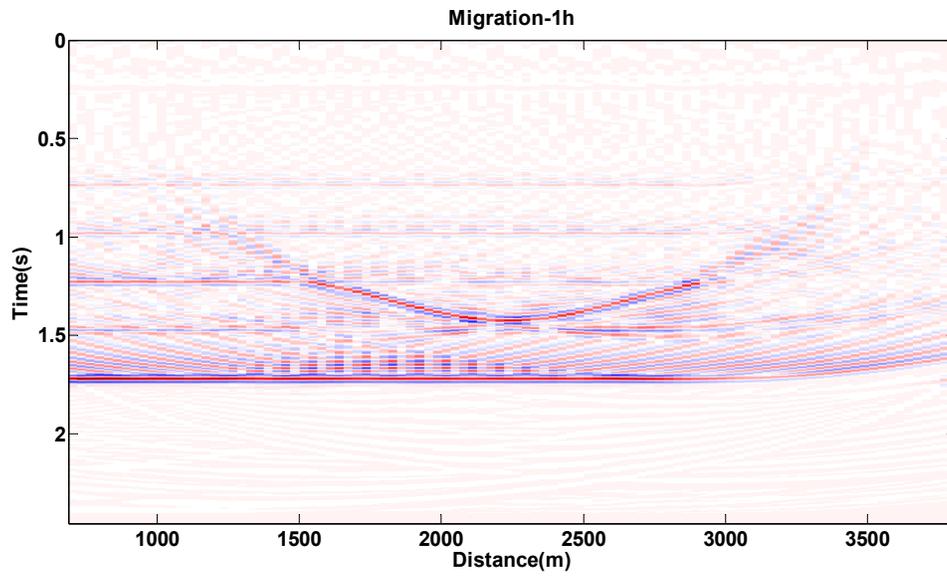


FIG. 25. Migration of first offset bin with true velocity.

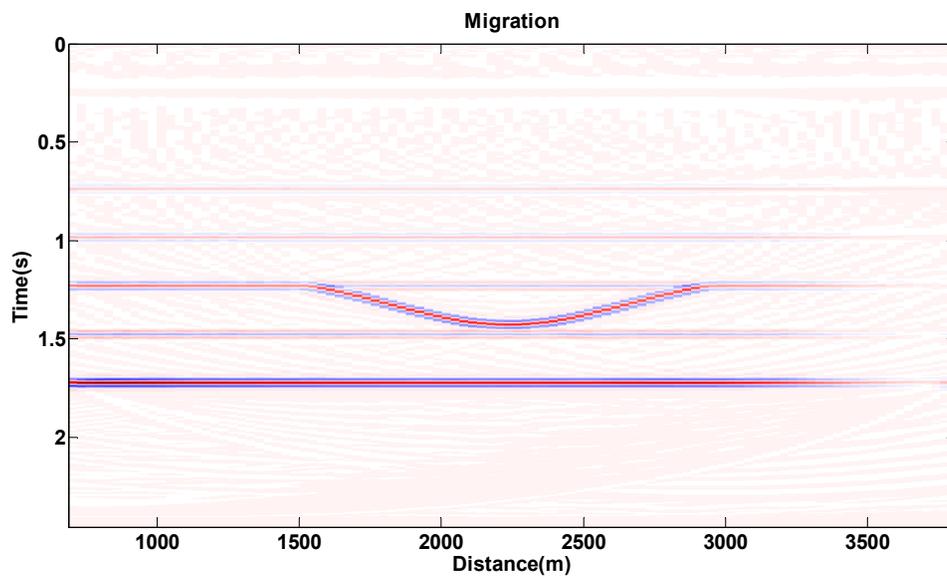


FIG. 26. Image resulted from stacking all offset bins migration images.

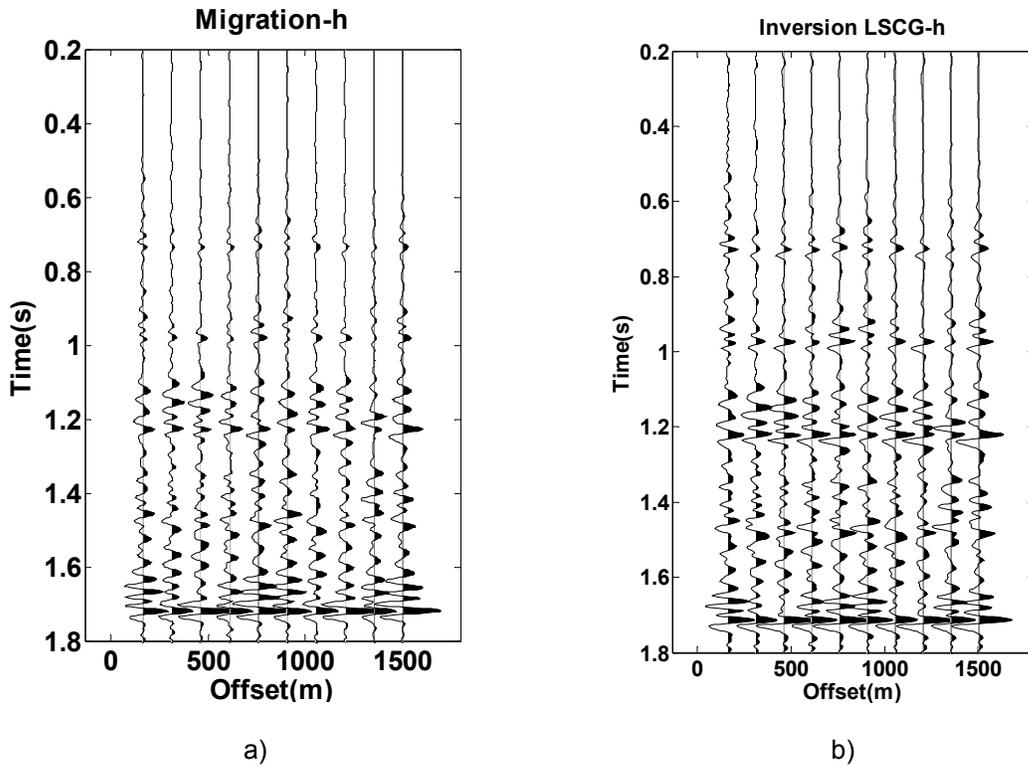


FIG. 27. Offset domain CIG at $x = 2250$ m with the true velocity from a) migration and b) LSPSM.

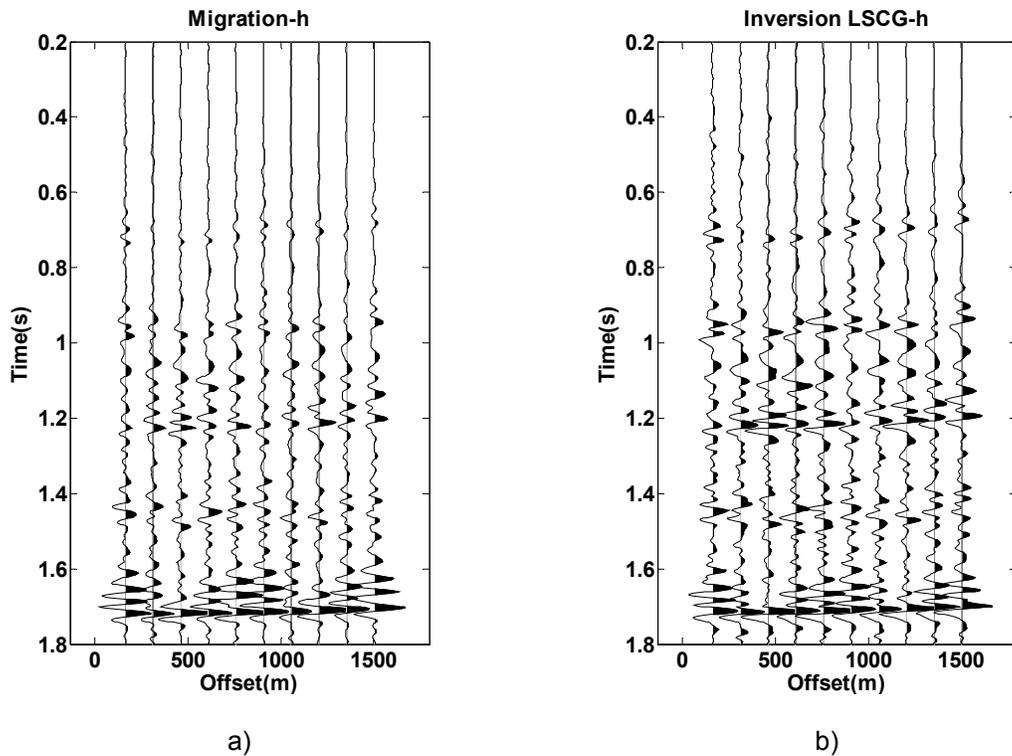


FIG. 28. Offset domain CIG at $x = 2250$ m with %10 lower velocity from a) migration and b) LSPSM.

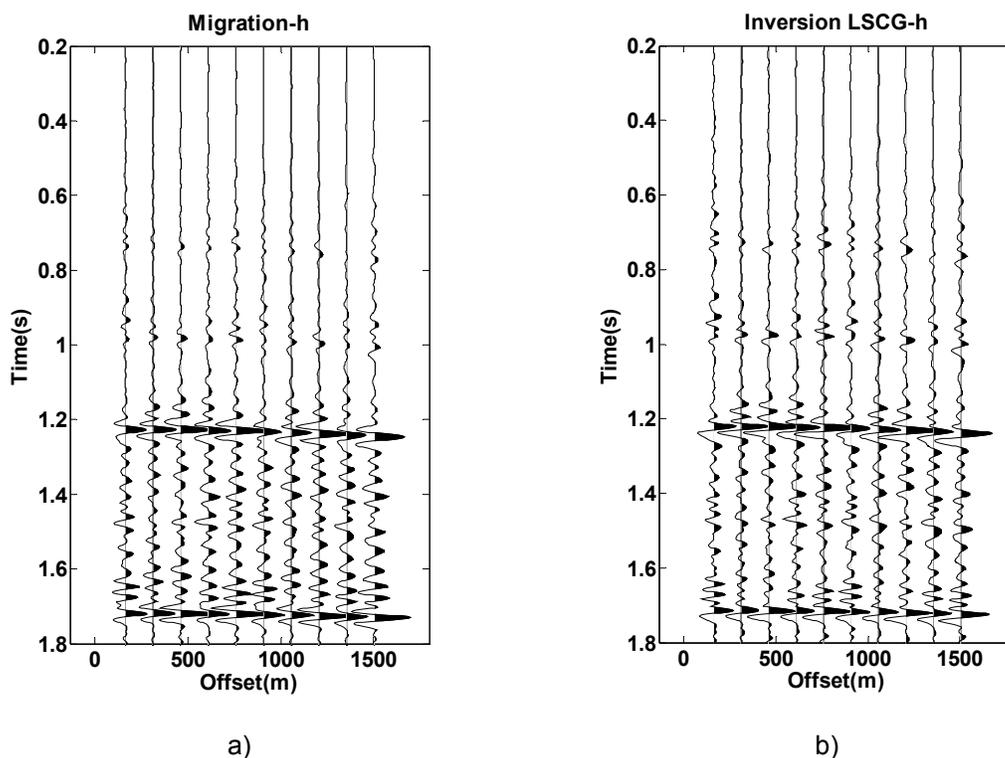


FIG. 29. Offset domain CIG at $x = 2250\text{ m}$ with %10 higher velocity from a) migration and b) LSPSM.

In this section we use LSPSM offset domain CIGs for velocity analysis and compare it with velocity analysis based on migration offset domain CIGs. For this reason first we performed velocity analysis on %50 regularly sampled data using semblance method on CMP gathers, offset domain CIG gathers from migration, and offset domain CIG gathers from LSPSM. The velocity range starts from 2000m/s and with the increment of 50m/s increases to 4500m/s.

Figures 30, 31, and 32 show the semblance spectrums at the position $x = 2250\text{ m}$ with %50 of data. Figure 30 shows the semblance spectrum for CMP gather, and Figures 31, and 32 show the semblance spectrum for offset domain CIG from migration and LSPSM, respectively. The difference between the spectrum resulted from LSPSM and the migration is negligible.

However the improvement in the semblance by using LSPSM instead of migration becomes more visible when most of the data are decimated. Figures 33, 34, and 35 show the semblance spectrum on data, on migration offset domain CIG, and on LSPSM offset domain CIG at the same position, $x = 2250\text{ m}$, when %90 of data regularly decimated.

Comparison between Figures 34 and 35 shows that offset domain CIG from LSPSM gives a better image for the velocity estimation than the migration offset domain CIGs.

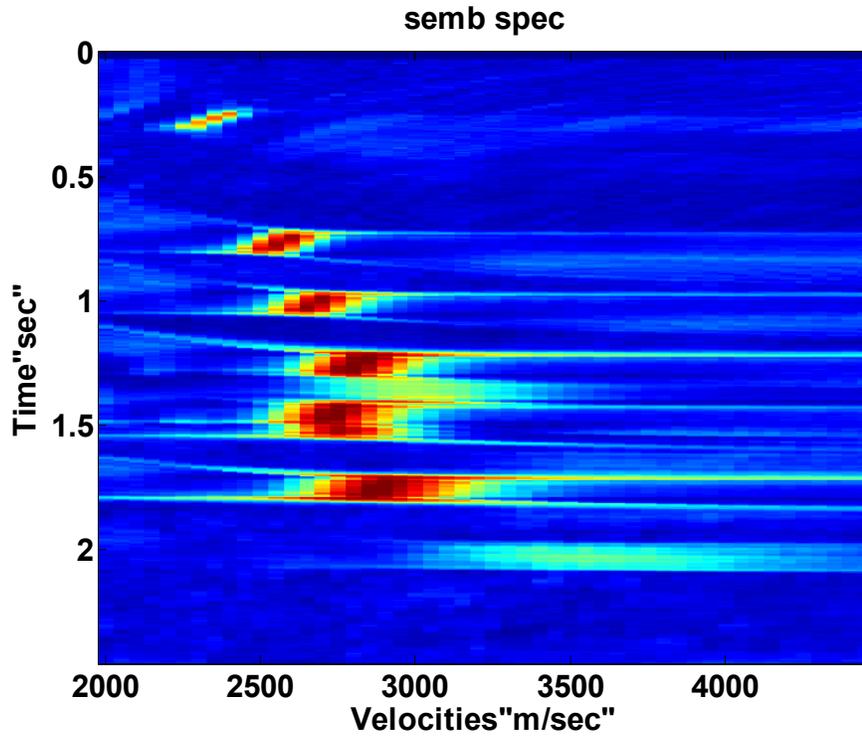


FIG. 30. Semblance spectrum for CMP at $x = 2250\text{ m}$ with %50 of data.

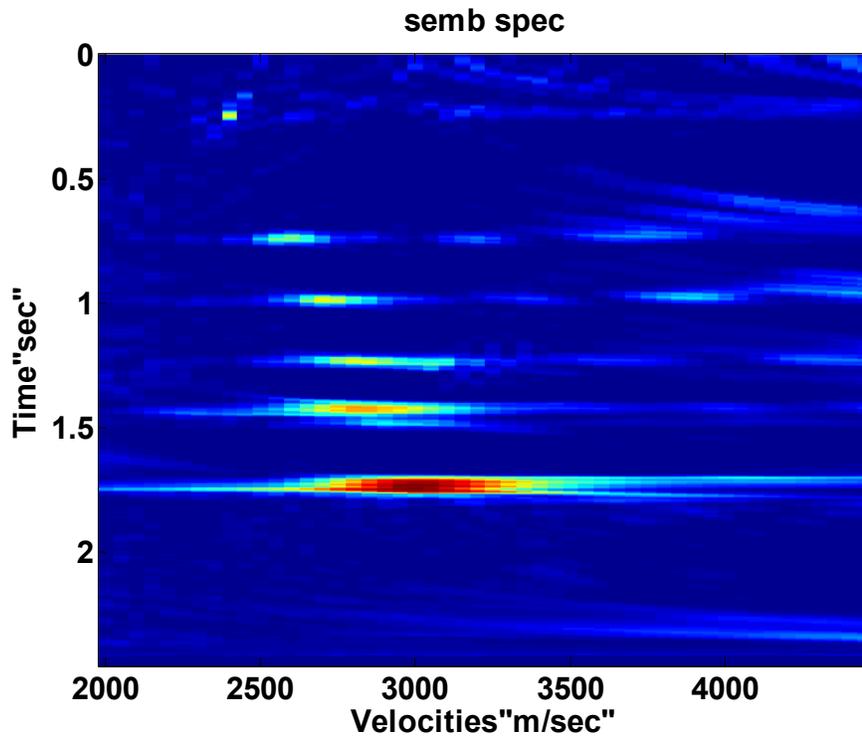


FIG. 31. Semblance spectrum for Migration offset domain CIG at $x = 2250\text{ m}$ with %50 of data.

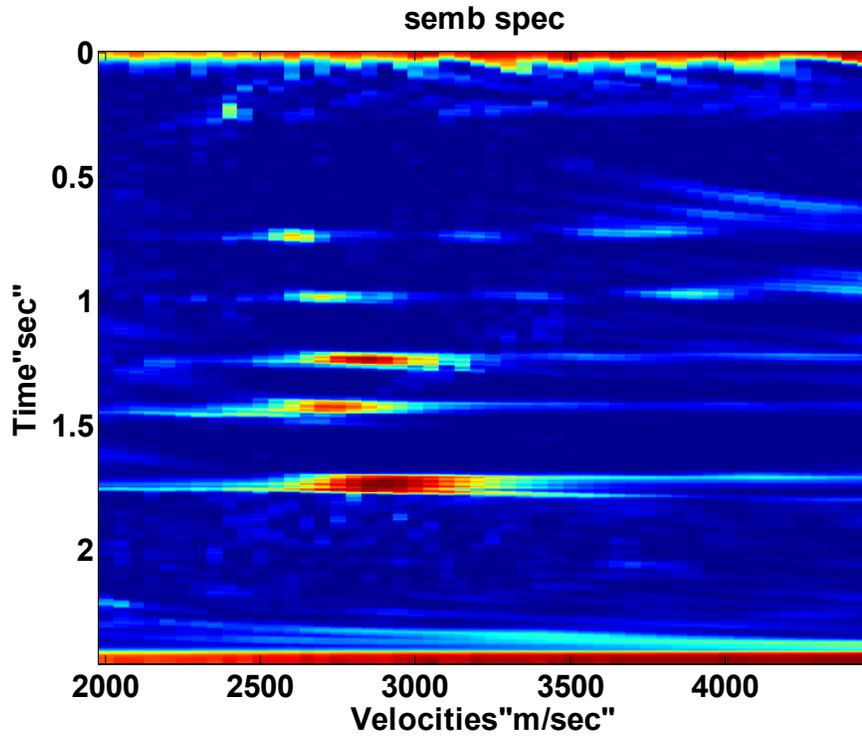


FIG. 32. Semblance spectrum for LSPSM offset domain CIG at $x = 2250\text{ m}$ with %50 of data.

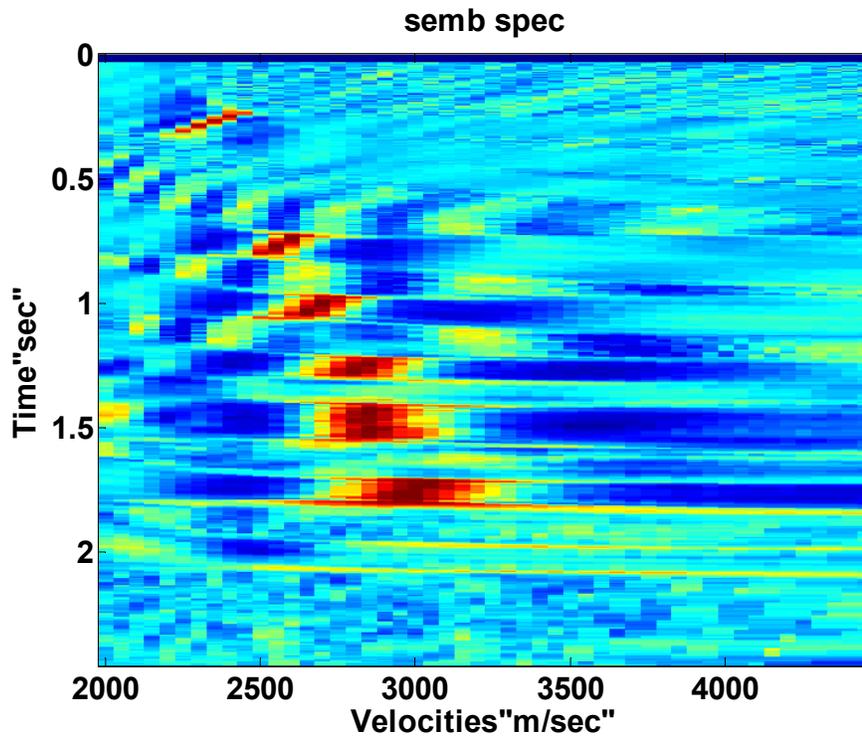


FIG. 33. Semblance spectrum for CMP at $x = 2250\text{ m}$ with %10 of data.

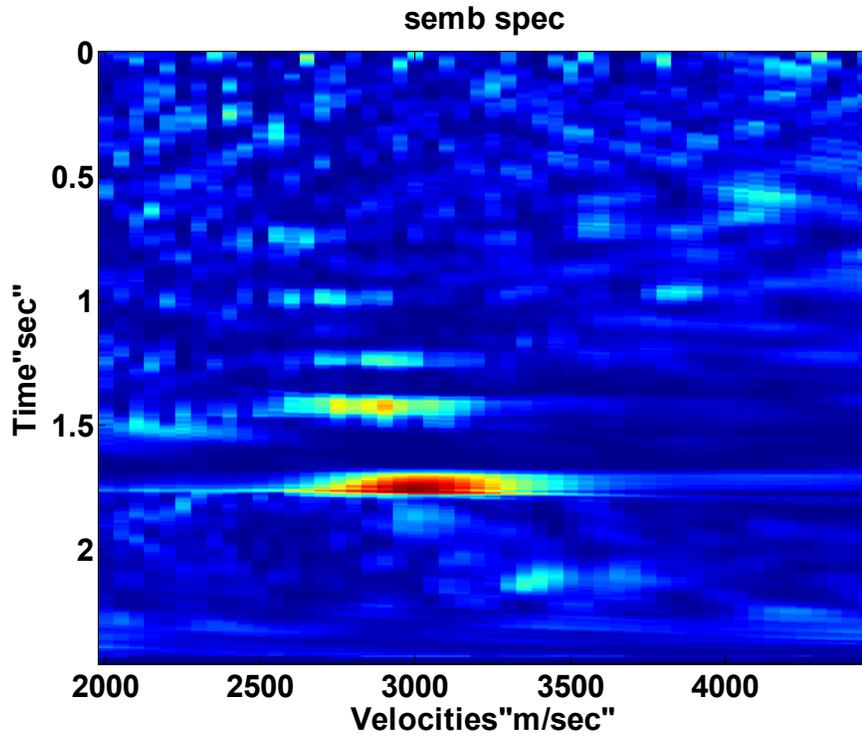


FIG. 34. Semblance spectrum for Migration offset domain CIG at $x = 2250\text{ m}$ with %10 of data.

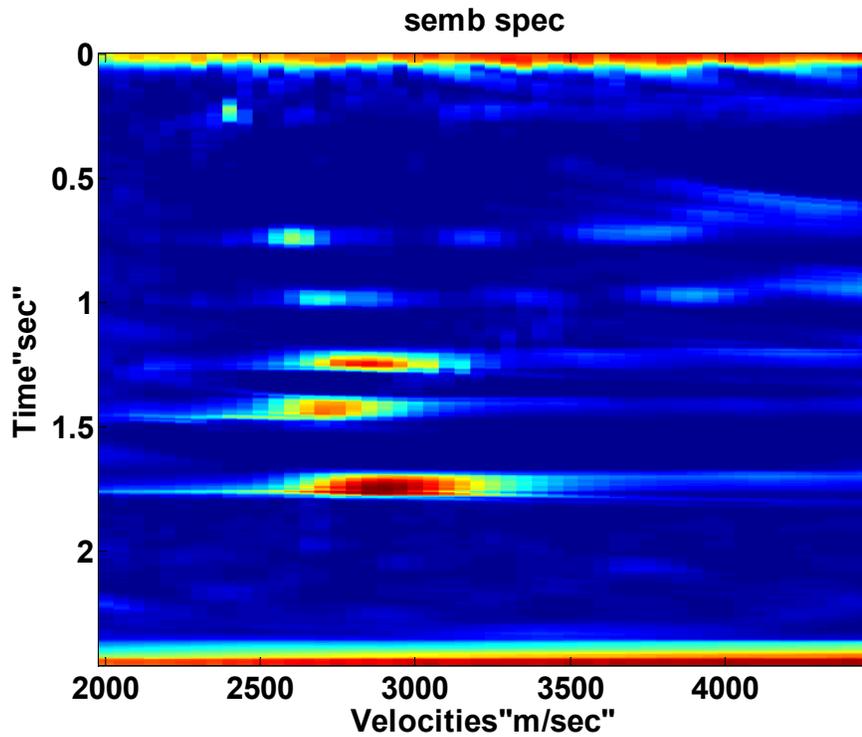


FIG. 35. Semblance spectrum for LSPSM offset domain CIG at $x = 2250\text{ m}$ with %10 of data.

With the velocity information extracted by this method on the %10 of regularly decimated data, decimated data are migrated. Figure 36 shows the migration result. Figure 37 shows the LSPSM result when the same data and velocity is used. Migration image is blurred but LSPSM is able to give a high resolution image with less blurriness.

This image is used for the reconstruction of %90 missing data. Figure 38 shows the result. In this figure, 38a shows the decimated data and 38b shows the reconstructed data. Difference between original and reconstructed data is shown in Figure 38c. Figure 38d shows when true velocity is used for data reconstruction. Comparison between Figures 38c and 38d shows that the data reconstruction was successful.

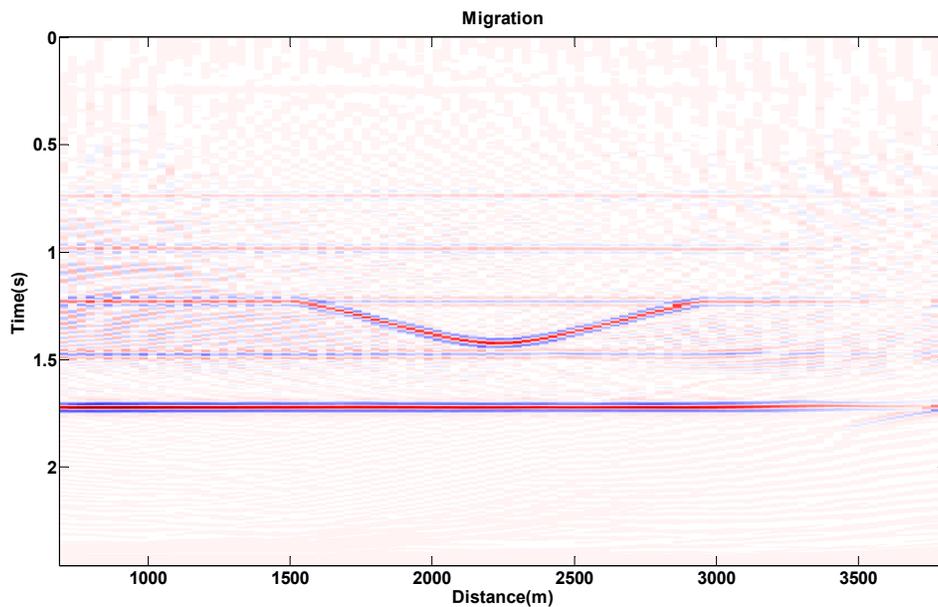


FIG. 36. Migration of %10 regularly decimated data with the velocity extracted with semblance analysis on LSPSTM offset domain CIGs.

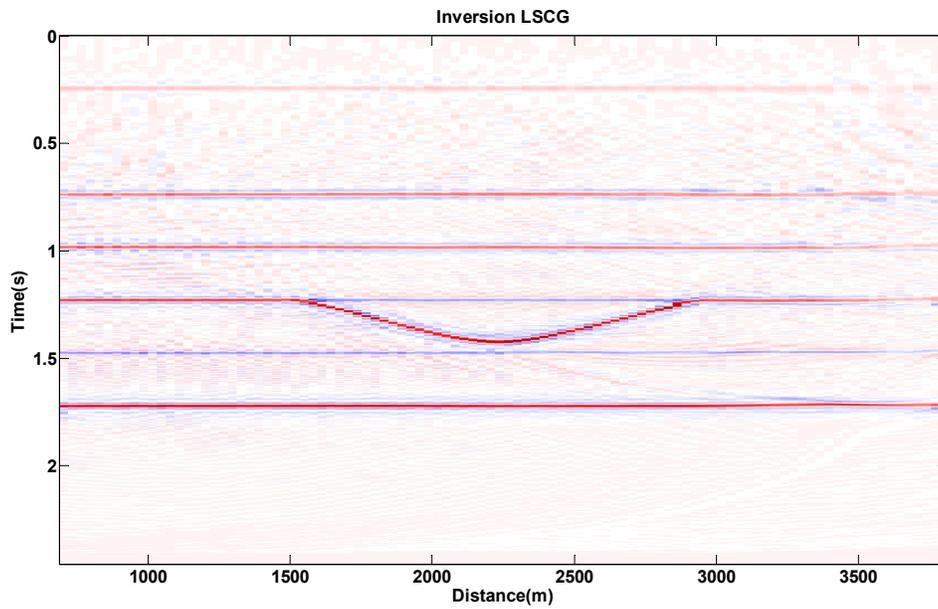


FIG. 37. LSPSTM of %10 regularly decimated data with the velocity extracted with semblance analysis on LSPSTM offset domain CIGs.

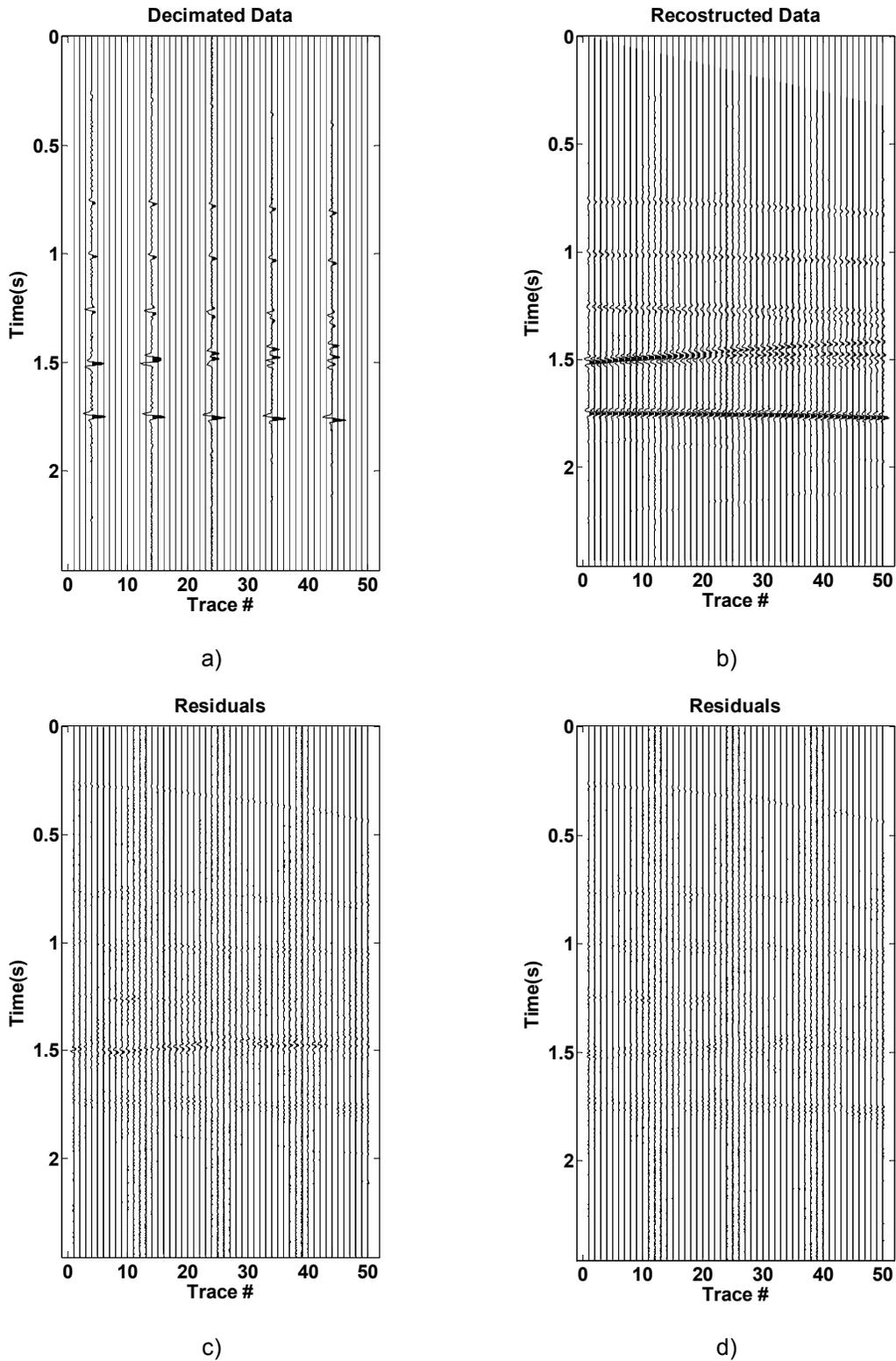


FIG. 38. Reconstruction of %90 data by LSPSM using extracted velocity information. a) decimated data, b) reconstructed data, c) residual using extracted velocity, d) residual with true velocity. First 50 traces from shotgather # 10 are shown.

To see accuracy of the extracted velocity information, it is also useful to compare the convergence rate of the LSPSM with true velocity and when the achieved velocity is implemented. Figure 39 shows the convergence when true, %10 more, %10 less, and the extracted velocity is used in LSPSM. In one iteration LSPSM with the extracted velocity converge to %31 of the original difference where this amount with true velocity is %27. Finally in 20 iterations, the extracted velocity cause the convergence rate goes down to %7, only %3 more than when the true velocity has been used for LSPSM.

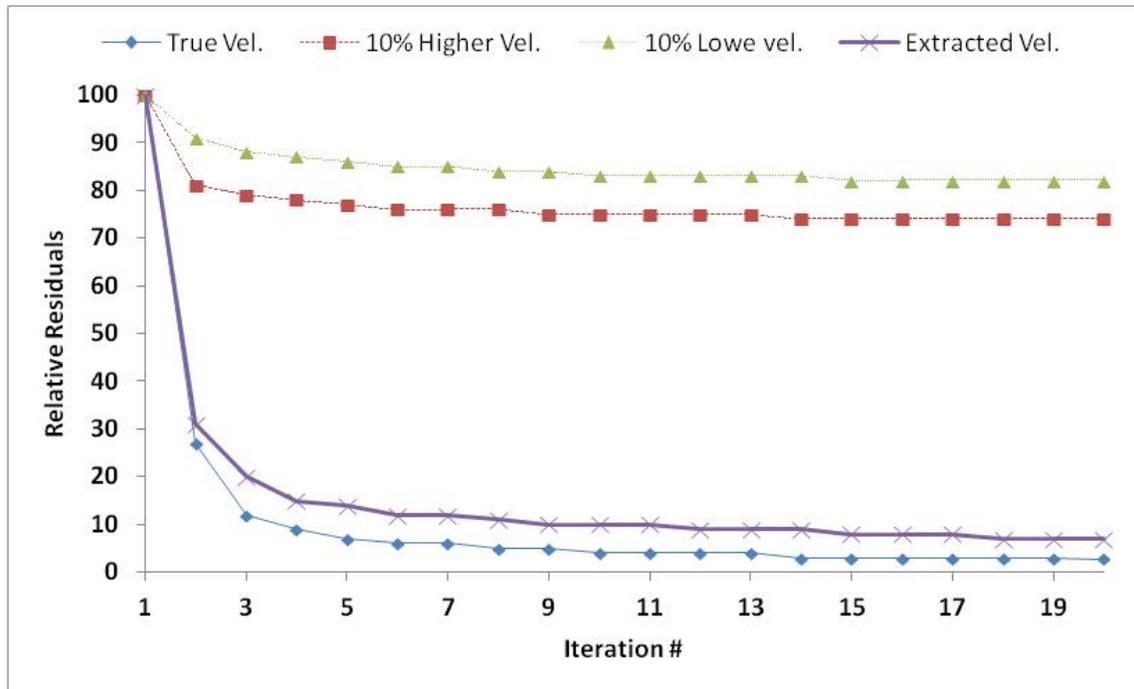


FIG. 39. Convergence of LSPSM in 20 iteration with true, %10 more, and %10 less and extracted velocity.

Velocity analysis on LSPSM Shot domain CIGs

In this section, the idea of velocity analysis on LSPSTM offset domain CIGs is extended to the LSPSM shot domain CIGs. We use LSPSM shot domain CIGs for velocity analysis and compare it with velocity analysis based on migration shot domain CIGs.

In the shot domain CIGs, the horizontal axis shows the distance between the source position and the corresponding image position. Therefore, the number of traces per CIG is equal to the number of shot points. The number of the shot points for a 2D seismic line can be relatively a large number, therefore, we limit the number of traces in the shot domain CIGs by binning the maximum source to image point distance to a reasonably small number. The number of bins depends on the shot distance and the complexity of

area. The best choice of the number of bins is when there are between two to five shot points stacked in one CIG bin.

For all analyses in this section, synthetic data created for the acquisition geometry shown in Figure 1 and the velocity model shown in Figure 15. The mentioned geometry includes 16 sources. We realized that 8 is a reasonable number for binning the source image point distance. Binning to the higher number introduces some artifacts to the migration of each bin. With this binning, and using the complete dataset, we performed migration and LSPSM.

Figure 40 shows two migration shot domain CIGs in the surface positions $x = 900m$ and $x = 2250m$. Figure 41 shows the two LSPSM shot domain CIGs from the same positions. Migration and LSPSM of the data in the shot domain CIGs are shown in Figures 42 and 43. LSPSM images show higher resolution than the migration image in both, the shot domain CIGs and the final images.

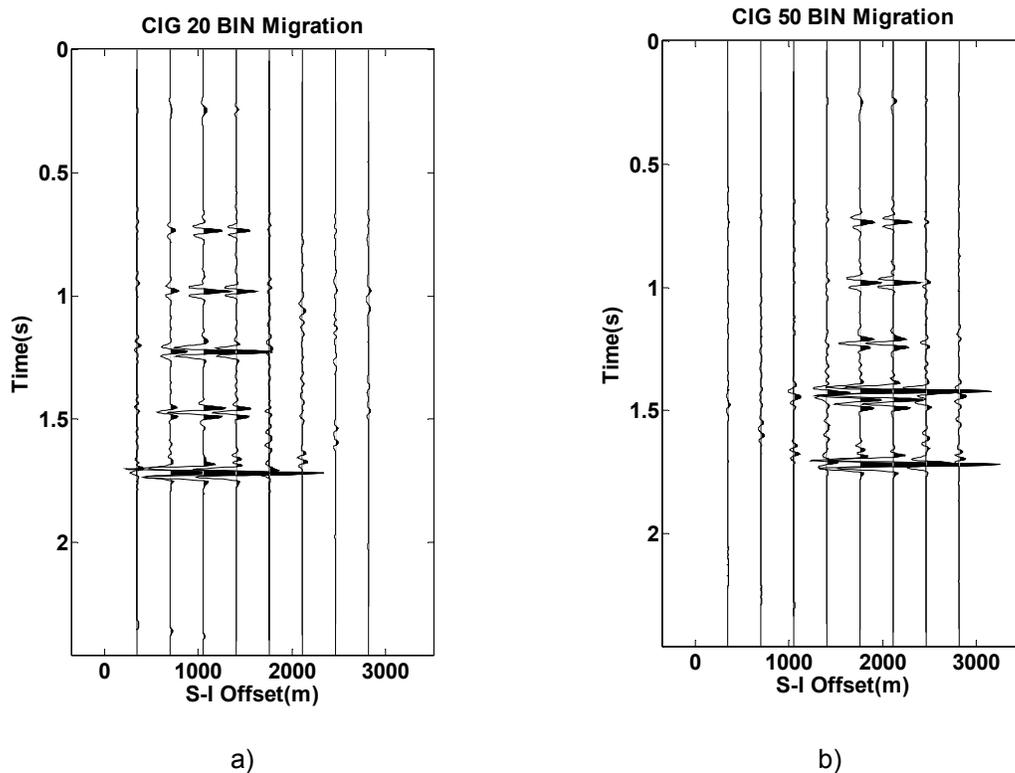


FIG. 40. Migration shot domain CIGs. a) at position $x = 900 m$, b) at position $x = 2250 m$.

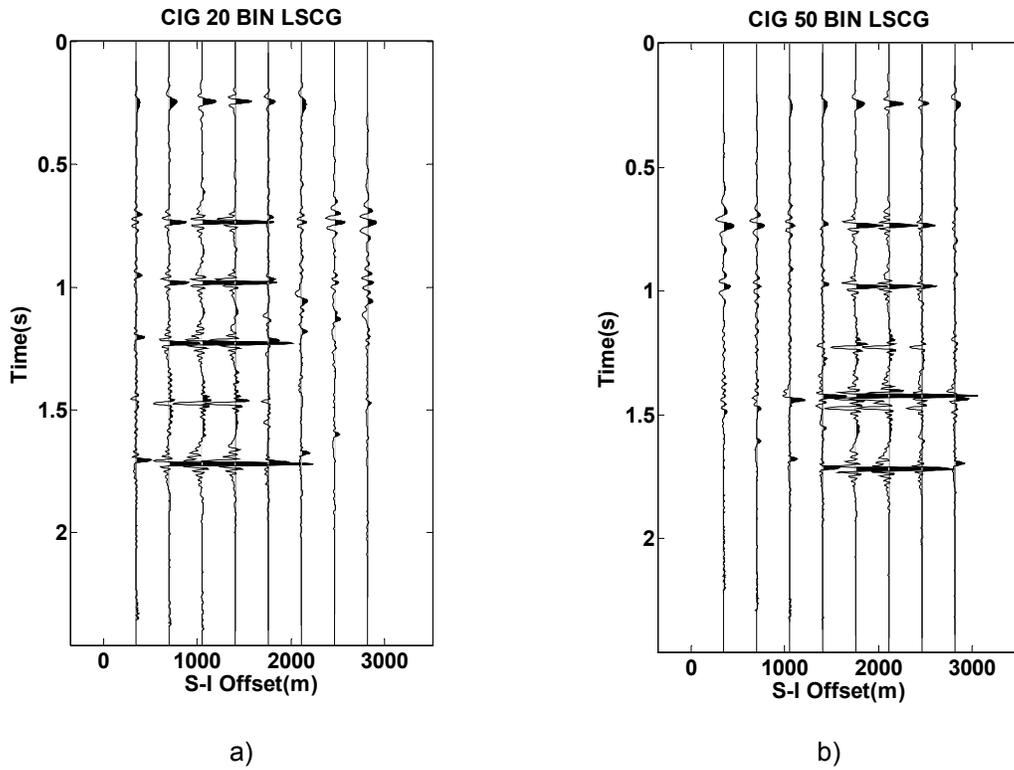


FIG. 41. LSPSM shot domain CIGs. a) at position $x = 900$ m, b) at position $x = 2250$ m.

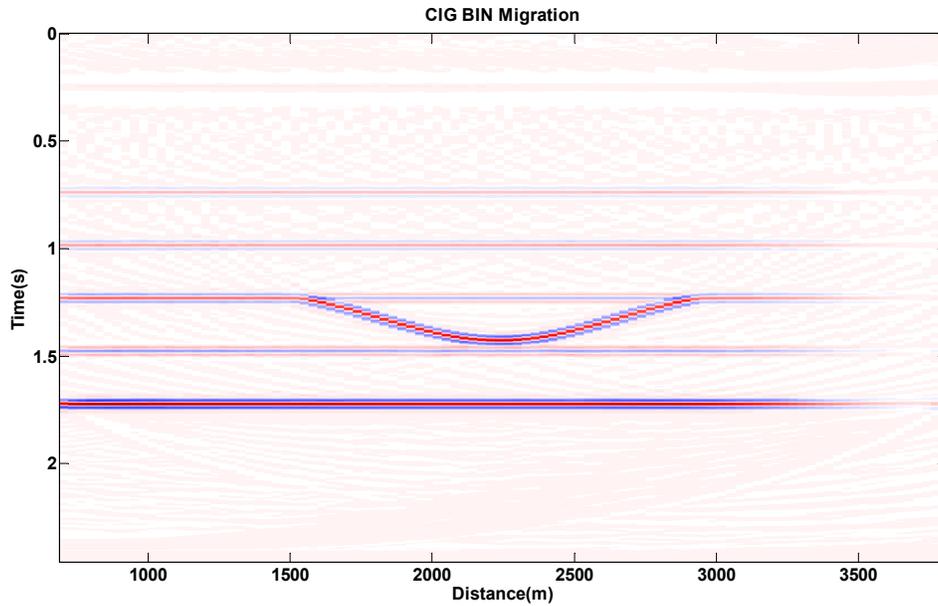


FIG. 42. Migration of shot domain binned synthetic data.

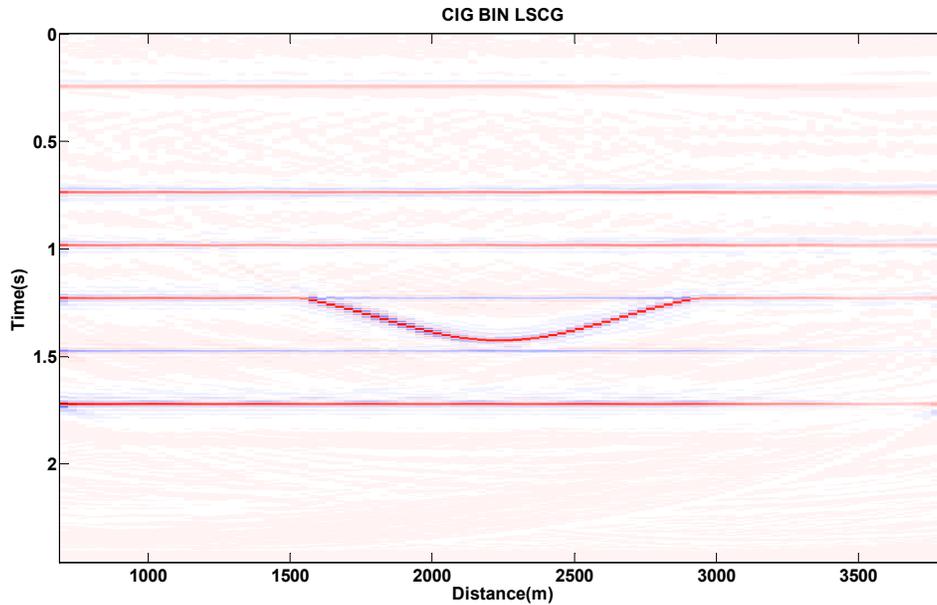


FIG. 43. LSPSM of shot domain binned synthetic data.

Now, using an incorrect velocity, migration and LSPSM is performed on the data. The resulted shot domain CIGs at the positions $x = 900m$ and $x = 2250m$ with a velocity that is %10 lower than the true velocity are shown in Figure 44. With this velocity, LSPSM shot domain CIGs at the same positions are shown in Figure 45.

With a velocity that is %10 higher than the true velocity, data migrated. Two CIGs in Figures 46 show the result at the positions $x = 900m$ and $x = 2250m$. Result of performing LSPSM with the velocity that is %10 more than the real velocity, for the same CIGs, is shown in Figure 47.

Comparing the CIGs show a significant difference between migration CIGs and LSPSM CIGs with incorrect velocities. LSPSM CIGs are noisier and the coherency between the traces in the CIGs is very low. The incoherency increases when data are highly decimated. Figure 48 shows the migration and LSPSM shot domain CIGs at the position $x = 2250m$ when %90 of data decimated regularly.

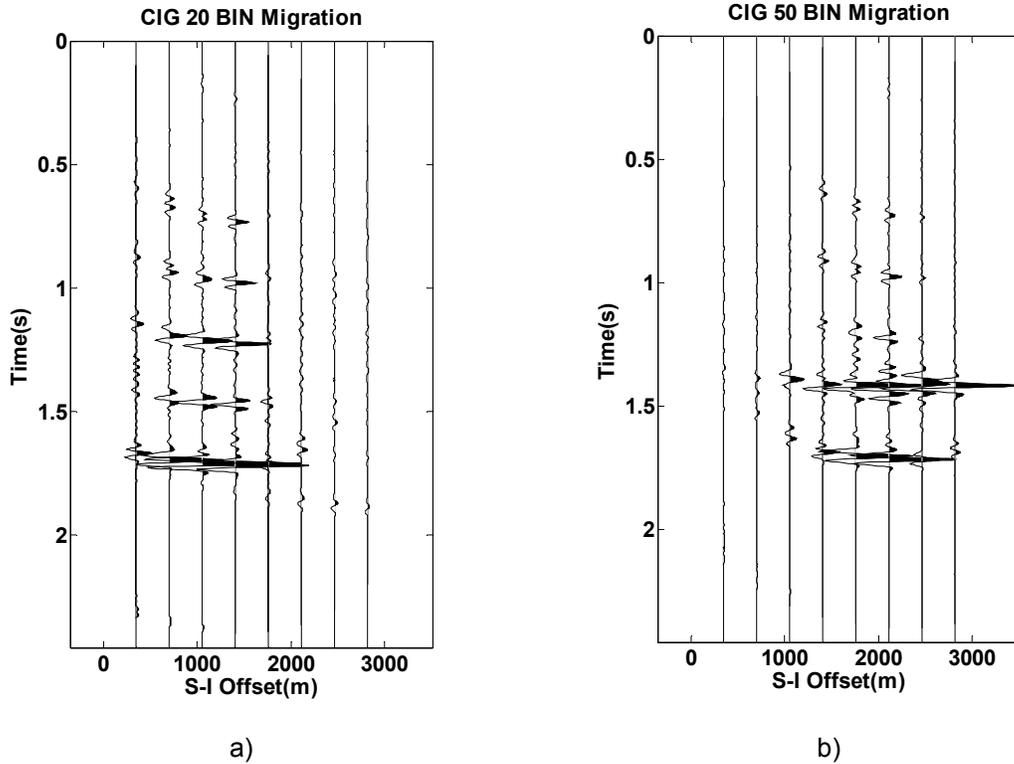


FIG. 44. Migration shot domain CIGs with a velocity %10 less than the true velocity. a) at position $x = 900\text{ m}$, b) at position $x = 2250\text{ m}$.

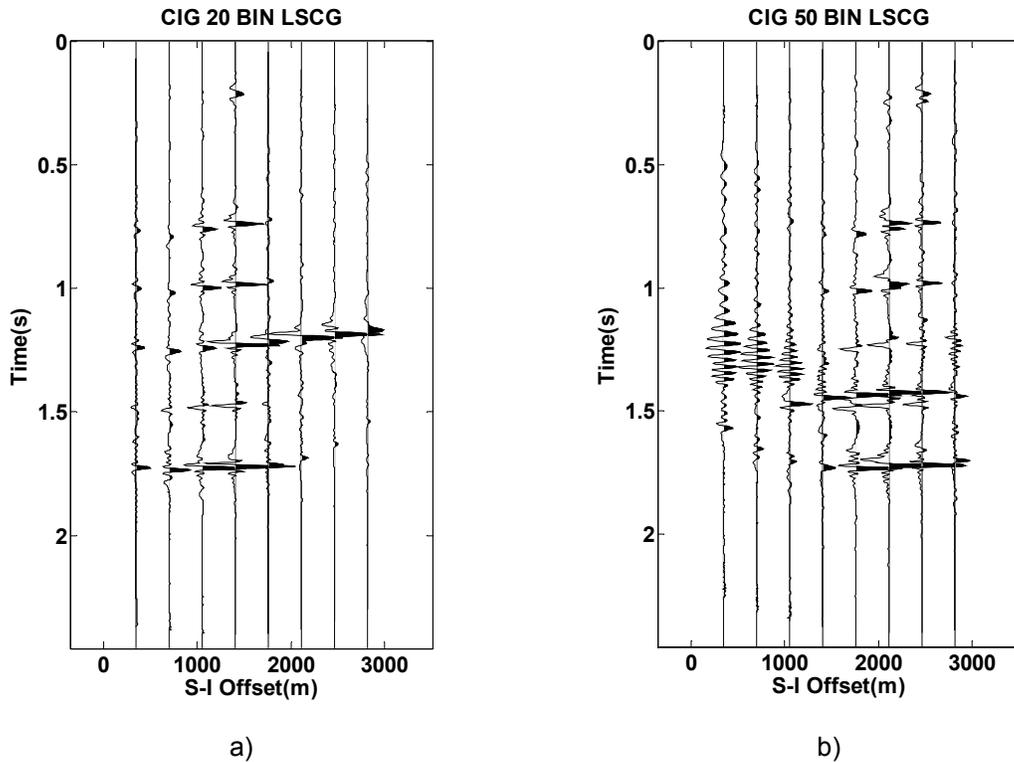


FIG. 45. LSPSM shot domain CIGs with a velocity %10 less than the true velocity. a) at position $x = 900\text{ m}$, b) at position $x = 2250\text{ m}$.

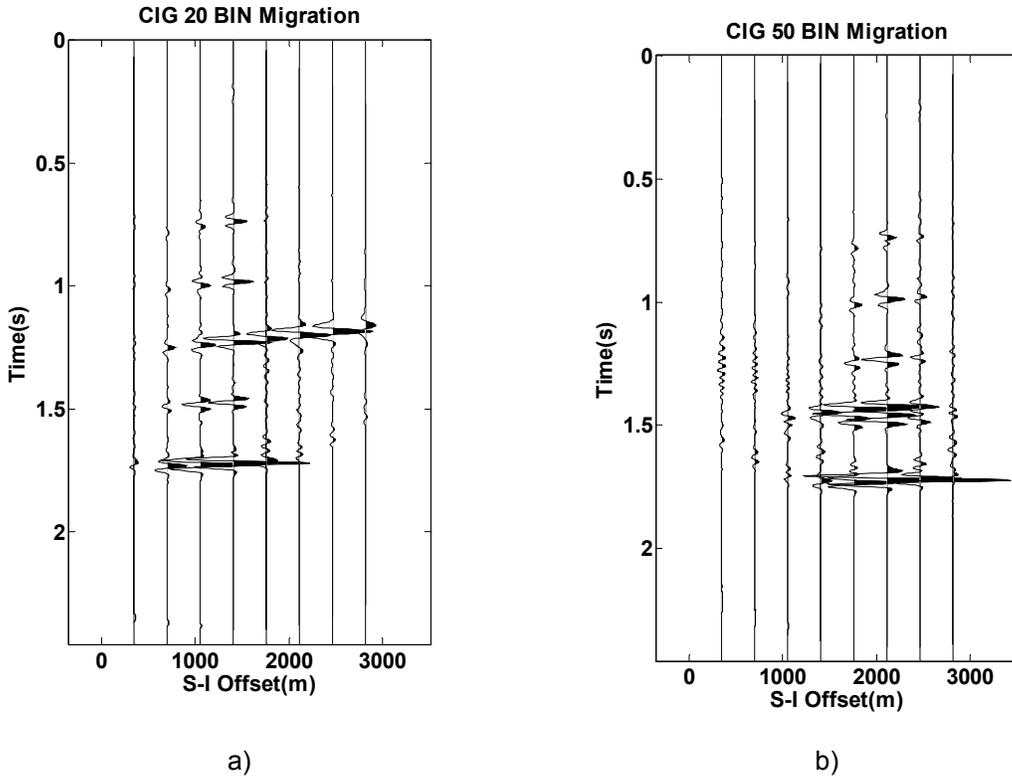


FIG. 46. Migration shot domain CIGs with a velocity %10 higher than the true velocity. a) at position $x = 900$ m, b) at position $x = 2250$ m.

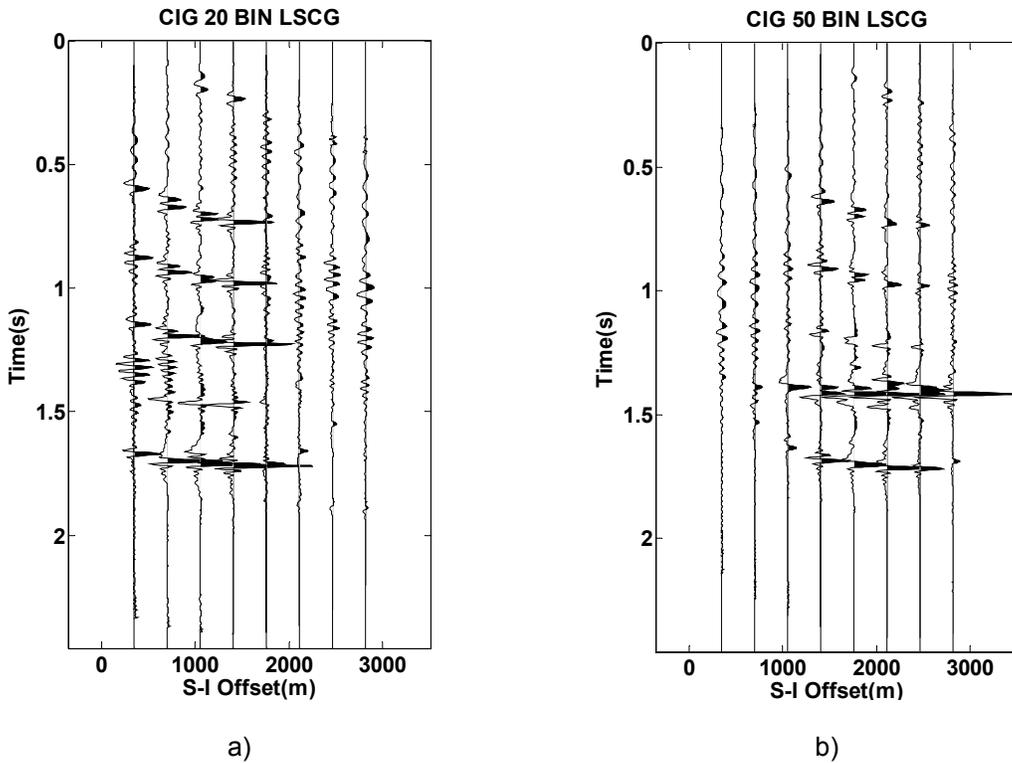


FIG. 47. LSPSM shot domain CIGs with a velocity %10 higher than the true velocity. a) at position $x = 900$ m, b) at position $x = 2250$ m.

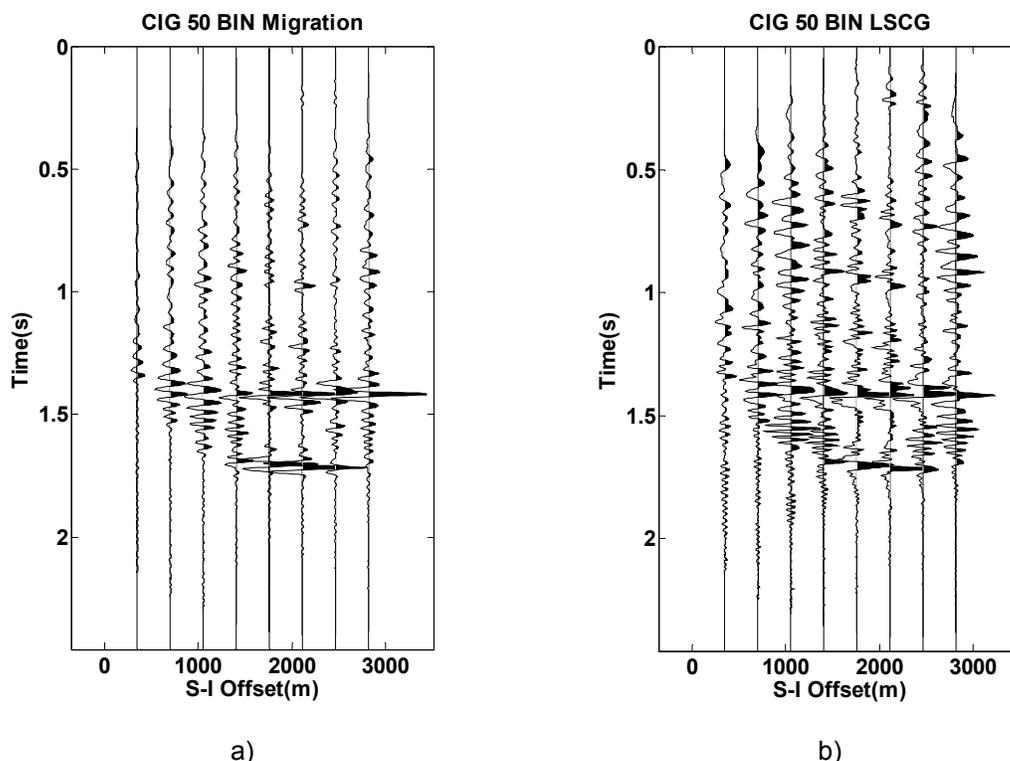


FIG. 48. Shot domain CIGs at the position $x = 2250\text{ m}$ with a velocity that is %10 lower than the true velocity and %90 of data decimated regularly. a) Migration, b) LSPSM.

The high sensitivity of the shot domain LSPSM CIGs to the velocity accuracy make it attractive for the velocity analysis. We perform velocity analysis on the shot domain CIGs in the same method that we used semblance analysis on the offset domain CIGs. Using constant velocities within a range that starts at 2000m/s and with the increment of 25m/s increases to 4500m/s, we performed constant velocity migration and LSPSM on the data and we achieved a semblance spectrum for each CIG.

Figure 49 shows the semblance spectrum on the migration shot domain CIG at the position $x = 2250\text{ m}$ when %50 of data regularly decimated. It is hard to choose the correct velocities for each time from this panel. This is due to strong incoherency of shot domain CIGs when the incorrect velocity is used. Fortunately, there are other methods of coherency measurements for velocity analysis. We found that the “unnormalized crosscorrelation sum” (Yilmaz, 2008), XC, is a useful quantity for measuring the coherency in the shot domain migration and LSPSM CIGs. XC is defined by,

$$XC = \frac{1}{2} \sum_t \left\{ \left[\sum_{i=1}^M f_{i,t(i)} \right]^2 - \sum_{i=1}^M f_{i,t(i)}^2 \right\}, \quad (18)$$

where $f_{i,t(i)}$ is the amplitude of i th trace at time $t(i)$ and M is the number of traces that contributes to the measurements in a defined gate (Yilmaz, 2008). Yilmaz (2008) interpreted XC is “half the difference between the output energy of the stack and the input energy”.

Figure 50 shows the XC spectrum for the migration shot domain CIG at the position $x = 2250m$ when %50 of data are regularly decimated. Comparison between figures 49 and 50 shows the effectiveness of using XC instead of semblance for velocity analysis on the shot domain CIGs.

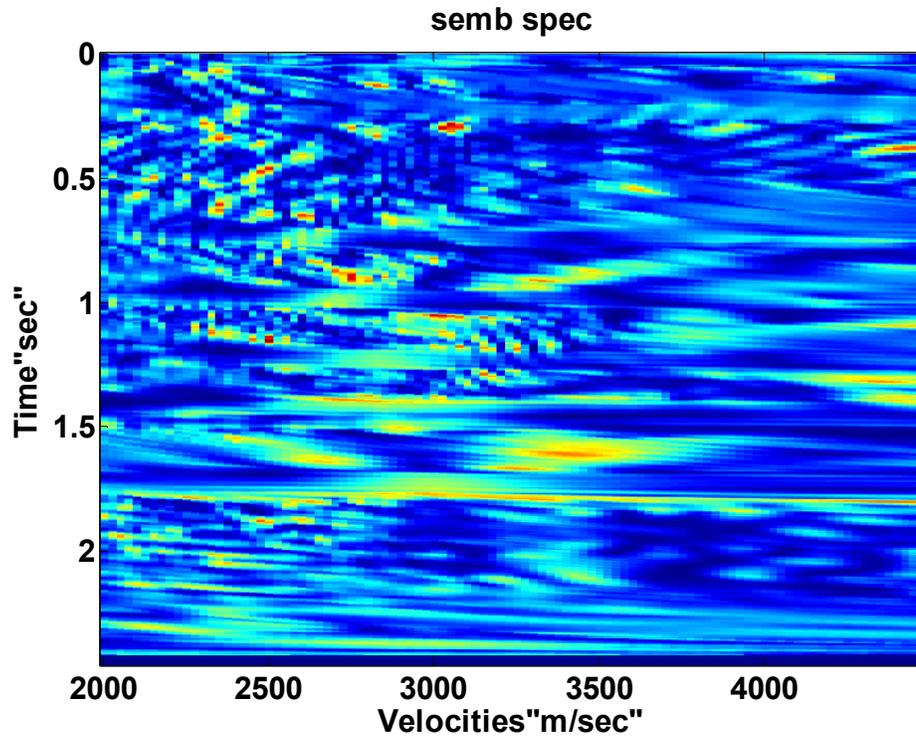


FIG. 49. Semblance spectrum for the migration shot domain CIG at $x = 2250 m$ with %50 of data.

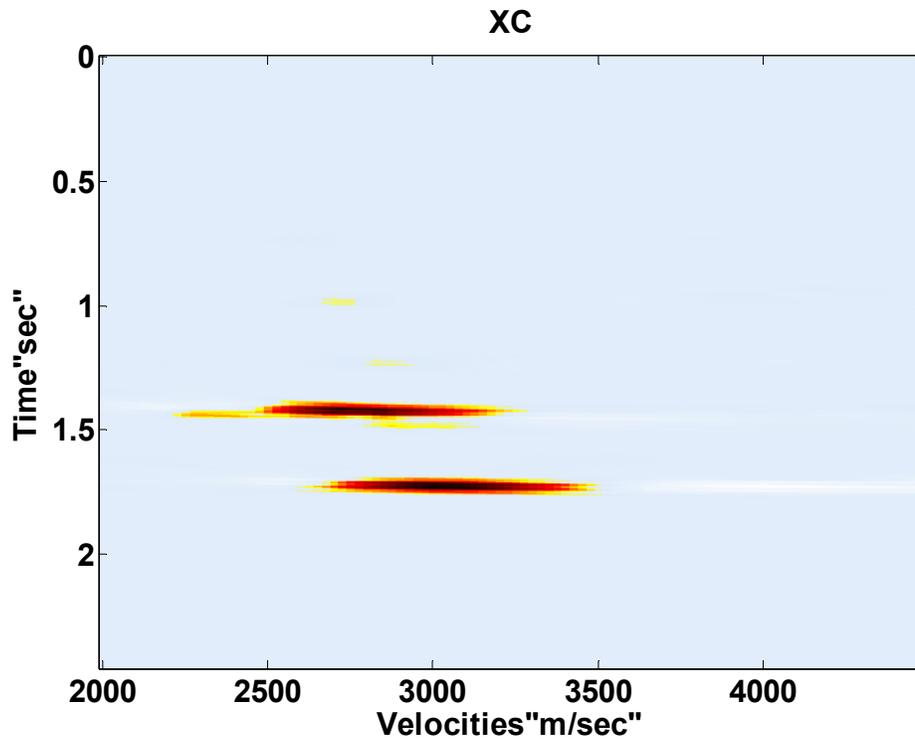


FIG. 50. XC spectrum for the migration shot domain CIG at $x = 2250\text{ m}$ with %50 of data.

Figures 51, and 52 show the XC spectrums for the migration and LSPSM shot domain CIGs, respectively, at the position $x = 900\text{ m}$ when only %10 of data is used. Figures 53 and 53 shows the same spectrum at the position $x = 2250\text{ m}$. The improvement in the XC resolution for velocity analysis by using LSPSM shot domain CIGs instead of migration shot domain CIGs is noticeable.

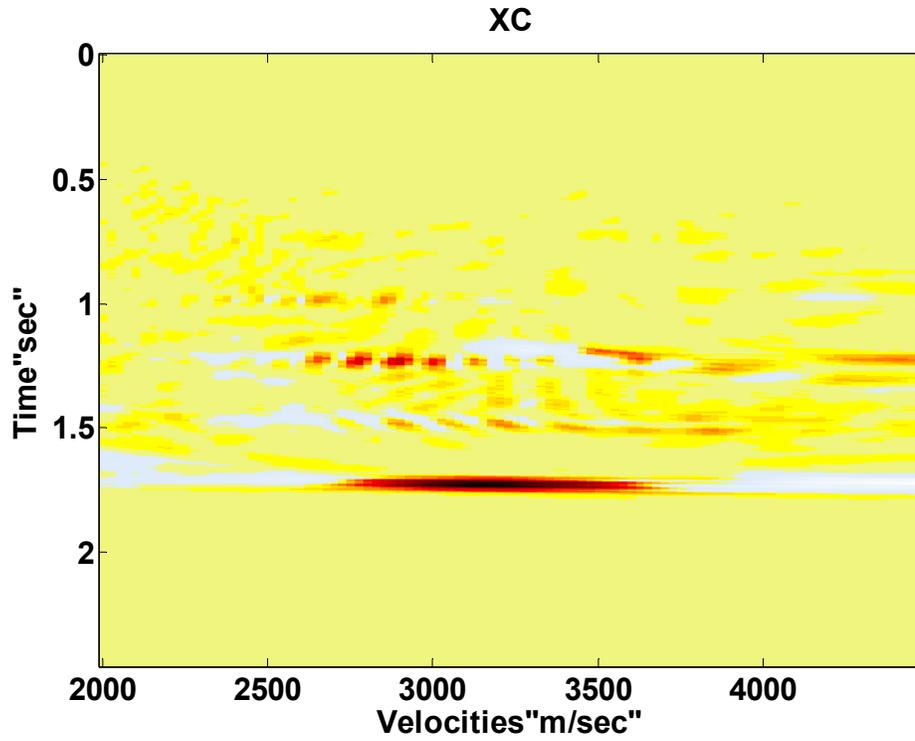


FIG. 51. XC spectrum for the migration shot domain CIG at $x = 900\text{ m}$ with %10 of data.

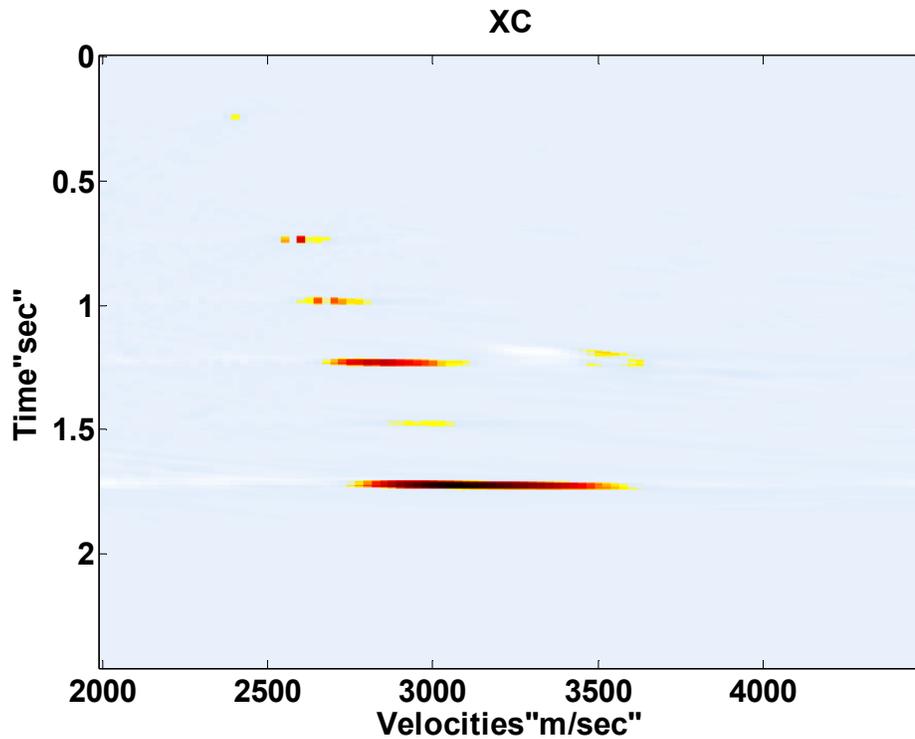


FIG. 52. XC spectrum for the LSPSM shot domain CIG at $x = 900\text{ m}$ with %10 of data.

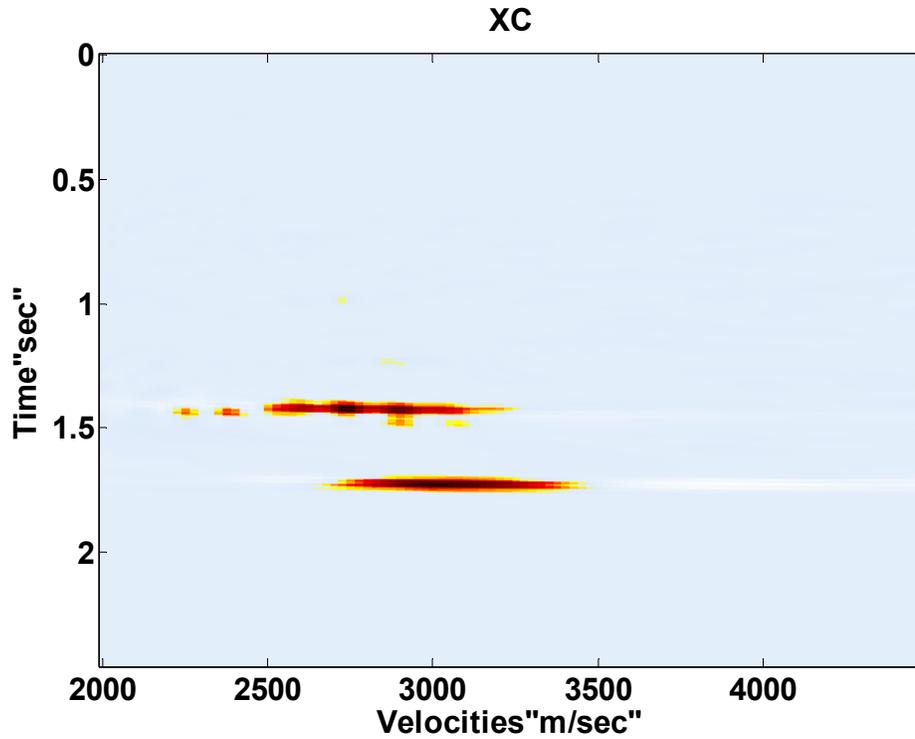


FIG. 53. XC spectrum for the migration shot domain CIG at $x = 2250$ m with %10 of data.

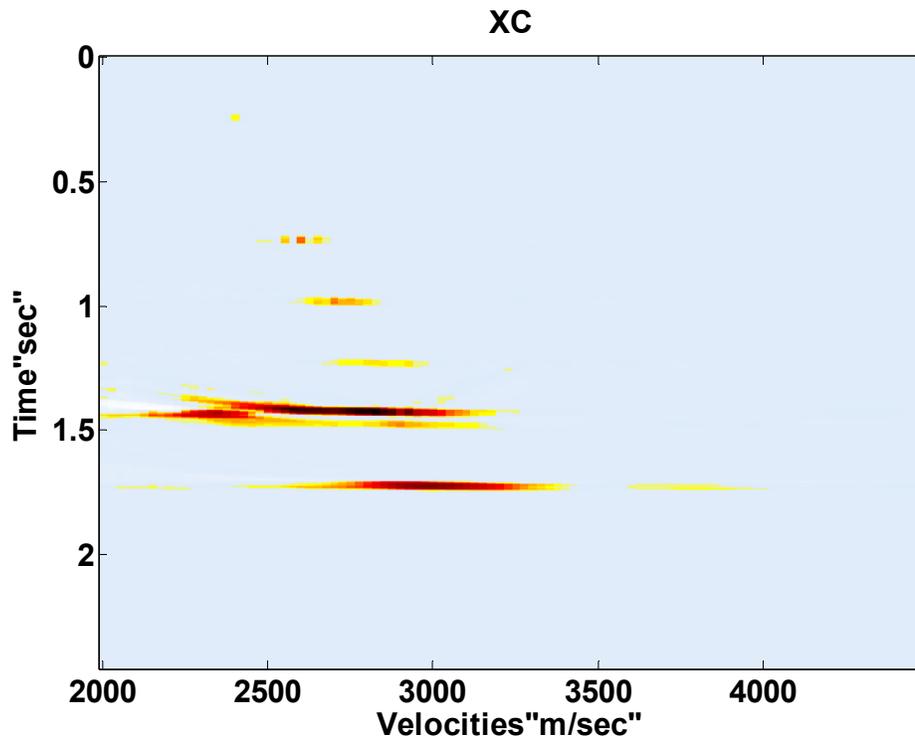


FIG. 54. XC spectrum for the LSPSM shot domain CIG at $x = 2250$ m with %10 of data.

True rms velocity function and the rms velocity function achieved by XC analysis on the LSPSM shot domain CIGs are shown in Figures 55 and 56, respectively. The extracted velocity is highly agreed with the true velocity. However, it is not able to reveal the low velocity channel at the middle of the model.

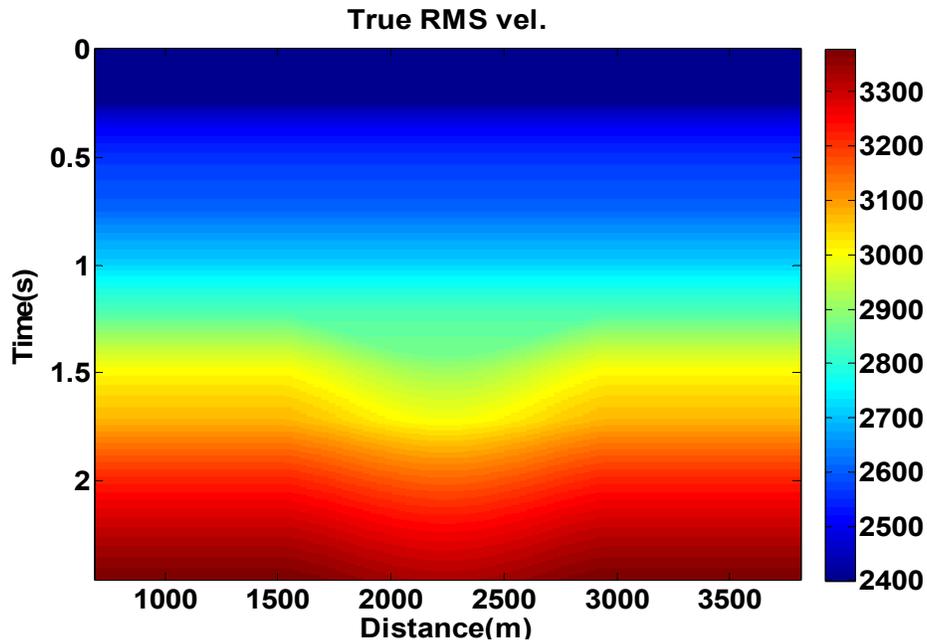


FIG. 55. RMS velocity model used for generation of synthetic data.

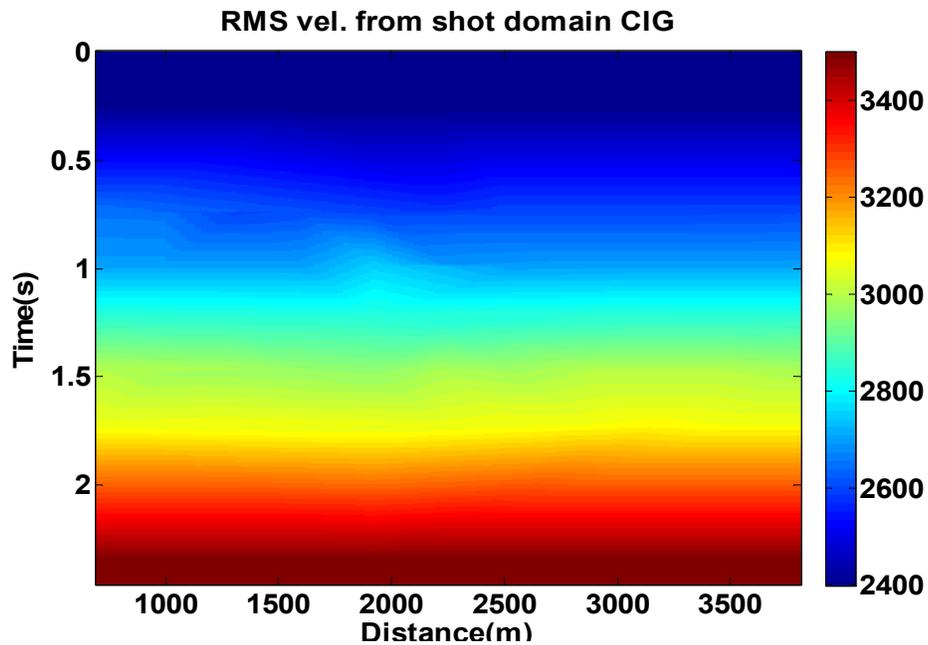


FIG. 56. RMS velocity model extracted using XC method on the LSPSM shot domain CIGs.

Since each iteration in the LSPSTM costs twice as the migration time, these methods seem to be a very costly procedure. However, doing 10 iterations on LSPSTM with %10

of data is only twice more expensive than the velocity analysis on migration CIGs using all data.

CONCLUSION

Kirchhoff is an efficient migration method for seismic imaging. However, Kirchhoff LSPSM with using proper velocity information gives a higher resolution image than the standard Kirchhoff migration and produces less acquisition footprints.

With dipping reflectors, semblance velocity analysis in the data domain fails to give the correct rms velocity. Semblance method on the offset domain CIGs or unnormalized crosscorrelation on the shot domain CIGs from migration can be used instead.

We showed that doing velocity analysis with semblance method on the migration offset domain CIGs, which is a robust tool for migration velocity analysis, can be extended to the LSPSM offset domain CIGs.

This gives better result when the data are irregularly or incompletely sampled. We showed that the velocity extracted by this method is accurate enough to give a high resolution image in LSPSM and works well for the data reconstruction. The same result is achieved by using an unnormalized crosscorrelation on the shot domain CIGs.

Replacement of migration CIGs with LSPSM CIGs for velocity analysis is a costly procedure. However, it gives better results with incomplete or irregular data.

ACKNOWLEDGMENTS

Authors wish to acknowledge the sponsors of CREWES project for their continuing support. They express their appreciation to the Dr. Helen Isaac and Dr. Hugh D. Geiger for their help and guidance. They also appreciate Kevin Hall and Rolf Maier in the CREWES project.

REFERENCES

- Baysal, E., Kosloff, D. D., and Sherwood, J. W. C., 1983, Reverse time migration: *Geophysics*, **48**, 1514-1524.
- Biondi, B. L., 2007, Concepts and applications in 3D seismic imaging: SEG publications.
- Claerbout, J. F., and Doherty, S. M., 1972, Downward continuation of moveout corrected seismograms: *Geophysics*, **37**, 741-768.
- Duquet, B., Marfurt, K. J., and Dellinger, J. A., 2000, Kirchhoff modelling, inversion for reflectivity, and subsurface illumination: *Geophysics*, **65**, 1195-1209.
- Gazdag, J., 1978, Wave equation migration with the phase shift method: *Geophysics*, **43**, 1342-1351.
- Gray, S. H., Etgen, J., Dellinger, J., and Whitmore, D., 2000, Seismic migration problems and solutions: *Geophysics*, **66**, No. 5, 1622-1640.
- Gulunay, N., 2003, Seismic trace interpolation in the Fourier transform domain: *Geophysics*, **68**, 355-369.
- Hagedoorn, J. G., 1954, A process of seismic reflection interpretation: *Geophysical Prospecting*, **2**, 85-127.
- Hestenes, M., and Steifel, E., 1952, Methods of conjugate gradients for solving linear systems: *Nat. Bur. Standards J. Res.*, **49**, 403-436.
- Hill, N. R., 1990, Gaussian beam migration: *Geophysics*, **55**, 1416-1428.
- Kuehl, H., and Sacchi, M. D., 2001, Split-step WKB least-squares migration/inversion of incomplete data: 5th Soc. Expl. Geophys. of Japan Internat. Symp. Imaging Technology.
- Nemeth, T., Wu, C., and Schuster, G. T., 1999, Least-squares migration of incomplete reflection data: *Geophysics*, **64**, 208-221.
- Porsani, M., 1999, Seismic trace interpolation using half-step prediction filters: *Geophysics*, **64**, 1461-1467.
- Sacchi, M. D., 2005, Inverse problems in exploration seismology: Short Course note, MITACS 2005, University of Alberta.
- Scales, J., 1987, Tomographic inversion via the conjugate gradient methods: *Geophysics*, **52**, 179-185.
- Schneider, W. A., 1978, Integral formulation for migration in two-dimensions and three-dimensions: *Geophysics*, **43**, 49-76.
- Sheriff, R. E., 2002, Encyclopedic dictionary of applied geophysics: SEG publications.
- Shewchuk, J. R., 1994, An Introduction to the Conjugate Gradient Method Without the Agonizing Pain: School of Computer Science, Carnegie Mellon University.
- Spitz, S., 1991, Seismic trace interpolation in the $f - x$ domain: *Geophysics*, **56**, 785-796.
- Stolt, R. H., 1978, Migration by Fourier transform: *Geophysics*, **43**, 23-48.
- Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: *Geophysics*, **49**, 1259-1266.
- Whitmore, N. D., 1983, Iterative depth migration by backward time propagation: 53rd Annual International Meeting, Society of Exploration Geophysicists, Expanded Abstract, 382-385.
- Yilmaz, Oz, 2008, Seismic data analysis: processing, inversion, and interpretation of seismic data: SEG publications.
- Yousefzadeh, A., 2008, Seismic Data Reconstruction with Regularized Least-Squares Prestack Kirchhoff Time Migration, M. Sc. thesis, Univ. of Alberta.