Long wavelength surface consistent solutions

John Millar and John C. Bancroft, CREWES

ABSTRACT

The surface consistent equations are a singular set of linear equations used to process land seismic data. They attempt to project components of a seismic signal to a particular source, or receiver coordinate. Midpoint and offset consistence may also be used.

Long wavelength errors present in a surface consistent statics solution turn up in the stack section as a smeared or even separated wavelet. Errors in surface consistent amplitudes will lead to problems with AVO. Deconvolution operators will also be affected. Conditions that will likely be problematic are areas with a sudden and persistent change in the near surface, such as shooting over lakes, or sand dunes.

In a typical application, the solutions are calculated using a series of Gauss-Seidel iterations, or using conjugate gradients. In this paper, we compare these results with a multigrid method. Our findings show the multigrid method is able to better resolve the long wavelengths with no significant increase in computer time.

INTRODUCTION

The surface consistent equations are a standard tool for any land seismic processing geophysicist. The underlying assumption behind the method is that strong effects on the seismic signal can be attributed to the near-surface conditions and coupling quality of a particular source or receiver. Long wavelength components of solutions to the surface consistent equations are notoriously difficult to solve for. They are considered a likely source of problems where near surface conditions change quickly and extend over large distances, such as near lakes, or sand dunes.

A large portion of the conventional commercial land processing job flow relies on the surface consistent equations (Taner and Koehler, 1981). Statics associated physically with a particular surface location improves our confidence that we are not arbitrarily changing the time structure of an event when applying static shifts. For AVO work, correcting for near surface absorption and other amplitude effects are both difficult and absolutely critical to the study. Surface consistent deconvolution is a robust and effective method to reject noise and whiten the spectrum of the data.

In this paper we discuss the importance of resolving the long wavelength solutions, and demonstrate that multigrid methods improve our ability to resolve the entire solution.

ITERATIVE METHODS

For many linear systems, direct solution may be impossible or take prohibitively long. An alternative approach is to apply an iterative correction to some initial estimate. There is a wide variety of methods available, more or less suited to different problems. A method is

said to *converge* to the correct solution if it is an appropriate way to solve the system.

Jacobi and Gauss-Seidel

One of the more simple methods to apply is the Jacobi method. It works by cycling through each unknown, and calculating the value of that unknown based on the other variables in the solution.

We wish to solve the linear system of equations,

$$\mathbf{Ax} = \mathbf{b}. \tag{1}$$

Expanding the Matrix multiplication for the n^{th} equation, out of N unknowns,

$$\begin{aligned} &A(n,1)x_i(1) + \cdots + A(n,n-1)x_i(n-1) \\ &\quad + A(n,n)x_{i+1}(n) + A(n,n+1)x_i(n+1) \\ &\quad + \cdots + A(n,N)x_i(N) = b(n). \end{aligned} \tag{2}$$

By rearranging equation 2, and solving for the $x(n)$ term, the i^{th} iteration of the Jacobi method is given,

$$\begin{aligned} x_{i+1}(n) = &\\ &b(n) - \frac{1}{A(n,n)}[A(n,1)x_i(1) + \cdots + A(n,n-1)x_i(n-1) \\ &+ A(n,n+1)x_i(n+1) + \cdots + A(n,N)x_i(N)]. \end{aligned} \tag{3}$$

Each estimated value is corrected by re-calculating its value using every other point in the trial solution.

The matrix $\mathbf{A}$ can be split into 3 separate matrices, one being a lower triangular, one upper triangular, and a diagonal matrix,

$$\mathbf{A} = \mathbf{L} + \mathbf{D} + \mathbf{U} \tag{4}$$

$$\begin{bmatrix} \nwarrow & & \mathbf{U} \\ & \mathbf{D} & \\ \mathbf{L} & & \searrow \end{bmatrix} = \begin{bmatrix} \nwarrow & & \mathbf{0} \\ & \mathbf{0} & \\ \mathbf{L} & & \searrow \end{bmatrix} + \begin{bmatrix} \nwarrow & & \mathbf{0} \\ & \mathbf{D} & \\ \mathbf{0} & & \searrow \end{bmatrix} + \begin{bmatrix} \nwarrow & & \mathbf{U} \\ & \mathbf{0} & \\ \mathbf{0} & & \searrow \end{bmatrix}.$$

Substituting 4 into 1, and rearranging,

$$\mathbf{Dx} = \mathbf{b} - (\mathbf{L} + \mathbf{U})\mathbf{x} \tag{5}$$

The Jacobi method can be written as a matrix operation,

$$\mathbf{x}_{i+1} = \mathbf{D}^{-1}[\mathbf{b} - (\mathbf{L} + \mathbf{U})\mathbf{x}_i]. \tag{6}$$

The smaller the $\mathbf{D^{-1}}(\mathbf{L} + \mathbf{U})$ term is, the less $\mathbf{x}_{i+1}$ depends on $\mathbf{x}_i$. This property is called diagonal dominance. When the diagonal component $\mathbf{D}$ is large compared to the sum of the off-diagonal terms, the solution depends less on the initial estimate (or the previous iteration), and convergence is faster. Each diagonal term must have an absolute value greater than the sum of all of the rest of the row for Jacobi to converge at all.

The Gauss-Seidel method is strongly related to the Jacobi method, only instead it takes advantage of each newly corrected unknown as it becomes available. Analogous to equation 3, the Gauss-Seidel method is

$$\begin{aligned} x_{i+1}(n) = & \\ & b(n) - \frac{1}{A(n,n)}[A(n,1)\underline{x_{i+1}}(1) + \cdots + A(n,n-1)\underline{x_{i+1}}(n-1) \\ & + A(n,n+1)x_i(n+1) + \cdots + A(n,N)x_i(N)]. \end{aligned} \tag{7}$$

The only difference between Gauss-Seidel (equation 7) and Jacobi (equation 3), is in the underlined subscripts, or iteration number.

The theoretical effect of each Gauss-Seidel iteration is found using the singular value decomposition of $[\mathbf{D} + \mathbf{L}]^{-\mathbf{1}}\mathbf{U}$. The *spectral radius* (Shewchuk, 2002) is the factor each wavelength of error is reduced with each pass. The singular values of $\mathbf{D^{-1}}[\mathbf{L} + \mathbf{U}]$ are used for Jacobi iterations.

Gauss-Seidel improves the speed of convergence over Jacobi by updating the unknown value, and using that value for each subsequent unknown. Convergence can be further improved by varying the pattern used, such as alternating between odd and even rows (Press et al., 1992).

Laplace equation example

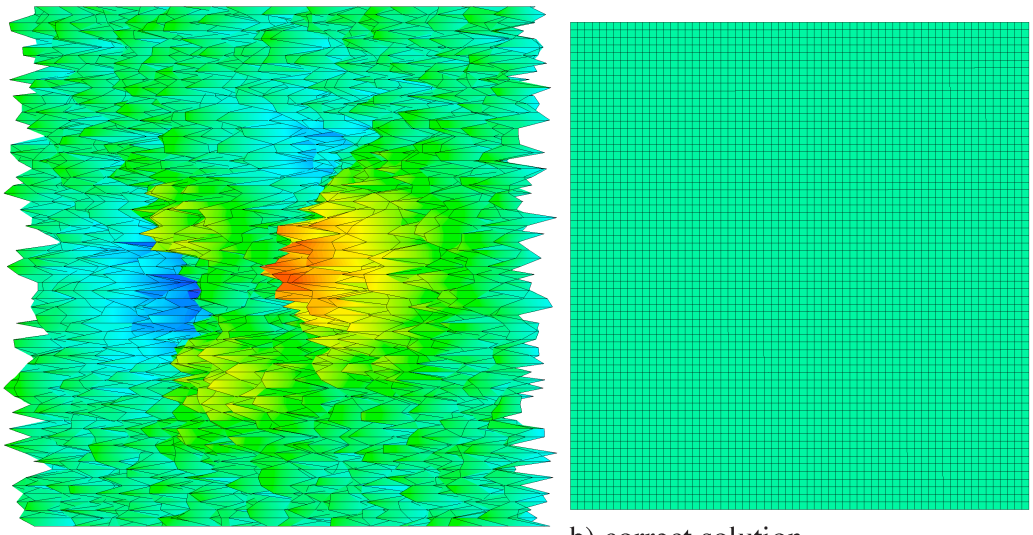
To show the effect of Gauss-Seidel *relaxation* we construct an example using the Laplace equation.

Laplace's equation takes the form

$$\nabla^2 p = 0. \tag{8}$$

This differential equation can be solved using finite differences, and expressing the problem as a linear system.

$$\mathbf{Lp} = \mathbf{0} \tag{9}$$



a) initial estimate b) correct solution

FIG. 1. The initial estimate and solution to the Laplace example. The Solution is converging to 0 throughout the domain.

The solution to this problem is zero everywhere. To test the ability of Gauss-Seidel relaxation, white noise is used as the initial estimate, and the solution is tracked after each iteration.

The initial estimate and the final solution we are aiming at are plotted in Figure 1. The solution after various numbers of iterations is plotted in Figure 3.

In Figure 3, the solution is converging slowly. If left for long enough the correct solution will eventually be reached within an error tolerance; however this is not very effective. Very quickly the high frequency spikes of the initial estimate are removed, but a smooth long wavelength error remains even after many iterations.

Figure 2 shows the spectral radius for the Gauss-Seidel operator, as compared with the actual convergence of the first few iterations. Each iteration of the method multiplies the frequency spectrum of the error by the function depicted. The long wavelengths near the left hand side of the plot are multiplied by a number very near, but just less than 1 and are virtually untouched by the Gauss-Seidel method. The high frequency error is attenuated strongly, by a factor of ≈ 0.3 with each pass. When the same operator is used on a different *grid* the shape of the spectral radius is left unchanged, only stretched or compressed along the wavenumber, k axis.

Multigrid

Multigrid methods generate solutions, by iterativly solving the equations, sampled on different size grids. By passing both the matrix operator and the source term ($\mathbf{A}$ and $\mathbf{b}$ from equation 1) through a series of low pass filters, the sample rate of the entire system

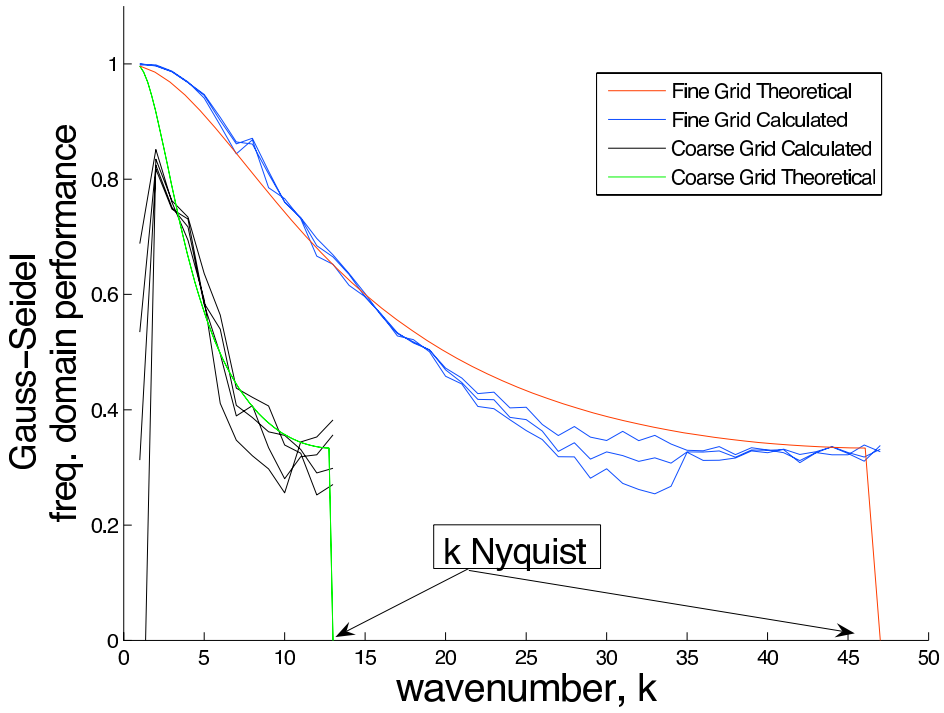


FIG. 2. The spectral efficiency of the Gauss-Seidel operator on different grid sizes for the Laplace equation.

can be lowered. This down sampling is called *rejection*, and the anti-alias filter coupled with rejection is called *restriction*.

Solving this reduced system iteratively takes far less computation time. As the number of unknowns decreases, so does the amount of work required to complete an iteration of a relaxation method.

The spectral radius of the Gauss-Seidel method attacks the long wavelength error far more effectively on a coarser grid (Figure 2). Once the coarse grid solution has converged, it is interpolated to the finer grid. This provides an initial estimate of that fine solution whose long wavelengths have already been calculated. This is where multigrid methods derive their strength. Interpolation tends to only introduce high frequency errors, quickly fixed by the relaxation method.

There are many ways to incorporate the use of multiple grids or meshes to solve a system of equations. A basic process is illustrated in Figure 4. A 9x9 system is reduced to a 5x5, then a 3x3 system. A solution is found by iteratively or directly solving the 3x3 system. The 3 variable solution is interpolated to 5 variables, and used as the initial estimate for an iteration method on this grid. Once the 5 unknowns have converged, we interpolate the solution to 9 variables, to be refined by a relaxation method on the 9x9 system. This process of interpolation and correction is repeated, until the desired accuracy is achieved.

Figure 5 shows the strength of using multiple grids to solve the sample problem with

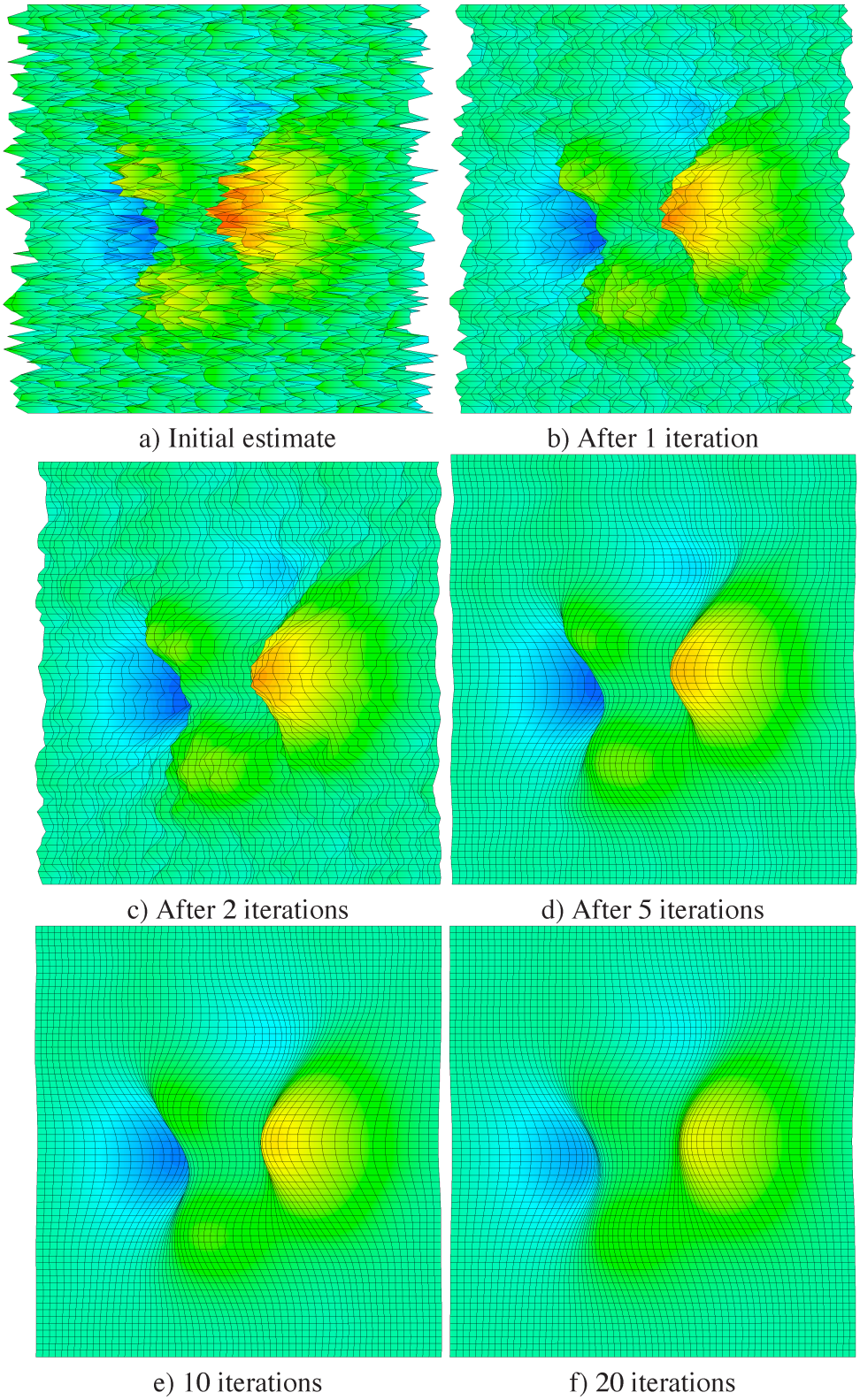


FIG. 3. The Gauss-Seidel solution to the Laplaces example. The Solution is converging to 0 throughout the domain, but the long wavelength errors are persistent.

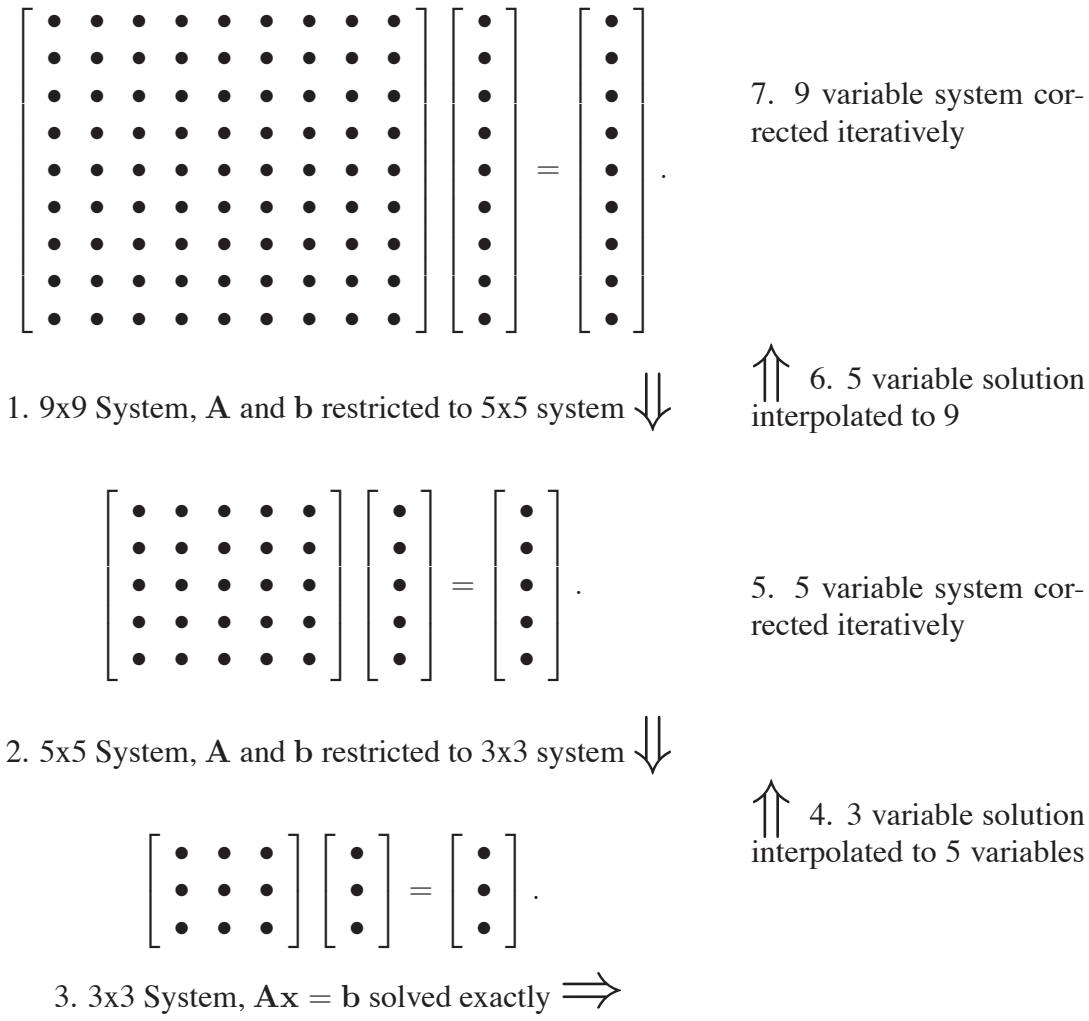


FIG. 4. A basic multigrid algorithm. A 9x9 system of equations is reduced to a 3x3 system using anti-alias filtering. This 3x3 system is solved exactly, and its solution is interpolated to a finer grid. The solution is repeatedly interpolated and corrected until the origional sample rate is acheived.

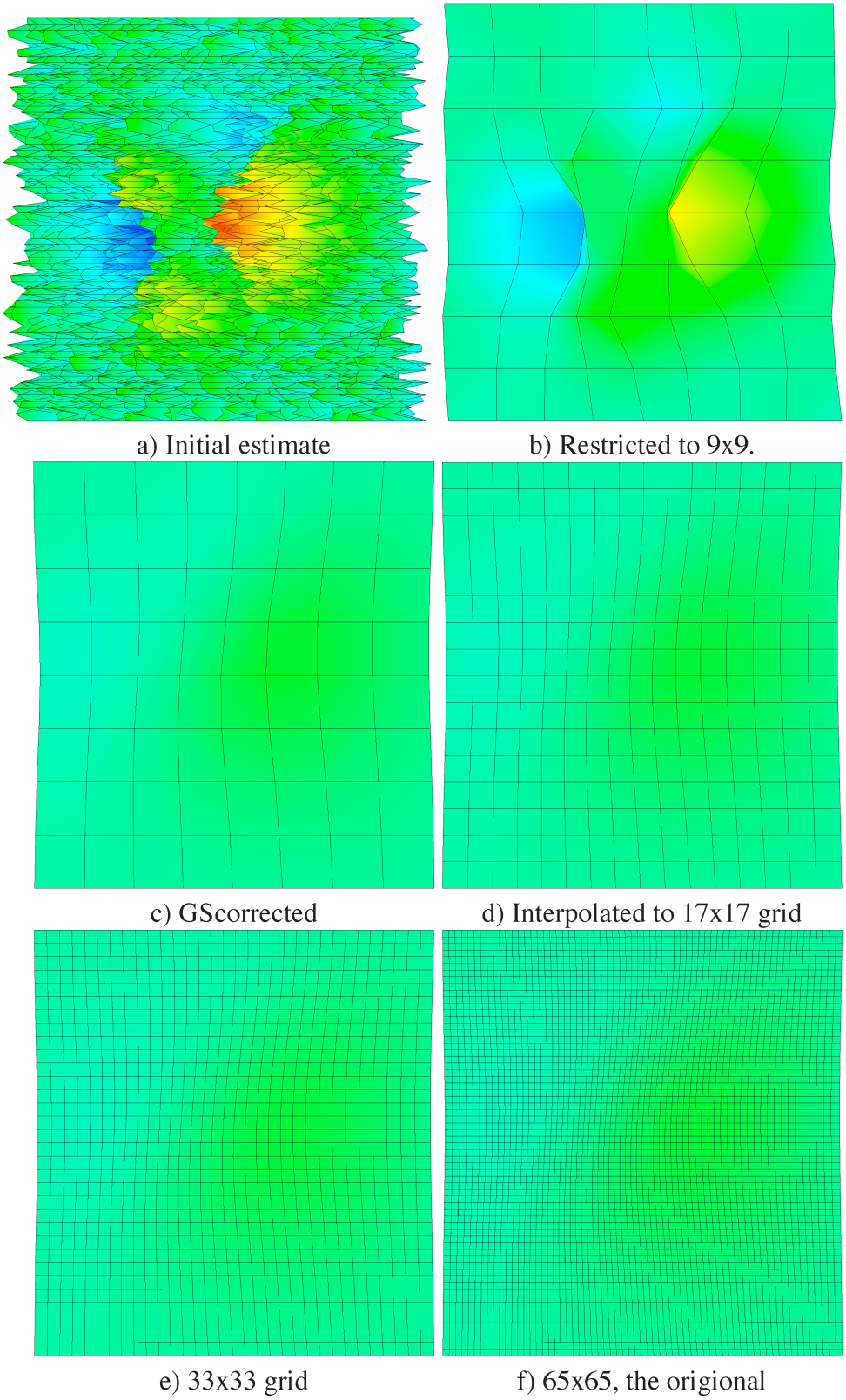


FIG. 5. The coarse grid corrected solution to the Laplace example.

Laplace equation. After only 2 corrections at each grid interval, the solution is far closer to zero than after many iterations using the Gauss-Seidel method. This coarse grid correction takes approximately the same calculation time as 2 Gauss-Seidel corrections.

THE EQUATIONS

We consider the problem of de-coupling separate source and receiver consistent components of a signal from a seismic trace. This could be an amplitude, a static as calculated by an auto-correlation with a model stack, or any number of features or attribute.

We assume a correction for each trace t_{ij} can be expressed as a contribution from the (both unknown) i^{th} source, S_i, and the j^{th} receiver,

$$t_{ij} = S_i + R_j. \tag{10}$$

In the case of deconvolution and amplitudes, the effect is a product,

$$t_{ij} = S_i \times R_j. \tag{11}$$

So we have to use the log of the spectrum to transform the multiplication into an addition.

$$\log t_{ij} = \log S_i + \log R_j. \tag{12}$$

Each trace contributes an equation to the linear system, which we express as a matrix operation of the form

$$\mathbf{As} = \mathbf{t}. \tag{13}$$

Here, $\mathbf{A}$ is a matrix of coefficients, $\mathbf{t}$ is a vector with all of the calculated *trace values* (static shift, amplitude etc.), and $\mathbf{s}$ is an unknown vector of the separated source and receiver consistent component of $\mathbf{t}$.

The form of $\mathbf{A}$ is a sparse rectangular matrix whose coefficients are determined by the geometry of the seismic survey. We assign a column of $\mathbf{A}$ to each unique shot, and one to each unique receiver. It has as many rows as traces, which is usually much greater than the number of columns. Its form is demonstrated best by partitioning it,

$$\mathbf{As} = [\mathbf{A}_s|\mathbf{A}_r][\frac{\mathbf{s}_s}{\mathbf{s}_r}]. \tag{14}$$

The unknowns $\mathbf{s}_s$ and $\mathbf{s}_r$ are vectors of source and receiver unknown values corresponding to equation 10. The n^{th} trace has a value of $\mathbf{t}_n$ associated with it. The n^{th} row of $\mathbf{A}_s$ is empty except for 1 in the column corresponding to the shot, and $\mathbf{A}_r$ contains only a 1 in the appropriate receiver column.

The system of equations that results from this problem is over-determined (more equations than unknowns), requiring a least squares solution,

$$\mathbf{A}^T\mathbf{A} = \mathbf{A}^T\mathbf{b}. \tag{15}$$

The classification of $\mathbf{A}^T\mathbf{A}$ would be that of a symmetric, positive indefinite matrix. It is not strictly diagonally dominant, as the sum of the off diagonals is equal to the diagonal term in each row. Trottenberg et al. (2001) implies that a necessary condition for using multigrid methods is that the matrix be an "M" matrix. One of the conditions of "M", is that it be non-singular, which is not true in this case, $\mathbf{A}^T\mathbf{A}$ is not "M". We can't guarantee multigrid will converge (Wesseling, 1992) .

2 TERMS

The above derivation is for what we call a *2 term* surface consistent reduction. The 2 terms are the source location and receiver location. We show the results for this reduction in Figure 6. The model is a random static distributed linearly, with a maximum of 8ms. A 50ms step function in the receiver consistent term. This model attempts to simulate a shallow bog covering the latter half of the survey. All of the shot holes were drilled into competent rock, but the receivers over the bog would suffer a large static shift relative to the receivers on dry land.

The synthetic survey has 200 shots, with 60 live geophones each shot. The cable moves 4 stations with each shot. There are 796 receiver stations.

The long wavelength component of the solution is not well imaged by the Gauss-Seidel method. The error (lower 2 panels of Figure 6) for the shot and receiver consistent statics form an 'S' shaped curve in either direction. The multigrid and conjugate gradient methods both appear to have a linear drift.

Both multigrid and conjugate gradients have a constant or linear residual error. Intermediate wavelengths are resolved very well.

When the error of Figure 6 is included in a stack section, the result is in Figure 7. A 40hz ricker wavelet is convolved with a spike at a constant time for each trace. Each trace has a static shift applied equal to the error, and is sorted by its midpoint and stacked. Only 100 midpoints either side of the center of the survey are plotted.

It appears as though both the multigrid and the conjugate gradients did a good job of removing the surface consistent static. The Gauss-Seidel solution has caused the wavelet to stack poorly, and smear the energy over 20ms. This is an important observation, demonstrating that the error is a resolution problem.

It is suspected that the very long wavelength error remaining in the multigrid and conjugate gradient solutions cannot be reduced. The equations are rank deficient by 1. Without external constraints the solution may always be off by a constant, or near linear amount. In practice this is not of concern, the resulting difference in the stack is a bulk shift of all traces, which does not effect the quality of the data.

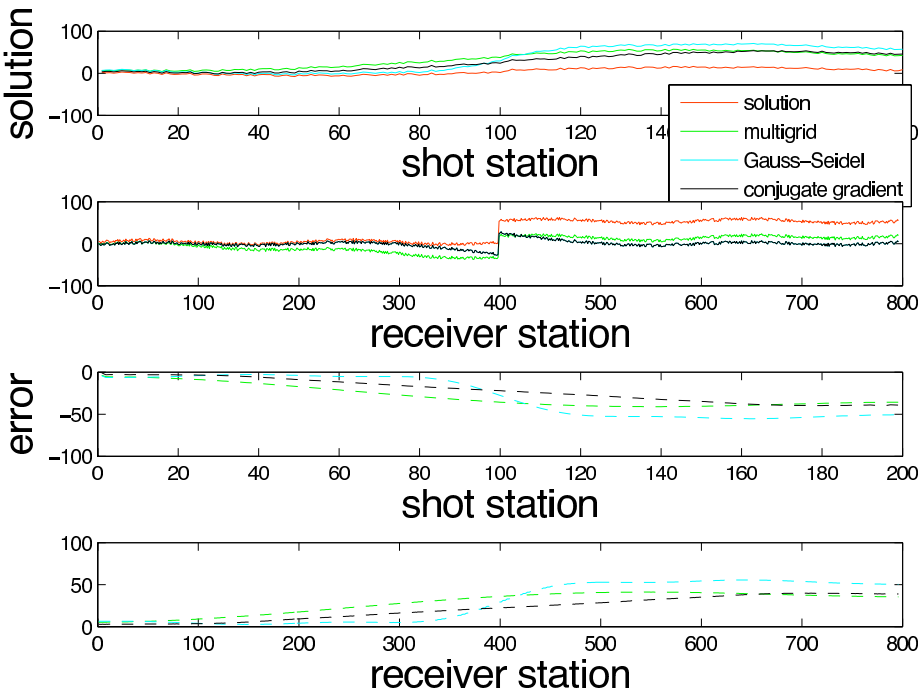


FIG. 6. The results from a 2 term surface consistent reduction. The top two axes show the calculated solutions for the shot and receiver components, the bottom axes show the error.

4 TERMS

When calculating surface consistent amplitudes, deconvolution operators, or when including residual NMO terms for statics, it is often good to use more terms. Most commonly, offset and midpoint bin components are calculated along with shot and receiver station. Ground roll is concentrated in the offset consistent component. Removing the offset consistent signal is a good noise reduction method. Forcing the midpoint consistent term to be smooth puts a realistic geologic constraint on the equations (Cary and Lorentz, 1993).

Introducing additional terms also introduces more singular values. The 4 term surface consistent equations for the same synthetic survey as Figure 6 are rank deficient by 13. Without a damping factor the solution becomes unstable, and grows exponentially. To stabilize the process, we introduce a damping factor, μ, and solve

$$[\mathbf{A}^T\mathbf{A} + \mu\mathbf{D}]\mathbf{x} = \mathbf{A}^T\mathbf{b}. \tag{16}$$

The $\mathbf{D}$ matrix is the main diagonal of $\mathbf{A}^T\mathbf{A}$, (see equation 4). With a small damping factor, the equation becomes diagonally dominant, so convergence to a minimum L-2 norm is provided. For this example $\mu = 0.01$ stabilizes the solution well.

Figure 8 shows the 4 terms of the solution, and Figure 9 shows the error. The conjugate gradient method does not seem to handle the new terms well. Multigrid performed better than other methods.

To show the effect the long wavelength error has on 4 term data, we assume the data comes from an amplitude correction. A surface consistent amplitude is calculated for each trace. The mid-point and offset consistent components are both identically 1, so there should be no amplitude anomalies in the corrected gathers.

In Figure 10 the wavelets depicted should all have the same amplitude. Each of the groups of lines correspond to traces that share the same midpoint. As we move across the surface consistent anomaly, the lines in each group diverge, indicating that the amplitude of that trace within the mid-point bin is not constant, as it should be. The lines corresponding to the multigrid gathers have the least spread, indicating that the multigrid solution provides the most accurate amplitude response.

Figure 11 shows some interesting statistics about the error in the amplitudes. The standard deviation of the error in each mid-point bin relates directly to the amount of possible error in the amplitude response. The multigrid method outperforms both Gauss-Seidel and conjugate gradient clearly in this test.

CONCLUSIONS

Multigrid methods have shown themselves to improve the resolution of long wavelengths in the surface consistent solutions. The quality of these solutions has a direct connection to the quality of the final seismic section. Two term solutions, when applied to statics, may reduce the number of iterations necessary for a processor to correct statics and velocities in the standard land processing flow. Resolving the solutions more accurately gives us more confidence that we are not arbitrarily assigning statics, and possibly harming structure. Improvement in the error relating to amplitudes may be reduced as much as 30-40 % when the 4 term reduction is applied using multigrid.

The multigrid method does not sacrifice any significant computer time compared to the Gauss-Seidel solutions. There is some overhead associated with multigrid approaches. A solution is not realized at all until after a calculation time close to that of 6 iterations of Gauss-Seidel. Once a solution is available, there is much less error compared with other methods.

To implement multigrid requires almost no change to the overall processing flow, making deployment for this purpose attractive.

The major test of the new method will be how it reacts to field seismic data. The results from the synthetic examples presented in this paper show good promise in our ability to improve seismic images using multigrid to calculate long wavelength surface consistent solutions.

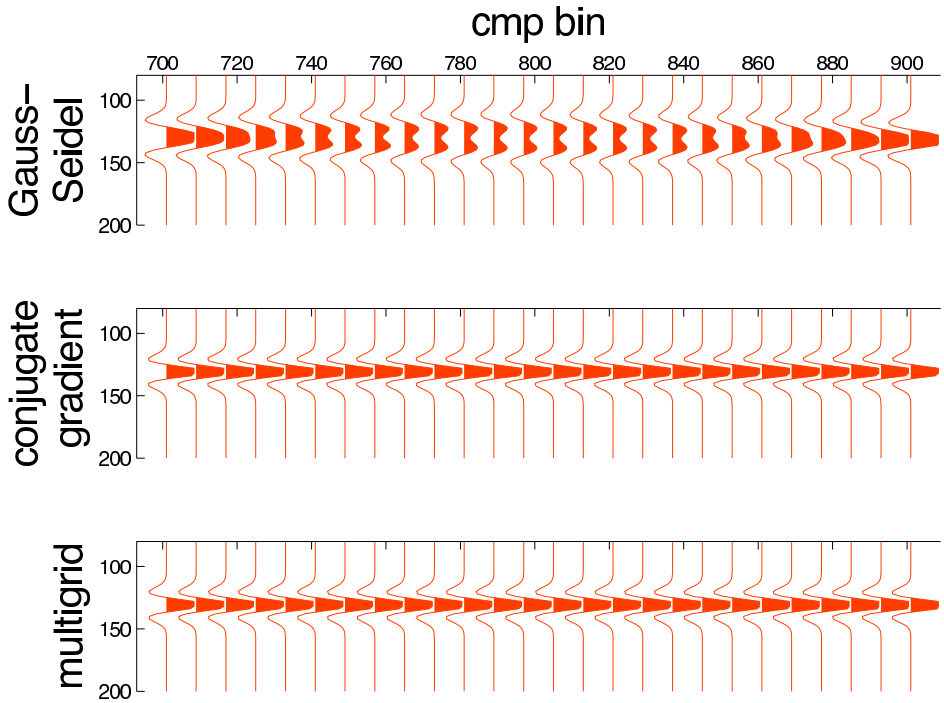


FIG. 7. The predicted stack section arising from using the various methods to solve the surface consistent equations for statics.

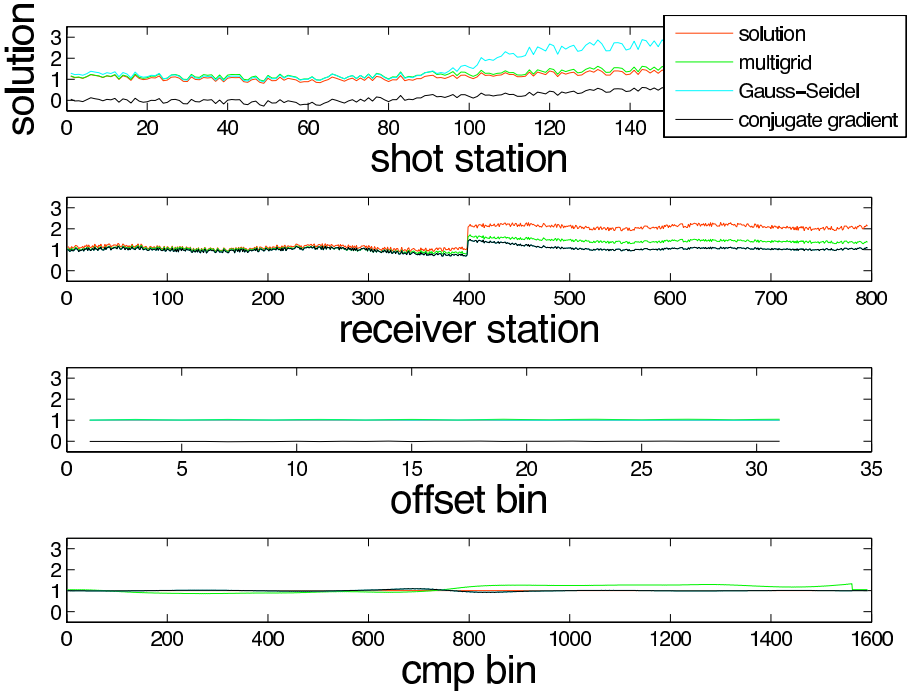


FIG. 8. The solutions for a 4 term reduction. Each term is given a separate axes.

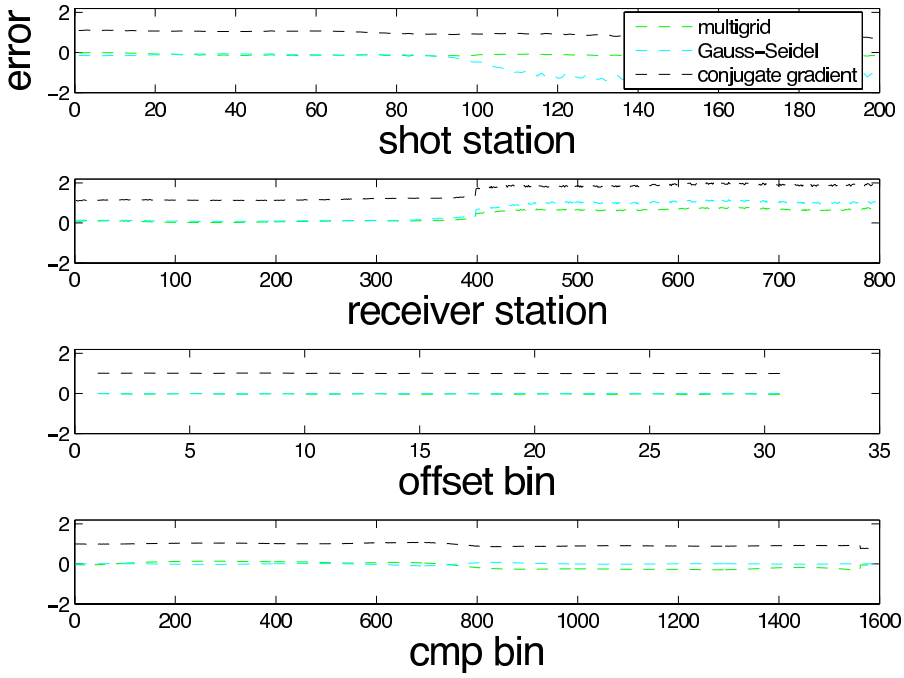


FIG. 9. The error for the 4 term reduction.

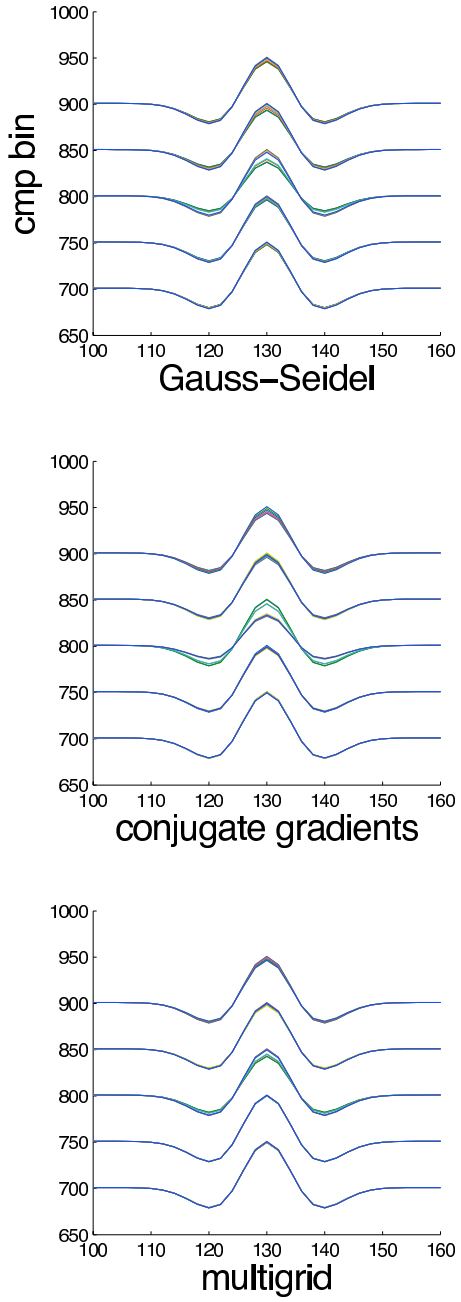


FIG. 10. A view of the pre-stack variability of the amplitudes generated using the 4 term reduction.

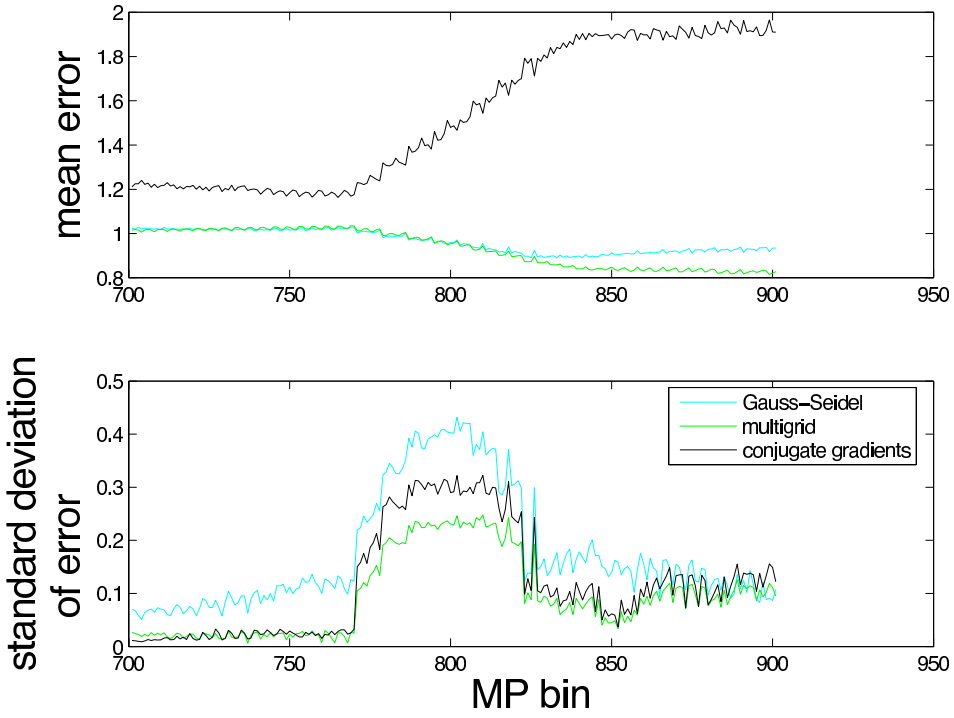


FIG. 11. Mean and standard deviation of the pre-stack variability of the amplitudes generated using the 4 term reduction. A wider standard deviation means a larger AVO error.

REFERENCES

Cary, P. W., and Lorentz, G. A., 1993, Four-component surface-consistent deconvolution: Geophysics, **58**, No. 3, 383–392.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., 1992, Numerical Recipes in C: Cambridge University Press, 2nd edn.

Shewchuk, J. R., 2002, An Introduction to the Conjugate Gradient Method Without the Agonizing Pain: unpublished.

Taner, M. T., and Koehler, F., 1981, Surface consistent equations: Geophysics, **46**, No. 1, 17–22.

Trottenberg, U., Oosterlee, C., and Schűller, A., 2001, Multigrid: Academic Press.

Wesseling, P., 1992, An Introduction to Multigrid Methods: John Wiley & Sons.