Stanford Exploration Project, Report 110, September 18, 2001, pages 1–91

Data dependent parameterization and covariance calculation for inversion of focusing operators

Barbara E. Cox[1]

ABSTRACT

The Common Focus Point (CFP) method makes it possible to convert conventional two way traveltimes to focusing operators that represent one way traveltimes. These focusing operators are inverted to obtain a velocity model. The under-determined nature of the inversion problem is addressed by the data dependent adjustment of the parameterization by means of the a posteriori covariance. This results in an efficient, non-laborious algorithm producing well determined inversion problems. The required a posteriori covariance is normally explicitly solved in the explicit matrix calculation during optimization. However, these explicit calculations are not feasible in larger problems. Fortunately, algorithms have been proposed to extract the a posteriori covariance from the more efficient approximate matrix inversion algorithms that are available.

INTRODUCTION

The quality of a seismic image is highly dependent on the accuracy of the velocity model of the subsurface. The CFP method (Berkhout, 1997a,b; Thorbecke, 1997) has proven to be an appropriate tool to estimate this velocity model, because the inversion of one way traveltime data (focusing operators) generated by this method is inherently simpler than the inversion of conventional two way traveltime data (Kabir and Verschuur, 2000; Hegge, 2000).

In this report, the inversion of focusing operators is also used to obtain a velocity model. As in all geophysical inversion problems, the under-determined nature is a problem that should be faced. Normally, this is handled by two methods; 1) choosing a well determined parameterization (global parameterization by user defined layers), or 2) regularizing the optimization (including a priori information and model covariance). A drawback of the first method is that it puts a constraint on the result and it can be laborious. Moreover, incorrect initial parameterization can lead to slow or non-convergence. A drawback of the second method is that the problem of over-parameterization is still not solved (e.g. when a regular grid is used). This might cause problems when the inversion problem is regularized; the model parameters need different levels of regularization, as certain regions in the model might consistently be more under-determined. In the data dependent parameterization shown in this report both methods will be combined; adjustment of the parameterization is based on the covariance after

[1]**email:** B.E.Cox@ctg.tudelft.nl

optimization. In the explicit least squares or singular value decomposition (SVD) optimization methods, the a posteriori covariance is generated during optimization. However, in more practical approximate optimization methods, obtaining the covariance becomes more difficult.

In this report, the concept of the CFP method and the focusing operator updating will be explained in the first section. Next, the inversion of the focusing operators, and the data dependent parameterization will be considered. The method will be applied to a synthetic example. The last part will compare the different approximate inversion algorithms for obtaining the inverse and the a posteriori covariance. Finally, some conclusions and plans for future work will be presented.

CFP METHOD

Focusing operators

For each point in the subsurface, seismic migration can be written in terms of two focusing steps: focusing in detection followed by focusing in emission. The focusing is done by a focusing operator, which can be considered as a one-way seismic response from one point in the subsurface to locations at the surface (Fig.1).

The first focusing step is performed by time-convolving a shot record with the focusing operator of the point under consideration. Next, all traces in the obtained move-out corrected shot record are summed (consider *shot 1* in Fig.1). When this first focusing step is applied to all shots, each shot record is transformed into a single trace by the focusing operator (consider *shot 1,2-n* in Fig.1). Together those traces define the so-called common focus point gather (CFP-gather). Each trace in the CFP gather is positioned at the source position of its corresponding shot record. One event in the CFP-gather is the focus point response. If the velocity model is correct, the focus point response and the time-reversed focusing operator (= Green's function) have equal traveltime: principle of equal traveltime (Fig.1).

Note that the first focusing step "transforms" conventional two way data to one way CFP data. The second focusing step is performed by applying the focusing operator again. This procedure transforms the focus point response into the seismic image of the subsurface grid point under consideration (CFP-stacking). In this report only the first step is considered, because this is the required step for obtaining focusing operators as will be shown in the next section.

Concept of focusing operator updating

The principle of equal traveltime formulates that the time-reversed focusing operator (based on the underlying velocity model) and the focus point response (based on the underlying seismic measurements) must have equal traveltime. In the situation of velocity errors both the operator and the response contain an error. The error in the operator has an opposite effect on the error in the response. Therefore, the exact operator is situated somewhere between the wrong

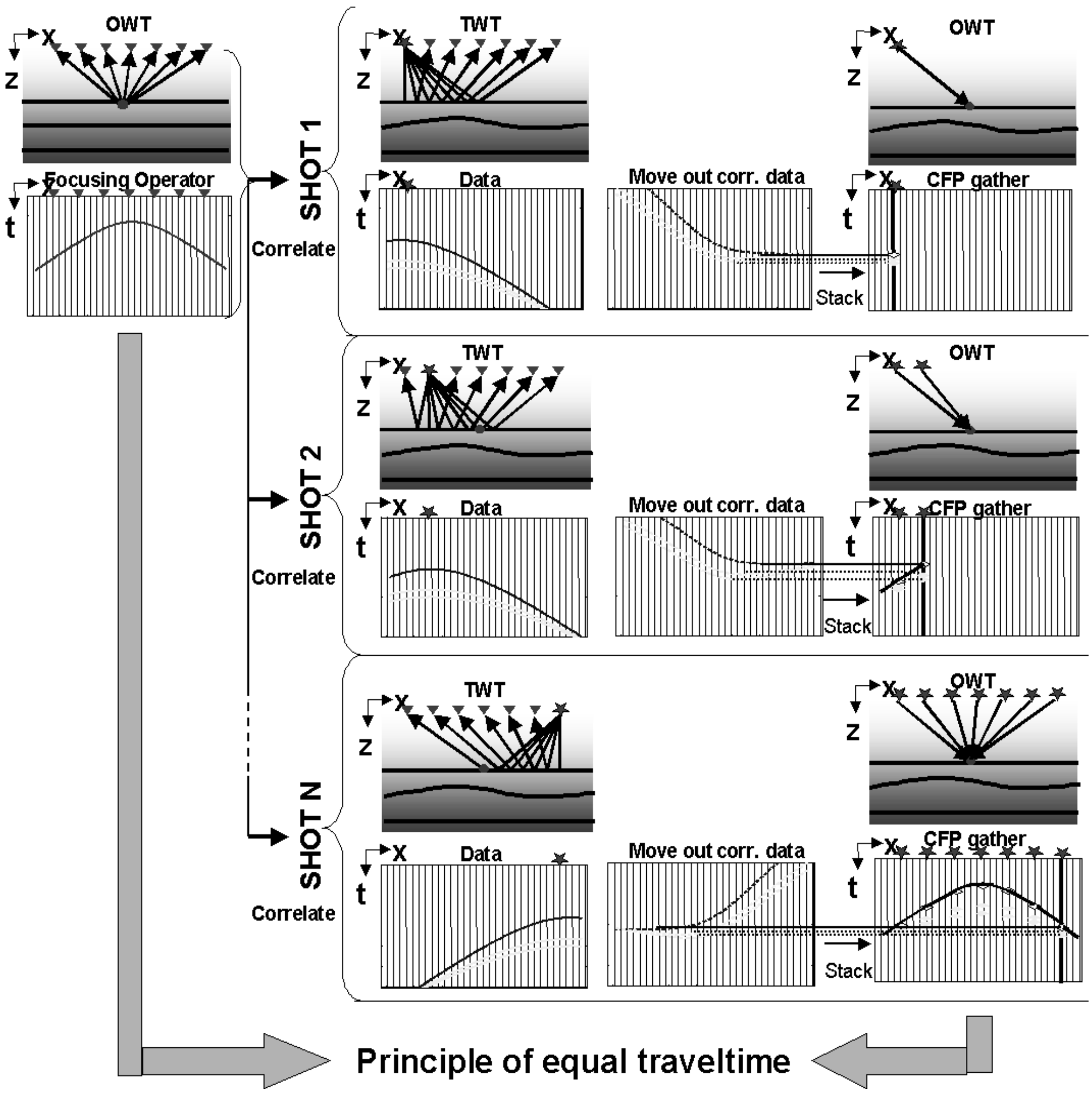


Figure 1: CFP method explained schematically, by both ray pictures and time sections; a modeled focusing operator (one way traveltime (owt)) is applied to a shot (two way traveltime data(twt)), resulting in move out corrected data. The stack of these data result in one trace of a CFP gather (owt). This is done for all (1,2......N) shots, generating a complete CFP gather. When a correct focusing operator is applied, the time-reversed focusing operators and the focus point response (one event in CFP gathers) have equal traveltime. barbara-fop [NR]

operator and the wrong response, depending on the velocity error and the local dip. A data driven update procedure (Berkhout, 1997a; Thorbecke, 1997; Bolte and Verschuur, 1998) for the focusing operator is as follows (Fig.2):

- Cross-correlate each trace of the time-reversed focusing operator with the corresponding trace of the CFP-gather. The result is a correlation panel, which contains stretch-free move-out corrected data around zero time;

- Update each trace of the focusing operator with half the corresponding time-shift in the correlation panel.

Finally, the correlation panel will show a flat event at zero time, indicating that the principle of equal traveltime is satisfied (Fig.2). This procedure will result in exact focusing operators, indicating the response from one point in the subsurface to acquisition locations at the surface. Note that at this stage, no velocity model is involved yet. By the inversion of the traveltimes of these focusing operators a velocity model of the subsurface may be derived .

FOCUSING OPERATOR INVERSION

The CFP velocity estimation that is described in this report is based on a separated inversion method, in which first the focusing operators are obtained from the conventional data (focusing operator updating as described in the previous section), and after that the velocity model is estimated by inverting these focusing operators. In the model, both the velocity and the exact locations of the focus points are unknown. The inversion of the focusing operators should result in a velocity model (that represents the velocity structure of the subsurface) and the exact location of the focus points (that corresponds to reflection energy). This inversion process is carried out with traveltime tomography.

Traveltime tomography

Traveltime tomography generates a velocity model of the subsurface from the observed traveltime response through the subsurface. In focusing operator inversion, the traveltime response is the traveltime of the focusing operator. Traveltime tomography contains three main steps: parameterization of the velocity in the subsurface; forward modeling through the subsurface model to obtain modeled focusing operators; and optimization in which the difference between the traveltimes of the observed and the modeled focusing operators is used to optimize the velocity model and the focus points positions. If the traveltimes of the observed and the modeled focusing operator correspond, an accurate velocity model is obtained. The strategy for focusing operator inversion is presented in Figure 3.

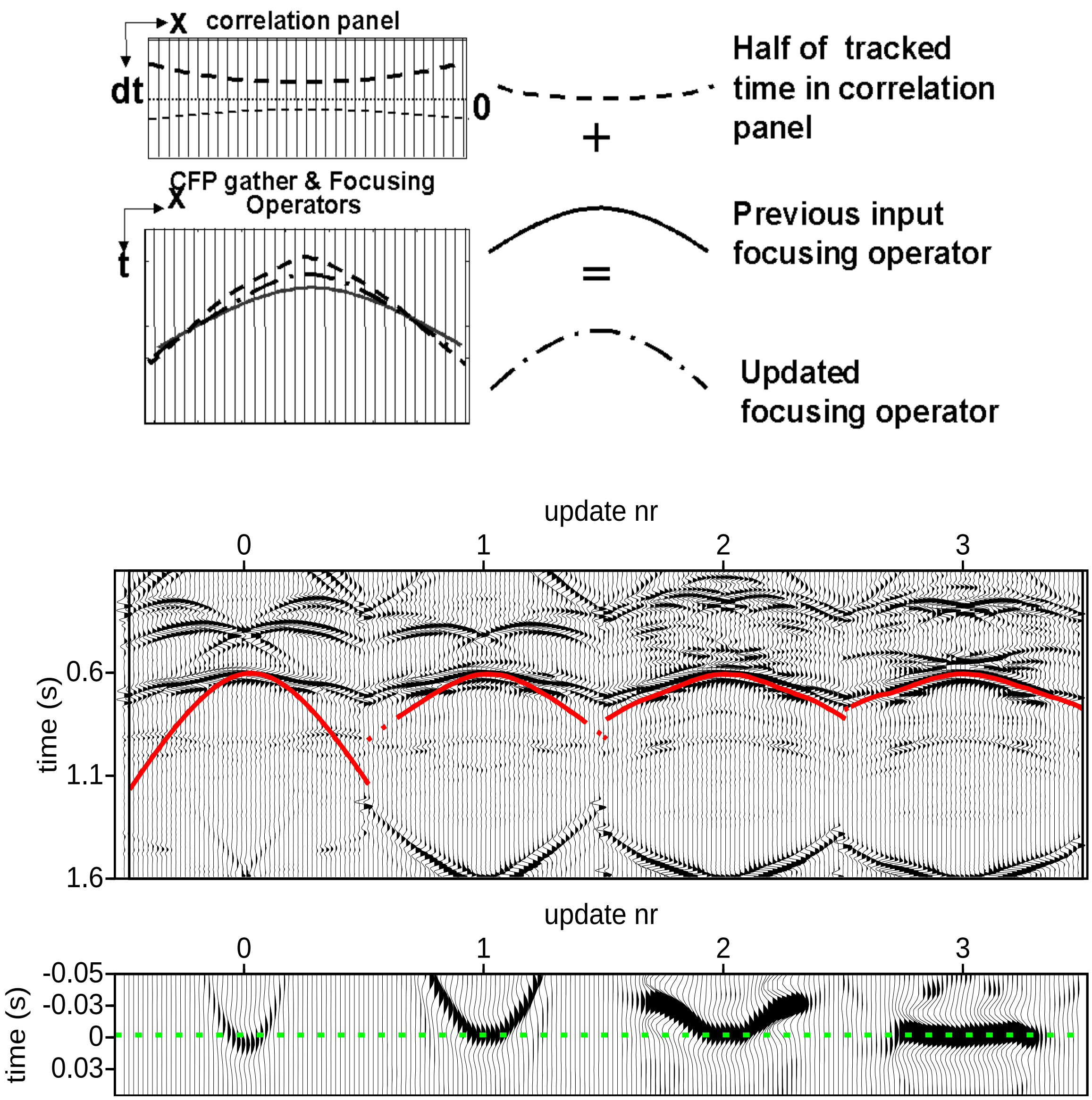


Figure 2: Focusing operator updating. *Top*: Adding half the time shift (dashed line) in correlation panel to the initial focusing operator (solid line in upper plot) results in an updated focusing operator (dash-dotted line in lower plot). Dashed line in lower plot represents the focus point response in the CFP gather. *Bottom*: 4 iterations of focusing operator updating; upper plots show CFP gather and focusing operator (gray line); lower plots show correlation panels. Note that after 4 iterations this panel is flat due to the principle of equal traveltime (*after Bolte (1998)*). barbara-fopup [NR]

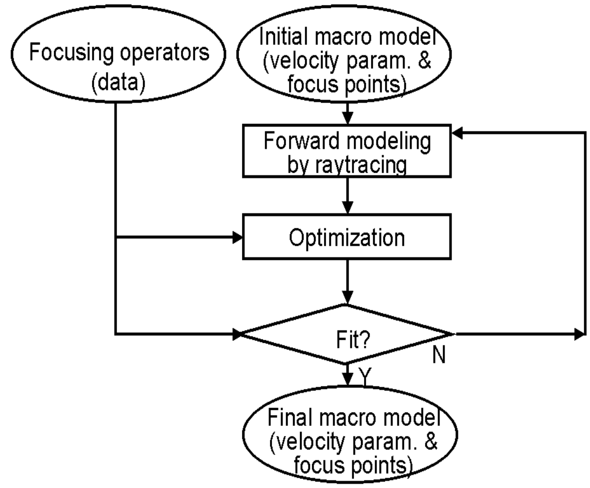


Figure 3: Strategy flowchart of inversion of focusing operators barbara-flow [NR]

Parameterization

There are several ways to describe the velocity model by a set of parameters. A layered parameterization described by global basis functions can be used, i.e. the velocity is estimated for regions and layers in the subsurface defined by the user. However, this parameterization constrains the possible range of solutions. Parameterization based on local basis functions is a more flexible description of the velocity model. The model can be described by cells or grid-points at which the values of the parameters (i.e. the values of the velocity field) are defined.

In the algorithm for tomographic inversion of focusing operators, Delaunay triangulations are used to construct cells (triangles) between grid-points (Fig.4a,b). The velocity is defined at the grid-points and the velocity within the triangles is calculated by a linear interpolation between the three grid-points defining each triangle (Fig.4b). In this way, every point can adopt an optimum velocity, and any kind of subsurface can be described. The focus points are defined independently of the velocity points. The focus points are related to positions where reflection energy is available. By parameterizing them independently of the velocity grid-points, the velocity changes are not dependent on, or constrained to, the reflectors.

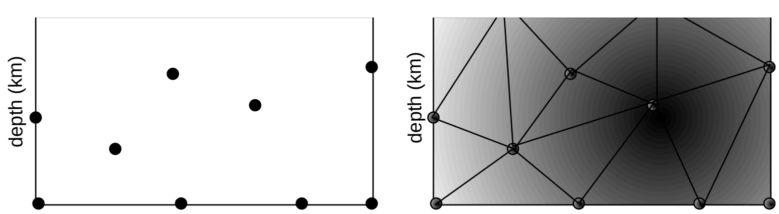


Figure 4: Parameterization of the velocity model. (a) The velocity is defined at grid-points (black dots) in the subsurface. (b) The grid-points are connected by Delaunay triangulation. The velocity within the triangles is calculated by linear interpolation. barbara-param [NR]

Forward Modeling

Forward modeling is done by ray tracing in the initial velocity model. An imaginary shot (focus point) is positioned in the subsurface and rays are traced towards receivers at the surface. In this way modeled focusing operators are obtained. The differences between the traveltimes of these operators and the real focusing operators are used to optimize the velocities in the model and the locations of the focus points. Moreover, the traced rays are needed to define the relation between the traveltime data and the model parameters (matrix **A**, in the next subsection).

Optimization

Optimization is done by calculating the updates of the model parameters by means of the difference between the traveltimes of the modeled and the observed focusing operators. The relation between the traveltime data and the model parameters is assumed to be linear:

$$\Delta \mathbf{d} = \mathbf{A} \Delta \mathbf{m}, \tag{1}$$

where $\Delta \mathbf{d}$ contains the difference in traveltimes, $\Delta \mathbf{m}$ contains the update in model parameters, and $\mathbf{A}$ describes the linear relation between these traveltimes and the model parameters. It might be very difficult to solve for the model vector $\Delta \mathbf{m}$ in terms of the data vector $\Delta \mathbf{d}$, as matrix $\mathbf{A}$ generally is not invertible. However, the generalized matrix inverses can be obtained by several methods. The generalized inverse obtained by these methods is defined by a dagger:

$$\Delta \mathbf{m} = \mathbf{A}^{\dagger} \Delta \mathbf{d}, \tag{2}$$

When a least squares method is chosen to obtain a generalized inverse, an update of $\Delta \mathbf{m}$ is calculated by:

$$\Delta \mathbf{m} = \left(\mathbf{A}^T \mathbf{A}\right)^{-1} \mathbf{A}^T \Delta \mathbf{d}, \tag{3}$$

where $\mathbf{A}^{\dagger} = \left(\mathbf{A}^T \mathbf{A}\right)^{-1} \mathbf{A}^T$ is the generalized inverse. Another way to obtain the generalized inverse is for example by singular value decomposition (SVD). Berryman (2001a) gives a profound overview of the available methods and their capacities. Several methods of approximating the generalized inverse will be presented in one of the next sections.

DATA DEPENDENT PARAMETERIZATION

As inversion problems are generally under-determined, they should be stabilized. Stabilization methods can be divided into two groups:

- Regularization of optimization;
- Adjustment of parameterization;

The first method can be performed by both regularization criteria (e.g. damping of the least squares method, or truncated SVD) and model regularization (e.g. including a priori information or conditional covariance matrices in the optimization). A drawback of this method, however, is that the problem of over-parameterization is still not solved, which might cause problems when the inversion problem is regularized. The second method, the adjustment of parameterization, can for example be done by using a global parameterization. However, as mentioned before, this parameterization constrains the possible range of solutions, and incorrect initial parameterization can lead to slow or non-convergence.

In the inversion of focusing operators an attempt is made to combine both stabilizing methods; the adjustment of parameterization is based on the covariance after optimization; the a posteriori covariance. Model parameters that have a high variance are removed, and in regions containing low variance parameters, extra parameters can be added. In this way, the inversion problem is well determined and over-parameterization is avoided. Moreover, all the available information within the data will be translated to the model, as the parameterization adapts to the data.

If the data is used to determine the parameterization, as a consequence the obtained parameterization tells something about the data. A coarse parameterization exposes for example that more focus points should positioned in that region (of course this is only possible if data related to the focus point is available). On the other hand, when updated focus points are positioned very close to each other, it might be wise to remove/shift them, as they will only duplicate information.

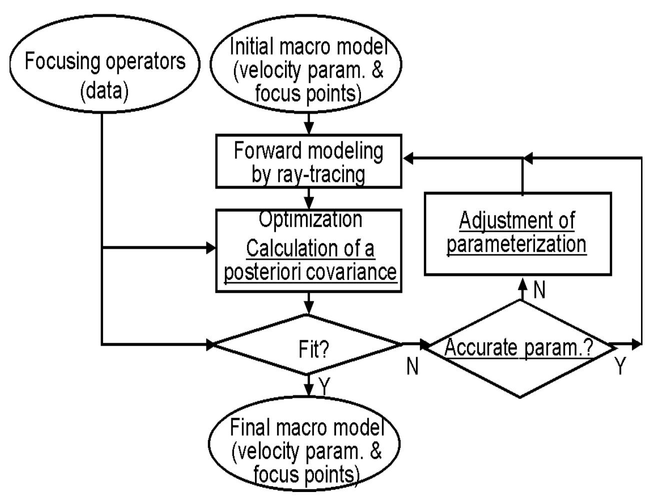


Figure 5: Strategy flowchart for inversion if focusing operators with data dependent parameterization. The extra steps w.r.t. Figure 3 are underlined. barbara-flowap [NR]

A posteriori covariance

The a posteriori covariance is an important tool to evaluate the solution of the optimization. Therefore it is useful to use it as a criteria for the adjustment of the parameterization. The covariance matrix can be calculated during optimization by (Menke, 1984):

$$\mathbf{C} = \mathbf{A}^{\dagger}(\mathbf{A}^{\dagger})^{T}, \tag{4}$$

When least squares optimization (shown in equation (3)) is used, the first term on the right hand side corresponds to the covariance matrix, so

$$\mathbf{C} = \left(\mathbf{A}^T \mathbf{A}\right)^{-1}. \tag{5}$$

The covariance matrix contains M diagonal elements that correspond to the (M) variances of the model parameters. These elements are used to adjust the parameterization after optimization (Fig.5).

In Figure 6 an example of the use of a posteriori covariance for data dependent parameterization is shown. For the velocity model shown in Figure 6a the traveltime data corresponding to the rays in Figure 6b are available. When the a posteriori covariance matrix is calculated and the diagonal elements of the matrix (variances) are plotted for each grid-point, this results in an image as shown in Figure 6c. In this image the dark values indicate low variance and light values indicate high variance. As mentioned before, high variance grid-points are removed, and in regions containing low variance grid-points extra parameters can be added. This finally results in the data dependent parameterization shown in Figure 6d.

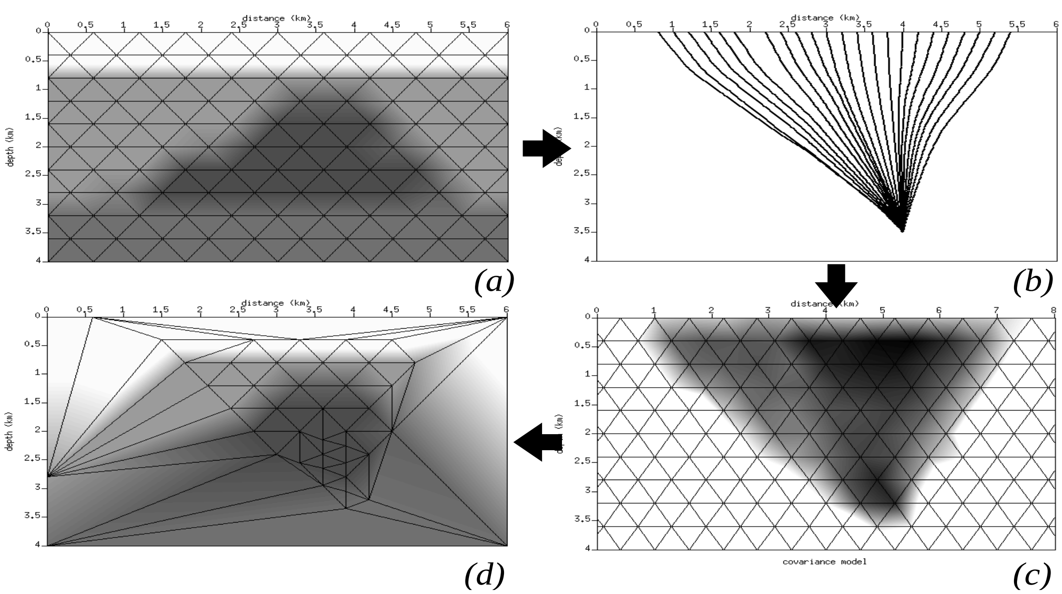


Figure 6: Example of data dependent parameterization. (a) Delaunay triangulation in model defined by a regular grid. (b) Ray-paths from a focus point at position (4.0km; 3.5km) to the surface. (c) Variances at the grid-points in the optimization with ray-paths in (b). Dark values indicate low variance, light values indicate high variance. (d) Adjusted parameterization based on the variances. barbara-datadep [NR]

A synthetic example

The algorithm for tomographic inversion of focusing operators with data dependent parameterization is applied to a synthetic 2D case. The model contains a salt dome, fault structures beneath the dome, a complex turbidite velocity structure with low velocities, and lateral and vertical velocity gradients within the layers (Fig.7a). This is the "real" velocity model of the subsurface, in which (in this synthetic case) the focusing operators are generated (Fig.7b). The positions of the focal points are represented as black dots in Fig.7a. An initial velocity model and an initial estimate of the focal point positions are needed to start the updating procedure (Fig.7c). From these initial locations focusing operators are modeled in the initial velocity model (Fig.7d). Note that for display purposes, a wavelet has been assigned to each operator, and the resulting wave fields have been shown. In the inversion algorithm, only the traveltimes are used. By the traveltime difference between the modeled operators and the real operators, the velocity model and focal point positions are updated. In each iteration, the parameterization will be adjusted and the velocity model and focal point positions will be updated.

The final result after 6 iterations is shown in Figure 8 (displayed in terms of squared-slowness). Note that at locations with high data information (i.e. originally low variances), the density of the grid-points is high. The corresponding velocity model of the final result is also shown in Figure 8. The positions of the focal points are similar to the real focal point positions in Figure 7a. The velocity model also resembles the real model, altough it is a smooth solution of the real case. However, the upper two layers (water layer, and layer below) are not well resolved, which is indicated by the higher velocity, and the downward shifted focus points. This artifact is caused by the severe ray distortion originated in the salt dome. When this velocity model is used for migration (Fig.8), most reflectors are clearly visible. Nevertheless, the inaccuracy of the upper two layers (water layer, and layer below) is also visible in this migration. In particular above the salt dome, where the reflectors are shifted downward.

APPROXIMATE COVARIANCE CALCULATION

For small scale problems, generalized inverses can be calculated easily by explicit matrix calculation during least squares optimization, or during SVD. However, for larger problems this becomes impractible. Several methods for approximation of the generalized inverse have been developed to circumvent this problem. In these methods no explicit matrix calculation is required. However, as the generalized inverse can not be solved explicitly, neither can the covariance matrix be provided by these methods. Several methods for obtaining the covariance matrix from the generalized inverse approximation are available. The approximation methods for the inverse and the covariance can be divided in two groups: iterative methods and model space weighting methods. Both approaches are considered in this section.

Iterative methods

The iterative methods for obtaining the generalized inverse are based on the SVD method. The available methods are among others Conjugate Gradient (CG) (Hestenes and Stiefel, 1952) and LSQR (Paige and Saunders, 1982). Berryman (2001b) provides a complete overview of the iterative methods, and an analysis of their capacities. All the iterative methods lead directly to

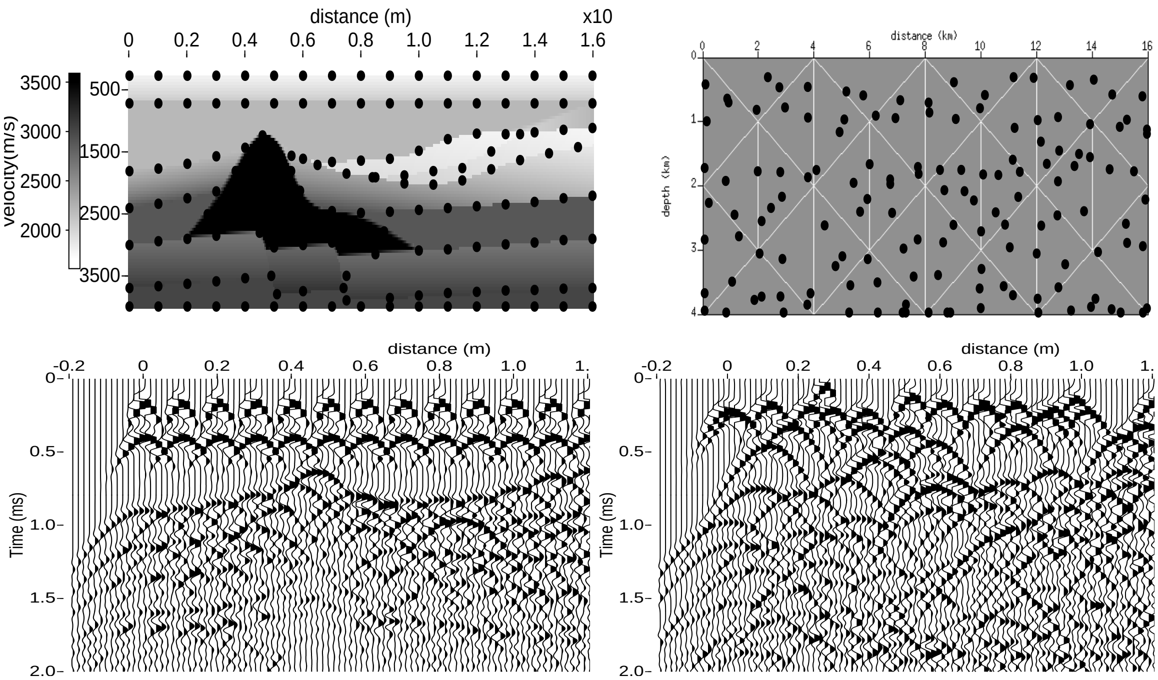


Figure 7: Velocity models of a synthetic case. Black dots represent focal points. (a) Real velocity model and focal points (b) Real focusing operators (modeled in (a), using a wavelet for display); (c) Initial model (squared slowness) and initial focal points; (d) Focusing operators modeled in (c) barbara-synth [NR]

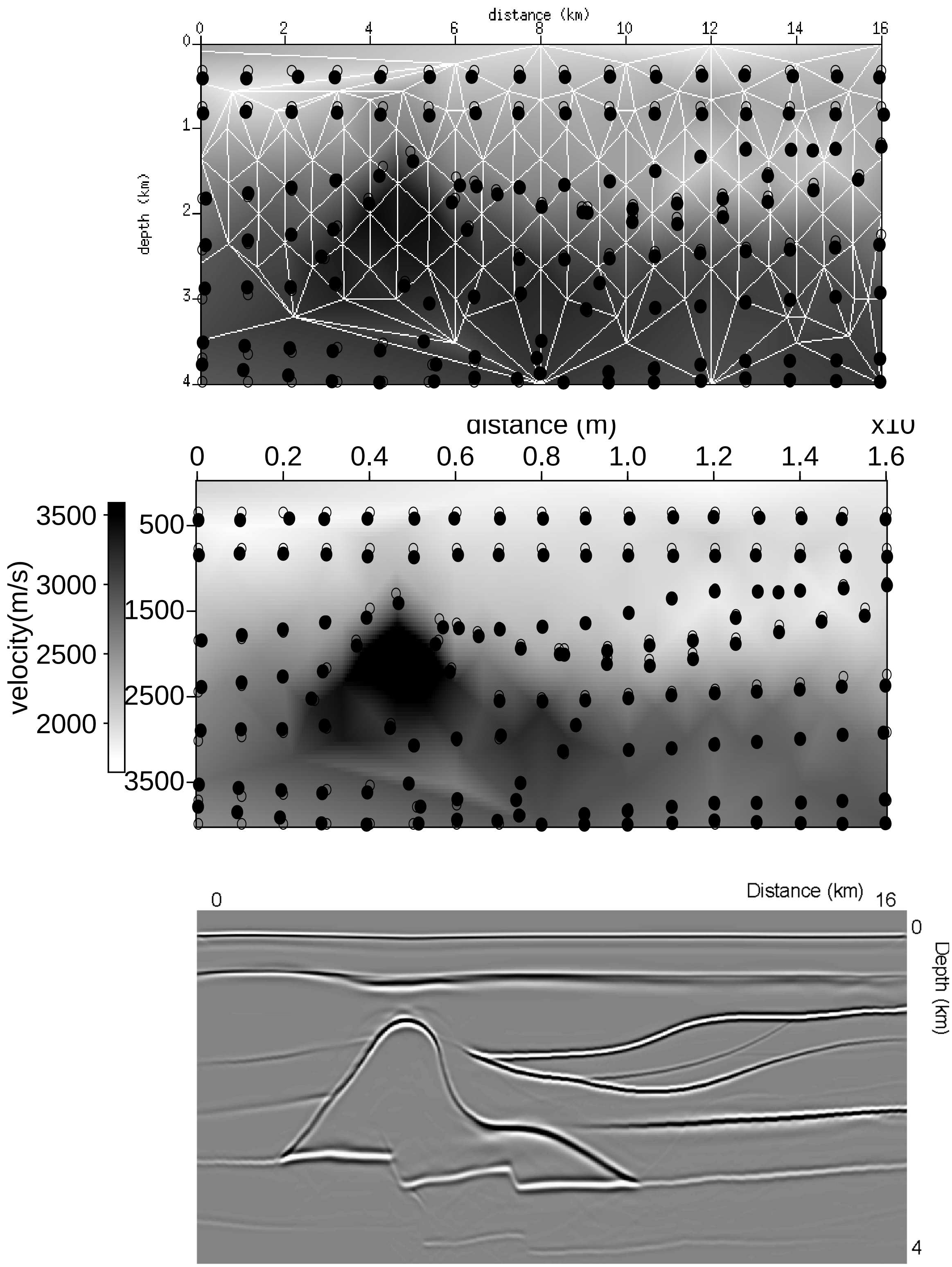


Figure 8: *Top*: Final (sixth) update of model (squared slowness) and focal points; open dots: real focus point locations, black dots: updated focus point location; *Middle*: Final update of model (velocity) and focal points. *Bottom*: Migration with velocity model displayed in figure above. barbara-res [NR]

approximations of the generalized inverse $\mathbf{A}^{\dagger}$. In general, carrying these iterative processes to completion will produce a result closely approximating $\mathbf{A}^{\dagger}$.

Yao et al. (1999) and Berryman (2001b) provide methods for calculating the a posteriori covariance for iterative methods based on SVD. For a thorough description of the methods I refer to these papers. The principle is as follows; each iteration provides one extra vector in the solution space. After K iterations these K vectors are used for calculating the a posteriori covariance. As a consequence, if more iterations are performed, not only the approximated inverse but also the a posteriori covariance becomes more accurate. Figure 9b,c shows a posteriori covariance matrices for $M = 33$ model parameters obtained by respectively CG and LSQR. Both matrices resemble the real a posteriori covariance matrix (Fig.9a). Note that both methods compute the complete covariance matrix, so also the non diagonal elements.

Model space weighting methods

Another approach of approximating the inverse and the a posteriori covariance matrix is provided by the work of Claerbout and Nichols (1994). Rather than trying to solve the explicit inverse problem the adjoint of $\mathbf{A}$ is applied to the data, and a diagonal operator $\mathbf{W_m^2}$ is constructed, such that

$$\Delta\mathbf{m} = \mathbf{W_m^2}\mathbf{A}^T\Delta\mathbf{d}, \tag{6}$$

in which $\mathbf{W_m^2}$ is related to the model space weighting. This term is needed to obtain the correct physical units. Note that this term also corresponds to the a posteriori covariance. The most significant properties of the model space weighting term can be explored by measuring its effects when applied to a reference model $\Delta\mathbf{m_{ref}}$.For example, if $\mathbf{W_m^2}$ is given by

$$\mathbf{W_m^2} = \frac{\mathbf{diag}\,(\Delta\mathbf{m_{ref}})}{\mathbf{diag}\left(\mathbf{A}^T\mathbf{A}\Delta\mathbf{m_{ref}})\right)} \approx \left(\mathbf{A}^T\mathbf{A}\right)^{-1}, \tag{7}$$

will have the same units as $\mathbf{A}^{\dagger}$, and the properties of this term can be explored. Two choices of $\Delta\mathbf{m_{ref}}$ will be shown here (Biondi, 1998):

- $\Delta\mathbf{m_{ref}}$ =unity vector
- $\Delta\mathbf{m_{ref}}$ =unit vector e^k in k direction in m-D ($k = 1,2,...,m$)

The first choice results in the sum of the elements in each column of $\mathbf{A}$ in the denominator. The second choice results in the sum of the square of the elements in each column of $\mathbf{A}$ in the denominator. The obtained vector forms the diagonal in the approximated $\left(\mathbf{A}^T\mathbf{A}\right)^{-1}$. Figure 9d,e shows a posteriori covariance matrices for $M = 33$ model parameters obtained by respectively the "sum" approximation and the "diagonal" approximation. Note that only the diagonal elements are calculated. These elements resemble the diagonal of the real a posteriori covariance matrix (Fig.9a).

Whether the iterative methods, or the model space weighting methods should be used for calculation of the covariance matrix depends also on the algorithm that is used for calculating the approximate inverse. In general, iterative methods are used for tomography, and the vectors needed for calculation of the covariance are obtained without any extra costs. The calculated covariance matrix will reflect the status of the inversion after the performed number of iterations. However, a drawback of these methods is that the solution space vectors need to be saved after each iteration. Moreover, not performing enough iterations might result in strange covariance matrix. Nevertheless, this matrix will still reflect the status of your inversion after the performed number of iterations. The model weight methods are very cheap, and give accurate results of the diagonal elements of the covariance. Nevertheless, the non-diagonal elements are not calculated. However, as for the re-parameterization only this diagonal is considered, these methods might be sufficient. When the covariance between the different parameters is going to play a role in the inversion too, the full covariance matrix obtained from the iterative methods should be used.

CONCLUSIONS

The proposed algorithm for inversion of focusing operators results in an accurate update of both velocities and focus point locations. The data dependent parameterization is an efficient, non-laborious algorithm resulting in well determined inversion problems. The a posteriori covariance that is needed in the algorithm can also be extracted from the approximate matrix inversion algorithms that are available.

FUTURE WORK

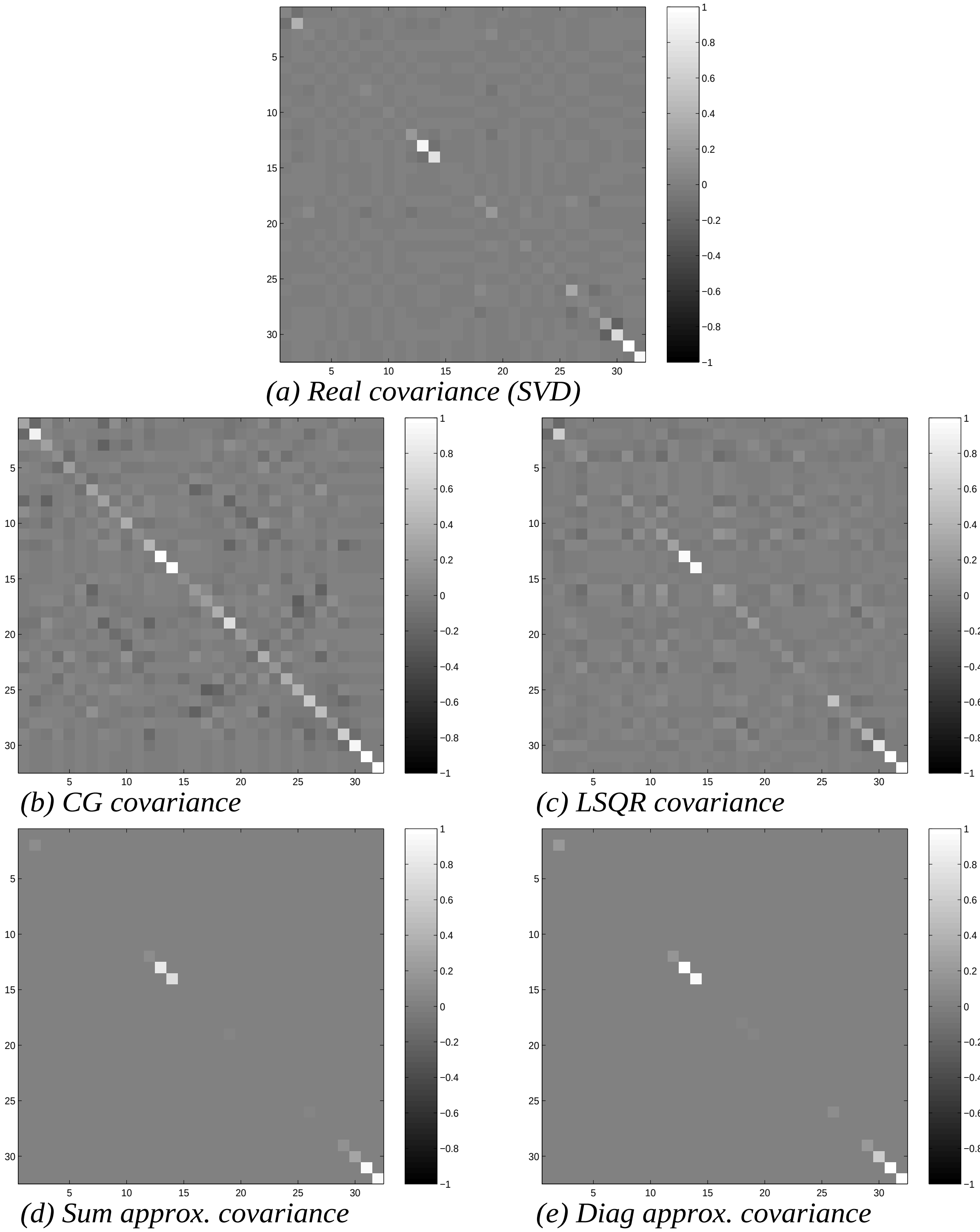
Although in the proposed method all the available information in the data can be translated to a model, this might still not result in a clear image of the subsurface. This can for example be caused by poor data quality or missing data in crucial regions. Therefore, it should always be possible to include a priori information about the subsurface into the inversion. Including a priori velocity information is already possible, however, it might also be necessary to include other information criteria in the inversion scheme.

ACKNOWLEDGMENTS

I would like to thank the SMAART consortium for making my visit to SEP possible, and the people of the CWP project for their Delaunay and ray-tracing software, which formed a base for the developed algorithm described in this report.

REFERENCES

Berkhout, A. J., 1997a, Pushing the limits of seismic imaging, part I: Prestack migration in terms of double dynamic focusing: Geophysics, **62**, no. 3, 937–953.



(a) Real covariance (SVD)

(b) CG covariance

(c) LSQR covariance

(d) Sum approx. covariance

(e) Diag approx. covariance

Figure 9: (a) Real covariance matrix calculated by SVD; (b)(c) Covariance matrices approximated by iterative methods; (d)(e) Covariance matrices approximated by model space weight methods. barbara-covariance [NR]

Berkhout, A. J., 1997b, Pushing the limits of seismic imaging, part II: Integration of prestack migration, velocity estimation, and avo analysis: Geophysics, **62**, no. 3, 954–969.

Berryman, J. G., 2001a, Analysis of approximate inverses in tomography I. Resolution analysis of common inverses: Optimization and Engineering, **1**, no. 1, 87–115.

Berryman, J. G., 2001b, Analysis of approximate inverses in tomography II. Iterative inverses: Optimization and Engineering, **1**, no. 1, 437–473.

Biondi, B., 1998, 3-D seismic imaging: SEP–**98**, 1–204.

Bolte, J. F. B., and Verschuur, D. J., 1998, Aspects of focusing operator updating: 68th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded abstracts, 1604–1607.

Claerbout, J., and Nichols, D., 1994, Spectral preconditioning: SEP–**82**, 183–186.

Hegge, R. F., 2000, Seismic macromodel estimation by inversion of focusing operators: Ph.D. thesis, Delft University of Technology.

Hestenes, M. R., and Stiefel, E., 1952, Methods of conjugate gradients for solving linear systems: J. Res. Nat. Bur. Stan., **B**, no. 49, 409–436.

Kabir, M. M. N., and Verschuur, D. J., 2000, A constrained parametric inversion for velocity analysis based on CFP technology: Geophysics, **65**, no. 4, 1210–1222.

Menke, W., 1984, Goephysical Data Analysis: Discrete Inverse Theory: Academic Press, Orlando.

Paige, C. C., and Saunders, M. A., 1982, LSQR: An algorithm for sparse linear equations and sparse least squares: ACM Transactions on Mathematical Software, **8**, no. 1, 43–71.

Thorbecke, J. W., 1997, Common Focus Point Technology: Ph.D. thesis, Delft University of Technology.

Yao, Z. S., Roberts, R. G., and Tryggvason, A., 1999, Calculating resolution and covariance matrices for seismic tomography with the LSQR method: Geophysical Journal International, **138**, 886–894.