A comparison of methodologies for the selection of regularization parameter in anisotropic traveltime tomography

Eduardo Telmo Fonseca Santos, IGEO/UFBA and CEFET/BA, Brazil.
Amin Bassrei, IF/UFBA and CPGG/UFBA, Brazil.
Jessé C. Costa, UFPA, Brazil.

Abstract

Since inverse problems are usually ill-posed it is necessary to use some method to reduce their deficiencies. The method that we choose is the regularization by derivative matrices. When a first derivative matrix is used the order is called the first. Then, second order regularization is when the matrix is formed by second order differences, and order zero means that the regularization matrix is the identity. There is a crucial problem in regularization, which is the selection of the regularization parameter λ. We use the L-curve as a tool for the selection of λ. L-curve was reintroduced in the literature of inverse problems by Hansen (1992a) and we use it in cross hole traveltime tomography. In this approach of tomography the goal is to obtain the 2-D velocity distribution from the measured values of traveltime between sources and receivers. Besides the L-curve, we also propose a new extension of it, which we called θ-curve. We present several simulation results with synthetic data and we validate the feasibility of regularization, as well as both parameter selection approaches.

Introduction

The main purpose of exploration geophysics for hydrocarbons is to provide trustworthy images of the subsurface, which could indicate potential hydrocarbon reservoirs. Exploration seismology, better known as seismics, is the area of applied geophysics most employed for the subsurface imaging in hydrocarbons reservoirs. And within seismics, tomography was incorporated as a suitable method of data inversion. In this work we use traveltime tomography where the input data is the acoustic traveltime measured at the receivers, and the velocity of the 2-D medium is the inversion output. For the forward modeling we compute the traveltime by acoustic ray tracing from a given 2-D velocity distribution. One common way to calculate inverse matrix is by the generalized inverse through singular value decomposition, but since geophysical tomography is an ill-posed inverse problem, it is necessary to use some tool to reduce this deficiency. The tool that we choose is the regularization of the inverse problem by derivative matrices, known in the literature by several names, specially as Tikhonov regularization. Regularization has an input parameter with crucial role, known as regularization parameter λ, which choice is already a problem. In this work we use the L-curve and an extension of it which we called θ-curve for the selection of regularization parameter in cross hole traveltime tomography.

Methodology

Consider a modeling process where the input of some system is described by certain parameters contained in ***m*** and the output is described as *Gm(=d)* which is a linear transformation on ***m***. If the vector ***d*** describes the observed output of the system, the problem is to "choose" the parameters ***m*** in order to minimize in some sense, the difference between the observed d and the prescribed output of the system *Gm*. If we measure this difference through the norm ||·||, our task is to find the value of **m** which minimizes

$$\left\| G\mathbf{m} - \mathbf{d} \right\|,$$

where the *M*×*N* matrix *G* and the data vector ***d*** with *M* elements are provided to the problem. This is called a least squares problem, which can be formally stated as follows. Considering the basic relationship

$$\mathbf{d} = G\mathbf{m},$$

we wish to minimize the error using the following objective function based on the work of Levenberg (1944) and Marquardt (1963):

$$\Phi(\mathbf{m}) = \mathbf{e}^T\mathbf{e} + \lambda L_2,$$

where the error is given by ***e*** = ***d*** - *G**m***, λ is a scalar called the damping parameter and $L_2 = \mathbf{m}^T \cdot \mathbf{m}$. The estimated solution, also called damped least squares (DLS) solution, is

$$\mathbf{m}^{est} = (G^T G + \lambda I)^{-1} G^T \mathbf{d}.$$

Model estimation can be solved using the method of conjugate gradient (CG), described by Hestenes and Stiefel (1952). This method was developed for the solution of the linear systems like $Ax = b$, where A is a symmetric positive defined matrix, which contains the coefficients of the linear system, b is the vector of observed data, and x is the vector of model parameters. For the solution of the tomographic inverse problem we make the following transformations:

$$A = G^T G,$$

$$\mathbf{b} = G^T \mathbf{d}.$$

The CG algorithm can be used together with regularization. In the below system of linear equations, which we adopted in this work, besides the inverse problem itself we have the first order regularization expressed by the identity matrix, and also the first order regularization represented by the matrix D_1:

$$\begin{bmatrix} A \\ \lambda_1 I \\ \lambda_2 D_1 \end{bmatrix} \mathbf{x} = \begin{bmatrix} \mathbf{b} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}.$$

In order to update the solution of the linear system we use the relation

$$G\Delta\mathbf{m} = \Delta\mathbf{d},$$

which leads to the minimum least squares method

$$\Delta\mathbf{m} = (G^T G)^{-1} G^T \Delta\mathbf{d}.$$

The above equation can be written as

$$\Delta\mathbf{x} = A^{-1}\Delta\mathbf{b}.$$

Thus one can use the CG algorithm to calculate the model residue update, rather than the estimated model itself. It has the advantage to use the first order regularization to minimize the energy of the updated model, making the inversion as linear as possible, which improves the stability in linearized inversion scheme.

Regularization, L-Curve and θ-curve

Least-squares solutions very often do not provide good solutions and sometimes they do not even exist. In order to solve this problem we use a tool of regularization or smoothing: the ill-conditioning of the matrix G is regularized and the unstable least-squares estimate $\mathbf{m}_{est}$ is consequently smoothed to greatly reduce the possibility of wild noise-induced fluctuation in $\mathbf{d}$, hopefully without distorting the resulting smoothed image too far from the true m (Titterington, 1985). The concept of regularization was introduced by Tikhonov in 1963 in order to improve the quality of the inversion. This theory was studied by many researchers, and we use the Twomey (1963) approach. See Bassrei and Rodi (1993) for a little bit more about names and history in regularization theory. Consider the following objective function:

$$\Phi(\mathbf{m}) = \lambda (D_l \mathbf{m})^T D_l \mathbf{m} + \mathbf{e}^T \mathbf{e},$$

where λ is the regularization parameter and D_l is the l^{th}-order derivative matrix. If $\partial\Phi(\mathbf{m})/\partial\mathbf{m} = 0$, then the estimated model is given by

$$\mathbf{m}^{est} = (G^T G + \lambda D^T D)^{-1} G^T \mathbf{d}.$$

Notice that if $\lambda = 0$ we obtain the standard least squares, and the method is said to be damped if $D^T D = I$ (order $l = 0$). If D is the first derivative matrix then the regularization is called to be first order and so on. The matrices D_1 and D_2 are expressed as follows:

$$D_1 = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & -1 & 1 \end{bmatrix},$$

$$D_2 = \begin{bmatrix} 1 & -2 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -2 & 1 \end{bmatrix}.$$

Since a method of regularization is chosen, we need to adopt a criterium to select the best λ. Farquharson and Oldenburg (2004) compared two automatic ways of estimating the best regularization parameter to non-linear inverse problems: GCV and L-curve. These criteria initially proposed for linear problems are applied to each iteration of linearized inverse problems, in a typical iterative process to obtain the linearized solution to the corresponding non-linear problem. Thus, the best λ is estimated for each linearized iteration. To ensure that the regularization parameter decreases along iterations, an attenuation factor is multiplied by the regularization parameter at last iteration to limit the next maximum allowable parameter.

In the present work, to our knowledge the first one in geophysical traveltime tomography using regularization with search for the optimum parameter, we employ the L-curve and one extension of it, which we call θ-curve. In the L-curve the *x* axis represents the error between the observed data and the calculated one, and the *y* axis represents the amount of regularization of the solution. L-curve was reintroduced in the literature of inverse problems by Hansen (1992a, 1998) and he also developed a toolbox (1992b). Hansen's book (1998) is a very good source of information for a more rigorous treatment of L-curve and also mentions the pioneering contributions in this field.

The L-curve knee represents a trade-off between smoother solutions with higher errors and rougher solutions with smaller errors. Thus, the knee detection at the L-curve is an heuristic criterion to select the most appropriate solution. Solutions near to the curve knee are also acceptable and possibly more physically meaningful. We applied the L-curve implementing an automatic method to initially select the best regularization parameter, but solutions with regularization parameter near to the selected one could be also considered. Thus, one can achieve a solution that simultaneously satisfy the criteria of error minimization, smoothness and also with physical meaning. The detection of the L-curve corner was performed using Hansen's toolbox. Considering this curve approximately L-shaped, one can find its knee searching the maximum curvature point. Although, secondary inflexions may occur, which may cause wrong detection of the best regularization parameter. Thus, the automatic method of knee detection adopted in this toolbox may lead sometimes to inadequate regularization parameters. Due to this problem, sometimes we needed to select the best regularization parameter by visual inspection of L-curve and the non-automatic detection of its knee.

We adopted a different criterion based on a curve representing the cosine of angles between adjacent segments of L-curve discrete representation, which we named Theta curve. Where the curve is locally straight, the angle tends to zero, leading the cosine of this angle to one. Near the L-curve knee, the angle tends to be greater than its neighbors, leading the cosine to values below one. Thus, smaller values of cosine are associated with inflexions of the curve, which lead us to inspect the minima of the Theta curve to find the knee of L-curve and consequently the best regularization parameter. The method developed to select the best regularization parameter is based on the detection of the first local minimum of the Theta curve. This minimum is automatically detected where the first derivative is close to zero and the second derivative is positive, adopting thresholds due to the discretization and arithmetic computer precision. Thus, the first occurrence of minimum at Theta curve is associated with the knee of L-curve, giving us the best regularization parameter. Further inflexions of the L-curve were discarded because only the first local minimum of Theta curve is associated with L-curve knee. This avoids the wrong regularization parameter detection described earlier, when one adopts the criterion of maximum curvature of the L-curve.

Anisotropy

Anisotropic materials have properties which depend on the considered direction. In geophysics, the stratified sub-surface can approximated as transversally isotropic due to preferred orientation of minerals. For an extent list of sedimentary rocks, anisotropy is weak, that is, values smaller than 1 % are common in nature (Thomsen, 1986). Besides, for sandstones and carbonates anisotropy can be treated as elliptic, making the inversion process simpler.

In the anisotropic inversion, it is necessary first to determine the elastic properties of the medium, and then to calculate the velocity field. In this section we review the relations between elastic parameters and seismic velocities. We use the Hooke's law to express the relation between the stress (σ) and the strain (ε):

$$\sigma_{ij} = c_{ijkl}\varepsilon_{kl},$$

where σ_{ij} and ε_{kl} are second order tensors and c_{ijkl} is a fourth order tensor. Using the symmetry properties of the three tensors we can express c_{ijkl} as a matrix, usually denoted by C. For the isotropic medium the matrix C is expressed by

$$C_{\alpha\beta} = \begin{bmatrix} C_{33} & (C_{33}-2C_{44}) & (C_{33}-2C_{44}) & & & \\ (C_{33}-2C_{44}) & C_{33} & (C_{33}-2C_{44}) & & & \\ (C_{33}-2C_{44}) & (C_{33}-2C_{44}) & C_{33} & & & \\ & & & C_{44} & & \\ & & & & C_{44} & \\ & & & & & C_{44} \end{bmatrix},$$

where the two coefficients C_{33} and C_{44} can be related to the so-called Lamé parameters, λ and μ, by the relations:

$C_{33} = \lambda+2\mu$,

$C_{44} = \mu$.

For a medium with weak elastic anisotropy the matrix C is given by

$$C_{\alpha\beta} = \begin{bmatrix} C_{11} & (C_{11}-2C_{66}) & C_{13} & & & \\ (C_{11}-2C_{66}) & C_{11} & C_{13} & & & \\ C_{13} & C_{13} & C_{33} & & & \\ & & & C_{44} & & \\ & & & & C_{44} & \\ & & & & & C_{66} \end{bmatrix},$$

From the above coefficients three parameters can be defined (Thomsen, 1986):

$\varepsilon = (C_{11} - C_{33}) / 2C_{33}$,

$\delta = [(C_{13} + C_{44})^2 - (C_{33} - C_{44})^2] / [2C_{33} (C_{33} - C_{44})]$,

$\gamma = (C_{66} - C_{44}) / 2C_{44}$.

If the anisotropy is weak the first parameter is small ($\varepsilon \ll 1$). The P wave and S wave velocities are, respectively,

$\alpha_0 = (C_{33} / \rho)^{1/2}$,

$\beta_0 = (C_{44} / \rho)^{1/2}$.

where the subscript denotes the angle in relation to the horizontal axis. With the above parameters it is possible to compute the P wave velocity as a function of the angle:

$$\alpha_\theta = \alpha_0(1 + \delta \sin^2\theta \cos^2\theta + \varepsilon \sin^4\theta).$$

For the particular case of elliptic anisotropy ($\delta = \varepsilon$):

$\varepsilon = (\alpha_{\pi/2} - \alpha_0) / \alpha_0$,

$\delta = (\beta_{\pi/2} - \beta_0) / \beta_0$.

In this work we used the mild elliptic anisotropy (Carrion et al., 1992) which does not imply in small or weak anisotropy, being than a more general approach. In this case the matrix of elastic coefficients is given by

$$C = \begin{bmatrix} C_{11} & (C_{11}-2C_{66}) & C_{13} & & & \\ (C_{11}-2C_{66}) & C_{11} & C_{13} & & & \\ C_{13} & C_{13} & C_{33} & & & \\ & & & C_{55} & & \\ & & & & C_{55} & \\ & & & & & C_{66} \end{bmatrix},$$

where the anisotropic parameter for the P wave is

$\varepsilon_P = (C_{11} - C_{33}) / (C_{11} + C_{33})$.

The horizontal and vertical P wave velocities are, respectively,

$\alpha_{\pi/2} = (C_{11})^{1/2}$,

$\alpha_{\pi/2} = (C_{33})^{1/2}$.

and the horizontal and vertical S wave velocities are expressed by

$\beta_{\pi/2} = \beta_0 = (c_{55})^{1/2}$.

There is a requirement that $C_{ii} > 0$ in order to obtain real values for the above expressions.

Results

We explore the L-curve and θ-curve approaches for the selection of the regularization parameter in a synthetic example, simulating a reef. The model space was discretized in 400 blocks (20 × 20), where the background medium is homogeneous and isotropic with 3,000 m/s, and the reef itself is homogeneous but anisotropic. Its horizontal velocity is 3,300 m/s and the vertical velocity is 5 % less, that is, 3,135 m/s. Figure 1 displays the true model with its vertical velocities and the horizontal velocities can be seen in Figure 2. The amount of anisotropy is showed in Figure 3. We call this amount the anisotropy factor, which is the percentage difference between the horizontal and the vertical velocities. The reef represents a hydrocarbon reservoir target. This model is a simplification of the geological configuration of the MIT's Earth Resources Laboratory test field, in the Michigan Basin. Since the inversion scheme is anisotropic, for each block there is a pair of model parameters, in such a way that the model parameter vector has 800 components. The adopted acquisition geometry is cross hole, with 31 sources and 31 receivers, equally spaced. This means 961 rays or equations and the inverse problem is thus over-determined, that is, there are more equations than unknowns. Since the problem is ill-posed one must use some kind of regularization. Some portions of the model are over-illumined, where there is a concentration of rays, and in other regions there is a poor illumination, that is, a shortage of rays. We used first order regularization in the horizontal direction, for both P wave velocities (α_0 and $\alpha_{\pi/2}$). Regularization will provide models that are smooth and also realistic in terms of geological stratification. It will also improve, although to a limited extension, the regions with poor illumination. Besides the first order regularization we also used the zero order with a constant regularization factor in the linearized inversion, in order to decrease the variation between iterations, avoiding abrupt changes in the iterative updating process that could divert beyond the linear behavior. We adopted 0.9 as the constant regularization factor for the zero order. Since this value is normalized and very close to one, it would also work fine for a different example. We noticed that for this example that value stabilized the inversion. The adoption of this value is not a problem by itself, since smaller values would lead to unstable solutions and larger values would only delay the convergence. For the choice of first order regularization parameter, one needs a better search procedure since this order of regularization is related to the smoothness of the recovered model in the horizontal direction. From one side, one wants to avoid models with a high smoothing and, from the other side models with abrupt changes in velocity are also undesired. We use the L-curve to obtain the compromise solution, that is, not too smooth and at the same time without unrealistic details.

The L-curve can be seen in Figure 4. We also used the θ-curve that we explained before, and it is displayed in Figure 5. Both curves provided the same "optimum parameter". The recovered vertical velocity is showed in Figure 6 and the horizontal one in Figure 7. We can see that the horizontal velocity is better recovered, due to geometry acquisition.

The anisotropy factor was also well recovered (Figure 8). It was well recovered showing the 5 % contrast of the anisotropic reef within the isotropic background. The velocities were also well recovered allowing the delineation of the potential hydrocarbon reservoir.

The above simulations were performed with noiseless data. In order to check the robustness of this approach we added Gaussian noise in such a level that the RMS error between the original vector traveltime and the corrupted one was around 1 %. The L-curve is showed in Figure 9 and the θ-curve in Figure 10. Here again the L-curve and the θ-curve provided the same "optimum" regularization parameter. Figures 11 and 12 display, respectively, the recovered vertical and horizontal velocities. Again we notice that the horizontal velocity is better recovered. The anisotropy factor is also very well recovered as seen in Figure 13.

Conclusions

From a set of overdetermined synthetic examples corrupted by noise and with an ill-conditioned kernel matrix we have shown that the regularization algorithm in question, together with its related approaches for the selection of the regularization parameter, is feasible in linearized geophysical traveltime tomography tomography. The comparison with non-regularized solution confirms the necessity of some kind of regularization. We considered first order regularization. One crucial aspect here is the selection of regularization parameter, usually choose by some trial and error approach. We used the L-curve and we proposed a variation of it, which we named θ-curve. The results were consistent, providing good to excellent approximations of the true model. Natural extensions of this work are the application of this formulation to a layered medium background, and application to cross hole real data.

Acknowledgements

This work is part of the project Multi-Dimensional and Spectral Seismic Processing in High Resolution Seismo-stratigraphy, sponsored by the edict CTPETRO/CNPq-FINEP 03/2001. The authors also thank the support from PETROBRAS.

References

Bassrei, A., Rodi, W. L., 1993. Regularization and inversion of linear geophysical data. 3rd International Congress of the Brazilian Geophysical Society, Rio de Janeiro, Brazil, volume I: 111-116.

Carrion P., Costa J., Pinheiro J. E. F, Schoenberg M., 1992. Cross-borehole tomography in anisotropic media. Geophysics, 57: 1194-1198.

Farquharson, C. G., Oldenburg, D. W., 2004. A comparison of automatic techniques for estimating the regularization parameter in non-linear inverse problems. Geophysical Journal International, 156: 411-425.

Hansen, P. C., 1992a. Analysis of discrete ill-posed problems by means of the L-curve. SIAM Review, 34: 561-580.

Hansen, P. C., 1992b. Regularization Tools. A Matlab Package for Analysis and Solution of Discrete Ill-Posed Problems. Technical University of Denmark, Lyngby, Denmark.

Hansen, P. C., 1998. Rank-Deficient and Discrete Ill-Posed Problems. SIAM, Philadelphia.

Hestenes, M. R., Stiefel, E., 1952. Methods of conjugate gradients for solving linear system. Journal of Research of National Bureau of Standards, 49: 409-436.

Levenberg, K. 1944. A method for the solution of certain non-linear problems in least squares. Quarterly of Applied Mathematics 2, 164-168.

Marquardt, D. W. 1963. An algorithm for least-squares of nonlinear parameters. Journal of SIAM 11: 431-441.

Thomsen, L., 1986. Weak elastic anisotropy. Geophysics, 51: 1954-1966.

Titterington, D. M., 1985. General structure of regularization procedures in image reconstruction. Astronomy and Astrophysics, 144: 381-387.

Twomey, S., 1963. On the numerical solution of Fredholm integral equations of the first kind by the inversion of the linear system produced by quadrature. Journal of the Association of Computing Machines, 10: 97-101.

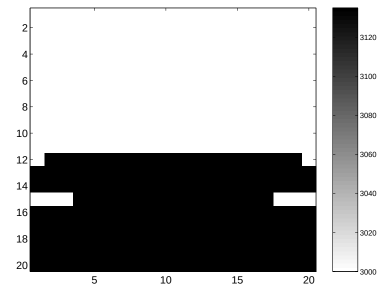

Figure 1 – True model.
Vertical velocity (m/s).

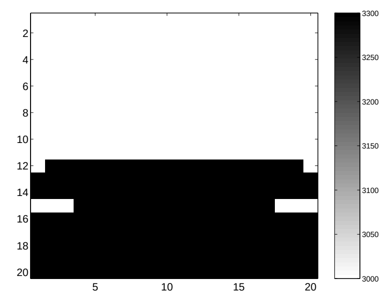

Figure 2 – True model.
Horizontal velocity (m/s).

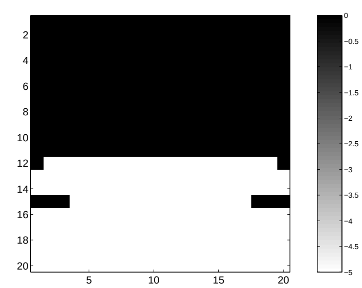

Figure 3 – True model.
Anisotropy factor (%).

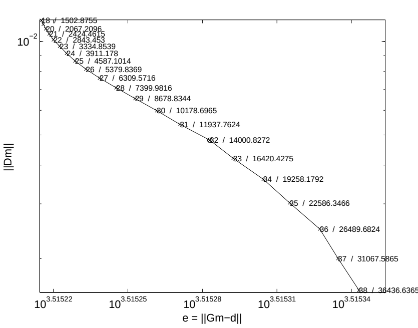

Figure 4 – L-curve for first order.
Noiseless data.

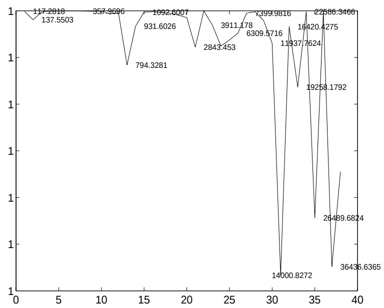

Figure 5 – θ-curve for first order. Noiseless data.

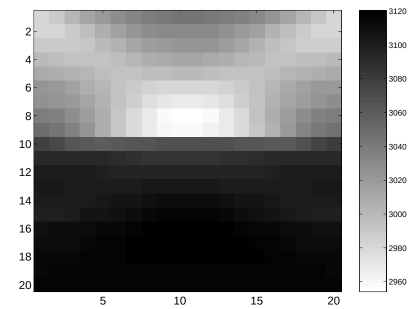

Figure 6 – Recovered vertical velocity (m/s). Noiseless data.

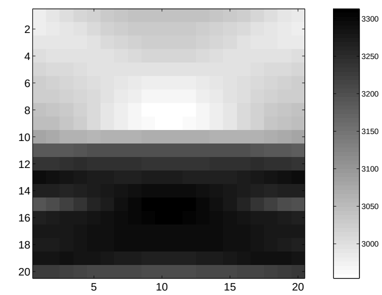

Figure 7 – Recovered horizontal velocity (m/s). Noiseless data.

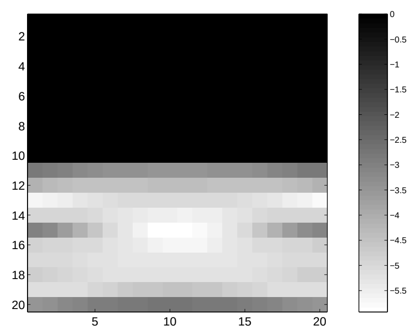

Figure 8 – Recovered anisotropy factor (%). Noiseless data.

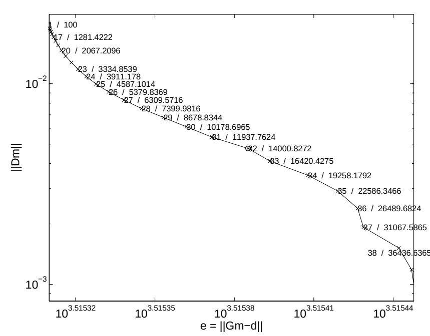

Figure 9 – L-curve for first order. Noisy data.

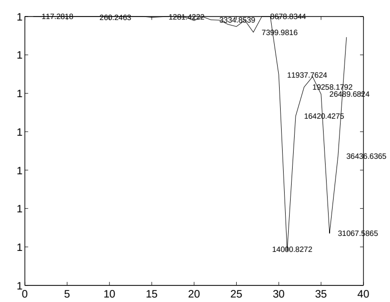

Figure 10 – θ-curve for first order. Noisy data.

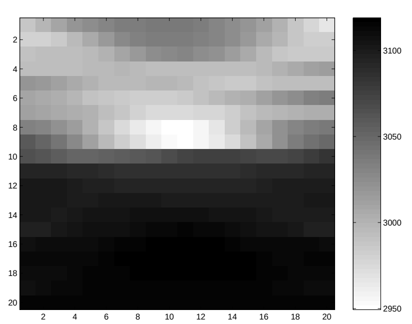

Figure 11 – Recovered vertical velocity (m/s). Noisy data.

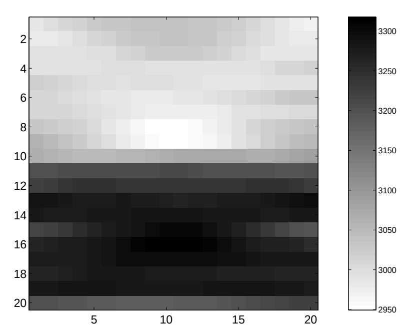

Figure 12 – Recovered horizontal velocity (m/s). Noisy data.

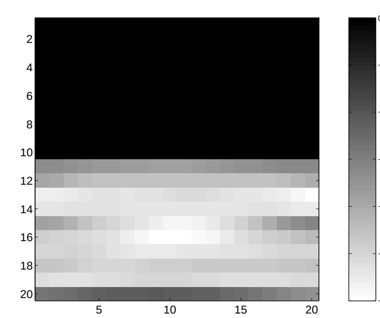

Figure 13 – Recovered anisotropy factor (%). Noisy data.