

CRITIQUE OF SOLUTIONS IN LINEARIZED INVERSE PROBLEMS: NUMERICAL EXPERIMENTS IN TRAVELTIME TOMOGRAPHY

Silvia L. Bejarano¹ and Amin Bassrei²

ABSTRACT. In this work, we evaluated the quality of the solution of numerical experiments in travelttime tomography, in linear and linearized cases, using singular value decomposition. The simulations were performed using a synthetic model, which has been discretized into uniform 2-D blocks, where the slowness in each block was considered to be constant. These simulations were performed with different numbers of singular values, and different levels of Gaussian noise were added to the traveltimes. Additionally, regularization by derivative matrices was used, known in literature as Tikhonov regularization. To evaluate the quality of the inversion, the behavior of the main diagonal of the data resolution matrix and the model resolution matrix was analyzed for different amounts of singular values, providing a clear indicator of the inversion success, either in the whole estimated model or in its parts. A second criterion used the concept of complementary solutions and the generation of a pseudo-constant solution, allowing the analysis of regions where the inversion was not successful.

Keywords: inverse problems, travelttime tomography, resolution matrices, Barbieri method, regularization.

RESUMO. Neste trabalho avaliamos a qualidade da solução em experimentos numéricos em tomografia de tempo de trânsito, nos casos linear e linearizado, utilizando o método de decomposição por valores singulares. As simulações foram feitas utilizando um modelo sintético, o qual foi discretizado em blocos 2-D, uniformes, considerando constantes as vagarosidades em cada bloco. Essas inversões foram efetuadas com diferentes números de valores singulares, assim como foi incorporado diferentes níveis de ruído gaussiano aos tempos de trânsito. Também foi utilizada a regularização por matrizes de derivada, conhecida na literatura como regularização de Tikhonov. Para avaliar a qualidade da inversão foi analisado o comportamento da diagonal principal da matriz de resolução de dado e da matriz de resolução de modelo, para diferentes quantidades de valores singulares, fornecendo um claro indicador do sucesso da inversão, seja no modelo estimado como um todo ou em partes. Um segundo critério utilizou o conceito da solução complementar e da geração da solução pseudo-constante, permitindo analisar em quais regiões do modelo a inversão não foi bem-sucedida.

Palavras-chave: problemas inversos, tomografia de tempos de trânsito, matrizes de resolução, método de Barbieri, regularização.

¹Formerly at Universidade Federal da Bahia, Institute of Geosciences, currently at Observatório Nacional, Rio de Janeiro, Brazil – E-mail: silviabermudez@on.br

²Universidade Federal da Bahia, Research Center in Geophysics and Geology, Institute of Geosciences, and National Institute of Science and Technology in Petroleum Geophysics (INCT-GP), Rua Barão de Jeremoabo, s/n, Ondina, 40170-115 Salvador, BA, Brazil. Phone: +55(71) 3283-8508 – E-mail: bassrei@ufba.br

INTRODUCTION

Many geophysicists dedicate their efforts to the solutions of inverse problems, but generally little attention is given to the criticism of the solution. In other words, once the solution is obtained, little emphasis is given to a quantitative analysis of the solution. The main objective of this work is to present new methodologies to evaluate the inverse problem solution in traveltime tomography.

Seismic tomography is an inverse problem, which allows us to estimate a function from a line integral. In traveltime tomography, the unknown of the problem is the slowness distribution, which is obtained from the values of the data parameter, in this case, the traveltimes between sources and receivers.

In this study, two criteria were used to evaluate the solutions of inverse problems. The first criterion is the resolution matrix, as defined in Jackson (1972). The closer the resolution matrix is to the identity matrix, the better the quality of the inversion.

The second criterion was originally suggested by Barbieri (1974) with an application in medical tomography. We consider that the estimated solution \mathbf{m}^{est} has an associate vector called the complementary estimated solution, expressed by $\mathbf{m}^{est,c}$. The sum of these two vectors is a third vector, denoted by vector \mathbf{w} . If the inverse problem is exact, the vector \mathbf{w} is constant.

To obtain the pseudo-inverse matrix \mathbf{G} , also called the Moore-Penrose inverse (Penrose, 1955), we use singular value decomposition, or SVD. When using SVD, it is necessary to avoid small singular values so as not to compromise the results because these small singular values act as if they were noise. Therefore, we used some criteria for the selection of singular values, such as the decay behavior of the singular values, the RMS error between the observed and calculated data parameters, and the RMS error between the true and estimated model parameters. These criteria were applied in traveltime tomography by Silva & Bassrei (2007) for the linear case and Silva & Bassrei (2008, 2009) for the linearized case.

Due to the illposedness of the tomographic problem we also used, for the linearized iterative inversion, regularization with derivative matrices. The results improved and validated the application of Barbieri method.

The region of interest was divided into square 2-D blocks with constant dimensions. The data acquisition geometry was cross-well, where the sources and receivers were uniformly distributed on the sides of a 30×20 grid. The application of the proposed methodology to synthetic data was necessary because it allows the calculation of the RMS error between the true model parameters \mathbf{m}^{true} and the estimated model parameters \mathbf{m}^{est} . In real data, \mathbf{m}^{true} is not available. We also calculated the RMS error

between the observed data parameters \mathbf{d}^{obs} and the calculated data parameters \mathbf{d}^{cal} . However, this estimator is only auxiliary because our inverse problem is usually ill-posed and non-unique and a small RMS error for data parameters is not usually associated with consistent solutions.

INVERSE PROBLEMS AND TRAVELTIME TOMOGRAPHY

Inverse problems are usually ill-conditioned; that is, its solutions do not exist, and if they exist, they are not unique and/or they are unstable. On the other hand, overdetermined problems do not have a unique solution. However, the least squares method enables the determination of a unique solution, which has a minimum RMS error. If the problem to be analyzed is underdetermined, then the system does not have enough information to determine a unique solution, and therefore there are infinitely many solutions.

In addition to these problems, there is also the issue of stability; that is, even if the solution exists and is unique, it may still be unstable. A small disturbance in the data (such as noise) implies a great disturbance in the model parameters to be estimated.

One way to overcome this problem is to regularize it to increase the stability of the solution. The instability of the inverse problem can be assessed by the condition number of the problem, defined as the ratio between the largest and smallest singular value of the matrix that relates the model parameters to the data parameters.

According to Menke (2012), one can classify the relationship between data and model parameters into two classes in the analysis of physical phenomena. The first class is direct or forward modeling, where the data are predicted from a given model. In the second class, called inverse modeling, the model parameters are determined from observed data. Menke (2012) also states that the linear, linear by parts or linearized inverse problems can be formulated as a system of linear equations:

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

where $\mathbf{d} = [d_1, d_2, \dots, d_M]^T$ is the vector of the data parameters and $\mathbf{m} = [m_1, m_2, \dots, m_N]^T$ is the vector of the model parameters. $\mathbf{G}_{M \times N}$ is a coefficient matrix that relates the M data parameters to the N model parameters.

In most geophysical problems, the matrix \mathbf{G} is an approximation of a non-linear operator g :

$$\mathbf{d} = g(\mathbf{m}). \quad (2)$$

Tomography is an image reconstruction technique with many applications in geophysics (Stewart, 1991; Lo & Inderwiesen,

1996). Travelttime tomography is an inverse problem with a kinematic approach that uses the traveltimes between sources and receivers as input data in the inverse procedure. Travelttime tomography finds application in reservoir geophysics because it is an appropriate technique used in reservoir characterization and monitoring.

The travelttime of a given ray in a path r is given by

$$t = \int_r s(x, z) dl, \tag{3}$$

where $s(x, z)$ is the 2-D slowness distribution and dl is an element of the ray path. By Fermat's principle, the ray path in the above integral is the one that has stationary value.

The travelttime equation (3) is a non-linear relationship like equation (2). In homogeneous media, the rays are straight, but the medium is non-homogeneous. Moreover, the ray path r depends on the slowness distribution $s(x, z)$ and the ray path will be curved. Before we proceed to linearization, we will use vector notation for the traveltimes t and the slowness distribution $s(x, z)$ in such a way that

$$\mathbf{t} = g[\mathbf{s}(x, z)]. \tag{4}$$

Using Taylor series, we can expand \mathbf{t} around a reference slowness \mathbf{s}_0 . We truncate the expansion by ignoring the terms with order greater than or equal to two:

$$\mathbf{t} - \mathbf{t}_0 = \left. \frac{\partial g}{\partial \mathbf{s}} \right|_{\mathbf{s}=\mathbf{s}_0} (\mathbf{s} - \mathbf{s}_0), \tag{5}$$

which can be written as

$$\Delta \mathbf{t} = \mathbf{G} \Delta \mathbf{s}, \tag{6}$$

where $\Delta \mathbf{t} = \mathbf{t} - \mathbf{t}_0$ is the travelttime residual and $\Delta \mathbf{s} = \mathbf{s} - \mathbf{s}_0$ is the slowness residual. Equation (6) can be adapted for an iterative procedure so that the travelttime residual can now be expressed as $\Delta \mathbf{t} = \mathbf{t}^{obs} - \mathbf{t}^{(k)}$, where \mathbf{t}^{obs} is the vector of the observed traveltimes and $\mathbf{t}^{(k)}$ is the k -th iteration vector of the calculated traveltimes, and $\Delta \mathbf{s}^{(k)} = \mathbf{s}^{(k+1)} - \mathbf{s}^{(k)}$, where $\mathbf{s}^{(k+1)}$ is the $k + 1$ -th iteration vector of the estimated slownesses and $\mathbf{s}^{(k)}$ refers to the k -th iteration.

We also used another kind of regularization to make the solution more stable is minimally limiting the variation of model parameters. One can use the difference between physically adjacent model parameters as an approximation of the first derivative. The sum of these values can be defined as flatness $\mathbf{1}_1$ (Menke, 2012) of the solution,

$$\mathbf{1}_1 = \begin{pmatrix} -1 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & -1 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} m_1 \\ m_2 \\ \vdots \\ m_N \end{pmatrix} = \mathbf{D}_1 \mathbf{m}. \tag{7}$$

The elements of matrix $\mathbf{G}_{M \times N}$ are the distances traveled by the rays. Each row of matrix \mathbf{G} is associated to one ray and each column to a specific portion of the medium. A simpler way to parameterize the slowness distribution is to divide the study area into small cells or blocks and assign constant values of slowness. The use of high-frequency sources allows a precise determination of the traveltimes, which will provide high-resolution images of velocity structures. Thus, the g_{ji} elements correspond to the j -th ray inside the i -th block. This matrix is sparse, for a given ray intercepts only a small part of the model.

The so called direct or forward modeling addresses the calculation of traveltimes, performed in this work by ray tracing and discussed in detail in Appendix A.

Properly speaking, the inversion is the determination of the final estimate of \mathbf{s} , the vector of the model parameters. This is done iteratively along with ray tracing. In each iteration k , the input data for the inversion is the vector $\Delta \mathbf{t} = \mathbf{t}^{obs} - \mathbf{t}^{(k)}$ and the outcome is the vector $\Delta \mathbf{s}^{(k)} = \mathbf{s}^{(k+1)} - \mathbf{s}^{(k)}$, which in turn allows the calculation of the final estimate of \mathbf{s} using the relationship $\mathbf{s}^{(k+1)} = \mathbf{s}^{(k)} + \Delta \mathbf{s}^{(k)}$. The inverse of \mathbf{G} is the generalized inverse \mathbf{G}^+ (Penrose, 1955), calculated in this work by singular value decomposition or SVD.

REGULARIZATION

For the numerical inversion we used the classical singular value decomposition (SVD), which is presented in Appendix B. One important issue in SVD is the existence of small singular values, which are always a problem because they usually degrade the solution of the inverse problem. Thus, some type of regularization is necessary, and we choose singular value selection, that is, our criterion was to check the amplitude of the singular values. Usually a SVD program provides the singular values in descending order, so that from a certain value, this information is regarded as noise that compromises the quality of inversion. Different amounts of singular values were tested, in order to make a selection of the number of singular values, denoted by n_{sv} , to be used in the SVD inversion.

Or we can choose to use the roughness \mathbf{l}_2 (Menke, 2012) of the model parameters by using an approximation of the second derivative matrix \mathbf{D}_2 :

$$\mathbf{l}_2 = \begin{pmatrix} 1 & -2 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & -2 & 1 \end{pmatrix} \begin{pmatrix} m_1 \\ m_2 \\ \vdots \\ m_N \end{pmatrix} = \mathbf{D}_2 \mathbf{m}. \tag{8}$$

The value of L_n , either flatness ($n = 1$) or roughness ($n = 2$), related to the model parameters is:

$$L_n = \|\mathbf{l}_n\|_2^2 = (\mathbf{D}_n \mathbf{m})^T (\mathbf{D}_n \mathbf{m}),$$

where n is the order of the derivative matrix. One can then define an objective function $S(\mathbf{m})$ as:

$$S(\mathbf{m}) = \mathbf{e}^T \mathbf{e} + \lambda L_n;$$

or

$$S(\mathbf{m}) = (\mathbf{d} - \mathbf{G}\mathbf{m})^T (\mathbf{d} - \mathbf{G}\mathbf{m}) + \lambda (\mathbf{D}_n \mathbf{m})^T (\mathbf{D}_n \mathbf{m}), \tag{9}$$

where λ is a positive constant called regularization parameter, which represents the regularization intensity that is applied to provide a satisfactory solution. When $\lambda = 0$ the above equation lies in the least squares method. If $\lambda \neq 0$ and the regularization order is zero; that is, $L_0 = \|\mathbf{l}_0\|_2^2 = (\mathbf{D}_0 \mathbf{m})^T (\mathbf{D}_0 \mathbf{m})$, we have the so called damped least squares method. Minimizing the objective function shown in equation (9), we get:

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

When $n = 0$, \mathbf{D}_n corresponds to the identity matrix resulting in a zero-order regularization, and the solution reduces to the damped least squares method. When $n = 1$, \mathbf{D}_n corresponds to the first derivative matrix, and the regularization is then called first-order. When $n = 2$, we have the second-order regularization, which uses the numerical approximation of the second derivative.

For the linearized process, we start with a homogeneous initial model \mathbf{m}^0 and updated the model parameters iteratively through the equation:

$$\Delta \mathbf{m}^k = [(\mathbf{G}^T \mathbf{G} + \lambda \mathbf{D}_n^T \mathbf{D}_n)^k]^+ (\mathbf{G}^T)^k \Delta \mathbf{d}^k,$$

where,

$$\mathbf{m}^{k+1} = \mathbf{m}^k + \Delta \mathbf{m}^k.$$

RESOLUTION MATRICES

We used the resolution matrix criterion for a quantitative analysis of the inverse problem solution. The closer the resolution matrix is to the identity matrix, the best the inversion quality (Jackson, 1972).

There are two resolution matrices. The model resolution matrix is a matrix that characterizes the relationship between the estimated model parameters and the parameters of the true model

(Menke, 2012). The model resolution matrix depends on the inverse problem's structure (model geometry and experiment geometry) and, eventually, some a priori information. The main use of the model resolution matrix is to provide a resolution measurement obtained from the data that is of the degree so that the resolution matrix approximates the identity matrix (Jackson, 1972):

$$\mathbf{m}^{est} = \mathbf{G}^+ \mathbf{d}^{obs} = \mathbf{G}^+ (\mathbf{G} \mathbf{m}^{true}) = \mathbf{R}_m \mathbf{m}^{true},$$

and

$$\mathbf{R}_m = \mathbf{G}^+ \mathbf{G}.$$

In turn, the data resolution matrix is the matrix that characterizes the relationship between the observed data and calculated data with a given model (Menke, 2012). This matrix describes how well the predictions match the data. The data resolution matrix is determined only by the kernel matrix and a priori information added to the problem:

$$\mathbf{d}^{cal} = \mathbf{G} \mathbf{m}^{est} = \mathbf{G} (\mathbf{G}^+ \mathbf{d}^{obs}) = \mathbf{R}_d \mathbf{d}^{obs},$$

and

$$\mathbf{R}_d = \mathbf{G} \mathbf{G}^+.$$

To perform a quantitative study of the inverse problem, two estimators were proposed: one related to the error of the diagonal

elements of the model resolution matrix,

$$\varepsilon_{R_m} = \frac{1}{N} \sqrt{\sum_{i=1}^N (1 - R_{m,ii})^2} \times 100\%, \quad (10)$$

and the other related to the error of the diagonal elements of the data resolution matrix,

$$\varepsilon_{R_d} = \frac{1}{M} \sqrt{\sum_{j=1}^M (1 - R_{d,jj})^2} \times 100\%, \quad (11)$$

BARBIERI METHOD

The second method was suggested by Barbieri (1974) and was initially used in medical tomography. Consider that \mathbf{m}^{est} is the solution of an inverse problem, and $\mathbf{m}^{est,c}$ is its complementary solution, so that the sum of these two vectors is given by a constant vector \mathbf{w} :

$$\mathbf{m} + \mathbf{m}^c = \mathbf{w}. \quad (12)$$

We can either work with vectors or matrices. In other words, a vector \mathbf{m} can be written or plotted as a matrix \mathbf{M} :

$$\begin{aligned} \mathbf{m} &\Leftrightarrow \mathbf{M}, \\ \mathbf{m}^c &\Leftrightarrow \mathbf{M}^c, \end{aligned}$$

and

$$\mathbf{w} \Leftrightarrow \mathbf{W}.$$

In the latter case, \mathbf{W} is a constant matrix given by $\mathbf{W} = (w_0)$; $w_0 \geq \max(m_i)$. The complementary solution \mathbf{m}^{est} can be obtained from \mathbf{d}^{obs} and \mathbf{G} using any inverse method; in this work, we used SVD.

We then calculate the vector of complementary observed data parameters $\mathbf{d}^{obs,c}$ using the relationship $\mathbf{d}^{obs,c} = \mathbf{G}\mathbf{w} - \mathbf{d}^{obs}$. A second inversion is now performed to compute the vector of complementary observed model parameters $\mathbf{m}^{est,c}$. We now add \mathbf{m}^{est} and $\mathbf{m}^{est,c}$ to be:

$$\mathbf{w}^{est} = \mathbf{m}^{est} + \mathbf{m}^{est,c}. \quad (13)$$

If the inverse problem is exact, then $\mathbf{w}^{est} = \mathbf{w}$. Because this usually does not occur, we can verify in which regions the inversion was not successful.

In the case of the linearized inversion, the pseudo-constant vector \mathbf{w}^{est} is checked at each k -th ray tracing iteration:

$$\mathbf{w}^{est,k} = \mathbf{m}^{est,k} + \mathbf{m}^{est,c,k},$$

where

$$\begin{aligned} \mathbf{m}^{est,k} &= \mathbf{m}^{est,k-1} + \Delta\mathbf{m}^{est,k-1}, \\ &= \mathbf{m}^{est,k-1} + \mathbf{G}^{+,k} \Delta\mathbf{d}^{k-1}, \\ &= \mathbf{m}^{est,k-1} + \mathbf{G}^{+,k} (\mathbf{d}^{obs} - \mathbf{d}^{cal,k-1}), \end{aligned}$$

and

$$\begin{aligned} \mathbf{m}^{est,c,k} &= \mathbf{m}^{est,c,k-1} + \Delta\mathbf{m}^{est,c,k-1}, \\ &= \mathbf{m}^{est,c,k-1} + \mathbf{G}^{+,k} \Delta\mathbf{d}^{c,k-1}, \\ &= \mathbf{m}^{est,c,k-1} + \mathbf{G}^{+,k} (\mathbf{d}^{obs,c} - \mathbf{d}^{cal,c,k-1}), \end{aligned}$$

Therefore, we have a clear view whether the inversion method was effective on the whole model or in parts of it. In this study, we conducted a quantitative analysis using an error estimator between a constant vector \mathbf{w} and the usually non-constant \mathbf{w}^{est} :

$$\varepsilon_w = \frac{\sqrt{\sum_{i=1}^N (w_i - w_i^{est})^2}}{\sqrt{\sum_{i=1}^N (w_i)^2}} \times 100\%. \quad (14)$$

The Barbieri method was applied in in geophysical tomography by Bassrei (2000) for a simple example in linear travel-time tomography. In the present study, besides exploring deeper the linear case we extended the application to the linearized case, taking into account the presence of noise.

SIMULATIONS IN LINEAR AND LINEARIZED TRAVELTIME TOMOGRAPHY

The synthetic model, shown in Figure 1, was discretized into 600 blocks, where each block is a 20 by 20 m square and has a constant velocity. The overall dimensions of the region of interest are 600 m in horizontal length and 400 m in depth from the surface, thus resulting in 30 blocks in the horizontal direction and 20 blocks in the vertical direction. The model's main feature is a layer that represents a reservoir with a P-wave propagation of 4400 m/s. The model has several horizontal layers where the velocity ranges from 3600 m/s to 4000 m/s. Figure 1(a) shows the velocity distribution, where the color bar indicates the velocities in m/s. Because the unknown in the inverse process is slowness and not the velocity, the same true model is shown in Figure 1(b), where the color bar now displays the slowness in ms/m.

The data acquisition geometry is well to well, where the sources and receivers are uniformly distributed on the sides of a 20 × 30 grid. We considered three configurations in the acquisition geometry: the first configuration has 20 sources and 20 receivers resulting in an underdetermined problem with 400 equations and 600 unknowns; the second one has 25 sources and 24 receivers, resulting in a determined problem with 600 equations

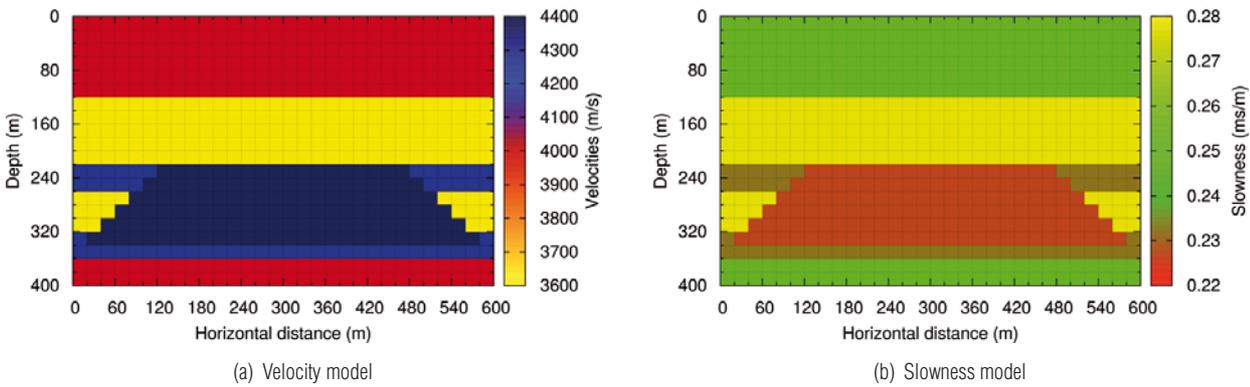


Figure 1 – Velocity model, where the color bar indicates the P-wave velocity in m/s in (a) and slowness in ms/m in (b).

and unknowns; finally, the third configuration, which is an overdetermined case, has the same number of unknowns but 900 equations resulting from 30 sources and 30 receivers.

The simulations were performed with data corrupted by different levels of Gaussian noise. The noise was added to the observed traveltimes by the equation $t_j^* = t_j(1 + ar_j)$, $j = 1, \dots, M$, where t_j^* represents the j -th component of the data parameter vector contaminated by noise, t_j represents the j -th component of the noise-free data parameter vector, a is the amplitude of the applied noise and r_j is the j -th element of a quasi-random number sequence.

The following two estimators were used to evaluate the inversion performance: (i) data parameter error, which is the difference between the observed and calculated model parameters and is expressed in terms of percentage:

$$\epsilon_d = \frac{\sqrt{\sum_{j=1}^M (t_j^{obs} - t_j^{cal})^2}}{\sqrt{\sum_{j=1}^M (t_j^{obs})^2}} \times 100\%. \quad (15)$$

where M is the number of rays or equations in the inversion, t_j^{obs} is the j -th observed traveltimes and t_j^{cal} is the calculated one, and (ii) model parameter error, which is the difference between the true and estimated model parameters and is expressed in terms of percentage:

$$\epsilon_m = \frac{\sqrt{\sum_{i=1}^N (s_i^{true} - s_i^{est})^2}}{\sqrt{\sum_{i=1}^N (s_i^{true})^2}} \times 100\%. \quad (16)$$

where N is the number of blocks or unknowns in the inversion, s_i^{true} is the i -th true slowness and s_i^{est} is the estimated one. Notice that this criterion can be used in only simulations with synthetic data because the N true slownesses s_i^{true} are never known.

Many simulations were performed, but only some results are displayed due to space limitations. All three situations were explored: underdetermined, determined and overdetermined. We used noise-free data and noisy data using different noise levels: $\alpha = 0.1, 0.05, 0.01, 0.005$ and 0.001 . Due to the harmful influence of small singular values over noisy data, different amounts of singular values were considered: $n_{sv} = 200, 250, 300, 350, 400, 450, 500$ and 550 . The relative RMS data error ϵ_d was computed by equation (15), and the relative RMS model error ϵ_m was computed by equation (16). The model resolution matrix error was computed using equation (10), and the data resolution matrix error was computed using equation (11). The pseudo-constant matrix error was computed using equation (14).

For Figure 2, the considered noise level was $\alpha = 0.1$ and the number of singular values used in the construction of the pseudo-inverse matrix was $n_{sv} = 200$. Most of the elements in the model resolution matrix's main diagonal are far from 1.0, as observed in Figure 2(c), and most of the elements in the pseudo-constant vector \mathbf{w}^{est} are far from the assumed value of 0.2 ms/m. Comparing Figure 2(d) to Figure 2(b) it is possible to see a good agreement between the yellow color (reference value of 0.20) in the new Figure 2(d) and the well recovered portions of the estimated model (Fig. 2(b)).

In Figure 3, the considered noise level was $\alpha = 0.01$ and the number of singular values used in the construction of the pseudo-inverse matrix was $n_{sv} = 400$. Most of the elements in the model resolution matrix's main diagonal are near 1, and most of the elements in the pseudo-constant matrix \mathbf{w}^{est} are near the assumed value of 0.2 ms/m. In other words, the inversion can be qualified as successful even without knowledge of the true model (Fig. 1(b)). The evaluation is done by analyzing Figure 3(c) and Figure 3(d).

Traveltome tomography is, by definition, a non-linear problem,

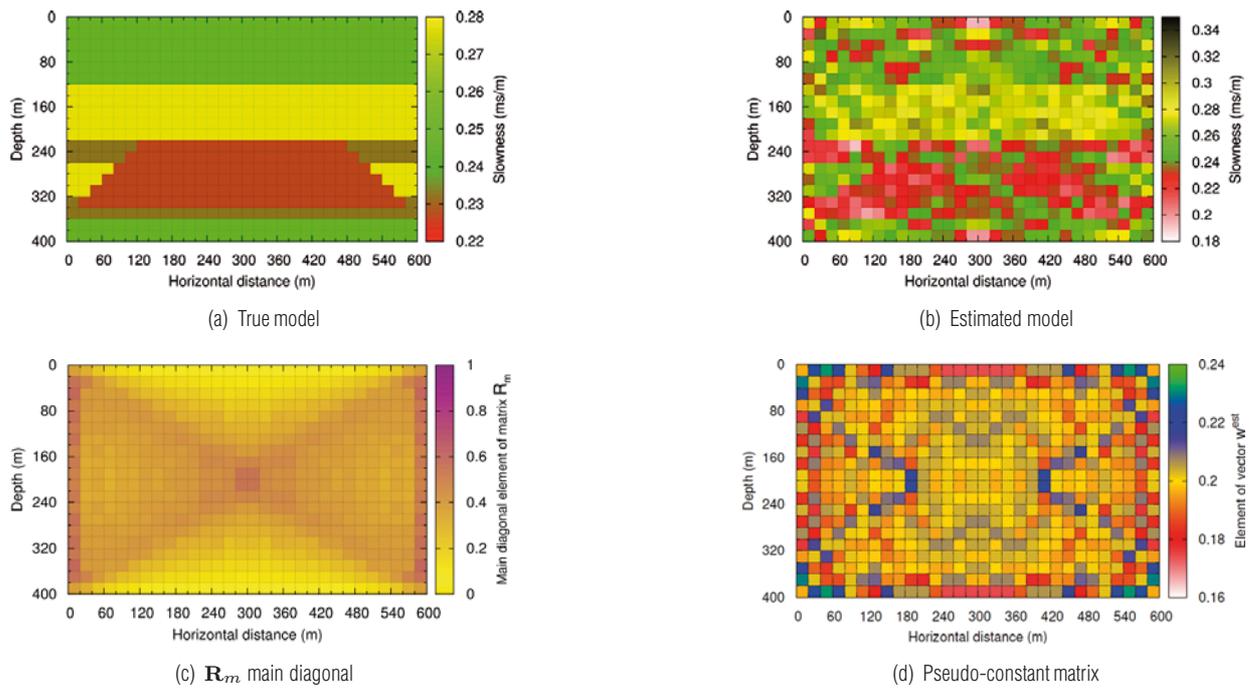


Figure 2 — Simulations for the overdetermined case. The noise level is $\alpha = 0.1$ and the number of singular values used in the construction of the pseudo-inverse matrix is $n_{sv} = 200$. (a) True model, where the color bar indicates the P-wave slowness in ms/m. (b) Estimated model. (c) The main diagonal of the model resolution matrix. (d) Pseudo-constant matrix where the color bar indicates the elements of \mathbf{w}^{est} in ms/m.

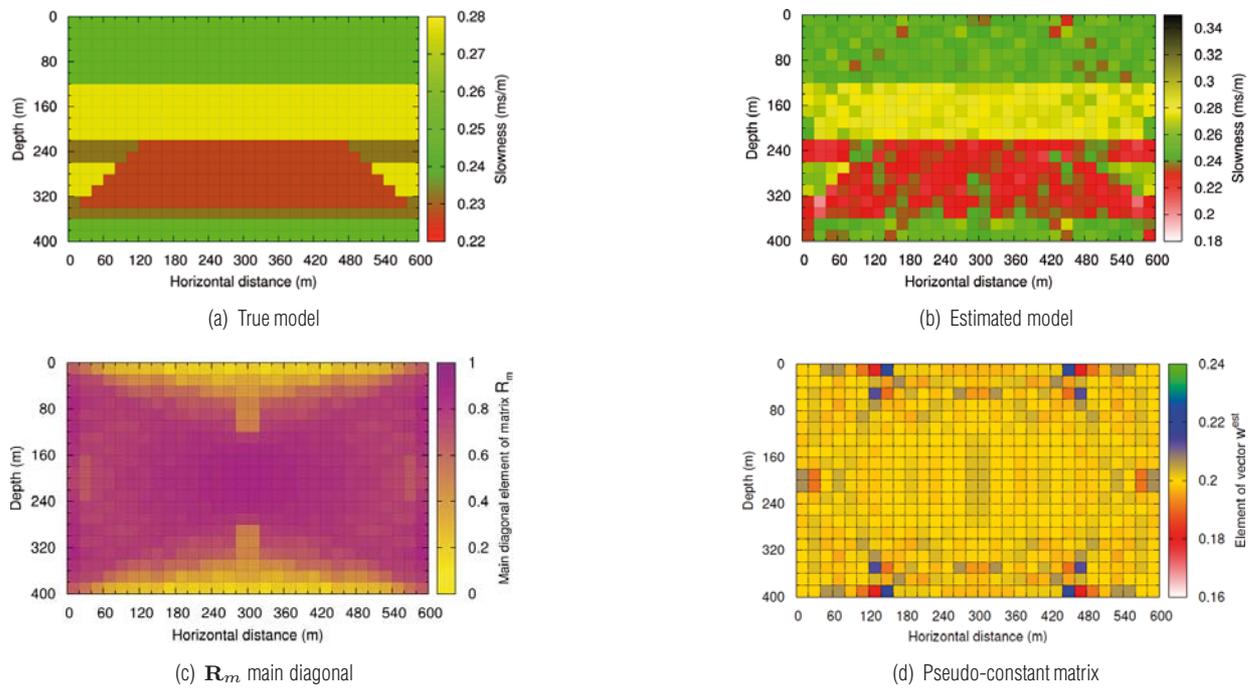


Figure 3 — Simulations for the overdetermined case. The noise level is $\alpha = 0.01$ and the number of singular values used in the construction of the pseudo-inverse matrix is $n_{sv} = 400$. (a) True model, where the color bar indicates the P-wave slowness in ms/m. (b) Estimated model. (c) The main diagonal of the model resolution matrix. (d) Pseudo-constant matrix where the color bar indicates the elements of \mathbf{w}^{est} in ms/m.

so the linear inversion is just an approximation that depends on the velocity variation in the subsurface, which may lead to inconsistent results. The linearized version is also an approximation but closer to physical world behavior.

Again, all three situations were explored: underdetermined, determined and overdetermined. Once more, we used noise-free data and noisy data using different noise levels: $\alpha = 0.1, 0.05, 0.01, 0.005$ and 0.001 . The harmful influence of small singular values over noisy data is even more evident in the linearized inversion scheme, and four amounts of singular values were considered: $n_{sv} = 200, 250, 300,$ and 350 . The estimators were computed, and analyzing the results of $\varepsilon_{R_m}, \varepsilon_{R_d}$ and ε_w , we can deduce that they decrease in value as the amount of singular values increases.

Figure 4 shows the result of the ε_m estimate as a function of iteration number for four different amounts of singular values n_{sv} and four different noise levels α , and Figure 5 presents the ε_d estimate for the conditions as in Figure 4. Notice that the ε_d estimate always converges because the inverse problem is ill-posed, but the same behavior is not valid for the ε_m estimate. In fact, the noise level expressed by the factor $\alpha = 0.1$ was very high, compromising the inversion result. In Figures 4 and 5 considered a maximum of 10 iterations, for the sake of comparison. However, the stopping criterion used in this work was the calculated data RMS residual between iterations k and $k + 1$, $\Delta\varepsilon_d^{k+1} = \varepsilon_d^{k+1} - \varepsilon_d^k \leq \varepsilon_{\min}$, where

$$\varepsilon_d^k = \frac{\sqrt{\sum_{j=1}^M (t_j^{obs} - t_j^{cal,k})^2}}{\sqrt{\sum_{j=1}^M (t_j^{obs})^2}} \times 100\%. \quad (17)$$

We adopted the value $\varepsilon_{\min} = 10^{-2}$.

Figure 6 presents the inversion results at different iterations for the overdetermined case with $n_{sv} = 250$ and $\alpha = 0.001$, using the above stopping criterion.

Comparing Figure 6(a), which is the estimated model parameter obtained at the first iteration to Figure 6(b), which is the pseudo-constant image, it is possible to see that most unsuccessful portions of the estimated solution are correspondent with deviations from the dark yellow color in Figure 6(b), like for instance the almost triangular symmetric artifact from 120 to 200 m depth. There is also the 'hole' in the reservoir, from 220 to 280 m depth. The same is true with the reservoir flanks. There are exceptions, like the last layer (380 to 400 m depth) which is well recovered in Figure 6(a) but the pseudo-constant image in Figure 6(b) indicates a large deviation in relation to reference (0.26 ms/m in relation to 0.20 ms/m). For the fifth iteration, the estimated model (Fig. 6(c)) is improved, and most of the colors in Figure 6(d) are

within the range around 0.20 ms/m. But again, the method was not successful for the last layer. One explanation is the fact that the ray coverage is poor both for the first and the last layers, and this inconsistency gets worse iteration after iteration. However for the first layer, the true model has constant velocity from the surface to 120 m depth, making the ray coverage uniform, despite not being dense. And for last layer (360 to 400 m depth), there is a slowness variation with the above layer (340 a 360 m), making the ray coverage less uniform. On should mention that the slowness range in the true model varies from 0.22 to 0.28 ms/s whereas in the estimated model it varies from 0.18 to 0.34 ms/s. This limitation comes from the inversion itself, that it, the illposedness together with the linearized approach.

The above simulations were repeated with the inclusion of regularization. Again the simulations were made for underdetermined, determined and overdetermined systems. As before, we made simulations using noise free data and data contaminated with different levels of noise, given by the parameters $\alpha = 0.00, 0.001, 0.005, 0.01, 0.05$ and 0.1 . All three orders of regularization were considered ($n = 0, 1, 2$). For the choice of regularization parameter we used the trial and error approach selecting, in the case of our synthetic data, $\lambda = 10^2, 10^4, 10^6$ and 10^8 . The introduction of regularization matrices improved the solutions, as can be seen in Table 1 which shows the results only for the linearized approach with regularization and for overdetermined case.

When compared to the non regularized simulations, the inclusion of regularization provided better results, as expected. Figure 7 shows different iterations of the linearized inversion through SVD method, for the overdetermined case and with first order regularization. Figure 7(a) shows the estimated model parameters for the first iteration and Figure 7(c) for the seventh iteration. The correspondence between Figure 7(a) and Figure 7(b) is very clear. There is also a good correspondence in relation to last iteration, in particular the reservoir flanks, not well recovered in Figure 7(c) and clearly identified in Figure 7(d), attesting the methodology effectiveness. When we compare Figure 6(c) to Figure 7(c) one can see that the regularization improves the solution quality, and the same is true when comparing Figure 6(d) to Figure 7(d). Overall the smoothing procedure 'cleans' the estimated model as well as the pseudo-constant image. The regularization parameter used was $\lambda = 10^4$ and the noise level was expressed by the factor $\alpha = 0.1$.

CONCLUSIONS

Seismic inversion is an inverse procedure to reconstruct the Earth's velocity model from seismic data and it is an important

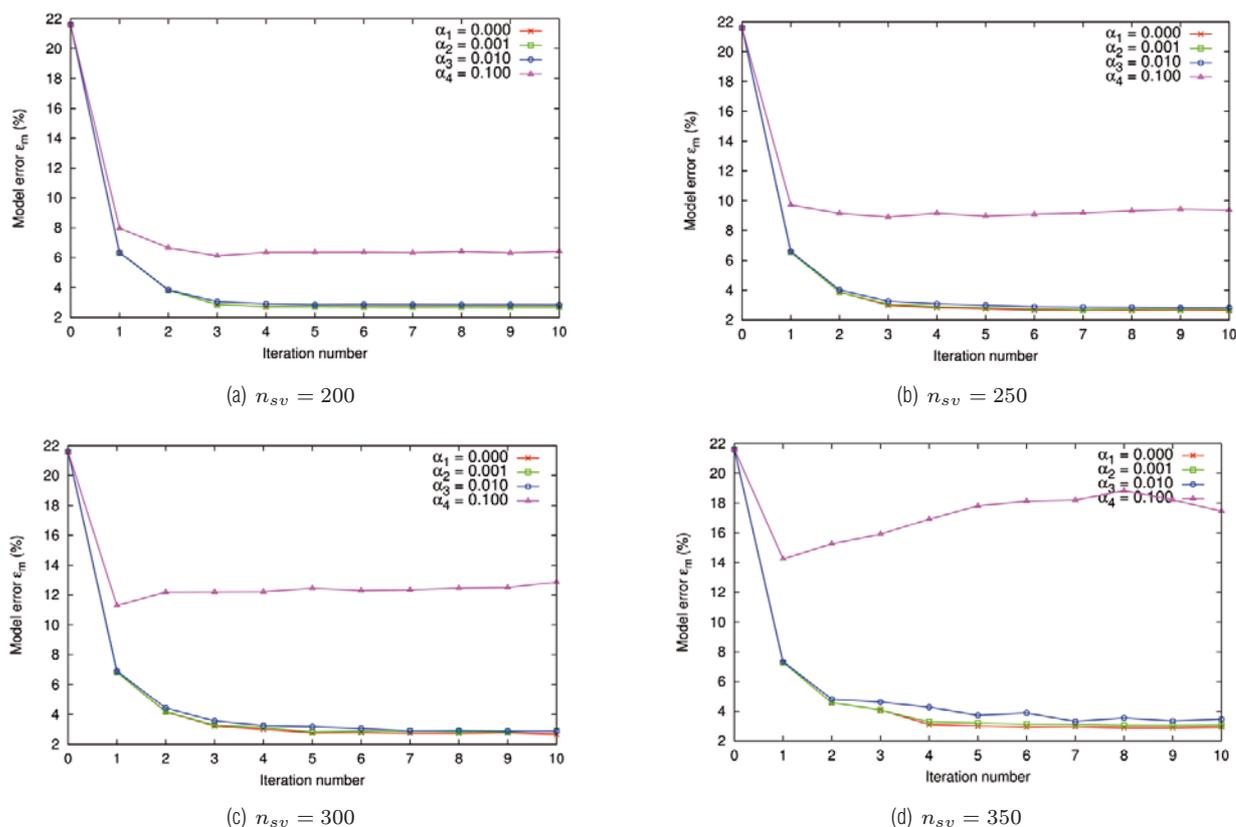


Figure 4 — Simulations for the linearized overdetermined case. The model parameter error ϵ_m is a function of the number of iterations for different noise levels α and different amounts of singular values used in the construction of the pseudo-inverse matrix n_{sv} : (a) $n_{sv} = 200$; (b) $n_{sv} = 250$; (c) $n_{sv} = 300$; (d) $n_{sv} = 350$.

tool for reservoir characterization, which provides a detailed image of the subsurface. On the other hand most of the geophysical problems are considered ill-posed and traveltime tomography, which is a nonlinear problem, does not escape this situation. To solve this ill-conditioning mathematical treatment is required, that could provide reliable solutions that address the problem non-linearity. In this sense we used singular value decomposition, making a cutoff study of the small singular values, and also Tikhonov regularization.

The presence of small singular values in the inversion procedure generates an anomalous increase in the model parameter error and data parameter error criteria. Thus, it was necessary to analyze some criteria to determine the regions with the optimal number of singular values. The results show that the application of data resolution matrix and model resolution matrix criteria show the regions of the 2-D model distribution where the inversion was unsuccessful.

Similarly, one can verify the inversion quality through the Barbieri method, where the sum of the vectors \mathbf{m}^{est} and $\mathbf{m}^{est,c}$ approaches a constant value \mathbf{w} ; as a result, one can observe and

evaluate the inversion quality, which is quantitatively confirmed by the RMS model parameter error. On the other hand, in the qualitative, visual approach, the Barbieri method indicates if the inversion was successful and/or in which parts of the model the inversion was successful. This can also be seen qualitatively, by checking the RMS model parameter error.

The numerical simulations performed in a synthetic model with 600 blocks, both for the linear and linearized cases, allowed to validate the criteria used in this study. The results were satisfactory, indicating the overall quality of the solution as well as the regions of the model where inversion was successful.

ACKNOWLEDGEMENTS

S. Bejarano thanks CNPq for a PhD scholarship. A. Bassrei thanks CNPq for the projects National Institute of Science and Technology in Petroleum Geophysics (INCT-GP) and 308690/2013-3 (research fellowship), PETROBRAS for sponsoring the project "Investigation on the Use of Crosswell Tomography as a Toll for Complex Reservoir Characterization", and FINEP for sponsoring the CTPETRO Network in Exploration Geophysics (Rede 01).

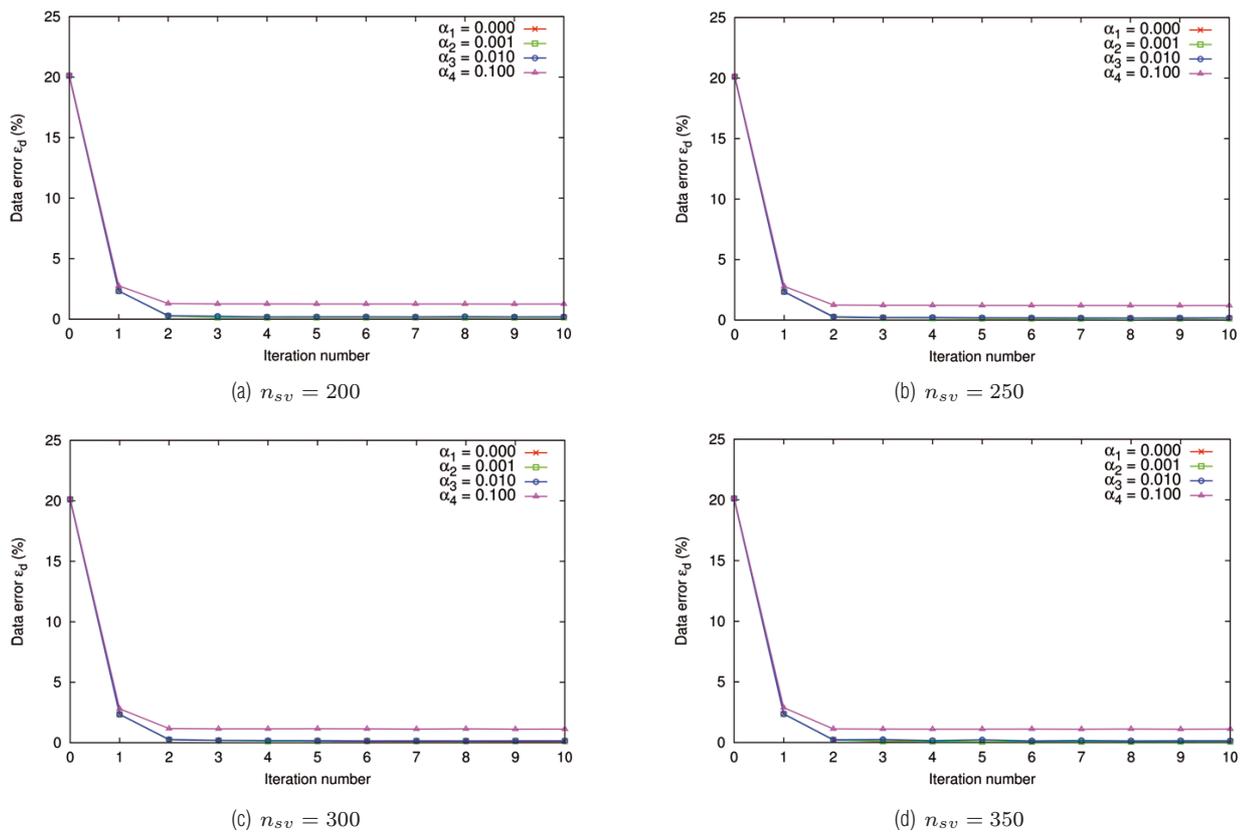


Figure 5 — Simulations for the linearized overdetermined case. The data parameter error ϵ_d is a function of the number of iterations for different noise levels α and different amounts of singular values used in the construction of the pseudo-inverse matrix n_{sv} : (a) $n_{sv} = 200$; (b) $n_{sv} = 250$; (c) $n_{sv} = 300$; (d) $n_{sv} = 350$.

REFERENCES

- ANDERSEN AH & KAK AC. 1982. Digital ray tracing in two-dimensional refractive fields. *Journal of Acoustical Society of America*, 72(5): 1593–1606.
- BARBIERI M. 1974. A criterion to evaluate three dimensional reconstructions from projections of unknown structures. *Journal of Theoretical Biology*, 48: 451–467.
- BASSREI A. 2000. Novel approaches for the solution and solution evaluation of linear and non-linear inverse problems in geophysics. *Boletim da Sociedade Brasileira de Matemática Aplicada e Computacional*, 4: 1–7.
- JACKSON DD. 1972. Interpretation of inaccurate, insufficient and inconsistent data. *Geophysical Journal of the Royal Astronomical Society*, 28: 97–109.
- LO T-W & INDERWIESEN P. 1996. *Fundamentals of Seismic Tomography*. Society of Exploration Geophysics, Tulsa, OK, 178 pp.
- MENKE W. 2012. *Geophysical Data Analysis: Discrete Inverse Theory*, 3rd ed., Academic Press, Orlando, FL, 293 pp.

PENROSE R. 1955. A generalized inverse for matrices. *Proceedings of the Cambridge Philosophical Society*, 51: 406–413.

SILVA JN & BASSREI A. 2007. Critérios de seleção de valores singulares em problemas inversos lineares: uma aplicação em tomografia de tempos de trânsito. *Sitientibus Série Ciências Físicas*, 3: 32–48.

SILVA CJMG & BASSREI A. 2008. Critérios de seleção de valores singulares em tomografia linearizada de tempos de trânsito. In: *III Simpósio Brasileiro de Geofísica*, Belém, PA, Brasil, 26 a 28 de novembro de 2008, 6 pp.

SILVA CJMG & BASSREI A. 2009. Regularização de problemas inversos por seleção de valores singulares: aplicação em tomografia iterativa de tempos de trânsito. In: *11th International Congress of the Brazilian Geophysical Society*, Salvador, BA, Brazil. CD-ROM.

STEWART RR. 1991. *Exploration Seismic Tomography: Fundamentals*. Society of Exploration Geophysicists, Tulsa, OK, 201 pp.

APPENDIX A: ACOUSTIC RAY TRACING

In the geometric acoustics approach, the energy may be transported along curves whose trajectories are orthogonal to the

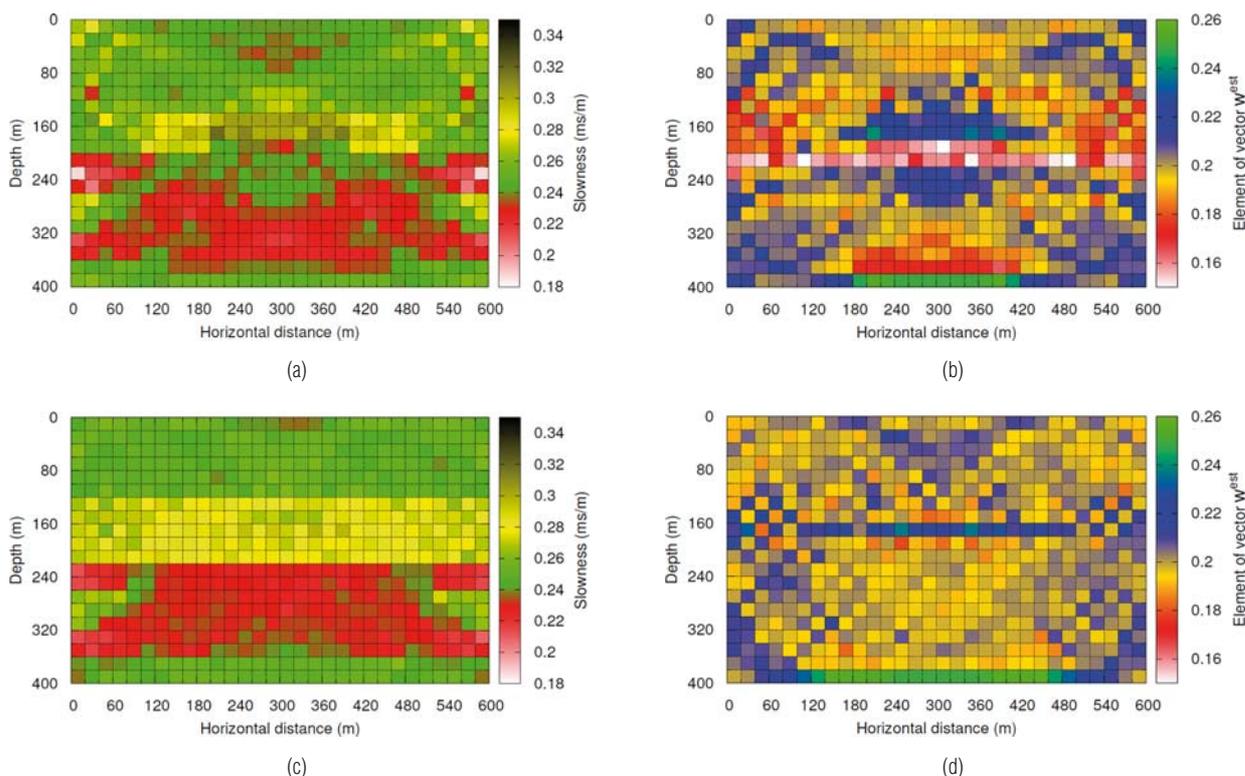


Figure 6 — Simulations for the linearized overdetermined case. The noise level is $\alpha = 0.001$ and the number of singular values used in the construction of the pseudo-inverse matrix is $n_{sv} = 250$. (a) and (c) show the estimated model parameter obtained, respectively, in the first and the fifth iterations, where the color bar indicates the P-wave slowness in ms/m. (b) and (d) show the pseudo-constant \mathbf{w}^{est} , respectively, for the first and the fifth iterations, where the color bar indicates the elements of \mathbf{w}^{est} in ms/m.

wavefront movement. A more logical way to analyze the ray trace without using the wavefront concept is through Fermat's principle. Several methods are described in the literature to determine the ray path between two points. The following steps describe the numerical algorithm proposed by Andersen & Kak (1982). Applying Fermat's principle and knowing that Euler's equation is a necessary condition for the existence of an extreme value of the integral $\int_{P_1}^{P_2} n ds$, we obtain the following differential equation for a non-homogeneous medium:

$$\frac{d}{ds} \left(n \frac{d\mathbf{r}}{ds} \right) = \nabla n, \tag{A1}$$

where $n(x, z)$ is the refraction index at position (x, z) , \mathbf{r} is the ray position vector, $d\mathbf{r}/ds$ is a vector tangent to the ray at (x, z) , ds is the length element in the ray trajectory, and $\nabla n = dn/d\mathbf{r}$ is the refraction index gradient. This differential equation is referred to as the ray equation of the radius, and its solution represents a family of rays with a shorter acoustic path for a certain regular neighborhood (where the refraction index varies smoothly)

(Andersen & Kak, 1982).

Developing the ray equation, we obtain

$$\left(\nabla n \cdot \frac{d\mathbf{r}}{ds} \right) \frac{d\mathbf{r}}{ds} + n \frac{d^2\mathbf{r}}{ds^2} = \nabla n. \tag{A2}$$

Expanding the position vector in Taylor series at the point $s + \Delta s$, where Δs is the ray increment, and considering only the first three terms, we have

$$\mathbf{r}(s + \Delta s) = \mathbf{r}(s) + \frac{d\mathbf{r}}{ds} \Delta s + \frac{1}{2} \frac{d^2\mathbf{r}}{ds^2} (\Delta s)^2. \tag{A3}$$

Isolating the curvature vector $d^2\mathbf{r}/ds^2$ in equation (A2) and substituting it into equation (A3), we obtain the following expression:

$$\mathbf{r}(s + \Delta s) = \mathbf{r}(s) + \frac{d\mathbf{r}}{ds} \Delta s + \frac{1}{2n} \left[\nabla n - \left(\nabla n \cdot \frac{d\mathbf{r}}{ds} \right) \frac{d\mathbf{r}}{ds} \right] (\Delta s)^2. \tag{A4}$$

The next point along the ray is estimated by the following

Table 1 – Simulation results with noisy synthetic data ($\alpha = 0.001$) using linearized inversion with SVD, different regularization orders and optimal parameters λ^{opt} . ε_m indicates the model error, ε_d the data error, ε_{R_m} the model resolution matrix error, ε_{R_d} the data resolution matrix error, and ε_w the pseudo-constant matrix error between iterations $k + 1$ and k .

α	λ	Iteration	Overdetermined ($n = 0$)				
			ε_m (%)	ε_d (%)	ε_{R_m} (%)	ε_{R_d} (%)	ε_w (%)
0,001	10^{+4}	0	21,617	20,119			
		1	6,227	2,438	0,112	0,112	6,122
		2	3,754	0,785	0,073	0,073	5,231
		3	2,793	0,303	0,059	0,059	5,609
		4	2,538	0,180	0,059	0,059	5,732
		5	2,420	0,155	0,073	0,073	5,818
		6	2,347	0,159	0,127	0,127	5,825
α	λ	Iteration	Overdetermined ($n = 1$)				
			ε_m (%)	ε_d (%)	ε_{R_m} (%)	ε_{R_d} (%)	ε_w (%)
0,001	10^{+4}	0	21,617	20,119			
		1	6,246	2,351	0,012	0,012	7,397
		2	3,602	0,762	0,012	0,012	4,013
		3	2,470	0,334	0,013	0,013	2,224
		4	2,120	0,288	0,008	0,008	1,840
		5	2,084	0,210	0,011	0,011	1,830
		6	2,031	0,121	0,011	0,011	1,911
		7	1,935	0,122	0,008	0,008	1,839
α	λ	Iteration	Overdetermined ($n = 2$)				
			ε_m (%)	ε_d (%)	ε_{R_m} (%)	ε_{R_d} (%)	ε_w (%)
0,001	10^{+6}	0	21,617	20,119			
		1	5,899	2,317	0,026	0,026	6,596
		2	3,443	0,680	0,012	0,012	3,490
		3	2,613	0,289	0,029	0,029	1,787
		4	2,490	0,236	0,029	0,029	2,235
		5	2,460	0,239	0,029	0,029	2,066

equations:

$$\begin{aligned}
 x_{k+1} &= x_k + \cos \alpha_k \Delta s \\
 &+ \frac{1}{2s_k} (s_{k,x} - d_k \cos \alpha_k) \Delta s^2, \\
 z_{k+1} &= z_k + \sin \alpha_k \Delta s \\
 &+ \frac{1}{2s_k} (s_{k,z} - d_k \sin \alpha_k) \Delta s^2,
 \end{aligned}
 \tag{A5}$$

where $s_{k,x}$ and $s_{k,z}$ are the slownesses in the x and z directions, respectively. d_k is defined as

$$d_k = s_{k,x} \cos \alpha_k + s_{k,z} \sin \alpha_k. \tag{A6}$$

Starting from a given initial point (x_0, z_0) that corresponds to the source position, one may obtain successive points along

the ray because the values of $\sin \alpha_k$ and $\cos \alpha_k$ are easily calculated. According to Andersen & Kak (1982), this method has some limitations as the errors caused by the discretization process or abrupt velocity transitions may be cumulative. To minimize this problem, one must adopt a grid with a sufficient resolution so that the medium is properly sampled, resulting in smoother velocity transitions. One can also smooth the velocity field and use the bilinear interpolation of the refraction index and its partial derivatives.

APPENDIX B: SINGULAR VALUE DECOMPOSITION

Consider a rectangular matrix $\mathbf{G}_{M \times N}$, with rank k , that has a singular value decomposition (SVD) as follows:

$$\mathbf{G} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \tag{B1}$$

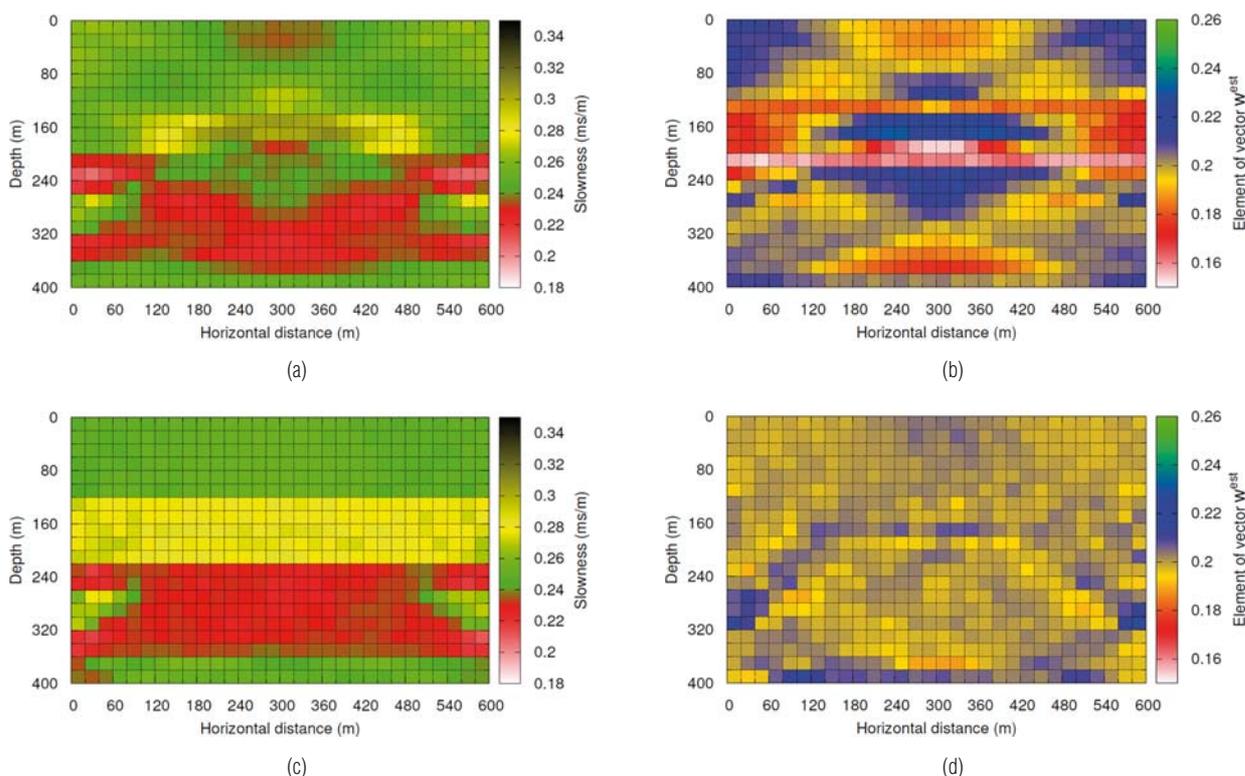


Figure 7 — Simulations in linearized inversion through SVD method, overdetermined case, with first order regularization. Regularization parameter $\lambda = 10^4$ and noise level $\alpha = 0.001$. (a) and (c) show the estimated model parameter obtained, respectively, in the first and the seventh iterations, where the color bar indicates the P-wave slowness in ms/m. (b) and (d) show the pseudo-constant \mathbf{w}^{est} , respectively, for the first and the seventh iterations, where the color bar indicates the elements of \mathbf{w}^{est} in ms/m.

such that (i) $\mathbf{V}_{N \times N}$ is the matrix containing the orthonormalized eigenvectors of $\mathbf{G}^T \mathbf{G}$; (ii) $\mathbf{\Sigma}_{M \times N}$ is the matrix containing the square roots of the eigenvalues (singular values) of \mathbf{G} , with these singular values displayed in decreasing order, that is, $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_k > 0$; and (iii) $\mathbf{U}_{M \times M}$ is the matrix containing the orthonormalized eigenvectors of $\mathbf{G} \mathbf{G}^T$.

Then, the pseudo-inverse or generalized inverse is a matrix equal to

$$\mathbf{G}^+ = \mathbf{V} \mathbf{\Sigma}^+ \mathbf{U}^T, \quad (\text{B2})$$

where $\mathbf{\Sigma}_{M \times N}^+$ is the matrix containing the reciprocal of the non-zero singular values of \mathbf{G} as follows:

$$\mathbf{\Sigma}^+ = \begin{pmatrix} \mathbf{E} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad (\text{B3})$$

and \mathbf{E} is a diagonal $k \times k$ matrix whose j -th diagonal element is $e_{ii} = (\sigma_i)^{-1}$ for $1 \leq i \leq k$.

If $\mathbf{G}_{M \times N}$ is a real matrix, the matrix $\mathbf{G}_{N \times M}^+$ will be its generalized inverse or a pseudo-inverse with the following prop-

erties (Penrose, 1955):

- (i) $\mathbf{G} \mathbf{G}^+ \mathbf{G} = \mathbf{G}$,
- (ii) $\mathbf{G}^+ \mathbf{G} \mathbf{G}^+ = \mathbf{G}^+$,
- (iii) $(\mathbf{G} \mathbf{G}^+)^T = \mathbf{G} \mathbf{G}^+$, and
- (iv) $(\mathbf{G} \mathbf{G}^+)^T = \mathbf{G}^+ \mathbf{G}$.

If those properties are satisfied, the generalized inverse or pseudo-inverse \mathbf{G}^+ will be unique.

If $\mathbf{G} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T$, then $\mathbf{G}^T = (\mathbf{U} \mathbf{\Sigma} \mathbf{V}^T)^T = \mathbf{V} \mathbf{\Sigma}^T \mathbf{U}^T$. The matrix $\mathbf{G}^T \mathbf{G}$ will be

$$\mathbf{G}^T \mathbf{G} = \mathbf{V} \mathbf{\Sigma}^T \mathbf{U}^T \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T. \quad (\text{B4})$$

If \mathbf{U} is an orthonormal matrix, that is, $\mathbf{U}^T \mathbf{U} = \mathbf{I}$, the above expression can be written as

$$\mathbf{G}^T \mathbf{G} = \mathbf{V} \mathbf{\Sigma}^T \mathbf{\Sigma} \mathbf{V}^T. \quad (\text{B5})$$

This last equation is an orthogonal transformation. It can be said that \mathbf{V} is the matrix containing the eigenvectors of $\mathbf{G}^T \mathbf{G}$

and \mathbf{V} is really orthogonal because the matrix $\mathbf{G}^T\mathbf{G}$ is always symmetric.

$\Sigma^T\Sigma$ is the matrix containing the eigenvalues of $\mathbf{G}^T\mathbf{G}$, or

singular values of \mathbf{G} . The same analysis can be done with the multiplication of \mathbf{G} by \mathbf{G}^T . In the case of square matrices of full rank, the classical inverse and pseudo-inverse are identical.

Recebido em 9 abril, 2015 / Aceito em 20 março, 2017

Received on April 9, 2015 / Accepted on March 20, 2017