Criterion for Evaluating Solutions in Traveltime Tomography

Marcelo do Prado Vieira* (PPGEOF/UFBA & INCT-GP), Amin Bassrei (CPGG/UFBA & INCT-GP)

Abstract

We present a criterion to evaluate the reliability of the velocity model obtained from traveltime tomographic inversion with Tikhonov regularization. The criterion is based on SVD filtering and on Barbieri's criterion. The proposed approach was able to improve the model in the RMS error sense, and was applied to a 2-D synthetic model based on the Dom João Field, Recôncavo Basin. Regularized inversion was iteratively performed by conjugate gradient method and regularization parameter in each step was selected by L-curve. Two levels of Gaussian noise were added to input data to simulate real conditions.

Introduction

Traveltime tomography is an imaging technique used to reconstruct velocity distribution from measurements of transmitted waves, in this case traveltimes related to paths that connect sources to receivers. Traveltime tomography has a key role in oil industry since it is a high-resolution tool capable of locating, characterizing and monitoring content evolution of hydrocarbon reservoirs (Lazaratos & Marion, 1997; Messud et al., 2017). Since the wave path, that is, the ray, is a function of the slowness (velocity reciprocal), the relation between traveltime and slowness is nonlinear. Also, field data is usually affected by noise and subsurface data coverage is usually low. Thus, solutions are very sensitive to small perturbations in the data, and tomographic inversion is said to be ill-conditioned (Tarantola, 2005).

Such problems can be numerically stabilized by regularization, which introduces some expectation about the solution (Menke, 2018). A widely used is the Tikhonov regularization. By applying a derivative operator to the solution, the objective function to be minimized includes both the data prediction error and the module of the smoothed solution in Euclidian length, weighted by an unknown regularization factor.

The idea is to iteratively solve the inverse problem with the appropriate parameter at each step. Several techniques to choose or to determine this parameter have been developed (Golub et al., 1979; Morozov, 1984; Hansen, 1992). And some of these approaches to select the regularization factor were applied to traveltime tomography (Bube & Langan, 1994; Yao & Roberts, 1999; Soupios et al., 2001).

Regularization produces a stable reduced-norm solution, but it does not solve the original problem. Assuming that it is not possible to obtain additional a priori information about the solution, at least one can evaluate particularities of the solution from the kernel matrix. The most common methods use the covariance matrix or data-resolution and model-resolution matrices (Menke, 2018). These matrices were applied to seismic tomography problems by Zhang & Thurber (2007) and Xia et al. (2008).

In a similar way, Barbieri (1974) suggested a heuristic criterion to identify artifacts introduced by the inversion algorithm. The criterion solves the inverse problem for a complementary model defined in such a way that, added to the original, it produces a solution with a constant value, if the inverse problem is exact. In terms of information, solution and the complementary solution are equivalent. Since the tomographic problem is, in practice, usually underdetermined, infinite solutions exist. Based on some norm or procedure, we select only one solution from the whole space of solutions. If any artifact is not present in two distinct solutions, one can presume it does not belong to the space of solutions (Gordon, 1974). By identifying the location and size of these artifacts, one can exploit this information to upgrade the solution.

Theory

Ray tracing

The ray represents the propagating wave energy path in a medium. In traveltime tomography, the ray is the propagating seismic-wave path in an acoustic or elastic medium. In an isotropic and nonhomogeneous medium, the acoustic path length along the ray between a source located at point S and a receiver at point R is given by

$$I = \int_S^R n \, dl,$$

where n is the refraction index distribution and dl is an infinitesimal arc length along the ray (Born & Wolf, 1999). Since Fermat's principle - which establishes that wave energy propagates along a path for which the traveltime has a minimum value - determines the path taken by a ray, one can apply the calculus of variations in I to derive Euler Equation, also known as ray equation in the context of geometrical optics,

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

whose solution, within a certain regular neighborhood, is the path taken by the ray with the shortest acoustic length in an isotropic and nonhomogeneous medium. The vector $\mathbf{r}$ is the position of any point along the ray, ∇n is the

gradient of n and $d\mathbf{r}/dl$ is the unit vector tangent to that point.

We consider the propagation of a seismic wave in terms of a ray, which is a high frequency approximation of the wave phenomenon, i.e., it considers negligible wavelength compared to scattering points. Ray equation-based ray tracing algorithms are useful as long as the refraction index varies smoothly.

The ray tracing algorithm used in the present paper is based on the algorithm of Andersen & Kak (1982), which is capable of modeling realistic medium in a geological sense. Their algorithm is based on the second-order Taylor approximation of the position of a point along the ray, $\mathbf{r}(l + \Delta l)$, where the second derivative is obtained by the Euler equation:

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

For a 2-D medium, the algorithm is given by the decomposition of $\mathbf{r}(l + \Delta l)$ into the horizontal (x-axis) and vertical (z-axis) directions:

$$x_{k+1} = x_k + \cos\alpha_k\,\Delta l + \frac{1}{2s_k}\left(s_{k,x} - d_k\cos\alpha_k\right)\Delta l^2,$$

and

$$z_{k+1} = z_k + \sin\alpha_k\,\Delta l + \frac{1}{2s_k}\left(s_{k,z} - d_k\sin\alpha_k\right)\Delta l^2.$$

s_K is the slowness at point $P_k = (x_k, z_k)$; $s_{k,x}$ and $s_{k,z}$ are the partial derivatives of s_k with respect to x and z, respectively; α_k is the angle between the x-axis and the tangent to the ray at P_k; $d_k = s_{k,x}\cos\alpha_k + s_{k,z}\sin\alpha_k$; Δl is the distance between P_k and the next point P_{k+1}, and $s = n/c$. c is the seismic velocity in a reference medium.

Given a slowness distribution, a point in the medium and an angle of incidence the ray path can be traced. Ray tracing is a critical part of tomographic reconstructions due to its high computational cost. Also, finding the angle of incidence that links source to receiver takes some time for complex models.

Forward modeling

The traveltime related to i-th ray, t_i, is given by the line integral of slowness s along the ray path (Scales, 1987). Let $t_i = g_i(s)$. One can linearize it by the first-order Taylor approximation around an initial slowness model s_0. For a set of M rays within the medium parameterized by a grid composed of N squares, the linearized problem is reduced to

$$\Delta\mathbf{t} = \mathbf{G}\Delta\boldsymbol{s},$$

where

$$G_{ij} = \left.\frac{\partial g_i}{\partial s_j}\right|_{\mathbf{s}=\mathbf{s}_0},$$

for $i = 1, \dots, M$ and $j = 1, \dots, N$. The elements of the sensitivity matrix $\mathbf{G}$, obtained by ray tracing algorithm, represent the path length taken by the i-th ray in the j-th square of the initial slowness model.

Thus, the nonlinear problem is replaced by a system of linear equations between perturbations both in traveltimes, $\Delta\mathbf{t}$, and in the initial slowness model, $\Delta\mathbf{s}$.

Regularized inversion

The exact solution $\Delta\mathbf{s} = \mathbf{G}^{-1}\Delta\mathbf{t}$ cannot be obtained because $\mathbf{G}$ is usually not full rank, even when it is a square matrix. Tikhonov regularization stabilizes the system numerically by applying a derivative operator $\mathbf{D}_n$ to the solution, where n is the order of the derivative. One can choose the first-order regularization if the model is expected to be flat; the second-order, if it is expected to be smooth, i.e., not rough (Menke, 2018).

The objective function to be minimized is the sum of the amount of data prediction error and the amount of smoothed solution in Euclidian length, which leads to

$$[\mathbf{G}^{\mathrm{T}}\mathbf{G} + \lambda\mathbf{D}_{\mathrm{n}}^{\mathrm{T}}\mathbf{D}_{\mathrm{n}}]\Delta\mathbf{s} = \mathbf{G}^{\mathrm{T}}\Delta\mathbf{t},$$

where λ is the regularization factor that controls data error and model smoothing. Greater λ produces smoother solutions and large data errors; the opposite occurs for smaller λ.

Determining an appropriate value for λ is crucial since it influences the solution. We use L-curve (Hansen, 1992), a graphical tool that evaluate the log-log plot of ordered pairs $(\|\Delta\mathbf{t} - \mathbf{G}\Delta\mathbf{s}(\lambda)\|_2, \|\mathbf{D}_{\mathrm{n}}\Delta\mathbf{s}(\lambda)\|_2)$ in the Euclidian plane for a pre-set λ values. Its name derives from its shape (see Figure 1), and it is expected that the appropriate λ corresponds to the point of maximum curvature in L-curve, since it indicates a balance between smoothness of solution and data fitting.

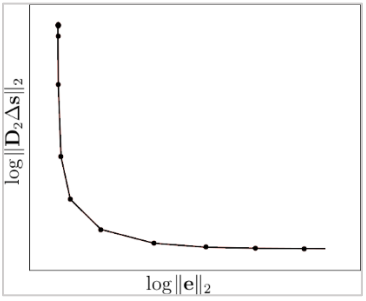

Figure 1 – An example of the parametric L-curve. The factor λ increases from the top of curve to the right and the appropriate value corresponds to the point of maximum curvature.

Method

Let $\mathbf{s}^{\mathrm{true}}$ be the exact solution for the linear system $\mathbf{Gs} = \mathbf{t}$, for a known data vector $\mathbf{t}$. Let $\mathbf{w}$ be a vector of same dimension as $\mathbf{s}$, with elements $w_j = \omega$, where ω is constant

for all j. Let $\mathbf{s}^{\mathrm{c}}$ be a complementary solution of $\mathbf{s}^{\mathrm{true}}$ such that

$$\mathbf{s}^{\mathrm{true}} + \mathbf{s}^{\mathrm{c}} = \mathbf{w}.$$

It is intuitive to obtain the complementary system as

$$\mathbf{G}\mathbf{s}^{\mathrm{c}} = \mathbf{G}\mathbf{w} - \mathbf{t},$$

where the right-hand side is called complementary data, $\mathbf{t}^{\mathrm{c}}$, and it is considered to be known. If we sum the complementary estimated solution, $\mathbf{s}^{\mathrm{c,est}}$, with the original estimated solution, $\mathbf{s}^{\mathrm{est}}$, both obtained by the same inversion algorithm, we obtain the pseudo-constant vector

$$\mathbf{w}^{\mathrm{est}} = \mathbf{s}^{\mathrm{est}} + \mathbf{s}^{\mathrm{c,est}}.$$

For an exact inversion, $\mathbf{w}^{\mathrm{est}} = \mathbf{w}$. Otherwise, $w_j \neq \omega$ indicates that the j-th element of the model is less reliable. Such a conclusion arises from the expectation that obtaining the complementary solution would not provide more knowledge than the original one since data and its complementary are equivalent in terms of information (Gordon, 1974).

As traveltime tomography is an ill-conditioned problem, the related complementary system will be as well and will demand a kind of regularization, for instance, Tikhonov regularization (Bejarano & Bassrei, 2016). The complementary regularized system is given by

$$\left[\mathbf{G}^{\mathrm{T}}\mathbf{G} + \lambda \mathbf{D}_{\mathrm{n}}{}^{\mathrm{T}}\mathbf{D}_{\mathrm{n}}\right]\Delta\mathbf{s}^{\mathrm{c,est}} = \mathbf{G}^{\mathrm{T}}\Delta\mathbf{t}^{\mathrm{c}}.$$

Let $\mathbf{r}$ be the unknown residuals defined so that

$$\mathbf{s}^{\mathrm{true}} = \mathbf{r} + \mathbf{s}^{\mathrm{est}}.$$

One can estimate $\mathbf{r}$ as follows. Let

$$\mathbf{p} = \mathbf{w} - \mathbf{w}^{\mathrm{est}}$$

be a pseudo-null vector since it is expected that nonzero elements represent the medium squares where the inversion behaved unexpectedly. Assuming that $\mathbf{s}^{\mathrm{c,est}} \approx \mathbf{s}^{\mathrm{c}}$, then $\mathbf{p} \approx \mathbf{r}$, and, therefore,

$$\mathbf{s}^{\mathrm{imp}} = \mathbf{p} + \mathbf{s}^{\mathrm{est}} \approx \mathbf{s}^{\mathrm{true}},$$

where $\mathbf{s}^{\mathrm{imp}}$ is the improved model by the proposed criterion.

One can represent the N-dimension vector $\mathbf{p}$ as the pseudo-null $N_z \times N_x$ matrix $\mathbf{P}$, i.e., each element of $\mathbf{P}$ is related to a square in the gridded medium. If one identifies each square of the medium by the ordered pair (p, q) in the Euclidian plane, then the element $P_{q,p}$ of the matrix belongs to square (p, q) and it is equivalent to the j-th element of $\mathbf{p}$, where $j = (q-1)N_x + p$, for $p = 1, \ldots, N_x$ and $q = 1, \ldots, N_z$.

Starting from the reasonable conjecture that regions of the medium where inversion was unsuccessful must be randomly located, it is possible to enhance such elements of $\mathbf{P}$ by attenuating every coherent signal. We suggest the power method (Golub & Loan, 1996) for it, by suppressing the first eigenimages since the corresponding singular values are much larger than the others, i.e., they retain laterally coherent information (Freire & Ulrych, 1988).

Results

We used a 2-D synthetic model (Figure 2) based on a reservoir from Dom João Field, Recôncavo Basin, in a crosswell acquisition geometry with 141 sources and 140 receivers evenly spaced, totalizing 19740 rays. Figure 3 shows traveltime diagram. The model covers an 215 × 410 m area and was parameterized by 3526 blocks of dimension 5 × 5 m, i.e., 43 squares horizontally and 82 vertically.

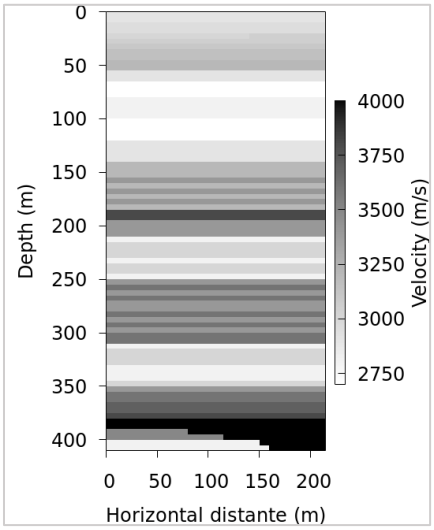

Figure 2 - True velocity model with 3526 blocks, based on real situation from Dom João Field, Recôncavo Basin.

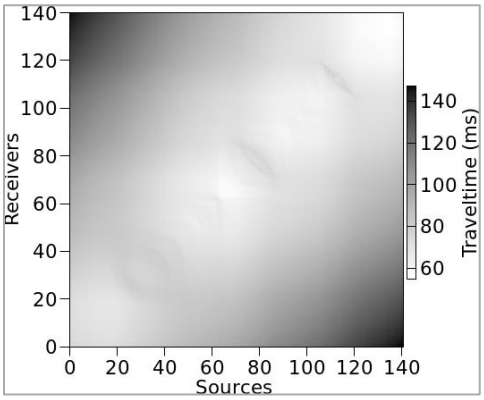

Figure 3 - Graphical representation of the traveltime of each one of the 19740 rays. Smaller numbers for sources and receivers correspond to smaller depths.

One must establish some measure of quality for any recovered model $\mathbf{s}$ in respect to true model. We use estimator

$$\varepsilon = 100\% \, \frac{\|\mathbf{s}^{\mathrm{true}} - \mathbf{s}\|_2}{\|\mathbf{s}^{\mathrm{true}}\|_2},$$

i.e., percentage RMS error of $\mathbf{s}$ with respect to $\mathbf{s}^{\mathrm{true}}$.

To better represent a real data condition, Gaussian noise was added to data. We denoted it as μ, the percentage RMS error of data with respect to noise-free data, and we choose two different levels: $\mu = 1\%$ and $\mu = 5\%$.

The inversion was performed by the conjugate gradient method optimized for sparse matrices (Scales, 1987), up to 8 iterations and it was regularized by the second-order derivative operator in both horizontal and vertical directions.

The regularization factor was selected by two approaches. The first one chose each iteration model as the model with minimum ε. The second approach selects the optimum regularization factor by the L-curve (Hansen, 1992). Since true model is unknown for real data, the former is a validation aproach for the latter. Table 1 shows both results. For all simulations, the processing time until convergence was up to 240 s and required no more than 5 iterations. The lowest number of required iterations occurred for $\mu = 5\%$, which is expected since very noisy data demand greater regularization.

μ (%)	Inversion		Proposed criteria
	Minimum ε	L-curve	
1	3.78	4.69	3.93
5	4.72	5.42	4.92

Table 1 – Estimator ε (in %) for the obtained models. The second and third columns are for inversion algorithm. The last column is for improved models by the proposed criterion. The decrease of ε for both noise levels added to data is significant even when compared to the reference model (minimum ε). These results are visually endorsed in Figure 6.

Figure 4 shows the ratio of consecutive singular values of the pseudo-null matrices. For the case of data with high noise ($\mu = 5\%$), dominance was verified and the corresponding eigenimage was suppressed. The procedure of improving the model was performed with the entire content of the pseudo-null matrix for the case $\mu = 1\%$.

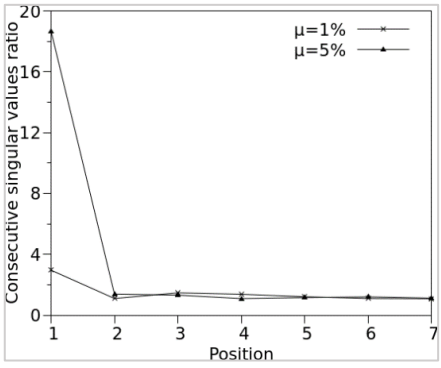

Figure 4 - Evolution of the first seven ratios between consecutive singular values of pseudo-null matrix **P**. Only for $\mu = 5\%$ there is high dominant coherence, where the first singular value is 18 times greater than the second one, and the corresponding eigenimage was suppressed.

The residual and pseudo-null matrices are placed side by side in Figure 5. For the case $\mu = 1\%$, **P** clearly maps almost all regions of malfunctions in the inversion algorithm. For $\mu = 5\%$, one can notice that some linear pattern in **R** seem to hide some random features that **P** exhibits. Figure 6 confirms that as it shows resolution enhance for both noise levels.

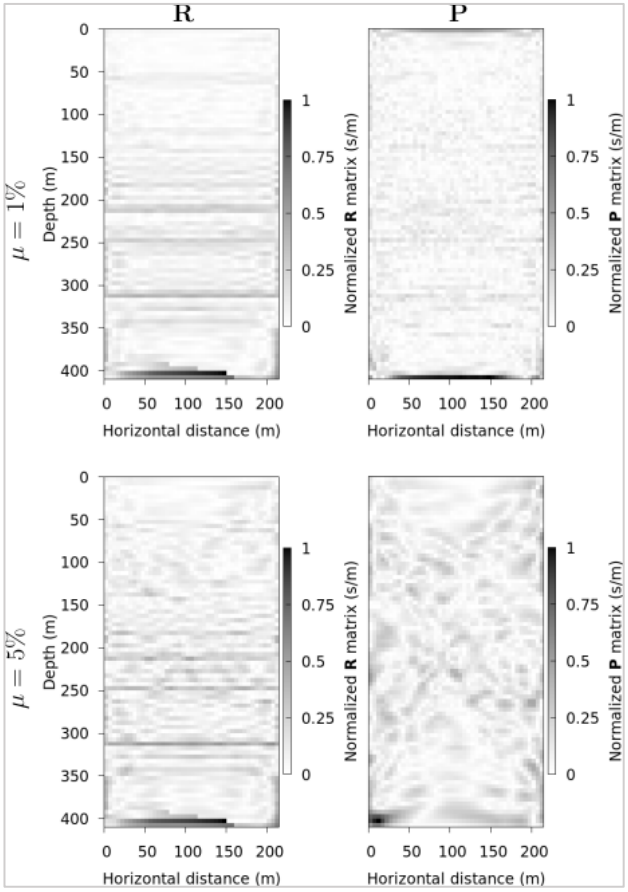

Figure 5 - Residual (**R**) and pseudo-null matrices (**P**) normalized by their maximum value and in absolute (for display purposes only). **R** is the matrix representation of residual vector **r** in the same way as it was done for **P**. Even for very noisy data ($\mu = 5\%$) **P** is capable of identifying regions of malfunction in the inversion algorithm that was hidden in horizontal pattern in **R**.

The regularization parameter required in the proposed criterion is very close to those ones selected in the inversion for $\mu = 1\%$, but greater regularization was required for $\mu = 5\%$. As the latter is an unlikely noise level, one can state that the proposed criterion maintains the inversion regularization parameter.

The last column of Table 1 shows estimator ε for the proposed criterion, and it complements the results shown in Figures 5 and 6 as it provides a good measure of the improvement in the model. Processing time for the proposed criterion was up to 30 s.

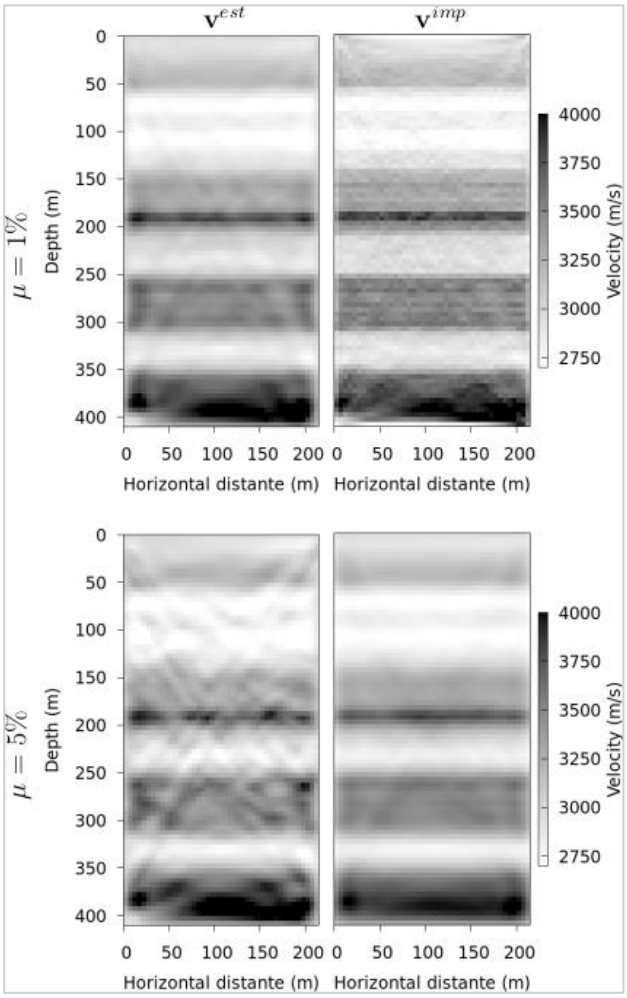

Figure 6 – Estimated velocity distribution in the inversion algorithm ($\boldsymbol{v}^{est}$) and the velocity distribution improved by the proposed criterion ($\boldsymbol{v}^{imp}$). For both noise levels the proposed criterion produced a model with higher resolution.

Conclusions

The proposed algorithm is a computational low-cost, easy-to-implement method effective for both mapping regularized tomographic inversion algorithms malfunctions and improving the solution in the RMS error sense, even in the presence of a high level of noise in the data. If there are dominant singular values, the corresponding eigenimages must be suppressed from the pseudo-null matrix with the purpose to enhance regularization features and to identify algorithm malfunctions.

Acknowledgments

The authors thank CNPq for supporting the Instituto Nacional de Ciência e Tecnologia de Geofísica do Petróleo (INCT-GP). M. Vieira thanks FAPESB for a PhD scholarship. A. Bassrei thanks CNPq for a research fellowship.

References

ANDERSEN, A. H.; KAK, A. C. (1982) Digital ray tracing in two-dimensional refractive fields, Journal of Acoustical Society of America, v. 72, 1593-1606.

BARBIERI, M. (1974) A criterion to evaluate three dimensional reconstructions from projections of unknown structures, Journal of Theoretical Biology, 48, 451-467.

BEJARANO, S. L.; BASSREI, A. (2016) Critique of solutions in linearized inverse problems: numerical experiments in traveltime tomography, Brazilian Journal of Geophysics, 34, 495-508.

BORN, M.; WOLF, E. (1999) Principles of Optics, 7th ed., Cambridge University Press, Cambridge.

BUBE, K. P.; LANGAN, R. T. (1994) A continuation approach to regularization for traveltime tomography, 64th Annual International SEG Meeting, 980-983.

FREIRE, S.; ULRYCH, T. (1988) Application of singular value decomposition to vertical seismic profiling, Geophysics, 53, 778-785.

GOLUB, G.; HEATH, M.; WAHBA, G. (1979) Generalized cross-validation as a method for choosing a good ridge parameter, Technometrics, 21, 215-223.

GOLUB, G.; LOAN, C. VAN (1996) Matrix Computations, Johns Hopkins University Press, Baltimore.

GORDON, R. (1974) A tutorial on ART (Algebraic Reconstruction Techniques), IEEE Transactions on Nuclear Science, 21, n. 3, 78-93.

HANSEN, P. C. (1992) Analysis of discrete ill-posed problems by means of the L-curve, Society for Industrial and Applied Mathematics Review, 34, 561-580.

LAZARATOS, S.; MARION, B. (1997) Crosswell seismic imaging of reservoir changes caused by CO2 injection, The Leading Edge, 16, 1300-1306.

MENKE, W. (2018) Geophysical Data Analysis: Discrete Inverse Theory, 4th ed., Academic Press, London.

MESSUD, J.; REINIER, M.; PRIGENT, H.; GUILLAUME, P.; COLÉOU, T.; MASCLET, S. (2017) Extracting seismic uncertainties from tomographic velocity inversion and their use in reservoir risk analysis, The Leading Edge, 36, 127-132.

MOROZOV, V. (1984) Methods for Solving Incorrectly Posed Problems, Springer, New York.

SCALES, J. A. (1987) Tomographic inversion via the conjugate gradient method, Geophysics, 52, 179-185.

SOUPIOS, P.; PAPAZACHOS, C.; JUHLIN, C.; TSOKAS, G. (2001) Nonlinear 3-D traveltime inversion of crosshole data with an application in the area of the Middle Ural Mountains, Geophysics, 66, 627–636.

TARANTOLA, A. (2005) Inverse Problem Theory and Methods for Model Parameter Estimation, SIAM, Philadelphia.

XIA, J.; MILLER, R.; XU, Y. (2008) Data resolution matrix and model resolution matrix of Rayleigh-wave inversion using a damped least-square method, Pure and Applied Geophysics, 165, 1227–1248.

YAO, Z.; ROBERTS, R. (1999) A practical regularization for seismic tomography, Geophysical Journal International, 138, 293–299.

ZHANG, H.; THURBER, C. (2007) Estimating the model resolution matrix for large seismic tomography problems based on Lanczos bidiagonalization with partial reorthogonalization, Geophysical Journal International, 170, 337-345.