Parameters Estimation of VTI Media, using Hybrid Genetic Algorithms

Samuel G. Huamán Bustamante, ICA/DEE/PUC-Rio, Brasil
Marco Aurélio C. Pacheco ICA/DEE/PUC-Rio, Brasil

This paper was prepared for presentation during the 11th International Congress of the Brazilian Geophysical Society held in Salvador, Brazil, August 24-28, 2009.

Abstract

At this work is proposed a sequential method based on hybrid genetic algorithms (HGA) composed by the genetic algorithm and the Nelder mead simplex algorithm to estimate parameters of a model from seismic reflexion. The model represents layers into underground. In this way, the parameters of a layer form a chromosome of the HGA and the set of layers forms a seismic model. In the inversion process of multiple layers, the representative velocities are the main problem to get the layer velocities. Consequently, the proposed method estimates sequentially the seismic parameters for solving this problem. The experiments use seismic models with horizontal layers with isotropy and vertical transverse isotropy (VTI). The proposed method is compared to the method of Medeiros (2005); that considers a chromosome for all seismic models. The results, when is applied the proposed method, are better than refereed method. The data with the travel times are synthetic and without noise. Other advantage is the reduction of the runtime of estimation algorithms; that allows processing of more quantity of data.

Introduction

Seismic methods are geophysics surveys, commonly used in engineering. The majority of studies about propagation of seismic wave are analogy of optic laws. There are two techniques: Seismic Reflection and Seismic Refraction showed in Dourado, (2001).

By the time, more complicated and exact equations were developed to estimate parameters and to describe forms undersurface, although many equations are simplifications to reduce time of computation, there are possibilities to develop new methods and to adapt techniques reducing time of computation and increasing accuracy.

In the bibliography can be found works that applie methods to estimate geophysics parameters with different techniques as in Margrave, (2003). Some of these methods use the technique of genetic algorithm (GA), but without great success (Medeiros, 2005).

The proposal of this work (Huamán, 2008) is to develop a sequential method for parameters estimation without assuming linear equations as in Gouveia, (1994) and using hybrid techniques to increase accuracy of estimated values.

The article is organized as follow: In the second part are explaining terms and definitions of data acquisition, preprocessing procedures and equations that were finding in the bibliographic references to analyze the travel times of seismic P-waves reflected on an interface between two layers undersurface. Below are described the part and parameters of GA and HGA techniques used to solve the problem of seismic parameter estimation. In the Results part is explained how is assumed seismic models of reference, and two cases are studied. Finally in the conclusions is mentioned the improvements with method and technique proposed in this article and the criterion to take before or when is running the estimation process.

Seismic Reflexion

In the seismic Reflection, a seismic section can be built directly on the seismogram (a set of seismic signal from aligned receptors), it represents a vertical section, but to do the quantitative analysis the section is represented by layers with characteristic modeled by equations. Before to execute this procedure, is necessary to do the next measurements and corrections.

Review of the available equations: consider it is a simple horizontal layer with low anisotropy and arbitrary symmetry (triclinic). Mensch and Rasolofosaon (1997) used an approximate disturbance of the first order of the elastic constant of the medium. They showed that the approximate phase of P-wave velocity is:

$$V^{(P)}(\theta,\lambda) = V_v\,(1+\sum_i T_i^{(P)}(\theta,\lambda)\,\alpha_i^{(P)})\,. \qquad (1)$$

Each term of the sum is a product of trigonometric function $T_i^{(P)}$ of the incidence angle (θ) and the azimuth (λ), and the factor weights $\alpha_i^{(P)}$, which is a function of the parameter of disturbance elastic. The use of the set of coefficients of $\alpha_i^{(P)}$ is according to the medium considered. The means analyzed in the bibliography are four: Triclinic, Monoclinic, Orthorhombic and transversally isotropic according Bazelaire, Bazelaire and Perroud, (1995).

Mensch and Rasolofosaon, (1997) also established that in the medium of low anisotropy, showed at the Figure 1, the P-wave velocity with arbitrary orientation of the unit ray vector n(θ, λ) is also given by (1). Also the expression of velocity can be used to obtain the travel time of the P wave in the medium considered as a function of x and offset the azimuth lambda:

$$t(x,\lambda) = \frac{2\left(\sqrt{(x/2)^2 + \left((V_v\ t_o)/2\right)^2}\right)}{V_v\,(1+\sum_i T_i^{(P)}(\theta,\lambda)\,\alpha_i^{(P)})}\,. \qquad (2)$$

The equation (2) is called the Exact Equations and it has ten (10) unknown parameters among them, V_v and t_o.

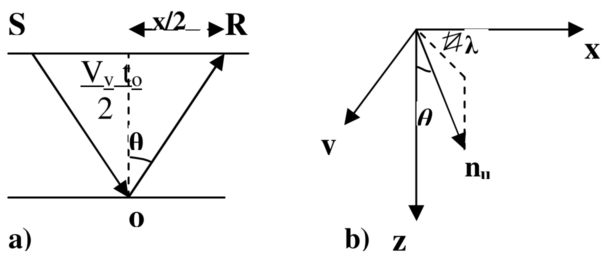

Figure 1. (a) Path of the reflected ray in the interface of the horizontal layer. (b) System of coordinate axis where n is the unit propagation vector with inclination theta and azimuth lambda.

Alkhalifah & Tsvankin (1995) proposed a model for the propagation of elastic waves in media VTI (transversely isotropic with a vertical symmetry axis). According to them, the events related to this type of medium can be approximate by a hyperbolic part, in which everything is happening as an isotropic medium, and a portion not hyperbolic. The model is represented by the equation below:

$$t(x)^2 = t_0^2 + \frac{x^2}{V_{rms}^{\ 2}} - \frac{2\eta x^4}{V_{rms}^{\ 2}\left(t_0^2 V_{rms}^{\ 2} + (1+2\eta)x^2\right)}, \quad (3)$$

where t_0 é o tempo zero-offset; $t(x)$ is the estimated time of the seismic event; x is a position of receptor; V_{rms} is the radix mean square velocity, at vertical directions; η is a parameter that measure the degree of anisotropy medium

The parameter η is a function of the Thomsen parameters δ and ε classic indicators of anisotropy and for VTI media can be calculated by (4):

$$\eta = \frac{\varepsilon - \delta}{1 + 2\delta} \quad (4)$$

The equation (3) is of pseudo-fourth degree, in which the third term is not hyperbolic, and that for η=0, the not hyperbole term is null and it represents an isotropic medium. Various studies show that if the anisotropy VTI is not taken into account in the analysis of velocity and therefore the correct pre-stacked, the final results of the seismic interpretation can be mistaken (Silva, 1995).

First deals with event on an interface and later this was generalized in case of more interfaces (or layers).

The V_{rms} can be calculated using (5), which is a function of the velocities of layers and zero-offset times of each layer above the interface analyzed. Some research works take as representative parameter the mean velocity, however V_{rms} is more used.

$$V^2_{rms} = \frac{\sum_{i=1}^{N} V_i^{\ 2}\ t_{oi}}{\sum_{i=1}^{N} t_{oi}}, \quad (5)$$

where V_{rms} is the mean square velocity at the vertical direction $V_{rms} = f\ (V_i,\ h_i)$; N is the number of layer considered; V_i, velocity in a 'i' layer; t_{oi}, is the zero offset time for 'i' layer.

The zero offset time of the layer N is computed with (6).

$$t_{oN} = \sum_{i}^{N} t_{oi} = 2\, h_1/V_1 + 2\, h_2/V_2 + \ldots\ldots\ 2\, h_N/V_N\ , \quad (6)$$

Where h_i is a thicknesses of layer 'i' and V_i is the velocity of layer 'i'.

The fact, which V_{rms} is used and not V_i (layer velocity) is because the wave travels by different mediums (layers) until arrive to receptor, it does that a layer can represent a set of layers with velocity equal to average of all velocities of the set.

Hybrid Genetic Algorithm

Genetic algorithm (GA) is a technique of search, parallel optimization and highly adaptive, based on the principle of natural selection comes from the evolutionary theory of Charles Darwin, associated with the notion of genetic reproduction studied in Goldberg, (1989). The genetic code is the identity of an individual, distinguishing it from all others, and is organized into chromosomes. According to the theory of Darwin, the principle of selection favors those individuals most able with greater longevity and a consequent higher probability of reproduction. With more offspring, these individuals are more likely to perpetuate their genetic codes in the next few generations. These are the principles that form the basis for the construction of computational algorithms that seek a better solution for a given problem through evolution in populations of solutions codified through artificial chromosomes.

One of the great advantages of GAs is that it is not necessary to describe *how* to find a good solution just what *is* a good solution. The actual evaluation of the individual serves as a compass indicating the direction in which they should be located the best solutions.

The following components characterize GA: The problem to optimize, Representation of the problem solutions, Decode chromosome, Evaluation, Selection, Genetic operator, Initialization of the population.

Each of these components must be adequate to solve the problem. The next section shows how these were adapted to the problem to estimate parameters.

Hybrid algorithms exploit the good properties of different methods by applying them to problems they can efficiently solve. For example, search is efficient when the problem has many solutions, while inference is efficient in proving unsatisfiability of overconstrained problems. That hybridizing is one possible way to build a competent GA that solves hard problems quickly, reliably and accurately without the need for any forms of human intervention (Gouveia, 1994; Wen and Lee, 2006).

Description of HGA to Estimate Parameter of VTI Media

The programs were developed in Matlab language using the Optimization Toolbox with genetic algorithm

(Mathworks, 2007). This allows easy edition, integration of functions and the possibility to combine GAs with other minimization algorithms.

The problem solutions are represented by real vector, where a chromosome, of three genes, represents properties of a layer and the genes are: the thickness of the (h), P-wave velocity (V) and anisotropic indicator (n), according to Table 1. In the case of several layers the general solution is the set of solutions of every layer.

Table 1. Structure of general solution.

General Solution			
Chromosome Structure	**Gen1**	**Gen2**	**Gen3**
1rt Layer	h_1	V_1	n_1
2nd Layer	h_2	V_2	n_2
3rd Layer	h_3	V_3	n_3

The value of every gene is real and limited by the empirical knowledge: Thickness h_i is in the interval of 0 to 700 meters; V_i (P-wave) is in the interval of 1500 m/s to 6000 m/s; n_i is in the interval of 0 to 1, without units.

These intervals are according to location and extension of seismic survey. They do not require any additional restrictions for that an individual can be valid.

The individuals are evaluated by the Error mean square (E_{rms}) of the travel time of the P wave expressed in (7). The evaluation function uses synthetic travel time as reference (or measurements) that is calculated with (3) following an ideal seismic model. This simplification is for reducing a great time of computation of synthetic seismograms and reducing error generated by simulator. Only the significant events of arrived times are compared as in Medeiros, (2005). The preview step is considered as a part of pre-processing of signals.

$$E_{rms\ i} = \sqrt{\frac{1}{j}\sum_{s=1}^{j}\left(\frac{t_{r\,is} - t_{is}}{t_{r\,is}}\right)^2}\ , \qquad (7)$$

Where t_{ris} is the travel time of reference and t_{is} is the travel time computed with evolved parameter. The superior limit 'j' of the sum is the number of receptors.

For solving an inversion problem, the equation used in the evaluation function to estimate parameter must be same to the equation used to produce data available. But if data is a set of measurements, then the equation used to represent the phenomenon, must get nearest values of the real values. In other case the estimation error of parameters may be great.

The selection of genitors (individuals than the operators transform in others) is by the roulette and selection stochastic uniform, with more probability of selection for individuals with more fitness (with less error).

Two operators are used: uniform mutation and arithmetic crossover Michalevicz, (1996).

The population was initialized with arbitrary genes for individuals. Random values were given for all thicknesses and anisotropy indicators and indicator. The velocities take values around 1500 m/s. This is based in knowledge of geological information, where the minimum P-wave velocity is in the water.

In this paper it is used a hybrid algorithm from optimizations toolboxes of Matlab software. A GA that does global search composes it and by subroutine called *fminsearch* based in The Nelder Mead simplex algorithm to local search; it is an enormously popular direct search method for multidimensional unconstrained minimization studied in Lagariasy, Reedsz, Wrightx and Wright, (1998).

The GA searches the optimum values (of exact values) for the global minimum value of the evaluation function, but if it is required accuracy of values, the *fminsearch* is used to find these values with lower grade of tolerance controlled by algorithm.

Results

The use of simulators as in margrave, (2003) must be considered when the equations are known or errors produced can be quantified. Then to avoid uncontrollable errors was used (3), to compute the synthetic references and fitness values.

The travel times, of the maximum amplitudes of p waves known as the more relevant events, were saved into a matrix as the Table 2.

Table 2. Matrix of travel times of reflected p-wave on three interfaces. The travel times are the times of large the amplitudes than are called as events.

		Events for seismic trace			
		Event 1	**Event 2**	**...**	**Event *i***
Seismic Shooting	**R_1**	$t_{r\,11}$	$t_{r\,21}$	...	$t_{r\,i1}$
	R_2	$t_{r\,12}$	$t_{r\,22}$	...	$t_{r\,i2}$
	R_3	$t_{r\,13}$	$t_{r\,23}$	...	$t_{r\,i3}$
	R_4	$t_{r\,14}$	$t_{r\,24}$	...	$t_{r\,i4}$
	...	...	...	...	...
	R_j	$t_{r\,1j}$	$t_{r\,2j}$	...	$t_{r\,ij}$

A regular seismic model is represented in the Figure 2, where the increment of the velocity layers was constant in the deep to differentiate layers during the test.

The number of receptors (RN) is 36, located each 20 meters in a line on the surface. The first receptor (R1) is to 20 meter of the source (S) and the last (RN) is to 720 m of the source.

The rays, in the Figure 2, represent the trajectory of P-wave that travel from source (S), reflected and refracted in each interface and when they are reflected to the surface, the receptors can record.

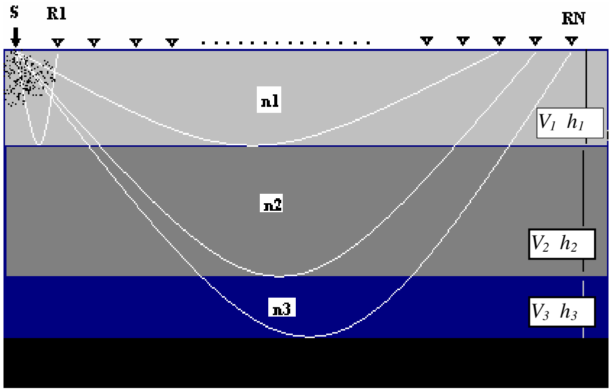

Figure 2. Distribution of three horizontal layers for VTI media and parameters considered.

The structure of the chromosome with seismic parameters was selected according to simplify the parameters estimation that solve a system of equations over determined. In other words, it can be seen as optimization problem.

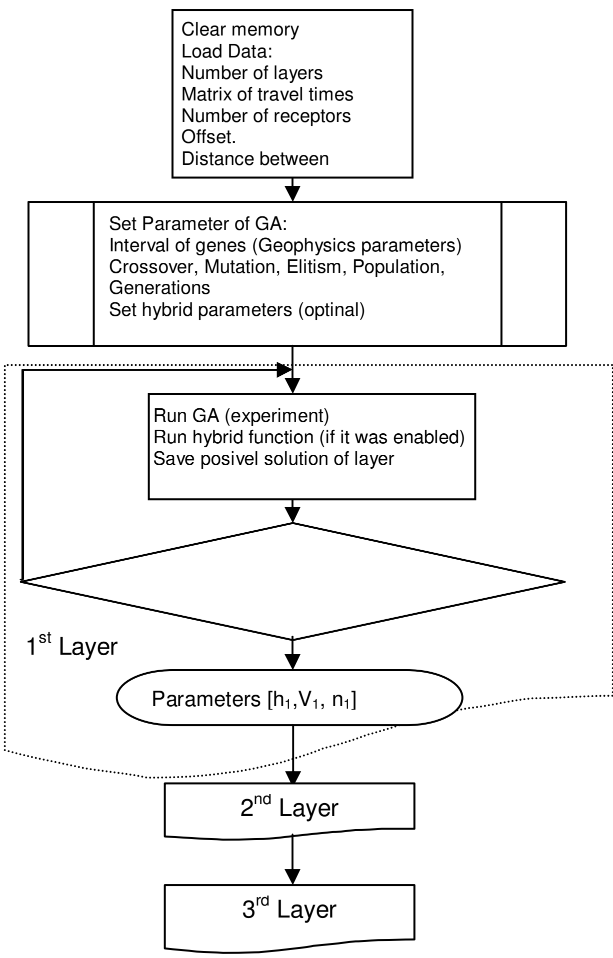

Figure 3. Blocks Diagram of program to estimate seismic parameter.

The first step estimate the parameter of the first layer, and sequentially the other layers using the solutions of previous layers used as constants, as it is showed in the Figure 3. In this way P-waves change velocity and direction when they cross different mediums.

In the case that the structure of chromosome has the parameters of all layers of seismic model into genes (Medeiros, 2005), the convergence is slower to find optimum values because the velocities of layer calculated for each layer are in a chromosome. It allows satisfying the equation of the velocity RMS, showed in (5) for the layers in the deep, with a great number of combinations of velocities. Other problem that can be seen is the difficult to extend the method for more parameters, because the method needs to be tuned for each parameter added.

Parameters estimation for isotropic layers: First, it is necessary to compare the proposed method with method proposed in Medeiros, (2005), where a chromosome has all parameters of seismic model. The simple isotropic model assumed is of two layers.

The dimensions used to generate travel times are: number of receptors equal to 100; Offset equal to 100 m; distance between receptors equal to 25 m.

The parameters of GA, to run the global experiment, are compared in the following Table 4, for one chromosome by model given in Medeiros, (2005) and for one Chromosome by layer proposed in this paper including anisotropy.

Table 3. Parameters of GA and computer resource of two structures of chromosomes for model of two isotropic layers.

	One Chromosome by Model OCM, GA	**One Chromosome by Layer OCL, GA**	**One Chromosome by Layer (OCL), HGA**
Size of the population	100	100	100
Number of generations	300	800; 300; 200	800; 300; 200
Arithmetic Crossover	0,8	0,8	0,8
Mutation	0,3	0,02	0,02
Selection	roulette	roulette	Roulette
GAP or Steady State	3	2	2
Number of experiments	10	2; 5	1; 1
Hybridization function - tolerance	-	-	Fminsearch, *$TolX = 10^{-4}$, $TolFun = 10^{-4}$*
Computer resources	Processor Pentium IV	Processor of 64 bits, 1.81 GHz and 1 GB of RAM	Processor of 64 bits, 1.81 GHz and 1 GB of RAM
Run Time (seconds)	-	160	32.4

In this case the differences for parameters of GA have few variations. Only the number of experiments decreases consistently.

The results of global experiments are compared in the Table 4, where the result of parameters estimation of Medeiros, (2005) is compared with results (of the best experiment of five others) using method of proposed in this paper: one chromosome by layer (OCL). This method uses two techniques: GA and the Nelder mead simplex algorithm.

Table 4. Results of parameters estimation for two different methods (OCM, one chromosome by model, and OCL, means one chromosome by layer).

Parameters of Layers	Real Value	Estimated with GA for OCM	Estimated with GA for OCL	Estimated with Hybrid for OCL
Thickness of Layer 1	1000 m	1331.7 m	1000.7 m	1000.0
Velocity of Layer 1	1500 m/s	1449.5 m/s	1500.9 m/s	1500
Anisotropy of Layer 1	-	-	0.0011	0.0000
Thickness of Layer 2	1000 m	1064.3 m	999.0 m	1000.0
Velocity of Layer 2	2800 m/s	2708 m/s	2798.9 m/s	2800.0
Anisotropy of Layer 2	-	-	4.0e-4	1e-7
msMAPE	-	4.50%	0.03%	4.34E-09

The results of parameters estimation for the method of one chromosome by layer are nearest of the real values compared with the method of one chromosome by model, in other words the modified symmetric mean absolute percentage errors (msMAPE). (Chen et al., 2004) are lower.

These good results must be confirmed in the same condition of computer resources and with more layers as in the next seismic model.

Parameters estimation for VTI layers: The next and main step consists of parameters estimation for a VTI medium described in Alkhalifah e Tsvankin, (1995).

The mediums VTI have horizontal layers and the model used has three horizontal layers. The first and the second layer are isotropic and the third layer has VTI properties as the Table 6.

The dimensions and number of receptors used to generate travel times are: number of receptors equal to 36; Offset equal to 20 m; distance between receptors equal to 20 m.

Initial tests were made to determine the optimal parameter of GA. The parameters for GA, to run the global experiment, are in the Table 5.

Variations of the parameters of GA that improved the behavior were the reduction of Number of generations and ratio of mutation. One reduces the run time and other improves the accuracy.

Table 5. Parameters of GA and computer resource of two structures of chromosomes for model of three layers with third layer VTI.

	OCM, GA	OCM, GA	OCM, HGA
Size of the population	100	100	100
Number of generations	2000	800; 300; 200	800; 300; 200
Heuristic Crossover	0.8	0.8	0.8
Tuning parameter of crossover	1	1	1
Uniform Mutation	0.08	0.05	0.02
Selection	Stochastic uniform	Stochastic uniform	Stochastic uniform
Steady State	5	2	2
Number of experiments	5	3; 2; 2	1; 1; 1
Hybrid function - Tolerance	-	-	Fminsearch, *TolX* $=10^{-4}$, *TolFun* $= 10^{-4}$
Computer resources	Processor of 64 bits, 1.81 GHz and 1 GB of RAM	Processor of 64 bits, 1.81 GHz and 1 GB of RAM	Processor of 64 bits, 1.81 GHz and 1 GB of RAM
Run Time (seconds)	2400	86	38.7

The results of global experiments are in the Table 6, to compare methods and variations of techniques.

Table 6. Results of parameters estimation for two different methods: OCM and OCL model of two isotropic layers and a VTI layer.

Parameters of Layers	Real Value	Estimated with GA for OCM	Estimated with GA for OCL	Estimated with HGA for OCL
Thickness of Layer 1	300 m	329.184 m	300.045 m	300.000 m
Velocity of Layer 1	1500 m/s	1621.121 m/s	1500.215 m/s	1500.000 m/s
Anisotropy of Layer 1	0	0.0018	0.0004	0.2e-7
Thickness of Layer 2	215 m	200.655 m	213.944 m	214.999 m
Velocity of Layer 2	2072 m/s	2021.3159 m/s	2064.064 m/s	2071.999 ms
Anisotropy of Layer 2	0	0.0214	0.0105	0.4e-7
Thickness of Layer 3	557 m	576.909 m/s	558.596 m	557.000 m
Velocity of Layer 3	2680 m/s	2759.958 m/s	2644.729 m/s	2640.000 m/s
Anisotropy of Layer 3	0.805	0.1916	0.8640	0.8749
msMAPE	-	3.574%	0.134%	$4.4\text{x}10^{-5}$%

The msMAPE of estimation of each parameter with the method of one chromosome by layer (OCL), are lower than the method of one chromosome by model (OCM).

When the second method with hybrid technique is used the accuracy is increased, too the run time is less, according the Table 5.

Conclusions

The use of OCL reduce the error and run time with respect to OCM, but the use of HGA improves significantly the accuracy of parameters estimation. The use of HGA allows reducing the time of syntonization of GA when accuracy is required and number of variables is increased. The unique condition is to get a best point (or vector) near of the global minimum, in opposite way the Nelder mead simplex can fall into local minimum.

Every time that is estimated parameters for a new layer it is equal to minimize a new function (or equation).

It is possible to find bad results when real data is used, because the noise can produce errors in the interpretations of travel times or because the layer are not horizontal exactly and finally the type of anisotropy can be different of type assumed in this paper.

There are a lot of approximated equations according to type of medium that can be applied in the evaluations functions. This implies to modify the parameters of GA to get optimum results and reducing the run time that increases with number of parameter and number of layers.

Bibliographic References

Alkhalifah, T.; Tsvankin, I., "Velocity Analysis in Transversely Isotropic Media". In: Geophysics, Vol 60, No 5, pp 1550 – 1566; 1995.

Bazelaire C.; Bazelaire E.; Perroud H. Analytical Inversion of a Stack of Weakly Anisotropic Layers. The Society of Explorations Geophysicists, 2000.

Chen, Z.; Yang, Y. Assessing Forecast Accuracy Measures. Iowa, USA: Iowa State University, March 2004.

Dourado, J. C. Sísmica de Reflexão 3, A Seção Sísmica. Aula 8, revisão 2001. Disponível em: < www.geologia.ufpr.br/graduacao/geofisica2007/reflexao3_aula8.pdf>.

Goldberg, D. E., "Genetic Algorithms in Search, Optimization, and Machine Learning". Addison-Wesley, 1989

Gouveia, W. Residual statics estimation by a hybrid distributed genetic algorithm. Consortium Project on Seismic Inverse Methods for Complex Structures, May 10, 1994. CWP-153. PDF File.

Huamán B., S. G. Inversão de parâmetros sísmicos em três dimensões a partir de dados de reflexão sísmica por algoritmos genéticos híbridos. Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro RJ, Brasil, agosto 2008.

Lagariasy, J. C.; Reedsz, J. A.; Wrightx, M. H.; Wright, P. E. Convergence Properties of The Nelder-Mead Simplex Method In Low Dimensions. Society for Industrial and Applied Mathematics Vol. 9, No. 1, pp. 112-147, 1998.

Margrave, G. F. Numerical Methods of Exploration Seismology with algorithms in MATLAB. Calgary, Canada: Department of Geology and Geophysics, University of Calgary, July 11, 2003.

Mathworks, Inc. Genetic Algorithm and Direct Search Toolbox ™ 2. User's Guide, copyright 2004-2007. Disponível em: <www.mathworks.com>.

Mensch, T.; Rasolofosaon, P. Elastic-wave velocities in anisotropic media of arbitrary symmetry - generalization of Thomsen's parameters ε, δ, and γ. Geophys. J. Int.128, 43-64, 1997.

Medeiros, S. C. D. Inversão de Parâmetros em Dados Sísmicos por Algoritmos Genéticos. 2005. 87p. Dissertação de mestrado. Departamento de Engenharia Elétrica – PUC-Rio, Rio de Janeiro, 2005.

Michalevicz, Z. Genetic Algorithms + data structures = evolution programs. Springer-Verlag, 1996.

Silva, M.B.C. "Influência da Anisotropia VTI na correção de Sobre tempo Normal em Dados Sísmicos e Análise de Velocidade por Gradiente Descendente", Tese de Doutorado, PUC-Rio, 2006

Tarantola, A. Inverse Problem Theory and Method for Model Parameter Estimation. Paris, France: Siam, 2005. 357p.

Yu T.; Wen X.; Lee S. "A Hybrid of Sequential-Self Calibration and Genetic Algorithm Inversion Technique for Geostatistical Reservoir Modeling" 2006 IEEE Congress on Evolutionary Computation, Vancouver, BC, Canada, July 16-21, 2006.