Analysis of one-dimensional seismic waveform inversion by Regularized Global Approximation

Gennady Ryzhikov, Marina Biryulina and Shmariahu Keydar

key words: global optimization, forward modeling, inversion, nonuniqueness, objective function, regularization, soft constraint.

Abstract

Direct analysis of normal incidence seismogram inversion with respect to a velocity profile is available now due to applying of a new global optimization algorithm. The latter is based upon regularized global approximation of an objective function which is not supposed to be differentiable. The new technique allows to see clearly a nonuniqueness of the inversion problem, no matter how high is a quality of the input data. It is induced by a few factors: a source wavelet is a function of a finite frequency band, an effective wave length of the sounding signal is increasing jointly with the velocity, and the power of a media response is decreasing with respect to the depth.
The nonuniqueness means that there is no inversion/processing enable to solve the problem if it does not take into account a priori information about the recovered velocity profile. It is shown how an a priori assumption about a trend of the profile can essentially reduce the nonuniqueness of the problem. The corresponding regularization has a form of a soft constrain on the misfit function and leads to an unbiased estimation of the velocity profile when the latter is monotonous function with respect to the depth. On the other hand the regularization suggested allows to reconstruct nonmonotonous functions as well which leads to a biased estimation of the velocity profile as like as any conventional regularized inversion does. Examples of computer experiments are given that yield an opportunity to reconstruct the images of nonregularized and regularized objective functions as well as to see an accuracy of corresponding solutions of the inverse problem.

Introduction

The problem of reconstruction of acoustic impedance from reflection seismic data, so called a normal incidence inversion has been widely studied by many authors (e.g. Kunetz(1963), Gjevik et al.(1976), Gray (1984), Burridge (1980)). As a rule, this inverse problem was treated as a dynamical one involving an inverse continuation of the wave field. A complete review of comparison of inverse methods can be found in the paper of Ursin and Berteussen (1986).
Another approach was proposed by Bamberger et al. (1982). It was formulated as an optimization problem: to find the acoustic impedance that minimizes the squared difference between the synthetic and recorded seismograms under additional constraint on the sum of the absolute value of the impedances. The disadvantage of this method is that it uses a gradient- type method of optimization and therefore succeeds to converge only in the vicinity of the true velocity model. This is because of the fact that the inverse problem is essentially nonlinear (the correspondence between the seismogram amplitudes and the elastic parameters is nonlinear) and the objective function of residuals is multimodal, i.e. it has many minima.
Recently, attempts were done to use Genetic and Simulated Annealing algorithms for 1-D inversion. These algorithms do not use the gradient of the objective function, but the number of necessary forward calculations can be significantly large (e.g. Stoffa (1994), where 80,000 iterations are required for Genetic Algorithm), even in the case where a number of unknown parameters is rather small. These algorithms are still in the preliminary stage.
We have applied a new method to estimate the velocity model from seismograms. It takes into account all the interbed multiples and does not require any special preliminary treatment of field record, as deconvolution or multiple suppression.

With the help of this method a 1-D waveform inversion can be analyzed. In its turn the method is based on a global optimization algorithm, so called Regularized Global Approximation algorithm, or *RGA-algorithm*) (Ryzhikov and Biryulina, 1995). The objective function is a weighted difference between the synthetic seismogram (given by one-dimensional forward modeling in a layered medium with a chosen velocity model) and an observed seismogram. It was shown that such a choice of the objective function leads to nonuniqueness of the solution. This means that well-separated velocity models can yield indistinguishable synthetic seismograms. To overcome this problem we use a new objective function with a regularization constraint. We explain the algorithm and show its implementation on synthetic data.

1 The forward problem

We consider normal incidence inversion; therefore the following assumptions have to be fulfilled:
1. there is only propagation of acoustic waves;
2. the propagation velocity depends on the depth only, i.e. we deal with one-dimensional medium model;
3. there is a plane wave controlled source.

The optimization requires a large number of evaluations of the objective function, which is based on calculation of the seismogram. Therefore, a fast and reliable forward modeling is crucial for the proposed inversion method. For this purpose we use a forward modeling based on algorithm suggested by Claerbout (1968). In this algorithm the reflection coefficients $(r_1, r_2..., r_t)$ in $[0-T]$ time interval for every sample are computed by the Levinson relation for the minimum-phase factorization of an autocorrelation function.

The forward problem can be defined as follows: assume the velocity and density distributions in depth and the wavelet $w(t)$ to be known, find the pressure p_t at every sample in the interval $[0, T]$ in time. The set of p_t obtained in this interval is called the seismogram. To get the seismogram the reflection coefficients obtained in the first step of Claerbout's algorithm are convolved with a given frequency band source wavelet $w(t)$.

2 The inverse problem

The inverse problem consists in estimating the velocity model from the set of obtained pressures in the time interval $[0-T]$. For simplicity we assume that the density model is known. Because of the noise-corrupted nature of the observation it is the classical way to define a solution of the inverse problem as the velocity model supplying with such a seismic response p_{cal} that fits the observed measurements p_{obs} best. Moreover we want the velocity we are looking for to match the a priori information we have about the solution. This can be achieved by constraining the solution into a compact set of velocity models. Therefore, the mathematical formulation of the inverse problem is as follows: given the observation pressures p_{obs} and the wavelet $w(t)$, find the velocity model $v(z)$ that minimizes the objective function $\mathcal{F}^p$. The function is the weighted difference between the observed seismogram p_{obs}, contaminated with noise ε. and synthetic seismogram $p_{cal} = p_{cal}(v)$ calculated for a velocity model $v(z)$:

$$\mathcal{F}^p = \langle \, p_{obs} - p_{cal} \mid p_{obs} - p_{cal} \rangle \, . \tag{1}$$

The weight is induced with the covariance matrix C_ε of the noise ε, i.e. the inner product (1) can be defined as:

$$\langle \, p^{'} \mid p^{''} \rangle = J^{-1} \sum_{i}^{J} \sum_{j}^{J} (C_\varepsilon^{-1})_{ij} p_i^{'} p_j^{''} \, , \tag{2}$$

where $p^{'}$, $p^{''}$ denote arbitrary seismograms with J samples, $p^{'} = \{p_1^{'}, p_2^{'}, \ldots, p_j^{'}, \ldots, p_J^{'}\}$, and the summation is done over all samples of the synthetic seismogram. In the following numerical examples we shall see that such a choice of objective function leads to nonuniqueness of the solution. This means that well-separated velocity models can yield indistinguishable synthetic seismograms. To overcome this problem we use a new objective function with a soft constraint, which provides a regularized solution. Assuming that the velocity model has a dominant increasing with depth, we change the objective function (1) in the following way:

$$\mathcal{F} = \mathcal{F}^p + \mathcal{F}^v = \mathcal{F}^p + \alpha \Big(\sum_{k}^{K-1} \Delta_k v \Big/ \sum_{k}^{K-1} |\Delta_k v| - 1 \Big)^2 \, , \tag{3}$$

where $\Delta_k v$ is the jump of velocity between layer k and layer $k+1$ and K is the entire number of layers. The positive factor α plays the role of the parameter of regularization, while the regularizing function $\mathcal{F}^v$ depends on a priori information concerning velocity model only. Note, that in terms of the Delphi [2] this kind of regularization means an a priori assumption that a velocity macromodel has nonnegative *trend* with respect to depth.
Soft constrains allow to deal with a priori information in a stochastic form. The solution v^* providing minimal value to the objective function (3) can be interpreted in terms of a Bayesian strategy. Namely, $\mathcal{F}^v$ can be treated as induced by Gaussian distribution of velocities $P(v)$:

$$P(v) \propto \exp\{-\mathcal{F}^v\} = \exp\{-\alpha[(\Delta v - \int |\nabla v(z)|dz)/\int |\nabla v(z)|dz]^2\} \qquad (4)$$

where $\Delta v = v_K - v_1$ denotes the value of velocity jump between the last (K-th) and the first layers. This means that the following a priori assumption is included into the inversion: to get a velocity model that increases with depth monotonously is more probable than a model with waveguides. Because of zero value of the regularizing function $\mathcal{F}^v$ for velocity models without waveguides the corresponding Bayesian strategy yields an unbiased estimate of the functions $\{v(z) \mid \partial_z v \geq 0\}$.

3 The solution of the optimization problem

The method of minimization of the objective function is a crucial stage of the inversion algorithm. Many algorithms have been proposed to solve this problem (Himmelblau (1972); Gill and Murray (1974); Gill, Murray and Wright 1981). However, none of these algorithms has proved to be a universal solution for nonlinear programming problems. We suggest a new algorithm of global optimization (Ryzhikov and Biryulina, 1995), that allows to treat nondifferentiable multimodal objective functions in the way of Simulated Annealing or Genetic algorithms. Unlike Simulated Annealing or Genetic algorithms, the suggested algorithm includes a predictor-corrector procedure which significantly reduces the number of necessary evaluations of the objective function. Besides it allows to deal with an objective function in a multidimensional parameter space (e.g. 20-D, as it is shown in this paper). Therefore it yields an opportunity to deal with a computationally expensive objective function. Global optimization algorithm is based on Regularized Global Approximation of a objective function (**RGA** -algorithm). Using values of the objective function (OF), calculated on a set of points, a differentiable global approximation of the OF is constructed. This approximate OF can be interpreted as prediction of OF-value at any point of the domain of searching. The next step consists in expansion of the initial set of points by extremal points of the OF-approximation. The OF is to be calculated at these points. The enriched set of OF values gives a renewing OF-approximation. Termination of the algorithm is defined by a condition that minimum of calculated values of OF is reached namely at the point of global minimum of OF-approximation provided the OF-value is well-predicted.
Let $\mathcal{F}$ be the objective function of variables $\mathbf{v} = \{v_1, v_2, \ldots, v_n\}$, where $\mathbf{v} \in \mathcal{V} \subset \mathbf{R^n}$ $(n \leq 20)$. It means that any velocity model with n layers is represented by the respective point in the domain of searching in the n -dimensional space, and a few models compose a subset of points. The solution $\mathbf{v}^*$ should provide a minimal value to the $\mathcal{F}$. In a certain sense the RGA-algorithm is a recursive procedure: starting from a subset of velocity models and respective OF-values on a j-th step of the recursion, the algorithm should indicate a set of new velocity models on the $j+1$-th step. The forward modeling allows to evaluate OF for these new models (i.e. for new points in the n-dimensional space). The subroutines of RGA-algorithm are following (Fig.1):

- **Starter**

  The initial subset of velocity models is generated in a random manner.

- **Predictor**

  This is the main part of algorithm, which constructs the approximation of the objective function on the base of already known values of OF at a few points. The OF-approximation

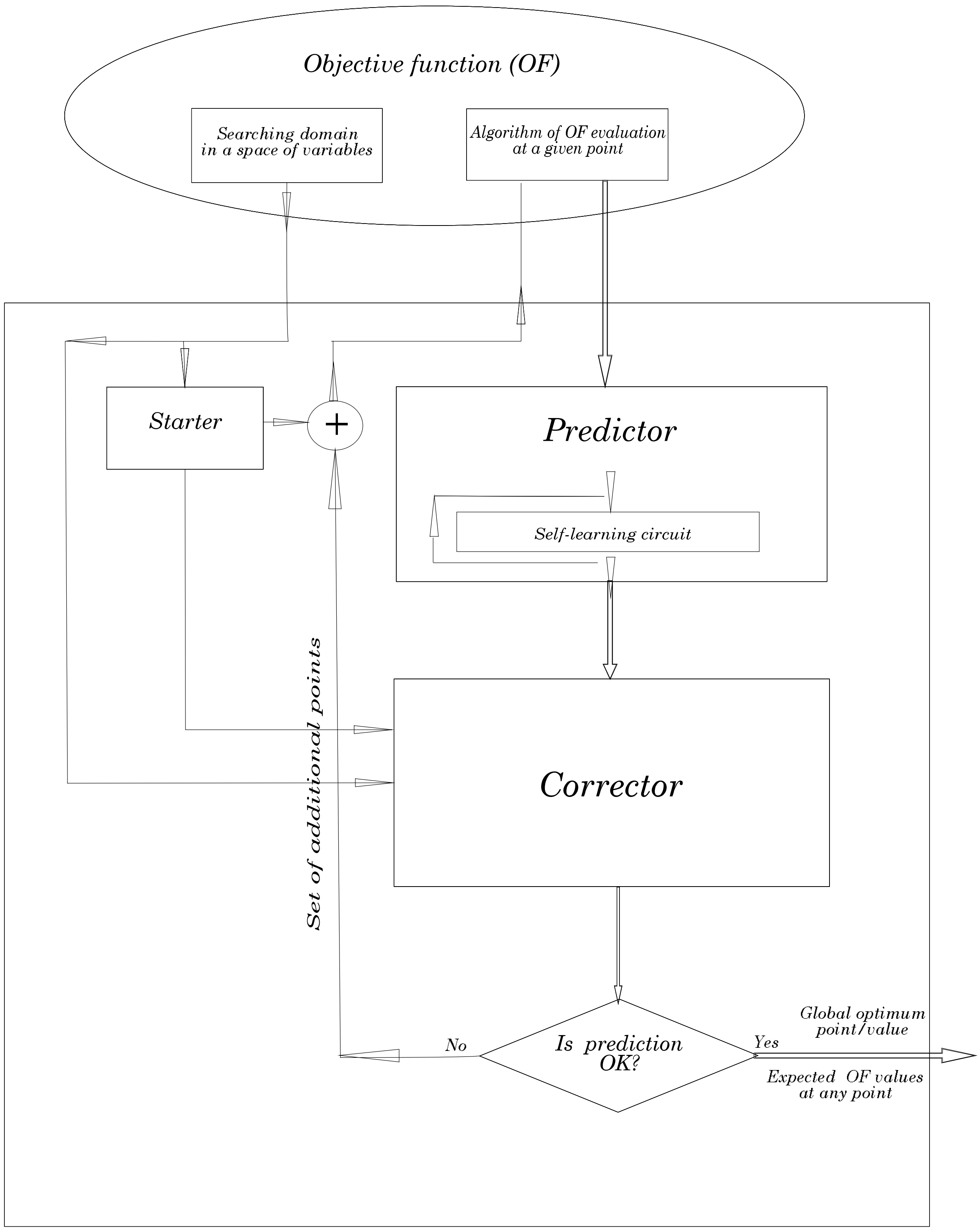

Figure 1: Block-scheme of the RGA- global optimization algorithm.

should match those values well and be fairly smooth everywhere. Note, the approximation constructed by *predictor* on a given step of recursion does not "remember" a proper approximation constructed on a previous step, and in this sense the algorithm is not an iterative one.

•• **self-learning circuit**

The 'trade-off' between 'well-fitting' of known OF-values and 'smoothness' of the entire OF-approximation is governed by parameter of regularization which is determined locally in the searching domain in accordance with a test-prediction.

- **Corrector**

  This subroutine indicates all extremal points of the OF-approximation, which yields the new velocity models for direct OF-evaluation.

- **Termination trigger**

  It compares the predicted location as well as value of global minimum of OF and the actual value of OF at the predicted point. Sure, the actual value should be minimal compared with the values at all other ponts where OF has been already evaluated. The algorithm is terminated if the result of comparison is OK,

Note, that the 2nd derivative of OF (3) is discontinuous, but it is not crucial for the algorithm being enable to operate even with discontinuous OF's.
Mathematical details of the RGA - global optimization algorithm can be found in [13]. See also results of application of RGA to real data in [11].

4 Computer experiments

4.1 Example 1

The purpose of this example is to show how the regularization constraint can reduce nonuniqueness. An artificial 'true' velocity model consists of 20 layers with equal thickness of 2 m. The equidistance of the true model is not crucial for our approach as far as we do not apply Fourier-transform and hence do not use the translation symmetry of the model.
Starting from the velocity model and a given high frequency (up to 130 Hz) Ricker wavelet (Figure 2) we compute the 'true' seismogram, using forward modeling (Claerbout, 1968). 'Target' zone consists of 5 last layers (Figure 3, true velocities are shown with the *dashed line*) while 'overburden' (15 upper layers) is supposed to be known. So, the unknown parameters are the velocities at each of the 5 layers. The objective function (1) is used for the optimization procedure without regularization and the function (3) for the regularized optimization correspondingly. The seismograms in the time window of interest are shown in Figure 4, 5 with dashed lines.
The optimization algorithm found the solution after about 250 calculations of synthetic seismograms and provided a value of the objective function with an error less than 1%. The obtained velocity model (Figure 3, left solid line) is quite different from the true (Figure 3, dashed line), while the part of the constructed seismogram from 200 ms-400 ms consists information for the unknown layers similar to the observed one (Figure 4).
In order to reduce the nonuniqueness we use the a priori knowledge about macromodel, namely that the velocities increase with depth. Corresponding to this information, we add a regularization constraint to the objective function (1), so the regularized objective function is represented with (3). The algorithm defined a solution after a several hundreds of forward modeling calculations. The solution of the velocity model is shown in Figure 3 with the right solid line with asterisks, and the respective seismogram is shown in Figure 5.
The constructed velocities are close to the 'true' with a maximum error less than 4%. To illustrate the behavior of the objective function in the 5-parametrical space (1) graphically, we calculated for

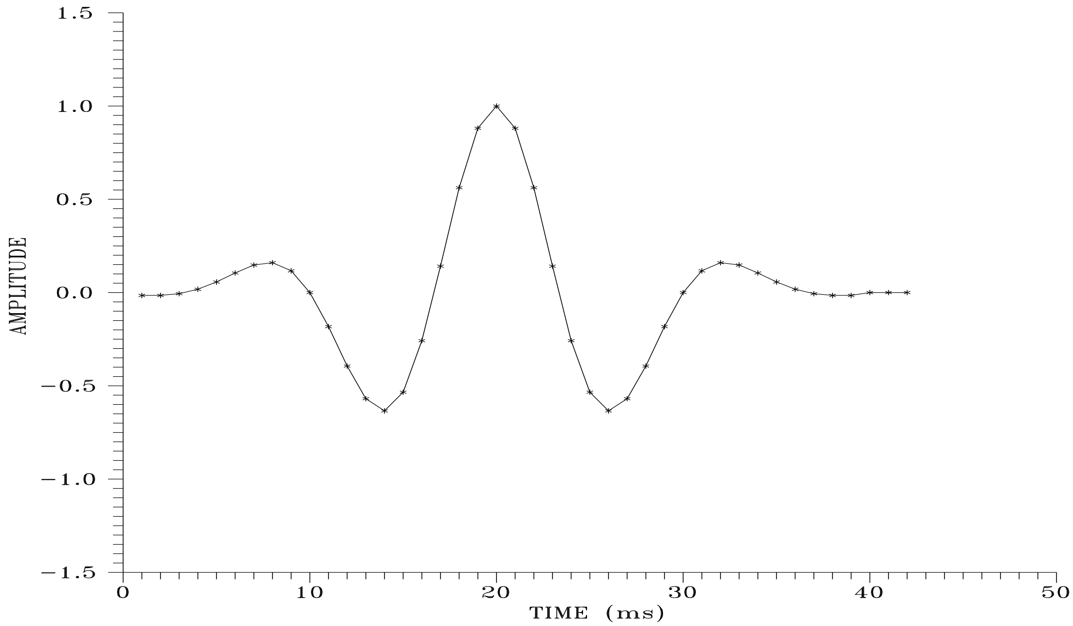

Figure 2: Source wavelet.

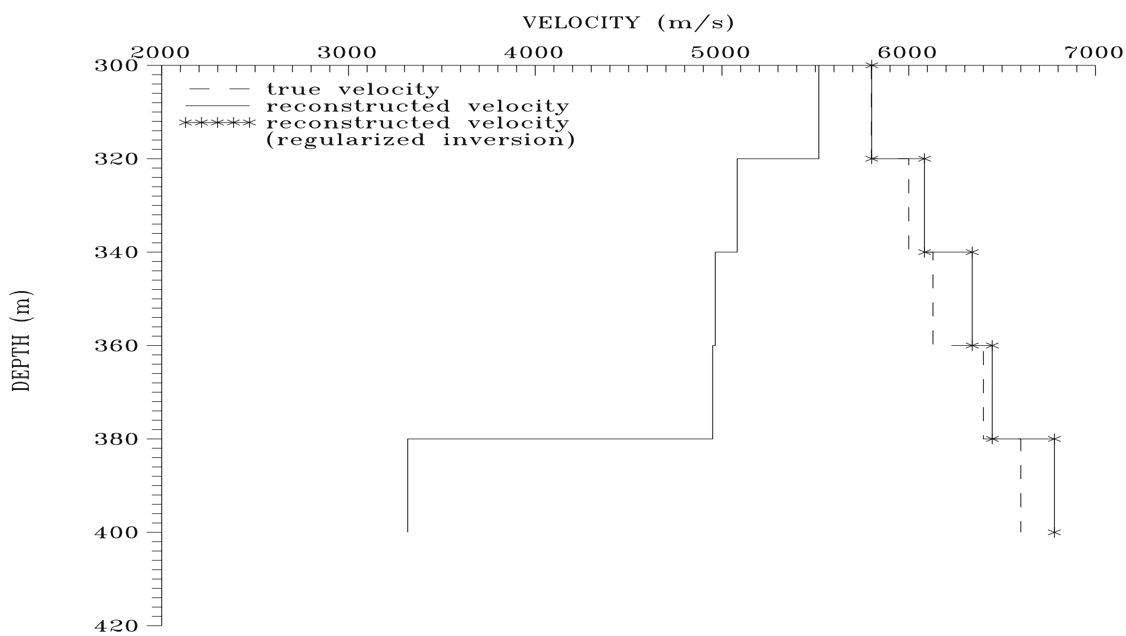

Figure 3: Model I. 'Target' zone consists of 5 last layers (true velocities are shown with the *dashed line*) while 'overburden' (15 upper layers) is supposed to be known. One of the results of unregularized reconstruction is shown by the *solid line* on the left . Regularized reconstruction is represented with the *solid line with asterisks.*

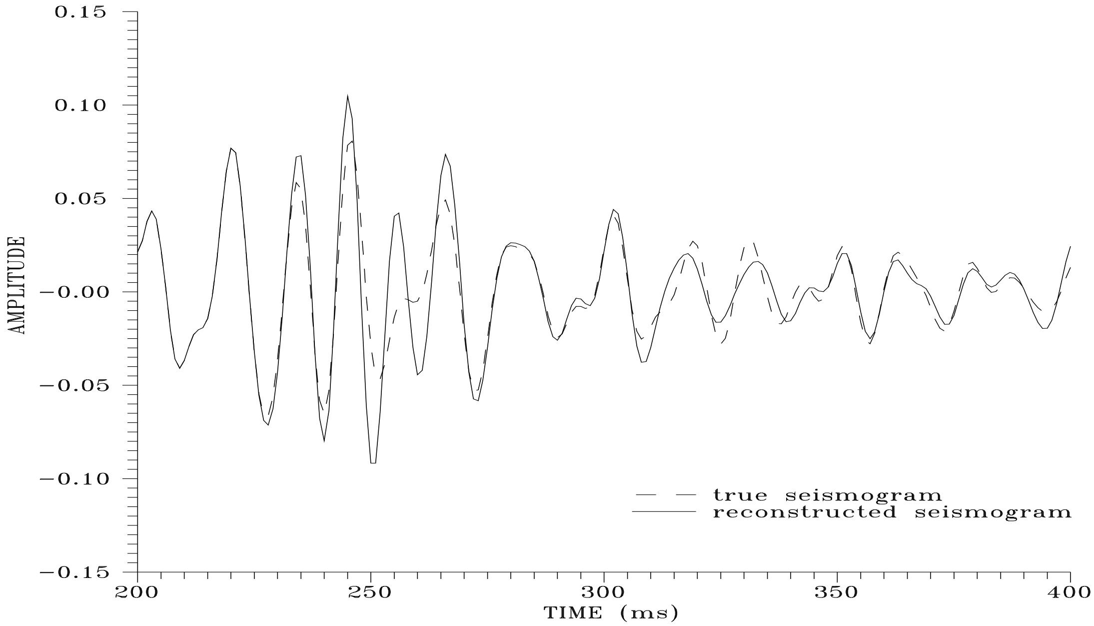

Figure 4: Model I. Comparison between original seismogram and seismogram induced by the unregularized reconstruction.

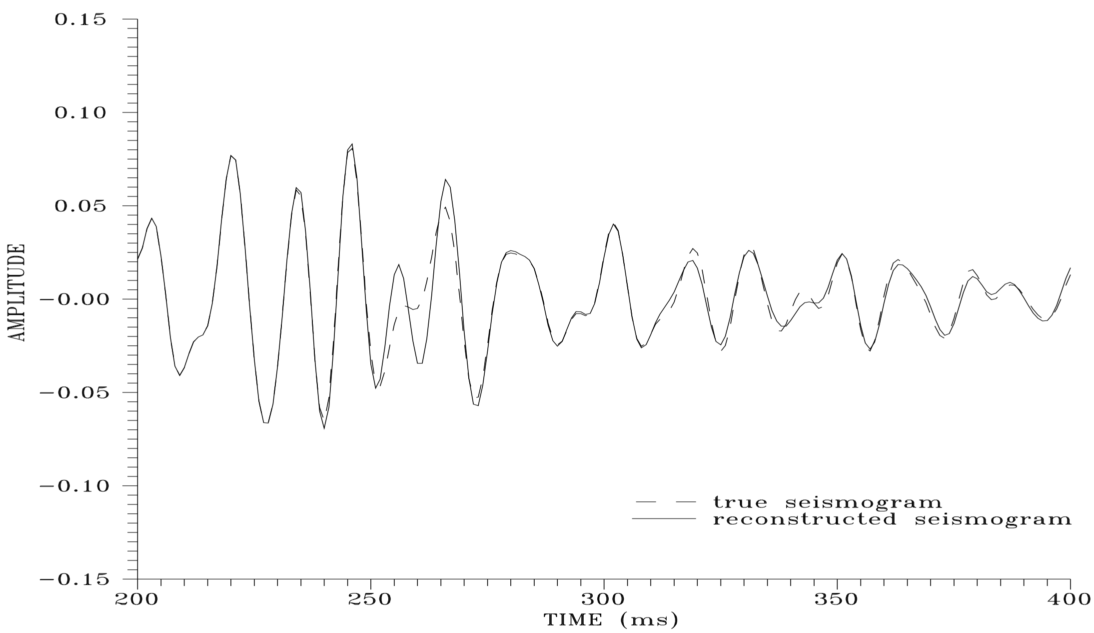

Figure 5: Model I. Comparison between original seismogram and seismogram induced by the reconstruction under the regularization constrain.

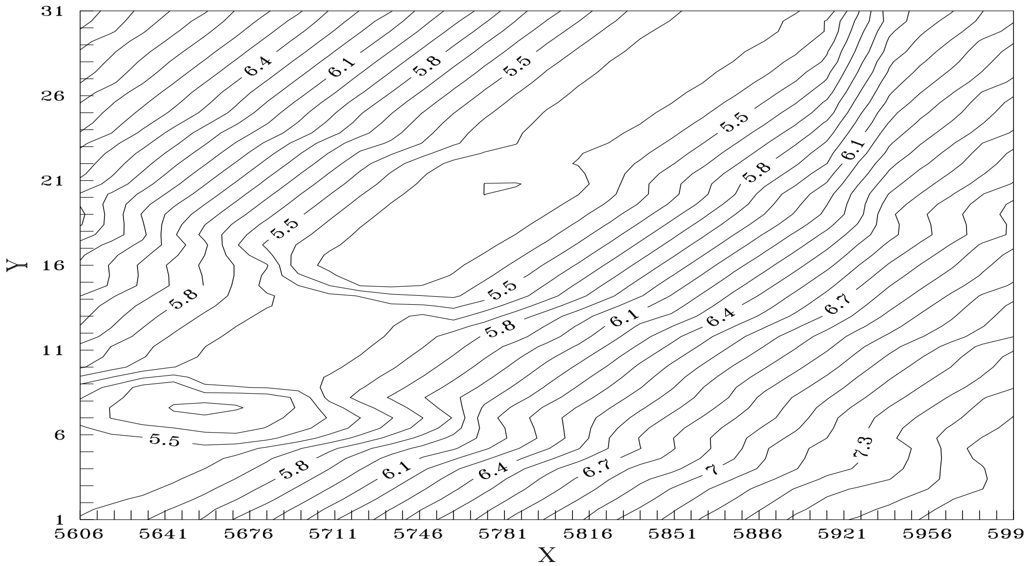

Figure 6: Model I. Contours of the objective function without regularization.

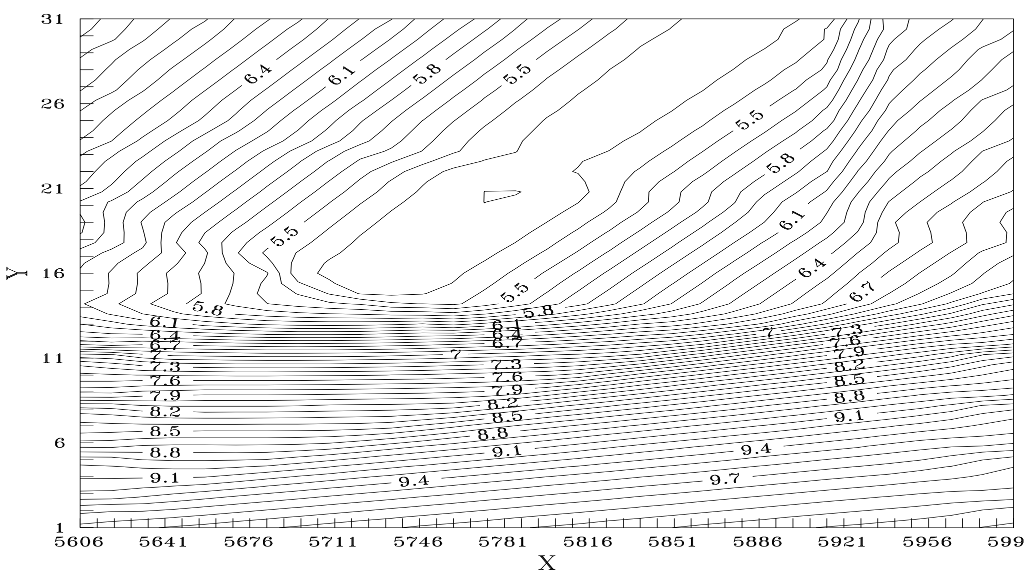

Figure 7: Model I. Contours of the objective function with regularization.

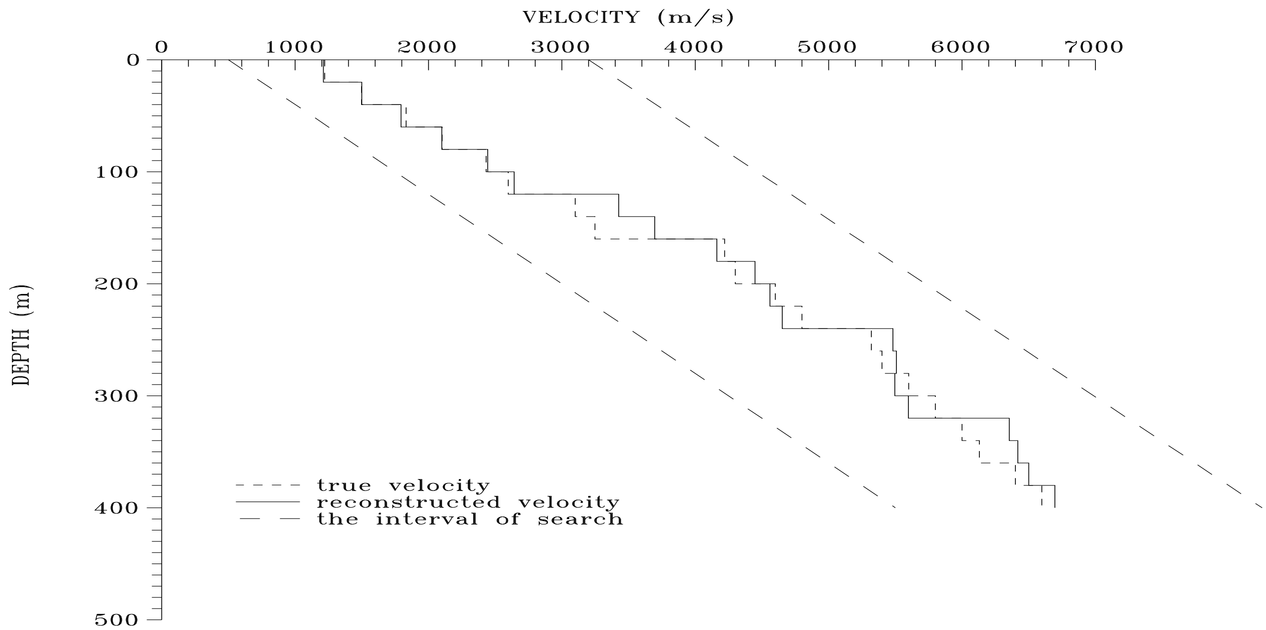

Figure 8: Model II. The searching parameters are all twenty velocities. The domain of searching is limited with an interval given by *broad dash line*. Result of a regularized reconstruction is shown with solid line.

the model, described above, the values of the OF in the section, containing both of the constructed solutions, without regularization constraint (Figure 6) and with it.
One can clearly see that the unregularized objective function has a few minima even in the vicinity of the true solution. But the objective function for the same velocity model in the same section of the 5-parametrical space has no secondary minimum after including the regularization constraint (3), as it is shown in Figure 7.

4.2 Example 2

The next run was carried out for the same velocity model but all the twenty values of interval velocities were treated as the unknown parameters. The comparison of the true and reconstructed velocities are shown in Figure 8.
We can see that the resolution decreases with depth because of the increasing effective wavelength of sounding signal and the loss of the signal energy.

4.3 Example 3

To illustrate that the *soft* regularizing constrain allows to reconstruct a velocity micromodel *details* [2] we carried out an experiment represented in Figures 9 and 10. The unknown parameters all of twenty parameters as like as in the previous run, but the true velocity model is such that the respective value of regularizing function is not equal zero. It means that in this case the constrain becomes *active*, that leads to the *biased* Bayesian estimation: as a rule negative jumps of the reconstructed velocity are less than true ones. Nevertheless a good agreement is visible between the final and true model. To get the result took a few hundred synthetic seismogram evaluations. The observed and the final seismograms are shown in Figure 10. Again, the fine structure of deep layers is substituted with thick layers without loosing of information contained in records.

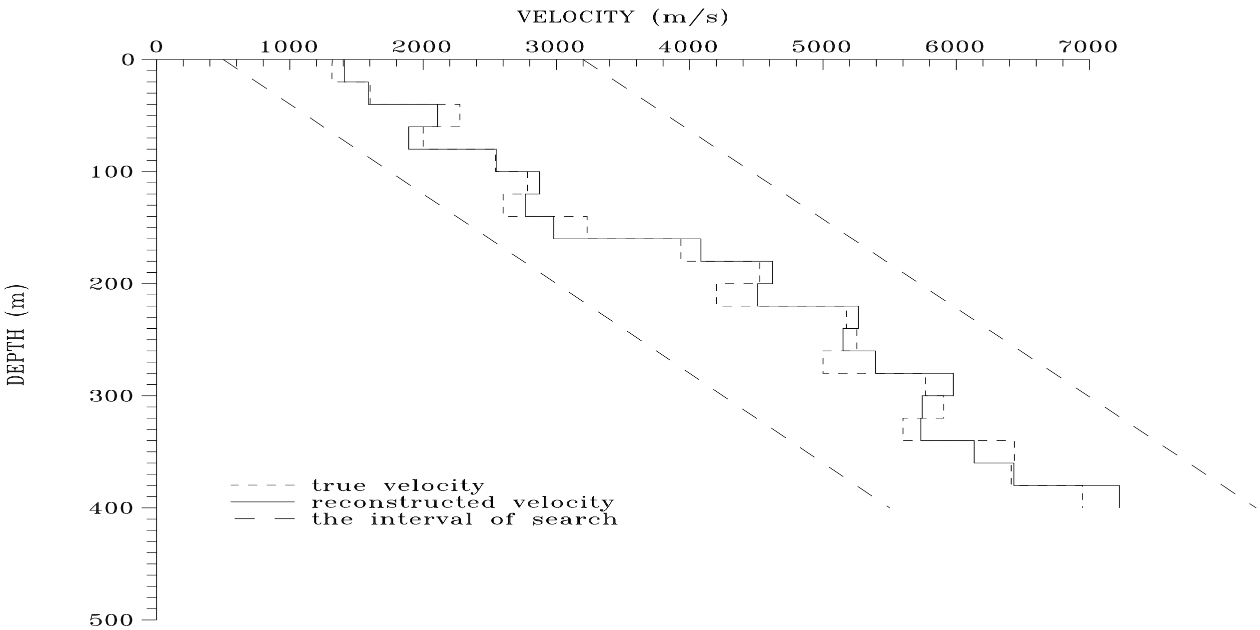

Figure 9: Model III. Result of reconstruction of a velocity model with waveguides.

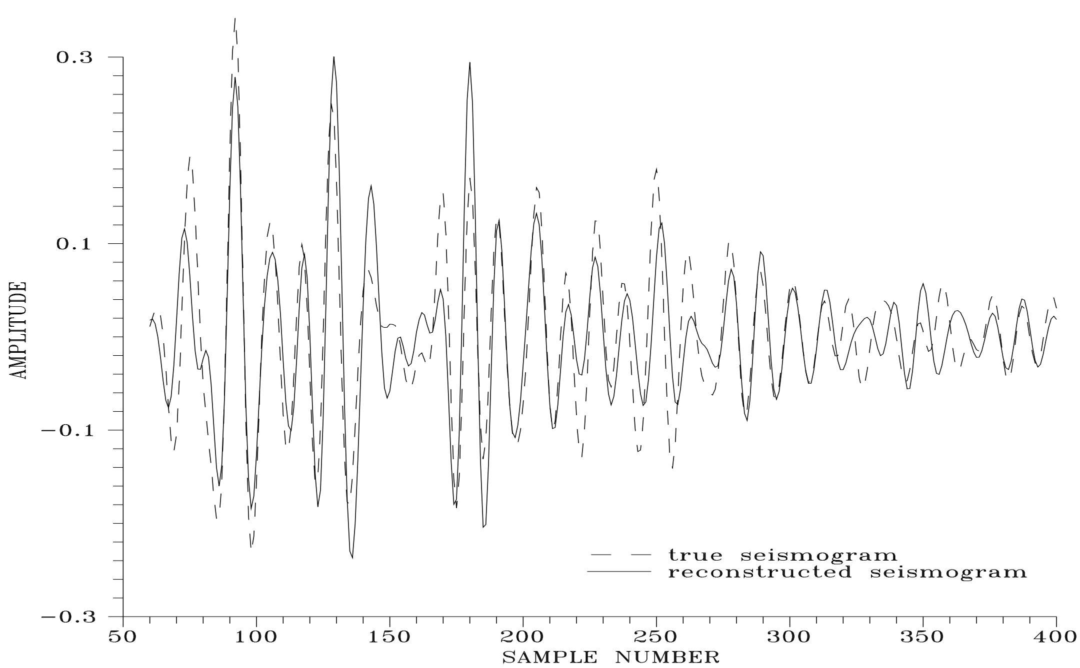

Figure 10: Model III. Comparison between 'observed' and synthetic seismograms.

5 Conclusion

1. Analysis of 1D waveform inversion as like as the proper inversion became available due to applying the RGA - global optimization algorithm. It is clear from Figure 6, that applying of a gradient-type of algorithm is useless in the general case: even in the 5-dimensional parameter space the inverse objective function is essentially multimodal in addition to the valley structure of the objective function at any local minimum.
2. Introducing of a priori information (regularization) can exclude secondary minima and lead to the true solution, provided the a priori information is adequate.
3. The results of inversion represented by Figures 8 and 10 illustrate that the vertical resolution is decreasing with the depth because of an increasing effective wavelength of the sounding signal and a loss of the signal energy. Therefore it is useful to choose the model in accordance with the characteristic wavelength of the sounding wave field: fine structure of deep layers can be substituted by rather rough structure, yielding dynamical response with great accuracy.
The latter can be interpreted as numerical proof of the existence of *dynamically equivalent models.*

Acknowledgments

The research was conducted under the EU projects "Seismic Tomography Based on Advanced Ray Tracing" (associated with the Geoscience II Project " Stratigraphic Modelling & Seismic Inversion") and "3D Asymptotic Seismic Imaging" (JOU2-93-0321). Both projects were financed by the Norwegian Scientific Council.
We would like to thank the anonymous reviewer for valuable remarks and Dr. T. J. Moser for proofreading the first version of the manuscript.

References

[1] Bamberger, A., Chavent, G., Hemon, C. J., and Lailly, P., 1982. Inversion of normal incidence seismograms: Geophysics, 47: 757-770.

[2] Berkhout, A. J. and Wapenaar, C.P.A., 1990. Delphi: Delft philosophy on acoustic and elastic inversion: Leading Edge, Feb. 1990: 30-33

[3] Biryulina, M. and Ryzhikov, G., 1995. Rytov-Born decomposition in 3-D reflection seismics, in *Extended Abstracts of EAEG and EAPG 57th Conference and Technical Exhibition, Glasgow,* Vol.1, E046.

[4] Burridge, R., 1980. The Gel'fand-Levitan , the Marchenko and the Gopinath-Sondhi integral equations of inverse impulse-response problem: Wave motion, 2: 305- 323.

[5] Claerbout, J.F., 1968. Synthesis of a layered medium from its acoustic transmission response: Geophysics, 33: 264-269.1

[6] Gill, P.E. and Murray, W. 1974. Numerical Methods for constrained optimization: Academic Press Inc.

[7] Gill, P.E. Murray, W. and Wright, M.M 1981. Practical optimization: Academic Press Inc.

[8] Gjevik, B., Nilsen A., and Hoyen J., 1976. An attempt of the inversion of reflection data: Geophys. Prosp., 24: 492-505.

[9] Gray, S. H, 1984. The relationship between 'direct, discrete' and 'iterative, continuous' one-dimensional inverse methods: Geophysics, 49: 54-59.

[10] Himmelblau, D. M. 1972. Applied nonlinear programming McGraw-Hill Book Co.

[11] Keydar, S., Landa, E., Ryzhikov, G. and Sorin, V., 1995, A new approach to coherency inversion by global optimization 65th SEG Meeting, Expanded Abstracts: 1081-1083

[12] Kunetz, G., 1963. Quelques examples d'analyse d'enregistrements sismiques: Geophys. Prosp., 11: 409-422

[13] Ryzhikov, G. and Biryulina, M., 1995. 3D nonlinear inversion by entropy of image contrast optimization: Nonlinear Processes in Geophysics, 2: 228-240.

[14] Ursin, B., and Berteussen, K., 1986. Comparison of some inverse methods for wave propagation in layered media: Proc. of the IEEE, 74, 3.