Chapter

FLOAT-ENCODED GENETIC ALGORITHM USED FOR THE INVERSION PROCESSING OF WELL-LOGGING DATA

Norbert P. Szabó[1*] and Mihály Dobróka[1,2]
[1]University of Miskolc, Department of Geophysics, Hungary
[2]MTA-ME Research Group for Earth Science and Engineering, Hungary

ABSTRACT

A Float-Encoded Genetic Algorithm is presented in this chapter for solving the well-logging inverse problem. The aim of the global inversion of well-logging data is to provide a robust and reliable estimate of petrophysical properties, such as porosity, water saturation, shale volume and mineral content, associated with geological structures. There are two possible ways to solve the interpretation problem. The first is a conventional inversion scheme, separately estimating the unknowns to different depths. In the forward modeling phase of the local inversion procedure the theoretical well-logging data set is calculated by using locally defined probe response functions, which are then fit to real data in order to estimate model parameters to one depth only. This procedure leads to a marginally over-determined inverse problem, which results in relatively poor parameter estimates. A further disadvantage of the latter technique is that some crucial quantities, such as the thickness of layered geological formations, cannot be extracted by inversion because they do

* Corresponding author's e-mail address: norbert.szabo.phd@gmail.com

not appear explicitly in local response equations. A new inversion methodology introduced by the authors gives much more freedom in choosing the inversion parameters. The so-called interval inversion method inverts all data measured from a greater depth interval in a joint inversion process. By a series expansion-based discretization of the petrophysical model, a highly over-determined inverse problem can be formulated, enabling an estimation of the petrophysical parameters including new unknowns, such as zone parameters and layer thicknesses, more accurately compared to local inversion methods. The authors give further references for several applications of the global inversion method. In this chapter, one synthetic and two field examples are presented to demonstrate the application of the Genetic Algorithm-based inversion method. It is shown that the combination of the new inversion strategy and global optimization tools forms a highly effective and adaptive algorithm for earth scientists who are interested in a more reliable calculation of the reserves of hydrocarbons and other mineral resources.

Keywords: Genetic Algorithm, well-logging, petrophysical model, interval inversion

1. INTRODUCTION

Well-logging methods are extensively used in mineral exploration for collecting high resolution *in-situ* information associated with subsurface structures. Surveying methods are based on different physical principles, which result in well logs of several measured quantities (Serra, 1984; Asquith and Krygowski, 2004). The main object of a log analysis is the lithologic separation of the succession of strata and the estimation of several geometrical (layer-thickness, formation dip, lateral changes of layer boundaries) and petrophysical properties of geologic formations (porosity, water saturation, mineral content, permeability). The most advanced data processing tools are based on geophysical inversion methods. Both the principles and computer implementation of the well-logging inversion procedure are detailed by Mayer and Sibbit (1980), Alberty and Hashmy (1984) and Ball et al. (1987).

The interpretation of well-logging data sets is traditionally solved by linearized inversion methods (Menke, 1984; Tarantola 2005). Linear optimization methods work properly and quickly, provided that we have satisfactory *a priori* information about the petrophysical model. Well-logging measurements are usually supported by laboratory data measured on core samples and independent results of geophysical measurements made on the

surface (Telford et al., 1990). Although there is an abundance of information available, the data noise and inherent ambiguity of the inverse problem often cause the failure of the linear optimization procedure as it is trapped in a local minimum solution. The problem can be effectively avoided by using global optimization methods that search for the absolute extreme of the multivariate objective function. Currently, the most preferred global optimization procedures are the very fast algorithms of the Simulated Annealing method (Metropolis et al., 1953; Sen and Stoffa, 1997) and the Float-Encoded Genetic Algorithm (Michalewicz, 1992).

The Genetic Algorithm (GA) was proposed by John Holland (1975) for solving optimization problems by using the analogy of natural selection in biology. The procedure improves a population of random models in an iteration process. Each individual of a given generation has a fitness value, which represents its survival capability. In the geophysical inverse problem the fitness function is connected to the distance between the observed data and theoretical data calculated on the actual petrophysical model. The most probable petrophysical model can be estimated by maximizing the fitness function (i.e., minimization of data misfits). During the genetic process the fittest individuals reproduce more successfully in the subsequent generations than those who have less fitness. In the last generation, the individual with the maximum fitness corresponds to the solution of the optimization problem. In the classical GA procedure the model parameters are encoded using a binary coding scheme, setting a limit to the resolution of the petrophysical model and the accuracy of the estimation results. The Float-Encoded GA computes with model parameters as real numbers, which is faster than binary coding, because it does not use coding-decoding phases and achieves higher resolution of the model space (Houck et al., 1996).

Conventional well-logging inversion methods process the data set of a certain measuring point along the borehole to determine the petrophysical model parameters only to that point. This local inversion technique represents a narrow type of the over-determined inverse problem because the number of probe types is slightly more than that of the unknowns. This inversion strategy leads to a set of separate inversion runs in adjacent depth points for the logging interval. The marginal over-determination sets a limit to the accuracy and reliability of the estimation, and does not support the determination of too many petrophysical parameters. A novel inversion methodology for solving the above problems was developed by the Department of Geophysics, University of Miskolc. The so-called interval inversion method describes the petrophysical model parameters in the form of a series expansion. By this

formulation the validity of probe response functions used in local forward modeling is extended to a greater depth interval. Thus, well-logging data of an optional depth interval can be inverted jointly to give an estimate for the model parameters to the same interval. The joint inversion procedure has at least an order of magnitude greater over-determination (data to unknowns) ratio compared to local inversion, which results in a significant improvement in the quality of interpretation.

The interval inversion method has been applied to different interpretation problems related to hydrocarbon exploration (Dobróka and Szabó, 2001; Szabó et al., 2003; Szabó, 2004; Dobróka and Szabó, 2005; Dobróka et al., 2005; Dobróka et al., 2007; Dobróka et al., 2009; Dobróka and Szabó, 2011; Dobróka et al. 2012; Dobróka and Szabó, 2012). Because of its high performance and adaptability we chose the GA for solving the interval inversion problem, which has proven to give more accurate estimates than derivative-based linear inversion procedures and results in a practically independent solution of the initial model. This chapter presents the GA-based interval inversion algorithm applied to a synthetic inversion experiment and two field cases.

2. THE INVERSION METHODOLOGY

2.1. Forward Modeling

In the forward modeling phase of the well-logging inversion procedure a theoretical data set is predicted from the petrophysical model. Consider the column vector of the model parameters

$$\mathbf{m} = \left[\Phi, S_{x0}, S_w, V_{sh}, V_{ma,l}\right]^T, \tag{2.1}$$

where Φ denotes the porosity of the geological formation, S_{x0} is the water saturation in the nearest zone around the borehole where pore space is invaded by drilling mud, S_w is the water saturation in the undisturbed zone filled with the original pore fluid (water and/or hydrocarbon), V_{sh} is the shale volume, and $V_{ma,l}$ is the volume of the l-th mineral component forming the rock matrix (superscript T denotes matrix transpose). Each parameter in (2.1) is a dimensionless quantity given in a fraction or per cent.

Well-logging surveys provide information associated with rock properties in the immediate vicinity of the borehole. Measurement types can be grouped into three classes (i.e., lithology, porosity and saturation sensitive logs). The vector for a possible data set measured from a certain depth is

$$\mathbf{d}^{(meas.)} = \left[SP, GR, \Phi_N, \rho_b, \Delta t, R_s, R_d\right]^T, \tag{2.2}$$

where *SP (mV)* denotes the spontaneous potential (lithology log), *GR (API)* is the natural γ-ray intensity (lithology log), Φ_N *(%)* is the neutron porosity (porosity log), Δt *(μs/m)* is the acoustic traveltime (porosity log), ρ_b *(g/cm^3)* is the formation density (porosity log), and R_s and R_d *(ohm-m)* are electric resistivities measured with shallow and deep penetration tools (saturation logs), respectively. The relationship between the model (2.1) and data (2.2) corrected for borehole environmental effects is called the probe response function (**g**). The data vector approximated by the calculation is:

$$\mathbf{d}^{(calc.)} = \mathbf{g}(\mathbf{m}). \tag{2.3}$$

In practice, equation (2.3) is given by a set of empirical equations including the petrophysical parameters and some local specific constants chosen according to the geological setting.

The following simplified response equations based on Poupon and Leveaux (1971) and Alberty and Hashmy (1984) can be used to calculate data in the forward modeling procedure in case of shaly-sand hydrocarbon reservoirs:

$$SP^{(calc.)} = SP_{sd} + V_{sh}\left(SP_{sh} - SP_{sd}\right), \tag{2.4}$$

$$GR^{(calc.)} = \Phi\left[GR_{mf} S_{x0} + GR_{hc}\left(1 - S_{x0}\right)\right] + V_{sh} GR_{sh} + V_{sd} GR_{sd}, \tag{2.5}$$

$$\Phi_N^{(calc.)} = \Phi\left[\Phi_{N,mf} S_{x0} + \Phi_{N,hc}\left(1 - S_{x0}\right)\right] + V_{sh} \Phi_{N,sh} + V_{sd} \Phi_{N,sd}, \tag{2.6}$$

$$\rho_b^{(calc.)} = \Phi\left[\rho_{b,mf} S_{x0} + \rho_{hc}\left(1 - S_{x0}\right)\right] + V_{sh} \rho_{b,sh} + V_{sd} \rho_{b,sd}, \tag{2.7}$$

$$\Delta t^{(calc.)} = \Phi\left[\Delta t_{mf} S_{x0} + \Delta t_{hc}\left(1 - S_{x0}\right)\right] + V_{sh} \Delta t_{sh} + V_{sd} \Delta t_{sd}, \tag{2.8}$$

$$\frac{1}{\sqrt{R_s^{(calc.)}}} = \left[\frac{V_{sh}^{(1-0.5V_{sh})}}{\sqrt{R_{sh}}} + \frac{\left(\sqrt{\Phi}\right)^m}{\sqrt{aR_{mf}}} \right] \left(\sqrt{S_{x0}}\right)^n, \tag{2.9}$$

$$\frac{1}{\sqrt{R_d^{(calc.)}}} = \left[\frac{V_{sh}^{(1-0.5V_{sh})}}{\sqrt{R_{sh}}} + \frac{\left(\sqrt{\Phi}\right)^m}{\sqrt{aR_w}} \right] \left(\sqrt{S_w}\right)^n, \tag{2.10}$$

where V_{sd} denotes the volume of quartz matrix (sand), m is the cementation exponent, n is the saturation exponent, and a is the tortuosity factor. Subscripts appearing in equations (2.4)-(2.10) denote physical properties of mud filtrate (*mf*), hydrocarbon (*hc*), shale (*sh*) and sand (*sd*), which can be found in literature or *a priori* information of the measurement area. Some response function constants can be determined by the interval inversion procedure (Dobróka and Szabó, 2011), but this chapter does not deal with that problem. Layer-thicknesses are not included explicitly in local response equations; thus, they can only be determined from a distinct procedure (e.g., by manual log analysis). Obviously, the above set of equations is non-linear with respect to the petrophysical model; thus, it is advantageous to solve the inverse problem by means of global optimization. Moreover, the simple structure of equations (2.4)-(2.10) enables the calculation of data in a relatively fast procedure, which is highly favorable for using GA.

2.2. The Inverse Problem

2.2.1. Local Inversion Methods

The non-linear well-logging inverse problem is conventionally solved by local inversion methods that separately give an estimate for the petrophysical model at adjacent depth points. The calculated data of the *j*-th well log in a given depth from equation (2.3) is

$$d_j^{(calc.)} = g_j\left(m_1,\ldots,m_M\right), \tag{2.11}$$

where M is the number of model parameters in equation (2.1). Response function constants are not indicated in equation (2.11). They are defined for a longer depth interval, describing rock formations of the same origin and pore-

fluid properties, thereby named zone parameters. Theoretical data can be calculated by equations (2.4)-(2.10). The difference between the observations and predictions is minimized in an iterative procedure. The objective function of the optimization problem is based on the Euclidean norm of the overall data misfit (Mayer and Sibbit, 1980). Since the uncertainty of data is different, it is advantageous to use weighting in the data space. The standard deviation of data, depending on the probe type and borehole conditions, can be found in the literature. The linear optimization approach is popular in industrial applications because it gives a quick solution to the inverse problem. On the other hand, in the case of the Gaussian (linear) least squares solution, the data and model variances can be related to each other; thus, the estimation errors of model parameters can be quantified (Menke, 1984). The quality check of inversion estimates is an important issue in well-log analysis.

2.2.2. Interval Inversion Method

In local inverse problems the amount of data is barely more than that of the unknowns, which results in a relatively noise sensitive inversion procedure. To preserve the over-determination of the inverse problem the zone parameters are required to be fixed in the procedure. The accuracy of the solution highly depends on the noise level of data and the initial guess of the petrophysical model. To improve the over-determination ratio we extend the validity of function (2.11) to a greater depth interval,

$$d_j^{(calc.)}(z) = g_j\left(m_1(z), m_2(z), \ldots, m_M(z)\right), \tag{2.12}$$

which can be used to calculate the *j*-th data at an arbitrary depth. In equation (2.12) model parameters are represented by continuous functions that have to be properly discretized. (In the general case zone parameters are also the function of depth.) The interval inversion method is based on the series expansion of petrophysical parameters (Dobróka, 1993):

$$m_i(z) = \sum_{q=1}^{Q_i} B_q^{(i)} \Psi_q(z), \tag{2.13}$$

where m_i denotes the *i*-th model parameter, B_q is the *q*-th expansion coefficient and Ψ_q is the *q*-th basis function (Q_i is the requisite number of expansion coefficients describing the *i*-th model parameter). Basis functions are assumed to be known quantities, which can be arbitrarily chosen. In this chapter, we

present the simplest case when layer-wise homogeneous layers build up a sequence. The following combination of unit step functions divides the processed depth interval into homogeneous segments

$$\Psi_q(z) = u(z - Z_{q-1}) - u(z - Z_q), \tag{2.14}$$

where Z_{q-1} and Z_q are the upper and lower depth coordinates of the q-th layer, respectively. Since $\Psi_q(z)$ is always zero, except in the q-th layer (where $\Psi_q(z)=1$), each petrophysical parameter can be described by one series expansion coefficient $m_i\left(z_{q-1} \le z \le z_q\right) = B_q^{(i)}$. The model parameter vector of the inverse problem is

$$\mathbf{m} = \left[B_1^{(1)},\ldots,B_Q^{(1)},\ldots\ldots,B_1^{(M)},\ldots,B_Q^{(M)},Z_1,\ldots,Z_{Q-1}\right]^T, \tag{2.15}$$

where Q is the number of layers. The advantage of the interval-wise homogeneous model approximation is that an amount of data significantly higher than that of the unknowns can be integrated into a greatly over-determined inversion procedure. The latter also provides the opportunity to treat layer-boundary coordinates as inversion unknowns, which is not possible to extract by local inversion. The inversion method developed for the automatic estimation of layer-thicknesses was introduced by Dobróka and Szabó (2012).

Well-logging data have different magnitudes and measurement units. We introduce a weighted objective function based on the least squares criterion for the interval inversion process:

$$E(\mathbf{m}) = \sum_{p=1}^{P}\sum_{j=1}^{N}\left(\frac{d_{pj}^{(meas.)} - d_{pj}^{(calc.)}}{d_{pj}^{(meas.)}}\right)^2 = \min, \tag{2.16}$$

where $d_{pj}^{(meas.)}$ and $d_{pj}^{(calc.)}$ are the j-th measured and calculated data in the p-th depth, respectively (P is the number of measuring points in the processed interval and N is the number of logging instruments). The optimal values of the series expansion coefficients in equation (2.15) are estimated by the optimization of function E, which are then substituted into equation (2.13) to produce well logs of petrophysical parameters.

3. GENETIC ALGORITHM OPTIMIZATION

The Genetic Algorithm search is based on an analogy to the process of natural selection of living populations (Holland, 1975). For optimization problems, the model can be considered as an individual of an artificial population, where the fitness characterizes the goodness of the solution. A real-valued GA approach, based on Michalewicz (1992), is applied for seeking the absolute minimum of the objective function of the well-logging inverse problem. The procedure operates with model parameters encoded as floating-point numbers that provide the highest precision estimation and best CPU time performance of all GAs.

The survival capability of an individual is characterized by a fitness value, which shows its chance of reproduction into the next generation. The fitness function is inversely connected to objective function (2.16):

$$F(\mathbf{m}) = \frac{1}{\sum_{p=1}^{P}\sum_{j=1}^{N}\left(\frac{d_{pj}^{(meas.)} - d_{pj}^{(calc.)}}{d_{pj}^{(meas.)}}\right)^2 + \varepsilon^2} = \max , \qquad (3.1)$$

where ε^2 is a small positive number used for setting a limit of model fitness. Function (3.1) is inversely proportional to the Euclidean norm of data misfits. In the case of non-Gaussian data sets, including outliers, the L_1-norm based fitness function gives the more robust solution,

$$F(\mathbf{m}) = \frac{1}{\sum_{p=1}^{P}\sum_{j=1}^{N}\left|\frac{d_{pj}^{(meas.)} - d_{pj}^{(calc.)}}{d_{pj}^{(meas.)}}\right| + \varepsilon^2} = \max . \qquad (3.2)$$

The iteration procedure is convergent if the average fitness of the individuals increases in successive generations. This is assured by the appropriate choice of genetic operations and their control parameters.

After setting up an initial population of petrophysical models, a random search is launched for finding the optimal solution. In the procedure, three types of genetic operations are performed on the current population: selection, crossover and mutation. At first, the fittest individuals are selected from the

population. We apply a normalized geometric ranking selection that requires the individuals to be sorted according to their fitness values. The rank of the best individual is 1 and that of the worst is S, which is the size of the population. The probability of selecting the x-th individual is

$$P_x = \frac{p_b}{1-(1-p_b)^S}(1-p_b)^{r_x-1}, \tag{3.3}$$

where r_x is the rank of the x-th individual, and p_b is the probability of selecting the best individual.

The latter quantity is a control parameter which has to be set in advance. The x-th individual is selected and copied into the new population when the following condition fulfills

$$C_{x-1} < U \le C_x, \tag{3.4}$$

where U is a uniform random number between 0 and 1, and

$$C_x = \sum_{y=1}^{x} P_y \tag{3.5}$$

is the cumulative probability of the population.

In the next step, a pair of individuals is chosen from the previously selected population and a partial information exchange is made between them. The simple crossover operator recombines the original individuals to create new ones:

$$\begin{aligned} \mathbf{m}^{(1,new)} &= \begin{cases} m_i^{(1,old)} & \text{if } i < k \\ m_i^{(2,old)} & \text{otherwise} \end{cases} \\ \mathbf{m}^{(2,new)} &= \begin{cases} m_i^{(2,old)} & \text{if } i < k \\ m_i^{(1,old)} & \text{otherwise} \end{cases} \end{aligned}, \tag{3.6}$$

where k is the position of the crossing point, and $m_i^{(old)}$ and $m_i^{(new)}$ denote the i-th model parameter before and after crossover, respectively.

The third genetic operator is mutation, which selects an individual from the population and changes one (or more) of its model parameters to a random number. A simple mutation operator modifies the original model as follows:

$$\mathbf{m}^{(new)} = \begin{cases} u & \text{if } i = k \\ m_i^{(old)} & \text{otherwise} \end{cases}, \tag{3.7}$$

where u is a uniform random number generated from the allowed range of the k-th model parameter. More types of real genetic operations are detailed in Houck et al. (1996).

A quality check of individuals during the inversion procedure is also required. We use two quantities for the characterization of a misfit. The relative data distance based on equation (2.16) is the measure of fit between measured and calculated data

$$D_d = \sqrt{\frac{1}{P \cdot N} \sum_{p=1}^{P} \sum_{j=1}^{N} \left(\frac{d_{pj}^{(meas.)} - d_{pj}^{(calc.)}}{d_{pj}^{(meas.)}} \right)^2} . \tag{3.8}$$

The relative model distance characterizes the goodness of the estimated model in cases of synthetic inversion experiments:

$$D_m = \sqrt{\frac{1}{QM} \sum_{q=1}^{Q} \sum_{i=1}^{M} \left(\frac{m_{qi}^{(ex.)} - m_{qi}^{(est.)}}{m_{qi}^{(ex.)}} \right)^2} , \tag{3.9}$$

where $m_{qi}^{(est.)}$ and $m_{qi}^{(ex.)}$ are the i-th estimated and exactly known model parameters in the q-th interval, respectively (Q is the number of homogeneous layers along the logging interval and M is the number of model parameters). Quantities (3.8) and (3.9) are usually multiplied by 100 to measure the misfit percentage.

Figure 1 illustrates the workflow of the GA-based inversion procedure. After collecting all measurement data and *a priori* information we set the searching intervals of the parameters of the petrophysical model. Normally, an initial population of 30-100 models is randomly generated from the search space. In the forward modeling phase of the inversion procedure, a synthetic

well-logging data set is calculated for each model. The data sets are then compared with real measurements. As long as the misfit between the observation and prediction is too high, the model population is improved by using the three genetic operations. The development of convergence is highly dependent on the proper setting of control parameters of genetic operators. The searching process stops when a termination criterion is met. This can be defined by the maximum number of iterations (generations) or by a specified threshold in the distance between measured and calculated data. In the last iteration step we accept the fittest individual of the generation as the optimal solution.

The above presented GA-based inversion scheme forms an effective algorithm, but it requires more CPU time than Simulated Annealing or linear optimization methods. This is realated to the processor speed, which can be improved by using stronger workstations. Another disadvantage of GA is the possibility of a lack of quality control from one inversion program run.

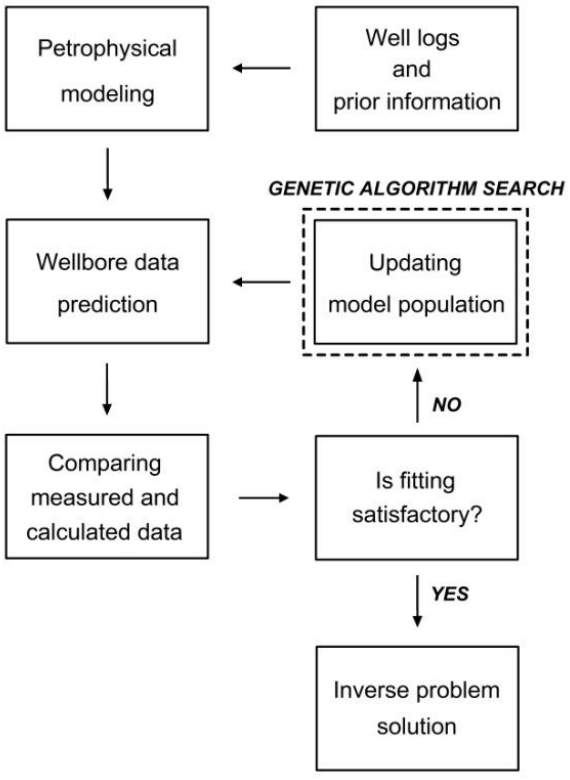

Figure 1. The scheme of GA-based inversion of well-logging data.

Linear optimization methods are based on the calculation of numerical derivatives of data with respect to model parameters, which can be applied to derive estimation errors of model parameters (Menke, 1984). These quantities cannot be extracted by a single GA process, but by using the subsequent combination of GA and a proper linear optimization method both the speed-up and accuracy quantification can be accomplished (Dobróka and Szabó, 2005).

4. INVERSION EXPERIMENT USING SYNTHETIC DATA

The aim of the inversion of synthetic data is to test the performance of the inversion procedure. The experiment assumes that the petrophysical model is exactly known. Synthetic well logs are calculated on the model, which are then contaminated with some amount of noise to simulate real measurements. This noisy data provides input for the GA-based inversion procedure. It can be studied how accurately and reliably the inversion parameters have returned back to the optimum, and how successfully the known model has been recovered. We used our own developed program code written in MATLAB to perform the GA inversion tests.

We modeled four shaly-sandy water and hydrocarbon-bearing (HC) formations with 6m, 2m, 8m and 4m thicknesses. The parameters of the layer-wise homogeneous model are the unknowns of the inverse problem (Table 1).

Table 1. The petrophysical model for GA inversion of synthetic data

Layer	Z(m)	Φ	S_{x0}	S_w	V_{sh}	V_{sd}	Lithology
1	6.0	0.2	0.8	0.4	0.3	0.5	Shaly sand HC reservoir
2	8.0	0.1	1.0	1.0	0.8	0.1	Shale with small sand
3	16.0	0.3	0.8	0.3	0.1	0.6	Sand HC reservoir with small shale
4	20.0	0.1	1.0	1.0	0.6	0.3	Sandy shale

The synthetic data set was generated by using equations (2.4)-(2.10), which was contaminated by 5% Gaussian distributed noise. The depth interval between adjacent observation points was 0.1 m. The noisy synthetic well logs are in Figure 2. In the case of local inversion we had 5 unknowns and 7 pieces of data in each measuring point, whereas the interval inversion treated 24 unknowns and 1,400 pieces of data in the 20m long interval. It was expected

that the significant increase of the data-to-unknowns (over-determination) ratio (from 7/5=1.4 to 1400/24=58.3) would make an improvement in the accuracy of inversion estimates.

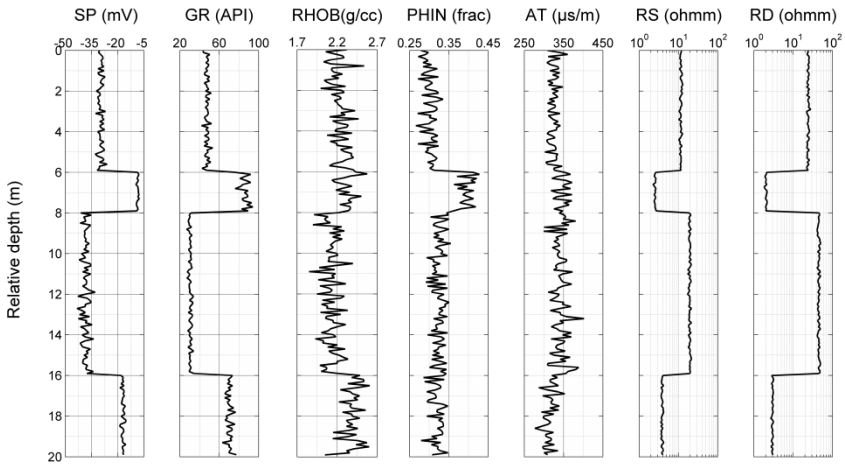

Figure 2. The synthetic well-logging data set contaminated with 5% Gaussian noise serving as input for the GA inversion experiment. Denotations: spontaneous potential (SP), natural γ-ray intensity (GR), density (RHOB), neutron porosity (PHIN), acoustic traveltime (AT), shallow resistivity (RS), deep resistivity (RD).

At first the well logs were processed by the GA-based local inversion method. The search interval of each model parameter was defined between 0 and 1. In each depth, the maximum number of generations was set to 3,500 (iteration steps). The initial population was set to 20 individuals. Real-valued genetic operators (3.3), (3.6), (3.7) were used to improve the petrophysical model population. The control parameters of genetic operators were the probability of selecting the best individual (p_b=0.08) and crossover retry (50) used when an individual is out of bounds, defined by the material balance equation and mutation probability (p_m=0.05). An elitism-based reproduction was applied as the individual with the maximum fitness value was automatically copied into the next generation. As an example, the relative data distances calculated by equation (3.8) in the first measuring point are listed in Table 2. In the first iteration step the average value of data distances was D_d=167.3% beside the standard deviation σ=360.1. Individual no. 12 represents a physically non-realistic model, which does not satisfy the material balance equation. These models are selected out of the population at once. In

the 150th generation the average data distance and standard deviation decreased to D_d=23.7% and σ=26.3, respectively. As the process converged towards the optimum, more and more individuals of high fitness values appeared in the population. The average data and model distance of the optimal model were D_d=5.6% and D_m=2.3%, respectively.

Table 2. The selected model generations in the local GA inversion procedure

Ordinal number of individual	Data distance (1st generation)	Data distance (150th generation)
1	96.4 %	10.2 %
2	29.5 %	10.2 %
3	30.3 %	111.0 %
4	96.1 %	17.6 %
5	30.3 %	6.0 %
6	98.9 %	13.6 %
7	83.8 %	7.6 %
8	89.7 %	64.5 %
9	156.7 %	7.8 %
10	122.8 %	7.8 %
11	112.4 %	46.1 %
12	1,690.2 %	46.1 %
13	156.3 %	38.4 %
14	96.0 %	16.4 %
15	75.0 %	16.4 %
16	71.7 %	17.7 %
17	80.9 %	6.0 %
18	68.6 %	6.0 %
19	80.9 %	18.3 %
20	80.9 %	6.0 %

The local inversion results are in Figure 3/a-b, which are as noisy as the input data. From the point of view of the hydrocarbon reserve calculation, further important petrophysical quantities can be derived from inversion estimates. The irreducible ($S_{hc,irr}$) and movable hydrocarbon saturations ($S_{hc,m}$) are:

$$S_{hc,irr} = 1 - S_{x0}\,, \tag{4.1}$$

$$S_{hc,m} = S_{hc} - S_{hc,irr} = (1 - S_w) - (1 - S_{x0}) = S_{x0} - S_w , \qquad (4.2)$$

which are also uncertain at low over-determination ratios. In Figure 3 the above quantities are also plotted. The hydrocarbon saturation is 0% when the pores are completely filled with water (S_{x0}=S_w=1).

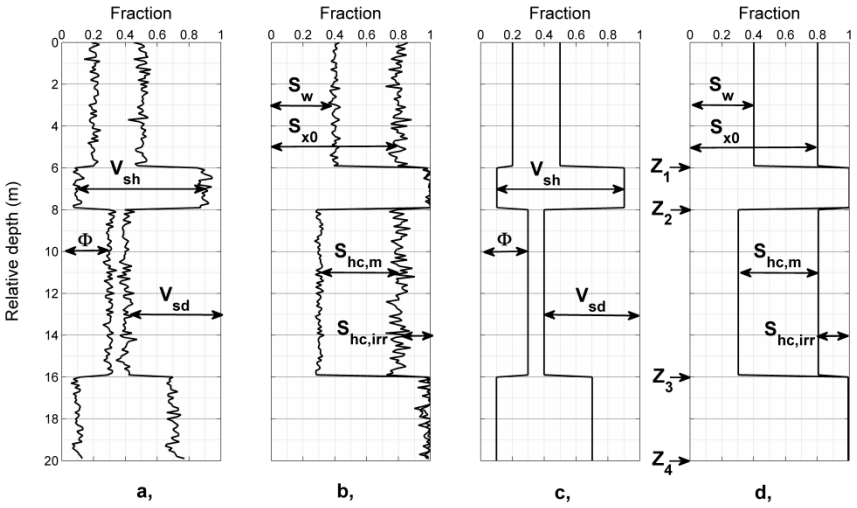

Figure 3. The GA inversion outputs for a layer-wise homogeneous petrophysical model: a-b, local inversion results, c-d, interval inversion results. Denotations: porosity (Φ), shale volume (V_{sh}), sand volume (V_{sd}), water saturation in the undisturbed zone (S_w), water saturation in the flushed zone (S_{x0}), movable hydrocarbon saturation ($S_{hc,m}$), irreducible hydrocarbon saturation ($S_{hc,irr}$), layer-boundary coordinate (Z).

In the next step, the interval inversion method was tested. GA was used to maximize function (3.1). We set the maximal number of generations to 30,000 and fixed the size of the model population to 20 individuals. Genetic operators (3.3), (3.6) and (3.7) were implemented with the same values of control parameters as in the local inversion procedure. Not only were the petrophysical parameters allowed to vary in the procedure, but layer-thicknesses were allowed to vary, too. The search space of layer-thicknesses was specified over the domain of real numbers between 0.1 and 10 (m).

The samples of petrophysical variables were chosen from the interval of 0 and 1. The gradual decrease of data and model distances can be seen in Figure

4. Around iteration step 90, the model distance curve calculated by equation (3.9) shows an escape from a local minimum region.

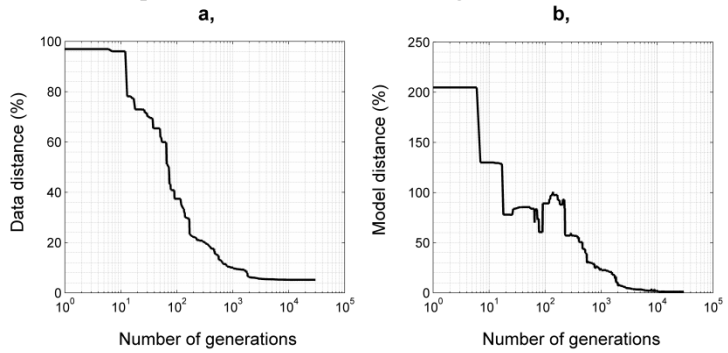

Figure 4. The convergence plots of the GA-based interval inversion procedure: a, data distance of the best individual vs. iteration step, b, model distance of the best individual vs. iteration step.

This event is connected to the finding of optimal layer boundary coordinates. After the 300th generation, only the petrophysical parameters were refined, which proved the stability of the interval inversion procedure.

The optimum was found in the last generation, where the average data distance stabilized around the noise level (D_d=5.6%) and at an average model distance D_m=1.1%. The interval inversion results, including reservoir parameters and formation boundary-coordinates, are shown in Figure 3/c-d. The estimated formation thicknesses were accurate to the third decimal place. Petrophysical model parameters were also obtained close to their exact values (Table 1). The target model represents a real petroleum geological situation, to which a similar field example can be found in Section 5.2. It can be mentioned that the over-determination ratio of the interval inversion problem decreases from 70 to 58.3 by treating formation thicknesses as an unknown model parameter during inversion, but it does not cause a significant quality loss in the overall estimation. This case is still more favorable than the marginally over-determined local inverse problem. By comparing local and interval inversion results, it can be stated that the latter gives more accurate estimates of movable and irreducible hydrocarbon reserves in the case of near-homogeneous reservoirs.

5. FIELD CASES

The test sites for inversion method development were chosen in the Great Hungarian Plain, Hungary. The study areas are part of the Pannonian Basin Province of Central Europe, a highly explored area, where many petroleum systems have been discovered. In the Pannonian Basin a thick Tertiary sedimentary sequence overlays the Mesozoic, Paleozoic and Precambrian basement. Most hydrocarbon reservoirs in the area are high or medium-porosity sandstones interbedded with clay, silt, marl and other layers. The oil- and gas-bearing formations situated mostly between 1,000-3,000m in depth represent a wide variety of structural, stratigraphical and combined traps in the province (Dolton 2006). Above the Tertiary basin-fill sequence, there are various Pannonian aged postrift sediments including thermal-water resources. The succeeding Quaternary sediments are characterized by a high variety of paludal, fluvial and delta-plain deposits including thick gravel and sand deposits saturated with fresh water. In this chapter, two case studies about the interval inversion processing of well-logging data originating from shallow water prospecting and deep hydrocarbon wells are presented.

5.1. Application to Water Prospecting

The first example shows the inversion processing of well-logging data collected from a water prospecting borehole (Well-1). As the Earth's crust is relatively thin compared to surrounding mountainous areas, there is a high heat flow in the region. The high amount of thermal water reserves is the source of a very significant and perspective geothermal energy. On the other hand, the same water is pumped into injection wells to improve oil recovery from the underlying hydrocarbon reservoirs. We processed the data of a short 19.4-meter interval from the thermal water zone. The well logs involved in the interval inversion procedure are in Figure 5. Measurement types can be found in equation (2.2). (The microlaterolog - *RMLL* - is a focused type of resistivity measurement with a very shallow radial depth of penetration.)

We assumed four beds by the GR log. Figure 6/a shows the result of manual separation of sedimentary layers. The lithology is also included in the figure, which was confirmed by prior geological knowledge of the area. We applied the interval inversion method to determine porosity, shale content, sand volume and layer boundaries. The GA search was performed with a population of 20 individuals and 12,000 generations. Genetic operators (3.3), (3.6), (3.7) were applied by the following control parameters: probability of selecting the best individual (p_b=0.08), crossover retry (200) and mutation

probability (p_m=0.05). The decrease of data distance during the inversion procedure can be seen in Figure 7.

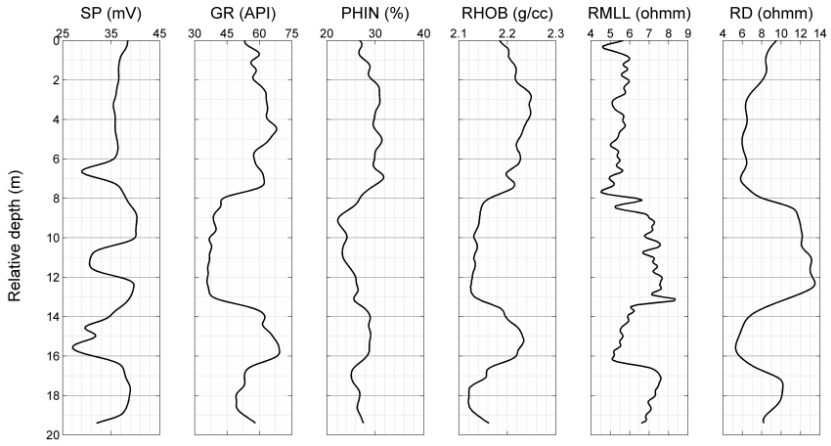

Figure 5. The well logs measured in Well-1 as input for the GA-based interval inversion procedure. Denotations: spontaneous potential (SP), natural γ-ray intensity (GR), density (RHOB), neutron porosity (PHIN), microlaterolog (RMLL), deep resistivity (RD).

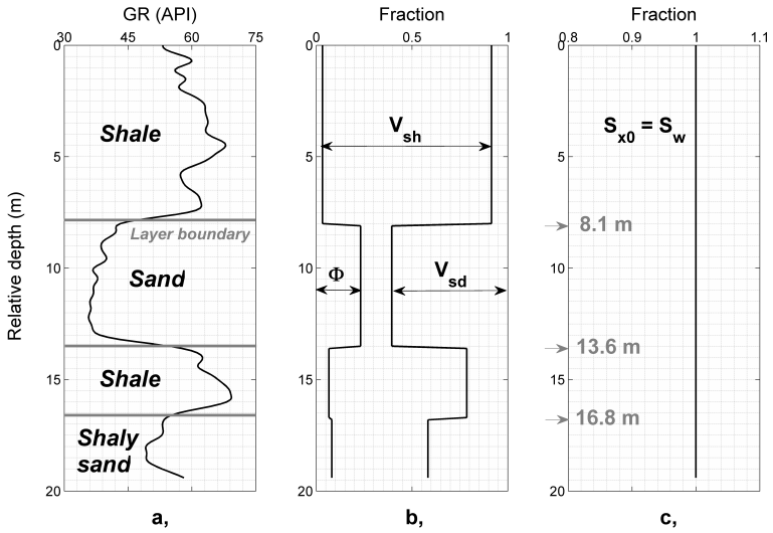

Figure 6. The results of the GA-based interval inversion procedure in Well-1: a, natural γ-ray log (GR) with lithology and manually traced out layer-boundaries; b, estimated

porosity (Φ), shale volume (V_{sh}), sand volume (V_{sd}); c, water saturation in the flushed (S_{x0}) and undisturbed zone (S_w), automatically estimated layer-boundary coordinates (indicated by arrows).

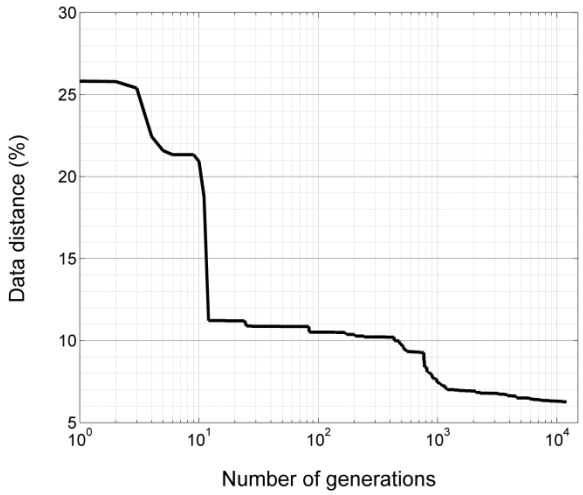

Figure 7. The convergence plot of the GA-based interval inversion procedure in Well-1: data distance of the best individual vs. iteration step.

At the end of the parameter search, a satisfactory fit in data space was achieved as the relative data distance was D_d=6.3%. (The model distance is impossible to compute for a lack of an exact model.)

The inversion results are in Figure 6/b-c. Among the shaly sandy layers there is a productive water-bearing reservoir with high porosity and permeability (second layer). The automatically estimated layer thicknesses are 8.1m, 5.5m, 3.2m and 2.6m, in close agreement with manual results. Since the pores are fully saturated with water along the entire interval, then S_{x0}=S_w=1. Shale particles dispersed in the pore space have large specific surface areas, which can absorb a high amount of water. Another type of clay incorporated in the rock matrix may also contain water in their mineral structure. Therefore, along the intervals of shale layers, a 100% water saturation is indicated.

5.1. Application to Hydrocarbon Exploration

The feasibility of the GA-based interval inversion method was tested by using another real well-logging data set originated from a deep Hungarian

hydrocarbon exploratory borehole (Well-2). The aim of the interpretation was to find and locate possible hydrocarbon reservoirs and estimate their storage capacity. We processed a 35-meter long interval of a hydrocarbon zone, where the sampling interval of well logging was 0.1 m. The inverted open-hole logs are in Figure 8.

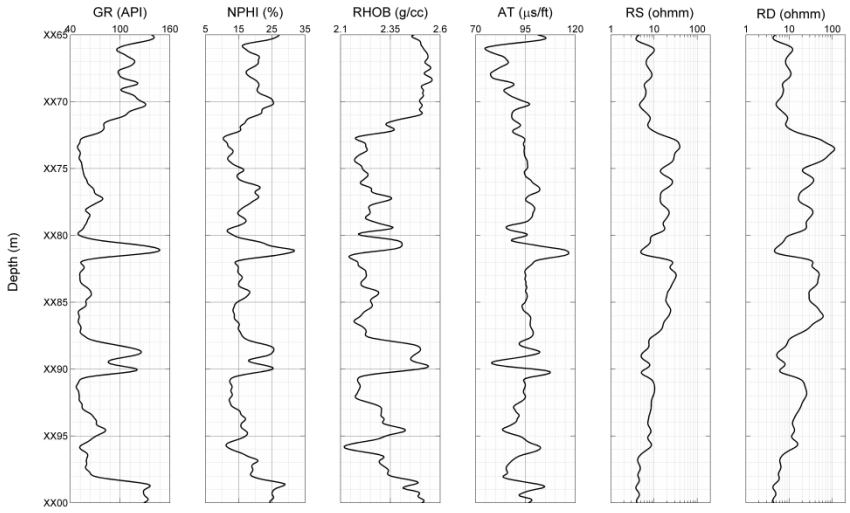

Figure 8. The well logs measured in Well-2 as input for the GA-based interval inversion procedure. Denotations: natural γ-ray intensity (GR), neutron porosity (NPHI), density (RHOB), acoustic traveltime (AT), shallow resistivity (RS), deep resistivity (RD).

Seven homogeneous formations were assumed by a preliminary study of well logs. The sequence consisted of gas-bearing shaly-sand formations with interbedded shale layers. The GA search required 10,000 generations for updating 30 individuals. The control parameters of genetic operators (3.3), (3.6), (3.7) were the probability of selecting the best individual (p_b=0.08), crossover retry (150) and mutation probability (p_m=0.03). The convergence of the inversion procedure was smooth and steady, as it is seen in Figure 9. The optimal solution was found at relative data distance D_d=7.2%.

The logs of estimated petrophysical parameters are in Figure 10/b-c. The method distinguished permeable and impermeable intervals and gave an estimate for rock interfaces at nearly the same places as they were inferred from the *GR* log.

Along the processed interval the evaluation revealed three hydrocarbon reservoirs with high porosity separated by shale beds (Figure 10/a). The estimated thicknesses of the layers were 7.3m, 7.8m, 1.6m, 6.0m, 2.0m and 8.1m.

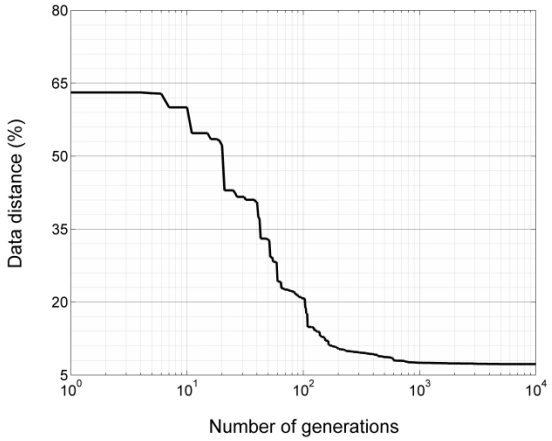

Figure 9. The convergence plot of the GA-based interval inversion procedure in Well-2: data distance of the best individual vs. iteration step.

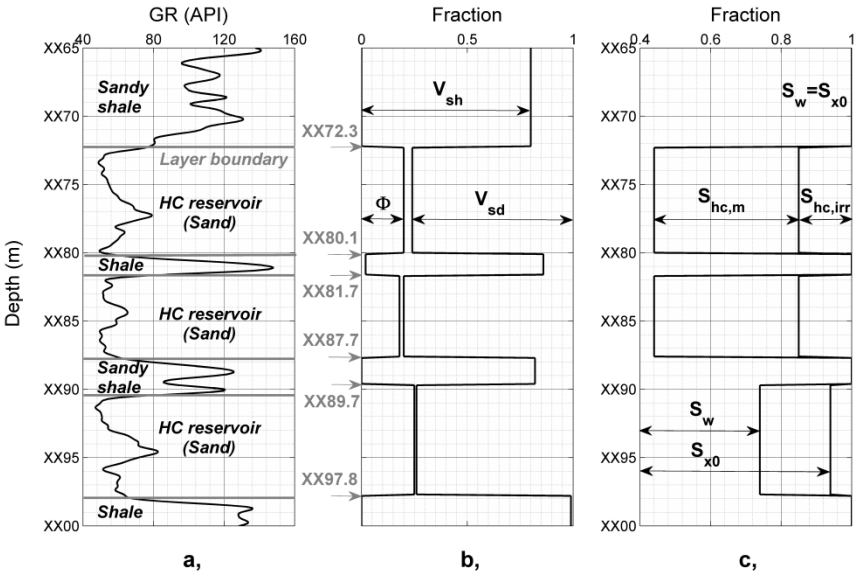

Figure 10. The GA-based interval inversion results in Well-2: a, natural γ-ray log with lithology and manually located layer-boundaries; b, estimated porosity (Φ), shale

volume (V_{sh}), sand volume (V_{sd}) and automatically estimated layer-boundary coordinates (indicated by arrows); c, water saturation in the flushed (S_{x0}) and undisturbed zones (S_w), movable hydrocarbon saturation ($S_{hc,m}$), irreducible hydrocarbon saturation ($S_{hc,irr}$).

The movable and irreducible gas saturation based on equations (4.1) and (4.2) were approximately 41% and 15% (first and second hydrocarbon reservoirs), 21% and 7% (third hydrocarbon reservoir), respectively.

The estimation of permeability as a derived parameter from inverse modeling was also made by using an empirical formula including porosity and other properly chosen constants (Timur, 1968). We calculated permeability values from 1,300 to 1,900 mDarcies in the porous and permeable rocks. It was concluded that the results of interpretation showed good quality hydrocarbon reservoirs.

CONCLUSION

In this chapter, the Float-Encoded Genetic Algorithm was presented as an effective global optimization tool used for the inversion processing of well-logging data. The interval inversion method gives an estimate for petrophysical parameters and the coordinates of formation boundaries at arbitrary depth intervals in a joint inversion procedure. The highly over-determined inverse problem results in an accurate, reliable and robust solution for the logged interval instead of making a prediction in separate measuring points. As the procedure of forward modeling is not very time-consuming, the CPU times take only several minutes by using a quad-core processor-based workstation, which can be further reduced by applying a subsequent combination of GA and a fast linear optimization method.

The purport of the interval inversion technique is to use series expansion for discretizing the model parameters. Assuming homogeneous layers, the interval inversion method enables a more accurate and reliable separation of movable and irreducible hydrocarbon volumes than by conventional (local) inversion methods. In the case of inhomogeneous layers, a higher resolution of petrophysical parameters can be achieved by using power or other appropriate types of basis functions. The interval inversion method can make an automatic prediction of layer boundary coordinates. The effective layer thickness is an important parameter in calculating the producible water or hydrocarbon reserves. The capability of giving an automatic estimate to geometrical quantities sets new perspectives to describe the morphology of geological

structures in more detail. The interval inversion algorithm can be developed for multi-well applications by using basis functions depending on more spatial coordinates. In that case, the formation boundaries can be approximated by appropriately chosen two or three dimensional basis functions. Thus, the surface and the volume of rock structures can be determined automatically. The method is supported by large amounts of wellbore information and large extents of over-determination. The latter is not reduced significantly by introducing some additional unknowns into the inverse problem. Therefore, some other zone parameters, routinely treated as local specific constants, can be estimated within the inversion procedure. As a result, the areal distributions of petrophysical parameters can be given from a more objective source that supports all fields of Earth science with useful *in-situ* information.

ACKNOWLEDGMENTS

The described work was carried out as part of the TÁMOP-4.2.2/B-10/1-2010-0008 project in the framework of the New Hungary Development Plan. The realization of this project is supported by the European Union, co-financed by the European Social Fund. The first author thanks to the support of the János Bolyai Research Fellowship of the Hungarian Academy of Sciences. The second author is grateful for the support of the MTA-ME Research Group for Earth Sciences.

REFERENCES

Alberty, M.; Hashmy, K. *SPWLA Symposium Transactions*. 1984, Z, 1-17.

Asquith, G.; Krygowski, D. Basic well log analysis; AAPG Methods in Exploration 16; *American Association of Petroleum Geologists,* Tulsa, 2004.

Ball, S. M.; Chace, D. M.; Fertl, W. H. Proc. *SPE Eastern Regional Meeting*. 1987, 17034, 61-85.

Dobróka, M. Scientific Report for the Hungarian Oil and Gas Company. 1993, *University of Miskolc.*

Dobróka, M.; Szabó, N. P. Geosciences, *Publications of the University of Miskolc*. 2001, 59, 115-137.

Dobróka, M.; Szabó, N. P. *Acta Geodaetica et Geophysica Hungarica*. 2005, 40, 203-214.

Dobróka, M.; Szabó, P. N.; Kiss, B.; Krasznavölgyi, *Proc. 67th EAGE Conference and Exhibition.* 2005, 215, 1-4.

Dobroka, M.; Kiss, B.; Szabo, N.; Toth, J.; Ormos, T. *Proc. 69th EAGE Conference and Exhibition*. 2007, 092, 1-5.

Dobróka, M.; Szabó, P. N.; Cardarelli, E.; Vass, P. *Acta Geodaetica et Geophysica Hungarica*. 2009, 44, 459-479.

Dobróka, M.; Szabó, N. P. Acta Geophysica. 2011, 59, 907-934.

Dobróka, M.; Szabó, N. P.; Turai, E. *Acta Geodaetica et Geophysica Hungarica*. 2012, 47, 172-184.

Dobróka, M.; Szabó, N. P. *Journal of Petroleum Science and Engineering*. 2012, 86–87, 144-152.

Dolton, G. L. USGS Bulletin 2204–B. 2006, 1-47.

Holland, J. H. Adaptation in natural and artificial systems; Univ. Michigan Press, *Ann Arbor*, 1975.

Houck, C. R.; Joines, J. A.; Kay M. G. *Technical report.* 1995, 1-14.

Mayer, C.; Sibbit, *A. Proc. SPE Annual Fall Technical Conference and Exhibition.* 1980, 9341, 1-14.

Menke, W. Geophysical data analysis: discrete inverse theory; Academic Press, New York, 1984.

Metropolis, N.; Rosenbluth, A.; Rosenbluth, M.; Teller, A.; Teller, E. *J. Chem. Phys.* 1953, 21, 1087–1092.

Michalewicz, Z. Genetic Algorithms + Data Structures = Evolution Programs; Springer-Verlag, New York, 1992.

Poupon, A.; Leveaux, J. *The Log Analyst*. 1971, 12, 3-8.

Sen M.; Stoffa P. L. In Seismic Exploration CD-ROM; Helbig, K.; Treitel, S.; Elsevier Science, 1997.

Serra, O. Fundamentals of well-log interpretation; Elsevier, Amsterdam, 1984.

Szabó, P. N.; Dobróka, M.; Hursán, L. Proc. *65th EAGE Conference and Exhibition*. 2003, 261, 1-4.

Szabó, N. P. *Geophysical Transactions*. 2004, 44, 313-329.

Tarantola, A. Inverse problem theory and methods for model parameter estimation; *Society for Industrial and Applied Mathematics*, Philadelphia, 2005.

Telford, W. M.; Geldart, L. P.; Sheriff, R. E. Applied Geophysics; Cambridge University Press, 1990.

Timur, A. *The Log Analyst.* 1968, 9, 8-17.

LCH