Math Geosci
https://doi.org/10.1007/s11004-018-9768-4

Series Expansion-Based Genetic Inversion of Wireline Logging Data

Norbert Péter Szabó[1,2] · **Mihály Dobróka**[1]

Received: 24 November 2017 / Accepted: 8 October 2018

Abstract An evolutionary approach is applied to solve the nonlinear well logging inverse problem. In the framework of the proposed interval inversion method, nuclear, sonic, and laterolog resistivity data measured at an arbitrary depth interval are jointly inverted, where the depth variation of porosity, water saturation, and shale volume is expanded into series using Legendre polynomials as basis functions. In the interval inversion procedure, the series expansion coefficients are estimated by using an adaptive float-encoded genetic algorithm. Since the solution of the inverse problem using traditional linear optimization tools highly depends on the selection of the initial model, a heuristic search is necessary to reduce the initial model dependence of the interval inversion procedure. The genetic inversion strategy used in interval inversion seeks the global extreme of the objective function and provides an estimate of the vertical distribution of petrophysical parameters, even starting the inversion procedure from extremely high distances from the optimum. For a faster computational process, after a couple of thousand generations, the genetic algorithm is replaced by some linear optimization steps. The added advantage of using the Marquardt algorithm is the possibility to characterize the accuracy of the series expansion coefficients and derived petrophysical properties. A Hungarian oil field example demonstrates the feasibility and stability of the improved interval inversion method. As a significance, the genetic inversion method does not require prior knowledge or strong restrictions on the

✉ Norbert Péter Szabó
norbert.szabo.phd@gmail.com

Mihály Dobróka
dobroka@uni-miskolc.hu

[1] Department of Geophysics, University of Miskolc, Miskolc-Egyetemváros 3515, Hungary

[2] MTA-ME Geoengineering Research Group, University of Miskolc, Miskolc-Egyetemváros 3515, Hungary

Journal: 11004 Article No.: 9768 TYPESET DISK LE CP Disp.:2018/10/16 Pages: 25 Layout: Small-X

values of petrophysical properties and gives highly reliable estimation results practically independent of the initial model and core information.

Keywords Float-encoded genetic algorithm · Interval inversion · Series expansion · Petrophysical parameter · Hungarian oil fields

1 Introduction

In oil field petrophysics, well logging inversion has been an integral part of computer-processed log analysis for more than three decades, the traditional method of which is based on the linearization of the functional relation between the measurements and the estimated petrophysical model (Mayer and Sibbit 1980). This linear optimization approach is widely used in formation evaluation, which offers a fast and satisfactory solution when one has appropriate prior information on the petrophysical characteristics of the hydrocarbon reservoir. The theory and geophysical aspects of linear inversion methods are detailed in Menke (1984) and Tarantola (2005). The strongest limitation of the gradient-based search techniques is that they usually find a local minimum of the objective function in the near vicinity of the initial model. To improve the performance of inverse modeling in terms of optimization, Sen and Stoffa (1995) recommended the use of global optimization methods such as simulated annealing and genetic algorithm. Since these optimization techniques do not require the calculation of the Jacobi matrix and look for the solution not only in the direction of the local minimum, they can wander freely in the parameter space and try many possible solutions. During their random search, not only the best models but also some of those with worse data misfit are temporarily accepted to avoid the localities and ambiguous solutions. Global optimization methods are able to solve constrained and strongly nonlinear inverse problems with high efficiency and adaptability, as they are capable of finding better solutions than local-optimization methods without any curvature information of the misfit function and a good starting model (Sambridge and Drijkoningen 1992; Mundim et al. 1998; Tlas and Asfahani 2008; Yang 2013; Toushmalani 2013). However, the global inversion techniques are computationally intensive; thus, they can be used most preferably in case of quickly computable forward problems and relatively small number of model parameters. For these reasons, their application provides several perspectives to well logging inversion (Szűcs and Civan 1996; Szabó 2004; Dobróka and Szabó 2012; Malkoti et al. 2017). The rapid development of computational power and capacity allows the ever-expanding use of global optimization techniques and similar metaheuristic approaches for solving complex optimization problems in petroleum geosciences (Cranganu et al. 2015). In well log analysis, for example, they have already been used in finite element modeling-assisted three-dimensional inversion of DC resistivity measurements (Gajda-Zagórska et al. 2015), in reservoir characterization using the combination of genetic algorithm and particle swarm optimization (Ali Ahmadi et al. 2013), in neural-network prediction of wireline logs from seismic attributes (Dorrington and Link 2004), in estimating the seismic velocities from well logs (Aleardi 2015), and in a genetic algorithm-assisted

Journal: 11004 Article No.: 9768 TYPESET DISK LE CP Disp.:2018/10/16 Pages: 25 Layout: Small-X

factor analysis to estimate the shale volume and permeability in hydrocarbon reservoirs (Szabó and Dobróka 2018).

Open-hole logging data are traditionally processed in separated inversion procedures at different depths. The inverse problem is solved in a given depth by minimizing the misfit between the predicted and observed data normalized by the data variances using the weighted least squares method (Mayer and Sibbit 1980). Several applications for this local inversion methodology have been published; for example, in the analysis of oil field well logs (Ball et al. 1987; Narayan and Yadav 2006; Woznicka et al. 2009) and direct-push logging data (Drahos 2005). Having barely more data than unknowns in the investigated depth, a set of marginally overdetermined inverse problems is to be run along the borehole. The narrow overdetermination sets a limit to the number of unknowns to be extracted and makes the inversion procedure highly sensitive to data noise. Moreover, the local inversion approach does not allow the estimation of layer thicknesses because they are not included explicitly in the local response functions, although this information is inherent in the entire well-logging data set. The marginal overdetermination of the local inverse problem is not favorable to the estimation of zone parameters. These parameters are normally fixed for the inversion process, which causes further uncertainty in the inverse problem solution.

Dobróka et al. (2016) suggested to increase the overdetermination (data-to-unknowns) ratio by using a series expansion-based inversion methodology to improve the quality of well logging inversion results. In the so-called interval inversion procedure, all data collected at a depth interval are jointly inverted to estimate the series expansion coefficients far less in number than data, which describes the vertical variation of petrophysical properties. The remarkably high overdetermination of the inverse problem significantly increases the noise suppression capability of the interval inversion method, as well as the accuracy and reliability of the inversion result. Moreover, the interval inversion procedure allows the automated estimation of such unknowns that have not been estimated by inversion; for instance, Archie's textural parameters, the physical properties of rock-forming minerals and fluids included in the sonde response functions (Dobróka and Szabó 2011), as well as the vertical coordinates of layer boundaries and their horizontal variation between the wells (Dobróka et al. 2009; Dobróka and Szabó 2012). Other statistical approaches different from the interval inversion method, which uses depth-dependent basis functions to describe the variation of petrophysical properties, are also applied for a more advanced analysis of well logs. The advantages of the introduction of a vertical continuity model in facies classification was shown by Eidsvik et al. (2004) and Lindberg and Grana (2015). They suggested modeling the continuous well-logging measurements as the result of a hidden Markov chain model for improved classification of hydrocarbon formations.

Similar to other inversion methods, the performance of the interval inversion procedure highly depends on the applied model discretization scheme. When choosing the number of unknowns, a compromise has to be made between the stability of the inversion procedure and the accuracy of the estimation results. In the first applications of interval inversion, the combination of Heaviside step (basis) functions were used to describe the petrophysical properties in sequential homogeneous layers (Dobróka et al. 2012). Since this was the case with the smallest possible number of unknowns, this approach resulted in a highly stable interval inversion procedure, but also a rough

Journal: 11004 Article No.: 9768 TYPESET DISK LE CP Disp.:2018/10/16 Pages: 25 Layout: Small-X

approximation of the variation of petrophysical properties along a borehole. In recent applications, orthogonal sets of polynomials have been selected as basis functions to model inhomogeneous formations (Dobróka et al. 2016; Szabó et al. 2018). The latter approach has proven to give higher spatial resolution of the petrophysical model and less correlated model parameters than using other basis functions; for instance, power functions or other non-orthogonal polynomials. However, it has been recently experienced that, by a large number of expansion coefficients as unknown, in certain cases, the stability of the inversion procedure is reduced and this phenomenon is highly dependent on the selection of the starting model. Although the inversion procedure can be numerically stabilized by regularization, it may still either trap in a local minimum or become divergent and continue to be sensitive to the initial model.

To avoid this difficulty, the interval inversion method is further developed by using a float-encoded genetic algorithm (FGA) suggested by Michalewicz (1992). This adaptive genetic inversion approach is used to determine the petrophysical parameters expanded into series using Legendre polynomials. In order to accelerate the time-consuming search for the global optimum, the genetic inversion process is coupled by some linear optimization steps using the damped least squares (DLSQ) method (Marquardt 1959). The sequential application of global and linear optimization techniques has been effectively used previously; for instance, in the joint inversion of geoelectric and seismic data (Chunduru et al. 1997). In the initialization phase of the genetic (interval) inversion procedure, the individuals of the initial population are generated randomly from a wide value range of model parameters (i.e., series expansion coefficients). The heuristic search improves the actual population of models, while it effectively avoids the local extrema of the objective function, defined as a norm of the difference between the measured and theoretical data vectors. At the end of the genetic process, a first approximation of the petrophysical parameters derived from the expansion coefficients is given, even wen starting the inversion procedure from extremely high distances from the global optimum. After a requisite number of generations, the FGA is replaced by the DLSQ algorithm, in which the vertical distribution of petrophysical parameters is further refined and the data misfit is quickly reduced to an optimal level. A Hungarian case study including the petrophysical evaluation of an unconsolidated hydrocarbon formation shows that the above suggested genetic inversion method is able to find the optimal values of series expansion coefficients and the related petrophysical parameters in a stable inversion procedure.

2 Theory and Methods

2.1 Forward Modeling

The interval inversion method allows the estimation of porosity, water saturation, shale content, and matrix volume in clastic hydrocarbon formations. In each generation of the genetic inversion procedure, theoretical well logs are calculated and compared to real data observed along the borehole. The synthetic logs are generated in the forward modeling procedure by assuming a known petrophysical model. For this purpose, the

Journal: 11004 Article No.: 9768 TYPESET DISK LE CP Disp.:2018/10/16 Pages: 25 Layout: Small-X

following sonde response functions are used for unconsolidated shaly-sand formations (Wyllie et al. 1956; Alberty and Hashmy 1984; Schlumberger 1989)

$$\rho_b = \Phi\left[\rho_{\mathrm{mf}} - 1.07\,(1 - S_{x0})\,(\alpha_0\rho_{\mathrm{mf}} - 1.24\rho_{\mathrm{hc}})\right] + V_{\mathrm{sh}}\rho_{\mathrm{sh}} + V_{\mathrm{sd}}\rho_{\mathrm{sd}}, \tag{1}$$

$$\mathrm{GR} = \rho_b^{-1}\left(V_{\mathrm{sh}}\mathrm{GR}_{\mathrm{sh}}\rho_{\mathrm{sh}} + V_{\mathrm{sd}}\mathrm{GR}_{\mathrm{sd}}\rho_{\mathrm{sd}}\right), \tag{2}$$

$$\Phi_{\mathrm{N}} = \Phi\left\{\begin{array}{l}\Phi_{\mathrm{N,mf}} - (1 - S_{x0})\,C_{\mathrm{cor}} - 2\Phi\,(1 - S_{x0})\,S_{\mathrm{hf}}\,(1 - 2.2\rho_{\mathrm{hc}}) \\ \cdot\left[1 - (1 - S_{x0})\,(1 - 2.2\rho_{\mathrm{hc}})\right]\end{array}\right\} + V_{\mathrm{sh}}\Phi_{\mathrm{N,sh}} + V_{\mathrm{sd}}\Phi_{\mathrm{N,sd}}, \tag{3}$$

$$\Delta t = \Phi\left[\Delta t_{\mathrm{mf}}S_{x0} + (1 - S_{x0})\,\Delta t_{\mathrm{hc}}\right]\,c_p + V_{\mathrm{sh}}\Delta t_{\mathrm{sh}} + V_{\mathrm{sd}}\Delta t_{\mathrm{sd}}, \tag{4}$$

$$R_{\mathrm{t}} = \left[\frac{\Phi^m S_{\mathrm{w}}^n}{a R_{\mathrm{w}}\,(1 - V_{\mathrm{sh}})} + \frac{V_{\mathrm{sh}}S_{\mathrm{w}}}{R_{\mathrm{sh}}}\right]^{-1}, \tag{5}$$

where the notations of calculated well log types are summarized in Table 1. The common parameters in Eqs. (1)–(5) are the volumetric fractions of the rock's fluid and solid constituents, such as shale-free porosity (Φ), shale volume (V_{sh}), sand volume (V_{sd}), and water saturation in the invaded zone (S_{x0}) and that of the virgin zone (S_{w}). In the framework of the traditional local (depth-by-depth) inversion, the physical properties of mud filtrate (mf), hydrocarbon (hc), shale (sh), and sand (sd) are treated as constants to solve a marginally overdetermined (local) inverse problem (Table 1). The textural parameters in Eq. (5), namely, the cementation exponent (m), saturation exponent (n), and tortuosity factor (a), are also fixed in local inversion (Archie 1942). These zone parameters can be set by crossplot techniques, core information, or from the literature. Some constants in Eqs. (1) and (3) come from empirical considerations used in the oil industry. Since these physical parameters do not vary as fast as the volumetric properties in a hydrocarbon formation or a zone, they are necessarily fixed for the interval inversion experiments. In this study, the genetic (interval) inversion method is suggested to determine only the volumetric properties; however, the highly overdetermined interval inversion method allows the estimation of some zone parameters simultaneously with the volumetric properties in an automated inversion procedure (Dobróka and Szabó 2011). The interval inversion framework also gives a possibility for the integration of new data types, the response functions of which can be freely chosen. The solution of the forward problem can be necessarily further improved, primarily in the case of a nonlinear resistivity response equation, with the more accurate account of borehole effects and mud invasion (Galsa et al. 2016; Jarzyna et al. 2016; Szijártó et al. 2017).

2.2 Genetic Inversion Approach

The evolutionary technique incorporating a large group of soft computing methods was first proposed by Holland (1975). The genetic algorithm is a highly adaptive and efficient optimization tool, which employs the mechanism of natural selection. In its meta-heuristic searching process, a population of models as artificial individuals is iteratively improved generation by generation. The quality of an individual is

Journal: 11004 Article No.: 9768 TYPESET DISK LE CP Disp.:2018/10/16 Pages: 25 Layout: Small-X

Table 1 Zone parameters chosen for genetic interval inversion of wireline logs measured in well H

Open-hole log	Zone parameter	Symbol	Constant	Unit
Gamma ray intensity (GR)	Sand	GR_{sd}	13	API
	Shale	GR_{sh}	160	API
Density (ρ_b)	Sand	ρ_{sd}	2.6	g/cm^3
	Shale	ρ_{sh}	2.47	g/cm^3
	Mud filtrate	ρ_{mf}	1.2	g/cm^3
	Gas	ρ_{hc}	0.135	g/cm^3
	Mud filtrate coefficient	α	1.11	–
Neutron porosity (Φ_N)	Sand	$\Phi_{N,sd}$	− 0.035	v/v
	Shale	$\Phi_{N,sh}$	0.3	v/v
	Mud filtrate	$\Phi_{N,mf}$	1	v/v
	Mud filtrate coefficient	C_{cor}	0.69	–
	Residual hydrocarbon coefficient	S_{hrf}	1.17	–
Acoustic interval time (Δt)	Sand	Δt_{sd}	56	µs/ft
	Shale	Δt_{sh}	108	µs/ft
	Mud filtrate	Δt_{mf}	188	µs/ft
	Gas	Δt_{hc}	211	µs/ft
	Compaction factor	c_p	1.08	–
True resistivity (R_t)	Shale	R_{sh}	1.0	Ωm
	Pore water	R_w	0.4	Ωm
	Mud filtrate	R_{mf}	0.28	Ωm
	Cementation exponent	m	1.4	–
	Saturation exponent	n	1.7	–
	Tortuosity factor	a	0.9	–

characterized by a fitness function specifying its survival capability. According to Darwin's theory, the fittest individuals are more likely to survive and reproduce, while others disappear from the population. The individuals are represented by chromosomes, the elements of which, called genes, are randomly exchanged and modified during the genetic computation process. The FGA is known as one of the most efficient global optimization methods, which applies real-valued operations directly on the genes for the best resolution and computational performance (Michalewicz 1992).

In this study, an evolutionary computation strategy is applied to solve the nonlinear problem of interval inversion. In the local inversion approach as a traditional method of well log analysis, the data set of a given depth is processed to estimate the petrophysical properties to the same depth in a quick linearized inversion procedure. In the interval inversion procedure, all data measured over a depth interval of arbitrary length

Journal: 11004 Article No.: 9768 TYPESET DISK LE CP Disp.:2018/10/16 Pages: 25 Layout: Small-X

is simultaneously inverted to estimate the vertical distribution of petrophysical parameters. For this purpose, the local response functions are extended to a depth-dependent form as

$$d_k^{(\mathrm{cal})}(z) = g_k\left(m_1(z), \ldots, m_i(z), \ldots, m_M(z)\right), \tag{6}$$

where m_i is the ith petrophysical property (M is the total number of petrophysical parameters) and g_k represents the response function of the kth sonde ($k = 1, 2, \ldots, K$, where K is the number of applied logging tools). The zone parameters are assumed constant (since their estimation is not the subject of the present study). Dobróka et al. (2016) introduced a series expansion technique for the discretization of model parameters, where the ith petrophysical parameter, such as porosity or shale volume, is approximated as

$$m_i(z) = \sum_{q=1}^{Q^{(i)}} B_q^{(i)} P_{q-1}(z), \tag{7}$$

where B_q is the qth expansion coefficient, $P_{q-1}(z) = P_l(z) = \left(2^l l!\right)^{-1} \frac{\mathrm{d}^l}{\mathrm{d}z^l}\left(z^2 - 1\right)^l$ is the lth degree Legendre polynomial as the basis function, and $Q^{(i)}$ is the number of expansion coefficients for describing the ith model parameter along the processed interval ($i = 1, 2, \ldots, M$). The Legendre polynomials form an orthonormal set of functions over the interval $[-1, 1]$, where the depth range of well logging measurements are linearly scaled. The Legendre polynomials up to the sextic degree are plotted in Fig. 1. In geophysical inversion, the problem of ambiguity primarily arises from the high correlation between the model parameters. Since there is no attachment between the Legendre polynomials of different orders, a relatively weak correlation is indicated between the estimated model parameters. This approach results in a unique solution and stable inversion procedure. Hence, the interval inversion method applying Legendre polynomials gives more accurate estimates to the series expansion coefficients than by using power functions or other non-orthogonal polynomials. The added advantage of implementing Eq. (7) is that the petrophysical model can be described by a significantly smaller number of expansion coefficients than data, which can be extracted in a greatly overdetermined inversion procedure.

Consider the combination of Eqs. (6) and (7), which expresses the calculation of the kth log type with the series expansion coefficients

$$d_k^{(\mathrm{cal})}(z) = g_k\left(B_1^{(1)}, \ldots, B_{Q^{(1)}}^{(1)}, \ldots, B_1^{(M)}, \ldots, B_{Q^{(M)}}^{(M)}, z\right). \tag{8}$$

The expansion coefficients forming independent variables of the above response function are collected into the column vector $\mathbf{B} = \left[B_1^{(1)}, \ldots, B_{Q^{(1)}}^{(1)}, \ldots, B_1^{(M)}, \ldots, B_{Q^{(M)}}^{(M)}\right]^{\mathrm{T}}$, which are the unknowns of the inverse problem. In the genetic inversion procedure, the population of individuals is formed by a number of series expansion coefficient vectors, which represent possible solutions to the inverse problem. First, one must integrate all well log types measured over the

Journal: 11004 Article No.: 9768 TYPESET DISK LE CP Disp.:2018/10/16 Pages: 25 Layout: Small-X

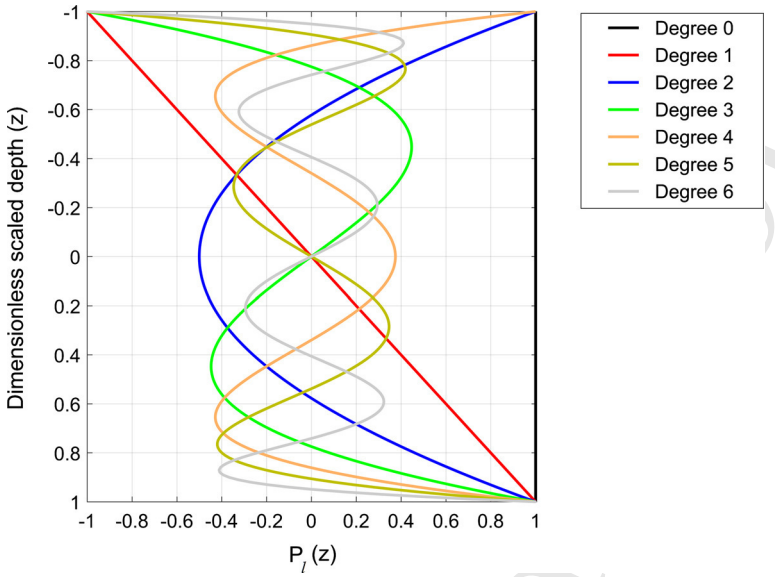

Fig. 1 Legendre polynomials of the lth degree ($l = 0, 1, \ldots, 6$) as depth functions plotted in the range $[-1, 1]$

processing interval into the data matrix $\mathbf{D}^{(\mathrm{obs})}$, the kth column of which includes all data of the kth observed quantity along the selected depth interval of the borehole. The matrix of theoretical well logging data is of equal size and calculated by Eq. (8). The misfit between the measured and calculated data is related to the fitness function

$$F\left(\mathbf{B}^{(s)}\right) = -\left[(NK)^{-1}\sum_{n=1}^{N}\sum_{k=1}^{K}\left(\frac{D_{nk}^{(\mathrm{obs})} - g_k\left(\mathbf{B}^{(s)}, z_n\right)}{D_{nk}^{(\mathrm{obs})}}\right)^2\right]^{1/2}, \tag{9}$$

where D_{nk} is the kth data measured in the nth depth and $\mathbf{B}^{(s)}$ is the sth series expansion vector ($s = 1, 2, \ldots, S$, where S is the population size). For maximizing the average fitness of the population, real genetic operations such as selection, crossover, and mutation are repeatedly applied on the individuals in consecutive generations (Houck et al. 1995).

At the beginning of the genetic inversion procedure, the fittest individuals are selected for reproduction by using the normalized geometric ranking operator. At first, the individuals are sorted by their fitness values calculated by Eq. (9). The rank of the fittest model is 1, while that of the worst adaptable one is S. The probability of selecting the ith expansion coefficient vector is expressed by

$$P\left(\mathbf{B}^{(i)}\right) = \frac{P_{\mathrm{b}}}{1-(1-P_{\mathrm{b}})^S}(1-P_{\mathrm{b}})^{r_i-1}, \tag{10}$$

where r_i is the rank of the ith expansion coefficient vector and P_{b} is the probability of selecting the best model. The cumulative probability of the ranked population is

Journal: 11004 Article No.: 9768 TYPESET DISK LE CP Disp.:2018/10/16 Pages: 25 Layout: Small-X

$C_i = \sum_{t=1}^{i} P\left(\mathbf{B}^{(t)}\right)$. If the condition fulfils $C_{i-1} < \beta \leq C_i$, the ith chromosome of series expansion coefficients is copied into the new population, where β is a random number from the uniform density on [0, 1]. In the next step, a pair of individuals $\mathbf{B}^{(1)}$ and $\mathbf{B}^{(2)}$ are chosen from the selected population to exchange some genes between them. The crossover operator normally cuts the chromosomes at one or more crossover points and swaps the expansion coefficients to the right of them. In this study, the heuristic crossover operation is applied that extrapolates two models as

$$\begin{aligned} \mathbf{B}^{*(1)} &= \mathbf{B}^{(1)} + \gamma\left(\mathbf{B}^{(1)} - \mathbf{B}^{(2)}\right), \\ \mathbf{B}^{*(2)} &= \mathbf{B}^{(1)}, \end{aligned} \tag{11}$$

where γ is a random number generated from U(0, 1). The condition of using Eq. (11) is that the fitness of $\mathbf{B}^{(1)}$ is higher than that of $\mathbf{B}^{(2)}$. If any value of $\mathbf{B}^{*(1)}$ is out of range, the generation of γ is to be repeated and Eq. (11) is recalculated. After some hundreds of failures, the values of expansion coefficients are changed to the old ones. The last genetic operator is uniform mutation, where the individual $\mathbf{B}^{*(1)}$ is selected and its vth expansion coefficient is substituted with the random number η chosen from its possible range

$$\mathbf{B}^{**(1)} = \begin{cases} \eta, & \text{if} \quad v = h \\ B_v^{*(1)}, & \text{otherwise} \end{cases}, \tag{12}$$

where $\mathbf{B}^{**(1)}$ is the mutated vector of series expansion coefficients. Equations (10)–(12) are applied repeatedly in successive generations until a termination criterion is met. In the last generation, the expansion coefficient vector with maximum fitness is regarded as the solution. At the end of the inversion procedure, the petrophysical properties are extracted from the expansion coefficients using Eq. (7) and the depth coordinates are rescaled to their original range.

2.3 Quality Check of the Inversion Result

The development of convergence of the interval inversion procedure can be significantly accelerated by switching the FGA around the global optimum to linearized optimization. In the last optimization phase, not just the estimation of the petrophysical model is finalized but also the Jacobi matrix is calculated, holding information on the accuracy of the model parameters. The calculation of the estimation error of petrophysical parameters would take an unreasonably high computational time when executing the genetic inversion procedure several times. Instead of doing so, the DLSQ method is used as a post-processing calculation method in the interval inversion workflow.

Menke (1984) showed that the model covariance matrix estimated by a linearized inversion method is directly proportional to the data covariance matrix including the

Journal: 11004 Article No.: 9768 TYPESET DISK LE CP Disp.:2018/10/16 Pages: 25 Layout: Small-X

uncertainties of observed data. This can be written for the series expansion coefficients as

$$\mathrm{cov}\mathbf{B} = \mathbf{G}^{-g}\mathrm{cov}\mathbf{d}^{(\mathrm{obs})}\left(\mathbf{G}^{-g}\right)^{\mathrm{T}}, \tag{13}$$

where $\mathbf{d}^{(\mathrm{obs})}$ denotes the observed data vector of the linear inverse problem and $\mathbf{G}^{-g}$ is the general inverse matrix of the DLSQ method (Marquardt 1959). The linearization of the nonlinear inverse problem produces an approximated uncertainty appraisal. To cast the genetic algorithm inversion in a solid probabilistic framework, the neighborhood algorithm (Sambridge 1999) or a hybrid genetic algorithm-Gibbs sampler approach (Aleardi and Mazzotti 2017), for instance, can be suggested. The linear approach of Menke (1984) is a standard method of uncertainty analysis in the oil field practice, based upon which a comparative study can be made between the traditional local inversion method and the suggested interval inversion procedure. The genetic algorithm phase of the interval inversion procedure is very time-consuming. In this study, global optimization is used for setting a proper initial model for the subsequent linear phase, in which the estimation errors can be calculated relatively quickly. By knowing the optimal values of the expansion coefficients, Eq. (7) is used to derive the depth function of petrophysical properties. The estimation error of petrophysical parameters derived from Eq. (13) requires the propagation of errors to be taken into account. Dobróka et al. (2016) introduced the depth-dependent model covariance matrix for the petrophysical parameters as

$$\left[\mathrm{cov}\mathbf{m}\left(z\right)\right]_{ij} = \sum_{w=1}^{Q^{(i)}}\sum_{t=1}^{Q^{(j)}} P_{w-1}\left(z\right)\left(\mathrm{cov}\mathbf{B}\right)_{uv} P_{t-1}\left(z\right), \tag{14}$$

where indices i and j run from 1 to M and indices u and v are $u = w + Q^{(1)} + Q^{(2)} + \cdots + Q^{(i-1)}$ and $v = t + Q^{(1)} + Q^{(2)} + \cdots + Q^{(j-1)}$. The main diagonal of the model covariance matrix includes the variances of the petrophysical quantities. The estimation error of the ith petrophysical parameter is derived from Eq. (14) as $\sigma\left[m_i\left(z\right)\right] = \left[\mathrm{cov}\mathbf{m}_{ii}\left(z\right)\right]^{1/2}$. It can be mentioned that, in case of asymmetric uncertainty ranges of observed variables, one can derive the upper and lower boundaries of the error intervals of model parameters based on the above methodology (Baker Atlas 1996). As another quality check parameter, the correlation coefficient between the ith and jth petrophysical parameters is calculated as $\left[\mathrm{corr}\ \mathbf{m}\left(z\right)\right]_{ij} = \left[\mathrm{cov}\ \mathbf{m}\left(z\right)\right]_{ij}\sigma\left[m_i\left(z\right)\right]^{-1}\sigma\left[m_j\left(z\right)\right]^{-1}$, which characterizes the reliability of the estimated model parameters for the entire processing interval. One can introduce the mean spread as a scalar for measuring the average correlation

$$R\left(\mathbf{m}\right) = \left\{\left[M\left(M-1\right)\right]^{-1}\sum_{i=1}^{M}\sum_{j=1}^{M}\left[\left[\mathrm{cov}\ \mathbf{m}\left(z\right)\right]_{ij} - \delta_{ij}\right]^2\right\}^{1/2}, \tag{15}$$

where δ denotes the Kronecker symbol (Menke 1984). From the point of view of the inversion procedure, a solution is unreliable when the estimated parameters strongly

Journal: 11004 Article No.: 9768 TYPESET DISK LE CP Disp.:2018/10/16 Pages: 25 Layout: Small-X

correlate. In that very case, model parameters are prevented from changing independently during the inversion process. Only poorly correlated parameters can be determined uniquely by inverse modeling referring to a reliable solution. As a third quantity, the relative data distance measures the overall misfit between the measured and calculated data matrices

$$D_{\mathrm{d}} = \left[(NK)^{-1} \sum_{n=1}^{N} \sum_{k=1}^{K} \left(\frac{D_{nk}^{(\mathrm{obs})} - D_{nk}^{(\mathrm{calc})}}{D_{nk}^{(\mathrm{obs})}} \right)^2 \right]^{1/2} \times 100\ (\%)\,, \tag{16}$$

which takes into account that the observed well logging parameters are of different magnitudes and dimensional units.

3 Results and Interpretation

3.1 Limitation of Linearized Optimization

As a preliminary numerical experiment, the Marquardt algorithm-based (linearized) interval inversion procedure is tested using a well logging data set measured in a Hungarian hydrocarbon exploratory borehole (well H). The formation investigated is composed of unconsolidated gas-bearing shaly-sand reservoirs of Pliocene age. The natural gamma ray intensity, density, neutron porosity, acoustic traveltime, and true resistivity logs are simultaneously inverted to estimate the expansion coefficients describing the porosity, invaded and virgin zone water saturation, and shale volume. The sand volume is calculated from the well-known material balance equation $\Phi + V_{\mathrm{sh}} + V_{\mathrm{sd}} = 1$. The actual values of zone parameters set in Eqs. (1)–(5) are given in Table 1. By a sampling distance of 0.1 m, $N = 194$ depth points, and $K = 5$ log types, 970 data are jointly inverted.

The petrophysical properties are discretized according to Eq. (7). Legendre polynomials of degree 44 as basis functions are used for describing each petrophysical parameter. Thus, the total number of unknowns is $Q^* = 180$ (since 45 expansion coefficients are necessary for a petrophysical parameter) and the data-to-unknowns ratio is 5.4 (it is only 1.25 for local inversion). The interval inversion procedure is run by using different starting models (Table 2). The initial values of the expansion coefficients of the Legendre polynomials of zero degree are chosen differently for each model. The expansion coefficient B_1 can be identified as the constant value of the relevant petrophysical parameter in the processed (equivalent homogeneous) depth interval. The coefficients of the higher degree basis functions are set to zero. All inversion runs over 20 iterations are plotted in Fig. 2. It is shown that the inversion process is highly stable, when model I is used (randomly) as the initial model. By applying Eq. (16), the relative data distance is calculated on the starting model as $D_{\mathrm{d},0} = 86\%$, which is continuously decreased to $D_{\mathrm{d}} = 2\%$. Let just one parameter of the initial model change; for example, the value of shale volume is increased by 0.1. For the new model (model II), the development of convergence fails even when introducing a non-zero damping factor (ε). Data predicted on models III and IV fall the farthest from the

Journal: 11004 Article No.: 9768 TYPESET DISK LE CP Disp.:2018/10/16 Pages: 25 Layout: Small-X

Table 2 Starting models used for linearized interval inversion of wireline logs observed in well H

Model no.	$B_1^{(\Phi)}$	$B_1^{(S_w)}$	$B_1^{(V_{sh})}$	$B_1^{(S_{x0})}$	ε	$D_{d,0}$ (%)	D_d (%)
I	0.20	0.40	0.15	0.80	0	86	2.05
II	0.20	0.40	0.25	0.80	0	63	Divergent
					1.0		Divergent
III	0.10	0.50	0.20	0.70	1.0	97	12.20
IV	0.20	0.20	0.10	0.50	1.0	320	Divergent

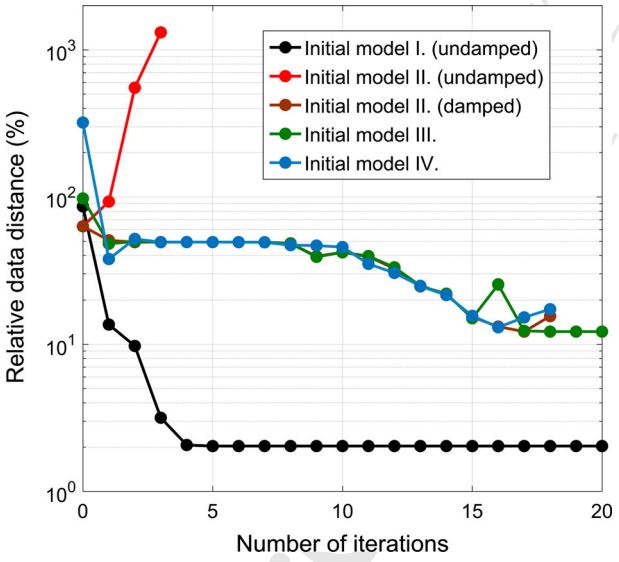

Fig. 2 Data distance versus iteration step curves obtained by linearized interval inversion of well logs using different starting models. The parameters of initial models I–IV are given in Table 2

observations (see data distances in Table 2). Despite regularization, they do not give an optimal solution. Although in the case of almost 100% data distance the application of model III leads to a convergent solution, the search is obviously trapped in a local minimum of the objective function, Eq. (9). Model IV with an extremely high data distance represents an ill-conditioned linear inverse problem (caused by a near-zero determinant of the inverted matrix), which does not give interpretable results. It is concluded that the success of the linear optimization-based interval inversion procedure critically depends on the selection of the initial model. Inappropriately chosen initial models lead to a divergent or not an optimal solution, which can be assisted by the use of genetic inversion.

Journal: 11004 Article No.: 9768 TYPESET DISK LE CP Disp.:2018/10/16 Pages: 25 Layout: Small-X

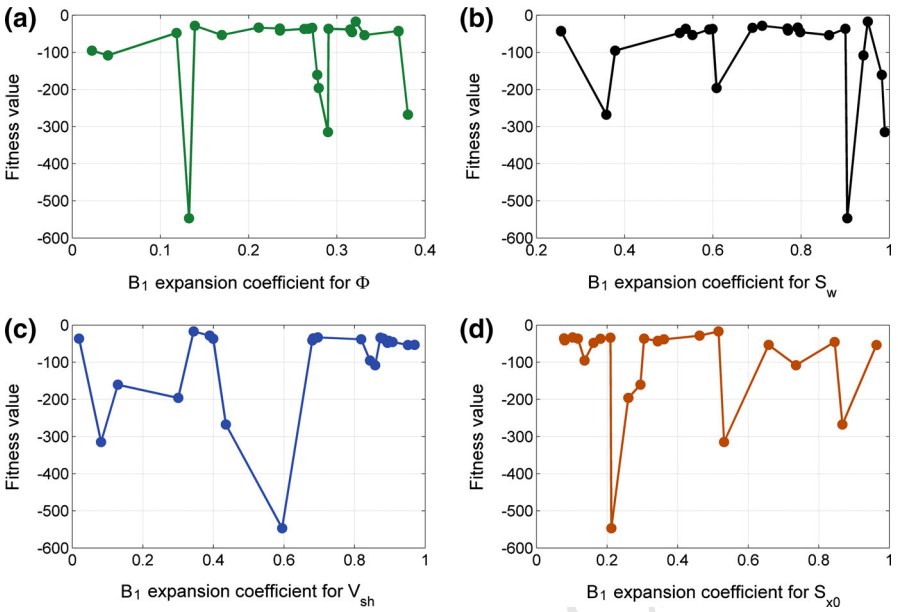

Fig. 3 Fitness values of series expansion coefficients of zeroth degree Legendre polynomials (B_1) contributing the depth function of: **a** porosity (Φ), **b** water saturation in uninvaded zone (S_w), **c** shale volume (V_{sh}), **d** water saturation in invaded zone (S_{x0})

3.2 Testing the Genetic Inversion Method

The FGA-based interval inversion method is applied in well H. For reducing the CPU time of the heuristic search, an approximate solution is given for the petrophysical parameters using Legendre polynomials of 14 degrees in series expansion. In the initialization of the FGA process, the upper and lower limits of expansion coefficients can be chosen within wide limits. The initial values of the expansion coefficients for the zeroth degree Legendre polynomials correspond to the full practical range of petrophysical parameters: $0 \le B_1^{(\Phi)} \le 0.4$, $0 \le B_1^{(S_w)} \le 1$, $0 \le B_1^{(S_{x0})} \le 1$, $0 \le B_1^{(V_{sh})} \le 1$. Based on the linear inversion experiments detailed in the previous section, the possible range of higher degree expansion coefficients are set as $-0.2 \le B_q^{(i)} \le 0.2$ ($i = 1, 2, 3, 4$ and $q = 1, 2, \ldots, 15$). Figure 3 shows that the fitness function defined in Eq. (9) plotted against the coefficients B_1 has several localities. The fitness values at minimum locations correspond to a relative data distance of approximately 800–1000%. The multidimensional fitness function plotted against the expansion coefficients of higher degree Legendre polynomials also has several extrema (Fig. 4), which justifies the use of the FGA method.

A population of 50 individuals is improved by the FGA procedure. When having four petrophysical properties as unknown, the number of expansion coefficients per individual (petrophysical model) is 60 and the total number of genes in the population is 3000. The material balance equation as a constraint is built in the forward problem solution. The FGA search normally takes 10^4–10^5 generations to find the global

Journal: 11004 Article No.: 9768 TYPESET DISK LE CP Disp.:2018/10/16 Pages: 25 Layout: Small-X

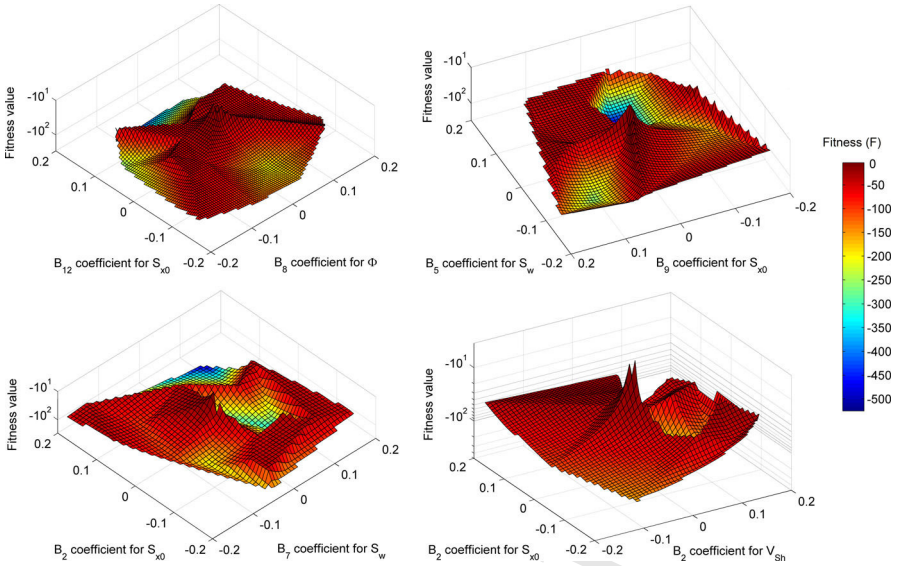

Fig. 4 Fitness functions of two arguments: F is the fitness value inversely related to data misfit, B_q is the series expansion coefficient of the $(q-1)$th degree Legendre polynomial as the basis function, porosity (Φ), water saturation in uninvaded zone (S_w), shale volume (V_{sh}), water saturation in invaded zone (S_{x0})

optimum in our applications. In contrast, the inversion procedure is run over 5000 iterations to achieve sufficient approximation and requires CPU time of <2 min. The control parameters of the genetic search are the probability of selecting the best model ($P_b = 0.03$), the number of crossover retries (100), and the probability of mutation (0.05). For the reproduction of the best models, the rule of elitism is also applied, by which the worst individual of the current generation is automatically replaced with the fittest one of the previous generation. The improvement of the average fitness of the model population from the initial generation to later generations can be seen in Fig. 5a. The negative value of the fitness is comparable in order of magnitude with the data distance defined in Eq. (16). In the initial population, the mean of the fitness is -7.4×10^9 ($\sim D_d = 10^9$–10^{10}%), which reduces to -53 at the end of the global optimization procedure. The best individual has a fitness value of -5.27, which corresponds to $D_d = 16.9\%$ relative data distance. The depth variation of the petrophysical parameters calculated from the estimated expansion coefficients is plotted in Fig. 6.

When the degree of Legendre polynomials is not changed in the linearized inversion (DLSQ) phase (remaining at 14), the data distance does not improve. Better resolution for the petrophysical model can be reached by increasing the number of expansion coefficients. By assuming a faster depth variation of petrophysical parameters, in the DLSQ process, the number of unknowns is increased from 60 to 180 using Legendre polynomials of degree 44 (Fig. 6). Since the condition number of the self-product of the Jacobi matrix is $\sim 10^6$, the inverse problem is ill posed. The inverse problem is regularized by setting a damping factor of 1 for the Marquardt algorithm, which is decreased to its actual value by 60% in each iteration. The maximal number of

Journal: 11004 Article No.: 9768 TYPESET DISK LE CP Disp.:2018/10/16 Pages: 25 Layout: Small-X

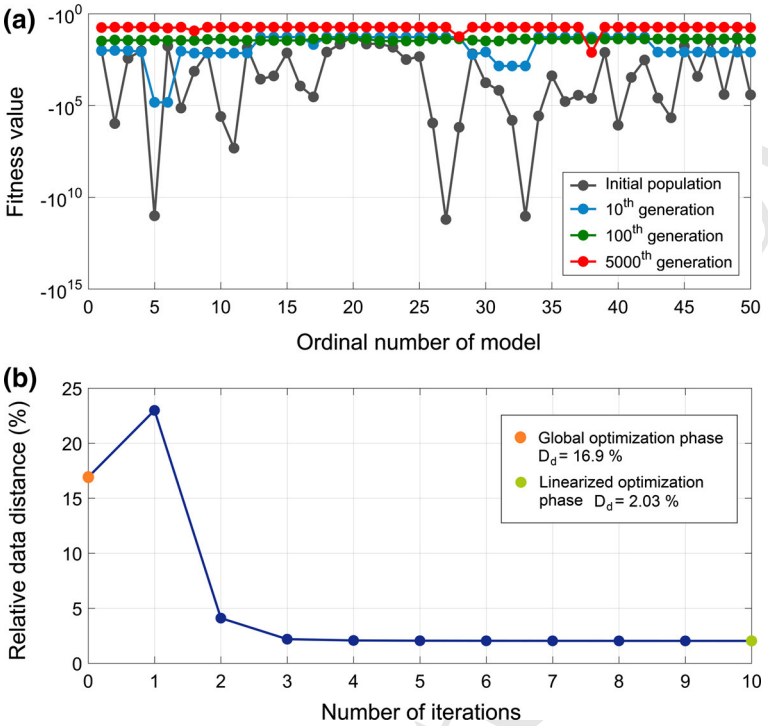

Fig. 5 Genetic interval inversion process including global and linear optimization phases in well H: **a** global inversion result for Legendre polynomials of degree 14, **b** global and linearized inversion result for Legendre polynomials of degree 44. D_d (%) is the relative data distance between the measured and calculated well logs

iterations is 20, at the end of which the regularization factor is 3.6×10^{-5}. The values of both the expansion coefficients and the data distance do not change after ten iterations; thus, Fig. 5b shows the results until that iteration step. The development of convergence is quick and steady. By the increased number of expansion coefficients, the data distance can be effectively decreased. In this case, the inversion procedure stabilizes at relative data distance $D_d = 2.03\%$. It should be noted that, by executing the genetic inversion procedure with initial populations of different sizes, the same optimum is given. The measured well logs are in close agreement with the theoretical logs calculated on the resultant model (see tracks 1–5 in Fig. 7). The fractional volume of water, hydrocarbon, shale, and quartz matrix is plotted in track 6. The irreducible and movable hydrocarbon volumes are directly derived from the inversion results using $V_{hc,mov} = \Phi(S_{x0} - S_w)$ and $V_{hc,irr} = \Phi(1 - S_{x0})$. In the last track of Fig. 7, the pore fluid concentrations are given for unit volume of rock: brine volume (blue color), irreducible hydrocarbon saturation (brown color), and movable hydrocarbon saturation (red color). The petrophysical parameters show strong correlations to those estimated by a linear optimization-based local inversion method (Fig. 8). The result of the regression analysis shows the indication of heteroscedasticity. The non-constant variance in the crossplots may be caused by the different vertical variability of the

Journal: 11004 Article No.: 9768 TYPESET DISK LE CP Disp.:2018/10/16 Pages: 25 Layout: Small-X

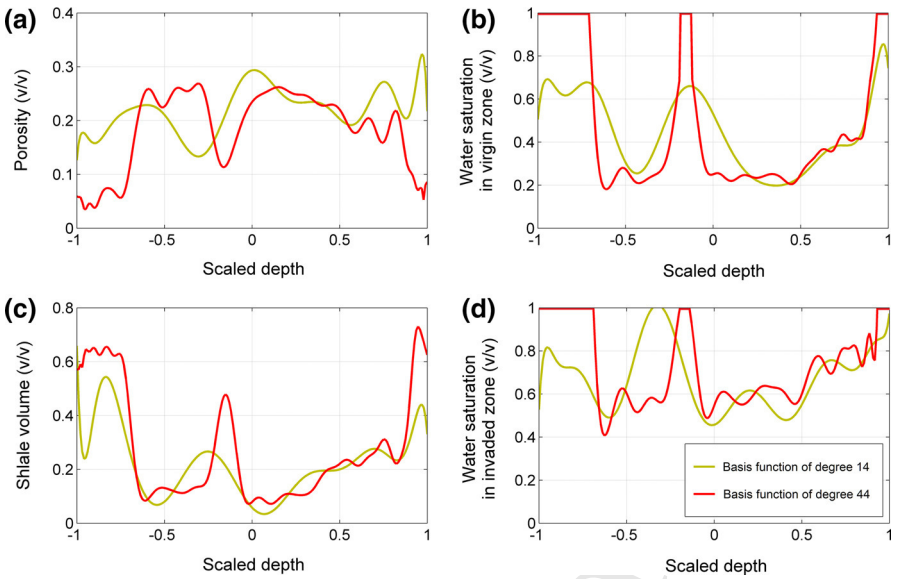

Fig. 6 Petrophysical parameter distributions derived from the series expansion coefficients estimated by the genetic interval inversion procedure in well H: **a** porosity, **b** uninvaded zone water saturation, **c** shale volume, **d** invaded zone water saturation estimated using Legendre polynomials of degree 14 (olive color) and degree 44 (red color)

petrophysical parameters described by the same number of series expansion coefficients. The genetic inversion procedure including also the linearized optimization phase requires a CPU time of 125 s using a quad-core processor workstation. Based on the inversion experiments, it is concluded that the genetic inversion procedure is stable and practically independent of the initial model.

4 Discussion

The feasibility of the linearized interval inversion method was studied previously by Dobróka et al. (2016). They proved that the estimation accuracy and reliability of the petrophysical parameters were significantly improved compared to the weighted least squares-based local inversion method. The results of interval inversion were also successfully compared to those of robust statistical methods (Szabó et al. 2018). The application of FGA increases its numerical stability and allows the start model to be provided within broader limits. The linear phase of the hybrid inversion algorithm gives a possibility of checking the quality and noise sensitivity of the interval inversion method. After substituting Eq. (13) into Eq. (14), it is seen that the covariance matrix of observed data is an indispensable input quantity for calculating the estimation accuracy. In this study, it is assumed that the inverted well logs have different uncertainties, the values of which are available in sonde specifications, user manuals, or the literature (Horváth 1973).

Journal: 11004 Article No.: 9768 TYPESET DISK LE CP Disp.:2018/10/16 Pages: 25 Layout: Small-X

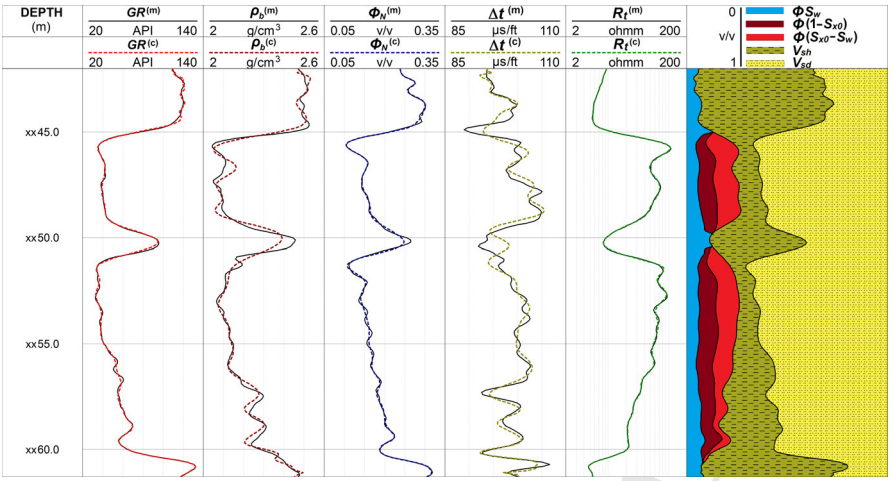

Fig. 7 Results of the genetic inversion procedure in well H: natural gamma ray (GR), density (ρ_b), neutron porosity (Φ_N), acoustic traveltime (Δt), true resistivity (R_t) logs. Superscripts (m) and (c) denote the measured and calculated data, respectively. Estimated petrophysical quantities are porosity (Φ), shale volume (V_{sh}), and water saturations (S_{x0} and S_w)

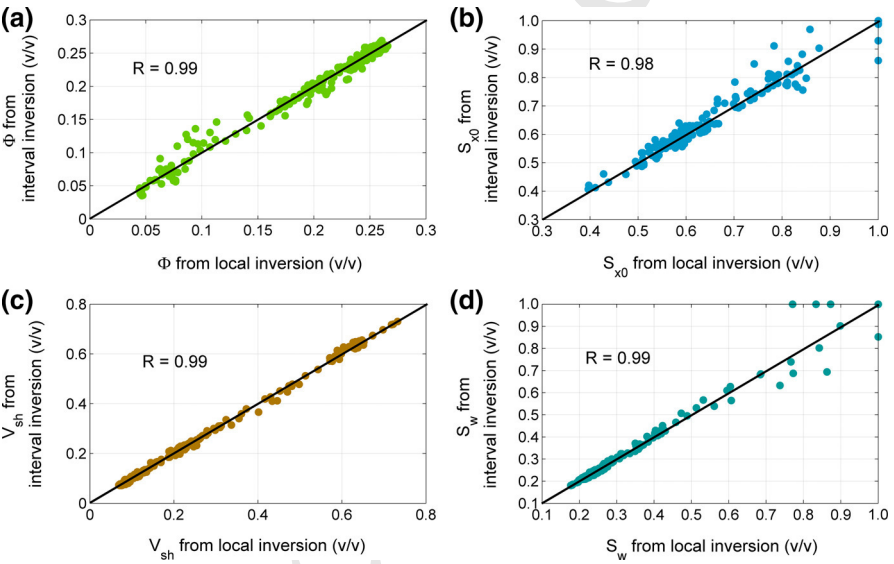

Fig. 8 Regression relation between the results of local inversion and genetic interval inversion in well H. Strong linear connection is indicated for: **a** porosity (Φ), **b** water saturation in uninvaded zone (S_w), **c** shale volume (V_{sh}), **d** water saturation in invaded zone (S_{x0}). R is Pearson's correlation coefficient

In order to test the quality of the interval inversion results, the standard deviation of well logging data are chosen as $\sigma_{GR} = 0.12$, $\sigma_{\rho_b} = 0.08$, $\sigma_{\Phi_N} = 0.13$, $\sigma_{\Delta t} = 0.10$, and $\sigma_{R_t} = 0.15$, which are even higher than those in Dobróka et al. (2016). At first, the standard deviation of the expansion coefficients are calculated using Eq. (13). The

Journal: 11004 Article No.: 9768 TYPESET DISK LE CP Disp.:2018/10/16 Pages: 25 Layout: Small-X

estimated expansion coefficients including their estimation errors are plotted in Fig. 9a. The expansion coefficients of the zeroth degree Legendre polynomials are estimated as $B_1^{(\Phi)} = 0.19$, $B_1^{(S_w)} = 0.50$, $B_1^{(V_{sh})} = 0.27$, and $B_1^{(S_{x0})} = 0.87$, while the values of higher degrees are represented with an empiric probability density function in Fig. 9b. As Fig. 9c shows, the estimation error of half of the expansion coefficients is smaller than 0.05. A maximum around 0.1 is mostly associated to the error of uninvaded zone water saturation, while that of 0.25 can be identified as the error of invaded zone water saturation. The average correlation of expansion coefficients calculated by Eq. (15) is $R(\mathbf{B}) = 0.17$, which indicates poorly correlated model parameters and highly reliable inversion result. Once the variances of expansion coefficients are calculated, one can derive the estimation error of the petrophysical parameters using Eq. (14). The error interval of the kth measured variable is calculated as $\Delta \mathrm{d_k} \left(1 \pm \sigma_{d_k}\right)$. The light brown shaded area in Fig. 10a shows the error intervals calculated from data variances. Tracks 1–4 of Fig. 10b show the estimation error of the petrophysical properties with olive color shaded area, the average values of which for the entire processing interval are given as $\bar{\sigma}_{\Phi} = 0.03$ v/v, $\bar{\sigma}_{S_w} = 2.6$ v/v, $\bar{\sigma}_{V_{sh}} = 0.16$ v/v, and $\bar{\sigma}_{S_{x0}} = 11.7$ v/v. The correlation between the petrophysical parameters is moderate (see track 5), the average being $R(\mathbf{m}) = 0.38$ (for local inversion, it is 0.59 in well H). From the standard deviations and Pearson's correlation coefficients, it is concluded that the porosity and shale volume are estimated more accurately than the water saturation in both the invaded and virgin zones. The latter two are more strongly correlated to each other than the previous ones, since the depth averages are $\overline{\mathrm{corr}}\,(\Phi, V_{sh}) = 0.19$ and $\overline{\mathrm{corr}}\,(S_w, S_{x0}) = 0.74$. The strong correlation between the water saturations causes relatively lower estimation accuracy for them. High correlation coefficients always result in the problem of ambiguity. The root of this problem is found in the structure of sonde response equations, which can be improved by revising the solution of the forward problem. From the point of view of inverse modeling, highly correlated model parameters should be estimated outside the inversion procedure. In conclusion, the advantage of the combined interval inversion approach is that it gives more accurate results than those of the marginally overdetermined local inversion.

The optimal number of basis functions can be found by seeking the minimum of the average correlation between the estimated expansion coefficients (Dobróka et al. 2016). For the combined interval inversion method, the dependency of the inversion result on the degree of the Legendre polynomials must be checked. By using the same well-logging data set and parameter settings, the inversion process is run by different numbers of basis functions selected between 1 and 44 in the genetic algorithm phase. The blue dots in Fig. 11a show that the interval inversion procedure combining the FGA and the DLSQ method always give the same optimum. In the genetic inversion phase, the relative data distance decreases significantly up to 16 basis functions. Further improvement can only be achieved by increasing the number of generations, which is currently 5000 steps for each inversion experimental run. Thus, in the case study, Legendre polynomials of 14 degrees in series expansion were chosen. The advantage of using a relatively small number of basis functions is the low computational cost and high estimation accuracy. As shown in Fig. 11b, the CPU time nonlinearly and highly increases with the number of polynomials. In Fig. 12, one can observe how the

Journal: 11004 Article No.: 9768 TYPESET DISK LE CP Disp.:2018/10/16 Pages: 25 Layout: Small-X

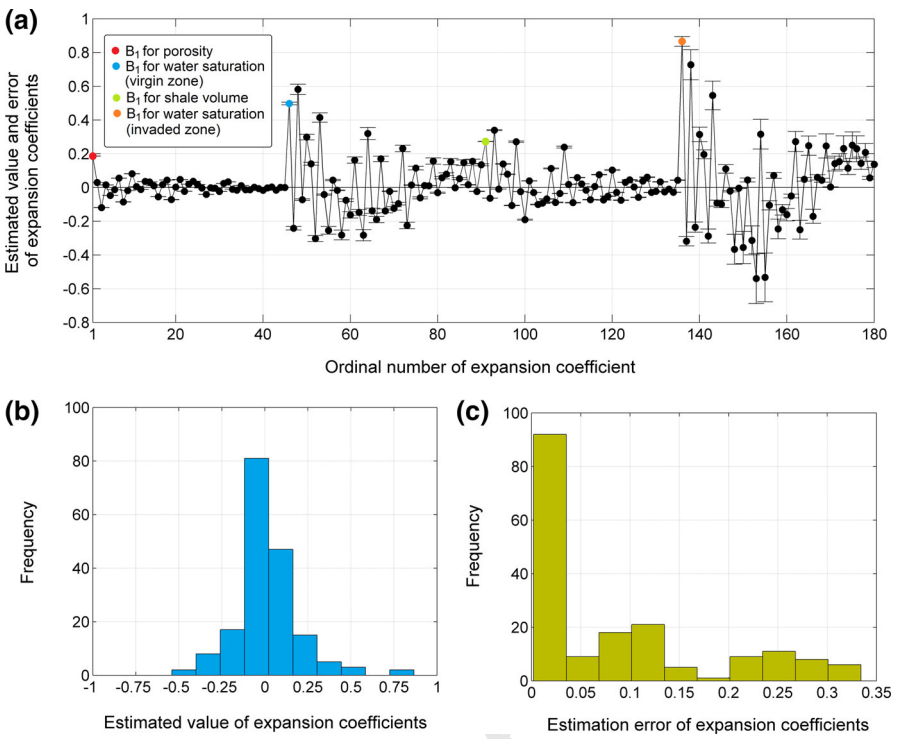

Fig. 9 Series expansion coefficients including their error bars estimated by genetic inversion in well H (**a**); histogram of the estimated expansion coefficients (**b**); histogram of the standard deviation of expansion coefficients (**c**)

FGA approximates the result of the combined interval inversion procedure (FGA + DLSQ) when using different numbers of basis functions. These tests reveal that the FGA provides a proper starting model for the DLSQ inversion, which gives stable predictions for different numbers of basis functions (Fig. 11a). In contrast, a single linearized inversion (DLSQ) can be highly affected by the starting model (Fig. 2). The difference between the results of the single and combined interval inversion procedures is explained as: (1) the data distance calculated for the starting models is around 100% or higher for the cases shown in Fig. 2, while in the case of using a fast FGA step, it is no more than 40%, as shown in Fig. 11a (i.e., the initial model is closer to the optimum); (2) the series expansion coefficients of higher order Legendre polynomials are set to zero (i.e., homogeneous starting model is assumed) before the DLSQ inversion in all examples of Fig. 2, whereas in the frame of the FGA + DLSQ procedure, the same coefficients are estimated first by FGA to give a more sophisticated (inhomogeneous) initial model for DLSQ inversion (Fig. 12). The estimated (non-zero) values of the expansion coefficients of these starting models are just further refined by DLSQ inversion, which brings stability to the combined interval inversion procedure. The above tests show that a trade-off should be made between the number of expansion coefficients as unknowns of the inverse problem (i.e., vertical resolution of petrophysical properties) and total CPU time.

Journal: 11004 Article No.: 9768 TYPESET DISK LE CP Disp.:2018/10/16 Pages: 25 Layout: Small-X

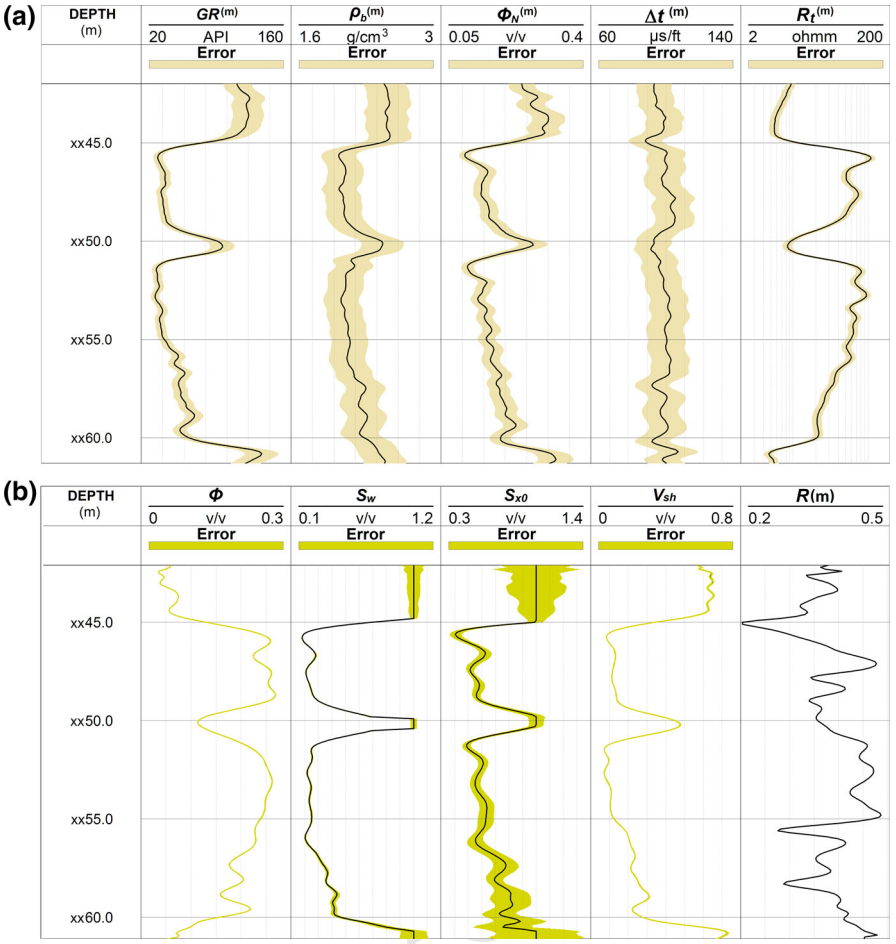

Fig. 10 Quality check of the interval inversion procedure in well H: observed natural gamma ray (GR), density (ρ_b), neutron porosity (Φ_N), acoustic traveltime (Δt), true resistivity (R_t) logs with accuracy ranges (**a**); inversion results with estimation errors: porosity (Φ), shale volume (V_{sh}), invaded and uninvaded zone water saturation (S_{x0} and S_w)

The feasibility of the combined interval inversion method can be further proven by making synthetic inversion experiments. Well logs are calculated by Eqs. (1)–(5) using a known petrophysical model. The depth variations of petrophysical properties of a hydrocarbon-saturated shaly-sand formation are constructed as 24 degree Legendre polynomials. To simulate real borehole measurements, different amounts of random noise can be added to the synthetic data. The kth noisy data at a given depth is calculated as $d'_k = d_k\,(1 + \mathrm{N}(\mu, \sigma))$, where $\mu = 0$ is the expected value and σ is the standard deviation of the desired noise level. The control parameters and iteration number of the genetic inversion process are chosen similarly as in Sect. 3.2. The zone parameters of the tool response equations are given in Table 1. The number of Legendre polynomials in Eq. (7) is chosen as 24 in both the genetic and linear phases.

Journal: 11004 Article No.: 9768 TYPESET DISK LE CP Disp.:2018/10/16 Pages: 25 Layout: Small-X

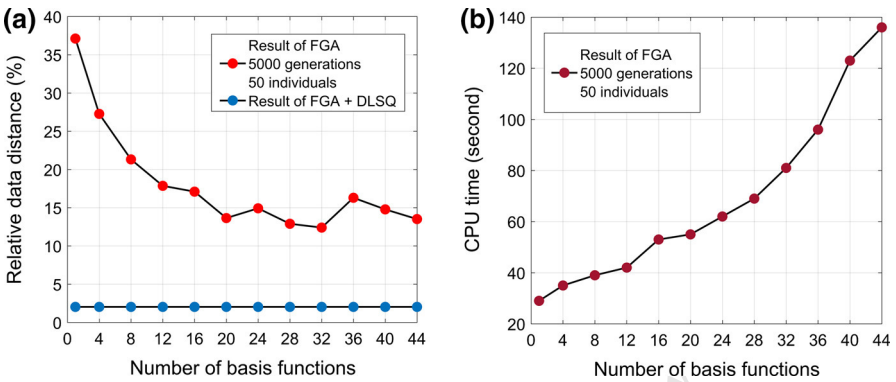

Fig. 11 Relative data distance versus the number of basis functions used in the genetic inversion phase of the combined interval inversion procedure (**a**); the CPU time requirement for the genetic algorithm versus the number of basis functions (**b**)

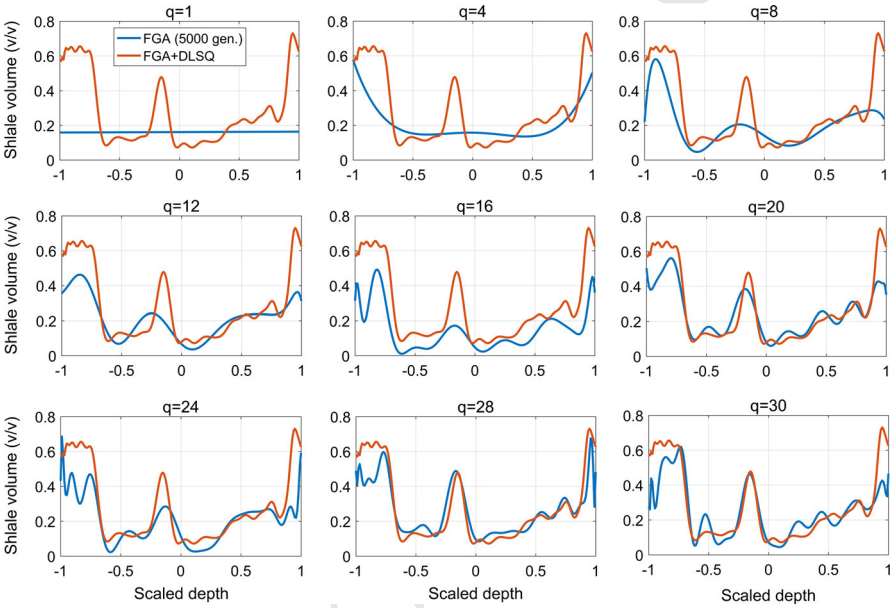

Fig. 12 Inversion results for different numbers of basis functions (q). The blue curves represent the results of genetic inversion using 5000 generations and the red curves show the results of the combined interval inversion procedure

The inversion results using a synthetic data set contaminated by 4% Gaussian noise is included in Fig. 13. The relative data distance between the noisy and calculated data is $D_d = 4.09\%$, which is proportional to the noise level of the input data (Fig. 13a). The petrophysical parameters of the true model can be directly compared to those of the predicted model in Fig. 13b. The *GR* image (track 1) shows the change of lithology, tracks 2 and 4 illustrate the saturation curves, while tracks 3 and 5 represent the results of compositional analysis. The overdetermination ratio of 12.5 results in

Journal: 11004 Article No.: 9768 TYPESET DISK LE CP Disp.:2018/10/16 Pages: 25 Layout: Small-X

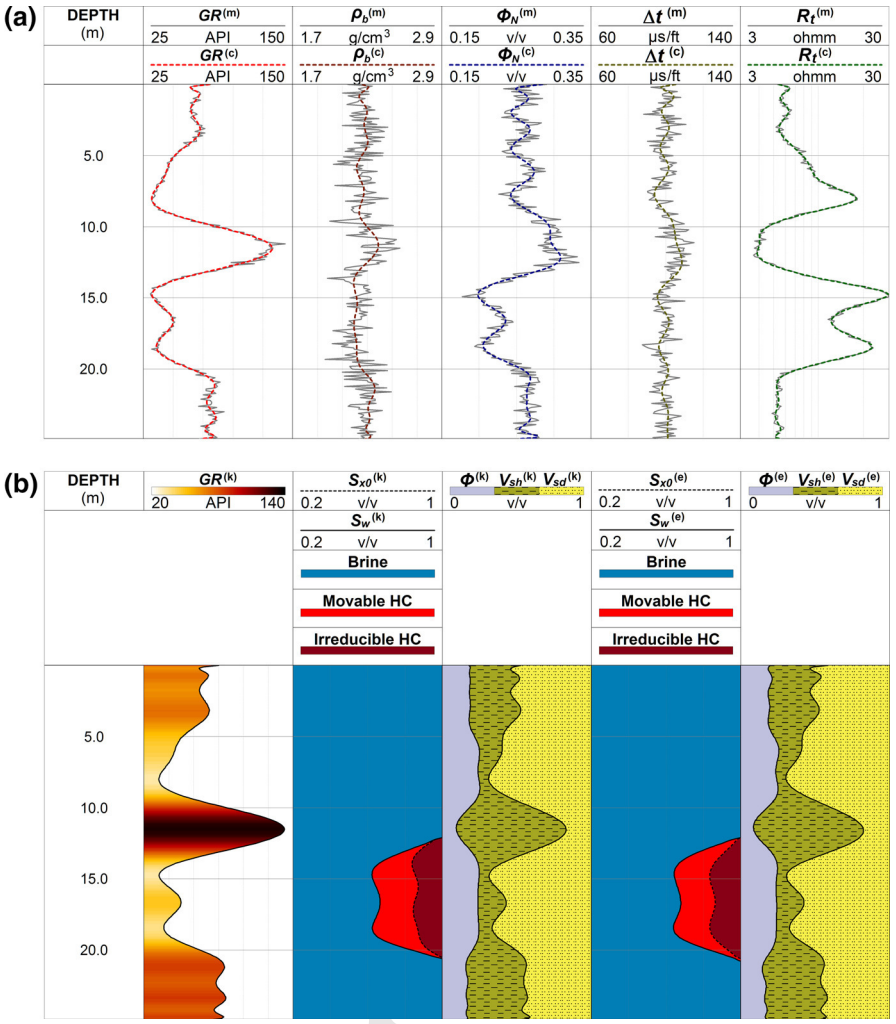

Fig. 13 Results of the combined interval inversion method using synthetic data contaminated with 4% Gaussian noise. Well logs of quasi measured data (m) and those of calculated data (c) on the estimated model (**a**); parameters of the known (k) model and estimated (e) model (**b**): porosity (Φ), water saturation in uninvaded zone (S_{w}) and in invaded zone (S_{x0}), and shale volume (V_{sh})

highly accurate estimation results. Assuming the data uncertainties given in the case study, the average estimation errors of petrophysical properties for the entire processing interval are obtained as $\bar{\sigma}_{\Phi} = 0.03$ v/v, $\bar{\sigma}_{S_{\mathrm{w}}} = 2.3$ v/v, $\bar{\sigma}_{V_{\mathrm{sh}}} = 0.14$ v/v, and $\bar{\sigma}_{S_{x0}} = 3.3$ v/v, which show nearly the same (slightly better) results as those of the field case. The average correlation between the series expansion coefficients is $R(\mathbf{B}) = 0.09$, which indicates practically uncorrelated model parameters and highly reliable inversion result. The CPU time requirement of the inversion run is approximately 2 min.

Journal: 11004 Article No.: 9768 TYPESET DISK LE CP Disp.:2018/10/16 Pages: 25 Layout: Small-X

5 Conclusions

A float-encoded genetic algorithm (FGA)-based interval inversion method is suggested to evaluate the petrophysical properties of inhomogeneous hydrocarbon formations. The global inversion approach implemented to solve the problem of interval inversion makes the inversion procedure computationally more effective and stable than linearized techniques. The feasibility is demonstrated in a gas-bearing shaly-sand formation and the results are compared to those of independent local inversion.

It is shown that the genetic inversion method is derivative-free and initial model-independent, including a quick forward problem solution. The coupled linear phase allows the fast calculation of the final solution, including the estimation error of both the series expansion coefficients and petrophysical parameters. The high rate of overdetermination originally ensures a low noise sensitivity; moreover, the FGA improves the fit between the measured and calculated data. For describing the geological structure, the basis functions in series expansion can be chosen arbitrarily. By selecting the number of expansion coefficients, one can control the CPU time of the inversion process, estimation accuracy, and the vertical resolution of the petrophysical model. Within reasonable limits, the genetic inversion procedure allows the integration of more well logs types and bigger data sets. This may require not only the development of new response functions or sophisticated forward modeling algorithms, but also the definition of the zone parameters from an objective source. The latter can be efficiently performed within the interval inversion procedure, which can be solved effectively by the interval inversion approach. The automated estimation of zone parameters sets a new perspective in well log analysis, which does not require prior knowledge of zone parameters for the inversion procedure. While preserving its stability, the efficiency of the interval inversion method can be increased by estimating some volumetric petrophysical parameters outside the inversion procedure. For instance, multivariate statistical techniques, like factor analysis, can be used to estimate the shale volume or other critical unknowns from an independent source. The resultant parameters can then be fixed during the interval inversion procedure. This may further increase the estimation accuracy for the remaining inversion unknowns.

The series expansion-based inversion methodology can be extended to cross-hole applications. By using Legendre polynomials of two or three variables as basis functions, the petrophysical parameters and the lateral (and, of course, the vertical) change of layer boundaries can be simultaneously estimated. In order to perform an initial model-independent two- or three-dimensional interval inversion of multi-borehole data, the genetic inversion method is preferably applied. In the future, the suggested genetic inversion workflow will be tested for the evaluation of unconventional hydrocarbon reservoirs, where the interval inversion method may be successfully applied for estimating an increasing number of unknowns due to multi-mineral rock matrix, organic content, and complex pore-space and fluid-saturation conditions.

Acknowledgements This research was supported by the GINOP-2.3.2-15-2016-00010 "Development of enhanced engineering methods with the aim at utilization of subterranean energy resources" project in the framework of the Széchenyi 2020 Plan, funded by the European Union, co-financed by the European Structural and Investment Funds. The first author is thankful for the support of the above GINOP project.

Journal: 11004 Article No.: 9768 TYPESET DISK LE CP Disp.:2018/10/16 Pages: 25 Layout: Small-X

References

Alberty M, Hashmy K (1984) Application of ULTRA to log analysis. In: SPWLA symposium transactions. Paper Z, pp 1–17

Aleardi M (2015) Seismic velocity estimation from well log data with genetic algorithms in comparison to neural networks and multilinear approaches. J Appl Geophys 117:13–22

Aleardi M, Mazzotti A (2017) 1D elastic full-waveform inversion and uncertainty estimation by means of a hybrid genetic algorithm–Gibbs sampler approach. Geophys Prospect 65:64–85

Ali Ahmadi M, Zendehboudi S, Lohi A, Elkamel A, Chatzis I (2013) Reservoir permeability prediction by neural networks combined with hybrid genetic algorithm and particle swarm optimization. Geophys Prospect 61:582–598

Archie GE (1942) The electrical resistivity log as an aid in determining some reservoir characteristics. SPE Trans AIME 146:54–62

Baker Atlas (1996) OPTIMA, eXpress reference manual. Baker Atlas, Western Atlas International, Inc., Houston

Ball SM, Chace DM, Fertl WH (1987) The Well Data System (WDS): an advanced formation evaluation concept in a microcomputer environment. In: Proceedings of the SPE eastern regional meeting, paper 17034, pp 61–85

Chunduru RK, Sen MK, Stoffa PL (1997) Hybrid optimization methods for geophysical inversion. Geophysics 62:1196–1207

Cranganu C, Luchian H, Breaban ME (2015) Artificial intelligent approaches in petroleum geosciences. Springer, Berlin

Dobróka M, Szabó NP (2011) Interval inversion of well-logging data for objective determination of textural parameters. Acta Geophys 59:907–934

Dobróka M, Szabó NP (2012) Interval inversion of well-logging data for automatic determination of formation boundaries by using a float-encoded genetic algorithm. J Pet Sci Eng 86–87:144–152

Dobróka M, Szabó PN, Cardarelli E, Vass P (2009) 2D inversion of borehole logging data for simultaneous determination of rock interfaces and petrophysical parameters. Acta Geod Geophys Hung 44:459–479

Dobróka M, Szabó NP, Turai E (2012) Interval inversion of borehole data for petrophysical characterization of complex reservoirs. Acta Geod Geophys Hung 47:172–184

Dobróka M, Szabó NP, Tóth J, Vass P (2016) Interval inversion approach for an improved interpretation of well logs. Geophysics 81:D155–D167

Dorrington KP, Link CA (2004) Genetic-algorithm/neural-network approach to seismic attribute selection for well-log prediction. Geophysics 69:212–221

Drahos D (2005) Inversion of engineering geophysical penetration sounding logs measured along a profile. Acta Geod Geophys Hung 40:193–202

Eidsvik J, Mukerji T, Switzer P (2004) Estimation of geological attributes from a well log: an application of hidden Markov chains. Math Geol 36:379–397

Gajda-Zagórska E, Schaefer R, Smołka M, Paszyński M, Pardo D (2015) A hybrid method for inversion of 3D DC resistivity logging measurements. Nat Comput 14:355–374

Galsa A, Herein M, Drahos D, Herein A (2016) Effect of the eccentricity of normal resistivity borehole tools on the current field and resistivity measurement. J Appl Geophys 134:281–290

Holland JH (1975) Adaptation in natural and artificial systems. University of Michigan Press, Ann Arbor

Horváth SB (1973) The accuracy of petrophysical parameters as derived by computer processing. Log Anal 14:16–33

Houck CR, Joines JA, Kay MG (1995) A genetic algorithm for function optimization: a Matlab implementation. NCSU-IE technical report 95-09, North Carolina State University, pp 1–14

Jarzyna JA, Cichy A, Drahos D, Galsa A, Bała MJ, Ossowski A (2016) New methods for modeling laterolog resistivity corrections. Acta Geophys 64:417–442

Lindberg DV, Grana D (2015) Petro-elastic log-facies classification using the expectation–maximization algorithm and hidden Markov models. Math Geosci 47:719–752

Malkoti A, Vedanti N, Bhattacharya AK (2017) Inversion of well log data using improved shale model for determination of petrophysical parameters. J Indian Geophys Union 21:96–104

Marquardt DW (1959) Solution of nonlinear chemical engineering models. Chem Eng Prog 55:65–70

Mayer C, Sibbit AM (1980) GLOBAL, a new approach to computer-processed log interpretation. In: Proceedings of the 55th SPE annual fall technical conference and exhibition, paper 9341, pp 1–14

Menke W (1984) Geophysical data analysis: discrete inverse theory. Academic Press, Cambridge

Journal: 11004 Article No.: 9768 TYPESET DISK LE CP Disp.:2018/10/16 Pages: 25 Layout: Small-X

Michalewicz Z (1992) Genetic algorithms + data structures = evolution programs. Springer, Berlin
Mundim KC, Lemaire TJ, Bassrei A (1998) Optimization of non-linear gravity models through generalized simulated annealing. Physica A 252:405–416
Narayan JP, Yadav L (2006) Application of adaptive processing technique for the inversion of open hole logs recorded in oil fields of Indian basins. In: Proceedings of the SPG 6th international conference and exposition on petroleum geophysics, pp 505–512
Sambridge M (1999) Geophysical inversion with a neighbourhood algorithm—I. Searching a parameter space. Geophys J Int 138:479–494
Sambridge M, Drijkoningen G (1992) Genetic algorithms in seismic waveform inversion. Geophys J Int 109:323–342
Schlumberger (1989) Log interpretation principles/applications. Seventh printing. Schlumberger, Houston
Sen MK, Stoffa PL (1995) Global optimization methods in geophysical inversion. Elsevier, Amsterdam
Szabó NP (2004) Global inversion of well log data. Geophys Trans 44:313–329
Szabó NP (2016) Hydrocarbon formation evaluation using an efficient genetic algorithm-based factor analysis method. In: Proceedings of the 15th European conference on the mathematics of oil recovery, paper Mo P071, pp 1–12
Szabó NP, Dobróka M (2013) Extending the application of a shale volume estimation formula derived from factor analysis of wireline logging data. Math Geosci 45:837–850
Szabó NP, Dobróka M (2018) Exploratory factor analysis of wireline logs using a float-encoded genetic algorithm. Math Geosci 50:317–335
Szabó NP, Balogh GP, Stickel J (2018) Most frequent value-based factor analysis of direct-push logging data. Geophys Prospect 66:530–548
Szijártó M, Balázs L, Drahos D, Galsa A (2017) Numerical sensitivity test of three-electrode laterolog borehole tool. Acta Geophys 65:701–712
Szűcs P, Civan F (1996) Multi-layer well log interpretation using the simulated annealing method. J Pet Sci Eng 14:209–220
Tarantola A (2005) Inverse problem theory and methods for model parameter estimation. Society of Industrial and Applied Mathematics (SIAM), Philadelphia
Tlas M, Asfahani J (2008) Using of the adaptive simulated annealing (ASA) for quantitative interpretation of self-potential anomalies due to simple geometrical structures. J King Abdulaziz Univ Earth Sci 19:99–118
Toushmalani R (2013) Comparison result of inversion of gravity data of a fault by particle swarm optimization and Levenberg–Marquardt methods. SpringerPlus 2:462
Woznicka U, Jarzyna J, Cichy A (2009) Forward modelling and inverse problem in the comprehensive interpretation of well logs. Geophys Prospect 57:99–109
Wyllie MRJ, Gregory AR, Gardner LW (1956) Elastic wave velocities in heterogeneous and porous media. Geophysics 21:41–70
Yang XS (2013) Metaheuristic algorithms for inverse problems. Int J Innov Comput Appl 5:76–84

Journal: 11004 Article No.: 9768 TYPESET DISK LE CP Disp.:2018/10/16 Pages: 25 Layout: Small-X