

## META-ALGORITHM ASSISTED INTERVAL INVERSION FOR PETROPHYSICAL PROPERTIES PREDICTION

**Moataz Mohamed Gomaa Abdelrahman** 

*PhD student, University of Miskolc, Department of Geophysics  
Ain shams University, Department of Geophysics  
3515 Miskolc, Miskolc-Egyetemváros, e-mail: [mh.geophysicist@sci.asu.edu.eg](mailto:mh.geophysicist@sci.asu.edu.eg)*

**Norbert Péter Szabó** 

*professor, University of Miskolc, Department of Geophysics  
MTA-ME Geoengineering Research Group  
3515 Miskolc, Miskolc-Egyetemváros, e-mail: [norbert.szabo.phd@gmail.com](mailto:norbert.szabo.phd@gmail.com)*

**Mihály Dobróka** 

*professor emeritus, University of Miskolc, Department of Geophysics  
3515 Miskolc, Miskolc-Egyetemváros, e-mail: [dobroka@uni-miskolc.hu](mailto:dobroka@uni-miskolc.hu)*

### **Abstract**

*This research proposed integration between the GSS method and the interval inversion for regularizing the prediction algorithm of borehole logging data. The proposed algorithm has been tested using synthetic borehole logging data set contaminated with 5 % Gaussian noise. The GSS-based inversion scheme showed a smoother data distance convergence. The GSS-based inversion could decrease the conditional number of the sensitivity matrix. Furthermore, the algorithm has been applied to real data of a tight sand reservoir. The results from the Damped Least Square (DLSQ) scheme have been compared to those of the GSS-based scheme. The data distance of the DLSQ showed a rough convergence data distance line, while it showed a smooth convergence of the data distance line in the case of the GSS-based inversion. The convergence of the data distance of both schemes has been stabilized at 9.5%.*

**Keywords:** *Golden Section Search method, interval inversion, MATLAB, Borehole geophysics.*

### **1. Introduction**

The main objective of the well-logging data evaluation is to calculate the petrophysical parameters within each layer-thickness. Nowadays, the joint inversion process can be considered the most common technique for petrophysical parameters calculation. The joint inversion technique is better than the normal inversion because it could optimize the different petrophysical parameters at the same depth using well-logging data of the same depth. The inversion process depends on solving the linear equation at depth point, which is called point by point inversion or local inversion. This type of inversion process can be an overdetermined inverse problem as in the case of the borehole logging inversion, also, it can be an underdetermined inverse problem as in the case of potential data inversion, and finally, the inverse problem can be mixed determined as in case of seismic inversion (Menke, 1984). The maximum overdetermined ratio of the borehole logging inversion can be more or less 1.5, depending on the

available logging data at this depth point (Dobróka et al., 2016). The use of the local inversion has a significant advantage which is fast convergence, but it is highly dependent on the initial model. Besides that, the point-by-point inversion has limited parameters optimization availability due to the overdetermined ratio. Therefore, (Szabó and Dobróka, 2020) and (Dobróka and Szabó, 2005) have proposed an improved inversion scheme that can optimize the petrophysical parameters of an interval of points. This concept is based on the optimization of the series expansions coefficients of the polynomial which represent the change of the parameter within a layer thickness. The interval inversion method increased the overdetermined ratio significantly and gave the availability to optimize more parameters within the inversion process. Therefore, the interval inversion could solve the problem of the limited overdetermined ratio but as the number of the predicted parameters increase or in case of the heterogeneous reservoirs ill posed sensitivity matrix problem can appear. To overcome the ill-posed problem the global inversion has been proposed, where in the global inversion there is no need for deriving the derivatives and too many priori information (Liu et al., 2018). But the drawback of the global inversion is, that it takes a very long time to find the global optima. Therefore, to use the advantage of the global inversion and the fast convergence of the linearized inversion, a hybrid inversion has been introduced by (Dobróka and Szabó, 2005) and (Szabó, 2018a). The hybrid algorithm has been introduced as two main loops where the outer one is the global optimization which can optimize the thickness of the layers, while the inner one is the linearized inversion for predicting the petrophysical parameters regarding each derived layer thickness.

One of the most important parameters in any optimization technique is the hyper parameters. These kinds of parameters are used to control the optimization problem. One of these hyper parameters is the damping factor, which is used to overcome the ill-posed problem. The physical meaning of the ill-posed problem can be explained as a set of equations with a high-condition number. The high condition number means that any small changes in the sensitivity matrix can cause a huge change in the predicted parameters. Such problems can lead to an unstable inversion procedure (Prabhu, 2018) and (Van Rijn and Hutter, 2018). This research proposed integration between the GSS optimization method with the interval inversion technique for providing a suitable damping scheme for improving the stability of the inversion procedure.

In any inversion procedure, there are two groups of parameters. The first group is the petrophysical parameters which can be considered the targeted parameters. These parameters change with depth so we can call them depth-dependent parameters (e.g., porosity, fluid content, and matrix volume). The other group of parameters so-called the zone parameters, such parameters can be considered as fixed parameters within a certain layer, but vary from one to another layer (e.g., cementation exponent, water resistivity, and shale resistivity).

## 2. Theoretical overview

### 2.1. The inversion problem

The concept of the inversion procedure relies on the relationship between the data and the estimated parameters. This relationship is called the response function. These response functions can be used to calculate the well-logging data from the petrophysical model. The process of data calculation from the petrophysical data is the so-called forward modeling phase. The initial model parameters vector can be shown as follow:

$$m = [\emptyset, V_{sh}, V_{Vsd}, S_w, S_{x0}]^T, \quad (1)$$

where  $\emptyset$  is the percentage of the pore spaces of the formation,  $V_{sh}$  and  $V_{Vsd}$  are the volumes of shale and sand within the formation, respectively,  $S_w$  and  $S_{x0}$  are the water saturation of the uninvasion and invasion zones respectively. The well logging data calculated from the petrophysical model parameters can be arranged in a column vector called the data vector as follow:

$$d^{(0)} = [GR, \emptyset_N, \rho_b, R_s, R_d, P_e]^T, \quad (2)$$

where GR is the Gamma Ray,  $\emptyset_N$  is neutron porosity,  $\rho_b$  is gamma-gamma,  $R_s$  and  $R_d$  are the shallow and deep resistivities respectively, and  $P_e$  is the photo-electric log. The T symbol denotes matrix transpose. The relationship between the model parameters and the calculated data can be shown as follows:

$$d^{(c)} = g(m), \quad (3)$$

where g is the response function used for calculating the well-logging data. The initial forward model can be constructed using the following equations (Alberty and Hashmy, 1984):

$$GR = \rho_b^{-1} \left( V_{sh} GR_{sh} \rho_{sh} + \sum_{i=1}^N V_{ma,i} GR_{ma,i} \rho_{ma,i} \right), \quad (4)$$

$$\rho_b = \emptyset [\rho_{mf} - 1.07(1 - S_{x0})(\alpha_0 \rho_{mf} - 1.24 \rho_h)] + V_{sh} \rho_{sh} + \sum_{i=1}^N V_{ma,i} \rho_{ma,i}, \quad (5)$$

$$\emptyset_N = \emptyset \left\{ \begin{array}{l} \emptyset_{N,mf} - (1 - S_{x0}) C_{cor} \\ -2\emptyset(1 - S_{x0}) - S_{hf}(1 - 2.2\rho_h) \\ \cdot [1 - (1 - S_{x0})(1 - 2.2\rho_h)] \end{array} \right\} + V_{sh} \rho_{sh} + \sum_{i=1}^N V_{ma,i} \emptyset_{ma,i}, \quad (6)$$

$$P_e = \frac{1.07}{\rho_b + 0.19} [\emptyset S_{x0} U_{mf} + \emptyset(1 - S_{x0}) U_h + V_{sh} U_{sh} + \sum_{i=1}^N V_{ma,i} U_{ma,i}], \quad (7)$$

$$\frac{1}{\sqrt{R_d}} = \left[ \frac{V_{sh}^{(1-0.5V_{sh})}}{\sqrt{R_{sh}}} + \frac{\sqrt{\emptyset}^m}{\sqrt{aR_w}} \right] \sqrt{S_w}^{-n}, \quad (8)$$

$$\frac{1}{\sqrt{R_s}} = \left[ \frac{V_{sh}^{(1-0.5V_{sh})}}{\sqrt{R_{sh}}} + \frac{\sqrt{\emptyset}^m}{\sqrt{aR_{mf}}} \right] \sqrt{S_{x0}}^{-n}, \quad (9)$$

$$\emptyset + V_{sh} + \sum_{i=1}^N V_{ma,i} = 1. \quad (10)$$

The fractional volume of  $i$ -th matrix constituent is represented by  $V_{ma, i}$  (v/v).  $\emptyset_{Ns}$  is the reading of the neutron log against a thick pure sandstone rock. The physical properties of the mud filtrate and the mud correction are expressed as  $mf$  and  $C_{cor}$  respectively. Parameters  $m$ ,  $n$ , and  $a$  are the Archie constants (cementation exponent, saturation exponent, and tortuosity factor). The zone parameters are considered

constant within the forward modelling phase. Special and conventional core analyses, drilling parameters and reports, and processing techniques can be used for deriving the zone parameters. The invaded and uninvaded zones' resistivities were calculated using the Indonesian formulae. The following table shows the zone parameters used for calculating the well-logging data in the forward modelling phase.

**Table 1.** Zone parameters are used for the traditional processing of borehole geophysics (Szabó and Dobróka, 2020).

Well log	Zone Parameter	Symbol	Value	Dimensional unit
Natural gamma-ray	Sand	$GR_{sd}$	10	API
	Shale	$GR_{sh}$	154	API
Gamma-gamma (density)	Sand	$\rho_{sd}$	2.65	g/cc
	Shale	$\rho_{sh}$	2.54	g/cc
	Mud filtrate	$\rho_{mf}$	1.02	g/cc
	Hydrocarbon(gas)	$\rho_h$	0.2	g/cc
Neutron porosity	Sand	$\Phi_{N,sd}$	-0.04	v/v
	Shale	$\Phi_{N,sh}$	0.31	v/v
	Mud filtrate	$\Phi_{N,mf}$	0.95	v/v
	mf correction coefficients	$C_{cor}$	0.69	-
	Residual hydrocarbon	$S_{hrf}$	1.2	-
Deep resistivity	Shale	$R_{sh}$	1	Ohm.m
	Pore water	$R_w$	0.5	Ohm.m
	Mud filtrate	$R_{mf}$	0.28	Ohm.m
	Cementation exponent	$m$	1.5	-
	Saturation exponent	$n$	1.8	-
	Tortuosity factor	$a$	1	-
Photo-electric Index	Sand	$U_{sd}$	4.8	Barn/cm <sup>3</sup>
	Shale	$U_{sh}$	9	Barn/cm <sup>3</sup>
	Mud filtrate	$U_{mf}$	0	Barn/cm <sup>3</sup>
	Hydrocarbon(gas)	$U_h$	0	Barn/cm <sup>3</sup>

Finally, the relative data distance between the calculated data from the forward modeling phase and the measured or actual data is as follows:

$$D_d = \sqrt{\left[ \frac{1}{N} \sum_{k=1}^N \left( \frac{d_k^{(obs)} - d_k^{(cal)}}{d_k^{(obs)}} \right)^2 \right]} * 100 (\%), \quad (11)$$

where the measured data expressed as  $d_k^{(obs)}$ , and the calculated data expressed as  $d_k^{(cal)}$ , while the symbol N is the number of inverted data. In the case of synthetic data, the model distance can be calculated as follows:

$$D_m = \sqrt{\left[ \frac{1}{M} \sum_{k=1}^M \left( \frac{m_k^{(obs)} - m_k^{(cal)}}{m_k^{(obs)}} \right)^2 \right]} * 100 (\%), \quad (12)$$

where the  $m_k^{(obs)}$  and  $m_k^{(cal)}$  are the actual synthetic model parameters and predicted or estimated model parameters respectively.

## 2.2. Interval inversion problem

The previous section shows that the overdetermined ratio is limited in the point-by-point inversion problem, which results in a noise-sensitive inversion procedure as well as a limited number of the optimized parameters (Szabó and Dobróka, 2020) and (Szabó and Dobróka, 2018). The interval inversion concept aims to increase the number of inverted data by representing the petrophysical parameters within a depth interval as a continuous function (Dobróka, 1993). Regarding this concept, the interval inversion problem counts on the discretized series expansion of petrophysical parameters. The data column vector can be expressed as a depth-dependent calculated data as follows:

$$d_j^{(c)}(z) = g_j(m_1(z), m_2(z), \dots, m_M(z)), \quad (13)$$

where  $M$  is the number of the predicted parameters at  $z$  depth,  $d_j^{(c)}$  is the  $j$ -th calculated data. Therefore, the discretization of the petrophysical parameters can be expressed as:

$$m_i(z) = \sum_{q=1}^{Q_i} B_q^{(i)} \Psi_q(z), \quad (14)$$

where  $m_i$  represents the  $i$ -th model parameter,  $B_q^{(i)}$  is the  $q$ -th expansion coefficient, and  $\Psi_q(z)$  is the basis function. Regarding the interval inversion concept, the term of the basis function can be considered as a known. By substituting the equation 14 in equation 3 the interval inversion problem can be expressed as:

$$d_k^c(z_k) = \sum_{i=1}^P \sum_{q=1}^{Q(i)} G_{ki} B_q^{(i)} \Psi_q(z_k), \quad (15)$$

To formalize the interval inversion in the layer-wise homogeneous layers sequence model can be represented using the step function or the Heaviside function. The Heaviside function divides the model into homogeneous segments (Szabó, 2018b; Szűcs et al., 2021), and (Dobróka et al., 2009). The layer-wise interval inversion can be expressed as:

$$\Psi_q(z) = u(z - Z_{q-1}) - u(z - Z_{q+1}). \quad (16)$$

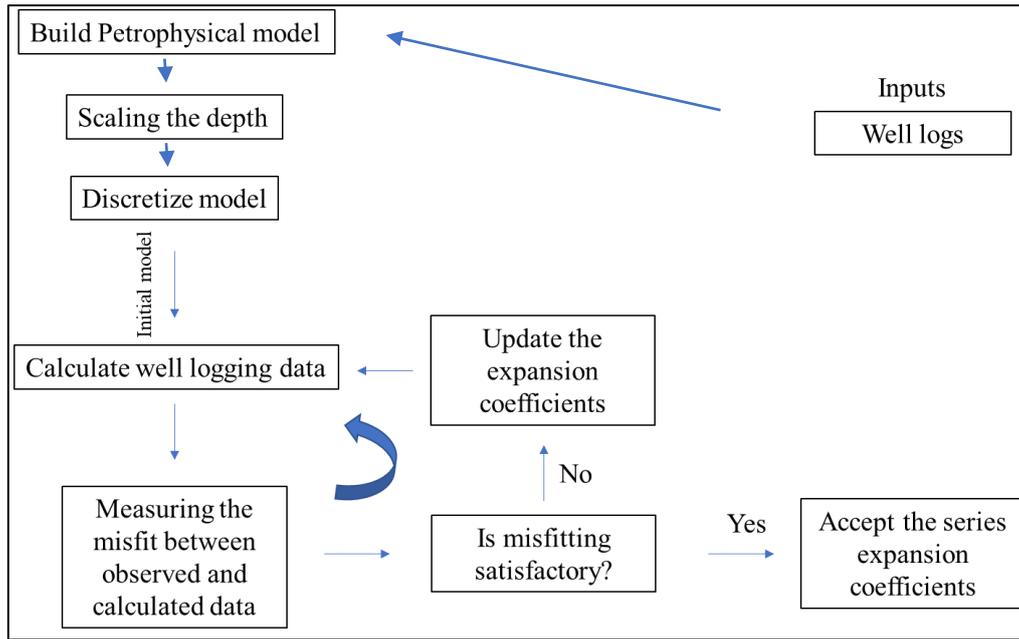
After understanding the interval concept the model vector can be expressed in the light of the series expansion concept as follow:

$$\mathbf{m} = [B_1^{(1)}, \dots, B_Q^{(1)}, \dots, B_1^{(M)}, \dots, B_Q^{(M)}, Z_1, \dots, Z_{Q-1}]^T. \quad (17)$$

According to (Szabó et al., 2021; Liu et al., 2018; Szabó and Dobróka, 2020), and (Dobróka et al., 2016), in the interval inversion procedure, the objective function that must be reduced is a weighted function based on the least-squares criteria.

$$E(\mathbf{m}) = \sum_{p=1}^P \sum_{j=1}^N \left( \frac{d_{pj}^m - d_{pj}^c}{d_{pj}^m} \right)^2 = \min. \quad (18)$$

The interval inversion workflow can be seen in the following figure.



**Figure 1.** Interval inversion Workflow.

DLSQ also can be called Levenberg-Marquardt (LM) which is a linear inversion method, in which the initial model is refined within an iterative algorithm. The criteria for stopping the algorithm could be a threshold value of the data distance or a certain number of the iteration. The model update equation in case of a point-by-point inverse problem is:

$$(\mathbf{G}(\mathbf{m}^i)^T \mathbf{G}(\mathbf{m}^i) + \lambda I) \delta \mathbf{m} = -\mathbf{G}(\mathbf{m}^i)^T \delta \mathbf{d}, \quad (19)$$

while in the case of interval inversion the model update equation can be as follow:

$$(\mathbf{G}(\boldsymbol{\beta}^i)^T \mathbf{G}(\boldsymbol{\beta}^i) + \lambda I) \delta \boldsymbol{\beta} = -\mathbf{G}(\boldsymbol{\beta}^i)^T \delta \mathbf{d}, \quad (20)$$

where  $\lambda$  is the damping factor.

### 2.3. Damping factor meaning

The physical meaning of the damping factor is introduced by (Levenberg, 1944), where the damping factor is used to conserve the stability of the inversion process. The instability comes from insufficient extracted information from the data or the noise sensitivity. (Menke and Eilon, 2015) mentioned that the

damping factor can be defined as prior information, which is implemented from the previous expectation of the solution behaviour. According to (Menke and Eilon, 2015) the stability of the inversion process can be obtained using two ideas. The first one is using a certain constraint. But by this way, we increase our influence on the solution by using a priori information. The second idea is to use the covariance matrix to assess the correlation of the different model parameters.

Usually, the damping factor is used as a positive and relatively high number which is decreased by the iterations until a value close to zero. The instability of the inversion problem can be called an ill-posed problem. The ill-posed problem can be represented mathematically using the conditional number as follow:

$$k = \frac{\lambda_{max}}{\lambda_{min}}, \quad (21)$$

where  $\lambda_{max}$  and  $\lambda_{min}$  are the maximum and minimum eigenvalue of the sensitivity matrix. The ill-posed problem refers to a high condition number while the well-posed indicates a small condition number of the set of solving equations. In the case of a high condition number, any small change in the sensitivity matrix cause a bigger change in the estimated parameters. Therefore, choosing the proper damping factor can avoid instability and improve the condition number. (Abdelrahman et al., 2021) and (Heriyanto and Srigutomo, 2017) shows the feasibility of extracting the damping factor from the decomposition of the sensitivity matrix, in which the damping factor is a function of the eigenvalues of the decomposed sensitivity factor.

#### 2.4. Golden Section Search method (GSS)

The GSS is one of the unimodal optimization algorithms. Besides that, it can be called optimization with interval reduction. Therefore, the unimodal function is a continuous function over an interval [a,b] if there is one minimum point that is included in the same [a,b] interval (Noroozi et al., 2022). The intuition behind the integration between the interval inversion algorithm and the GSS is that the constructed algorithm is a hybrid algorithm. In other words, the MATLAB program is based on two loops, the first one is the outer loop which contains the GSS that will optimize the damping factor. The second loop is the inner loop which applies the chosen damped factor from GSS and applies the interval inversion procedure. Furthermore, the acceptance criterion is based on the data distance calculated based on the optimized damped factor (Chakraborty and Panda, 2016) and (Tsai et al., 2010).

Regarding the GSS, there are two goals, the first one is finding the optimal damped factor for the interval, while the second one is finding the optimal value with a minimum number of function calls. The theoretical background of the interval reduction optimization problems such as the bisection method or equal interval method can be explained mathematically as follows:

- Choose a mid-point between a and b,  $m = (a+b)/2$  in case of bisection method and a small  $d > 0$
- Let  $X_1 = (m-d)/2$  and  $X_2 = (m+d)/2$ .
- If  $f(X_1) < f(X_2)$  then reduce the interval into  $[a, X_1]$
- Else the reduced interval will be  $[X_2, b]$

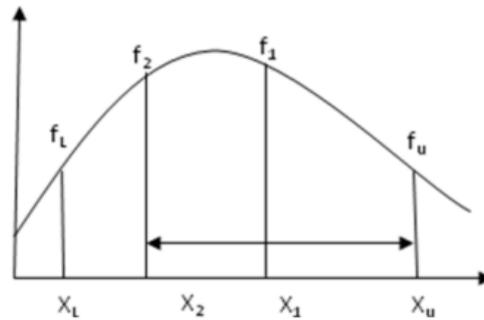


Figure 2. Equal interval method intuition (Chang, 2015).

The weak point of the equal interval method is it will take a long time to find the optimum point. Therefore, the GSS is recommended in the case of hybrid algorithms. The Golden Section choose three points  $X_L$ ,  $X_1$  and  $X_u$  ( $X_L < X_1 < X_u$ ) along the  $x$ -axis with corresponding values of the function  $f(X_L)$ ,  $f(X_1)$ , and  $f(X_u)$ , respectively. Since  $f(X_1) > f(X_L)$  and  $f(X_1) > f(X_u)$  (Yazıcı et al., 2021). The maximum must lie between  $X_1$  and  $X_u$ . Now a fourth point denoted by  $x_2$  is chosen to be between the larger of the two intervals of  $[X_L, X_1]$  and  $[X_1, X_u]$ . The intermediate points  $X_1$  and  $X_2$  are chosen such that, the ratio of the distance from these points to the boundaries of the search region is equal to the golden ratio (Pejic and Arsic, 2019).

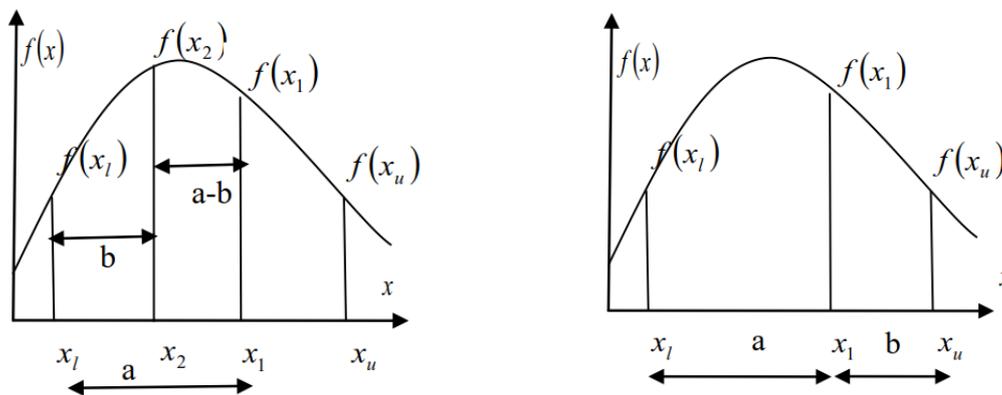


Figure 3. (Right) Determining the first intermediate point, (left) Determining the second intermediate point.

The algorithm of the GSS can be concluded as follow:

- Initialize the known boundaries  $[X_L, X_u]$
- Determine two intermediate points  $X_1$  and  $X_2$  using the golden ratio  $d$  where  $d = \frac{\sqrt{5}-1}{2}(x_u - x_L)$
- Then evaluate the inversion loop and extract the data distance of the lower, upper, and intermediate points.
- Finally, if  $X_u - X_L < \text{threshold value}$  then the damped factor can be chosen as  $(X_u+X_L)/2$  stop the outer loop and accept the model of the optimal damping factor.

### 3. Results

The proposed algorithm was carried out on synthetic and real data. The synthetic data was built based on a sand-shale model with water and hydrocarbon fluids filling the pore spaces, while the real data represents a heterogeneous reservoir of hydrocarbon field. The workflow of the proposed algorithm can be shown in figure (4).

#### 3.1. Synthetic data

The GSS-based interval inversion was carried out on synthetic data contaminated with 5 % Gaussian noise. The synthetic model consists of a sandstone formation filled with a hydrocarbon that overlies a shale layer which overlies another sandstone layer filled with water. A 20-degree Legendre polynomial was used to fit the suggested synthetic model. The forward well-logging data vector calculated using equations 3 to 8 consists of gamma-ray, density, neutron, compressional sonic, shallow, and deep resistivities, and photoelectric logs. The Oil-Water Contact (OWC) suggested being located at 6 m.

The calculated data from the response functions were inverted using DLSQ and GSS-based interval inversion algorithms. Figure 5 shows the synthetic data and the initial model of the inversion procedure, while Figure 6 shows the actual model parameters with the initial model parameters used for the inversion procedure. The model parameters of the synthetic dataset were constructed based on 5 model parameters (porosity, volume of sand, volume of shale, water saturation of uninvaded zone, and water saturation of the invaded zone). So, by taking into consideration the number of the parameters and the degree of the polynomial, the number of the unknowns of the digitized model parameters is  $5 \cdot (20+1) = 105$  unknowns. At the same time, the number of the overdetermined ratio can be calculated from the number of the logging data at each depth  $N_L$  multiplied by the number of the data points (depth points)  $N$ ,  $7 \cdot 200 = 1400$  data points, therefore, the overdetermined ratio is  $1400/105 \approx 13$  in case of the suggested synthetic model.

Figure (7) shows the fitting between the synthetic data and the predicted data using the GSS-based interval inversion, while figure (8) shows the predicted model parameters fitted to the actual model parameters of the synthetic model.

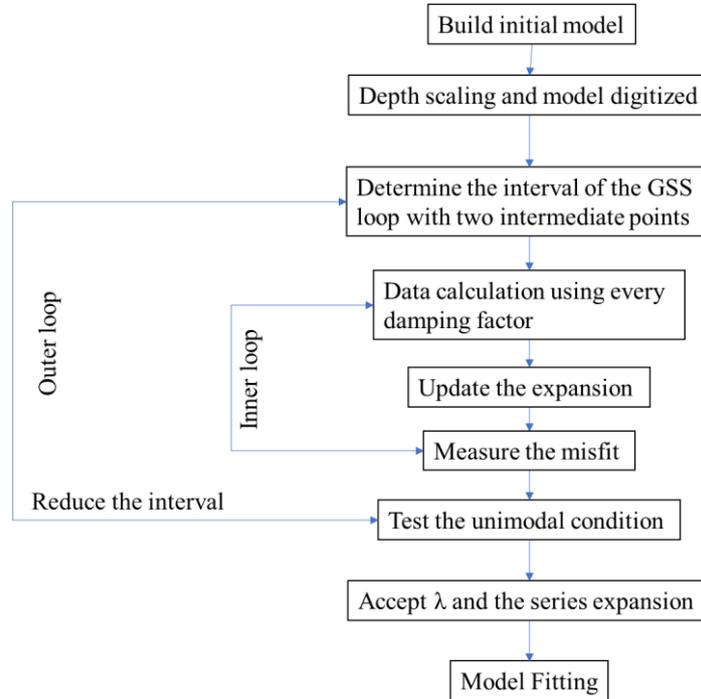


Figure 4. The GSS and interval inversion integration workflow.

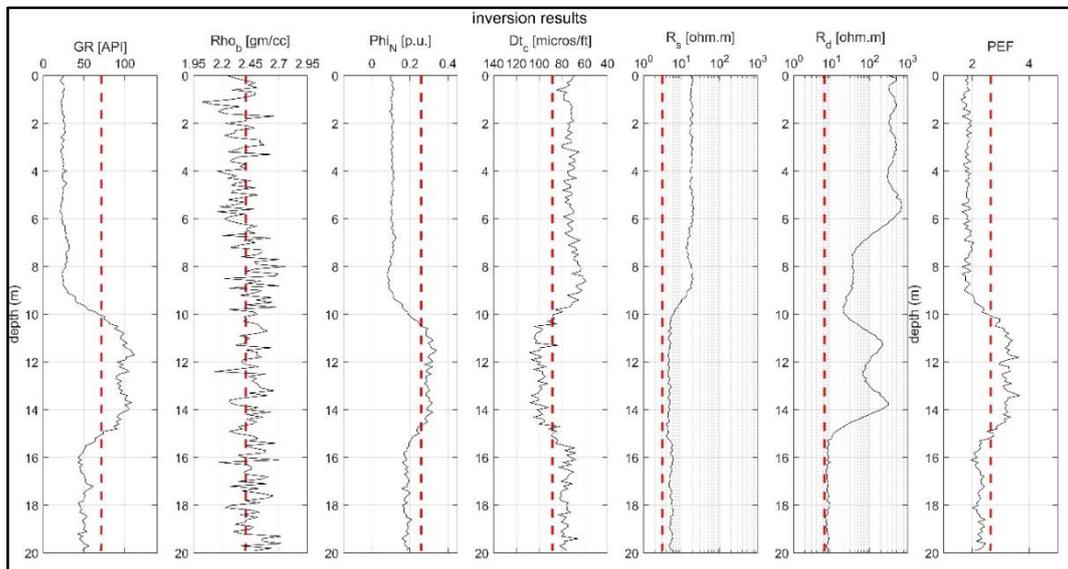


Figure 5. 5% Gaussian noise-contaminated synthetic data; the red dashed lines represent the initial model, while the solid black lines represent the synthetic data.

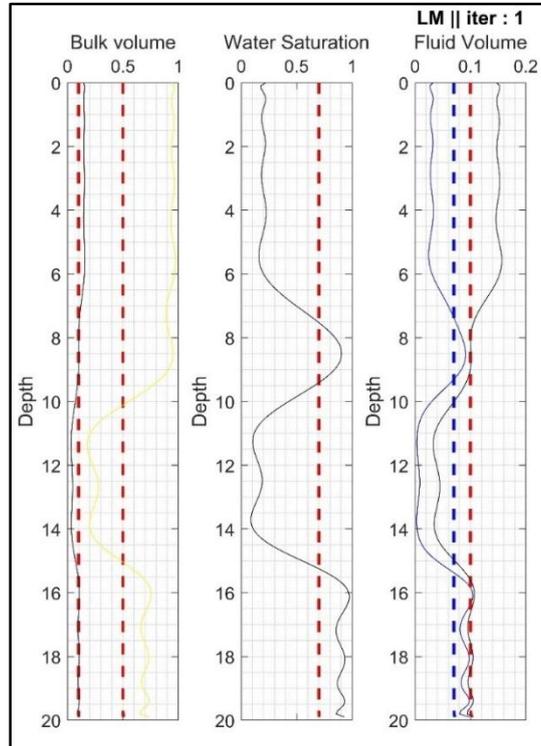


Figure 6. The model parameters of the synthetic data and the initial model parameters; the red dashed lines represent the initial model, while the solid black lines represent the synthetic data.

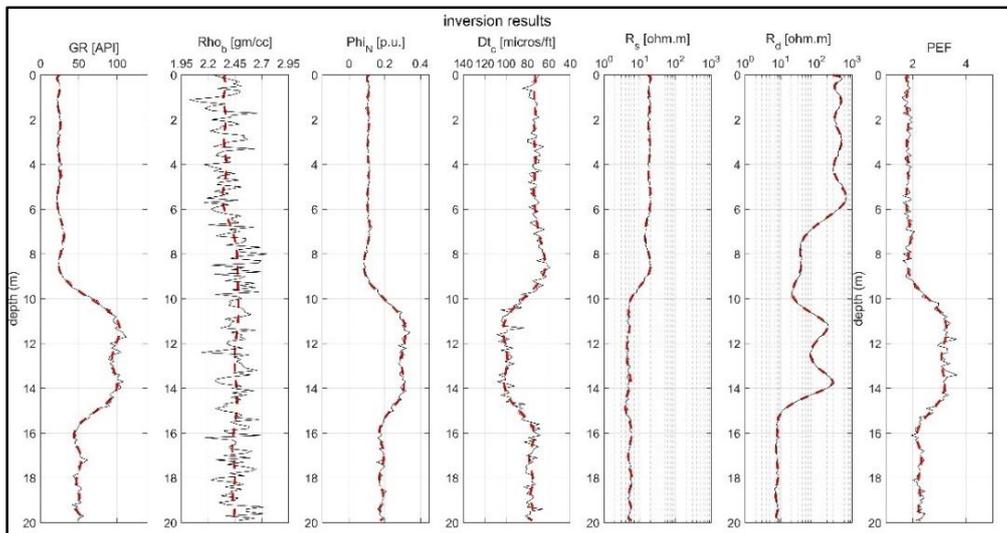
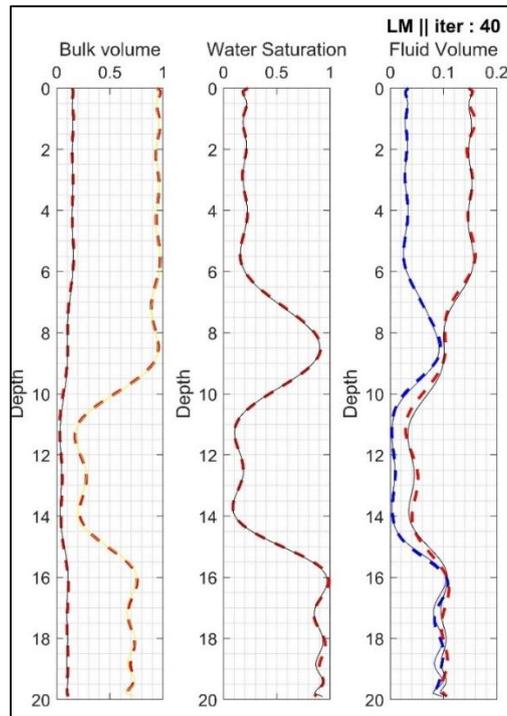
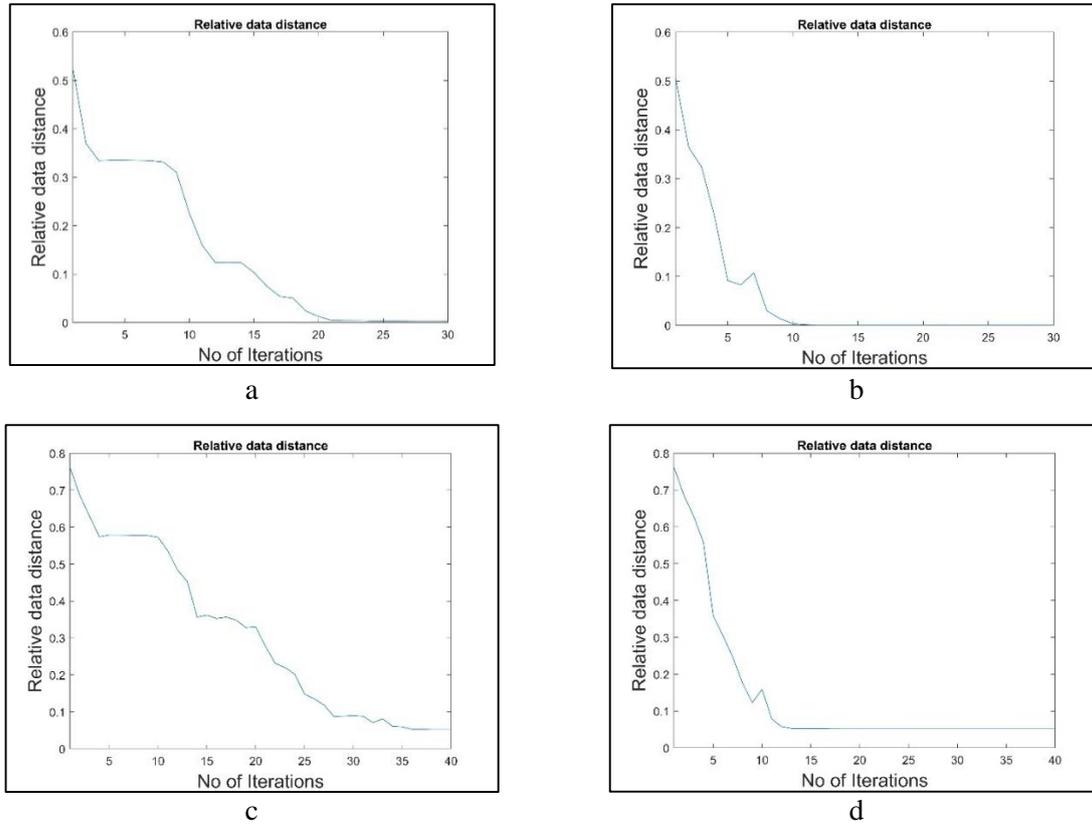


Figure 7. Fitting between the synthetic data and the calculated data (iteration 40); the red dashed lines represent the initial model, while the solid black lines represent the synthetic data.



**Figure 8.** The synthetic and the predicted model parameters (iteration 40); the red dashed lines represent the initial model, while the solid black lines represent the synthetic data.

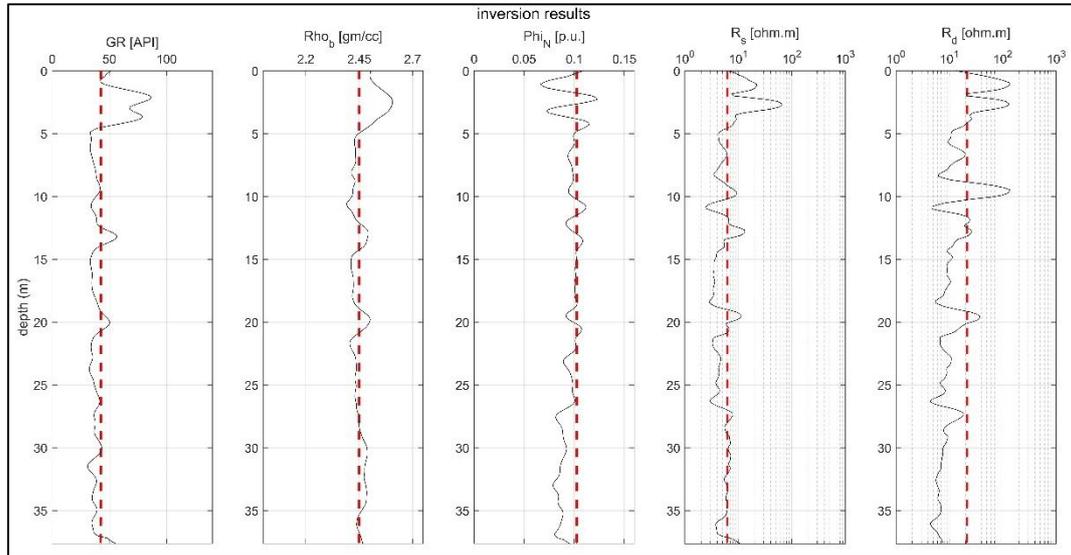
The difference between the GSS-based interval inversion and the conventional interval inversion can be seen from the data distance convergence curve (figure 9). The data distance of the conventional interval inversion shows what we can call a plateau data distance, this plateau data distance causes no change in the optimization process and increase the number of the iteration. This research attributed the reason for this effect to the choice of the suitable damped factor, which directly affects the condition number of the sensitivity matrix and causes a higher damp effect. However, the GSS-based interval shows a rapid convergence of the data distance with a smooth behavior. The data distance in the case of the conventional interval of inversion shows a rough convergence behavior. In figure 9, the data distance of the noise free dataset and 5% Gaussian noise-contaminated datasets can be seen. Where the plateau effect can be significantly detected. There is a difference in the iteration of the convergence stability in the case of the conventional interval inversion noise free and the contaminated dataset, while the GSS one shows the same iteration of convergence stability. The data distance convergence of the conventional interval inversion started to stabilize at iteration 35 in the 5% Gaussian noise dataset and iteration 22 in the noise free dataset, while in the case of the GSS-based interval inversion the data distance convergence started to stabilize around iteration 10 in both cases.



**Figure 9.** Data distance of conventional and GSS-based interval inversion, (a) and (b) noise free dataset, (c) and (d) are the contaminated datasets.

### 3.2. Field data

After examination of the conventional and the GSS-based interval inversion using synthetic data, the proposed algorithm has been applied to the real data. The real well logging data was acquired from a hydrocarbon field located in the northeastern part of Egypt. According to the regional geology of the field, the target reservoir is a Jurassic reservoir. It is considered a tight sandstone layer with a highly heterogeneous degree because of the diagenesis process that is responsible for transforming the kaolinite into illite by reducing both the permeability and the porosity. The data vector of this well consists of gamma-ray, neutron, density, and shallow and deep resistivity logs. Figures (10) show the well-logging dataset and the initial model.

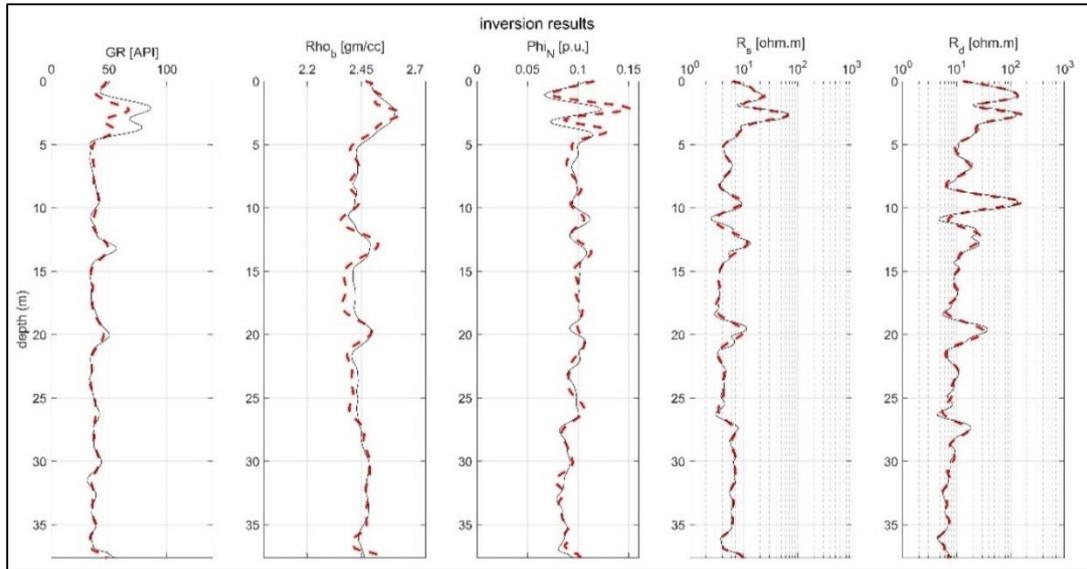


**Figure 10.** The well logging dataset (black curve) and the initial model (red dashed line).

The initial model was constructed based on the porosity of 0.1, the volume of sand of 0.8, and water saturation of 0.2. To reduce the number of unknowns the balance equation was used to derive the volume of shale and the water saturation of the invaded zone can be calculated from the water saturation of the invaded zone. Therefore, the overdetermined ratio can be calculated as follow:

$$O = \frac{N_p * (N_g + 1)}{N} = \frac{3 * (50 + 1)}{1240} = 8.1, \quad (22)$$

where  $N_p$  is the number of parameters, and  $N_g$  is the degree of the polynomials. Figure (11) shows the fitting between the predicted and the real datasets, while figure (12) shows the predicted model parameters. The inversion results show that the reservoir consisting mainly of sandstone layers underlies the thin shale layer. Furthermore, the inversion procedure shows that the sandstone layer filled with hydrocarbon starts from 5m to 27m with a thickness of 22m, where the contact between hydrocarbon and water lies at depth of 27. The results show also that there is a layer of shale with around 5 m thickness. The resulted parameters can successfully show the variation of the quality of the sand which is reflected in the amount of hydrocarbon saturation.



**Figure 11.** Fitting between the real and the predicted datasets (iteration 60); (black curve) and the initial model (red dashed line).

In the case of the conventional interval inversion, the data distance curve shows a rough convergence curve with significant peaks at iterations from 22 to 35, before reaching its stability at iteration 40. While data distance of the GSS-based interval inversion algorithm shows a smooth convergence behavior. The GSS-based inversion could overcome the problem of the presence of a sudden increase in the data distance and increase the stability of the inversion procedure. Furthermore, the convergence of the data distance started to be stable at the iteration of 20. The stable data distance of the conventional interval inversion and the GSS-based interval inversion is around 9.75%. Figure (13) shows the difference between the data distance behavior in the case of the conventional interval inversion and the GSS-based interval inversion. The stopping criterion of the algorithms is a threshold number of iterations, which have been chosen to be 60 iterations.

To study the convergence behavior of the proposed algorithm, the conditional number was calculated for each iteration. Figure (14) shows the difference between the conditional number in the case of the conventional and the GSS inversion schemes. The calculated conditional number is scaled using the extent scaling to represent which iterations have a higher conditional number and which have a low. The red curve is related to the GSS-based interval inversion, while the blue one is related to the conventional interval inversion. The GSS-based interval inversion shows that using the GSS to optimize the damping factor successfully conserves the low conditional number (which avoids the ill-posed problem) during the initial model is far from the solution or in case of rapid increase of the data distance. A significant increase in the conditional number in the case of the GSS-based interval inversion at iteration 20 aid the inversion procedure to avoid the increase in the data distance that can be observed in the case of the conventional interval inversion. For understanding the distribution of the conditional number behavior between low and high, Gaussian fitting was performed on the conditional number of the conventional and the GSS interval inversion (figure 15).

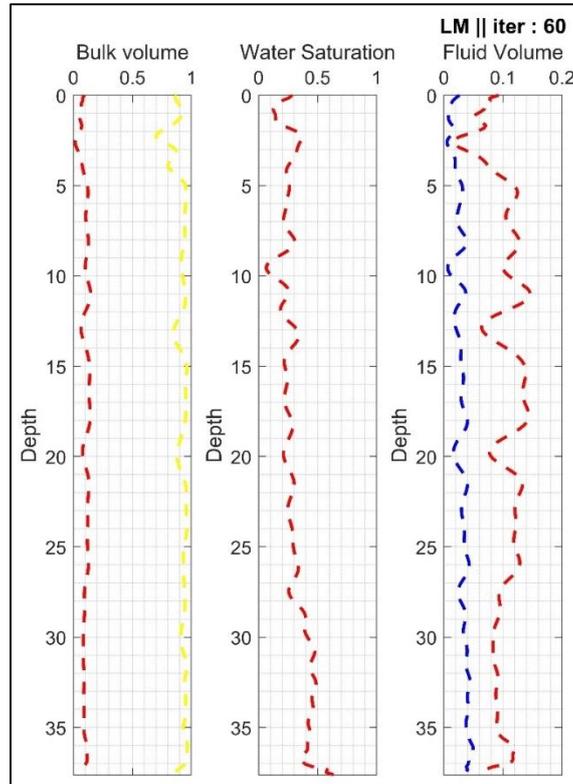


Figure 12. The predicted parameter of the real data (iteration 60); (black curve) and the initial model (red dashed line).

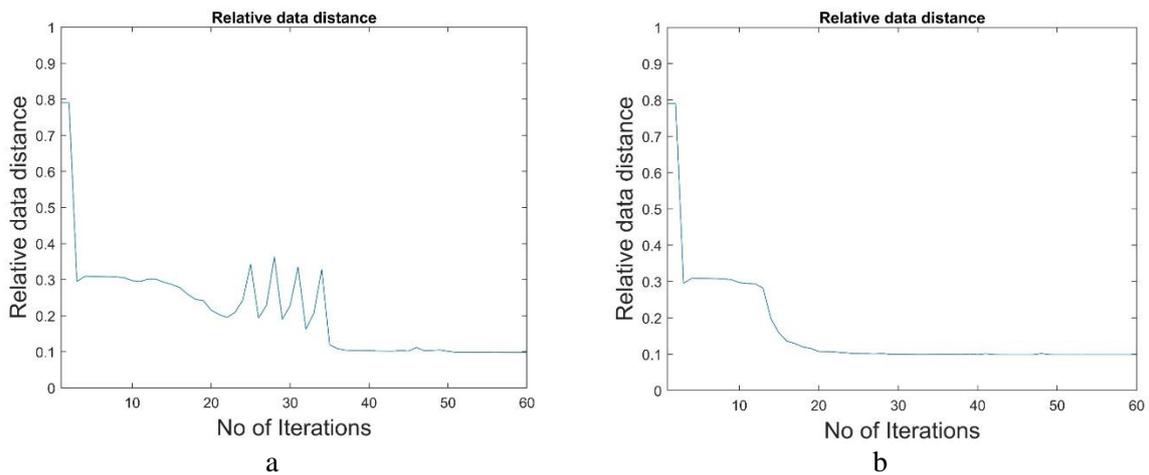
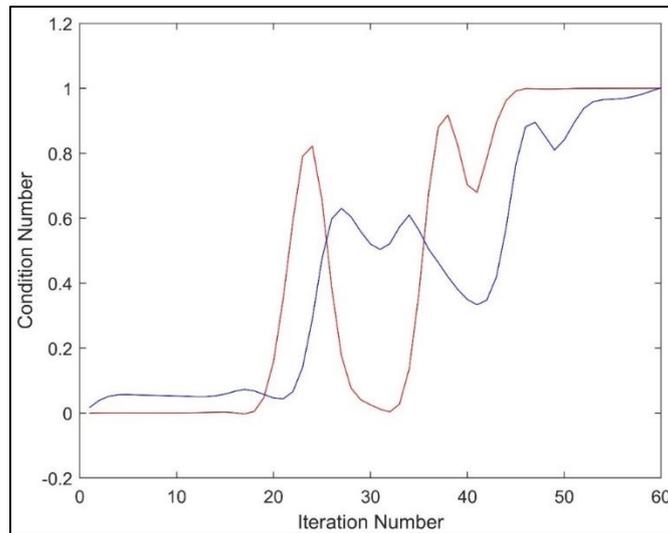


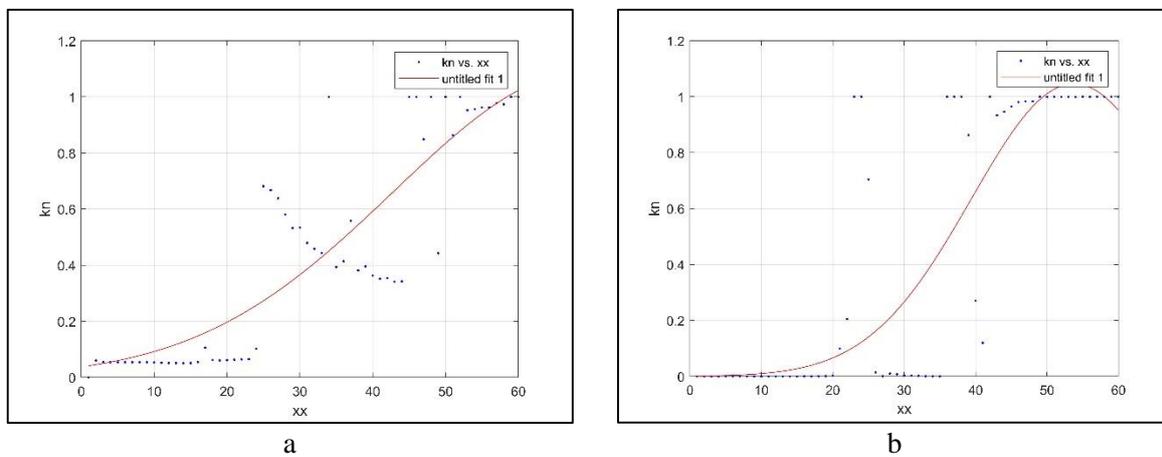
Figure 13. Data distance curves, (a) conventional interval inversion, (b) GSS- based interval inversion.

The Gaussian fitting can show that the conventional interval inversion follows a gradual increase in the conditional number without taking into the consideration any change in the data distance behavior,

while the GSS-based interval inversion there is a half of Gaussian distribution to show the decrease in the early iterations and high in the late iterations. The GSS-based interval inversion shows a nonlinear decrease of the damping factor by optimal choosing of its value, while the conventional one just decreases the damping factor within a linear relationship.



**Figure 14.** Scaled conditional number, (blue) is the conventional interval inversion, (red) is the GSS-based interval inversion.



**Figure 15.** Gaussian fitting of both inversion schemes, (a) conventional interval inversion, (b) GSS-based interval inversion.

#### 4. Conclusion

The conventional interval inversion and the GSS-based interval inversion can be used in the case of synthetic and heterogeneous real data. The damping factor reduces as a linear relationship in the case of

the conventional inversion, while the damping process follows a nonlinear relationship in the case of the GSS-based interval. the damping factor is the key to the stability of the inversion process.

The conventional inversion can suffer from the proposed term ‘plateau data distance’, this effect can increase the number of the iterations. The heterogeneous dataset shows a rough convergence curve related to the linear relationship of the damping factor reducing process. The GSS could overcome the plateau effect in the case of synthetic data, which reduces the number of iterations needed for convergence stability. Besides that, the GSS can be used in the case of the heterogeneous real dataset of well logging. Also, it could guarantee the stability and smoothness of the convergence curve of the data distance. The GSS-based interval inversion shows a stable inversion process in the case of synthetic datasets with free noise and 5% Gaussian noise.

The real data shows that the reservoir consists of sandstone varies in quality, which affects the hydrocarbon saturation and storage capacity. Both algorithms show a stable convergence of the data distance at 9.75%.

## Reference

- [1] Menke, W. (1984). *Geophysical data analysis: Discrete inverse theory*. Academic Press Inc.
- [2] Dobróka, M., Szabó, N. P., Tóth, J., and Vass, P. (2016). Interval inversion approach for an improved interpretation of well logs. *Geophysics*, 81(2), D155–D167. <https://doi.org/10.1190/geo2015-0422.1>
- [3] Szabó, N. P. and Dobróka, M. (2020). Interval inversion as innovative well log interpretation tool for evaluating organic-rich shale formations. *J. Pet. Sci. Eng.*, 186(November), 106696. <https://doi.org/10.1016/j.petrol.2019.106696>
- [4] Dobróka, M. and Szabó, P. N. (2005). Combined global / Linear inversion of well-logging data in layer-wise. *Acta Geophys.*, 40(2), 203–214. <https://doi.org/10.1556/AGeod.40.2005.2.7>
- [5] Liu, S., Liang, M., and Hu, X. (2018). Particle swarm optimization inversion of magnetic data: Field examples from iron ore deposits in China. *Geophysics*, 83(4), J43–J59. <https://doi.org/10.1190/geo2017-0456.1>
- [6] Szabó, N. P. (2018a). A genetic meta-algorithm-assisted inversion approach: hydrogeological study for the determination of volumetric rock properties and matrix and fluid parameters in unsaturated formations. *Hydrogeol. J.*, 26(6), 1935–1946. <https://doi.org/10.1007/s10040-018-1749-7>
- [7] Prabhu (2018). *Understanding Hyperparameters and its Optimisation techniques*. <https://towardsdatascience.com/understanding-hyperparameters-and-its-optimisation-techniques-f0debba07568>
- [8] Van Rijn, J. N. and Hutter, F. (2018). Hyperparameter importance across datasets. In: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Jul. (pp. 2367–2376). <https://doi.org/10.1145/3219819.3220058>
- [9] Alberty, M. and Hashmy, K. (1984). Application of ULTRA to log analysis. In: *SPWLA 25th Annual Logging Symposium 1984*, (pp. 1–17).
- [10] Szabó, N. P. and M. Dobróka, M. (2018). Exploratory factor analysis of wireline logs using a float-encoded genetic algorithm. *Math. Geosci.*, 50(3), 317–335. <https://doi.org/10.1007/s11004-017-9714-x>
- [11] Dobróka, M. (1993). Scientific report for the Hungarian Oil and Gas Company.
- [12] Szabó, N. P. (2018). A genetic meta-algorithm-assisted inversion approach: hydrogeological study for the determination of volumetric rock properties and matrix and fluid parameters in

- unsaturated formations. *Hydrogeol. J.*, 26(6), 1935–1946. <https://doi.org/10.1007/s10040-018-1749-7>
- [13] Szűcs, P., Szabó, N. P., Zubair, M., and S. Szalai, S. (2021). Innovative hydrogeophysical approaches as aids to assess Hungarian groundwater bodies. *Appl. Sci.*, 11(5), 2099. <https://doi.org/10.3390/app11052099>
- [14] Dobróka, M., Szabó, N. P., Cardarelli, E., and Vass, P. (2009). 2D inversion of borehole logging data for simultaneous determination of rock interfaces and petrophysical parameters. *Acta Geod. Geophys. Hungarica*, 44(4), 459–479. <https://doi.org/10.1556/AGeod.44.2009.4.7>
- [15] Szabó, N. P., Braun, B. A., Abdelrahman, M. M. G., and Dobróka, M. (2021). Improved well logs clustering algorithm for shale gas identification and formation evaluation. *Acta Geod. Geophys.*, 56(4), 711–729. <https://doi.org/10.1007/s40328-021-00358-0>
- [16] Levenberg, K. (1944). A method for the solution of certain non-linear problems in least squares. *Q. Appl. Math.*, 1(278), 536–538. <https://doi.org/10.1090/qam/10666>
- [17] Menke, W. and Eilon, Z. (2015). Relationship between data smoothing and the regularization of inverse problems. *Pure Appl. Geophys.*, 172(10), 2711–2726. <https://doi.org/10.1007/s00024-015-1059-0>
- [18] Abdelrahman, M. M. G., Szabó, N. P., and Dobróka, M. (2021). Petrophysical parameters estimation using Levenberg-Marquardt and Singular Value Decomposition inversion schemes. *Multidiszcip. tudományok*, 11(5), 24–38. <https://doi.org/10.35925/j.multi.2021.5.3>
- [19] Heriyanto, M. and Srigutomo, W. (2017). 1-D DC resistivity inversion using singular value decomposition and Levenberg-Marquardt's Inversion Schemes. *Journal of Physics: Conference Series*, 877(1), 012066. <https://doi.org/10.1088/1742-6596/877/1/012066>
- [20] Noroozi, M., Mohammadi, H., Efatinasab, E., Lashgari, A., Eslami, M., and Khan, B. (2022). Golden search optimization algorithm. *IEEE Access*, 10(April), 37515–37532. <https://doi.org/10.1109/ACCESS.2022.3162853>
- [21] Chakraborty, S. K. and G. Panda, G. (2016). Golden section search over hyper-rectangle : a direct search method. *Int. J. Math. Oper. Res.*, 8(3), 279–292. <https://doi.org/10.1504/IJMOR.2016.075517>
- [22] Tsai, C. H., Kolibal, J., and Li, M. (2010). The golden section search algorithm for finding a good shape parameter for meshless collocation methods. *Eng. Anal. Bound. Elem.*, 34(8), 738–746. <https://doi.org/10.1016/j.enganabound.2010.03.003>
- [23] Chang, K.-H. (2015). Chapter 3 - Design Optimization. In: K.-H. Chang (Ed.) *Design Theory and Methods Using CAD/CAE* (pp. 103–210). Boston: Academic Press. <https://doi.org/10.1016/B978-0-12-398512-5.00003-7>
- [24] Yazıcı, İ., Yaylacı, E. K., and Yalçın, F. (2021). Modified golden section search based MPPT algorithm for the WECS. *Eng. Sci. Technol. an Int. J.*, 24(5), 1123–1133. <https://doi.org/10.1016/j.jestch.2021.02.006>
- [25] Pejic, D. and Arsic, M. (2019). Minimization and maximization of functions: Golden-section search in one dimension. Springer International Publishing. [https://doi.org/10.1007/978-3-030-13803-5\\_3](https://doi.org/10.1007/978-3-030-13803-5_3)