An automatic inversion tool for geoelectrical resistivity data using supervised learning algorithm of adaptive neuro fuzzy inference system (ANFIS)

A. Stanley Raj¹ · D. Hudson Oliver² · Y. Srinivas³

Received: 8 December 2014 / Accepted: 11 May 2015 / Published online: 28 May 2015
© Springer International Publishing Switzerland 2015

Abstract Estimation of subsurface parameters of earth need an efficient and knowledge based algorithm to enthrall the real world truth clearly. Implementing the adaptive neuro fuzzy inference system (ANFIS) is worthwhile in this case of non-linear parametric approach. The ambiguous property of the conventional inversion technique results can be prevailing over by implementing the soft computing tool. The coalesce behavior of neural networks logics and fuzzy sets with certain rule based logics will concise the inversion technique to obtain the preferred result. In the present study, ANFIS algorithm was applied in direct inversion approach and the most prominent of this approach is supervised learning techniques adapted in the algorithm specially to enroll the concepts of inverting the geoelectrical data in a systematic way. The subsurface parameters of earth are mysteriously identified by sounding or direct bore techniques. Sounding method in geophysics plays the prominent role in understanding the subsurface features of earth. But major part of the sounding method relies on inversion techniques. Since the data obtained from the earth is non-linear and heterogeneous it is difficult to estimate the parameters more clearly. Thus apart from using any conventional inversion techniques which are mainly focusing on

initial model layer parameters. If the initial layer parameters are not given in the particular range, then the forward modeling solution tends dissimilarity of observed bore hole/litholog data. Thus direct inversion dominates in estimating the parameters with the help of soft computing inversion techniques. The proposed technique solves most of the subsurface problems since it depends on the trained knowledge. The supervised learning technique has been validated with Tuticorin and Kanyakumari coastal region data and found to be successful.

Keywords ANFIS · Vertical electrical sounding · Resistivity inversion · Layer model

Introduction

Inversion of electrical resistivity data can be done by various means viz., straightforward inversion scheme, Occam's or Zohdy's method etc., This paper projects a novel approach used for the inversion of electrical resistivity data with the efficient adaptive neuro fuzzy inference system (ANFIS). Many numbers of tools has been applied by earlier researchers (Flathe 1955; Kosinky and Kelly 1981; Singh et al. 2010; Srinivas et al. 2010, 2012a, b; Sri Niwas and Singhal 1981; Sugeno 1985 and Zadeh 1965). But ANFIS-soft computing tool provides a different approach in interpreting geoelectrical resistivity data. ANFIS network was developed by Jang (1993) and found that this technique has the capability of adaptive nature. Electrical resistivity data was obtained from the vertical electrical sounding (VES) method and the resulting interpretation thus supports the adaptive nature of the system. ANFIS algorithm has been applied here for interpretation.

✉ A. Stanley Raj
stanleyraj_84@yahoo.co.in

¹ Department of Physics, Vel Tech University, Avadi, Chennai 600 062, India

² Department of Physics, Senthamarai College of Arts and Science, Palkalai Nagar, Madurai 625 021, India

³ Centre for Geotechnology, Manonmaniam Sundaranar University, Tirunelveli, Tamil Nadu 627 012, India

Geophysical method

Direct current resistivity methods of geophysical exploration are in extensive use globally for aquifer mapping and estimation of aquifer parameters viz., resistivity and thickness (Kosinky and Kelly 1981; Mazac et al. 1985, Sri Niwas and Singhal 1981; Yadav and Abolfazli 1998). The physical basis of the resistivity method is based on the relative distribution of impressed current in the earth controlled by subsurface resistivity distribution. The electrical resistivity method still proves the most powerful technique and analysing the data is less tedious and is economical (Ako and Olorunfemi 1989; Batte et al. 2008, Ekine and Osobonye 1996; Zohdy et al. 1974). Due to the excessive computational requirement and the interpretation of VES data has been interpreted using different computational methods (Flathe 1955; Ghosh 1971; Mooney et al. 1966; Van Dam 1964).

The Geophysical method consisting of VES survey is used to know the variation of resistivity of the aquifer parameters (Rijo et al. 1977). Schlumberger electrode array is used to study the electrical resistivity distribution of the subsurface in order to understand the groundwater conditions such as resistivity, thickness and depth. Usually the depth of penetration is proportional to the separation between the electrodes and varying the electrode separation provides information about the stratification of the ground. The apparent resistivity value depends on the electrical conductivities of different rocks and minerals available in the subsurface. Thus electrical prospecting can be carried out to understand the subsurface earth. The data collected from the field has been interpreted using ANFIS algorithm. The ANFIS algorithm provides the necessary database needed for interpretation. Moreover, the best model of the trained database fits with the apparent resistivity curve. The corresponding layer model will be produced as an output with lowest root mean square error in particular number of epochs.

ANFIS—architecture and theoretical background

ANFIS is a combined fuzzy logic neural network systems. This kind of inference system has the adaptive nature to rely on the situation it trained. Thus it has a lot of advantages from learning to validating the output. Takagi–Sugeno fuzzy model is shown in Fig. 1.

As shown in Fig. 1, the ANFIS system consists of 5 layers. Each layer is symbolized by the box that is adaptive. Meanwhile, symbolized by the circle is fixed. Each output of each layer is symbolized by a sequence of nodes. Each output of each layer is symbolized by $O_{l,i}$

Fig. 1 Takagi–Sugeno fuzzy model for ANFIS 5 layer architecture

Fig. 2 ANFIS flowchart for geoelectrical data inversion

with i is a sequence of nodes and 1 is the sequence showing the lining. Here is an explanation for each layer, namely:

Layer 1

Layer 1 serves to raise the degree of membership with input variables of each node. Generally, in this layer membership function changes its form corresponding to the value of the parameters of fuzzy sets.

$$O_{1,i} = \mu_A(x), \quad i = 1, 2. \quad (1)$$

and

$$O_{1,i} = \mu_B(y), \quad i = 1, 2. \quad (2)$$

with x and y are the input for the i -th node

$$\mu_A(x) = 1 / [1 + (\det(x - c_i) / a_i)^2 b_i],$$

by $\{a_i, b_i$ and $c_i\}$ are the parameters of membership function or called as a parameter premise.

Layer 2

In this layer, every node is fixed and the output is the product of all incoming signals. Each node represents the firing strength of a rule.

$$O_{2,i} = w_i = \mu_A(x) \mu_B(y), \quad i = 1, 2. \quad (3)$$

Layer 3

Normalize the firing strength has been done in this layer.

$$O_{3,i} = \bar{w}_i = \frac{w_i}{w_1 + w_2}, \quad i = 2. \quad (4)$$

Layer 4

Layer 4 is an adaptive node, where normalized firing strength of layer 3 is the parameter set of this node. Parameters in this layer are referred to as consequent parameters. Calculating the output based on the parameters of the rule consequent $\{p_i, q_i$ and $r_i\}$.

$$O_{4,i} = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i). \quad (5)$$

Layer 5

Finally, the layer 5 computes the overall output by summing up all the incoming signals. Thus the constructed adaptive network is functionally equivalent to a Sugeno fuzzy model.

Counting the ANFIS output signal by summing all incoming signals will produce

Fig. 3 Graphical User Interface for ANFIS geoelectrical data inversion

$$\sum_i \bar{w}_{if_i} = \frac{\sum_i w_{if_i}}{\sum_i w_i} \quad (6)$$

ANFIS algorithm description and application

In the present algorithm, ANFIS training has been classified into two major parts Fig. 2a and b.

Primary training

The input data obtained from the user has been processed in the primary training (Fig. 2a). In the application for geoelectrical resistivity inversion, the vertical electrical sounding data (AB/2 and apparent resistivity) is fed to the primary training, smoothing the raw data obtained from the field is done by applying random weights to each input with certain controlling

Fig. 4 Profile showing the VES data locations

Fig. 5 ANFIS inverted 2d pseudo section and resistivity cross sections

Table 1 Longitudinal Resistivity variations in the profile taken for Kanyakumari district **Fig. 6** Geological cross section of the profile taken under study (after GSI 2005)

parameters. The number of synthetic data set which is proportional to the number of epochs is produces after the training. Initially, the multilayer models will be obtained for each synthetic data after training using the slope variation method.

Slope variation method

Slope variation refers to the basic method for obtaining the trend of the curve changing with the subsurface layers obtained from the field curve. Whenever the curve changes its

trend, it is considered as a line segment. Each line segment is treated as a separate layer and the corresponding subsurface layer parameters are obtained. Mathematically, each line segment is a linear curve and is represented by the straight line equation $y = mx + c$, where m represents the slope of the linear segment. Whenever the tangent of the slope varies, corresponding subsurface layer parameters can be obtained. Initially the multilayer model will be obtained since each and every point obtained from the field may not be linear. Figure 4 shows the representation of slope variation model.

Slope is normally described by the ratio of the “rise” divided by the “run” between two points on a line. The line may be practical—a set obtained by the AB/2 and apparent resistivity data values. Whenever the curve changes its behavior then slope changes.

The change for the y-axis can be depicted as $y_2 - y_1$ i.e., Δy . Similarly for the x-axis it is $x_2 - x_1$ i.e., Δx . Thus slope m of the line can be expressed as

$$m = \frac{y_2 - y_1}{x_2 - x_1}.$$

The concept of slope applies directly to grades or gradients in geography, geophysics and civil engineering. Through trigonometry, the grade m of a curve is related to its angle of incline θ by

$$m = \tan \theta.$$

Thus at the end of the primary training, we have more number of synthetic datasets with corresponding multilayer models.

(a)

(b)

(c)

Fig. 7 ANFIS generalized algorithm inverted model for VES 2 (S31) with litholog sections

Major class training

In the major class training (Fig. 2b), the synthetic data obtained from the primary training session is fed as an input to the ANFIS network. After this training, multilayer model is obtained for the corresponding synthetic data. Moreover, other parameter limitations follow the same rule as that of parameters used in the previous session.

In testing phase, the trained datasets were tested with the original field data and the output model will be linearly regressed with each multilayer and provide a compressed layer model as an output. The output model parameters viz., true resistivity and depth were plotted in the framed GUI and the user can save both the models individually. The specially designed algorithm provides many models in every iterations and the user can able to fix certain model.

ANFIS-RESINV tool

Based upon the algorithm and their properties presented in the previous section, we have developed a new Graphical User

Interface (GUI) on the MATLAB platform (Fig. 3). The program has a single graphical user interface (Fig. 3) that used to handle input files, to perform the computational operations, and to save the output results to a created file. The main GUI frame is subdivided into three components that represent (1) the input data features (1 & 2), (2) the computational results (results panel) and the graphical visualizations shown in 3, 4 & 5 of Fig. 3. The upper part of the GUI window is composed of the fuzzy membership functions and its architecture. The ANFIS RESINV program is realized in the Matlab environment (version R2010a and higher), using basic GUI properties.

Resistivity inversion and validation

This research work was validated with various field data sets and the resultant approach confirms the evident of efficient artificial intelligent algorithm for geoelectrical resistivity inversion techniques. This platform can perform the competent model based approach. The developed algorithm was subjected to different kind of datasets obtained from different geological background of the world as follows.

Fig. 8 ANFIS generalized algorithm inverted model for VES 3 (S32) with litholog sections

Coastal Data

- Kanyakumari
- Malaysia (Samsudin et al. 2008)

Hard Rock data

- South East Nigeria (Edet and Okereke 1997)

Glacier data

- Thulagi Glacier Lake, Nepal (Pant and Reynolds 2000)

Kanyakumari district is located in the southern tip of India. It lies between the latitude $77^{\circ}18'45''$ – $77^{\circ}35'15''$ E and $8^{\circ}4'$ – $8^{\circ}13'45''$ N longitude. The study area is underlain by the crystalline rocks like gneiss and charnockite of Archaean age. Along the coast the sands of recent origin are noticed. The geology map (Fig. 6) of the study area is obtained from Geological survey of India (GSI 2005). The peninsular gneisses occupy the largest area in the district. The general trend of the strike of this area is in the N–NW to S–SE direction. Garnetiferous sillimanite, graphite gneiss and garnet biotite gneiss are the two major group identified in Kanyakumari district. The charnockite group rocks are well exposed around Rajakkamangalam areas. Charnockite group mainly consists of charnockite,

(a)

(b)

(a)

(b)

Fig. 9 Applied ANFIS membership functions (gaussian) for inverting VES 2 data

Fig. 10 Applied ANFIS membership functions (gaussian) for inverting VES 3 data

pyroxene granulite and their associated migmatites. Charnokites are also exposed within the gneiss as bands and lenses. Near Kanyakumari calcareous limeshell of sub recent origin is noticed. The general sand types seen along the coast are bay deposits or lateral deposits of sand, zircon, rutile, illemanite and garnet. The straight west coast line continuing without any break is itself suggestive of faulted one and the faulting would have taken place during the Pliocene period. The pediments and the structural hills are run off zones and hence have poor potential regions. The valleys have a good infiltration—recharge zone has a medium groundwater potential zones. The coastal plains are characterized by beaches and sand dunes comprising of medium to fine sandy windblown particles, which is also a good groundwater potential zone. For agricultural developments almost the entire shallow aquifer zone is tapped in the study area. The groundwater occurs in almost all the geological formations like crystalline rocks, sedimentary formations and quaternary alluvium and beach sands. The groundwater occurrence in hard rock region is limited to the weathered mantle of thickness 10–35 m below ground level. The weathered thickness in hard rock regions is discontinuous both in space and depth. Hence the groundwater potentiality is influenced by the intensity of weathering. In the sedimentary formations having alluvial deposits the water table is very shallow which is up to a maximum depth of 10 m (PWD 2005).

Figure 4 shows the profile of VES data locations. Figure 5 shows the ANFIS inverted 2d pseudo section and resistivity cross sections. Longitudinal resistivity variations was calculated using the formula

$$\text{Longitudinal resistivity } \rho = H/S \quad (7)$$

where H is the total thickness and $S = 1/R = \sum h_i/\rho_i$ Longitudinal unit conductance.

Table 1 shows the variations of longitudinal resistivity along the profile.

Figure 6 shows the geological cross section of the profile taken under study. VES 2 (S31) and VES 3 (S32) are inverted using the ANFIS generalized algorithm and was validated with litholog sections available near the site (Figs. 7, 8). Figures 9 and 10 shows the membership functions and its ANFIS architecture for VES 2 and VES 3 respectively.

Figure 11 shows the ANFIS inverted model for Malaysian data with respective validation of litholog section as shown in Fig. 12. The groundwater resource in Malaysia is primarily based on alluvial and hard rock aquifers. Groundwater is currently a major source of water supply in north Kelantan, Malaysia. The first aquifer at the surface is about 4 m thick, with an apparent resistivity of more than 100 ohm m. The second aquifer, which is separated by the clay layer, is about 18–33 m below the

ground surface, with an apparent resistivity of less than 45 ohm m (Samsudin et al. 2008).

Figure 13 shows the ANFIS inverted model for South East Nigerian data. The validation based on the membership functions incorporated in the inversion steps of the algorithm. Figure 14 shows the ANFIS inverted results.

Figures 15 and 16 show the ANFIS inverted model for Thulagi glacier data. Thulagi glacier lake in the Gorkha district of western Nepal is surrounded by buried glacier ice and permafrost zones. So the investigation of electrical surveys yields very high resistivity values ($>20,000$ and >5000 ohm-m). Water saturated glacial sediments have values of electrical resistivity less than 3000 ohm-m. The materials that overlie the dead ice is thought to contain neoglacial and supra-glacial lacustrine deposits (Pant and Reynolds 2000). The interpretation results shows that the unsaturated moraine having high resistivities occupies the first layer ($4-8 \times 10^3$ ohm-m) and the second layer comprises of saturated moraine deposits ($2-4 \times 10^3$ ohm-m) and the third layer composed of dead ice having the resistivity range of 10^5 ohm-m. The validation based on the membership functions incorporated in the inversion steps of the algorithm. Figures 15 and 16 show the ANFIS inverted results.

Results and discussions

The system is said to be adaptable only when it have the capability to withstand in the situation where the problem is going to deviate from original track of the result. Though ANFIS tussles in bringing the appropriate result at the initial stepwise regression, it provides the exemplary favorable result while increasing the number of epochs. Increase in number of epochs will definitely brings out the correct result instead of minimizing it.

Earlier works on ANFIS was based on the local optima. This research work lays the platform for attempting the problem through ANFIS has global optima. The result proves that this tool can be used as a generalized optimization algorithm for solving any non-linear problem. This novel approaches enlighten and bring forth the generalization applied in the geoelectrical data inversion. Data obtained from different geological settings were processed and was validated successfully. This kind of approach will be very helpful in noisy data processing as well as non-linearity involved in data processing. Thus, this could be one of the best artificial intelligent techniques used for inverting geoelectrical data. Moreover, other algorithms will completely rely on the tuning parameters. But in this approach the tuning parameters were automatically fixed by the algorithm initially in the primary class training. So the best accomplishment of the result can be produced without any stay back or hangout in between the data processing.

Fig. 11 ANFIS inverted model for Malaysian data (*Data source* Samsudin et al. 2008)

(a)

(b)

(d)

(e)

(f)

Fig. 12 Validated ANFIS model with corresponding litholog section for Malaysian data

Fig. 13 ANFIS inverted model for South East Nigerian data (Data source Edet and Okereke 1997)

Fig. 14 Validated ANFIS model with corresponding litholog section for SE Nigerian data

The preprocessing of data was done in the primary class training itself. When the data is entering the major class training, it is free from any noises or disturbances. So the data

with appropriate type of curve is interpreted using the ANFIS algorithm with the help of slope-variation model. The solution obtained depends on the number of epochs and error percent.

Fig. 15 ANFIS inverted model for Thulagi glacier data (*Data source* Pant and Reynolds 2000)

(a)

(b)

(c)

(d)

(e)

(f)

Fig. 16 Validated ANFIS model with corresponding litholog section for Thulagi glacier data (*Data source* Pant and Reynolds 2000)

Advantages of using the algorithm

1. All types of the curve (A, K, H, Q, AK, AH, AQ, KH, KQ, KA, AHK, HQK, KHQ., etc.) can be interpreted using the ANFIS algorithm. Especially the ANFIS algorithm was designed to perform the solution for all kind of geological dataset.
2. If noisy data arises, this ANFIS algorithm will adapt itself to train the non-linear data by adjusting the initialization of parameters necessary to fix the result.
3. If any data is missing, the algorithm interpolates the data based on the nearby data values thus producing the exact curve for interpretation.
4. Above all, this is the direct inversion technique, no need to worry about the initial layer parameters. ANFIS itself will handle the entire case of the problem and will produce more successful layer model parameter.

References

- Ako BD, Olorunfemi MO (1989) Geoelectric survey for groundwater in the newer basalts of Vom plateau state. *Nig J Min Geol* 25:247–450
- Batte AG, Muwanga A, Sigrist PW, Owor M (2008) Vertical electrical sounding as an exploration technique to improve on the certainty of groundwater yield in the fractured crystalline basement aquifers of eastern Uganda. *Hydrogeol J* 16:1683–1693
- Edet AE, Okereke CS (1997) Assessment of hydrogeological conditions in basement aquifers of the Precambrian Oban massif, southeastern Nigeria. *J Appl Geophys* 36:195–204
- Ekine AS, Osobonye GT (1996) Surface geoelectric sounding for the determination of aquifer characteristics in parts of Bonny local government area of river state. *Nig J Phys* 85:93–97
- Flathe H (1955) A practical method of calculating geoelectrical model graphs for horizontally stratified media. *Geophys Prospect* 3: 268–294
- Ghosh DP (1971) Inverse filter coefficients for the computation of the apparent resistivity standard curves for horizontally stratified earth. *Geophys Prospect* 19:769–775
- GSI (2005) Geology and mineral map of Kanyakumari district, compiled by Sundaram R, Ranganathan M and Vasudevan D and published by Geological Survey of India
- Jang JSR (1993) Adaptive-network based fuzzy inference system. *IEEE Trans. Syst Man Cybern* 23:665–685
- Kosinky WK, Kelly WE (1981) Geoelectrical sounding for predicting aquifer properties. *Groundwater* 19:163–171
- Mazac O, Kelly WE, Landa I (1985) A hydrophysical model for relation between electrical and hydraulic properties of aquifer. *J Hydrol* 79:1–19
- Mooney HM, Orellana E, Pickett H, Tornheim L (1966) A resistivity computation method for layered earth model. *Geophysics* 31:192–203
- Niwas Sri, Singhal DC (1981) Estimation of aquifer transmissivity from Dar-Zarrouk parameters in porous media. *J Hydrol* 50:393–399
- Pant SR, Reynolds JM (2000) Application of electrical imaging techniques for the investigation of natural dams: an example from the Thulagi glacier lake, Nepal. *J Nepal Geol Soc* 22:211–218
- PWD (2005) Groundwater perspectives: a profile of Kanyakumari district. Tamil Nadu Public Works Department, Tamil Nadu
- Rijo L, Pelton W, Feitosa E, Wars S (1977) Interpretation of apparent resistivity data from Apodi Valley, Rio Grande de Norte, Brasil. *Geophysics* 42:811–822
- Samsudin AR, Haryono A, Hamzah U, Rafek AG (2008) Salinity mapping of coastal groundwater aquifers using hydrogeochemical and geophysical methods: a case study from north Kelantan, Malaysia. *Environ Geol* 55:1737–1743
- Singh UK, Singh DK, Singh H (2010) Application of NeuroFuzzy pattern recognition method in borehole geophysics. *Acta Geodaetica et Geophysica Hungarica* 45:417–425
- Srinivas Y, Stanley Raj A, Hudson Oliver D, Muthuraj D, Chandrasekar N (2010) An application of artificial neural network for the interpretation of three layer electrical resistivity data using feed forward back propagation algorithm. *Curr Dev Artif Intell* 1:1–11
- Srinivas Y, Stanley Raj A, Hudson Oliver D, Muthuraj D, Chandrasekar N (2012a) Estimation of sub surface strata of earth using adaptive neuro fuzzy inference system (ANFIS). *Acta Geodaetica et Geophysica Hungarica* 47(1):78–89
- Srinivas Y, Stanley Raj A, Hudson Oliver D, Muthuraj D, Chandrasekar N (2012b) A robust behaviour of feed forward back propagation algorithm of artificial neural networks in the application of vertical electrical sounding data inversion. *Geosci Front* 3(5):729–736
- Sugeno M (1985) Industrial applications and fuzzy control. Elsevier, New York
- Van Dam JC (1964) A simple method for the calculation of standard graphs to be used in geoelectrical prospecting, Ph. D. Thesis, Delft Technological University, The Netherlands.
- Yadav GS, Abolfazli H (1998) Geoelectrical soundings and their relationships to hydraulic parameters in semi arid regions of Jalore, North West India. *J Appl Geophys* 39:35–51
- Zadeh L (1965) Fuzzy sets. *Inf Control* 8:338–353
- Zohdy AAR, Eaton GP, Mabey DR (1974) Application of surface geophysics to groundwater investigations. USGS techniques of water resource investigations. Book 2. Chap. DI, pp 116