A robust behavior of Feed Forward Back propagation algorithm of Artificial Neural Networks in the application of vertical electrical sounding data inversion

Y. Srinivas\textsuperscript{a}, A. Stanley Raj\textsuperscript{a,}\textsuperscript{*}, D. Hudson Oliver\textsuperscript{a}, D. Muthuraj\textsuperscript{b}, N. Chandrasekar\textsuperscript{a}

\textsuperscript{a}Centre for Geotechnology, Manonmaniam Sundaranar University, Tirunelveli, Tamil Nadu-627012, India
\textsuperscript{b}M.D.T. Hindu college, Tirunelveli, Tamil Nadu, India

Received 23 June 2011; received in revised form 20 January 2012; accepted 8 February 2012
Available online 22 February 2012

KEYWORDS
Artificial neural networks (ANN); Resistivity inversion; coastal aquifer parameters; Layer model

Abstract The applications of intelligent techniques have increased exponentially in recent days to study most of the non-linear parameters. In particular, the behavior of earth resembles the non-linearity applications. An efficient tool is needed for the interpretation of geophysical parameters to study the subsurface of the earth. Artificial Neural Networks (ANN) perform certain tasks if the structure of the network is modified accordingly for the purpose it has been used. The three most robust networks were taken and comparatively analyzed for their performance to choose the appropriate network. The single-layer feed-forward neural network with the back propagation algorithm is chosen as one of the well-suited networks after comparing the results. Initially, certain synthetic data sets of all three-layer curves have been taken for training the network, and the network is validated by the field datasets collected from Tuticorin Coastal Region (78°7’30”E and 8°48’45”N), Tamil Nadu, India. The interpretation has been done successfully using the corresponding learning algorithm in the present study. With proper training of back propagation networks, it tends to give the resistivity and thickness of the subsurface layer model
1. Introduction

Artificial Neural Networks (ANN) have strong fundamentals framed based on the biological neural network and perform computations similar to the natural neural networks. Earth science’s non-linearity can be studied using this efficient tool. Network capabilities can be very well proven to be good for noise immune system and fault tolerant cases (Sreekanth et al., 2009). The ANN inversion has been used for many years to interpret electrical resistivity data (e.g. Dey and Morrison, 1979; Smith and Vozoff, 1984; Tripp et al., 1984; Constable et al., 1987; Uchida and Murakami, 1990; Poulton et al., 1992; Griffith and Barker, 1993; Lake and Barker, 1996; Calderon-Macias et al., 2000; Vander Baan and Jutten, 2000; El-Qady and Ushijima, 2001; Stephen et al., 2004; Neyamadpour et al., 2010). ANNs are also being increasingly applied in the field of engineering geophysics (Brown and Poulton, 1996).  

Aquifer parameters such as thickness and resistivity can be studied by direct current resistivity methods mainly used in the field of geophysical exploration (Kosinsky and Kelly, 1981; Sri Niwas and Singhal, 1981; Mazac et al., 1985; Yadav and Abolfazli, 1998). Vertical Electrical Sounding (VES) data, can be interpreted using curve matching technique, is well useful for estimating the subsurface geology (Flathe, 1955; Mooney et al., 1966; Ghosh, 1971). Due to the advancement in computational techniques, the interpretation can be done with soft computing tools, one such tool is ANNs. The raw data collected from the field should be smoothened for interpretation because of many noises and deviations. Here, electrical resistivity data of the coastal region near Tuticorin (Longitude from 78°3′30″E to 78°11′45″E and Latitude from 8°40′N to 8°55′N) Tamil Nadu, India, has been taken to interpret the layer model of the subsurface earth using ANN.

2. Geophysical method

The geophysical method consisting of VES survey has been carried out in the Tuticorin area from the coast to inland in order to know the variation of resistivity of the aquifer and to find the depth of the ground water level (Rijo et al., 1977). A total number of seventeen VES were carried out in the study area. The Schlumberger electrode array is used throughout the survey in order to study groundwater conditions (Fig. 1). The field procedure involves the following parameters. The potential electrodes (M and N) remain fixed, and the current electrodes (A and B) are expanded symmetrically from the center of the spread. Maximum half current electrode (AB/2) separation used in these surveys is 100 m. 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 (Telford et al., 2010). The VES locations are shown as S1—S17 in Fig. 2.

3. Geology and hydrogeology of the study area

The prominent geomorphic units identified in the district are (1) fluvial, (2) marine, (3) fluvio-marine, (4) aeolian and (5) erosional landforms depending on the environment of formation. The important aquifer systems in the district are constituted by (i) unconsolidated and semi-consolidated formations and (ii) weathered and fractured crystalline rocks. The maximum thickness of alluvium is 45 m, whereas the average thickness is about 25 m. The productive zones are encountered in the depth range of 29.5—62 m. The water-bearing properties of crystalline formations which lack primary porosity depend on the extent of development of secondary inter-granular porosity. The occurrence and movement of ground water in these rocks are under unconfined conditions in the joints and fissures and dependent on the nature and extent of pores and interconnection of the fracture zones. The aquifer parameters of the wells show wide variation, both in crystalline and sedimentary formations (Balachandran, 2009). Groundwater occurs in the weathered and fractured zones of gneisses and sandy aquifers in sedimentary and alluvial formations. The shallow ground water has been recorded in bore wells located in the industrial complex and in open wells near the coast. Deep water table conditions (<14 m) observed in wells at localized pocket sand are mainly due to the exploitation of ground-water for domestic/irrigation purposes. The well inventory data indicates that the seasonal rainfall causes 0.49—5.95 m water level rise with an average value of 1.5 m in the area (Mondal et al., 2009).

4. Massaging resistivity data for Artificial Neural Network

Many algorithms can be used to interpret the resistivity data. However, the network performance of each algorithm should be
validated. It is therefore essential to check which network performance is best suited for our application before going into the interpretation part.

4.1. Best network performance

The Radial basis networks, Generalized Regression Networks and Feed Forward Back propagation (FFBP) algorithms are being chosen for testing the network performance.

The function newrbe (new radial basis function used in MATLAB) takes matrices of input vectors $P$ and target vectors $T$. The constant $spread$ is the radial basis function, and returns a network with weights and biases such that the outputs are exactly $T$ when the inputs are $P$. The function newrbe creates as many neurons as there are input vectors in $P$, and sets the first-layer weights to $P'$. Thus, we have a layer of neurons in which each neuron acts as a detector for a different input vector. If there are $Q$ input vectors, then there will be $Q$ neurons. The function newrbe then creates a network with nearly zero errors on training vectors.

A Generalized Regression Neural Network (GRNN) is often used for function approximation. GRNN is a universal approximator used to smooth functions, so it should be able to solve any smooth function-approximation problem given enough data. The main drawback of GRNN is that it suffers badly from the curse of dimensionality. GRNN cannot ignore irrelevant inputs without major modifications to the basic algorithm. The function will be called as follows.

Net = newgrnn ($P$, $T$, NEURONS).

Newgrnn function creates Generalized Regression Neural Network with $P$ inputs and $T$ targets having a particular number of neurons.

FFBP method used to compute the network parameters by an iterative Levenberg-Marquardt algorithm (LMA): from an initial set of parameters, LMA computes step by step a final parameter set approximating criterion's minima. Because the FFBP has multiple local minima and there is no proof for the existence of a unique global minimum, the output parameters depend on the choice of the input parameters. The back propagation learning technique will be the most efficient technique for obtaining good results.

FFBP is defined as follows:

Net = newff ($P$, $T$, NEURONS)

Here in our program the initial parameters were fixed as mentioned below.

Net.trainparam.show = 2;
Net.trainparam.lr = 0.01;

Figure 2  Vertical Electrical Sounding locations in the study area.
net.trainparam.mc = 0.9;
net.trainparam.epochs = 100;
net.trainparam.goal = 1 × 10⁻⁶.

For choosing the best network, 10 synthetic sample datasets were taken for testing with the networks and the results are shown in (Fig. 3). Performance of each training algorithm was checked, and it is represented in Table 1. Thus, the best algorithm trainlm is used for testing the field data. The feed forward layer network is constructed to analyze the observed field data and while analyzing it smoothens the data by updating weights and biases of subsequent ANN network layers. Linear transfer function (purelin) is used to make the nonlinear data to be linear by adjusting certain weights used as the parameters in the program. Network training function trainlm (Levenberg-Marquardt optimization) is used to updates weight and bias values. The specified transfer function will affect the changes in training, and the optimization of the observed data weight and bias values. The specified transfer function will affect the changes in training, and the optimization of the observed data will then be adapted to the network and fits with the activation function of the network. The network completes the training after the performance goal has been met with the particular number of epochs. We are able to see the performance of the network using perform function.

4.2. Robustness of feed forward Neural Networks

Feed forward neural network is a non-parametric estimation of statistical models for extracting nonlinear relations of the input data. The training algorithm involves two phases (Rumelhart et al., 1986).

I. Forward Phase: The free parameters of the networks are fixed and the input signal is propagated through the network during this phase. It finishes with the computation of an error signal.

\[ e_i = d_i - y_i \]  \hspace{1cm} (1)

Where \( d_i \) is the desired response and \( y_i \) is the actual output produced by the network in response to the input \( x_i \).

II. Backward Phase: During this second phase, the error signal \( e_i \) is propagated through the network in the backward direction, hence the name of the algorithm. It is during this phase that the adjustments applied to the free parameters of the network to minimize the error \( e_i \) in a statistical sense. The back propagation learning algorithm is simple to implement and computationally efficient.

The set of training examples is split into two parts:

- Estimation subset used for the training of the model.
- Validation subset used for evaluating the model performance.

In general, the network is finally tuned using the entire set of training examples and then tested on test data (Haykin, 1999). In the present study, we have adopted Feed Forward Back propagation algorithm for the interpretation of the subsurface layer model. The flow chart in Fig. 4 shows the brief description of the program after checking the best network performance.

5. Data preprocessing using Artificial Neural Network

The network is trained with 18 synthetic electrical resistivity sounding data sets in the present study. Training large amount of data will provide us good results of the problem and make the hidden units to train the data more effectively.

The back propagation algorithm updates neuronal activations in the network for the input layer as

\[ \delta(x^k_i) = x^k_i, \ i = 1, \ldots, n \]  \hspace{1cm} (2)

\[ \delta(x^0_k) = x^0_k = 1 \]  \hspace{1cm} (3)

Where \( x^k_i \) is the \( k \)th component of the input vector presented in the network, and \( \delta(x^0_k) \) is the input layer bias neuron signal that is independent of iteration index.

And for the hidden layer the network activation will be

\[ Z^h_k = \sum_{i=0}^{n} w^h_{ik} \delta(x^0_i) = \sum_{i=0}^{n} w^h_{ik} x^k_i, \ h = 1, \ldots, q \]  \hspace{1cm} (4)

\[ \delta(Z^h_k) = 1 / \left(1 + e^{-Z^h_k}\right), \ h = 1, \ldots, q \]  \hspace{1cm} (5)

\[ \delta(Z^0_k) = 1 \]  \hspace{1cm} (6)

Where \( w^h_{ik} \) are the biases of the hidden neurons, and \( \delta(Z^0_k) \) is the hidden layer bias neuron signal which is independent of the iteration index.

The output layer neuronal activations for the back propagation will be

\[ y^j_i = \sum_{h=0}^{n} w^j_{ih} \delta(Z^h_k), \ j = 1, \ldots, q \]  \hspace{1cm} (7)

\[ \delta(y^j_i) = 1 / \left(1 + e^{-y^j_i}\right), \ j = 1, \ldots, q \]  \hspace{1cm} (8)

Where \( w^j_{ih} \) are the biases of the output neurons. The output functions are differentiable, which is essential for the back propagation of errors (Yegnanarayana, 2005).

The learning rate \( \eta \) in the back propagation algorithm has to be kept small in order to maintain a smooth trajectory in weight space, because large learning rates can lead to oscillations during learning.

\[ \Delta w^j_{ij} = \eta \sum_{I=t}^{k} \alpha^{k-j} \delta^j \delta^i S^j = - \sum_{I=t}^{k} \alpha^{k-j} \frac{\partial E_t}{\partial w^j_{ij}} \]  \hspace{1cm} (9)
The above equation generalizes the weight change at the $k^{th}$ iteration in terms of the weight gradient at each of the previous iterations (Satish Kumar, 2007).

The actual output for a given input training pattern is determined by computing the outputs of units for each hidden layer in the forward pass of the input data (Yegnanarayana, 2005).

6. Interpretation and representation of aquifer parameters

After getting the field data as an input, the program checks the curve type and proceeds for further learning. The learning rate, momentum, bias, epochs will be chosen according to the data. Weights will be updated while crossing through the appropriate function used in the Feed Forward Back propagation algorithm. The apparent resistivity data given as input will be processed and the program analyses the synthetic database then represents the corresponding aquifer parameters. The output representation explains the curve type and layer parameters along with error percentage. Thus, the inversion of the field data by ANN reveals the nature of the water quality. The resistivity of the aquifer in some locations is as low as 0.2083 $\Omega \cdot m$, which indicate highly saline water, and unsuitable for drinking purpose. The interpreted subsurface layer parameters viz., resistivity and thickness using ANN with error percentage is presented in the Table 2.

7. Results and discussion

The training stops whenever the goal is achieved in a particular number of epochs. The regression plot of the synthetic trained data will be well fitted to achieve the target. The test using trained data indicates that the ANN system can converge to the target rapidly and accurately. 1D resistivity inversion procedure using the ANN system was carried out because the procedure works well for the observed data. The synthetic ANN trained data along with layer model is plotted and shown in Fig. 5. The interpreted layer model of the field data is shown in Fig. 6. For interpreting the layer model the network calls the memory associated with the layer parameters. If the model parameters’ matches with the synthetic trained data stored in the network memory, then it produces the corresponding model parameters with respective error percentage. The well-trained network will increase the performance level of the output parameters. The inversion thus represents the layer parameters of subsurface geology viz., thickness and resistivity. The results of interpretation of the near-subsurface features by this ANN technique are satisfactory and are more efficient, and the error has been checked with the conventional method. VES 2 apparent resistivity data have been taken into account for representing the graphs. Fig. 5 shows the ANN training for synthetic data and its corresponding layer model. Fig. 6 shows the ANN training for observed field data and its corresponding layer model. The VES interpreted results and error percentage in comparison with the conventional method is shown in Table 2. The spatial analysis of groundwater level for the 17 soundings has been presented in Fig. 7. The virtual projection of groundwater table depth for the 17 VES locations is shown in Fig. 8, and the circle marker represents the VES locations.

### Table 1: Performance of different Artificial Neural Networks training algorithm.

<table>
<thead>
<tr>
<th>Training algorithms</th>
<th>Mean square error</th>
<th>Performance accuracy</th>
<th>Computational time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>traingd</td>
<td>0.6759</td>
<td>99.3241</td>
<td>721.974680</td>
</tr>
<tr>
<td>traingdm</td>
<td>2.7143</td>
<td>97.2857</td>
<td>117.627901</td>
</tr>
<tr>
<td>trainlm</td>
<td>$1.6577 \times 10^{-9}$</td>
<td>~ 100</td>
<td>3.758382</td>
</tr>
<tr>
<td>trainoss</td>
<td>3.1613</td>
<td>96.8387</td>
<td>58.392031</td>
</tr>
<tr>
<td>trainbfg</td>
<td>7.1943</td>
<td>92.8057</td>
<td>39.021104</td>
</tr>
<tr>
<td>trainscg</td>
<td>1.9638</td>
<td>98.0362</td>
<td>43.834609</td>
</tr>
<tr>
<td>traincgb</td>
<td>0.1717</td>
<td>99.8283</td>
<td>23.653276</td>
</tr>
<tr>
<td>traincgp</td>
<td>1.1077</td>
<td>98.8923</td>
<td>22.346944</td>
</tr>
<tr>
<td>traincgf</td>
<td>7.7202</td>
<td>92.2798</td>
<td>36.042076</td>
</tr>
<tr>
<td>trainrp</td>
<td>6.2140</td>
<td>93.7860</td>
<td>60.860422</td>
</tr>
<tr>
<td>traindx</td>
<td>2.6063</td>
<td>97.3937</td>
<td>30.353578</td>
</tr>
<tr>
<td>trainda</td>
<td>2.1812</td>
<td>97.8188</td>
<td>29.966193</td>
</tr>
</tbody>
</table>

Figure 4 Flow chart showing the description of the ANN-FFBP program.
Table 2  Interpreted layer model of field data using ANN along with Error percentage.

<table>
<thead>
<tr>
<th>Field data</th>
<th>Curve types</th>
<th>Model parameters ($\rho = \text{True resistivity of subsurface layers in } \Omega \cdot \text{m}, T = \text{Thickness of layers in } \text{m}$)</th>
<th>Error percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>VES 1</td>
<td>H type</td>
<td>$\rho_1 = 34.6, \rho_2 = 5.08, \rho_3 = 31.9$  $T_1 = 1.3, T_2 = 8$</td>
<td>0.0667</td>
</tr>
<tr>
<td>VES 2</td>
<td>H type</td>
<td>$\rho_1 = 34.61, \rho_2 = 0.385, \rho_3 = 4.56$  $T_1 = 1.2, T_2 = 17$</td>
<td>0.0110</td>
</tr>
<tr>
<td>VES 3</td>
<td>H type</td>
<td>$\rho_1 = 36.45, \rho_2 = 14.65, \rho_3 = 33.7$  $T_1 = 3.0, T_2 = 8$</td>
<td>1.8554</td>
</tr>
<tr>
<td>VES 4</td>
<td>H type</td>
<td>$\rho_1 = 163.70, \rho_2 = 5.726, \rho_3 = 32.6$  $T_1 = 1.5, T_2 = 8.6$</td>
<td>0.7060</td>
</tr>
<tr>
<td>VES 5</td>
<td>H type</td>
<td>$\rho_1 = 23.35, \rho_2 = 7.27, \rho_3 = 494.25$  $T_1 = 4.0, T_2 = 10.2$</td>
<td>2.2567</td>
</tr>
<tr>
<td>VES 6</td>
<td>H type</td>
<td>$\rho_1 = 34.6, \rho_2 = 0.231, \rho_3 = 492.0$  $T_1 = 1.3, T_2 = 1.6$</td>
<td>0.0614</td>
</tr>
<tr>
<td>VES 7</td>
<td>H type</td>
<td>$\rho_1 = 36.26, \rho_2 = 6.68, \rho_3 = 493.6$  $T_1 = 2.9, T_2 = 9.6$</td>
<td>1.6678</td>
</tr>
<tr>
<td>VES 8</td>
<td>H type</td>
<td>$\rho_1 = 34.91, \rho_2 = 5.33, \rho_3 = 492.31$  $T_1 = 1.5, T_2 = 8.2$</td>
<td>0.3173</td>
</tr>
<tr>
<td>VES 9</td>
<td>A type</td>
<td>$\rho_1 = 0.28, \rho_2 = 0.97, \rho_3 = 428$  $T_1 = 0.9, T_2 = 3.4$</td>
<td>$\sim 0 \left(3.33 \times 10^{-5}\right)$</td>
</tr>
<tr>
<td>VES 10</td>
<td>H type</td>
<td>$\rho_1 = 165.42, \rho_2 = 15.22, \rho_3 = 494.42$  $T_1 = 3.2, T_2 = 8.6$</td>
<td>2.4216</td>
</tr>
<tr>
<td>VES 11</td>
<td>H type</td>
<td>$\rho_1 = 27, \rho_2 = 16.8, \rho_3 = 147$  $T_1 = 2.9, T_2 = 1.5$</td>
<td>$\sim 0 \left(6.6317 \times 10^{-15}\right)$</td>
</tr>
<tr>
<td>VES 12</td>
<td>H type</td>
<td>$\rho_1 = 22.0027, \rho_2 = 5.9, \rho_3 = 492.9$  $T_1 = 2.6, T_2 = 8.8$</td>
<td>0.9027</td>
</tr>
<tr>
<td>VES 13</td>
<td>A type</td>
<td>$\rho_1 = 17.1, \rho_2 = 81.4, \rho_3 = 273.20$  $T_1 = 4.2, T_2 = 20.1$</td>
<td>1.3043</td>
</tr>
<tr>
<td>VES 14</td>
<td>H type</td>
<td>$\rho_1 = 35.685, \rho_2 = 6.10, \rho_3 = 32.98$  $T_1 = 2.3, T_2 = 9.0$</td>
<td>1.0854</td>
</tr>
<tr>
<td>VES 15</td>
<td>H type</td>
<td>$\rho_1 = 36.5, \rho_2 = 6.97, \rho_3 = 493.95$  $T_1 = 3.1, T_2 = 9.9$</td>
<td>1.9546</td>
</tr>
<tr>
<td>VES 16</td>
<td>H type</td>
<td>$\rho_1 = 22.01, \rho_2 = 5.93, \rho_3 = 492.91$  $T_1 = 2.6, T_2 = 8.8$</td>
<td>0.9154</td>
</tr>
<tr>
<td>VES 17</td>
<td>H type</td>
<td>$\rho_1 = 37.34, \rho_2 = 7.76, \rho_3 = 494.74$  $T_1 = 3.9, T_2 = 10.9$</td>
<td>2.7400</td>
</tr>
</tbody>
</table>

Figure 5  Trained synthetic data layer model plot.

Figure 6  Interpreted field data layer model plot.
8. Conclusion

The present study reveals that FFBP algorithm of Artificial Neural network can be better employed for the interpretation of resistivity data in terms of resistivity and thickness. The results indicated that the concept also useful for studying the groundwater quality. The significance of the study shows that the ground water table depth in the Tuticorin study site seems to be unstable. It may be due to the variations in the rainfall, and natural discharge of the ground water. The reason for having the lowest resistivity values such as 0.385 for VES 2, 0.231 for VES 6, 0.28 for VES 9 is may be due to contamination.

Figure 7  Vertical Electrical Sounding locations and spatial analysis of groundwater occurrence.

Figure 8  Discrete modeling of groundwater table depth.
Acknowledgements

The field data was collected under the project seed money and the first author is grateful to the Vice Chancellor, Manonmaniam Sundaranar University, Tirunelveli for providing the financial support under this scheme. The corresponding author expresses sincere thanks to Raghu Ram Associates, Porur, Chennai, Tamilnadu, for providing the computational facility needed for this research work. Furthermore, the corresponding author acknowledged Dr. Franz Kohlbeck, Institute of Geodesy and Geophysics Vienna University of Technology, Gusshausstrasse, Vienna, Austria for his help rendered by him during computations. Moreover, the authors earnestly thank the anonymous reviewers for their valuable review comments to improve the manuscript to its present form.

References

Balachandran, A., Scientist D, Central Ground Water Board, Govt. of India, Ministry of water resources, South eastern coastal region, Chennai District ground water brochure technical report series, Thoothukudi district, Tamil Nadu, March 2009, 21p.


