

Genetic Algorithm Coupled with Neural Networks to Guesstimate the Subsurface Features of the Earth

A. Stanley Raj^{1,*}, Y. Srinivas², R. Damodharan², B. Chendhoor², M. Sanjay Vimal²

¹Department of Physics, Loyola College, Chennai

²Centre for GeoTechnology, Manonmaniam Sundaranar University, Tirunelveli, Tamil Nadu-627012, India.

Abstract

Electrical resistivity method is often used to estimate the subsurface structure of the earth. Many inversion algorithms are available to estimate the subsurface features. However, predicting the exact parameter in the non-linear subsurface of the earth is difficult because of its complex composition. Soft computing tools can approximate the subsurface parameters more clearly. Each soft computing tool has certain advantages and disadvantages. A hybrid formation of algorithms will make the decision more appropriate than depending on a single tool. Here in our study the data obtained through Vertical Electrical Sounding has been used to determine the sub surface characteristics of earth viz., true resistivity and thickness. Artificial Neural Networks (ANN) requires certain optimizing procedures. Here in this paper, Genetic Algorithm (GA) is applied to optimize Artificial Neural Networks (ANN). This coupled approach is tested with the field data. Error percentage of algorithm nearly mimics the behavior of earth and is verified. The best performance result shows that this technique can be implemented to estimate the non-linear characteristics of the earth more noticeably.

Corresponding author: A. Stanley Raj, Department of Physics, Loyola College, Chennai, Phone: +919940120058; Fax No: +91462 2334363

Keywords: Genetic Algorithm, Neural Networks, Optimization, Soft computing, Geo-resistivity

Received: Jun 22, 2020

Accepted: Jul 23, 2020

Published: Jul 24, 2020

Editor: Masoud Vakili, Iran University Science and Technology, Iran.

Introduction

In recent times, people moved towards advanced techniques to model the complex behavior system associated with nonlinearity, high-order dynamics, time-varying behavior, and imprecise measurements.

One of such non-linear characteristic behavior that we know is the field of Earth Sciences. Because of the complex structural behavior of rocks and soil formations, it is difficult to fix each parameter in a certain range. So the efficient soft computing tool is needed to estimate the subsurface parameters of the earth that are pertinent to the real world.

Geophysical inverse problems almost are nonlinear in nature and possess irregular objective function which is of multimodal type [1]. The traditional approaches over such geophysical problems can lead to failure due to the initial model that are closer to some local minimum which will attract it. As these problems possess the objective function that shows several local maxima and minima, they can be optimized through sophisticated approach by global optimization techniques. Evolutionary Algorithms which are generally known as global optimization technique, are appropriate for providing solutions to the irregular nonlinear problems of geophysics. [1,2,3] have successfully implemented this technique to solve many types of problems. The possible nature of this technique paves ways to locate the global optimum with the actual presence of many local optima, they are having major advantage to evolutionary computation is that it can easily get parallelized to run on multiple modes of processors simultaneously. In a single generation as many as several hundred objective function evaluations can be made, but each of them has been independently as a result of any other, so the process of parallelizing of the algorithm can theoretically increase efficiency by the means of orders of magnitude.

In recent years, Genetic Algorithm plays a significant role in optimizing various geophysical problems which includes Seismic waveform inversion, Marine Electromagnetic data inversion, Inversion of data obtained through surface waves and that have been proved by many researchers. [4,5,6,7,8,9]

Here in our study Artificial Neural Networks (ANN) and Genetic Algorithm (GA) were combined to perform a best unification algorithm for predicting the subsurface features of the earth. Neural networks can approximate the result based on the training. It can exhibit mapping capabilities easily between the input and output patterns. Here synthetic apparent resistivity data were trained with ANN algorithm and tested with the field data. As ANN can learn by examples, the training dataset was developed synthetically and tested. The layer model provides the information about the true resistivity and the thickness of the subsurface layer with error percentage.

Genetic algorithm (GA) is a search algorithm based on the mechanics of natural selection and natural genetics. The central theme of research on genetic algorithms has been robustness, the balance between efficiency and efficacy necessary for survival in many different environments. It is a computationally powerful tool that seeks to reproduce mathematically the mechanics of natural selection and natural genetics, according to the biological processes of survival and adaptation [10]. GAs excel in solving complex nonlinear problems by making use of enormous parallel information to search more efficiently. As they are highly capable of finding the global optimum with the presence of local optima, discontinuities and noises, so this stands as a major advantage for optimizing geophysical problems rather than other techniques.

Many applications were done on the recent years based on artificial intelligent techniques to control, prediction and inference studies [11,12,13,14,15]. The fuzzy modeling was first explored by Takagi and Sugeno [14] and later ANFIS network was developed by Jang in 1993 [16]. An adaptive network is a multilayer feed forward network in which each node performs a particular function on giving the input to the network. Electrical resistivity data collected from the VES (Vertical Electrical Sounding) data is used for training the data set. Trained data set will be the reference data for interpreting the subsurface layer parameters of the earth. This dataset will be used to train ANN that best model this data. Layer parameters viz., apparent resistivity and Thickness of earth's subsurface can be estimated using this tool with less error percentage.

Genetic Algorithm coupled Neural Networks (GA-ANN) approaches the combined behavior of genetically optimized neural network system. Input data are categorized to seek the objective function using genetic algorithm which is one of the best global search optimization technique can able to find the global optimum minimum. They can be used to modeling or control linear or non-linear systems using ANN. ANN can generate the synthetic data required for geoelectrical inversion. GA searches the best-fit optimization generated the synthetic data and interpretation was done by the ANN model. The best-fitting individual of the population in GA will predict the output with less error percentage

Our Study area comprises of Rajapalayam-Alangulam Belt of Virudhunagar District , Tamil Nadu State in Indian subcontinent. This area lies in the southern part of the state in the Eastern slopes of Western Ghats. The Study area is endowed with various mineral and water resources. To study the layered structure of the Earth with the delineation of aquifer parameters , the major problem that has to be optimized in this area is the mixed and complex nature of geological formation in that area . Vertical Electrical Sounding method provides a good solution for that. A wide variety of optimization techniques have been used to interpret Vertical Electrical Sounding(VES) data of geophysical prospecting but this coupled approach of GA-ANN shows better results in obtaining information regarding layered structure of the Earth

Geophysical Method

Geophysical exploration techniques are vibrant and powerful tools that play a vital role in the delineation of aquifer parameters in different geological formations. In particular, The Geophysical method consisting of vertical electrical sounding (VES) survey has been proved to know the variation of resistivity of the aquifer parameters [17]. Schlumberger electrode array having principle advantage over several types of arrays , (Fig. 1) is used to study the electrical resistivity distribution of the subsurface in order to understand the groundwater conditions. The Vertical Electrical Sounding provides a non-destructive,fast and economic way to study properties of aquifers. An important advantage of

VES method is that quantitative modelling is possible using either model curves or software. The resulting models provides accurate estimates of electrical resistivity, thickness and depth of subsurface strata of the Earth. This array is a powerful tool in delineation of groundwater potentials because of its simplicity nature and cost effective. The field procedure involves, among the four electrodes that are used, the potential electrodes (M and N) remain fixed and the current electrodes (A and B) are expanded symmetrically about the centre of the spread (Fig. 1). With very large values of current electrodes, however, it is necessary to increase the potential electrodes. Maximum half current electrodes (AB/2) separation used in this survey is 100 meters. 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.

Artificial Neural Network Architecture

Artificial neural network can play a major role to dig out the mystery of earth science by means of the previous learned examples. These networks have self-learning capability and are fault-tolerant as well as noise-immune and have applications in various fields [18].

A very important feature of these networks is their adaptive nature, whose 'learning by example' replaces "programming" in solving problems. This feature makes such computational models very appealing in application domains where one has little or incomplete understanding of the problem to be solved but where training data is readily available. ANNs are now being increasingly recognized in the area of classification and prediction, where regression model and other related statistical techniques have traditionally been employed. Introducing the layer of neurons involved in the training is shown in (Fig 2). The most widely used learning algorithm in an ANN is the back propagation algorithm. There are various types of ANNs like multilayered perception, radial basis function and kohonen networks. In fact majority of the network are more closely related to the traditional, mathematical and/or statistical models such as non-parametric pattern classifiers, clustering algorithms, non-linear filters, and

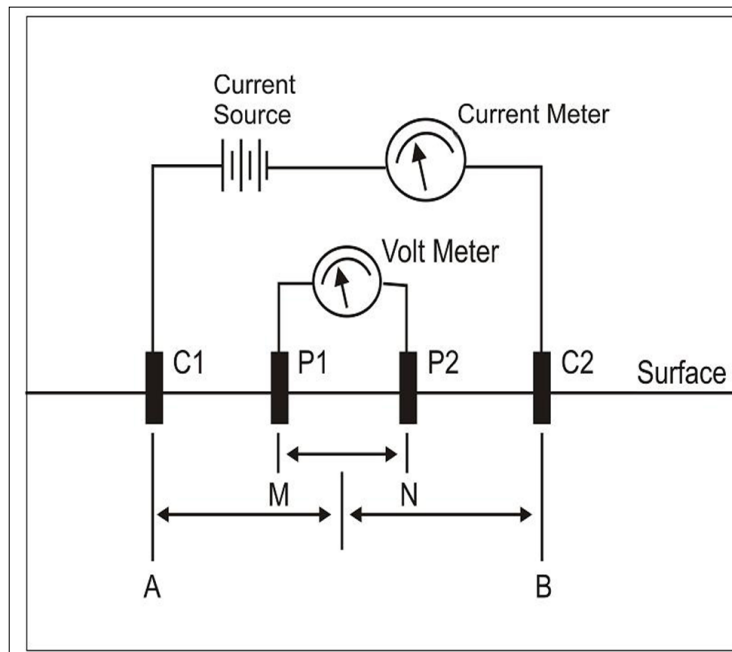


Figure 1. Schlumberger Electrode Configuration

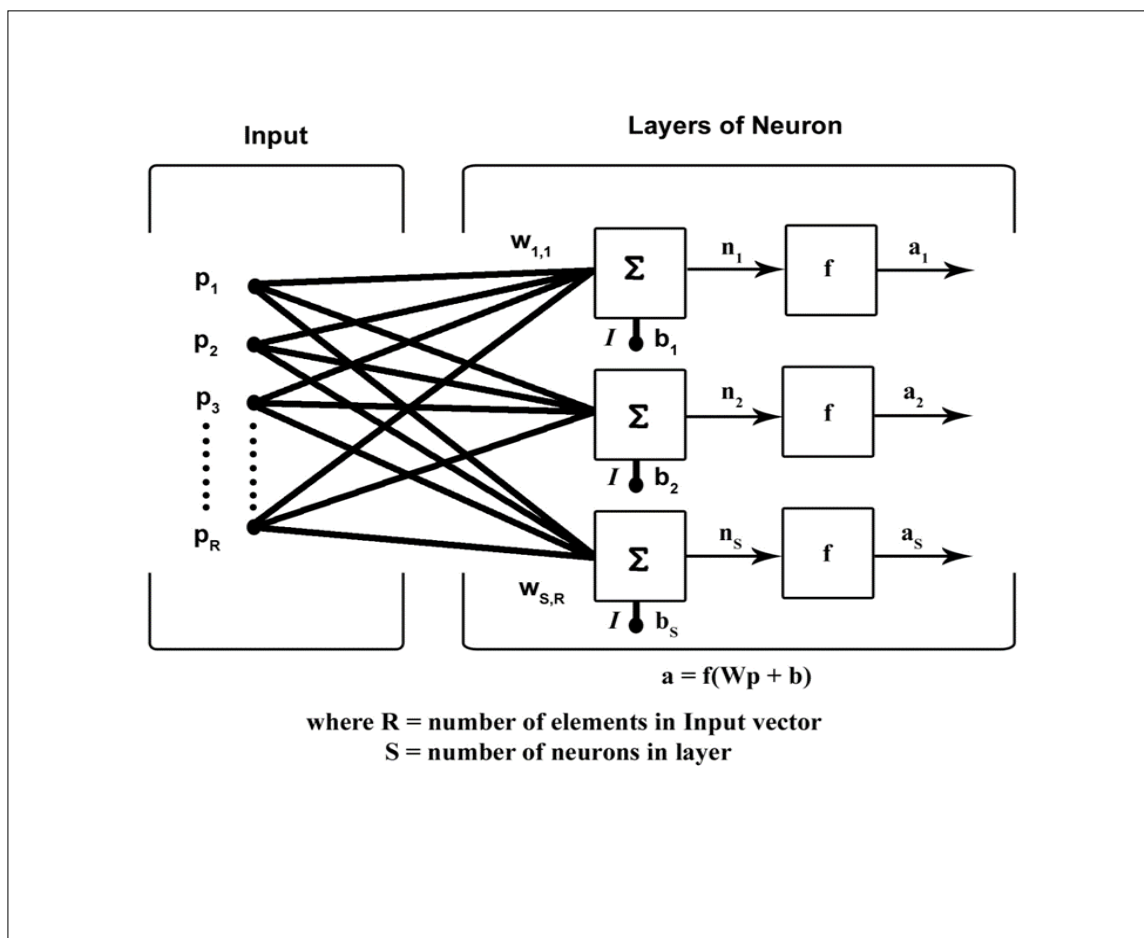


Figure 2. Input element vector of ANN with layer of neurons having certain weights and biases

statistical regression models. ANNs have been used for a wide variety of applications in the field of non-linearity.

Here the sigmoid function is used as activation function. It can produce outputs with reasonable discriminating power and its output functions are differentiable, which is essential for the back propagation of errors [19]. During FNN training, each hidden and output neuron process inputs by multiplying each input by its weights. The products are summed and processed using the activation function. The expression for logistic sigmoid function and its derivative is given below.

$$f(x) = 1 / (1 + e^{-x})$$

$$f'(x) = f(x)[1 - f(x)]$$

The number of neurons in the hidden layer varies as per the requirement of optimum performance, which could be decided on trial and error basis. Initial weights are assigned at random in the range suitable for the activation function at neurons [20].

In this paper an iterative non-linear feed forward network backpropagation algorithm and the performance was tested.

Genetic Algorithm

Various optimization problems that can be solved by the method of Genetic Algorithm (GA), which is based on natural selection particularly of biological evolution. GA is generally mimicking the process of natural evolution, which generates useful solutions for optimization problems using techniques inspired by natural evolution, such as inheritance, mutation, selection and crossover [10]. The genetic algorithms are inspired by Darwin's theory of evolution that

To ensure their survival, the respective individual Organism must be over reproductive.

1. In the next generation, some minor stochastic changes are seen.
2. The more successful will be better adaptive individuals which will be able to pass their genes to the successive generations.

To optimize a particular problem GA has five major steps, that are as follows;

A random initial population of the problem that has to be optimized is generated by the means of random

generator. Initial population is represented by the set of parameter values which describes the problem. These set of parameters are then transformed into a string of binary bits (0's and 1's) of specific length. At each and every generations, the chromosomes are tested for fitness using fitness function which is achieved by decoding binary strings into set of parameter values

By the means of various ranking schemes which results that good chromosomes have higher chances to be in the next generation. Rank selection schemes performs the operation over the population on the basis of fitness values. In particular over various ranking schemes tournament selection is preferable [10]. In tournament selection process, a set of random individuals is selected from the existing generation and the individuals which processes high fitness values are forwarded to next generation.

The process of cross over and mutation is performed over existing generation to create new generation which is based on crossover and mutation probability.

The new generation is formed by replacing the currently existing generation. The children of the successive generations are decided based upon the fitness value through fitness function provided.

The process of GA stops when at least one among the various stopping criteria such as Maximum number of generations, Maximum time limit, Maximum fitness limit etc., met.

Conventional search methods are not robust. The robustness behavior of GA has been used successfully in many applications. Far more desirable performance would be obtained in this Robust Scheme. Simple genetic algorithm that yields good results in many practical problems is composed of three operators. Fig. 3 represents the genetic algorithm coupled with neural networks optimization algorithm.

1. Reproduction
2. Crossover
3. Mutation

Reproduction is a process in which individual strings are copied according to their objective function values. The operator, of course, is an artificial version of natural selection, a Darwinian survival of the fittest

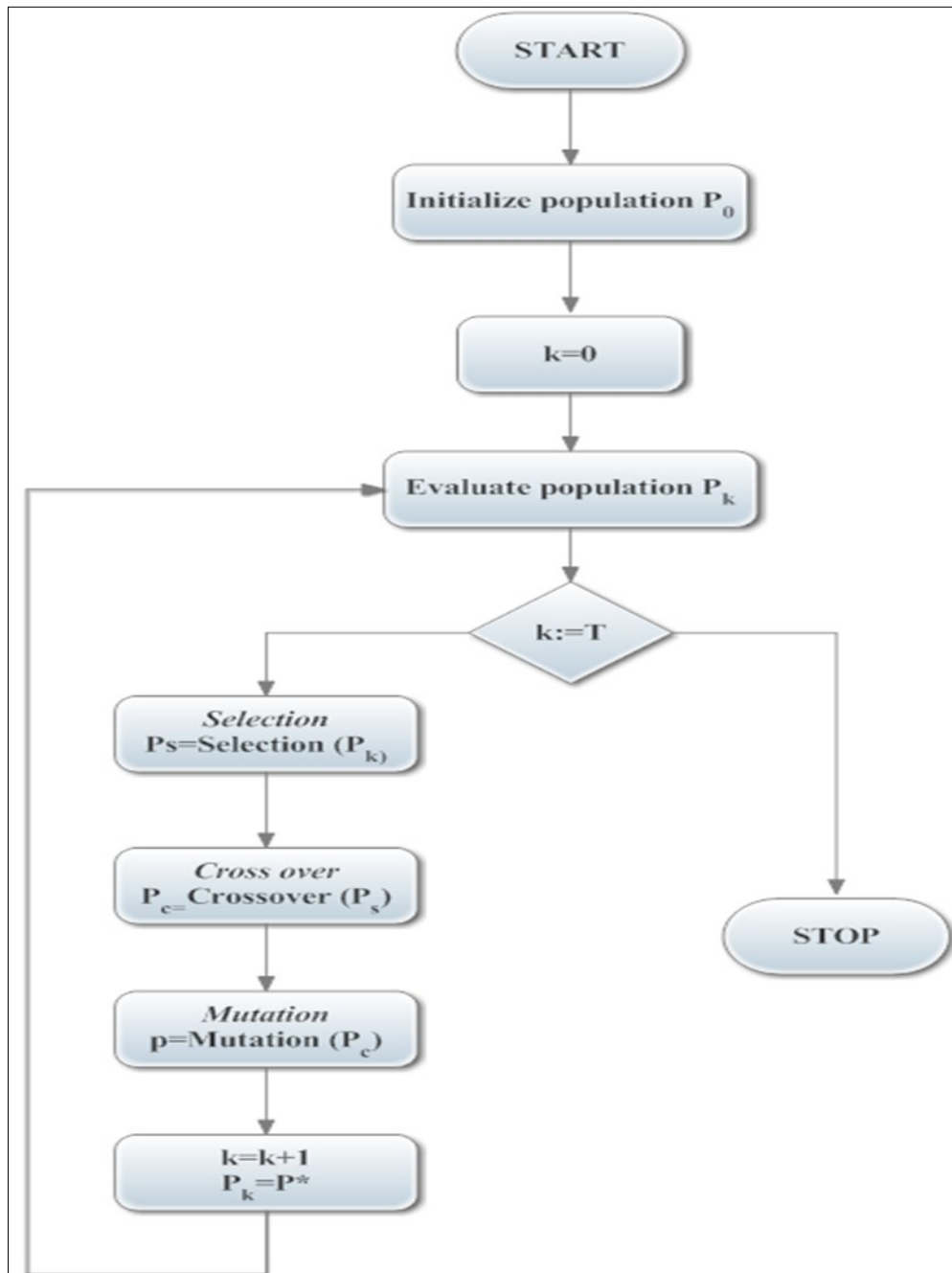


Figure 3. Genetic algorithm optimization algorithm

among string creatures. In natural populations fitness is determined by a creature's ability to survive predators, pestilence, and the other obstacles to adulthood and subsequent reproduction. In our case the apparent resistivity data has been chosen. The best fit optimized true resistivity data was the output that better survives in the selection operator. Reproduction process with individual strings is on the basis of apparent resistivity data. The dataset provides the necessary population generation needed for the natural selection of the algorithm.

- Genetic Algorithm (GA) description:
- step1: Generation = 0;
- step2: Initialize M (Generation);
- Evaluate M (Generation);
- step3: While (GA has not converged or terminated)
- Generation = Generation + 1;
- Select M (Generation) from M (Generation - 1);
- Crossover M (Generation);
- Mutate M (Generation);
- Evaluate M (Generation);
- End (while)
- step4: Terminate the GA.

Hybrid Genetic Algorithm Optimization on Neural Networks

Many optimization algorithms can be used to interpret the Vertical Electrical Sounding data. But most of the optimization algorithm doesn't behave in the way to represent the exact lithological features. But when the optimization algorithm concepts are integrated, the performance of the architecture has been improved. One of such method is integrating the GA and ANN. Many works based on Geoelectrical resistivity data has been going on to converge the interpreted results to a unique solution. Earlier in 1930's many researchers investigated the inverse problem [21,22,23,24]. Later the improvement of computational facilities in 1980s there has been considerable improvement in the data collection and interpretation. The interpretation of vertical electric sounding data was greatly marked by

three events: (i) the linear filter theory [25], (ii) the widespread use of digital computers, and (iii) the application of general linear inverse theory. With the latter, the concept of automatic inversion and analysis became popular, in which computer programs resulting from this theory were able to find the best-fitting model automatically, provided they were given a reasonably accurate initial model or a priori information [25, 26, 27, 28, 29, 30,33].

Since the last decades the optimization techniques, especially global optimization techniques, are widely used for the inversion of resistivity data and other geophysical related problems.

Sen et al. in 1993 [3] introduced the application of nonlinear global optimization method namely simulated annealing to the direct interpretation of resistivity sounding data. Chundururu et al, in 1995 [31] applied three non-linear algorithms metropolis simulated annealing, very fast simulated annealing and a genetic algorithm (GA) to invert synthetic resistivity data for the classical models embedded in a uniform matrix of contrasting resistivity. The GA technique has also been applied for solving the layer parameter estimation and ground water problems [32].(Fig. 4, 5, 6 and 7) shows the VES interpretation for different stations in Rajapalayam-Alangulam belt. Stanley Raj, et al. [34, 35] proved the multiplayer model for evaluating the subsurface parameters easily. Thus regression of multilayer model will provide the exact layer parameters of the earth. Table 1 shows the comparison of conventional method to the GA-ANN model output which proves that the minor subsurface layers also can be extracted using this algorithm. If there is a small deviation in the curve of VES will produce a discriminate change in the layer parameter output. (Fig. 8) shows the genetic algorithm population and fitness for Neural networks optimization function. (Fig. 9, 10,11 and 12) shows the 3d plot for genetic algorithm optimizing ANN function and the performance in terms of reduced error percent is depicted.

Results and Discussion

For the field data validation, VES data collected from Rajapalayam-Alangulam belt has been chosen for evaluating the performance of GA coupled ANN algorithm. Table 1 shows the interpretation of VES data

Table 1. Comparison of conventional and GA-ANN optimization results on geoelectrical resistivity data

Stations	Location	Conventional Method		Genetically optimized ANN	
		True Resistivity	Depth	True Resistivity	Depth
Station 1	Longitude 9.39494 Latitude 77.611	68 53.8 936	1.5 22.4	27.90 44.46 48.35 71.59 82.42 405.50 120.8 148.36	2 10 25 30 50 60 80
Station 2	Latitude 77.6463 Longitude 9.37541	20.9 40.3 218	1.24 9.29	4.7 30.33 37.09 105.83	2 10.5 30
Station 3	Longitude 9.36231 Latitude 77.6926	15 49.4 253	4.53 14.4	132.47 86.21 85.71 96.16 106.76 128.1 168.5	2 15 30 45 50 80
Station 4	Longitude 9.47087 Latitude 77.6269	32.5 57.9 426	1.61 21.8	32.8 65 366	2 25

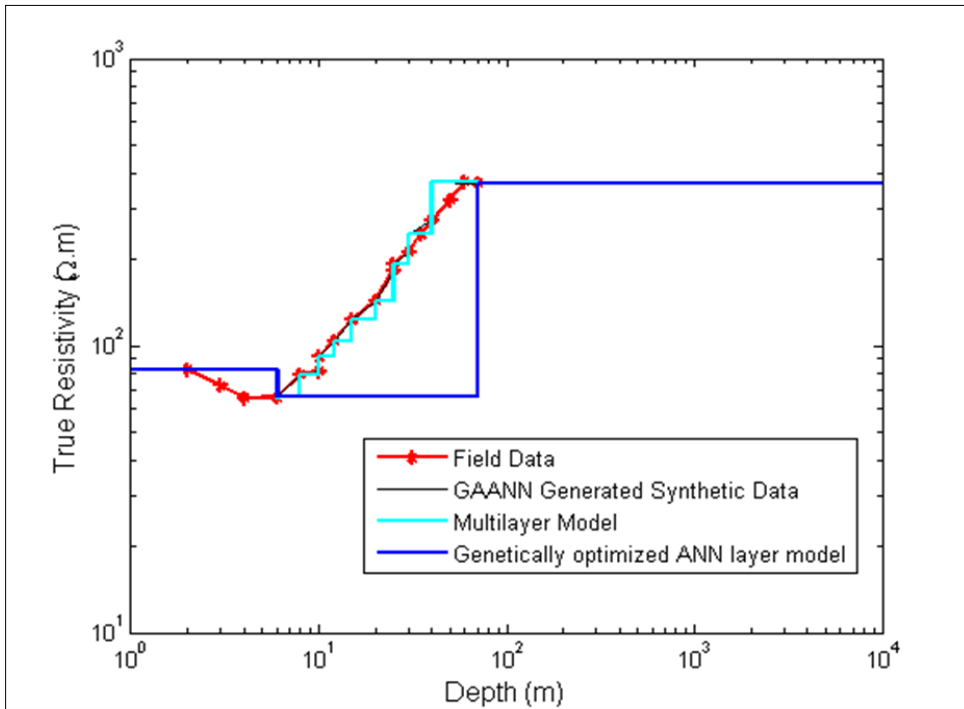


Figure 4. Genetically optimized Neural Networks geoelectrical inverted layer model for Station1

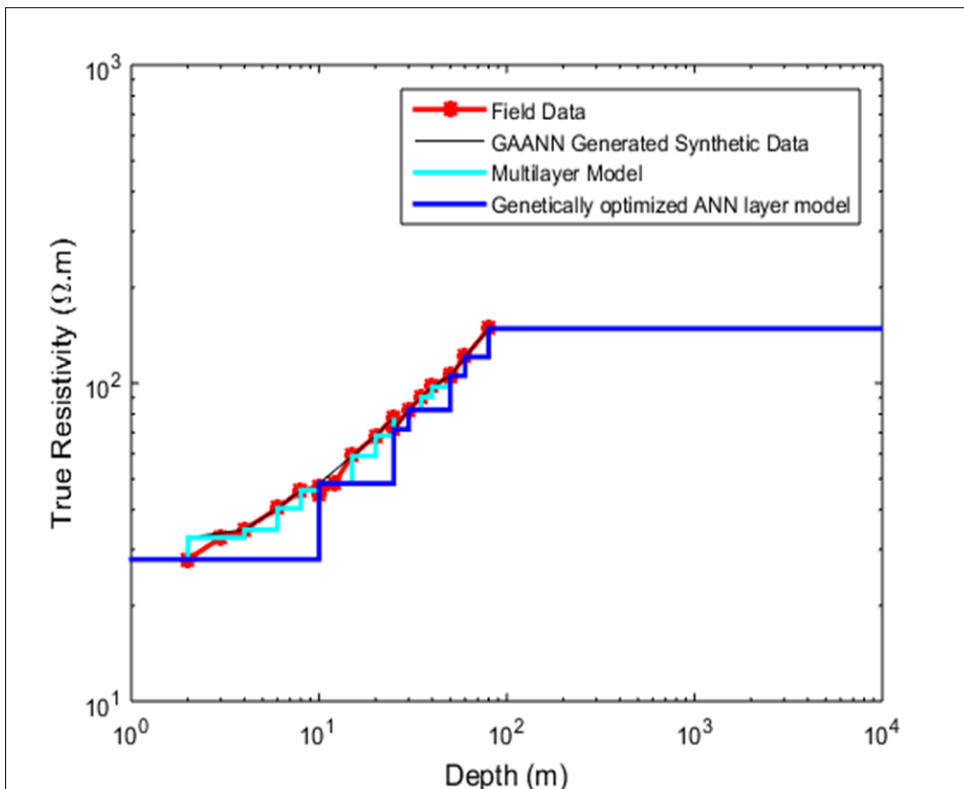


Figure 5. Genetically optimized Neural Networks geoelectrical inverted layer model for Station 2

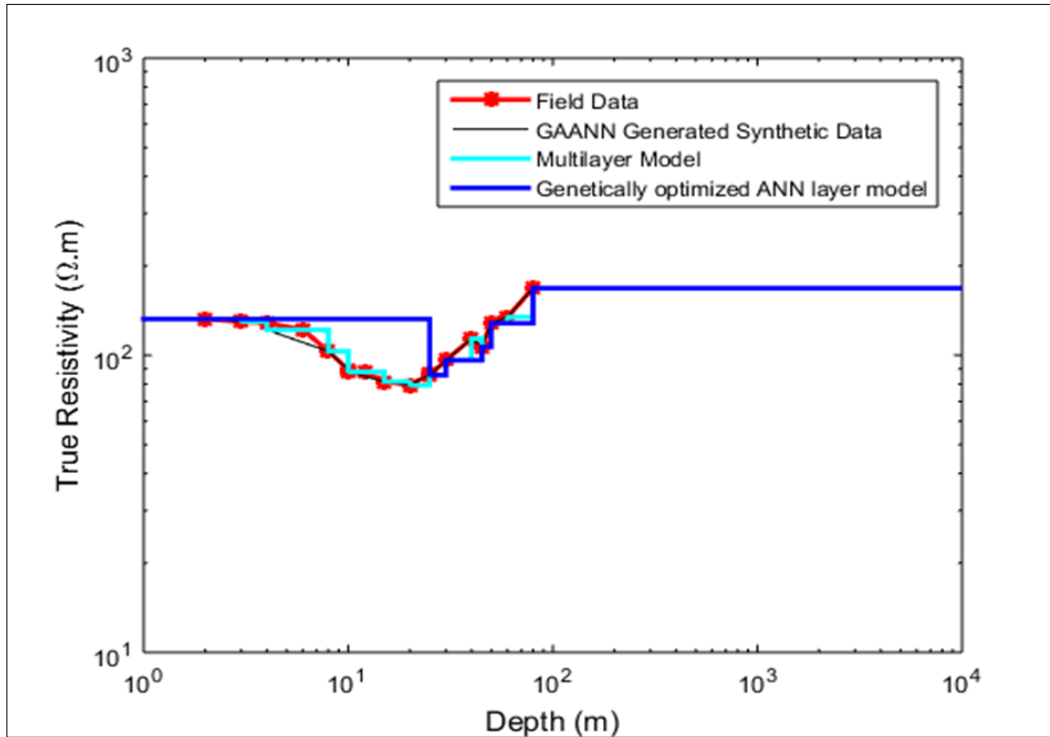


Figure 6. Genetically optimized Neural Networks geoelectrical inverted layer model Station 3

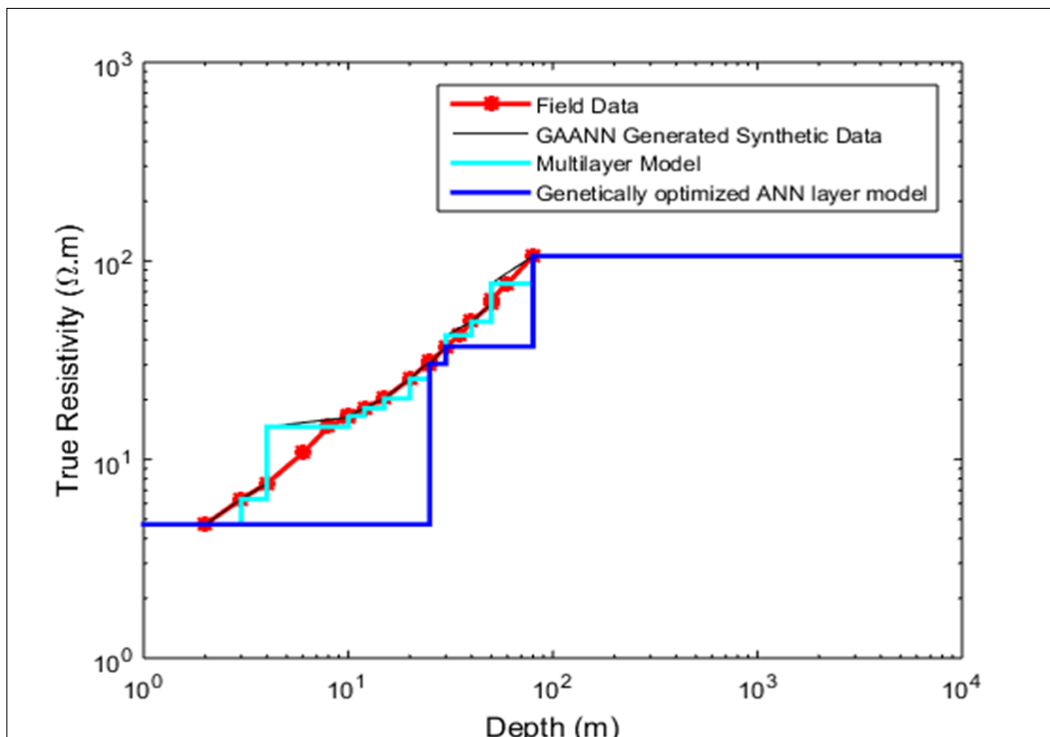


Figure 7. Genetically optimized Neural Networks geoelectrical inverted layer model Station 4

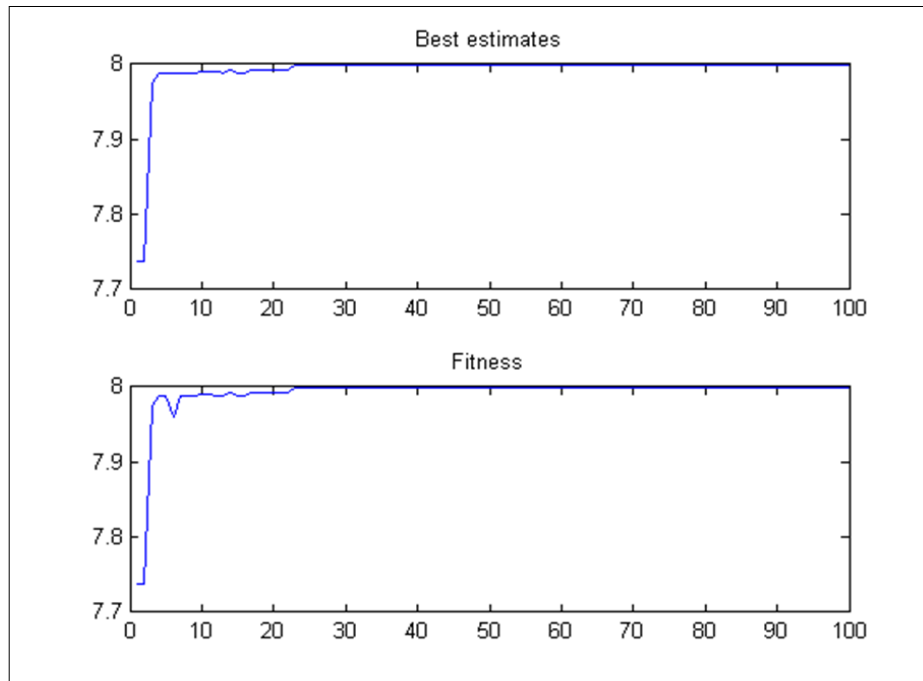


Figure 8. Graph showing the Genetic algorithm best estimates and fitness

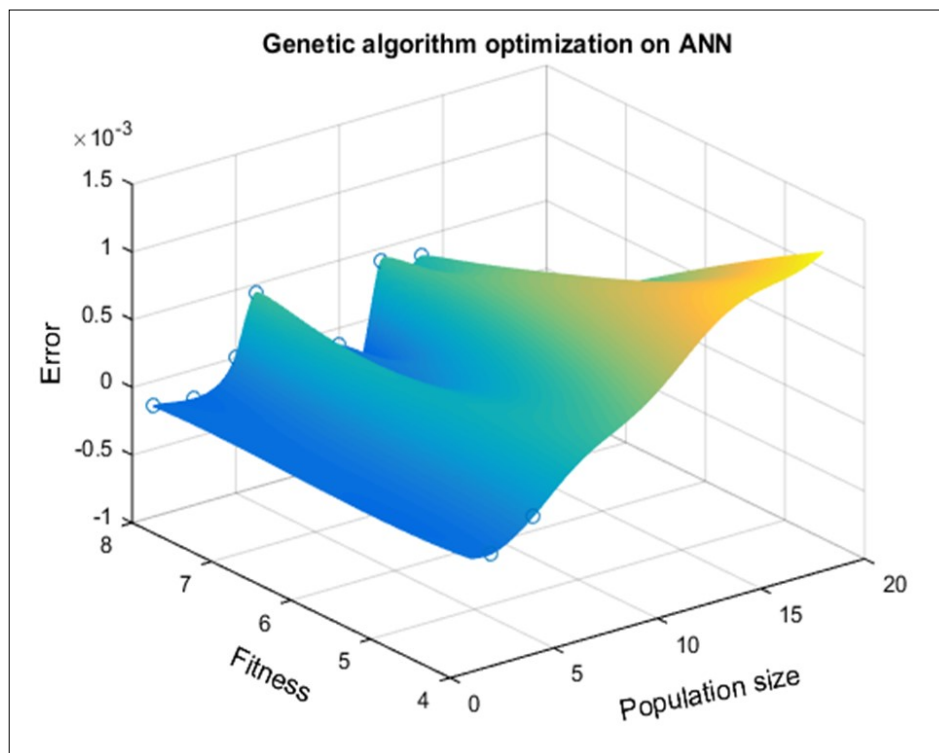


Figure 9. Three-dimensional model for population size, fitness and error percent of GA-ANN for geoelectrical resistivity inversion for station 1

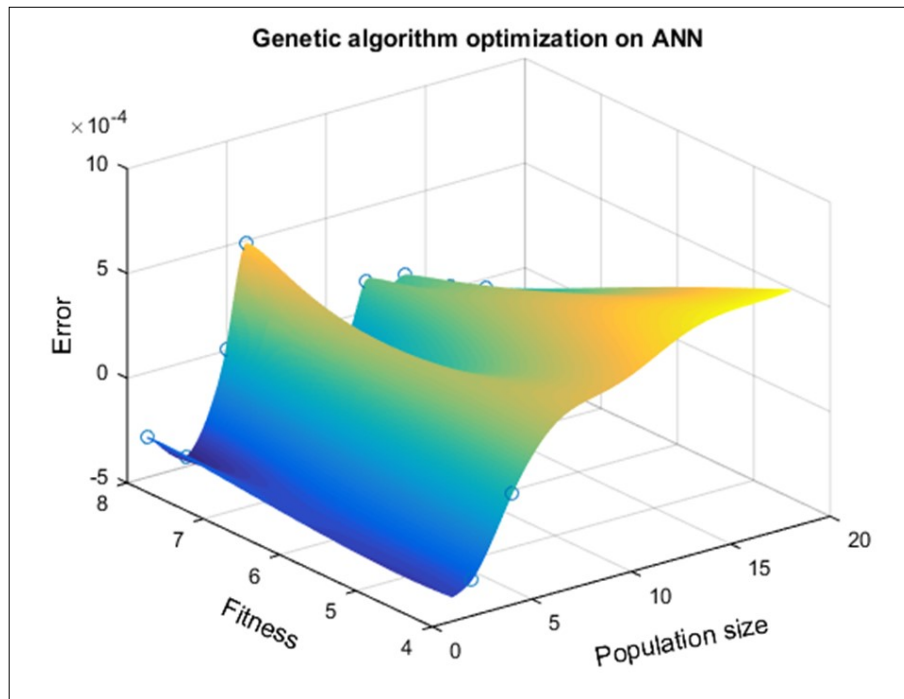


Figure 10. Three-dimensional model for population size, fitness and error percent of GA-ANN for geoelectrical resistivity inversion for station 2

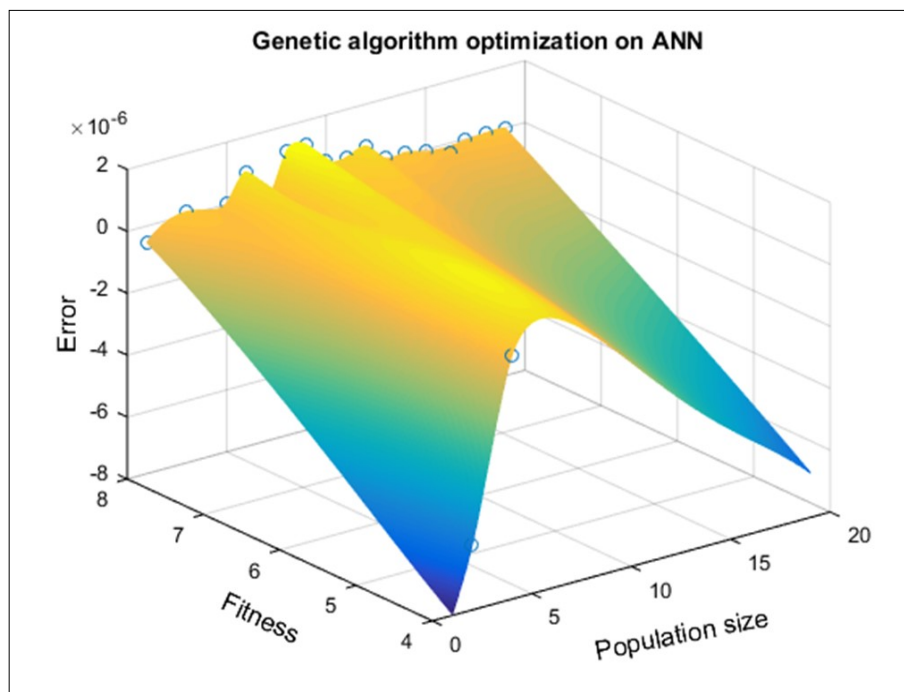


Figure 11. Three-dimensional model for population size, fitness and error percent of GA-ANN for geoelectrical resistivity inversion for station 3

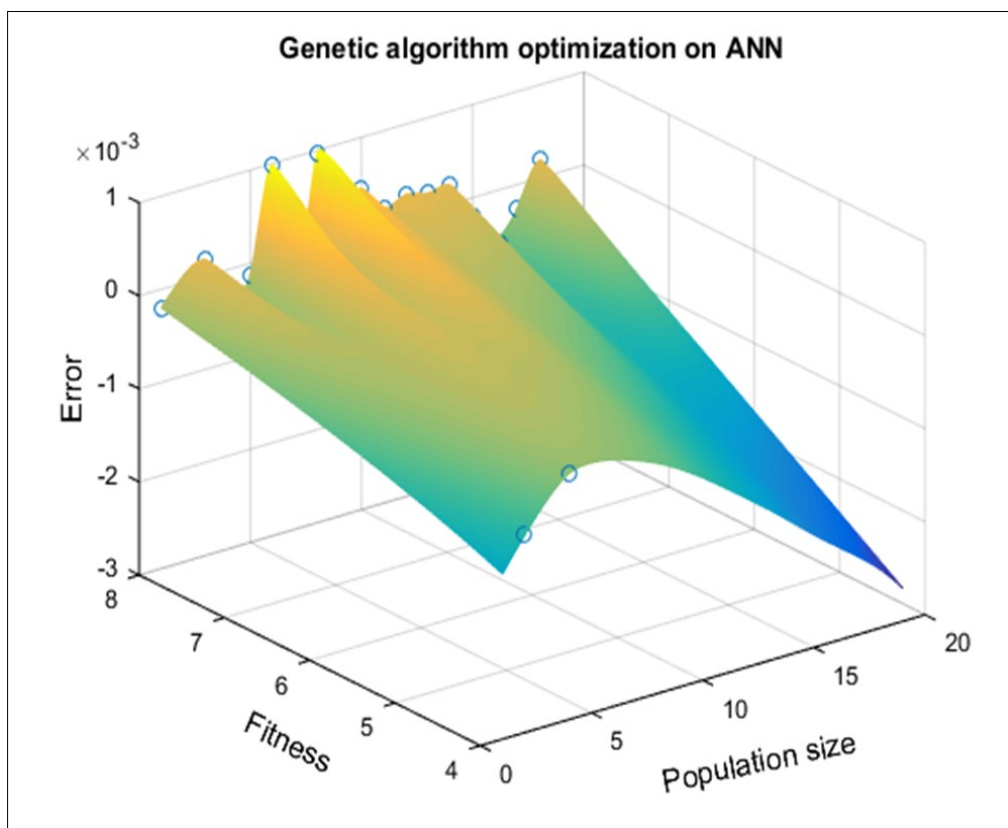


Figure 12. Three-dimensional model for population size, fitness and error percent of GA-ANN for geoelectrical resistivity inversion for station 4

using conventional and GA-ANN method. It reveals that using GA-ANN technique we can extract very minor layer information of earth's subsurface and it is proved to be efficient. If the geology is more complex, conventional algorithm cannot be expected to have better performance and we need to apply such artificial intelligent techniques.

Conclusion

Thus, on summarizing hybrid GA-ANN model always surpasses other algorithms, especially the converging technique. The predicted accuracy of GA-ANN expresses best performance result of more than 90%. Finally, the output layer parameters were modeled. In this model, it converges more quickly and reduces the complexity involves in the non-linear problem. Specialized behavior of GA-ANN algorithm proved that this can be applied to complex geological data and the results are satisfactory and also it can be employed to all kinds of non-linear geophysical inversion problems.

GA-ANN coupled algorithm addresses the geoelectrical data inversion using hybrid approach. In this study, optimization of neural networks is done by genetic algorithm and the local convergence is made using multi node neural networks. On comparing with the traditional approach, this hybrid algorithm has a suitability to predict the subsurface layer parameters effectively. Errors and noises are easily removed through this hybrid approach because of the optimization of objective function of neural networks is made by genetic algorithm.

References

1. Stoffa, P. A. and M. K. Sen, 1991, Nonlinear multiparameter optimization using genetic algorithms: Inversion of plane-wave seismograms, *Geophysics*, vol. 56, no. 11, 1794-1810.
2. Shibutani, T. A., M. Sambridge and B. Kennet, 1996, Genetic Algorithm Inversion for Receiver Functions with Application to the Crust and Uppermost Mantle Structure Beneath Eastern Australia, *Geophysical*

- Resecirch Letters, 23, 1829-1832.
3. Sen M.K., Bhattacharya B.B., Stoffa P.L. (1993) Nonlinear inversion of resistivity sounding data. 58 (4), 496-507.
 4. Pierini, S., Aleardi, M., & Mazzotti, A. (2019). A method to attenuate genetic drift in genetic-algorithm optimizations: Applications to analytic objective functions and two seismic optimization problems. *Geophysics*, 84(2), R295-R310.
 5. Rubaiyn, A., Priyono, A., & Yudistira, T. (2019, August). Near-surface S-wave Estimation base on Inversion of Rayleigh Wave Dispersion Curve Using Genetic Algorithm. In *IOP Conference Series: Earth and Environmental Science* (Vol. 318, No. 1, p. 012013). IOP Publishing.
 6. Putra, A. S., Srigutomo, W., Hidayat, Y., & Lesmana, E. (2019, April). A Comparative Study of Simulated Annealing and Genetic Algorithm Method in Bayesian Framework to the 2D-Gravity Data Inversion. In *Journal of Physics: Conference Series* (Vol. 1204, No. 1, p. 012079). IOP Publishing.
 7. Ayani, M., MacGregor, L., & Mallick, S. (2020). Inversion of marine controlled source electromagnetic data using a parallel non-dominated sorting genetic algorithm. *Geophysical Journal International*, 220(2), 1066-1077.
 8. Yan, X., Li, P., Tang, K., Gao, L., & Wang, L. (2020). Clonal selection based intelligent parameter inversion algorithm for prestack seismic data. *Information Sciences*, 517, 86-99.
 9. Gunawan, I., Wahyudi, E. J., Alawiyah, S., Kadir, W. G. A., & Fauzi, U. (2020). Annotation of Using Borehole Time-Lapse Gravity by Genetic Algorithm Inversion for Subsurface Modeling. *Journal of Engineering & Technological Sciences*, 52(2).
 10. Goldberg D.E. (1988) Genetic algorithms in search optimization and machine learning. Addison-Wesley Publishing company Inc. 1-45.
 11. Flathe H. (1955) A practical method of calculating geoelectrical model graphs for horizontally stratified media. 3(8), 268-294.
 12. Kosinky W.K., Kelly W.E. (1981) Geoelectrical sounding for predicting aquifer properties. 19(2), 163-171.
 13. Niwas S., Singhal D.C. (1981) Estimation of aquifer transmissivity from Dar-Zarrouk parameters in porous media. 50, 393-399.
 14. Sugeno M. (1985) Industrial Applications and Fuzzy Control. Elsevier, New York.
 15. Zadeh L. (1965) Fuzzy sets. 8(3), 338-353.
 16. Jang J.S.R. (1993) Adaptive-network based fuzzy inference system. 23(3), 665-685.
 17. Rijo L., Pelton W.H., Feitosa E.C., Ward S.H. (1977) Interpretation of apparent resistivity data from Apodi Valley, Rio Grande do Norte, Brazil. 42(4), 811-822.
 18. Sreekanth P.D., Geethanjali N., Sree Devi P.D., et al. (2009) Forecasting Groundwater level using artificial Neural networks. 96(7), 933-939.
 19. Yegnanarayana B. (2004) Artificial Neural Networks. Prentice Hall of India Private Limited, New Delhi. 461.
 20. Stephen J., Manoj C., Singh S.B. (2004) A Direct inversion scheme for deep resistivity sounding data using artificial neural networks. 113(1), 49-66.
 21. Langer R.E. (1933) An inverse problem in differential equations. 39(10), 814-820.
 22. Pekeris C.L. (1940) Direct method of interpretation in resistivity prospecting. 5(1), 31-42.
 23. Stichter L.B. (1933) The Interpretation of the resistivity prospecting method for horizontal structures. 4(9), 307-322.
 24. Werbos P. (1974) Beyond regression: New tools for prediction and analysis in the behavioral sciences. Ph. D. dissertation, Harvard University.
 25. Ghosh D.P. (1971) Inverse filter coefficients for the computation of the apparent resistivity standard curves for horizontally stratified earth. 19(4), 769-775.
 26. Inman J. R., Ryu J., Ward S.H. (1973) Resistivity inversion. 38(6), 1021-1240.
 27. Johansen H.K. (1977) A man/computer interpretation system for resistivity soundings over horizontally stratified earth. 25(4), 667-691.
 28. Kunetz G., Rocroi J.P. (1970) Automatic processing of

- electrical soundings. 18(2),157-198.
29. Mogren S., Batayneh A., Elawadi E., Al-Bassam A., Ibrahim E., Qaisy S. (2011) Aquifer boundaries explored by geoelectrical measurements in the Red Sea coastal plain of Jazan area, Southwest Saudi Arabia. 6(15), 3688-3696.
 30. Zohdy A.A.R. (1989) A new method for automatic interpretation of schlumberger and Wenner sounding curves. 54(2), 148-273.
 31. Chunduru R.K., Sen M.K., Stoffa P.L., Nagendra R. (1995) Non-linear inversion of resistivity profiling data for some regular geometrical bodies. 43(8), 979-1003.
 32. Jha M.K., Kumar S., Chowdhury A. (2008) Vertical electrical sounding survey and resistivity inversion using genetic algorithm optimization technique. 359 (1-2), 71-87.
 33. Srinivas Y., Stanley Raj A., Hudson Oliver D., Muthuraj D., Chandrasekar N., (2012) A Robust behaviour of Feed Forward Back propagation algorithm of Artificial Neural Networks in the application of vertical electrical sounding data inversion. *Geoscience Frontiers*, 3(5), 729-736
 34. Stanley Raj, A., Y.Srinivas, D.Hudson Oliver D. Muthuraj, (2014) A novel and generalized approach in the inversion of geoelectrical resistivity data using Artificial Neural Networks (ANN). *Journal of Earth System Sciences* 123 (2), 395-411.
 35. Stanley Raj A., Hudson Oliver D., Srinivas Y (2015) An automatic inversion tool for geoelectrical resistivity data using supervised learning algorithm of adaptive neuro fuzzy inference system (ANFIS). *Model Earth Syst Environ*. 1(6).