Prediction of transmissivity of aquifer from geoelectric data using artificial neural network

Oluseun Adetola SANUADE[1]*, Kehinde David OYEYEMI[2], Joel Olayide AMOSUN[3], Julius O. FATOBA[3], Olaide S. HAMMED[4]

[1]Geosciences Department, King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia
[2]Department of Industrial Physics, Covenant University, Ota, Ogun State, Nigeria
[3]Department of Geophysics, Federal University Oye-Ekiti, Ekiti State, Nigeria
[4]Department of Physics, Federal University Oye-Ekiti, Ekiti State, Nigeria.

*sheunsky@gmail.com

Abstract. We have used artificial neural network (ANN) for the prediction of transmissivity (T) from geoelectric parameter in an aquifer of a typical basement complex, Southwestern, Nigeria. The study area is composed of migmatite-gneiss, charnockite and granite gneiss. A geophysical investigation involving Vertical Electrical Sounding (VES) was carried out in the study area. Twenty VES were acquired using Campus Ohmegaresistivity meter. VES curves were interpreted quantitatively by using the technique of partial curve matching and with 1D forward modeling by means of computer using *WinResist*software. The interpreted data were used to determine transverse resistance, T_R. Transmissivity were measured in the boreholes drilled at the 20 VES locations. T_R and measured T were subjected to ANN analysis using *MATLAB 2017a software* in order to predict T of the aquifer. Root mean square error (RMSE) was used to test the performances of our model. The results show that T_R range from 133.00 - 381.20 Ωm^2, T also vary from 0.4 to 4.1 m^2/day. However, ANN model was able to predict T values with coefficient of correlation (R) values of 0.97, 1.00, 0.99 and 0.94 for training, test, validation and all network models respectively. RMSE value for the ANN model was found to be 0.085 which implies high performance of our model. A linear relationship was suggested for the ANN analysis to predict T. It can therefore be concluded that with ANN model, it is possible to predict T of aquifer in the study area where geoelectric data such as T_R is known and T values unknown.
Keywords:Aquifer, transmissivity. ANN, geoelectric, prediction, MATLAB.

1. Introduction

Hydrogeological parameters including transmissivity, hydraulic conductivity, recharge, and so on, can be determined in an aquifer by means of classical hydrogeological techniques such as pumping tests, grain size analysis, tracer techniques. Most of these techniques are expensive. However geophysical methods such as electrical, seismic, gravity, electromagnetic have been used for groundwater investigation to determine the geometry of aquifer, subsurface electrical conductivity, depth to basement and so on [1]. These methods, when compared with hydrogeological methods, are cost effective, relatively fast and non-destructive. There are established relationships between different geophysical and hydrogeological parameters. For instance, a relationship exists between electrical

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

IOP Conf. Series: Earth and Environmental Science **173** (2018) 012025 doi:10.1088/1755-1315/173/1/012025

resistivity (ρ) and transmissivity (T). This relationship is made possible because both parameters are associated with the structure of pore spaces and heterogeneity [2-3]. Therefore, the combination of aquifer parameter estimated from borehole and electrical resistivity parameter estimated from surface electrical resistivity measurement could be a useful tool of estimating hydrogeological properties and potentials of aquifers.

Several studies have been done to know the relationship between hydrogeological and geophysical parameters [4-7]. [8] calculated the transmissivity of an aquifer using Dar Zarrouck parameters by taking into consideration two important laws which are Darcy's and Ohm's laws. The authors made use of the correlation that exists between the two parameters. [9] also reviewed the relationship that exist between transverse resistance and transmissivity by means of geostatistical analysis. [10] determined the correlation between aquifer resistivity that were normalized and hydrogeological parameter and normalized transverse resistance and transmissivity of aquifer in a weathered granite geological environment that are overlain by sand dunes in Jalore, northwestern part of India.

However, most of the previous studies on the relationship between hydrogeological and geophysical parameters used correlation studies and geostatistical approaches. The objective of this research, therefore, is to predict the transmissivity of aquifer from geoelectric data (transverse resistance) in a typical basement complex environment using artificial neural network (ANN).

The unique attribute of ANN is that it is data driven, which provides a unique opportunity to predict transmissivity effectively from transverse resistance as opposed to previous studies that finds a relationship between hydrogeological parameters and Dar Zarrouck parameters using other approaches. This will reduce non-uniqueness of parameter estimation and lead to improved transmissivity estimation.

The study area is in southwest area of Nigeria (Fig. 1). It lies within longitude 748600 – 749900 and latitude 850100 – 850600. Fig 2 shows a simplified geological map of the study area. The area is situated in the Precambrian Basement Complex where Migmatite-Gneiss, Charnockite and Granite Gneiss are the rock types in the area.

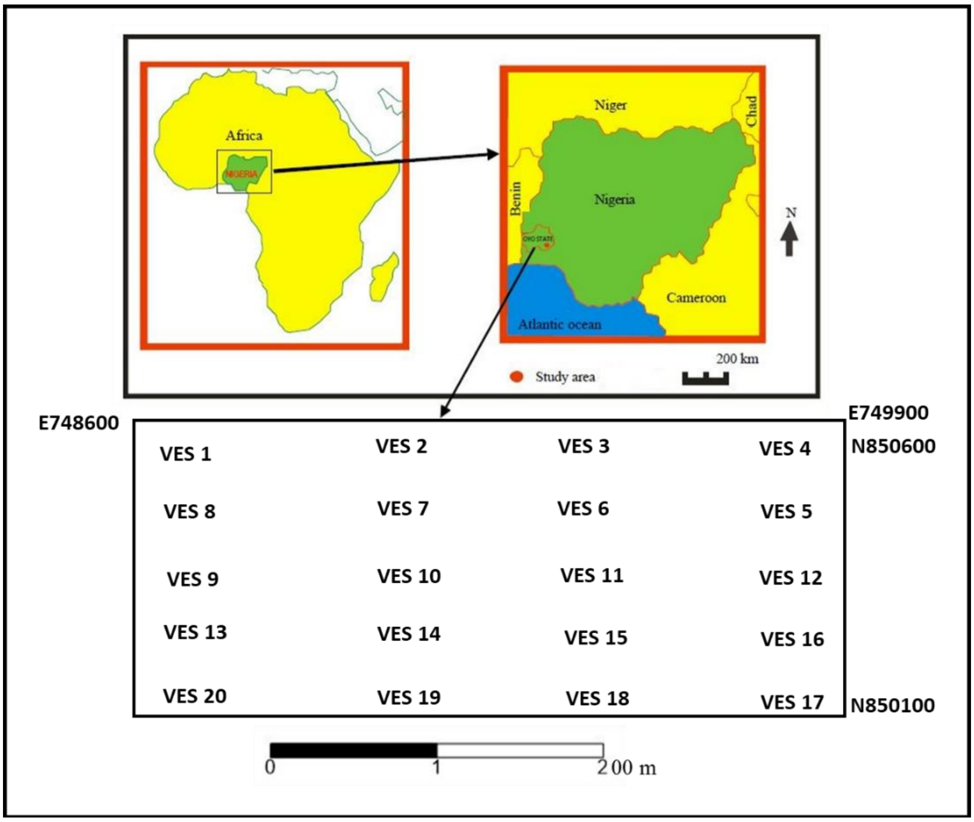

Fig. 1: Map of the study area

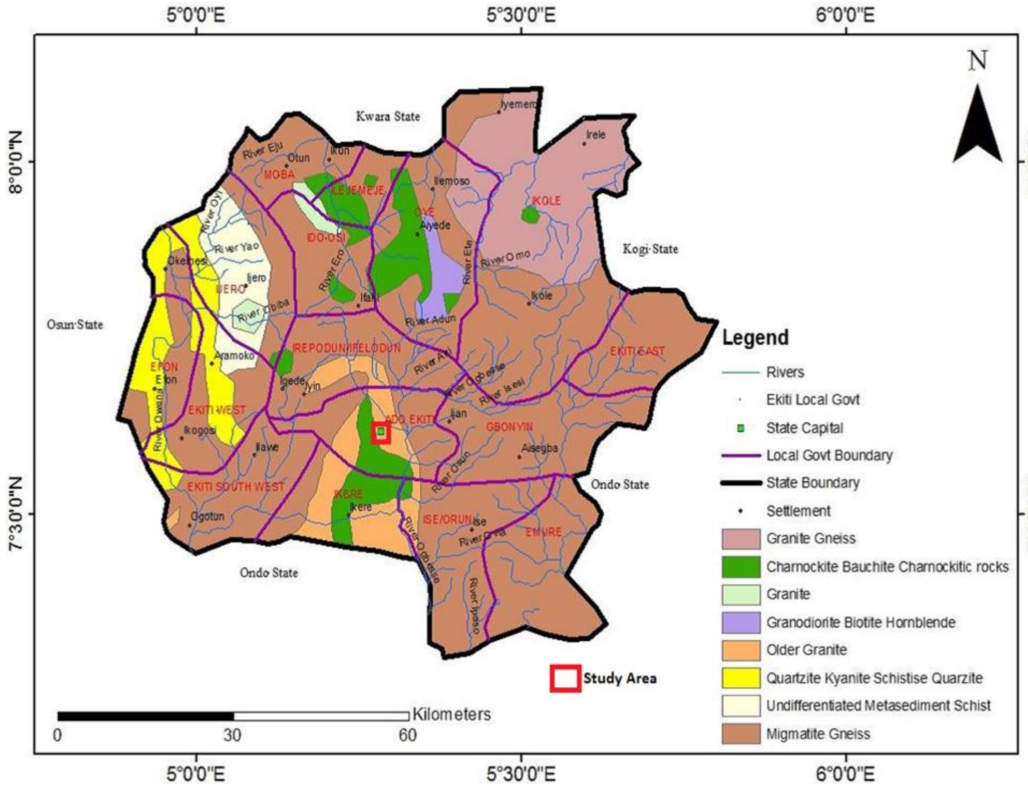

Fig. 2: Geological map of the study area [11]

.

2. Data acquisition and processing

2.1 Geophysical investigations

Geophysical investigations involving the Vertical Electrical Sounding (VES) was carried out to assess the groundwater potentials in the study area. Twenty (20) VES were acquired using Campus Ohmegaresistivity meter. The Schlumberger configuration, having current electrode spacing (AB/2) that varies from 1 to 100 m, was used during the data acquisition. VES curves were interpreted quantitatively using the technique of partial curve matching and 1D forward modeling using computer with the *WinResist* version 1.0 software. The processed and interpreted data were then used to determine aquifer resistivity, aquifer thickness and Dar Zarrouck parameter (transverse resistance (T_R) for the aquifer using equation 1 [12].

$$T_R = h.\rho \qquad (1)$$

Where h and ρ are the thickness and resistivity of the aquifer respectively

2.1 Hydrogeological investigations

Twenty (20) boreholes were drilled to depth of about 35 m at the VES points. Pumping test was carried out at the 20 VES points. The test was done for about 3 to 6 hours using constant discharge rate method. Transmissivity (T) of aquifer was determined from the pumping test [13]. The data obtained from both geoelectrical (Transverse resistance) and hydrogeological data (transmissivity) were subjected to ANN analysis for the prediction of transmissivity of the aquifer from the geoelectric data.

2.3 Artificial neural network (ANN)

ANN is an artificial intelligence method, which mimics the human brain's behaviour by acquiring knowledge through a learning process [14-15]. ANN is capable of modeling both non-linear and linear systems without assuming implicit functions as it is in the unlike most conventional statistical approaches. ANN has been utilized in several areas of science and engineering for classification, regression analysis, time-series forecasting, clustering and dynamic system modelling [16-17]. In this research, the multi layered perception (MLP) supervised feed-forward Levenberg-Marquardt based backpropagation algorithm (LMA) was used to predict transmissivity in the aquifer. The network consist of three layers (Fig.3): (i) input layer (transmissivity), which accepts and presents the input data to the network through a set of neurons (ii) hidden layer, which runs a set of algorithms to compute non-linear transformation of the input data (iii) output layer, which computes desired result through a set of linear output neurons in an iterative process. The output (T) depends on the weighted sum (w) of input variable (T) plus additional term called bias (b) to ensure numerical stability [16]. *MATLAB* software was used to implement the procedures. The procedures [15] is expressed as shown in equations 2–3:

$$p_j(n) = f\left(U_j(n)\right) = f\left(\sum_{i=0}^{n_i} \omega_{ji} x_i(n) + b_j\right) \; for\; j = 1, 2, 3, \dots\dots\dots. n_j \qquad (2)$$

$$q_k(n) = f\left(U_k(n)\right) = f\left(\sum_{j=0}^{n_j} \omega_{kj} p_j(n) + b_k\right) \; for\; k = 1, 2, 3, \dots\dots\dots. n_k \qquad (3)$$

where $x_i(n)$ = input to neuron i, $p_j(n)$ is the parameter estimated by neuron j of the hidden layer, $q_k(n)$ is the output calculated by the neuron k, ω_{ji} and ω_{kj} are parameters that control connection strength between input neuron i and hidden neuron j and j and k respectively. We used a non-linear hyperbolic tangent sigmoid transfer function (equation 4) and a linear transfer function in both the hidden and output layers.

$$f(U_j) = \frac{1 - e^{-2U_j}}{1 + e^{-2U_j}} \qquad (4)$$

whereU_j ε [-∞ ∞] and f (U_j) is bounded on (-1 1).

We used trial and error method of ANN for the development of our models for the estimation of the hidden layer's neutrons. We increased the numbers of neurons at each stage of analysis in order to optimize the ANN model. Back-propagation method is used for the iteration of the network learning process so as to ensure the error between the input and output data is minimized. The gradient vector of the error surface is then calculated and the networks are updated by as described in equations 5 and 6 [14].

$$\Delta\omega(n+1) = -\eta \frac{\partial f}{\partial \omega} |\omega = \omega(n) \qquad (5)$$

$$\omega(n+1) = \omega(n) - \eta \frac{\partial f}{\partial \omega} |\omega = \omega(n) \qquad (6)$$

where ω = strength that connects nodes of input and hidden layer and hidden and output layer nodes, n = epoch that means the amount of learning iterations, η = rate of learning which determines the number of change to connection weights of the network, Δω = increment in weight and f = sigmoid activation function.

Levenberg-Marquardt Algorithm (LMA) used in the network is a kind of backpropagation algorithm, which is a combination of two minimization techniques, namely, Gauss-Newton and gradient descent algorithms. This quality makes it sophisticated and powerful. In LMA, weights are calculated by using equation 7 [18].

$$\Delta\omega(n+1) = (J_k^T J_k + \mu I)^{-1} J_k^T e \qquad (7)$$

I is identity unit matrix, k = the output nodes and J = Jacobian of input errors.

The parameter μ is adjusted automatically at each iteration of Jacobian matrix computation. When μ is small, LMA turns to Gauss-Newton algorithm, and when it is very large, it transform to gradient descent algorithm. To prevent the process of learning being trapped in local minimal as a result of assigning random weights, we trained our ANN network for a minimum of 20 times and we chose the best model based on the goodness of fit. The process of training were terminated automatically when the sum of mean square error (MSE) is found within a predefined tolerance error. The neural network structure was employed to construct the ANN model as given in equation 8 and the structures of the ANN model are also provided in Table 1.

$$T = function\ (T_R) \qquad (8)$$

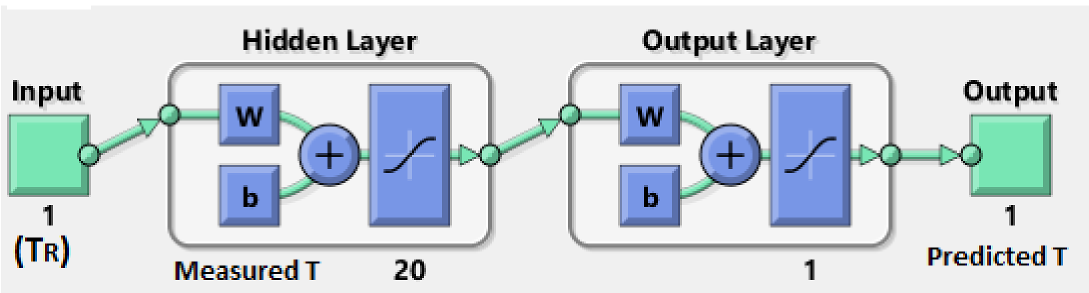

Fig. 3: Artificial neural network architecture

Table 1: The structures of the ANN model

Input neuron number	Hidden neuron number	Output neuron number	Type of network	Transfer function	Training parameter	Algorithm used for training
1	20	1	Feed-forward back-propagation	Tansig	Learning rate epochs 114	LM (Levenberg-Marquardt

3. Data interpretation

3.1 Geoelectric interpretation

The interpreted electrical resistivity survey shows that the study area is made up of 3 to 4 geoelectric layers. The dominant curve is the QH curve (Fig. 4). The typical curve types are given in Fig 5. The lithology is generally consist of top soil, laterite/clay, weathered layer and fresh basement rock.

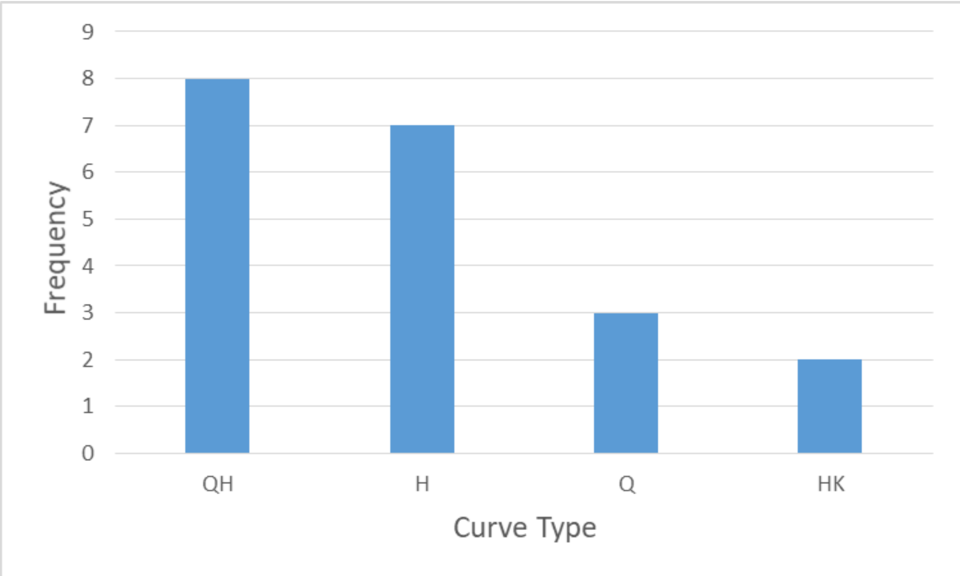

Fig. 4: The frequency of occurrence of the various curve types.

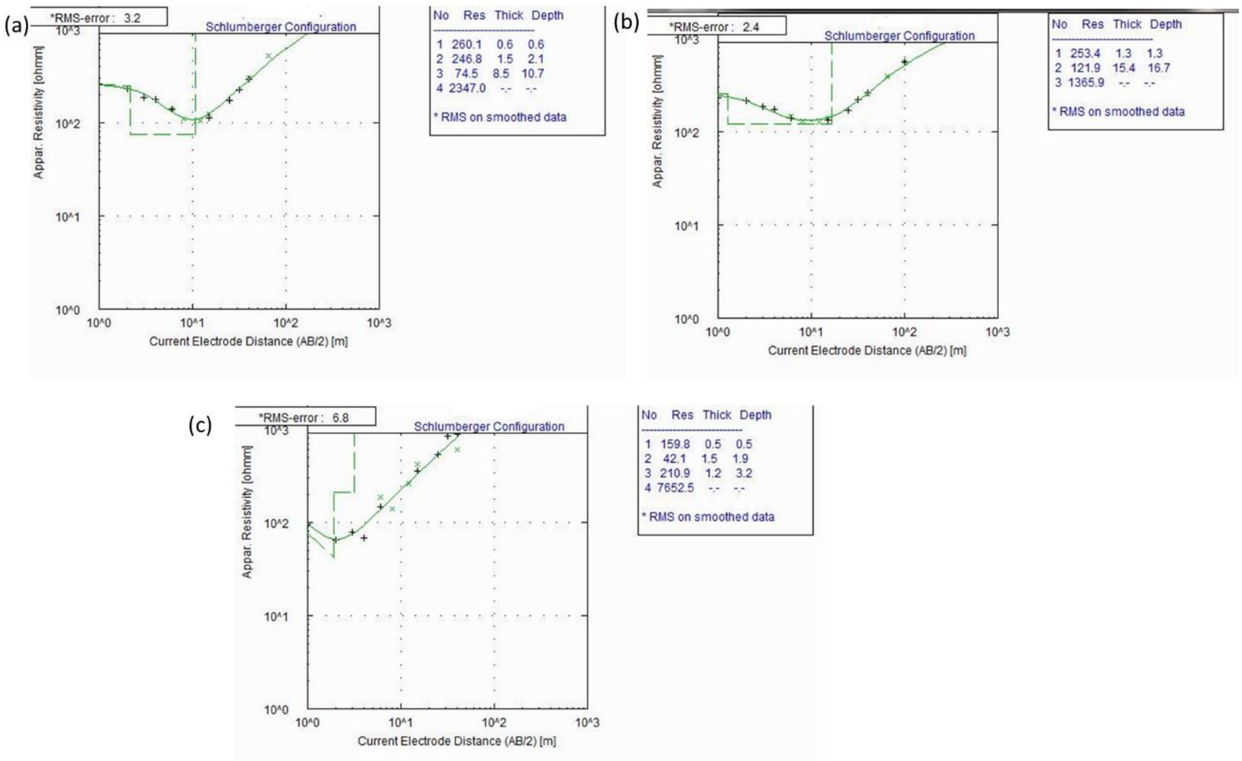

Fig. 5: Typical VES curve types in the study area (a) QH-type (b) H-type (c) HK-type

*3.1.1 Transverse Resistance.*The transverse resistance (T_R) is a parameter that describes the areas with high potentials for groundwater exploration [19]. T_R for each VES stations is presented in Table 2 as well as the measured transmissivity (T). High values of T_R indicates that the area have high potential for groundwater exploration and vice versa. T_R values vary from 133.00 to 381.20 Ωm^2 across the 20 VES whereas T range from 0.4 to 4.1 m^2/day. The highest T_R value was observed at VES station 3.

Table 2: Calculated Transverse resistance

VES Station	Transverse Resistance (T_R) (Ohm.m^2)	Measured Transmissivity (m^2/day)
1	202.10	0.8
2	276.15	0.6
3	381.20	1.5
4	168.60	0.8
5	218.90	0.5
6	133.20	0.4
7	261.50	3.9
8	172.20	4.1
9	242.20	2.2
10	261.10	3.5
11	202.10	0.8
12	276.15	0.6
13	281.00	1.5
14	136.00	4.2
15	218.90	0.5
16	130.20	0.4
17	261.50	3.9
18	172.20	4.1
19	242.20	1.2
20	261.10	2.2

3.2 Artificial neural network (ANN)

The regression model for ANN implementation is shown in Figs 6-9. Fig 6 shows the regression model for training data. The training data compute gradient, estimate and update appropriate weights and biases to capture the pattern/non-linear correlation between the input and output data. The degree of correlation of predicted transmissivity with ANN and measured values of transmissivity was found to be high for the training data with coefficient of correlation (R) that is very close to unity (R = 0.97).

IOP Conf. Series: Earth and Environmental Science **173** (2018) 012025 doi:10.1088/1755-1315/173/1/012025

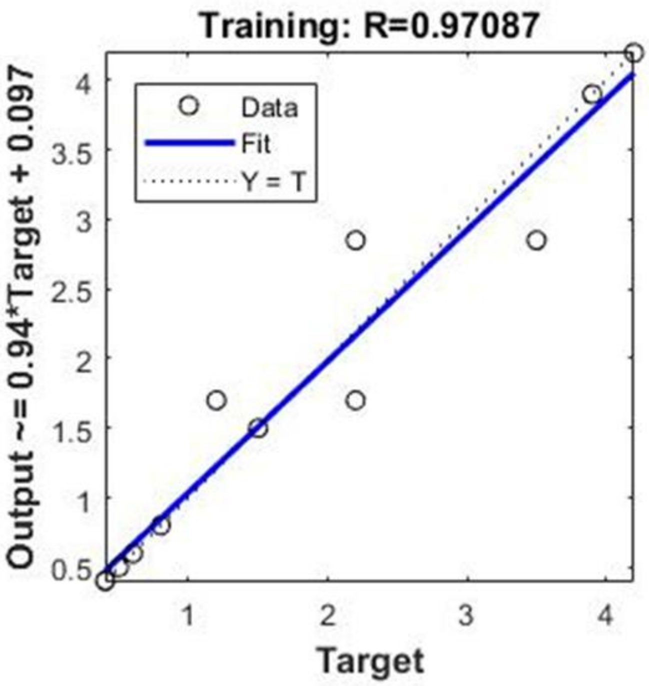

Fig. 6: Regression model for the training network

Fig 7 shows the regression model for validation. This model computes input and output parameters using a subset of the input data. It is an indication of checking the performance of prediction of the model [20-21]. The result of this model gives the correlation coefficient between the predicted and actual transmissivity to be perfect with R of 0.99. The capability of ANN to simulate transmissivity of aquifer was proven on the testing data set.

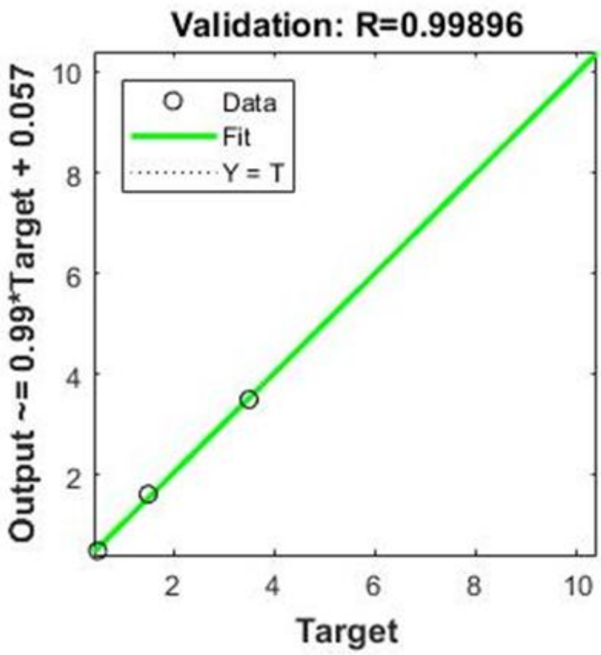

Fig. 7: Regression model for the validation network

Also Fig 8 shows the plot of testing data which has correlation coefficient of 1.0. We therefore compared the values of measured transmissivity with that of predicted T from ANN analysis as seen in Table 3. From the table, there are some deviation of the predicted values of T from the values measured on the field. However, the deviation (percentage error) is within an acceptable limit of less than 5 % [21].

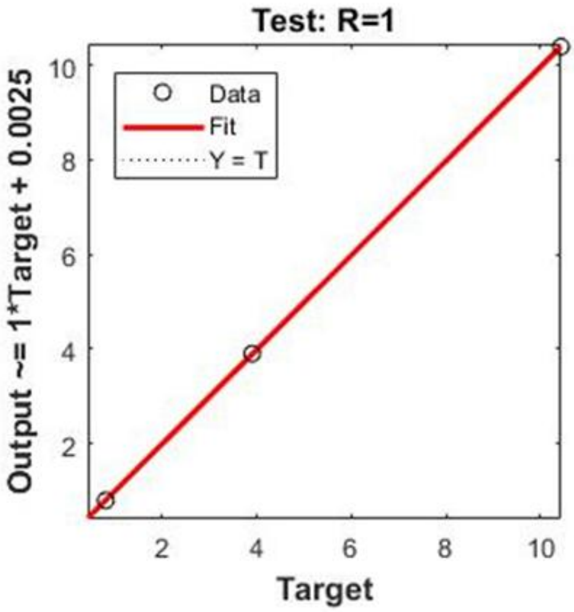

Fig. 8: Regression model for the testing network

Table 3: Measured and Predicted transmissivity

VES Station	Measured Transmissivity (m^2/day)	Predicted Transmissivity (m^2/day)	Error (%)
1	0.80	0.79	+1.25
2	0.60	0.60	0.00
3	1.50	1.51	-0.67
4	0.80	0.80	0.00
5	0.50	0.50	0.00
6	0.40	0.39	+2.5
7	3.90	3.89	+0.26
8	4.10	4.10	0.00
9	2.20	2.20	0.00
10	3.50	3.50	0.00
11	0.80	0.79	+1.25
12	0.60	0.60	0.00
13	1.50	1.50	0.00
14	4.20	4.22	-0.48
15	0.50	0.51	-2.00
16	0.40	0.40	0.00
17	3.90	3.90	0.00
18	4.10	4.10	0.00
19	1.20	1.20	0.00
20	2.20	2.21	-0.45

Moreover, the ‘all’ network phase of the model combined training, testing and validation plots as seen in Fig 9. The network phase shows a very good relationship between the measured and predicted T with coefficient of correlation (R) of 0.94.

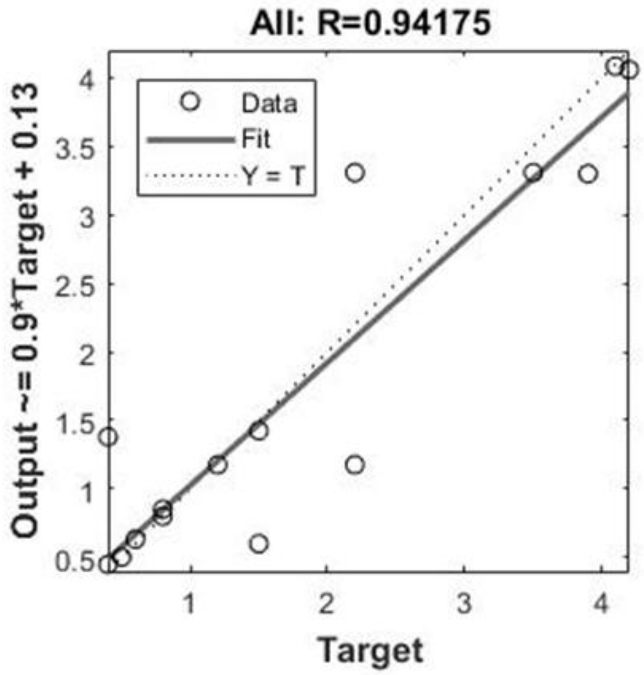

Fig. 9: Regression model for network phase

The plot of the neural network model generated is given in Fig. 10. This plot shows the relationship between the measured transmissivity and the predicted transmissivity using the ANN model developed in this study. The figure shows a very good correlation between the values of our predicted and measured T (R^2 = 1.00).

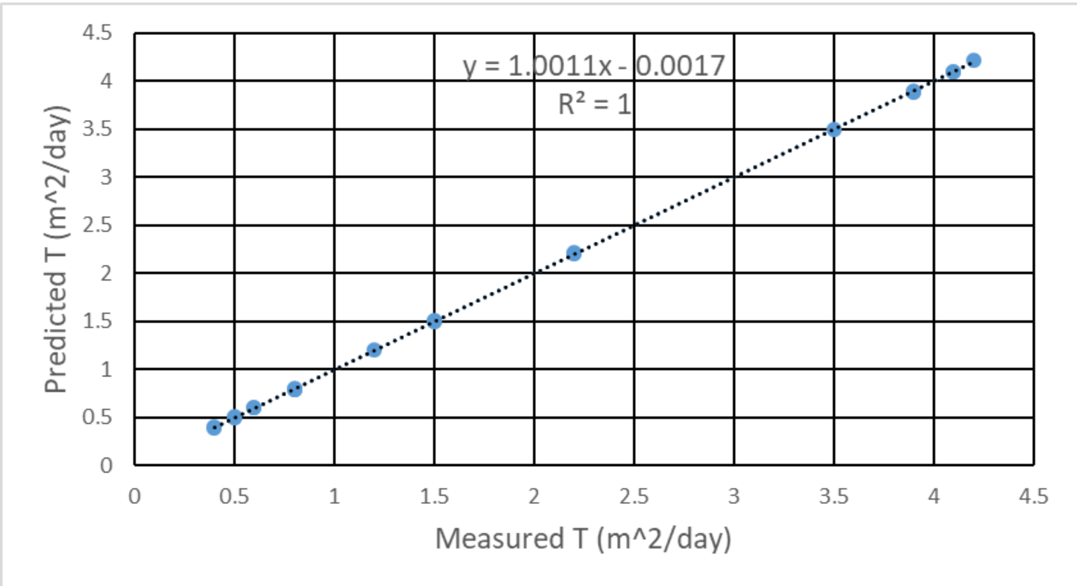

Fig. 10: Plot of the predicted transmissivity versus measured transmissivity

We also estimated root mean square error (RMSE) to examine the effectiveness of the predictive ANN model (equation 9).

$$RMSE = \sqrt{\frac{\sum(y - y')^2}{N}} \quad (9)$$

where y represents the measured parameter, y′ represents the predicted parameters and N is the total data points which is 20 in this study.

The implication of this statistical index is that a good model should have low values of RMSE. However, the calculated RMSE for our model shows a value of 0.085. The low value of the calculated RMSE indicated high prediction performances for our mode

4. Conclusions

In this study, ANN have been used to predict T from geoelectric data in a basement complex area of Southwest, Nigeria. The measurement of the goodness of the fit statistics of our ANN model reveals that the ANN model generated in the study area are reliable in predicting T. The determination of correlation coefficient between measured and predicted T values was 1.0. Also, the RMSE value for the ANN model was found to be 0.085 which indicate high performances of our model. A linear relationship was suggested for the ANN analysis to predict T. This would lead to reduction in cost, effort and time to carry out standard laboratory procedure or pumping test to determine T of aquifer from several discrete points along several transects. However, it is worthy to note that the predicted equation derived by the authors are valid only for understudied aquifer with similar characteristics or geology as the study area. Further studies are needed for the verification of these relationships.

Acknowledgments

The authors thank KFUPM for the provision of valuable software.

References

[1] Chandra S, Kumar D, Ahmed S, Perrin J, Dewandel B (2008)“Contribution of geophysical methods in exploration and assessment of groundwater in hard rock aquifers”. The 3rd International Conference on Water Resources and Arid Environments and the 1st Arab Water Forum.

[2] Mazac O, Kelly WE, Landa I (1985) A hydrogeological model for relations between electrical and hydraulic properties of aquifers. J. Hydrol. 79: 1-19.

[3] Soupios PM, Kouli M, Vallianatos F, Vafidis A, Stavroulakis G (2007) Estimation of aquifer hydraulic parameters from surficial geophysical methods: a case study of Keritis Basin in Chania (Crete-Greece). Journal of Hydrology 338: 122-131. doi:10.1016/j.jhydrol.2007.02.028.

[4] Sri Niwas, Lima OAL (2003) Aquifer parameter estimation from surface resistivity data. Ground Water 41(1): 95-99.

[5] Chand R, Chandra S, Rao VA, Singh VS, Jain SC (2004) Estimation of natural recharge and its dependency on subsurface geoelectric parameters. J. Hydrol. 299, 67-83.

[6] Singh KP (2005) Nonlinear estimation of aquifer parameters from surficial resistivity measurements. Hydol. Earth Syst. Sci. Discuss. 2: 917-938

[7] Chandra S (2006) Contribution of geophysical properties in estimating hydrogeological parameters of an aquifer. PhD, BHU, Varanasi.

[8] Sri Niwas, Singhal DC (1981) Estimation of aquifer transmissivity from Dar-Zarroukparameters in porous media. Journal of Hydrology 50: 393-399. doi:10.1016/0022-1694(81)90082-2.

[9] Ahmed S, Marsily GD, Talbot A (1988) Combined use of hydraulic and electrical properties of an aquifer in a geostatiscal estimation of transmissivity. Ground Water 26(1): 78-86.

[10] Yadav GS, Abolfazli H (1998) Geoelectrical soundings and their relationship to hydraulic

parameters in semi-arid regions of Jalore, Northwestern India. Jour. Appl..Geophys. 39: 35-51.

[11] Ademilua, O. L. (1997) A Geoelectric and geologic evaluation of groundwater potential of Ekiti and Ondo States, Southwestern, Nigeria. Unpublished M.Sc. Thesis, Dept. of Geology, Obafemi Awolowo University, Ile-Ife, Nigeria. pp. 1- 67

[12] Zohdy AAR, Eaton GP, Mabey DR (1974) Application of surface geophysics to Ground-Water Investigation. Techniques of Water Resources Investigation, US Geological Survey.

[13] Marechal JC, Dewandel B, Subrahmanyam K (2004) Use of hydraulic test at different scales to characterize fracture network properties in the weathered-fractured layer of hard rock aquifer. Water Resour. Res. 40 (W11508): 1-17.

[14] Haykin SS (2009) Neural networks and learning machines. Pearson Education Upper Saddle River, vol. 3.

[15] Singh UK, Tiwari R, Singh S (2013) Neural network modeling and prediction of resistivity structures using VES Schlumberger data over a geothermal area. Computers & Geosciences 52: 246–257.

[16] Atkinson PM, Tatnall A (1997) Introduction neural networks in remote sensing. International Journal of Remote Sensing 18(4): 699–709.

[17] Chantasut N, Charoenjit C, Tanprasert C (2004) Predictive mining of rainfall predictions using artificial neural networks for Chao Phraya River. 4th International Conference of the Asian Federation of Information Technology in Agriculture and the 2nd World Congress on Computers in Agriculture and Natural Resources, Bangkok, Thailand, August, pp. 9–12.

[18] Ranganathan A (2004) Thelevenberg-marquardt algorithm. Tutorial on LM algorithm 11(1): 101–110.

[19] Worthington PF (1977) Geophysical investigation of groundwater resources on the Kalibari Basin. Geophysics 42 (4): 838-849.

[20] Gokceoglu C, Zorlu K. (2004) A fuzzy model to predict the uniaxial compressive strength and the modulus of elasticity of a problematic rock. Engineering Applications of Artificial Intelligence 17(1): 61–72.

[21] SanuadeOA**,**AdesinaRB, AmosunJO, FajanaAO, Olaseeni OG (2017) Using artificial neural network to predict dry density of soil from thermal conductivity. RMZ – M&G 64:.1–12.