Prediction of Permeability Using Group Method of Data Handling (GMDH) Neural Network from Well Log Data

Baraka Mathew Nkurlu [1,2], **Chuanbo Shen** [1,2,*], **Solomon Asante-Okyere** [1,2], **Alvin K. Mulashani** [2,3], **Jacqueline Chungu** [2] **and Liang Wang** [1,2]

[1] Key Laboratory of Tectonics and Petroleum Resources, Ministry of Education, China University of Geosciences, Wuhan 430074, China; barakankurlu@hotmail.com (B.M.N.); pseulo@cug.edu.cn (S.A.-O.); wangl@cug.edu.cn (L.W.)

[2] Department of Petroleum Geology, Faculty of Earth Resources, China University of Geosciences, Wuhan 430074, China; alvinmulashani@yahoo.com (A.K.M.); jaychungu85@yahoo.com (J.C.)

[3] Department of Geoscience and Mining Technology, College of Engineering and Technology, Mbeya University of Science and Technology, Mbeya 00225, Tanzania

* Corresponding author: cbshen@cug.edu.cn

Abstract: Permeability is an important petrophysical parameter that controls the fluid flow within the reservoir. Estimating permeability presents several challenges due to the conventional approach of core analysis or well testing, which are expensive and time-consuming. On the contrary, artificial intelligence has been adopted in recent years in predicting reliable permeability data. Despite its shortcomings of overfitting and low convergence speed, artificial neural network (ANN) has been the widely used artificial intelligent method. Based on this, the present study conducted permeability prediction using the group method of data handling (GMDH) neural network from well log data of the West arm of the East African Rift Valley. Comparative analysis of GMDH permeability model and ANN methods of the back propagation neural network (BPNN) and radial basis function neural network (RBFNN) were further explored. The results of the study showed that the proposed GMDH model outperformed BPNN and RBFNN as it achieved R/root mean square error (RMSE) value of 0.989/0.0241 for training and 0.868/0.204 for predicting, respectively. Sensitivity analysis carried out revealed that shale volume, standard resolution formation density, and thermal neutron porosity were the most influential well log parameters when developing the GMDH permeability model.

Keywords: permeability; group method of data handling; artificial neural network; well logs; sensitivity analysis

1. Introduction

Petrophysical properties such as permeability play an important role in understanding the subsurface characteristics of a reservoir, in order to estimate oil and gas reserves available. To date, it has become a difficult task to fully capture and make accurate estimations of the complex heterogeneity of permeability in the petroleum industry. Permeability, by definition, is the capability of the rock to convey fluid under specific pressure difference [1–5] and it is reliant on the interaction of fluid, pore, lithology, and porosity composition [6]. Reservoir permeability is either obtained from direct measurement of core and also well testing analysis or from empirical equations of well log parameters. On the contrary, a direct measurement from core and well test have shown to be costly and time-consuming method if carried out on all wells while empirical equation obtained from well logs cannot fully account for the heterogeneity of permeability in all reservoir conditions [7,8].

In line with this, numerous computational intelligent methods have been successfully adopted to model reservoir permeability. The extensively applied intelligent method in permeability prediction with improved accuracy over that of statistical methods is the artificial neural network (ANN) [3,9]. This is as a result of its ability to discover patterns in non-linear and complex data systems [10–13]. However, standard ANN is known to exhibit several drawbacks such as overfitting, converging at local minima, and low computational speed. In addition, training of neural network models requires the input of user-defined parameters. When training a standard neural network, the number of the hidden neurons and hidden layers (or spread parameter) are specified by the user through trial and error means. It is necessary to note that the execution of these neural network models is based on network structure [14–16].

Therefore, several attempts to improve the standard ANN have been developed [17–21]. In an attempt to address the issue of model bias faced by ANN, group method of data handling (GMDH), automatically synthesize network from the database of inputs and outputs. This process which is called self-organization of input models removes the user from specifying the network architecture in advance and hence removes biases in the model [22–24]. Reduced computational time is another advantage of the GMDH method as compared to the standard ANN. This is made possible by short analyses in the model synthesizing process. The algorithms for the GMDH were first developed by Professor Alexey Grigorevich Ivakhnenko in 1968 with the goal of identifying the affiliation between the involvement layer and output layer in nonlinear systems [25–27]. According to Ayoub et al. [28], GMDH may also be represented as a polynomial neural networks as well as an algorithm modeling method to define nonlinear input-output variable relationships. Ma et al. [29] explained that GMDH algorithm operates by linking sets of neurons connected to quadratic polynomials, resulting in new sets of neurons in the subsequent layer.

Over the years, the GMDH concept has been widely applied in different areas, Menad et al. [30] applied GMDH and gene expression programming (GEP) to establish reliable correlations for estimating temperature-based oil-water relative permeability. Shen et al. [31], developed a GMDH model that can detect the various lithofacies using pre-processing techniques composed of dimensionality reduction (DR) and wavelet analysis (WA). Ayoub et al. [28] in their publication, utilized GMDH to model the oil compressibility below the bubble point pressure. Teng et al. [32] proposed a GMDH model to predict China's transport energy demand, the proposed model showed very satisfactory predictions. Nguyen et al. [33] proposed the GMDH network for geometrically nonlinear problems of computational mechanics. Safari et al. [34] used the generalized structure of GMDH to model sediment transport in rigid boundary open channels. Rostami et al. [35] in their study used a GMDH for estimating the ionic liquids heat capacities. GMDH algorithms have adopted three dissimilar types of elementary functions such as probabilistic graphs, second-order polynomials, and Bayes' formulas [36–38].

The fast GMDH theory development which includes the comprehensive spectrum of algorithms has led to, unlike the category approach in which methods of GMDH can be clustered into parametric algorithm category and non-parametric algorithm category [39].

Parametric algorithms can be classified in accordance with the nature of activation function such as partial descriptions which is a set of characteristics by which it can be recognized or the model complexity structure type. Single-layer self-organizing algorithms which also known as combinatorial algorithms, perform and achieve fully comprehensive quest among all models. Multilayer/iterative algorithms employ an iterative procedure, that increases the complexity of the model whereas an exterior criterion classifies these models to be progressed after developed to the following next predetermined layer. In the case of multilayer algorithms, no fully comprehensive quest of entire aspirant models nevertheless the time of computation is abridged and the independent amount of variables in the line of being processed turn out to be larger. Regarding the nature of making it active and effective function, the algorithms of the GMDH network can be illustrious into harmonic, multiplicative-additive, fuzzy and polynomials [40,41].

In this study, GMDH was proposed for the prediction of permeability using well log data. To further ascertain the performance of the GMDH permeability model, its results were compared with

ANN algorithms of back propagation neural network (BPNN) and radial basis function neural network (RBFNN). Finally, a sensitivity analysis was carried out to assign the impact of each well log parameter used to develop the GMDH model.

2. Data and Methods

2.1. Data Descriptions

East African Rift Valley (Gregory Rift) comprises of the major Ethiopian Rift, which runs in the eastward that from an Afary Triple Junction, the valley prolonged its wonder which continues to the south which is called Kenyan Rift Valley [42]. Spreading its branches the western Rift Valley contains Albertine Rift, beyond its part of South with the valley of Lake Malawi which is found in the south-west of Tanzania. This rift Valley follows one of known two paths to the North of the so-called Afar Triple Junction, while in the East to Ridge of Aden found in Gulf of Aden and West is direct to Red Sea Rift [43,44].

In the present study, we considered three wells Mpyo 1, Mpyo 2, and Mpyo 3 from Mpyo field located in the Northern part of the Lake Albert rift basin. The field is aligned along the Eastern Graben Trend in the northern part of Albertine Graben and extends from the Victoria Nile area towards the North Eastern extremity of the basin margin (Figure 1). The depositional environment is the Victoria Nile Delta, where sediments consist of alternating fluvial/deltaic sands and lacustrine shales. The study area has five formations, Alluvial Sands/ Semanya which comprises alluvial sands with some claystone interbeds, Upper Paraa Formation which continued from the sands of the alluvium and Semanya, Lower Paraa Formation, which is marked with the appearance of the claystone. Pacego Formation, the top of this formation is marked by coarsening upward sand. Below the sand is predominantly claystone with some sand/silt interbeds and Wangkwar Formation, which contains the reservoir sands interbedded with claystones and siltstones. The stratigraphic diagram of the studied formations is described in Figure 2.

Well log and permeability data of 366 from Mpyo 1 and Mpyo 2 wells were used to train the model. Afterward, the trained models were tested on the well log and permeability data of 280 from Mpyo 3 well. Well log inputs of standard resolution formation density (RHOZ), apparent resistivity focusing mode 5 (RLA5), standard gamma ray (SGR), apparent resistivity focusing mode 1(RLA1) thermal neutron porosity (TNPH), rock's shale volume (VSH) as represented in (Figure 3) were selected and used in development of the GMDH and ANN models. Apart from other well logs, the VSH is a dependent log calculated from the gamma-ray log. The statistical analysis of the data used for this research is summarized in Table 1.

The entire well log and permeability data were normalized within the range of 0 and 1 using the min-max normalization approach to ensure that the artificial intelligent models treat and handle the input and output data in an unbiased manner.
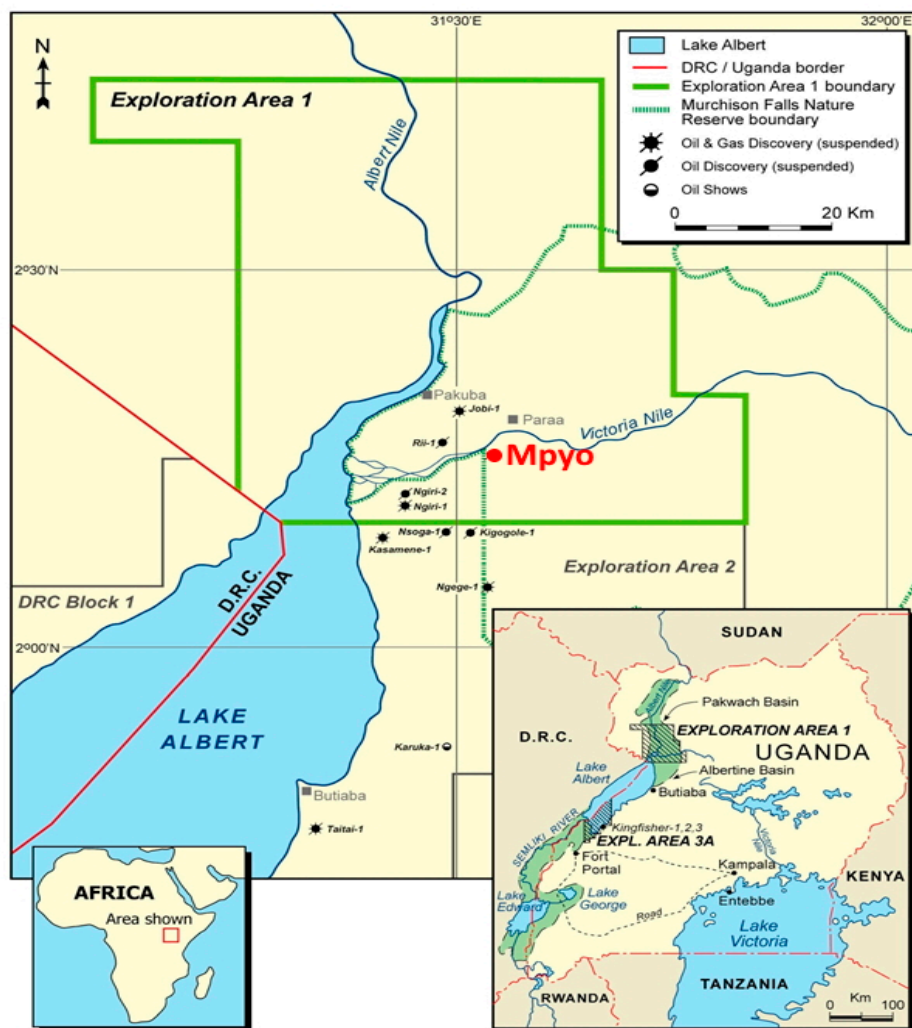
Figure 1. Location of Mpyo oil field found in the East African Rift Valley.

Table 1. Statistical analysis of the data used.

| Wells		RHOZ	RLA1	RLA5	SGR	TNPH	VSH
Mpyo 1	Max	3.4864	34.759	78.1588	72.7292	1.0351	0.9113
	Min	1.9273	0.0867	0.0558	13.6988	0.2521	0.0142
	Average	2.1685	8.021	11.4722	42.1452	0.3757	0.3094
	STDEV	0.1284	5.7368	11.4411	15.5787	0.1017	0.2408
Mpyo 2	Max	4.097	139.3818	125.6737	74.8748	1.6155	0.9121
	Min	1.6833	0.0375	0.0277	16.2123	0.1231	0.0198
	Average	2.481	4.6441	4.9643	44.5679	0.7406	0.3462
	STDEV	0.5874	17.2156	17.9904	11.2071	0.3482	0.219
Mpyo 3	Max	3.4877	61.7264	83.4196	60.8802	1.0336	0.8132
	Min	1.8848	0.0084	0.0057	8.4259	0.1788	0.0016
	Average	2.2449	6.0888	7.516	34.0213	0.3347	0.2184
	STDEV	0.1794	5.1768	7.2833	10.8559	0.1288	0.1515
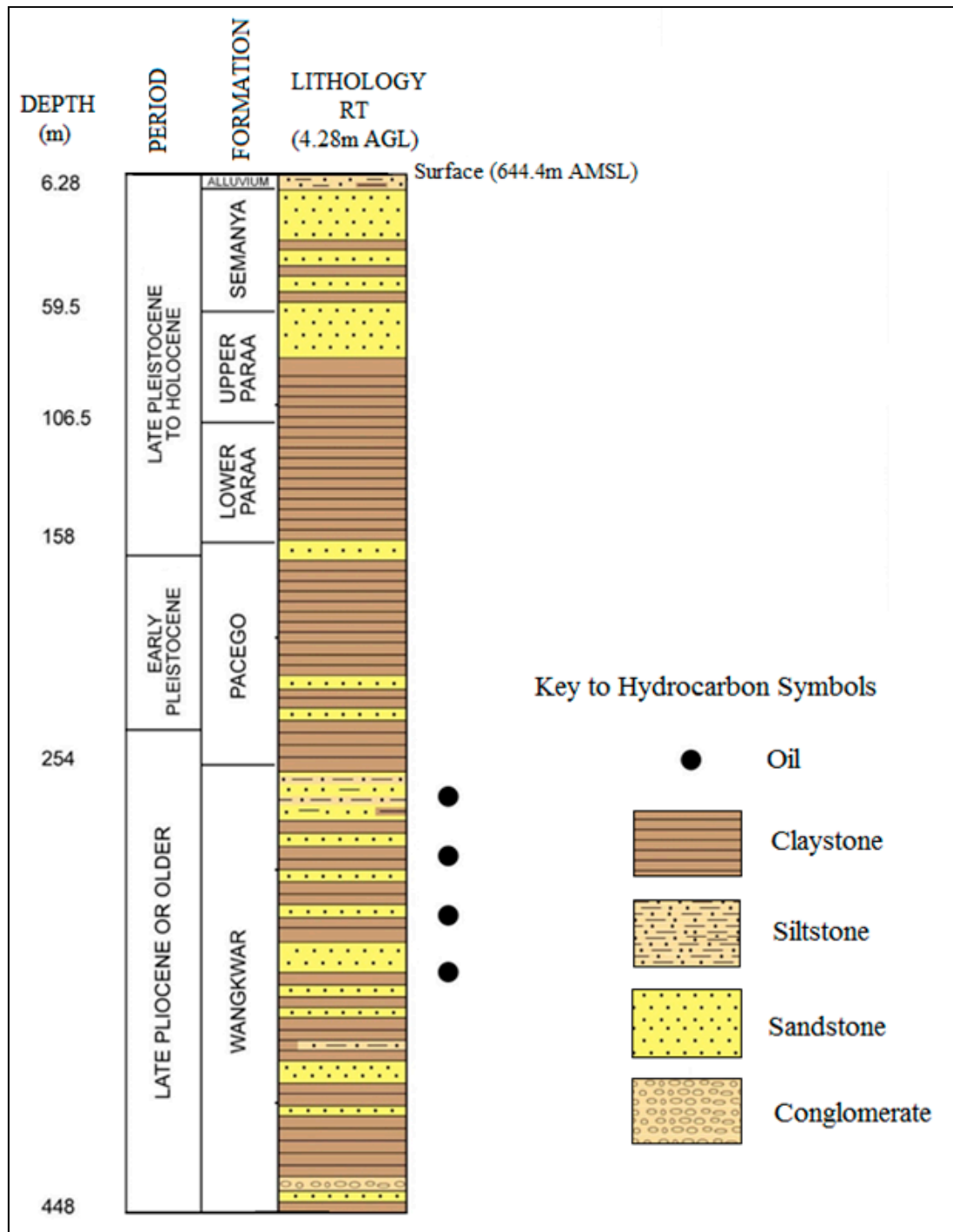
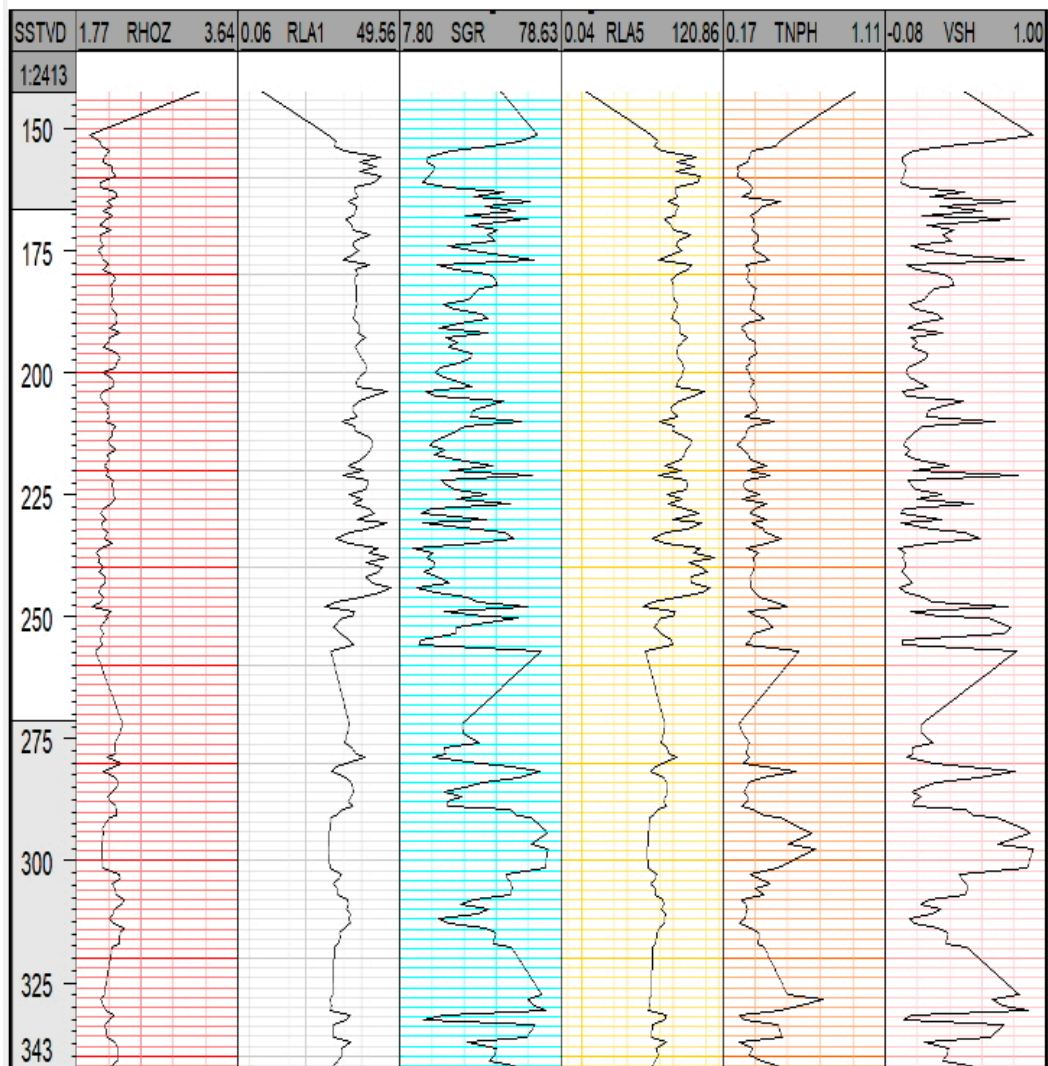**Figure 2.** Stratigraphic diagram of the study area.

Figure 3. Sample geophysical well logs of standard resolution formation density (RHOZ), apparent resistivity focusing mode 5 (RLA5), standard gamma-ray (SGR), apparent resistivity focusing mode 1(RLA1), thermal neutron porosity (TNPH), rock's shale volume (VSH).

2.2. Group Method of Data Handling (GMDH) Model

The combination of multi-layer algorithms in which a network of nodes and layers is engendered via numbers of selected input from the set of designed data being assessed is known as Group Method of Data Handling (GMDH). The idea of this type of neural network was first presented by G. Ivakhnenko during the 1960s as a method of identifying and recognizing nonlinear relationships between the variables of input and the output variables [17,22].

The GMDH neural networks algorithms apply the benefits of both self-organizing standards and multilayer neural networks to choose the best associations between variables. Firstly, GMDH algorithm automatically discovers the inter-relations between variables and selects the best model to fit data. Secondly, the GMDH neural network, such as artificial neural network, combines the concepts of black-box, biological neuron technique, inductive technique, probability concept, and many other approaches [45]. When generating a GMDH network, a set of combinatorial inputs is developed and wholly passed into the first layer of the network. This layer generates outputs which are classified and carefully selected to give another set of combinations as inputs into the succeeding layer [46]. This process persists until when results from a succeeding layer $(k + 1)$ are no longer good as compared to results from its predecessor, layer (k).

The GMDH builds general linking between input and output parameters mathematically, which can be also termed as a reference point. The model corresponding to the minimum is the required nonphysical model, whose correctness is estimated by the criterion of the estimating error variation as follows

$$E_i^2 = \frac{\sum_1^N (y_i - Y_i)^2}{\sum_1^N (y_i - \overline{y})^2} \rightarrow \min,$$

(1)

Where: y_i is the actual parameter value, the value of the predicted parameter is given by Y_i, and $\overline{y}$ is the mean value of the predicted parameters, *E* can often be called relative root mean square error. With a minimum value *E* and with extended but always same samples, we can apply a deductive method, i.e., we can select best physical models by an internal degree in each group and entrust the ending of the iteration to an expert [47].

The network connection among the input variables and output of several multiple inputs with the only single output of self-organizing neural networks are formulated and expressed through the complex discrete formula in form of Volterra functional polynomial series known as Kolmogorov–Gabor (VKG) of the form [40,46].

$$y = \sum_{i=1}^M a_i x_i + \sum_{i=1}^M \sum_{j=1}^M a_{ij} x_i x_j + \sum_{i=1}^M \sum_{j=1}^M \sum_{k=1}^M a_{ijk} x_i x_j x_k + a_0$$

(2)

where $A = (a_0, a_i, a_{ij}, a_{ijk})$ is a vector of weights or coefficients, *y* is a consistent output vector value, and $X = (x_1, x_2, x_3, x_M)$ is the input variables vector. The full mathematical explanation can be denoted by the polynomial equation. The system of entire equations can be denoted and represented in the form of matrix *X* in Equation (3).

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & \cdots & x_{1M} \\ x_{21} & x_{22} & \cdots & \cdots & x_{2M} \\ \cdots & \cdots & \cdots & x_{ij} & x_{iM} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{N1} & x_{N2} & \cdots & \cdots & x_{NM} \end{bmatrix}, y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{N1} \end{bmatrix}.$$

(3)

For the sake of easiness Equation (2) can be substituted by the system of a partial polynomial as shown by Equation (4)

$$y = a_5 x_j^2 + a_4 x_i^2 + a_3 x_i x_j + a_2 x_j + a_1 x_i + a_0,$$

Where $j, i = 1, 2, 3...., M; j \neq i$

(4)

The GMDH neural network algorithms track some procedures and steps systematically (Figure 4) to model the intrinsic connection between parameters of inputs and targeted output [48]. Data sample given by *N* observations and *M* autonomous variables are needed; the data set are divided into predicting data and training data.

A new regression polynomial is similar in representation as that in Equation (4) is initially created by taking combinatorial possibilities of all independent variables of Equation (3) (matrix equation represented by *X*) as a couple of two in time and assume *q* and *p* as columns of the matrix *X*.

$$y_i = a_{qp} + b_{qp}x_{ip} + c_{qp}x_{iq} + f_{qp}x_{ip}x_{iq} + e_{qp}x_{ip}^2 + d_{qp}x_{iq}^2,$$

Here,

$$\begin{cases} p = 1,2,3,.....,M, p \neq q, \\ q = 1,2,3,.....,M, p > q, \\ i = 1,2,3.......,N, \end{cases} \tag{5}$$

The data for training set A and Equation (5) are employed in calculating a set of constants of regression for wholly partial functions by parameter approximation technique [49]. A new matrix U stores the newly calculated regression constants (see Equation (6)).

$$U_i = a_{qp} + b_{qp} + c_{qp} + d_{qp} + e_{qp} + f_{qp}, \begin{cases} i = 1,2,3.......,N, \\ q = 1,2,3,.....,M, q < p, \\ p = 1,2,3,.....,M, q \neq p, \end{cases} \tag{6}$$

Mathematically, the number of combinations (c) of pairs of input is determined by

$$\text{Number of combinations (c)} = \frac{M}{2}(M-1) \tag{7}$$

The polynomial at each data points N is assessed in order to compute new estimation $y_{i,\,qp}$ as follows:

$$y_{i,pq} = a_{pq} + b_{pq}x_{ip} + c_{pq}x_{iq} + f_{pq}x_{ip}x_{iq.} + d_{pq}x_{ip}^2 + e_{pq}x_{iq}^2 \tag{8}$$

In an iterative manner, the procedure will be repeated until entire pairs are estimated to produce the novel pairs of regression which would be kept into another novel matrix termed as Y matrix in Equation (10). These newly regression pairs generated could be taken as the newly enhanced variables which have the best predictability compared to that of data set X (Equation (3))

$$Y = \{y_{ji}\}, \begin{cases} j=1,2,3,.....,\frac{M}{2}(M-1), \\ i=1,2,3,.....,N, \end{cases} \tag{9}$$

$$Y = \begin{bmatrix} y_{11} & y_{12} & \cdots & \cdots & y_1, & \frac{M}{2}(M-1) \\ y_{21} & y_{22} & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & y_{ji} & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{N1} & y_{N2} & \cdots & \cdots & y_N, & \frac{M}{2}(M-1) \end{bmatrix}. \tag{10}$$

Quality procedures of these functions are evaluated objectively in accordance to the law adopted via a set of predicting data and this can be completed by individually comparing each column of the newly created matrix Y with that dependent variable y_i.

Therefore, this whole process is repeated up to when error variation is kept smaller than the error obtained in the previous layer. A GMDH-model for the designed data sets is calculated by backtracking entire route of polynomials which matches to the lowest relative root mean square error and under each and every layer. Figure 4 is the schematic diagram which illustrates a self-organizing GMDH algorithm.
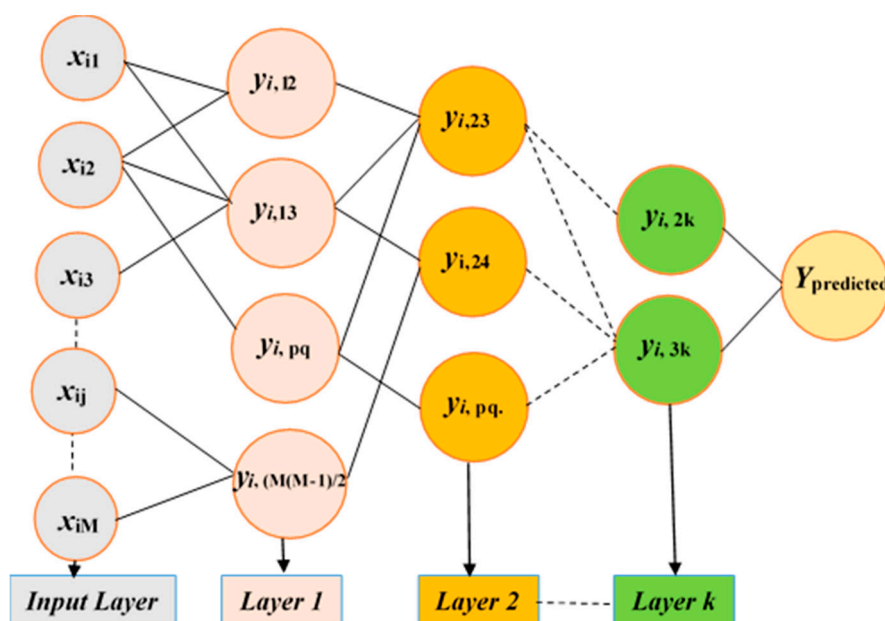
Figure 4. Network of self-organizing group method of data handling (GMDH) algorithm with *M* inputs and *k* layers.

In this present paper, selection of the input parameters follows the function of the GMDH-type network, which enhances the ability of filtering variables. The steps involved in the process include:

Step 1. The well log and core permeability data for the training and predicting were standardized. The uses of the training data are to estimate the weights of GMDH neurons whereas the predicting data are used to evaluate the network architectures.

Step 2. Evaluate the regression polynomial by means of Equation (4), for individually couple of input variables which are x_i and x_j with the corresponding output y of training data set that fits the rely on remarks *M* in the training data set

Step 3. The polynomial for entire observations of *N* for an individual regression is computed. New observations N are stored into different matrix Y. By using a similar method, the new columns of *Y* are computed. Matrix *Y* can be understood as novel enhanced variables with improved predictability than those variables from the original generation of x_1, x_2, x_3,....., x_M.

Step 4. The fourth step is screening out the last variables which are effective.

Step 5. According to the increasing of regularity criterion, order the columns of *Y*, then choose the *Y* columns which satisfies regularity criterion <S where S is a minimum residual value prescribed by the user in order to substitute that previous original column of matrix X.

Step 6. Steps 1 to 5 are repeated until the final estimates are automatically determined when the least error is achieved based on the metaheuristic and self-organizing nature of GMDH algorithm. Plot and compare the smallest estimated error calculated in each generation with the smallest estimating error of the current generation until it starts to give an increasing trend. All of the steps have been summarized in the flow-chart shown in Figure 5.
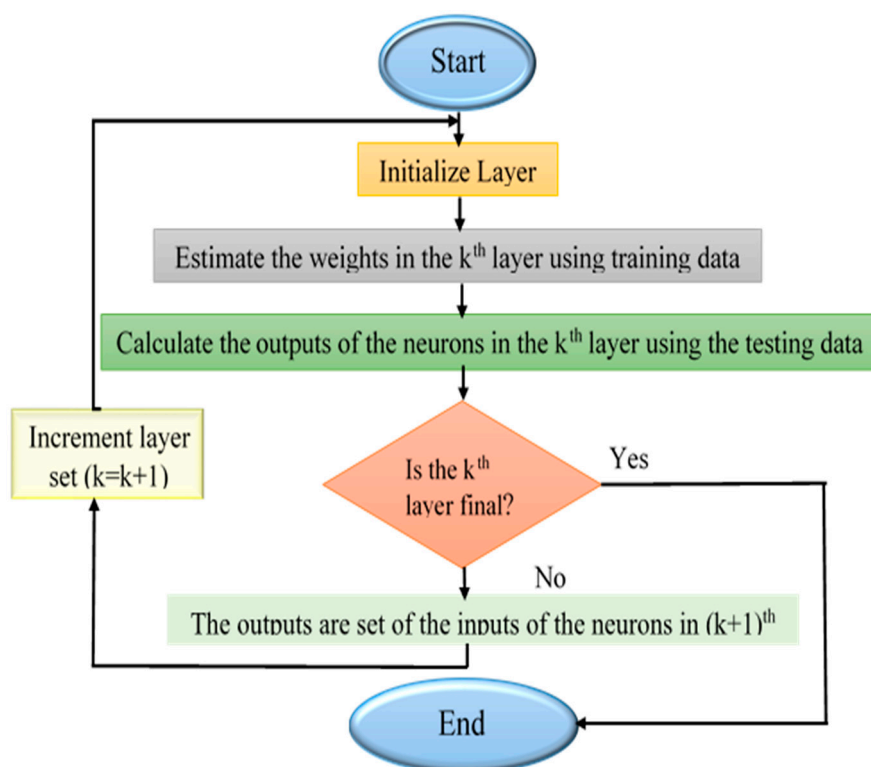
Figure 5. GMDH network flow chart.

3. Results and Discussion

3.1. Performance Indicators

The statistical parameters such as correlation coefficient (R) and root mean square error (RMSE) was adopted to compare the performance of the GMDH and ANN permeability models. The mathematical expression for *R* and RMSE are provided in Equations (11) and (12), respectively [50] [51].

$$R = \frac{\sum\limits_{i=1}^{N}(y_i - \overline{y})(Y_i - \overline{Y}_i)}{\left(\sqrt{\sum\limits_{i=1}^{N}(y_i - \overline{y}_i)^2}\right)\left(\sqrt{\sum\limits_{i=1}^{N}(Y_i - \overline{Y}_i)^2}\right)} \tag{11}$$

$$RMSE = \sqrt{\frac{1}{N}\sum\limits_{i=1}^{n}(y_i - Y_i)^2} \tag{12}$$

Where y_i is the actual or measured parameter value, *N* represents the overall number of data points, $\overline{y}_i$ is the average value of measured parameters, the value of the predicted parameter is denoted by Y_i and $\overline{Y}_i$ is the average value of the predicted parameters,

3.2. GMDH Model Development

The GMDH model was implemented in MATLAB R2016b. The output of the developed GMDH consisted of seven hidden layers and one output layer. It was identified that every hidden layer was made up of one hidden neuron which has been labelled (z_1, z_2, z_3,..., z_7) and *Y* for the output layer. As explained early, these unitless hidden neurons are only statistically responsible for correlating an

input layer and the output layer. This configuration was accomplished after a sequence of adjusting progressions by monitoring an actual performance of the GMDH network until the accomplishment of optimum neural network structure. The schematic diagram shown in Figure 6 depicts in detail the GMDH permeability model developed. The equations for the layers of the GMDH model network that were needed in order to generate permeability prediction are be given in Equations (11)–(18). The GMDH model developed has eight (8) total number of layers:

Layer #1

　　Number of neurons: One *neuron* (z_1)

$$z_1 = 116.22 - 630.4x_6 + 636.597x_5 - 730.37x_5x_6 + 603.017x_6^2 - 81.896x_5^2 \quad (13)$$

Layer #2

　　Number of neurons: One *neuron* (z_2)

$$z_2 = 45.07 + 1.698z_1 - 464.07x_1 - 6.078x_1z_1 + 0.006z_1^2 + 1100.198x_1^2 \quad (14)$$

Layer #3

　　Number of neurons: One *neuron* (z_3)

$$z_3 = -38.874 + 0.818z_2 + 87.99x_4 + 0.802x_4z_2 - 0.00011z_2^2 - 24.597x_4^2 \quad (15)$$

Layer #4

　　Number of neurons: One (*neuron z_4*)

$$z_4 = -27.04 + 1.1677z_3 + 170.033x_6 - 0.556x_6z_3 - 8.174e - 05z_3^2 - 200.529x_6^2 \quad (16)$$

Layer #5

　　Number of neurons: One *neuron* (z_5)

$$z_5 = -9.4475 + 1.359z_4 + 31.066x_1 - 1.494x_1z_4 - 9.23607e - 06z_4^2 - 8.8648x_1^2 \quad (17)$$

Layer #6

　　Number of neurons: One *neuron* (z_6)

$$z_6 = 27.0166 + 0.807z_5 - 132.229x_4 + 0.539x_4z_5 - 3.7408e - 05z_5^2 + 142.044x_4^2 \quad (18)$$

Layer #7

　　Number of neurons: One *neuron* (z_7)

$$z_7 = -11.683 + 1.0684z_6 + 73.8461x_5 - 0.3010x_5z_6 + 0.00015z_6^2 - 70.7582x_5^2 \quad (19)$$

Layer #8 (Output layer)

　　Number of neurons: One *neuron* (Y)

$$Y = -0.418 + 1.0092z_7 + 2.662x_3 - 0.759x_3z_7 - 8.084e - 06z_7^2 + 351.589x_3^2 \quad (20)$$

where x_1 is standard resolution formation density (RHOZ), x_3 is standard gamma ray (SGR), x_4 is apparent resistivity focusing mode 5 (RLA5), x_5 is thermal neutron porosity (TNPH) and x_6 is rock's shale volume (VSH) which is a dependent log calculated from the gamma ray log.
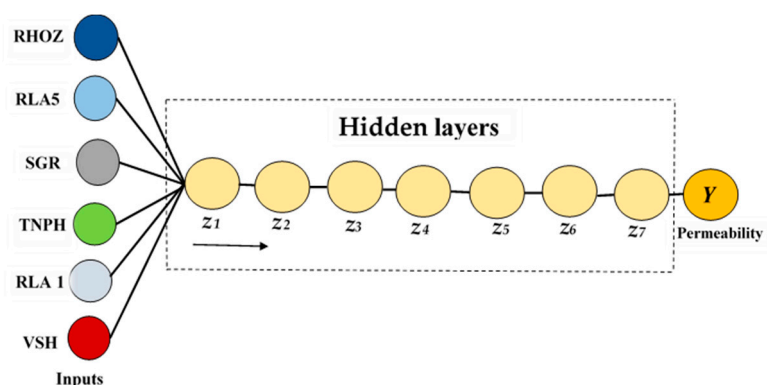
Figure 6. Illustration of the proposed GMDH-type neural network.

3.3. Comparing with ANN

In this section, the results of GMDH are presented with other neural network algorithms. Back propagation neural network (BPNN) and radial basis function neural network (RBFNN) were further subjected to the same training and prediction dataset 366 and 280, respectively. As a result of the non-linearity of the dataset, the hidden and output layers of BPNN used the hyperbolic tangent sigmoid and linear transfer function. BPNN models were trained using the Levenberg–Marquardt learning algorithm [32] while the gradient descent learning algorithm was to train RBFNN [33]. It is important to note that the performance of BPNN and RBFNN are highly dependent on the model structure which was achieved in this study through a sequential trial and error method.

In determining permeability, the optimal BPNN model structure that obtained the highest correlation and the least RMSE was made up of 3 hidden nodes i.e., 6 inputs, 1 hidden layer with 3 hidden nodes, and 1 output as shown in Table 2. While the optimal RBFNN consisted of 6 inputs, a hidden layer with a maximum of 21 hidden nodes having a width parameter of 1 and 1 output as shown in Table 3.

Table 2. Results from tuning back propagation neural network (BPNN) structure.

Hidden Neuron	R		RMSE	
	Train	Test	Train	Test
1	0.618	0.618	0.416	0.428
2	0.985	0.822	0.025	0.206
3	**0.982**	**0.823**	**0.031**	**0.205**
4	0.992	0.611	0.02	0.445
5	0.983	0.813	0.032	0.214
6	0.978	0.824	0.034	0.209
7	0.976	0.822	0.037	0.208

Table 3. Results from tuning radial basis function neural network (RBFNN) structure.

Spread Parameter	R		RMSE	
	Train	Test	Train	Test
0.3	0.973	0.68	0.034	0.407
0.5	0.97	0.696	0.034	0.41
0.7	0.981	0.809	0.03	0.211
0.9	0.982	0.821	0.03	0.208
1	**0.988**	**0.822**	**0.026**	**0.206**

Table 4 shows the obtained results of the performance indices for all the developed permeability models. All the results of *R* and RMSE for both training and predicting revealed that GMDH displayed an improved model performance than ANN models of BPNN and RBFNN. When training, GMDH produced R and RMSE values of 0.989 and 0.0241, respectively, which were better than BPNN and RBFNN as summarized in Table 4. However, BPNN and RBFNN attained 0.982/0.0313 and 0.988/0.026 as *R*/RMSE values, respectively. The predicting process also saw GMDH generate significantly better results of 0.868 and 0.204 as *R* and RMSE respectively (Table 4). From Table 4, we can observe that the generalization capability of BPNN and RBFNN permeability models were comparatively worse as they achieved *R*/RMSE score of 0.823/0.2053 and 0.822/0.206 respectively.

Depending on the computational speed, the GMDH produced the results using a computational time of 1.44seconds, while BPNN and RBFNN produced results in a time of 13.86 s and 6.45 s respectively (Table 4) on a Windows machine with AMD Ryzen 5 @ 2 GHz. The performance of the GMDH and ANN predictive models when compared with the measured permeability data are presented in Figures 7 and 8.

Table 4. Statistical indicators of the developed permeability models.

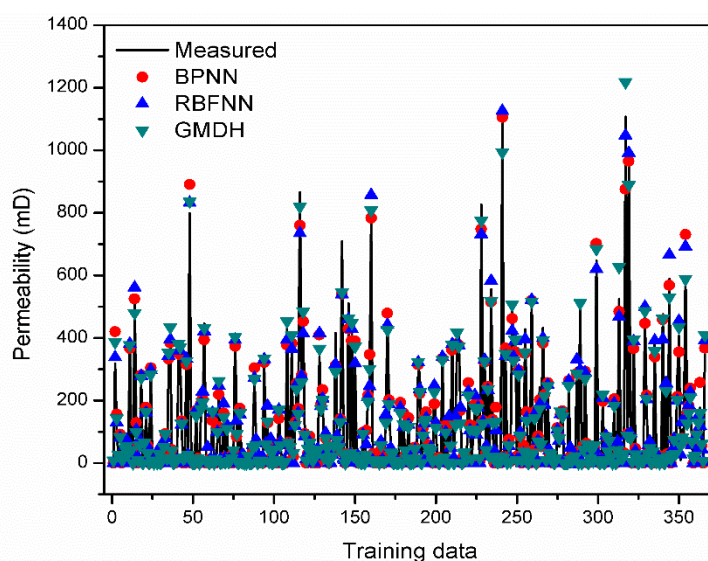|	Training		Predicting		
Model	**R**	**RMSE**	**R**	**RMSE**	**Computational Time**
GMDH	0.989	0.0241	0.868	0.204	1.44 s
BPNN	0.982	0.0313	0.823	0.2053	13.86 s
RBFNN	0.988	0.026	0.822	0.206	6.45 s

Figure 7. GMDH and artificial neural network (ANN) prediction in comparison with measured permeability data during training.
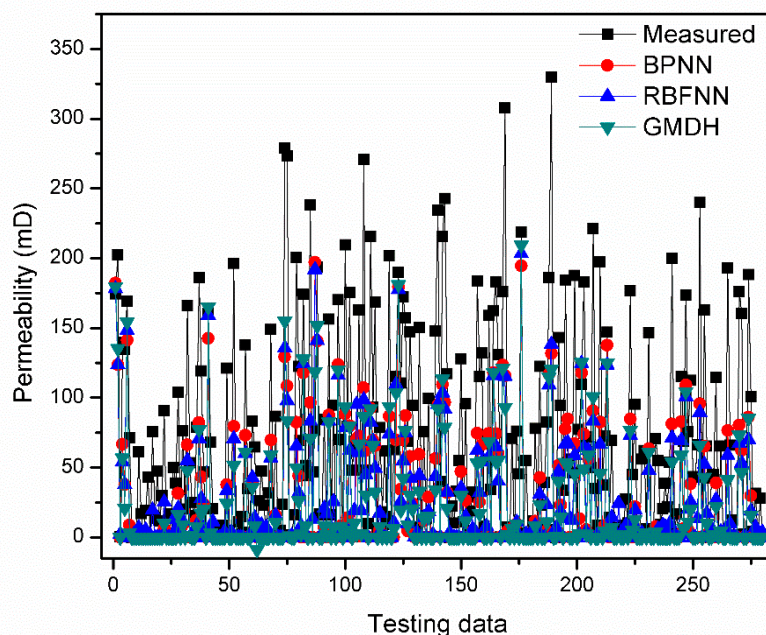
Figure 8. GMDH and ANN prediction in comparison with measured permeability data during training.

3.4. Sensitivity Analysis for the GMDH Model

Sensitivity analysis will examine the contribution made by each input well log in determining permeability values. Sensitivity analysis was carried out for every input well log used to develop the GMDH permeability model. Sensitivity analysis of the well log input data was conducted using Equation (22):

$$S = \frac{1}{N} \sum_{i=1}^{N} \left(\frac{percent\ change\ in\ output}{percent\ change\ in\ input} \right)_i \times 100 \tag{20}$$

where the percentage change in output and input means the variation of parameters from the minimum to the maximum value. It is significant to note that when estimating permeability, GMDH automatically selected five out of the six input well logs. These were RHOZ, RLA5, SGR, TNPH, and VSH. It is, therefore, important to identify how each of the selected well log input contributed to the development of the GMDH permeability model. The higher the *S* value, the more that parameter of well log affected the predicted outcome of permeability. The lower the *S* value means that input value cannot heavily affect the generated permeability value.

From Figure 9, RHOZ observed great contribution on permeability prediction with an *S* value of 5.787%. VSH and TNPH similarly had a great impact of 4.14% and 2.143%, respectively. Both SGR and RLA5 had less impact on permeability prediction model than all parameters with an *S* value of 0.1% and 0.04%, respectively.
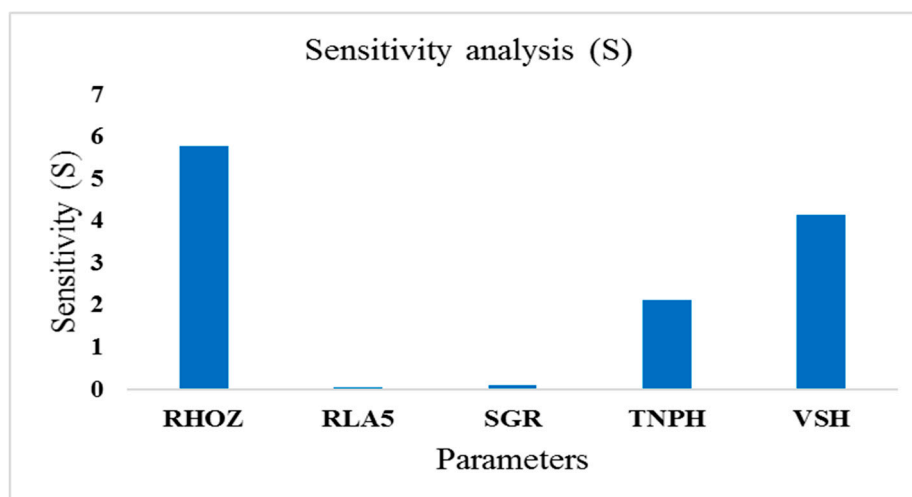
Figure 9. Sensitivity analysis of GMDH input values.

4. Conclusions

The purpose of this study was to explore the potential of the application of the group method of data handling (GMDH) neural network in estimating the permeability using well log data. Based on this, well log and core permeability data of the West arm of the East African Rift valley were considered for the present research.

The optimal GMDH permeability model was identified to be composed of 6 inputs and 7 hidden layers each having one node. The statistical performance of the developed GMDH model was better when compared with ANN algorithms of back propagation neural network (BPNN) and radial basis function neural network (RBFNN). During training, GMDH produced estimates having R/RMSE score of 0.989/0.0241 while BPNN and RBFNN recorded 0.982/0.0313 and 0.988/0.026, respectively. Additionally, GMDH generalized better than the ANN models as it generated R/RMSE of 0.868/0.204 as compared to 0.823/0.2053 and 0.822/0.206 for BPNN and RBFNN, respectively.

The outcome of the sensitivity analysis on the automatically selected well log input parameters revealed that shale volume, standard resolution formation density, and thermal neutron porosity had a high influence on the performance of the GMDH permeability model.

Author Contributions: C.N was responsible for the overall supervision of the research team and edited the final manuscript. B.M.N. prepared and analyzed the data whilst prepared the first draft, S.A.-O. Prepared and presented the algorithm and editing the manuscript. A.K.M. was responsible for well log data interpretation and analysis. J.C. Review writing L.W. Contributed to the validation of the design model.

Conflicts of Interest: The authors declare no conflict of interest

References

1. Chehrazi, A.; Rezaee, R. A systematic method for permeability prediction, a Petro-Facies approach. *J. Pet. Sci. Eng.* **2012**, *82*, 1–16.
2. Gholami, R.; Moradzadeh, A.; Maleki, S.; Amiri, S.; Hanachi, J. Applications of artificial intelligence methods in prediction of permeability in hydrocarbon reservoirs. *J. Pet. Sci. Eng.* **2014**, *122*, 643–656.

3.　Moussa, T.; Elkatatny, S.; Mahmoud, M.; Abdulraheem, A. Development of new permeability formulation from well log data using artificial intelligence approaches. *J. Energy Resour. Technol.* **2018**, *140*, 072903–072908.

4.　Mohaghegh, S.; Balan, B.; Ameri, S. Permeability determination from well log data. *SPE Form. Eval.* **1997**, *12*, 170–174.

5.　Rafik, B.; Kamel, B. Prediction of permeability and porosity from well log data using the nonparametric regression with multivariate analysis and neural network, Hassi R'Mel Field, Algeria. *Egypt J. Pet.* **2017**, *26*, 763–778, doi:10.1016/j. ejpe 2016.10.013.

6.　Anifowose, F.A.; Labadin, J.; Abdulraheem, A. Hybrid intelligent systems in petroleum reservoir characterization and modeling: The journey so far and the challenges ahead. *J. Pet. Explor. Prod. Technol.* **2017**, *7*, 251–263.

7.　Elkatatny, S.; Mahmoud, M.; Tariq, Z.; Abdulraheem, A. New insights into the prediction of heterogeneous carbonate reservoir permeability from well logs using artificial intelligence network. *Neural Comput. Appl.* **2018**, *30*, 2673–2683.

8.　Ahmadi, M.;Saemi, M.;Asghari, K., "Estimation of the reservoir permeability by petrophysical information using intelligent systems," *Petroleum science and technology*, **2008.**, 26, 1656-1667,

9.　Ali Ahmadi, M.; Zendehboudi, S.; Lohi, A.; Elkamel, A.; Chatzis, I. Reservoir permeability prediction by neural networks combined with hybrid genetic algorithm and particle swarm optimization. *Geophys. Prospect.* **2013**, *61*, 582–598, doi:10.1111/j.1365-2478.2012.01080.x.

10.　Guo, J.; Wang, H.; Guo, F.; Huang, W.; Yang, H.; Yang, K.; Xie, H. The back propagation based on the modified group method of data-handling network for oilfield production forecasting. *J. Pet. Explor. Prod. Technol.* **2019**, *9*, 1285–1293.

11.　Samarasinghe, S. *Neural Networks for Applied Sciences and Engineering: From Fundamentals to Complex Pattern Recognition*; Auerbach publications, New York USA: 2016.

12.　Ahmadi, M.A.; Zendehboudi, S.; Dusseault, M.B.; Chatzis, I. Evolving simple-to-use method to determine water–oil relative permeability in petroleum reservoirs. *Petroleum* **2016**, *2*, 67–78.

13.　Paul, A.; Bhowmik, S.; Panua, R.; Debroy, D. Artificial Neural Network-Based Prediction of Performances-Exhaust Emissions of Diesohol Piloted Dual Fuel Diesel Engine Under Varying Compressed Natural Gas Flowrates. *J. Energy Resour. Technol.* **2018**, *140*, 112201.

14.　Chen, W.; Wilson, J.; Tyree, S.; Weinberger, K.; Chen, Y. Compressing neural networks with the hashing trick. In Proceedings of the 32nd International Conference on Machine Learning, Lille, France, 6–11 July 2015; pp. 2285–2294.

15.　Iturrarán-Viveros, U.; Parra, J.O. Artificial neural networks applied to estimate permeability, porosity and intrinsic attenuation using seismic attributes and well-log data. *J. Appl. Geophys.* **2014**, *107*, 45–54.

16.　Singh, S.; Kanli, A.I.; Sevgen, S. A general approach for porosity estimation using artificial neural network method: A case study from Kansas gas field. *Studia Geophys. Geod.* **2016**, *60*, 130–140.

17.　Ebtehaj, I.; Bonakdari, H.; Khoshbin, F.; Bong, C.H.J.; Ghani, A.B. Development of group method of data handling based on genetic algorithm to predict incipient motion in rigid rectangular storm water channel. *Sci. Iran. Trans. A Civ. Eng.* **2017**, *24*, 1000–1009.

18.　Baziar, S.; Tadayoni, M.; Nabi-Bidhendi, M.; Khalili, M. Prediction of permeability in a tight gas reservoir by using three soft computing approaches: A comparative study. *J. Nat. Gas Sci. Eng.* **2014**, *21*, 718–724.

19.　Bhatt, A.; Helle, H.B. Committee neural networks for porosity and permeability prediction from well logs. *Geophys. Prospect.* **2002**, *50*, 645–660.

20.　Karimpouli, S.; Fathianpour, N.; Roohi, J. A new approach to improve neural networks' algorithm in permeability prediction of petroleum reservoirs using supervised committee machine neural network (SCMNN). *J. Pet. Sci. Eng.* **2010**, *73*, 227–232.

21.　Lim, J.-S.; Kim, J. Reservoir porosity and permeability estimation from well logs using fuzzy logic and neural networks. In Proceedings of the SPE Asia Pacific Oil and Gas Conference and Exhibition, Perth, Australia, 18–20 October 2004.

22.　Najafzadeh, M.; Barani, G.-A.; Hessami-Kermani, M.-R. Evaluation of GMDH networks for prediction of local scour depth at bridge abutments in coarse sediments with thinly armored beds. *Ocean Eng.* **2015**, *104*, 387–396.

23.　Porkhial, S.; Salehpour, M.; Ashraf, H.; Jamali, A. Modeling and prediction of geothermal reservoir temperature behavior using evolutionary design of neural networks. *Geothermics* **2015**, *53*, 320–327.

24. Shaghaghi, S.; Bonakdari, H.; Gholami, A.; Ebtehaj, I.; Zeinolabedini, M. Comparative analysis of GMDH neural network based on genetic algorithm and particle swarm optimization in stable channel design. *Appl. Math. Comput.* **2017**, *313*, 271–286.

25. Seliverstova, A.V.; Pavlova, D.A.; Tonoyan, S.A.; Gapanyuk, Y.E. The Time Series Forecasting of the Company's Electric Power Consumption. In Proceedings of the International Conference on Neuroinformatics, Moscow, Russia, 8–12 October 2018; pp. 210–215.

26. Jacob, T.; Usman, U.A.; Bemdoo, S.; Susan, A.A. Short-term electrical energy consumption forecasting using GMDH-type neural network. *J. Electr. Electron. Eng.* **2015**, *3*, 42–47.

27. Yu, J.; Park, J.; Kim, S. A New Input Selection Algorithm Using the Group Method of Data Handling and Bootstrap Method for Support Vector Regression Based Hourly Load Forecasting. *Energies* **2018**, *11*, 2870.

28. Ayoub, M.; Almansour, A.; Hassan, A. A Novel Formula for Estimating Oil Compressibility Below Bubble Point Pressure Using Group Method of Data Handling: A Comparative Approach. In Proceedings of the SPE/IATMI Asia Pacific Oil & Gas Conference and Exhibition, Bali, Indonesia, 29–31 October 2019.

29. Ma, S.-W.; Kou, C.-H.; Chen, L.; Wang, A.-P. Application of Group Method of Data Handling to Strean-Way Transition. In Proceedings of the 2009 International Joint Conference on Artificial Intelligence, Pasadena, CA, USA, 11–17 July 2009; pp. 301–304.

30. Menad, N.A.; Noureddine, Z.; Hemmati-Sarapardeh, A.; Shamshirband, S.; Mosavi, A.; Chau, K.-w. Modeling temperature dependency of oil-water relative permeability in thermal enhanced oil recovery processes using group method of data handling and gene expression programming. *Eng. Appl. Comput. Fluid Mech.* **2019**, *13*, 724–743.

31. Shen, C.; Asante-Okyere, S.; Yevenyo Ziggah, Y.; Wang, L.; Zhu, X. Group Method of Data Handling (GMDH) Lithology Identification Based on Wavelet Analysis and Dimensionality Reduction as Well Log Data Pre-Processing Techniques. *Energies* **2019**, *12*, 1509.

32. Teng, G.; Xiao, J.; He, Y.; Zheng, T.; He, C. Use of group method of data handling for transport energy demand modeling. *Energy Sci. Eng.* **2017**, *5*, 302–317.

33. Nguyen, T.N.; Lee, S.; Nguyen-Xuan, H.; Lee, J. A novel analysis-prediction approach for geometrically nonlinear problems using group method of data handling. *Comput. Methods Appl. Mech. Eng.* **2019**, *354*, 506–526.

34. Safari, M.J.S.; Ebtehaj, I.; Bonakdari, H.; Es-haghi, M.S. Sediment transport modeling in rigid boundary open channels using generalize structure of group method of data handling. *J. Hydrol.* **2019**, *577*, 123951.

35. Rostami, A.; Hemmati-Sarapardeh, A.; Karkevandi-Talkhooncheh, A.; Husein, M.M.; Shamshirband, S.; Rabczuk, T. Modeling heat capacity of ionic liquids using group method of data handling: A hybrid and structure-based approach. *Int. J. Heat Mass Transf.* **2019**, *129*, 7–17.

36. Lambert, R.S.; Lemke, F.; Kucherenko, S.S.; Song, S.; Shah, N. Global sensitivity analysis using sparse high dimensional model representations generated by the group method of data handling. *Math. Comput. Simul.* **2016**, *128*, 42–54.

37. Ivakhnenko, A. Heuristic self-organization in problems of engineering cybernetics. *Automatica* **1970**, *6*, 207–219.

38. Ivakhnenko, A.G. Polynomial theory of complex systems. *IEEE Trans. Syst. Man Cybern.* **1971**, 364-378.

39. Ivakhnenko, G. Short-Term Process Forecasting by Analogues Complexing GMDH Algorithm. In Proceedings of the 2nd Int. Conf. on Inductive Modeling, Kyiv, Ukraine, 15–19 September 2008; pp. 241–245.

40. Park, H.-S.; Park, B.-J.; Kim, H.-K.; Oh, S.-K. Self-organizing polynomial neural networks based on genetically optimized multi-layer perceptron architecture. *Int. J. Control Autom. Syst.* **2004**, *2*, 423–434.

41. Saemi, M.; Ahmadi, M. Integration of genetic algorithm and a coactive neuro-fuzzy inference system for permeability prediction from well logs data. *Transp. Porous Media* **2008**, *71*, 273–288.

42. Committee, I.U.M. *East African Rift System: Upper Mantle Committee-UNESCO Seminar, Nairobi, April 1965*; University College, Nairobi Kenya: 1965.

43. Tiberi, C.; Ebinger, C.; Ballu, V.; Stuart, G.; Oluma, B. Inverse models of gravity data from the Red Sea-Aden-East African rifts triple junction zone. *Geophys. J. Int.* **2005**, *163*, 775–787.

44. Barberi, F.; Varet, J. The Afar rift junction. In *Petrology and Geochemistry of Continental Rifts*; Springer: Berlin/Heidelberg, Germany, 1978; pp. 55–69.

45. Mueller, J.-A.; Ivachnenko, A.; Lemke, F. GMDH algorithms for complex systems modelling. *Math. Comput. Model. Dyn. Syst.* **1998**, *4*, 275–316.

46. Howland, J.C.; Voss, M.S. Natural gas prediction using the group method of data handling. In Proceedings of the 7th International Conference on Artificial Intelligence and Soft Computing, Banff, AB, Canada, 27–30 June 2004.

47. Abdel-Aal, R. Abductive networks: A new modeling tool for the oil and gas industry. In Proceedings of SPE Asia Pacific Oil and Gas Conference and Exhibition, Melbourne, Australia, 8–10 October 2002.

48. Madala, H.R.; Ivakhnenko, A.G. *Inductive Learning Algorithms for Complex Systems Modeling*; CRC press Boca Raton, Florida, USA: 1994; Volume 368.

49. Guo, J.; Huang, W.; Mao, Q.; Wang, X.; Wang, X.; Song, T. Modified GMDH networks for oilfield production prediction. *Geosyst. Eng.* **2018**, *21*, 217–225.

50. Asante-Okyere, S.; Shen, C.; Yevenyo Ziggah, Y.; Moses Rulegeya, M.; Zhu, X. Investigating the Predictive Performance of Gaussian Process Regression in Evaluating Reservoir Porosity and Permeability. *Energies* **2018**, *11*, 3261.

51. Semenov, A.A.; Oshmarin, R.A.; Driller, A.; Butakova, A. Application of Group Method Of Data Handling For Geological Modeling of Vankor Field. In Proceedings of the North Africa Technical Conference and Exhibition, Cairo, Egypt, 14–17 February 2010.

.