

Received November 13, 2020, accepted December 9, 2020, date of publication December 22, 2020, date of current version January 11, 2021.

Digital Object Identifier 10.1109/ACCESS.2020.3046698

# A Comprehensive Review of Soft Computing Models for Permeability Prediction

MUBARAK SAAD ALMUTAIRI<sup>1</sup>, (Member, IEEE)

CSET Department, Hafr Al Batin Community College, University of Hafr Al Batin, Hafr Al Batin 31991, Saudi Arabia

e-mail: [almutairi@uhb.edu.sa](mailto:almutairi@uhb.edu.sa)

**ABSTRACT** Crude oil is a vital and valuable commodity in the energy industry. In order to maintain continuous, stable, and reasonably priced supplies, oil producers need cheaper exploration and extraction techniques. Permeability is one of the formation parameters that is a key interest to petroleum engineers in determining the economic worth and yield of crude deposits, yet permeability prediction remains a difficult problem. Many approaches have been applied to solving this important issue. Soft computing has been deployed to predict permeability. In this paper, we present an extensive review of the existing research that has been conducted on applications of soft computing for permeability prediction. This paper finds out that traditional approaches for permeability prediction are still relevant in the oil and gas industry. Soft computing methods in particular are worthy of addition to this interesting area. This extensive review is intended to be an entry point for further exploration of other approaches that have received little or no attention from researchers.

**INDEX TERMS** Permeability prediction, soft computing models, oil and gas.

## I. INTRODUCTION

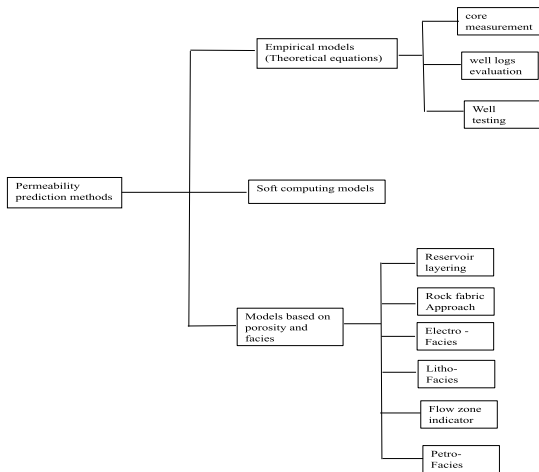
Crude oil is central to the world economy by meeting the energy demand and it is the engine room of most economies. Players in the oil market need cheaper extraction methods and techniques to maintain a reasonable price and stability of the world economy. Permeability is one of the formation parameters that is pivotal to petroleum engineers because it is essential in determining the economic worth and yield of crude deposits. Permeability is a measure to determine the ability of fluid to flow through materials [1]. It is fundamental in reservoir management, in the choice of optimal drainage port, and in perforation. Formations permeability are usually measured from the core sample and well logging in the laboratories which are often tedious and expensive. Researchers have agreed that permeability prediction is a complex problem in oil industries [1]–[3]. The difficult and dynamic nature of the problem has aroused the interest of researchers over the years. The cost of direct measurement has led researchers to show interest in other ways of predicting this vital parameter. One of the methods that have been harnessed is soft computing applications in predicting permeability. The successes of soft computing techniques in other application areas have led to its popularity in oil and gas exploration.

The associate editor coordinating the review of this manuscript and approving it for publication was Tawfik Al-Hadhrani<sup>1</sup>.

We classify all the methods for predicting permeability as shown in Fig. 1. In this paper, we are concentrating on soft computing techniques. This study is intended to be the starting point for any researcher that is delving into the application of artificial intelligence techniques in permeability prediction. We conduct a systematic and in-depth analysis of the methods available in the literature by exposing their strengths and weaknesses. The remaining part of this paper is organized as follows. Section II presents an overview of soft computing models. In Section III, the formation studied in literature and data availability are discussed. Section IV presents statistical quality measures that are employed in the literature for the performance evaluation of proposed models. Section V classifies all the works in the literature on the applications of computational intelligence into single and hybrid models while Section VI concludes the paper.

## II. OVERVIEW OF SOFT COMPUTING MODELS

Soft computing, artificial intelligence, computational intelligence, and machine learning are used interchangeably in the literature. Though, there are subtle differences between some of them, for the sake of readers from diverse research disciplines. We live it at that. For the sake of consistency, throughout this paper, we will be using soft computing techniques. Soft computing is an aspect of computing that attempts to learn from data and adjust itself with experience where necessary. There are several soft computing techniques available



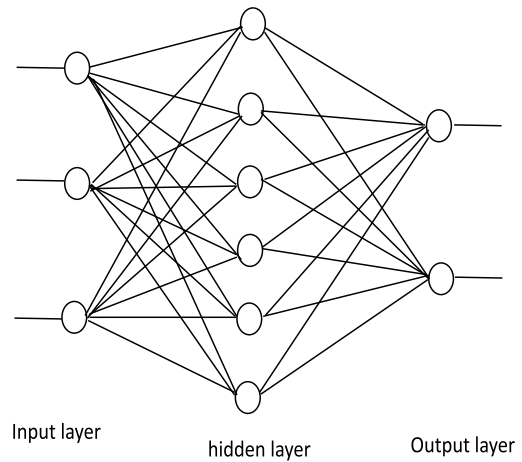
**FIGURE 1.** Classification of permeability predictions methods.

in the literature. Some of them are nature-inspired such as genetic algorithm, cuckoo search, ant colony, particle swarm optimization, and so on while others mimic the human neural system like artificial neural networks. Others are purely mathematical models such as extreme learning machines, support vector machines, etc. The learning process for all the algorithms are usually classified as supervised, unsupervised, and semi-supervised or reinforcement learning. In this section, we will provide an overview of the soft computing models/techniques that are mostly applied for permeability prediction in the literature.

### A. ARTIFICIAL NEURAL NETWORK

Artificial Neural Network (ANN) is a mathematical model designed to emulate the human brain. It attempts to mimic the biological neural system of using neurons and nodes to process information [4]. A neural network is a black box universal approximator that can approximate difficult and complex functions. Its details working are like human reasoning with the ability to learn and adapt from data provided to it. ANN is made up of four major components; information is processed with the aid of neurons. Each link connects one or more neurons. Every link between neurons is associated with adjustable weight and an activation function is applied to its input to regulate its outputs. There are no strict rules on the number of hidden layers. It is usually based on trial and error [5]. Structurally, the basic ANN is made up of an input layer, an output layer and in between these two layers are one or more hidden layers as shown in Fig. 2. The operations of ANN involve training, validation, and testing. The network can be trained with data, which means finding an appropriate relationship between the inputs and outputs after adjusting weight and biases. The “reasonable” relationship depends on the predefined error threshold.

There are many training algorithms for ANN [6]. Backpropagation (BP) is used to train ANN by adjusting weight and learn from the data. Backpropagation is a supervised learning algorithm where a set of data is provided against a set of outputs. During training, in every iteration, it tries



**FIGURE 2.** Structure of feedforward neural networks.

to minimize the error by adjusting weight biases. There are three training termination criteria; They are: (a) if the gradient performance falls below a predefined threshold (b) predefined error limit is reached and (c) if the maximum iterations are reached [7]. ANN can handle many problems. It can process nonlinear problems independent of the initial assumption. However, it has its drawbacks. ANN is restrictive and can get stuck in local minima. To address some of these challenges faced by the basic neural network which is a Feedforward Neural Network (FFNN) that gives multi-layer perceptron or backpropagation learning algorithms. There are many modifications and extensions to ANN such as Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Deep Learning Network (DLN), Radial Basis Neural Network (RBNN), Higher-order Networks, Probabilistic Networks, Fuzzy Neural Networks (FNN), Wavelength Neural Networks (WNN), and Generalized Regression Neural Network (GRNN) [8]. RBNN is similar in structure to FFNN, that is, it has input, hidden and output layer. However, the hidden layer of RBNN is made up of a radial basis function. Its input bypasses the weight computation at the input layer and goes directly to the hidden layer with activation of radially symmetric basis function. RBNN has a high learning speed and local approximation [9]. RNN is more structurally complex compared to RBNN and FFNN. In RNN, there are intermediate inputs connected directly to the input at the hidden layer as can be seen from Fig. 3. In the use of GRNN, investigating the structure and iterative training is not necessary as in backpropagation [8].

### B. ADAPTIVE NEURO-FUZZY INFERENCE SYSTEMS

Adaptive Neuro-Fuzzy Inference Systems (ANFIS) is a neuro-fuzzy technique that combines the neural network and the fuzzy inference system. The relational structure of the artificial neural networks and its learning ability is fused with the decision-making mechanism of fuzzy logic. The neural network is highly adaptable while fuzzy logic handles imprecision and uncertainty of the system. The disadvantages of black-box nature and indeterminable weight of artificial

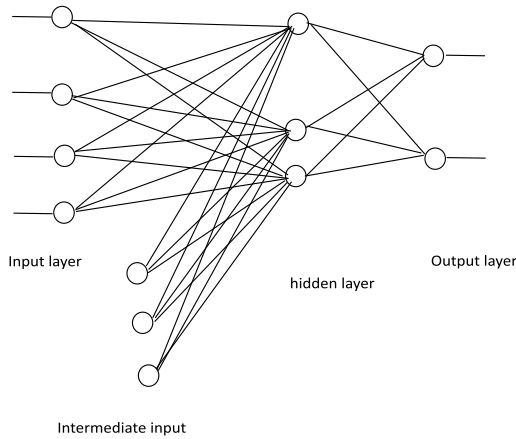


FIGURE 3. Structure of recurrent neural networks.

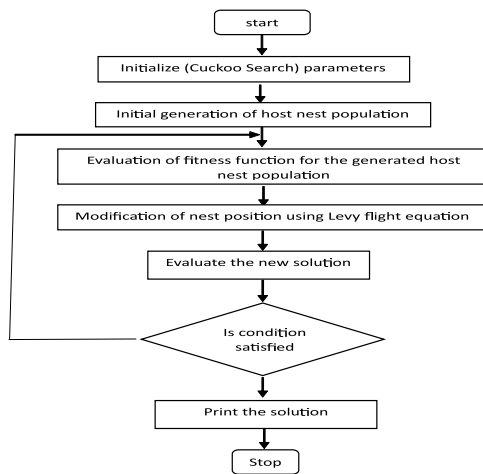


FIGURE 4. Flowchart of cuckoo search algorithm.

neural network is eliminated by the structure of ANFIS in the fuzzy inference system. Detailed descriptions of the ANFIS model can be found in [10].

**C. CUCKOO SEARCH ALGORITHM**

Cuckoo Search Algorithm (CSA) is inspired by the lifestyle and behavior of a bird species called cuckoo. Due to the peculiar reproductive behavior of laying their eggs in another bird’s nest. Cuckoo can detect new next and its egg, even when mixed with the host’s bird’s eggs. Some host birds might retaliate by either rejecting the cuckoo eggs or move to a new location to build a new nest. For simplicity, each egg in a nest represents a solution and each cuckoo egg represents a new one [11]. The cuckoo search procedure is presented below and its flowchart is shown in Fig. 4.

- Each cuckoo lays one egg at a time and puts it in a randomly chosen nest.
- The best nest with a high-quality egg is carried over to the next generation.
- The number of available host nest is constant and a host has the probability  $P_a \in [0, 1]$  of discovering an alien egg.
- Consequently, the host bird can either throw away the cuckoo egg or move to a new location to build a new nest.

**D. EXTREME LEARNING MACHINE**

Extreme Learning Machine (ELM) is a learning algorithm for a single hidden layer feedforward neural network (SLFN) with  $\hat{N}$  hidden neurons and  $\hat{N} < N$ , where  $N$  is the number of training samples. The learning process as proposed in [12], and is described as follows:

With  $N$  training samples  $(\mathbf{x}_i, \mathbf{t}_i)$ , where  $\mathbf{x}_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T \in \mathbf{R}^n$  and  $\mathbf{t}_i = [t_{i1}, t_{i2}, \dots, t_{im}]^T \in \mathbf{R}^m$ . Hence, a standard SLFN with  $\hat{N}$  hidden neurons and activation function  $g(x)$  is defined as:

$$\sum_{i=1}^{\hat{N}} \beta_i g(\mathbf{w}_i \cdot \mathbf{x}_j + b_i) = \mathbf{o}_j, \quad j = 1, 2, \dots, N, \quad (1)$$

where  $\mathbf{w}_i = [w_{i1}, w_{i2}, \dots, w_{im}]^T$  is the connecting weight vector between the  $i^{th}$  neuron and the input neuron and  $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{im}]^T$  is the connecting weight vector between the  $i^{th}$  hidden neuron and the output neuron.  $b_i$  is the threshold of  $i^{th}$  hidden neurons.  $\mathbf{w}_i \cdot \mathbf{x}_j$  is the inner product of  $\mathbf{w}_i$  and  $\mathbf{x}_j$ .

The standard SLFN with  $\hat{N}$  hidden neurons and activation function  $g(x)$  is to approximate these  $N$  samples with zero error that is  $\sum_{j=1}^{\hat{N}} \|\mathbf{o}_j - \mathbf{t}_j\| = 0$ , there exist  $\beta_i$ ,  $\mathbf{w}_i$ , and  $b_i$  such that:

$$\sum_{i=1}^{\hat{N}} \beta_i g(\mathbf{w}_i \cdot \mathbf{x}_j + b_i) = \mathbf{t}_j, \quad j = 1, 2, \dots, N. \quad (2)$$

The above equation can be re-written as

$$\mathbf{H}\beta = \mathbf{T}, \quad (3)$$

where

$$\mathbf{H}(\mathbf{w}_1, \dots, \mathbf{w}_N, b_1, \dots, b_N, \mathbf{x}_1, \dots, \mathbf{x}_N) = \begin{bmatrix} g(\mathbf{w}_1 \cdot \mathbf{x}_1 + b_1) \dots g(\mathbf{w}_N \cdot \mathbf{x}_1 + b_N) \\ \vdots \\ g(\mathbf{w}_1 \cdot \mathbf{x}_N + b_1) \dots g(\mathbf{w}_N \cdot \mathbf{x}_N + b_N) \end{bmatrix}_{N \times N}, \quad (4)$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_N^T \end{bmatrix}_{N \times m}, \quad (5)$$

$$\mathbf{T} = \begin{bmatrix} \mathbf{t}_1^T \\ \vdots \\ \mathbf{t}_N^T \end{bmatrix}_{N \times m}. \quad (6)$$

As proposed in [13],  $\mathbf{H}$  is called the hidden layer output matrix of the neural network; the  $i^{th}$  column of  $\mathbf{H}$  is the  $i^{th}$  hidden neuron’s output vector.

**E. FUNCTIONAL NETWORKS**

Functional Networks (FN) is an extension of the neural networks for efficient generalization and a combination of both domain and data. In FN, there is no weight associated with links and it can select the best function by learning from the data and attempt to minimize MSE. The learning process in FN involves both parametric and structural. Parametric is the estimation of the neural function while structural is based on the topology of the network [14].

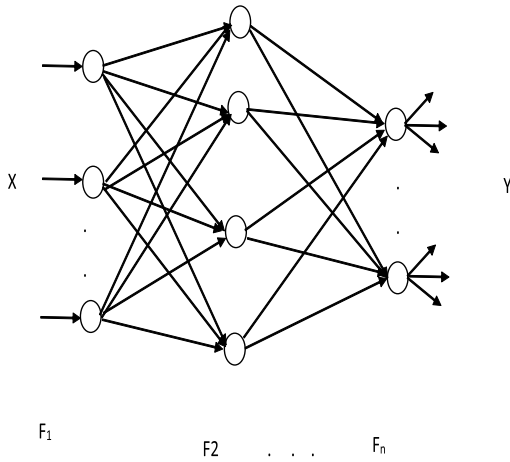


FIGURE 5. A general functional network structure.

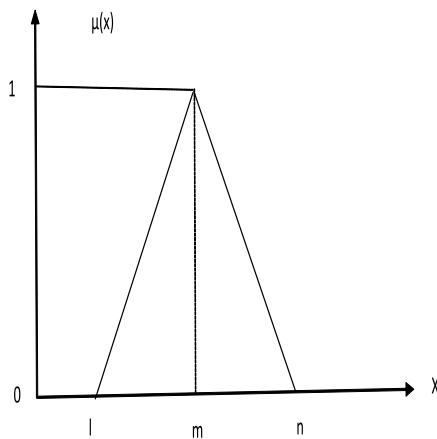


FIGURE 6. Triangular membership function for type I fuzzy.

FN has the following components (a) input layer, an output layer, and one or several layers of processing unit. (b) None, one or several layers of intermediate units that store intermediate results as shown in Fig. 5.

Generally,  $n$  layer FN model is as follows:

$$Y = F_n \cdot F_{n-1} \dots F_1(X), \tag{7}$$

where  $Y \in \mathbf{R}^m$  represents output of the system,  $X \in \mathbf{R}^n$  represents input to  $F_1, F_2, \dots, F_n$ .

**F. FUZZY LOGIC**

Type I fuzzy set was introduced in [15]. It is used to model uncertain processes such as temperature and pressure. The membership function  $U_x(x)$  is chosen based on individual experience from [0 1]. A simple example of a triangular membership function for type I fuzzy is shown in Fig. 6. Expert experience is always needed. Type I fuzzy set membership function is crisp. It has been applied in several applications such as data mining, time series modeling, reservoir characterization, and permeability prediction.

Despite the successes of type I, there are limitations to the level of uncertainties that can be modeled by type I. Hence, type II is an extension of type I. The membership function

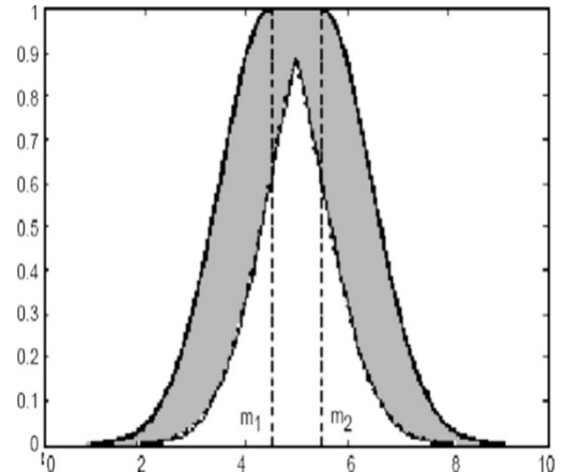


FIGURE 7. Gaussian membership function.

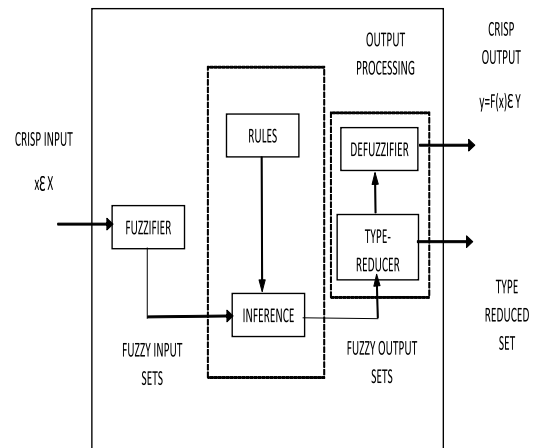


FIGURE 8. Type II fuzzy logic system.

$U_x(x)$  in type II is itself fuzzy. A special case of type II fuzzy logic is called interval type II which is widely used. The membership function in interval type II is an interval. For example, consider the Gaussian membership function in Fig. 7. Interval  $U_x(5) = [0.7, 1]$ . Type II is bounded above and below by two type I fuzzy sets. They are called Upper Membership Function (UMF) and Lower Membership Function (LMF) and the shaded area is called the footprint of uncertainty. The membership function can be constructed from surveys or using an optimization algorithm [16]. Type II fuzzy logic system is shown in Fig. 8. The fuzzifier maps the crisp input into the fuzzy set. The rules are of the form:  $R : \text{if } x_1 \text{ is } f_1^l \text{ and } x_2 \text{ is } f_2^l \text{ and } \dots x_p \text{ is } f_p^l \text{ then } y \text{ is } G^l$ . Note that it is not a necessary condition for all the antecedent/consequent to be type II. If either one antecedent or consequent is type II, hence it all becomes a type II. The inference engine combines rules and gives the mapping from input type II fuzzy to output type II fuzzy set. The defuzzifier produces an extension principle to produce the type I defuzzification method. This process is called type reduction.

**G. SUPPORT VECTOR MACHINE**

Support Vector Machine (SVM) is a statistical learning model developed in [17]. SVM is a supervised learning model that

can receive input data and output function that can be used to predict the features of future data [18]. It constructs nonlinear decision functions by training a classifier to perform a linear separation in a high-dimensional space which is nonlinearly related to the input space. SVM has been applied in solving several problems such as classification, regression, function approximation, text categorization, pattern recognition. The ability of SVM to overcome getting stuck in local maxima and ease of generalization makes it a model of choice for many applications. Given samples, the SVM model is formulated as a minimization of Vapnik’s  $\epsilon$ -insensitive loss function with the target value ( $y$ ) as:

$$|y - f(x)|_\epsilon = \max\{0, |y - f(x) - \epsilon|\}. \quad (8)$$

To estimate a linear regression

$$f(x) = (w \cdot x) + b, \quad (9)$$

where  $x$  is the input vector,  $w$  is the weight, and  $b$  is the bias term. The objective is to minimize

$$\frac{1}{2} \|w\|^2 + C \sum_{i=1}^m |y - f(x)|'_\epsilon. \quad (10)$$

To ensure that the margin  $\epsilon$  is maximized and error of classification is minimized,  $C$  is introduced as a trade-off parameter. Considering the set of constraints, the problem can be formulated as an optimization problem as:

$$\text{minimize } \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i + \xi'_i, \quad (11)$$

$$\text{subject to: } y_i - w^T \cdot x - b \leq \xi_i + \epsilon, \quad (12)$$

$$w^T \cdot x + b - y_i \leq \xi'_i + \epsilon, \quad (13)$$

$$\xi_i, \xi'_i, x_i \geq 0. \quad (14)$$

According to the constraints (11) and (12), if the error is less than  $\epsilon$  will not enter the objective function and does not require positive  $\xi_i$  or  $\xi'_i$  [19]. SVM basis function could be polynomial, radial basis function, and sigmoid function.

### H. GENETIC ALGORITHM

Genetic Algorithm (GA) is an evolutionary search algorithm that attempts to mimic Charles Darwin’s theory of survival of the fittest in biology. GA is used to solve optimization problems. It uses three standard processes of selection, crossover, and mutation. Since it is a search problem, the optimal solution is always our objective. The algorithm is summarized in Fig. 9. We start off by randomly selecting an initial population that is likely to have the optimal solution. The population evolves with time as an individual and as a group. During selection, the best performing chromosomes are selected and the weaker ones dropped. The process continues iteratively until an optimal solution is reached.

### I. GENETIC PROGRAMMING

Genetic Programming (GP) is an example of an evolutionary algorithm inspired by biological evolution. It is used to discover solutions to problems humans do not know how to solve directly. GP implements algorithms that use random

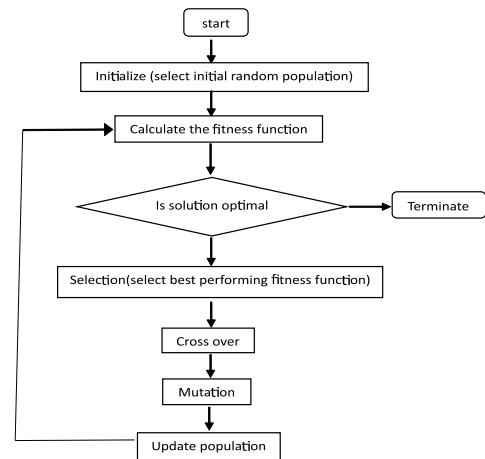


FIGURE 9. Genetic algorithm flowchart.

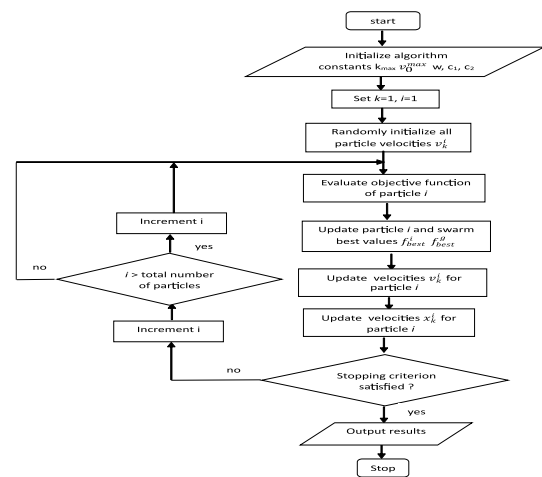


FIGURE 10. Particle swarm optimization algorithm flowchart.

mutation, crossover, fitness function, and multiple generations of evolution to resolve a user-defined task [20]. It carries out the continuous improvement of an initial random population of programs. These improvements are made possible by stochastic variations of programs and selections. Solutions are reached according to predefined criteria for judging an acceptable solution [21].

### J. PARTICLE SWARM OPTIMIZATION

Particle Swarm Optimization (PSO) is inspired by the social behavior and movement dynamics of fish, birds, and insects. It is a stochastic search method suitable for continuous-variable problems. PSO has several advantages such as an efficient global search algorithm, simple implementation, and very few algorithm parameters. However, one of its drawbacks is a weak local search [22]. The PSO algorithm is described as follows and the flowchart is as shown in Fig. 10:

- 1) Create a population of particles (agents) uniformly distributed over  $x$ .
- 2) Evaluate each particle’s position according to the object function.

TABLE 1. Statistical properties of the Middle Eastern well data

Predictors	Meaning	Mean	Standard Deviation	Min	Max
MSFL	Microspherically focused log	1.1762	0.4575	0.5377	2.4368
NPHI	Neutron porosity log	0.1371	0.0517	0.0300	0.2611
PHIT	Porosity log	0.1529	0.0679	0.0335	0.2909
RHOB	Density log	2.4371	0.1419	2.1810	2.6682
SWT	Water saturation	0.1703	0.1783	0.0400	1.0000
CALI	Caliper	8.4105	0.1038	8.1558	8.4891
CT	Conductivity	0.0495	0.0305	0.0001	0.1121
DRHO	Density correction log	0.0570	0.0277	0.0030	0.1298
GR	Gamma ray log	14.7926	3.9779	6.0399	31.0351
RT	Deep resistivity	1.3108	3.3131	0.0089	10.0000
k	Core permeability	59.6300	163.2500	0.0150	966.8900

- 3) If the particle current position is better than the previous position, then update.
- 4) Determine the best particle based on its previous position. Update the particle velocity using

$$V_i^{t+1} = V_i^t + \varphi_1 U_1^T (Pb_l^t - X_i^t) + \varphi_2 U_2^T (gb_l^t - X_i^t), \quad (15)$$

where  $V_i^{t+1}$  is velocity,  $Pb_l^t$  is the best-remembered individual particle position,  $gb_l^t$  is the best-remembered swarm position,  $X_i^{t+1} = X_i^t + V_i^{t+1}$ ,  $\varphi_1, \varphi_2$  are cognitive and social parameter, and  $U_1, U_2$  are random numbers between 0 and 1.

- 5) Move to step 2 until stopping criteria are satisfied.

### III. FORMATION STUDIED AND DATA AVAILABILITY

The reservoir formations data used in literature are mostly obtained from carbonate reservoirs in the Middle Eastern region, Iran (Mansouri, Sarawak, Kangan, South Pars), and a few other formations from other countries such as the Hassi field in Algeria, Western Sichuan Basin and Mesozoic strata Gaoqing in China, and Washakie Basin in the USA. Others are the North sea and Ula field in Norway and Venture gas field in offshore Canada. The datasets named the Middle East with the highest number of the published papers consist of a total of 356 observations from well logging with core permeability measurements. The statistical descriptions of the predictors and the corresponding core permeability are given in Table 1. Most times, the predictors are of different units, the majority of the researchers normalized their data to the range of [0,1] while the permeability output was mostly expressed in the logarithmic scale. The data sets were divided randomly into 70% for the training of the models and 30% for testing in most of the studies.

### IV. PERFORMANCE EVALUATION METHODS

In this section, we present statistical quality measures that are employed in the literature for the performance evaluation of proposed models.

- 1) **Correlation Coefficient (r):** The correlation coefficient  $r$  measures the strength and the direction of a linear relationship between two variables. Strong positive linear relationships have values of  $r$  closer to 1. If  $x$  and  $y$  have a strong positive linear correlation,  $r$  is close to +1. An  $r$  value of exactly +1 indicates a

perfect positive fit. Positive values indicate a relationship between  $x$  and  $y$  variables such that as values for  $x$  increases, values for  $y$  also increase. Negative correlation: If  $x$  and  $y$  have a strong negative linear correlation,  $r$  is close to -1. An  $r$  value of exactly -1 indicates a perfect negative fit. Negative values indicate a relationship between  $x$  and  $y$  such that as values for  $x$  increase, values for  $y$  decrease. No correlation: If there is no linear correlation or a weak linear correlation,  $r$  is close to 0. A value near zero means that there is a random, nonlinear relationship between the two variables. Note that  $r$  is a dimensionless quantity; that is, it does not depend on the units employed. A perfect correlation of  $\pm 1$  occurs only when the data points all lie exactly on a straight line. A correlation greater than 0.8 is generally described as strong, whereas a correlation less than 0.5 is generally described as weak. These values can vary based upon the type of data being examined. A study utilizing medical data may require a stronger correlation than a study using social experiment data.

$$r = \frac{\sum_{i=1}^N (y_a - \hat{y}_a)(y_p - \hat{y}_p)}{\sqrt{\sum (y_a - \hat{y}_a)^2 \sum (y_p - \hat{y}_p)^2}}, \quad (16)$$

where  $y_a$  and  $y_p$  are the actual and predicted values and  $\hat{y}_a$  and  $\hat{y}_p$  are the mean of the actual and predicted values.

- 2) **Root Mean Square Error (RMSE):** RMSE is the standard deviation of the residuals (prediction errors). These deviations are called residuals when the calculations are performed over the data sample that was used for estimation and are called errors (or prediction errors) when computed out-of-sample. Residuals are a measure of how far from the regression line data points are; RMSE is a measure of how spread out these residuals are. In other words, it tells you how concentrated the data is around the line of best fit.

$$RMSE = \sqrt{\frac{1}{N} \sum_{k=1}^N (y_k - \hat{y}_k)^2}, \quad (17)$$

where  $y_k$  and  $\hat{y}_k$  are the actual and predicted values and  $N$  is the number of data samples.

- 3) **Mean Absolute Percentage Error (MAPE):** The MAPE measures the size of the error in percentage terms.

It is calculated as the average of the unsigned percentage error. The MAPE is scale-sensitive and not suitable for low-volume data. Notice that because actual is in the denominator of the equation, the MAPE is undefined when actual demand is zero. Furthermore, when the actual value is not zero, but quite small, the MAPE will often take on extreme values. This scale sensitivity of MAPE makes it unsuitable as an error measure for low-volume data.

$$MAPE = \frac{1}{N} \sum_{k=1}^N |y_k - \hat{y}_k| * 100, \quad (18)$$

where  $y_k$  and  $\hat{y}_k$  are the actual and predicted values and  $N$  is the number of data samples.

#### 4) Mean Absolute Error (MAE):

$$MAE = \frac{1}{N} \sum_{k=1}^N |y_k - \hat{y}_k|, \quad (19)$$

where  $y_k$  and  $\hat{y}_k$  are the actual and predicted values and  $N$  is the number of data samples.

#### 5) Mean Square Error (MSE):

$$MSE = \sum_{k=1}^N \left| \frac{(y_k - \hat{y}_k)^2}{y_k} \right| * 100, \quad (20)$$

where  $y_k$  and  $\hat{y}_k$  are the actual and predicted values and  $N$  is the number of data samples.

#### 6) Average Absolute Percentage Relative Error ( $E_a$ ):

$$E_a = \frac{1}{N} \sum_{k=1}^N (y_k - \hat{y}_k)^2, \quad (21)$$

where  $y_k$  and  $\hat{y}_k$  are the actual and predicted values and  $N$  is the number of data samples.

#### 7) Standard Deviation (SD): SD is a measure that is used to quantify the amount of variation or dispersion of a set of data values from the mean.

$$SD = \sqrt{\frac{1}{N} \sum_{k=1}^N (y_k - \bar{y})^2}, \quad (22)$$

where  $y$  is the actual value,  $\bar{y}$  is the mean value, and  $N$  is the number of data samples.

## V. APPLICATIONS OF SOFT COMPUTING FOR PERMEABILITY PREDICTION

As the quest for effective permeability prediction techniques continues, soft computing techniques were harnessed as it proves to be able to handle complex problems. In this section, we provide a detailed analysis of the works in the literature.

### A. SINGLE MODEL APPLICATION

Single model applications are the type of applications that uses only one soft computing technique in their applications. All techniques have their limitations, soft computing is not an exception. Though with a lot of successes is solving linear, nonlinear, and complex problems such as classification,

regression, prediction there are drawbacks. For example, one of the widely used neural networks is multilayer perception. It could easily get stuck in local minima, it is slow in training, determining the optimal architecture using trial and error is difficult.

In this section, we present works that use only a single model for the prediction of permeability. As stated earlier, traditionally there are many methods for estimating permeability indirectly from rock properties acquired from well log measurements. In [2], it is shown that permeability is correlated to the pore (micro-scale) model, pore characteristics, and statistical (percolation and fractal techniques). Singh [23] combined the estimates from well logs and core data using buckle methods to estimate permeability. Early work on the application of soft computing techniques for permeability prediction shows that is promising. It is inexpensive, non-interruption of production, and the speedy result of the investigation. Soft computing model applications have expanded over the years. According to [3], the initial well-known models for permeability prediction are ANN, fuzzy logic (FL), and neuro-fuzzy. Each of these models has its strength and weaknesses. Mohaghegh *et al.* [24] proposed multivariate regression analysis as a useful tool for permeability correlation. Furthermore, in [24], the authors proposed a virtual measurement using ANN. Due to the poor structural design of ANN, GRNN was used to design the optimal architecture and used backpropagation for the prediction. GRNN reaches a stable state in a short time but lacks generalization ability while ANN generalizes better i.e., it can predict new data reasonably. Their result shows that ANN performed reasonably well in prediction permeability of well log it was not used to.

Huang *et al.* [25] used ANN to model the relationship between the spatial position and permeability of six wells in the Venture gas field in offshore Canada. Helle *et al.* [26] used ANN to predict the porosity and permeability directly from well logs using data from the North sea formation. Singh [27] used a conventional log in FFNN and back-propagation to estimate permeability. Ben-Awuah and Padmanabhan [28] developed ANN model using facies instead of conventional logs. The authors used porosity as input data to the model and permeability as the target output. The correlation coefficient raised to 99%. In [29], Abdideh expressed the relationship between porosity and permeability with linear regression. The author created ANN model by dividing the dataset into different well zones. The correlation coefficient of each zone ranges from 0.73 to 0.85. Due to complex architecture design and slow training encountered with ANN, Saljooghi and Hezarkhani [30] combined wavelet theory and ANN to form a wave network (wavenet). The authors applied different wavelet as the activation function to the ANN. Though, wavelet parameters such as dilation and translation were kept constant. The wavenet outperform ANN with  $R^2$  of 92 compared to 89. In [31], Mohebbi *et al.* noted that the distinct feature of high pressure, heterogeneity in the Iran oil field inhibits the performance of other studies using ANN. They modified their work by zoning the reservoir

geological characteristic before applying ANN. The estimation of permeability from porosity, specific surface area, and irreducible water saturation was proposed in [32], while the authors in [33] examined the relationship between porosity, permeability, and depth. Depth and porosity were passed as input to the ANN to predict permeability. The researchers in [34] estimate a log derived permeability using ANN.

Researchers in [35] opined that generally, permeability prediction is a complex problem. The problem becomes more challenging in a tight sand formation with strong heterogeneity. This is the case of middle Jurassic Shaximiao formation in Western Sichuan Basin China studied by the authors. They investigated porosity using single linear regression (SLR) and permeability with multiple linear regression (MLR), multi-layer perceptron (MLP), and support vector regression (SVR) with multiple inputs using porosity and well logs. The result of their investigation indicated that MLR performed better than SLR while MLP and SVR outperformed SLR and MLR. The result of this work affirmed the assertion that though standard statistical models are still invaluable to petroleum engineers soft computing techniques have improved prediction accuracy. Hamada and Elshafei [36] combined well logs with nuclear magnetic resonance (NMR) for predicting gas-sand permeability using ANN. The permeability is derived from the empirical relationship between NMR porosity and mean value of T2 time. This model was tested with the combination of data from two wells from the Middle Eastern reservoir formation. The achieved correlation coefficient of 0.978 and 0.961 for training and testing respectively are very close to values obtained from core logs. Elkhatatny *et al.* [37] proposed a reduction in the number of inputs to ANN for permeability prediction. They used three logs namely; neutron porosity, bulk density, and resistivity as inputs to the ANN to predict permeability. In addition, they proposed a term called the mobility index from studying the interrelationship among the logs. The mobility index shows a high correlation with permeability values from the core. Their results were compared to that of ANFIS and SVM. The researchers in [38] used data from Mesaverde tight gas sandstone in Washakie Basin in the USA to predict permeability with the aid of MLP, SVM, and CANFIS (coactive neuro-fuzzy inference system).

In [39], the authors proposed that in order to overcome the slow training, high computational cost, and getting stuck in local minima associated with MLP, there is a need for a modification in the structure of ANN. According to the authors, the human brain is modular and massively parallel. These two features make them work independently. This modularity concept can be applied to MLP to improve its performance. The authors used a modular neural network (MNN) in permeability prediction in Persian gulf Iranian offshore, a formation with high heterogeneity. Different inputs were selected for different indicators in the formation. Spectral gamma-ray was used for the shale region, electricity resistivity and water saturation were considered for the permeable region while total and secondary porosity were also applied.

Different architectures of MNN were examined and the result shows that MNN reduced training time and CPU time as well as improving performance. Bagheripour [40] proposed the formation of a committee neural network for permeability prediction. First, the overlapping of data was removed using principal component analysis (PCA). Three committees were formed with MLP, RBF, and GRNN. The output from each member of the committee was fed as input into the committee neural network. The results show the committee performed better than the individual member. Furthermore, Irani and Nasimi [41] proposed an improvement to ANN by using GA to select optimal values for the weights and biases. The GA is used for its global search capability to overcome the weakness of ANN that usually gets stuck in local minima. The results of the optimized ANN with GA are better than the prediction without weight optimization. Aifa *et al.* [42] developed a nonlinear regression by using fuzzy logic to select the best input to the ANN model. Fuzzy logic was used to calibrate the permeability and prediction model with linear regression and backpropagation was constructed and compared. Olatunji *et al.* [43] applied ELM to build a model by assuming a nonlinear relationship of permeability with other rock properties using Middle Eastern well data. The ELM surpasses ANN and SVM in terms of speed of training, RMSE, and coefficient correlation. An extension of ANN, functional networks [44] was used in [45] for permeability prediction. FN allows neurons to be multivariate, multi-argument, and different learning function [44]. The authors used the least square (LS) method to estimate the activation functions for the network. The activation function can be based on the least square, steepest descent, and mini-max. The authors compared their results to that obtained from NN, linear, and nonlinear regression and fuzzy inference systems. FN performed better than ANN and ANFIS.

Alfaouri *et al.* [15] observed that most of the earlier work on permeability prediction was done on sandy reservoirs. They applied FL on the carbonate reservoir by modifying the defuzzification part of the fuzzy logic system. Fuzzy C-means was used for rock type clustering in [46]. Then ANN was applied to verify the result of the model. In [47], the authors proposed the use of FL for permeability prediction. The authors used ANFIS to learn the rules from the data, thereafter, applied the least square method and backpropagation gradient descent to train the fuzzy inference system (FIS). Rules were extracted using grid partitioning. Grid partitioning though easy to use, it may give rise to rules explosion. Subtractive clustering was used to generate input data clusters. Wang *et al.* [48] posited that permeability in tight and heterogeneous sand is more challenging. They proposed the idea of feature engineering for the optimization of FL. This involves the application of the Student-Newman-Keuls method to the sample before applying it to FL. They compared the results to an ordinary regression model without feature engineering. The optimized model performed better. In [49], the author maintained that different litho-facies exhibit different features. Also, log porosity may



TABLE 2. Summary of published papers by techniques and data sources

References	Technique	Source of data
[25]	ANN + Multiple Linear Regression + Multiple Nonlinear Regression	Venture Gas Field Offshore Canada
[29]	ANN + Regression Analysis	Southwest Iranian Oil Field
[28]	ANN	Sarawak Foreland Basin
[30]	ANN + Wavelet Theory	Not disclosed
[45]	FN	Middle Eastern Oil Well
[43]	ELM	Middle Eastern Oil Well
[42]	ANN + FL	Hassi Oil Field Algeria
[51]	SVM	South Pars Field Iran
[24]	ANN + GRNN	Gramy Greek Field in Western Virginia
[27]	ANN	Uinta Basin
[52]	ANN	Iranian Oil Field
[34]	ANN	Abu Dhabi
[32]	ANN	Unknown
[33]	ANN	Asman Oil Field in Southern Iran
[26]	ANN	North Sea
[35]	ANN, MLP, MLR, SVR	Middle Jurassic Shaximiao, Western Sichuan Basin China
[36]	ANN	Middle Eastern Oil Well
[39]	ANN	Persian Gulf Iranian Offshore
[37]	ANN	Middle Eastern Oil Well
[40]	MLP, RBF, GRNN	Kangan and Dallan Formation (South Pars Field Iran)
[41]	Evolving ANN + GA	Mansuri Bangestan (Ahwaz) Iran
[38]	MLP + SVM + CANFIS	Mesaverde Tight Gas, Washakie Basin, USA
[15]	FL	Sarvak and Asmari Formation (Iran)
[46]	Fuzzy c-means + ANN	Kangan Iran (Offshore Gas Field)
[47]	FL	Middle Eastern Oil Well
[48]	FL	Mesozoic Strata Gaoqing
[49]	FL	North Sea and Ula Field South West of Norway
[50]	T2FL	Middle Eastern Oil Well

be associated with permeability but some are more likely than others. They assigned a data bin to each litho-type for FL to learn rules from. Furthermore, a comparison of FL with other methods utilized was made. That ANN needs the right conditions and architecture to perform well. Least square regression (LSR) cannot predict extreme values but FL can while FL cannot predict any value outside the data point, in this case, LSR helps out. Olatunji *et al.* [50] expressed that the level of uncertainties in real life is higher than the one that might be handled by type I fuzzy logic (T1FL). The authors proposed the use of type II fuzzy logic (T2FL) to handle the high uncertainties in real life well data. Table 2 summarizes the published papers by techniques and data sources. Table 3 summarizes the published papers by comparison to other techniques.

**B. MULTI MODELS APPLICATION**

This section discusses works that include multiple models. In literature, some referred to the combination of multiple models as hybrid while others called it ensemble. Hybrid computational intelligence is defined as an effective combination of intelligence techniques that perform superior or in a competitive manner compared to the single technique [43]. The ensemble on the other hand involves learning and integration of multiple models in order to improve the final prediction. Ensemble methods help to reduce the chance of error while increasing the overall reliability and confidence of the model [53]. Different combination approaches are as shown in Figs. 11 and 12. Improving soft computing capability through a combination of models has gained popularity lately [43], [54], [55]. In Fig. 11, models were created from

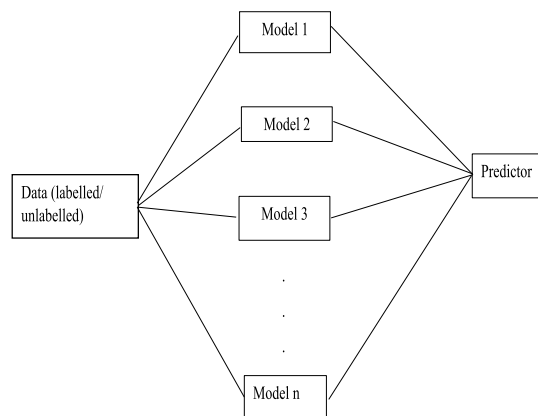


FIGURE 11. Ensembles with single data multiple models.

the overall dataset. Thereafter, each model contributes to the prediction. Several methods abound in the literature for model switching or selection. On the other hand, in Fig. 12, the data is segmented or clustered. Model clusters are then created for each cluster, then they are combined to give the final prediction.

In [54], the authors asserted that determining the structure of ANN is a difficult problem, to overcome this weakness, there is a need for an optimization algorithm. The authors proposed multi-gene genetic programming (MGGP) a variant of GP to obtain the optimal structure of ANN. Although, GP represents a problem with a tree structure. The work combined the selection ability of GP and the estimation power of regression. There is always a tradeoff between the complexity and

TABLE 3. Summary of published papers by comparison to other techniques

References	Technique	Comparison
[45]	FN	Linear and Nonlinear Regression, ANN and ANFIS
[43]	ELM	SVM, ANN
[42]	SVM	GRNN
[51]	SVM	ANN
[35]	ANN, MLP, MLR, SVR	MLR Vs SLR, MLP Vs SVR
[41]	Evolving ANN + GA	Unoptimized ANN + GA
[38]	MLP + SVM + CANFIS	MLP, SVM, CANFIS

TABLE 4. Summary of published papers using hybrid techniques

References	Technique	Source
[54]	MGGP + ANN + ANFIS	Iran Oil Reserve
[64]	T2FL + ELM	Middle Eastern Oil Well
[43]	T2FL + SBLLM	Middle Eastern Oil Well
[56]	PSO + SVM	Middle Eastern Oil Well
[55]	LS-FN + T2FL	North American and Middle Eastern Oil Well
[61]	FN + SVM + T2FL	Middle Eastern Oil Well
[62]	FN + SVM + T2FL	Seismic Properties + Middle Eastern Oil Well
[58]	GA + PSO + ANN	Mansouri Oil Field Iran
[59]	GA-BP, PSO-BP	Mansouri Oil Field Iran

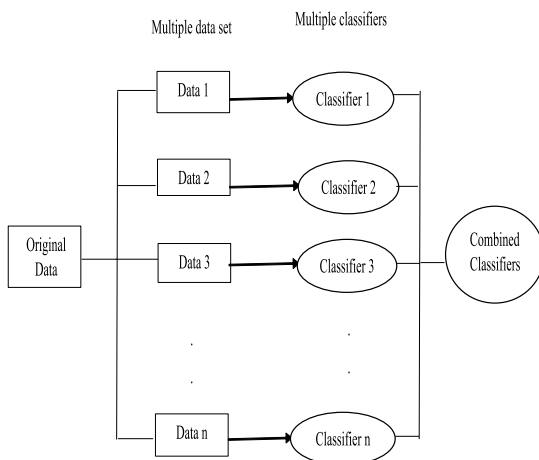


FIGURE 12. Ensembles with multiple data multiple classifiers.

success of the algorithm. The proposed model was applied to a porous media and compared to another model with ANN, ANFIS, and GP only. Similarly, to improve the generalization of ELM, Olatunji *et al.* [43] introduced T2FL to handle uncertainties. As ELM is fast, with better generalization, and avoid local minima, the combined model is used to predict permeability in the Middle Eastern reservoir. A comparison with results obtained from T2FL, ELM, SVM, and ANN as a single model applied to the same data show that the multiple models performed better. In [55], the researchers opined that T2FL is complex and performed poorly with small data. They proposed the hybrid of least square functional and T2FL and their model was tested with North American and Middle Eastern reservoirs. But the researchers in [56] used PSO to select optimal hyperparameters for SVM (PSO-SVM) and applied them to the same data. It shows that PSO-based SVM outperformed the ordinary SVM. In related research, the authors in [57] posited that FIS performed poorly without optimization. They combined FL, LS-SVM, and GA to form two distinctive models made up of GA-FL and GA-LS-SVM. The authors achieved a correlation coefficient

of 0.96 and 0.97 for both models, respectively. Furthermore, Ahmadi *et al.* [58] stated that despite the successes of GA as an optimization algorithm. It is computationally expensive in a large scale optimization problems. The authors proposed the combination of GA and PSO for the benefit of both to optimized ANN for permeability prediction. A similar model was proposed in [59], where GA and PSO were used to optimized BP to generate GA-BP and PSO-BP, respectively. The new hybrids combined the local search ability of BP with the global search abilities of GA and PSO for effective permeability prediction.

Anifowose and Abdurraheem [60] proposed two hybrids from SVM, FN, and FL. The data were divided using stratified sampling methods and processed using the least square fitting algorithm. The two hybrids are FN-T2FL-SVM and FN-SVM-T2FL. The difference is in the pattern of model combination during training and testing. Earlier in [61], the authors used the same approaches in [60] by using FN to select the best predictor variables. Both Anifowose and Abdurraheem [60] and Helmy *et al.* [61] were applied to well logs. However, in [62], the authors deviated a bit by integrating five seismic properties and six oil well logs and applied the same hybrids in [60], [61]. Olatunji *et al.* [43] combined T2FL and sensitivity based linear learning method (SBLLM) [63]. The T2FL was used to handle uncertainties and cleaning of data and rule extraction, thereafter, SBLLM is used for prediction. Table 4 summarizes the published papers using hybrid techniques.

## VI. CONCLUSION

This paper presents a comprehensive review of the available research in the literature on the application of soft computing in solving the permeability prediction problem in the oil and gas industry. The formations studied in the literature were mostly from the Middle East specifically oil fields in Saudi Arabia and Iran. Others are few oil fields from Algeria, Abu Dhabi, China, Canada, and Norway. It is observed that there is limited data available for research in this area. This is

unconnected with the fact that oil companies keep their data close to their chest.

Traditional methods of permeability prediction based on empirical models and theoretical equations as well as models based on porosity and facies are still useful to petroleum engineers. Moreover, soft computing methods are worthy of addition to the solution of permeability prediction. In addition, the cost and time-consuming nature of other approaches such as well logging make soft computing methods to be appealing to oil and gas companies. Although there are no approaches without limitation, soft computing is not an exception. Traditional methods are used as a baseline for measurement with other approaches.

Early works in soft computing techniques focus mainly on ANN and FL, later SVM but recently, there are many research efforts that harnessed optimization techniques to complement ANN in weight selection to improve the performance of ANN. Furthermore, optimization algorithms have been used to improve the performance of SVM through optimal parameter selection.

Ensembles and hybrid computational techniques have received tremendous attention from researchers. This is evident from the number of recent publications on the use of hybrid soft computing models for the prediction of permeability. Nonetheless, more work needs to be done on the standardization of the methods for ensemble integration.

This review provides a comprehensive work in the application of soft computing in permeability prediction. The work serves as an entry point for researchers wanting to delve into this area and reference points for practitioners in oil and gas exploration.

## ACKNOWLEDGMENT

The author gratefully acknowledges the support provided by the University of Hafr Al Batin.

## REFERENCES

- [1] U. Ahmed, S. F. Crary, and G. R. Coates, "Permeability estimation: The various sources and their interrelationships," *J. Petroleum Technol.*, vol. 43, no. 5, pp. 1–10, May 1991.
- [2] T. Babadagli and S. Al-Salmi, "Improvement of permeability prediction for carbonate reservoirs using well log data," in *Proc. SPE Asia Pacific Oil Gas Conf. Exhib.*, Melbourne, VIC, Australia, 2002, pp. 1–17.
- [3] A. Chehrizi and R. Rezaee, "A systematic method for permeability prediction, a petro-facies approach," *J. Petroleum Sci. Eng.*, vols. 82–83, pp. 1–16, Feb. 2012.
- [4] S. Haykin, *Neural Networks: A Comprehensive Foundation*. Stuttgart, Germany: Macmillan Coll Div, Jan. 1994.
- [5] H. Chiroma, S. Abdul-kareem, A. Shukri Mohd Noor, A. I. Abubakar, N. S. Safa, L. Shuib, M. F. Hamza, A. Y. Gital, and T. Herawan, "A review on artificial intelligence methodologies for the forecasting of crude oil price," *Intell. Autom. Soft Comput.*, vol. 22, no. 3, pp. 449–462, Jan. 2016.
- [6] M. Yaghini, M. M. Khoshraftar, and M. Fallahi, "A hybrid algorithm for artificial neural network training," *Eng. Appl. Artif. Intell.*, vol. 26, no. 1, pp. 293–301, Jan. 2013.
- [7] J. Cheng and Q. S. Li, "Reliability analysis of structures using artificial neural network based genetic algorithms," *Comput. Methods Appl. Mech. Eng.*, vol. 197, nos. 45–48, pp. 3742–3750, Aug. 2008.
- [8] J.-S. Chou and Y. Tai, "Generalized regression neural nets in estimating the high-tech equipment project cost," in *Proc. 2nd Int. Conf. Comput. Eng. Appl.*, Bali Island, ID, USA, 2010, pp. 281–284.
- [9] J. Gong and H. Shu, "A time series prediction method based on a modified radial basis function," *Int. J. Simul., Syst., Sci. Technol.*, vol. 17, no. 46, pp. 1–5, 2016.
- [10] F. O. Karray and C. W. De Silva, *Soft Computing and Intelligent Systems Design: Theory, Tools and Applications*. Reading, MA, USA: Addison-Wesley, Jun. 2004.
- [11] A. R. Yildiz, "Cuckoo search algorithm for the selection of optimal machining parameters in milling operations," *Int. J. Adv. Manuf. Technol.*, vol. 64, nos. 1–4, pp. 55–61, Mar. 2012.
- [12] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: A new learning scheme of feedforward neural networks," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, Budapest, Hungary, Jul. 2004, pp. 985–990.
- [13] G.-B. Huang and H. A. Babri, "Upper bounds on the number of hidden neurons in feedforward networks with arbitrary bounded nonlinear activation functions," *IEEE Trans. Neural Netw.*, vol. 9, no. 1, pp. 224–229, Jan. 1998.
- [14] G. Zhou, Y. Zhou, H. Huang, and Z. Tang, "Functional networks and applications: A survey," *Neurocomputing*, vol. 335, pp. 384–399, Mar. 2019.
- [15] S. Alfaouri, M. A. Riahi, N. Alizadeh, and M. Rezaei, "Permeability prediction enhancement in carbonate reservoirs by proposing a new fuzzy logic approach," *J. Appl. Sci.*, vol. 9, no. 18, pp. 3263–3274, Sep. 2009.
- [16] Q. Liang and J. M. Mendel, "Interval type-2 fuzzy logic systems: Theory and design," *IEEE Trans. Fuzzy Syst.*, vol. 8, no. 5, pp. 535–550, Oct. 2000.
- [17] V. N. Vapnik, *Statistical Learning Theory*. Hoboken, NJ, USA: Wiley, 1998.
- [18] N. Cristianini and B. Schölkopf, "Support vector machines and kernel methods: The new generation of learning machines," *AI Mag.*, vol. 23, no. 3, pp. 31–41, Sep. 2002.
- [19] H.-J. Lin and J. P. Yeh, "Optimal reduction of solutions for support vector machines," *Appl. Math. Comput.*, vol. 214, no. 2, pp. 329–335, Aug. 2009.
- [20] M. T. Ahvanooy, Q. Li, M. Wu, and S. Wang, "A survey of genetic programming and its applications," *KSII Trans. Internet Inf. Syst.*, vol. 13, no. 4, pp. 1765–1794, Apr. 2019.
- [21] J. R. Koza, *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. A Bradford Book, Dec. 1992.
- [22] S. C. Lhee, R. R. A. Issa, and I. Flood, "Using particle swarm optimization to predict cost contingency on transportation construction projects," *J. Inf. Technol. Construct.*, vol. 21, pp. 504–516, Dec. 2016.
- [23] N. P. Singh, "Permeability prediction from wireline logging and core data: A case study from assam-arakan basin," *J. Petroleum Explor. Prod. Technol.*, vol. 9, no. 1, pp. 297–305, Jun. 2018.
- [24] S. Mohaghegh, B. Balan, and S. Ameri, "State-of-the-art in permeability determination from well log data: Part 2-verifiable, accurate permeability predictions, the touch-stone of all models," in *Proc. SPE Eastern Regional Meeting*, Morgantown, WV, USA, 1995, pp. 1–5.
- [25] Z. Huang, J. Shimeld, M. Williamson, and J. Katsube, "Permeability prediction with artificial neural network modeling in the Venture gas field, offshore eastern Canada," *Geophysics*, vol. 61, no. 2, pp. 422–436, Mar. 1996.
- [26] H. B. Helle, A. Bhatt, and B. Ursin, "Porosity and permeability prediction from wireline logs using artificial neural networks: A north sea case study," *Geophys. Prospecting*, vol. 49, no. 4, pp. 431–444, Jul. 2001.
- [27] S. Singh, "Permeability prediction using artificial neural network (ANN): A case study of uinta basin," in *Proc. SPE Annu. Tech. Conf. Exhib.*, Dallas, TX, USA, 2005, pp. 1–8.
- [28] J. Ben-Awuah and E. Padmanabhan, "An enhanced approach to predict permeability in reservoir sandstones using artificial neural networks (ANN)," *Arabian J. Geosci.*, vol. 10, no. 7, pp. 1–15, Apr. 2017.
- [29] M. Abdideh, "Estimation of permeability using artificial neural networks and regression analysis in an iran oil field," *Int. J. Phys. Sci.*, vol. 7, no. 34, pp. 5308–5313, Sep. 2012.
- [30] B. Shokoooh Saljooghi and A. Hezarkhani, "A new approach to improve permeability prediction of petroleum reservoirs using neural network adaptive wavelet (wavenet)," *J. Petroleum Sci. Eng.*, vol. 133, pp. 851–861, Sep. 2015.
- [31] A. Mohebbi, R. Kamalpour, K. Keyvanloo, and A. Sarrafi, "The prediction of permeability from well logging data based on reservoir zoning, using artificial neural networks in one of an iranian heterogeneous oil reservoir," *Petroleum Sci. Technol.*, vol. 30, no. 19, pp. 1998–2007, Jul. 2012.
- [32] B. Basbug and Z. T. Karpyn, "Estimation of Permeability from Porosity, Specific Surface Area, and Irreducible Water Saturation using an Artificial Neural Network," in *Proc. Latin Amer. Caribbean Petroleum Eng. Conf.*, Buenos Aires, Argentina, pp. 1–10, Apr. 2007.

- [33] M. Jamialahmadi and F. G. Javadpour, "Relationship of permeability, porosity and depth using an artificial neural network," *J. Petroleum Sci. Eng.*, vol. 26, nos. 1–4, pp. 235–239, May 2000.
- [34] V. Badarinadh, K. Suryanarayana, F. Z. Youssef, K. Sahouh, and A. Valle, "Log-derived permeability in a heterogeneous carbonate reservoir of middle east, abu dhabi, using artificial neural network," in *Proc. Abu Dhabi Int. Petroleum Exhib. Conf.*, Villahermosa, Mexico, 2002, pp. 1–13.
- [35] G. Zhang, Z. Wang, H. Li, Y. Sun, Q. Zhang, and W. Chen, "Permeability prediction of isolated channel sands using machine learning," *J. Appl. Geophys.*, vol. 159, pp. 605–615, Dec. 2018.
- [36] G. M. Hamada and M. A. Elshafei, "Neural network prediction of porosity and permeability of heterogeneous gas sand reservoirs," in *Proc. SPIE Saudi Arabia Sect. Tech. Symp.*, Al-Khobar, Saudi Arabia, 2009, pp. 1–12.
- [37] S. Elkattany, M. Mahmoud, Z. Tariq, and A. Abdulraheem, "New insights into the prediction of heterogeneous carbonate reservoir permeability from well logs using artificial intelligence network," *Neural Comput. Appl.*, vol. 30, no. 9, pp. 2673–2683, Jan. 2017.
- [38] S. Baziar, M. Tadayoni, M. Nabi-Bidhendi, and M. Khalili, "Prediction of permeability in a tight gas reservoir by using three soft computing approaches: A comparative study," *J. Natural Gas Sci. Eng.*, vol. 21, pp. 718–724, Nov. 2014.
- [39] P. Tahmasebi and A. Hezarkhani, "A fast and independent architecture of artificial neural network for permeability prediction," *J. Petroleum Sci. Eng.*, vols. 86–87, pp. 118–126, May 2012.
- [40] P. Bagheripour, "Committee neural network model for rock permeability prediction," *J. Appl. Geophys.*, vol. 104, pp. 142–148, May 2014.
- [41] R. Irani and R. Nasimi, "Evolving neural network using real coded genetic algorithm for permeability estimation of the reservoir," *Expert Syst. Appl.*, vol. 38, no. 8, pp. 9862–9866, Aug. 2011.
- [42] T. Aïfa, R. Baouche, and K. Baddari, "Neuro-fuzzy system to predict permeability and porosity from well log data: A case study of hassi RE5Mel gas field, algeria," *J. Petroleum Sci. Eng.*, vol. 123, pp. 217–229, Nov. 2014.
- [43] S. Olatunji and A. Selamat, "Extreme learning machines based model for predicting permeability of carbonate reservoir," *Int. J. Digit. Content Technol. Appl.*, vol. 7, no. 1, pp. 450–459, Jan. 2013.
- [44] E. A. El-Sebakhy, "Functional networks training algorithm for statistical pattern recognition," in *Proc. 9th Int. Symp. Comput. Commun.*, Alexandria, Egypt, 2004, pp. 92–97.
- [45] E. A. El-Sebakhy, O. Asparouhov, A.-A. Abdulraheem, A.-A. Al-Majed, D. Wu, K. Latinski, and I. Raharja, "Functional networks as a new data mining predictive paradigm to predict permeability in a carbonate reservoir," *Expert Syst. Appl.*, vol. 39, no. 12, pp. 10359–10375, Sep. 2012.
- [46] A. K. Ilkhchi, M. Rezaee, and S. A. Moallemi, "A fuzzy logic approach for estimation of permeability and rock type from conventional well log data: An example from the kangan reservoir in the iran offshore gas field," *J. Geophys. Eng.*, vol. 3, no. 4, pp. 356–369, Dec. 2006.
- [47] A. Abdulraheem, E. A. El-Sebakhy, M. Ahmed, A. Vantala, P. D. Raharja, and G. Korvin, "Estimation of permeability from wireline logs in a middle eastern carbonate reservoir using fuzzy logic," in *Proc. SPIE Middle East Oil Gas Show Conf.*, Manama, Bahrain, Mar. 2007, pp. 1–11.
- [48] X. Wang, S. Yang, Y. Wang, Y. Zhao, and B. Ma, "Improved permeability prediction based on the feature engineering of petrophysics and fuzzy logic analysis in low porosity–permeability reservoir," *J. Petroleum Explor. Prod. Technol.*, vol. 9, no. 2, pp. 869–887, Oct. 2018.
- [49] S. J. Cuddy, "Litho-facies and permeability prediction from electrical logs using fuzzy logic," *SPE Reservoir Eval. Eng.*, vol. 3, no. 4, pp. 1–6, Aug. 2000.
- [50] S. O. Olatunji, A. Selamat, and A. A. A. Raheem, "Predicting correlations properties of crude oil systems using type-2 fuzzy logic systems," *Expert Syst. Appl.*, vol. 38, no. 9, pp. 10911–10922, Sep. 2011.
- [51] R. Gholami, A. Moradzadeh, S. Maleki, S. Amiri, and J. Hanachi, "Applications of artificial intelligence methods in prediction of permeability in hydrocarbon reservoirs," *J. Petroleum Sci. Eng.*, vol. 122, pp. 643–656, Oct. 2014.
- [52] H. Kaydani, A. Mohebbi, and A. Baghaie, "Permeability prediction based on reservoir zonation by a hybrid neural genetic algorithm in one of the iranian heterogeneous oil reservoirs," *J. Petroleum Sci. Eng.*, vol. 78, no. 2, pp. 497–504, Aug. 2011.
- [53] R. Polikar, *Ensemble Learning*. Springer, 2012, pp. 1–34.
- [54] H. Kaydani and A. Mohebbi, "A comparison study of using optimization algorithms and artificial neural networks for predicting permeability," *J. Petroleum Sci. Eng.*, vol. 112, pp. 17–23, Dec. 2013.
- [55] F. Anifowose, J. Labadin, and A. Abdulraheem, "A least-square-driven functional networks type-2 fuzzy logic hybrid model for efficient petroleum reservoir properties prediction," *Neural Comput. Appl.*, vol. 23, no. 1, pp. 179–190, Jan. 2013.
- [56] K. O. Akande, T. O. Owolabi, S. O. Olatunji, and A. AbdulRaheem, "A hybrid particle swarm optimization and support vector regression model for modelling permeability prediction of hydrocarbon reservoir," *J. Petroleum Sci. Eng.*, vol. 150, pp. 43–53, Feb. 2017.
- [57] M.-A. Ahmadi, M. R. Ahmadi, S. M. Hosseini, and M. Ebadi, "Connectionist model predicts the porosity and permeability of petroleum reservoirs by means of petro-physical logs: Application of artificial intelligence," *J. Petroleum Sci. Eng.*, vol. 123, pp. 183–200, Nov. 2014.
- [58] M. Ali Ahmadi, S. Zendejboudi, A. Lohi, A. Elkamel, and I. Chatzis, "Reservoir permeability prediction by neural networks combined with hybrid genetic algorithm and particle swarm optimization," *Geophys. Prospecting*, vol. 61, no. 3, pp. 582–598, May 2013.
- [59] R. Nasimi and R. Irani, "Combining a neural network with a genetic algorithm and particle swarm optimization for permeability estimation of the reservoir," *Energy Sour., A, Recovery, Utilization, Environ. Effects*, vol. 37, no. 4, pp. 384–391, Jan. 2015.
- [60] F. Anifowose and A. Abdulraheem, "Fuzzy logic-driven and SVM-driven hybrid computational intelligence models applied to oil and gas reservoir characterization," *J. Natural Gas Sci. Eng.*, vol. 3, no. 3, pp. 505–517, Jul. 2011.
- [61] T. Helmy, A. Fatai, and K. Faisal, "Hybrid computational models for the characterization of oil and gas reservoirs," *Expert Syst. Appl.*, vol. 37, no. 7, pp. 5353–5363, Jul. 2010.
- [62] F. Anifowose, S. Adeniyi, and A. Abdulraheem, "Recent advances in the application of computational intelligence techniques in oil and gas reservoir characterisation: A comparative study," *J. Experim. Theor. Artif. Intell.*, vol. 26, no. 4, pp. 551–570, Jun. 2014.
- [63] A. Selamat, S. O. Olatunji, and A. A. A. Raheem, "A hybrid model through the fusion of Type-2 fuzzy logic systems and sensitivity-based linear learning method for modeling PVT properties of crude oil systems," *Adv. Fuzzy Syst.*, vol. 2012, pp. 1–19, 2012.
- [64] S. O. Olatunji, A. Selamat, and A. Abdulraheem, "A hybrid model through the fusion of type-2 fuzzy logic systems and extreme learning machines for modelling permeability prediction," *Inf. Fusion*, vol. 16, pp. 29–45, Mar. 2014.



**MUBARAK SAAD ALMUTAIRI** (Member, IEEE) received the B.S. degree in systems engineering from the King Fahd University of Petroleum and Minerals (KFUPM), Dhahran, Saudi Arabia, in 1997, the M.S. degree in industrial and systems engineering from the University of Florida, Gainesville, FL, USA, in 2003, and the Ph.D. degree in systems design engineering from the University of Waterloo, Waterloo, ON, Canada, in 2007. From 1997 to 2000, he was an Industrial Engineer with Saudi Arabian Oil Company, Saudi Aramco. He was a recipient of three research grants in the field of ITC. He has edited two books and published 22 journal, conference papers, and book chapters. He served as a member of the KFUPM board, from 2011 to 2015 and serving as a member of the University of Hafr Al Batin board, since 2015. He is a member of IEEE (USA), Senior Member IIE (USA), ASNT (USA), ASEE (USA), and SCE (Saudi Arabia). He was also a recipient of the 2007 Distinguished Teaching Award at the University of Waterloo. He is currently an Associate Professor with the Computer Science and Engineering Department, University of Hafr Al Batin (UHB), Hafr Al Batin, Saudi Arabia. He is also currently serving as the Dean of Hafr Al Batin Community College, UHB. His research interests include decision analysis, expert systems, risk assessment, information security, fuzzy logic, and mobile government applications.

...